Reliable Autonomy Under Uncertainty: From Learning-Based to Non-rational Control

Thesis by
Taylan Kargin

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

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2025
Defended May 16, 2025

© 2025

Taylan Kargin
ORCID: 0000-0001-6744-654X

All rights reserved except where otherwise noted

ACKNOWLEDGEMENTS

First and foremost, I owe an immeasurable debt of gratitude to my advisor, Professor Babak Hassibi, for granting me the extraordinary privilege of pursuing this research under his mentorship at one of the world's premier research institutions. I am profoundly grateful for the intellectual liberty he afforded me, encouraging me to indulge my curiosity wherever it led while always offering insightful guidance and unwavering support when I needed direction. His remarkable breadth and depth of knowledge, visionary outlook, incisive questions, and rare ability to grasp the big picture while upholding deep technical rigor and meticulous attention to detail have been a continual source of inspiration. These qualities have profoundly shaped my growth as a scholar, an influence clearly reflected throughout the pages of this dissertation.

Beyond the technical dimension, I have been equally fortunate to benefit from Professor Hassibi's infinite forbearance, generosity, and encouragement throughout the many ups and downs of my doctoral journey. His unwavering support extended far beyond research, from reading countless drafts with care to his efforts in facilitating meaningful collaborations, and creating valuable professional opportunities. I am especially thankful for his generous support in helping me share our work at numerous international conferences, which greatly broadened my horizons and strengthened my academic network. For his wisdom, forbearance, and steadfast belief in my potential, I remain deeply and sincerely grateful.

ABSTRACT

Autonomous systems are profoundly reshaping our societies, industries, and daily lives, delivering unprecedented levels of efficiency, innovation, and adaptability. From self-driving vehicles navigating dense urban traffic and coordinated swarms of search-and-rescue robots operating in hazardous environments, to next-generation intelligent power grids and high-precision industrial automation, these systems are increasingly deployed in safety-critical and high-stakes settings where they are routinely entrusted with split-second decisions that carry profound economic and lethal consequences. In such contexts, the imperative for reliability, safety, and robustness is paramount: a single unanticipated failure within a power distribution network can trigger extensive blackouts, and a momentary lapse in decision-making or perception by an autonomous vehicle can endanger lives.

Despite their remarkable capabilities, securing such reliability guarantees faces formidable and multifaceted challenges. The environments in which these systems operate are characterized by unprecedented complexity, vast scale, and pervasive uncertainty as they frequently interact with numerous external entities such as humans or other autonomous agents whose behaviors may be volatile, adversarial, or fundamentally unknown. Explicitly and exhaustively modeling this complexity a priori is practically infeasible, compelling systems to infer, adapt, and respond to the novel environments by learning from data. Although contemporary machine-learning models afford expressive representations, their assurances are limited by the scope and fidelity of their training data. Consequently, such models remain vulnerable to distribution shifts, rare events, or unmodeled edge cases, which can precipitate catastrophic failure.

Further complicating matters, real-world applications frequently impose stringent resource constraints, including limited computation, memory, communication, and power. These constraints demand principled trade-offs between competing performance objectives and operational constraints such as safety, stability, robustness, and efficiency, especially in high-stakes and uncertainty-laden settings. This dissertation addresses these challenges by contributing fundamental theoretical results and practical computational tools towards provably reliable, resource-efficient, and scalable autonomy.

Operating safely in dynamic and a priori unknown environments poses a fundamental

challenge for autonomous systems: balancing *exploration*, *i.e.*, the pursuit of long-term optimality by probing uncertain policy landscape at the risk of degraded safety, against *exploitation*, *i.e.*, leveraging current knowledge to ensure short-term performance and stability at the expense of settling for a suboptimal policy. In Part I, we study online reinforcement learning approaches for unknown linear dynamical systems to address this challenge. We present computationally efficient algorithms for online learning and control in both state-feedback and measurement-feedback settings that operate safely without any prior knowledge of the system. We rigorously establish their feasibility through finite-time guarantees on performance, computational complexity, and stability, matching the fundamental theoretical bounds.

Statistical models underlie every layer of an autonomous system, serving as representations of complex data-generating phenomena. Typically constructed from empirical data through a blend of explicit modeling, machine learning, and simulation, these models are vulnerable to distribution shift, *i.e.*, discrepancies between design and deployment conditions, which can jeopardize both performance and safety. In Part II, we investigate distributionally robust optimization (DRO) methods for control, prediction, communication, and unsupervised learning to guard against model misspecification and distribution shifts. DRO blends average-case optimality with worst-case guarantees: by maximizing expected performance against the least-favorable statistical model consistent with the available data, it strikes a balanced trade-off between robustness and performance informed by data.

Autonomous control systems must often balance several performance goals, such as cost efficiency, robustness, risk tolerance, and stability, while meeting practical constraints such as suitability for real-time implementation and scalability. Because these design problems are inherently infinite-dimensional, only a handful of special cases admit exact, tractable solutions (e.g., Linear-Quadratic-Gaussian, $\mathscr{H}_\infty$-optimal, or regret-optimal control) while widely studied formulations like mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control remain unresolved. In Part III we present *non-rational control*, a unified framework that makes many such problems both solvable and implementable. The key is an optimize-then-approximate strategy that delivers provably near-optimal, stabilizing, finite-order (rational) controllers even when the true optimum resides in an infinite-dimensional (non-rational) policy space.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] J. Hajar*, T. Kargin*, V. Malik, and B. Hassibi, "Distributionally robust kalman filtering over an infinite-horizon," in *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025, pp. 1–5. DOI: `10.1109/ICASSP49660.2025.10889742`,
T.K. is the co-first author of this paper. He participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[2] J. Hajar, T. Kargin, V. Malik, and B. Hassibi, "The distributionally robust infinite-horizon lqr," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, 2024, pp. 6539–6546. DOI: `10.1109/CDC56724.2024.10886462`,
T.K. participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[3] T. Kargin*, J. Hajar*, V. Malik, and B. Hassibi, *Distributionally Robust Kalman Filtering over Finite and Infinite Horizon*, arXiv:2407.18837, Jul. 2024. [Online]. Available: `http://arxiv.org/abs/2407.18837`,
T.K. is the co-first author of this paper. He participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[4] T. Kargin*, J. Hajar*, V. Malik*, and B. Hassibi, "Infinite-Horizon Distributionally Robust Regret-Optimal Control," in *Proceedings of the 41st International Conference on Machine Learning*, R. Salakhutdinov, Z. Kolter, K. Heller, *et al.*, Eds., ser. Proceedings of Machine Learning Research, vol. 235, PMLR, Jul. 2024, pp. 23 187–23 214. [Online]. Available: `https://proceedings.mlr.press/v235/kargin24a.html`,
T.K. is the co-first author of this paper. He participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[5] T. Kargin*, J. Hajar*, V. Malik*, and B. Hassibi, "Wasserstein distributionally robust regret-optimal control over infinite-horizon," in *Proceedings of the 6th Annual Learning for Dynamics & Control Conference*, A. Abate, M. Cannon, K. Margellos, and A. Papachristodoulou, Eds., ser. Proceedings of Machine Learning Research, vol. 242, PMLR, Jul. 2024, pp. 1688–1701. [Online]. Available: `https://proceedings.mlr.press/v242/kargin24a.html`,
T.K. is the co-first author of this paper. He participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[6] V. Malik, T. Kargin, V. Kostina, and B. Hassibi, "A Distributionally Robust Approach to Shannon Limits using the Wasserstein Distance," in *2024 IEEE International Symposium on Information Theory (ISIT)*, Athens, Greece: IEEE, Jul. 2024, pp. 861–866, ISBN: 9798350382846. DOI: `10.1109/ISIT57864.2024.10619597`. [Online]. Available: `https:`

//ieeexplore.ieee.org/document/10619597/,
T.K. participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[7] V. Malik*, T. Kargin*, J. Hajar, and B. Hassibi, "Optimal infinite-horizon mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control," in *2024 60th Annual Allerton Conference on Communication, Control, and Computing*, 2024, pp. 1–8. DOI: 10.1109/Allerton63246.2024.10735276,
T.K. is the co-first author of this paper. He participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[8] J. Hajar, T. Kargin, and B. Hassibi, "Wasserstein distributionally robust regret-optimal control under partial observability," in *2023 59th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2023, pp. 1–6. DOI: 10.1109/Allerton58177.2023.10313386,
T.K. participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[9] T. Kargin, S. Lale, K. Azizzadenesheli, A. Anandkumar, and B. Hassibi, "Thompson Sampling for Partially Observable Linear-Quadratic Control," in *2023 American Control Conference (ACC)*, San Diego, CA, USA: IEEE, May 2023, pp. 4561–4568, ISBN: 9798350328066. DOI: 10.23919/ACC55779.2023.10156461. [Online]. Available: https://ieeexplore.ieee.org/document/10156461/,
T.K. participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[10] T. Kargin*, J. Hajar*, V. Malik*, and B. Hassibi, *Wasserstein distributionally robust regret-optimal control in the infinite-horizon*, 2023. arXiv: 2312.17376 [eess.SY]. [Online]. Available: https://arxiv.org/abs/2312.17376,
T.K. participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

[11] T. Kargin*, S. Lale*, K. Azizzadenesheli, A. Anandkumar, and B. Hassibi, "Thompson sampling achieves $\tilde{O}(\sqrt{T})$ regret in linear quadratic control," in *Proceedings of Thirty Fifth Conference on Learning Theory*, P.-L. Loh and M. Raginsky, Eds., ser. Proceedings of Machine Learning Research, vol. 178, PMLR, Jul. 2022, pp. 3235–3284. [Online]. Available: https://proceedings.mlr.press/v178/kargin22a.html,
T.K. is the co-first author of this paper. He participated in the conception of the project, the analysis of the problem, and the writing of the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

The progression of modern society is increasingly driven by the advancement of autonomous systems, which are rapidly transforming technological, industrial, and societal landscapes. These systems, characterized by their ability to perceive, decide, and act with minimal human intervention, cohesively integrate computational intelligence with physical components to execute complex tasks in real time. Their growing adoption is driving unprecedented gains in efficiency, productivity, and economic growth across sectors such as transportation, aerospace, manufacturing, energy, and infrastructure.

Indeed, the scale of this transformation is staggering: the International Federation of Robotics reports a record 4.28 million industrial robots on factory floors worldwide in 2024 [110]. Mobility is undergoing a comparable transformation. McKinsey & Company projects that by 2035, autonomous driving could generate $300 billion to $400 billion in annual revenue [164]. Energy sector likewise relies increasingly on autonomous and intelligent technologies. According to IoT Analytics counts, utilities had installed about 1.06 billion smart electricity meters by late 2023, roughly 43% of all global meters [112].

These trends underscore a profound transition: Increasingly, real-time decisions with *high-stakes* physical, economic, and societal consequences are being delegated to algorithms operating in *complex, uncertain*, and *resource-constrained* environments. As the autonomy and influence of these systems grow, so too does the need for rigorous frameworks that ensure their safe, reliable, and efficient operation.

- **High Stakes and Safety Critical.** Autonomous systems increasingly interact with humans and critical infrastructure, or frequently function in environments where safety is non-negotiable, and failures can be catastrophic. The U.S. National Highway Traffic Safety Administration's Standing General Order on crash reporting has documented 1612 crashes involving Level-2 driver-assistance ("autopilot") systems across more than a dozen automakers between July 2021 and mid-2024, 40 of which resulted in fatalities [169]. A misconfigured state estimator triggered a widespread power outage on August 14, 2003 that cut electricity to nearly 55 million people in large swathes of United States and Canada, causing an estimated $10 billion in losses [229]. These examples highlight the critical necessity for assurances of safety, reliability and robustness to prevent consequential outcomes.

- **Complex and Pervasive Uncertainties.** Delegating increasingly complex and high-stakes tasks to autonomous systems exposes them to pervasive and multifaceted uncertainties. Real-world operating conditions are typically dynamic, volatile, and only partly observable. Consider urban traffic, where self-driving vehicles must anticipate the intentions of human drivers and pedestrians while coping with sudden adverse shifts in weather or road conditions. At the very extreme, search-and-rescue robots navigating collapsed structures or planetary rovers traversing uncharted Martian terrain encounter environments that are entirely unknown and unstructured. Such complexity stretches the limits of traditional, explicit model-based design and engineering. Therefore, operating reliably under these circumstances demands the capacity to learn from sparse interactions and available data to explore safely and adapt the response in real time.

  Learning-based models offer far greater expressive power for capturing rich and complex uncertainties than hand-crafted, explicit mathematical models. Yet that flexibility comes at a price: formal guarantees of reliability and safety are far harder to establish. Since performance hinges on the quality and representativeness of the training data, these models can be brittle under variations in the development and deployment environments or when confronted with rare, safety-critical edge cases absent from the dataset. Even a small deviation in the model can propagate through downstream tasks and precipitate catastrophic failures. A stark illustration is the 2018 fatal crash involving an autonomous SUV operated by Uber in Tempe, Arizona, where the perception network misclassified a jay-walking pedestrian at night as "vehicle" or "bicycle" and withheld emergency braking until just 0.2 seconds before impact [170].

- **Performance and Resource Constraints.** Increasingly complex real-world tasks rarely fit into a single performance metric. Instead, they often involve multiple, distinct, and often conflicting performance criteria and operational constraints that must be carefully balanced through informed trade-offs. A good illustration is modern power-grid operation: grid operators must (i) ensure power supply robustly meets highly fluctuating demand even under adverse weather events, (ii) minimize operating costs for generators and consumers, and (iii) reduce CO_2 emissions by integrating intermittent renewable sources and routing power efficiently through the network.

These demands are further compounded in large-scale cyber-physical systems (CPS) by structural and resource constraints that hinder real-time implementation. Limitations in computational power, memory, data availability, energy, and communication bandwidth or latency must all be carefully managed. For instance, drones in coordinated fleets for surveillance or disaster response must balance coverage, tracking, and monitoring while trading off speed, accuracy, and energy. Operating under limited battery, on-board computation, and noisy sensing, they also face intermittent, bandwidth-limited communication, requiring decentralized decisions based on local data.

In light of these omnipresent, inseparable challenges, this dissertation is guided by a single, overarching aim:

> *To design **provably reliable**, **resource-efficient**, and **scalable** autonomous systems that can **learn and adapt** to novel environments by leveraging **data-driven** methods to balance robustness, performance, and competing objectives through principled **reasoning about uncertainty, risk, and trade-offs**.*

Guided by this objective, the dissertation is organized into three interrelated parts, each addressing a distinct but complementary facet of this overarching goal:

- **Part I: Learning and Control.** Operating safely in unknown and dynamic environments requires autonomous systems to balance exploration and exploitation. In Part I, we develop computationally efficient reinforcement learning algorithms for controlling unknown linear systems, both with full and partial observations. These methods achieve finite-time guarantees on performance, stability, and complexity, enabling safe real-time operation without prior system knowledge.
- **Part II: Distributionally Robust Optimization.** Statistical models, central to autonomous systems, often face distribution shifts between training and deployment, threatening performance and safety. In Part II, we develop distributionally robust optimization (DRO) methods for control, prediction,

communication, and learning, which hedge against model uncertainty by optimizing expected performance under the worst-case distribution consistent with the data, offering a principled and data-informed balance between worst-case robustness and average-case performance.

- **Part III: Non-rational Control.** We introduce a unified framework for stabilizing and scalable controller synthesis tailored to diverse infinite-horizon control objectives, subsuming distributionally robust, risk-sensitive, mixed criteria, etc. An "optimize-then-approximate" pipeline converts these infinite-dimensional formulations into finite-dimensional controllers with quantifiable sub-optimality, enabling scalable real-time implementation while preserving the rigorous guarantees demanded in safety-critical autonomy.

The remainder of this chapter outlines the primary contributions of the thesis, organized by each of its three main parts.

1.1 Outline and Scope of Part I: Learning and Control

In many real-world applications, the governing dynamics of the system are not fully known in advance. This lack of prior knowledge fundamentally complicates the real-time control task, as effective decision-making must proceed in tandem with learning the system itself. In such scenarios, control and learning become inherently coupled: the autonomous agent must actively gather information about the environment through interaction while simultaneously striving to achieve its control objectives.

This dual necessity gives rise to the classic *exploration-exploitation dilemma*. On one hand, an autonomous agent must *explore* the unknown system to acquire information about its dynamics and thereby adapt and improve its policy in the long run. However, reckless exploration without regard for control objectives can compromise safety and degrade performance before any gains are realized. Conversely, the agent must also *exploit* its acquired knowledge to maintain safety and acceptable performance, but over-reliance on exploitation risks becoming trapped in a suboptimal policy, never discovering more effective policies. Ultimately, the agent's challenge is to strike the right balance between these competing objectives: gathering enough information to drive future improvements while ensuring safe, reliable operation and high performance at every step.

Part III of this thesis is concerned with this dilemma in the context of controlling a priori unknown linear dynamical systems in real time. In general, these systems can be described as a state-space model as

$$x_{t+1} = A_\star x_t + B_\star u_t + w_t, \tag{1.1a}$$

$$y_t = C_\star x_t + v_t, \tag{1.1b}$$

where x_t denotes the system state, u_t the control input, y_t the observation, and w_t, v_t the process and observation noise, respectively. The system matrices $(A_\star, B_\star, C_\star)$ are unknown to the agent in advance. The overarching goal of the autonomous agent is to execute control actions in closed-loop with an unknown system that (i) optimize long-term performance, (ii) ensure system stability and safety, and (iii) do so with minimal computational and sample complexity, all while contending with model uncertainty and navigating the fundamental trade-off between exploration and exploitation to meet these objectives.

A central question in this setting is to understand the fundamental performance limits of learning-based control. A widely used metric for quantifying performance is *regret*: the cumulative excess cost incurred by a learning agent relative to a baseline policy, typically the optimal policy of an oracle agent with full knowledge of the system dynamics. Regret thus captures the price of uncertainty, reflecting how much suboptimality is suffered by learning on the fly. Ideally, one seeks regret that grows sublinearly with the time horizon T, the number of agent-environment interactions, so that the average per-step cost approaches optimality over time. For the case of unknown linear dynamical systems, the minimax optimal regret rate has been shown to scale as $\sqrt{T}$ [210] for state-feedback control (*i.e.*, x_t observed). While several algorithms have been proposed achieving this optimal performance, some of these work under the assumption that a stabilizing controller is known a priori [159], [210], a requirement often not met in real-world situations, circumventing the need for balancing exploration and exploitation.

Alternatively, the Optimism in the Face of Uncertainty (OFU) principle has been shown to achieve both optimal regret and system stabilization without relying on any prior knowledge of the dynamics [2], [140]. Widely used in online decision-making under uncertainty (e.g., in bandit problems), this approach constructs confidence sets for the unknown system parameters based on observed data and then selects the control policy corresponding to the most optimistic model, *i.e.*, the one with the lowest predicted control cost, within the confidence set. While OFU enjoys strong theoretical guarantees, including regret optimality and stability, it suffers from significant computational intractability, as identifying the optimistic model requires solving a non-convex, NP-hard optimization problem, making it impractical for any real-time implementation. This leads to a fundamental question at the heart of learning-based control [4]:

> *Can we design a computationally feasible (i.e.,* poly*-time) state-feedback control algorithm that provably stabilizes any a priori unknown linear dynamical system while achieving the order-optimal $\sqrt{T}$ regret rate?*

In **Chapter 2**, we introduce Thompson Sampling-based Adaptive Control (TSAC), a provably efficient (*i.e.*, poly-time) and stabilizing state-feedback control algorithm with optimal $\sqrt{T}$ regret for any a priori unknown linear dynamical system, thereby affirmatively answering this question. Rather than performing an expensive search for the optimistic model, this algorithm executes the stabilizing optimal policy for a randomly selected model from a confidence set, requiring only a constant time compute. Similar to OFU, it balances exploration and exploitation by progressively reducing uncertainty in the system estimates, thereby biasing the sampling distribution toward models that yield lower control cost.

Although TSAC attains the order-optimal $\sqrt{T}$ regret, it does so by trading computational efficiency for a multiplicative correction factor in the regret, which depends inversely on the probability of sampling an optimistic model. Previous work demonstrated a Thompson-sampling approach with similar guarantees [4], but only in the special case of scalar systems. One of our key contributions is to show that the additional regret incurred due to randomized sampling, rather than explicitly searching for the most optimistic model, remains bounded in arbitrary multidimensional systems. In particular, TSAC samples sufficiently optimistic models with high enough frequency to match the optimal $\sqrt{T}$ regret without compromising performance.

In **Chapter 3**, we extend the setting from fully observed state-feedback systems to partially observed, noise-corrupted systems by introducing Thompson Sampling under Partial Observability (TSPO). This algorithm leverages a novel closed-loop system identification procedure to iteratively refine both the parameter estimates and their associated confidence intervals. TSPO then employs Thompson Sampling guided by these intervals to design control policies for unknown Linear-Quadratic-Gaussian

(LQG) systems. We show that this method is not only computationally efficient but also achieves the optimal $\sqrt{T}$ regret rate while ensuring system stabilization.

1.2 Outline and Scope of Part II: Distributionally Robust Optimization

Statistical models form the backbone of modern autonomous decision-making systems by providing probabilistic representations of the external variables and processes that influence a system's behavior. In stochastic optimization, one uses these models to choose decision rules that optimize average-case performance with respect to the probability distribution implied by the model. This framework underpins a wide range of tasks, including optimal control under stochastic noise and uncertainty, predictive modeling and estimation in sensing and perception, learning from data, and reliable data processing, compression, and transmission over noisy communication channels.

Most downstream tasks, from stabilizing a drone in turbulence to filtering sensor noise in autonomous vehicles, rely critically on how faithfully those representations mirror reality. In practice, these models are typically constructed from empirical data (e.g., via machine learning), from high-fidelity simulations (e.g., robotic motion planning and climate modeling), from first-principles or hand-crafted models, or some combination thereof. Yet the world encountered in deployment rarely matches the conditions seen during design: discrepancies arise when the operational data-generating processes differ from the one used during model construction, a phenomenon known as distribution shift. If left unaddressed, such distribution shifts and model misspecifications can have severe consequences in high-stakes and safety-critical applications. For example, in autonomous driving systems, a model trained predominantly on clear daytime driving scenarios may perform poorly under nighttime or adverse weather conditions, potentially leading to unsafe decisions [170].

To ensure safety and reliability, it is essential that autonomous systems account for the risks posed by model misspecification and distribution shifts. Historically, robustness in decision-making has been addressed through the lens of adversarially robust optimization, which forgoes probabilistic models entirely and instead seeks to optimize performance under the worst-case realization of uncertainty. While this approach provides strong worst-case guarantees, it often proves overly conservative: discarding prior statistical information sacrifices average-case performance in exchange for blanket robustness. Bridging the gap between epistemic optimism of stochastic optimization and caution of adversarially robust optimization remains a central challenge in designing dependable data-driven decision systems.

Against this backdrop, distributionally robust optimization (DRO) has emerged as a principled remedy. At its core, DRO retains an ambiguity set of plausible statistical models consistent with the available data and optimizes expected performance under the worst-case distribution in that set. Concretely, the canonical problem setup of DRO can be formulated as a minimax optimization:

$$\inf_{\theta \in \Theta} \sup_{\mathbb{P} \in \mathbb{B}_r(\mathbb{P}_\circ)} \mathbb{E}_{\xi \sim \mathbb{P}} \left[\ell(\theta, \xi) \right], \tag{1.2}$$

where $\theta \in \Theta$ is the decision variable, $\xi \in \Xi$ is the uncertain variable, and $\ell : \Theta \times \Xi \to \mathbb{R}$ is the loss function. In practice, the decision-maker typically possesses a nominal model $\mathbb{P}_\circ$ for the uncertain variable, but lacks confidence in its accuracy in capturing the true, unknown data-generating distribution. To account for this model uncertainty, the decision maker constructs an ambiguity set as a "ball" of probability distributions $\mathbb{B}_r(\mathbb{P}_\circ)$ centered at the nominal $\mathbb{P}_\circ$ with a specified radius to contain all distributions deemed plausibly close to the nominal, including the true distribution. The precise form of this set depends on the chosen notion of distributional similarity, which may be defined via a divergence (e.g., Kullback-Leibler) or a distance metric (e.g., Wasserstein, total variation).

The choice of radius effectively reflects the decision-maker's confidence in the fidelity of the nominal model. As the radius approaches zero, DRO reduces to standard stochastic optimization under the nominal distribution. Conversely, as the radius increases, the formulation becomes increasingly conservative, eventually converging to the adversarial robust optimization setting. This flexibility enables a data-driven trade-off between robustness and performance, seamlessly bridging the gap between stochastic and worst-case decision-making.

Part II of this thesis focuses on the safe and reliable mitigation of risks arising from distribution shifts and model uncertainty in control, prediction, communication, and unsupervised learning, through the lens of distributionally robust optimization (DRO). A central emphasis is placed on ambiguity sets defined via Wasserstein distances (also known as optimal transport metrics). In contrast to divergence-based ambiguity sets, which are restricted to distributions supported on the same set as the nominal model, Wasserstein ambiguity sets offer greater expressive power, respect the underlying geometric structure of the data space, and lend themselves to tractable convex reformulations [74], [134].

A particularly challenging setting arises in the control and prediction of partially observed dynamical systems subject to complex, temporally correlated, and poorly understood disturbances. In real-world environments, such systems are continually influenced by disturbances with rich temporal dependencies and structural complexities that are poorly captured by oversimplified probabilistic models, such as Gaussian noise. As a result, standard tools like Linear-Quadratic-Gaussian (LQG) control and the Kalman filter, which rely heavily on assumptions of Gaussianity and temporal independence, can become both ineffective and unreliable. The challenge is further compounded when system parameters are imperfectly known, as modeling errors introduce additional structured disturbances, effectively amplifying uncertainty.

In **Chapter 4** and **Chapter 5**, we develop distributionally robust control and filtering frameworks for partially observed linear systems operating under correlated disturbances. Departing from most prior work that assumes independence across time [153], [203], [220], [248], our approach leverages the Wasserstein distance to model uncertainty in the *joint* distribution of the *entire disturbance trajectory* over a *finite horizon*, enabling reliable control and prediction under rich, temporally correlated uncertainty.

Another critical capability for autonomous systems is the reliable real-time transmission of high-dimensional data—such as images and videos—over long distances. This is especially vital in networked cyber-physical systems with multiple interconnected agents, such as search-and-rescue drones or robotic teams, where safe and effective coordination depends on timely and accurate communication. However, the unpredictability of both data sources and communication channels undermines the reliability of classical methods rooted in Shannon theory [206], which presuppose perfect knowledge of the source and channel distributions.

In **Chapter 6**, we characterize the fundamental limits of reliable communication under distributional uncertainty, focusing on two key trade-offs: (i) data compression rate and distortion error, and (ii) transmission rate and power allocation over a noisy channel. Uncertainty in both the data source and channel noise distributions is modeled using Wasserstein balls centered at nominal distributions. These problems mark a significant departure from standard DRO formulations, as the objectives involve information-theoretic quantities, such as mutual information and entropy, rather than simple expected costs. Nonetheless, we derive tractable convex reformulations based on linear matrix inequalities (LMIs), enabling efficient computation and robustness in communication systems design.

Perception and environmental understanding in autonomous systems often rely on unsupervised learning and pattern recognition methods, such as segmentation and clustering to transform large volumes of raw, unlabeled sensor data into discrete, interpretable objects. However, classical clustering methods are often vulnerable to outliers, distribution shifts, and limited sample sizes. In **Chapter 7**, we develop a distributionally robust k-means clustering algorithm that minimizes the mean-squared distortion error for quantizing the worst-case distribution within a Wasserstein-2 ball centered at the empirical data distribution. Our formulation naturally yields a soft-clustering scheme during training, replacing rigid cluster boundaries with smoothly weighted regions, resulting in significantly improved robustness to outliers and generalization under distribution shifts.

1.3 Outline and Scope of Part III: Non-rational Control

Major advances in control theory have historically been driven by practical challenges that existing methods failed to address, for example, Kalman's state-space theory [121] and Bellman's dynamic programming [16], [17] emerged to handle stochastic disturbances during the space race, while Zames's $\mathscr{H}_\infty$ control [255] addressed model uncertainty overlooked by classical stochastic control. Today, with control systems underpinning complex, high-stakes autonomous operations in uncertain and resource-constrained environments, we face a similar shift. The growing reliance on learning-based controllers demands new approaches that can manage the unreliability of data-driven models while ensuring stability, robustness, safety, and real-time performance under computational and communication constraints.

In response to these emerging demands, several new paradigms have been proposed, including distributionally robust, risk-sensitive, and multi-criteria control. While these frameworks offer elegant theoretical formulations, translating them into practical, real-time deployable controllers remains a significant analytical and computational challenge. For instance, finite-horizon formulations, such as the Wasserstein distributionally robust controller studied in this thesis, typically result in high-dimensional optimization problems that scale poorly with the time horizon, rendering real-time implementation impractical. Receding horizon strategies, such as model predictive control (MPC), partially alleviate this by solving a sequence of short-horizon optimization problems rather than a single long-horizon one. However, these approaches often lack formal stability guarantees and can exhibit non-smooth, erratic, or myopic behavior, ultimately compromising long-term system performance.

A more principled approach lies in infinite-horizon controller synthesis, which provides provably stable policies with performance guarantees and enables efficient real-time implementation. However, designing optimal infinite-horizon controllers for emerging paradigms, often involving sophisticated, nonstandard objective functions, remains fundamentally challenging, as it typically results in infinite-dimensional optimization problems. These difficulties are further compounded by structural and information constraints, such as causality, communication delays, and requirements for distributed or decentralized implementation. To date, exact closed-form solutions have been found only for a limited set of problems such as LQG/$\mathscr{H}_2$ [121], $\mathscr{H}_\infty$ [51], and regret-optimal control [191]. In contrast, for many other practically relevant formulations, including mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control, no exact solution has yet been discovered. This raises the following fundamental question:

> *How can we systematically and efficiently synthesize stabilizing, scalable and exact optimal controllers for generalized, nonstandard performance objectives beyond the classical formulations?*

Another major challenge in this generalized setting is that optimal controllers for nonstandard performance objectives are typically infinite-dimensional (i.e., *non-rational*) which presents a fundamental barrier to practical, real-time, and scalable implementation. In contrast to classical settings such as LQG/$\mathscr{H}_2$ and $\mathscr{H}_\infty$, where optimal controllers admit finite-dimensional state-space realizations, generalized formulations often lack this property. This phenomenon arises even in well-studied problems like mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control [165]. Indeed, in Chapter 9, we show that the optimal infinite-horizon Wasserstein distributionally robust controller is itself non-rational. While approximation strategies, such as restricting to finite impulse response (FIR) controllers, can yield tractable formulations, they often fail to capture long-range dependencies or incur significant suboptimality unless one adopts an impractically large FIR length. This raises another fundamental question:

> *How can we practically implement infinite-dimensional (i.e., non-rational) optimal controllers for general objectives?*

These realities underscore the need for a new generation of practical controller-synthesis techniques that

i. accommodate a wide array of performance metrics,

ii. ensure closed-loop stability,

iii. impose minimal computational overhead during real-time implementation,

iv. scale efficiently to large-scale systems, and

v. achieve near-optimal performance with provably negligible suboptimality gaps.

In **Part III**, we address these challenges by proposing a unified framework: *non-rational control*. Embracing the infinite-dimensional nature of control problems, this unified framework offers new analytical and computational tools that render otherwise intractable controller design tasks both solvable and practically implementable. Crucially, it adopts an *optimize-then-approximate* paradigm, enabling synthesis of provably near-optimal, stabilizing finite-dimensional (rational) state-space controllers tailored to diverse objectives, even when the true optimum resides in an infinite-dimensional (non-rational) policy class.

More concretely, the central object of our study is a generalized infinite-horizon control problem posed as an infinite-dimensional optimization: find a causal and admissible controller $\mathcal{K}$ that minimizes a performance objective specified by a function f of the squared closed-loop transfer operator $\mathcal{T}_{\mathcal{K}}^{*}\mathcal{T}_{\mathcal{K}}$:

$$\inf_{\text{causal } \mathcal{K}} f(\mathcal{T}_{\mathcal{K}}^{*}\mathcal{T}_{\mathcal{K}}) \quad \text{subject to} \quad \mathcal{K} \in \mathscr{K}_{\text{admissible}} \tag{1.3}$$

Here, the closed-loop transfer operator $\mathcal{T}_{\mathcal{K}}$ maps external disturbances to regulated outputs, the objective function f encodes the desired performance criteria, and $\mathscr{K}_{\text{admissible}}$ encodes the subset of admissible controllers. This general formulation encompasses a wide class of control problems, including the classical and emerging examples discussed earlier. Our framework is built on the following key components:

1. **Infinite-dimensional convex duality.** By formulating the control objective at the operator level and invoking convex duality, the original design problem with generalized performance criteria is recast as a tractable dual optimization problem.

2. **Efficient numerical solution.** Exploiting the Fourier-domain (transfer-function) representation of the dual variable allows the use of standard, scalable optimization algorithms (e.g., first-order methods) to compute the exact infinite-dimensional optimum.

3. **Rational controller synthesis.** A novel rational-approximation scheme translates the infinite-dimensional solution into finite-dimensional controllers that are guaranteed to be stabilizing and within a quantifiable performance gap, enabling practical real-time deployment without sacrificing performance.

The non-rational control framework integrates and extends $\mathscr{H}_2$, $\mathscr{H}_\infty$, distributionally robust, risk-sensitive, regret-optimal, and multi-objective control paradigms into a cohesive framework that enables scalable real-time implementation. The underlying numerical optimization and rational controllers synthesis algorithms are highly efficient in terms of computational complexity and horizon independent, as opposed to finite-horizon formulations which scale with the time-horizon. Moreover, the resulting near-optimal rational controllers significantly outperform those derived from restrictive policy classes, such as those obtained from FIR approximation.

In **Chapter 8**, we provide a comprehensive background on optimization-based methods for controller synthesis, introduce the necessary notation, and review key concepts from linear systems theory. The chapter concludes with a summary of our main contributions and an overview of the non-rational control framework developed in the subsequent chapters.

In **Chapter 9, Chapter 10, and Chapter 11**, we present a range of infinite-horizon control and filtering problems unified under our framework, spanning both classical formulations and emerging paradigms such as distributionally robust, risk-sensitive, and multi-objective control. In particular, we derive the infinite-horizon formulations of the Wasserstein distributionally robust control and filtering problems, extending the finite-horizon versions introduced in Part II, and discuss mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control as a canonical example of the multi-objective setting.

In **Chapter 12**, we develop the duality theory for the generalized infinite-horizon control problems introduced earlier. Our main tool is the Fenchel–Legendre conjugate f^* of the objective function f, together with the strong duality framework established by Fenchel and Rockafeller. A key advantage of working with the dual formulation is that it enables a tractable reformulation of the original problem: the dual problem becomes a max–min optimization, where the inner minimization is a weighted $\mathscr{H}_2$ objective, and all structural and information constraints, such as causality and decentralization, are absorbed into this simpler inner problem. The weighting factor in this setting is precisely the dual optimization variable.

This reformulation admits a compelling interpretation as a minimax game between the controller and an adversarial disturbance process: the controller minimizes its expected infinite-horizon quadratic cost under a colored (correlated) Gaussian disturbance, whose auto-covariance kernel serves as the dual variable, while the disturbance aims to maximize this cost but is penalized via the conjugate function f^* evaluated on its auto-covariance kernel. In this sense, the dual problem can be viewed as an infinite-horizon stochastic control problem under a generalized risk measure induced by f. We conclude the chapter by applying this duality framework to the specific control problems introduced in the preceding section.

In **Chapter 13**, we present the necessary and sufficient conditions for the existence of a saddle-point solution. Our approach proceeds in two stages. First, we show how the optimal controller can be derived by reducing the problem to a weighted $\mathscr{H}_2$ optimization using the classical Wiener–Hopf technique [239], [250]. This step highlights the strength of our framework: a complex and seemingly intractable control problem is reformulated as a structured weighted $\mathscr{H}_2$ problem that is analytically and computationally tractable. However, determining the optimal controller still requires solving the dual problem to identify the optimal weighting, which is achieved by deriving the Karush-Kuhn-Tucker (KKT) optimality conditions.

We apply this methodology to the examples introduced in the preceding chapters, leading to explicit closed-form expressions for the optimal solutions in the infinite-horizon Wasserstein distributionally robust control and filtering problems, as well as for the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control problem. These solutions reveal the inherent non-rational structure of the resulting optimal controllers.

A key insight of our approach is that, although the optimal controllers may be non-rational, they can often be fully characterized by a finite-dimensional parameterization, *i.e.*, the solution lies within a structured family of non-rational transfer functions. This structure enables the development of efficient numerical algorithms that operate in the Fourier domain using first-order gradient information. In **Chapter 14**, we present adaptations of several well-known optimization methods, including proximal gradient and Frank–Wolfe algorithms, where gradients are evaluated at a finite (but sufficiently dense) set of frequency points, enough to accurately approximate the underlying infinite-dimensional objects.

While the optimal non-rational controller may be theoretically well-defined, its infinite-dimensional nature renders it impractical for real-time implementation as a finite-dimensional state-space controller. In **??**, we introduce a computationally efficient rational approximation algorithm to synthesize finite-dimensional, stabilizing, near-optimal controllers with a guaranteed suboptimality gap and prescribed state order. Rather than directly approximating the non-rational controller, which leads to a non-convex and generally NP-hard problem, our method targets the optimal dual variable, a positive-definite transfer function, and approximates it using rational functions formed as ratios of positive polynomials in the $\mathscr{H}_\infty$ norm. The positivity of both the dual variable and its approximants enables a convex reformulation of the approximation problem. Finally, by applying the Wiener–Hopf technique, we construct a finite-dimensional, stabilizing state-space controller from the rational approximation of the dual variable.

Part I

Learning and Control

Chapter 2

LEARNING TO CONTROL FULLY OBSERVED LINEAR DYNAMICAL SYSTEMS

2.1 Introduction

There has been a significant development in data-driven methods for controlling dynamical systems in recent years due to the development of novel reinforcement learning approaches and techniques [108]. Adaptive control of unknown linear dynamical systems has been the main focus due to its simplicity and its ability to capture the crux of the problem and give insights on more challenging tasks [187]. Among linear dynamical systems, Linear Quadratic Regulators (LQRs) are the canonical settings with quadratic regulatory costs to design desirable controllers and have been studied in an array of prior works [2], [4], [36], [62], [140], [159], [210]. These works provide finite-time performance guarantees of adaptive control algorithms in terms of *regret*, which is the difference between the attained cumulative cost and the expected cost of the optimal controller. In particular, they show that $\tilde{O}(\sqrt{T})$ regret after T time steps is optimal in adaptive control of LQRs. They utilize several different paradigms for algorithm design such as Certainty Equivalence, Optimism or Thompson Sampling, yet, they suffer either from the inherent algorithmic drawbacks or limited applicability in practice.

Certainty equivalent control and its challenges: Certainty equivalent control (CEC) is one of the most straightforward paradigms for control design in adaptive control of dynamical systems. In CEC, an agent obtains a nominal estimate of the system, and executes the optimal control law for this estimated system. Even though Mania, Tu, and Recht [159] and Simchowitz and Foster [210] show that this simple approach attains optimal regret in LQRs, the proposed algorithms have several drawbacks. First and foremost, CEC is sensitive to model mismatch and requires significantly small model estimation error to a point that exploration of the system dynamics is not required. Since this level of refinement is challenging to obtain for an unknown system, these methods rely on access to an initial stabilizing controller to enable a long exploration. In practice, such a priori known controllers may not be available, which hinders the deployment of these algorithms.

Optimism-based control and its challenges: Optimism is one of the most prominent methods to effectively balance exploration and exploitation in adaptive control [21]. In optimism-based control, an agent executes the optimal policy for the model with the lowest cost within a set of plausible models. In Abbasi-Yadkori and Szepesvári [2] and Faradonbeh, Tewari, and Michailidis [62], the authors use optimism-based control design to achieve $\tilde{O}(\sqrt{T})$ regret with exponential dimension dependency. Both algorithms solve a non-convex optimization problem to find the optimistic controllers, which is an NP-hard problem in general [6]. Unfortunately, this computational inefficiency severely limits their practicality. Recently, Abeille and Lazaric [5] proposed a relaxation to the optimistic controller computation, which makes the optimism-based controllers efficient. However, their algorithm also requires a significantly well-refined model estimate and a given initial stabilizing policy, similar to CEC.

Restricted LQR settings in the prior works: In our work, we study the stabilizable multi-dimensional LQR setting. Stabilizability is necessary and sufficient condition to have a well-posed LQR control problem [120]. On the contrary, prior works usually consider the controllable LQR setting, which is a subclass of stabilizable LQRs [36], [40]. While the controllability condition simplifies the learning and control problem, it is also often violated in many real-world control systems [69]. Recently, Lale, Azizzadenesheli, Hassibi, *et al.* [140] proposed an adaptive control algorithm that does not need an initial stabilizing controller and achieves optimal regret in stabilizable LQRs. However, their method relies on optimism, and unfortunately inherits the aforementioned computational complexity of optimistic methods.

Thompson Sampling and its challenges: Thompson Sampling (TS) is one of the oldest strategies to balance the exploration vs. exploitation trade-off [222]. In TS, the agent samples a model from a distribution computed based on prior control input and observation pairs, and then takes the optimal action for this sampled model and updates the distribution based on its novel observation. Since it relies solely on sampling, this approach provides polynomial-time algorithms for adaptive control. Therefore, it is a promising alternative to overcome the computational burden faced in optimismic control design. For this reason, Abeille and Lazaric [3], [4] propose adaptive control algorithms using TS. In particular, Abeille and Lazaric [4] provide the first TS-based adaptive control algorithm for LQRs that attains optimal regret of $\tilde{O}(\sqrt{T})$. However, their result *only holds for scalar* stabilizable systems, since they were able to show that TS samples optimistic parameters with constant probability in only scalar systems. Further, they conjecture that this is true

Table 2.1: Comparison with the prior works that attain $\tilde{O}(\sqrt{T})$ regret on LQR, $\dagger =$ 1-dim LQRs

Work	**Setting**	**Stabilizing Controller**	**Computation**
[4]	Stabilizable†	Not Required	Feasible
[159]	Controllable	Required	Feasible
[210]	Stabilizable	Required	Feasible
[36]	Controllable	Not required	Feasible
[140]	Stabilizable	Not required	Infeasible
This work	Stabilizable	Not required	Feasible

in multidimensional systems as well and TS-based adaptive control can provide optimal regret in multidimensional LQRs, and provide a simple numerical example to support their claims.

Contributions

In this work, we give an affirmative answer to the conjecture posed in Abeille and Lazaric [4]:

- We propose an efficient adaptive control algorithm, **T**hompson **S**ampling-based **A**daptive **C**ontrol (TSAC), that attains $\tilde{O}(\sqrt{T})$ regret in multidimensional stabilizable LQRs. This makes TSAC the first efficient adaptive control algorithm to achieve order-optimal regret in all stabilizable LQRs without the prior knowledge of a stabilizing policy (Table 2.1).
- We empirically demonstrate the performance of TSAC and compare to the optimism (heuristic) and TS-based methods that do not require initial stabilizing policy in flight control of Boeing 747 with linearized dynamics. We show that TSAC effectively explores the system to find a stabilizing policy and achieves the competitive regret performance, while being computationally feasible.

The design of TSAC and our regret guarantee hinge on three important pieces missing in prior works: Fixed policy update rule, improved exploration in early stages of adaptive control, and a novel lower bound that shows TS samples optimistic parameters with non-zero probability in multidimensional LQRs. Unlike the frequent policy update rule of Abeille and Lazaric [4] in scalar LQRs, TSAC updates its policy with fixed time periods. This policy update rule prevents fast policy changes that would cause state blow-ups in stabilizable LQRs. In the beginning of agent-environment interaction, TSAC focuses on quickly finding a stabilizing controller to avoid state blow-ups due to lack of a known initial stabilizing policy. By using isotropic exploration in the early stages along with the exploration of TS policy, we show that TSAC achieves fast stabilization.

After stabilizing the unknown system dynamics, TSAC relies on the effective exploration of the TS to find desirable controllers. In particular, we show that the TS samples optimistic parameters with a constant probability in any LQR setting. This novel lower bound shows that the TS is an efficient alternative to optimism in all adaptive control problems in LQRs. Combining this lower bound with the fixed policy update rule, we derive the optimal regret guarantee for TSAC.

2.2 Preliminaries

Notation: We denote the Euclidean norm of a vector x as $\|x\|_2$. For a matrix $A \in \mathbb{R}^{n\times d}$, we denote $\rho(A)$ as the spectral radius of A, $\|A\|_F$ as its Frobenius norm and $\|A\|$ as its spectral norm. $\mathrm{tr}(A)$ denotes its trace, $A^\intercal$ is the transpose. For any positive definite matrix V, $\|A\|_V = \|V^{1/2}A\|_F$. For matrices $A, B \in \mathbb{R}^{n\times d}$, $A \bullet B = \mathrm{tr}(AB^\intercal)$ denotes their Frobenius inner product. The j-th singular value of a rank-n matrix A is $\sigma_j(A)$, where $\sigma_{\max}(A) := \sigma_1(A) \geq \ldots \geq \sigma_{\min}(A) := \sigma_n(A)$. I represents the identity matrix with the appropriate dimensions. $\mathbb{M}_n = \mathbb{R}^{n\times n}$ denotes the set of n-dimensional square matrices. $\mathcal{N}(\mu, \Sigma)$ denotes normal distribution with mean μ and covariance Σ. $Q(\cdot)$ denotes the Gaussian Q-function. $O(\cdot)$ and $o(\cdot)$ denote the standard asymptotic notation and $f(T) = \omega(g(T))$ is equivalent to $g(T) = o(f(T))$. $\tilde{O}(\cdot)$ presents the order up to logarithmic terms.

Setting

Suppose we are given a discrete time linear time-invariant system with the following dynamics,

$$x_{t+1} = A_* x_t + B_* u_t + w_t, \tag{2.1}$$

where $x_t \in \mathbb{R}^n$ is the state of the system, $u_t \in \mathbb{R}^d$ is the control input, $w_t \in \mathbb{R}^n$ is i.i.d. process noise at time t. At each time step t, the system is at state x_t where the agent observes the state. Then, the agent applies a control input u_t and the system evolves to x_{t+1} at time $t+1$. The underlying system (2.1) can be represented as $x_{t+1} = \Theta_*^\intercal z_t + w_t$, where $\Theta_*^\intercal = [A_* \;\; B_*]$ and $z_t = [x_t^\intercal \;\; u_t^\intercal]^\intercal$. In this work, we consider stabilizable linear dynamical systems Θ_*, such that there exists a controller K where $\rho(A_* + B_* K) < 1$. More precisely, the systems with the following property:

Assumption 2.2.1 (Bounded and (κ, γ)-stabilizable System)**.** The unknown system Θ_* is a member of a set $\mathcal{S}$ such that $\mathcal{S} \subseteq \big\{\Theta' = [A', B'] \,\big|\, \Theta'$ is (κ,γ)-stabilizable, $\|\Theta'\|_F \leq$

$S\}$ for some $\kappa \geq 1$ and $0 < \gamma \leq 1$. In particular, for the underlying system Θ_*, we have $\|K(\Theta_*)\| \leq \kappa$ and there exists L and $H \succ 0$ such that $A_* + B_* K(\Theta_*) = HLH^{-1}$, with $\|L\| \leq 1 - \gamma$ and $\|H\|\|H^{-1}\| \leq \kappa$.

Note that the stabilizability condition is necessary and sufficient condition to define the optimal control problem [120] and it is weaker than the controllability assumption considered in prior works [2], [36], [40]. In particular, the set of stabilizable systems subsumes the set of controllable systems. Moreover, (κ, γ)-stabilizability is merely a quantification of stabilizability for the finite-time analysis and it is adopted in recent works [32], [39], [40]. One can show that any stabilizable system is also (κ, γ)-stabilizable for some κ and γ, conversely, (κ, γ)-stabilizability implies stabilizability (Lemma B.1 Cohen, Hassidim, Koren, *et al.* [39]). We have the following assumption on w_t.

Assumption 2.2.2 (Gaussian Process Noise)**.** There exists a filtration $\mathcal{F}_t$ such that for all $t \geq 0$, x_t, z_t are $\mathcal{F}_t$-measurable and $w_t | \mathcal{F}_t = \mathcal{N}(0, \sigma_w^2 I)$ for some known $\sigma_w > 0$.

Note that this assumption is standard in literature and adopted for simplicity of exposure. The following results can be extended to sub-Gaussian process noise setting using the techniques developed in Lale, Azizzadenesheli, Hassibi, *et al.* [140]. At each time step, the regulating cost is $c_t = x_t^\intercal Q x_t + u_t^\intercal R u_t$, where $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{d \times d}$ are known positive definite matrices such that $\|Q\|, \|R\| < \bar{\alpha}$ and $\sigma_{\min}(Q), \sigma_{\min}(R) > \underline{\alpha} > 0$. The goal is to minimize the average expected cost

$$J(\Theta_*) = \lim_{T \to \infty} \min_{u = [u_1, \ldots, u_T]} \frac{1}{T} \mathbb{E}\Big[\sum\nolimits_{t=1}^{T} x_t^\intercal Q x_t + u_t^\intercal R u_t\Big], \tag{2.2}$$

by designing control inputs based on past observations. This problem is the canonical infinite horizon linear quadratic regulator (LQR) problem. If the underlying system Θ_* is known, the solution of the optimal control problem is a linear feedback control $u_t = K(\Theta_*) x_t$ with $K(\Theta_*) = -(R + B_*^\intercal P(\Theta_*) B_*)^{-1} B_*^\intercal P(\Theta_*) A_*$, where $P(\Theta_*)$ is the unique positive definite solution to

$$P(\Theta_*) = A_*^\intercal P(\Theta_*) A_* + Q - A_*^\intercal P(\Theta_*) B_* (R + B_*^\intercal P(\Theta_*) B_*)^{-1} B_*^\intercal P(\Theta_*) A_*, \tag{2.3}$$

i.e., the discrete algebraic Riccati equation (DARE), and $J(\Theta_*) = \sigma_w^2 \operatorname{Tr}(P(\Theta_*))$. Note that since the system is stabilizable, $J(\Theta_*) < \infty$. In fact, using Assumption 2.2.1, one can show that $\|P(\Theta')\| \leq D := \bar{\alpha}\gamma^{-1}\kappa^2(1 + \kappa^2)$ for all $\Theta' \in \mathcal{S}$, including Θ_* (Lemma 2.1 of Lale, Azizzadenesheli, Hassibi, *et al.* [140]).

Finite-Time Adaptive Control Problem

In this work, we consider the adaptive control setting, where Θ_* is *unknown*. The goal in the finite-time adaptive control problem is to minimize the cumulative cost *i.e.*, $\sum_{t=0}^{T} c_t$. In order to design a controller that achieves this goal, the controlling agent needs to interact with the system to learn the Θ_* that governs the dynamics. However, due to a lack of knowledge of model dynamics, the agent takes sub-optimal actions. In this work, we use *regret*, R_T, as the metric to evaluate the finite-time performance of the controlling agent. The regret quantifies the difference between the performance of the agent and the expected performance of the optimal controller, $R_T = \sum_{t=0}^{T}(c_t - J(\Theta_*))$.

Learning the System Dynamics

For any given input and state pairs up to time t, Θ_* can be estimated using regularized least squares (RLS) for some $\mu > 0$: $\min_\Theta \sum_{s=0}^{t-1} \operatorname{tr}\left((x_{s+1} - \Theta^\intercal z_s)(x_{s+1} - \Theta^\intercal z_s)^\intercal\right) + \mu\|\Theta\|_F^2$. The solution is given as $\hat{\Theta}_t = V_t^{-1} \sum_{s=0}^{t-1} z_s x_{s+1}^\intercal$ where $V_t = \mu I + \sum_{s=0}^{t-1} z_s z_s^\intercal$. Using Theorem 1 of Abbasi-Yadkori and Szepesvári [2], for any $\delta \in (0,1)$, for all $0 \le t \le T$, the underlying parameter Θ_* lives in $\mathcal{E}_t^{\mathrm{RLS}}(\delta)$ with probability at least $1-\delta$ where $\mathcal{E}_t^{\mathrm{RLS}}(\delta) = \{\Theta : \|\Theta - \hat{\Theta}_t\|_{V_t} \le \beta_t(\delta)\}$ for $\beta_t(\delta) = \sigma_w \sqrt{2n \log((\det(V_t)^{1/2})/(\delta \det(\mu I)^{1/2}))} + \sqrt{\mu} S$.

2.3 TSAC Framework

In this section, we present TSAC, a sample efficient TS-based adaptive control algorithm for the unknown stabilizable LQRs. The algorithm is summarized in Algorithm 1. It has two phases: 1) TS with improved exploration and 2) Stabilizing TS.

TS with Improved Exploration Due to lack of a priori known stabilizing controller, TSAC focuses on rapidly learning stabilizing controllers in the early stages of the algorithm. To achieve this, TSAC explores the system dynamics effectively in this phase. At any time-step t, given the RLS estimate $\hat{\Theta}_t$ and the design matrix V_t as described in Section 2.2, TSAC samples a perturbed model parameter $\tilde{\Theta}_t = \mathcal{R}_{\mathcal{S}}(\hat{\Theta}_t + \beta_t(\delta) V_t^{-1/2} \eta_t)$, where $\mathcal{R}_{\mathcal{S}}$ denotes the rejection sampling operator associated with the set $\mathcal{S}$ given in Assumption 2.2.1 and $\eta_t \in \mathbb{R}^{(n+d)\times n}$ is a matrix with independent standard normal entries. Here $\mathcal{R}_{\mathcal{S}}$ guarantees that $\tilde{\Theta}_t \in \mathcal{S}$ and $\beta_t(\delta) V_t^{-1/2} \eta_t$ randomizes the sampled parameter coherently with the RLS estimate and the uncertainty associated with it. Using this sampled model parameter, TSAC constructs the optimal linear controller $\bar{u}_t = K(\tilde{\Theta}_t)x_t$ for $\tilde{\Theta}_t$.

However, to obtain stabilizing controllers for an unknown linear dynamical system, one needs to explore the state-space in all directions (Lemma 4.2 of [140]). Unfortunately, due to lack of reliable estimates in the early stages, deploying the policy achieved via TS, $\bar{u}_t$, may not achieve such effective exploration. Therefore, in the early stages of interactions with the underlying system, TSAC deploys isotropic perturbations along with the sampled policy. In particular, for the first T_w time-steps, TSAC uses $u_t = \bar{u}_t + \nu_t$ as the control input where $\nu_t \sim \mathcal{N}(0, 2\kappa^2\sigma_w^2 I)$. This improved exploration policy effectively excites and explores all dimensions of the system to certify the design of stabilizing controllers. TSAC sets T_w such that all the sampled controllers $K(\tilde{\Theta}_t)$ are guaranteed to stabilize the underlying system Θ_* for all $t > T_w$ (Appendix 2.B).

Unlike most of the popular RL strategies that follow lazy updates, TSAC updates its sampled policy in every fixed τ_0 steps, *i.e.*, the same sampled policy $K(\tilde{\Theta}_t)$ is deployed for τ_0 time-steps. This update rule is carefully chosen such that TSAC samples enough optimistic policies to reduce the cumulative regret and avoids too frequent policy changes which would cause state blow-ups.

Stabilizing TS After guaranteeing the design of stabilizing policies with improved exploration in the first phase, TSAC starts the adaptive control with only TS. In particular, for the remaining time-steps, TSAC deploys $u_t = K(\tilde{\Theta}_t)x_t$ for $\tilde{\Theta}_t = \mathcal{R}_{\mathcal{S}}(\hat{\Theta}_t + \beta_t(\delta)V_t^{-1/2}\eta_t)$ and updates the sampled model parameter in every τ_0 time-steps. Note that, even though all the policies during this phase are stabilizing, frequent policy changes can still cause undesirable state growth. TSAC prevents this possibility by applying the same control policy for τ_0 time-steps in this phase as well. During this phase, TSAC decays the possible state blow-ups in the first phase and maintains stable dynamics.

2.4 Theoretical Analysis

In this section, we study the theoretical guarantees of TSAC. The following states the first order-optimal frequentist regret bound for TS in multidimensional stabilizable LQRs, our main result.

Theorem 2.4.1 (Regret of TSAC)**.** *Suppose Assumptions 2.2.1 and 2.2.2 hold and set $\tau_0 = 2\gamma^{-1}\log(2\kappa\sqrt{2})$ and $T_0 = \mathsf{poly}(\log(1/\delta), \sigma_w^{-1}, n, d, \bar{\alpha}, \gamma^{-1}, \kappa)$. Then, for long enough T, TSAC achieves the regret $R_T = \widetilde{\mathrm{O}}\left((n+d)^{(n+d)}\sqrt{T\log(1/\delta)}\right)$ w.p.*

Algorithm 1 TSAC

Input: κ, γ, Q, R, σ_w^2 , $V_0 = \mu I$, $\hat{\Theta}_0 = 0$
for $i = 0, 1, \ldots$ **do**
  Estimate $\hat{\Theta}_i$ & Sample $\tilde{\Theta}_i = \mathcal{R}_{\mathcal{S}}(\hat{\Theta}_i + \beta_t V_t^{-1/2}\eta_t)$
  for $t = i\tau_0, \ldots, (i+1)\tau_0 - 1$ **do**
    if $t \leq T_w$ **then**
      Deploy $u_t = K(\tilde{\Theta}_i)x_t + \nu_t$ {TS with Improved Exploration}
    else
      Deploy $u_t = K(\tilde{\Theta}_i)x_t$ {Stabilizing TS}
    end if
  end for
end for

at least $1 - 10\delta$*, if* $T_w = \max\left(T_0,\ c_1(\sqrt{T}\log T)^{1+o(1)}\right)$ *for a constant* $c_1 > 0$. *Furthermore, if the closed loop matrix of the optimally controlled underlying system,* $A_{c,*} \coloneqq A_* + B_* K_*$*, is non-singular, w.p. at least* $1 - 10\delta$*, TSAC achieves the regret* $R_T = \widetilde{\mathrm{O}}\left(\mathsf{poly}(n,d)\sqrt{T\log(1/\delta)}\right)$ *if* $T_w = \max\left(T_0,\ c_2(\log T)^{1+o(1)}\right)$ *for a constant* $c_2 > 0$.

This makes TSAC the *first efficient* adaptive control algorithm that achieves optimal regret in adaptive control of all LQRs *without an initial stabilizing policy*. To prove this result, we follow similar approach as the existing methods in literature, and define the high probability joint event $E_t = \hat{E}_t \cap \tilde{E}_t \cap \bar{E}_t$, where $\hat{E}_t$ states that the RLS estimate $\hat{\Theta}$ concentrates around Θ_*, $\tilde{E}_t$ states that the sampled parameter $\tilde{\Theta}$ concentrates around $\hat{\Theta}$, and $\bar{E}_t$ states that the state remains bounded respectively (Appendix 2.C). Conditioned on this event, we decompose the frequentist regret as, $R_T \mathbb{1}_{E_T} \leq R_{T_w}^{\text{exp}} + R_T^{\text{RLS}} + R_T^{\text{mart}} + R_T^{\text{TS}} + R_T^{\text{gap}}$, where $R_{T_w}^{\text{exp}}$ accounts for the regret attained due to improved exploration, R_T^{RLS} represents the difference between the value function of the true next state and the predicted next state, R_T^{mart} is a martingale with bounded difference, R_T^{TS} measures the difference in optimal average expected cost between the true model Θ_* and the sampled model $\tilde{\Theta}$, and R_T^{gap} measures the regret due to policy changes. The decomposition and expressions are given in Appendix 2.E. In the analysis, we bound each term separately (Appendix 2.F). Before discussing the details of the analysis, we first consider the prior works that use TS for adaptive control of LQRs and discuss their shortcomings. Further, we highlight the challenges in adaptive control of multidimensional stabilizable LQRs using TS and present our approaches to overcome these.

Prior Work on TS-based Adaptive Control and Challenges

For the frequentist regret minimization problem given in Section 2.2, the state-of-the-art adaptive control algorithm that uses TS is Abeille and Lazaric [4]. They consider the "contractible" LQR systems, i.e. $|A_* + B_*K(\Theta_*)| < 1$, and provide $\tilde{O}(\sqrt{T})$ regret upper bound for scalar LQRs, i.e. $n = d = 1$. Notice that the set of contractible systems is a small subset of the set $\mathcal{S}$ defined in Assumption 2.2.1 and they are only equivalent for scalar systems since $\rho(A_* - B_*K(\Theta_*)) = |A_* - B_*K(\Theta_*)|$. This simplified setting allow them to reduce the regret analysis into the trade-off between $R_T^{\text{TS}} = \sum_{t=0}^{T}\{J(\tilde{\Theta}_t) - J(\Theta_*)\}$ and $R_T^{\text{gap}} = \sum_{t=0}^{T} \mathbb{E}[x_{t_1}^\intercal(P(\tilde{\Theta}_{t+1}) - P(\tilde{\Theta}_t)x_{t+1} \,|\, \mathcal{F}_t]$.

These regret terms are central in the analysis of several adaptive control algorithms. In the certainty equivalent control approaches, R_T^{TS} is bounded by the quadratic scaling of model estimation error after a significantly long exploration with a known stabilizing controller [159], [210]. In the optimism-based algorithms, R_T^{TS} is bounded by 0 by design [2], [62]. Similarly, in Bayesian regret setting, [175] assume that the underlying parameter Θ_* comes from a known prior that the expected regret is computed with respect to. This true prior yields $\mathbb{E}[R_T^{\text{TS}}] = 0$ in certain restrictive LQRs. The conventional approach in the analysis of R_T^{gap} is to have lazy policy updates, *i.e.*, $O(\log T)$ policy changes, via doubling the determinant of V_t [3], [140] or exponentially increasing epoch durations [32], [63].

On the other hand, Abeille and Lazaric [4] bound R_T^{TS} by showing that TS samples the optimistic parameters, $\tilde{\Theta}_t$ such that $J(\tilde{\Theta}_t) \leq J(\Theta_*)$, with a constant probability, which reduces the regret of non-optimistic steps. Unlike the conventional policy update approaches, the key idea in Abeille and Lazaric [4] is to update the control policy every time-steps via TS, which increases the amount of optimistic policies during the execution. They show that while this frequent update rule reduces R_T^{TS}, it only results with $R_T^{\text{gap}} = \widetilde{\mathrm{O}}(\sqrt{T})$. However, they were only able to show that this constant probability of optimistic sampling holds for scalar LQRs.

The difficulty of the analysis for the probability of optimistic parameter sampling lies in the challenging characterization of the optimistic set. Since $J(\tilde{\Theta}) = \sigma_w^2 \operatorname{tr}(P(\tilde{\Theta}))$, one needs to consider the spectrum of $P(\tilde{\Theta})$ to define optimistic models, which makes the analysis difficult. In particular, decreasing the cost along one direction may be result in an increase in other directions. However, for the scalar LQR setting considered in Abeille and Lazaric [4], $J(\tilde{\Theta}) = P(\tilde{\Theta})$ and using standard perturbation results on DARE suffices. As mentioned in Abeille and Lazaric [4], one can naively consider the surrogate set of being optimistic in all directions, *i.e.* $P(\tilde{\Theta}) \preccurlyeq P(\Theta_*)$.

Nevertheless, this would result in probability that decays linear in time and does not yield sub-linear regret. In this work, we propose new surrogate sets to derive a lower bound on the probability of having optimistic samples, and show that TS in fact samples optimistic model parameters with constant probability.

In designing TS-based adaptive control algorithms for multidimensional stabilizable LQRs, one needs to maintain bounded state. In bounding the state, Abeille and Lazaric [4] rely on the fact that the underlying system is contractive, $\|\tilde{A} + \tilde{B}K(\tilde{\Theta})\| < 1$. However, under Assumption 2.2.1, even if the optimal policy of the underlying system is chosen by the learning agent, the closed-loop system may not be contractive since for any symmetric matrix M, $\rho(M) \leq \|M\|$. Thus, to avoid dire consequences of unstable dynamics, TS-based adaptive control algorithms should focus on finite-time stabilization of the system dynamics in the early stages.

Moreover, the lack of contractive closed-loop mappings in stabilizable LQRs, prevent frequent policy changes used in Abeille and Lazaric [4]. From the definition of (κ, γ)-stabilizability (Assumption 2.2.1), for any stabilizing controller K', we have that $A_* + B_* K' = H' L H'^{-1}$, with $\|L\| < 1$ for some similarity transformation H'. Thus, even if all the policies are stabilizing, changing the policies at every time step could cause couplings of these similarity transformations and result in linear growth of state over time. Thus, TS-based adaptive control algorithms need to find the balance in rate of policy updates, so that frequent policy switches are avoided, yet, enough optimistic policies are sampled. In light of these observations, our results hinge on the following:

- Improved exploration of TSAC, which allows fast stabilization of the system dynamics,
- Fixed policy update rule of TSAC, which prevents state blow-up and reduces R_T^{gap} and R_T^{TS},
- A novel result that shows TS samples optimistic model parameters with a constant probability for multidimensional LQRs and gives a novel bound on R_T^{TS}.

Details of the analysis

The improved exploration along with TS in the early stages allows TSAC to effectively explore the state-space in all directions. The following shows that for a long enough improved exploration phase, TSAC achieves consistent model estimates and guarantees the design of stabilizing policies.

Lemma 2.4.2 (Model Estimation Error and Stabilizing Policy Design)**.** *Suppose Assumptions 2.2.1 and 2.2.2 hold. For $t \geq 200(n+d)\log\frac{12}{\delta}$ time-steps of TS with improved exploration, with probability at least $1-2\delta$, TSAC obtains model estimates such that $\|\hat{\Theta}_t - \Theta_*\|_2 \leq 7\beta_t(\delta)/(\sigma_w\sqrt{t})$. Moreover, after $T_w \geq T_0 \coloneqq \mathrm{poly}(\log(1/\delta), \sigma_w^{-1}, n, d, \bar{\alpha}, \gamma^{-1}, \kappa)$ length TS with improved exploration phase, with probability at least $1-3\delta$, TSAC samples controllers $K(\tilde{\Theta}_t)$ such that the closed-loop dynamics on Θ_* is $(\kappa\sqrt{2}, \gamma/2)$ strongly stable for all $t > T_w$, i.e. there exists L and $H \succ 0$ such that $A_* + B_* K(\tilde{\Theta}_t) = HLH^{-1}$, with $\|L\| \leq 1-\gamma/2$ and $\|H\|\|H^{-1}\| \leq \kappa\sqrt{2}$.*

The proof and the precise expression of T_w can be collected in Appendix 2.B. In the proof, we show that the inputs $u_t = K(\tilde{\Theta}_i)x_t + \nu_t$ for $\nu_t \sim \mathcal{N}(0, 2\kappa^2\sigma_w^2 I)$ guarantees persistence of excitation with high probability, *i.e.*, the smallest eigenvalue of the design matrix V_t scales linearly over time. Combining this result, with the confidence set construction given in Section 2.2, we derive the first result. Using the first result and the fact that there exists a stabilizing neighborhood around the model parameter Θ_*, such that all the optimal linear controllers of the models within this region stabilize Θ_*, we derive the final result. Due to early improved exploration, TSAC stabilizes the system dynamics after T_w samples and starts stabilizing adaptive control with only TS. Using the stabilizing controllers for fixed $\tau_0 = 2\gamma^{-1}\log(2\kappa\sqrt{2})$ time-steps, TSAC decays the state magnitude and remedy possible state blow-ups in the first phase. To study the boundedness of state, define $T_r = T_w + (n+d)\tau_0\log(n+d)$. The following shows that the state is bounded and well-controlled.

Lemma 2.4.3 (Bounded states)**.** *Suppose Assumptions 2.2.1 & 2.2.2 hold. For given T_w and T_r, TSAC controls the state such that $\|x_t\| = O((n+d)^{n+d})$ for $t \leq T_r$, with probability at least $1-3\delta$ and $\|x_t\| \leq (12\kappa^2 + 2\kappa\sqrt{2})\gamma^{-1}\sigma_w\sqrt{2n\log(n(t-T_w)/\delta)}$ for $T \geq t > T_r$, with probability at least $1-4\delta$.*

The proof is given in Appendix 2.C, but here we provide a proof sketch. To bound the state for $t \leq T_r$, we show that deploying the same policy for τ_0 time-steps in the first phase maintains a well-controlled state except $n+d$ time-steps, under the high probability event of $\hat{E}_t \cap \tilde{E}_t$. Moreover, we show that this slow policy change prevents further state blow-ups due to non-contractive system dynamics in stabilizable systems.

To bound the state for $t > T_r$, we show that, with the given choice of τ_0, all the controllers during the stabilizing TS phase halves the magnitude of the state at the end of their control period. Thus, we prove that after $(n+d)\log(n+d)$ policy updates the state is well-controlled and brought to an equilibrium as shown in Lemma 2.4.3. This result shows that the joint event $E_t = \hat{E}_t \cap \tilde{E}_t \cap \bar{E}_t$ holds with probability at least $1-4\delta$ for all $t \leq T$.

Conditioned on this event, we individually analyze the regret terms individually (Appendix 2.F). We show that with probability at least $1-\delta$, $R^{\text{exp}}_{T_w}$ yields $\widetilde{\mathrm{O}}((n+d)^{n+d}T_w)$ regret due to isotropic perturbations. R^{RLS}_T and R^{mart}_T are $\widetilde{\mathrm{O}}((n+d)^{n+d}\sqrt{T_r} + \mathsf{poly}(n,d)\sqrt{T-T_r})$ with probability at least $1-\delta$ due to standard arguments based on the event E_T. More importantly, conditioned on the event E_T, we prove that $R^{\text{gap}}_T = \widetilde{\mathrm{O}}((n+d)^{n+d}\sqrt{T_r} + \mathsf{poly}(n,d)\sqrt{T-T_r})$ with probability at least $1-2\delta$, and $R^{\text{TS}}_T = \widetilde{\mathrm{O}}(nT_w + \mathrm{poly}(n,d)\sqrt{T-T_w})$ with probability at least $1-2\delta$, whose analyses require several novel fundamental results.

To bound on R^{gap}_T, we extend the results in Abeille and Lazaric [4] to multidimensional stabilizable LQRs and incorporate the slow update rule and the early improved exploration. We show that while TSAC enjoys well-controlled state with polynomial dimension dependency on regret due to slow policy updates, it also maintains the desirable $\widetilde{\mathrm{O}}(\sqrt{T})$ regret of frequent updates with only a constant τ_0 scaling. As discussed in Section 2.4, bounding R^{TS}_T requires selecting optimistic models with constant probability, which has been an open problem in the literature for multidimensional systems. In this work, we provide a solution to this problem and show that TS indeed selects optimistic model parameters with a constant probability for multidimensional LQRs. The precise statement of this result and its proof outline are given in Section 2.5. Leveraging this result, we derive the upper bound on R^{TS}_T. Combining all these terms yields the regret upper bound of TSAC given in Theorem 2.4.1.

2.5 Proof Outline of Sampling Optimistic Models with Constant Probability

In this section, we provide the precise statement that the probability of sampling an optimistic parameter is lower bounded by a fixed constant with high probability. Then we give the proof outline with the main steps. The complete proof with the intermediate results are given in Appendix 2.D.

Theorem 2.5.1 (Optimistic probability). *Let $\mathcal{F}^{cnt}_t := \sigma(F_{t-1}, x_t)$ be the information available to the controller up to time t. Denote the optimistic set by*

$\mathcal{S}^{\mathsf{opt}} := \left\{\Theta \in \mathbb{R}^{(n+d)\times n} \mid J(\Theta) \leq J(\Theta_*)\right\}$. *If* $T_w = cn^2(\sqrt{T}\log T)^{1+o(1)}$ *for a constant* $c > 0$*, then under the event* E_T *for large enough* T*, we have that* $p_t^{\mathsf{opt}} := \mathbb{P}\left\{\tilde{\Theta}_t \in \mathcal{S}^{opt} \mid \mathcal{F}_t^{cnt}, \hat{E}_t\right\} \geq \frac{Q(1)}{1+o(1)}$ *for any* $T_r < t \leq T$*. Furthermore, if the closed-loop matrix,* $A_{c,*} = A_* + B_* K_*$*, is non-singular, then the bound above still holds when* $T_w = c(\log T)^{1+o(1)}$ *for a constant* $c > 0$.

Surrogate Set Definition

First, we define a surrogate subset $\mathcal{S}^{\text{surr}}$ to the optimistic set $\mathcal{S}^{\mathsf{opt}}$. The construction of $\mathcal{S}^{\text{surr}}$ is important as the geometry of $\mathcal{S}^{\mathsf{opt}}$ is complicated to study due to (2.3) that controls the spectrum of $P(\Theta)$.

Lemma 2.5.2 (Surrogate set)**.** *Let* $J(\Theta, K) := \operatorname{tr}((Q + K^\intercal R K)\Sigma(\Theta, K))$ *be the expected average cost of controlling a system* $\Theta \in \mathcal{S}$ *by a fixed stabilizing control policy* $K \in \mathbb{R}^{d\times n}$ *where* $\Sigma(\Theta, K) := \lim_{t\to\infty} \mathbb{E}\left[x_t x_t^\intercal\right]$ *is the covariance of the state. The following surrogate set is a subset of* $\mathcal{S}^{\mathsf{opt}}$*:*

$$\mathcal{S}^{surr} := \left\{\Theta = (A,\ B)^\intercal \in \mathbb{R}^{(n+d)\times n} \,\middle|\, J(\Theta, K(\Theta_*)) \leq J(\Theta_*, K(\Theta_*)) = J(\Theta_*)\right\} \subset \mathcal{S}^{\mathsf{opt}}. \tag{2.4}$$

Note that $\Sigma(\Theta, K)$ satisfies the Lyapunov equation $\Sigma(\Theta, K) – \Theta^\intercal H_K \Sigma(\Theta, K) H_K^\intercal \Theta = \sigma_w^2 I$, where $H_K^\intercal := [I,\ K^\intercal]$, and $\Theta^\intercal H_K = A + BK$, given that K stabilizes the system Θ. We can analytically express $\Sigma(\Theta, K)$ as a converging infinite sum $\Sigma(\Theta, K) = \sigma_w^2 \sum_{t=0}^{\infty} (A+BK)^t (A^\intercal + K^\intercal B^\intercal)^t$ [120]. Using the properties of the trace operator, one can write $J(\Theta, K(\Theta_*)) = L(\Theta^\intercal H_*)$, where $L(A_c) := \sigma_w^2 \sum_{t=0}^{\infty} \left\|A_c^t\right\|_{Q_*}^2$ for any stable matrix A_c, $Q_* := Q + K(\Theta_*)^\intercal R K(\Theta_*)$, and $H_*^\intercal := [I, K(\Theta_*)^\intercal]$. Therefore, we can lower bound the probability of being optimistic as

$$\begin{aligned} p_t^{\mathsf{opt}} &\geq \mathbb{P}\left\{\tilde{\Theta}_t \in \mathcal{S}^{\text{surr}} \mid \mathcal{F}_t^{\text{cnt}}, \hat{E}_t\right\} = \mathbb{P}\left\{L(\tilde{\Theta}_t^\intercal H_*) \leq L(\Theta_*^\intercal H_*) \mid \mathcal{F}_t^{\text{cnt}}, \hat{E}_t\right\} \\ &\geq \min_{\hat{\Theta} \in \mathcal{E}_t^{\text{RLS}}} \mathbb{P}_t\{L(\hat{\Theta}^\intercal H_* + \eta^\intercal \beta_t V_t^{-\frac{1}{2}} H_*) \leq L(\Theta_*^\intercal H_*)\} \qquad (2.5) \\ &= \min_{\hat{\Theta} \in \mathcal{E}_t^{\text{RLS}}} \mathbb{P}_t\{L(\hat{\Theta}^\intercal H_* + \Xi\sqrt{F_t}) \leq L(\Theta_*^\intercal H_*)\} \qquad (2.6) \end{aligned}$$

where $\mathbb{P}_t\{\cdot\} := \mathbb{P}\{\cdot \mid \mathcal{F}_t^{\text{cnt}}\}$, $F_t := \beta_t^2 H_*^\intercal V_t^{-1} H_*$ and Ξ is a matrix of size $n \times n$ with iid $\mathcal{N}(0, 1)$ entries. Here (2.5) considers the worst possible estimate within $\mathcal{E}_t^{\text{RLS}}$ and (2.6) is the whitening transformation.

Reformulation in Terms of Closed-Loop Matrix

In the second step, we reformulate the probability of sampling optimistic parameters in terms of closed-loop system matrix $\tilde{A}_c := \tilde{\Theta}^\intercal H_* = \tilde{A} + \tilde{B} K(\Theta_*)$ of the sampled system $\tilde{\Theta} = (\tilde{A}, \tilde{B})^\intercal$ driven by the policy $K(\Theta_*)$. Transitioning to the closed-loop formulation allows tighter bounds on the optimistic probability. To complete this reformulation, we need to construct an estimation confidence set for the closed-loop system matrix $\hat{A}_c \coloneqq \hat{\Theta}^\intercal H_* = \hat{A} + \hat{B}K(\Theta_*)$ of the RLS-estimated system $\hat{\Theta} = (\hat{A}, \hat{B})^\intercal$ and show that the constructed confidence set is a super set to $\mathcal{E}_t^{\mathrm{RLS}}$.

Lemma 2.5.3 (Closed-loop confidence)**.** *Let* $F_t(\delta) \coloneqq \beta_t^2(\delta) H_*^\intercal V_t^{-1} H_*$. *For any* $t \geq 0$, *define by*

$$\mathcal{E}_t^{cl}(\delta) \coloneqq \left\{ \hat{\Theta} \in \mathbb{R}^{(n+d)\times n} \;\middle|\; \mathrm{tr}\left[(\hat{\Theta}^\intercal H_* - \Theta_*^\intercal H_*) F_t^{-1}(\delta) (\hat{\Theta}^\intercal H_* - \Theta_*^\intercal H_*)^\intercal \right] \leq 1 \right\}. \tag{2.7}$$

the closed-loop confidence set. Then, for all times $t \geq 0$ *and* $\delta \in (0,1)$, *we have that* $\mathcal{E}_t^{RLS}(\delta) \subseteq \mathcal{E}_t^{cl}(\delta)$.

Note that the definition of $\mathcal{E}_t^{\mathrm{cl}}(\delta)$ *only* involves closed-loop matrices $\hat{A}_c \coloneqq \hat{\Theta}^\intercal H_*$ and $A_{c,*} \coloneqq \Theta_*^\intercal H_*$. We can use the result of Lemma 2.5.3 to reformulate the probability of sampling optimistic parameters, $\tilde{\Theta} = (\tilde{A}, \tilde{B})$, as sampling optimistic closed-loop system matrices, $\tilde{A}_c$. We bound p_t^{opt} from below as

$$\begin{aligned}
p_t^{\mathsf{opt}} &\geq \min_{\hat{\Theta} \in \mathcal{E}_t^{\mathrm{cl}}} \mathbb{P}_t\{ L(\hat{\Theta}^\intercal H_* + \Xi\sqrt{F_t}) \leq L(A_{c,*}) \} && (2.8)\\
&= \min_{\hat{A}_c \,:\, \|\hat{A}_c^\intercal - A_{c,*}^\intercal\|_{F_t^{-1}} \leq 1} \mathbb{P}_t\{ L(\hat{A}_c + \Xi\sqrt{F_t}) \leq L(A_{c,*}) \} && (2.9)\\
&= \min_{\hat{\Upsilon} \,:\, \|\hat{\Upsilon}\|_F \leq 1} \mathbb{P}_t\{ L(A_{c,*} + \hat{\Upsilon}\sqrt{F_t} + \Xi\sqrt{F_t}) \leq L(A_{c,*}) \}, && (2.10)
\end{aligned}$$

where (2.8) is due to Lemma 2.5.3 and (2.9) follows from the fact that H_* has full column rank. Observe that, in equation (2.10), $\hat{\Upsilon}$ is a unit Frobenius norm matrix of size $n \times n$ and the term $A_{c,*} + \hat{\Upsilon}\sqrt{F_t}$ accounts for the confidence ellipsoid for the estimated closed-loop matrix, $\hat{A}_c$. The event in (2.10) corresponds to finding the closed-loop matrix, $A_{c,*} + (\Xi + \hat{\Upsilon})\sqrt{F_t}$ of the TS sampled system in the sublevel manifold $\mathcal{M}_* \coloneqq \left\{ A_c \in \mathbb{M}_n \mid L(A_c) \leq L(A_{c,*}) \right\}$ as illustrated in Figure 2.1.

Local Geometry of Optimistic Set under Perturbations

Next, we further simplify the form of the probability in (2.10) by exploiting the local geometric structure of the function $L : A_c \mapsto \sigma_w^2 \sum_{t=0}^{\infty} \left\| A_c^t \right\|_{Q_*}^2$ defined over the set of (Schur-)stable matrices, $\mathcal{M}_{\mathsf{Schur}} \coloneqq \{ A_c \in \mathbb{M}_n \mid \rho(A_c) < 1 \}$. The following lemma characterizes perturbative properties of L.

Lemma 2.5.4 (Perturbations). *The function $L : \mathcal{M}_{\mathit{Schur}} \to \mathbb{R}_+$ defined as $L(A_c) = \sigma_w^2 \sum_{t=0}^{\infty} \left\|A_c^t\right\|_{Q_*}^2$ is smooth in its domain. For any $A_c \in \mathcal{M}_{\mathit{Schur}}$, there exists $\epsilon > 0$ such that for any perturbation $\|G\|_F \leq \epsilon$, the function L admits a quadratic Taylor expansion as*

$$L(A_c + G) = L(A_c) + \nabla L(A_c) \bullet G + \frac{1}{2} G \bullet \mathcal{H}_{A_c + sG}(G) \tag{2.11}$$

for an $s \in [0, 1]$ where $\mathcal{H}_{A_c} : \mathbb{M}_n \to \mathbb{M}_n$ is the Hessian operator evaluated at a point $A_c \in \mathcal{M}_{\mathit{Schur}}$. In particular, we have that $\nabla L(A_{c_*}) = 2P(\Theta_*)A_{c,*}\Sigma_*$. Furthermore, there exists a constant $r > 0$ such that $|G \bullet \mathcal{H}_{A_c+sG}(G)| \leq r\|G\|_F^2$ for any $s \in [0, 1]$ and $\|G\|_F \leq \epsilon$.

Lemma 2.5.4 guarantees that if a perturbation is sufficiently small, the perturbed function can be locally expressed as a quadratic function of the perturbation. Since the set of stable matrices, $\mathcal{M}_{\mathrm{Schur}}$, is globally non-convex and Taylor's theorem only holds in convex domains, we restrict the perturbations in a ball of radius $\epsilon > 0$. The fact that there is a neighborhood of stable matrices around a matrix A_c enables us to apply Taylor's theorem in this neighborhood.

Given the optimal closed-loop system matrix $A_{c,*}$, let $\epsilon_* > 0$ be chosen such that the expansion in (2.11) holds for perturbations $\|G\|_F \leq \epsilon_*$ around $A_{c,*}$. Denote the perturbation due to Thompson sampling and estimation error as $G_t = (\Xi + \hat{\Upsilon})\sqrt{F_t}$ and let $\|G_t\|_F \leq \epsilon_*$. Then, we can write

$$\begin{aligned} L(A_{c,*} + G_t) &= L(A_{c,*}) + \nabla L(A_{c,*}) \bullet G_t + \frac{1}{2} G_t \bullet \mathcal{H}_{A_{c,*}+sG_t}(G_t) \\ &\leq L(A_{c,*}) + \nabla L(A_{c,*}) \bullet G_t + \frac{r_*}{2}\|G_t\|_F^2 \end{aligned} \tag{2.12}$$

where $r_* > 0$ is a constant due to Lemma 2.5.4. Using (2.12), we have the following lower bound on (2.10),

$$p_t^{\mathsf{opt}} \geq \min_{\hat{\Upsilon} : \|\hat{\Upsilon}\|_F \leq 1} \mathbb{P}_t \left\{ \frac{r_*}{2} \|(\Xi + \hat{\Upsilon}) F_t^{\frac{1}{2}}\|_F^2 + \nabla L_* \bullet (\Xi + \hat{\Upsilon}) F_t^{\frac{1}{2}} \leq 0, \text{ and } \|(\Xi + \hat{\Upsilon}) F_t^{\frac{1}{2}}\|_F \leq \epsilon_* \right\}, \tag{2.13}$$

where $\nabla L_* \coloneqq \nabla L(A_{c,*})$. The event in (2.13) corresponds to finding $A_{c,*} + (\Xi + \hat{\Upsilon})\sqrt{F_t}$ at the intersection of the stable ball $\mathcal{B}_* \coloneqq \left\{A_c \in \mathbb{M}_n \mid \|A_c - A_{c,*}\|_F \leq \epsilon_*\right\}$ and the sublevel manifold $\mathcal{M}_*^{\mathrm{qd}} \coloneqq \left\{A_c \in \mathbb{M}_n \mid \|A_c - A_{c,*} + r_*^{-1}\nabla L_*\|_F \leq \|r_*^{-1}\nabla L_*\|_F\right\}$ as illustrated in Figure 2.1.

The intersection $\mathcal{M}_*^{\mathrm{qd}} \cap \mathcal{B}_* \subset \mathcal{M}_*$ serves as another surrogate to sublevel manifold $\mathcal{M}_*$. Switching to the new surrogate $\mathcal{M}_*^{\mathrm{qd}}$ helps us overcome the issue of working with intractable and complicated geometry of $\mathcal{M}_*$ due to infinite sum in $L(A_{\mathrm{c}})$. We can utilize techniques relating to Gaussian probabilities as the geometry of $\mathcal{M}_*^{\mathrm{qd}}$ is described by a quadratic form.

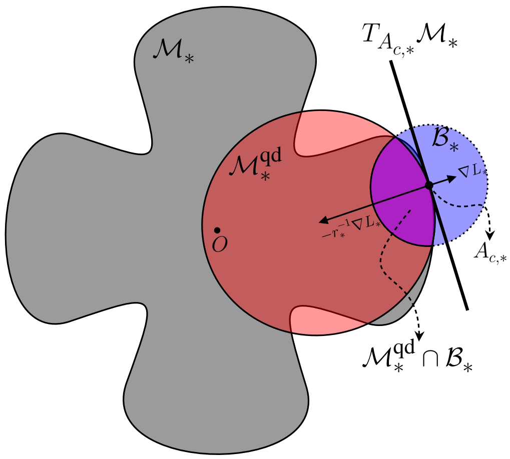


Figure 2.1: A visual representation of sublevel manifold $\mathcal{M}_*$. O is the origin and $A_{c,*}$ is the optimal closed-loop system matrix. $T_{A_{c,*}}\mathcal{M}_*$ is the tangent space to the manifold $\mathcal{M}_*$ at the point $A_{c,*}$ and ∇L_* is the Jacobian of the function L at $A_{c,*}$. $\mathcal{M}_*^{\mathrm{qd}}$ is the sublevel manifold of the quadratic approximation to L and $\mathcal{B}_*$ is a small ball of stable matrices around $A_{c,*}$. The intersection $\mathcal{M}_*^{\mathrm{qd}} \cap \mathcal{B}_*$ is a subset of $\mathcal{M}_*$.

Final Bound

Equipped with the preceding results, we can bound the optimism probability tractably from below by the probability of a TS sampled closed-loop system matrix lying inside the intersection of two balls $\mathcal{M}_*^{\mathrm{qd}} \cap \mathcal{B}_*$ as given in (2.13). By bounding the weighted Frobenius norms in (2.13) from above by $\lambda_{\max,t}$, the maximum eigenvalue of F_t, and normalizing the matrix $\nabla L_* \sqrt{F_t}$, we can write

$$\begin{aligned} p_t^{\mathsf{opt}} &\geq \min_{\|\hat{\Upsilon}\|_F \leq 1} \mathbb{P}_t \left\{ \frac{r_*}{2} \lambda_{\max,t} \|\Xi + \hat{\Upsilon}\|_F^2 + (\nabla L_* \sqrt{F_t}) \bullet (\Xi + \hat{\Upsilon}) \leq 0, \text{ and } \lambda_{\max,t} \|\Xi + \hat{\Upsilon}\|_F^2 \leq \epsilon_*^2 \right\} \\ &= \min_{\|\hat{\Upsilon}\|_F \leq 1} \mathbb{P}_t \left\{ \frac{(\nabla L_* F_t^{1/2}) \bullet (\Xi + \hat{\Upsilon})}{\|\nabla L_* F_t^{1/2}\|_F} \leq \frac{-\lambda_{\max,t} r_* \|\Xi + \hat{\Upsilon}\|_F^2}{2\|\nabla L_* F_t^{1/2}\|_F}, \text{ and } \|\Xi + \hat{\Upsilon}\|_F^2 \leq \frac{\epsilon_*^2}{\lambda_{\max,t}} \right\}. \end{aligned} \tag{2.14}$$

Observe that the inner product $(\nabla L_* F_t^{1/2}) \bullet \hat{\Upsilon}$ is maximized by $\Upsilon_\# := \frac{(\nabla L_* F_t^{1/2})}{\|\nabla L_* F_t^{1/2}\|_F}$ subject to $\|\hat{\Upsilon}\|_F \leq 1$. Since the probability distribution of $\|\Xi + \hat{\Upsilon}\|_F^2$ is invariant under orthogonal transformation of Ξ and $\hat{\Upsilon}$, (2.14) also attains its minimum at $\Upsilon_\#$.

Thus, we can rewrite (2.14) as

$$\begin{aligned} p_t^{\mathsf{opt}} &\geq \mathbb{P}_t\left\{\frac{(\nabla L_* F_t^{1/2})\bullet\Xi}{\|\nabla L_* F_t^{1/2}\|_F}+1\leq\frac{-\lambda_{\max,t}r_*}{2\|\nabla L_* F_t^{1/2}\|_F}\left\|\Xi+\Upsilon_\#\right\|_F^2,\ \text{and}\ \left\|\Xi+\Upsilon_\#\right\|_F^2\leq\frac{\epsilon_*^2}{\lambda_{\max,t}}\right\} \\ &= \mathbb{P}_t\left\{\xi+1\leq-\frac{\lambda_{\max,t}r_*}{2\|\nabla L_* F_t^{1/2}\|_F}\left((\xi+1)^2+X\right),\ \text{and}\ (\xi+1)^2+X\leq\frac{\epsilon_*^2}{\lambda_{\max,t}}\right\}, \end{aligned} \tag{2.15}$$

where $\xi \sim \mathcal{N}(0,1)$ and $X \sim \chi^2_{n^2-1}$ are independent standard normal and chi-squared distributions, and (2.15) is derived by rotating Ξ so that its first element is along the direction of $\nabla L_* F_t^{1/2}$. We use the following lemma to characterize the eigenvalues of F_t and control the lower bound (2.15) on p_t^{opt}.

Lemma 2.5.5 (Bounded eigenvalues)**.** *Suppose $T_w = O((\sqrt{T})^{1+o(1)})$. Denote the minimum and maximum eigenvalues of F_t by $\lambda_{\min,t}$ and $\lambda_{\max,t}$, respectively. Under the event E_T, for large enough T, we have that $\lambda_{\max,t} \leq C\frac{\log T}{T_w}$ and $\frac{\lambda_{\max,t}}{\lambda_{\min,t}} \leq C\frac{T\log T}{T_w}$ for any $T_r < t \leq T$ for a constant $C = \mathsf{poly}(n, d, \log(1/\delta))$.*

Lemma 2.5.5 states that maximum eigenvalue and the condition number of F_t are controlled inversely by the length of initial exploration phase T_w and proportionally by $\log T$ and $T\log T$ given that exploration time is bounded by a certain amount. The length of initial exploration T_w relative to the horizon T is critical in guaranteeing asymptotically constant optimistic probability p_t^{opt}. Although more lengthy initial exploration will lead to better convergence to constant optimistic probability, it also incurs higher asymptotic regret due to linear scaling of exploration regret with T_w.

Using the relation $\|\nabla L_* F_t^{\frac{1}{2}}\|_F \geq \max(\sigma_{\min,*}\|F_t^{\frac{1}{2}}\|_F,\ \lambda_{\min,t}^{\frac{1}{2}}\|\nabla L_*\|_F)$ where $\sigma_{\min,*}$ is the minimum singular value of ∇L_*, we can further bound (2.15) from below. From Lemma 2.5.4, we can write $\nabla L_* = 2P(\Theta_*)A_{c,*}\Sigma_*$ where $P(\Theta_*) \succ 0$ is the solution to the DARE in (2.3) and $\Sigma_* = \Sigma(\Theta_*, K_*) \succ 0$ is the stationary state covariance matrix. Notice that the minimum singular value of ∇L_* is positive (*i.e.*∇L_* is full-rank) if and only if the closed-loop system matrix, $A_{c,*}$, is non-singular.

In general, $A_{c,*}$ can be singular. Assuming that $T_w = O((\sqrt{T})^{1+o(1)})$, under the event E_T, we can use $\|\nabla L_* F_t^{\frac{1}{2}}\|_F \geq \sqrt{\lambda_{\min,t}}\|\nabla L_*\|_F$ to obtain the following lower bound on p_t^{opt} for $T_r < t \leq T$:

$$p_t^{\mathsf{opt}} \geq \mathbb{P}_t \left\{ \xi + 1 \leq -\frac{\sqrt{\lambda_{\max,t}}}{2\rho_*} \sqrt{\frac{\lambda_{\max,t}}{\lambda_{\min,t}}} \left((\xi+1)^2 + X \right), \text{ and } (\xi+1)^2 + X \leq \frac{\epsilon_*^2}{\lambda_{\max,t}} \right\},$$

$$\geq \mathbb{P} \left\{ \xi + 1 \leq -\frac{C}{2\rho_*} \frac{\sqrt{T}\log T}{T_w} \left((\xi+1)^2 + X \right), \text{ and } (\xi+1)^2 + X \leq \frac{\epsilon_*^2 T_w}{C \log T} \right\},$$

where $\rho_* := \|r_*^{-1} \nabla L_*\|_F$. Choosing the exploration time as $T_w = \omega(\sqrt{T}\log T)$ makes the coefficients $\frac{\sqrt{T}\log T}{T_w} = o(1)$ to be very small and $\frac{T_w}{\log T}$ to be very large, leading to constant lower bound on limiting optimistic probability $\liminf_{T\to\infty} p_T^{\mathsf{opt}} \geq \mathbb{P}\{\xi + 1 \leq 0\} =: Q(1)$.

On the other hand, if $A_{c,*}$ is non-singular, then we can use the alternative bound $\|\nabla L_* \sqrt{F_t}\|_F \geq \sigma_{\min,*} \|\sqrt{F_t}\|_F \geq \sigma_{\min,*}\sqrt{\lambda_{\max,t}}$ to obtain the following lower bound for $T_r < t \leq T$:

$$p_t^{\mathsf{opt}} \geq \mathbb{P}_t \left\{ \xi + 1 \leq -\frac{\sqrt{\lambda_{\max,t}}}{2\sigma_{\min,*}} \left((\xi+1)^2 + X \right), \text{ and } (\xi+1)^2 + X \leq \frac{\epsilon_*^2}{\lambda_{\max,t}} \right\},$$

$$\geq \mathbb{P} \left\{ \xi + 1 \leq -\frac{\sqrt{C}}{2\sigma_{\min,*}} \sqrt{\frac{\log T}{T_w}} \left((\xi+1)^2 + X \right), \text{ and } (\xi+1)^2 + X \leq \frac{\epsilon_*^2 T_w}{C \log T} \right\}.$$

Similarly, choosing the exploration time as $T_w = \omega(\log T)$ makes the coefficients $\sqrt{\frac{\log T}{T_w}} = o(1)$ to be very small and $\frac{T_w}{\log T} = \omega(1)$ to be very large, leading to constant lower bound on limiting optimistic probability $\liminf_{T\to\infty} p_T^{\mathsf{opt}} \geq Q(1)$.

In both cases, the optimistic probability achieves a constant lower bound for large enough T as $p_T^{\mathsf{opt}} \geq Q(1)(1 + o(1))^{-1}$. This result can be interpreted in a geometric way as follows. As the time passes, the estimates of the system become more accurate in the sense that the confidence region of the estimate shrinks very quickly as controlled by the eigenvalues of F_t. Similarly, the high-probability region of TS samples also shrink very fast controlled by the covariance matrix F_t. Therefore, for large enough T, the confidence region of the model estimate and the high-probability region of TS samples get significantly smaller compared to the surrogate optimistic set $\mathcal{M}_*^{\text{qd}} \cap \mathcal{B}_*$. This size difference effectively reduces the probability of finding a sampled system in $\mathcal{M}_*^{\text{qd}} \cap \mathcal{B}_*$ to the probability of finding a sampled system in the half-space separated by the tangent space $T_{A_{c,*}} \mathcal{M}_*$.

2.6 Numerical Experiments

Finally, we evaluate the performance of TSAC in longitudinal flight control of Boeing 747 with linearized dynamics [113]. We compare TSAC with three adaptive control

Table 2.2: Regret and Maximum State Norm After 200 Time Steps in Boeing 747 Flight Control

Algorithm	Average Regret	Top 95%	Top 90%	Average $\max \|x\|_2$	Top 95%	Top 90%
TSAC	$\mathbf{4.58 \times 10^7}$	$\mathbf{1.43 \times 10^5}$	$\mathbf{9.49 \times 10^4}$	$\mathbf{1.23 \times 10^3}$	$\mathbf{1.07 \times 10^2}$	$\mathbf{9.77 \times 10^1}$
StabL	1.34×10^4	1.05×10^3	9.60×10^3	3.38×10^1	3.14×10^1	2.98×10^1
OFULQ	1.47×10^8	4.19×10^6	9.89×10^5	1.62×10^3	5.21×10^2	2.78×10^2
TS-LQR	5.63×10^{11}	3.07×10^7	5.33×10^6	6.26×10^4	1.08×10^3	6.39×10^2

algorithms in literature that do not require an initial stabilizing policy:

(i) OFULQ of Abbasi-Yadkori and Szepesvári [2];

(ii) TS-LQR of Abeille and Lazaric [4];

(iii) StabL of Lale, Azizzadenesheli, Hassibi, *et al.* [140].

We perform 200 independent runs for 200 time-steps for each algorithm and report their average, top 95% and top 90% regret and maximum state norm performances. Note that, since optimistic control design is computationally intractable, we use projected gradient descent to *heuristically* find optimistic models in OFULQ and StabL. For fair comparison, we also adopt slow policy updates in OFULQ and TS-LQR and report the best results of each algorithm. Further details are in Appendix 2.H. The results are presented in Table 2.2. Notice that TSAC achieves the second best performance after StabL. As expected, StabL outperforms TSAC since it performs much heavier computations to find the optimistic controller in the confidence set, whereas TSAC samples optimistic parameters only with some fixed probability. However, TSAC compares favorably against both OFULQ and TS-LQR, making it the best performing computationally efficient algorithm.

2.7 Conclusion and Future Directions

We present the first efficient adaptive control algorithm, TSAC, that attains optimal regret of $\tilde{O}(\sqrt{T})$ in stabilizable LQRs without an initial stabilizing policy. We design TSAC to quickly stabilize the system and avoid state blow-ups via careful policy updates. Building on these design choices, the main technical contribution of this work is to show that TS samples optimistic parameters with constant probability in all LQRs, thereby resolving the conjecture in Abeille and Lazaric [4].

This result highlights that a simple sampling strategy provides effective exploration to recover low-cost achieving controllers in adaptive control of LQRs which yields order optimal regret. An important future direction is to investigate whether TS achieves optimal regret in partially observable LTI systems, e.g. [136], [139]. Moreover, to obtain constant probability of sampling optimistic parameters for general LQRs, TSAC requires $T_w = \omega(\sqrt{T}\log T)$ time-steps of improved exploration (Theorem 2.5.1), which causes the regret to be dominated by this phase. This long exploration is avoided in LQRs with non-singular optimal closed-loop matrix, which results in regret that scales polynominally in system dimensions (Theorem 2.4.1). It remains an open problem whether this polynomial dimension dependency in regret can be achieved via TS in general LQRs.

2.A Organization and Notations

In Appendix 2.A, we provide the notation tables for the paper. In Appendix 2.B, we provide the system identification and stabilization guarantees of TSAC. In particular, we give the proof of Lemma 2.4.2 and give the precise duration of the TS with improved exploration phase T_w. In Appendix 2.C, we show that under the joint event of $\hat{E}_t \cap \tilde{E}_t$ the state stays bounded as described in Lemma 2.4.3 with high probability. In Appendix 2.E, we provide the precise regret decomposition and discuss the individual terms in the regret upper bound. In Appendix 2.D, we provide the complete proof of Theorem 2.5.1, as well as the intermediate results discussed in the main text. Appendix 2.F comprises the analysis of individual terms in the regret decomposition. In particular, Appendix 2.F studies $R^{\text{exp}}_{T_w}$, Appendix 2.F studies R^{RLS}_T, Appendix 2.F studies R^{mart}_T, Appendix 2.F considers R^{TS}_T, Appendix 2.F bounds R^{gap}_T, and finally we combine these results to prove the regret upper bound of TSAC in Appendix 2.F. In Appendix 2.G, we give the technical theorems and lemmas used in the proofs. Finally, in Appendix 2.H, we give the implementation details of all algorithms. Before proceeding the next section, we define the following high probability events which are standard in TS-based algorithms. First recall the RLS confidence ellipsoid given in Section 2.2:

$$\mathcal{E}^{\text{RLS}}_t(\delta) = \{\Theta : \|\Theta - \hat{\Theta}_t\|_{V_t} \leq \beta_t(\delta)\},$$

for $\beta_t(\delta) = \sigma_w\sqrt{2n\log((T\det(V_t)^{1/2})/(\delta\det(\mu I)^{1/2}))} + \sqrt{\mu}S$. Further define

$$\mathcal{E}^{\text{TS}}_t(\delta) = \{\Theta : \|\Theta - \hat{\Theta}_t\|_{V_t} \leq \upsilon_t(\delta)\},$$

for $\upsilon_t(\delta) = \beta_t(\delta)n\sqrt{(n+d)\log(n(n+d)/\delta)}$. Define the events

$$\hat{E}_t = \{\forall s \leq t, \Theta_* \in \mathcal{E}^{\text{RLS}}_t(\delta)\} \tag{2.16}$$

$$\tilde{E}_t = \{\forall s \leq t, \tilde{\Theta}_s \in \mathcal{E}^{\text{TS}}_t(\delta)\}. \tag{2.17}$$

As described in Section 2.4, $\hat{E}_t$ defines the event that RLS estimates $\hat{\Theta}_t$ concentrate around Θ_* and $\tilde{E}_t$ defines the event that the sampled model parameter concentrates around $\hat{\Theta}_t$. From standard Gaussian tail bound and the self-normalized estimation error, we have that $\hat{E} \cap \tilde{E}$ for all $t \leq T$, with probability at least $1 - 2\delta$. Here the time dependency dropped since $\hat{E} := \hat{E}_T \subset \ldots \subset \hat{E}_1$ and $\tilde{E} := \tilde{E}_T \subset \ldots \subset \tilde{E}_1$. These events will be key in providing all the technical results starting from stabilization guarantees to final regret upper bound.

This section contains two tables which list the notations used throughout the paper for improving readability. In particular, Table 2.3 provides the system dependent notations and the useful notations for presenting the design of TSAC. In Table 2.4, we present the notation used in deriving theoretical results, namely, the regret analysis and the lower bound on the probability of selecting optimistic parameters. Further details are also referenced to the related parts of the paper.

Table 2.3: Useful Notations for the Design of TSAC

System Not.	**Definition**
Θ_*	Unknown discrete-time LTI system with dynamics of (2.1); $[A_* \;\; B_*]^\intercal$
x_t	State of the system $\in \mathbb{R}^n$
u_t	Input to the system $\in \mathbb{R}^d$
w_t	Process noise as defined in Assumption 2.2.2; $\mathcal{N}(0, \sigma_w^2 I)$
z_t	Stack of current state and input; $[x_t^\intercal \;\; u_t^\intercal]^\intercal$
Q, R	Known cost matrices; $\|Q\|, \|R\| < \bar{\alpha}$ and $\sigma_{\min}(Q), \sigma_{\min}(R) > \underline{\alpha} > 0$
c_t	Quadratic cost at time t; $x_t^\intercal Q x_t + u_t^\intercal R u_t$
$\mathcal{S}$	Set of (κ, γ)-stabilizable and bounded systems that Θ_* belongs (Assumption 2.2.1)
$P(\Theta)$	Unique p.d. solution to DARE (2.3) for a stabilizable system $\Theta = [A \;\; B]^\intercal$
$K(\Theta)$	Optimal controller for Θ; $-(R + B^\intercal P(\Theta) B)^{-1} B^\intercal P(\Theta) A$
$J(\Theta)$	Average expected cost of system Θ; $\sigma_w^2 \operatorname{Tr}(P(\Theta))$
κ	Bound over all possible optimal controllers in $\mathcal{S}$; $\sup_{\Theta \in \mathcal{S}} K(\Theta)$
D	Bound over all possible solutions to (2.3) in $\mathcal{S}$; $\bar{\alpha}\gamma^{-1}\kappa^2(1+\kappa^2)$
TSAC Not.	
$\hat{\Theta}$	Least squares estimate of Θ_* using the history of inputs and states; $[\hat{A} \;\; \hat{B}]^\intercal$
μ	Regularizer for least squares; set to $(1+\kappa^2)X_s^2$
V_t	Regularized design matrix; $\mu I + \sum_{s=0}^{t-1} z_s z_s^\intercal$
η_t	Random matrix with iid standard normal entries used for sampling systems
$\mathcal{R}_{\mathcal{S}}(\cdot)$	Rejection sampling to make sure that sampled system belongs to $\mathcal{S}$
$\tilde{\Theta}$	System obtained via TS; $\mathcal{R}_{\mathcal{S}}(\hat{\Theta}_t + \beta_t(\delta) V_t^{-1/2} \eta_t)$
ν_t	Improved exploration; $u_t = K(\tilde{\Theta}_t) x_t + \nu_t$ for $\nu_t \sim \mathcal{N}(0, 2\kappa^2 \sigma_w^2 I)$
Quantities	
δ	Fixed probability to define high probability events; $(0,1)$
T	Time horizon
T_w	Duration of TS with improved exploration; defined in Theorem 2.4.1
X_s	Upper bound on state after stabilization; $\|x_t\| \leq X_s$ for $t > T_r$ w.h.p.
S	Upper bound on the Frobenius norm of Θ_*
$\beta_t(\delta)$	Size of the RLS confidence ellipsoid at time t; $\sigma_w \sqrt{2n \log\left(\frac{\det(V_t)^{1/2}}{\delta \det(\mu I)^{1/2}}\right)} + \sqrt{\mu} S$
$\upsilon_t(\delta)$	Size of the sampling ellipsoid at time t; $\beta_t(\delta) n \sqrt{(n+d)\log(n(n+d)/\delta)}$
τ_0	Fixed duration for each sampled policy; $2\gamma^{-1}\log(2\kappa\sqrt{2})$
T_0	Number of samples required to identify a stabilizing controller; (2.19)
T_r	Time required to control the state w.h.p.; $T_w + (n+d)\tau_0 \log(n+d)$

Table 2.4: Useful Notations for the Analysis

Regret Analy.	
R_T	Regret of TSAC at until time T; $R_T = \sum_{t=0}^{T}(c_t - J(\Theta_*))$
$\mathcal{F}_t$	Filtration such that for all $t \geq 0$, x_t, z_t are $\mathcal{F}_t$-measurable
$\mathcal{F}_t^{\text{cnt}}$	Information available to the controller up to time t; $\sigma(F_{t-1}, x_t)$
R_T^{exp}	Regret attained due to improved exploration (Appendix 2.E)
R_T^{RLS}	Cost-to-go difference of the true and predicted next states (Appendix 2.E)
R_T^{mart}	Martingale with bounded difference (Appendix 2.E)
R_T^{TS}	Difference in $J(\Theta_*)$ and $J(\tilde{\Theta})$ (Appendix 2.E)
R_T^{gap}	Regret due to policy changes (Appendix 2.E)
$\mathcal{E}_t^{\text{RLS}}(\delta)$	Regularized least squares confidence ellipsoid; $\{\Theta : \|\Theta - \hat{\Theta}_t\|_{V_t} \leq \beta_t(\delta)\}$
$\mathcal{E}_t^{\text{TS}}(\delta)$	Confidence ellipsoid for sampled system; $\{\Theta : \|\Theta - \hat{\Theta}_t\|_{V_t} \leq \upsilon_t(\delta)\}$
$\hat{E}_t$	Event of $\{\forall s \leq t, \Theta_* \in \mathcal{E}_t^{\text{RLS}}(\delta)\}$
$\tilde{E}_t$	Event of $\{\forall s \leq t, \tilde{\Theta}_s \in \mathcal{E}_t^{\text{TS}}(\delta)\}$
$\bar{E}_t$	Event of $\{\forall t \leq T_r, \|x_t\| \leq c'(n+d)^{n+d}$ and $\forall t > T_r, \|x_t\| \leq X_s\}$
E_t	$E_t = \hat{E}_t \cap \tilde{E}_t \cap \bar{E}_t$
Optimism Analy.	
$\mathcal{S}^{\mathsf{opt}}$	Optimistic set; $\left\{\Theta = (A,\ B)^\intercal \in \mathbb{R}^{(n+d)\times n} \mid J(\Theta) \leq J(\Theta_*)\right\}$
p_t^{opt}	Probability of selecting optimistic system; $\mathbb{P}\left\{\tilde{\Theta}_t \in \mathcal{S}^{\text{opt}} \mid \mathcal{F}_t^{\text{cnt}}, \hat{E}_t\right\}$
$J(\Theta, K)$	Average expected cost of controlling Θ with a stabilizing controller K
$\Sigma(\Theta, K)$	Covariance matrix of the state in system Θ under controller K
$H_*^\intercal$	Concatenation of identity and optimal controller $K(\Theta_*)$; $[I, K(\Theta_*)^\intercal]$
Q_*	$Q + K(\Theta_*)^\intercal R K(\Theta_*)$
$L(A_c)$	Function that maps any stable matrix A_c to $\sigma_w^2 \sum_{t=0}^{\infty} \left\|A_c^t\right\|_{Q_*}^2$
F_t	Confidence interval for estimated closed-loop system; $\beta_t^2 H_*^\intercal V_t^{-1} H_*$
$\lambda_{\max,t}$	Maximum eigenvalue of F_t
$\lambda_{\min,t}$	Minimum eigenvalue of F_t
Ξ	Random matrix of size $n \times n$ with iid $\mathcal{N}(0,1)$ entries
$A_{c,*}$	Closed-loop system matrix of the Θ_* driven by $K(\Theta_*)$; $\Theta_*^\intercal H_*$
$\hat{A}_c$	Closed-loop system matrix of the $\hat{\Theta}$ driven by $K(\Theta_*)$; $\hat{\Theta}^\intercal H_*$
$\tilde{A}_c$	Closed-loop system matrix of the $\tilde{\Theta}$ driven by $K(\Theta_*)$; $\tilde{\Theta}^\intercal H_*$
$\mathcal{E}_t^{\text{cl}}(\delta)$	Closed-loop confidence set that is super set to $\mathcal{E}_t^{\text{RLS}}(\delta)$; (2.7)
$\hat{\Upsilon}$	Unit Fro. norm matrix s.t. $\hat{\Upsilon}\sqrt{F_t}$ is the h.p. confidence ellipsoid on $A_{c,*}$
$\mathcal{M}_n$	Manifold of square matrices of dimension n; $\mathbb{R}^{n\times n}$
$\mathcal{M}_{\text{Schur}}$	Manifold of (Schur-)stable matrices in $\mathcal{M}_n$; $\{A_c \in \mathbb{M}_n \mid \rho(A_c) < 1\}$
$\mathcal{M}_*$	Sublevel manifold in $\mathcal{M}_{\text{Schur}}$ s.t. $\left\{A_c \in \mathcal{M}_{\text{Schur}} \mid L(A_c) \leq L(A_{c,*})\right\}$
G_t	Perturbation around $A_{c,*}$; $(\Xi + \hat{\Upsilon})\sqrt{F_t}$
∇L_*	Jacobian operator of $L(\cdot)$ evaluated at $A_c \in \mathcal{M}_{\text{Schur}}$
$\sigma_{\min,*}$	Minimum singular value of ∇L_*
$\mathcal{H}_{A_c}$	Hessian operator of $L(\cdot)$ evaluated at $A_c \in \mathcal{M}_{\text{Schur}}$
$\mathcal{B}_*$	Stable ball for some constant ϵ_*; $\left\{A_c \in \mathbb{M}_n \mid \|A_c - A_{c,*}\|_F \leq \epsilon_*\right\}$
$\mathcal{M}_*^{\text{qd}}$	Sublevel manifold; $\left\{A_c \in \mathbb{M}_n \mid \|A_c - A_{c,*} + r_*^{-1}\nabla L_*\|_F \leq \|r_*^{-1}\nabla L_*\|_F\right\}$

2.B System Identification and Stabilization Guarantees

In this section, we show that improved exploration of TSAC provides persistently exciting inputs, which will be used to enable reaching a stabilizing neighborhood around Θ_*. From Assumption 2.2.2, we have that $\mathbb{E}[x_{t+1}x_{t+1}^\intercal \mid \mathcal{F}_t] \succcurlyeq \sigma_w^2 I$. Thus, with the input $u_t = K(\tilde{\Theta}_t)x_t + \nu_t$ for $\nu_t \sim \mathcal{N}(0, 2\kappa^2\sigma_w^2 I)$, we have that $\mathbb{E}[z_{t+1}z_{t+1}^\intercal \mid \mathcal{F}_t] \succcurlyeq \frac{\sigma_w^2}{2}I$. Using Lemma 2.G.5, we have that $V_t \succcurlyeq t\frac{\sigma_w^2}{40}I$ for $t \geq 200(n+d)\log\frac{12}{\delta}$ with probability at least $1-\delta$. Using the RLS estimate error bound given in Section 2.2, *i.e.*, under the event of $\hat{E}_t$ we have

$$\|\hat{\Theta}_t - \Theta_*\|_2 \leq \frac{\beta_t}{\sqrt{\lambda_{\min}(V_t)}} \tag{2.18}$$

with probability at least $1-\delta$. Plugging in the $\lambda_{\min}(V_t)$ in its place yields the first result.

For the second result, we use Lemma 4.2 of Lale, Azizzadenesheli, Hassibi, *et al.* [140]. Recall that $D = \overline{\alpha}\gamma^{-1}\kappa^2(1+\kappa^2)$. Lemma 4.2 of Lale, Azizzadenesheli, Hassibi, *et al.* [140] states that for any (κ,γ)-stabilizable system Θ_* and for any $\varepsilon \leq \min\{\sqrt{(\sigma_w^2 n)/(142D^7)}, 1/(54D^5\}$, such that $\|\Theta' - \Theta_*\| \leq \varepsilon$, $K(\Theta')$ produces $(\kappa\sqrt{2}, \gamma/2)$-stable closed-loop dynamics on Θ_* such that there exists L and $H \succ 0$ such that $A_* + B_*K(\Theta') = H'LH'^{-1}$, with $\|L\| \leq 1-\gamma/2$ and $\|H'\|\|H'^{-1}\| \leq \kappa\sqrt{2}$. Under the event of $\hat{E} \cap \tilde{E}$, we have $\|\tilde{\Theta}_t - \Theta_*\|_2 \leq \frac{\beta_t(\delta)+\upsilon_t(\delta)}{\sqrt{\lambda_{\min}(V_T)}}$. Under the event of $\hat{E} \cap \tilde{E}$, this yields $\|\tilde{\Theta}_t - \Theta_*\|_2 \leq \frac{7(\beta_t(\delta)+\upsilon_t(\delta))}{\sigma_w\sqrt{t}}$ with probability $1-\delta$. Combining this result with the required ε for finding the stabilizing neighborhood, for TS with exploration duration of

$$T_w \geq T_0 \coloneqq \frac{49(\beta_T(\delta)+\upsilon_T(\delta))^2}{\sigma_w \min\{(\sigma_w^2 n)/(142D^7), 1/(54^2D^{10}\}}, \tag{2.19}$$

TSAC achieves $(\kappa\sqrt{2}, \gamma/2)$-stable closed-loop dynamics on Θ_*, with probability at least $1-3\delta$.

2.C Boundedness of State, Proof of Lemma 2.4.3

In this section, we show that under the joint event of $\hat{E} \cap \tilde{E}$ and the stabilization guarantee of the previous section, the state is bounded at all times during TSAC and it is well-controlled during the stabilizing TS phase, *i.e.*provide the proof of Lemma 2.4.3. We first consider the evolution of state for $t \leq T_w$. To bound the state for the first phase, we adapt the state bounding strategy given in Section 4.1 of Abbasi-Yadkori and Szepesvári [2] for contractible systems to the stabilizable systems via the slow policy changes of TSAC. To this end, define the following

$$\bar{\alpha}_t = \frac{18\kappa^3}{\gamma(8\kappa-1)}\bar{\eta}^{n+d}\left[GZ_t^{\frac{n+d}{n+d+1}}\beta_t(\delta)^{\frac{1}{2(n+d+1)}} + (\|B_*\|\sigma_\nu + \sigma_w)\sqrt{2n\log\frac{nt}{\delta}}\right],$$

for

$$\bar{\eta} \geq \sup_{\Theta\in\mathcal{S}}\|A_* + B_*K(\Theta)\|, \qquad Z_T = \max_{1\leq t\leq T_r}\|z_t\|, \quad U = \frac{U_0}{H}$$
$$G = 2\left(\frac{2S(n+d)^{n+d+1/2}}{\sqrt{U}}\right)^{1/(n+d+1)}, \quad U_0 = \frac{1}{16^{n+d-2}\max\left(1,\ S^{2(n+d-2)}\right)}$$

and where H is any number satisfying

$$H > \max\left(16,\ \frac{4S^2M^2}{(n+d)U_0}\right), \text{ where } \quad M = \sup_{Y\geq 1}\frac{\left(\sigma_w\sqrt{n(n+d)\log\left(\frac{1+TY/\lambda}{\delta}\right)} + \lambda^{1/2}S\right)}{Y}.$$

Under the joint event of $\hat{E}_t \cap \tilde{E}_t$, Abbasi-Yadkori and Szepesvári [2] show that the norm of the state is well-controlled except $n+d$ times at most in any horizon T_r. Denoting the set of time-steps that the state is not well-controlled by $\mathcal{T}_t$, the following lemma formalizes this argument:

Lemma 2.C.1 (Lemma 18 of Abbasi-Yadkori and Szepesvári [2])**.** *We have that for any* $0 \leq t \leq T$*,*

$$\max_{s\leq t, s\notin T_t}\left\|(\Theta_* - \tilde{\Theta}_s)^\intercal z_s\right\| \leq GZ_t^{\frac{n+d}{n+d+1}}(\beta_t(\delta) + \upsilon(\delta))^{\frac{1}{2(n+d+1)}}.$$

Notice that this lemma is updated for TS. Moreover, it does not depend neither on the contractibility of the underlying system on the standard basis nor on the stabilizability. Equipped with this result, we write the closed loop system as

$$x_{t+1} = \Gamma_t x_t + r_t$$

where

$$\Gamma_t = \begin{cases} \tilde{A}_{t-1} + \tilde{B}_{t-1}K(\tilde{\Theta}_{t-1}) & t \notin \mathcal{T}_{T_w} \\ A_* + B_*K(\tilde{\Theta}_{t-1}) & t \in \mathcal{T}_{T_w} \end{cases} \text{ and } r_t = \begin{cases} (\Theta_* - \tilde{\Theta}_{t-1})^\intercal z_t + B_*\nu_t + w_t & t \notin \mathcal{T}_{T_w} \\ B_*\nu_t + w_t & t \in \mathcal{T}_{T_w} \end{cases} \tag{2.20}$$

Starting from $x_0 = 0$, we obtain the following roll out for the state,

$$\begin{aligned}
x_t &= \Gamma_{t-1}x_{t-1} + r_{t-1} = \Gamma_{t-1}\left(\Gamma_{t-2}x_{t-2} + r_{t-2}\right) + r_t \\
&= \Gamma_{t-1}\Gamma_{t-2}\Gamma_{t-3}x_{t-3} + \Gamma_{t-1}\Gamma_{t-2}r_{t-2} + \Gamma_{t-1}r_{t-1} + r_t \\
&= \Gamma_{t-1}\Gamma_{t-2}\dots\Gamma_{t-(t-1)}r_1 + \dots + \Gamma_{t-1}\Gamma_{t-2}r_{t-2} + \Gamma_{t-1}r_{t-1} + r_t \\
&= \sum_{k=1}^{t}\left(\prod_{s=k}^{t-1}\Gamma_s\right)r_k
\end{aligned} \tag{2.21}$$

Recall that the sampled model is an element of $\mathcal{S}$ due to rejection sampling, thus, it is (κ, γ)-stabilizable by its optimal controller (Assumption 2.2.1):

$$1 - \gamma \geq \max_{t \leq T} \rho\left(\tilde{A}_t + \tilde{B}_t K(\tilde{\Theta}_t)\right). \tag{2.22}$$

Notice that multiplication of the closed-loop system matrices are not guaranteed to be contractive without a similarity transformation. Therefore, unlike Abbasi-Yadkori and Szepesvári [2] that bounds the rollout terms via contractive mappings due to their assumption of contractive systems, we need to make sure that the policy changes does not cause unexpected growth in the magnitude of the state. The slow policy update schedule, *i.e.*, using all the sampled controllers for fixed τ_0 time-steps, allows us to prevent such undesirable outcomes, In particular, by setting $\tau_0 = 2\gamma^{-1}\log(2\kappa\sqrt{2})$, we have that

$$\|x_t\| \leq \frac{18\kappa^3\bar{\eta}^{n+d}}{\gamma(8\kappa-1)}\left(\max_{1\leq k\leq t}\|r_k\|\right) \tag{2.23}$$

Moreover, we have that $\|r_k\| \leq \left\|(\Theta_* - \tilde{\Theta}_{k-1})^\intercal z_k\right\| + \|B_*\nu_k + w_k\|$ when $k \notin \mathcal{T}_T$, and $\|r_k\| = \|B_*\nu_k + w_k\|$, otherwise. Hence,

$$\max_{k\leq t}\|r_k\| \leq \max_{k\leq t, k\notin\mathcal{T}_t}\left\|(\Theta_* - \tilde{\Theta}_{k-1})^\intercal z_k\right\| + \max_{k\leq t}\|B_*\nu_k + w_k\|$$

The first term is bounded by the Lemma 2.C.1. The second term involves summation of independent $\|B_*\|\sigma_\nu$ and σ_w Gaussian vectors. Using standard Gaussian tail inequalities, for all $k \leq t$, we have $\|B_*\nu_k + w_k\| \leq (\|B_*\|\sigma_\nu + \sigma_w)\sqrt{2n\log\frac{nt}{\delta}}$ with probability at least $1-\delta$. Therefore, on the joint event of $\hat{E} \cap \tilde{E}$,

$$\|x_t\| \leq \frac{18\kappa^3\bar{\eta}^{n+d}}{\gamma(8\kappa-1)}\left[GZ_t^{\frac{n+d}{n+d+1}}(\beta_t(\delta) + \upsilon(\delta))^{\frac{1}{2(n+d+1)}} + (\|B_*\|\sigma_\nu + \sigma_w)\sqrt{2n\log\frac{nt}{\delta}}\right] \tag{2.24}$$

for $t \leq T_w$ with probability $1-\delta$. Notice that this bound depends on Z_t and $\beta_t(\delta)$ which in turn depends on x_t. Using Lemma 5 of [2], one can obtain the following bound

$$\|x_t\| \leq c'(n+d)^{n+d}. \tag{2.25}$$

for some constant c' for all $t \leq T_w$ with probability $1-3\delta$, which gives the first advertised result.

To bound the state for $t > T_w$, we show that, with the given choice of τ_0, all the controllers during the stabilizing TS phase halves the magnitude of the state at the end of their control period. In particular, during the stabilizing TS phase, the closed-loop system dynamics can be written as $x_{t+1} = (A_* + B_* K(\tilde{\Theta}_t))x_t + w_t = \Theta_*^\intercal H_{K(\tilde{\Theta}_t)} + w_t$. From the choice of T_w for the stabilizable systems, we have that $\Theta_*^\intercal H_{K(\tilde{\Theta}_t)}$ is $(\kappa\sqrt{2}, \gamma/2)$-strongly stable. Thus, we have $\rho(\Theta_*^\intercal H_{K(\tilde{\Theta}_t)}) \leq 1-\gamma/2$ for all $t > T_w$ and $\|H_t\|\|H_t^{-1}\| \leq \kappa\sqrt{2}$ for $H_t \succ 0$, such that $\|L_t\| \leq 1-\gamma/2$ for $\Theta_*^\intercal H_{K(\tilde{\Theta}_t)} = H_t L_t H_t^{-1}$. Then for $T > t > T_w$, if the same policy, $\Theta_*^\intercal H_{K(\tilde{\Theta})}$ is applied starting from the state x_{T_w}, we have the following state roll-out on the event of $\hat{E}_t \cap \tilde{E}_t$

$$\|x_t\| = \left\| \prod_{i=T_w+1}^{t} \Theta_*^\intercal H_{K(\tilde{\Theta})} x_{T_w} + \sum_{i=T_w+1}^{t} \left(\prod_{s=i}^{t-1} \Theta_*^\intercal H_{K(\tilde{\Theta})} \right) w_i \right\| \tag{2.26}$$

$$\leq \kappa\sqrt{2}(1-\gamma/2)^{t-T_w}\|x_{T_w}\| + \max_{T_w < i \leq T} \|w_i\| \left(\sum_{i=T_w+1}^{t} \kappa\sqrt{2}(1-\gamma/2)^{t-i+1} \right) \tag{2.27}$$

$$\leq \kappa\sqrt{2}(1-\gamma/2)^{t-T_w}\|x_{T_w}\| + \frac{2\kappa\sigma_w\sqrt{2}}{\gamma}\sqrt{2n\log(n(t-T_w)/\delta)} \tag{2.28}$$

with probability at least $1-\delta$. Since $\tau_0 = 2\gamma^{-1}\log(2\kappa\sqrt{2})$, we have $\kappa\sqrt{2}(1-\gamma/2)^{\tau_0} \leq 1/2$. Therefore, at the end of each controller period the effect of previous state is halved. Using this fact, at the ith policy change after T_w, we get

$$\|x_{t_i}\| \leq 2^{-i}\|x_{T_w}\| + \sum_{j=0}^{i-1} 2^{-j}\frac{2\kappa\sigma_w\sqrt{2}}{\gamma}\sqrt{2n\log(n(t-T_w)/\delta)}$$

$$\leq 2^{-i}\|x_{T_w}\| + \frac{4\kappa\sigma_w\sqrt{2}}{\gamma}\sqrt{2n\log(n(t-T_w)/\delta)}$$

For all $i > (n+d)\log(n+d) - \log(\frac{2\kappa\sigma_w\sqrt{2}}{\gamma}\sqrt{2n\log(n(t-T_w)/\delta)})$, at policy change i, we get

$$\|x_{t_i}\| \leq \frac{6\kappa\sigma_w\sqrt{2}}{\gamma}\sqrt{2n\log(n(t-T_w)/\delta)}.$$

Finally, from (2.28), we have that

$$\|x_t\| \leq \frac{(12\kappa^2 + 2\kappa\sqrt{2})\sigma_w}{\gamma}\sqrt{2n\log(n(t-T_w)/\delta)}, \tag{2.29}$$

with probability $1-4\delta$ for all $t > T_r := T_w + ((n+d)\log(n+d))\,\tau_0$. Based on this result, let $X_s = \frac{(12\kappa^2+2\kappa\sqrt{2})\sigma_w}{\gamma}\sqrt{2n\log(n(T-T_w)/\delta)}$. We define our final good event,

$$\bar{E}_t = \{\forall t \leq T_r, \|x_t\| \leq c'(n+d)^{n+d} \text{ and } \forall t > T_r, \|x_t\| \leq X_s\}. \tag{2.30}$$

Notice that the joint event $E_t = \hat{E}_t \cap \tilde{E}_t \cap \bar{E}_t$ holds with probability at least $1-4\delta$. This event will be the key conditioning in the regret decomposition and the analysis.

2.D Constant Probability of Sampling Optimistic Models

In this section, we give the proof of the main technical contribution of this work, showing that TS samples optimistic model parameters with constant probability (Theorem 2.5.1). The proof follows the outline provided in Section 2.5. We first provide the proofs of each lemma in Section 2.5. In particular Lemma 2.5.2 is proven in Appendix 2.D, Lemma 2.5.3 is studied in Appendix 2.D, Lemma 2.5.4 in Appendix 2.D, and Lemma 2.5.5 in 2.D. Finally, we combine these results to prove Theorem 2.5.1 in Appendix 2.D.

Proof of Lemma 2.5.2

Given a stabilizable system $\Theta = (A, B)^\intercal$, and a stabilizing linear feedback controller K, we can find the LQR cost as follows

$$\begin{aligned} J(\Theta, K) &= \lim_{T\to\infty} \frac{1}{T}\,\mathbb{E}\left[\sum\nolimits_{t=1}^{T} x_t^\intercal Q x_t + u_t^\intercal R u_t\right], && (2.31)\\ &= \lim_{T\to\infty} \frac{1}{T}\,\mathbb{E}\left[\sum\nolimits_{t=1}^{T} \operatorname{tr}\left((Q + K^\intercal R K) x_t x_t^\intercal\right)\right], && (2.32)\\ &= \lim_{T\to\infty} \operatorname{tr}\left((Q + K^\intercal R K)\frac{1}{T}\sum\nolimits_{t=1}^{T} \mathbb{E}\left[x_t x_t^\intercal\right]\right), && (2.33)\\ &= \operatorname{tr}\left((Q + K^\intercal R K)\Sigma(\Theta, K)\right) && (2.34) \end{aligned}$$

where $\Sigma(\Theta, K) := \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[x_t x_t^\intercal\right]$ is the stationary state covariance of the closed-loop system. In (2.32), we used the feedback control policy relation $u_t = Kx_t$ and trace trick for inner products of vectors. Note that the closed-loop system evolves as

$$x_{t+1} = (A + BK)x_t + w_t. \tag{2.35}$$

The covariance of the state at time t can be written as a recursive relation

$$\mathbb{E}\left[x_{t+1}x_{t+1}^{\intercal}\right] = \mathbb{E}\left[((A+BK)x_t + w_t)((A+BK)x_t + w_t)^{\intercal}\right] \tag{2.36}$$

$$= (A+BK)\,\mathbb{E}\left[x_t x_t^{\intercal}\right](A+BK)^{\intercal} + \mathbb{E}\left[w_t w_t^{\intercal}\right] \tag{2.37}$$

$$= (A+BK)\,\mathbb{E}\left[x_t x_t^{\intercal}\right](A+BK)^{\intercal} + \sigma_w^2 I \tag{2.38}$$

where (2.37) is because $\mathbb{E}\left[w_t\right] = 0$ and w_t and x_t are independent. Since $\rho(A+BK) < 1$, the above iteration converges to a finite fixed-point. Furthermore, we have the following relation

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[x_{t+1}x_{t+1}^{\intercal}\right] = (A+BK)\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[x_t x_t^{\intercal}\right](A+BK)^{\intercal} + \sigma_w^2 I \tag{2.39}$$

Denoting by $\Sigma_T(\Theta,K) \coloneqq \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[x_t x_t^{\intercal}\right]$ the finite averaged state covariance, we have the following

$$\Sigma_T(\Theta,K) + \frac{\mathbb{E}\left[x_{T+1}x_{T+1}^{\intercal}\right] - \mathbb{E}\left[x_1 x_1^{\intercal}\right]}{T} = (A+BK)\Sigma_T(\Theta,K)(A+BK)^{\intercal} + \sigma_w^2 I \tag{2.40}$$

Taking the limit of both sides as $T \to \infty$ and noting that $\mathbb{E}\left[x_{T+1}x_{T+1}^{\intercal}\right]$ has a finite value at the limit, we obtain the following Lyapunov equation

$$\Sigma(\Theta,K) = (A+BK)\Sigma(\Theta,K)(A+BK)^{\intercal} + \sigma_w^2 I \tag{2.41}$$

whose solution is given by the following convergent infinite sum

$$\Sigma(\Theta,K) = \sum_{t=0}^{\infty}(A+BK)^t \sigma_w^2 I\left((A+BK)^{\intercal}\right)^t \tag{2.42}$$

It is well known that the optimal control policy of infinite-horizon LQR systems can be achieved by stationary linear feedback controllers [19]. Therefore, we can find the optimal LQR cost of a stabilizable system by minimizing its closed-loop cost among all stabilizing stationary linear feedback controllers.

Suppose $\Theta \in \mathcal{S}^{\text{surr}}$, *i.e.*, $J(\Theta, K(\Theta_*)) \leq J(\Theta_*, K(\Theta_*))$. Then, the optimal LQR cost of Θ is given as

$$J(\Theta) = J(\Theta, K(\Theta)) = \min_{K \in \mathbb{R}^{d\times n}} J(\Theta, K) \tag{2.43}$$

$$\leq J(\Theta, K(\Theta_*)) \overset{(a)}{\leq} J(\Theta_*, K(\Theta_*)) = J(\Theta_*) \tag{2.44}$$

where (a) is due to $\Theta \in \mathcal{S}^{\text{surr}}$. Thus, $\Theta \in \mathcal{S}^{\text{opt}}$. ■

Proof of Lemma 2.5.3

The following lemma will be used as the backbone for Lemma 2.5.3.

Lemma 2.D.1. *Let $V_1, V_2 \in \mathbb{R}^{n\times n}$ be symmetric positive semi-definite matrices. Define two ellipsoids as*

$$\mathcal{E}_1 \coloneqq \left\{\Theta \in \mathbb{R}^{n\times m} \;\middle|\; \mathrm{tr}\,(\Theta^\intercal V_1 \Theta) \leq 1\right\} \quad \textit{and} \quad \mathcal{E}_2 \coloneqq \left\{\Theta \in \mathbb{R}^{n\times m} \;\middle|\; \mathrm{tr}\,(\Theta^\intercal V_2 \Theta) \leq 1\right\} \tag{2.45}$$

Then, $\mathcal{E}_1 \subseteq \mathcal{E}_2$ if and only if $V_1 \succcurlyeq V_2$.

Proof. For the forward direction, assume $V_1 - V_2$ has a negative eigenvalue, *i.e.*, there exist $\lambda < 0$ and a unit vector $\theta \in \mathbb{R}^n/\{0\}$ such that $(V_1 - V_2)\theta = \lambda\theta$. Construct $\Theta = [\theta, \theta, \ldots, \theta] \in \mathbb{R}^{n\times m}$. Observe that $\mathrm{tr}\,(\Theta^\intercal V_1 \Theta) = m\theta^\intercal V_1 \theta$ and $\mathrm{tr}\,(\Theta^\intercal V_2 \Theta) = m\theta^\intercal V_2 \theta$. Therefore, we have the relationship $\mathrm{tr}\,(\Theta^\intercal V_2 \Theta) = \mathrm{tr}\,(\Theta^\intercal V_1 \Theta) - m\lambda$.

If $V_1\theta = 0$, then $\mathrm{tr}(\Theta^\intercal V_1 \Theta) = 0 \leq 1$ and therefore for any scalar $\alpha > 0$, $\alpha\Theta \in \mathcal{E}_1$. On the other hand, $\mathrm{tr}\,(\Theta^\intercal V_2 \Theta) = -m\lambda > 0$ and therefore, one can find a scalar $\alpha > 0$ such that $\mathrm{tr}\,((\alpha\Theta)^\intercal V_2(\alpha\Theta)) = -m\lambda\alpha^2 > 1$, *i.e.*$\alpha\Theta \notin \mathcal{E}_2$. If $V_1\theta \neq 0$, then define $\Theta' = \frac{1}{\sqrt{m\theta^\intercal V_1\theta}}\Theta$ and observe that $\mathrm{tr}\left(\Theta'^\intercal V_1 \Theta'\right) = 1$, *i.e.*, $\Theta' \in \mathcal{E}_1$. On the other hand, $\mathrm{tr}\left(\Theta'^\intercal V_2 \Theta'\right) = 1 - \frac{\lambda}{\theta^\intercal V_1 \theta} > 1$, *i.e.*, $\Theta' \notin \mathcal{E}_2$. Therefore, we have that if $\mathcal{E}_1 \subseteq \mathcal{E}_2$ then $V_1 \succcurlyeq V_2$.

For the reverse direction, assume that $V_1 \succcurlyeq V_2$ and $\Theta \in \mathcal{E}_1$. Then, $\mathrm{tr}\,(\Theta^\intercal(V_1 - V_2)\Theta) \geq 0$ and $\mathrm{tr}\,(\Theta^\intercal V_2 \Theta) \leq \mathrm{tr}\,(\Theta^\intercal V_1 \Theta) \leq 1$. Therefore, $\Theta \in \mathcal{E}_2$. ■

Proof of Lemma 2.5.3. Let us rewrite the the ellipsoids. For the time being, we will drop δ dependence for simplicity.

$$\mathcal{E}_t^{\mathrm{RLS}} = \left\{\hat{\Theta} \in \mathbb{R}^{(n+d)\times n} \;\middle|\; \mathrm{tr}\left((\hat{\Theta} - \Theta_*)^\intercal \beta_t^{-1} V_t (\hat{\Theta} - \Theta_*)\right) \leq 1\right\}, \tag{2.46}$$

$$\mathcal{E}_t^{\mathrm{cl}} = \left\{\hat{\Theta} \in \mathbb{R}^{(n+d)\times n} \;\middle|\; \mathrm{tr}\left((\hat{\Theta} - \Theta_*)^\intercal H_* F_t^{-1} H_*^\intercal (\hat{\Theta} - \Theta_*)\right) \leq 1\right\}. \tag{2.47}$$

In order to prove the lemma, it is necessary and sufficient to show $\beta_t^{-1} V_t \succcurlyeq H_* F_t^{-1} H_*^\intercal$ by Lemma 2.D.1. Eliminating b_t terms from both sides and multiplying by $V_t^{-\frac{1}{2}}$ from left and right, we obtain the equivalent condition,

$$I \succcurlyeq V_t^{-\frac{1}{2}} H_* (H_*^\intercal V_t^{-1} H_*)^{-1} H_*^\intercal V_t^{-\frac{1}{2}} = V_t^{-\frac{1}{2}} H_* (H_*^\intercal V_t^{-1} H_*)^{-\frac{1}{2}} (H_*^\intercal V_t^{-1} H_*)^{-\frac{1}{2}} H_*^\intercal V_t^{-\frac{1}{2}}, \tag{2.48}$$

In other words, we have that $\mathcal{E}_t^{\text{RLS}} \subseteq \mathcal{E}_t^{\text{cl}}$ if and only if $\|(H_*^\intercal V_t^{-1} H_*)^{-\frac{1}{2}} H_*^\intercal V_t^{-\frac{1}{2}}\|_2 \leq 1$. Notice that

$$\|(H_*^\intercal V_t^{-1} H_*)^{-\frac{1}{2}} H_*^\intercal V_t^{-\frac{1}{2}}\|_2^2 = \sigma_1\left((H_*^\intercal V_t^{-1} H_*)^{-\frac{1}{2}} H_*^\intercal V_t^{-\frac{1}{2}}\right)^2, \tag{2.49}$$

$$= \lambda_{\max}\left(V_t^{-\frac{1}{2}} H_* (H_*^\intercal V_t^{-1} H_*)^{-1} H_*^\intercal V_t^{-\frac{1}{2}}\right), \tag{2.50}$$

$$= \lambda_{\max}\left((H_*^\intercal V_t^{-1} H_*)^{-\frac{1}{2}} H_*^\intercal V_t^{-1} H_* (H_*^\intercal V_t^{-1} H_*)^{-\frac{1}{2}}\right), \tag{2.51}$$

$$= \lambda_{\max}(I) = 1, \tag{2.52}$$

where we used the fact that $\sigma_1(A) = \sqrt{\lambda_{\max}(A^\intercal A)} = \sqrt{\lambda_{\max}(AA^\intercal)}$. This is true for any time t and δ and therefore completes the proof. ■

Proof of Lemma 2.5.4

The following lemma guarantees existence of a stable neighborhood around any stable matrix.

Lemma 2.D.2. *Let $A_c \in \mathcal{M}_{Schur}$, i.e., $\rho(A_c) < 1$. Then, there exists $\epsilon > 0$ such that for any $\Delta \in \mathbb{M}_n$ with $\|\Delta\|_F \leq 1$, we have that $A_c + \epsilon\Delta \in \mathcal{M}_{Schur}$, i.e., $\rho(A_c + \epsilon\Delta) < 1$.*

Proof. Per Gelfand's formula, we have that for any $\delta > 0$, there exists $N_\delta \in \mathbb{N}$ such that

$$\rho(A_c) \leq \|A_c^k\|_F^{1/k} < \rho(A_c) + \delta \tag{2.53}$$

for any $k \geq N_\delta$. Since the mapping $A_c \mapsto \|A_c^k\|_F^{1/k}$ is smooth for any $k \in \mathbb{N}$, we can write the following expansion by Taylor's theorem for any $t \in \mathbb{R}$

$$\|(A_c + t\Delta)^k\|_F^{1/k} = \|A_c^k\|_F^{1/k} + t\frac{\mathrm{d}}{\mathrm{d}t}\|(A_c + t\Delta)^k\|_F^{1/k}\Big|_{\lambda t} \tag{2.54}$$

where $\lambda \in [0, 1]$. For a given $t \in \mathbb{R}$, there exists a constant $M_{k,t} > 0$ such that for any $\|\Delta\|_F \leq 1$, we have that $|\frac{\mathrm{d}}{\mathrm{d}t}\|(A_c + t\Delta)^k\|_F^{1/k}| \leq M_{k,t}$ by Taylor's theorem. Then, we can write the following upper bound

$$\|(A_c + t\Delta)^k\|_F^{1/k} \leq \|A_c^k\|_F^{1/k} + |t| M_{t,k} \tag{2.55}$$

Using the relation (2.53) and the upper bound (2.55), we have that for any $\delta > 0$, $t > 0$, and $\|\Delta\|_F \leq 1$, there exists $N_\delta \in \mathbb{N}$ and $M_{t,N_\delta} > 0$ such that

$$\rho(A_c + t\Delta) \leq \|(A_c + t\Delta)^{N_\delta}\|_F^{1/N_\delta} \leq \|A_c^{N_\delta}\|_F^{1/N_\delta} + tM_{t,N_\delta} \tag{2.56}$$

$$< \rho(A_c) + \delta + tM_{t,N_\delta} \tag{2.57}$$

Fix a $\delta > 0$ such that $\rho(A_c) + \delta < 1$ and fix a $t > 0$. Then, we can find $0 < \epsilon \leq t$ such that $\rho(A_c) + \delta + \epsilon M_{t,N_\delta} < 1$ and thus

$$\rho(A_c + \epsilon\Delta) < \rho(A_c) + \delta + \epsilon M_{t,N_\delta} < 1 \tag{2.58}$$

for any $\|\Delta\|_F \leq 1$ by (2.57). ■

Proof of Lemma 2.5.4. For any $A_c \in \mathbb{M}_{\text{Schur}}$, there exists a constant $\epsilon > 0$, such that for any $\|\Delta\|_F \leq 1$, we have that $A_c + \epsilon\Delta \in \mathbb{M}_{\text{Schur}}$ by Lemma 2.D.2. To see smoothness of L, we write $A_t := A_c + t\Delta$ and $L(A_t) = \operatorname{tr}(Q_*\Sigma_t)$ for any $|t| \leq \epsilon$ and $\|\Delta\|_F \leq 1$ where Σ_t solves the following Lyapunov equation

$$\Sigma_t - A_t\Sigma_t A_t^\intercal = \sigma_w^2 I \text{ and } \Sigma_0 - A_c\Sigma_0 A_c^\intercal = \sigma_w^2 I \tag{2.59}$$

Note that, $\rho(A_t) < 1$ for any $|t| \leq \epsilon$ and therefore both equations in (2.59) have unique solutions for any $|t| \leq \epsilon$. The Jacobian $\nabla L(A_c) \in \mathbb{M}_n$ satisfies $\nabla L(A_c) \bullet \Delta = \frac{\mathrm{d}}{\mathrm{d}t} L(A_t)\big|_{t=0} = \operatorname{tr}(Q_*\dot{\Sigma}_0)$ for any $\|\Delta\|_F \leq 1$ where $\dot{\Sigma}_t$ is the derivative of Σ_t and satisfies the following Lyapunov equation

$$\dot{\Sigma}_t - A_t\dot{\Sigma}_t A_t^\intercal = \Delta\Sigma_t A_t^\intercal + A_t\Sigma_t\Delta^\intercal \text{ and } \dot{\Sigma}_0 - A_c\dot{\Sigma}_0 A_c^\intercal = \Delta\Sigma_0 A_c^\intercal + A_c\Sigma_0\Delta^\intercal \tag{2.60}$$

Similarly, both equations in (2.60) have unique solutions for any $|t| \leq \epsilon$ and therefore $\nabla L(A_c)$ exists for any A_c. To find the Jacobian, we have that $\dot{\Sigma}_0 = \sum_{k=0}^{\infty} A_c^k \left(\Delta\Sigma_0 A_c^\intercal + A_c\Sigma_0\Delta^\intercal\right) (A_c^\intercal)^k$ and

$$\operatorname{tr}(Q_*\dot{\Sigma}_0) = \operatorname{tr}\left(Q_* \sum_{k=0}^{\infty} A_c^k \left(\Delta\Sigma_0 A_c^\intercal + A_c\Sigma_0\Delta^\intercal\right) (A_c^\intercal)^k\right) \tag{2.61}$$

$$= 2\operatorname{tr}\left(\sum_{k=0}^{\infty} (A_c^\intercal)^k Q_* A_c^k A_c\Sigma_0\Delta^\intercal\right) = 2\sum_{k=0}^{\infty} (A_c^\intercal)^k Q_* A_c^k A_c\Sigma_0 \bullet \Delta \tag{2.62}$$

Therefore, $\nabla L(A_c) = 2\sum_{k=0}^{\infty}(A_c^\intercal)^k Q_* A_c^k A_c\Sigma_0$. In particular, in the case of $A_{c,*}$, we have that $\sum_{k=0}^{\infty}(A_{c,*}^\intercal)^k Q_* A_{c,*}^k = P_*$, the solution to the Riccati equation, and thus $\nabla L(A_{c,*}) = 2P_*A_{c,*}\Sigma_*$. Repeating the same process, one can see that $L(A_t)$ is infinitely differentiable and thus we conclude L is a smooth function.

Denote by $\mathcal{B}_\epsilon := \{A \in \mathbb{M}_n \mid \|A - A_c\|_F \leq \epsilon\} \subset \mathbb{M}_{\text{Schur}}$ the ball of radius $\epsilon > 0$ around $A_c \in \mathbb{M}_{\text{Schur}}$. Consider the function L restricted to the domain $\mathcal{B}_\epsilon$. Since

$\mathcal{B}_\epsilon$ is a convex set, we can apply Taylor's theorem to L around A_c in this domain to obtain

$$L(A_c + \epsilon\Delta) = L(A_c) + \nabla L(A_c) \bullet \epsilon\Delta + \frac{1}{2}\epsilon\Delta \bullet \mathcal{H}_{A_c+s\Delta}(\epsilon\Delta) \tag{2.63}$$

for $\|\Delta\|_F \leq 1$ and for some $s \in [0, \epsilon]$. Here, $\mathcal{H}_{A_c} : \mathbb{M}_n \to \mathbb{M}_n$ is the Hessian operator evaluated at a point $A_c \in \mathcal{M}_{\text{Schur}}$ and satisfies the following relationship

$$\Delta \bullet \mathcal{H}_{A_c}(\Delta) = \frac{\mathrm{d}^2}{\mathrm{d}t^2} L(A_c + t\Delta)\Big|_{t=0} \tag{2.64}$$

for any $\|\Delta\|_F \leq 1$. Finally, there exists a constant $r > 0$, such that for any $G \in \mathbb{M}_n$, we have that $|G \bullet \mathcal{H}_{A_c+s\Delta}(G)| \leq r\|G\|_F^2$ for any $s \in [0, \epsilon]$ and $\|\Delta\|_F \leq 1$ by Taylor's theorem . ■

Proof of Lemma 2.5.5

In this section, we will assume that Assumptions 2.2.1 & 2.2.2 hold. First, we need to show the boundedness of the stacked state and control input vector, z_t.

Lemma 2.D.3. *Define the terms*

$$Z'_{T_w} \coloneqq (1+\kappa)c'(n+d)^{n+d} + \kappa\sigma_w\sqrt{4d\log(dT_w/\delta)} \tag{2.65}$$

$$Z''_T \coloneqq (1+\kappa)(12\kappa^2 + 2\kappa\sqrt{2})\gamma^{-1}\sigma_w\sqrt{2n\log(n(T-T_w)/\delta)} \tag{2.66}$$

Then, the following holds w.p. at least $1 - 4\delta$*,*

$$\|z_t\| \leq \begin{cases} Z'_{T_w}, & \textit{for } t \leq T_r \\ Z''_T, & \textit{for } T_r < t \leq T \end{cases} \tag{2.67}$$

Proof. From Lemma 2.4.3, we know that $\|x_t\| \leq c'(n+d)^{n+d}$ with $c' > 0$ a constant for $t \leq T_r$ and $\|x_t\| \leq (12\kappa^2 + 2\kappa\sqrt{2})\gamma^{-1}\sigma_w\sqrt{2n\log(n(t-T_w)/\delta)}$ for all $T_r < s \leq T$ w.p. at least $1 - 4\delta$. Furthermore, under the event of E_t, we have that $\|u_t\| \leq \kappa\|x_t\| + \|v_t\| \leq \kappa\|x_t\| + \kappa\sigma_w\sqrt{4d\log(dT_w/\delta)}$ for all $0 \leq t \leq T_w$. Observing that $\|z_t\| = \sqrt{\|x_t\|^2 + \|u_t\|^2} \leq \|x_t\| + \|u_t\|$, one can reach the desired result by substituting the appropriate bounds on $\|x_t\|$ and $\|u_t\|$ and considering the maximal case achieved when $t = T$. ■

The following lemma will be used to bound V_t.

Lemma 2.D.4. *Let $V_t = \mu I + \sum_{s=0}^{t-1} z_s z_s^\intercal$. On the event of $E_T = \hat{E} \cap \tilde{E} \cap \bar{E}$, we have*

$$\lambda_{\max}(V_t) \leq \begin{cases} \mu + t{Z'}_{T_w}^2, & \text{for } t \leq T_r \\ \mu + T_r {Z'}_{T_w}^2 + (t - T_r){Z''_T}^2, & \text{for } T_r < t \leq T \end{cases} \tag{2.68}$$

$$\text{and} \quad \lambda_{\min}(V_t) \geq \begin{cases} \mu + t\frac{\sigma_w^2}{40}, & \text{for } 200(n+d)\log\frac{12}{\delta} \leq t \leq T_w \\ \mu + T_w \frac{\sigma_w^2}{40}, & \text{for } T_w < t \leq T \end{cases} \tag{2.69}$$

Proof. Recall that on the event E_T, the RLS estimates, TS sampled systems are concentrated and the state is bounded, *i.e.*, Lemma 2.4.3. Conditioned on this event, we will start with bounding $\lambda_{\max}(V_t)$. For any time $0 \leq t \leq T$, triangle inequality gives $\lambda_{\max}(V_t) = \|\mu I + \sum_{s=0}^{t-1} z_s z_s^\intercal\|_2 \leq \mu + \sum_{s=0}^{t-1} \|z_s\|^2$. Using the bounds on $\|z_t\|$ given in Lemma 2.D.3, we can write $\lambda_{\max}(V_t) \leq \mu + t{Z'}_{T_w}^2$ for $t \leq T_r$ and $\lambda_{\max}(V_t) \leq \mu + T_r {Z'}_{T_w}^2 + (t - T_r){Z''_T}^2$ for $T_r < t \leq T$. For the lower bound, note that we have that $\mathbb{E}[z_{t+1} z_{t+1}^\intercal \mid \mathcal{F}_t] \succcurlyeq \frac{\sigma_w^2}{2} I$. Using Lemma 2.G.5, on the event E_T, we have that $V_t \succcurlyeq \mu I + t\frac{\sigma_w^2}{40} I$ for $200(n+d)\log\frac{12}{\delta} \leq t \leq T_w$. Since $V_{t+1} = V_t + z_t z_t^\intercal$, we have that $V_t \succcurlyeq V_{T_w} \succcurlyeq \mu I + T_w \frac{\sigma_w^2}{40} I$ for $T_w < t \leq T$. ■

Finally, we will use the following lemma to bound $\beta_t(\delta) = \sigma_w \sqrt{2n \log\left(\frac{\det(V_t)^{1/2}}{\delta \det(\mu I)^{1/2}}\right)} + \sqrt{\mu} S$

Lemma 2.D.5. *On the event of E_T, we have the following upper bound on $\beta_T(\delta)$:*

$$\beta_T(\delta) \leq 4\sigma_w^2 n \log\left(\frac{1}{\delta}\right) + 2\sigma_w^2 n(n+d)\log\left(1 + \frac{T_r {Z'}_{T_w}^2 + (T - T_r){Z''_T}^2}{(n+d)\mu}\right) + 2\mu S^2 \tag{2.70}$$

Proof. Following a similar approach pursued in Lemma 10 of [2], we can bound the log-determinant of V_t as

$$\log\frac{\det(V_T)}{\det(\mu I)} \leq (n+d)\log\left(1 + \frac{T_r {Z'}_{T_w}^2 + (T - T_r){Z''_T}^2}{(n+d)\mu}\right)$$

by Lemma 2.D.4. This leads to the following upper bound on $\beta_t(\delta)$

$$\begin{aligned} \beta_T(\delta)^2 &\leq \left(\sigma_w \sqrt{2n \log\left(\frac{1}{\delta}\right) + n(n+d)\log\left(1 + \frac{T_r {Z'}_{T_w}^2 + (T - T_r){Z''_T}^2}{(n+d)\mu}\right)} + \sqrt{\mu} S\right)^2 \\ &\leq 4\sigma_w^2 n \log\left(\frac{1}{\delta}\right) + 2\sigma_w^2 n(n+d)\log\left(1 + \frac{T_r {Z'}_{T_w}^2 + (T - T_r){Z''_T}^2}{(n+d)\mu}\right) + 2\mu S^2. \end{aligned}$$

■

Proof of Lemma 2.5.5. We will first show the desired bounds on $\lambda_{\min}(F_t)$ and $\lambda_{\max}(F_t)$. Recall that the event E_T holds with probability at least $1-4\delta$. Noting that $H_*^\intercal H_* = I + K_*^\intercal K_*$, it is clear that $F_t \succcurlyeq \beta_t^2 \lambda_{\min}(V_t^{-1}) H_*^\intercal H_* \succcurlyeq \frac{\beta_t^2}{\lambda_{\max}(V_t)} I$. Thus, from Lemma 2.D.4, for $T_r < t \leq T$, we have that $\lambda_{\min,t} \geq \frac{\beta_t^2}{\lambda_{\max}(V_t)} \geq \frac{\beta_t^2}{\mu + T_r Z'^2_{T_w} + (t-T_r) Z''^2_T}$.

On the other hand, $F_t \preccurlyeq \beta_t^2 \lambda_{\max}(V_t^{-1}) H_*^\intercal H_* \preccurlyeq \frac{\beta_t^2(1+\kappa^2)}{\lambda_{\min}(V_t)} I$. Again using Lemma 2.D.4, for $T_r < t \leq T$, we have that $\lambda_{\max,t} \leq \frac{(1+\kappa^2)\beta_t^2}{\lambda_{\min}(V_t)} \leq \frac{(1+\kappa^2)\beta_t^2}{\mu + T_w \frac{\sigma_w^2}{40}}$. Since $t \mapsto \beta_t$ is increasing, $t \mapsto \lambda_{\max,t}$ is increasing as well. The condition number $\kappa_t := \frac{\lambda_{\max,t}}{\lambda_{\min,t}} \leq \frac{\mu + T_r Z'^2_{T_w} + (t-T_r) Z''^2_T}{(1+\kappa^2)^{-1}(\mu + T_w \frac{\sigma_w^2}{40})}$ is increasing for $T_r < t \leq T$.

If $T_w = O(\sqrt{T}^{1+o(1)})$, then we have that $\lambda_{\max}(V_T) \leq O(\mathsf{poly}(n, d, \log(1/\delta)) T \log T)$ and $\beta_T(\delta) \leq O(\mathsf{poly}(n, d, \log(1/\delta)) \log T)$. Thus, there are positive constants $C = \mathsf{poly}(n, d, \log(1/\delta))$ and $c = \mathsf{poly}(n, d, \log(1/\delta))$ such that $\lambda_{\max,T} \leq C \frac{\log T}{T_w}$ and $\kappa_t = \frac{\lambda_{\max,T}}{\lambda_{\min,t}} \leq c \frac{T \log T}{T_w}$ for $T_r < t \leq T$ for large enough T. Choosing the larger between C and c yields the desired result. ■

Proof of Theorem 2.5.1

Defining by $p_t^{\mathsf{opt}} := \mathbb{P}\left\{\tilde{\Theta}_t \in \mathcal{S}^{\mathsf{opt}} \mid \mathcal{F}_t^{\mathsf{cnt}}, \hat{E}_t\right\}$ the optimistic probability, and by $\mathbb{P}_t\{\cdot\} := \mathbb{P}\{\cdot \mid \mathcal{F}_t^{\mathsf{cnt}}\}$ conditional probability measure, we can write

$$
\begin{aligned}
p_t^{\mathsf{opt}} &\geq \mathbb{P}\left\{\tilde{\Theta}_t \in \mathcal{S}^{\mathsf{surr}} \mid \mathcal{F}_t^{\mathsf{cnt}}, \hat{E}_t\right\} && (2.71)\\
&= \mathbb{P}\left\{L(\tilde{\Theta}_t^\intercal H_*) \leq L(\Theta_*^\intercal H_*) \mid \mathcal{F}_t^{\mathsf{cnt}}, \hat{E}_t\right\} && (2.72)\\
&\geq \min_{\hat{\Theta} \in \mathcal{E}_t^{\mathsf{RLS}}} \mathbb{P}_t\{L(\hat{\Theta}^\intercal H_* + \eta^\intercal \beta_t V_t^{-\frac{1}{2}} H_*) \leq L(\Theta_*^\intercal H_*)\} && (2.73)\\
&= \min_{\hat{\Theta} \in \mathcal{E}_t^{\mathsf{RLS}}} \mathbb{P}_t\{L(\hat{\Theta}^\intercal H_* + \Xi\sqrt{F_t}) \leq L(\Theta_*^\intercal H_*)\} && (2.74)
\end{aligned}
$$

where (2.71) is by Lemma 2.5.2, (2.73) is a worst-case estimation bound within high-probability confidence region, and (2.74) is because $\eta^\intercal \beta_t V_t^{-\frac{1}{2}} H_*$ and $\Xi\sqrt{F_t}$ have the same distributions with $\eta \in \mathbb{R}^{(n+d)\times n}$ and $\Xi \in \mathbb{R}^{n \times n}$ being i.i.d. standard normal random matrices.

The bound in (2.74) can be further lower bounded by minimizing over a larger confidence set as

$$p_t^{\mathsf{opt}} \geq \min_{\hat{\Theta}\in\mathcal{E}_t^{\mathrm{cl}}} \mathbb{P}_t\{L(\hat{\Theta}^\intercal H_* + \Xi\sqrt{F_t}) \leq L(A_{c,*})\} \tag{2.75}$$

$$= \min_{\hat{\Upsilon}\,:\,\|\hat{\Upsilon}\|_F\leq 1} \mathbb{P}_t\{L(A_{c,*} + (\Xi+\hat{\Upsilon})\sqrt{F_t}) \leq L(A_{c,*})\}, \tag{2.76}$$

where (2.75) is by Lemma (2.5.3) and (2.76) is because H_* is full column rank and therefore we can minimize over closed-loop matrices instead of open-loop system parameters.

Denoting by $G_t = (\Xi + \hat{\Upsilon})\sqrt{F_t}$ the perturbation due to estimation and sampling, Lemma 2.5.4 suggests that there exists constants $\epsilon_* > 0$ and $r_* > 0$ such that

$$L(A_{c,*} + G_t) = L(A_{c,*}) + \nabla L_* \bullet G_t + \frac{1}{2} G_t \bullet \mathcal{H}_{A_{c,*}+sG_t}(G_t) \tag{2.77}$$

$$\leq L(A_{c,*}) + \nabla L_* \bullet G_t + \frac{r_*}{2}\|G_t\|_F^2 \tag{2.78}$$

whenever $\|G_t\|_F \leq \epsilon_*$. Substituting (2.78) into (2.76) leads to the following lower bound

$$p_t^{\mathsf{opt}} \geq \min_{\hat{\Upsilon}\,:\,\|\hat{\Upsilon}\|_F\leq 1} \mathbb{P}_t\{L(A_{c,*}+G_t) \leq L(A_{c,*})\} \tag{2.79}$$

$$\geq \min_{\hat{\Upsilon}\,:\,\|\hat{\Upsilon}\|_F\leq 1} \mathbb{P}_t \left\{ \begin{array}{c} L(A_{c,*}) + \nabla L_* \bullet G_t + \frac{r_*}{2}\|G_t\|_F^2 \leq L(A_{c,*}), \\ \text{and} \quad \|G_t\|_F \leq \epsilon_* \end{array} \right\} \tag{2.80}$$

$$= \min_{\hat{\Upsilon}\,:\,\|\hat{\Upsilon}\|_F\leq 1} \mathbb{P}_t \left\{ \begin{array}{c} \frac{r_*}{2}\|(\Xi+\hat{\Upsilon})\sqrt{F_t}\|_F^2 + \nabla L_* \bullet (\Xi+\hat{\Upsilon})\sqrt{F_t} \leq 0, \\ \text{and} \quad \|(\Xi+\hat{\Upsilon})\sqrt{F_t}\|_F \leq \epsilon_* \end{array} \right\} \tag{2.81}$$

Noting that $\|(\Xi+\hat{\Upsilon})\sqrt{F_t}\|_F \leq \sqrt{\lambda_{\max,t}}\|\Xi+\hat{\Upsilon}\|_F$ where $\lambda_{\max,t} := \lambda_{\max}(F_t)$, we can further relax the lower bound (2.81) as

$$p_t^{\mathsf{opt}} \geq \min_{\hat{\Upsilon}\,:\,\|\hat{\Upsilon}\|_F\leq 1} \mathbb{P}_t \left\{ \begin{array}{c} \frac{\lambda_{\max,t} r_*}{2}\|\Xi+\hat{\Upsilon}\|_F^2 + (\nabla L_*\sqrt{F_t}) \bullet (\Xi+\hat{\Upsilon}) \leq 0, \\ \text{and} \quad \sqrt{\lambda_{\max,t}}\|\Xi+\hat{\Upsilon}\|_F \leq \epsilon_* \end{array} \right\} \tag{2.82}$$

$$= \min_{\hat{\Upsilon}\,:\,\|\hat{\Upsilon}\|_F\leq 1} \mathbb{P}_t \left\{ \begin{array}{c} \left\|\Xi+\hat{\Upsilon}+\frac{\nabla L_*\sqrt{F_t}}{\lambda_{\max,t} r_*}\right\|_F^2 \leq \left\|\frac{\nabla L_*\sqrt{F_t}}{\lambda_{\max,t} r_*}\right\|_F^2, \\ \text{and} \quad \|\Xi+\hat{\Upsilon}\|_F^2 \leq \frac{\epsilon_*^2}{\lambda_{\max,t}} \end{array} \right\} \tag{2.83}$$

where (2.83) is obtained by completion of squares. Let $\mathcal{U} : \mathbb{M}_n \to \mathbb{M}_n$ be an orthogonal transformation such that $\mathcal{U}\left(\hat{\Upsilon} + \frac{\nabla L_*\sqrt{F_t}}{\lambda_{\max,t} r_*}\right) = \left\|\hat{\Upsilon} + \frac{\nabla L_*\sqrt{F_t}}{\lambda_{\max,t} r_*}\right\|_F E_{11}$ where $E_{11} \in \mathbb{M}_n$ has 1 in its $(1,1)$ entry and zeros elsewhere. Since Frobenius norm and the probability density of Ξ are invariant under orthogonal transformations, (2.83) can be rewritten as

$$p_t^{\mathsf{opt}} \geq \min_{\hat{\Upsilon}\,:\,\|\hat{\Upsilon}\|_F \leq 1} \mathbb{P}_t \left\{ \begin{array}{c} \left\| \mathcal{U}\left(\Xi + \hat{\Upsilon} + \frac{\nabla L_* \sqrt{F_t}}{\lambda_{\max,t} r_*} \right) \right\|_F^2 \leq \left\| \frac{\nabla L_* \sqrt{F_t}}{\lambda_{\max,t} r_*} \right\|_F^2, \\ \text{and} \quad \|\mathcal{U}(\Xi + \hat{\Upsilon})\|_F^2 \leq \frac{\epsilon_*^2}{\lambda_{\max,t}} \end{array} \right\} \tag{2.84}$$

$$= \min_{\hat{\Upsilon}\,:\,\|\hat{\Upsilon}\|_F \leq 1} \mathbb{P}_t \left\{ \begin{array}{c} \left\| \Xi + \left\| \hat{\Upsilon} + \frac{\nabla L_* \sqrt{F_t}}{\lambda_{\max,t} r_*} \right\|_F E_{11} \right\|_F^2 \leq \left\| \frac{\nabla L_* \sqrt{F_t}}{\lambda_{\max,t} r_*} \right\|_F^2, \\ \text{and} \quad \|\Xi + \mathcal{U}(\hat{\Upsilon})\|_F^2 \leq \frac{\epsilon_*^2}{\lambda_{\max,t}} \end{array} \right\} \tag{2.85}$$

$$= \min_{\hat{\Upsilon}\,:\,\|\hat{\Upsilon}\|_F \leq 1} \mathbb{P}_t \left\{ \begin{array}{c} \left(\Xi_{11} + \left\| \hat{\Upsilon} + \frac{\nabla L_* \sqrt{F_t}}{\lambda_{\max,t} r_*} \right\|_F \right)^2 + \sum_{i,j \neq 1,1} \Xi_{ij}^2 \leq \left\| \frac{\nabla L_* \sqrt{F_t}}{\lambda_{\max,t} r_*} \right\|_F^2, \\ \text{and} \quad \|\Xi + \mathcal{U}(\hat{\Upsilon})\|_F^2 \leq \frac{\epsilon_*^2}{\lambda_{\max,t}} \end{array} \right\} \tag{2.86}$$

Notice that the probability in (2.86) is described by the intersection of two balls whose centers are far apart by $\frac{\|\nabla L_* \sqrt{F_t}\|_F}{\lambda_{\max,t} r_*}$ and hence the intersection has a fixed shape. Choosing $\hat{\Upsilon}$ along the direction of $\|\nabla L_* \sqrt{F_t}\|_F$ moves the center of the first ball furthest possible from the origin which leads to the intersection of the balls to move furthest away from the origin as well. Therefore, the probability in (2.86) attains its minimum at $\hat{\Upsilon}_\# := \frac{\nabla L_* \sqrt{F_t}}{\|\nabla L_* \sqrt{F_t}\|_F}$ and (2.86) can be equivalently expressed by

$$p_t^{\mathsf{opt}} \geq \mathbb{P}_t \left\{ \begin{array}{c} \left(\Xi_{11} + 1 + \frac{\|\nabla L_* \sqrt{F_t}\|_F}{\lambda_{\max,t} r_*} \right)^2 + \sum_{i,j \neq 1,1} \Xi_{ij}^2 \leq \frac{\|\nabla L_* \sqrt{F_t}\|_F^2}{\lambda_{\max,t}^2 r_*^2}, \\ \text{and} \quad \|\Xi + E_{11}\|_F^2 \leq \frac{\epsilon_*^2}{\lambda_{\max,t}} \end{array} \right\} \tag{2.87}$$

$$= \mathbb{P}_t \left\{ \begin{array}{c} \left(\xi + 1 + \frac{\|\nabla L_* \sqrt{F_t}\|_F}{\lambda_{\max,t} r_*} \right)^2 + X \leq \frac{\|\nabla L_* \sqrt{F_t}\|_F^2}{\lambda_{\max,t}^2 r_*^2}, \\ \text{and} \quad (\xi + 1)^2 + X \leq \frac{\epsilon_*^2}{\lambda_{\max,t}} \end{array} \right\} \tag{2.88}$$

where $\xi \sim \mathcal{N}(0,1)$ and $X \sim \chi^2_{n^2-1}$ are independent normal and chi-squared random variables, respectively. Denoting by $a_t := \frac{\|\nabla L_* \sqrt{F_t}\|_F}{\lambda_{\max,t} r_*}$ and $b_t = \frac{\epsilon_*}{\sqrt{\lambda_{\max,t}}}$ the radii of the balls, we can rewrite (2.88) as

$$p_t^{\mathsf{opt}} \geq \mathbb{P}_t\left\{(\xi+1+a_t)^2 + X \leq a_t^2, \text{ and } (\xi+1)^2 + X \leq b_t^2\right\} \tag{2.89}$$

$$= \mathbb{P}_t\left\{\begin{array}{c} |\xi+1+a_t| \leq \sqrt{a_t^2 - X}, \text{ and } |\xi+1| \leq \sqrt{b_t^2 - X}, \\ \text{and } X \leq \min(a_t^2, b_t^2) \end{array}\right\} \tag{2.90}$$

$$= \int_0^{\min(a_t^2, b_t^2)} \mathbb{P}_t\left\{|\xi+1+a_t| \leq \sqrt{a_t^2 - x}, \text{ and } |\xi+1| \leq \sqrt{b_t^2 - x}\right\} f_{n^2-1}(x)\mathrm{d}x \tag{2.91}$$

$$= \int_0^{\min(a_t^2, b_t^2)} \mathbb{P}_t\left\{\begin{array}{c} 1+a_t - \sqrt{a_t^2 - x} \leq \xi \leq 1 + a_t + \sqrt{a_t^2 - x}, \\ \text{and } 1 - \sqrt{b_t^2 - x} \leq \xi \leq 1 + \sqrt{b_t^2 - x}, \end{array}\right\} f_{n^2-1}(x)\mathrm{d}x \tag{2.92}$$

where $f_k(x) := \left(2^{\frac{k}{2}}\Gamma(\frac{k}{2})\right)^{-1} x^{\frac{k}{2}-1}e^{-\frac{x}{2}}$ is the probability density function of the chi-squared distribution with $k \in \mathbb{N}$ degrees of freedom. (2.91) is derived from law of total probability. Notice that the probability inside the integral in (2.92) is determined by the intersection of two intervals. This probability will have a non-zero value only for a fixed interval of x depending on the relation between a_t and b_t. We will investigate three cases:

i. $\boldsymbol{0 \leq b_t \leq \sqrt{2}a_t}$ **:** There is a non-empty intersection if and only if $0 \leq x \leq b_t^2\left(1 - \frac{b_t^2}{4a_t^2}\right)$ and the integral (2.92) becomes

$$p_t^{\mathsf{opt}} \geq \int_0^{b_t^2\left(1-\frac{b_t^2}{4a_t^2}\right)} \mathbb{P}_t\left\{1 + a_t - \sqrt{a_t^2 - x} \leq \xi \leq 1 + \sqrt{b_t^2 - x}\right\} f_{n^2-1}(x)\mathrm{d}x \tag{2.93}$$

$$= \int_0^{b_t^2\left(1-\frac{b_t^2}{4a_t^2}\right)} \left[Q\left(1 + a_t - \sqrt{a_t^2 - x}\right) - Q\left(1 + \sqrt{b_t^2 - x}\right)\right] f_{n^2-1}(x)\mathrm{d}x \tag{2.94}$$

where Q is the Gaussian Q-function. Notice that for fixed values of b_t, (2.94) is monotonically increasing with respect to a_t and vice versa.

ii. $\boldsymbol{\sqrt{2}a_t \leq b_t \leq 2a_t}$ **:** There is a non-empty intersection if and only if $0 \leq x \leq a_t^2$ and the integral (2.92) becomes

$$p_t^{\mathsf{opt}} \geq \int_0^{a_t^2} \mathbb{P}_t \left\{ 1 + a_t - \sqrt{a_t^2 - x} \leq \xi \leq 1 + \sqrt{b_t^2 - x} \right\} f_{n^2-1}(x)\mathrm{d}x \tag{2.95}$$

$$= \int_0^{a_t^2} \left[Q\left(1 + a_t - \sqrt{a_t^2 - x}\right) - Q\left(1 + \sqrt{b_t^2 - x}\right) \right] f_{n^2-1}(x)\mathrm{d}x \tag{2.96}$$

Notice that for fixed values of b_t, (2.96) is monotonically increasing with respect to a_t and vice versa.

iii. $\boldsymbol{2a_t \leq b_t}$ **:** There is a non-empty intersection if and only if $0 \leq x \leq a_t^2$ and the integral (2.92) becomes

$$p_t^{\mathsf{opt}} \geq \int_0^{a_t^2} \mathbb{P}_t \left\{ 1 + a_t - \sqrt{a_t^2 - x} \leq \xi \leq 1 + a_t + \sqrt{a_t^2 - x} \right\} f_{n^2-1}(x)\mathrm{d}x \tag{2.97}$$

$$= \int_0^{a_t^2} \left[Q\left(1 + a_t - \sqrt{a_t^2 - x}\right) - Q\left(1 + a_t + \sqrt{a_t^2 - x}\right) \right] f_{n^2-1}(x)\mathrm{d}x \tag{2.98}$$

Notice that for fixed values of b_t, (2.98) is monotonically increasing with respect to a_t and vice versa.

As seen from all three case, the integral in (2.92) is monotonically increasing with respect to both a_t, and b_t regardless of their relative relation. Therefore, we will consider tight lower bounds of $a_t = \frac{\|\nabla L_* \sqrt{F_t}\|_F}{\lambda_{\max,t} r_*}$ so that the relation $b_t \geq 2a_t$ holds for large enough $t \geq 0$. Noting that $\nabla L_* = 2P_* A_{c,*} \Sigma_*$ by Lemma 2.5.4 and $P_* \succ 0$, $\Sigma_* \succ 0$, we will consider two cases.

1. Singular $\boldsymbol{A_{c,*}}$ **:** In this case, the Jacobian matrix ∇L_* becomes singular as well. Then, we can bound a_t from below as $a_t = \frac{\|\nabla L_* \sqrt{F_t}\|_F}{\lambda_{\max,t} r_*} \geq \sqrt{\lambda_{\min,t}} \frac{\|\nabla L_*\|_F}{\lambda_{\max,t} r_*} = \sqrt{\frac{\lambda_{\min,t}}{\lambda_{\max,t}}} \frac{\|r_*^{-1} \nabla L_*\|_F}{\sqrt{\lambda_{\max,t}}}$. Furthermore, choosing $T_w = O((\sqrt{T})^{1+o(1)})$, we can use upper bounds for $\frac{\lambda_{\max,t}}{\lambda_{\min,t}}$ and $\lambda_{\max,t}$ from Lemma 2.5.5 to write down, $a_t \geq \frac{T_w}{\sqrt{T}\log T} \frac{\|\nabla L_*\|_F}{C r_*} =: a_{1,T}$ and $b_t \geq \sqrt{\frac{T_w}{\log T}} \frac{\epsilon_*}{\sqrt{C}} =: b_{1,T}$ for all $T_r < t \leq T$ under the event E_T for large enough T. Therefore, replacing a_t and b_t with $a_{1,T}$ and $b_{1,T}$ in (2.92) gives a lower bound to (2.92). Noting that the ratio $\frac{b_{1,T}}{a_{1,T}} = \sqrt{\frac{T \log T}{T_w}} \frac{\epsilon_* r_* \sqrt{C}}{\|\nabla L_*\|_F}$ can be made to be greater than or equal to 2 by an appropriate choice of T_w leading to the case (iii) bound

$$p_t^{\mathsf{opt}} \geq \int_0^{a_{1,T}^2} \left[Q\left(1 + a_{1,T} - \sqrt{a_{1,T}^2 - x}\right) - Q\left(1 + a_{1,T} + \sqrt{a_{1,T}^2 - x}\right) \right] f_{n^2-1}(x)\mathrm{d}x \tag{2.99}$$

for all $T_r < t \leq T$ for large enough T.

2. Nonsingular $\boldsymbol{A_{c,*}}$: In this case, the Jacobian matrix ∇L_* becomes nonsingular as well. Then, we can bound a_t from below as $a_t = \frac{\|\nabla L_* \sqrt{F_t}\|_F}{\lambda_{\max,t} r_*} \geq \sigma_{\min,*} \frac{\|\sqrt{F_t}\|_F}{\lambda_{\max,t} r_*} \geq \frac{\sigma_{\min,*}}{r_* \sqrt{\lambda_{\max,t}}}$. Choosing $T_w = O((\sqrt{T})^{1+o(1)})$, we can use the upper bound for $\lambda_{\max,t}$ from Lemma 2.5.5 to write the lower bound, $a_t \geq \sqrt{\frac{T_w}{\log T}} \frac{\min(\sigma_{\min,*}, \epsilon_* r_*/2)}{\sqrt{C} r_*} =: a_{2,T}$ and $b_t \geq \sqrt{\frac{T_w}{\log T}} \frac{\epsilon_*}{\sqrt{C}} =: b_{2,T}$ for all $T_r < t \leq T$ under the event E_T for large enough T. Therefore, replacing a_t and b_t with $a_{2,T}$ and $b_{2,T}$ in (2.92) gives a lower bound to (2.92) for $T_r < t \leq T$. Noting that the ratio $\frac{\beta_{2,T}}{a_{2,T}} = \frac{\epsilon_* r_*}{\min(\sigma_{\min,*}, \epsilon_* r_*/2)} = \max\left(\frac{\epsilon_* r_*}{\sigma_{\min,*}}, 2\right) \geq 2$, we can use the case (iii) bound

$$p_t^{\mathsf{opt}} \geq \int_0^{a_{2,T}^2} \left[Q\left(1 + a_{2,T} - \sqrt{a_{2,T}^2 - x}\right) - Q\left(1 + a_{2,T} + \sqrt{a_{2,T}^2 - x}\right) \right] f_{n^2-1}(x)\mathrm{d}x \tag{2.100}$$

for all $T_r < t \leq T$ for large enough T.

In both cases, our focus will be on the following probability with a parameters $a > 0$, and $k \in \mathbb{N}$

$$p_k(a) := \int_0^{a^2} \left[Q(1 + a - \sqrt{a^2 - x}) - Q(1 + a + \sqrt{a^2 - x}) \right] f_k(x)\mathrm{d}x \tag{2.101}$$

The following lemma summarizes some of the important properties of the function $a \mapsto p_k(a)$.

Lemma 2.D.6. *The non-negative real valued function $a \mapsto p_k(a)$ is monotonically increasing with respect to $a \geq 0$. Furthermore, we have that $\frac{1}{p_k(a)} \leq \frac{1}{Q(1)}\left(1 + \frac{Ck}{a^{1/2}}\right)$ for $a \geq ck$ for problem independent constants $c, C > 0$.*

Proof. Notice that for a fixed value of $0 \leq x \leq a^2$, the functions $a \mapsto 1+a-\sqrt{a^2 - x}$ and $a \mapsto 1 + a + \sqrt{a^2 - x}$ are monotonically decreasing and monotonically increasing, respectively. As Q-function is monotonically decreasing, the function

$a \mapsto Q(1+a-\sqrt{a^2-x}) - Q(1+a+\sqrt{a^2-x})$ is monotonically increasing for fixed $0 \leq x \leq a^2$. Therefore, the function $a \mapsto p_k(a)$ is also monotonically increasing.

In order to obtain the desired asymptotic bound, let $\epsilon \in (0,1)$ and we can write

$$\begin{aligned}
p_k(a) &= \int_0^{a^2} \left[Q(1+a-\sqrt{a^2-x}) - Q(1+a+\sqrt{a^2-x})\right] f_k(x)\mathrm{d}x \\
&\geq \int_0^{\epsilon a^2} \left[Q(1+a-\sqrt{a^2-x}) - Q(1+a+\sqrt{a^2-x})\right] f_k(x)\mathrm{d}x \\
&\geq \int_0^{\epsilon a^2} \min_{0\leq x' \leq \epsilon a^2} \left[Q(1+a-\sqrt{a^2-x'}) - Q(1+a+\sqrt{a^2-x'})\right] f_k(x)\mathrm{d}x \\
&= \left[Q(1+a(1-\sqrt{1-\epsilon})) - Q(1+a(1+\sqrt{1-\epsilon}))\right] F_k(\epsilon a^2)
\end{aligned}$$

where $F_k(x) \coloneqq 1 - \frac{\Gamma(k/2,\, x/2)}{\Gamma(k/2)}$ is the cumulative distribution function of chi-square distribution and $(s,x) \mapsto \Gamma(s,x) \coloneqq \int_x^\infty t^{s-1}e^{-t}\mathrm{d}t$ and $s \mapsto \Gamma(s) \coloneqq \int_o^\infty t^{s-1}e^{-t}\mathrm{d}t$ are upper incomplete Gamma and ordinary Gamma functions respectively. Notice that the functions $(s,x) \mapsto \Gamma(s,x)$ and $x \mapsto Q(x)$ are monotonically decreasing with increasing $x > 0$. Therefore, for large enough $\epsilon a^2 \gg 1$ and large enough $a \gg 1$, we can claim that $\Gamma(k/2,\, \epsilon a^2/2) \ll 1$ and $Q(1+a) \ll 1$ are small enough. Furthermore, for small enough $\epsilon \ll 1$, we can use Taylor expansion to see that $1-\sqrt{1-\epsilon} = \frac{\epsilon}{2}\sum_{k=0}^\infty \frac{\epsilon^k}{2^k}(2k-1)! \leq c_1\epsilon$ for a problem independent constant $c_1 > 0$. Then, for small enough $\epsilon \ll 1$, we have that

$$\begin{aligned}
p_k(a) &\geq \left[Q(1+a(1-\sqrt{1-\epsilon})) - Q(1+a(1+\sqrt{1-\epsilon}))\right] \left(1 - \frac{\Gamma(k/2,\, \epsilon a^2/2)}{\Gamma(k/2)}\right) \\
&\geq \left[Q(1+c_1\epsilon a) - Q(1+a)\right] \left(1 - \frac{\Gamma(k/2,\, \epsilon a^2/2)}{\Gamma(k/2)}\right)
\end{aligned}$$

Furthermore, for small enough $\epsilon a \ll 1$, we have that $Q(1+c_1\epsilon a) \geq Q(1) - c_2\epsilon a$ by Taylor's theorem where c_2 is a problem independent constant. Using these bounds, we can bound the inverse of $p_k(a)$ from above for small enough $\epsilon \ll 1$, small enough $\epsilon a \ll 1$, large enough $a \gg 1$ and large enough $\epsilon a^2 \gg 1$ as

$$\begin{aligned}
\frac{1}{p_k(a)} &\leq \frac{1}{Q(1) - c_2\epsilon a - Q(1+a)} \frac{1}{1 - \frac{\Gamma(k/2,\, \epsilon a^2/2)}{\Gamma(k/2)}} \\
&= \frac{1}{Q(1)} \frac{1}{(1 - c_2\epsilon a - Q(1+a))\left(1 - \frac{\Gamma(k/2,\, \epsilon a^2/2)}{\Gamma(k/2)}\right)} \\
&\leq \frac{1}{Q(1)}\left[1 + 2C\left(c_2\epsilon a + Q(1+a) + \frac{\Gamma(k/2,\, \epsilon a^2/2)}{\Gamma(k/2)}\right)\right] \qquad (2.102)
\end{aligned}$$

where we used the Taylor expansion $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \leq 1 + Cx$ for small enough $x \ll 1$ with $C > 0$ being a problem independent constant.

The assumption $\epsilon a^2 \gg 1$ can be used to write the asymptotic expansion of incomplete Gamma function $\Gamma(k/2, \epsilon a^2/2) = (\epsilon a^2/2)^{k/2-1} e^{-\epsilon a^2/2} \left[1 + O\left((\epsilon a^2/2)^{-1}\right)\right]$. Noting that the Q function is always bounded as $Q(1+a) \leq \frac{e^{-\frac{(1+a)^2}{2}}}{\sqrt{2\pi}(1+a)}$, we claim that choosing $\epsilon = \frac{k}{2ea^{1+1/2}}$, for $\alpha \geq c''k$ with a constant $c'' > 0$ guarantees that $\epsilon a = \frac{k}{2e} a^{-1/2} \ll 1$ and $\epsilon a^2 = \frac{k}{2e} a^{1-1/2} \gg 1$. Therefore, the upper bound (2.102) is valid for $\alpha \geq c''k$. Furthermore, the term ϵa decays slower than both $Q(1+a)$ and $\frac{\Gamma(k/2, \epsilon a^2/2)}{\Gamma(k/2)}$ and thus ϵa dominates as

$$\frac{1}{p_k(a)} \leq \frac{1}{Q(1)} \left(1 + \frac{Ck}{2e} a^{-1/2}\right)$$

for a problem independent constant $C > 0$. ■

Based on Lemma 2.D.6, the integrals in (2.99) and (2.100) are asymptotically constant if both $a_{1,T}$ and $a_{2,T}$ are asymptotically large enough. This can be achieved if $a_{1,T} = \frac{T_w}{\sqrt{T}\log T} \frac{\|\nabla L_*\|_F}{Cr_*} = \omega(1)$ for singular $A_{c,*}$ and $a_{2,T} = \sqrt{\frac{T_w}{\log T}} \frac{\min(\sigma_{\min,*}, \epsilon_* r_*/2)}{\sqrt{C}r_*} = \omega(1)$ for non-singular $A_{c,*}$. In other words, choosing $T_w = n^2\omega(\sqrt{T}\log T)$ for singular $A_{c,*}$ and $T_w = n^2\omega(\log T)$ for non-singular $A_{c,*}$ yields the desired bound

$$p_t^{\text{opt}} \geq \frac{Q(1)}{1 + o(1)}.$$

for $T_r < t \leq T$ for large enough T. Combined with the upper $T_w = O((\sqrt{T})^{1+o(1)})$, the proposed choices of T_w satisfy the asymptotic conditions.

2.E Regret Decomposition

Denote the optimal expected average cost of an LQR system Θ with process noise covariance W by $J_*(\Theta, W) = \mathrm{tr}(P(\Theta)W)$. Note that during the initial exploration period, we have that $u_t = \bar{u}_t + \nu_t$ for $t \leq T_w$ and after the initial exploration, we have that $u_t = \bar{u}_t$ for $t > T_w$ where we denote by $\bar{u}_t := K(\tilde{\Theta}_t)x_t$ the optimal control action assuming the system $\tilde{\Theta}_t$. Since initial exploration period injects independent random perturbations through the optimal control input, $\bar{u}_t$, for sampled system, $\tilde{\Theta}_t$, the state dynamics can be reformulated in order to take the external perturbations into account by adding it to the process noise:

$$x_{t+1} = A_* x_t + B_* \bar{u}_t + \zeta_t, \tag{2.103}$$

where $\bar{u}_t = K(\tilde{\Theta}_t)x_t$, $\zeta_t = B_*\nu_t + w_t$ for $t \leq T_w$, and $\zeta_t = w_t$ for $t > T_w$. We can write the regret explicitly as

$$R_T = \sum\nolimits_{t=0}^{T} \left\{ x_t^\intercal Q x_t + u_t^\intercal R u_t - J_*(\Theta_*, \sigma_w^2 I) \right\} = R_{T_w}^{\text{exp}} + R_T^{\text{noexp}}, \tag{2.104}$$

where

$$R_{T_w}^{\text{exp}} := \sum\nolimits_{t=0}^{T_w} \left(2\bar{u}_t^\intercal R \nu_t + \nu_t^\intercal R \nu_t \right), \quad \& \quad R_T^{\text{noexp}} := \sum\nolimits_{t=0}^{T} \left\{ x_t^\intercal Q x_t + \bar{u}_t^\intercal R \bar{u}_t - J_*(\Theta_*, \sigma_w^2 I) \right\}$$

Since $E_s \subset E_t$ for any $0 \leq s \leq t$, we have that

$$\begin{aligned} R_T^{\text{noexp}} \mathbb{1}_{E_T} &= \sum_{t=0}^{T} \left\{ x_t^\intercal Q x_t + \bar{u}_t^\intercal R \bar{u}_t - J_*(\Theta_*, \sigma_w^2 I) \right\} \mathbb{1}_{E_T} \\ &\leq \sum_{t=0}^{T} \left\{ x_t^\intercal Q x_t + \bar{u}_t^\intercal R \bar{u}_t - J_*(\Theta_*, \sigma_w^2 I) \right\} \mathbb{1}_{E_t}, \end{aligned} \tag{2.105}$$

$$R_{T_w}^{\text{exp}} \mathbb{1}_{E_T} = \sum\nolimits_{t=0}^{T_w} \left(2\bar{u}_t^\intercal R \nu_t + \nu_t^\intercal R \nu_t \right) \mathbb{1}_{E_T} \leq \sum\nolimits_{t=0}^{T_w} \left(2\bar{u}_t^\intercal R \nu_t + \nu_t^\intercal R \nu_t \right) \mathbb{1}_{E_t}. \tag{2.106}$$

From Bellman optimality equations [19], we obtain

$$\begin{aligned} &J_*(\tilde{\Theta}_t, \mathrm{Cov}[\zeta_t]) + x_t^\intercal P(\tilde{\Theta}_t) x_t \\ &\quad = \min_u \left\{ x_t^\intercal Q x_t + u^\intercal R u + \mathbb{E}\left[(\tilde{A}_t x_t + \tilde{B}_t u + \zeta_t)^\intercal P(\tilde{\Theta}_t)(\tilde{A}_t x_t + \tilde{B}_t u + \zeta_t) \,\middle|\, \mathcal{F}_t \right] \right\}, \\ &\quad = x_t^\intercal Q x_t + \bar{u}_t^\intercal R \bar{u}_t + \mathbb{E}\left[(\tilde{A}_t x_t + \tilde{B}_t \bar{u}_t + \zeta_t)^\intercal P(\tilde{\Theta}_t)(\tilde{A}_t x_t + \tilde{B}_t \bar{u}_t + \zeta_t) \,\middle|\, \mathcal{F}_t \right], \\ &\quad = x_t^\intercal Q x_t + \bar{u}_t^\intercal R \bar{u}_t + \mathbb{E}\left[(\tilde{A}_t x_t + \tilde{B}_t \bar{u}_t)^\intercal P(\tilde{\Theta}_t)(\tilde{A}_t x_t + \tilde{B}_t \bar{u}_t) \,\middle|\, \mathcal{F}_t \right] + \mathbb{E}\left[\zeta_t^\intercal P(\tilde{\Theta}_t) \zeta_t \,\middle|\, \mathcal{F}_t \right], \\ &\quad = x_t^\intercal Q x_t + \bar{u}_t^\intercal R \bar{u}_t + \mathbb{E}\left[(\tilde{A}_t x_t + \tilde{B}_t \bar{u}_t)^\intercal P(\tilde{\Theta}_t)(\tilde{A}_t x_t + \tilde{B}_t \bar{u}_t) \,\middle|\, \mathcal{F}_t \right] \\ &\qquad + \mathbb{E}\left[x_{t+1}^\intercal P(\tilde{\Theta}_t) x_{t+1} \,\middle|\, \mathcal{F}_t \right] - \mathbb{E}\left[(A_* x_t + B_* \bar{u}_t)^\intercal P(\tilde{\Theta}_t)(A_* x_t + B_* \bar{u}_t) \,\middle|\, \mathcal{F}_t \right], \\ &\quad = x_t^\intercal Q x_t + \bar{u}_t^\intercal R \bar{u}_t + \mathbb{E}\left[x_{t+1}^\intercal P(\tilde{\Theta}_t) x_{t+1} \,\middle|\, \mathcal{F}_t \right] \\ &\qquad + (\tilde{A}_t x_t + \tilde{B}_t \bar{u}_t)^\intercal P(\tilde{\Theta}_t)(\tilde{A}_t x_t + \tilde{B}_t \bar{u}_t) - (A_* x_t + B_* \bar{u}_t)^\intercal P(\tilde{\Theta}_t)(A_* x_t + B_* \bar{u}_t), \\ &\quad = x_t^\intercal Q x_t + \bar{u}_t^\intercal R \bar{u}_t + \mathbb{E}\left[x_{t+1}^\intercal P(\tilde{\Theta}_t) x_{t+1} \,\middle|\, \mathcal{F}_t \right] + \bar{z}_t^\intercal \tilde{\Theta}_t P(\tilde{\Theta}_t) \tilde{\Theta}_t^\intercal \bar{z}_t - \bar{z}_t^\intercal \Theta_* P(\tilde{\Theta}_t) \Theta_*^\intercal \bar{z}_t, \end{aligned}$$

where $\bar{z}_t^\intercal = [x_t^\intercal, \bar{u}_t^\intercal]$. Rearranging the terms and subtracting the optimal expected average cost of the true system, we obtain the following for each term in (2.105),

$$\begin{aligned} &\left\{ x_t^\intercal Q x_t + \bar{u}_t^\intercal R \bar{u}_t - J_*(\Theta_*, \sigma_w^2 I) \right\} \mathbb{1}_{E_t} \\ &= \left\{ J_*(\tilde{\Theta}_t, \mathrm{Cov}[\zeta_t]) - J_*(\Theta_*, \sigma_w^2 I) \right\} \mathbb{1}_{E_t} + \left\{ \bar{z}_t^\intercal \Theta_* P(\tilde{\Theta}_t) \Theta_*^\intercal \bar{z}_t - \bar{z}_t^\intercal \tilde{\Theta}_t P(\tilde{\Theta}_t) \tilde{\Theta}_t^\intercal \bar{z}_t \right\} \mathbb{1}_{E_t}, \\ &\quad + x_t^\intercal P(\tilde{\Theta}_t) x_t \mathbb{1}_{E_t} - \mathbb{E}\left[x_{t+1}^\intercal P(\tilde{\Theta}_t) x_{t+1} \mathbb{1}_{E_t} \,\middle|\, \mathcal{F}_t \right]. \end{aligned}$$

Note that, $\mathbb{1}_{E_t}\mathbb{1}_{E_{t+1}} = \mathbb{1}_{E_{t+1}}$ since $E_{t+1} \subset E_t$. Since $P(\tilde{\Theta}_t) \succ 0$, we obtain

$$\begin{aligned}
&\mathbb{E}\left[x_{t+1}^{\intercal} P(\tilde{\Theta}_t) x_{t+1}\mathbb{1}_{E_t} \,\middle|\, \mathcal{F}_t\right] \\
&= \mathbb{E}\left[x_{t+1}^{\intercal} P(\tilde{\Theta}_t) x_{t+1}\mathbb{1}_{E_t}\left(\mathbb{1}_{E_{t+1}} + \mathbb{1}_{E_{t+1}^c}\right) \,\middle|\, \mathcal{F}_t\right], \\
&= \mathbb{E}\left[x_{t+1}^{\intercal} P(\tilde{\Theta}_t) x_{t+1}\mathbb{1}_{E_{t+1}} \,\middle|\, \mathcal{F}_t\right] + \mathbb{E}\left[x_{t+1}^{\intercal} P(\tilde{\Theta}_t) x_{t+1}\mathbb{1}_{E_t}\mathbb{1}_{E_{t+1}^c} \,\middle|\, \mathcal{F}_t\right], \\
&\geq \mathbb{E}\left[x_{t+1}^{\intercal} P(\tilde{\Theta}_t) x_{t+1}\mathbb{1}_{E_{t+1}} \,\middle|\, \mathcal{F}_t\right], \\
&= \mathbb{E}\left[x_{t+1}^{\intercal}\left(P(\tilde{\Theta}_t) - P(\tilde{\Theta}_{t+1})\right) x_{t+1}\mathbb{1}_{E_{t+1}} \,\middle|\, \mathcal{F}_t\right] + \mathbb{E}\left[x_{t+1}^{\intercal} P(\tilde{\Theta}_{t+1}) x_{t+1}\mathbb{1}_{E_{t+1}} \,\middle|\, \mathcal{F}_t\right].
\end{aligned}$$

Therefore,

$$\begin{aligned}
\left\{x_t^{\intercal} Q x_t + \bar{u}_t^{\intercal} R \bar{u}_t - J_*(\Theta_*, \sigma_w^2 I)\right\}\mathbb{1}_{E_t} \leq &\left\{J_*(\tilde{\Theta}_t, \mathrm{Cov}[\zeta_t]) - J_*(\Theta_*, \sigma_w^2 I)\right\}\mathbb{1}_{E_t}, \\
&+ \left\{\bar{z}_t^{\intercal}\Theta_* P(\tilde{\Theta}_t)\Theta_*^{\intercal}\bar{z}_t - \bar{z}_t^{\intercal}\tilde{\Theta}_t P(\tilde{\Theta}_t)\tilde{\Theta}_t^{\intercal}\bar{z}_t\right\}\mathbb{1}_{E_t}, \\
&+ \left\{x_t^{\intercal} P(\tilde{\Theta}_t) x_t\mathbb{1}_{E_t} - \mathbb{E}\left[x_{t+1}^{\intercal} P(\tilde{\Theta}_{t+1}) x_{t+1}\mathbb{1}_{E_{t+1}} \,\middle|\, \mathcal{F}_t\right]\right\}, \\
&+ \mathbb{E}\left[x_{t+1}^{\intercal}\left(P(\tilde{\Theta}_{t+1}) - P(\tilde{\Theta}_t)\right) x_{t+1}\mathbb{1}_{E_{t+1}} \,\middle|\, \mathcal{F}_t\right]
\end{aligned} \tag{2.107}$$

Notice that $\mathrm{Cov}[\zeta_t] = \sigma_\nu^2 B_* B_*^{\intercal} + \sigma_w^2 I$ for $t \leq T_w$ and $\mathrm{Cov}[\zeta_t] = \sigma_w^2 I$ for $t > T_w$ and therefore

$$J_*(\tilde{\Theta}_t, \mathrm{Cov}[\zeta_t]) = \mathrm{Tr}(P(\tilde{\Theta}_t)\,\mathrm{Cov}[\zeta_t]) = \begin{cases} \sigma_\nu^2\,\mathrm{Tr}(P(\tilde{\Theta}_t) B_* B_*^{\intercal}) + \sigma_w^2\,\mathrm{Tr}(P(\tilde{\Theta}_t)) & t \leq T_w \\ \sigma_w^2\,\mathrm{Tr}(P(\tilde{\Theta}_t)) & t > T_w \end{cases} \tag{2.108}$$

Summing the terms in (2.107) upto time T and adding the $R_{T_w}^{\text{exp}}$ term, we obtain

$$R_T\mathbb{1}_{E_T} = R_{T_w}^{\text{exp}}\mathbb{1}_{E_T} + R_T^{\text{noexp}}\mathbb{1}_{E_T} \leq R_{T_w}^{\text{exp,1}} + R_{T_w}^{\text{exp,2}} + R_T^{\text{TS}} + R_T^{\text{RLS}} + R_T^{\text{mart}} + R_T^{\text{gap}} \tag{2.109}$$

where

$$R_{T_w}^{\text{exp,1}} = \sum\nolimits_{t=0}^{T_w}(2\bar{u}_t^{\intercal} R\nu_t + \nu_t^{\intercal} R\nu_t)\mathbb{1}_{E_t}, \tag{2.110}$$

$$R_{T_w}^{\text{exp,2}} = \sum\nolimits_{t=0}^{T_w}\sigma_\nu^2\,\mathrm{Tr}(P(\tilde{\Theta}_t) B_* B_*^{\intercal})\mathbb{1}_{E_t}, \tag{2.111}$$

$$R_T^{\text{TS}} = \sum\nolimits_{t=0}^{T}\left\{J_*(\tilde{\Theta}_t, \sigma_w^2 I) - J_*(\Theta_*, \sigma_w^2 I)\right\}\mathbb{1}_{E_t}, \tag{2.112}$$

$$R_T^{\text{RLS}} = \sum\nolimits_{t=0}^{T}\left\{\bar{z}_t^{\intercal}\Theta_* P(\tilde{\Theta}_t)\Theta_*^{\intercal}\bar{z}_t - \bar{z}_t^{\intercal}\tilde{\Theta}_t P(\tilde{\Theta}_t)\tilde{\Theta}_t^{\intercal}\bar{z}_t\right\}\mathbb{1}_{E_t}, \tag{2.113}$$

$$R_T^{\text{mart}} = \sum\nolimits_{t=0}^{T}\left\{x_t^{\intercal} P(\tilde{\Theta}_t) x_t\mathbb{1}_{E_t} - \mathbb{E}\left[x_{t+1}^{\intercal} P(\tilde{\Theta}_{t+1}) x_{t+1}\mathbb{1}_{E_{t+1}} \,\middle|\, \mathcal{F}_t\right]\right\}, \tag{2.114}$$

$$R_T^{\text{gap}} = \sum\nolimits_{t=0}^{T}\mathbb{E}\left[x_{t+1}^{\intercal}\left(P(\tilde{\Theta}_{t+1}) - P(\tilde{\Theta}_t)\right) x_{t+1}\mathbb{1}_{E_{t+1}} \,\middle|\, \mathcal{F}_t\right]. \tag{2.115}$$

In the next section, we will give upper bounds to each term.

2.F Regret Analysis

In this section, we bound each term in regret decomposition individually. In particular, $R^{\text{exp}}_{T_w}$ is studied in Appendix 2.F, R^{RLS}_{T} is studied in Appendix 2.F, R^{mart}_{T} in Appendix 2.F, R^{TS}_{T} in Appendix 2.F, and R^{gap}_{T} in Appendix 2.F. Finally, in Appendix 2.F, we combine these results to obtain the regret upper bound of TSAC as stated in Theorem 2.4.1.

Bounding $R^{\text{exp,1}}_{T_w}$ and $R^{\text{exp,2}}_{T_w}$

The following gives an upper bound on the regret attained due to isotropic perturbations in the TS with improved exploration phase of TSAC.

Lemma 2.F.1 (Direct Effect of Improved Exploration on Regret)**.** *The following holds with probability at least* $1-\delta$*,*

$$R^{\textit{exp,1}}_{T_w} = \sum_{t=0}^{T_w} \{2\bar{u}_t^\intercal R\nu_t + \nu_t^\intercal R\nu_t\}\, \mathbb{1}_{E_t} \leq d\sigma_\nu\sqrt{B_\delta} + d\|R\|\sigma_\nu^2\left(T_w + \sqrt{T_w}\log\frac{4dT_w}{\delta}\sqrt{\log\frac{4}{\delta}}\right)$$

where

$$B_\delta = 8\left(1+T_w\kappa^2\|R\|^2(n+d)^{2(n+d)}\right)\log\left(\frac{4d}{\delta}\left(1+T_w\kappa^2\|R\|^2(n+d)^{2(n+d)}\right)^{1/2}\right).$$

Furthermore, we have $R^{\textit{exp,2}}_{T_w} \leq \sigma_\nu^2 D\,\|B_*\|_F^2 T_w$.

Proof. First we will study $R^{\text{exp,1}}_{T_w}$. Let $q_t^\intercal = \bar{u}_t^\intercal R\,\mathbb{1}_{E_t}$. The first term can be written as

$$2\sum_{t=0}^{T_w}\sum_{i=1}^{d} q_{t,i}\nu_{t,i} = 2\sum_{i=1}^{d}\sum_{t=0}^{T_w} q_{t,i}\nu_{t,i}$$

Let $M_{t,i} = \sum_{k=0}^{t} q_{k,i}\nu_{k,i}$. By Theorem 2.G.1 on some event $G_{\delta,i}$ that holds with probability at least $1-\delta/(2d)$, for any $t\geq 0$,

$$M_{t,i}^2 \leq 2\sigma_\nu^2\left(1+\sum_{k=0}^{t} q_{k,i}^2\right)\log\left(\frac{2d}{\delta}\left(1+\sum_{k=0}^{t} q_{k,i}^2\right)^{1/2}\right)$$

Note that $\|q_k\| = \|R\bar{u}_t\|\,\mathbb{1}_{E_t} \leq \kappa\|R\|(n+d)^{n+d}$, thus $q_{k,i} \leq \kappa\|R\|(n+d)^{n+d}$. Using union bound we get, for probability at least $1-\frac{\delta}{2}$,

$$\sum_{t=0}^{T_w} 2\nu_t^\intercal R\bar{u}_t \mathbb{1}_{R_t} \leq$$

$$d\sqrt{8\sigma_\nu^2\left(1+T_w\kappa^2\|R\|^2(n+d)^{2(n+d)}\right)\log\left(\frac{4d}{\delta}\left(1+T_w\kappa^2\|R\|^2(n+d)^{2(n+d)}\right)^{1/2}\right)} \tag{2.116}$$

Let $W = \sigma_\nu\sqrt{2d\log\frac{4dT_w}{\delta}}$. Define $\Psi_t = \nu_t^\intercal R\nu_t - \mathbb{E}\left[\nu_t^\intercal R\nu_t|\mathcal{F}_{t-1}\right]$ and its truncated version $\tilde{\Psi}_t = \Psi_t\mathbb{I}_{\left\{\Psi_t\leq 2DW^2\right\}}$.

$$\mathbb{P}\left(\sum_{t=1}^{T_w}\Psi_t > 2\|R\|W^2\sqrt{2T_w\log\frac{4}{\delta}}\right) \leq$$
$$\mathbb{P}\left(\max_{1\leq t\leq T_w}\Psi_t > 2\|R\|W^2\right) + \mathbb{P}\left(\sum_{t=1}^{T_w}\tilde{\Psi}_t > 2\|R\|W^2\sqrt{2T_w\log\frac{4}{\delta}}\right)$$

Using Lemma 2.G.4 with union bound and Theorem 2.G.2, summation of terms on the right hand side is bounded by $\delta/2$. Thus, with probability at least $1-\delta/2$,

$$\sum_{t=0}^{T_w}\nu_t^\intercal R\nu_t \leq dT_w\sigma_\nu^2\|R\| + 2\|R\|W^2\sqrt{2T_w\log\frac{4}{\delta}}. \tag{2.117}$$

Combining (2.116) and (2.117) gives the statement of lemma for the regret of external exploration noise. Next, we consider $R_{T_w}^{\text{exp},2}$. Due to rejection sampling $\mathcal{R}_{\mathcal{S}}(\cdot)$, a new model sample is redrawn until it lies on the set $\mathcal{S}$ at every TS step, *i.e.*, $\tilde{\Theta}_t \in \mathcal{S}$ for every time step $t \geq 0$. By Assumption 2.2.1, we have $\|P(\tilde{\Theta}_t)\|_F \leq D = \bar{\alpha}\gamma^{-1}\kappa^2(1+\kappa^2)$. Thus, we have

$$\begin{aligned}
R_{T_w}^{\text{exp},2} &= \sum_{t=0}^{T_w}\sigma_\nu^2\operatorname{Tr}(P(\tilde{\Theta}_t)B_*B_*^\intercal)\mathbb{1}_{E_t},\\
&\leq \sum_{t=0}^{T_w}\sigma_\nu^2\|P(\tilde{\Theta}_t)\|_F\|B_*\|_F^2\mathbb{1}_{E_t},\\
&\leq \sigma_\nu^2 D\,\|B_*\|_F^2\sum_{t=0}^{T_w}\mathbb{1}_{E_t} \leq \sigma_\nu^2 D\,\|B_*\|_F^2 T_w.
\end{aligned} \tag{2.118}$$

■

Bounding $R_T^{\mathbf{RLS}}$

Bounding this term is achieved by manipulating the similar bounds in Abbasi-Yadkori and Szepesvári [2] and Abeille and Lazaric [3] to our setting and TS algorithm. We first have the following result from regularized least squares estimate.

Lemma 2.F.2. *On the event of E_T, for $X_s = \frac{(12\kappa^2+2\kappa\sqrt{2})\sigma_w}{\gamma}\sqrt{2n\log(n(T-T_w)/\delta)}$, we have,*

$$\sum_{t=0}^{T}\|(\Theta_* - \tilde{\Theta}_t)^\intercal z_t\|^2 \leq 2(\beta_T(\delta)+\upsilon_T(\delta))^2\Bigg(\left(1+\frac{(1+\kappa^2)(n+d)^{2(n+d)}}{\mu}\right)^{\tau_0+1}\log\frac{\det(V_{T_r})}{\det(\mu I)} + \left(1+\frac{(1+\kappa^2)X_s^2}{\mu}\right)^{\tau_0+1}\log\frac{\det(V_T)}{\det(V_{T_r})}\Bigg).$$

Proof. Let $\tau \leq t$ be the last time step before t, when the policy was updated. Using Cauchy-Schwarz inequality, we have:

$$\sum_{t=0}^{T}\|(\Theta_*-\tilde{\Theta}_t)^\intercal z_t\|^2 \leq \sum_{t=0}^{T}\|V_t^{\frac{1}{2}}(\tilde{\Theta}_t-\Theta_*)\|^2\|z_t\|^2_{V_t^{-1}} \leq \sum_{t=0}^{T}\frac{\det(V_t)}{\det(V_\tau)}\|V_\tau^{\frac{1}{2}}(\tilde{\Theta}_\tau-\Theta_*)\|^2\|z_t\|^2_{V_t^{-1}}. \tag{2.119}$$

Note that $t-\tau\leq\tau_0$ due to policy update rule. Moreover, we have

$$\det(V_t) = \det(V_\tau)\prod_{i=0}^{t-\tau}(1+\|z_t\|^2_{V_{t-i}^{-1}}) \leq \det(V_\tau)\left(1+\frac{\|z_t\|^2}{\mu}\right)^{\tau_0}.$$

Combining this with (2.119), on the event of E_T, for $t\leq T_r$, we have:

$$\sum_{t=0}^{T_r}\|(\Theta_* - \tilde{\Theta}_t)^\intercal z_t\|^2 \leq \sum_{t=0}^{T_r}\left(1+\frac{(1+\kappa^2)(n+d)^{2(n+d)}}{\mu}\right)^{\tau_0}\|V_\tau^{1/2}(\tilde{\Theta}_\tau-\Theta_*)\|^2\|z_t\|^2_{V_t^{-1}} \tag{2.120}$$

$$\leq \sum_{t=0}^{T_r}\left(1+\frac{(1+\kappa^2)(n+d)^{2(n+d)}}{\mu}\right)^{\tau_0}(\beta_T(\delta)+\upsilon_T(\delta))^2\|z_t\|^2_{V_t^{-1}}, \tag{2.121}$$

$$\leq \frac{2(1+\kappa^2)(n+d)^{2(n+d)}}{\mu}\left(1+\frac{(1+\kappa^2)(n+d)^{2(n+d)}}{\mu}\right)^{\tau_0}(\beta_T(\delta)+\upsilon_T(\delta))^2\log\left(\frac{\det(V_{T_r})}{\det(\mu I)}\right) \tag{2.122}$$

where in (2.121) we used the fact that on the event of E_T, using triangle inequality, we have $\|\tilde{\Theta}_\tau-\Theta_*\|_{V_\tau} \leq \|\tilde{\Theta}_\tau-\hat{\Theta}_\tau\|_{V_\tau}+\|\hat{\Theta}_\tau-\Theta_*\|_{V_\tau} \leq \upsilon_\tau(\delta)+\beta_\tau(\delta) \leq \upsilon_T(\delta)+\beta_T(\delta)$

and in (2.122) we used used the upper bound of $\|z_t\|_{V_t^{-1}}$ to utilize Lemma 10 of Abbasi-Yadkori and Szepesvári [2]. Similarly, on the even of E_t, for $t > T_r$, we get:

$$\sum_{t=T_r+1}^{T} \|(\Theta_* - \tilde{\Theta}_t)^\intercal z_t\|^2 \leq \frac{2(1+\kappa^2)X_s^2}{\mu}\left(1+\frac{(1+\kappa^2)X_s^2}{\mu}\right)^{\tau_0}(\beta_T(\delta)+\upsilon_T(\delta))^2 \log\left(\frac{\det(V_T)}{\det(V_{T_r})}\right)$$

■

Lemma 2.F.3 (Bounding R_T^{RLS} for TSAC)**.** *Let R_T^{RLS} be as defined by* (2.113)*. Under the event of E_T, setting $\mu = (1+\kappa^2)X_s^2$, we have*

$$\left|R_T^{RLS}\right| = \tilde{O}\left((n+d)^{(\tau_0+2)(n+d)+1.5}\sqrt{n}\sqrt{T_r} + (n+d)n\sqrt{T-T_r}\right).$$

Proof.

$$\left|R_T^{\text{RLS}}\right| \leq \sum_{t=0}^{T}\left|\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\tilde{\Theta}_t^\intercal z_t\right\|^2 - \left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\Theta_*^\intercal z_t\right\|^2\right| \tag{2.123}$$

$$= \sum_{t=0}^{T_r}\left|\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\tilde{\Theta}_t^\intercal z_t\right\|^2 - \left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\Theta_*^\intercal z_t\right\|^2\right| + \sum_{t=T_r}^{T}\left|\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\tilde{\Theta}_t^\intercal z_t\right\|^2 - \left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\Theta_*^\intercal z_t\right\|^2\right|$$

$$\leq \left(\sum_{t=0}^{T_r}\left(\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\tilde{\Theta}_t^\intercal z_t\right\| - \left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\Theta_*^\intercal z_t\right\|\right)^2\right)^{\frac{1}{2}}\left(\sum_{t=0}^{T_r}\left(\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\tilde{\Theta}_t^\intercal z_t\right\| + \left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\Theta_*^\intercal z_t\right\|\right)^2\right)^{\frac{1}{2}}$$
$$+ \left(\sum_{t=T_r}^{T}\left(\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\tilde{\Theta}_t^\intercal z_t\right\| - \left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\Theta_*^\intercal z_t\right\|\right)^2\right)^{\frac{1}{2}}\left(\sum_{t=T_r}^{T}\left(\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\tilde{\Theta}_t^\intercal z_t\right\| + \left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\Theta_*^\intercal z_t\right\|\right)^2\right)^{\frac{1}{2}} \tag{2.124}$$

$$\leq \left(\sum_{t=0}^{T_r}\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\left(\tilde{\Theta}_t - \Theta_*\right)^\intercal z_t\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=0}^{T_r}\left(\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\tilde{\Theta}_t^\intercal z_t\right\| + \left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\Theta_*^\intercal z_t\right\|\right)^2\right)^{\frac{1}{2}}$$
$$+ \left(\sum_{t=T_r}^{T}\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\left(\tilde{\Theta}_t - \Theta_*\right)^\intercal z_t\right\|^2\right)^{\frac{1}{2}}\left(\sum_{t=T_r}^{T}\left(\left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\tilde{\Theta}_t^\intercal z_t\right\| + \left\|P(\tilde{\Theta}_t)^{\frac{1}{2}}\Theta_*^\intercal z_t\right\|\right)^2\right)^{\frac{1}{2}} \tag{2.125}$$

where (2.123) and (2.125) follow from triangle inequality, and (2.124) follows from Cauchy Schwarz inequality. Note that for $t \leq T_r$, we have $\|z_t\|^2 \leq (1+\kappa^2)(n+d)^{2(n+d)}$ and for $t > T_r$ we have $\|z_t\|^2 \leq (1+\kappa^2)X_s^2$. Moreover, since $\tilde{\Theta}$ belongs to

$\mathcal{S}$ by construction of the rejection sampling, we get

$$\begin{aligned}
\left|R_T^{\text{RLS}}\right| \leq &\left(D\sum_{t=0}^{T_r}\left\|\left(\tilde{\Theta}_t-\Theta_*\right)^\intercal z_t\right\|^2\right)^{\frac{1}{2}}\sqrt{4T_rD(1+\kappa^2)S^2(n+d)^{2(n+d)}}\\
&+\left(D\sum_{t=T_r}^{T}\left\|\left(\tilde{\Theta}_t-\Theta_*\right)^\intercal z_t\right\|^2\right)^{\frac{1}{2}}\sqrt{4(T-T_r)D(1+\kappa^2)S^2X_s^2}\\
\leq &\frac{\sqrt{8T_r}DS(1+\kappa^2)(n+d)^{2(n+d)}(\beta_T(\delta)+\upsilon_T(\delta))}{\sqrt{\mu}}\left(1+\frac{(1+\kappa^2)(n+d)^{2(n+d)}}{\mu}\right)^{\frac{\tau_0}{2}}\sqrt{\log\left(\frac{\det(V_{T_r})}{\det(\mu I)}\right)}\\
&+\frac{\sqrt{8(T-T_r)}DS(1+\kappa^2)X_s^2\left(\beta_T(\delta)+\upsilon_T(\delta)\right)}{\sqrt{\mu}}\left(1+\frac{(1+\kappa^2)X_s^2}{\mu}\right)^{\frac{\tau_0}{2}}\sqrt{\log\left(\frac{\det(V_T)}{\det(V_{T_r})}\right)}
\end{aligned} \tag{2.126}$$

From Lemma 10 of Abbasi-Yadkori and Szepesvári [2], we have that $\log(\frac{\det(V_{T_r})}{\det(\mu I)}) \leq (n+d)\log(1+\frac{T_r(1+\kappa^2)(n+d)^{2(n+d)}}{\mu(n+d)})$ and $\log(\frac{\det(V_T)}{\det(V_{T_r})}) \leq (n+d)\log(1+\frac{T_r(1+\kappa^2)(n+d)^{2(n+d)}+(T-T_r)X_s^2}{\mu(n+d)})$.

After inserting these quantities into (2.126), we have the dimension dependency of $(n+d)^{2(n+d)} \times \sqrt{n(n+d)} \times (n+d)^{(n+d)\tau_0} \times (n+d)$ on the first term where $\sqrt{n(n+d)}$ is due to $\beta_T(\delta)+\upsilon_T(\delta)$. For the second term, for large enough T, we have the dimension dependency of $n \times \sqrt{n(n+d)} \times n^{(\tau_0/2)} \times \sqrt{n+d}$, where n comes from X_s^2. Thus, we achieve the following bound for $\left|R_T^{\text{RLS}}\right|$:

$$\left|R_T^{\text{RLS}}\right| = \tilde{O}\left((n+d)^{(\tau_0+2)(n+d)+1.5}\sqrt{n}\sqrt{T_r}+(n+d)n^{1.5+\tau_0/2}\sqrt{T-T_r}\right).$$

With the choice of $\mu=(1+\kappa^2)X_s^2$, the dependency of $n^{(\tau_0/2)}$ on the second term can be converted to a scalar multiplier of $\sqrt{2}^{\tau_0}$ and reduces the dependency of X_s^2 to X_s, which gives the advertised bound. ■

Bounding R_T^{mart}

Notice that this term is very similar to corresponding term in Abbasi-Yadkori and Szepesvári [2] and Abeille and Lazaric [4], besides the difference of early improved exploration. Following the same analysis, while including the effect of improved exploration gives the upper bound on R_T^{mart}. A similar analysis is also conducted in Lale, Azizzadenesheli, Hassibi, *et al.* [140], yet we provide it for completeness.

Lemma 2.F.4 (Bounding R_T^{mart}). *Let R_T^{mart} be as defined by* (2.114). *Under the event* *of E_T, with probability at least $1-\delta$, for $t > T_r$, we have*

$$\begin{aligned}
R_T^{mart} \leq\ & k_{s,1}(n+d)^{n+d}(\sigma_w + \|B_*\|\sigma_\nu)n\sqrt{T_r}\log((n+d)T_r/\delta) \\
&+ \frac{k_{s,2}(12\kappa^2 + 2\kappa\sqrt{2})}{\gamma}\sigma_w^2 n\sqrt{n}\sqrt{T-T_w}\log(n(T-T_w)/\delta) \\
&+ k_{s,3} n\sigma_w^2\sqrt{T-T_w}\log(nT/\delta) + k_{s,4} n(\sigma_w + \|B_*\|\sigma_\nu)^2\sqrt{T_w}\log(nT/\delta),
\end{aligned}$$

for some problem dependent coefficients $k_{s,1}, k_{s,2}, k_{s,3}, k_{s,4}$.

Proof. Let $f_t = A_* x_t + B_* u_t$. One can decompose R_T^{mart} as

$$R_1 = x_0^\intercal P(\tilde{\Theta}_0)x_0 - x_{T+1}^\intercal P(\tilde{\Theta}_{T+1})x_{T+1} + \sum_{t=1}^{T} x_t^\intercal P(\tilde{\Theta}_t)x_t - \mathbb{E}\left[x_t^\intercal P(\tilde{\Theta}_t)x_t \middle| \mathcal{F}_{t-2}\right]$$

Since $P(\tilde{\Theta}_0)$ is positive semidefinite and $x_0 = 0$, the first two terms are bounded above by zero. Recall that $\zeta_t = B_*\nu_t + w_t$ for $t \leq T_w$, and $\zeta_t = w_t$ for $t > T_w$. The second term is decomposed as follows

$$\begin{aligned}
&\sum_{t=1}^{T} x_t^\intercal P(\tilde{\Theta}_t)x_t - \mathbb{E}\left[x_t^\intercal P(\tilde{\Theta}_t)x_t \middle| \mathcal{F}_{t-2}\right] \\
&\qquad = \sum_{t=1}^{T} f_{t-1}^\intercal P(\tilde{\Theta}_t)\zeta_{t-1} + \sum_{t=1}^{T}\left(\zeta_{t-1}^\intercal P(\tilde{\Theta}_t)\zeta_{t-1} - \mathbb{E}\left[\zeta_{t-1}^\intercal P(\tilde{\Theta}_t)\zeta_{t-1} \middle| \mathcal{F}_{t-2}\right]\right)
\end{aligned}$$

Let $R_{1,1} = \sum_{t=1}^{T} f_{t-1}^\intercal P(\tilde{\Theta}_t)\zeta_{t-1}$, $R_{1,2} = \sum_{t=1}^{T}\left(\zeta_{t-1}^\intercal P(\tilde{\Theta}_t)\zeta_{t-1} - \mathbb{E}\left[\zeta_{t-1}^\intercal P(\tilde{\Theta}_t)\zeta_{t-1} \middle| \mathcal{F}_{t-2}\right]\right)$, and $v_{t-1}^\intercal = f_{t-1}^\intercal P(\tilde{\Theta}_t)$. Then one can write $R_{1,1}$. Let can be written as

$$R_{1,1} = \sum_{t=1}^{T}\sum_{i=1}^{n} v_{t-1,i}\zeta_{t-1,i} = \sum_{i=1}^{n}\sum_{t=1}^{T} v_{t-1,i}\zeta_{t-1,i}.$$

Let $M_{t,i} = \sum_{k=1}^{t} v_{k-1,i}\zeta_{k-1,i}$. By Theorem 2.G.1 on some event $G_{\delta,i}$ that holds with probability at least $1-\delta/(2n)$, for any $t \geq 0$,

$$\begin{aligned}
M_{t,i}^2 \leq\ & 2(\sigma_w^2 + \|B_*\|^2\sigma_\nu^2)\left(1 + \sum_{k=1}^{T_r} v_{k-1,i}^2\right)\log\left(\frac{2n}{\delta}\left(1 + \sum_{k=1}^{T_r} v_{k-1,i}^2\right)^{1/2}\right) \\
&+ 2\sigma_w^2\left(1 + \sum_{k=T_r+1}^{t} v_{k-1,i}^2\right)\log\left(\frac{2n}{\delta}\left(1 + \sum_{k=T_r+1}^{t} v_{k-1,i}^2\right)^{1/2}\right) \quad \text{for } t > T_r.
\end{aligned}$$

Notice that TSAC stops additional isotropic perturbation after $t = T_w$, and the state starts decaying until $t = T_r$. For simplicity of presentation we treat the time between T_w and T_r as TS with improved exploration while sacrificing the tightness of the result. On E_T, $\|v_k\| \leq DS(n+d)^{n+d}\sqrt{1+\kappa^2}$ for $k \leq T_r$ and $\|v_k\| \leq \frac{(12\kappa^2+2\kappa\sqrt{2})DS\sigma_w\sqrt{1+\kappa^2}}{\gamma}\sqrt{2n\log(n(t-T_w)/\delta)}$ for $k > T_r$. Thus, $v_{k,i} \leq DS(n+d)^{n+d}\sqrt{1+\kappa^2}$ and $v_{k,i} \leq \frac{(12\kappa^2+2\kappa\sqrt{2})DS\sigma_w\sqrt{1+\kappa^2}}{\gamma}\sqrt{2n\log(n(t-T_w)/\delta)}$ respectively for $k \leq T_r$ and $k > T_r$. Using union bound we get, for probability at least $1-\frac{\delta}{2}$, for $t > T_r$,

$$
\begin{aligned}
R_{1,1} \leq n&\sqrt{2(\sigma_w^2+\|B_*\|^2\sigma_\nu^2)\left(1+T_rD^2S^2(n+d)^{2(n+d)}(1+\kappa^2)\right)}\times \\
&\qquad\sqrt{\log\left(\frac{4n}{\delta}\left(1+T_rD^2S^2(n+d)^{2(n+d)}(1+\kappa^2)\right)^{1/2}\right)} \\
&+n\sqrt{2\sigma_w^2\left(1+\frac{2(t-T_r)(12\kappa^2+2\kappa\sqrt{2})^2D^2S^2n\sigma_w^2(1+\kappa^2)}{\gamma^2}\log(n(T-T_w)/\delta)\right)}\times \\
&\sqrt{\log\left(\frac{4n}{\delta}\left(1+\frac{2(t-T_r)(12\kappa^2+2\kappa\sqrt{2})^2D^2S^2n\sigma_w^2(1+\kappa^2)}{\gamma^2}\log(n(T-T_w)/\delta)\right)\right)}.
\end{aligned}
$$

Let $\mathcal{W}_{exp} = (\sigma_w+\|B_*\|\sigma_\nu)\sqrt{2n\log\frac{4nT}{\delta}}$ and $\mathcal{W}_{noexp} = \sigma_w\sqrt{2n\log\frac{4nT}{\delta}}$. Define $\Psi_t = \zeta_{t-1}^\intercal P(\tilde{\Theta}_t)\zeta_{t-1} - \mathbb{E}\left[\zeta_{t-1}^\intercal P(\tilde{\Theta}_t)\zeta_{t-1}|\mathcal{F}_{t-2}\right]$ and its truncated version $\tilde{\Psi}_t = \Psi_t\mathbb{I}_{\left\{\Psi_t\leq 2DW_{exp}^2\right\}}$ for $t \leq T_w$ and $\tilde{\Psi}_t = \Psi_t\mathbb{I}_{\left\{\Psi_t\leq 2DW_{noexp}^2\right\}}$ for $t > T_w$. Notice that $R_{1,2} = \sum_{t=1}^T \Psi_t$.

$$
\begin{aligned}
&\mathbb{P}\left(\sum_{t=1}^{T_w}\Psi_t > 2DW_{exp}^2\sqrt{2T_w\log\frac{4}{\delta}}\right)+\mathbb{P}\left(\sum_{t=T_w+1}^{T}\Psi_t > 2DW_{noexp}^2\sqrt{2(T-T_w)\log\frac{4}{\delta}}\right) \\
&\leq \mathbb{P}\left(\max_{1\leq t\leq T_w}\Psi_t > 2DW_{exp}^2\right)+\mathbb{P}\left(\max_{T_w+1\leq t\leq T}\Psi_t > 2DW_{noexp}^2\right) \\
&+\mathbb{P}\left(\sum_{t=1}^{T_w}\tilde{\Psi}_t > 2DW_{exp}^2\sqrt{2T_w\log\frac{4}{\delta}}\right)+\mathbb{P}\left(\sum_{t=T_w+1}^{T}\tilde{\Psi}_t > 2DW_{noexp}^2\sqrt{2(T-T_w)\log\frac{4}{\delta}}\right)
\end{aligned}
$$

By Lemma 2.G.4 with union bound and Theorem 2.G.2, summation of terms on the right hand side is bounded by $\delta/2$. Thus, with probability at least $1-\delta/2$, for $t > T_w$,

$$
R_{1,2} \leq 4nD\sigma_w^2\sqrt{2(t-T_w)\log\frac{4}{\delta}}\log\frac{4nT}{\delta}+4nD(\sigma_w+\|B_*\|\sigma_\nu)^2\sqrt{2T_w\log\frac{4}{\delta}}\log\frac{4nT}{\delta}.
$$

Combining $R_{1,1}$ and $R_{1,2}$ gives the statement. ■

Bounding R_T^{TS}

Lemma 2.F.5 (Bounding R_T^{TS} for TSAC)**.** *Let R_T^{TS} be as defined by* (2.112)*. Under the event of E_T, we have that*

$$\left|R_T^{TS}\right| \leq \tilde{O}\left(\sqrt{n}T_w + \mathsf{poly}(n, d, \log(1/\delta))\sqrt{T - T_w}\right).$$

with probability at least $1 - 2\delta$ if $T_w = \omega(\sqrt{T}\log T)$ for singular $A_{c,*}$ and $T_w = \omega(\log T)$ for non-singular $A_{c,*}$.

Proof. We decompose R_T^{TS} into two pieces as

$$R_T^{\text{TS}} = \underbrace{\sum_{t=0}^{T_w}\left\{J_*(\tilde{\Theta}_t, \sigma_w^2 I) - J_*(\Theta_*, \sigma_w^2 I)\right\}\mathbb{1}_{E_t}}_{R_{T_w}^{\text{TS,exp}}} + \underbrace{\sum_{t=T_w+1}^{T}\left\{J_*(\tilde{\Theta}_t, \sigma_w^2 I) - J_*(\Theta_*, \sigma_w^2 I)\right\}\mathbb{1}_{E_t}}_{R_T^{\text{TS,noexp}}}$$

Since every sampled system is in set $\mathcal{S}$, we have that $\|P(\tilde{\Theta}_t)\|_F \leq D$ and therefore

$$\begin{aligned}
R_{T_w}^{\text{TS,exp}} &\leq \sum_{t=0}^{T_w}|J_*(\tilde{\Theta}_t, \sigma_w^2 I) - J_*(\Theta_*, \sigma_w^2 I)|\,\mathbb{1}_{E_t} \\
&\leq \sum_{t=0}^{T_w}\left(|J_*(\tilde{\Theta}_t, \sigma_w^2 I)| + |J_*(\Theta_*, \sigma_w^2 I)|\right) && (2.127) \\
&\leq \sqrt{n}\sigma_w^2\sum_{t=0}^{T_w}\left(\|P(\tilde{\Theta}_t)\|_F + \|P(\Theta_*)\|_F\right) \leq 2\sqrt{n}\sigma_w^2 D T_w && (2.128)
\end{aligned}$$

where we used the relation $\mathrm{tr}(P) \leq \sqrt{n}\|P\|_F$ in (2.127). Considering the number of times a new TS sample is drawn, the second term in R_T^{TS} can be written as

$$R_K^{\text{TS,noexp}} = \sum_{k=0}^{K}\tau_0\left\{J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I) - J_*(\Theta_*, \sigma_w^2 I)\right\}\mathbb{1}_{E_{t_k}}$$

where $t_k = T_w + 1 + k\tau_0$ and $K = \left\lceil\frac{T-T_w}{\tau_0}\right\rceil$. Denoting the information available to the controller up to time $t \geq 0$ via $\mathcal{F}_t^{\text{cnt}} := \sigma\left(\mathcal{F}_{t-1}, x_t\right)$, $R_K^{\text{TS,noexp}}$ can be further decomposed into two pieces as

$$\begin{aligned}
R_K^{\text{TS,noexp}} &= \underbrace{\sum_{k=0}^{K}\tau_0\left\{J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I) - \mathbb{E}\left[J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I)\,\middle|\,\mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]\right\}\mathbb{1}_{E_t},}_{R_K^{\text{TS,1}}} \\
&\quad + \underbrace{\sum_{k=0}^{K}\tau_0\left\{\mathbb{E}\left[J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I)\,\middle|\,\mathcal{F}_t^{\text{cnt}}, E_{t_k}\right] - J_*(\Theta_*, \sigma_w^2 I)\right\}\mathbb{1}_{E_{t_k}}}_{R_K^{\text{TS,2}}}
\end{aligned}$$

We will investigate each term in order under the event of E_T.

Bounding $R_K^{\text{TS,1}}$. Notice that $\{R_K^{\text{TS,1}}\}_{K\geq 0}$ is a martingale sequence with $|R_K^{\text{TS,1}} - R_{K-1}^{\text{TS,1}}| \leq 2\tau_0\sigma_w^2\sqrt{n}D$. Therefore it can be bounded by Azuma's inequality (Lemma 2.G.2) w.p. at least $1-\delta$ as

$$R_K^{\text{TS,1}} \leq \sigma_w^2 D\sqrt{8n\tau_0^2 K\log(2/\delta)} \leq \sigma_w^2 D\sqrt{8n\tau_0(T-T_w)\log(2/\delta)} \qquad (2.129)$$

Bounding $R_K^{\text{TS,2}}$. Denoting by $\mathcal{S}^{\text{opt}} := \left\{\Theta \in \mathbb{R}^{(n+d)\times n} \mid J_*(\Theta, \sigma_w^2 I) \leq J_*(\Theta_*, \sigma_w^2 I)\right\}$ the set of optimistic parameters and defining $R_k^{TS,2} := \left\{\mathbb{E}\left[J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I) \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right] - J_*(\Theta_*, \sigma_w^2 I)\right\}\mathbb{1}_{E_{t_k}}$. Notice that, for any $\Theta \in \mathcal{S}^{\text{opt}}$, we can write

$$\begin{aligned}
R_k^{TS,2} &\leq \left\{\mathbb{E}\left[J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I) \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right] - J_*(\Theta, \sigma_w^2 I)\right\}\mathbb{1}_{E_{t_k}} \\
&\leq |J_*(\Theta, \sigma_w^2 I) - \mathbb{E}\left[J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I) \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]|\,\mathbb{1}_{E_{t_k}}
\end{aligned}$$

As the above bound holds for any $\Theta \in \mathcal{S}^{\text{opt}}$, we can replace the right hand side with an expectation over the optimistic set $\mathcal{S}^{\text{opt}}$. Specifically, we choose an i.i.d. copy of $\tilde{\Theta}_{t_k}$, that is, we choose a random variable $\tilde{\Theta}'_{t_k}$ which has the same distribution as $\tilde{\Theta}_{t_k}$ and independent from it. Then, we have that

$$\begin{aligned}
R_k^{TS,2} &\leq \mathbb{E}\left[|J_*(\tilde{\Theta}'_{t_k}, \sigma_w^2 I) - \mathbb{E}\left[J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I) \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]|\,\mathbb{1}_{E_{t_k}} \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}, \tilde{\Theta}'_{t_k} \in \mathcal{S}^{\text{opt}}\right] \\
&= \frac{\mathbb{E}\left[|J_*(\tilde{\Theta}'_{t_k}, \sigma_w^2 I) - \mathbb{E}\left[J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I) \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]|\,\mathbb{1}_{E_{t_k}}\,\mathbb{1}_{\tilde{\Theta}'_{t_k} \in \mathcal{S}^{\text{opt}}} \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]}{\mathbb{P}\left(\tilde{\Theta}'_{t_k} \in \mathcal{S}^{\text{opt}} \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, \hat{E}_{t_k}\right)}
\end{aligned}$$

Denoting by $p_t^{\mathsf{opt}} = \mathbb{P}\left(\tilde{\Theta}'_t \in \mathcal{S}^{\text{opt}} \,\middle|\, \mathcal{F}_t^{\text{cnt}}, \hat{E}_t\right)$ the probability of drawing cost optimistic TS samples, we can write further bounds on $R_k^{TS,2}$ as

$$\begin{aligned}
R_k^{TS,2} &\leq \frac{1}{p_{t_k}^{\mathsf{opt}}}\,\mathbb{E}\left[|J_*(\tilde{\Theta}'_{t_k}, \sigma_w^2 I) - \mathbb{E}\left[J_*(\tilde{\Theta}_{t_k}, \sigma_w^2 I) \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]| \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right] \\
&= \frac{\sigma_w^2}{p_{t_k}^{\mathsf{opt}}}\,\mathbb{E}\left[|\mathrm{Tr}\left(P(\tilde{\Theta}'_{t_k}) - \mathbb{E}\left[P(\tilde{\Theta}_{t_k}) \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]\right)| \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right] \\
&\leq \frac{n\sigma_w^2}{p_{t_k}^{\mathsf{opt}}}\,\mathbb{E}\left[\left\|P(\tilde{\Theta}'_{t_k}) - \mathbb{E}\left[P(\tilde{\Theta}_{t_k}) \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]\right\|_2 \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right] \qquad (2.130)
\end{aligned}$$

where we used the relation $|\mathrm{tr}(A)| \leq n\|A\|_2$. Denoting $P_k := \mathbb{E}\left[P(\tilde{\Theta}_{t_k}) \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]$, the following definition will be used in the rest of the section to understand the behavior of $R_k^{TS,2}$

$$\Delta_k := \mathbb{E}\left[\|P(\tilde{\Theta}_{t_k}) - P_k\|_2 \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right] \qquad (2.131)$$

The following lemma will be used to bound Δ_k from above.

Lemma 2.F.6. *For any $\Theta \in \mathcal{S}$, any positive definite matrix $V \in \mathbb{R}^{(n+d)\times(n+d)}$, and for any $i, j \in [n]$,*

$$\|\nabla P_{ij}(\Theta)\|_V \le \Gamma \|H(\Theta)\|_V,$$

where $\Gamma \geq 0$ is a problem dependent constant.

Proof. Let $\delta P(\Theta, \delta\Theta)$ be the differential of $P(\Theta)$ in the direction $\delta\Theta$. Then, we have that

$$\begin{aligned}\delta P(\Theta, \delta\Theta) = A_c(\Theta)^\intercal \delta P(\Theta, \delta\Theta) A_c(\Theta) \\ + A_c(\Theta)^\intercal P(\Theta)\delta\Theta^\intercal H(\Theta) + H(\Theta)^\intercal \delta\Theta P(\Theta) A_c(\Theta)\end{aligned} \tag{2.132}$$

where $A_c(\Theta) = \Theta^\intercal H(\Theta)$ is the closed-loop matrix. We know that $P(\Theta)$ satisfies the Riccati equation as

$$P - A_c^\intercal P A_c = Q + K^\intercal R K \succ 0 \implies \left(P^{\frac{1}{2}} A_c P^{-\frac{1}{2}}\right)^\intercal P^{\frac{1}{2}} A_c P^{-\frac{1}{2}} \prec I$$

where we dropped Θ dependence for simplicity. Therefore, similarity transformation of the closed-loop matrix $\bar{A}_c := P^{\frac{1}{2}} A_c P^{-\frac{1}{2}}$ is a contraction, *i.e.*, $\|P^{\frac{1}{2}} A_c P^{-\frac{1}{2}}\|_2 =: \sigma_\Theta < 1$. Multiplying both sides of (2.132) by $P^{-\frac{1}{2}}$ we obtain

$$\delta\bar{P}(\delta\Theta) = \bar{A}_c^\intercal \delta\bar{P}(\delta\Theta)\bar{A}_c + \bar{A}_c^\intercal P^{\frac{1}{2}}\delta\Theta^\intercal H P^{-\frac{1}{2}} + P^{-\frac{1}{2}} H^\intercal \delta\Theta P^{\frac{1}{2}} \bar{A}_c$$

where $\delta\bar{P}(\delta\Theta) = P^{-\frac{1}{2}}\delta P(\delta\Theta)P^{-\frac{1}{2}}$. Taking the spectral norm of both sides and using sub-multiplicativity of spectral norm as well as equivalence of matrix norms, we have that

$$\begin{aligned}\|\delta\bar{P}(\delta\Theta)\|_2 &\le \|A_c\|_2^2 \|\delta\bar{P}(\delta\Theta)\|_2 + 2\|\bar{A}_c\|_2 \|P^{\frac{1}{2}}\delta\Theta^\intercal H P^{-\frac{1}{2}}\|_2 \\ &\le \|A_c\|_2^2 \|\delta\bar{P}(\delta\Theta)\|_2 + 2\|\bar{A}_c\|_2 \|P^{\frac{1}{2}}\delta\Theta^\intercal H P^{-\frac{1}{2}}\|_F \\ &= \sigma_\Theta^2 \|\delta\bar{P}(\delta\Theta)\|_2 + 2\sigma_\Theta \|\delta\Theta^\intercal H\|_F\end{aligned}$$

By rearranging the inequality and using the property $\|\delta\Theta^\intercal H\|_F \le \|\delta\Theta\|_{V^{-1}} \|H\|_V$, we obtain

$$\|\delta\bar{P}(\delta\Theta)\|_2 \le \frac{2\sigma_\Theta}{1-\sigma_\Theta^2} \|\delta\Theta\|_{V^{-1}} \|H\|_V$$

Observing that $\|\delta P(\delta\Theta)\|_2 = \|P^{\frac{1}{2}}\delta\bar{P}(\delta\Theta)P^{\frac{1}{2}}\|_2 \le \|P\|_2 \|\delta\bar{P}(\delta\Theta)\|_2 \le D\|\delta\bar{P}(\delta\Theta)\|_2$ and noting that $\|\nabla P_{ij}(\Theta)\|_V = \sup_{\|\delta\Theta\|_{V^{-1}}=1} |\delta P_{ij}(\delta\Theta)| \le \sup_{\|\delta\Theta\|_{V^{-1}}=1} \|\delta P(\delta\Theta)\|_2$, one can get

$$\|\nabla P_{ij}(\Theta)\|_V \le \frac{2D\sigma_\Theta}{1-\sigma_\Theta^2} \|H(\Theta)\|_V$$

Observing that the function $\sigma_\Theta : \mathcal{S} \to \mathbb{R}_+$ is continuous on $\mathcal{S}$ and $\sigma_* := \max_{\Theta \in \mathcal{S}} \sigma_\Theta < 1$ as $\mathcal{S}$ is compact, we can further bound the scalar from above by Θ independent constant $\Gamma = \frac{2D\sigma_*}{1-\sigma_*^2} > 0$. ■

The following lemma gives a useful upper bound on Δ_k.

Lemma 2.F.7. *Let Δ_k be defined as in* (2.131)*. Then, for all $k \geq 0$, we have that*

$$\Delta_k \leq 2n^2 \upsilon_{t_k} \Gamma\, \mathbb{E}\left[\|H(\tilde{\Theta}_{t_k})\|_{V_{t_k}^{-1}} \,\middle|\, \mathcal{F}_{t_k}^{cnt}, E_{t_k} \right].$$

Proof. The proof follows directly from applying the bound in Lemma 2.F.6 to Equation 11 in [4]. ■

Finally, we are ready to give a bound on the summation of Δ_k terms

Lemma 2.F.8. *Let Δ_k be defined as in* (2.131) *for any $k \geq 0$. Then, the following bound holds with probability at least $1 - \delta$*

$$\begin{aligned}\sum_{k=0}^{K} \Delta_k &\leq \frac{16n^2 \alpha \upsilon_T \Gamma}{1 + \frac{1}{\beta_T}} \left(\sum_{t=T_w+1}^{T} \|z_t\|_{V_t^{-1}} + 2\alpha \sqrt{2 \frac{T - T_w}{\tau_0} \frac{1+\kappa^2}{\mu} \log\left(\frac{2}{\delta}\right)} \right) \\ &\leq \tilde{O}(\mathsf{poly}(n, d, \log(1/\delta)) \sqrt{T - T_w})\end{aligned}$$

where $\alpha = (1 + 1/\beta_0^2)(\sqrt{2n \log(3n)} + \upsilon_T + (1+\kappa) S X_s)$.

Proof. Define $\bar{\Theta}_{t_k} = \hat{\Theta}_{t_k} + \beta_{t_k} V_{t_k}^{-\frac{1}{2}} \eta_{t_k}$. Using Proposition 9 in [4], we have that

$$\begin{aligned}\|H(\bar{\Theta}_{t_k})\|_{V_{t_k}^{-1}} &\leq \frac{8}{1 + \frac{1}{\beta_{t_k}}} \left\| H(\bar{\Theta}_{t_k})\, \mathbb{E}\left[x_{t_k} x_{t_k}^\intercal \mathbb{1}_{\|x_{t_k}\| \leq \alpha} \,\middle|\, \mathcal{F}_{t_k - 1}, E_{t_k - 1}, \bar{\Theta}_{t_k} \right] \right\|_{V_{t_k}^{-1}} \\ &\leq \frac{8}{1 + \frac{1}{\beta_{t_k}}} \left\| \mathbb{E}\left[H(\bar{\Theta}_{t_k}) x_{t_k} x_{t_k}^\intercal \mathbb{1}_{\|x_{t_k}\| \leq \alpha} \,\middle|\, \mathcal{F}_{t_k - 1}, E_{t_k - 1}, \bar{\Theta}_{t_k} \right] \right\|_{V_{t_k}^{-1}} \\ &\leq \frac{8\alpha}{1 + \frac{1}{\beta_{t_k}}}\, \mathbb{E}\left[\left\| H(\bar{\Theta}_{t_k}) x_{t_k} \right\|_{V_{t_k}^{-1}} \mathbb{1}_{\|x_{t_k}\| \leq \alpha} \,\middle|\, \mathcal{F}_{t_k - 1}, E_{t_k - 1}, \bar{\Theta}_{t_k} \right]\end{aligned}$$

By Lemma 2.F.7 and the preceding bound, we can write

$$\begin{aligned}
\Delta_k &\leq 2n^2 \upsilon_{t_k} \Gamma \, \mathbb{E}\left[\|H(\tilde{\Theta}_{t_k})\|_{V_{t_k}^{-1}} \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right] = 2n^2 \upsilon_{t_k} \Gamma \frac{\mathbb{E}\left[\|H(\bar{\Theta}_{t_k})\|_{V_{t_k}^{-1}} \mathbb{1}_{\bar{\Theta}_{t_k} \in \mathcal{S}} \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]}{\mathbb{P}\left\{\bar{\Theta}_{t_k} \in \mathcal{S} \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right\}} \\
&\leq \frac{16n^2 \alpha \upsilon_{t_k} \Gamma}{1 + \frac{1}{\beta_{t_k}}} \frac{\mathbb{E}\left[\mathbb{E}\left[\|H(\bar{\Theta}_{t_k}) x_{t_k}\|_{V_{t_k}^{-1}} \mathbb{1}_{\|x_{t_k}\| \leq \alpha} \,\middle|\, \mathcal{F}_{t_k - 1}, E_{t_k - 1}, \bar{\Theta}_{t_k}\right] \mathbb{1}_{\bar{\Theta}_{t_k} \in \mathcal{S}} \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right]}{\mathbb{P}\left\{\bar{\Theta}_{t_k} \in \mathcal{S} \,\middle|\, \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}\right\}} \\
&= \frac{16n^2 \alpha \upsilon_{t_k} \Gamma}{1 + \frac{1}{\beta_{t_k}}} \underbrace{\mathbb{E}[\mathbb{E}[\|\underbrace{H(\tilde{\Theta}_{t_k}) x_{t_k}}_{z_{t_k}}\|_{V_{t_k}^{-1}} \mathbb{1}_{\|x_{t_k}\| \leq \alpha} \mid \mathcal{F}_{t_k - 1}, E_{t_k - 1}, \tilde{\Theta}_{t_k}] \mid \mathcal{F}_{t_k}^{\text{cnt}}, E_{t_k}]}_{=:Y_k}
\end{aligned}$$

Notice that $\mathbb{E}\left[Y_k \middle| \mathcal{F}_{t_{k-1}}\right] = \mathbb{E}\left[\|z_{t_k}\|_{V_{t_k}^{-1}} \mathbb{1}_{\|x_{t_k}\| \leq \alpha} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]$ by law of iterated expectations and $\|z_{t_k}\|_{V_{t_k}^{-1}} \mathbb{1}_{\|x_{t_k}\| \leq \alpha} \leq \frac{1}{\sqrt{\mu}} \|H(\tilde{\Theta}_{t_k}) x_{t_k}\| \, \mathbb{1}_{\|x_{t_k}\| \leq \alpha} \leq \sqrt{\frac{1+\kappa^2}{\mu}} \alpha$.

Therefore, the sequence $\left\{Y_k - \|z_{t_k}\|_{V_{t_k}^{-1}} \mathbb{1}_{\|x_{t_k}\| \leq \alpha}\right\}_{k \geq 0}$ is a bounded martingale difference sequence. By Azuma's inequality, we have that with probability at least $1 - \delta$,

$$\sum_{k=0}^{K} \left(Y_k - \|z_{t_k}\|_{V_{t_k}^{-1}} \mathbb{1}_{\|x_{t_k}\| \leq \alpha}\right) \leq 2\alpha \sqrt{2 \frac{T - T_w}{\tau_0} \frac{1 + \kappa^2}{\mu} \log\left(\frac{2}{\delta}\right)}$$

We can bound the sum of $\|z_{t_k}\|_{V_{t_k}^{-1}}$ terms using Lemma 10 of [1] and Hölder's inequality as

$$\begin{aligned}
\sum_{k=0}^{K} \|z_{t_k}\|_{V_{t_k}^{-1}} &\leq \sum_{k=0}^{K} \|z_{t_k}\|_{V_{t_k}^{-1}} + \sum_{k=0}^{K} \sum_{t = t_k + 1}^{t_{k+1} - 1} \|z_t\|_{V_t^{-1}} \\
&= \sum_{t = T_w + 1}^{T} \|z_t\|_{V_t^{-1}} \leq \sqrt{T - T_w} \log \frac{\det(V_T)}{\det(V_{T_w})}
\end{aligned}$$

Combining these results, we obtain the desired bound

$$\sum_{k=0}^{K} \Delta_k \leq \frac{16n^2 \alpha \upsilon_T \Gamma}{1 + \frac{1}{\beta_T}} \left(\sum_{t = T_w + 1}^{T} \|z_t\|_{V_t^{-1}} + 2\alpha \sqrt{2 \frac{T - T_w}{\tau_0} \frac{1 + \kappa^2}{\mu} \log\left(\frac{2}{\delta}\right)}\right)$$

■

Now, we are ready to bound $R_K^{TS,2}$. Under the event E_T Theorem 2.5.1 suggests that $1/p_{t^{\text{opt}}} \leq O(1)$ if $T_w = \omega(\sqrt{T} \log T)$ for singular $A_{c,*}$ and $T_w = \omega(\log T)$ for non-singular $A_{c,*}$. Using this result together with Lemma 2.F.8, we have that

$$R_K^{TS,2} = \sum_{k=0}^{K} \tau_0 R_k^{TS,2} \leq n\sigma_w^2 \tau_0 \sum_{k=0}^{K} \frac{\Delta_k}{p_{t_k^{\text{opt}}}} \leq \tilde{O}(\mathsf{poly}(n,d)\sqrt{(T-T_w)\log(1/\delta)}) \tag{2.133}$$

with probability at least $1-\delta$. Combining the above with (2.128) and (2.129), we obtain the desired bound. ■

Bounding R_T^{gap}

Lemma 2.F.9 (Bounding R_T^{gap} for TSAC)**.** *Let R_T^{gap} be as defined by* (2.115)*. Under the event of E_T, we have that*

$$|R_T^{gap}| = \tilde{O}\left(\mathsf{poly}(n,d)\sqrt{T\log(1/\delta)}\right).$$

with probability at least $1-2\delta$ for large enough T.

Proof.

$$R_T^{\text{gap}} = \sum_{t=0}^{T} \mathbb{E}\left[x_{t+1}^{\intercal}\left(P(\tilde{\Theta}_{t+1}) - P(\tilde{\Theta}_t)\right) x_{t+1} \mathbb{1}_{E_{t+1}} \,\middle|\, \mathcal{F}_t\right] \tag{2.134}$$

$$= \sum_{t=0}^{K} \mathbb{E}\left[x_{t_k+1}^{\intercal}\left(P(\tilde{\Theta}_{t_k+1}) - P(\tilde{\Theta}_{t_k})\right) x_{t_k+1} \mathbb{1}_{E_{t_k+1}} \,\middle|\, \mathcal{F}_{t_k}\right] \tag{2.135}$$

Separating the duration of TSAC into two parts at $t = T_r$, we obtain two same term achieved in [4]. Note that in Abeille and Lazaric [4], the authors follow frequent update rule and TSAC updates every τ_0 time-steps. The proof of these terms similarly follow Section 5.2 in [4] and using Lemma 2.F.8 we obtain $\widetilde{\mathrm{O}}((n+d)^{n+d}\sqrt{T_r}+$ poly$(n,d)\sqrt{T-T_r})$. Note that there is an additional τ_0 factor in these bounds, due to "relatively slower" update of TSAC. For large enough T such that the second term dominates the overall upper bound, we obtain the advertised guarantee. ■

Proof of Theorem 2.4.1

Collecting the regret terms derived in subsections of Appendix 2.F, for large enough T, under the event E_T, we have that

$$\begin{aligned}
R_{T_w}^{\text{exp}} &= \tilde{O}\left((n+d)^{n+d}T_w\right), \quad \text{w.p. } 1-\delta \\
R_T^{\text{RLS}} &= \tilde{O}\left((n+d)^{n+d}\sqrt{T_r} + \mathsf{poly}(n,d,\log(1/\delta))\sqrt{T-T_r}\right), \\
R_T^{\text{mart}} &= \tilde{O}\left((n+d)^{n+d}\sqrt{T_r} + \mathsf{poly}(n,d,\log(1/\delta))\sqrt{T-T_w}\right), \quad \text{w.p. } 1-\delta \\
R_T^{\text{gap}} &= \tilde{O}\left((n+d)^{n+d}\sqrt{T_r} + \mathsf{poly}(n,d,\log(1/\delta))\sqrt{T-T_r}\right), \quad \text{w.p. } 1-2\delta
\end{aligned}$$

and choosing $T_w = \omega(\sqrt{T}\log T)$ for singular $A_{c,*}$ and $T_w = \omega(\log T)$ for non-singular $A_{c,*}$ gives

$$R_T^{\text{TS}} = \tilde{O}\left(\mathsf{poly}(n,d)T_w + \mathsf{poly}(n,d,\log(1/\delta))\sqrt{T-T_r}\right), \quad \text{w.p. } 1-2\delta.$$

Recall that the event E_T is true with probability at least $1-4\delta$. Combining all these bounds, we have the overall regret bound as

$$R_T = \tilde{O}\left((n+d)^{n+d}T_w + \mathsf{poly}(n,d,\log(1/\delta))\sqrt{T-T_w}\right), \quad \text{w.p. } 1-10\delta. \tag{2.136}$$

Notice that R_T is linear in the initial exploration time T_w with an exponential dimension dependency. Also note that $T_w \geq T_0 := \mathsf{poly}(\log(1/\delta), \sigma_w^{-1}, n, d, \bar{\alpha}, \gamma^{-1}, \kappa)$ guarantees a stabilizing controller by Lemma 2.4.2. In order to control the growth of R_T by $\tilde{O}(\sqrt{T})$, the initial exploration time can maximally be in the order of $(\sqrt{T})^{1+o(1)}$ where $T^{o(1)}$ hides all multiplicative sub-polynomial growths, *i.e.*, $T_w = O\left((\sqrt{T})^{1+o(1)}\right) = \tilde{O}(\sqrt{T})$.

On the other hand, Theorem 2.5.1 puts strict lower bounds on the growth of T_w in order to maintain asymptotically constant optimistic probability. In particular, for singular $A_{c,*}$, this condition is stated as $T_w = \omega(\sqrt{T}\log T)$. Combined with the required upper bound $O\left((\sqrt{T})^{1+o(1)}\right)$, it must be that $T_w = \max\left(T_0,\, c(\sqrt{T}\log T)^{1+o(1)}\right)$ for a constant $c > 0$ for large enough T. Inserting this result in (2.136) gives us

$$R_T = \tilde{O}\left((n+d)^{n+d}\sqrt{T}\right), \quad \text{w.p. } 1-10\delta$$

for large enough T. Observe that exponential dimension dependence is unavoidable in this case as the system is excited with isotropic noise in every direction long enough to dominate with exponential dimension.

For non-singular $A_{c,*}$, the lower bound is stated as $T_w = \omega(\log T)$. For large enough T, choosing $T_w = \max\left(T_0,\, c(\log T)^{1+o(1)}\right)$ for a constant $c > 0$ is sufficient to satisfy both the upper and lower bounds on T_w. Inserting this result in (2.136) gives us

$$R_T = \tilde{O}\left(\mathsf{poly}(n,d,\log(1/\delta))\sqrt{T}\right), \quad \text{w.p. } 1-10\delta$$

for large enough T. Observe that the exponential dimension dependence is not dominant anymore since logarithmically large T_w is sufficient to guarantee asymptotically constant optimistic probability.

2.G Technical Theorems

Theorem 2.G.1 (Theorem 1 of Abbasi-Yadkori, Pál, and Szepesvári [1])**.** *Let* $(\mathcal{F}_t; k \geq 0)$ *be a filtration and* $(m_k; k \geq 0)$ *be an* $\mathbb{R}^d$*-valued stochastic process adapted to* $(\mathcal{F}_k)$, $(\eta_k; k \geq 1)$ *be a real-valued martingale difference process adapted to* $(\mathcal{F}_k)$. *Assume that* η_k *is conditionally sub-Gaussian with constant* R. *Consider the martingale*

$$S_t = \sum_{k=1}^{t} \eta_k m_{k-1}$$

and the matrix-valued processes

$$V_t = \sum_{k=1}^{t} m_{k-1} m_{k-1}^\intercal, \quad \overline{V}_t = V + V_t, \quad t \geq 0$$

Then for any $0 < \delta < 1$*, with probability* $1 - \delta$

$$\forall t \geq 0, \quad \|S_t\|^2_{\overline{V}_t^{-1}} \leq 2R^2 \log \left(\frac{\det \left(\overline{V}_t\right)^{1/2} \det(V)^{-1/2}}{\delta} \right)$$

Theorem 2.G.2 (Azuma's inequality)**.** *Assume that* X_s *is a supermartingale and* $|X_s - X_{s-1}| \leq c_s$ *almost surely for* $s \geq 0$. *Then for all* $t > 0$ *and all* $\epsilon > 0$,

$$P\left(|X_t - X_0| \geq \epsilon\right) \leq 2 \exp\left(\frac{-\epsilon^2}{2 \sum_{s=1}^{t} c_s^2} \right)$$

Lemma 2.G.3 (Lemma 10 of Abbasi-Yadkori and Szepesvári [2])**.** *The following holds for any* $t \geq 1$ *:*

$$\sum_{k=0}^{t-1} \left(\|z_k\|^2_{V_k^{-1}} \wedge 1 \right) \leq 2 \log \frac{\det\left(V_t\right)}{\det(\lambda I)}$$

Further, when the covariates satisfy $\|z_t\| \leq c_m, t \geq 0$ *with some* $c_m > 0$ *w.p. 1 then*

$$\log \frac{\det\left(V_t\right)}{\det(\lambda I)} \leq (n + d) \log \left(\frac{\lambda(n+d) + t c_m^2}{\lambda(n+d)} \right)$$

Lemma 2.G.4 (Norm of Subgaussian vector)**.** *Let* $v \in \mathbb{R}^d$ *be a entry-wise* R*-subgaussian random variable. Then with probability* $1 - \delta$, $\|v\| \leq R\sqrt{2d \log(d/\delta)}$.

Lemma 2.G.5 (Theorem 20 of Cohen, Koren, and Mansour [40])**.** *Let* $z_t \in \mathbb{R}^{n+d}$ *for* $t = 0, 1, \ldots$ *be a sequence random variables that is adapted to a filtration* $\{\mathcal{F}_t\}_{t=0}^{\infty}$. *Suppose that* z_t *are conditionally Gaussian on* $\mathcal{F}_{t-1}$ *and that* $\mathbb{E}\left[z_t z_t^T \mid \mathcal{F}_{t-1}\right] \succeq \sigma_z^2 I$ *for some fixed* $\sigma_z^2 > 0$. *Then for* $t \geq 200(n+d) \log \frac{12}{\delta}$ *we have that with probability at least* $1 - \delta$

$$\sum_{s=1}^{t} z_s z_s^T \succeq \frac{t \sigma_z^2}{40} I.$$

2.H Implementation Details of Numerical Experiments

The LQR problem for the longitudinal flight control of Boeing 747 with linearized dynamics [113] is given as

$$A_* = \begin{bmatrix} 0.99 & 0.03 & -0.02 & -0.32 \\ 0.01 & 0.47 & 4.7 & 0 \\ 0.02 & -0.06 & 0.4 & 0 \\ 0.01 & -0.04 & 0.72 & 0.99 \end{bmatrix}, \; B_* = \begin{bmatrix} 0.01 & 0.99 \\ -3.44 & 1.66 \\ -0.83 & 0.44 \\ -0.47 & 0.25 \end{bmatrix}, \; Q = I, \; R = I, \; w \sim \mathcal{N}(0, I). \tag{2.137}$$

This system has been studied in [140], [191]. It corresponds to the dynamics for level flight of Boeing 747 at the altitude of 40000ft with the speed of 774ft/sec, for a discretization of 1 second. The first element of the state corresponds to the velocity of aircraft along body axis, the second is the velocity of aircraft perpendicular to body axis, the third is the angle between body axis and horizontal and the fourth is the angular velocity of aircraft. The system takes two dimensional inputs, where the first is the elevator angle and the second one is thrust.

For this task we deploy 4 different adaptive control algorithms that do not require initial stabilizing controller: (i) TSAC, (ii) StabL of Lale, Azizzadenesheli, Hassibi, *et al.* [140], (iii) TS-LQR of Abeille and Lazaric [4], and (iv) OFULQ of Abbasi-Yadkori and Szepesvári [2]. Each algorithm has certain hyperparamters and we tune each parameter in terms of its effect on refret and present the performance of the best performing hyperparameter choices. We use the actual estimation errors in the algorithm design. Note that this has been observed to have negligible effect on the performance [44].

To have fair comparison in the regret performance in a stabilizable system like (2.137), we follow fixed update rule in TS-LQR in parallel with TSAC, and add an additional minimum policy duration constraint to the standard design matrix determinant doubling of OFULQ. Moreover, in the implementation of optimistic parameter search we deploy projected gradient descent (PGD). Even though this approach works efficiently for the small dimensional problems such as (2.137), it becomes computationally challenging as the dimensionality of the system grows. Nevertheless, our results show that PGD is effective to find optimistic parameters and as observed in Lale, Azizzadenesheli, Hassibi, *et al.* [140] yields the superior performance of StabL with a small margin between TSAC. This difference is in parallel with the predictions of theory. As we show in our analysis, TS samples an optimistic model with a fixed probability. However, an effective way of solving the optimistic control design problem yields optimistic controllers at every time-step and gives more effective control over exploration vs. exploitation trade-off.

Chapter 3

LEARNING TO CONTROL PARTIALLY OBSERVED LINEAR DYNAMICAL SYSTEMS

3.1 Introduction

In this work, we study the adaptive control of an unknown partially observable (measurement-feedback) linear dynamical system with quadratic cost and Gaussian disturbances, i.e. the LQG control problem. This problem is central in control theory and reinforcement learning (RL) since it captures the crux of the challenges in policy design for real-world systems with unknown model dynamics [31]. In such systems, the controlling agent does not have access to the latent state of the system and observes the dynamics via noisy measurements. Since the state of the system is not directly observable, the challenges in system identification, and balancing the exploration vs. exploitation trade-off in policy design, are especially magnified.

In recent years, there have been several developments in algorithmic design and statistical learning guarantees in adaptive control [136]–[138], [159], [176], [211], [223], [261]. These studies primarily focus on improving the performance of the adaptive control algorithms in terms of regret (the excess cost against the optimal policy that knows the system dynamics) and on computational efficiency. The prior works that consider the regret minimization problem in adaptive control of unknown LQG control systems mainly adopt three different paradigms for policy design: certainty equivalence [159], online learning [136], [211], and the optimism principle [137], [138]. Even though these methods provide a variety of algorithms with strong theoretical regret guarantees, they suffer either from limited applicability in practice or inherent algorithmic drawbacks (see Section 3.7).

Among these methods, the optimism principle provides the most sophisticated strategy to handle the exploration vs. exploitation trade-off via selecting the model with the lowest cost within the set of possible models and executing the optimal policy for this model [117]. This strategy encourages exploration of rarely visited regions of the state space and can be shown to converge to the optimal policy asymptotically [21]. Thus, the optimism principle has been the central policy design option in the prior works on adaptive control in unknown LQG control systems [137], [138]. These works have established that using optimism, the adaptive control algorithms can attain $\widetilde{\mathrm{O}}(\sqrt{T})$

regret after T time steps, which is the best-known performance guarantee.[1] However, finding the model with the lowest cost, i.e., the optimistic model, requires solving a non-convex optimization problem and it is NP-hard in general [6]. This computational intractability severely limits the efficiency and applicability of these algorithms.

Thompson Sampling (TS) is a promising alternative to overcome the computational burden of finding the optimistic policy [218]. In TS, the agent samples at random from the posterior distribution of models computed from a given prior distribution and the observed data and executes the corresponding LQG-optimal policy for this model [222]. This approach replaces the cumbersome optimization in optimism with straightforward sampling and results in a polynomial-time method. It should be noted that in RL there are several empirical studies that demonstrate the efficacy of TS-based methods [12], [34], [173]. Motivated by these computational and empirical advantages, [4] proposed TS-based adaptive control algorithms for fully observable (state-feedback) linear quadratic (LQ) control systems, *i.e.*, LQRs. Their algorithm attains optimal performance only for scalar systems. More recently, [126] developed a TS-based adaptive control algorithm that attains optimal $\widetilde{\mathrm{O}}(\sqrt{T})$ regret for all stabilizable LQRs with arbitrary input and state dimensions. However, until this, there have been no theoretical or empirical studies of TS for the more challenging problem of adaptive control in unknown partially observable LQG control systems.

Our contributions: In this work, we theoretically and empirically study TS in adaptive control of unknown partially observable LQG control systems. In particular, we propose an efficient TS-based adaptive control algorithm, **T**hompson **S**ampling under **P**artial **O**bservability, TSPO, for learning and controlling unknown LQG control systems. We show that TSPO attains $\widetilde{\mathrm{O}}(\sqrt{T})$ regret after T time steps, which makes TSPO the first efficient adaptive control algorithm to achieve this regret rate for adaptive control of partially observable LQ control systems with convex cost (Table 3.1). Furthermore, we empirically study the performance of TSPO in the measurement-feedback control of a $2^{\text{nd}}-$order SISO system. We show that TSPO effectively explores the model dynamics and achieves competitive regret performance in a computationally efficient way.

TSPO starts with a short warm-up period to gather data to generate an initial model estimate. It then interacts with the system in epochs where it uses a fixed controller throughout each epoch. At the beginning of each epoch, TSPO uses a closed-loop system identification method (via a predictor-form state-space representation) and

[1] Here $\widetilde{\mathrm{O}}(\cdot)$ presents order up to logarithmic terms.

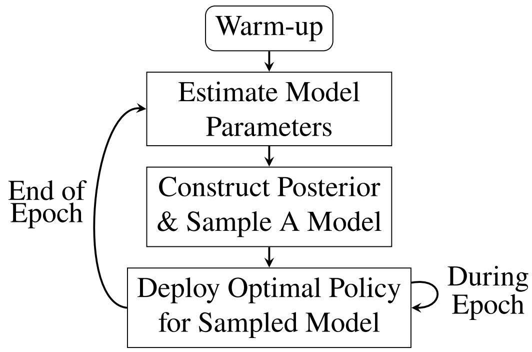


Figure 3.1: TSPO Framework

estimates the underlying model parameters along with confidence intervals. Using these estimates and associated uncertainties, TSPO constructs a posterior distribution on the model parameters and randomly samples a model from it. Throughout the epoch, it uses the optimal policy LQG for this sampled model. TSPO uses epochs with doubling length, and adaptively improves the model estimates and the controllers. The outline of TSPO is given in Figure 3.1. Conceptually, the TSPO method may not seem very surprising. What is surprising is that the simple TSPO achieves $\widetilde{\mathcal{O}}(\sqrt{T})$ regret. The main technical challenge of this paper is to establish this fact. To do so, we first show that the regret of a fixed TS policy scales linearly over time with respect to the estimation error in the model parameters (Theorem 3.4.4). Further, we prove that TS policies maintain stable system dynamics and bounded inputs/outputs provided a long enough warm-up duration (Theorem 3.4.1). Finally, we show that model the estimates and the TS samples jointly concentrate around the true model parameters over time (Theorem 3.4.2). Combining these results with the logarithmic policy updates of TSPO, we prove that the regret of TSPO is $\widetilde{\mathcal{O}}(\sqrt{T})$.

Due to space constraints, some of the proofs are given as sketches in this manuscript. The details and full proofs can be found in the extended version of this work online.

3.2 Problem Setting

Let $\Theta_\star := (A_\star, B_\star, C_\star)$ with $A_\star \in \mathbb{R}^{n\times n}$, $B_\star \in \mathbb{R}^{n\times d}$, $C_\star \in \mathbb{R}^{m\times n}$ be model parameters of an *environment* modelled as a linear time-invariant dynamical system in state-space form

$$\begin{aligned} x_{t+1} &= A_\star x_t + B_\star u_t + w_t, \\ y_t &= C_\star x_t + v_t, \end{aligned} \tag{3.1}$$

where $w_t \sim \mathcal{N}(0, W)$ and $v_t \sim \mathcal{N}(0, V)$ are independent process and measurement noise sequences, respectively, each with i.i.d. normal distribution with positive definite covariance matrices $W \in \mathbb{R}^{n\times n}$ and $V \in \mathbb{R}^{m\times m}$.

Table 3.1: Comparison with prior works on adaptive control in partially observable LQ control systems (CE = Certainty Equivalent, GD = Gradient Descent). S.Convex stands for strongly convex cost, *i.e.*, positive definite Q, R in (3.2)

Work	Method	Regret	Cost	Complexity
[159]	CE	$\sqrt{T}$	S.Convex	Polynomial
[211]	Online GD	$\sqrt{T}$	S. Convex	Polynomial
[136]	Online GD	polylog(T)	S. Convex	Polynomial
[137]	Optimism	$T^{2/3}$	Convex	NP-hard
[138]	Optimism	$\sqrt{T}$	Convex	NP-hard
[211]	Online GD	$T^{2/3}$	Convex	Poly
Our Work	**TS**	$\sqrt{\mathbf{T}}$	**Convex**	**Polynomial**

At each time step $t \geq 0$, an *agent* observes the output $y_t \in \mathbb{R}^m$ when the system is at (hidden) state $x_t \in \mathbb{R}^n$. Based on the knowledge of past output observations and control inputs, the agent then exerts a control input $u_t \in \mathbb{R}^d$ and suffers an instantaneous cost

$$c_t := y_t^\intercal Q y_t + u_t^\intercal R u_t, \tag{3.2}$$

where $Q \in \mathbb{R}^{m\times m}$ and $R \in \mathbb{R}^{d\times d}$ are positive semidefinite and positive definite matrices, respectively. After taking the control input u_t, the state evolves to x_{t+1}.

The goal of the agent is to reduce the cumulative cost $\sum_{t=0}^{T} c_t$ by deploying control actions after $T \geq 0$ number of interactions with the environment. This can be achieved by finding the best *control policy* that minimizes the average expected cost subject to the dynamical constraints in (3.1) as

$$J_\star := \limsup_{T\to\infty} \inf_{u_0,\dots,u_T} \frac{1}{T} \mathbb{E}\left[\sum\nolimits_{t=0}^{T} c_t\right] \text{ s.t. (3.1)}, \tag{3.3}$$

where $J_\star$ is the *optimal average expected cost* of the system $\Theta_\star$. Note that the control input u_t at a time $t \geq 0$ can be designed based on the past input-output pairs, $\mathcal{H}_t := \sigma(y_t, ..., y_0, u_{t-1}, ..., u_0)$. Any control policy that attains the cost of $J_\star$ in (3.3) is called an *optimal control policy*.

In the adaptive control setting, the agent is assumed to be unacquainted with the model parameters $\Theta_\star$ and has to design control inputs without knowing the underlying system. In this case, the agent can learn the model parameters of the underlying system from past interactions with the environment and can design control inputs at the same time accordingly based on past observations and inputs. Due to uncertainty in the true system, the agent deploys a suboptimal control policy even after several interactions. We measure the finite-time performance of the agent by *regret* defined as

$$R(T) \coloneqq \sum\nolimits_{t=0}^{T}(c_t - J_\star) \tag{3.4}$$

which is the difference between the cumulative cost the agent suffers after T interactions and the optimal steady-state cost attained after T steps with the perfect knowledge of $\Theta_\star$.

In order to assure that $J_\star$ attains a finite value, we assume that the underlying system, $\Theta_\star$ is controllable and observable. The optimal control policy of the system $\Theta_\star$ is given by $u_t = -K_\star \widehat{x}_{t|t}(\Theta_\star)$ with optimal feedback matrix

$$K_\star \coloneqq (R + B_\star^\intercal P_\star B_\star)^{-1} B_\star^\intercal P_\star A_\star, \tag{3.5}$$

where $P_\star$ is the unique positive semidefinite solution to the following discrete algebraic Riccati equation (DARE):

$$P_\star = A_\star^\intercal P_\star A_\star + C_\star^\intercal Q C_\star - A_\star^\intercal P_\star B_\star (R + B_\star^\intercal P_\star B_\star)^{-1} B_\star^\intercal P_\star A_\star. \tag{3.6}$$

The term $\widehat{x}_{t|t}(\Theta_\star)$ is the minimum mean squared error (MMSE) estimate of the underlying state x_t assuming system parameters $\Theta_\star$ and given the past information defined as $\mathcal{H}_t \coloneqq \sigma(y_t, \dots, y_0, u_{t-1}, \dots, u_0)$. The estimates can be computed efficiently by Kalman filter recursions given as

$$\begin{aligned} \widehat{x}_{t|t}(\Theta_\star) &= (I - L_\star C_\star)\widehat{x}_{t|t-1}(\Theta_\star) + L_\star y_t, \\ \widehat{x}_{t|t-1}(\Theta_\star) &= A_\star \widehat{x}_{t-1|t-1}(\Theta_\star) + B_\star u_{t-1}, \end{aligned} \tag{3.7}$$

with initial condition $\widehat{x}_{0|-1}(\Theta_\star) = 0$ where

$$L_\star \coloneqq \Sigma_\star C_\star^\intercal (C_\star \Sigma_\star C_\star^\intercal + V)^{-1}, \tag{3.8}$$

is the optimal Kalman gain and $\Sigma_\star$ is the unique positive semidefinite solution to the following DARE:

$$\Sigma_\star = A_\star \Sigma_\star A_\star^\intercal + W - A_\star \Sigma_\star C_\star^\intercal (C_\star \Sigma_\star C_\star^\intercal + V)^{-1} C_\star \Sigma_\star A_\star^\intercal. \tag{3.9}$$

The optimal average expected cost of controlling $\Theta_\star$ takes a finite value and can be computed as

$$J(\Theta_\star) = \mathrm{tr}(Q(C_\star \Sigma_\star C_\star^\intercal + V)) + \mathrm{tr}(P_\star(\Sigma_\star - S_\star)) \tag{3.10}$$

where $S_\star := \Sigma_\star - \Sigma_\star C_\star^\intercal (C_\star \Sigma_\star C_\star^\intercal + V)^{-1} C_\star \Sigma_\star$ is the covariance of the error $\widehat{x}_{t|t}(\Theta_\star) - x_t$. The dynamical system $\Theta_\star$ depicted in the state-space formulation in (3.1) can be equivalently represented as

$$\begin{aligned} \widehat{x}_{t+1} &= \bar{A}_\star \widehat{x}_t + B_\star u_t + F_\star y_t \\ y_t &= C_\star \widehat{x}_t + e_t \end{aligned} \tag{3.11}$$

where $F_\star := A_\star L_\star$ is the predictor gain, $\bar{A}_\star := A_\star - F_\star C_\star$, and $\{e_t\}$ is the zero mean and white innovation process. This equivalent representation is commonly known as the predictor form [120], [121] and the state $\widehat{x}_t$ can considered to be equivalent to $\widehat{x}_{t|t-1}(\Theta_\star)$, the MMSE estimate of state x_t given $(y_{t-1}, \ldots, y_0, u_{t-1}, \ldots, u_0)$. Since Kalman filter (3.7) converges to the steady state exponentially fast [172], the innovations process in the predictor form (3.11) attains the steady-state distribution $e_t \sim \mathcal{N}(0, C_\star \Sigma_\star C_\star^\intercal + V)$ and therefore the current output y_t is described by the history of inputs and outputs with an i.i.d. normal disturbance, e_t.

In our study, we will use the notion of strong stability introduced in [39] to quantify the stability of a matrix.

Definition 3.2.1 (Strong stability [39])**.** A matrix $A \in \mathbb{R}^{n \times n}$ is (κ, γ)-stable for $\kappa > 0$ and $\gamma \in (0, 1]$ if there exists a similarity transformation $A = S \Lambda S^{-1}$ such that $\|S\| \|S^{-1}\| \leq \kappa$ and $\|\Lambda\| \leq 1 - \gamma$.

Before stating the assumptions on $\Theta_\star$, we define the following metric to quantify the mismatch between model parameters which is invariant under similarity transformation as these transformations preserve the input-output dynamics.

Definition 3.2.2 (Model Mismatch Pseudometric)**.** Given two model parameters $\Theta_1 = (A_1, B_1, C_1)$, $\Theta_2 = (A_2, B_2, C_2)$, we define the following pseudo-metric

$$\rho(\Theta_1, \Theta_2) := \min_{\mathbf{T}, \mathbf{S} \in \mathsf{GL}_n} \max \left\{ \begin{array}{c} \|\mathbf{T}^{-1} A_1 \mathbf{T} - \mathbf{S}^{-1} A_2 \mathbf{S}\|, \\ \|\mathbf{T}^{-1} B_1 - \mathbf{S}^{-1} B_2\|, \\ \|C_1 \mathbf{T} - C_2 \mathbf{S}\| \end{array} \right\}$$

which is invariant under similarity transformations.

Assumption 3.2.3. The system $\Theta_\star = (A_\star, B_\star, C_\star)$ lies in a set $\mathcal{S}$ such that

$$\mathcal{S} \subseteq \left\{ \Theta = (A, B, C) \;\middle|\; \begin{array}{c} A \text{ is } (\kappa_1, \gamma_1)\text{-stable}, \\ (A, C) \text{ is observable}, \\ (A, B) \text{ is controllable}, \\ (A, F(\Theta)) \text{ is controllable}, \\ \max(\|A\|, \|B\|, \|C\|) \leq D \end{array} \right\},$$

Algorithm 2 TSPO

Input: (n, m, d), (Q, R), T_w, H, $\delta > 0$, $\mu > 0$, $D > 0$

—— WARM-UP ——

for $t = 0, 1, \ldots, T_w$ **do**

  Deploy $u_t \sim \mathcal{N}(0, \sigma_u^2 I)$ and store $\mathcal{D}_0 = \{y_t, u_t\}_{t=1}^{T_w}$

end for

—— ADAPTIVE CONTROL ——

for $i = 0, 1, \ldots$ **do**

  Compute $(\widehat{\mathcal{M}}_i, V_i) \leftarrow \mathrm{RLS}(\mathcal{D}_i = \{y_t, u_t\}_{t=0}^{2^i T_w}, \mu)$

  Sample $\widetilde{\mathcal{M}}_i \leftarrow \mathcal{R}_{\mathcal{S}}(\widehat{\mathcal{M}}_i + \beta_i V_i^{-\frac{1}{2}} \Xi)$, $[\Xi]_{ij} \sim \mathcal{N}(0, 1)$

  $\widetilde{\Theta}_i \leftarrow \textsc{SysId}\,(\widetilde{\mathcal{M}}_i, H, n)$

  for $t = 2^i T_w + 1, \ldots, 2^{i+1} T_w$ **do**

    Execute the optimal controller for $\widetilde{\Theta}_i$

  end for

end for

where $D > 0$, $\kappa_1 > 0$, and $\gamma_1 \in (0, 1]$. Furthermore, $\mathcal{S}$ consists of strongly stable systems, *i.e.*, there exist constants $\kappa_2, \kappa_3 > 0$ and $\gamma_2, \gamma_3 \in (0, 1]$ such that $A - BK(\Theta)$ is (κ_2, γ_2)-stable and $A - F(\Theta)C$ is (κ_3, γ_3)-stable for all $\Theta \in \mathcal{S}$.

The above assumptions are standard in system identification settings in order to ensure the possibility of accurate estimation of the system parameters [136]–[138], [159], [176], [198], [211], [224], [261].

Remark 3.2.4. By assuming controllability and observability of the underlying system with state dimension n, we implicitly assume the order of the underlying system is also n, *i.e.*, the system is in its minimal representation. We adopt this assumption for ease of presentation. There are several efficient algorithms that find the order of an unknown linear dynamical system [198]. Using these techniques, we can lift the assumption on the order of the system without jeopardizing any performance guarantees.

3.3 Thompson Sampling under Partial Observability (TSPO)

In this section, we present our proposed algorithm TSPO, the first computationally efficient and regret optimal RL algorithm for partially observable linear-quadratic control systems with convex instantaneous cost. TSPO is provided in Algorithm 2. It consists of two phases: (i) Warm-up period for pure exploration, (ii) Adaptive control using TS.

Warm-up: In the early stages, TSPO excites the system by injecting i.i.d.isotropic

Gaussian noise $u_t \sim \mathcal{N}(0, \sigma_u^2)$ for a duration of $T_w \geq 0$ and collects samples of observed output and control input, $\mathcal{D}_0 = \{(y_t, u_t)\}_{t=0}^{T_w}$. By exciting the system with $\mathrm{i.i.d.}$noise, TSPO explores the system effectively and generates a reliable initial estimate of the underlying model using the data collected. The warm-up duration, T_w, is set to meet a desired estimation accuracy so that any policy designed from the confidence set is guaranteed to stabilize and persistently excite the underlying system. We provide formal guarantees for stabilization in Theorem 3.4.1 and for estimation accuracy in Theorem 3.4.2.

Adaptive Control: After guaranteeing the design of stabilizing and persistently exciting policies during the warm-up phase, TSPO proceeds to the adaptive control phase. In this phase, TSPO cycles through epochs of fixed-policy control with doubling duration. At the beginning of each epoch, TSPO updates its policy based on input-output data gathered up to that time. The policy design involves three steps.

In the first step, TSPO deploys subroutine RLS to perform a closed-loop model estimation from the collected input-output data using regularized least squares. Consider the predictor form of system $\Theta_\star$ given in (3.11). Rolling back the state evolution $H > 0$ time steps back, we can write the observation at time $t \geq H$ as follows

$$y_t = \sum_{s=0}^{H-1} C_\star \bar{A}_\star^s \begin{bmatrix} F_\star & B_\star \end{bmatrix} \begin{bmatrix} y_{t-s-1} \\ u_{t-s-1} \end{bmatrix} + e_t + C_\star \bar{A}_\star^H \widehat{x}_{t-H}.$$

Since $\bar{A}_\star$ is stable by Assumption 3.2.3, the last term decays exponentially fast and is negligible for large enough H. Using this definition and following [136], the output y_t can be written compactly as follows

$$y_t = \mathcal{M}_\star \phi_t + e_t + C_\star \bar{A}_\star^H \widehat{x}_{t-H} \tag{3.12}$$

where $\mathcal{M}_\star \in \mathbb{R}^{m \times (m+d)H}$ is the H-truncated matrix of *predictor Markov parameters* defined as

$$\mathcal{M}_\star := \begin{bmatrix} M_\star^{(0)} & \ldots & M_\star^{(H-1)} \end{bmatrix}, \tag{3.13}$$

with $M_\star^{(s)} := C_\star \bar{A}_\star^s [F_\star \quad B_\star]$. The vector $\phi_t \in \mathbb{R}^{(m+d)H}$ is the truncated history of input-output data defined as

$$\phi_t := \begin{bmatrix} y_{t-1}^\intercal & \ldots & y_{t-H}^\intercal & u_{t-1}^\intercal & \ldots & u_{t-H}^\intercal \end{bmatrix}^\intercal \tag{3.14}$$

Thus, any input-output trajectory $\mathcal{D} = \{y_s, u_s\}_{s=0}^t$ up to time $t \geq H$ can be represented as

$$Y_t = \Phi_t \mathcal{M}_\star^\intercal + E_t + N_t \tag{3.15}$$

for $Y_t = \begin{bmatrix} y_H & y_{H+1} & \dots & y_t \end{bmatrix}^\intercal$, $\Phi_t = \begin{bmatrix} \phi_H & \phi_{H+1} & \dots & \phi_t \end{bmatrix}^\intercal$, $E_t = \begin{bmatrix} e_H & e_{H+1} & \dots & e_t \end{bmatrix}^\intercal$, $N_t = C_\star \bar{A}_\star^H \begin{bmatrix} \widehat{x}_0 & \widehat{x}_1 & \dots & \widehat{x}_{t-H} \end{bmatrix}^\intercal$.

The subroutine RLS takes any input-output trajectory data $\mathcal{D} = \{y_s, u_s\}_{s=0}^t$ for $t \geq H$, and constructs the data matrices Y_t and Φ_t. Provided with a regularization parameter $\mu > 0$, RLS obtains an estimate of the unknown truncated ARX model $\mathcal{M}_\star$ by solving the following regularized least square problem first introduced in [136],

$$\widehat{\mathcal{M}}_t \coloneqq \arg\min_{\mathbf{M}} \|Y_t - \Phi_t \mathbf{M}^\intercal\|_F^2 + \mu \|\mathbf{M}\|_F^2. \tag{3.16}$$

Denoting the *design matrix* by $V_t \coloneqq \mu I + \Phi_t^\intercal \Phi_t$, the solution to (3.16) is given by $\widehat{\mathcal{M}}_t = Y_t \Phi_t V_t^{-1}$. In Lemma 3.3.1, we give a self-normalized finite-sample estimation error for the closed-loop estimate $\widehat{\mathcal{M}}_t$ following [136, Thm. 3].

Lemma 3.3.1 (Closed-Loop Estimation, [136])**.** *Fix a horizon $T \geq H$. For all $t \in [H, T]$ and $\delta \in (0,1)$, true ARX model $\mathcal{M}_\star$ lies in the set $\widehat{\mathcal{C}}_t$ defined as*

$$\widehat{\mathcal{C}}_t \coloneqq \left\{ \mathbf{M} \mid \operatorname{tr}((\widehat{\mathcal{M}}_t - \mathbf{M}) V_t (\widehat{\mathcal{M}}_t - \mathbf{M})^\intercal) \leq \beta_t^2 \right\} \tag{3.17}$$

with probability at least $1-\delta$ where

$$\beta_t \coloneqq \sqrt{m \Sigma_e \log\left(\frac{\det(V_t)^{1/2}}{\delta \det(\mu I)^{1/2}} \right)} + \|\mathcal{M}_\star\|_F \sqrt{\mu} + \frac{t\sqrt{H}}{T^2}$$

for $\Sigma_e \coloneqq \|C_\star \Sigma_\star C_\star^\intercal + V\|_F$, as long as $H \geq H_c \coloneqq \Omega(\log T)$.

In the second step, TSPO calls subroutine TS to further explore the unknown system by sampling a random model from a distribution incorporating the estimated model and the associated uncertainty in the estimation. Given the estimate $\widehat{\mathcal{M}}_t$ and the design matrix V_t at time $t \geq T_w$, TS samples a perturbed truncated ARX model $\widetilde{\mathcal{M}}_t$ as follows

$$\widetilde{\mathcal{M}}_t = \mathcal{R}_{\mathcal{S}}(\widehat{\mathcal{M}}_t + \beta_t V_t^{-\frac{1}{2}} \Xi) \tag{3.18}$$

where $\mathcal{R}_{\mathcal{S}}$ denotes the rejection sampling operator associated with the set $\mathcal{S}$ given in Assumption 3.2.3, β_t is the confidence ellipsoid bound in Lemma 3.3.1, and $\Xi \in \mathbb{R}^{m \times (m+d)H}$ is the random perturbation matrix with i.i.d.standard normal entries, $[\Xi]_{ij} \sim \mathcal{N}(0,1)$.

The perturbation $\beta_t V_t^{-\frac{1}{2}} \Xi$ randomizes the RLS estimate coherently with the uncertainty conveyed by the design matrix. The rejection sampling operator $\mathcal{R}_{\mathcal{S}}$

keeps sampling independent random perturbations Ξ until the perturbed model $\widehat{\mathcal{M}}_t + \beta_t V_t^{-\frac{1}{2}} \Xi$ lies in set $\mathcal{S}$. The following lemma gives the confidence set for the sampled $\widetilde{\mathcal{M}}_t$.

Lemma 3.3.2 (TS confidence set)**.** *For all $t \geq H$, the sampled ARX model $\widetilde{\mathcal{M}}_t$ lies in the set $\widetilde{\mathcal{C}}_t$ defined as*

$$\widetilde{\mathcal{C}}_t \coloneqq \left\{ \mathbf{M} \mid \operatorname{tr}((\widehat{\mathcal{M}}_t - \mathbf{M}) V_t (\widehat{\mathcal{M}}_t - \mathbf{M})^\intercal) \leq \nu_t^2 \right\} \tag{3.19}$$

with probability at least $1 - \delta$ where

$$\nu_t \coloneqq \beta_t m \sqrt{2(m+d)H \log\left(2m(m+d)HT\delta^{-1}\right)} \tag{3.20}$$

Proof. We bound the probability of belonging to $\widetilde{\mathcal{C}}_t$ as

$$\begin{aligned}
\mathbb{P}(\forall t \leq T, \widetilde{\mathcal{M}}_t \in \widetilde{\mathcal{C}}_t) &= 1 - \mathbb{P}(\exists t \leq T, \widetilde{\mathcal{M}}_t \notin \widetilde{\mathcal{C}}_t) &(3.21)\\
&\geq 1 - \sum\nolimits_{i=0}^{T} \mathbb{P}(\widetilde{\mathcal{M}}_t \notin \widetilde{\mathcal{C}}_t) &(3.22)\\
&\geq 1 - \sum\nolimits_{i=0}^{T} \mathbb{P}(\|\Xi\|_F \geq \nu_t / \beta_t) &(3.23)\\
&\geq 1 - \delta &(3.24)
\end{aligned}$$

where (3.22) is due to union bound, (3.23) is due to rejection sampling and (3.24) is due to Gaussian norm bound. ■

In the last step of policy design, TSPO deploys subroutine SysId to obtain a state-space realization from the sampled truncated ARX matrix using a system identification method introduced by [136]. SysId is a subspace identification algorithm and a variation of well-known Ho-Kalman method [105]. By taking in $\widetilde{\mathcal{M}}_t$, SysId constructs corresponding block Hankel matrices using sampled Markov parameters $\tilde{M}_t^{(s)}$ and uses SVD and Assumption 3.2.3 to recover model parameters $\widetilde{\Theta}_t \coloneqq (\widetilde{A}_t, \widetilde{B}_t, \widetilde{C}_t)$ realizing the system governed by the truncated ARX model $\widetilde{\mathcal{M}}_t$. A detailed description of SysId can be found in [136]. In the following, we show that propagation of error from truncated ARX model to the state-space realization designed by SysId is linear.

Lemma 3.3.3 (Error propagation in SysId [136])**.** *Suppose Assumption 3.2.3 holds. Let $\Theta'_\star = (A'_\star, B'_\star, C'_\star)$ and $\widetilde{\Theta} = (\widetilde{A}, \widetilde{B}, \widetilde{C})$ be the model realizations obtained from $\mathcal{M}_\star$ and $\widetilde{\mathcal{M}}$ using SysId, respectively. The error between $\Theta'_\star$ and $\widetilde{\Theta}$ as measured in ρ can be bounded as follows*

$$\rho(\widetilde{\Theta}, \Theta'_\star) \leq \ell \|\widetilde{\mathcal{M}} - \mathcal{M}_\star\| \tag{3.25}$$

if $\|\widetilde{\mathcal{M}}-\mathcal{M}_\star\|\leq\epsilon_M$, *for positive problem dependent constants* $\ell=\mathsf{poly}(n,H,D,\kappa_2,\gamma_2)$ *and* $\epsilon_M=\mathsf{poly}(n,H,D,\kappa_2,\gamma_2)$.

In the rest of the epoch, TSPO deploys the optimal control policy of the sampled model $\widetilde{\Theta}_t$ given as $u_t = -\widetilde{K}_t\widehat{x}_{t|t}(\widetilde{\Theta}_t)$ where $\widetilde{K}_t$ is the optimal feedback matrix of $\widetilde{\Theta}_t$ and $\widehat{x}_{t|t}(\widetilde{\Theta}_t)$ is the MMSE estimate of the state assuming system $\widetilde{\Theta}_t$. Repeating this, TSPO keeps collecting samples during each epoch and uses the gathered data for refined model estimation, uncertainty quantification, and uncertainty-informed model sampling to further improve controller design in the next epoch. Due to reliable model estimation from the warm-up period, the controller designed right after the warm-up and all subsequently designed controllers stabilize and persistently excite the underlying model (Theorem 3.4.2)

3.4 Algorithmic Guarantees

In this section, we derive the algorithmic guarantees promised in Section 3.3. In particular, we formally state the guarantees pertaining to stability and persistence of excitation of the system under TSPO. We eventually show that the model mismatch error, *i.e.*, error between the sampled model and the underlying model decays as $\widetilde{\mathrm{O}}(1/\sqrt{t})$.

In the rest of this manuscript, we use asymptotic notation and hide problem-dependent constants to streamline the exposition as we are mainly interested in the regret rate with respect to the horizon, T. We also note that all the constants in this manuscript, where some are omitted to ease the presentation, have polynomial dependence in the problem-dependent constants.

As shown in [138, Lem. 3.1], the underlying system is persistently excited, *i.e.*, $\sigma_{\min}(V_{T_w}) = \Omega(T_w)$ during the warm-up period by injection of Gaussian input. Using the concentration result for the closed-loop estimation in Lemma 3.3.1, the estimation error at the end of warm-up period is bounded as

$$\|\widehat{\mathcal{M}}_{T_w} - \mathcal{M}_\star\|_F \leq \frac{\beta_{T_w}}{\sigma_{\min}(V_{T_w})} \leq \widetilde{\mathrm{O}}\left(\frac{1}{\sqrt{T_w}}\right) \tag{3.26}$$

Similarly, for the perturbation error in TS, we have $\|\widetilde{\mathcal{M}}_{T_w}-\widehat{\mathcal{M}}_{T_w}\|_F=\widetilde{\mathrm{O}}(1/\sqrt{T_w})$. This gives $\|\widetilde{\mathcal{M}}_{T_w}-\mathcal{M}_\star\|_F=\widetilde{\mathrm{O}}(1/\sqrt{T_w})$. By Lemma 3.3.3, we obtain the final model mismatch as $\rho(\widetilde{\Theta}_{T_w},\Theta_\star)=\widetilde{\mathrm{O}}(1/\sqrt{T_w})$ for $T_w\geq H+\tilde{\Omega}(1/\epsilon_M^2)$.

Our next objective is to find guarantees for stabilizing and persistently exciting policy design right after the warm-up period. Our strategy is to set T_w so that model mismatch error after the warm-up $\rho(\widetilde{\Theta}_{T_w}, \Theta_\star)$ is small enough to yield the desired policy. The following lemma shows that the underlying system can be stabilized by the optimal controller of another model with small model mismatch error.

Lemma 3.4.1 (Stability and Persistence of Excitation)**.** *Suppose that system $\Theta_\star \in \mathcal{S}$ is controlled by the optimal policy of a model $\widetilde{\Theta} \in \mathcal{S}$ for a duration of $\tau \geq 0$. For all $t \leq \tau$ and $\delta \in (0,1)$, with probability $1-\delta$, we have that*

$$\begin{aligned} \|x_t\| \leq \bar{X}_\tau, \qquad \|y_t\| \leq \bar{Y}_\tau, \\ \|\widehat{x}_{t|t}(\widetilde{\Theta})\| \leq \bar{\mathcal{X}}_\tau, \quad \|u_t\| \leq \bar{U}_\tau, \end{aligned} \tag{3.27}$$

where $\bar{X}_\tau, \bar{Y}_\tau, \bar{\mathcal{X}}_\tau, \bar{U}_\tau = O(\sqrt{\log(\tau/\delta)})$ whenever the model mismatch error is small as $\rho(\widetilde{\Theta}, \Theta_\star) \leq \epsilon_s$ for a problem dependent constant $\epsilon_s > 0$. Moreover, assuming the system $\Theta_\star$ is persistently excited by its optimal controller, we have the following $\sigma_{\min}(V_t) \geq t\sigma_p^2 \min(\sigma_{\min}^2(W), \sigma_{\min}^2(V))/16$, with probability $1-\delta$ whenever $\rho(\widetilde{\Theta}, \Theta_\star) \leq \epsilon_p$ for problem dependent constants $\epsilon_p, \sigma_p > 0$ [138].

Due to space constraints, we provide a proof sketch. Given the optimal control policy of $\widetilde{\Theta}$, we construct a $2n$-dimensional autonomous linear dynamical system of joint evolution of the state x_t and the Kalman filter $\widehat{x}_{t|t}(\widetilde{\Theta})$. By showing that the joint evolution is stable when the system $\Theta_\star$ is controlled by its own optimal controller, we can create a neighborhood (ρ-ball) around $\Theta_\star$ such that any model in the proximity yields a (κ', γ')-stable joint evolution with $\kappa', \gamma' = \mathsf{poly}(\kappa_1, \kappa_2, \kappa_3, \gamma_1, \gamma_2, \gamma_3, D)$. Similar to prior work in finding a stabilizing neighborhood, *e.g.* [140], we deduce the estimation error that we can tolerate such that the TS controllers for the systems sampled within the confidence sets, stabilize the underlying system. The proof of persistence of excitation also follows similarly by constructing a neighborhood of $\Theta_\star$ such that any controller from that neighborhood persistently excites as well. A detailed version of this proof can be found in [138, Lem. 3.2].

Here we assume that there exists a $\sigma_p > 0$ such that the underlying system is persistently excited with its own optimal controller. The necessary conditions for this is given in [138]. Following Lemma 3.4.1, $T_w \geq H + \tilde{\Omega}\left(\max(1/\epsilon_M^2, 1/\epsilon_s^2, 1/\epsilon_p^2)\right)$ guarantees that TSPO stabilizes and persistently excites the underlying system right after the warm-up during the first epoch in adaptive control period. Noticing that refining the model estimation by collecting more data in the subsequent epochs does not decrease the design matrix, we can argue that the model mismatch error in the next epochs does not increase. Therefore, all controllers designed by TSPO in the subsequent epochs stabilize and persistently excite the underlying system by Lemma 3.4.1. This leads to the following end-to-end result guaranteeing improving model mismatch error and stability throughout the adaptive control period.

Theorem 3.4.2 (End-to-End Guarantee)**.** *Fix a time horizon $T \geq T_w$. Denote by $\widetilde{\Theta}_i$ the model parameter obtained by TSPO at the beginning of the i^{th} epoch and by $T_i := 2^i T_w$ the time passed until the beginning of the i^{th} epoch. For all $i = 0, 1, \ldots, \lfloor \log(T/T_w) \rfloor$ and $\delta \in (0,1)$, the model mismatch error decays with probability at least $1-\delta$ as*

$$\rho(\widetilde{\Theta}_i, \Theta_\star) \leq \widetilde{\mathrm{O}}(T_i^{-1/2}) \tag{3.28}$$

Moreover, $(x_t, y_t, \widehat{x}_{t|t}, u_t)$ are bounded with high probability as in (3.27) *throughout the adaptive control phase.*

Proof. Observing $V_t \succcurlyeq V_{T_w}$ for any $t \geq T_w$, we can argue for the chosen T_w that $\rho(\widetilde{\Theta}_1, \Theta_\star) \leq \rho(\widetilde{\Theta}_0, \Theta_\star)$. Therefore, the requirements of Lemma 3.4.1 are satisfied and the system is stabilized and persistently excited in the next epoch as well. This yields $\rho(\widetilde{\Theta}_1, \Theta_\star) \leq \widetilde{\mathrm{O}}(1/\sqrt{2T_w})$. Following the same argument recursively, we conclude that the desired results hold for all the subsequent epochs. ■

We end this section with regret bounds for fixed policies. The following meta-theorem gives a regret upper bound for deploying a i.i.d.Gaussian excitation for a fixed period.

Theorem 3.4.3 (Regret of Gaussian Excitation, [137])**.** *Suppose system $\Theta_\star \in \mathcal{S}$ with dynamics* (3.1) *is driven by an* i.i.d.*normal Gaussian input process, $u_t \sim \mathcal{N}(0, \sigma_u^2 I)$. For $\delta \in (0,1)$, the regret incurred after $\tau \geq 0$ steps is bounded as*

$$R(\tau) \leq \mathsf{poly}(\sigma_u)\tau + \mathsf{poly}\left(\sigma_u, \log\left(\frac{1}{\delta}\right)\right) \widetilde{\mathrm{O}}(\sqrt{\tau}). \tag{3.29}$$

with probability at least $1-\delta$.

The following meta-theorem provides an upper bound on the regret of controlling a system $\Theta_\star$ by deploying the optimal policy of another system $\widetilde{\Theta}$ for a fixed period of time. This result shows that inaccuracies in due to model mismatch are propagated linearly in regret with linear-time growth. By controlling the model mismatch error in each fixed-policy epoch, we can reduce the regret to a desired level.

Theorem 3.4.4 (Regret of Model Mismatch)**.** *Suppose that system $\Theta_\star \in \mathcal{S}$ is controlled by the optimal policy of a model $\widetilde{\Theta} \in \mathcal{S}$ for a duration of $\tau \geq 0$. For $\delta \in (0,1)$, with probability $1-\delta$, the regret incurred due to model mismatch after $\tau \geq 0$ steps is bounded as*

$$R_{\widetilde{\Theta}}(\tau) \leq \widetilde{\mathrm{O}}(\rho(\widetilde{\Theta}, \Theta_\star)\tau). \tag{3.30}$$

with probability at least $1-\delta$ whenever the model mismatch error is small as $\rho(\widetilde{\Theta}, \Theta_\star) \leq \min(\epsilon_s, \epsilon_r)$ for a problem dependent constant $\epsilon_s > 0$.

Proof. We split the regret as follows:

$$R_{\widetilde{\Theta}}(\tau) = \sum\nolimits_{t=0}^{\tau}(c_t - \widetilde{J}) + \tau(\widetilde{J} - J_\star), \tag{3.31}$$

where $\widetilde{J}$ is the optimal average expected cost of $\widetilde{\Theta}$. Note that the dynamical variables are all bounded by Lemma 3.4.1 as $\rho(\widetilde{\Theta}, \Theta_\star) \leq \epsilon_s$. Therefore, following the analysis of [137, Thm. 4.1], we can bound the first term as $\widetilde{\mathrm{O}}(\rho(\widetilde{\Theta}, \Theta_\star)\tau)$. For the second regret term, consider $\delta\Theta := \widetilde{\Theta} - \Theta_\star$ the difference between models. Without loss of generality, we can argue

$$\epsilon := \max(\|\delta A\|_F, \|\delta C\|_F, \|\delta C\|_F) = \rho(\widetilde{\Theta}, \Theta_\star). \tag{3.32}$$

Notice that the optimal average expected cost function, $J(\Theta)$ given in (3.10) is a smooth function of its parameters, Θ within the highly non-convex domain $\mathcal{S}$. In order to obtain an error bound on the difference $\widetilde{J} - J_\star$, we can use linearized Taylor expansion in the close vicinity of $\Theta_\star$. In other words, there exists a problem dependent constant $\epsilon_r > 0$ such that for $\epsilon \leq \epsilon_r$, we have that

$$\begin{aligned} \widetilde{J} - J_\star &= \nabla_\Theta J(\Theta) \bullet \delta\Theta \\ &\leq \max(\|\nabla_A J(\Theta)\|, \|\nabla_B J(\Theta)\|, \|\nabla_C J(\Theta)\|)\epsilon \end{aligned}$$

where $\Theta_1 \bullet \Theta_2 := \mathrm{tr}(A_1 A_2^\intercal) + \mathrm{tr}(B_1 B_2^\intercal) + \mathrm{tr}(C_1 C_2^\intercal)$ is the Euclidean inner product and $\Theta = \Theta_\star + t\delta\Theta$ for $t \in [0,1]$. Taking the supremum of the last inequality over all $\Theta \in \mathcal{S}$ and noting that $\nabla_\Theta J(\Theta)$ is a continuous function over the compact set $\mathcal{S}$, we obtain the error bound $\widetilde{J} - J_\star \leq \Gamma_{\mathcal{S}}\,\epsilon$ where $\Gamma_{\mathcal{S}}$ is the maximum norm of $\nabla_\Theta J(\Theta)$ attained in $\mathcal{S}$. Substituting this result into the regret decomposition yields the desired regret bound. ∎

3.5 Regret Analysis

In this section, we formally state our main regret result for TSPO utilizing the algorithmic guarantees and fixed-policy regret bounds developed in Section 3.3. Theorem 3.4.3 shows that random exploration in the warm-up period incurs linear regret, *i.e.*, $O(T_w)$. In order to analyze the regret incurred during the adaptive control period, the system should be stabilized and the model mismatch error should decay. For a fixed horizon $T \geq T_w$, we define the following events

$$\bar{E}_T := \left\{\forall i \in [0, i_T], \|\widehat{x}_{t|t}(\widetilde{\Theta}_i)\| \leq \bar{\mathcal{X}}, \|y_t\| \leq \bar{Y}\right\} \tag{3.33}$$

$$\tilde{E}_T := \left\{\forall i \in [0, i_T], \rho(\widetilde{\Theta}_i, \Theta_\star) \leq \widetilde{\mathrm{O}}\left(\frac{1}{\sqrt{T_i}}\right)\right\} \tag{3.34}$$

where $\bar{\mathcal{X}}, \bar{Y} = O(\sqrt{\log T})$, $i_T := \lfloor \log(T/T_w) \rfloor$, and $T_i := 2^i T_w$ is the total time passed until the beginning of the i^{th} epoch. It is clear from Theorem 3.4.2 that the intersection $\bar{E}_T \cap \tilde{E}_T$ holds with high probability under the Assumption 3.2.3 if $T_w \geq H + \tilde{\Omega}\left(\frac{1}{\min(\epsilon_M^2, \epsilon_s^2, \epsilon_p^2, \epsilon_r^2)}\right)$. This result is critical for the regret analysis as it shows that inaccuracies in the estimation, as well as the random perturbations from TS are refining with the order of data collected in the past epochs and these errors, do not cause explosions in the system. With these results in hand, we give an upper bound on the overall regret of TSPO.

Theorem 3.5.1 (Regret of TSPO)**.** *Suppose Assumption 3.2.3 holds. Fixing a horizon* $T > 0$*, let* $H = \max(2n+1, \Omega(\log T))$ *and* $T_w \geq H + \Omega\left(\frac{1}{\min(\epsilon_M^2, \epsilon_s^2, \epsilon_p^2, \epsilon_r^2)}\right)$*. The regret incurred by TSPO up to horizon* T *is bounded with high probability as*

$$R_{\textit{TSPO}}(T) = \widetilde{\mathrm{O}}(\sqrt{T}). \tag{3.35}$$

Proof. We split the overall regret into individual regrets incurred during the warm-up period and each of the epochs in the adaptive control period as

$$R_{\text{TSPO}}(T) = R(T_w) + \sum_{i=0}^{i_T} R_i(T_{i+1} - T_i) \tag{3.36}$$

where R_i is the regret incurred during i^{th} epoch. From Theorem 3.4.3, we have that $R(T_w) = \widetilde{\mathrm{O}}(T_w)$. From Theorem 3.4.4, we can bound each regret term as

$$R_i(T_{i+1} - T_i) \leq \widetilde{\mathrm{O}}(\rho(\widetilde{\Theta}_i, \Theta_\star)(T_{i+1} - T_i)) \tag{3.37}$$

Noting that $\rho(\widetilde{\Theta}_i, \Theta_\star) \leq \widetilde{\mathrm{O}}(1/\sqrt{T_i})$ by Theorem 3.4.2, we have

$$R_i(T_{i+1} - T_i) \leq \widetilde{\mathrm{O}}\left(\frac{T_{i+1} - T_i}{\sqrt{T_i}}\right) = \widetilde{\mathrm{O}}\left(\sqrt{2^i T_w}\right) \tag{3.38}$$

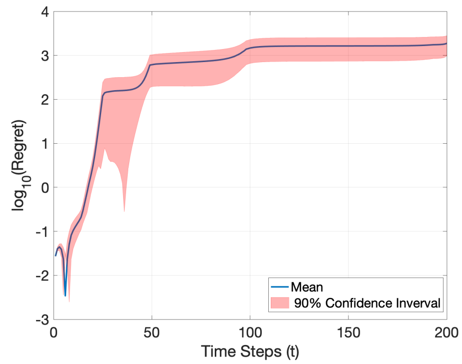


Figure 3.2: Regret Performance

Summing all these terms for $i_T = \lfloor \log(T/T_w) \rfloor$, we obtain $\sum_{i=0}^{i_T} R_i(T_{i+1} - T_i) = \widetilde{O}(\sqrt{T})$. ■

For the systems whose optimal policy is not a persistently exciting controller, we provide the following regret bound.

Theorem 3.5.2 (Regret without PE)**.** *If the underlying system $\Theta_\star$ is not persistently excited with its optimal policy, TSPO incurs the following regret with high probability,*

$$R_{TSPO}(T) = \widetilde{O}\left(T_w + \frac{T - T_w}{\sqrt{T_w}}\right) \tag{3.39}$$

Thus, setting $T_w = O(T^{2/3})$ gives $R_{TSPO}(T) = \widetilde{O}(T^{2/3})$.

Proof. Similar to the proof of Theorem 3.5.1, TSPO incurs $\widetilde{O}(T_w)$ regret during warm-up. Since the system is not guaranteed to be persistently excited, the best error bound for model mismatch error is attained right after warm-up. In other words, $\rho(\widetilde{\Theta}_i, \Theta_\star) \leq \widetilde{O}(1/\sqrt{T_w})$ for all epochs. By substituting this error result in the regret decomposition by invoking Theorem 3.4.4, the desired bound is obtained. Substituting $T_w = O(T^{2/3})$ yields the specified bound. ■

3.6 Numerical Simulations

In this section, we evaluate the performance of TSPO in a simulated adaptive measurement-feedback control task. In the simulations, we used state-space parameters given as

$$A_\star = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.7 \end{bmatrix}, \; B_\star = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, C_\star = \begin{bmatrix} 2 & 1 \end{bmatrix} \tag{3.40}$$

with $Q = R = I$ and isotropic Gaussian process and measurement noise with standard deviations as $\sigma_w = \sigma_v = 0.05$. We set the hyperparameters of TSPO as follows: ARX model truncation length $H = 10$, warm-up period $T_w = 12$, Gaussian excitation covariance $\sigma_u = 0.01$, RLS regularization parameter $\mu = 0.01$, and $\delta = 0.05$.

We perform 100 independent runs for 200 time-steps for TSPO and report their average and 90% confidence interval. The results are presented in Figure 3.2. The simulation results demonstrate that the regret over time almost stabilizes for the given system and the growth is sub-quadratic, matching the theoretical findings.

3.7 Related Works

Our work mainly relates to the works at the intersection of statistical learning and control. Recently, there have been considerable efforts to give finite-time regret guarantees for adaptive control algorithms in linear dynamical systems. TSPO fits into this line of work and overcomes some of the drawbacks of prior algorithms.

Fully Observable LQ Control (LQR): Due to their simplicity, state-feedback LQ control problems have been the primary focus in prior work [4], [63], [64], [126], [140], [174], [210]. These works have established that $\tilde{O}(\sqrt{T})$ regret is optimal in this setting. Some of these methods rely on certainty equivalent (CE) control which is sensitive to model mismatch and requires a priori knowledge of a stabilizing controller [210]. Some of them utilize optimism and avoid the need for a stabilizing controller, yet suffer from the inefficient (generally NP-hard) algorithmic procedure [64], [140]. TS-based methods overcome these drawbacks and have recently been shown to provide the first efficient adaptive control algorithm to achieve optimal regret in all stabilizable LQRs [126]. TS is also shown to be efficient in the control of continuous-time systems [61].

Partially Observable LQ Control: The statistical learning literature on measurement-feedback systems is more sparse due to the challenges of partial observability [136]–[138], [159], [211]. Among these, CE-based method in [159] attains $\tilde{O}(\sqrt{T})$ regret if the quadratic cost is *strongly convex* ($Q, R \succ 0$), which is only a subset of systems studied in this work. Similarly, under strongly convex cost condition, [211] show that $\tilde{O}(\sqrt{T})$ regret is attainable using online learning (gradient descent), and [136] further prove that optimal polylogarithmic regret is achievable in this setting. However, these results non-trivially rely on the strong convexity of the cost. Until now, the only efficient algorithm that provides regret guarantees in the setting of convex cost is given in [211], which also uses online learning but attains sub-optimal regret of $\tilde{O}(T^{2/3})$. TSPO and its guarantees match this result in the most general setting (Theorem 3.5.2), *i.e.* if the underlying system is not persistently excited by its optimal policy, and show that $\widetilde{\mathrm{O}}(\sqrt{T})$ regret is also attainable efficiently (Theorem 3.5.1).

3.8 Conclusion

In this work, we provide the first efficient adaptive control algorithm, TSPO, that attains $\widetilde{\mathrm{O}}(\sqrt{T})$ regret in partially observable LQ control systems with quadratic cost. We show that TSPO provides consistent estimates of the model parameters and designs controllers that stabilize the underlying system. Moreover, we show that the regret performance of controllers designed via TS improves linearly with respect to model estimation error, which allows us to derive our regret guarantees. One of the most important future directions is to further investigate the role of persistence of excitation (PE). In LQR literature, [126] shows that without PE, one can attain $\widetilde{\mathrm{O}}(\sqrt{T})$ using a self-normalized construction in the analysis. It remains an open problem if this result could be extended to the LQG control problem. Another important direction is to see whether TSPO can achieve polylogarithmic regret under strongly convex cost conditions.

3.A Proof of Lemma 3.4.1 (Stability and PE)

2. Proof of Lemma 3.3.1 (closed-loop estimation)

3. Proof of Lemma 3.3.2 (TS confidence set)

4. Proof of Lemma 3.3.3 (SysId error propagation)

5. Proof of Lemma 3.4.1 (Stability and PE)

6. Proof of Theorem 3.4.2 (End-to-End Guarantee)

7. Proof of Theorem 3.4.3 (Regret of Gaussian Excitation)

8. Proof of Theorem 3.4.4 (Regret of Model Mismatch)

9. Proof of Theorem 3.5.1 (Regret of TSPO)

10. Proof of Theorem 3.5.2 (Regret without PE)

Lemma 3.A.1 (Strong stability of perturbation [225, Prop. 4.0.1])**.** *Suppose the matrix $A \in \mathbb{R}^{n\times n}$ is (κ,γ)-stable for $\kappa \geq 1$ and $\gamma \in [0,1)$. For $\gamma' \in [\gamma,1)$ and perturbation $\Delta \in \mathbb{R}^{n\times n}$, the perturbed matrix $A+\Delta$ is (κ,γ')-stable whenever $\|\Delta\| \leq \kappa^{-1}(\gamma'-\gamma)$.*

Lemma 3.A.2 (Bounded state [32, Lem. 38])**.** *Suppose the matrix $A \in \mathbb{R}^{n\times n}$ is (κ,γ)-stable for $\kappa \geq 1$ and $\gamma \in [0,1)$ and the matrix $B \in \mathbb{R}^{n\times m}$ has bounded norm, $\|B\| \leq D$ for $D \geq 0$. Consider the linear dynamical system*

$$x_{t+1} = Ax_t + Bw_t, \;\; \text{for all } t \geq 0. \tag{3.41}$$

Given two time steps $0 \leq t_0 \leq t_1 < \infty$, we have that

$$\|x_t\| \leq \kappa\gamma^{t-t_0}\|x_{t_0}\| + \frac{\kappa D}{1-\gamma}\max_{s\in[t_0,t_1]}\|w_s\|, \tag{3.42}$$

for all $t \in [t_0, t_1]$.

Proof.

$$x_t = A^{t-t_0}x_{t_0} + \sum\nolimits_{s=t_0}^{t-1} A^{t-s-1}Bw_s \tag{3.43}$$

$$\begin{aligned}
\|x_t\| &\leq \|A^{t-t_0}\|\|x_{t_0}\| + \sum\nolimits_{s=t_0}^{t-1}\|A^{t-s-1}\|\|B\|\|w_s\| && (3.44)\\
&\leq \kappa\gamma^{t-t_0}\|x_{t_0}\| + \kappa D\max_{s\in[t_0,t_1]}\|w_s\|\sum\nolimits_{s=t_0}^{t-1}\gamma^{t-s-1} && (3.45)\\
&\leq \kappa\gamma^{t-t_0}\|x_{t_0}\| + (1-\gamma)^{-1}\kappa D\max_{s\in[t_0,t_1]}\|w_s\| && (3.46)
\end{aligned}$$

■

Lemma 3.A.3 (Strong stability under model mismatch)**.** *Suppose that a system $\Theta := (A,B,C) \in \mathcal{S}$ is controlled by the optimal policy of a model $\widetilde{\Theta} := (\widetilde{A},\widetilde{B},\widetilde{C}) \in \mathcal{S}$. The closed-loop dynamics is $(\kappa',\gamma')-$stable if $\mathbf{dist}(\widetilde{\Theta},\Theta) \leq \epsilon_{\mathit{stab}}$ where*

$$\kappa' := \frac{4\,\kappa_c \vee \kappa_o\|BK(I-LC)\|}{1-\gamma_c\vee\gamma_o}, \quad \gamma' := \frac{3+\gamma_c\vee\gamma_o}{4}, \tag{3.47}$$

$$\epsilon_{\mathit{stab}} := \epsilon_{K,L} \wedge \frac{(1-\gamma_c\vee\gamma_o)^2}{16\,\kappa_c\vee\kappa_o\|BK(I-LC)\|c_{\mathbf{\Phi}}}, \tag{3.48}$$

and $c_{\mathbf{\Phi}} > 0$ is a problem-dependent polynomial constant.
$\epsilon_{K,L}$ is the maximum mismatch bound required to obtain first-order perturbation bounds on K and L. Replace $\|BK(I-LC)\|$ with a constant independent of models but only on $\mathcal{S}$

Proof. use lemma D1, lemma D3, and Theorem J3

$$\begin{aligned}
x_{t+1} &= Ax_t + Bu_t + w_t\\
y_t &= Cx_t + v_t
\end{aligned} \tag{3.49}$$

$$\begin{aligned}\widehat{x}_{t+1|t} &= (A-BK)(I-LC)\widehat{x}_{t|t-1} + (A-BK)Ly_t && (3.50)\\ u_t &= -K(I-LC)\widehat{x}_{t|t-1} - KLy_t && (3.51)\end{aligned}$$

$$\begin{bmatrix} x_{t+1} \\ \widehat{x}_{t+1|t} \end{bmatrix} = \underbrace{\begin{bmatrix} A-BKLC & -BK(I-LC) \\ (A-BK)LC & (A-BK)(I-LC) \end{bmatrix}}_{\boldsymbol{\Phi}} \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} I & -BKL \\ 0 & (A-BK)L \end{bmatrix}}_{\boldsymbol{\Xi}} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \tag{3.52}$$

$$\begin{bmatrix} y_t \\ u_t \end{bmatrix} = \underbrace{\begin{bmatrix} C & 0 \\ -KLC & -K(I-LC) \end{bmatrix}}_{\boldsymbol{\Psi}} \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & I \\ 0 & -KL \end{bmatrix}}_{\boldsymbol{\Upsilon}} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \tag{3.53}$$

Defining $s_t \coloneqq x_t - \widehat{x}_{t|t-1}$

$$\begin{bmatrix} x_t \\ s_t \end{bmatrix} = \underbrace{\begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}}_{\boldsymbol{T}} \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1} \end{bmatrix}, \qquad \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1} \end{bmatrix} = \underbrace{\begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}}_{\boldsymbol{T}^{-1}} \begin{bmatrix} x_t \\ s_t \end{bmatrix}, \tag{3.54}$$

$$\begin{aligned} \begin{bmatrix} x_{t+1} \\ s_{t+1} \end{bmatrix} &= \underbrace{\begin{bmatrix} A-BK & BK(I-LC) \\ 0 & A-FC \end{bmatrix}}_{\boldsymbol{T\Phi T}^{-1}} \begin{bmatrix} x_t \\ s_t \end{bmatrix} + \underbrace{\begin{bmatrix} I & -BKL \\ I & -AL \end{bmatrix}}_{\boldsymbol{T\Xi}} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \\ \begin{bmatrix} y_t \\ u_t \end{bmatrix} &= \underbrace{\begin{bmatrix} C & 0 \\ -K & K(I-LC) \end{bmatrix}}_{\boldsymbol{\Psi T}^{-1}} \begin{bmatrix} x_t \\ s_t \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & I \\ 0 & -KL \end{bmatrix}}_{\boldsymbol{\Upsilon}} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \end{aligned} \tag{3.55}$$

$$\boldsymbol{T\Phi T}^{-1} = (A-BK) \oplus (A-FC) + \begin{bmatrix} 0 & BK(I-LC) \\ 0 & 0 \end{bmatrix} \tag{3.56}$$

The direct sum $(A-BK) \oplus (A-FC)$ is $(2\max(\kappa_c, \kappa_o),\, \max(\gamma_c, \gamma_o))$-stable by Lemma 3.C.4. The remaining block matrix on the right-hand side has all its eigenvalues at zero and its powers are exactly the zero matrices. Therefore, we claim it is $\left(\frac{2\|BK(I-LC)\|}{1-\max(\gamma_c,\gamma_o)}, \frac{1-\max(\gamma_c,\gamma_o)}{2}\right)$-stable.

$$\kappa_\star := \frac{4\max(\kappa_c,\kappa_o)\|BK(I-LC)\|}{1-\max(\gamma_c,\gamma_o)}, \quad \gamma_\star := \frac{1+\max(\gamma_c,\gamma_o)}{2} \tag{3.57}$$

Therefore, the joint closed-loop dynamics of the underlying system together with its optimal controller is $(\kappa_\star,\gamma_\star)$-stable.

$$\widehat{x}_{t+1|t}(\widetilde{\Theta}) = (\widetilde{A}-\widetilde{B}\widetilde{K})(I-\widetilde{L}\widetilde{C})\widehat{x}_{t|t-1}(\widetilde{\Theta}) + (\widetilde{A}-\widetilde{B}\widetilde{K})Ly_t \tag{3.58}$$

$$u_t = -\widetilde{K}(I-\widetilde{L}\widetilde{C})\widehat{x}_{t|t-1}(\widetilde{\Theta}) - \widetilde{K}\widetilde{L}y_t \tag{3.59}$$

$$\begin{bmatrix} x_{t+1} \\ \widehat{x}_{t+1|t}(\widetilde{\Theta}) \end{bmatrix} = \underbrace{\begin{bmatrix} A-B\widetilde{K}\widetilde{L}C & -B\widetilde{K}(I-\widetilde{L}\widetilde{C}) \\ (\widetilde{A}-\widetilde{B}\widetilde{K})\widetilde{L}C & (\widetilde{A}-\widetilde{B}\widetilde{K})(I-\widetilde{L}\widetilde{C}) \end{bmatrix}}_{\widetilde{\mathbf{\Phi}}} \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1}(\widetilde{\Theta}) \end{bmatrix} + \underbrace{\begin{bmatrix} I & -B\widetilde{K}\widetilde{L} \\ 0 & (\widetilde{A}-\widetilde{B}\widetilde{K})\widetilde{L} \end{bmatrix}}_{\widetilde{\mathbf{\Xi}}} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \tag{3.60}$$

$$\begin{bmatrix} y_t \\ u_t \end{bmatrix} = \underbrace{\begin{bmatrix} C & 0 \\ -\widetilde{K}\widetilde{L}C & -\widetilde{K}(I-\widetilde{L}\widetilde{C}) \end{bmatrix}}_{\widetilde{\mathbf{\Psi}}} \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1}(\widetilde{\Theta}) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & I \\ 0 & -\widetilde{K}\widetilde{L} \end{bmatrix}}_{\widetilde{\mathbf{\Upsilon}}} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \tag{3.61}$$

Defining $\tilde{s}_t := x_t - \widehat{x}_{t|t-1}(\widetilde{\Theta})$

$$\begin{bmatrix} x_t \\ \tilde{s}_t \end{bmatrix} = \underbrace{\begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}}_{\boldsymbol{T}} \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1}(\widetilde{\Theta}) \end{bmatrix}, \qquad \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1}(\widetilde{\Theta}) \end{bmatrix} = \underbrace{\begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}}_{\boldsymbol{T}} \begin{bmatrix} x_t \\ \tilde{s}_t \end{bmatrix}, \tag{3.62}$$

$$\begin{aligned} \begin{bmatrix} x_{t+1} \\ \tilde{s}_{t+1} \end{bmatrix} &= \boldsymbol{T}\widetilde{\mathbf{\Phi}}\boldsymbol{T}^{-1} \begin{bmatrix} x_t \\ \tilde{s}_t \end{bmatrix} + \boldsymbol{T}\widetilde{\mathbf{\Xi}} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \\ \begin{bmatrix} y_t \\ u_t \end{bmatrix} &= \widetilde{\mathbf{\Psi}}\boldsymbol{T}^{-1} \begin{bmatrix} x_t \\ \tilde{s}_t \end{bmatrix} + \widetilde{\mathbf{\Upsilon}} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \end{aligned} \tag{3.63}$$

$$\widetilde{\mathbf{\Phi}} - \mathbf{\Phi} = \mathbf{\Phi}_\Delta^{(1)} + \mathbf{\Phi}_\Delta^{(2)} + \mathbf{\Phi}_\Delta^{(3)} + \mathbf{\Phi}_\Delta^{(4)} \tag{3.64}$$

$$\mathbf{\Phi}_\Delta^{(1)} := \begin{bmatrix} -B(K_\Delta - KLC_\Delta) & -(KLC_\Delta + KL_\Delta C - K_\Delta(I-LC)) \\ -A_\Delta + B_\Delta K + ALC_\Delta & (A_\Delta - B_\Delta K)(I-LC) - A(LC_\Delta + L_\Delta C) \end{bmatrix} \tag{3.65}$$

$$\begin{aligned}
\|\mathbf{\Phi}_\Delta^{(1)}\| &\le 2\|B(K_\Delta - KLC_\Delta)\| + 2\|A_\Delta - B_\Delta K - ALC_\Delta\| \\
&\qquad + \|B(KL_\Delta + K_\Delta L)C\| + \|(A_\Delta - B_\Delta K)LC - AL_\Delta C\| \\
&\le (2 + \|LC\|)\|A_\Delta\| + (2\|K\| + \|KLC\|)\|B_\Delta\| + 2(\|AL\| + \|BKL\|)\|C_\Delta\| \\
&\qquad + (2 + \|LC\|)\|B\|\|K_\Delta\| + (\|A\| + \|BK\|)\|C\|\|L_\Delta\| \\
&\le c_{\mathbf{\Phi}}^{(1)} \text{ for } \Delta \le \epsilon_{K,L}
\end{aligned}$$

where

$$\begin{aligned}
c_{\mathbf{\Phi}}^{(1)} &:= 2 + \|LC\| + 2\|K\| + \|KLC\| + 2(\|AL\| + \|BKL\|) \\
&\qquad + (2 + \|LC\|)\|B\|c_K + (\|A\| + \|BK\|)\|C\|c_L
\end{aligned}$$

$$\mathbf{\Phi}_\Delta^{(2)} := \begin{bmatrix} B(KL_\Delta + K_\Delta L)C_\Delta & -B(KL_\Delta C_\Delta + K_\Delta LC_\Delta + K_\Delta L_\Delta C) \\ (A_\Delta - B_\Delta K)LC_\Delta + AL_\Delta C_\Delta + B_\Delta K_\Delta & -(A_\Delta - B_\Delta K)(LC_\Delta + L_\Delta C) - AL_\Delta C_\Delta - B_\Delta K_\Delta(I - LC \end{bmatrix} \tag{3.66}$$

$$\begin{aligned}
\|\mathbf{\Phi}_\Delta^{(2)}\| &\le 2\|B(KL_\Delta + K_\Delta L)C_\Delta\| + 2\|(A_\Delta - B_\Delta K)LC_\Delta + AL_\Delta C_\Delta + B_\Delta K_\Delta\| \\
&\qquad + \|BK_\Delta L_\Delta C\| + \|(A_\Delta - B_\Delta K)L_\Delta C - B_\Delta K_\Delta LC\| \\
&\le 2\|BK\|\|L_\Delta C_\Delta\| + 2\|B\|\|L\|\|K_\Delta\|\|C_\Delta\| + 2\|L\|\|A_\Delta\|\|C_\Delta\| \\
&\qquad + 2\|KL\|\|B_\Delta\|\|C_\Delta\| + 2\|A\|\|L_\Delta C_\Delta\| + 2\|B_\Delta K_\Delta\| + \|B\|\|C\|\|K_\Delta L_\Delta\| \\
&\qquad + \|C\|\|A_\Delta L_\Delta\| + \|K\|\|C\|\|B_\Delta\|\|L_\Delta\| + \|LC\|\|B_\Delta K_\Delta\| \\
&\le c_{\mathbf{\Phi}}^{(2)}\Delta^2 \text{ for } \Delta \le \epsilon_{K,L}
\end{aligned}$$

where

$$\begin{aligned}
c_{\mathbf{\Phi}}^{(2)} &:= 2\|BK\|c_L + 2\|B\|\|L\|c_K + 2\|L\| + 2\|KL\| + 2\|A\|c_L \\
&\qquad + 2c_K + \|B\|\|C\|c_K c_L + \|C\|c_L + \|K\|\|C\|c_L + \|LC\|c_K
\end{aligned}$$

$$\mathbf{\Phi}_\Delta^{(3)} := \begin{bmatrix} BK_\Delta L_\Delta C_\Delta & -BK_\Delta L_\Delta C_\Delta \\ (A_\Delta - B_\Delta K)L_\Delta C_\Delta - B_\Delta K_\Delta LC_\Delta & -(A_\Delta - B_\Delta K)L_\Delta C_\Delta + B_\Delta K_\Delta(LC_\Delta + L_\Delta C) \end{bmatrix} \tag{3.67}$$

$$\begin{aligned}\|\mathbf{\Phi}_\Delta^{(3)}\| &\leq 2\|BK_\Delta L_\Delta C_\Delta\| + 2\|(A_\Delta - B_\Delta K)L_\Delta C_\Delta - B_\Delta K_\Delta L C_\Delta\| + \|B_\Delta K_\Delta L_\Delta C\| \\ &\leq 2\|B\|\|K_\Delta L_\Delta C_\Delta\| + 2\|A_\Delta L_\Delta C_\Delta\| + 2\|K\|\|B_\Delta\|\|L_\Delta C_\Delta\| \\ &\quad + 2\|L\|\|B_\Delta K_\Delta\|\|C_\Delta\| + \|C\|\|B_\Delta K_\Delta L_\Delta\| \\ &\leq c_{\mathbf{\Phi}}^{(3)}\Delta^3 \text{ for } \Delta \leq \epsilon_{K,L}\end{aligned}$$

where

$$c_{\mathbf{\Phi}}^{(3)} := 2\|B\|c_K c_L + 2c_L + 2\|K\|c_L + 2\|L\|c_K + \|C\|c_L c_K$$

$$\mathbf{\Phi}_\Delta^{(4)} := \begin{bmatrix} 0 & 0 \\ -B_\Delta K_\Delta L_\Delta C_\Delta & B_\Delta K_\Delta L_\Delta C_\Delta \end{bmatrix} \tag{3.68}$$

$$\begin{aligned}\|\mathbf{\Phi}_\Delta^{(4)}\| &\leq 2\|B_\Delta K_\Delta L_\Delta C_\Delta\| && (3.69)\\ &\leq 2c_K c_L \Delta^4 \text{ for } \Delta \leq \epsilon_{K,L} && (3.70)\\ &\leq c_{\mathbf{\Phi}}^{(4)}\Delta^4 \text{ for } \Delta \leq \epsilon_{K,L} && (3.71)\end{aligned}$$

$$\begin{aligned}\|\widetilde{\mathbf{\Phi}} - \mathbf{\Phi}\| &\leq \|\mathbf{\Phi}_\Delta^{(1)}\| + \|\mathbf{\Phi}_\Delta^{(2)}\| + \|\mathbf{\Phi}_\Delta^{(3)}\| + \|\mathbf{\Phi}_\Delta^{(4)}\| && (3.72)\\ &\leq c_{\mathbf{\Phi}}^{(1)}\Delta + c_{\mathbf{\Phi}}^{(2)}\Delta^2 + c_{\mathbf{\Phi}}^{(3)}\Delta^3 + c_{\mathbf{\Phi}}^{(4)}\Delta^4 \text{ for } \Delta \leq \epsilon_{K,L} && (3.73)\\ &\leq c_{\mathbf{\Phi}}\Delta \text{ for } \Delta \leq \epsilon_{K,L} && (3.74)\end{aligned}$$

where

$$c_{\mathbf{\Phi}} := c_{\mathbf{\Phi}}^{(1)} + c_{\mathbf{\Phi}}^{(2)}\epsilon_{K,L} + c_{\mathbf{\Phi}}^{(3)}\epsilon_{K,L}^2 + c_{\mathbf{\Phi}}^{(4)}\epsilon_{K,L}^3$$

For $c_{\mathbf{\Phi}}\Delta \leq \frac{1-\gamma_\star}{2\kappa_\star}$, we have that $\widetilde{\mathbf{\Phi}}$ is $(\kappa_\star, \frac{1+\gamma_\star}{2})-$stable by Lemma 3.A.1. Therefore,

$$\Delta \leq \epsilon_{\text{stab}} := \min\left(\epsilon_{K,L}, \frac{1-\gamma_\star}{2\kappa_\star c_{\mathbf{\Phi}}}\right) \tag{3.75}$$

guarantees the desired result. ■

Lemma 3.A.4 (Precise Stability Statement of Lemma 3.4.1)**.** *Suppose that a system $\Theta_\star \in \mathcal{S}$ is controlled by the optimal policy of a model $\widetilde{\Theta} \in \mathcal{S}$ for a duration of $\tau \geq 0$. For $\delta \in (0,1)$, the closed-loop dynamics is bounded for all $t \leq \tau$ as*

$$\|x_t\| \leq \bar{X}_\tau :=, \tag{3.76}$$
$$\|y_t\| \leq \bar{Y}_\tau :=, \tag{3.77}$$
$$\|\widehat{x}_{t|t}(\widetilde{\Theta})\| \leq \bar{\mathcal{X}}_\tau :=, \tag{3.78}$$
$$\|u_t\| \leq \bar{U}_\tau := O(\sqrt{\log(\tau/\delta)}) \tag{3.79}$$

with probability $1-\delta$ whenever the model mismatch error is small as $\mathbf{dist}(\widetilde{\Theta}, \Theta_\star) \leq \epsilon_{bdd}$, where

$$\epsilon_{bdd} := \tag{3.80}$$

Proof. The closed-loop dynamics is given as

$$\begin{bmatrix} x_{t+1} \\ \widehat{x}_{t+1|t}(\widetilde{\Theta}) \end{bmatrix} = \underbrace{\begin{bmatrix} A_\star - B_\star \widetilde{K}\widetilde{L}C_\star & -B_\star \widetilde{K}(I - \widetilde{L}\widetilde{C}) \\ (\widetilde{A} - \widetilde{B}\widetilde{K})\widetilde{L}C_\star & (\widetilde{A} - \widetilde{B}\widetilde{K})(I - \widetilde{L}\widetilde{C}) \end{bmatrix}}_{\widetilde{\mathbf{\Phi}}} \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1}(\widetilde{\Theta}) \end{bmatrix} + \underbrace{\begin{bmatrix} I & -B_\star \widetilde{K}\widetilde{L} \\ 0 & (\widetilde{A} - \widetilde{B}\widetilde{K})\widetilde{L} \end{bmatrix}}_{\widetilde{\mathbf{\Xi}}} \begin{bmatrix} w_t \\ v_t \end{bmatrix}$$

$$\begin{bmatrix} y_t \\ u_t \end{bmatrix} = \underbrace{\begin{bmatrix} C_\star & 0 \\ -\widetilde{K}\widetilde{L}C_\star & -\widetilde{K}(I - \widetilde{L}\widetilde{C}) \end{bmatrix}}_{\widetilde{\mathbf{\Psi}}} \begin{bmatrix} x_t \\ \widehat{x}_{t|t-1}(\widetilde{\Theta}) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & I \\ 0 & -\widetilde{K}\widetilde{L} \end{bmatrix}}_{\widetilde{\mathbf{\Upsilon}}} \begin{bmatrix} w_t \\ v_t \end{bmatrix}$$

Observe that for $\mathbf{dist}(\widetilde{\Theta}, \Theta_\star) \leq \epsilon_{\text{stab}}$, the closed-loop evolution matrix $\widetilde{\mathbf{\Phi}}$ is $(\kappa', \gamma')-$stable by Lemma 3.A.3.

Denoting by $A_\Delta := \widetilde{A} - A_\star$, $B_\Delta := \widetilde{B} - B_\star$, $C_\Delta := \widetilde{C} - C_\star$, $K_\Delta := \widetilde{K} - K_\star$, $L_\Delta := \widetilde{L} - L_\star$ the mismatch errors, we have the following perturbation results for closed-loop dynamics.

■

3.B Proof of Theorem 3.4.2 (End-to-End Guarantee)

Lemma 3.B.1 (Bounded state in time-varying dynamics [32, Lem. 39])**.** *Suppose $\{A_i\}_{i=0}^{I-1} \subset \mathbb{R}^{n\times n}$ are (κ, γ)-stable matrices for $\kappa \geq 1$ and $\gamma \in [0,1)$, and $\{B_i\}_{i=0}^{I-1} \subset \mathbb{R}^{n\times m}$ are bounded as $\max_{0\leq i\leq I-1} \|B_i\| \leq D$ for $D \geq 0$. For a fixed horizon $T \geq 0$,* *let* $\{t_i\}_{i=0}^{I} \subset \mathbb{N}$ *such that* $0 \leq t_0 \leq \cdots \leq t_I \leq T$. *Consider the following time-varying linear dynamical system*

$$\forall i < I,\, \forall t \in [t_i, t_{i+1}),\; x_{t+1} = A_i x_t + B_i w_t. \tag{3.81}$$

Suppose further that

$$\tau \coloneqq \min_{i<I} |t_{i+1} - t_i| \geq \frac{\log(\kappa/\rho)}{\log(1/\gamma)} \tag{3.82}$$

for $\rho \in (0,1)$. *Then, we have that,*

$$\|x_t\| \leq (\kappa+1)\max\left(\|x_{t_0}\|, \frac{\kappa D W_T}{(1-\gamma)(1-\rho)}\right), \quad \text{for all } t \in [t_0, t_I], \tag{3.83}$$

where $W_T \coloneqq \max_{t\in[0,T]}\|w_t\|$.

strong stability of the closed-loop system (value of κ) depends on the similarity transformations. So this is not a well-defined property for a closed-loop system. Consider defining an alternative strong stability constant for Markov parameters.

Theorem 3.B.2 (Precise End-to-End Guarantee of Theorem 3.4.2)**.** *Fix a time horizon* $T \geq T_w$. *Denote by* $\widetilde{\Theta}_i$ *the model parameter obtained by TSPO at the beginning of the* i^{th} *epoch and by* $T_i \coloneqq 2^i T_w$ *the time passed until the beginning of the* i^{th} *epoch. For all* $i = 0, 1, \ldots, \lfloor \log(T/T_w) \rfloor$ *and* $\delta \in (0,1)$*, the model mismatch error decays with probability at least* $1-\delta$ *as*

$$\mathbf{dist}(\widetilde{\Theta}_i, \Theta_\star) \leq \widetilde{\mathrm{O}}(T_i^{-1/2}) \tag{3.84}$$

Moreover, $(x_t, y_t, \widehat{x}_{t|t}, u_t)$ *are bounded with high probability as in* (3.27) *throughout the adaptive control phase.*

Proof. 1. Long enough warm-up period T_w to guarantee a small model mismatch error (TS+RLS) at the end of the warm-up
2. Using the small model mismatch error by the end of the warm-up, guarantee strong stability of the closed-loop, boundedness, and PE during the first epoch
3. By stability, boundedness, and PE during the first epoch, guarantee smaller model mismatch error by the end of the first epoch.
4. Using the small model mismatch error by the end of the first epoch, guarantee strong stability of the closed-loop, boundedness, and PE during the second epoch
5. Repeat the same argument for every epoch to obtain the strong stability of closed-loop dynamics, boundedness, PE, and decaying model mismatch error rate

■

3.C Technical Theorems

Theorem 3.C.1 (sub-Gaussian tail inequality)**.** *Let $B \in \mathbb{R}^{n \times m}$ be a matrix and define $W := B^\intercal B$. Suppose that $w \in \mathbb{R}^m$ is a zero-mean sub-Gaussian random vector with parameter $\sigma > 0$, i.e.,*

$$\mathbb{E}\left[\exp(\lambda^\intercal w)\right] \leq \exp(\|\lambda\|^2 \sigma^2/2), \quad \textit{for all } \lambda \in \mathbb{R}^m. \tag{3.85}$$

For any $\delta \in (0,1)$, we have that

$$\|Bw\|^2 \leq \sigma^2 \left(\|W\|_1 + 2\|W\|_2 \sqrt{\log\left(\frac{1}{\delta}\right)} + 2\|W\|_\infty \log\left(\frac{1}{\delta}\right) \right) \tag{3.86}$$

with probability at least $1 - \delta$.

Theorem 3.C.2 (sub-Gaussian tail inequality v2)**.** *Suppose that $w \in \mathbb{R}^n$ is a zero-mean sub-Gaussian random vector with parameter $\sigma > 0$, i.e.,*

$$\mathbb{E}\left[\exp(\lambda^\intercal w)\right] \leq \exp(\|\lambda\|^2 \sigma^2/2), \quad \textit{for all } \lambda \in \mathbb{R}^n. \tag{3.87}$$

For any $\delta \in (0,1)$, we have that

$$\|w\| \leq \sigma \sqrt{2n \log\left(\frac{2n}{\delta}\right)} \tag{3.88}$$

with probability at least $1 - \delta$.

Theorem 3.C.3 (independent sub-Gaussian vectors)**.** *Suppose that $\{w_t\}_{t=1}^T \subset \mathbb{R}^n$ is a collection of $T > 0$ independent and zero-mean sub-Gaussian random vectors with parameter $\sigma > 0$. For any $\delta \in (0,1)$, we have that*

$$W_T := \max_{0 \leq t \leq T} \|w_t\| \leq \sigma \sqrt{2n \log\left(\frac{2nT}{\delta}\right)} \tag{3.89}$$

with probability at least $1 - \delta$.

Proof.

$$\begin{aligned} \mathbb{P}\left\{ \max_{0 \leq t \leq T} \|w_t\| \leq W \right\} &= \prod_{t=1}^T \mathbb{P}\left\{ \|w_t\| \leq W \right\} && (3.90) \\ &= \left(1 - \mathbb{P}\left\{ \|w_1\| > W \right\}\right)^T && (3.91) \\ &\geq 1 - T\ \mathbb{P}\left\{ \|w_1\| > W \right\} && (3.92) \end{aligned}$$

∎

Lemma 3.C.4. *Suppose that $A, B \in \mathbb{R}^{n\times n}$ are (κ_A, γ_A) and (κ_B, γ_B) stable matrices. Then, we have that*

i. $A + B$ *is* $(\kappa_A\kappa_B, \gamma_A + \gamma_B)$*-stable,*

ii. $A \oplus B$ *is* $(2\max(\kappa_A, \kappa_B), \max(\gamma_A, \gamma_B))$*-stable.*

Proposition 3.C.5 (Proper Transfer Functions)**.** *Let $G : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}^{m\times d}$ be a matrix-valued complex function such that the $(i, j)^{th}$ entry is a rational function with real coefficients, i.e., $G_{ij} \in \mathbb{R}(z)$. G is proper if and only if $G(z) = C(zI - A)^{-1}B + D$ for some $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times d}, C \in \mathbb{R}^{m\times n}, D \in \mathbb{R}^{m\times d}$. Such a quadruple (A, B, C, D) is called a realization of G with dimension $n \in \mathbb{N}$*

Definition 3.C.6 (Minimal realization)**.** Suppose $G(z)$ is a proper transfer matrix. A realization (A, B, C, D) of G with $A \in \mathbb{R}^{n\times n}$ is called a *minimal realization* if and only if there is no other realization with dimension no less than n.

Proposition 3.C.7. *Suppose $G(z)$ is a proper transfer matrix. A realization (A, B, C, D) of G is minimal if and only if (A, B) is controllable and (A, C) is observable.*

Proposition 3.C.8. *Suppose $G(z)$ is a proper transfer matrix. The following statement hold.*

i. If (A, B, C, D) is a realization of G, then $(TAT^{-1}, TB, CT^{-1}, D)$ is also a realization for any invertible matrix (of appropriate size) T.

ii. $G(z)$ *is strictly proper if and only if* $D = 0$*.*

Part II

Distributionally Robust Optimization

Chapter 4

FINITE-HORIZON DISTRIBUTIONALLY ROBUST CONTROL

4.1 Introduction

Regret-optimal control [48], [82], [161], [191], [219], is a new approach in control theory that focuses on minimizing the regret associated with control actions in uncertain systems. The regret measures the cumulative difference between the performance achieved by a causal control policy and the performance achieved by an optimal policy that could have been chosen in hindsight. In regret-optimal control, the worst-case regret over all ℓ_2-norm-bounded disturbance sequences is minimized.

Distributionally robust control [219], [220], [247], [258], on the other hand, addresses uncertainty in system dynamics and disturbances by considering a set of plausible probability distributions rather than relying on a single distribution as in LQG control, or on a worst-case disturbance, such as in H_∞ or RO control. This approach seeks to find control policies that perform well across all possible distributions within the uncertainty set, thereby providing robustness against model uncertainties and ensuring system performance in various scenarios. The size of the uncertainty set allows one to control the amount of desired robustness so that, unlike H_∞ controllers, say, the controller is not overly conservative. The uncertainty set is most often taken to be the set of disturbances whose distributions are within a given Wasserstein-2 distance of the nominal disturbance distribution. The reason is that, for quadratic costs, the supremum of the expected cost over a Wasserstein ball reduces to a tractable semi-definite program (SDP).

The current paper considers and extends the framework introduced in [219] that applied distributionally robust (DR) control to the regret-optimal (RO) setting. In the full-information finite-horizon setting, the authors of [219] reduce the DR-RO problem to a tractable SDP. In this paper, we extend the results of [219] to partially observable systems where, unlike the full-information setting, the controller does not have access to the system state. Instead, it only has access to partial information obtained through noisy measurements. This is often called the measurement feedback (MF) problem. Of course, the solution to the measurement feedback problem in LQG and H_∞ control is classical. The measurement-feedback setting for DR control has been studied in [220], [93], and for RO control in [81].

In the finite-horizon case, we reduce the DR-RO control problem with measurement feedback to an SDP similar to the full-information case studied in [219]. Furthermore, we validate the effectiveness and performance of our approach through simulations, showcasing its applicability in real-world control systems.

The organization of the paper is as follows. In section 4.2, we review the LQG and regret optimal control formulation in the measurement-feedback setting. In section 4.3, we present the distributionally robust regret-optimal with measurement feedback (DR-RO-MF) problem formulation, in section 4.4 we reformulate the problem as a tractable SDP, and in section 4.5 we show numerical results for controlling the flight of a Boeing 747 [114].

4.2 Preliminaries

Notations

$\mathbb{R}$ denotes the set of real numbers, $\mathbb{N}$ is the set of natural numbers, $\|\cdot\|$ is the 2-norm, $\mathbb{E}_{(\cdot)}$ is the expectation over $(\cdot)$, $\mathcal{M}(\cdot)$ is the set of probability distributions over $(\cdot)$ and Tr denotes the trace.

A Linear Dynamical System

We consider the following state-space model of a discrete-time, linear time-invariant (LTI) dynamical system:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t, \\ y_t &= Cx_t + v_t. \end{aligned} \tag{4.1}$$

Here, $x_t \in \mathbb{R}^n$ represents the state of the system, $u_t \in \mathbb{R}^m$ is the control input, $w_t \in \mathbb{R}^n$ is the process noise, while $y_t \in \mathbb{R}^p$ represents the noisy state measurements that the controller has access to, and $v_t \in \mathbb{R}^p$ is the measurement noise. The sequences $\{w_i\}$ and $\{v_i\}$ are considered to be randomly distributed according to an unknown joint probability measure P which lies in a specified compact ambiguity set, $\mathcal{P}$. For simplicity, we take x_0 to be zero.

In the rest of this paper, we adopt an operator form representation of the system dynamics (4.1). To this end, assume a horizon of $N \in \mathbb{N}$, and let us define

$$x := \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} \in \mathbb{R}^{Nn} \quad , \quad u := \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N-1} \end{bmatrix} \in \mathbb{R}^{Nm}$$

and similarly for $y \in \mathbb{R}^{Np}$, $w \in \mathbb{R}^{Nn}$, and $v \in \mathbb{R}^{Np}$. Using these definitions, we can represent the system dynamics (4.1) equivalently in operator form as

$$\begin{aligned} x &= Fu + Gw, \\ y &= Ju + Lw + v, \end{aligned} \tag{4.2}$$

where $F \in \mathbb{R}^{Nn \times Nm}$, $G \in \mathbb{R}^{Nn \times Nn}$, $J \in \mathbb{R}^{Np \times Nm}$, and $L \in \mathbb{R}^{Np \times Nn}$ are strictly causal time-invariant operators (i.e, strictly lower triangular block Toeplitz matrices) corresponding to the dynamics (4.1).

We consider the Linear-Quadratic Gaussian (LQG) cost given as

$$J(u, w, v) \coloneqq x^T Q x + u^T R u \tag{4.3}$$

where $Q, R \succ 0$ are positive definite matrices of the appropriate dimensions. In order to simplify the notation, we redefine x and u as $x \leftarrow Q^{\frac{1}{2}} x$, and $u \leftarrow R^{\frac{1}{2}} u$, so that (4.3) becomes

$$J(u, w, v) = \|x\|^2 + \|u\|^2. \tag{4.4}$$

Controller Design

We consider a linear controller that has only access to the measurements:

$$u = Ky, \quad K \in \mathcal{K}, \tag{4.5}$$

where $\mathcal{K} \subseteq \mathbb{R}^{Nm \times Np}$ is the space of causal (i.e., lower triangular) matrices. Then, the closed-loop state measurement becomes

$$y = (I - JK)^{-1}(Lw + v). \tag{4.6}$$

As in [81], let

$$E = K(I - JK)^{-1}, \tag{4.7}$$

be the Youla parametrization, so that

$$K = (I + EJ)^{-1} E. \tag{4.8}$$

The closed-loop LQG cost (4.4) can then be written as:

$$J(K, w, v) = \begin{bmatrix} w^T & v^T \end{bmatrix} T_K^T T_K \begin{bmatrix} w \\ v \end{bmatrix}, \tag{4.9}$$

where T_K is the transfer operator associated with K that maps the disturbance sequences $\begin{bmatrix} w \\ v \end{bmatrix}$ to the state and control sequences $\begin{bmatrix} x \\ u \end{bmatrix}$:

$$T_K \coloneqq \begin{bmatrix} FEL + G & FE \\ EL & E \end{bmatrix}. \tag{4.10}$$

Regret-Optimal Control with Measurement-Feedback

Given a noncausal controller $K_0 \in \mathcal{K}$, we define the regret as:

$$R(K,w,v) \coloneqq J(K,w,v) - J(K_0,w,v), \tag{4.11}$$

$$= \begin{bmatrix} w^T & v^T \end{bmatrix} (T_K^T T_K - T_{K_0}^T T_{K_0}) \begin{bmatrix} w \\ v \end{bmatrix}, \tag{4.12}$$

which measures the excess cost that a causal controller suffers by not knowing the future. In other terms, regret is the difference between the cost accumulated by a causal controller and the cost accumulated by a benchmark noncausal controller that knows the complete disturbance trajectory. The problem of minimizing regret in the measurement-feedback setting is referred to as (RO-MF) and is formulated as:

$$\inf_{K \in \mathcal{K}} \sup_{w,v} \frac{R(K,w,v)}{\|w\|^2 + \|v\|^2}, \tag{4.13}$$

which is solved suboptimally by reducing it to a level-1 suboptimal Nehari problem [81].

4.3 Distributionally Robust Regret-Optimal Control

In this section, we introduce the **distributionally robust regret-optimal** (DR-RO) control problem **with measurement feedback**, which we refer to as **DR-RO-MF**.

In this setting, the objective is to find a controller $K \in \mathcal{K}$ that minimizes the maximum expected regret among all joint probability distributions of the disturbances in an ambiguity set $\mathcal{P}$. This can be formulated formally as

$$\inf_{K \in \mathcal{K}} \sup_{P \in \mathcal{P}} \mathbb{E}_P[R(K,w,v)], \tag{4.14}$$

where the disturbances $\begin{bmatrix} w \\ v \end{bmatrix}$ are distributed according to $P \in \mathcal{P}$.

To solve this problem, we first need to characterize the ambiguity set $\mathcal{P}$ and explicitly determine a benchmark noncausal controller K_0. As in [219], we choose $\mathcal{P}$ to be the set of probability distributions that are at a distance of at most $r > 0$ to a nominal probability distribution, $P_0 \in \mathcal{M}(\mathbb{R}^{N(n+p)})$. Here, the distance is chosen to be the type-2 Wasserstein distance defined as [197]:

$$W_2^2(P_1,P_2) := \inf_{\pi \in \Pi(P_1,P_2)} \int_{\mathbb{R}^n \times \mathbb{R}^n} \|z_1 - z_2\|^2 \, \pi(dz_1, dz_2),$$

where the set $\Pi(P_1, P_2)$ comprises all joint distributions that have marginal distributions P_1 and P_2. Then, $\mathcal{P}$ can be written as:

$$\mathcal{P} := \{P \in \mathcal{M}(\mathbb{R}^{N(n+p)}) \,|\, W_2(P_0, P) \leq r\}. \tag{4.15}$$

Unlike the full-information case, we know from Theorem 1 in [81] that in the measurement feedback case, there is no optimal noncausal controller that dominates every other controller for every disturbance. Therefore, we will choose K_0 as the optimal noncausal controller that minimizes the Frobenius norm of T_K. Theorem 3 in [81] shows that such a controller can be found as:

$$K_0 = (I + E_0 J)^{-1} E_0, \tag{4.16}$$

where the associated operator, T_{K_0} is:

$$T_{K_0} = \begin{bmatrix} FE_0L + G & FE_0 \\ E_0L & E_0 \end{bmatrix}, \tag{4.17}$$

with

$$E_0 := -T^{-1}F^T G L^T U^{-1}, \tag{4.18}$$

$$T := I + F^T F, \tag{4.19}$$

$$U := I + LL^T. \tag{4.20}$$

4.4 Tractable Formulation

In this section, we introduce a tractable reformulation of the DR-RO-MF control problem (4.14).

DR-RO-MF Control Problem

Defining

$$\mathcal{C}_K := T_K^T T_K - T_{K_0}^T T_{K_0}, \tag{4.21}$$

we can rewrite the DR-RO-MF control problem (4.14) as

$$\inf_{K \in \mathcal{K}} \sup_{P \in \mathcal{P}} \mathbb{E}_P \left[\begin{bmatrix} w^T & v^T \end{bmatrix} \mathcal{C}_K \begin{bmatrix} w \\ v \end{bmatrix} \right]. \tag{4.22}$$

The following theorem gives the dual problem of inner maximization and characterizes the worst-case distribution.

Theorem 4.4.1. *[adapted from Theorems 2 and 3 in [219]]. Suppose P_0 is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^N$ and $\begin{bmatrix} w_0 \\ v_0 \end{bmatrix} \sim P_0$. The optimization problem:*

$$\sup_{P \in \mathcal{P}} \mathbb{E}_P \left[\begin{bmatrix} w^T & v^T \end{bmatrix} \mathcal{C}_K \begin{bmatrix} w \\ v \end{bmatrix} \right] \tag{4.23}$$

where $\begin{bmatrix} w \\ v \end{bmatrix} \sim P$ and $\mathcal{C}_K \in \mathbb{S}^{N(n+p)}$, with $\lambda_{max}(\mathcal{C}_K) \neq 0$, has a finite solution and is equivalent to the convex optimization problem:

$$\inf_{\substack{\gamma \geq 0, \\ \gamma I \succ \mathcal{C}_K}} \gamma(r^2 - \operatorname{Tr}(M_0)) + \gamma^2 \operatorname{Tr}(M_0(\gamma I - \mathcal{C}_K)^{-1}), \tag{4.24}$$

where $M_0 := \mathbb{E}_{P_0} \left[\begin{bmatrix} w \\ v \end{bmatrix} \begin{bmatrix} w^T & v^T \end{bmatrix} \right]$. Furthermore, the disturbance that achieves the worst-case regret is $\begin{bmatrix} w^* \\ v^* \end{bmatrix} \sim P^*$, where $\begin{bmatrix} w^* \\ v^* \end{bmatrix} = \gamma^*(\gamma^* I - \mathcal{C}_K)^{-1} \begin{bmatrix} w_0 \\ v_0 \end{bmatrix}$, and γ^* is the optimal solution of (4.24)*, which also satisfies the algebraic equation:*

$$\operatorname{Tr}((\gamma(\gamma I - \mathcal{C}_K)^{-1} - I)^2 M_0) = r^2 \tag{4.25}$$

Proof. The proof follows from Theorems 2 and 3 in [219] and is omitted for brevity here. ∎

We highlight two remarks pertaining to the presented theorem.

Remark 1: Notice that the supremum of the quadratic cost depends on P_0 only though its covariance matrix M_0. Note further that as $r \to \infty$, the optimal γ reaches its smallest possible value (since r^2 multiplies γ in (4.24)). The smallest possible value that γ can take is simply the operator norm of $\mathcal{C}_K$, which means that the DR-RO-MF controller approaches the regret-optimal controller as $r \to \infty$.

Remark 2: Notice that the worst-case disturbance takes on a Gaussian distribution when the nominal disturbance is Gaussian. This is not immediately evident as the ambiguity set $\mathcal{P}$ contains non-Gaussian distributions. Note further that the worst-case disturbance is correlated even if the nominal distribution has white noise.

Assuming the covariance of the nominal distribution to be

$$M_0 = \mathbb{E}_{P_0} \left[\begin{bmatrix} w \\ v \end{bmatrix} \begin{bmatrix} w^T & v^T \end{bmatrix} \right] = I. \tag{4.26}$$

so that $\mathrm{Tr}(M_0) = N(n+p)$, the optimization problem (4.22) can be cast equivalently using Theorem 4.4.1 as

$$\begin{aligned} &\inf_{K\in\mathcal{K}} \inf_{\gamma\geq 0} \gamma(r^2 - N(n+p)) + \gamma^2 \,\mathrm{Tr}((\gamma I - \mathcal{C}_K)^{-1}) \\ &\text{s.t.} \begin{cases} \gamma I \succ \mathcal{C}_K \\ \mathcal{C}_K = T_K^T T_K - T_{K_0}^T T_{K_0} \end{cases} \end{aligned} \tag{4.27}$$

As in [81], define the unitary matrices Ψ and Θ:

$$\Theta = \begin{bmatrix} S^{-\frac{1}{2}} & 0 \\ 0 & T^{-\frac{T}{2}} \end{bmatrix} \begin{bmatrix} I & -F \\ F^T & I \end{bmatrix} \tag{4.28}$$

$$\Psi = \begin{bmatrix} I & L^T \\ -L & I \end{bmatrix} \begin{bmatrix} V^{-\frac{1}{2}} & -0 \\ 0 & U^{-\frac{T}{2}} \end{bmatrix} \tag{4.29}$$

where T and U are as in (4.19) and (4.20), and

$$S = I + FF^T \tag{4.30}$$

$$V = I + L^T L. \tag{4.31}$$

and $S^{\frac{1}{2}}, T^{\frac{1}{2}}, U^{\frac{1}{2}}$, and $V^{\frac{1}{2}}$ are (block) lower triangular matrices, such that $S = S^{\frac{1}{2}} S^{\frac{T}{2}}$, $T = T^{\frac{T}{2}} T^{\frac{1}{2}}$, $U = U^{\frac{1}{2}} U^{\frac{T}{2}}$, $V = V^{\frac{T}{2}} V^{\frac{1}{2}}$. Then, the optimization problem (4.27) is equivalent to:

$$\begin{aligned} &\inf_{\substack{K\in\mathcal{K},\\ \gamma\geq 0,\\ \gamma I \succ \bar{\mathcal{C}}_K}} \gamma(r^2 - N(n+p)) + \gamma^2 \,\mathrm{Tr}((\gamma I - \bar{\mathcal{C}}_K)^{-1}) \\ &\text{s.t.} \begin{cases} \bar{\mathcal{C}}_K = (\Theta T_K \Psi)^T \Theta T_K \Psi - (\Theta T_{K_0} \Psi)^T \Theta T_{K_0} \Psi \end{cases} \end{aligned} \tag{4.32}$$

which holds true since trace is invariant under unitary Θ and Ψ. By introducing an auxiliary variable $X \succeq \gamma^2(\gamma I - \bar{\mathcal{C}}_K)^{-1}$ and leveraging the Schur complement theorem as in [219], the problem (4.32) can be recast as

$$\begin{aligned} &\inf_{\substack{K\in\mathcal{K},\\ \gamma\geq 0,\\ X\succeq 0}} \gamma(r^2 - N(n+p)) + \mathrm{Tr}(X) \\ &\text{s.t.} \begin{cases} \begin{bmatrix} X & \gamma I \\ \gamma I & \gamma I - \bar{\mathcal{C}}_K \end{bmatrix} \succeq 0 \\ \gamma I - \bar{\mathcal{C}}_K \succ 0 \\ \bar{\mathcal{C}}_K = (\Theta T_K \Psi)^T \Theta T_K \Psi - (\Theta T_{K_0} \Psi)^T \Theta T_{K_0} \Psi \end{cases} \end{aligned} \tag{4.33}$$

In the following lemma, we establish some of the important identities that are utilized to convert problem (4.33) to a tractable convex program.

Lemma 4.4.2. *[adapted from [81]]. The following statements hold:*

1.
$$\gamma I - \bar{\mathcal{C}}_K = \begin{bmatrix} \gamma I & -PZ \\ -Z^T P^T & \gamma I - Z^T Z \end{bmatrix} \tag{4.34}$$

where

$$Z = T^{\frac{1}{2}} E U^{\frac{1}{2}} - W \tag{4.35}$$
$$W = -T^{-\frac{T}{2}} F^T G L^T U^{-\frac{T}{2}} \tag{4.36}$$
$$P = V^{-\frac{T}{2}} G^T F T^{-\frac{1}{2}} \tag{4.37}$$

and E, T, U and V are as defined in 4.7, 4.19, 4.20 and 4.31 respectively.

2.
$$\gamma I - \bar{\mathcal{C}}_K \succ 0 \Leftrightarrow \|Y - W_{-,\gamma}\|_2 \leq 1 \tag{4.38}$$

where

$$\gamma^{-1} I + \gamma^{-2} P^T P = M_\gamma^T M_\gamma \tag{4.39}$$
$$M_\gamma = \left(\gamma^{-1} I + \gamma^{-2} P^T P\right)^{\frac{1}{2}} \tag{4.40}$$
$$W_\gamma = M_\gamma W \tag{4.41}$$
$$Y = M_\gamma T^{\frac{1}{2}} E U^{\frac{1}{2}} - W_{+,\gamma} \tag{4.42}$$

and $W_{+,\gamma}$ and $W_{-,\gamma}$ are the causal and strictly anticausal parts of W_γ. Here, M_γ is lower triangular, and positive-definite.

3. *Y is causal iff E is causal, where E can be found as follows:*

$$E = T^{-\frac{1}{2}} M_\gamma^{-1} (Y + W_{+,\gamma}) U^{-\frac{1}{2}} \tag{4.43}$$

4. *The condition in (4.38) is recognized as a level-1 suboptimal Nehari problem that approximates a strictly anticausal matrix $W_{-,\gamma}$ by a causal matrix Y.*

Proof. The proof follows from Theorem 4 in [81] and is omitted for brevity here. ∎

Using Lemma 4.4.2, problem (4.33) can be reformulated as a tractable optimization program:

$$\inf_{\substack{Z,Y\in\mathcal{K},\\ \gamma\geq 0,\\ X\succeq 0}} \gamma(r^2 - N(n+p)) + \mathrm{Tr}(X)$$

$$\text{s.t.} \begin{cases} \begin{bmatrix} X_{11} & X_{12} & \gamma I & 0 \\ X_{12}^T & X_{22} & 0 & \gamma I \\ \gamma I & 0 & \gamma I & -PZ \\ 0 & \gamma I & -Z^T P^T & \gamma I - Z^T Z \end{bmatrix} \succeq 0 \\ \|Y - W_{-,\gamma}\|_2 \leq 1 \end{cases} \tag{4.44}$$

$$= \inf_{\substack{Z,Y\in\mathcal{K},\\ \gamma\geq 0,\\ X\succeq 0}} \gamma(r^2 - N(n+p)) + \mathrm{Tr}(X)$$

$$\text{s.t.} \begin{cases} \begin{bmatrix} X_{11} & X_{12} & \gamma I & 0 & 0 \\ X_{12}^T & X_{22} & 0 & \gamma I & 0 \\ \gamma I & 0 & \gamma I & -PZ & 0 \\ 0 & \gamma I & -Z^T P^T & \gamma I & Z^T \\ 0 & 0 & 0 & Z & I \end{bmatrix} \succeq 0 \\ \|Y - W_{-,\gamma}\|_2 \leq 1 \end{cases} \tag{4.45}$$

where the last step follows from the Schur complement. Using (4.35), (4.43), and

$$H_\gamma = M_\gamma^{-1} W_{+,\gamma} - W \tag{4.46}$$

we establish our main theorem.

Theorem 4.4.3 (Tractable Formulation of DR-RO-MF)**.** *The distributionally robust* *regret-optimal control problem in the measurement feedback setting* (4.14) *reads:*

$$\begin{aligned}
&\inf_{\substack{Y\in\mathcal{K},\\ \gamma\geq 0,\\ X\succeq 0}} \gamma(r^2 - N(n+p)) + \mathrm{Tr}(X) \\
&\textit{s.t.} \begin{cases}
\begin{bmatrix} X_{11} & X_{12} & \gamma I & 0 & 0 \\ X_{12}^T & X_{22} & 0 & \gamma I & 0 \\ \gamma I & 0 & \gamma I & -P(*) & 0 \\ 0 & \gamma I & -(*)^T P^T & \gamma I & (*)^T \\ 0 & 0 & 0 & (*) & I \end{bmatrix} \succeq 0 \\
(*) = M_\gamma^{-1} Y + H_\gamma \\
\begin{bmatrix} I & (Y - W_{-,\gamma})^T \\ Y - W_{-,\gamma} & I \end{bmatrix} \succ 0
\end{cases}
\end{aligned} \tag{4.47}$$

The optimal controller K^* *is then obtained using* (4.8) *and* (4.43).

Sub-Optimal Problem

For a given value of γ, problem (4.47) can be simplified into a tractable SDP. In practical implementations, we can solve problem (4.47) by optimizing the objective function with respect to the variables Y and X while fixing γ, thus transforming the problem into an SDP, which can be solved using standard convex optimization packages. We then iteratively refine the value of γ until it converges to the optimal solution γ^*. This iterative process ensures that we obtain the best possible value for γ that minimizes the objective function in problem (4.47).

LQG and RO-MF Control Problems as Special Cases

Interestingly, LQG and RO control in the measurement feedback setting can be recovered from the DR-RO-MF control by varying the radius r which represents the extent of uncertainty regarding the accuracy of the nominal distribution in the ambiguity set. When $r \to 0$, the ambiguity set transforms into a singular set comprising solely the nominal distribution. Consequently, the problem simplifies into a stochastic optimal control problem under partial observability:

$$\inf_{K\in\mathcal{K}} \mathbb{E}_{P_0}[J(K, w, v)] \tag{4.48}$$

As $r \to \infty$, the ambiguity set transforms into the set of any disturbance generated adversarially and the optimal γ reaches its smallest possible value which is the operator norm of $\mathcal{C}_K$. This means that the problem reduces to the RO-MF control problem which we discussed in section 4.2.

4.5 Simulations

Flight Control

We focus on the problem of controlling the longitudinal flight of a Boeing 747 which pertains to the linearized dynamics of the aircraft, as presented in [114]. The linear dynamical system provided describes the aircraft's dynamics during level flight at an altitude of 7.57 miles and a speed of 593 miles per hour, with a discretization interval of 0.1 second. The state variables of the system encompass the aircraft's velocity along the body axis, velocity perpendicular to the body axis, angle between the body axis and the horizontal plane, and angular velocity. The inputs to the system are the elevator angle and thrust. The process noise accounts for variations caused by external wind conditions. The discrete-time state space model is:

$$A = \begin{bmatrix} 0.9801 & 0.0003 & -0.0980 & 0.0038 \\ -0.3868 & 0.9071 & 0.0471 & -0.0008 \\ 0.1591 & -0.0015 & 0.9691 & 0.0003 \\ -0.0198 & 0.0958 & 0.0021 & 1.000 \end{bmatrix}$$

$$B = \begin{bmatrix} -0.0001 & 0.0058 \\ 0.0296 & 0.0153 \\ 0.0012 & -0.0908 \\ 0.0015 & 0.0008 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We conduct all experiments using MATLAB, on a PC with an Intel Core i7-1065G7 processor and 16 GB of RAM. The optimization problems are solved using the CVX package [84].

We limit the horizon to $N = 10$. We take the nominal distribution P_0 to be Gaussian with mean $\mu_0 = 0$ and covariance $\Sigma_0 = I$, and we investigate various values for the radius r, specifically:

$$r \in \{0, 0.2, 0.4, 0.6, 0.8, 1, 1.5, 2, 4, 8, 16, 32, 126\}.$$

For each value of r, we solve the sub-optimal problem described in section 4.4, iterating over γ until convergence to γ^*.

To assess the performance of the controller, we compute the worst-case disturbance, which lies at a Wasserstein distance r from P_0, as discussed in theorem 4.4.1. Finally,

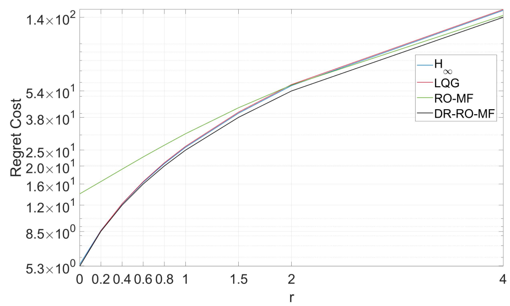


Figure 4.1: Controller costs for $r \in 0, 0.2, 0.4, 0.6, 0.8, 1, 1.5, 2, 4$.
At $r = 0$, the top-performing controllers are DR-RO-MF and LQG, exhibiting regret costs of 5.4. They are followed by H$_\infty$ with a regret cost of 5.9, and finally RO-MF with a regret cost of 13.8. The ranking of the controllers based on regret costs is: **DR-RO-MF=LQG**=5.34 $<$ **H$_\infty$**=5.47 $<$ **RO-MF**=13.8.
As r increases to 4, DR-RO-MF remains the best-performing controller with a regret of 141. It is followed by RO-MF with a regret of 144, H$_\infty$ with a regret of 154, and finally H2 with a regret of 156. The ranking of the controllers at $r = 4$ based on regret costs is: **DR-RO-MF**=141 $<$ **RO-MF**=144 $<$ **H$_\infty$**=154 $<$ **LQG**=156.

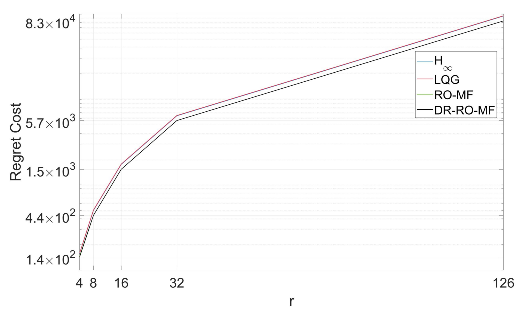


Figure 4.2: Controller costs for $r \in 4, 8, 16, 32, 126$.
At $r = 8$, the best-performing controller is DR-RO-MF with a regret of 437, which is closely comparable to the regret of the RO-MF controller of 438. They are followed by H$_\infty$ with a regret of 499, and finally H2 with a regret of 505. The ranking of controllers based on regret costs is as follows: **DR-RO-MF**=437 $\lesssim$ **RO-MF**=438 $<$ **H$_\infty$**=499 $<$ **LQG**=505.
When r increases to 126, which approximates the behavior of r approaching infinity, the order of the best-performing controllers remains unchanged: **DR-RO-MF=RO-MF**=8.33×10^4 $<$ **H$_\infty$**=9.50×10^4 $<$ **LQG**=9.57×10^4. DR-RO-MF and RO-MF controllers exhibit similar performance in this regime.

we compare the regret cost of the DR-RO-MF controller with that of the LQG, H_∞ [101], and RO-MF [81] controllers while considering the worst-case disturbance corresponding to the DR-RO-MF controller. The results are shown in Figures 4.1 and 4.2.

The DR-RO-MF controller achieves the minimum cost under worst-case disturbance conditions for any given value of r. When r is sufficiently small (less than 0.2), the cost of the DR-RO-MF controller closely approximates that of the LQG controller (figure 4.1). Conversely, for sufficiently large values of r (greater than 8), the cost of the DR-RO-MF controller closely matches that of the RO-MF controller (figure 4.2). These observations align with theoretical findings as elaborated in section 4.4.

Furthermore, it is worth noting that for large values of r (figure 4.2), the LQG controller yields the poorest results. Conversely, for small values of r (figure 4.1), the LQG controller performs on par with the DR-RO-MF controller, emerging as the best choice, as mentioned earlier. This discrepancy is expected since LQG control accounts only for disturbances drawn from the nominal distribution, assuming uncorrelated noise. On the other hand, RO-MF exhibits inferior performance when r is small (figure 4.1), but gradually becomes the top-performing controller alongside DR-RO-MF as r increases. This behavior arises from the fact that RO-MF is specifically designed for sufficiently large r. Lastly, note that the H_∞ cost lies between the costs of the other controllers, interpolating their respective costs.

Performance Under Adversarially Chosen Distribution

For any given causal controller K_c, an adversary can choose the worst-case distribution of disturbances for a fixed r as

$$\arg\max_{P\in\mathcal{P}} \mathbb{E}_P R(K_c, w, v) =: P_c, \tag{4.49}$$

where R is the regret as in (4.11). Denoting by $K_{\text{DR-RO-MF}}$ the optimal DR-RO-MF controller and by $P_{\text{DR-RO-MF}}$ the worst-case (adversarial) distribution corresponding to $K_{\text{DR-RO-MF}}$, we have that

$$\begin{aligned}
\mathbb{E}_{P_c} R(K_c, w, v) &= \max_{P\in\mathcal{P}} \mathbb{E}_P R(K_c, w, v), && (4.50)\\
&\geq \min_{K\in\mathcal{K}} \max_{P\in\mathcal{P}} \mathbb{E}_P R(K, w, v), && (4.51)\\
&= \mathbb{E}_{P_{\text{DR-RO-MF}}} R(K_{\text{DR-RO-MF}}, w, v), && (4.52)\\
&\geq \mathbb{E}_{P_c} R(K_{\text{DR-RO-MF}}, w, v), && (4.53)
\end{aligned}$$

where the first equality follows from (4.49) and the last inequality is due to the fact that $P_{\text{DR-RO-MF}}$ is the worst-case distribution for $K_{\text{DR-RO-MF}}$. In other words, DR-RO-MF controller is robust to adversarial changes in distribution as it yields smaller expected regret compared to any other causal controller K_c when the disturbances are sampled from the worst-case distribution P_c corresponding to K_c.

r	0.2	1	2	4	16	32
LQG(%)	0.860	8.17	14.8	21.9	28.5	29.3
RO-MF(%)	56.6	43.0	32.3	17.2	1.95	0.465

Table 4.1: Relative difference in % (as in (4.57)) between the expected regret of LQG/RO-MF and of DR-RO-MF controllers, under the worst-case disturbance of LQG/RO-MF, respectively, as in (4.49) for different values of r

The simulation results presented in Subsection 4.5 show that DR-RO-MF outperforms RO-MF, H_∞, and LQG (designed assuming disturbances are sampled from P_0) controllers under the worst-case distribution of the DR-RO-MF controller $P_{\text{DR-RO-MF}}$, i.e

$$\mathbb{E}_{P_{\text{DR-RO-MF}}} R(K_c, w, v) \geq \mathbb{E}_{P_{\text{DR-RO-MF}}} R(K_{\text{DR-RO-MF}}, w, v). \tag{4.54}$$

This directly implies that the theoretically expected inequality

$$\mathbb{E}_{P_c} R(K_c, w, v) \geq \mathbb{E}_{P_c} R(K_{\text{DR-RO-MF}}, w, v) \tag{4.55}$$

is validated and positively exceeded following the inequalities (4.53) and

$$\mathbb{E}_{P_c} R(K_c, w, v) \geq \mathbb{E}_{P_{\text{DR-RO-MF}}} R(K_c, w, v). \tag{4.56}$$

To further support our claims, we assess the performance of LQG and RO-MF controllers by measuring the relative reduction in expected regret when DR-RO-MF controller is utilized under the worst-case distributions corresponding to LQG and RO-MF controllers, respectively:

$$\frac{\mathbb{E}_{P_c} R(K_c, w, v) - \mathbb{E}_{P_c} R(K_{\text{DR-RO-MF}}, w, v)}{\mathbb{E}_{P_c} R(K_c, w, v)} \times 100, \tag{4.57}$$

where K_c is either LQG or RO-MF controller and P_c is the corresponding worst-case distribution. The results are shown in Table 4.1 for $r \in \{0.2, 1, 2, 4, 16, 32\}$.

Limitations

In our scenario with a relatively short planning horizon of $N = 10$, the cost reduction achieved by employing DR-RO-MF control, in comparison to traditional controllers such as LQG and H_∞, is moderate. However, it is anticipated that this reduction would become more pronounced with the utilization of a longer planning horizon. Unfortunately, in our experimental setup, we were restricted to using $N = 10$ due to computational limitations. Solving semi-definite programs involving large matrices is computationally inefficient, necessitating this constraint. In practice, this limitation can be overcome by implementing the controller in a receding horizon fashion, where the controller is updated every x time steps.

4.6 Conclusion

In conclusion, this paper extended the distributionally robust approach to regret-optimal control by incorporating the Wasserstein-2 distance [219] to handle cases of limited observability. The proposed DR-RO-MF controller demonstrated superior performance compared to classical controllers such as LQG and H_∞, as well as the RO-MF controller, in simulations of flight control scenarios. The controller exhibits a unique interpolation behavior between LQG and RO-MF, determined by the radius r that quantifies the uncertainty in the accuracy of the nominal distribution. As the time horizon increases, solving the tractable SDP to which the solution reduces, becomes more challenging, highlighting the practical need for a model predictive control approach. Overall, the extended distributionally robust approach presented in this paper holds promise for robust and effective control in systems with limited observability.

Chapter 5

FINITE-HORIZON DISTRIBUTIONALLY ROBUST KALMAN FILTERING

5.1 Problem Setup

In this section, we formulate the distributionally robust filtering problem for both finite and infinite horizon settings. To this end, consider the following state-space model:

$$\begin{aligned} x_{t+1} &= Ax_t + Bw_t, \\ y_t &= C_y x_t + v_t, \\ s_t &= C_s x_t, \end{aligned} \tag{5.1}$$

At time $t \in \mathbb{N}$, let $x_t \in \mathbb{R}^{\mathsf{d}_x}$ denote the unobserved *latent state*, $y_t \in \mathbb{R}^{\mathsf{d}_y}$ the *measurement*, $s_t \in \mathbb{R}^{\mathsf{d}_s}$ the unobserved *target signal* to be estimated, $w_t \in \mathbb{R}^{\mathsf{d}_w}$ the *process noise*, and $v_t \in \mathbb{R}^{\mathsf{d}_v}$ the *measurement noise*. The combined process-measurement noise sequence constitutes the *exogenous disturbance*. The setup presented above is quite general and widely adopted in the estimation and filtering literature [101], [120]. The usual state estimation problem is a specific instance of this setup with $C_s = I$. Moreover, we assume that (A, C_y) and (A, C_s) are detectable and (A, B) is controllable.

We take a global view of the dynamics (5.1) by treating the entire signal trajectories over a fixed time horizon $T > 0$ as large column vectors. Concretely, we define the measurements vector $\mathbf{y}_T \coloneqq [y_0; y_1; \dots; y_{T-1}] \in \mathbb{R}^{T\mathsf{d}_y}$, the target signal vector $\mathbf{s}_T \coloneqq [s_0; s_1; \dots; s_{T-1}] \in \mathbb{R}^{Td_s}$, the process noise vector $\mathbf{w}_T \coloneqq [x_0; w_0; \dots; w_{T-2}] \in \mathbb{R}^{\mathsf{d}_x + (T-1)\mathsf{d}_w}$, and the measurement noise vector $\mathbf{v}_T \coloneqq [v_0; v_1; \dots; v_{T-1}] \in \mathbb{R}^{T\mathsf{d}_v}$. Notice that the initial state x_0 is considered unknown and included in the vector of process noise, $\mathbf{w}_T$, for convenience. With the prevailing notation, the state-space dynamics can be represented compactly as a *causal linear measurement model*:

$$\begin{aligned} \mathbf{y}_T &= \mathcal{H}_T \mathbf{w}_T + \mathbf{v}_T, \\ \mathbf{s}_T &= \mathcal{L}_T \mathbf{w}_T, \end{aligned} \tag{5.2}$$

where $\mathcal{H}_T$ and $\mathcal{L}_T$ are both *block causal* (*i.e.*, block lower-triangular) matrices. These matrices can be constructed easily from the state-space parameters (A, B, C_y, C_u) (see **??**). This representation is quite general and can be extended immediately for time-varying state-space models with appropriately constructed matrices $\mathcal{H}_T$ and $\mathcal{L}_T$.

Letting the stacked column vector $\boldsymbol{\xi}_T := [\mathbf{w}_T; \mathbf{v}_T] \in \Xi_T$ denote the combined disturbances where $\Xi_T := \mathbb{R}^{\mathsf{d}_x + (T-1)\mathsf{d}_w + T\mathsf{d}_v}$, the disturbances $\boldsymbol{\xi}_T$ are distributed according to an unknown distribution $\mathbb{P}_T \in \mathscr{P}_2(\Xi_T)$. Note that the disturbances can be arbitrarily correlated in general. Wlog, we will assume $\boldsymbol{\xi}_T$ to be zero-mean for convenience. Our main assumption for $\mathbb{P}_T$ is as follows:

Assumption 5.1.1. The true distribution $\mathbb{P}_T$ of disturbances $\boldsymbol{\xi}_T$ resides in a W_2-ball,

$$\mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T) := \left\{ \mathbb{P}_T \in \mathscr{P}_2(\Xi_T) \mid \mathsf{W}_2(\mathbb{P}_T,\, \mathbb{P}_{\circ,T}) \leq \rho_T \right\}, \tag{5.3}$$

where $\rho_T > 0$ is a specified radius and $\mathbb{P}_T^\circ \in \mathscr{P}_2(\Xi_T)$ is a given nominal disturbance distribution.

Remark 5.1.2. Although the state-space parameters (A, B, C_y, C_s) are assumed to be known perfectly, uncertainty in them can be incorporated into the disturbances without loss of generality.

The Finite-Horizon Distributionally Robust Filtering

A *filtering policy* $\pi_T := \{\pi_t \mid t = 0, \dots, T-1\}$ is a sequence of mappings that generate estimates $\widehat{s}_t$ of s_t from the past and present measurement as $\widehat{s}_t = \pi_t(y_t, y_{t-1}, \dots, y_0)$. In particular, we focus on *linear filtering policies* $\mathcal{K}_T : \mathbf{y}_T \mapsto \widehat{\mathbf{s}}_T$ such that $\widehat{\mathbf{s}}_T = \mathcal{K}_T \mathbf{y}_T$ where $\widehat{\mathbf{s}}_T := [\widehat{s}_0; \widehat{s}_1; \dots; \widehat{s}_{T-1}]$ is the the column vector of estimates. We denote the class of all such policies by $\mathscr{K}_T$, defined as

$$\mathscr{K}_T \triangleq \left\{ \mathcal{K}_T \in \mathbb{R}^{T\mathsf{d}_s \times T\mathsf{d}_y} \mid \mathcal{K}_T \text{ is block lower-triangular} \right\}. \tag{5.4}$$

The restriction to linear filters is a common strategy in estimation literature, as general nonlinear estimators can be challenging to compute [120]. Additionally, linear filters are optimal for Gaussian processes. In Theorem 5.2.1, we establish the optimality of linear filters when the nominal is Gaussian.

For a filtering policy $\mathcal{K}_T$, let $\mathbf{e}_T(\boldsymbol{\xi}_T, \mathcal{K}_T) \triangleq \widehat{\mathbf{s}}_T - \mathbf{s}_T = \mathcal{T}\mathcal{K}_T \boldsymbol{\xi}_T$ be the estimation error where $\mathcal{T}_{\mathcal{K}_T} : \boldsymbol{\xi}_T \mapsto \mathbf{e}_T$ is the *error transfer operator* defined as

$$\mathcal{T}_{\mathcal{K}_T} \triangleq \begin{bmatrix} \mathcal{K}_T \mathcal{H}_T - \mathcal{L}_T & \mathcal{K}_T \end{bmatrix}. \tag{5.5}$$

Given that the true distribution of disturbances is unknown, we focus on minimizing the *worst-case mean-squared error (MSE)* across all distributions within the ambiguity set $\mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)$, namely,

$$\mathsf{E}_T(\mathcal{K}_T, \rho_T) \triangleq \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \mathcal{K}_T)\|^2\right]. \tag{5.6}$$

where $\mathbb{E}_{\mathbb{P}_T}$ denotes the expectation under the distribution $\mathbb{P}_T$. We state the distributionally robust Kalman filtering problem for the finite-horizon setting as follows:

Problem 5.1.3 (W_2-DR-KF over a finite-horizon)**.** For a given time-horizon $T>0$ and a radius $\rho_T > 0$, find a casual filtering policy, $\mathcal{K}_T \in \mathscr{K}_T$, that minimizes the worst-case MSE defined in (5.6), *i.e.*,

$$\inf_{\mathcal{K}_T \in \mathscr{K}_T} \mathsf{E}_T(\mathcal{K}_T, \rho_T) = \inf_{\mathcal{K}_T \in \mathscr{K}_T} \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \mathcal{K}_T)\|^2\right]. \tag{5.7}$$

Remark 5.1.4. The causality constraint on the estimates $\widehat{\mathbf{s}}_T$ is crucial for filtering. Without causality enforced, Problem 5.1.3 essentially reduces to a standard estimation problem as the nominal non-causal estimator is optimal for any $\rho_T>0$ (Lemma 5.2.3).

5.2 Tractable Convex Reormulation

In this section, we provide tractable formulations for the finite horizon W_2-DR-KF problem. In Theorem 5.2.4, we present an SDP formulation for the finite-horizon problem 5.1.3. We also characterize the optimal estimator and the worst-case distribution. The proofs of the theorems presented in this section are deferred to the Appendix.

Before proceeding with the main theorems, we present a minimax theorem establishing the optimality of linear filtering policies for Gaussian nominal distributions.

Theorem 5.2.1 (Minimax duality)**.** *Let $T>0$ be a fixed horizon and Π_T be the class of non-linear causal estimators. Suppose that the nominal $\mathbb{P}_T^\circ$ is Gaussian. Then, the following holds:*

$$\inf_{\pi_T \in \Pi_T} \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right] = \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \inf_{\pi_T \in \Pi_T} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right], \tag{5.8}$$

Moreover, (5.8) *admits a saddle point $(\pi_T^\star, \mathbb{P}_T^\star)$ such that the worst-case distribution $\mathbb{P}_T^\star$ is Gaussian and the optimal causal filter $\pi_T^\star$ is linear, i.e., $\pi_T^\star \in \mathscr{K}_T$.*

For simplicity and clarity, we make the following assumption for the remainder of this paper.

Assumption 5.2.2. The nominal disturbances are uncorrelated, *i.e.*, $\mathbb{E}_{\mathbb{P}_T^\circ}[\boldsymbol{\xi}_T \boldsymbol{\xi}_T^*] = \mathcal{I}_T$ for any $T>0$.

An SDP for the Finite-Horizon Filtering

In this section, we state the SDP formulation of Problem 5.1.3 for a fixed horizon $T > 0$. To this end, we first state the following lemma identifying the optimal non-causal estimator.

Lemma 5.2.3. *Under the Assumption 5.2.2, $\mathcal{K}_T^\circ \triangleq \mathcal{L}_T\mathcal{H}_T^*(\mathcal{I}_T+\mathcal{H}_T\mathcal{H}_T^*)^{-1}$ is the unique, optimal, non-causal estimator minimizing the worst-case MSE in* (5.6) *for any $\rho_T \geq 0$.*

This result highlights the triviality of non-causal estimation as opposed to causal estimation. In Theorem 5.2.4, we demonstrate that the finite-horizon W_2-DR-KF problem 5.1.3 reduces to an SDP.

Theorem 5.2.4 (An SDP formulation for finite-horizon W_2-DR-KF)**.** *Let the horizon $T>0$ be fixed and denote $\mathcal{T}_{\mathcal{K}_T^\circ}\mathcal{T}_{\mathcal{K}_T^\circ}^* := \mathcal{L}_T(\mathcal{I}_T+\mathcal{H}_T^*\mathcal{H}_T)^{-1}\mathcal{L}_T^*$. Then, the Problem 5.1.3 reduces to the following SDP*

$$\inf_{\substack{\mathcal{K}_T\in\mathscr{K}_T,\\ \gamma\geq 0,\, \mathcal{X}_T\in\mathbb{S}_+^{T\mathsf{d}_s}}} \gamma(\rho_T^2-\mathrm{Tr}(\mathcal{I}_T))+\mathrm{Tr}(\mathcal{X}_T) \quad \text{s.t.} \quad \begin{bmatrix} \mathcal{X}_T & \gamma\mathcal{I}_T & 0 \\ \gamma\mathcal{I}_T & \gamma\mathcal{I}_T-\mathcal{T}_{\mathcal{K}_T^\circ}\mathcal{T}_{\mathcal{K}_T^\circ}^* & \mathcal{K}_T-\mathcal{K}_T^\circ \\ 0 & (\mathcal{K}_T-\mathcal{K}_T^\circ)^* & (\mathcal{I}_T+\mathcal{H}_T\mathcal{H}_T^*)^{-1} \end{bmatrix} \succcurlyeq 0.$$

Moreover, the worst-case disturbance $\boldsymbol{\xi}_T^\star$ can be identified from the nominal disturbances $\boldsymbol{\xi}_T^\circ$ as

$$\boldsymbol{\xi}_T^\star = (\mathcal{I}_T - \gamma_\star^{-1}\mathcal{T}_{\mathcal{K}_T^\star}^*\mathcal{T}_{\mathcal{K}_T^\star})^{-1}\boldsymbol{\xi}_T^\circ, \tag{5.9}$$

where $\gamma_\star > 0$ and $\mathcal{K}_T^\star$ are the optimal solutions.

Remark 5.2.5. As $\rho_T \to \infty$, the ambiguity set covers all bounded energy disturbances, and the optimal W_2-DR-KF policy, $\mathcal{K}_T^\star$, recovers the $\mathscr{H}_\infty$-filter. Conversely, as $\rho_T \to 0$, the ambiguity set reduces to the singleton $\mathbb{P}_T^\circ$, and $\mathcal{K}_T^\star$ recovers the Kalman filter. Thus, adjusting ρ_T allows the DR filter to interpolate between the conservative $\mathscr{H}_\infty$-filter and the nominal Kalman filter.

Notice that the variable dimension of the SDP in Theorem 5.2.4 scales with the horizon T, which can be prohibitive for practical implementation for longer horizons.

Corollary 5.2.6. *The time complexity of interior-point method for solving the SDP in Theorem 5.2.4 with accuracy $\epsilon > 0$ is $\widetilde{O}(\max(\mathsf{d}_y, \mathsf{d}_s)^6\, T^6 \log(1/\epsilon))$.*

Chapter 6

DISTRIBUTIONALLY ROBUST APPROACH TO SHANNON LIMITS

6.1 Introduction

Literature context

Rapid technological advancements in recent years have spurred a significant increase in data production and usage. Major technological forces driving this data era include 5G/6G, the Internet of Things, machine learning, and autonomous vehicles. Core to these technologies are compression and reconstruction algorithms, which provide essential means for data storage and transmission. For example, video streaming services require efficient data compression, yet the source distribution can vary greatly. Similarly, autonomous vehicles depend on sensor data, which is prone to distributional shifts that can lead to critical errors if not adequately addressed. In machine learning, models are often trained on a dataset that has undergone compression and decompression. These applications underscore the need for compression and error correction schemes that are resilient to distributional changes.

Variable-length universal codes that adapt their encoding rate to the source / noise distribution they observe have been a topic of longstanding interest in the information theory community. Perhaps the most famous of these is the Lempel-Ziv family of lossless codes [236], [262], [263] that learn the source distribution on the fly, while asymptotically achieving the entropy rate of the source. Lempel-Ziv codes have also been extended to apply to lossy compression [245], [246]. An approach to universal lossy compression that utilizes Markov chain Monte Carlo with the reconstruction alphabet changing over time is proposed in [13] (continuous sources),[118] (finite alphabet sources). Variable-rate universal codes for channel coding are presented in [27], [54], [155], [209], [221].

The idea that enables the aforementioned coding strategies [13], [27], [54], [118], [155], [209], [221], [236], [245], [246], [262], [263] is that of variable-length coding, i.e., the length of the codeword being transmitted is adapted to the data that is being seen. However, many physical systems and much of coding theory require all codewords to be of the same length. In that scenario, the relevant figure of merit is the worst-case performance over a family of source / noise distributions, known as *compound* rate-distortion function / channel capacity-cost function, and the corresponding asymptotic fundamental limits are given by minimax convex optimization problems [195], [22]. In the context of compression, [194] considers the distributional class of stationary random processes with bounded 4^{th} order moment, whereas [183] examines a class of sources determined by spectral capacities. In the context of error correction, [23] extends the compound channel capacity formula of [22] to random coding, to arbitrarily varying channels (AVCs), and to adversarial channels, [190] shows the worst-case capacity of a class of arbitrarily varying Gaussian channels, and [73] studies compound capacity of broadcast channels; see [142] for a survey.

In this work, we employ Wasserstein-2 (W_2) distance as the measure of distributional shifts from a nominal distribution to define a distributional family. The W_2-distance between distributions P_1, P_2 on $\mathbb{R}^d$ is defined as [231]

$$\mathsf{W}_2(P_1, P_2) := \left(\inf_{P_{XY} \in \Pi(P_1, P_2)} \mathbb{E}\left[\|X - Y\|^2\right] \right)^{\frac{1}{2}}, \tag{6.1}$$

where $\Pi(P_1, P_2)$ denotes the set of all joint distributions with marginals P_1 and P_2.

In contrast to other commonly used statistical distances such as total-variation distance, Kullback–Leibler divergence, Hellinger distance, the W_2-distance incorporates information from the geometric structure of the underlying domain, making it more suitable for handling structured real-world data. Furthermore, the W_2-distance accounts for the cost of transporting mass from one probability distribution to another. This makes the W_2-distance a particularly suitable tool for quantifying robustness in communication systems, as it reflects the potential impact of variations in signal distributions on the fidelity of communication. For instance, in scenarios where small perturbations in a signal could lead to disproportionately large errors, the W_2-distance provides a measure of how these perturbations affect system performance. This makes it a natural choice for modeling scenarios where the goal is to maintain high fidelity in the presence of uncertain or varying signal distributions, thereby enhancing the reliability and efficiency of communication systems under distributionally robust frameworks.

Owing to its geometric interpretability and tractable formulation, W_2-distance has recently gained popularity as a statistical distance in diverse fields such as control [125], filtering [203], and machine learning [134]. In [219], a controller that minimizes the worst-case regret across all disturbance distributions within a W_2-ambiguity set is proposed. The controller's parameters are obtained via a solution to a semi-definite program (SDP), which is formulated leveraging the tractability of W_2-distance. Unlike prior robust control approaches, such as H_∞ [101], where the controller is designed to minimize worst-case cost against an adversarially generated disturbance, and regret-optimal control [80], where the controller is designed to minimize the regret with respect to a hypothetical optimal noncausal controller, the controller in [219] is designed to attain the desired level of robustness via adjusting the size of the ambiguity set. Additionally, [219] demonstrates that the worst-case disturbance follows a Gaussian distribution if the nominal distribution is Gaussian. Subsequent contributions [89] and [125] extend the framework in [219] to control scenarios involving measurement noise and infinite horizon, respectively. In the domain of lossy source coding, [144] trains a deep neural network compressor that achieves distributional robustness by incorporating W_2 transportation cost into the optimization problem. The W_2-distance has also been used to measure the perceptual quality of the decompressed message [26] using what is called the Rate-Distortion-Perception (RDP) function. The work [202] characterizes the analytical bounds for the Gaussian RDP function.

Contributions

We investigate the minimax rate-distortion function (RDF) formulated in [195] and the minimax capacity formulated in [22] for an unknown source / noise distribution residing within a W_2-ambiguity set centered at a given nominal distribution. In the case of source compression, we assume that the source distributions in the ambiguity set have a finite $2+\epsilon$-th moment for an arbitrarily small $\epsilon > 0$. The finiteness of $(2+\epsilon)$-th moment is a common assumption on the source distributions in rate-distortion theory to control the growth of the mean square error distortion (see [182, Sec. 23.3]).

In the case of a Gaussian nominal source distribution, we show that the worst-case source / noise distribution in the W_2 ambiguity set is also Gaussian. To do so, we leverage the Gelbrich bound [78, Thm. 2.1] on the W_2 distance and the Gaussian saddle point properties of mutual information [182, Thm. 5.11]. Our analysis reveals an expression of the minimax rate-distortion function / channel capacity solely in terms of the covariance matrices of distributions within the ambiguity set and the radius of the W_2-ball. Adjusting the radius of the W_2-ball enables a gradual interpolation from a known nominal source distribution (at a radius of 0) to an uncertain distribution as the radius increases. In the scalar case, we derive closed-form expressions demonstrating the effect of the varying radius on those fundamental limits.

Notation. The set of real numbers is denoted by $\mathbb{R}$. For a real-valued function f, we denote $f^+(\cdot) \coloneqq \max\{0, f(\cdot)\}$. For a vector $x \in \mathbb{R}^d$, we denote by $\|x\|$, the Euclidean norm of x. For a matrix $A \in \mathbb{R}^{d\times m}$, its transpose is denoted by $A^T \in \mathbb{R}^{m\times d}$. For square matrices, $\mathrm{tr}(\cdot)$ and $|\cdot|$ are the trace and the determinant operation. The notations $\mathbb{S}_+^d$ and $\mathbb{S}_{++}^d$ represent the sets of $d \times d$ symmetric, positive semidefinite and positive definite matrices, respectively. For two positive semidefinite matrices, $\succeq$ denotes the Löwner order. The set of all probability distributions over $\mathbb{R}^n$ with bounded p^{th}-moment is denoted by $\mathcal{P}_p(\mathbb{R}^n)$. For a random variable X over $\mathbb{R}^n$, we denote its probability distribution by P_X. The conditional distribution of a random variable Y given X is denoted by $P_{Y|X}$. The expectation with respect to the distribution P_X is denoted by $\mathbb{E}_{P_X}[\cdot]$. The Gaussian distribution with mean $\mu \in \mathbb{R}^d$ and covariance $\Sigma \in \mathbb{S}_+^d$ is denoted by $\mathcal{N}(\mu, \Sigma)$. The mutual information between random variables X and Y is denoted by $I(X;Y)$. The differential entropy of X is denoted by $h(X)$, and the conditional differential entropy of X given Y is given by $h(X|Y)$.

6.2 Main results

In this section, we state our main results: Theorem 6.2.1 shows an SDP for the worst-case RDF defined in (6.2), below, and Theorem 6.2.3 shows an SDP for the worst-case capacity, defined in (6.14).

Source Compression

Fix $\epsilon > 0$. Consider a source that generates i.i.d. symbols $X_i \in \mathbb{R}^d$ following an unknown probability distribution $P_X \in \mathcal{S}$, where $\mathcal{S} \subseteq \mathscr{P}_{2+\epsilon}(\mathbb{R}^d)$ is a distributional family. An n-block of source symbols $X^n = (X_1, \ldots, X_n)$ is mapped into one of the $\exp(nR)$ distinct codewords $\widehat{X}^n = (\widehat{X}_1, \ldots, \widehat{X}_n)$ while satisfying the average distortion constraint $\frac{1}{n}\sum_{i=1}^n \mathbb{E}[\|X_i - \widehat{X}_i\|^2] \leq D$.

The minimum universally achievable coding rate $R_{\mathcal{S},n}(D)$ is defined as the minimum R at which such a mapping exists, regardless of the $P_X \in \mathcal{S}$. The minimum asymptotically achievable universal coding rate is the operational compound RDF [195], $R_{\mathcal{S}}(D) = \limsup_{n\to\infty} R_{\mathcal{S},n}(D)$. Sakrison [195] established the following single-letter formula for the compound RDF:

$$R_{\mathcal{S}}(D) = \inf_{\substack{P_{\widehat{X}|X}\colon \mathbb{R}^d \mapsto \mathbb{R}^d \\ \mathbb{E}\left[\|X-\widehat{X}\|^2\right] \le D}} \sup_{P_X \in \mathcal{S}} I(X;\widehat{X}). \tag{6.2}$$

The achievability result of the coding theorem in (6.2) assumes that the class of distributions $\mathcal{S}$ is compact, a notion defined by [195] as follows: a class of distributions $\mathcal{S}$ is compact if, for any $\epsilon > 0$, there is a totally bounded set $\mathcal{K} \subset \mathbb{R}^d$ and a function $f\colon \mathbb{R}^d \to \mathbb{R}^d$ with range $\mathcal{K}$ such that,

$$\sup_{P_X \in \mathcal{S}} \mathbb{E}_{P_X}\left[\|X - f(X)\|^2\right] \le \epsilon. \tag{6.3}$$

Any class of distributions with bounded $2+\epsilon$ order moments for $\epsilon > 0$ satisfies the compactness definition given in [195]. However, the distributions in the W_2-ball do not necessarily satisfy this condition. Therefore, to ensure that the function on the right side of (6.2), whose computation constitutes one of our main results, has an operational meaning, we consider only those distributions in the W_2-ball that have bounded $2+\epsilon$ order moments.

Our main result for source compression computes the compound rate-distortion function (6.2) for the W_2-ambiguity set of distributions $\mathcal{S} = \left\{P \in \mathscr{P}_{2+\epsilon}(\mathbb{R}^d) \mid \mathsf{W}_2(P,\, P_\circ) \le r\right\}$. For brevity, we denote the compound rate-distortion function with such a choice of the ambiguity set by

$$R_{P_\circ}(D, r) \coloneqq R_{\mathcal{S}}(D). \tag{6.4}$$

If $r = 0$, (6.4) reduces to the rate-distortion function of the known distribution $P_\circ$. In our analysis, we assume a Gaussian nominal distribution.

To set the stage for our result, we denote the Bures-Wasserstein metric on the positive semi-definite cone $\mathbb{S}^d_+$ [20] as,

$$\mathsf{BW}(\Sigma, \Sigma_\circ) \coloneqq \Big(\operatorname{tr}(\Sigma) + \operatorname{tr}(\Sigma_\circ) - 2\operatorname{tr}(\Sigma^{\frac{1}{2}}\Sigma_\circ\Sigma^{\frac{1}{2}})\Big)^{1/2}. \tag{6.5}$$

Moreover, note that the RDF of a $d-$dimensional multivariate Gaussian vector $X \sim \mathcal{N}(0,\Sigma)$ is given by [133, Eqn. (13)]:

$$R_{\mathcal{N}(0,\Sigma)}(D) = \inf_{\substack{A \in \mathbb{R}^{d\times d}, \Sigma_Z \succeq 0: \\ \operatorname{tr}((A-I)\Sigma(A-I)^T + \Sigma_Z) \le D}} \frac{1}{2}\log\frac{\left|A\Sigma A^T + \Sigma_Z\right|}{|\Sigma_Z|}. \tag{6.6}$$

The reconstructed message $\widehat{X}$ follows the forward law $\widehat{X} = AX + Z$, where $Z \sim \mathcal{N}(0, \Sigma_Z)$ is independent of X. The solution to (6.6) is given by reverse-waterfilling on the eigenvalues [41, Thm. 10.3.3]. If $d = 1$, the RDF in (6.6) takes the form (6.9), below.

Theorem 6.2.1 (Compound RDF for W_2 ambiguity set)**.** *The compound RDF* (6.4) *with a Gaussian center* $P_\circ = \mathcal{N}(0, \Sigma_\circ)$ *is given as,*

$$R_{P_\circ}(D, r) = \sup_{\substack{\Sigma \succeq 0, \\ \mathsf{BW}(\Sigma, \Sigma_\circ) \leq r}} R_{\mathcal{N}(0,\Sigma)}(D), \tag{6.7}$$

where the function $R_{P_\circ}(D, r)$ *is achieved by a Gaussian* P_X.

Proof. We first refer to [195] to establish strong duality, see Lemma 6.2.2, below. We then present an upper bound on the compound RDF and show that it is achieved by a Gaussian source distribution in the W_2-ball if the nominal source distribution $P_\circ$ is Gaussian. This extends the classical result [133, Eqn. (13)] showing that, among all distributions with a fixed second-order moment, a Gaussian source achieves the largest single source RDF (See Section 6.3 for details). ■

Lemma 6.2.2 (Strong duality of the compound RDF)**.** *The compound RDF* (6.4) *admits a dual formulation given by*

$$R_{P_\circ}(D, r) = \sup_{\substack{P_{X|X_\circ} : \mathbb{R}^d \mapsto \mathbb{R}^d \\ P_X \in \mathscr{P}_{2+\epsilon}(\mathbb{R}^d) \\ \mathbb{E}\left[\|X - X_\circ\|^2\right] \leq r^2}} \inf_{\substack{P_{\hat{X}|X} : \mathbb{R}^d \mapsto \mathbb{R}^d \\ \mathbb{E}\left[\|X - \widehat{X}\|^2\right] \leq D}} I(X; \widehat{X}), \tag{6.8}$$

where $X_\circ \sim P_\circ$.

Proof. The result follows via an application of [195, Thm. Source Encoding] to mean square error and a W_2 ambiguity set. ■

In (6.8), the nominal source $X_\circ \sim P_\circ$, the worst-case source X and the reconstructed signal $\widehat{X}$ can be regarded as a Markov chain:

$$X_\circ \xrightarrow{P_{X|X_\circ}} X \xrightarrow{P_{\hat{X}|X}} \widehat{X}.$$

We can interpret (6.8) as a Nash equilibrium of a zero-sum game between two competing channels: while the channel $P_{X|X_\circ}$ adversarially maximizes the compression rate from X to $\widehat{X}$ under the transportation cost $\mathbb{E}\|X - X_\circ\|^2 \leq r^2$ from the nominal source to "deceive" the encoder-decoder channel $P_{\hat{X}|X}$, the channel $P_{\hat{X}|X}$ minimizes the compression rate from X to $\widehat{X}$ under the distortion constraint $\mathbb{E}\|X-\widehat{X}\|^2 \leq D$ by assuming the worst-case source from $P_{X|X_\circ}$.

To elucidate the tradeoff exposed in Theorem 6.2.1, consider a scalar-source RDF with an unknown distribution from a W_2-ambiguity set centered at $P_\circ = \mathcal{N}(0, \sigma_\circ^2)$. First, recall the RDF of a scalar Gaussian source [41, Eqn. (10.36)]:

$$R_{\mathcal{N}(0,\sigma^2)}(D) = \frac{1}{2}\log^+ \frac{\sigma^2}{D}. \tag{6.9}$$

Second, observe that the Bures-Wasserstein distance (6.5) between scalars is,

$$\mathsf{BW}(\sigma^2, \sigma_\circ^2) = |\sigma - \sigma_\circ|. \tag{6.10}$$

We now apply Theorem 6.2.1. The compound RDF for the scalar case, using (6.7), is given as,

$$R_{\mathcal{N}(0,\sigma_\circ^2)}(D, r) = \sup_{|\sigma-\sigma_\circ|\leq r} R_{\mathcal{N}(0,\sigma^2)}(D) \tag{6.11}$$

$$= \sup_{|\sigma-\sigma_\circ|\leq r} \frac{1}{2}\log^+ \frac{\sigma^2}{D} \tag{6.12}$$

$$= \frac{1}{2}\log^+ \frac{(\sigma_\circ + r)^2}{D}. \tag{6.13}$$

The expression (6.13) is immediate due to the fact that the logarithm is monotonic. Note that for $r = 0$, the compound RDF is the Shannon RDF for $P_\circ$, and for $r > 0$, it is the Shannon RDF for $P_X = \mathcal{N}(0, (\sigma_\circ + r)^2)$. Thus, as the radius of the ambiguity set increases, the required number of bits to achieve the same distortion increases.

Channel Coding

Consider transmission of an equiprobable message $W \in \{1, \ldots, \exp(nR)\}$ over an i.i.d. additive noise channel, which, upon receiving $X_i \in \mathbb{R}^d$, outputs $Y_i \in \mathbb{R}^d$, where $Y_i = HX_i + Z_i$, with $Z_i \sim P_Z$ independent of each other and of $X^i = (X_1, \ldots, X_i)$, and H is fixed. Here, $P_Z \in \mathcal{S}$ is fixed but unknown. The encoder injectively maps W to a codeword X^n under the average power constraint $\frac{1}{n}\sum_{i=1}^n \mathbb{E}\|X_i\|^2 \leq B$. The codeword is transmitted over the channel, and, upon receiving an n-block of channel outputs $Y^n = (Y_1, \ldots, Y_n)$, the decoder outputs $\widehat{W}$, an estimate of W.

The maximum universally achievable coding rate $C_{\mathcal{S},n}(B, \epsilon)$ compatible with average error probability $\mathbb{P}\left[W \neq \widehat{W}\right] \leq \epsilon$ is defined as the maximum rate R for which such an encoder-decoder mapping exists, regardless of $P_Z \in \mathcal{S}$. The maximum asymptotically

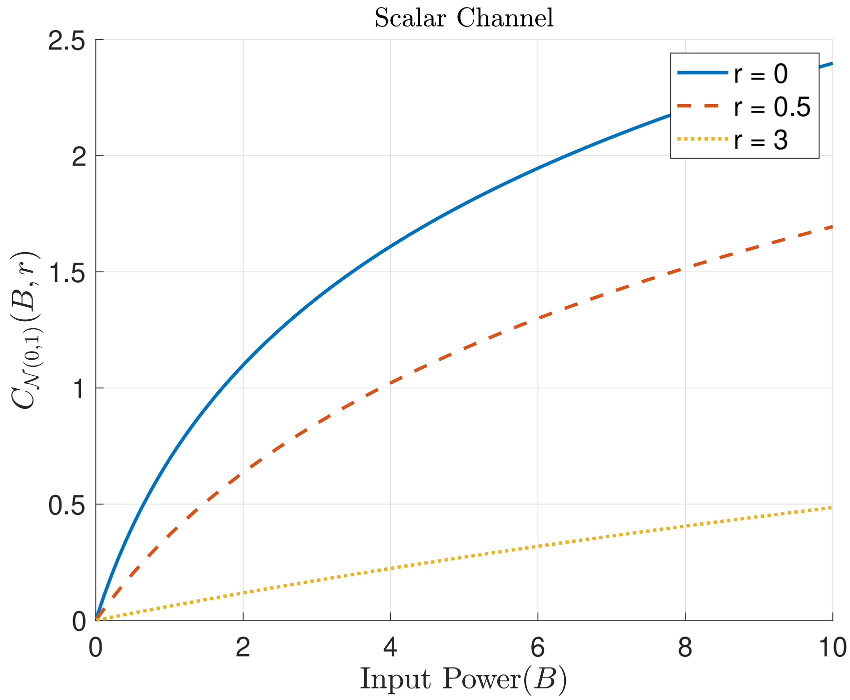


Figure 6.1: The scalar compound capacity for $P_\circ = \mathcal{N}(0,1)$ for $r = 0$. The scalar compound capacity for $r = 0$ is the Shannon capacity for $P_\circ$ and for $r > 0$ is the Shannon capacity for $P_Z = \mathcal{N}(0,(1+r)^2)$. As the radius of the ambiguity set increases, the required input power B to achieve the same capacity increases.

achievable universal coding rate is the operational compound capacity-cost function [22], $C_{\mathcal{S},n}(B) = \lim_{\epsilon\to 0} \liminf_{n\to\infty} C_{\mathcal{S},n}(B,\epsilon)$. Blackwell [22] established a single-letter formula for the compound capacity-cost function of a discrete memoryless channel, which was extended to continuous alphabet domain in [42], [154]:

$$C_{\mathcal{S}}(B) = \sup_{\substack{P_X \text{ on } \mathbb{R}^d:\\ \mathrm{tr}(\Sigma)\le B}} \inf_{P_Z\in\mathcal{S}} I(X; HX+Z), \tag{6.14}$$

where Σ is the covariance of $X \sim P_X$. Our main result for channel coding computes the compound capacity (6.14) for the W_2-ambiguity set of additive channel noises

$$\mathcal{S} = \left\{ P \in \mathscr{P}_2(\mathbb{R}^d) \mid \mathsf{W}_2(P, P_\circ) \le r \right\}. \tag{6.15}$$

Here, we assume that the nominal distribution $P_\circ \in \mathscr{P}_2(\mathbb{R}^d)$ to ensure weak compactness of $\mathcal{S}$ [253, Thm. 1]. This assumption entails $P \in \mathscr{P}_2(\mathbb{R}^d)$ [253, Lem. 1] for all P in the W_2 ball. For brevity, we denote the compound capacity-cost function with such a choice of the ambiguity set by

$$C_{P_\circ}(B,r) \coloneqq C_{\mathcal{S}}(B). \tag{6.16}$$

Our problem setting satisfies the regularity conditions [42, H1-H4]:

1. The input and output alphabets as well as the set of states (which correspond to the probability measures in the W_2 ball equipped with the W_2 distance) are separable metric spaces, and the output alphabet is complete.

2. The channel depends continuously on the input and states, *i.e.*, for a sequence of inputs $X_n \to X$ and states $P_{Z_n} \to P_Z$, $P_{Y|X,P_Z}(\cdot|X_n, P_{Z_n})$ converges weakly to $P_{Y|X,P_Z}(\cdot|X, P_Z)$. Here, $P_{Y|X,P_Z}(\cdot|X, P_Z)$ represents a channel with a given input X and a fixed noise distribution P_Z in the W_2 ball. This follows from the fact that the W_2-ball is weakly compact whenever the nominal distribution has a finite 2-nd order moment [253, Thm. 1].

3. The constraint function $\|\cdot\|^2$ on the input is Borel measurable.

4. There exists a sequence of input alphabets that satisfy the given power constraint.

The validity of the Blackwell formula (6.14) ensures that the function on the right side of (6.14), whose computation constitutes one of our main results, has an operational meaning.

Note that the capacity-cost function for a d-dimensional multivariate Gaussian noise vector $Z_g \sim \mathcal{N}(0, \Sigma_Z)$ is given by (See [57, Thm. 9.1] and Lemma 6.B.1 in Appendix 6.B below),

$$C_{\mathcal{N}(0,\Sigma_Z)}(B) = \sup_{\substack{\Sigma\succeq 0:\\ \mathrm{tr}(\Sigma)\leq B}} \log \frac{|\Sigma_Z + H\Sigma H^T|}{|\Sigma_Z|}. \tag{6.17}$$

Theorem 6.2.3 (Compound capacity-cost function for W_2 ambiguity set)**.** *The compound capacity* (6.16) *with a Gaussian center $P_\circ = \mathcal{N}(0, \Sigma_\circ)$ is given as,*

$$C_{P_\circ}(B, r) = \inf_{\substack{\Sigma_Z\succeq 0\\ \mathsf{BW}(\Sigma_Z,\Sigma_\circ)\leq r}} C_{\mathcal{N}(0,\Sigma_Z)}(B). \tag{6.18}$$

Proof. The proof of Theorem 6.2.3 is along the same lines as that of Theorem 6.2.1. We leverage the fact that Gaussian noise minimizes capacity for a given noise covariance (See Lemma 6.B.2 in Appendix 6.B below). This helps us write the compound capacity in terms of the second-order statistics of the distributions in the W_2 ball. The operational meaning of Gaussian being the worst-case noise is discussed in [141]. See Appendix 6.B for details. Also note that a Gaussian nominal distribution ensures that distributions in the W_2-ball have bounded 2-nd order moments [253, Lem. 1]. ■

Lemma 6.2.4. *The compound capacity in* (6.14) *admits a dual formulation given by*

$$C_{P_\circ}(B,r) = \inf_{\substack{P_{Z|Z_\circ}:\,\mathbb{R}^d \mapsto \mathbb{R}^d \\ P_Z \in \mathscr{P}_2(\mathbb{R}^d) \\ \mathbb{E}\left[\|Z-Z_\circ\|^2\right] \le r^2}} \sup_{\substack{P_X \text{ on } \mathbb{R}^d: \\ \mathbb{E}\left[\|X\|^2\right] \le B}} I(X; HX+Z). \tag{6.19}$$

where the $\inf$ *is achieved by a Gaussian noise distribution.*

Proof. The result follows from [42, Thm. 5], where the authors prove the above result for the case of AVCs, which is a more general case of the setting that we consider. ■

Consider the compound capacity of a scalar additive channel with an unknown additive noise distribution drawn from a W_2-ambiguity set centered at $P_\circ = \mathcal{N}(0, \sigma_\circ^2)$. First, recall the capacity of a scalar Gaussian channel [41, Eqn. (9.17)]:

$$C_{\mathcal{N}(0,\sigma_\circ^2)}(B) = \frac{1}{2}\log\left(1 + \frac{B}{\sigma^2}\right). \tag{6.20}$$

We now apply Theorem 6.2.3. Plugging (6.10) and (6.20) into (6.18) and using the fact that logarithm is monotonic, we write the worst case capacity for $P_\circ = \mathcal{N}(0, \sigma_\circ^2)$ as,

$$\begin{aligned} C_{\mathcal{N}(0,\sigma_\circ^2)}(B,r) &= \inf_{|\sigma-\sigma_\circ| \le r} C_{\mathcal{N}(0,\sigma_\circ^2)}(B) && (6.21)\\ &= \inf_{|\sigma-\sigma_\circ| \le r} \frac{1}{2}\log\left(1+\frac{B}{\sigma^2}\right) && (6.22)\\ &= \frac{1}{2}\log\left(1+\frac{B}{(\sigma+r)^2}\right). && (6.23) \end{aligned}$$

Naturally, increasing the radius of ambiguity r diminishes the compound channel capacity. This effect can be observed in Fig 6.1.

6.3 Proof of Theorem 6.2.1

Useful Results

If the source distribution $X \sim P_\circ$ is known, (6.2) reduces to the standard RDF [41, Eqn. (10.12)]

$$R_{P_\circ}(D) = R_{P_\circ}(D,0) = \inf_{\substack{P_{\widehat{X}|X}:\,\mathbb{R}^d \mapsto \mathbb{R}^d \\ \mathbb{E}\left[\|X-\widehat{X}\|^2\right] \le D}} I(X;\widehat{X}). \tag{6.24}$$

The next lemma brings to light the significance of the Gaussian RDF in (6.6) as the worst-case RDF among sources with the same covariance matrix.

Lemma 6.3.1 (Gaussian is the hardest to encode, [133, Eqn. (13)]). *Assume that the source distribution $P_\circ$ has a covariance $\Sigma_\circ \in \mathbb{S}_+^d$. The RDF $R_{P_\circ}(D)$ satisfies*

$$R_{P_\circ}(D) \leq R_{\mathcal{N}(0,\Sigma_\circ)}(D), \tag{6.25}$$

and the upper bound is attained by $P_\circ = \mathcal{N}(0, \Sigma_\circ)$.

Proof. This result is due to Kolmogorov [133, Eqn. (13)]. We provide a short proof in Appendix 6.A for completeness. ■

Noting that

$$\mathsf{W}_2(P, P_\circ)^2 \leq r \iff \mathbb{E}[\|X - X_\circ\|^2] \leq r, \tag{6.26}$$

we rewrite the dual formulation of the compound RDF in Lemma 6.2.2 as,

$$R_{P_\circ}(D, r) = \sup_{\substack{P \in \mathscr{P}_{2+\epsilon}(\mathbb{R}^d): \\ \mathsf{W}_2(P, P_\circ) \leq r}} R_P(D). \tag{6.27}$$

Proof of Theorem 6.2.1: upper bound

Consider two distributions $P_\circ, P \in \mathscr{P}_{2+\epsilon}(\mathbb{R}^d)$ with means $\mu_\circ, \mu \in \mathbb{R}^d$ and covariances $\Sigma_\circ, \Sigma \in \mathbb{S}_+^d$, respectively. The W_2-distance between them satisfies the Gelbrich bound [78, Thm. 2.1],

$$\mathsf{W}_2(P, P_\circ)^2 \geq \mathsf{BW}(\Sigma, \Sigma_\circ)^2 + \|\mu_\circ - \mu\|^2, \tag{6.28}$$

where equality is attained if both $P_\circ$ and P are Gaussian distributions. It follows that the W_2-distance between distributions upper-bounds the Bures-Wasserstein distance between their covariance matrices:

$$\mathsf{W}_2(P,\, P_\circ) \geq \mathsf{BW}(\Sigma,\, \Sigma_\circ). \tag{6.29}$$

Applying (6.29) to (6.27), we obtain an upper bound on the worst-case RDF as:

$$R_{P_\circ}(D, r) \leq \sup_{\substack{P \in \mathscr{P}_{2+\epsilon}(\mathbb{R}^d): \\ \mathsf{BW}(\Sigma, \Sigma_\circ) \leq r}} R_P(D) \tag{6.30}$$

$$\leq \sup_{\mathsf{BW}(\Sigma, \Sigma_\circ) \leq r} R_{\mathcal{N}(0,\Sigma)}(D), \tag{6.31}$$

where (6.31) is by Lemma 6.3.1. Plugging (6.6) into the right side of (6.31) yields the $\leq$ direction of (6.7).

Proof of Theorem 6.2.1: lower bound

By Lemma 6.3.1, equality is achieved in (6.31) by $P = \mathcal{N}(0, \Sigma)$. Since both $P_\circ$ and P are Gaussian, equality is achieved in (6.29), and, by extension, in (6.30) as well.

6.4 Conclusion

In this paper, we studied Sakrison's [195] compound RDF and Blackwell's compound capacity [22], focusing on a scenario where the source / noise distribution belongs to a Wasserstein-2 ambiguity set. Our key findings include the identification of the Gaussian distribution as the worst-case scenario for encoding within this set (Theorem 6.2.1, Theorem 6.2.3). The compound RDF (Theorem 6.2.1) and capacity (Theorem 6.2.3) are expressed in terms of the covariance matrices of Gaussian distributions. Future work could explore ambiguity sets defined using distances beyond Wasserstein-2, tradeoffs between coding and learning the distribution to decrease the size r of the ambiguity set, extensions to multiterminal settings, and to causal source and channel coding.

6.A Proof of Lemma 6.3.1

Without loss of generality, assume that X is zero-mean. We first state the following useful lemma, which extends [182, Thm. 5.11] to a vector channel.

Lemma 6.A.1 (Gaussian input maximizes the mutual information in an AWGN channel)**.** *Let X be a random channel input in $\mathbb{R}^d$ with known covariance $\Sigma \in \mathbb{S}^d_+$ and let $A \in \mathbb{R}^{d\times d}$ be a fixed channel matrix. Let $Z_g \sim \mathcal{N}(0, \Sigma_Z)$ be the additive Gaussian channel noise independent of X, with known covariance $\Sigma_Z \in \mathbb{S}^d_+$. We have that;*

$$I(X,\ AX + Z_g) \leq I(X_g,\ AX_g + Z_g), \tag{6.32}$$

where $X_g \sim \mathcal{N}(0, \Sigma)$ is independent of Z_g, and equality holds if $X \sim \mathcal{N}(0, \Sigma)$.

Proof. The mutual information is given by,

$$\begin{aligned} I(X,\, AX + Z_g) &= h(AX + Z_g) - h(AX + Z_g|X), & (6.33)\\ &= h(AX + Z_g) - h(Z_g). & (6.34) \end{aligned}$$

The second term in the last equality does not depend on X, while the first term, by the maximum entropy property of Gaussian random variables [41, Thm. 8.6.5], is bounded above by

$$h(AX + Z_g) \leq \frac{1}{2}\log\left|2\pi \mathrm{e}\,\widehat{\Sigma}\right|, \tag{6.35}$$

where $\widehat{\Sigma} \coloneqq \mathrm{Cov}(AX + Z_g) = A\Sigma A^T + \Sigma_Z$ is the covariance of matrix of $AX + Z_g$. This upper bound is achieved by a Gaussian input $X_g \sim \mathcal{N}(0, \Sigma)$ independent of the channel noise Z_g. ■

Consider the RDF for a known source distribution P_X given in (6.24). Restricting the infimization to linear mappings of the form

$$\widehat{X} = AX + Z_g, \tag{6.36}$$

where $A \in \mathbb{R}^{d\times d}$, and $Z_g \sim \mathcal{N}(0, \Sigma_Z)$ is independent of X, we get an upper bound on $R_{P_X}(D)$ as,

$$\begin{aligned}
R_{P_X}(D) &\leq \inf_{\substack{A\in\mathbb{R}^{d\times d},\Sigma_Z\succeq 0:\\ \mathbb{E}\left[\|X-\widehat{X}\|^2\right]\leq D}} I(X; AX + Z_g) && (6.37)\\
&\leq \inf_{\substack{A\in\mathbb{R}^{d\times d},\Sigma_Z\succeq 0:\\ \mathbb{E}\left[\|X_g-\widehat{X}_g\|^2\right]\leq D}} I(X_g; AX_g + Z_g) && (6.38)\\
&= R_{\mathcal{N}(0,\Sigma)}(D), && (6.39)
\end{aligned}$$

where $X_g \sim \mathcal{N}(0, \Sigma)$ and $\widehat{X}_g \coloneqq AX_g + Z_g$. Inequality (6.38) is by Lemma 6.A.1. Since the condition $\mathbb{E}\left[\|X_g - \widehat{X}_g\|^2\right] \leq D$ can be written as $\mathrm{tr}((A-I)\Sigma(A-I)^T + \Sigma_Z) \leq D$, and the mutual information between two Gaussian random vectors can be written as,

$$I(X_g; AX_g + Z_g) = \frac{1}{2}\log\frac{\left|A\Sigma A^T + \Sigma_Z\right|}{|\Sigma_Z|}, \tag{6.40}$$

the right side of (6.38) is equal to the right side of (6.6). It remains to show (6.39). Applying the argument in (6.37), (6.38) to $P_X = \mathcal{N}(0, \Sigma)$ yields the $\geq$ in (6.39). To show $\leq$, fix any $P_{\widehat{X}|X_g}$. Using standard arguments, [41, Proof of Th. 10.3.2], we have,

$$\begin{aligned}
I(X_g, \widehat{X}) &= h(X_g) - h(X_g|\widehat{X}) && (6.41)\\
&= \frac{1}{2}\log|2\pi e\Sigma| - h(X_g - \widehat{X}|X_g) && (6.42)\\
&\geq \frac{1}{2}\log|2\pi e\Sigma| - h(X_g - \widehat{X}) && (6.43)\\
&\geq \frac{1}{2}\log|2\pi e\Sigma| - \frac{1}{2}\log\left|2\pi e\Sigma_{X_g-\widehat{X}}\right|, && (6.44)
\end{aligned}$$

where $\Sigma_{X_g-\widehat{X}}$ is the covariance of $X_g - \widehat{X}$. Here, (6.43) holds because conditioning decreases entropy [41, Thm. 8.6.1 Cor. 2], and (6.44) is due to (6.35). The expression on the right-hand side of (6.44) is the mutual information between the jointly Gaussian pair $(X_g, \widehat{X})$. It follows that $R_{\mathcal{N}(0,\Sigma)}(D)$ is lower bounded by the right side of (6.38).

6.B Proof of Theorem 6.2.3

We first state the following useful lemmas. In Lemma 6.B.1, we state the capacity of a Gaussian vector channel. To get Lemma 6.B.1, we slightly modify [57, Thm. 9.1] for our use case. In Lemma 6.B.2, we show that for an additive channel, Gaussian noise, among all noises with the same covariance, minimizes the channel capacity. A similar result in the setting of stochastic processes is established in [111].

Lemma 6.B.1 (Capacity of an AWGN channel)**.** *Consider the channel coding setting defined in the first paragraph of Section 6.2. The capacity-cost function for a d-dimensional multivariate Gaussian noise vector $Z_g \sim \mathcal{N}(0, \Sigma_Z)$ is given by,*

$$C_{\mathcal{N}(0,\Sigma_Z)}(B) = \sup_{\substack{\Sigma \succeq 0: \\ tr(\Sigma) \leq B}} \log \frac{|\Sigma_Z + H\Sigma H^T|}{|\Sigma_Z|}. \tag{6.45}$$

Proof. The capacity for $Z_g \sim \mathcal{N}(0, I)$, is given by [57, Thm. 9.1],

$$C_{\mathcal{N}(0,I)}(B) = \sup_{\substack{\Sigma \succeq 0: \\ \mathrm{tr}(\Sigma) \leq B}} \log |I + H\Sigma H^T|. \tag{6.46}$$

The capacity for $Z_g \sim \mathcal{N}(0, \Sigma_Z)$ can be obtained by considering the channel matrix $\tilde{H} = \Sigma_Z^{-1/2} H$ [57, Rem. 9.1].

$$C_{\mathcal{N}(0,\Sigma_Z)}(B) = \sup_{\substack{\Sigma \succeq 0: \\ \mathrm{tr}(\Sigma) \leq B}} \log |I + \Sigma_Z^{-1/2} H\Sigma H^T \Sigma_Z^{-T/2}|. \tag{6.47}$$

Multiplying $I + \Sigma_Z^{-1/2} HQH^T \Sigma_Z^{-T/2}$ by $\Sigma_Z^{1/2}$ on the left and $\Sigma_Z^{T/2}$ on the right, and dividing the argument of $\log$ in (6.47) by $|\Sigma_Z|$, we get (6.45). ■

Lemma 6.B.2 (Gaussian noise minimizes the channel capacity of an additive channel)**.** *Let Z be a random channel noise in $\mathbb{R}^d$ with known covariance $\Sigma_Z \in \mathbb{R}^{d\times d}$ and let $H \in \mathbb{R}^{d\times d}$ be a fixed channel matrix. Let $Z_g \sim \mathcal{N}(0, \Sigma_Z)$ be the additive Gaussian channel noise, with the same covariance as Z, independent of the channel input X. Then,*

$$C_{\mathcal{N}(0,\Sigma_Z)}(B) \leq C_P(B), \tag{6.48}$$

and equality is achieved by $P = \mathcal{N}(0, \Sigma_Z)$.

Proof. Consider $X_g \sim \mathcal{N}(0, \Sigma)$ where Σ is the covariance of X. Using (6.32) and then [49, Lem. II.2], we have the following set of inequalities,

$$I(X,\ HX + Z_g) \leq I(X_g,\ HX_g + Z_g), \tag{6.49}$$

$$\leq I(X_g,\ HX_g + Z). \tag{6.50}$$

Now note that,

$$\sup_{\substack{\Sigma \succeq 0: \\ \mathbb{E}\left[\|X_g\|^2\right] \leq B}} I(X_g,\ HX_g + Z) \leq \sup_{\substack{P \text{ on } \mathbb{R}^d: \\ \mathbb{E}\left[\|X\|^2\right] \leq B}} I(X, HX + Z). \tag{6.51}$$

The expression on the right-hand side of (6.51) is $C_P(B)$. The expression on the left-hand side of (6.51) is lower bounded by $C_{\mathcal{N}(0,\Sigma_Z)}(B)$. Indeed, After applying inequalities (6.50) and then (6.49) to lower-bound the left side of (6.51) and taking supremum over all input distributions leads to (6.45). ■

Proof of Theorem 6.2.3:

Converse

Consider the dual formulation of the compound capacity (6.19) and a noise distribution in the W_2 ball P with mean μ and covariance Σ. We re-write the dual formulation as,

$$C_{P_\circ}(B, r) = \inf_{\substack{P \in \mathscr{P}_2(\mathbb{R}^d): \\ \mathsf{W}_2(P, P_\circ) \leq r}} C_P(B). \tag{6.52}$$

By the Gelbrich bound [78, Thm. 2.1] we have (6.29). Applying (6.29) to (6.52), we obtain a lower bound on the compound capacity as:

$$C_{P_\circ}(B, r) \geq \inf_{\substack{P \in \mathscr{P}_2(\mathbb{R}^d): \\ \mathsf{BW}(\Sigma, \Sigma_\circ) \leq r}} C_P(B) \tag{6.53}$$

$$\geq \inf_{\substack{P \in \mathscr{P}_2(\mathbb{R}^d): \\ \mathsf{BW}(\Sigma, \Sigma_\circ) \leq r}} C_{\mathcal{N}(0,\Sigma)}(B), \tag{6.54}$$

where (6.54) is by Lemma 6.B.2. Plugging (6.45) into the right side of (6.54) yields the $\geq$ direction of (6.18).

Achievability

Equality in (6.54) is achieved by Gaussian P (6.48). Since we assume that $P_\circ$ is Gaussian, equality is achieved in (6.29), and, by extension, in (6.53) as well.

Chapter 7

DISTRIBUTIONALLY ROBUST CLUSTERING

7.1 Introduction

In recent years, the widespread availability of large-scale, high-dimensional datasets has driven significant interest in clustering algorithms that are both computationally efficient and robust to distributional shifts. The classical clustering method, k-means, can be seen as an application of the Lloyd-Max quantization algorithm, where the distribution which is being quantized is the empirical distribution of the points that need to be clustered. This empirical distribution will be different from the *true* underlying distribution, especially when the number of points to be clustered is small. This leads to distributional shift, which can further occur in many real world settings, such as image segmentation, biological data analysis, and sensor networks, due to noise variations, sensor inaccuracies, or environmental changes. Distributional shifts can severely impact the performance of clustering algorithms, leading to degraded cluster assignments and unreliable downstream analysis.

The field of clustering has a rich history. One of the most popular algorithms in this field is the k-means algorithm, introduced by [156]. The k-means algorithm has various variants, such as k-means++ [10], which uses a special seeding method to improve convergence characteristics. Robust clustering has also been explored in the literature, with notable works including [79], [148], [50], [45], and [46]. Noise-robust clustering has also been studied, with [106] providing a comprehensive survey. High-dimensional, low-data clustering is often encountered in biological data. In this context, high-dimensional clustering has been studied separately in [256] and [38]. [150] further explored the problem of high-dimensional clustering using a minimax optimization criterion. For an extensive survey of clustering methods, readers can refer to [75].

Fundamentally, the k-means algorithm can be seen as an application of the Lloyd-Max algorithm [152], [163] when the distribution is the empirical distribution based on the observed data points. Lloyd [152] introduced two iterative approaches to design a locally optimal quantizer, referred to as Method I and Method II. Max [163] derived equations for the necessary conditions for a quantizer to be locally optimal and independently proposed Lloyd's Method II. This is now recognized as the Lloyd-Max algorithm. It should be noted that the generalized clustering problem of finding the optimal quantization points and regions for any given rate and dimensionality is NP-hard [76]

Moreover, the k-means algorithm has many practical applications and several variations of its original version are in use today. For example, [232] explores the use of k-means in the field of image compression. [47] proposes a variant of k-means for image segmentation. Moreover, k-means has extensively been used to learn large-scale representation of images [37]. The wide ranging applications of the classical k-means, makes it even more important to consider robustness to distributional shifts in the original algorithm, especially in the low data regime when the empirical distribution is not an accurate estimate of the underlying distribution.

In this work, we use the Wasserstein-2 (W_2) distance as a measure of distributional shifts to define a family of distributions. The W_2-distance between distributions P_1 and P_2 on $\mathbb{R}^d$ is defined as [231]

$$\mathsf{W}_2(P_1, P_2) := \left(\inf_{P_{XY} \in \Pi(P_1, P_2)} \mathbb{E}\left[\|X - Y\|^2\right] \right)^{\frac{1}{2}}, \tag{7.1}$$

where $\Pi(P_1, P_2)$ denotes the set of all joint distributions with marginals P_1 and P_2. Unlike other common statistical distances such as Total-Variation distance, Kullback–Leibler (KL) divergence, or Hellinger distance, the W_2-distance incorporates information about the geometric structure of the underlying domain, making it particularly suitable for handling structured real-world data. Additionally, unlike KL divergence, it can quantify distances between distributions with different support. Due to its geometric interpretability and tractable formulation, W_2-distance has recently gained popularity as a statistical distance in diverse fields such as control [125], filtering [203], machine learning [134], and data compression [157].

Contributions

In this paper, we propose a variant of the classic k-means algorithm that is robust to distributional shifts. That is, the algorithm minimizes the worst-case error among the W_2 family of distributions. We show that the proposed algorithm is a descent method and find the necessary conditions for the optimal placement of K cluster centers that best represent the W_2 family of distributions for a given nominal distribution $P_\circ$ and ambiguity radius r. Finally, we present numerical simulations to analyze the performance of the algorithm. The organization of the paper is as follows. In Section 7.2, we define the notations used in the paper and formally define the optimization problem. In Section 7.3, we give the necessary conditions on the optimal solution to the problem. In Section 7.4, we formalize the iterative method. In Section 7.5, we provide numerical simulations of our algorithm on synthetic and real world datasets and show that in the low data regime, our algorithm performs better than k-means++. We measure our performance using the worst-case error in the W_2 ball and the misclassification rate for a certain dataset corrupted with outliers.

Notations: The letters $\mathbb{N}$ and $\mathbb{R}$ denote the set of natural and real numbers, respectively. We denote the set of natural numbers upto $N \in \mathbb{N}$ as $[N] \coloneqq \{1, \ldots, N\}$. The set of positive real numbers is denoted by $\mathbb{R}_+$. The set $\Delta_K \coloneqq \{\boldsymbol{\pi} \in \mathbb{R}_+^K \mid \sum_{i\in[K]} \pi_i = 1\}$ denotes the $K-$probability simplex for $K \in \mathbb{N}$. Indicator function of set $\Phi \subseteq \mathbb{R}^d$ is represented as $\mathbb{1}_\Phi : \mathbb{R}^d \to \{0, 1\}$ such that $\mathbb{1}_\Phi(x) = 1$ only if $x \in \Phi$ and zero otherwise. We reserve the boldface capital letters (e.g., $\mathbf{X}, \mathbf{M}$, etc.) for matrices. The expressions $\|x\|$ and $x^\intercal$ denote the Euclidean norm and the transpose of a vector $x \in \mathbb{R}^d$. We use the letter $\mathbb{P}$ and $\mathbb{Q}$ to denote probability distributions, and $\mathbb{E}_\mathbb{P}$ to denote expectation under a distribution $\mathbb{P}$. The set of probability distributions over a domain $\mathbb{R}^d$ is denoted by $\mathscr{P}(\mathbb{R}^d)$ whereas $\mathscr{P}_p(\mathbb{R}^d)$ denotes the set of distributions with finite p^{th} moment.

7.2 Problem Setup

Let $X_\sharp \sim \mathbb{P}_\sharp$ be a random vector that encapsulates the true data population of interest, drawn from an *unknown* probability distribution $\mathbb{P}_\sharp \in \mathscr{P}(\mathbb{R}^d)$. Our objective is to partition this population into $K \in \mathbb{N}$ distinct clusters. Formally, we seek a partition of the domain $\mathbb{R}^d$ into K pairwise disjoint subsets, *i.e.*, $\Phi_1, \ldots, \Phi_K \subseteq \mathbb{R}^d$ such that

$$\Phi_k \cap \Phi_l = \emptyset, \text{ for } k \neq l, \quad \text{and} \quad \cup_{k=1}^K \Phi_k = \mathbb{R}^d. \tag{7.2}$$

Additionally, each cluster is associated with a representative center, *i.e.*, $\mu_k \in \Phi_k$ for $k \in [K]$. For any such pair of cluster centers and partitions $(\mathbf{M}, \boldsymbol{\Phi})$ where $\mathbf{M} \coloneqq \begin{bmatrix}\mu_1 & \ldots & \mu_K\end{bmatrix} \in \mathbb{R}^{d\times K}$ and $\boldsymbol{\Phi} \coloneqq \{\Phi_k\}_{k=1}^K$, the associated clustering map $\mathcal{Q}_{\mathbf{M},\boldsymbol{\Phi}} : \mathbb{R}^d \to \mathbb{R}^d$ that maps a given data $x \in \mathbb{R}^d$ to the center representative of the cluster it belongs to, namely,

$$\mathcal{Q}_{\mathbf{M},\boldsymbol{\Phi}}(x) \coloneqq \sum_{k=1}^K \mu_k \, \mathbb{1}_{\Phi_k}(x). \tag{7.3}$$

Standard K-Means Clustering as Empirical Risk Minimization

In order to assess the performance of a clustering scheme $(\mathbf{M}, \boldsymbol{\Phi})$ for an arbitrary random vector $X \sim \mathbb{P} \in \mathscr{P}(\mathbb{R}^d)$, we consider the expected squared L_2-distance between the random vector X and its corresponding cluster center $\mathcal{Q}_{\mathbf{M},\mathbf{\Phi}}(X)$, namely:

Definition 7.2.1 (Clustering Risk)**.** The L_2-risk of a clustering scheme $(\mathbf{M}, \mathbf{\Phi})$ under a distribution $\mathbb{P} \in \mathscr{P}_2(\mathbb{R}^d)$ is defined as

$$\mathsf{Risk}(\mathbf{M}, \mathbf{\Phi}, \mathbb{P}) \coloneqq \mathbb{E}_{\mathbb{P}}\left[\|X - \mathcal{Q}_{\mathbf{M},\mathbf{\Phi}}(X)\|^2\right]. \tag{7.4}$$

We seek to find a clustering scheme $(\mathbf{M}, \mathbf{\Phi})$ that minimizes its population risk under the true underlying distribution $\mathbb{P}_\sharp$, namely,

$$\inf_{(\mathbf{M},\mathbf{\Phi})} \left\{\mathsf{Risk}(\mathbf{M}, \mathbf{\Phi}, \mathbb{P}_\sharp) = \mathbb{E}_{\mathbb{P}_\sharp}\left[\|X_\sharp - \mathcal{Q}_{\mathbf{M},\mathbf{\Phi}}(X_\sharp)\|^2\right]\right\}. \tag{7.5}$$

In reality, we almost never have access to the true distribution to begin with. Instead, usually a finite dataset $\mathcal{D}_N \coloneqq \{x_n\}_{n=1}^N \subset \mathbb{R}^d$ of $N \in \mathbb{N}$ points is given, presumably drawn from the underlying but unknown distribution $\mathbb{P}_\sharp$. In this case, the empirical distribution

$$\widehat{\mathbb{P}}_N \coloneqq N^{-1} \sum_{n=1}^N \delta_{x_n} \tag{7.6}$$

of the dataset $\mathcal{D}_N$ can be used as a proxy to $\mathbb{P}_\sharp$, leading to the *empirical clustering risk* minimization:

Problem 7.2.2 (Standard K-Means as Empirical Risk Minimization)**.** Given a dataset $\mathcal{D}_N$, find a pair of K cluster centroids and regions $(\mathbf{M}, \mathbf{\Phi})$ that minimizes the empirical clustering risk, *i.e.*,

$$\inf_{(\mathbf{M},\mathbf{\Phi})} \left\{\mathsf{Risk}(\mathbf{M}, \mathbf{\Phi}, \widehat{\mathbb{P}}_N) = \mathbb{E}_{\widehat{\mathbb{P}}_N}\left[\|X - \mathcal{Q}_{\mathbf{M},\mathbf{\Phi}}(X)\|^2\right] = \frac{1}{N}\sum_{n=1}^N \|x_n - \mathcal{Q}_{\mathbf{M},\mathbf{\Phi}}(x_n)\|^2\right\}. \tag{7.7}$$

This is the standard objective used in traditional K-means clustering. Although this objective is convex in cluster centers $\mathbf{M}$ alone, the joint minimization over the cluster regions $\mathbf{\Phi}$ makes it highly non-convex in general. Nevertheless, the necessary optimality conditions are obtained as follows:

Theorem 7.2.3 (Optimality of K-Means [152], [163])**.** *Given a fixed set of centroids* $\mathbf{M}$*, the optimal partitioning* $\mathbf{\Phi}^\star$ *is the nearest-neighbor partition, i.e.,*

$$\Phi_k^\star = \left\{x \in \mathbb{R}^d \mid \|x - \mu_k\| \le \|x - \mu_l\| \quad \forall l \neq k\right\}, \quad k = 1, \ldots, K. \tag{7.8}$$

Furthermore, the optimal centroids $\mathbf{M}^\star$ *for a dataset* $\mathcal{D}_N$ *satisfy the following necessary conditions:*

$$\mu_k^\star = \frac{\sum_{n=1}^N x_n \, \mathbb{1}_{\Phi_k^\star}(x_n)}{\sum_{n=1}^N \mathbb{1}_{\Phi_k^\star}(x_n)} = \mathbb{E}_{\widehat{\mathbb{P}}_N}\left[X \mid X \in \Phi_k^\star\right], \quad k = 1, \ldots, K. \tag{7.9}$$

The infamous Lloyd-Max algorithm is essentially developed as a fixed-point computing locally optimal centroids.

$$\Phi_k^{(t)} \leftarrow \left\{ x \in \mathbb{R}^d \mid \|x - \mu_k^{(t)}\| \leq \|x - \mu_l^{(t)}\| \quad \forall l \neq k \right\} \tag{7.10}$$

$$\mu_k^{(t+1)} \leftarrow \frac{\sum_{n=1}^N x_n \, \mathbb{1}_{\Phi_k^{(t)}}(x_n)}{\sum_{n=1}^N \mathbb{1}_{\Phi_k^{(t)}}(x_n)} = \mathbb{E}_{\widehat{\mathbb{P}}_N}\left[X \mid X \in \Phi_k^{(t)} \right] \tag{7.11}$$

The Lloyd-Max algorithm has per-iteration complexity $O(dNK)$, and the empirical risk is monotonically decreasing at each iteration, namely,

$$\mathsf{Risk}(\mathbf{M}^{(t+1)}, \mathbf{\Phi}^{(t+1)}, \widehat{\mathbb{P}}_N) \leq \mathsf{Risk}(\mathbf{M}^{(t)}, \mathbf{\Phi}^{(t)}, \widehat{\mathbb{P}}_N),$$

converging to a local minimum. While this makes it an easy-to-implement algorithm, the accuracy of the resulting clustering scheme on the true population highly depends on the quality and representativeness of the dataset of the true underlying population.

Distributionally Robust K-Means Clustering

As discussed in the preceding section, the standard K-means clustering on a dataset $\mathcal{D}_N$ is an empirical proxy to clustering the underlying true population with distribution $\mathbb{P}_\sharp$. When the dataset is generated directly from the population, say in $\mathrm{i.i.d.}$ fashion, the representativeness of the empirical distribution $\widehat{\mathbb{P}}_N$ as a proxy to the population $\mathbb{P}_\sharp$ diminishes severely in high-dimensional data-starve settings, such that $\log(N) \ll d$, due to curse of dimensionality.

Moreover, distribution shifts in the data-generating process, *i.e.*, deviations from the underlying population due to reasons like corruption by noise, outliers, or other unknown mechanisms, further degrade the accuracy of clusters obtained from standard K-means on the underlying population. However, the standard K-means does not have a priori performance guarantee for the underlying population when the dataset is subject to distribution shifts, making it highly susceptible to cluster misrepresentation under such phenomena.

To overcome these limitations, we propose a distributionally robust K-means clustering method. In this setting, we explicitly incorporate our uncertainty about the true population distribution $\mathbb{P}_\sharp$ into the design of K-means clusters through an ambiguity set of plausible probability distributions consistent with the dataset $\mathcal{D}_N$, to which the true distribution $\mathbb{P}_\sharp$ belongs. We can describe this notion of plausibility by a statistical distance on the space of probability measures $\mathscr{P}(\mathbb{R}^d)$. Since the clustering risk is assessed by the L_2-distance, a natural statistical distance to use is the Wasserstein-2 distance.

Definition 7.2.4 (Wasserstein-2 Ambiguity Sets)**.** The W_2-ambiguity set of radius $r > 0$ around a nominal distribution $\mathbb{P}_\circ \in \mathscr{P}(\mathbb{R}^d)$ is defined as

$$\mathbb{B}_r^{\mathrm{w}}(P_\circ) := \left\{ \mathbb{P} \in \mathscr{P}(\mathbb{R}^d) \mid \mathsf{W}_2(\mathbb{P}, \mathbb{P}_\circ) \le r \right\}. \tag{7.12}$$

In this paper, we take the empirical distribution $\widehat{\mathbb{P}}_N$ as the nominal. Notably, the Wasserstein-2 ambiguity set offers the significant advantage of encompassing continuous, mixed, and discrete distributions. Accordingly, we impose the following assumption for the remainder of our paper:

Assumption 7.2.5. Given $\mathcal{D}_N$ and $r > 0$, the true distribution $\mathbb{P}_\sharp$ of the underlying population belongs to the ambiguity set $\mathbb{B}_r(\widehat{\mathbb{P}}_N)$.

The radius $r > 0$ essentially determines the level of uncertainty about and/or deviation from the unknown population $\mathbb{P}_\sharp$. When the dataset samples are drawn $\mathrm{i.i.d.}$ from the population $\mathbb{P}_\sharp$, it is indeed possible to determine an upper bound on r that guarantees that Assumption 7.2.5 holds with high probability.

Theorem 7.2.6 (High-Confidence Ambiguity Sets [66, Thm. 2])**.** *Let $d > 4$, $\varepsilon \in (0,1)$, and $\mathbb{E}_{\mathbb{P}_\sharp}[\exp(\|X\|^\alpha)] < +\infty$ for an $\alpha > 2$. When $\mathcal{D}_N$ is drawn* $\mathrm{i.i.d.}$ *from $\mathbb{P}_\sharp$, the true density $\mathbb{P}_\sharp$ resides in the ambiguity set $\mathbb{B}_{r(N,\varepsilon)}(\widehat{\mathbb{P}}_N)$ with probability at least $1 - \varepsilon$, i.e.,*

$$\mathbb{P}_\sharp^{\otimes N} \left\{ \mathsf{W}_2(\mathbb{P}_\sharp, \widehat{\mathbb{P}}_N) \le r(N,\varepsilon) \right\} \ge 1 - \varepsilon \ \ \textit{where} \ \ r(N,\varepsilon) := \begin{cases} \left(\frac{\log(C/\varepsilon)}{cN}\right)^{\frac{2}{d}}, & \textit{if}\, N \ge \frac{\log(C/\varepsilon)}{c}, \\ \left(\frac{\log(C/\varepsilon)}{cN}\right)^{\frac{2}{\alpha}}, & \textit{if}\, N < \frac{\log(C/\varepsilon)}{c}, \end{cases} \tag{7.13}$$

where $C, c > 0$ are constants depending on α and d.

The above theorem simply provides a priori high-confidence upper bound on the W_2 distance between the unknown true distribution $\mathbb{P}_\sharp$ and the empirical distribution $\widehat{\mathbb{P}}_N$ sampled from $\mathbb{P}_\sharp$. Similar a priori bounds can also be derived for distribution shifts induced by feature noise, data corruption, or outlier injection.

Assumption 7.2.5 allows establishing performance bounds on the population risk of any clustering scheme through the worst-case risk attained among all plausible distributions in the ambiguity set, defined as follows:

Definition 7.2.7 (Worst-Case Risk)**.** Given a dataset $\mathcal{D}_N$ and radius $r > 0$, the worst-case risk of a quantization scheme $(\mathbf{M}, \mathbf{\Phi})$ is defined as

$$\mathsf{wRisk}_{N,r}(\mathbf{M}, \mathbf{\Phi}) := \sup_{\mathbb{P} \in \mathbb{B}_r(\widehat{\mathbb{P}}_N)} \left\{ \mathsf{Risk}(\mathbf{M}, \mathbf{\Phi}, \mathbb{P}) = \mathbb{E}_{\mathbb{P}} \left[\|X - \mathcal{Q}_{\mathbf{M},\mathbf{\Phi}}(X)\|^2 \right] \right\}. \tag{7.14}$$

Although the true distribution is not known, the worst-case risk provides an upper bound on the population risk of a clustering scheme $(\mathbf{M}, \mathbf{\Phi})$,

$$\inf_{\mathbf{M}', \mathbf{\Phi}'} \mathsf{Risk}(\mathbf{M}', \mathbf{\Phi}', \mathbb{P}_\sharp) \leq \mathsf{Risk}(\mathbf{M}, \mathbf{\Phi}, \mathbb{P}_\sharp) \leq \mathsf{wRisk}_{N,r}(\mathbf{M}, \mathbf{\Phi}). \tag{7.15}$$

Thus, minimizing the worst-case clustering risk among all plausible distributions in the ambiguity set instead of directly minimizing the empirical risk provides a proxy with a guaranteed population in the face of uncertainties and distribution shifts between empirical training data and the true population. This is formalized in the following problem.

Problem 7.2.8 (Distributionally Robust K-Means)**.** Given a dataset $\mathcal{D}_N$ and a ambiguity set radius $r > 0$, find a clustering scheme $(\mathbf{M}, \mathbf{\Phi})$ that minimizes the worst-case clustering risk, *i.e.*,

$$\inf_{(\mathbf{M}, \mathbf{\Phi})} \sup_{\mathbb{P} \in \mathbb{B}_r(\widehat{\mathbb{P}}_N)} \left\{ \mathsf{Risk}(\mathbf{M}, \mathbf{\Phi}, \mathbb{P}) = \mathbb{E}_{\mathbb{P}} \left[\|X - \mathcal{Q}_{\mathbf{M},\mathbf{\Phi}}(X)\|^2 \right] \right\}. \tag{7.16}$$

7.3 Optimality Conditions for Distributionally Robust Clusters

In this section, we establish the necessary optimality conditions for distributionally robust K-means clustering problem in 7.2.8. Our approach is grounded in the strong duality result (Theorem 7.3.2) from the Wasserstein distributionally robust optimization (DRO) literature and employs the Karush-Kuhn-Tucker (KKT) conditions to characterize local optimality in Theorem 7.3.4.

We start by establishing the optimality of the nearest neighbor partition in (7.8) for fixed centroids over any population distribution.

Lemma 7.3.1 (Optimality of Nearest-Neighbor Partition)**.** *For any fixed centroids* $\mathbf{M} = \begin{bmatrix} \mu_1 & \ldots & \mu_K \end{bmatrix} \in \mathbb{R}^{d \times K}$*, the nearest neighbor partition* $\mathbf{\Phi}^\star$ *in* (7.8) *universally dominates clustering risk such that*

$$\mathsf{Risk}(\mathbf{M}, \mathbf{\Phi}, \mathbb{P}) \geq \mathsf{Risk}(\mathbf{M}, \mathbf{\Phi}^\star, \mathbb{P}) = \underbrace{\mathbb{E}_{\mathbb{P}} \left[\min_{k \in [K]} \|X - \mu_k\|^2 \right]}_{:=\mathsf{Risk}(\mathbf{M},\mathbb{P})}, \quad \textit{for all} \quad \mathbb{P} \in \mathscr{P}(\mathbb{R}^d). \tag{7.17}$$

Proof. See Section 7.A. ■

This result implies that the distributionally robust clusters for fixed centroids will be the nearest neighbor clusters as well, *i.e.*,

$$\inf_{(\mathbf{M},\mathbf{\Phi})} \sup_{\mathbb{P}\in\mathbb{B}_r(\widehat{\mathbb{P}}_N)} \mathsf{Risk}(\mathbf{M},\mathbf{\Phi},\mathbb{P}) = \inf_{\mathbf{M}\in\mathbb{R}^{d\times K}} \underbrace{\sup_{\mathbb{P}\in\mathbb{B}_r(\widehat{\mathbb{P}}_N)} \mathbb{E}_{\mathbb{P}}\left[\min_{k\in[K]} \|X-\mu_k\|^2\right]}_{:=\mathsf{wRisk}_{N,r}(\mathbf{M})}, \tag{7.18}$$

where $\mathsf{wRisk}_{N,r}(\mathbf{M})$ is the worst-case clustering risk of centroids $\mathbf{M}$ with nearest neighbor clusters. In this formulation, evaluating $\mathsf{wRisk}_{N,r}(\mathbf{M})$ entails solving an infinite-dimensional optimization problem over the space of probability distributions. By leveraging the strong duality property of distributionally robust optimization as established in Gao and Kleywegt [74, Thm. 1], we derive a tractable reformulation in Theorem 7.3.2 that effectively reduces the problem to a single-variable optimization over the nominal empirical distribution.

Theorem 7.3.2 (Strong Dual)**.** *The worst-case clustering risk* $\mathsf{wRisk}_{N,r}(\mathbf{M})$ *incurred by any given set of* K *centroids* $\mathbf{M}\in\mathbb{R}^{d\times K}$ *is equivalent to following dual formulation:*

$$\mathsf{wRisk}_{N,r}(\mathbf{M}) = \inf_{\gamma>1} \gamma r^2 + \frac{1}{N}\sum\nolimits_{n=1}^{N} e_n(\gamma,\mathbf{M}), \tag{7.19}$$

where

$$e_n(\gamma,\mathbf{M}) := \sup_{x\in\mathbb{R}^d} \min_{k\in[K]} \|x-\mu_k\|^2 - \gamma\|x-x_n\|^2. \tag{7.20}$$

Proof. See Section 7.B ■

While the primal problem in (7.18) is reduced to a tractable finite-dimensional dual optimization in (7.19), it requires solving a non-convex max-min optimization for each data point x_n. The following lemma provides a convex-concave min-max reformulation of the inner max-min problem and a tractable expression for the worst-case source in terms of the nominal.

Lemma 7.3.3 (Inner Saddle Point and the Worst-Case)**.** *For* $\gamma>1$*, the objective* $e_n(\gamma,\mathbf{M})$ *in* (7.20) *is equivalent to the following convex-concave min-max optimization:*

$$\begin{aligned} e_n(\gamma,\mathbf{M}) &= \min_{\boldsymbol{\pi}\in\Delta_K} \max_{x\in\mathbb{R}^d} \sum\nolimits_{k=1}^{K} \|x-\mu_k\|^2\pi_k - \gamma\|x-x_n\|^2, &(7.21)\\ &= \min_{\boldsymbol{\pi}\in\Delta_K} \sum\nolimits_{k=1}^{K} \|x_n-\mu_k\|^2\pi_k + \frac{1}{\gamma-1}\|x_n-\mathbf{M}\boldsymbol{\pi}\|^2. &(7.22)\end{aligned}$$

Furthermore, this min-max optimization admits a saddle point $(x_n^\star, \boldsymbol{\pi}_n^\star) \in \mathbb{R}^d \times \Delta_K$ *such that*

$$x_n^\star := x_n + \frac{x_n - \mathbf{M}\boldsymbol{\pi}_n^\star}{\gamma - 1}, \quad \textit{and} \quad \boldsymbol{\pi}_n^\star \in \operatorname*{arg\,min}_{\boldsymbol{\pi}\in\Delta_K} \sum\nolimits_{k=1}^{K} \|x_n^\star - \mu_k\|^2 \pi_k. \tag{7.23}$$

Proof. See Section 7.C ■

Note that the optimal probability mass function (pmf) $\boldsymbol{\pi}_n^\star \in \Delta_K$ is a function of the data point x_n. This pmf can be interpreted as a stochastic cluster assignment for x_n for the given set of centroids $\mathbf{M}$, namely, $\pi_{kn}^\star$ represents the conditional probability of assigning the data point x_n to k^{th} cluster.

Observe that the term $\sum_{k=1}^{K} \|x_n - \mu_k\|^2 \pi_{kn}$ in (7.22) represents the expected squared distance between data point x_n and the centroids $\mathbf{M}$ under the stochastic assignment rule $\boldsymbol{\pi}_n$. In contrast, $\mathbf{M}\boldsymbol{\pi}$ denotes the weighted average of the cluster centers, so that $\|x_n - \mathbf{M}\boldsymbol{\pi}_n\|^2$ quantifies the bias associated with the data point x_n under the stochastic cluster assignment scheme $(\mathbf{M}, \boldsymbol{\Pi})$, where

$$\boldsymbol{\Pi} := \begin{bmatrix} \boldsymbol{\pi}_1 & \dots & \boldsymbol{\pi}_N \end{bmatrix} \in \mathbb{R}^{K\times N}.$$

Thus, the second term in the optimization problem (7.22) acts as a regularizer. This regularization, governed by the parameter $\gamma > 1$, mediates the trade-off between the expected squared distance of x_n to the cluster centroids and the bias introduced by the stochastic cluster assignment.

Furthermore, $x_n^\star$ represents the worst-case data point associated with x_n and given centroids $\mathbf{M}$, so that the worst-case distribution becomes $\mathbb{P}^\star = N^{-1}\sum_{n=1}^{N} \delta_{x_n^\star}$. These worst-case data points are simply shifted away from the original data x_n to the direction of averaged cluster center $\mathbf{M}\boldsymbol{\pi}_n$, introducing a bias scaled by the inverse of $\gamma - 1$. As the radius of the ambiguity set diminishes, *i.e.*, $r \to 0$, we have $\gamma \to \infty$. In this limiting case, both the regularization and perturbation terms vanish, thereby recovering the standard K-means over the dataset $\mathcal{D}_N$.

We conclude this section with Theorem 7.3.4, which formally states the main result of this section, *i.e.*, the local optimality conditions for the distributionally robust centroids.

Theorem 7.3.4 (Conditions for Local Optimality)**.** *Given a fixed dataset* $\mathcal{D}_N$ *and a radius* $r > 0$*, globally optimal centroids* $\mathbf{M}^\star$ *solving the Wasserstein-2 distributionally* *robust K-means clustering problem 7.2.8 must satisfy*

$$\mu_k^\star = \frac{\sum_{n=1}^N x_n^\star \pi_{kn}^\star}{\sum_{n=1}^N \pi_{kn}^\star}, \quad k = 1, \dots, K. \tag{7.24}$$

where $\mathbf{X}^\star := \begin{bmatrix} x_1^\star & \dots & x_N^\star \end{bmatrix} \in \mathbb{R}^{d\times N}$ *and* $\Pi^\star \in \mathbb{R}^{K\times N}$ *are as in* (7.23)*, and the optimal* $\gamma^\star$ *satisfies*

$$\gamma^\star = 1 + r^{-1}\sqrt{\sum_{n=1}^N \|x_n - \mathbf{M}^\star \boldsymbol{\pi}_n^\star\|^2}. \tag{7.25}$$

These conditions are sufficient for local optimality.

Proof. See Section 7.D. ∎

7.4 An Efficient Algorithm for Distributionally Robust Centroids

In this section, we present a practical algorithm inspired by the theoretical insights developed in Section 7.3 to identify distributionally robust centroids. The core of our approach lies in formulating an iterative method that computes locally optimal centroidal regions through a fixed-point iteration à la Lloyd-Max. These fixed points are characterized by the KKT conditions in Theorem 7.3.4.

Before presenting the algorithm, we derive a closed-form expression for the saddle point $(\mathbf{X}^\star, \mathbf{\Pi}^\star)$ under a fixed $\gamma > 1$ and $\mathbf{M}$ in the scalar setting ($d = 1$). This derivation yields valuable insights into the structure of the stochastic assignment rule.

Lemma 7.4.1 (Closed-form in Scalar Setting)**.** *Given* $\mu_0 < \mu_1 < \mu_2 < \cdots < \mu_n < \mu_{n+1}$ *where* $\mu_0 := -\infty$ *and* $\mu_{n+1} := +\infty$ *and* $\gamma > 1$*, define the disjoint intervals*

$$\Phi_{\gamma,k} := \mu_k + (1-\gamma^{-1}) \left]\frac{\mu_{k-1}-\mu_k}{2},\ \frac{\mu_{k+1}-\mu_k}{2}\right[, \tag{7.26}$$

$$\Phi_{\gamma,k+\frac{1}{2}} := \mu_{k+\frac{1}{2}} + \gamma^{-1}\left[-\frac{\mu_{k+1}-\mu_k}{2},\ \frac{\mu_{k+1}-\mu_k}{2}\right], \tag{7.27}$$

for $k \in [K]$ *with* $\Phi_{\gamma,\frac{1}{2}} = \Phi_{\gamma,n+\frac{1}{2}} = \emptyset$*, where* $\mu_{k+\frac{1}{2}} := \frac{\mu_k+\mu_{k+1}}{2}$*. Defining the ratio,*

$$q_{k+\frac{1}{2},n} := \frac{x_n - \mu_{k+\frac{1}{2}}}{(\mu_{k+1}-\mu_k)/2}, \ \forall k \in [n-1], \tag{7.28}$$

the closed-form expressions for the worst-case source and the corresponding stochastic decoding rule given the nominal source are, respectively,

$$x_n^\star = \begin{cases} x_n + \frac{x_n - \mu_k}{\gamma - 1}, & x_n \in \Phi_{\gamma,k}, \\ \mu_{k+\frac{1}{2}}, & x_n \in \Phi_{\gamma,k+\frac{1}{2}}, \end{cases} \quad \textit{and} \quad \pi_{kn}^\star = \begin{cases} \frac{1}{2}(1+\gamma q_{k-\frac{1}{2},n}), & x_n \in \Phi_{\gamma,k-\frac{1}{2}}, \\ 1, & x_n \in \Phi_{\gamma,k}, \\ \frac{1}{2}(1-\gamma q_{k+\frac{1}{2},n}), & x_n \in \Phi_{\gamma,k+\frac{1}{2}}. \end{cases} \tag{7.29}$$

Proof. See Section 7.E ■

This result demonstrates that any data point x_n residing in the interval $\Phi_{\gamma,k+\frac{1}{2}}$ around the boundary point, $\mu_{k+\frac{1}{2}}$, between adjacent clusters invariably yields a worst-case data point $x_n^\star$ located precisely at the boundary point $\mu_{k+\frac{1}{2}}$. Consequently, the construction of this worst-case data point necessitates a non-deterministic weighting of the centroids.

Algorithm 3 Distributionally Robust K-Means Clustering

input: $\gamma > 1$, dataset $\mathcal{D}_N$, initialization $\mathbf{M}^{(0)}$, convergence tolerance $\varepsilon > 0$,

repeat

$\boldsymbol{\pi}_n^{(t)} \leftarrow \underset{\boldsymbol{\pi}_n \in \Delta_K}{\arg\min} \sum_{k=1}^{K} \|x_n - \mu_k^{(t)}\|^2 \pi_{kn} + \dfrac{1}{\gamma - 1}\|x_n - \mathbf{M}^{(t)}\boldsymbol{\pi}_n^{(t)}\|^2$ for each $n = 1, \dots, N$,

$\mathbf{M}^{(t+1)} \leftarrow \underset{\mathbf{M} \in \mathbb{R}^{d\times K}}{\arg\min} \dfrac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{K} \|x_n - \mu_k\|^2 \pi_{kn}^{(t)} + \dfrac{1}{\gamma - 1}\|x_n - \mathbf{M}\boldsymbol{\pi}_n^{(t)}\|^2$,

Increment $t \leftarrow t + 1$

until $\dfrac{\max_k |\mu_k^{(t+1)} - \mu_k^{(t)}|}{\max_l |\mu_l^{(t)}|} \leq \varepsilon$

return $\mathbf{M}^{(t)}$

We present Algorithm 3, a novel distributionally robust K-means clustering algorithm. For clarity, the algorithm is introduced assuming a fixed parameter $\gamma > 1$. However, it can be easily adapted to operate with a fixed radius $r > 0$ by computing the optimal $\gamma^\star > 1$ that satisfies the equation (7.25) using a straightforward bisection method.

Given a dataset $\mathcal{D}_N$ and an initial set of centroids $\mathbf{M}^{(0)} \in \mathbb{R}^{d\times K}$, the algorithm iteratively refines the centroids through a two-step process. First, it computes the stochastic assignment probabilities $\mathbf{\Pi}^{(t)} \in \mathbb{R}^{K\times N}$ corresponding to the current centroids $\mathbf{M}^{(t)}$ by solving (7.22), a convex quadratic program. Next, the centroids are updated, with $\mathrm{M}^{(t+1)}$ being the minimizer of the following objective for fixed $\mathbf{\Pi}^{(t)}$,

$$\mathbf{M}^{(t+1)} \leftarrow \operatorname*{arg\,min}_{\mathbf{M}\in\mathbb{R}^{d\times K}} \frac{1}{N}\sum\nolimits_{n=1}^{N}\sum\nolimits_{k=1}^{K}\|x_n-\mu_k\|^2\pi_{kn}^{(t)}+\frac{1}{\gamma-1}\|x_n-\mathbf{M}\boldsymbol{\pi}_n^{(t)}\|^2. \tag{7.30}$$

This is an unconstrained convex quadratic objective and admits closed form solution. Notably, this algorithm generalizes the classical Lloyd-Max algorithm, which emerges as a limiting case when $\gamma \to \infty$. In this limit, the fuzzy midpoint regions $\Phi_{k+\frac{1}{2}}^{(t)}$ vanish, and the stochastic decoder transitions to the deterministic nearest-neighbor decoder, thereby eliminating ambiguity about the source distribution. We conclude this section with the following theorem on the convergence of this algorithm.

Theorem 7.4.2 (Monotonic Convergence)**.** *The iterates* $\{\mathrm{M}^{(t)}\}_{t=0}^{\infty}$ *generated by Algorithm 3 monotonically decrease the worst-case clustering risk, i.e.,*

$$\mathsf{wRisk}_{N,r}(\mathbf{M}^{(t+1)}) \leq \mathsf{wRisk}_{N,r}(\mathbf{M}^{(t)}),\ \textit{for all } t \geq 0. \tag{7.31}$$

Proof. See Section 7.F. ■

The observed monotonic decrease mirrors that of the classical Lloyd–Max algorithm for standard K-means clustering. In particular, although it may not be immediately apparent, the iterative updates that alternate between optimizing the centroids and adjusting the stochastic assignments can be interpreted as a Expectation-Maximization (EM) or coordinate descent procedure.

7.5 Numerical Experiments

In this section, we provide an empirical comparison of the Distributionally Robust K-Means (DRKM) and K-Means++(KM). For better convergence, we initialize our algorithm with the output of K-Means++. Since our algorithm is better suited for the data starved regime, we study the worst-case performance and its dependence on the number of data points (Experiments 1 and 3). We also show that the DRKM is robust to outliers in terms of classification error (Experiment 2). In this section, we utilize a mix of both synthetic and real-world datasets.

Ambiguity Radius

In order to have a meaningful radius of the W_2 ball across varying N, we set,

$$r = \alpha \left(\frac{1}{N} \right)^{\frac{1}{d}}, \tag{7.32}$$

where d is the dimension, N is the number of data points and α is a user-specified parameter. The reasoning behind this choice of r is that as the number of data points increases, our uncertainty in the true distribution decreases; hence the ambiguity radius should adjust accordingly. We chose (7.32) to maintain a constant confidence in our estimate of the true distribution with varying N. For details, see Theorem 2 in [66].

Synthetic Datasets

In Experiment 1, we consider a Gaussian mixture model (GMM) with 3 components with weights $\{0.2, 0.26, 0.53\}$ and randomly chosen centers in $d = 7$ dimensions. We chose $\alpha = 10$ and $K = 3$, where each cluster center corresponds to a unique Gaussian component. We observe in Figure 7.1 that as the number of data points N increases from 5 to 50, the difference in the worst-case performance between the DRKM and KM decreases. This is because a higher N better represents the underlying data distribution, making it easier for KM to obtain a better estimate of the underlying distribution.

In Experiment 2, we consider a dataset which is corrupted by an outlier cluster. The original dataset, as represented in Figure 7.2, consists of two clusters (Cluster A and Cluter B) which are normally distributed N_A and N_B points, with means $(-2, -2)$ and $(2, 2)$ respectively. The outlier cluster is also a set of N_o normally distributed points centered at $(8, 8)$. The misclassification error of DRKM and KM for different data points is given in Table 7.1. The values are averaged over 200 trial runs. We see that the DRKM performs better than KM in the low data regime.

Real World Dataset

In Experiment 3, we consider the worst-case performance on the Shuttle dataset of the UCI repository [160]. The original dataset consists of $43,500$ data points with $d = 9$ features. We sparsify the dataset to study the worst-case performance of DRKM and KM in the low-data regime. We set $\alpha = 100$ in this experiment. The results are shown in Figure 7.3. As observed in Figure 7.1, we see that as the number of data points N increases from 5 to 50, the difference in the worst-case performance between the DRKM and KM decreases.

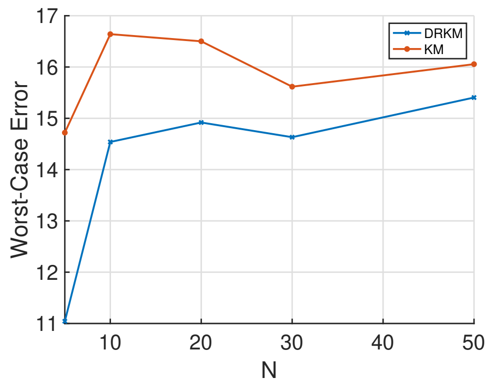


Figure 7.1: The trend of the worst-case error in Experiment 1, averaged over 30 trials for each N. We see that as N increases, the difference in the performance of DRKM and KM decreases, signifying the performance improvement of DRKM over KM in the low data regime.

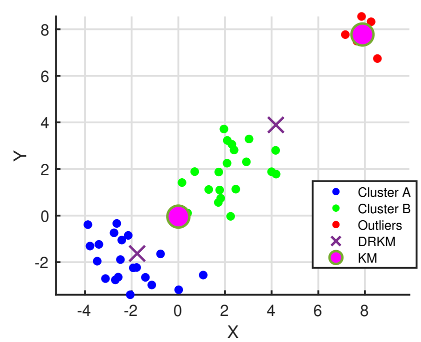


Figure 7.2: A realization of the dataset in Experiment 2. The blue points are in Cluster A and the green points are in Cluster B. We see that the cluster center of KM is shifted to the outlier cluster whereas the cluster center of DRKM, is close to the true cluster (Cluster B) even in the presence of an outlier class and even when we initialize DRKM with the output of KM.

7.6 Conclusion

We have introduced a distributionally robust variant of the classical k-means algorithm that explicitly accounts for distributional shifts by minimizing the worst-case error over a W_2 family of distributions. Our formulation leverages a Wasserstein-2 ambiguity set centered at the nominal distribution $P_\circ$ with a specified radius r, ensuring that the resulting cluster placements are optimal under uncertainty. We have established that the proposed algorithm operates as a descent method and derived the necessary conditions for the optimal positioning of K clusters. Numerical simulations confirm the efficacy of our approach, demonstrating its performance in data starved regime where traditional k-means may falter due to distributional

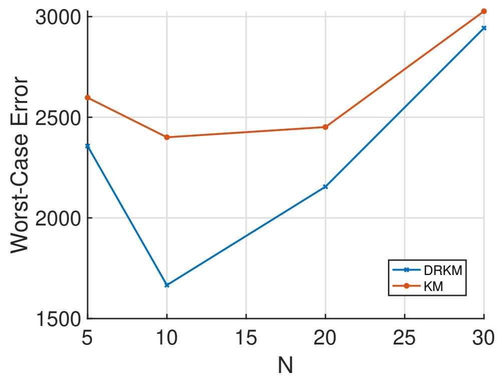


Figure 7.3: The trend of the worst-case error in Experiment 3, averaged over 30 trials for each N. We randomly pick N data points from the dataset in our experiment. We see that as N increases, the difference in the performance of DRKM and KM decreases, signifying the performance improvement of DRKM over KM in the low data regime.

N_A	N_B	N_o	Misclassification Rate (%) (DRKM)	Misclassification Rate (%) (KM)	r
20	20	5	13.6	21.9	2.25
10	10	2	7.8	20.4	2.5

Table 7.1: The Missclassification Rate of DRKM and KM for different data samples. We see that DRKM noticeably outperforms KM.

deviations. Future directions include studying the convergence gaurantees of the algorithm.

7.A Proof of Lemma 7.3.1

Since $\mathbf{\Phi}$ is a partition, for any $x \in \mathbb{R}^d$, there is exactly one $k^\star \in [K]$ for which $\mathbb{1}_{\mathbf{\Phi}^\star}(x) = 1$. Then, we have that:

$$\|x - \mathcal{Q}_{\mathbf{M},\mathbf{\Phi}}(x)\|^2 = \left\|\sum_{k=1}^{K} \mathbb{1}_{\Phi_k}(x)(x-\mu_k)\right\|^2, \tag{7.33}$$

$$= \sum_{k=1}^{K} \mathbb{1}_{\Phi_k}(x)\|x-\mu_k\|^2, \tag{7.34}$$

$$\geq \min_{k\in[K]} \|x-\mu_k\|^2. \tag{7.35}$$

Since this holds for any $x \in \mathbb{R}^d$, it also holds under the expectation operator over any distribution $\mathbb{P}$:

$$\mathbb{E}_{\mathbb{P}}\left[\|x - \mathcal{Q}_{\mathbf{M},\mathbf{\Phi}}(x)\|^2\right] \geq \mathbb{E}_{\mathbb{P}}\left[\min_{k\in[K]} \|x-\mu_k\|^2\right]. \tag{7.36}$$

Finally, the nearest neighbor partition for the given centroids $\mathbf{M}$ achieve this lower bound. ∎

7.B Proof of Theorem 7.3.2

Notice that the worst-case risk in (7.18) is in the form of standard Wasserstein DRO problem as

$$\sup_{P \in \mathbb{B}_r^{\mathrm{w}}(P_\circ)} \mathbb{E}_{X \sim P}\left[f(x \mid \mathbf{M})\right] \tag{7.37}$$

where $f(x \mid \mathbf{M}) \coloneqq \min_{k \in [K]} \|x - \mu_k\|^2$ is distance of a single point $x \in \mathbb{R}^d$ to the nearest cluster centroid. Applying the strong duality of Wasserstein DRO in Gao and Kleywegt [74, Thm. 1], we obtain

$$\inf_{\gamma \geq 0} \gamma r^2 + \mathbb{E}_{X_\circ \sim \widehat{\mathbb{P}}_N}\left[\sup_{x \in \mathbb{R}^d} f(x \mid \mathbf{M}) - \gamma \|x - X_0\|^2\right], \tag{7.38}$$

$$= \inf_{\gamma \geq 0} \gamma r^2 + \frac{1}{N}\sum\nolimits_{n=1}^{N}\left[\sup_{x \in \mathbb{R}^d} \min_{k \in [K]} \|x - \mu_k\|^2 - \gamma \|x - x_n\|^2\right], \tag{7.39}$$

Notice that any choice of $\gamma \leq 1$ would make the inner supremum over $x \in \mathbb{R}^d$ infinity, thus restricting $\gamma > 1$. ■

7.C Proof of Lemma 7.3.3

Note that the minimization over the indices inside the expectation in the primal problem (7.18) can be reexpressed as a minimization over the probability simplex:

$$\min_{i \in [N]} (x - \mu_i)^2 = \min_{\pi \in \Delta_K} \sum_{i \in [N]} (x - \mu_i)^2 \pi_i \tag{7.40}$$

$$= \min_{\pi \in \Delta_K} \mathbb{E}_{I \sim \pi}\left[(x - \mu_I)^2\right], \tag{7.41}$$

where I is a random index in $[N]$ distributed according to the discrete probability mass function $\pi \in \Delta_K$. Substituting this back to the definition in (7.20), we get

$$e_\gamma(x_n, \mathbf{M}) = \sup_{x \in \mathbb{R}^d} \min_{\pi \in \Delta_K} \mathbb{E}_{I \sim \pi}\left[(x - \mu_I)^2\right] - \gamma (x - x_\circ)^2,$$

Note that the objective is strictly concave in $x \in \mathbb{R}^d$ for $\gamma > 1$, affine in $\pi \in \Delta_K$. As the simplex $\Delta_K \subset \mathbb{R}^n$ is convex and compact, and assuming $\mathbb{R}^d \subseteq \mathbb{R}$ is convex, the $\sup$ and the $\min$ can be exchanged by Sion's minimax theorem [212, Cor. 3.3] to yield (7.21). Furthermore, for $\mathbb{R}^d = \mathbb{R}$, the inner min-max problem in (7.21) admits a saddle point $(x^\star, \pi^\star) \in \mathbb{R}^d \times \Delta_K$ such that

$$e_\gamma = \sup_{x \in \mathbb{R}^d} \mathbb{E}_{I^\star \sim \pi^\star}\left[(x - \mu_{I^\star})^2\right] - \gamma (x - x_\circ)^2, \tag{7.42}$$

$$= \min_{\pi \in \Delta_K} \mathbb{E}_{I \sim \pi}\left[(x^\star - \mu_I)^2\right] - \gamma (x^\star - x_\circ)^2. \tag{7.43}$$

By strict concavity and smoothness of the objective (7.42) in x, the supremum is achieved by the unique stationary point $x^\star \in \mathbb{R}^d$ that vanishes the gradient wrt x. Furthermore, any $\pi^\star \in \Delta_K$ that achieves the minimum in (7.43) is a saddle point.

The expression in (7.21) can be further simplified by explicitly taking the supremum of the quadratic objective over x, resulting in the equivalent expression

$$\min_{\pi \in \Delta_K} \mathbb{E}_{I\sim\pi}\left[(x_n - \mu_I)^2\right] + \frac{1}{\gamma - 1}(x_\circ - \mathbb{E}_{I\sim\pi}[\mu_I])^2 \tag{7.44}$$

■

7.D Proof of Theorem 7.3.4

Proof. Using (7.22), we rewrite the dual optimization problem as follows:

$$\inf_{\substack{\mu_1,\ldots,\mu_n \in \mathbb{R}^d, \\ \pi(\cdot):\mathbb{R}^d \to \Delta_K, \\ \gamma > 1}} \gamma r^2 + \mathbb{E}_{\substack{X_\circ \sim \mathbb{P}_\circ, \\ I\sim\pi(X_\circ)}}\left[(x_n - \mu_I)^2 + \frac{(x_n - \mathbb{E}[\mu_I \mid X_\circ])^2}{\gamma - 1}\right]$$

Denoting the objective by $F_\gamma(\mathbf{M}, \pi)$, note that it is differentiable and convex with respect to $\mu_1, \ldots, \mu_n$ for fixed $\pi(\cdot) : \mathbb{R}^d \to \Delta_K$ and $\gamma > 1$. Therefore, the minimum wrt $\mu_1, \ldots, \mu_n$ is attained at the stationary points $\{\mu_i^\star\}_{i=1}^n$ satisfying $\partial_k F_\gamma(\{\mu_i^\star\}_{i=1}^n, \pi) = 0$, which is equivalent to:

$$\mathbb{E}_{X_\circ\sim\mathbb{P}_\circ}\left[-2\pi_k(X_\circ)\left(x_n - \mu_k^\star - \frac{x_n - \mathbb{E}[\mu_I^\star \mid X_\circ]}{\gamma - 1}\right)\right] = 0.$$

Rearranging the terms, we get

$$\mu_k^\star = \frac{\mathbb{E}_{X_\circ\sim\mathbb{P}_\circ}[x^\star(X_\circ)\pi_k(X_\circ)]}{\mathbb{E}_{X_\circ\sim\mathbb{P}_\circ}[\pi_k(X_\circ)]} = \mathbb{E}_{\substack{x_n\sim\mathbb{P}_\circ, \\ I\sim\pi(X_\circ)}}[x^\star(X_\circ) \mid I = i].$$

The minimum wrt π for fixed $\mathbf{M}$ and $\gamma > 1$ is achieved by (**??**) according to Lemma 7.3.3. Finally, the minimum wrt $\gamma > 1$ when the rest of the variables are fixed is achieved by the stationary point $\partial_\gamma F_{\gamma^\star}(\{\mu_i^\star\}_{i=1}^n, \pi^\star) = 0$ or equivalently:

$$r^2 - \frac{\mathbb{E}_{\substack{X_\circ\sim\mathbb{P}_\circ, \\ I^\star\sim\pi^\star(X_\circ)}}\left[(x_n - \mathbb{E}[\mu_I^\star \mid X_\circ])^2\right]}{(\gamma^\star - 1)^2} = 0. \tag{7.45}$$

■

7.E Proof of Lemma 7.4.1

Proof. It is clear that $\pi^\star_\gamma(x_\circ) = \delta_k$ if and only if $x^\star_\gamma(x_\circ) = x_\circ + (\gamma-1)^{-1}(x_\circ - \widehat{x}_k)$ and $|x^\star_\gamma(x_\circ) - \widehat{x}_k| < |x^\star_\gamma(x_\circ) - \widehat{x}_l|$ for $l \neq k$ by Theorem 7.3.4. Rearranging the last term, one can obtain the equivalence of the last condition to $x_\circ \in \Phi_{\gamma,k}$. Note that since centroids are distinct, any given point can have, at most, two closest points. This means that $\pi_\gamma(x_\circ)$ cannot be a mixture of more than two centroids. In that case, $\pi^\star_\gamma(x_\circ) = p\delta_k + (1-p)\delta_{k+1}$ for $p \in (0,1)$ if and only if $x^\star_\gamma(x_\circ) = x_\circ + (\gamma-1)^{-1}(x_\circ - p\widehat{x}_k - (1-p)\widehat{x}_{k+1})$ and $|x^\star_\gamma(x_\circ) - \widehat{x}_k| = |x^\star_\gamma(x_\circ) - \widehat{x}_{k+1}| < |x^\star_\gamma(x_\circ) - \widehat{x}_l|$ for $l \neq k, k+1$ by Theorem 7.3.4. This is possible if and only if $x^\star_\gamma(x_\circ) = \widehat{x}_{k+\frac{1}{2}}$, *i.e.*, the midpoint between two consecutive centroids, and thus $x_\circ = \widehat{x}_{k+\frac{1}{2}} + \gamma^{-1}\left(p\widehat{x}_k + (1-p)\widehat{x}_{k+1} - \widehat{x}_{k+\frac{1}{2}}\right)$. The range of $p \in [0,1]$ allows $x_\circ$ to take values only in $\Phi_{\gamma,k+\frac{1}{2}}$. Rearranging this relationship between $p \in [0,1]$ and $x_\circ \in \Phi_{\gamma,k+\frac{1}{2}}$ to express p in terms of $x_\circ$ yields $p = \frac{1}{2}(1 - \gamma q_{k+\frac{1}{2}}(x_\circ))$. ■

7.F Proof of Theorem 7.4.2

Proof. Denote by $F_\gamma(\{\widehat{x}_i\}_{i=1}^n, \pi)$ the objective used in the proof of Theorem 7.3.4. Then, given $\{\widehat{x}_i^{(t)}\}_{i=1}^n$ at iteration $t \in \mathbb{N}$, we have that

$$\pi^{(t)} \in \operatorname*{arg\,min}_{\pi(\cdot):\mathcal{X}\to\Delta_n} F_\gamma(\{\widehat{x}_i^{(t)}\}_{i=1}^n, \pi). \tag{7.46}$$

Thus, we get

$$F_\gamma(\{\widehat{x}_i^{(t)}\}_{i=1}^n, \pi^{(t)}) = \mathrm{QE}_r(\{\widehat{x}_k^{(t)}\}_{k\in[N]}). \tag{7.47}$$

Similarly, given $\pi^{(t)}$, the next iterate of centroids $\{\widehat{x}_i^{(t+1)}\}_{i=1}^n$ are given by:

$$\{\widehat{x}_i^{(t+1)}\}_{i=1}^n = \operatorname*{arg\,min}_{\widehat{x}_1,\ldots,\widehat{x}_n} F_\gamma(\{\widehat{x}_i\}_{i=1}^n, \pi^{(t)}). \tag{7.48}$$

Therefore, we get

$$\begin{aligned}
\mathrm{QE}_r(\{\widehat{x}_k^{(t+1)}\}_{k\in[N]}) &= \inf_\pi F_\gamma(\{\widehat{x}_i^{(t+1)}\}_{i=1}^n, \pi) \\
&\leq F_\gamma(\{\widehat{x}_i^{(t+1)}\}_{i=1}^n, \pi^{(t)}) \\
&\leq F_\gamma(\{\widehat{x}_i^{(t)}\}_{i=1}^n, \pi^{(t)}) = \mathrm{QE}_r(\{\widehat{x}_k^{(t)}\}_{k\in[N]}).
\end{aligned}$$

■

Part III

Non-rational Control

Chapter 8

INTRODUCTION AND MOTIVATION

8.1 Introduction

Autonomous control systems now permeate across critical domains, from self-driving cars and power grids to manufacturing and rescue robotics, where their decisions carry major economic and safety implications. Reliability is essential: a failure in control can lead to major blackouts or endanger lives. Yet, these systems face environments of unprecedented complexity and uncertainty, including interacting subsystems, adversarial disturbances, noisy sensors, and incomplete models. Simultaneously, they must meet stringent performance demands under resource constraints, requiring trade-offs between objectives like cost, robustness, and safety.

Traditional design methods fall short. Stochastic approaches (e.g., LQG, $\mathcal{H}_2$) optimize average-case performance but fail under model mismatch. Robust control (e.g., $\mathcal{H}_\infty$) guards against worst-case scenarios but at the cost of excessive conservatism, ignoring useful statistical information. Recent data-driven methods enabled by machine learning and reinforcement learning show strong empirical results but remain brittle under distribution shifts and lack formal guarantees, risking unsafe behavior in novel situations.

To address these limitations, a third paradigm has emerged: distributionally robust control (DRC). DRC optimizes performance against worst-case distributions within data-informed ambiguity sets, bridging the gap between stochastic and robust designs. Complementary risk-aware methods (e.g., CVaR, entropic penalties) and multi-criteria formulations (e.g., mixed $\mathcal{H}_2/\mathcal{H}_\infty$, chance constraints) offer further tools to explicitly manage uncertainty, tail risk, and competing performance objectives.

Challenges with Controller Synthesis

Despite their promise, turning these elegant formulations into practical, real-time deployable controllers poses severe analytical and computational hurdles. Finite-horizon formulations lead to high-dimensional and intractable optimization problems that scale poorly with the time horizon, making real-time implementation infeasible. While receding horizon strategies (e.g., model predictive control) offer a workaround, they often lack rigorous stability guarantees and may result in non-smooth, erratic, or myopic behaviors that hinder long-term performance.

A more principled solution lies in infinite-horizon controller synthesis, which offers provably stable and performance-guaranteed policies with efficient online implementation. However, designing optimal infinite-horizon controllers for generalized objectives remains a deeply challenging problem. To date, exact closed-form solutions are known only for a few special cases, such as LQG, $\mathcal{H}_\infty$, regret-optimal, and entropic risk-sensitive control, while other widely studied formulations, including mixed $\mathcal{H}_2/\mathcal{H}_\infty$, has so far been unresolved.

Furthermore, unlike LQG and $\mathcal{H}_\infty$ problems, many of these problems, including the mixed $\mathcal{H}_2/\mathcal{H}_\infty$ and Wasserstein DRC, generally admit only infinite-dimensional (i.e., non-rational) optimal controllers without finite-state-space representations, posing major barriers to practical real-time deployment. Although approximation strategies—such as constraining to finite impulse response (FIR) controllers or closed-loop responses—can reduce the problem to a tractable form, they introduce new issues: they may fail to capture long-range dependencies or incur significant suboptimality unless the memory length is taken prohibitively large.

Non-rational Control Framework

These realities underscore the need for a new generation of practical controller-synthesis techniques that

- accommodate a wide array of performance metrics, including risk mitigation and robustness under uncertainty, and other multi-objective criteria;
- ensure closed-loop stability,
- impose minimal computational overhead in real-time implementations,
- scale efficiently to high-dimensional, large-scale systems, and
- achieve near-optimal performance with provably negligible suboptimality gaps.

Recent breakthroughs including our own results on infinite-horizon Wasserstein distributionally robust control [90], [124], [125], filtering [92], [122] and the exact infinite-dimensional solution of the mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control [158] reveal a unifying lens: *non-rational control framework*.

Embracing the infinite-dimensional nature of control problems, this unified framework offers new analytical and computational tools that render otherwise intractable controller design tasks both solvable and practically implementable. Crucially, it adopts an optimize-then-approximate paradigm, enabling synthesis of provably near-optimal, stabilizing finite-dimensional (rational) state-space controllers, even when the true optimum resides in an infinite-dimensional (non-rational) policy class. The framework is built on the following key components:

1. **Infinite-dimensional convex duality.** By formulating the control objective at the operator level and invoking convex duality, the original design problem with generalized performance criteria is recast as a tractable optimization program.

2. **Efficient numerical solution.** Exploiting the Fourier-domain (transfer-function) representation of the dual variables allows the use of standard, scalable optimization algorithms (e.g., first-order methods) to compute the exact infinite-dimensional optimum.

3. **Rational controller synthesis.** A novel rational-approximation scheme translates the infinite-dimensional solution into finite-order controllers that are guaranteed to be stabilizing and within a quantifiable performance gap, enabling practical real-time deployment without sacrificing performance.

The non-rational control framework integrates and extends $\mathcal{H}_2$, $\mathcal{H}_\infty$, distributionally robust, risk-sensitive, regret-optimal, and multi-objective control paradigms into a cohesive framework that enables scalable real-time implementation. The underlying numerical optimization and rational controllers synthesis algorithms are highly efficient in terms of computational complexity and horizon independent, as opposed to finite-horizon formulations which scale with the time-horizon. Moreover, the resulting near-optimal rational controllers significantly outperform those derived from restrictive policy classes, such as those obtained from Finite Impulse Response (FIR) approximation.

8.2 Infinite-Horizon Control via Closed-Loop System Responses

Consider a linear time-invariant (LTI) dynamical system in discrete-time:

$$x_{t+1} = Ax_t + B_u u_t + B_w w_t, \tag{8.1}$$

where $x_t \in \mathbb{R}^{d_x}$ is the *state*, $u_t \in \mathbb{R}^{d_u}$ is the *control input*, and $w_t \in \mathbb{R}^{d_w}$ is the *disturbance* at time t. Here, (A, B_u) is stabilizable and (A, B_w) is controllable, namely there exists a matrix $K_0 \in \mathbb{R}^{d_u \times d_x}$ such that $A - B_u K_0$ is stable and the controllability matrix $\begin{bmatrix} B_w & AB_w & \dots & A^{d_x-1}B_w \end{bmatrix}$ is full-rank.

We restrict attention to strictly causal LTI controllers that observe past disturbances to generate control inputs[1]. These controllers take the form:

$$u_t = \sum_{s \le t} K_{t-s} w_s, \text{ with } K_0 = 0 \tag{8.2}$$

where $(K_t)_{t=0}^{\infty}$ are the Markov parameters (or the impulse response) of the controller.

We are interested in the evolution of the state and control input trajectories as functions of the control policy $(K_s)_{s=0}^{\infty}$ and the disturbance process (w_t). When only a finite time horizon $T > 0$ is considered, and assuming $x_0 = 0$ for simplicity, the relationship between the state, control input, and disturbance trajectories can be compactly expressed as:

$$\mathbf{x}_T = \mathcal{P}_{xu,T}\mathbf{u}_T + \mathcal{P}_{xw,T}\mathbf{w}_T, \tag{8.3a}$$

$$\mathbf{u}_T = \mathcal{K}_T \mathbf{w}_T, \tag{8.3b}$$

where the stacked trajectories are defined as:

$$\mathbf{x}_T := \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_T \end{bmatrix}, \mathbf{u}_T := \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_T \end{bmatrix}, \mathbf{w}_T := \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_T \end{bmatrix}. \tag{8.4}$$

For $i \in \{u, w\}$, the matrices $\mathcal{P}_{xi,T}$ and $\mathcal{K}_T$ are strictly block lower-triangular and defined by:

$$\mathcal{P}_{xi,T} := \begin{bmatrix} 0 & & & & \\ B_i & 0 & & & \\ AB_i & B_i & 0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ A^{T-1}B_i & A^{T-2}B_i & \dots & B_i & 0 \end{bmatrix}, \tag{8.5}$$

$$\mathcal{K}_T := \begin{bmatrix} 0 & & & & \\ K_1 & 0 & & & \\ K_2 & K_1 & 0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ K_T & K_{T-1} & \dots & K_1 & 0 \end{bmatrix}. \tag{8.6}$$

[1]This is closely related to the standard state-feedback policies, as one can recover w_{t-1} from x_t, x_{t-1}, u_{t-1} when B_w is full column rank, via $w_{t-1} = (B_w^T B_w)^{-1} B_w^T (x_t - Ax_{t-1} - B_u u_{t-1})$.

These block-Toeplitz matrices succinctly encode two convolutional operations: one maps past inputs and disturbances to the state trajectory (via the system dynamics), and the other maps past disturbances to the control inputs (via the strictly causal controller).

As the horizon extends to infinity (*i.e.*, $T \to \infty$), the finite-length trajectories $\mathbf{x}_T, \mathbf{u}_T, \mathbf{w}_T$ become infinite-dimensional sequences $\mathbf{x}, \mathbf{u}, \mathbf{w}$ and the block-Toeplitz matrices $\mathcal{P}_{xu,T}, \mathcal{P}_{xw,T}, \mathcal{K}_T$ naturally extend to infinite-dimensional block-Toeplitz operators $\mathcal{P}_{xu}, \mathcal{P}_{xw}, \mathcal{K}$. In this limit, the relationship between state, control input, and disturbance trajectories analogously satisfy:

$$\mathbf{x} = \mathcal{P}_{xu}\mathbf{u} + \mathcal{P}_{xw}\mathbf{w}, \tag{8.7a}$$

$$\mathbf{u} = \mathcal{K}\mathbf{w}, \tag{8.7b}$$

Observe that (8.7) constitutes an infinite system of linear equations in which the disturbance process $\mathbf{w}$ serves as the independent input. Eliminating $\mathbf{u}$ via substituting the control law $\mathbf{u} = \mathcal{K}\mathbf{w}$ into the state equation (8.7a) yields the closed-loop transfer operator mapping disturbances to the resulting state and control trajectories:

$$\mathcal{T}_{\mathcal{K}} : \mathbf{w} \mapsto \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix}, \quad \mathcal{T}_{\mathcal{K}} \coloneqq \begin{bmatrix} \mathcal{P}_{xu}\mathcal{K} + \mathcal{P}_{xw} \\ \mathcal{K} \end{bmatrix} \tag{8.8}$$

The block-operator $\mathcal{T}_{\mathcal{K}}$ characterizes the closed-loop behavior induced by the controller $\mathcal{K}$, *i.e.*, describing how disturbances propagate through both the plant and the controller. We will similarly use the notation $\mathcal{T}_{\mathcal{K}_T} : \mathbf{w}_T \mapsto \begin{bmatrix} \mathbf{x}_T \\ \mathbf{u}_T \end{bmatrix}$ to denote the closed-loop transfer matrix over a finite horizon $T > 0$.

Many control problems can be cast as optimization tasks, where the goal is to design a controller $\mathcal{K}$, typically subject to constraints such as stability, sparsity, or structural requirements, that optimizes a performance objective defined in terms of the closed-loop transfer operator $\mathcal{T}_{\mathcal{K}}$. More formally,

$$\inf_{\text{s.causal } \mathcal{K}} f(\mathcal{T}_{\mathcal{K}}) \quad \text{subject to} \quad \mathcal{K} \in \mathscr{K}. \tag{8.9}$$

where f is the performance objective to be minimized and $\mathscr{K}$ is the set of admissible control policies. The formulation naturally incorporates both constraints on the controller $\mathcal{K}$ through the admissible set $\mathscr{K}$ and on the induced closed-loop behavior captured by $\mathcal{T}_{\mathcal{K}}$ through f.

At this level of generality, however, it is difficult to characterize or compute optimal solutions, even when the objective function f and the admissible set $\mathscr{K}$ are both convex, due to the infinite-dimensional nature of the problem. Consequently, much of the research in control theory has historically focused on specific problem classes of significant practical importance, most notably $\mathscr{H}_2$, $\mathscr{H}_\infty$, and regret optimal control [191].

In the following subsections, we review several standard control problems and outline their corresponding solution strategies. We conclude with a discussion of various other control problems, including mixed objective criteria and Wasserstein distributionally robust control, highlighting the key challenges associated with their implementation and tractable synthesis.

$\mathscr{H}_2$-Optimal Control via Wiener-Hopf Technique

Problem 8.2.1 ($\mathscr{H}_2$-Optimal Control)**.** Find a strictly causal LTI control policy $\mathcal{K}$ that minimizes the $\mathscr{H}_2$ norm of the closed-loop transfer operator $\mathcal{T}_{\mathcal{K}_T}$, *i.e.*,

$$\inf_{\text{s.causal } \mathcal{K}} \|\mathcal{T}_{\mathcal{K}}\|_2^2 = \operatorname{tr}(\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}), \tag{H$_2$}$$

Here, the trace of an infinite-dimensional block Toeplitz operator $\mathcal{X}$ is defined by the Fourier integral

$$\operatorname{tr}(\mathcal{X}) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \operatorname{Tr}(X(e^{j\omega}))d\omega \tag{8.10}$$

where Tr is the usual trace of matrices and $X(z) := \sum_{k=-\infty}^{\infty} X_k z^{-k}$ at $z \in \mathbb{C}$ is the transfer function representation of the Toeplitz operator $\mathcal{X}$.

The $\mathscr{H}_2$ objective naturally arises in various operational contexts, most notably in the stochastic Linear-Quadratic Regulator (LQR) problem. Given positive-semidefinite state and control weighting matrices $Q \in \mathbb{S}^n+$ and $R \in \mathbb{S}_{++}^{d_u}$, the controller seeks to minimize the average expected cost:

$$\lim_{T \to \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=0}^{T} x_t^\top Q x_t + u_t^\top R u_t \right] \tag{8.11}$$

Here, the disturbances are modeled as random variables drawn from a known probability distribution, and the expectation is taken with respect to this distribution. Specifically, suppose disturbances constitute a white Gaussian noise process, *i.e.*, $(w_t) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \Sigma)$ where $\Sigma \succ 0$.

For notational convenience, one can take the positive-definite matrices Q, R and Σ to be identity matrices without loss of generality[2] by reparameterizing the dynamical variables as $x_t \mapsto Q^{1/2}x_t$, $u_t \mapsto R^{1/2}u_t$, $w_t \mapsto \Sigma^{1/2}w_t$ and the state-space parameters as $A \mapsto Q^{1/2}AQ^{-1/2}$, $B_u \mapsto Q^{1/2}B_uR^{-1/2}$ and $B_w \mapsto Q^{1/2}B_w\Sigma^{-1/2}$.

The cumulative LQR cost upto time T can be simplified as

$$\mathbb{E}\left[\sum_{t=0}^{T} x_t^\top Q x_t + u_t^\top R u_t\right] = \mathbb{E}\left[\|\mathbf{x}_T\|^2 + \|\mathbf{u}_T\|^2\right]. \tag{8.12}$$

Using the finite-horizon closed-loop mapping, we can write the cumulative cost in terms of the controller $\mathcal{K}$ and the disturbances $\mathbf{w}_T$ as

$$\begin{aligned}\mathbb{E}\left[\mathbf{w}_T^* \mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T} \mathbf{w}_T\right] &= \mathbb{E}\left[\mathrm{Tr}(\mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T} \mathbf{w}_T \mathbf{w}_T^*)\right], \\ &= \mathrm{Tr}(\mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T}\, \mathbb{E}[\mathbf{w}_T \mathbf{w}_T^*]),\end{aligned}$$

where we used the cyclical property of trace and the linearity of expectation and trace. The expectation $\mathbb{E}[\mathbf{w}_T\mathbf{w}_T^*]$ is simply the autocorrelation matrix of the disturbance process, which is the identity matrix. Thus, the cumulative expected cost incurred upto time T is the squared Frobenius norm of the finite-horizon closed-loop mapping: $\|\mathcal{T}_{\mathcal{K}_T}\|_F^2 = \mathrm{Tr}(\mathcal{T}_{\mathcal{K}_T}^*\mathcal{T}_{\mathcal{K}_T})$. Therefore, the squared $\mathscr{H}_2$-norm in the infinite-horizon setting can be though of as the limit of horizon-normalized squared Frobenius norm:

$$\|\mathcal{T}_{\mathcal{K}}\|_2^2 = \lim_{T\to\infty} \frac{1}{T}\|\mathcal{T}_{\mathcal{K}_T}\|_F^2 = \lim_{T\to\infty} \frac{1}{T}\,\mathrm{Tr}(\mathcal{T}_{\mathcal{K}_T}^*\mathcal{T}_{\mathcal{K}_T}). \tag{8.13}$$

Before we outline the solution strategy for the $\mathscr{H}_2$ problem, it is worthwhile to investigate the quadratic term $\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}$. Expanding this term and completing the squares, the controller term $\mathcal{K}$ can be isolated as follows:

$$\begin{aligned}\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}} &= (\mathcal{P}_{xu}\mathcal{K} + \mathcal{P}_{xw})^*(\mathcal{P}_{xu}\mathcal{K} + \mathcal{P}_{xw}) + \mathcal{K}^*\mathcal{K}, \\ &= (\mathcal{K} - \mathcal{K}_\circ)^*(\mathcal{I} + \mathcal{P}_{xu}^*\mathcal{P}_{xu})(\mathcal{K} - \mathcal{K}_\circ) \\ &\quad + \mathcal{P}_{xw}(\mathcal{I} - \mathcal{P}_{xu}(\mathcal{I} + \mathcal{P}_{xu}^*\mathcal{P}_{xu})^{-1}\mathcal{P}_{xu}^*)\mathcal{P}_{xw},\end{aligned}$$

where $\mathcal{K}_\circ$ is defined as[3]

$$\mathcal{K}_\circ := -(\mathcal{I} + \mathcal{P}_{xu}^*\mathcal{P}_{xu})^{-1}\mathcal{P}_{xu}^*\mathcal{P}_{xw}. \tag{8.14}$$

[2]While this transformation requires $Q \succ 0$, a similar reduction can be performed even when Q is singular by defining a new output variable $s_t = Q^{1/2}x_t$ and considering the closed-loop mapping to s instead of the state.

[3]One must exercise particular caution when working with inverses or other analytic functions of infinite-dimensional operators, as the resulting operators may not belong to the same space (e.g., trace class) unlike the finite-dimensional case, where such operations are typically well-behaved. We set aside these technical considerations for now and focus on the core conceptual insights.

Clearly, the last term in the expansion of $\mathcal{T}_{\mathcal{K}}^{*}\mathcal{T}_{\mathcal{K}}$ does not depend on $\mathcal{K}$ and therefore is an additive constant in after the taking the trace. When the strict causality constraint is ignored, the squared $\mathscr{H}_2$ objective $\mathcal{K} \mapsto \operatorname{tr}(\mathcal{T}_{\mathcal{K}}^{*}\mathcal{T}_{\mathcal{K}})$ admits its unique minimum at $\mathcal{K}_\circ$, which we call as the optimal non-causal controller, and the constant simply becomes

$$\mathcal{T}_{\mathcal{K}_\circ}^{*}\mathcal{T}_{\mathcal{K}_\circ} = \mathcal{P}_{xw}(\mathcal{I} - \mathcal{P}_{xu}(\mathcal{I} + \mathcal{P}_{xu}^{*}\mathcal{P}_{xu})^{-1}\mathcal{P}_{xu}^{*})\mathcal{P}_{xw} \tag{8.15}$$

One approach to synthesizing the strictly causal $\mathscr{H}_2$-optimal controller is to explicitly enforce the strict causality constraint via Lagrange multipliers and derive the associated Karush–Kuhn–Tucker (KKT) conditions. However, we instead present a significantly more elegant method, originally developed by Wiener and Hopf [239].

Theorem 8.2.2 (Solution of $\mathscr{H}_2$ Control via Wiener-Hopf Technique)**.** *The $\mathscr{H}_2$-optimal control problem* (H$_2$) *admits the following unique optimal solution:*

$$\mathcal{K}_2 \coloneqq \Delta^{-1}\{\Delta\mathcal{K}_\circ\}_{+}, \tag{8.16}$$

where Δ is causal with bounded and causal inverse Δ^{-1} and satisfies the canonical spectral factorization $\Delta^{*}\Delta = \mathcal{I} + \mathcal{P}_{xu}^{*}\mathcal{P}_{xu}$.

Here, the notation $\{\mathcal{X}\}_{+}$ and $\{\mathcal{X}\}_{-}$ refer to the strictly causal (*i.e.*, strictly lower-block-triangular) and anticausal (*i.e.*, upper-block-triangular) components of the block Toeplitz operator $\mathcal{X}$, respectively. These projections take a particularly simple form when $\mathcal{X}$ is expressed in its power series representation:

$$\{\mathcal{X}\}_{+}(z) = \sum_{t=1}^{\infty} X_t z^{-t}, \quad \{\mathcal{X}\}_{-}(z) = \sum_{t=-\infty}^{0} X_t z^{-t}. \tag{8.17}$$

Below, we provide a proof sketch of Theorem 8.2.2 in the infinite-horizon setting. To keep the presentation streamlined, we suppress some of the finer technical details related to spectral factorization, simply viewing it as the infinite-dimensional analogue of Cholesky decomposition.

Proof. Since $\mathcal{I} + \mathcal{P}_{xu}^{*}\mathcal{P}_{xu}\Delta^{*}\Delta$ is strictly positive-definite and bounded below by the identity map $\mathcal{I}$, it admits a canonical spectral factorization $\Delta^{*}\Delta = \mathcal{I} + \mathcal{P}_{xu}^{*}\mathcal{P}_{xu}$ where Δ is causal and non-singular, and its inverse Δ^{-1} is causal and bounded. We modify the expansion of $\mathcal{T}_{\mathcal{K}}^{*}\mathcal{T}_{\mathcal{K}}$ by distributing the operators Δ and Δ^{*}:

$$\mathcal{T}_{\mathcal{K}}^{*}\mathcal{T}_{\mathcal{K}} = (\Delta\mathcal{K} - \Delta\mathcal{K}_\circ)^{*}(\Delta\mathcal{K} - \Delta\mathcal{K}_\circ) + \mathcal{T}_{\mathcal{K}_\circ}^{*}\mathcal{T}_{\mathcal{K}_\circ} \tag{8.18}$$

As the last term is independent of $\mathcal{K}$, the optimal solution of Problem 8.2.1 is equivalent to the solution of the following:

$$\inf_{\text{s.causal } \mathcal{K}} \|\Delta\mathcal{K} - \Delta\mathcal{K}_\circ\|_2^2. \tag{8.19}$$

Notice that the multiplication $\Delta\mathcal{K}$ is always causal whereas $\Delta\mathcal{K}_\circ$ is non-causal. The squared $\mathscr{H}_2$ norm of any operator $\mathcal{X}$ can be decomposed[4] as the sum of squared $\mathscr{H}_2$ norms of its s. causal $\{\mathcal{X}\}_+$ and anticausal projections $\{\mathcal{X}\}_-$, namely

$$\|\mathcal{X}\|_2^2 = \|\{\mathcal{X}\}_+\|_2^2 + \|\{\mathcal{X}\}_-\|_2^2. \tag{8.20}$$

Therefore, we can decompose the objective as

$$\inf_{\text{s.causal } \mathcal{K}} \|\Delta\mathcal{K} - \{\Delta\mathcal{K}_\circ\}_+\|_2^2 + \|\{\Delta\mathcal{K}_\circ\}_-\|_2^2. \tag{8.21}$$

Setting $\Delta\mathcal{K} = \{\Delta\mathcal{K}_\circ\}_+$ and noting that the inverse Δ^{-1} of the causal spectral factor is also causal, one gets the unique optimal controller by inverting $\mathcal{K} = \Delta^{-1}\{\Delta\mathcal{K}_\circ\}_+$. ■

This procedure, often referred to as the Wiener–Hopf technique, amounts to orthogonally projecting the non-causal term $\Delta\mathcal{K}_\circ$ onto the subspace of strictly causal operators, thereby yielding the $\mathscr{H}_2$-optimal controller.

$\mathscr{H}_\infty$ and Regret Optimal Control via Nehari Problem

The $\mathscr{H}_\infty$ norm of $\mathcal{T}_\mathcal{K}$ corresponds to its operator norm as a mapping from $\ell_2(\mathbb{Z})$ to $\ell_2(\mathbb{Z})$, *i.e.*,

$$\|\mathcal{T}_\mathcal{K}\|_\infty := \sup_{\mathbf{w}\in\ell_2/\{0\}} \frac{\|\mathcal{T}_\mathcal{K}\mathbf{w}\|}{\|\mathbf{w}\|} = \max_{\omega\in[0,2\pi)} \overline{\sigma}(T_K(\mathrm{e}^{\jmath\omega})) \tag{8.22}$$

Generalized Control Objectives

In this section, we introduce a broad class of optimal control problems that can be formulated as minimization of an objective function defined over the squared closed-loop transfer operator, $|\mathcal{T}_\mathcal{K}|^2 = \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}$. We show that various well-known control problems fall into this class, including standard $\mathscr{H}_2$ and $\mathscr{H}_\infty$ control as well as risk-sensitive, Wasserstein distributionally robust, and mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control.

To be more concrete, consider the following general optimal control problem:

[4]For intuition, simply think of the $\mathscr{H}_2$ norm as the infinite-dimensional analogue of Frobenius norm

Problem 8.2.3 (General Control Problem)**.** Given a convex function $f : \mathscr{L}_\infty \to \overline{\mathbb{R}}$, find a causal and stabilizing LTI controller $\mathcal{K} \in \mathscr{K}$, that minimizes the objective $\mathcal{K} \mapsto f(\mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K})$, *i.e.*,

$$\boxed{p_\star := \inf_{\mathcal{K} \in \mathscr{K}} \; f(\mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K}),} \tag{P}$$

where $p_\star \in \overline{\mathbb{R}}$ is the optimal value.

It should be noted that (P) may not be convex program even though the function $\mathcal{C} \mapsto f(\mathcal{C})$ and the constraint $\mathcal{K} \in \mathscr{K}$ are convex. A trivial counter example is $f(\mathcal{C}) = -\operatorname{tr}(\mathcal{C})$, in which case, (P) becomes concave and the optimal value is $p_\star = -\infty$.

Therefore, we make the following assumption to ensure (P) is a well-posed convex program.

Assumption 8.2.4. The convex function $f : \mathscr{L}_\infty(\mathbb{S}_+^{d_w}) \to \overline{\mathbb{R}}$ is monotonically non-decreasing over the psd cone, *i.e.*, $f(\mathcal{C}_1) \leq f(\mathcal{C}_2)$ for $0 \preccurlyeq \mathcal{C}_1 \preccurlyeq \mathcal{C}_2$.

This sufficient condition on f implies the convexity of $\mathcal{K} \mapsto f(\mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K})$ since the quadratic mapping $\mathcal{K} \mapsto \mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K}$ is convex with respect to the positive-definite order of operators, *i.e.*, $\mathcal{T}_\lambda^* \mathcal{T}_\lambda \preccurlyeq \lambda \mathcal{T}_0^* \mathcal{T}_0 + (1-\lambda) \mathcal{T}_1^* \mathcal{T}_1$ where $\mathcal{T}_{\mathcal{K}_\lambda} = \mathcal{T}_{\mathcal{K}_0} + \lambda(\mathcal{T}_{\mathcal{K}_1} - \mathcal{T}_{\mathcal{K}_0})$ for any $\mathcal{K}_0$, $\mathcal{K}_1$ and $\lambda \in [0,1]$

We first illustrate various control problems subsumed by this class of objectives. Henceforth, we use the notation

$$\langle \mathcal{C}, \mathcal{M} \rangle := \frac{1}{2\pi} \int_{-\pi}^{\pi} \operatorname{Tr}(C(e^{j\omega})^* M(e^{j\omega})) d\omega \tag{8.23}$$

Example 8.2.5 (Mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control).

$$\inf_{\text{s.causal } \mathcal{K}} \|\mathcal{T}_\mathcal{K}\|_2^2 \quad \text{subject to} \quad \|\mathcal{T}_\mathcal{K}\|_\infty \leq \gamma. \tag{8.24}$$

with the associated objective function

$$f_{\gamma, \mathscr{H}_2/\mathscr{H}_\infty}(\mathcal{C}) := \operatorname{tr}(\mathcal{C}) + \begin{cases} 0, & \|\mathcal{C}\|_\infty \leq \gamma^2, \\ +\infty, & \text{o.w.} \end{cases} \tag{8.25}$$

Unlike the pure $\mathscr{H}_2$ and pure $\mathscr{H}_\infty$ controllers, it was proved by [165] that the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller is non-rational whenever the $\mathscr{H}_\infty$ constraint is active. Since non-rational functions do not admit finite-dimensional state-space realizations, researchers mostly focused either on finite-dimensional approximations or more tractable auxiliary objectives.

Example 8.2.6 (**H_{2p} optimal control**)**.** For $p \in [1, \infty]$

$$\inf_{\text{s.causal } \mathcal{K}} \|\mathcal{T}_{\mathcal{K}}\|_{2p}^2 = \|\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}\|_p \tag{8.26}$$

$$f_{H_{2p}}(\mathcal{C}) := \|\mathcal{C}\|_p \tag{8.27}$$

Example 8.2.7 (**Risk-sensitive control**)**.** The risk-sensitive control objective aims to minimize an exponential cost, formulated below

$$\inf_{\text{s.causal } \mathcal{K}} \gamma \log \left(\mathbb{E}_{\mathbf{w} \sim \mathbb{P}_\circ} \left[\mathrm{e}^{\gamma^{-1} \mathbf{w}^* \mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}} \mathbf{w}} \right] \right), \tag{8.28}$$

where $\gamma > 0$ is the risk parameter and $\mathbb{P}_\circ$ is a nominal probability distribution of the disturbances. The expectation above should be understood formally as the time-averaged limit of finite-horizon risk-sensitive costs. The convex function corresponding to this problem is given by

$$f_{\gamma,\text{RS}}(\mathcal{C}) := \gamma \log \left(\mathbb{E}_{\mathbf{w} \sim \mathbb{P}_\circ} \left[\mathrm{e}^{\gamma^{-1} \mathbf{w}^* \mathcal{C} \mathbf{w}} \right] \right) \tag{8.29}$$

With the decreasing value of γ, The risk-sensitive objective resolves the gap between smaller and larger cost values. It penalizes higher cost levels relatively more than the smaller values as γ decreases. This essentially incentivizes the controller to be more risk-averse to reduce the chances of yielding higher costs.

In the special case of the nominal distribution $\mathbb{P}_\circ$ of disturbances forming a stationary Gaussian process with auto-covariance operator $\mathcal{M}_\circ \succ 0$, the risk-sensitive objective simplifies further to

$$\inf_{\text{s.causal } \mathcal{K}} -\frac{\gamma}{2} \operatorname{logdet}(\mathcal{I} - 2\gamma^{-1} \mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}} \mathcal{M}_\circ), \tag{8.30}$$

where $\operatorname{logdet}(\cdot)$ should be understood as $\operatorname{tr}(\log(\cdot))$.

The corresponding convex function then becomes

$$f_{\gamma,\text{RS}}(\mathcal{C}) := -\frac{\gamma}{2} \operatorname{logdet}(\mathcal{I} - 2\gamma^{-1} \mathcal{C} \mathcal{M}_\circ) \tag{8.31}$$

Example 8.2.8 (**Wasserstein distributionally robust control**)**.** When the ambiguity set of plausible probability distributions of disturbances is constructed as a Wasserstein-2 ball, the distributionally robust controller can be obtained by solving the following primal optimization problem:

$$\inf_{\substack{\text{s.causal } \mathcal{K}, \\ \gamma > \|\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}\|_\infty}} \gamma r^2 + \gamma \operatorname{tr} \left[\left((\mathcal{I} - \gamma^{-1} \mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}})^{-1} - \mathcal{I} \right) \mathcal{M}_\circ \right], \tag{8.32}$$

where $\mathcal{M}_\circ \in \mathscr{L}_1^+$ is the auto-covariance operator of the nominal disturbance process, which is assumed to be weakly stationary, and γ is a Lagrange multiplier determined by the desired radius of the Wasserstein-2 ball, $r > 0$. The corresponding convex function for this optimization problem then becomes:

$$f_{\mathsf{W}_2}(\mathcal{C}) := \inf_{\gamma > \|\mathcal{C}\|_\infty} \gamma r^2 + \gamma \operatorname{tr}\left[\left((\mathcal{I} - \gamma^{-1}\mathcal{C})^{-1} - \mathcal{I}\right)\mathcal{M}_\circ\right]. \tag{8.33}$$

The suboptimal problem is

$$\inf_{\text{s.causal } \mathcal{K}} \gamma \operatorname{tr}\left[(\mathcal{I} - \gamma^{-1}\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K})^{-1}\right] \tag{8.34}$$

The corresponding objective function is

$$f_{\gamma,\mathsf{W}_2}(\mathcal{C}) := \gamma \operatorname{tr}\left[(\mathcal{I} - \gamma^{-1}\mathcal{C})^{-1}\right] \tag{8.35}$$

Remark 8.2.9 (Norm interpretation).

$$\sup_{\mathcal{M} \in \mathscr{M}} \sqrt{\langle \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}, \mathcal{M}\rangle} \tag{8.36}$$

is a norm for $\mathcal{T}_\mathcal{K}$ whenever $\mathscr{M} \cap \operatorname{int}(\mathscr{L}_1(\mathbb{S}_+)) \neq \varnothing$, that is, there exists a strictly positive definite element $\mathcal{M} \succ 0$ of $\mathscr{M}$.

Summary of Non-Rational Control Approach

Step 1: Derive the dual problem using Fenchel conjugate $f^*(\mathcal{M}) := \sup_{\mathcal{C} \succcurlyeq 0} \langle \mathcal{C}, \mathcal{M}\rangle - f(\mathcal{C})$ of the objective:

$$\sup_{\mathcal{M} \succcurlyeq 0} \inf_{\mathcal{K} \in \mathscr{K}} \langle \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}, \mathcal{M}\rangle - f^*(\mathcal{M}) \tag{8.37}$$

Step 2: Solve the inner minimization over $\mathcal{K}$ for a given $\mathcal{M} \succ 0$ using the Wiener-Hopf technique:

$$\mathcal{K}_\star = \Delta^{-1}\{\Delta\mathcal{K}_\circ\mathcal{L}\}_+\mathcal{L}^{-1} \tag{8.38}$$

where $\mathcal{L}\mathcal{L}^* = \mathcal{M}$ is the canonical spectral factorization.

Step 3: Derive the expression for the subgradient $\partial f^*(\mathcal{M})$ of the conjugate function f^* and solve for the saddle point $(\mathcal{K}_\star, \mathcal{M}_\star)$ to get the optimal controller:

$$\mathcal{T}_{\mathcal{K}_\star}^*\mathcal{T}_{\mathcal{K}_\star} \in \partial f^*(\mathcal{M}_\star) \tag{8.39}$$

$$\mathcal{K}_\star = \Delta^{-1}\{\Delta\mathcal{K}_\circ\mathcal{L}_\star\}_+\mathcal{L}_\star^{-1} \tag{8.40}$$

where $\mathcal{L}_\star\mathcal{L}_\star^* = \mathcal{M}_\star$ is the canonical spectral factorization.

Step 4: To solve numerically, derive the gradient $\partial f^*(\mathcal{M})$ in the Fourier domain, and simply implement a first-order gradient based method by propagating thee gradient information for each individual frequency over a large number of frequency samples on $[0, 2\pi]$

Step 5: Obtain a finite-dimensional approximation solution via rational approximation of the positive definite transfer function $M_\star(e^{j\omega})$ and then find the approximate controller via Wiener-Hopf technique

Notations

The letters $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ denote the set of natural numbers, integers, real, and complex numbers, respectively. $\mathbb{T}$ denotes the complex unit circle. For $z \in \mathbb{C}$, $|z|$ is its magnitude, and z^* is the conjugate. $\mathbb{S}_+^n$ denotes the set of positive semidefinite (psd) matrices of size $n \times n$. Bare calligraphic letters ($\mathcal{K}$, $\mathcal{M}$, etc.) are reserved for operators. $\mathcal{I}$ is the identity operator with a suitable block size. For an operator $\mathcal{M}$, its adjoint is $\mathcal{M}^*$. For a matrix A, its transpose is $A^\intercal$, and its Hermitian conjugate is A^*. For psd operators/matrices, $\succcurlyeq$ denotes the Löwner order. For a psd operator $\mathcal{M}$, both $\sqrt{\mathcal{M}}$ and $\mathcal{M}^{\frac{1}{2}}$ denote the PSD square-root. $\{\mathcal{M}\}_+$ and $\{\mathcal{M}\}_-$ denote the causal and strictly anti-causal parts of an operator $\mathcal{M}$. $M(z)$ denotes the z-domain transfer function of a Toeplitz operator $\mathcal{M}$. $\mathrm{tr}(\cdot)$ denotes the trace of operators and matrices. $\|\cdot\|$ is the usual Euclidean norm. $\|\cdot\|_\infty$ and $\|\cdot\|_2$ are the $\mathscr{H}_\infty$ operator) and $\mathscr{H}_2$ (Frobenius) norms, respectively. Probability distributions are denoted by $\mathbb{P}$. $\mathscr{P}_p(\mathbb{R}^d)$ denotes the set of distributions with finite p^{th} moment over a $\mathbb{R}^d$. $\mathbb{E}$ denotes the expectation. The Wasserstein-2 distance between distributions $\mathbb{P}_1, \mathbb{P}_2 \in \mathbb{R}^d$ is denoted by $\mathsf{W}_2(\mathbb{P}_1, \mathbb{P}_2)$ such that

$$\mathsf{W}_2(\mathbb{P}_1, \mathbb{P}_2) \triangleq \left(\inf\ \mathbb{E}\left[\|\mathbf{w}_1 - \mathbf{w}_2\|^2\right]\right)^{1/2}, \tag{8.41}$$

where the infimum is over all joint distributions of $(\mathbf{w}_1, \mathbf{w}_2)$ with marginals $\mathbf{w}_1 \sim \mathbb{P}_1$ and $\mathbf{w}_2 \sim \mathbb{P}_2$.

8.3 A Primer on Operator Theory and Functional Analysis

Throughout this thesis we fix $p \in [1, \infty]$ and two finite–dimensional real inner–product spaces $\mathbb{V} \cong \mathbb{R}^d$ and $\mathbb{W} \cong \mathbb{R}^n$, each endowed with their Euclidean norm $\|\cdot\|$. Sequences indexed by $\mathbb{Z}, \mathbb{N} := \{0, 1, 2, \dots\}$ taking values in $\mathbb{V}$ are collected in the standard ℓ_p spaces,

$$\ell_p(\mathbb{Z}; \mathbb{V}) \text{ and } \ell_p(\mathbb{N}; \mathbb{V}), \qquad 1 \le p \le \infty,$$

all equipped with their usual norms. Boldface letters (e.g. $\mathbf{v}$) denote sequences, while capital calligraphic letters (e.g. $\mathcal{L}$) denote linear operators acting on such sequences. Given a Banach space X, we write $\mathscr{B}X$ for the space of bounded linear operators on X and $\mathrm{Hom}(\mathbb{V}, \mathbb{W})$ for the space of linear maps from $\mathbb{V}$ to $\mathbb{W}$.

Fundamental shift and projection operators.

- **(Szegő projection)** Szegő: $\ell_p(\mathbb{Z}; \mathbb{V}) \to \ell_p(\mathbb{N}; \mathbb{V})$ denotes the *causal projection* $(\text{Szegő}\,\mathbf{v})_n \coloneqq \mathbf{v}_n\,\mathbf{1}_{\{n\geq 0\}}$. It is a bounded idempotent operator (Szegő^2 = Szegő) for every $p \in (1, \infty)$[180].
- **(Bilateral shift)** The operator $\mathcal{Z} : \ell_p(\mathbb{Z}; \mathbb{V}) \to \ell_p(\mathbb{Z}; \mathbb{V})$ is defined by $(\mathcal{Z}\mathbf{v})_n \coloneqq \mathbf{v}_{n+1}$. For $p = 2$, $\mathcal{Z}$ is a unitary isomorphism and hence an isometry on ℓ_2.
- **(Unilateral shift)** The operator $\mathcal{Z}_+ : \ell_p(\mathbb{N}; \mathbb{V}) \to \ell_p(\mathbb{N}; \mathbb{V})$ acts as $(\mathcal{Z}_+\mathbf{v})_n \coloneqq \mathbf{v}_{n+1}$. Its adjoint on ℓ_2 satisfies $\mathcal{Z}_+^*\mathcal{Z}_+ = \mathcal{I}$ but $\mathcal{Z}_+\mathcal{Z}_+^* \neq \mathcal{I}$; hence $\mathcal{Z}_+$ is an isometry, yet not unitary.

Fourier transforms. Let $\mathbb{T} \coloneqq \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle equipped with normalised Haar measure $\frac{\mathrm{d}\omega}{2\pi}$.

- $\mathcal{F} : \ell_2(\mathbb{Z}; \mathbb{V}) \xrightarrow{\sim} L_2(\mathbb{T}; \mathbb{V}),\ (\mathcal{F}\mathbf{v})(e^{\mathrm{i}\omega}) \coloneqq \sum_{n\in\mathbb{Z}} \mathbf{v}_n e^{-\mathrm{i}\omega n}$.
- $\mathcal{F}_+ : \ell_2(\mathbb{N}; \mathbb{V}) \xrightarrow{\sim} H_2(\mathbb{T}; \mathbb{V})$ is the restriction of $\mathcal{F}$ to non–negative indices, where H_2 denotes the Hardy space of square–integrable analytic functions on $\mathbb{T}$.

Both maps are unitary and extend to ℓ_p–L_p isomorphisms via the Hausdorff–Young inequality.

Laurent operators and transfer matrices. An operator $\mathcal{L} \in \mathscr{B}\ell_2(\mathbb{Z}; \mathbb{V}), \ell_2(\mathbb{Z}; \mathbb{W})$ is *Laurent* (a.k.a. convolution or bi–infinite Toeplitz) if it commutes with the bilateral shift: $\mathcal{L}\mathcal{Z}_{\mathbb{V}} = \mathcal{Z}_{\mathbb{W}}\mathcal{L}$. Equivalently, there exists a bounded measurable function $L : \mathbb{T} \to \mathrm{Hom}(\mathbb{V}, \mathbb{W})$—the *transfer matrix*—such that

$$(\mathcal{F}_{\mathbb{W}}\,\mathcal{L}\,\mathcal{F}_{\mathbb{V}}^* x)(z) = L(z)x(z), \qquad x \in L_2(\mathbb{T}; \mathbb{V}).$$

We denote the corresponding Banach space by

$$\mathscr{L}_p(\mathbb{V}, \mathbb{W}) \coloneqq \big\{\mathcal{L} \text{ Laurent} :\ L \in L_p(\mathbb{T}; \mathrm{Hom}(\mathbb{V}, \mathbb{W}))\big\},$$

endowed with $\|\mathcal{L}\|_{\mathscr{L}_p} := \|L\|_{L_p}$. When $\mathbb{V} = \mathbb{W}$ we abbreviate $\mathscr{L}_p(\mathbb{V}) \equiv \mathscr{L}_p$, and in the scalar case we write $\mathscr{L}_p(\mathbb{R}^{n\times d})$ or, for self–adjoint operators, $\mathscr{L}_p(\mathbb{S}^d)$.

Positivity and order. A self–adjoint $\mathcal{M} \in \mathscr{L}_p$ is *positive* (written $\mathcal{M} \succcurlyeq 0$) if $\langle \mathbf{w}, \mathcal{M}\mathbf{w}\rangle \geq 0$ for all $\mathbf{w} \in \ell_2(\mathbb{Z}; \mathbb{V})$. Equivalently, its transfer matrix $M(z)$ satisfies $M(z) \succcurlyeq 0$ a.e. on $\mathbb{T}$. The cone $\mathscr{L}_p(\mathbb{S}^d_+)$ is closed and convex; the Loewner order is given by $\mathcal{M} \succcurlyeq \mathcal{N} \;\Leftrightarrow\; \mathcal{M} - \mathcal{N} \succcurlyeq 0$.

Hardy sub–algebras. The *causal* Laurent operators are those with analytic transfer matrix: $L \in H_p(\mathbb{T})$; their collection is the sub–algebra $\mathscr{H}_p(\mathbb{V}, \mathbb{W}) \subset \mathscr{L}_p(\mathbb{V}, \mathbb{W})$. The strictly anti–causal operators form $\mathscr{H}_p^-(\mathbb{V}, \mathbb{W})$.

Szegő and Cauchy projections. For $\mathcal{L} \in \mathscr{L}_p$ we write $\{\mathcal{L}\}_+$ (resp. $\{\mathcal{L}\}_-$) for its causal (resp. strictly anti–causal) part obtained by projecting the Fourier series of L onto non–negative (resp. negative) indices. The Szegő projection $\{\mathcal{L}\}_{+\frac{1}{2}}$ further splits the zeroth coefficient $\widehat{L}_0$ symmetrically; its kernel is the classical Hilbert transform operator $\mathcal{H}$. For $1 < p < \infty$ these projections are bounded on $\mathscr{L}_p$ by the Marcinkiewicz–Pichorides theorem[180].

Orthogonal decomposition. In $\mathscr{L}_2$ the causal projection is orthogonal, yielding the direct sum decomposition $\mathscr{L}_2 = \mathscr{H}_2 \oplus \mathscr{H}_2^{\perp}$.

Signal and system norms. For $L \in L_p(\mathbb{T}; \mathbb{V})$ we set

$$\|L\|_p := \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \|L(e^{\mathrm{i}\omega})\|^p \,\mathrm{d}\omega\right)^{1/p}, \quad 1 \leq p < \infty, \qquad \|L\|_\infty := \operatorname*{ess\,sup}_{z\in\mathbb{T}} \|L(z)\|.$$

For $\mathcal{L} \in \mathscr{L}_p(\mathbb{R}^{n\times d})$ we define the (trace) $\mathscr{L}_p$ norm by

$$\|\mathcal{L}\|_p := \big(\operatorname{tr}|\mathcal{L}|^p\big)^{1/p} = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{Tr}|L(e^{\mathrm{i}\omega})|^p \,\mathrm{d}\omega\Big)^{1/p},$$

with the usual modification for $p = \infty$. Duality is given by the bilinear form

$$\langle \mathcal{C}, \mathcal{M}\rangle := \operatorname{tr}\big(\mathcal{C}^*\mathcal{M}\big) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{Tr}\big(C(e^{\mathrm{i}\omega})^* M(e^{\mathrm{i}\omega})\big)\,\mathrm{d}\omega.$$

For $1 < p < \infty$ we have the Banach duality $(\mathscr{L}_p)^* \cong \mathscr{L}_q$ with $1/p + 1/q = 1$, while $\mathscr{L}_1$ and $\mathscr{L}_\infty$ are non–reflexive.

Analytic and Wiener algebras. We write $W(\mathbb{T})$ for the Wiener algebra of absolutely summable Fourier series and $A(\overline{\mathbb{D}})$ ("disk algebra") for functions analytic in the open unit disk $\mathbb{D}$ and continuous on $\mathbb{T}$. Inclusions $W(\mathbb{T}) \subset C(\mathbb{T}) \subset L_\infty(\mathbb{T})$ are standard.

Any causal $G(z) = \sum_{n=0}^{\infty} \widehat{G}_n z^{-n} \in H^{n\times m}(\overline{\mathbb{D}}_{-}^{c})$ has exponentially decaying Markov parameters and admits an exponentially stable (possibly infinite–dimensional) state–space realization; a constructive proof is included verbatim from the original draft .

The sequel relies on a handful of classical facts from topological vector space theory; we record them for completeness.

Definition 8.3.1 (Topological Vector Space)**.** A *topological vector space* (tvs) is a vector space $\mathscr{X}$ over $\mathbb{R}$ or $\mathbb{C}$ endowed with a Hausdorff topology τ making addition $(x, y) \mapsto x + y$ and scalar multiplication $(\lambda, x) \mapsto \lambda x$ jointly continuous.

Definition 8.3.2 (Local convexity, Fréchet, Banach, Hilbert)**.** A tvs is *locally convex* if 0 admits a neighbourhood basis of convex sets. A *Fréchet* space is a complete, metrizable, locally convex tvs. A *Banach* space is a complete normed space, and a *Hilbert* space is a Banach space whose norm arises from an inner product.

Definition 8.3.3 (Weak and weak* topologies)**.** Let $\mathscr{X}$ be a Banach space with dual $\mathscr{X}^*$. The $\sigma(\mathscr{X}, \mathscr{X}^*)$ topology—the coarsest making all elements of $\mathscr{X}^*$ continuous—is called the *weak topology*. On $\mathscr{X}^*$, the $\sigma(\mathscr{X}^*, \mathscr{X})$ topology is the *weak* topology*.

Theorem 8.3.4 (Banach–Alaoglu)**.** *The closed unit ball of $\mathscr{X}^*$ is compact in the weak* topology.*

Theorem 8.3.5 (Sequential Banach–Alaoglu)**.** *For a Banach space $\mathscr{X}$ the following are equivalent:*

i. $\mathscr{X}$ is separable;

ii. the dual unit ball is weak–metrizable;*

iii. the dual unit ball is sequentially compact in the weak topology.*

Theorem 8.3.6 (Mazur)**.** *If $x_n \rightharpoonup x$ weakly in a Banach space, then there exists a sequence of convex combinations $y_n \in \operatorname{conv}\{x_1, \dots, x_n\}$ such that $\|y_n - x\| \to 0$.*

Theorem 8.3.7 (Krein–Smulian – weak* closed convex sets)**.** *Let $\mathscr{X}$ be Banach, $\mathscr{B}^*$ the dual unit ball. A convex $\mathscr{C} \subset \mathscr{X}^*$ is weak* closed iff $\mathscr{C} \cap r\mathscr{B}^*$ is weak* closed for every $r > 0$.*

Theorem 8.3.8 (Krein–Smulian – weak compact convex hulls)**.** *If $\mathscr{C} \subset \mathscr{X}$ is norm (resp. weak) compact, then its norm (resp. weak) closed convex hull is norm (resp. weak) compact.*

Theorem 8.3.9 (Goldstine)**.** *In a Banach space $\mathscr{X}$, the canonical embedding $\mathscr{X} \hookrightarrow \mathscr{X}^{**}$ maps the unit ball densely (for the weak* topology) inside the bidual unit ball.*

Theorem 8.3.10 (Eberlein–Smulian)**.** *In a Banach space, relative compactness, sequential compactness, and countable compactness coincide for subsets endowed with the weak topology.*

Theorem 8.3.11 (Bishop–Phelps)**.** *In a Banach space every continuous linear functional can be approximated in norm by functionals attaining their norm on a given closed, bounded, convex set. Consequently, support functionals are norm–dense in the barrier cone and support points are dense on the boundary.*

Theorem 8.3.12 (James reflexivity criterion)**.** *A Banach space $\mathscr{X}$ is reflexive iff every $x^* \in \mathscr{X}^*$ attains its norm on the closed unit ball of $\mathscr{X}$.*

Theorem 8.3.13 (Brondsted–Rockafellar)**.** *For a proper, convex, lower–semicontinuous $f : \mathscr{X} \to \mathbb{R}_{\text{ext}}$ the points where the subdifferential $\partial f(x)$ is non–empty are dense in* $\operatorname{dom} f$.

Theorem 8.3.14 (Interior of the positive cone)**.** *For $1 \le p < \infty$ the cone $\mathbf{Pos}(\mathscr{L}_p) := \{\mathcal{X} \in \mathscr{L}_p : \mathcal{X} = \mathcal{X}^*,\ \mathcal{X} \succcurlyeq 0\}$ has empty norm–interior, whereas $\mathbf{Pos}(\mathscr{L}_\infty)$ has non–empty interior.*

Definition 8.3.15 (Non–commutative $\mathscr{L}_p$)**.** Let $(\mathcal{M}, \tau)$ be a finite, faithful, normal, tracial von Neumann algebra. For $1 \le p < \infty$ the non–commutative space $L_p(\mathcal{M})$ is the completion of $\mathcal{M}$ under the norm $\|x\|_p := \tau(|x|^p)^{1/p}$; we set $L_\infty(\mathcal{M}) \equiv \mathcal{M}$.

The particular case $\mathcal{M} = L_\infty(\mathbb{T}, M_d(\mathbb{C}))$ is used implicitly when dealing with matrix–valued Laurent operators.

Theorem 8.3.16 (Marcel Riesz)**.** *For $1 < p < \infty$ the Hilbert transform $\mathcal{H} : L_p(\mathbb{T}) \to L_p(\mathbb{T})$ is bounded, and so is the Szegő projection $\{\cdot\}_{+\frac{1}{2}}$.*

Theorem 8.3.17 (Paley–Wiener)**.** *A function* $f \in H_2(\mathbb{T})$ *extends analytically to* $|z| > 1$ *if and only if its Laurent coefficients decay exponentially.*

Theorem 8.3.18 (Douglas lemma)**.** *Given* $A, B \in \mathscr{B}\mathcal{H}$ *between Hilbert spaces, the operator equation* $AA^* \succeq BB^*$ *admits a contraction* C *with* $B = CA$ *if and only if* $\operatorname{ran} B \subseteq \operatorname{ran} A$.

Theorem 8.3.19 (von Neumann bicommutant)**.** *A unital *–subalgebra* $\mathcal{N} \subset \mathscr{B}\mathcal{H}$ *is a von Neumann algebra if and only if* $\mathcal{N} = \mathcal{N}''$.

Theorem 8.3.20 (Non–commutative Jensen inequality)**.** *Let* $f : \mathbb{R} \to \mathbb{R}_{\text{ext}}$ *be convex,* $\mathcal{M} \in \mathscr{L}_1^+$ *with* $\operatorname{tr} \mathcal{M} = 1$, *and* $\mathcal{C} \in \mathscr{L}_\infty$ *such that* $\operatorname{spec} \mathcal{C} \subset \operatorname{dom} f$. *Then* $f(\langle \mathcal{C}, \mathcal{M} \rangle) \le \langle \mathcal{M}, f(\mathcal{C}) \rangle$.

8.4 A Primer on Linear Systems Theory

Consider the following discrete-time linear and time-invariant (LTI) state-space model:

$$\begin{aligned} x_{t+1} &= A\,x_t \;+ B_u u_t \;\; + B_w w_t, \\ y_t &= C_y x_t + D_{yu} u_t + D_{yw} w_t, \\ z_t &= C_z x_t + D_{zu} u_t + D_{zw} w_t. \end{aligned} \tag{8.42}$$

Here, $x_t \in \mathbb{R}^{d_x}$ denotes the latent state, $u_t \in \mathbb{R}^{d_u}$ and $w_t \in \mathbb{R}^{d_w}$ are respectively the control input and exogenous disturbance, and $y_t \in \mathbb{R}^{d_y}$ and $z_t \in \mathbb{R}^{d_z}$ are respectively the observed output and the regulated output at time $t \in \mathbb{Z}$. The matrices (A, B_i, C_o, D_{oi}) are of appropriate dimensions for $i \in \{u, w\}$, $o \in \{y, z\}$.

Input-Output Representation

The state-space representation in (8.42) can be represented more succinctly as linear operators from the space of input sequences to the space of output sequences. Concretely, we introduce the sequences $\mathbf{x} := \{x_t\}_{t\in\mathbb{Z}}$ for the states, $\mathbf{u} := \{u_t\}_{t\in\mathbb{Z}}$ for the control inputs, $\mathbf{w} := \{w_t\}_{t\in\mathbb{Z}}$ for the exogenous disturbances, $\mathbf{y} := \{y_t\}_{t\in\mathbb{Z}}$ for the observed outputs, and $\mathbf{z} := \{z_t\}_{t\in\mathbb{Z}}$ for the regulated outputs. We do not yet impose the exact spaces where these sequences live other than $\mathbb{R}^{\mathbb{Z}}$.

The relationship between the inputs $(\mathbf{u}, \mathbf{w})$ and the outputs $(\mathbf{y}, \mathbf{z})$ is captured by the open-loop plant transfer operator $\mathcal{P}$ as follows:

$$\begin{aligned} \mathcal{P} : \begin{bmatrix} \mathbf{u} \\ \mathbf{w} \end{bmatrix} \mapsto \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} \triangleq \begin{bmatrix} \mathcal{P}_{yu} & \mathcal{P}_{yw} \\ \mathcal{P}_{zu} & \mathcal{P}_{zw} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{w} \end{bmatrix} \\ \text{where} \quad \begin{bmatrix} \mathcal{P}_{yu} & \mathcal{P}_{yw} \\ \mathcal{P}_{zu} & \mathcal{P}_{zw} \end{bmatrix} = \left[\begin{array}{c|cc} A & B_u & B_w \\ \hline C_y & D_{yu} & D_{yw} \\ C_z & D_{zu} & D_{zw} \end{array} \right] \end{aligned} \tag{8.43}$$

Here, for each input $i \in \{u, w\}$ and output $o \in \{y, z\}$, the notation $\mathcal{P}_{oi} : \mathbf{i} \mapsto \mathbf{o}$ represents the transfer operator mapping the input $\mathbf{i}$ to the output $\mathbf{o}$. As the underlying system is causal and time-invariant, $\mathcal{P}_{oi}$ is a causal $(d_o \times d_i)$-block Laurent operator, *i.e.*, $\mathcal{P}_{oi} \in$ Laurent for $i \in \{u, w\}$, $o \in \{y, s\}$. Moreover, the notation in the right-hand side of (8.43) means that the z-domain transfer matrix function corresponding to the transfer operator $\mathcal{P}_{oi}$ is given as:

$$P_{oi}(z) = C_o(zI - A)^{-1}B_i + D_{oi}, \text{ for } i \in \{u, w\},\, o \in \{y, s\} \tag{8.44}$$

Closed-Loop Control

We consider causal LTI controllers that map observed outputs up to time $t \in \mathbb{Z}$ to the current control input u_t via a convolution:

$$u_t \coloneqq \sum_{s \le t} \widehat{K}_{t-s} y_s, \quad \forall t \in \mathbb{Z}, \tag{8.45}$$

where $\{\widehat{K}\}_{t \ge 0} \subseteq \mathbb{R}^{d_u \times d_y}$ are the Markov parameters (aka impulse response) of the controller. Similar to the dynamics of the open-loop plant, this relationship can be captured globally as

$$\mathcal{K} : \mathbf{y} \mapsto \mathbf{u} \coloneqq \mathcal{K}\mathbf{y}, \tag{8.46}$$

where $\mathcal{K}$ is a causal $(d_u \times d_y)$-block Laurent operator, *i.e.*, $\mathcal{K} \in$ Laurent.

Notice that, we avoid making assumptions regarding the internal structure of the class of controllers as opposed to many prior works on controller synthesis. This is rather intentional as several optimal controllers of interest do not possess such a priori internal state-space structures. Therefore, we keep the class of controllers as general as possible by considering causal LTI controllers.

For a feedback loop between the plant $\mathcal{P}$ and a controller $\mathcal{K}$ to be well defined, a controller must ensure invertibility of $I - D_{yu}\widehat{K}_0 \in \mathbb{R}^{d_y \times d_y}$.

Given a feedback systems $(\mathcal{P}, \mathcal{K})$, the closed-loop transfer function $\mathcal{T}_{\mathcal{K}} : \mathbf{w} \mapsto \mathbf{z}$ from disturbances to regulated output is given by

$$\mathcal{T}_{\mathcal{K}} \coloneqq \mathcal{P}_{zu}\mathcal{K}(\mathcal{I} - \mathcal{P}_{yu}\mathcal{K})^{-1}\mathcal{P}_{yw} + \mathcal{P}_{zw}. \tag{8.47}$$

Note that, invertablity of $I - D_{yu}\widehat{K}_0$ ensures invertibility of $\mathcal{I} - \mathcal{P}_{yu}\mathcal{K}$.

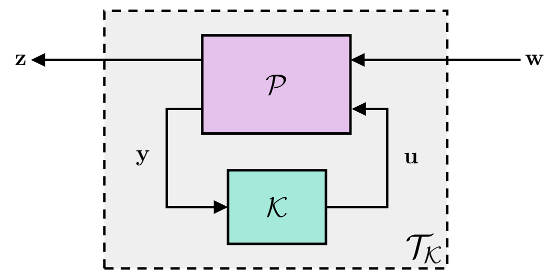


Figure 8.1: Closed-loop Plant-Controller System

Stability and Controller Parametrization

It is crucial that a synthesized controller "stabilizes" a plant in the closed-loop feedback. The notion of stability in dynamical systems is usually understood to be zero-input asymptotic stability of the internal state, also called as *internal stability*. For a meaningful definition of internal stability, a meaningful notion of internal state is needed. However, when an internal structure is not imposed a priori, and the system in consideration is only known by its input-output mapping, then internal stability may not be a well-defined notion. For instance, a stable system known only through by its transfer function might have might have unstable poles which are canceled by zeros. However, in many cases, either the underlying state-space structure is known or a system known only by its transfer function is assumed to be realized by its minimal degree state-space realization if it exists. Rational transfer function are special in the sense that, they always admit a finite order state-space realization, and therefore, one can still talk about internal stability of such input/output systems via its minimal degree realizations.

However, when the transfer function of the input/output system is not a rational function, then the underlying system cannot be realized by a finite degree state-space model. Such systems may still posses an a priori notion of internal state-space structure, although infinite dimensional one, such as PDEs or continuous time-delay systems. If such an a priori notion of state-space is available, it is still possible to talk about internal stability meaningfully, although there might be various notions of internal stability due to various notions convergence in infinite dimensional spaces.

However, unless such a state-space structure is given a priori, internal stability is not a meaningful notion for input/output system with non-rational transfer function.

While we assumed the underlying plant is equipped with a priori state-space structure, an arbitrary causal LTI controller with a non-rational transfer function does not admit such a natural state-space structure, which makes the notion of internal stability an ill-posed notion for the closed-loop system $\mathcal{T}_\mathcal{K}$ as no internal closed-loop state can be identified.

Therefore, we shift our focus to external stability, which is a purely a property of systems as input/outputs maps as opposed to internal stability. A system can b be externally stable if inputs of certain type always lead to outputs of certain type. In this paper, we will mostly be concerned with the bounded energy signals, therefore, we define a system to be stable of ℓ_2-norm bounded signals are always mapped to ℓ_2-norm bounded outputs. For causal LTI systems, ℓ_2/ℓ_2-external stability coincides with bounded $\mathscr{H}_\infty$-norm of the underlying transfer function of the system. We simply call a causal Laurent operator $\mathcal{T}$ stable if $\|\mathcal{T}\|_{\mathscr{H}_\infty} < +\infty$.

Assumption 8.4.1. The subsystem $\mathcal{P}_{yu} : \mathbf{u} \mapsto \mathbf{y}$ is stabilizable and detectable, *i.e.*, (A, B_u) is stabilizable and (A, C_y) is detectable.

Theorem 8.4.2 (Youla Parametrization [251, Lem. 3][213, Thm. 1][67, Ch. 4])**.** *Under Assumption 8.4.1, the following statements hold:*

i. The causal rational transfer operator $\mathcal{P}_{yu}$ admits a doubly-coprime factorization $\mathcal{P}_{yu} = \mathcal{N}_r\mathcal{D}_r^{-1} = \mathcal{D}_l^{-1}\mathcal{N}_l$ *such that*

$$\begin{bmatrix} \mathcal{Y}_l & -\mathcal{X}_l \\ -\mathcal{N}_l & \mathcal{D}_l \end{bmatrix} \begin{bmatrix} \mathcal{D}_r & \mathcal{X}_r \\ \mathcal{N}_r & \mathcal{Y}_r \end{bmatrix} = \begin{bmatrix} \mathcal{I}_{d_u} & \mathbf{0}_{d_u \times d_y} \\ \mathbf{0}_{d_y \times d_u} & \mathcal{I}_{d_y} \end{bmatrix}, \tag{8.48}$$

where $\mathcal{N}_r, \mathcal{N}_l \in \mathscr{RH}_\infty^{d_y \times d_u}$, $\mathcal{X}_r, \mathcal{X}_l \in \mathscr{RH}_\infty^{d_u \times d_y}$, $\mathcal{D}_r, \mathcal{Y}_l \in \mathscr{RH}_\infty^{d_u \times d_u}$, $\mathcal{D}_l, \mathcal{Y}_r \in \mathscr{RH}_\infty^{d_y \times d_y}$ are rational and stable transfer operators.

ii. A causal controller $\mathcal{K}$ is input/output (resp. internally) stabilizing the plant $\mathcal{P}$ if and only if there exists a Youla parameter $\mathcal{Q} \in \mathscr{H}_\infty^{d_u \times d_y}$ *(resp. $\mathscr{RH}_\infty^{d_u \times d_y}$) such that*[5]

$$\mathcal{K} = (\mathcal{X}_r - \mathcal{D}_r\mathcal{Q})(\mathcal{Y}_r - \mathcal{N}_r\mathcal{Q})^{-1}. \tag{8.49}$$

[5] equivalently $\mathcal{K} = (\mathcal{Y}_l - \mathcal{Q}\mathcal{N}_l)^{-1}(\mathcal{X}_l - \mathcal{Q}\mathcal{D}_l)$

iii. The closed-loop transfer operator (8.47) *under an input/output (resp. internally) stabilizing controller can be re-expressed in terms of its Youla parameter* $\mathcal{Q} \in \mathscr{H}_\infty^{d_u \times d_y}$ *(resp.* $\mathscr{R}\mathscr{H}_\infty^{d_u \times d_y}$*) as*

$$\mathcal{T}_\mathcal{Q} = \mathcal{T}_1 \mathcal{Q} \mathcal{T}_2 + \mathcal{T}_3 \in \mathscr{H}_\infty \iff \mathcal{Q} \in \mathscr{H}_\infty \tag{8.50}$$

where $\mathcal{T}_1 \coloneqq -\mathcal{P}_{zu}\mathcal{D}_r \in \mathscr{R}\mathscr{H}_\infty^{d_z \times d_u}$, $\mathcal{T}_2 \coloneqq \mathcal{D}_l \mathcal{P}_{yw} \in \mathscr{R}\mathscr{H}_\infty^{d_y \times d_w}$*, and* $\mathcal{T}_3 \coloneqq \mathcal{P}_{zw} + \mathcal{P}_{zu}\mathcal{X}_r \mathcal{D}_l \mathcal{P}_{yw} \in \mathscr{R}\mathscr{H}_\infty^{d_z \times d_w}$ *are rational and stable transfer operators.*

The theorem is true for any integral domains, including exponentially stable holomorphic functions, disk algebra, $\mathscr{H}_\infty$, $\mathscr{R}\mathscr{H}_\infty$

Chapter 9

INFINITE-HORIZON DISTRIBUTIONALLY ROBUST CONTROL

9.1 Introduction

Addressing uncertainty is a core challenge in decision-making. Control systems inherently encounter various uncertainties, such as external disturbances, measurement errors, model disparities, and temporal variations in dynamics [53], [88]. Neglecting these uncertainties in policy design can result in considerable performance decline and may lead to unsafe and unintended behavior [196].

Traditionally, the challenge of uncertainty in control systems has been predominantly approached through either stochastic or robust control frameworks [51], [121], [255]. Stochastic control (e.g., Linear–Quadratic–Gaussian (LQG) or H_2-control) aims to minimize an expected cost, assuming disturbances follow a known probability distribution [101]. However, in practical scenarios, the true distribution is often estimated from sampled data, introducing vulnerability to inaccurate models. On the other hand, robust control minimizes the worst-case cost across potential disturbance realizations, such as those with bounded energy or power (H_∞ control) [260]. While this ensures robustness, it can be overly conservative. Two recent approaches have emerged to tackle this challenge.

Regret-Optimal (RO) Control. Introduced by [82], [191], this framework offers a promising strategy to tackle both stochastic and adversarial uncertainties. It defines regret as the performance difference between a causal control policy and a clairvoyant, non-causal one with perfect knowledge of future disturbances. In the full-information setting, RO controllers minimize the worst-case regret across all bounded energy disturbances [82], [191]. The infinite-horizon RO controller also takes on a state-space form, making it conducive to efficient real-time computation [191].

Extensions of this framework have been investigated in various settings, including measurement-feedback control [81], [91], dynamic environments [80], safety-critical control [48], [161], filtering [82], [193], and distributed control [162]. While these controllers effectively emulate the performance of non-causal controllers in worst-case disturbance scenarios, they may exhibit excessive conservatism when dealing with stochastic ones.

Distributionally Robust (DR) Control. In contrast to traditional approaches such as H_2 or H_∞ and RO control that focus on a single distribution or worst-case disturbance realization, the DR framework addresses uncertainty in disturbances by considering ambiguity sets – sets of plausible probability distributions [7], [8], [30], [93], [131], [220], [247]. This methodology aims to design controllers with robust performance across all probability distributions within a given ambiguity set. The size of the ambiguity set provides control over the desired robustness against distributional uncertainty, ensuring that the resulting controller is not excessively conservative.

The controller's performance is highly sensitive to the chosen metric for quantifying distributional shifts. Common choices include the total variation (TV) distance [227], [228], the Kullback-Leibler (KL) divergence [60], [149], and the Wasserstein-2 (W_2) distance [7], [30], [89], [123], [219], [220]. The controllers derived from KL-ambiguity sets [60], [178] have been linked to the well-known risk-sensitive controller [115], [216], [237], which minimizes an exponential cost (see [96] and the references therein). However, distributions in a KL-ambiguity set are restricted to be absolutely continuous with respect to the nominal distribution [109], significantly limiting its expressiveness.

In contrast, W_2-distance, which quantifies the minimal cost of transporting mass between two probability distributions, induces a Riemannian structure on the space of distributions [231] and allows for ambiguity sets containing distributions with both discrete and continuous support. Thanks to this versatility and the rich geometric framework, it has found widespread adoption across various fields, including machine learning [9], computer vision [151], [177], estimation and filtering [153], [186], [203], data compression [26], [144], [157], and robust optimization [24], [74], [134], [257]. Moreover, the W_2-distance has emerged as a theoretically appealing statistical distance for DR linear-quadratic control problems [219] due to its compatibility with quadratic objectives and the resulting tractability of the associated optimization problems [74].

Contributions

This paper explores the framework of Wasserstein-2 distributionally robust regret-optimal (W_2-DR-RO) control of linear dynamical systems in the infinite-horizon setting. Initially introduced by [219] for the full-information setting, W_2-DR-RO control was later adapted to the partially observable case by [89]. Similarly, [220] derived a DR controller for the partially observed linear-quadratic-Gaussian (LQG) problem, assuming time-independent disturbances. These prior works, focusing on the finite-horizon setting, are hampered by the requirement to solve a semi-definite program (SDP) whose complexity scales with the time horizon, prohibiting their applicability for large horizons.

Our work addresses this limitation by considering the infinite-horizon setting where the probability distribution of the disturbances over the entire time horizon is assumed to lie in a W_2-ball of a specified radius centered at a given nominal distribution. We seek a linear time-invariant (LTI) controller that minimizes the worst-case expected regret for distributions adversarially chosen within the W_2-ambiguity set. Our contributions are summarized as follows.

1. Stabilizing time-invariant controller. As opposed to the finite-horizon controllers derived in [8], [89], [93], [219], [220], the controllers obtained in the infinite-horizon setting stabilize the underlying dynamics (Corollary 9.3.4)

2. Robustness against non-iid disturbances. In contrast to several prior works that assume time-independence of disturbances [7], [8], [93], [131], [220], [247], [258], our approach does not impose such assumptions, thereby ensuring that the resulting controllers are robust against time-correlated disturbances.

3. Characterization of the optimal controller. We cast the W_2-DR-RO control problem as a max-min optimization and derive the worst-case distribution and the optimal controller using KKT conditions (Theorem 9.3.2). While the resulting controller is non-rational, lacking a finite-order state-space realization (Corollary 13.4.3), we show it admits a finite-dimensional parametric form (Theorem 13.4.2).

4. Efficient computation of the optimal controller. Utilizing the finite-dimensional parametrization, we propose an efficient algorithm based on the Frank-Wolfe method to compute the optimal non-rational W_2-DR-RO controller in the frequency-domain with arbitrary fidelity (Algorithm 4).

5. Near-optimal state-space controller. We introduce a novel convex program that finds the best rational approximation of any given order for the non-rational controller in the $\mathscr{H}_\infty$-norm (Theorem 9.5.5). Therefore, our approach enables efficient real-time implementation using a near-optimal state-space controller (Lemma 9.5.7).

Notations: The letters $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ denote the set of natural numbers, integers, real, and complex numbers, respectively. $\mathbb{T}$ denotes the complex unit circle. For $z \in \mathbb{C}$, $|z|$ is its magnitude, and z^* is the conjugate. $\mathbb{S}_+^n$ denotes the set of positive semidefinite (psd) matrices of size $n \times n$. Bare calligraphic letters ($\mathcal{K}$, $\mathcal{M}$, etc.) are reserved for operators. $\mathcal{I}$ is the identity operator with a suitable block size. For an operator $\mathcal{M}$, its adjoint is $\mathcal{M}^*$. For a matrix A, its transpose is $A^\intercal$, and its Hermitian conjugate is A^*. For psd operators/matrices, $\succcurlyeq$ denotes the Loewner order. For a psd operator $\mathcal{M}$, both $\sqrt{\mathcal{M}}$ and $\mathcal{M}^{\frac{1}{2}}$ denote the PSD square-root. $\{\mathcal{M}\}_+$ and $\{\mathcal{M}\}_-$ denote the causal and strictly anti-causal parts of an operator $\mathcal{M}$. $M(z)$ denotes the z-domain transfer function of a Toeplitz operator $\mathcal{M}$. $\operatorname{tr}(\cdot)$ denotes the trace of operators and matrices. $\|\cdot\|$ is the usual Euclidean norm. $\|\cdot\|_\infty$ and $\|\cdot\|_2$ are the $\mathscr{H}_\infty$ operator) and $\mathscr{H}_2$ (Frobenius) norms, respectively. Probability distributions are denoted by $\mathbb{P}$. $\mathscr{P}_p(\mathbb{R}^d)$ denotes the set of distributions with finite p^{th} moment over a $\mathbb{R}^d$. $\mathbb{E}$ denotes the expectation. The Wasserstein-2 distance between distributions $\mathbb{P}_1, \mathbb{P}_2 \in \mathbb{R}^d$ is denoted by $\mathsf{W}_2(\mathbb{P}_1, \mathbb{P}_2)$ such that

$$\mathsf{W}_2(\mathbb{P}_1, \mathbb{P}_2) \triangleq \left(\inf\, \mathbb{E}\left[\|\mathbf{w}_1 - \mathbf{w}_2\|^2\right]\right)^{1/2}, \tag{9.1}$$

where the infimum is over all joint distributions of $(\mathbf{w}_1, \mathbf{w}_2)$ with marginals $\mathbf{w}_1 \sim \mathbb{P}_1$ and $\mathbf{w}_2 \sim \mathbb{P}_2$.

9.2 Preliminaries

Linear-Quadratic Control

Consider a discrete-time, linear time-invariant (LTI) dynamical system expressed as a state-space model given by:

$$x_{t+1} = Ax_t + B_u u_t + B_w w_t, \quad s_t = Cx_t. \tag{9.2}$$

Here, $x_t \in \mathbb{R}^{d_x}$ is the *state*, $s_t \in \mathbb{R}^{d_s}$ is the *regulated output*, $u_t \in \mathbb{R}^{d_u}$ is the *control input*, and $w_t \in \mathbb{R}^{d_w}$ is the *exogenous disturbance* at time t. The state-space parameters (A, B_u, B_w, C) are known with stabilizable (A, B_u), controllable (A, B_w), and observable (A, C). The disturbances are generated from an unknown stochastic process.

We focus on the infinite-horizon setting, where the time index spans from the infinite past to the infinite future, taking values in $\mathbb{Z}$[1]. Defining the doubly-infinite column vectors of regulated output $\mathbf{s} := (s_t)_{t\in\mathbb{Z}}$, control input $\mathbf{u} := (u_t)_{t\in\mathbb{Z}}$, and disturbance process $\mathbf{w} := (w_t)_{t\in\mathbb{Z}}$ trajectories, we can express the temporal interactions between these variables globally by representing the dynamics (9.2) as a *causal linear input/output model*, described by:

$$\mathbf{s} = \mathcal{F}\mathbf{u} + \mathcal{G}\mathbf{w}, \tag{9.3}$$

[1]The doubly-infinite horizon is chosen for simplicity in derivations, but the results are extendable to a semi-infinite horizon.

where $\mathcal{F}$ and $\mathcal{G}$ are *strictly causal* (*i.e.*, strictly lower-triangular) and doubly-infinite $d_s \times d_u$ and $d_s \times d_w$-block *Toeplitz operators*, respectively. These operators describe the influence of the control input and disturbances on the regulated output through convolution with the *impulse response* of the dynamical system (9.2), which are completely determined by the model parameters (A, B_u, B_w, C).

Control Policy. We restrict our attention to the full-information setting where the control input u_t at time $t \in \mathbb{Z}$ has access to the past disturbances $(w_s)_{s=-\infty}^{t}$. In particular, we consider linear time-invariant (LTI) *disturbance feedback control*[2] (DFC) policies that map the disturbances to the control input via a causal convolution sum:

$$u_t = \sum\nolimits_{s=-\infty}^{t} \widehat{K}_{t-s} w_s, \quad \text{for all} \quad t \in \mathbb{Z}. \tag{9.4}$$

The sequence $\{\widehat{K}_t\}_{t=0}^{\infty}$ of $d_u \times d_w$ matrices are known as the *Markov parameters* of the controller. Similar to the causal linear model in (9.3), the controller equation in (9.4) can be expressed globally by $\mathbf{u} = \mathcal{K}\mathbf{w}$, where $\mathcal{K}$ is a bounded, *strictly causal*, $d_u \times d_w$-block *Toeplitz operator* with lower block-diagonal entries given by the Markov parameters. The set of causal DFC policies is denoted by $\mathscr{K}$.

Cost. At each time step, the control inputs and disturbances incur a quadratic instantaneous cost $s_t^\intercal s_t + u_t^\intercal R u_t$, where $R \succ 0$. Without loss of generality, we take $R = I$ by redefining $B_u R^{-\frac{1}{2}} \to B_u$ and $R^{\frac{1}{2}} u_t \to u_t$. By defining the truncated sequences $\mathbf{s}_T \coloneqq (s_t)_{t=0}^{T-1}$ and $\mathbf{u}_T \coloneqq (u_t)_{t=0}^{T-1}$ the *cumulative cost* over a horizon of $T \in \mathbb{N}$ is simply given by

$$\mathrm{cost}_T(\mathbf{u}, \mathbf{w}) \coloneqq \|\mathbf{s}_T\|^2 + \|\mathbf{u}_T\|^2. \tag{9.5}$$

The Regret-Optimal Control Framework

We aim to design controllers that reduce the regret against the best offline sequence of control inputs selected in hindsight. For a horizon T, the *cumulative regret* is given by

$$\textsc{Regret}_T(\mathbf{u},\mathbf{w}) \coloneqq \mathrm{cost}_T(\mathbf{u},\mathbf{w}) - \min_{\mathbf{u}'_T} \mathrm{cost}_T(\mathbf{u}',\mathbf{w}). \tag{9.6}$$

We highlight that the minimization on the right-hand side is among all control input sequences, including *non-causal* (offline) ones. The regret-optimal (RO) control framework, introduced by [191], aims to craft a causal and time-invariant controller

[2] Youla parametrization enables the conversion between a DFC controller and a state-feedback controller [251].

$\mathcal{K} \in \mathscr{K}$ that minimizes the steady-state worst-case regret across all *bounded energy disturbances*. This can be formally cast as

$$\gamma_{\mathrm{RO}} := \inf_{\mathcal{K}\in\mathscr{K}} \limsup_{T\to\infty} \frac{1}{T} \sup_{\|\mathbf{w}_T\|^2 \le 1} \textsc{Regret}_T(\mathcal{K}\mathbf{w}, \mathbf{w}). \tag{9.7}$$

In the full-information setting, the best sequence of control inputs selected in hindsight is given by $\mathbf{u}_\circ = \mathcal{K}_\circ \mathbf{w}$ where

$$\mathcal{K}_\circ := -(\mathcal{I} + \mathcal{F}^*\mathcal{F})^{-1}\mathcal{F}^*\mathcal{G}, \tag{9.8}$$

is the optimal non-causal policy [101]. Since a non-causal controller lacks physical realization, the optimal RO controller, $\mathcal{K}_{\mathrm{RO}}$ represents the "best" causal policy, attaining performance levels akin to the optimal non-causal policy $\mathcal{K}_\circ$, which enjoys complete access to the disturbance trajectory in advance.

Exploiting the time-invariance of dynamics in (9.2) and the controller $\mathcal{K} \in \mathscr{K}$, Sabag, Goel, Lale, *et al.* [191] demonstrates the equivalence of (9.7) to the following:

$$\inf_{\mathcal{K}\in\mathscr{K}} \sup_{\|\mathbf{w}\|^2 \le 1} \mathbf{w}^*\mathcal{R}_{\mathcal{K}}\mathbf{w} = \inf_{\mathcal{K}\in\mathscr{K}} \|\mathcal{R}_{\mathcal{K}}\|_\infty, \tag{9.9}$$

where $\|\mathbf{w}\|$ is the ℓ_2-norm, and $\mathcal{R}_{\mathcal{K}}$, which we call *the regret operator*, is given as

$$\mathcal{R}_{\mathcal{K}} := (\mathcal{K} - \mathcal{K}_\circ)^*(\mathcal{I} + \mathcal{F}^*\mathcal{F})(\mathcal{K} - \mathcal{K}_\circ). \tag{9.10}$$

The resulting controller closely mirrors the non-causal controller's performance under the worst-case disturbance sequence but may be conservative for stochastic disturbances.

Distributionally Robust Regret-Optimal Control

This paper investigates distributionally robust regret-optimal control, seeking to devise a causal controller minimizing the worst-case expected regret within a Wasserstein-2 (W_2) ambiguity set of disturbance probability distributions. The W_2-ambiguity set $\mathscr{W}_T(\mathbb{P}_\circ, r)$ for horizon T is defined as a W_2-ball of radius of $r_T > 0$ centered at a nominal distribution $\mathbb{P}_{\circ,T} \in \mathscr{P}(\mathbb{R}^{Td_w})$, namely:

$$\mathscr{W}_T(\mathbb{P}_\circ, r_T) := \left\{ \mathbb{P} \in \mathscr{P}(\mathbb{R}^{Td_w}) \,|\, \mathsf{W}_2(\mathbb{P},\, \mathbb{P}_\circ) \le r_T \right\}. \tag{9.11}$$

In contrast to (9.7), which addresses the worst-case regret across all bounded energy disturbances, our focus is on the *worst-case expected regret* across all distributions within the W_2-ambiguity set, as defined by Taha, Yan, and Bitar [219]

$$\mathsf{Reg}_T(\mathcal{K}, r_T) := \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_{\circ,T}, r_T)} \mathbb{E}_{\mathbb{P}_T}\left[\textsc{Regret}_T(\mathcal{K}\mathbf{w}, \mathbf{w})\right],$$

where $\mathbb{E}_{\mathbb{P}_T}$ denotes the expectation such that $\mathbf{w}_T \sim \mathbb{P}_T$. In the infinite-horizon case, this cumulative quantity diverges to infinity. Therefore, we focus on the *steady-state worst-case expected regret*, as defined by [123]:

Definition 9.2.1. The *steady-state worst-case expected regret* suffered by a policy $\mathcal{K} \in \mathscr{K}$ is given by the ergodic limit of the cumulative worst-case expected regret, *i.e.*,

$$\overline{\mathsf{Reg}}_\infty(\mathcal{K}, r) \coloneqq \limsup_{T\to\infty} \frac{1}{T} \, \mathsf{Reg}_T(\mathcal{K}, r_T). \tag{9.12}$$

To ensure the limit in (9.12) is well-defined, the asymptotic behavior of the ambiguity set must be specified. For this purpose, we make the following assumption.

Assumption 9.2.2. The nominal disturbance process $\mathbf{w}_\circ \coloneqq (w_{\circ,t})_{t\in\mathbb{Z}}$ forms a zero-mean weakly stationary random process with an auto-covariance operator $\mathcal{M}_\circ \coloneqq (\widehat{M}_{\circ,t-s})_{t,s\in\mathbb{Z}}$, *i.e.*, $\mathbb{E}_{\mathbb{P}_\circ}[w_{\circ,t} w_{\circ,s}^\intercal] = \widehat{M}_{\circ,t-s}$. Moreover, the size of the ambiguity set for horizon T scales as $r_T \sim r\sqrt{T}$ for a $r > 0$.

The choice of $r_T \propto \sqrt{T}$ aligns with the fact that the W_2-distance between two random vectors of length T, each sampled from two different iid processes, scales proportionally to $\sqrt{T}$.

While the limit (9.12) is well-defined under Assumption 9.2.2, it can still be infinite depending on the chosen controller $\mathcal{K}$. Notably, a finite value for (9.12) implies closed-loop stability. In Problem 9.2.3, we formally state the infinite-horizon Wasserstein-2 W_2-DR-RO problem.

Problem 9.2.3 (Distributionally Robust Regret-Optimal Control). Find a causal LTI controller, $\mathcal{K} \in \mathscr{K}$, that minimizes the steady-state worst-case expected regret (9.12), *i.e.*,

$$\inf_{\mathcal{K}\in\mathscr{K}} \overline{\mathsf{Reg}}_\infty(\mathcal{K}, r) = \inf_{\mathcal{K}\in\mathscr{K}} \limsup_{T\to\infty} \frac{1}{T} \, \mathsf{Reg}_T(\mathcal{K}, r_T). \tag{9.13}$$

In Section 9.3, we provide an equivalent max-min optimization formulation of Problem 9.2.3.

9.3 A Saddle-Point Problem

This section presents a tractable convex reformulation of the infinite-horizon W_2-DR-RO problem. Concretely, Theorem 9.3.1 introduces an equivalent single-variable variational characterization of the steady-state worst-case expected regret

(9.12) incurred by a fixed time-invariant controller. Exploiting this, we show in Theorem 9.3.2 that Problem 9.2.3 reduces to a convex program over positive-definite operators via duality. Moreover, we characterize the optimal controller and the worst-case distribution via KKT conditions. All the proofs of the subsequent theorems are deferred to the Appendix.

Two major challenges are present in solving the Problem 9.2.3: ergodic limit in (9.12) and causality constraint in (9.13). Firstly, the ergodic limit definition of the worst-case expected regret for a fixed policy $\mathcal{K} \in \mathscr{K}$ requires successively solving optimization problems with ever-increasing dimensions. To address this challenge, we leverage the asymptotic convergence properties of Toeplitz matrices and derive an equivalent formulation of (9.12) as an optimization problem over a single decision variable as in Kargin*, Hajar*, Malik*, *et al.* [123]. Similar to the time-domain derivations of $\mathscr{H}_2$ and risk-sensitive controllers in the infinite horizon, the resulting formulations involve the Toeplitz operators $\mathcal{R}_{\mathcal{K}}$. This result is presented formally in the subsequent theorem.

Theorem 9.3.1 (**A Variational Formula for $\overline{\mathsf{Reg}}_\infty$ [123, Thm.5]).** *Under Assumption 9.2.2, the steady-state worst-case expected regret $\overline{\mathsf{Reg}}_\infty(\mathcal{K}, r)$ incurred by a causal policy $\mathcal{K} \in \mathscr{K}$ is equivalent to the following:*

$$\inf_{\gamma \geq 0,\ \gamma\mathcal{I} \succ \mathcal{R}_{\mathcal{K}}} \gamma \operatorname{tr}\left[((\mathcal{I} - \gamma^{-1}\mathcal{R}_{\mathcal{K}})^{-1} - \mathcal{I})\mathcal{M}_\circ\right] + \gamma r^2. \tag{9.14}$$

which takes a finite value whenever $\mathcal{R}_{\mathcal{K}}$ is bounded. Additionally, the worst-case disturbance is obtained from $w_\star := (\mathcal{I} - \gamma_\star^{-1}\mathcal{R}_{\mathcal{K}})^{-1} w_\circ$ where $\gamma_\star$ is the optimal solution of (9.14) *satisfying* $\operatorname{tr}\left[((\mathcal{I} - \gamma_\star^{-1}\mathcal{R}_{\mathcal{K}})^{-1} - \mathcal{I})^2\mathcal{M}_\circ\right] = r^2$.

Notice that the optimization in (9.14) closely mirrors the finite-horizon version presented by Taha, Yan, and Bitar [219, Thm. 2], with the key difference being the substitution of finite-horizon matrices with Toeplitz operators.

The second challenge is addressing the causality constraint on the controller. When the causality assumption on the controller is lifted, the non-causal policy $\mathcal{K}_\circ$ achieves zero worst-case expected regret since $\mathcal{R}_{\mathcal{K}}$ becomes zero and so the worst-case regret by Theorem 9.3.1. While this example illustrates the triviality of non-causal W_2-DR-RO problem, the minimization of worst-case expected regret objective in (9.14) over causal policies is, in general, not a tractable problem.

Leveraging Fenchel duality of the objective in (9.14), we address the causality constraint by reformulating Problem 9.2.3 as a concave-convex saddle-point problem in Theorem 9.3.2 so that the well-known Wiener-Hopf technique [120], [239] can be used to obtain the optimal W_2-DR-RO controller (see Lemma 9.C.2 for details). To this end, let $\Delta^*\Delta = \mathcal{I}+\mathcal{F}^*\mathcal{F}$ be the *canonical spectral factorization*[3], where both Δ and its inverse Δ^{-1} are causal operators. We also introduce the *Bures-Wasserstein* (BW) distance for positive-definite (pd) operators defined as

$$\mathsf{BW}(\mathcal{M}_1,\mathcal{M}_2)^2 := \operatorname{tr}\left[\mathcal{M}_1+\mathcal{M}_2-2(\mathcal{M}_2^{\frac{1}{2}}\mathcal{M}_1\mathcal{M}_2^{\frac{1}{2}})^{\frac{1}{2}}\right].$$

where $\mathcal{M}_1,\mathcal{M}_2 \succ 0$ with finite trace [20].

Theorem 9.3.2 (**A saddle-point problem for W_2-DR-RO**). *Under Assumption 9.2.2, Problem 9.2.3 reduces to a feasible concave-convex saddle-point problem given as*

$$\sup_{\mathcal{M}\succ 0}\inf_{\mathcal{K}\in\mathscr{K}} \operatorname{tr}(\mathcal{R}_{\mathcal{K}}\mathcal{M}) \quad \text{s.t.} \quad \mathsf{BW}(\mathcal{M},\mathcal{M}_\circ)\le r. \tag{9.15}$$

Letting $\mathcal{K}_{\mathscr{H}_2} := \Delta^{-1}\{\Delta\mathcal{K}_\circ\}_+$ be the $\mathscr{H}_2$ controller, the unique saddle point $(\mathcal{K}_\star,\mathcal{M}_\star)$ of (9.15) *satisfies:*

$$\mathcal{K}_\star = \mathcal{K}_{\mathscr{H}_2} + \Delta^{-1}\left\{\{\Delta\mathcal{K}_\circ\}_-\mathcal{L}_\star\right\}_+\mathcal{L}_\star^{-1}, \tag{9.16a}$$

$$\mathcal{M}_\star = (\mathcal{I}-\gamma_\star^{-1}\mathcal{R}_{\mathcal{K}_\star})^{-1}\mathcal{M}_\circ(\mathcal{I}-\gamma_\star^{-1}\mathcal{R}_{\mathcal{K}_\star})^{-1}, \tag{9.16b}$$

where $\mathcal{L}_\star\mathcal{L}_\star^* = \mathcal{M}_\star$ is the canonical spectral factorization with causal and unique[4] *$\mathcal{L}_\star$ and $\mathcal{L}_\star^{-1}$, and $\gamma_\star>0$ uniquely satisfies* $\operatorname{tr}\left[((\mathcal{I}-\gamma_\star^{-1}\mathcal{R}_{\mathcal{K}_\star})^{-1}-\mathcal{I})^2\mathcal{M}_\circ\right]=r^2$.

This result demonstrates that the optimal W_2-DR-RO controller integrates the $\mathscr{H}_2$ controller with an additional correction term that accounts for the time correlations in the worst-case disturbance, $\mathbf{w}_\star$, which are encapsulated by the auto-covariance operator $\mathcal{M}_\star$.

Remark 9.3.3. As $r\to\infty$, the optimal $\gamma_\star$ approaches the lower bound $\gamma_{\mathrm{RO}} = \inf_{\mathcal{K}\in\mathscr{K}}\|\mathcal{R}_{\mathcal{K}}\|_\infty$ and $\mathcal{K}_\star$ recovers the regret-optimal (RO) controller. Conversely, as $r\to 0$, the ambiguity set collapses to the nominal model as $\gamma_\star\to\infty$ and $\mathcal{K}_\star$ recovers the $\mathscr{H}_2$ controller when $\mathcal{M}_\circ=\mathcal{I}$. Thus, adjusting r facilitates the W_2-DR-RO controller to interpolate between the RO and $\mathscr{H}_2$ controllers.

We conclude this section by asserting the closed-loop stability of (9.2) under the optimal W_2-DR-RO controller, $\mathcal{K}_\star$. This stability directly results from the saddle-point problem (9.15) achieving a finite optimal value.

Corollary 9.3.4. *$\mathcal{K}_\star$ stabilizes the closed-loop system.*

[3] Analogues to Cholesky factorization of finite matrices.

[4] See the note in Section 9.B about the uniqueness of $\mathcal{L}$

9.4 An Efficient Algorithm

In this section, we introduce a numerical method to compute the saddle-point $(\mathcal{K}_\star, \mathcal{M}_\star)$ of the max-min problem in (9.15). While both $(\mathcal{K}_\star, \mathcal{M}_\star)$ are non-rational, *i.e.*, do not admit a finite order state-space realization, Theorem 13.4.2 states that $\mathcal{M}_\star$ possesses a *finite-dimensional parametric* form in the frequency domain. Exploiting this fact, we conceive Algorithm 4, a procedure based on the Frank-Wolfe method, to compute the optimal $\mathcal{M}_\star$ in the frequency domain. Furthermore, we devise a novel approach to approximate the non-rational $\mathcal{M}_\star$ in $\mathscr{H}_\infty$-norm by positive rational functions, from which a near-optimal state-space W_2-DR-RO controller can be computed using (9.16a). We leave the discussion on the rational approximation method to Section 9.5.

To enhance the clarity of our approach, we assume for the remainder of this paper that the nominal disturbances are uncorrelated, *i.e.*, $\mathcal{M}_\circ = \mathcal{I}$. Additionally, we utilize the frequency-domain representation of Toeplitz operators as transfer functions, denoting $\mathcal{M}$ as $M(z)$, $\mathcal{L}$ as $L(z)$, $\mathcal{K}$ as $K(z)$, and similarly for other operators, where $z \in \mathbb{C}$.

An Iterative Optimization in the Frequency Domain

Although the problem is concave, its infinite-dimensional nature complicates the direct application of standard optimization tools. To address this challenge, we employ frequency-domain analysis via transfer functions, allowing for the adaptation of standard optimization techniques. Specifically, we utilize a variant of the Frank-Wolfe method [68], [116]. Our approach is versatile and can be extended to other methods, such as projected gradient descent [83] and the fixed-point method in [123]. Furthermore, the convergence of our method to the saddle point $(\mathcal{K}_\star, \mathcal{N}_\star)$ can be demonstrated using standard tools in optimization. Detailed pseudocode is provided in Algorithm 4 in Section 9.E.

Frank-Wolfe: We define the following function and its (Gateaux) gradient [43]:

$$\Phi(\mathcal{M}) \triangleq \inf_{\mathcal{K} \in \mathscr{K}} \operatorname{tr}(\mathcal{R}_{\mathcal{K}} \mathcal{M}) \tag{9.17}$$

$$\nabla \Phi(\mathcal{M}) = \mathcal{L}^{-*} \{\Delta \mathcal{K}_\circ \mathcal{L}\}_-^* \{\Delta \mathcal{K}_\circ \mathcal{L}\}_- \mathcal{L}^{-1}. \tag{9.18}$$

where $\mathcal{L}\mathcal{L}^* = \mathcal{M}$ is the spectral factorization. Rather than directly solving the optimization (9.15), the Frank-Wolfe method solves a linearized subproblem in consecutive steps. Namely, given the k^{th} iterate $\mathcal{M}_k$, the next iterate $\mathcal{M}_{k+1}$ is obtained via

$$\widetilde{\mathcal{M}}_k = \operatorname*{arg\,max}_{\mathcal{M} \succcurlyeq \mathcal{I},\, \mathsf{BW}(\mathcal{M}, \mathcal{I}) \leq r} \operatorname{tr}(\nabla \Phi(\mathcal{M}_k)\, \mathcal{M}) \tag{9.19a}$$

$$\mathcal{M}_{k+1} = (1-\eta_k)\mathcal{M}_k + \eta_k\widetilde{\mathcal{M}}_k, \tag{9.19b}$$

where $\eta_k \in [0,1]$ is a step-size, commonly set to $\eta_k = \frac{2}{k+2}$ [116]. Letting $\mathcal{R}_k :=$ $\nabla\Phi(\mathcal{M}_k)$ be the gradient as in (9.17), Frank-Wolfe updates can be expressed equivalently using spectral densities as:

$$\widetilde{M}_k(z) = (I - \gamma_k^{-1} R_k(z))^{-2} \tag{9.20}$$

$$M_{k+1}(z) = (1-\eta_k)M_k(z) + \eta_k\widetilde{M}_k(z), \quad \forall z \in \mathbb{T} \tag{9.21}$$

where $\gamma_k > 0$ solves $\operatorname{tr}\left[((\mathcal{I} - \gamma_k^{-1}\mathcal{R}_k)^{-1} - \mathcal{I})^2\right] = r^2$. See Section 9.E for a closed-form $R_k(z)$.

Discretization: Instead of the continuous unit circle $\mathbb{T}$, we use its uniform discretization with N points, $\mathbb{T}_N := \{\mathrm{e}^{j2\pi n/N} \mid n = 0, \ldots, N-1\}$. Updating $M_{k+1}(z)$ at a frequency z using the gradient $R_k(z)$ at the same z requires $M_k(z')$ at all frequencies $z' \in \mathbb{T}$ due to spectral factorization. Thus, $M_{k+1}(z)$ depends on $M_k(z')$ across the entire circle. This can be addressed by finer discretization.

Spectral Factorization: For the non-rational spectral densities $M_k(z)$, we can only use an *approximate* factorization [199]. Consequently, we use the DFT-based algorithm from Rino [188], which efficiently factorizes scalar densities (*i.e.*, $d_w = 1$), with errors diminishing rapidly as N increases. Matrix-valued spectral densities can be factorized using various other algorithms [58], [243]. See Section 9.E for a pseudocode.

Bisection: We use bisection method to find the $\gamma_k > 0$ that solves $\operatorname{tr}\left[((\mathcal{I} - \gamma_k^{-1}\mathcal{R}_k)^{-1} - \mathcal{I})^2\right] =$ r^2 in the Frank-Wolfe update (9.20). See Section 9.E for a pseudocode.

Remark 9.4.1. The gradient $R_k(z)$ requires the computation of the finite-dimensional parameter via (13.30), which can be performed using N-point trapezoidal integration. See Section 9.E for details.

We conclude this section with the following convergence result due to [116], [135].

Theorem 9.4.2 (Convergence of $\mathcal{M}_k$). *There exists constants $\delta_N > 0$, depending on discretization N, and $\kappa > 0$, depending only on state-space parameters* (9.2) *and r, such that, for a large enough N, the iterates in* (9.19) *satisfy*

$$\Phi(\mathcal{M}_\star) - \Phi(\mathcal{M}_k) \leq \frac{2\kappa}{k+2}(1+\delta_N). \tag{9.22}$$

9.5 Rational Approximation

The preceding section determined that the optimal solution, denoted as $\mathcal{N}_\star$, is non-rational and lacks a state-space representation. Nevertheless, Algorithm 4 introduced in Section 9.4 can effectively approximate it in the frequency domain. Indeed, after convergence, the algorithm returns the optimal finite parameter, $\Gamma_\star$, which can be used to compute $N_\star(z)$ at *any arbitrary frequency* using Theorem 13.4.2, and thus $K_\star(z)$ (see Algorithm 4 in Section 9.E). However, a state-space controller must be devised for any practical real-time implementation.

This section introduces an efficient method to obtain state-space controllers approximating the non-rational optimal controller. Instead of directly approximating the controller itself, our method involves an initial step of *approximating the power spectrum* $N_\star(z)$ *of the worst-case disturbance* to minimize the $\mathscr{H}_\infty$-norm of the approximation error using positive rational functions. While problems involving rational function approximation generally do not admit a convex formulation, we show in Theorem 9.5.5 that approximating positive power spectra by a ratio of positive fixed order polynomials can be cast as a convex feasibility problem. After finding a rational approximation of $N_\star(z)$, we compute a state-space controller according to (9.16a). For the sake of simplicity, we focus on scalar disturbances, *i.e.*, $d_w = 1$.

State-Space Models from Rational Power Spectra

As established in Theorem 9.3.2, the derivation of a optimal controller $K_\star$ is achieved through the positive operator $\mathcal{N}_\star = \mathcal{L}_\star^* \mathcal{L}_\star$ using the Wiener-Hopf technique. Specifically, we have $\mathcal{K}_\star = \mathcal{K}_{\mathscr{H}_2} + \Delta^{-1} \left\{ \{\Delta \mathcal{K}_\circ\}_- \mathcal{L}_\star \right\}_+ \mathcal{L}_\star^{-1} \mathcal{L}_\star^{-1}$. Since other controllers of interest, including $\mathscr{H}_2$, $\mathscr{H}_\infty$, and RO, can all be formulated this way, we focus on obtaining approximations to positive power spectra.

It is worth noting that a positive and symmetric rational approximation $\widehat{N}(z)$ of order $m \in \mathbb{N}$ can be represented as a ratio $\widehat{N}(z) = P(z)/Q(z)$ of two positive symmetric polynomials $P(z) = p_0 + \sum_{k=1}^m p_k (z^k + z^{-k})$, and $Q(z) = q_0 + \sum_{k=1}^m q_k (z^k + z^{-k})$. When such $P(z), Q(z)$ exist, we can obtain a rational spectral factorization of $\widehat{N}(z)$ by obtaining spectral factorization for $P(z)$, and $Q(z)$.

Finally, we end this section by stating an exact characterization of positive trig. polynomials. While verifying the positivity condition for general functions might pose challenges, the convex cone of positive symmetric trigonometric polynomials, $\mathscr{T}_{m,+}$, possess a characterization through a linear matrix inequality (LMI), as outlined below:

Lemma 9.5.1 (Trace parametrization of $\mathscr{T}_{m,+}$ [55, Thm. 2.3])**.** *For* $k = [-m, m]$*, let* $\mathbf{\Theta}_k \in \mathbb{R}^{(m+1)\times(m+1)}$ *be the primitive Toeplitz matrix with ones on the* k^{th} *diagonal and zeros everywhere else. Then,* $P(z) = p_0 + \sum_{k=1}^{m} p_k(z^k + z^{-k}) > 0$ *if and only if there exists a real positive definite matrix* $\mathbf{P} \in \mathbb{S}_+^{m+1}$ *such that*

$$p_k = \mathrm{tr}(\mathbf{P}\mathbf{\Theta}_k), \;\; k = 0, \ldots, m. \tag{9.23}$$

According to Lemma 9.5.1, any positive trig. polynomial of order at most m can be expressed (non-uniquely) as $P(z) = \sum_{k=-r}^{r} \mathrm{tr}(\mathbf{P}\mathbf{\Theta}_k) z^{-1} = \mathrm{tr}\left(\mathbf{P}\mathbf{\Theta}(z)\right)$. Here, $\mathbf{\Theta}(z) := \sum_{k=-r}^{r} \mathbf{\Theta}_k z^{-1}$.

Rational Approximation using $\mathscr{H}_\infty$-norm

In this context, we present a novel and efficient approach for deriving rational approximations of non-rational power spectra. Our method bears similarities to the flexible uniform rational approximation approach described in [207], which approximates a function with a rational form while imposing the positivity of the denominator of the rational form as a constraint. Our method uses $\mathscr{H}_\infty$-norm as criteria to address the approximation error effectively. First, consider the following problem:

Problem 9.5.2 (Rational approximation via H_∞-norm minimization)**.** Given a positive spectrum $\mathcal{N}$, find the best rational approximation of order at most $m \in \mathbb{N}$ with respect to H_∞ norm, *i.e.*,

$$\inf_{\mathcal{P}, \mathcal{Q} \in \mathscr{T}_{m,+}} \|\mathcal{P}/\mathcal{Q} - \mathcal{N}\|_\infty \text{ s.t. } \mathrm{tr}(\mathcal{Q}) = 1 \tag{9.24}$$

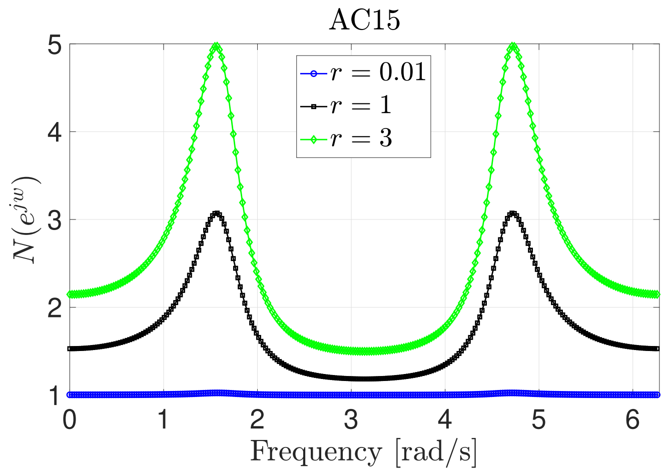
Note that the constraint $\mathrm{tr}(\mathcal{Q}) = 1$, equivalent to $q_0 = 1$, eliminates redundancy in the problem since the fraction $\mathcal{P}/\mathcal{Q}$ is scale invariant.

While the objective function in Equation (9.24) is convex with respect to $\mathcal{P}$ and $\mathcal{Q}$ individually, *it is not jointly convex in* $(\mathcal{P}, \mathcal{Q})$. In this form, Problem 9.5.2 is not amenable to standard convex optimization tools.

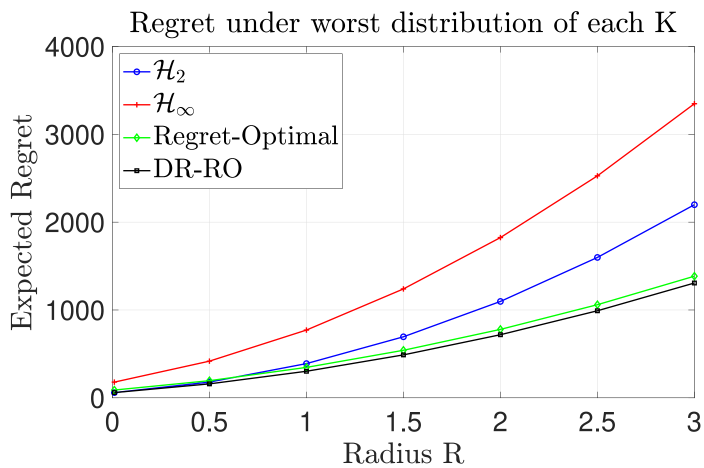
To circumvent this issue, we instead consider the sublevel sets of the objective function in Equation (9.24).

Definition 9.5.3. For a given $\epsilon > 0$ approximation bound, the ϵ-sublevel set of Problem 9.5.2 is defined as

$$\mathscr{S}_\epsilon := \left\{(\mathcal{P}, \mathcal{Q}) \mid \|\mathcal{P}/\mathcal{Q} - \mathcal{N}\|_\infty \leq \epsilon, \; \mathrm{tr}(\mathcal{Q}) = 1\right\}.$$



(a) The frequency domain representation of $\mathcal{N}$ for $r = 0.01, 1, 3$ for system [AC15].



(b) The worst-case expected regret of different controllers for the system [AC15].

Figure 9.1: Variation of $\mathcal{N}$ with r and the performance of the W$_2$-DR-RO controller versus the H_2, H_∞, and RO controller.

By applying the definition of $\mathscr{H}_\infty$-norm, we have that

$$\begin{aligned} &\|\mathcal{P}/\mathcal{Q}-\mathcal{N}\|_\infty = \max_{z\in\mathbb{T}} |P(z)/Q(z) - N(z)| \leq \epsilon \\ \iff &\begin{cases} P(z)-(N(z)+\epsilon)\,Q(z) \leq 0, \\ P(z)-(N(z)-\epsilon)\,Q(z) \geq 0, \end{cases} \end{aligned} \tag{9.25}$$

where the last set of inequalities hold for all $z \in \mathbb{T}$. Notice that the inequalities in Equation (9.25) and the positivity constraints on $\mathcal{P}$, $\mathcal{Q}$ are jointly affine in $(\mathcal{P}, \mathcal{Q})$. Moreover, the equation $\operatorname{tr}(\mathcal{Q}) = 1$ is an affine equality constraint. Therefore, we have the following claim.

Lemma 9.5.4. *The set $\mathscr{S}_\epsilon$ is jointly convex in* $(\mathcal{P}, \mathcal{Q})$.

Unlike its non-convex optimization counterpart in Problem 9.5.2, a membership oracle for the convex set $\mathscr{S}_\epsilon$ offers a means to obtain accurate rational approximations for non-rational functions. According to Lemma 9.5.1, the positive trig. polynomials $(\mathcal{P}, \mathcal{Q}) \in \mathscr{S}_\epsilon$ can be parameterized by psd matrices $\mathbf{P}$ and $\mathbf{Q}$. This allows the equality constraint $\mathrm{tr}(\mathcal{Q})$ and the affine inequalities in (9.25) to be expressed as Linear Matrix Inequalities (LMIs) in terms of $\mathbf{P}$ and $\mathbf{Q}$. The resulting theorem characterizes the ϵ-sublevel sets.

Theorem 9.5.5 (Feasibility of $\mathscr{S}_\epsilon$)**.** *Let $\epsilon > 0$ be a given accuracy level, and $m \in \mathbb{N}$ is a fixed order. The trig. polynomials $\mathcal{P}$ and $\mathcal{Q}$ of order m belong to the ϵ-sublevel set, $(\mathcal{P}, \mathcal{Q}) \in \mathscr{S}_\epsilon$ if and only if there exists $\mathbf{P}, \mathbf{Q} \in \mathbb{S}_+^{m+1}$ such that $\mathrm{tr}\,(\mathbf{Q}) = 1$ and for all $z \in \mathbb{T}$,*

$$1)\ \mathrm{tr}\,(\mathbf{P}\boldsymbol{\Theta}(z)) - (N(z)+\epsilon)\,\mathrm{tr}\,(\mathbf{Q}\boldsymbol{\Theta}(z)) \leq 0, \tag{9.26}$$

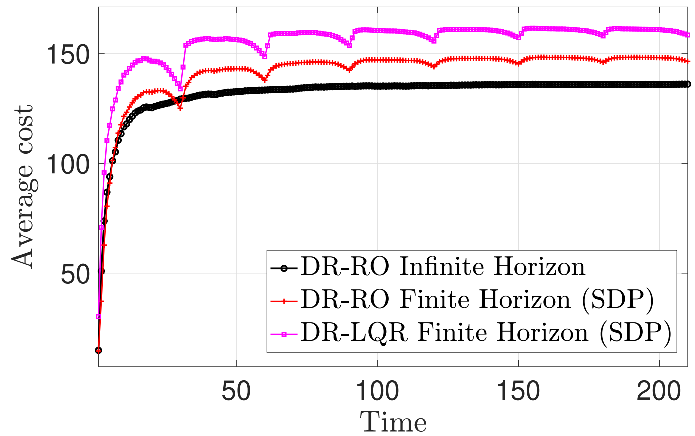
$$2)\ \mathrm{tr}\,(\mathbf{P}\boldsymbol{\Theta}(z)) - (N(z)-\epsilon)\,\mathrm{tr}\,(\mathbf{Q}\boldsymbol{\Theta}(z)) \geq 0. \tag{9.27}$$

The sole limitation in this approach arises from the fact that for a non-rational $N(z)$, the set of infinitely many inequalities in (9.25) cannot be precisely characterized by a finite number of constraints, as seen in the trace parametrization of positive polynomials. To overcome this challenge, one can address the inequalities in (9.25) solely for a finite set of frequencies, such as $\mathbb{T}_N = \{\mathrm{e}^{j2\pi n/N} \mid n = 0, \ldots, N-1\}$ for $N \gg m$. While this introduces an approximation, the method's accuracy can be enhanced arbitrarily by increasing the frequency samples. By taking this approach, the problem of rational function approximation can be reformulated as a convex feasibility problem involving LMIs and a finite number of affine (in)equality constraints.

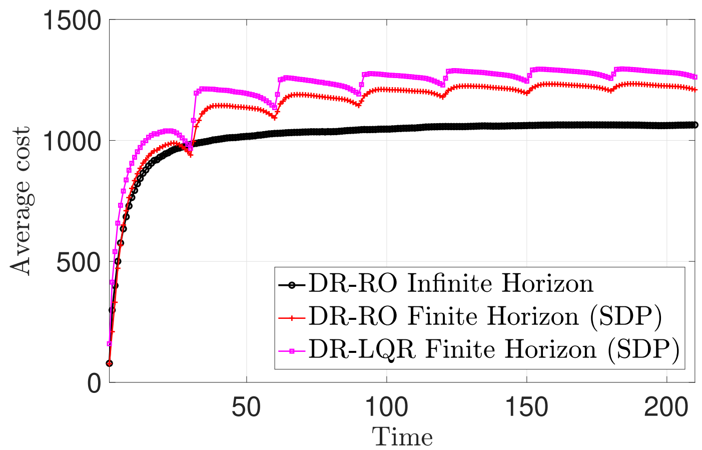
It is crucial to note that our algorithm can be used in the following two operational modes. These modes highlight the algorithm's adaptability for the given two use cases.

1. **Best Precision for a given degree** By adjusting the parameter ϵ, which signifies our tolerance for deviations from $M(e^{jw})$, we can refine the approximation's accuracy. This method is particularly valuable when finding the best possible polynomial representation of $M(e^{jw})$ for a given degree.

2. **Lowest Degree for a given precision** In contrast, we can ask for the lowest degree polynomial, which achieves a certain precision level ϵ. This mode is



(a) White noise



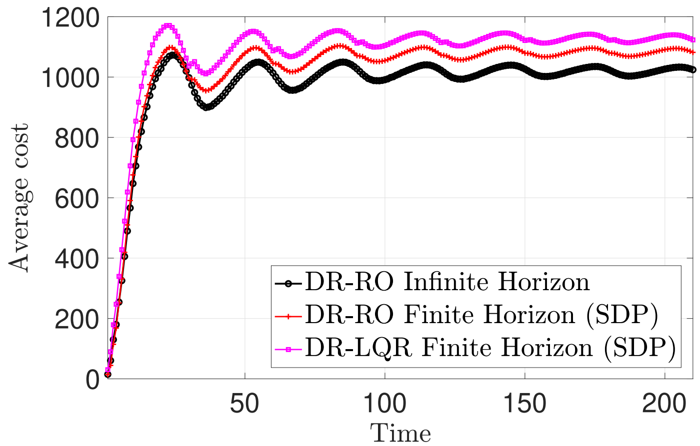
(b) Worst disturbance for W_2-DR-RO, infinite horizon

Figure 9.2: The control costs of different DR controllers under (a) white noise and (b) worst disturbance for W_2-DR-RO in infinite horizon, for system [AC15]. The finite-horizon controllers are re-applied every $s = 30$ steps. The infinite horizon W_2-DR-RO controller achieves the lowest average cost compared to the finite-horizon controllers.

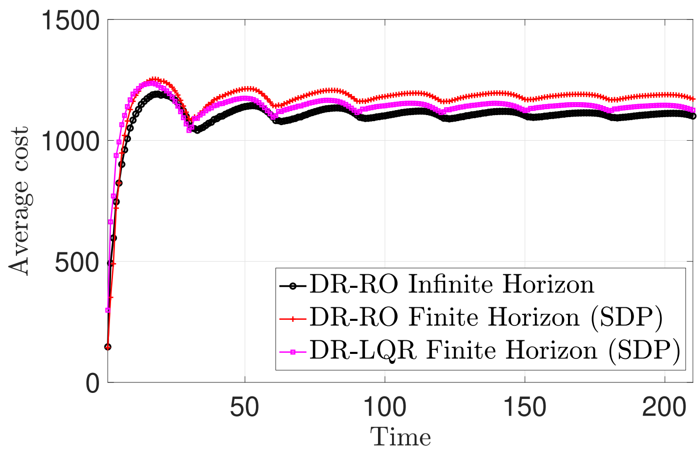
advantageous when the priority is to minimize computational overhead or when we need a simpler polynomial approximation, as long as the approximation remains within acceptable accuracy bounds

Obtaining State-Space Controllers

Note that given the polynomial z-spectra, we require its spectral factorization to obtain the state-space controller that approximates the W_2-DR-RO controller. The following Lemma introduces a simple way to obtain such an approximation



(a) Worst disturbance for W_2-DR-RO, finite horizon



(b) Worst disturbance for DR-LQR, finite horizon

Figure 9.3: The control costs of different DR controllers under (a) worst disturbances for W_2-DR-RO in finite horizon and (b) worst disturbances for DR-LQR in finite horizon, for system [AC15]. The finite-horizon controllers are re-applied every $s = 30$ steps. Despite being designed to minimize the cost under specific disturbances, the finite horizon DR controllers are outperformed by the infinite horizon W_2-DR-RO controller.

Lemma 9.5.6 (Canonical factor of polynomial z-spectra [199, Lem. 1])**.** *Consider a Laurent polynomial of degree* m*,* $P(z) = \sum_{k=-m}^{m} p_k z^{-k}$*, with* $p_k = p_{-k} \in \mathbb{R}$*, such that* $P(z) > 0$*. Then, there exists a canonical factor* $L(z) = \sum_{k=0}^{m} \ell_k z^{-k}$ *such that* $P(z) = |L(z)|^2$ *and* $L(z)$ *has all of its root in* $\mathbb{T}$*.*

Using Lemma 9.5.6, we can compute spectral factors by factorizing the symmetric positive polynomials and multiplying all the factors with stable roots together. Consequently, this rational spectral factor enables the derivation of a rational controller, denoted as $K(z)$ (refer to Section 9.5).

Now we present the W_2-DR-RO controller in state-space form.

Lemma 9.5.7. *Let $\tilde{L}(z)$ be the rational factor of the spectral factorization $\tilde{N}(z) = \tilde{L}(z)^* \tilde{L}(z) = P(z)/Q(z)$ of a degree m rational approximation $P(z)/Q(z)$. The controller obtained from $\tilde{L}(z)$ using* (9.16)*, i.e., $K(z) = K_{\mathscr{H}_2}(z) + \Delta(z)^{-1} \left\{ \{\Delta(z) K_\circ(z)\}_- \tilde{L}(z) \right\}_+ \tilde{L}(z)^{-1}$ is rational and can be realized as a state-space controller as follows:*

$$e(t+1) = \widetilde{F} e(t) + \widetilde{G} w(t), \quad u(t) = \widetilde{H} e(t) + \widetilde{J} w(t)) \tag{9.28}$$

where e_t is the controller state, and $(\widetilde{F}, \widetilde{G}, \widetilde{H}, \widetilde{J})$ are determined from (A, B_u, B_w) and $\tilde{L}(z)$.

9.6 Numerical Experiments

In this section, we present the performance of the W_2-DR-RO controller, compared to H_2, H_∞, regret-optimal and other finite-horizon DR controllers. We present frequency domain and time-domain evaluations, and we showcase the performance of the rational approximation method. We employ benchmark models such as [REA4], [AC15], and [HE3] from [145]. In the frequency domain simulations, results for [REA4] and [HE3] are presented. In the time domain simulations for the aircraft model [AC15] are presented, with additional simulations provided in Appendix 9.H. The [REA4] is a chemical reactor model and [HE3] is a helicopter model with 8 states each. The [AC15] is an aircraft model with 4 states. We perform all experiments using MATLAB, on an Apple M1 processor with 8 GB of RAM. We specify the nominal distribution as a Gaussian, with zero mean and identity covariance.

Frequency Domain Evaluations

We investigate the behaviour of the W_2-DR-RO controller and its rational approximation for various values of the radius r.

To show the behavior of the worst-case disturbance we plot its power spectrum $N(e^{j\omega})$ for three different values of the radius r for the [AC15] system in Figure 9.1a. As can be seen for $r = 0.01$, the worst-case disturbance is almost white, since that is the case for the nominal disturbance. As r increases, the time correlation of the worst-case disturbance increases, and the power spectrum becomes peaky.

For the [AC15] system, the worst-case expected regret cost, as outlined in (9.2.1), for W_2-DR-RO, the H_2, H_∞, and RO controllers. are depicted in Figure 9.1b. We observe that for smaller r, the W_2-DR-RO performs close to the H_2 controller.

However, as r increases, the worst-case regret is close to the regret achieved by the RO controller. Throughout the variation in r, the W_2-DR-RO achieves the lowest worst-case expected regret among all the other mentioned controllers.

To implement the W_2-DR-RO controller in practice, we need a rational controller. We find the rational approximation of $N(\mathrm{e}^{\jmath\omega})$ as $\frac{P(\mathrm{e}^{\jmath\omega})}{Q(\mathrm{e}^{\jmath\omega})}$ using the method of Section 9.5 for [AC15] and degrees $m = 1, 2, 3$. The performance of the resulting rational controllers is compared to the non-rational W_2-DR-RO in Table 9.1. As can be seen, the rational approximation with an order greater than 2 achieves an expected regret that well matches that of the non-rational for all values of r.

	r=0.01	**r=1**	**r=1.5**	**r=2**	**r=3**
$DRRO$	59.16	302.08	488.57	718.20	1307.12
RA(1)	60.49	33394.74	4475.70	9351.89	2376.77
RA(2)	59.58	303.33	491.75	723.96	1318.98
RA(3)	59.57	302.41	489.49	719.72	1309.85

Table 9.1: The worst-case expected regret of the non-rational W_2-DR-RO controller, compared to the rational controllers RA(1), RA(2), and RA(3), obtained from degree 1, 2, and 3 rational approximations to $N(e^{j\omega})$.

Time Domain Evaluations

We compare the time-domain performance of the infinite horizon W_2-DR-RO controller to its finite horizon counterparts, namely W_2-DR-RO and DR-LQR, as outlined in [219]. The latter controllers are computed through an SDP whose dimension scales with the time horizon. We plot the average LQR cost over 210 time steps, aggregated over 1000 independent trials. Figure 9.2a illustrates the performance of DR controllers under white Gaussian noise, while 9.2b, 9.3a, and 9.3b demonstrate responses to worst-case noise scenarios dictated by each of the controllers, using $r = 1.5$. For computational efficiency, the finite horizon controllers operate over a horizon of only $s = 30$ steps and are re-applied every s steps. Their worst-case disturbances in 9.3a and 9.3b are also generated every s steps, resulting in correlated disturbances only within each s steps. Our findings highlight the infinite horizon W_2-DR-RO controller's superior performance over all four scenarios. Note that extending the horizon of the SDP for longer horizons to come closer to the infinite horizon performance is extremely computationally inefficient. These underscore the advantages of using the infnite horizon W_2-DR-RO controller.

9.7 Future Work

Our work presents a complete framework for solving the DR control problem in the full-information setting. Future generalizations would address our limitations. One is to extend the rational approximation method from single to multi input systems. Another is to extend the results to partially observable systems where the state is not directly accessible. Finally, it would be useful to incorporate adaptation as the controller learns disturbance statistics through observations.

9.A Organization of the Appendix

This appendix is organized into several sections:

First, Section 9.B provides notations, definitions, and remarks about the problem formulation and uniqueness of the spectral factorization.

Next, Section 9.C contain proofs of the duality and optimality theorems in Section 9.3.

Subsequently, Section 9.D is dedicated to proofs of lemmas and theorems related to the efficient algorithm discussed in Section 9.4, and Section 9.E describes the pseudo-code of the algorithm.

Further, Section 9.F contains the proof of the state-space representation of the controller presented in Section 9.5.

Finally, additional simulation results are presented in Section 9.H.

9.B Notations, Definitions and Remarks

Notations

In the paper, we use the notations in Table 9.2 for brevity.

Explicit Form of Finite-Horizon State Space Model

Consider the restrictions of the infinite-horizon dynamics in (9.3) to the finite horizon as

$$\mathbf{s}_T = \mathcal{F}_T \mathbf{u}_T + \mathcal{G}_T \mathbf{w}_T. \tag{9.29}$$

Symbol	Description
x_t	State at time t
s_t	Regulated output at time t
u_t	Control input at time t
w_t	Exogenous disturbance at time t
A	State transition matrix
B_u	Control input matrix
B_w	Disturbance input matrix
C	Regulated output matrix
R	Control input cost matrix
$\mathcal{F}_T$	Finite-horizon operator for control input
$\mathcal{G}_T$	Finite-horizon operator for disturbance
$\mathcal{F}$	Infinite-horizon operator for control input
$\mathcal{G}$	Infinite-horizon operator for disturbance
$\\|\cdot\\|$	Euclidean norm
$\\|\cdot\\|_2$	$\mathscr{H}_2$ (Frobenius) norm
$\\|\cdot\\|_\infty$	H_∞ (operator) norm
$\mathbb{E}$	Expectation
$\mathscr{K}$	Set of causal (online) and time-invariant DFC policies
$\mathscr{K}_T$	Set of causal DFC policies over a horizon T
$\mathcal{R}_{\mathcal{K}}$	Regret operator
$\mathcal{M}$	Auto-covariance operator for disturbances
$\mathcal{L}$	Unique, causal and causally invertible spectral factor of $\mathcal{M} = \mathcal{L}\mathcal{L}^*$
$\mathcal{N}$	The unique positive definite operator equal to $\mathcal{L}^*\mathcal{L}$
W_2	Wasserstein-2 metric
$\mathcal{S}$	Symmetric positive polynomial matrix
$\mathrm{tr}(\cdot)$	Trace of a Toeplitz operator
$\{\cdot\}_+$	Causal part of an operator
$\{\cdot\}_-$	Strictly anti-causal part of an operator
$\sqrt{\mathcal{M}}$, or $\mathcal{M}^{\frac{1}{2}}$	Symmetric positive square root of an operator or matrix
$\mathbb{S}_n^+$	The set of positive semidefinite matrices
$\mathscr{T}_{m,+}$	The set of positive trigonometric polynomials of degree m

Table 9.2: Notation Table

The causal linear measurement model for the finite-horizon case in (9.29) can be stated explicitly as follows:

$$\underbrace{\begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_T \end{bmatrix}}_{\mathbf{s}_T} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ B_u & 0 & 0 & \dots & 0 \\ AB_u & B_u & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A^{T-1}B_u & A^{T-2}B_u & A^{T-3}B_u & \ddots & 0 \end{bmatrix}}_{\mathcal{F}_T} \underbrace{\begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_T \end{bmatrix}}_{\mathbf{u}_T} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ B_w & 0 & 0 & \dots & 0 \\ AB_w & B_w & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A^{T-1}B_w & A^{T-2}B_w & A^{T-3}B_w & \ddots & 0 \end{bmatrix}}_{\mathcal{G}_T} \underbrace{\begin{bmatrix} u \\ u \\ u \\ \vdots \\ w \end{bmatrix}}_{\mathbf{w}} \tag{9.30}$$

A Note about Robustness to Disturbances vs Robustness to Model Uncertainties

In our approach, we consider distributional robustness against disturbances, which provides flexibility, adaptability, and dynamic responses to unforeseen events. Although we do not explicitly address model uncertainties, these uncertainties can be effectively lumped together as disturbances—a technique known as uncertainty/disturbance lumping. This approach is particularly effective when the model uncertainties are relatively small. By treating parameter uncertainties as disturbances, we simplify system analysis and ensure that the controller is robust not only to known uncertainties but also to unexpected variations and modeling errors.

A note about the Uniqueness of the spectral factor $\mathcal{L}$

In Theorem 9.3.2, given that $\mathcal{L}$ is the causal and causally invertible spectral factor of $\mathcal{M} = \mathcal{L}\mathcal{L}^*$, it is unique up to a unitary transformation of its block-elements from the right. Fixing the choice of the unitary transformation in the spectral factorization (eg. positive-definite factors at infinity [59]) results in a unique $\mathcal{L}$.

9.C Proof of Optimality Theorems

Proof of Theorem 9.3.1

This result is proven in detail in Kargin*, Hajar*, Malik*, *et al.* [123, Appendix A1]. Due to its length, we provide only a brief sketch here. Interested readers can refer to Kargin*, Hajar*, Malik*, *et al.* [123] for the complete proof. For completeness, we provide the following proof sketch.

First, we provide a finite-horizon counterpart of the strong duality result from [219]. Then, we reformulate the objective functions of both the finite-horizon and infinite-horizon dual problems using normalized spectral measures. We demonstrate the pointwise convergence of the finite-horizon dual objective function to the infinite-horizon objective by analyzing the limiting behavior of the spectrum of Toeplitz matrices. Finally, we show that the infinite-horizon dual problem attains a finite value and that the limits of the optimal values (and solutions) of the finite-horizon dual problem coincide with those of the infinite-horizon dual problem.

Proof of Theorem 9.3.2

The proof involves four main steps.

- **Reformulation using Lemma 9.C.1**: Using Lemma 9.C.1, we reformulate the original optimization problem. This lemma allows the expression of the convex mapping $\mathcal{X} \mapsto \operatorname{tr}(\mathcal{X}^{-1})$ using Fenchel duality, which transforms the objective function into a form that involves the supremum over a positive semi-definite matrix $\mathcal{M}$ and the only term depending on $\mathcal{K}$ remains $\operatorname{tr}(\mathcal{R}_{\mathcal{K}}\mathcal{M})$.

- **Application of Wiener-Hopf Technique**: We then Lemma 9.C.2, which provides a method to approximate a non-causal controller by a causal one, minimizing the cost $\operatorname{tr}(\mathcal{R}_{\mathcal{K}}\mathcal{M})$. The optimal causal controller $\mathcal{K}_\star$ is derived using the Weiner-Hopf Technique.

- **Karush-Kuhn-Tucker (KKT) Conditions**: We then find the conditions on the optimal $\mathcal{M}$. This involves simplifying the objective function and finding the optimal $\mathcal{M}_\gamma$ and $\mathcal{K}_\gamma$ for the level γ.

- **Final Reformulation and Duality**: We further simplify the problem and apply strong duality to achieve the final form. The optimal $\mathcal{K}_\star$ is then derived from the Wiener-Hopf technique, with $\gamma_\star$ and $\mathcal{M}_\star$ obtained through duality arguments.

Before proceeding with the proof, we first state two useful lemmas

Lemma 9.C.1. *The convex mapping $\mathcal{X} \mapsto \operatorname{tr} \mathcal{X}^{-1}$ for $\mathcal{X} \succ 0$ can be expressed via Fenchel duality as*

$$\sup_{\mathcal{M} \succ 0} -\operatorname{tr}(\mathcal{X}\mathcal{M}) + 2\operatorname{tr}(\sqrt{\mathcal{M}}) = \begin{cases} \operatorname{tr}(\mathcal{X}^{-1}), & \textit{if } \mathcal{X} \succ 0 \\ +\infty, & \textit{o.w.} \end{cases} \tag{9.31}$$

Proof. Observe that the objective $-\operatorname{tr}(\mathcal{X}\mathcal{M}) + 2\operatorname{tr}(\sqrt{\mathcal{M}})$ is concave in $\mathcal{M}$, and the expression on the right-hand side can be obtained by solving for $\mathcal{M}$. When $\mathcal{X} \not\succeq 0$, *i.e.*, $\mathcal{X}$ may have negative eigenvalues, then the expression $\operatorname{tr}(\mathcal{X}\mathcal{M})$ can be made arbitrarily negative, and $\operatorname{tr}(\sqrt{\mathcal{M}})$ arbitrarily large, by chosing an appropriate $\mathcal{M}$. ∎

The following lemma will be useful in the proof of Theorem 9.3.2.

Lemma 9.C.2 (Wiener-Hopf Technique [120])**.** *Consider the problem of approximating a non causal controller $\mathcal{K}_\circ$ by a causal controller $\mathcal{K}$, such that $\mathcal{K}$ minimises the cost $\operatorname{tr}(\mathcal{R}_{\mathcal{K}}\mathcal{M})$, i.e.,*

$$\inf_{\mathcal{K} \in \mathscr{K}} \operatorname{tr}(\mathcal{R}_{\mathcal{K}}\mathcal{M}) \tag{9.32}$$

where $\mathcal{M} \succ 0$, $\mathcal{R}_{\mathcal{K}} = (\mathcal{K} - \mathcal{K}_\circ)^* \Delta^* \Delta (\mathcal{K} - \mathcal{K}_\circ)$ *and* $\mathcal{K}_\circ$ *is the non-causal controller that makes the objective defined above zero. Then, the solution to this problem is given by*

$$\mathcal{K}_\star = \Delta^{-1} \{\Delta \mathcal{K}_\circ \mathcal{L}\}_+ \mathcal{L}^{-1}, \tag{9.33}$$

where $\mathcal{L}$ *is the unique causal and causally invertible spectral factor of* $\mathcal{M}$ *such that* $\mathcal{M} = \mathcal{L}\mathcal{L}^*$ *and* $\{\cdot\}_+$ *denotes the causal part of an operator. Alternatively, the controller can be written as,*

$$\mathcal{K}_\star = \mathcal{K}_{\mathscr{H}_2} + \Delta^{-1} \{\{\Delta \mathcal{K}_\circ\}_- \mathcal{L}_\star\}_+ , \tag{9.34}$$

where $\mathcal{K}_{\mathscr{H}_2} := \Delta^{-1}\{\Delta \mathcal{K}_\circ\}_+$.

Proof. Let $\mathcal{L}$ be the unique causal and causally invertible spectral factor of $\mathcal{M}$, *i.e.* $\mathcal{M} = \mathcal{L}\mathcal{L}^*$. Then, using the cyclic property of tr, the objective can be written as,

$$\inf_{\mathcal{K} \in \mathscr{K}} \mathrm{tr}(\Delta (\mathcal{K} - \mathcal{K}_\circ) \mathcal{M} (\mathcal{K} - \mathcal{K}_\circ)^* \Delta^*) = \inf_{\mathcal{K} \in \mathscr{K}} \mathrm{tr}((\Delta\mathcal{K} - \Delta\mathcal{K}_\circ) \mathcal{L}\mathcal{L}^* (\Delta\mathcal{K} - \Delta\mathcal{K}_\circ)^*) \tag{9.35}$$

$$= \inf_{\mathcal{K} \in \mathscr{K}} \mathrm{tr}((\Delta\mathcal{K}\mathcal{L} - \Delta\mathcal{K}_\circ\mathcal{L}) (\Delta\mathcal{K}\mathcal{L} - \Delta\mathcal{K}_\circ\mathcal{L})^*) \tag{9.36}$$

$$= \inf_{\mathcal{K} \in \mathscr{K}} \|\Delta\mathcal{K}\mathcal{L} - \Delta\mathcal{K}_\circ\mathcal{L}\|_2^2 . \tag{9.37}$$

Since $\Delta, \mathcal{K}$ and $\mathcal{L}$ are causal, and $\Delta\mathcal{K}_\circ\mathcal{L}$ can be broken into causal and non-causal parts, it is evident that the controller that minimizes the objective is the one that makes the term $\Delta\mathcal{K}\mathcal{L} - \Delta\mathcal{K}_\circ\mathcal{L}$ strictly anti-causal, cancelling off the causal part of $\Delta\mathcal{K}_\circ\mathcal{L}$. This means that the optimal controller satisfies,

$$\Delta\mathcal{K}_\star\mathcal{L} = \{\Delta\mathcal{K}_\circ\mathcal{L}\}_+ . \tag{9.38}$$

Also, since $\mathcal{L}^{-1}$ and Δ^{-1} are causal, the optimal causal controller is given by (9.33). Finally, using the fact that $\Delta\mathcal{K}_\circ = \{\Delta\mathcal{K}_\circ\}_+ + \{\Delta\mathcal{K}_\circ\}_-$ and simplifying, we get (9.34). ■

Proof of Theorem 9.3.2. We first simplify our optimization problem (9.13) using Lemma 9.C.1. We then find the conditions on the optimal optimization variables using Karush-Kuhn-Tucker (KKT) conditions. Using Lemma 9.C.1, we can write,

$$\inf_{\substack{\mathcal{K}\in\mathscr{K},\\ \gamma\mathcal{I}\succ\mathcal{R}_\mathcal{K}}} \operatorname{tr}((\mathcal{I}-\gamma^{-1}\mathcal{R}_\mathcal{K})^{-1}\mathcal{M}_\circ) = \inf_{\mathcal{K}\in\mathscr{K}} \sup_{\mathcal{M}\succ 0} -\operatorname{tr}((\mathcal{I}-\gamma^{-1}\mathcal{R}_\mathcal{K})\mathcal{M}) + 2\operatorname{tr}\left(\sqrt{\mathcal{M}_\circ^{\frac{1}{2}}\mathcal{M}\mathcal{M}_\circ^{\frac{1}{2}}}\right)$$

$$= \sup_{\mathcal{M}\succ 0} -\operatorname{tr}(\mathcal{M}) + 2\operatorname{tr}\left(\sqrt{\mathcal{M}_\circ^{\frac{1}{2}}\mathcal{M}\mathcal{M}_\circ^{\frac{1}{2}}}\right) + \inf_{\mathcal{K}\in\mathscr{K}} \gamma^{-1}\operatorname{tr}(\mathcal{R}_\mathcal{K}\mathcal{M})$$

Fixing $\gamma \geq 0$, we focus on the reduced subproblem of (9.14),

$$\sup_{\mathcal{M}\succ 0} -\gamma\operatorname{tr}(\mathcal{M}) - \gamma\operatorname{tr}(\mathcal{M}_\circ) + 2\gamma\operatorname{tr}\left(\sqrt{\mathcal{M}_\circ^{\frac{1}{2}}\mathcal{M}\mathcal{M}_\circ^{\frac{1}{2}}}\right) + \inf_{\mathcal{K}\in\mathscr{K}} \operatorname{tr}(\mathcal{R}_\mathcal{K}\mathcal{M}). \quad (9.39)$$

Using the definition of the Bures-Wasserstein distance, we can reformulate (9.39) as

$$\sup_{\mathcal{M}\succ 0} \inf_{\mathcal{K}\in\mathscr{K}} \operatorname{tr}(\mathcal{R}_\mathcal{K}\mathcal{M}) - \gamma\,\mathsf{BW}(\mathcal{M},\mathcal{M}_\circ)^2 := \sup_{\mathcal{M}\succ 0} \Phi(\mathcal{M}). \quad (9.40)$$

Thus, the original formulation in (9.14) can be expressed as

$$\inf_{\gamma\geq 0} \sup_{\mathcal{M}\succ 0} \inf_{\mathcal{K}\in\mathscr{K}} \operatorname{tr}(\mathcal{R}_\mathcal{K}\mathcal{M}) + \gamma\left(r^2 - \mathsf{BW}(\mathcal{M},\mathcal{M}_\circ)^2\right). \quad (9.41)$$

Note that the objective above is affine in $\gamma \geq 0$ and strictly concave in $\mathcal{M}$. Moreover, primal and dual feasibility hold, enabling the exchange of $\inf_{\gamma\geq 0}\sup_{\mathcal{M}\succ 0}$ resulting in

$$\sup_{\mathcal{M}\succ 0} \inf_{\mathcal{K}\in\mathscr{K}} \inf_{\gamma\geq 0} \operatorname{tr}(\mathcal{R}_\mathcal{K}\mathcal{M}) + \gamma\left(r^2 - \mathsf{BW}(\mathcal{M},\mathcal{M}_\circ)^2\right), \quad (9.42)$$

where the inner minimization over γ reduces the problem to its constrained version in Equation (9.15).

Finally, the form of the optimal $\mathcal{K}_\star$ follows from the Wiener-Hopf technique in Lemma 9.C.2 and the optimal $\gamma_\star$ and $\mathcal{M}_\star$ can be obtained using the strong duality result in Section 9.C. To see the optimal form of $\mathcal{M}_\star$, consider the gradient of $\Phi(\mathcal{M})$ in (9.40) with respect to $\mathcal{M}$ and setting it to 0. Using Danskin theorem [43], we have,

$$\nabla\Phi(\mathcal{M}) = \mathcal{M}_\circ^{\frac{1}{2}}\left(\mathcal{M}_\circ^{\frac{1}{2}}\mathcal{M}_\star\mathcal{M}_\circ^{\frac{1}{2}}\right)^{-\frac{1}{2}}\mathcal{M}_\circ^{\frac{1}{2}} - \mathcal{I} + \gamma^{-1}\mathcal{R}_{\mathcal{K}_\star} = 0. \quad (9.43)$$

Taking inverse on both sides, we get,

$$\mathcal{M}_\circ^{-\frac{1}{2}}\left(\mathcal{M}_\circ^{\frac{1}{2}}\mathcal{M}_\star\mathcal{M}_\circ^{\frac{1}{2}}\right)^{\frac{1}{2}}\mathcal{M}_\circ^{-\frac{1}{2}} = \left(\mathcal{I} - \gamma^{-1}\mathcal{R}_{\mathcal{K}_\star}\right)^{-1}. \quad (9.44)$$

We can now obtain two equations. First, by right multiplying by $\mathcal{M}_\circ^{\frac{1}{2}}$ and second, by left multiplying by $\mathcal{M}_\circ^{\frac{1}{2}}$. On multiplying these two equations together and simplifying, we get (9.16b).

9.D Proofs related to the Efficient Algorithm in Section 9.4

Proof of Lemma 13.4.1

With $\mathbb{T} \coloneqq \{\Delta\mathcal{K}_\circ\}_-$, the optimality condition in (9.16) takes the equivalent form:

$$i.\ \mathcal{M}_\star = \left(\mathcal{I} - \gamma_\star^{-1}\mathcal{R}_{\mathcal{K}_\star}\right)^{-2}, \tag{9.45a}$$

$$ii.\ \mathcal{R}_{\mathcal{K}_\star} = \mathcal{L}_\star^{-*}\left\{\mathbb{T}\mathcal{L}_\star\right\}_-^*\left\{\mathbb{T}\mathcal{L}_\star\right\}_-\mathcal{L}_\star^{-1}, \tag{9.45b}$$

$$iii.\ \operatorname{tr}\left[\left((\mathcal{I} - \gamma_\star^{-1}\mathcal{R}_{\mathcal{K}_\star}(z))^{-1} - \mathcal{I}\right)^2\right] = r^2, \tag{9.45c}$$

Using the spectral factorization $\mathcal{M}_\star\mathcal{L}_\star\mathcal{L}_\star^*$, the conditions *i.* and *ii.* can be equivalently re-expressed as

$$i.\ (\mathcal{L}_\star\mathcal{L}_\star^*)^{-1/2} = I - \gamma_\star^{-1}\mathcal{R}_{\mathcal{K}_\star} \tag{9.46}$$

$$ii.\ \mathcal{R}_{\mathcal{K}_\star} = \mathcal{L}_\star^{-*}\left\{\mathbb{T}\mathcal{L}_\star\right\}_-^*\left\{\mathbb{T}\mathcal{L}_\star\right\}_-\mathcal{L}_\star^{-1} \tag{9.47}$$

By plugging *ii.* into *i.*, we get

$$0 = \mathcal{I} - (\mathcal{L}_\star\mathcal{L}_\star^*)^{-1/2} - \gamma_\star^{-1}\left(\mathcal{L}_\star^{-*}\left\{\mathbb{T}\mathcal{L}_\star\right\}_-^*\left\{\mathbb{T}\mathcal{L}_\star\right\}_-\mathcal{L}_\star^{-1}\right) = 0, \tag{9.48}$$

Multiplying by $\mathcal{L}_\star^*$ from the left and by $\mathcal{L}_\star$ from the right, we get

$$0 = \mathcal{L}_\star^*\mathcal{L}_\star - (\mathcal{L}_\star^*\mathcal{L}_\star)^{1/2} - \gamma_\star^{-1}\left\{\mathbb{T}\mathcal{L}_\star\right\}_-^*\left\{\mathbb{T}\mathcal{L}_\star\right\}_-,$$

where we used the identity $\mathcal{L}_\star^*(\mathcal{L}_\star\mathcal{L}_\star^*)^{-1/2}\mathcal{L}_\star = (\mathcal{L}_\star^*\mathcal{L}_\star)^{1/2}$. Letting $\mathcal{N}_\star = \mathcal{L}_\star^*\mathcal{L}_\star$, this expression can be solved for $\mathbb{N}_\star$, yielding the following implicit equation,

$$\mathcal{N}_\star = \mathcal{L}_\star^*\mathcal{L}_\star = \frac{1}{4}\left(\mathcal{I} + \sqrt{\mathcal{I} + 4\gamma_\star^{-1}\left\{\mathbb{T}\mathcal{L}_\star\right\}_-^*\left\{\mathbb{T}\mathcal{L}_\star\right\}_-}\right)^2, \tag{9.49}$$

implying thus (13.29), with $\gamma_\star > 0$ satisfying $\operatorname{tr}\left[((\mathcal{I} - \gamma_\star^{-1}\mathcal{R}_{\mathcal{K}_\star})^{-1} - \mathcal{I})^2\right] = r^2$ (or equivalently, $\mathsf{BW}(\mathcal{L}_\star\mathcal{L}_\star^*, \mathcal{I}) = r$). ■

Note on Frequency Domain Representation of Toeplitz Operators

We start this section of the appendix by justifying our choice of working out our results in the frequency domain.

Let $\mathcal{V} = [V_{ij}]_{i,j=-\infty}^{\infty}$ be a doubly infinite block matrix, *i.e.*a *Toeplitz* operator, which represents a discrete, linear, time-invariant system (*i.e.*$V_{ij} = V_{i-j}$), and which maps a sequence of inputs to a sequence of outputs.

In this case of a time-invariant system, the representation of the operator in the z-domain (or the so-called bilateral z-transform) is

$$V(z) = \sum_{i=-\infty}^{\infty} V_i z^{-i}, \tag{9.50}$$

defined for the regions of the complex plane where the above series converges absolutely, known as the ROC: region of convergence. $V(z)$ is also known as the transfer matrix. The causality of $\mathcal{V}$ can be readily given in terms of $V(z)$. Indeed, we have the following: $\mathcal{V}$ is causal if and only if $V(z)$ is analytic in the exterior of some annulus, $|z| > \alpha > 0$. Likewise, $\mathcal{V}$ is anticausal if and only if $V(z)$ is analytic in the interior of some annulus, $|z| < \alpha < 0$. Moreover, $\mathcal{V}$ is strictly causal (anticausal) if and only if it is causal (anticausal) and $V(\infty) = 0(V(0) = 0)$.

We also define the trace of a Toeplitz operator $\mathcal{M}$ as follows

$$\mathrm{tr}(\mathcal{M}) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{Tr}(M(\mathrm{e}^{\jmath\omega}))d\omega. \tag{9.51}$$

In the coming sections, we use the frequency domain counterparts of our Toeplitz operators (such as $\mathcal{F}, \mathcal{G}, \mathcal{M}$...) by setting $z = e^{j\omega}$ for $\omega \in [0, 2\pi)$.

Frequency-Domain Characterization of the Optimal Solution of Problem 9.2.3

We present the frequency-domain formulation of the saddle point $(\mathcal{K}_\star, \mathcal{M}_\star)$ derived in Theorem 9.3.2 to reveal the structure of the solution. We first introduce the following useful results:

Denoting by $M_\star(z)$ and $R_{K_\star}(z)$ the transfer functions corresponding to the optimal $\mathcal{M}_\star$ and $\mathcal{R}_{\mathcal{K}_\star}$, respectively, the optimality conditions in (9.16) and (13.29) take the equivalent forms:

$$\textit{i. } M_\star(z) = \left(I - \gamma_\star^{-1} R_{K_\star}(z)\right)^{-2}, \tag{9.52a}$$

$$\textit{ii. } R_{K_\star}(z) = L_\star(z)^{-*}\left\{\mathbb{T}\mathcal{L}_\star\right\}_-(z)^*\left\{\mathbb{T}\mathcal{L}_\star\right\}_-(z)L_\star(z)^{-1}, \tag{9.52b}$$

$$\textit{iii. } \mathrm{tr}\left[\left((I - \gamma_\star^{-1} R_{K_\star}(z))^{-1} - I\right)^2\right] = r^2,, \tag{9.52c}$$

$$\textit{iv. } N_\star(z) = L_\star(z)^* L_\star(z) = \frac{1}{4}\left(I + \sqrt{I + 4\gamma^{-1}\left\{\mathbb{T}\mathcal{L}_\star\right\}_-(z)^*\left\{\mathbb{T}\mathcal{L}_\star\right\}_-(z)}\right)^2, \tag{9.52d}$$

where

$$L_\star(z) = \sum_{t=0}^{\infty} \widehat{L}_{\star,t} z^{-t} \tag{9.53}$$

is the transfer function corresponding to the causal canonical factor $\mathcal{L}_\star$ and $\mathbb{T} = \{\Delta\mathcal{K}_\circ\}_-$ is the strictly anticausal operator where its transfer function, $T(z)$, is found from the following:

Lemma 9.D.1 (Adapted from lemma 4 in [191])**.** *The transfer function $\Delta(z)K_\circ(z)$ can be written as the sum of a causal and strictly anticausal transfer functions:*

$$\Delta(z)K_\circ(z) = T(z) + U(z) \tag{9.54}$$

$$T(z) = \{\Delta\mathcal{K}_\circ\}_-(z) = \overline{C}(z^{-1}I - \overline{A})^{-1}\overline{B} \tag{9.55}$$

$$U(z) = \{\Delta\mathcal{K}_\circ\}_+(z) = \Delta(z)K_{H_2}(z) = \overline{C}P(A(zI - A)^{-1} + I)B_w \tag{9.56}$$

where $K_{H_2}(z) = \Delta^{-1}\{\Delta\mathcal{K}_\circ\}_+(z)$, and $K_{lqr} := (I + B_u^* P B_u)^{-1} B_u^* P A$, with $P \succ 0$ is the unique stabilizing solution to the LQR Riccati equation $P = Q + A^* P A - A^* P B_u (I + B_u^* P B_u)^{-1} B_u^* P A$, $Q = C^\intercal C$, $A_k = A - B_u K_{lqr}$, and

$$\overline{A} = A_k^*, \quad \overline{B} = A_k^* P B_w, \quad \overline{C} = -(I + B_u^* P B_u)^{-*/2} B_u^*. \tag{9.57}$$

Notice that given the causal $L(z)$ and strictly anti-causal $T(z)$, the strictly anti-causal part $\{T(z)L(z)\}_-$ has a state space representation, shown in the following lemma.

Lemma 9.D.2. *Let $\mathcal{L}$ be a causal operator. The strictly anti-causal operator $\{\mathbb{T}\mathcal{L}\}_-$ possesses a state space representation as follows:*

$$\{\mathbb{T}\mathcal{L}\}_-(z) = \overline{C}(z^{-1}I - \overline{A})^{-1}\Gamma, \tag{9.58}$$

where

$$\Gamma = \frac{1}{2\pi}\int_0^{2\pi} (I - e^{\jmath\omega}\overline{A})^{-1}\overline{B}L(e^{\jmath\omega})d\omega. \tag{9.59}$$

Proof. Let $L(z) = \sum_{t=0}^\infty \widehat{L}_t z^{-t}$ be the transfer function of $\mathcal{L}$. Using equations (9.54) and (9.55),(9.56), $S(z) := \{\Delta\mathcal{K}_\circ\mathcal{L}\}_-(z)$, can be written as:

$$S(z) = \{TL\}_-(z) + \{UL\}_-(z) \tag{9.60}$$

$$\stackrel{(a)}{=} \left\{\overline{C}(zI - \overline{A})^{-1}\overline{B}L(z)\right\}_- \tag{9.61}$$

$$\stackrel{(b)}{=} \left\{\overline{C}\sum_{t=0}^\infty z^{(t+1)}\overline{A}^t\overline{B}\sum_{m=0}^\infty \widehat{L}_m z^{-m}\right\}_- \tag{9.62}$$

$$\stackrel{(c)}{=} \overline{C}\left(\sum_{t=0}^\infty z^{(t+1)}\overline{A}^t\right)\left(\sum_{m=0}^\infty \overline{A}^m\overline{B}\widehat{L}_m\right) \tag{9.63}$$

$$\stackrel{(d)}{=} \overline{C}(z^{-1}I - \overline{A})^{-1}\Gamma \tag{9.64}$$

Here, (a) holds because both $U(z)$ and $L(z)$ are causal, so the strictly anticausal part of $U(z)L(z)$ is zero. (b) holds as we do the Neumann series expansion of $(zI - \overline{A})$

and replace $L(z)$ by its equation (9.53). (c) holds as we take the anticausal part of expression (9.62) to be the strictly positive exponents of z. (d) completes the result by using the Neuman series again, defining $\Gamma := \sum_{t=0}^{\infty} \overline{A}^t \overline{B} \widehat{L}_t$, and leveraging Parseval's theorem to conclude the equation of the finite parameter

$$\Gamma = \frac{1}{2\pi} \int_0^{2\pi} (I - e^{\jmath\omega} \overline{A})^{-1} \overline{B} L(e^{\jmath\omega}) d\omega. \tag{9.65}$$

■

Proof of Theorem 13.4.2: Using Lemma 9.D.2, and plugging (9.58) into (9.52d), the frequency-domain optimality equation (9.52d) can be reformulated explicitly as follows

$$N_\star(z) = L_\star(z)^* L_\star(z) = \frac{1}{4} \left(I + \sqrt{I + 4\gamma_\star^{-1} \Gamma_\star^* (z^{-1} I - \overline{A})^{-*} \overline{C}^* \overline{C} (z^{-1} I - \overline{A})^{-1} \Gamma_\star} \right)^2 \tag{9.66}$$

where $\Gamma_\star$ as in (9.65), and $\gamma_\star > 0$ satisfying $\operatorname{tr}\left[((\mathcal{I} - \gamma_\star^{-1} \mathcal{R}_{\mathcal{K}_\star})^{-1} - \mathcal{I})^2 \right] = r^2$ (or equivalently, $\mathsf{BW}(\mathcal{L}_\star \mathcal{L}_\star^*, \mathcal{I}) = r$), which gives the desired result. ■

Proof of Corollary 13.4.3: Notice that the rhs of (9.66) involves the positive definite square-root of the rational term $\Gamma_\star^* (z^{-1} I - \overline{A})^{-*} \overline{C}^* \overline{C} (z^{-1} I - \overline{A})^{-1} \Gamma_\star$. The square root does not preserve rationality in general, implying the desired result. ■

Proof of Theorem 9.4.2

Before proceeding with the proof, we state the following useful lemma.

Lemma 9.D.3. *For a positive-definite Toeplitz operator $\mathcal{M} \succ 0$ with $\operatorname{tr}(\mathcal{M}) < \infty$ and $\operatorname{tr}(\log(\mathcal{M})) > -\infty$, let $\mathcal{M} \mapsto \Phi(\mathcal{M})$ be a mapping defined as*

$$\Phi(\mathcal{M}) \triangleq \inf_{\mathcal{K} \in \mathscr{K}} \operatorname{tr}(\mathcal{R}_{\mathcal{K}} \mathcal{M}). \tag{9.67}$$

Denote by $\mathcal{M} = \mathcal{L}\mathcal{L}^*$ and $\Delta\Delta^* = \mathcal{I} + \mathcal{F}^*\mathcal{F}$ the canonical spectral factorizations where $\mathcal{L}$, Δ as well as their inverses $\mathcal{L}^{-1}$, Δ^{-1} are causal operators. The following statements hold:

i. The function Φ can be written in closed form as

$$\Phi(\mathcal{M}) = \operatorname{tr}\left[\{\Delta \mathcal{K}_\circ \mathcal{L}\}_-^* \{\Delta \mathcal{K}_\circ \mathcal{L}\}_- \right]. \tag{9.68}$$

ii. The gradient of Φ has the following closed form

$$\nabla\Phi(\mathcal{M}) = \mathcal{R}_{\mathcal{K}} = \mathcal{L}^{-*} \{\Delta \mathcal{K}_\circ \mathcal{L}\}_-^* \{\Delta \mathcal{K}_\circ \mathcal{L}\}_- \mathcal{L}^{-1}. \tag{9.69}$$

iii. The function Φ is concave, positively homogeneous, and

$$\Phi(\mathcal{M}) = \mathrm{tr}(\mathcal{M}\,\nabla\Phi(\mathcal{M})). \tag{9.70}$$

Proof of Theorem 9.4.2. Our proof of convergence follows closely from the proof technique used in [116]. In particular, since the unit circle is discretized and the computation of the gradients are approximate, the linear suboptimal problem is solved upto an approximation, δ_N which depends on the problem parameters and the discretization level N. Namely, for a large enough N, we have

$$\mathrm{tr}(\nabla\Phi(\mathcal{M}_k)\widetilde{\mathcal{M}}_{k+1}) \geq \sup_{\mathcal{M}\in\Omega_r} \mathrm{tr}(\nabla\Phi(\mathcal{M}_k)\mathcal{M}) - \delta_N \tag{9.71}$$

where

$$\Omega_r := \{\mathcal{M} \succ 0 \mid \mathrm{tr}(\mathcal{M} - 2\sqrt{\mathcal{M}} + \mathcal{I}) \leq r^2\}, \tag{9.72}$$

Therefore, using Theorem 1 of [116], we obtain

$$\Phi(\mathcal{M}_\star) - \Phi(\mathcal{M}_k) \leq \frac{2\kappa}{k+2}(1+\delta_N). \tag{9.73}$$

where $\kappa > 0$ is the so-called curvature constant associated with the problem which is defined as follows

$$\begin{aligned}
\kappa :=& \sup_{\substack{\mathcal{M},\widetilde{\mathcal{M}}\in\Omega_r \\ \eta\in[0,1] \\ \mathcal{M}'=\mathcal{M}+\eta(\widetilde{\mathcal{M}}-\mathcal{M})}} \frac{2}{\eta^2}\left[-\Phi(\mathcal{M}') + \Phi(\mathcal{M}) + \mathrm{tr}(\nabla\Phi(\mathcal{M})\,(\mathcal{M}'-\mathcal{M}))\right], && (9.74)\\
=& \sup_{\substack{\mathcal{M},\widetilde{\mathcal{M}}\in\Omega_r \\ \eta\in[0,1] \\ \mathcal{M}'=\mathcal{M}+\eta(\widetilde{\mathcal{M}}-\mathcal{M})}} \frac{2}{\eta^2}\left(\mathrm{tr}(\mathcal{M}'\nabla\Phi(\mathcal{M})) - \Phi(\mathcal{M}')\right), && (9.75)\\
=& \sup_{\substack{\mathcal{M},\widetilde{\mathcal{M}}\in\Omega_r \\ \eta\in[0,1] \\ \mathcal{M}'=\mathcal{M}+\eta(\widetilde{\mathcal{M}}-\mathcal{M})}} \inf_{\mathcal{K}\in\mathscr{K}} \frac{2}{\eta^2}\,\mathrm{tr}\left(\mathcal{M}'(\nabla\Phi(\mathcal{M}) - \mathcal{R}_{\mathcal{K}})\right) && (9.76)
\end{aligned}$$

where the last two equalities follow from Lemma 9.D.3.

9.E Algorithms

Pseudocode for Frequency-domain Iterative Optimization Method Solving (9.15)

The pseudocode for Frequency-domain tterative optimization method is presented in Algorithm 4.

Algorithm 4 Frequency-domain iterative optimization method solving (9.15)

Input: Radius $r > 0$, state-space model (A, B_u, B_w), discretizations $N > 0$ and $N' > 0$ tolerance $\epsilon > 0$
Compute $(\overline{A}, \overline{B}, \overline{C})$ from (A, B_u, B_w) using (9.57)
Generate frequency samples $\mathbb{T}_N := \{e^{j2\pi n/N} \mid n = 0, \ldots, N-1\}$
Initialize $M_0(z) \leftarrow I$ for $z \in \mathbb{T}_N$, and $k \leftarrow 0$
repeat
  Set the step size $\eta_k \leftarrow \frac{2}{k+2}$
  Compute the spectral factor $L_k(z) \leftarrow \texttt{SpectralFactor}(M_k)$ (see Section 9.E)
  Compute the parameter $\Gamma_k \leftarrow \dfrac{1}{N} \sum_{z \in \mathbb{T}_N} (I - z\overline{A})^{-1} \overline{B} L_k(z)$. (see Section 9.E)
  Compute the gradient for $z \in \mathbb{T}_N$ (see Section 9.E)
    $R_k(z) \leftarrow L_k(z)^{-*} \left\{\Delta \mathcal{K}_\circ \mathcal{L}_k\right\}_- (z)^* \left\{\Delta \mathcal{K}_\circ \mathcal{L}_k\right\}_- (z) L_k(z)^{-1}$
  Solve the linear subproblem (9.19a) via bisection (see Section 9.E)
    $\widetilde{M}_k(z) \leftarrow (I - \gamma_k^{-1} R_k(z))^{-2}$ for $z \in \mathbb{T}_N$ and γ_k through $\texttt{Bisection}$
  Set $M_{k+1}(z) \leftarrow (1 - \eta_k) M_k(z) + \eta_k \widetilde{M}_k(z)$ for $z \in \mathbb{T}_N$.
  Increment $k \leftarrow k + 1$
until $\|M_{k+1} - M_k\| / \|M_k\| \le \epsilon$
Compute $N_k(z) = \frac{1}{4} \left(I + \sqrt{I + 4\gamma_k^{-1} \Gamma_k^* (z^{-1} I - \overline{A})^* \overline{C}^* \overline{C} (z^{-1} I - \overline{A})^{-1} \Gamma_k} \right)^2$ for $z \in \mathbb{T}_{N'} := \{e^{j2\pi n/N'} \mid n = 0, \ldots, N'-1\}$
Compute $K(z) \leftarrow \texttt{RationalApproximate}(N_k(z))$ (see Section 9.E)

Additional Discussion on the Computation of Gradients

By the Wiener-Hopf technique discussed in Lemma 9.C.2, the gradient $\mathcal{R}_k = \nabla \Phi(\mathcal{M}_k)$ can be obtained as

$$R_k(z) = L_k(z)^{-*} \left\{\Delta \mathcal{K}_\circ \mathcal{L}_k\right\}_- (z)^* \left\{\Delta \mathcal{K}_\circ \mathcal{L}_k\right\}_- (z) L_k(z)^{-1}, \tag{9.77}$$

where $\mathcal{L}_k \mathcal{L}_k^* = \mathcal{M}_k$ is the unique spectral factorization. Furthermore, using (9.64),(9.65), we can reformulate the gradient $R_k(z)$ more explicitly as

$$R_k(z) = L_k(z)^{-*} \Gamma_k^* (I - z\overline{A})^{-*} \overline{C}^* \, \overline{C} (I - z\overline{A})^{-1} \Gamma_k L_k(z)^{-1}, \tag{9.78}$$

where $\Gamma_k = \frac{1}{2\pi} \int_0^{2\pi} (I - e^{j\omega} \overline{A})^{-1} \overline{B} L_k(e^{j\omega}) d\omega$ as in (9.65). Here, the spectral factor $L_k(z)$ is obtained for $z \in \mathbb{T}_N$ by Section 9.E. Similarly, the parameter Γ_k can be computed numerically using trapezoid rule over the discrete domain $\mathbb{T}_N$, *i.e.*,

$$\Gamma_k \leftarrow \frac{1}{N} \sum_{z \in \mathbb{T}_N} (I - z\overline{A})^{-1} \overline{B} L_k(z). \tag{9.79}$$

The gradient $R_k(z)$ can thus be efficiently computed for $z \in \mathbb{T}_N$.

Implementation of Spectral Factorization

To perform the spectral factorization of an irrational function $M(z)$, we use a spectral factorization method via discrete Fourier transform, which returns samples of the spectral factor on the unit circle. First, we compute $\Lambda(z)$ for $z \in \mathbb{T}_N$, which is defined to be the logarithm of $M(z)$, then we take the inverse discrete Fourier transform λ_k for $k = 0, \ldots, N-1$ of $\Lambda(z)$ which we use to compute the spectral factorization as

$$L(z_n) \leftarrow \exp\left(\frac{1}{2}\lambda_0 + \sum_{k=1}^{N/2-1} \lambda_k z_n^{-k} + \frac{1}{2}(-1)^n \lambda_{N/2}\right)$$

for $k = 0, \ldots, N-1$ where $z_n = \mathrm{e}^{j2\pi n/N}$.

The method is efficient without requiring rational spectra, and the associated error term, featuring a purely imaginary logarithm, rapidly diminishes with an increased number of samples. It is worth noting that this method is explicitly designed for scalar functions.

Algorithm 5 `SpectralFactor`: Spectral Factorization via DFT

1: **Input:** Scalar positive spectrum $M(z) > 0$ on $\mathbb{T}_N \coloneqq \{\mathrm{e}^{j2\pi n/N} \mid n{=}0, \ldots, N{-}1\}$
2: **Output:** Causal spectral factor $L(z)$ of $M(z) > 0$ on $\mathbb{T}_N$
3: Compute the cepstrum $\Lambda(z) \leftarrow \log(M(z))$ on $z \in \mathbb{T}_N$.
4: Compute the inverse DFT
$\lambda_k \leftarrow \mathrm{IDFT}(\Lambda(z))$ for $k = 0, \ldots, N{-}1$
5: Compute the spectral factor for $z_n = \mathrm{e}^{j2\pi n/N}$
$$L(z_n) \leftarrow \exp\left(\frac{1}{2}\lambda_0 + \sum_{k=1}^{N/2-1} \lambda_k z_n^{-k} + \frac{1}{2}(-1)^n \lambda_{N/2}\right), \quad n = 0, \ldots, N{-}1$$

Implementation of Bisection Method

To find the optimal parameter γ_k that solves $\operatorname{tr}\left[((I - \gamma_k^{-1}\mathcal{R}_k)^{-1} - I)^2\right] = r^2$ in the Frank-Wolfe update (9.20), we use a bisection algorithm. The pseudo code for the bisection algorithm can be found in Algorithm 6. We start off with two guesses of γ *i.e.*$(\gamma_{left}, \gamma_{right})$ with the assumption that the optimal γ lies between the two values (without loss of generality).

Algorithm 6 `Bisection Algorithm`

Input: $h(\gamma), \gamma_{right}, \gamma_{left}$
Compute the gradient at γ_{right}: $\nabla h(\gamma)|_{\gamma_{right}}$
while $| \gamma_{right} - \gamma_{left} | > \epsilon$ **do**
  Calculate the midpoint γ_{mid} between γ_{left} and γ_{right}
  Compute the gradient at γ_{mid}: $\nabla h(\gamma)|_{\gamma_{mid}}$
  if $\nabla h(\gamma)|_{\gamma_{mid}}$ = 0 **then**
    return γ_{mid}
  else if $\nabla h(\gamma)|_{\gamma_{mid}} > 0$ **then**
    Update γ_{right} to γ_{mid}
  else
    Update γ_{left} to γ_{mid}
  end if
end while
return the average of γ_{left} and γ_{right}

Implementation of Rational Approximation

We present the pseudocode of `RationalApproximation`.

Algorithm 7 `RationalApproximation`

Input: Scalar positive spectrum $N(z) > 0$ on $\mathbb{T}_{N'} := \{\mathrm{e}^{j2\pi n/N'} \mid n=0,\ldots,N'-1\}$, and a small positive scalar ϵ
Output: Causal controller $K(z)$ on $\mathbb{T}_{N'}$
Get $P(z), Q(z)$ by solving the convex optimization in (9.24), for *fixed* ϵ, given $M(z)$, *i.e.*:

$$\min_{p_0,\ldots p_m \in \mathbb{R}, q_0,\ldots q_m \in \mathbb{R}, \varepsilon \geq 0} \varepsilon \tag{9.80}$$
$$\text{s.t.} \quad q_0 = 1, \quad P(z), Q(z) > 0, \quad P(z)\,(N(z)+\epsilon)\,Q(z) \leq 0, \quad P(z)\,(N(z)\epsilon)\,Q(z) \geq 0 \quad \forall z \in \mathbb{T}_{N'}$$

Get the rational spectral factors of $P(z), Q(z)$, which are $S_P(z), S_Q(z)$ using the canonical Factorization method in [199]
Get $L^r(z)$,the rational spectral factor of$N(z)$, as $S_P(z)/S_Q(z)$
Get $K(z)$ from the formulation in (9.28),(9.99)

9.F Proof of the State-Space Representation of the Controller

Proof of Lemma 9.5.7

Let the spectral factor $\tilde{L}(z)$ in rational form be given as

$$\tilde{L}(z) = (I + \tilde{C}(zI - \tilde{A})^{-1}\tilde{B})\tilde{D}^{1/2}, \tag{9.81}$$

with its inverse given by:

$$\tilde{L}^{-1}(z) = \tilde{D}^{-1/2}(I - \tilde{C}(zI - (\tilde{A} - \tilde{B}\tilde{C}))^{-1}\tilde{B}), \tag{9.82}$$

and its operator form denoted by $\tilde{\mathcal{L}}$.

We write the DR-RO controller, $K(z)$, as a sum of causal functions:

$$\begin{aligned} K(z) &= \Delta^{-1}(z)\{\Delta\mathcal{K}_\circ\tilde{\mathcal{L}}\}_+(z)\tilde{L}^{-1}(z) && (9.83) \\ &= \Delta^{-1}(z)\left(\{\Delta\mathcal{K}_\circ\}_+(z)\tilde{L}(z) + \{\{\Delta\mathcal{K}_\circ\}_-\tilde{\mathcal{L}}\}_+(z)\right)\tilde{L}^{-1}(z) && (9.84) \\ &= \Delta^{-1}(z)\{\Delta\mathcal{K}_\circ\}_+(z) + \Delta^{-1}\{\{\Delta\mathcal{K}_\circ\}_-\tilde{\mathcal{L}}\}_+(z)\tilde{L}^{-1}(z). && (9.85) \end{aligned}$$

From Lemma 4 in [192], we have:

$$\{\Delta\mathcal{K}_\circ\}_-(z) = -\bar{R}B_u^*(z^{-1}I - A_k^*)^{-1}A_k^*PB_w \tag{9.86}$$

where the LQR controller is defined as $K_{lqr} = (I + B_u^*PB_u)^{-1}B_u^*PA$ and the closed-loop matrix $A_K = A - B_uK_{lqr}$ with $P \succ 0$ is the unique stabilizing solution to the LQR Riccati equation $P = Q + A^*PA - A^*PB_u(I + B_u^*PB_u)^{-1}B_u^*PA$, $Q = C^\intercal C$, and with $\bar{R} = (I + B_u^*PB_u)^{-*/2}$.

Multiplying equation (9.86) with $\tilde{L}$, and taking its causal part, we get:

$$\{\{\Delta\mathcal{K}_\circ\}_-\tilde{\mathcal{L}}\}_+(z) = \{-\bar{R}B_u^*(z^{-1}I - A_k^*)^{-1}A_k^*PB_w\tilde{C}(zI - \tilde{A})^{-1}\tilde{B}\tilde{D}^{1/2} - \bar{R}B_u^*(z^{-1}I - A_k^*)^{-1}A_k^*PB_w\tilde{D}^{1} \tag{9.87}$$

Given that the term $\bar{R}B_u^*(z^{-1}I - A_k^*)^{-1}A_k^*PB_w\tilde{D}^{1/2}$ is strictly anticausal, and considering the matrix $\tilde{U}$ which solves the lyapunov equation: $A_k^*PB_w\tilde{C} + A_k^*\tilde{U}A = \tilde{U}$, we get $\{\{\Delta\mathcal{K}_\circ\}_-\tilde{\mathcal{L}}\}_+(z)$ as:

$$\{\{\Delta\mathcal{K}_\circ\}_-\tilde{\mathcal{L}}\}_+(z) = \{-\bar{R}B_u^*((z^{-1}I - A_k^*)^{-1}A_k^*\tilde{U} + \tilde{U}\tilde{A}(zI - \tilde{A})^{-1} + \tilde{U})\tilde{B}\tilde{D}^{1/2}\}_+ \tag{9.88}$$

$$\begin{aligned} &= -\bar{R}B_u^*\tilde{U}(\tilde{A}(zI - \tilde{A})^{-1} + I)\tilde{B}\tilde{D}^{1/2} && (9.89) \\ &= -z\bar{R}B_u^*\tilde{U}(zI - \tilde{A})^{-1}\tilde{B}\tilde{D}^{1/2} && (9.90) \end{aligned}$$

Now, multiplying equation (9.90) by the inverse of $\tilde{L}$ (9.82), we get:

$$\{\{\Delta\mathcal{K}_\circ\}_-\tilde{\mathcal{L}}\}_+(z)\tilde{L}^{-1}(z) = -z\bar{R}B_u^*\tilde{U}(zI-\tilde{A})^{-1}\tilde{B}(I+\tilde{C}(zI-\tilde{A})^{-1}\tilde{B})^{-1} \tag{9.91}$$

$$= -z\bar{R}B_u^*\tilde{U}(zI-\tilde{A})^{-1}(I+\tilde{B}\tilde{C}(zI-\tilde{A})^{-1})^{-1}\tilde{B} \tag{9.92}$$

$$= -z\bar{R}B_u^*\tilde{U}(zI-\tilde{A}_k)^{-1}\tilde{B} \tag{9.93}$$

$$= -\bar{R}B_u^*\tilde{U}(I+(zI-\tilde{A}_k)^{-1}\tilde{A}_k)\tilde{B} \tag{9.94}$$

where $\tilde{A}_k = \tilde{A} - \tilde{B}\tilde{C}$.

The inverse of Δ is given by $\Delta^{-1}(z) = (I - K_{lqr}(zI-A_k)^{-1}B_u)\bar{R}^*$, and we know from lemma 4 in [192] that $\{\Delta\mathcal{K}_\circ\}_+(z) = -\bar{R}B_u^*PA(zI-A)^{-1}B_w - \bar{R}B_u^*PB_w$.

Then we can get the 2 terms of equation (9.85):

$$\Delta^{-1}(z)\{\Delta\mathcal{K}_\circ\}_+(z) = -K_{lqr}(zI-A_k)^{-1}(B_w - B_u\bar{R}^*\bar{R}B_u^*PB_w) - \bar{R}^*\bar{R}B_u^*PB_w \tag{9.95}$$

and

$$\Delta^{-1}(z)\{\{\Delta\mathcal{K}_\circ\}_-\tilde{\mathcal{L}}\}_+(z)\tilde{L}^{-1}(z) = -(I-K_{lqr}(zI-A_k)^{-1}B_u)\bar{R}^*\bar{R}B_u^*\tilde{U}(zI-\tilde{A}_k)^{-1}\tilde{A}_k\tilde{B} \tag{9.96}$$

$$+ K_{lqr}(zI-A_k)^{-1}B_u\bar{R}^*\bar{R}B_u^*\tilde{U}\tilde{B} \tag{9.97}$$

$$- \bar{R}^*\bar{R}B_u^*\tilde{U}\tilde{B} \tag{9.98}$$

Finally, summing equations (9.95) and (9.96), we get the controller $K(z)$ in its rational form:

$$K(z) = \underbrace{-\begin{bmatrix}\bar{R}^*\bar{R}B_u^* & -K_{lqr}\end{bmatrix}}_{\tilde{H}}(zI - \underbrace{\begin{bmatrix}\tilde{A}_K & 0\\ B_u\bar{R}^*\bar{R}B_u^* & A_k\end{bmatrix}}_{\tilde{F}})^{-1}\underbrace{\begin{bmatrix}\tilde{A}_K\tilde{B}\\ -B_w + B_u\bar{R}^*\bar{R}B_u^*(PB_w+U_1\tilde{B})\end{bmatrix}}_{\tilde{G}}$$
$$\underbrace{-\bar{R}^*\bar{R}B_u^*(PB_w+U_1\tilde{B})}_{\tilde{J}} \tag{9.99}$$

which can be explicitly rewritten as in (9.28).

9.G SDP Formulation for the Finite Horizon from [219]

In this section, we state the SDP formulation of the finite-horizon DR-RO control problem for a fixed horizon $T>0$ presented in [219], which is the main controller we compare against, to showcase the value of the infinite-horizon setting. This result highlights the triviality of non-causal estimation as opposed to causal estimation. In Theorem 9.G.2, we demonstrate that the finite-horizon DR-RO problem reduces to an SDP.

Problem 9.G.1 (Distributionally Robust Regret-Optimal (W_2-DR-RO) Control in the Finite Horizon). Find a casual and time-invariant controller, $\mathcal{K}_T \in \mathscr{K}_T$, that minimizes the worst-case expected regret in the finite horizon (9.2), *i.e.*,

$$\inf_{\mathcal{K}_T \in \mathscr{K}_T} R(\mathcal{K}_T, r) \tag{9.100}$$

Theorem 9.G.2 (Adapted from [219]. An SDP formulation for finite-horizon DR-RO)**.** *Let the horizon $T>0$ be fixed and given the noncausal controller $\mathcal{K}_{\circ,T} := -(\mathcal{I}_T + \mathcal{F}_T^* \mathcal{F}_T)^{-1} \mathcal{F}_T^* \mathcal{G}_T$, the Problem 9.G.1 reduces to the following SDP*

$$\inf_{\substack{\mathcal{K}_T \in \mathscr{K}_T, \\ \gamma \geq 0,\, \mathcal{X}_T \succ 0}} \gamma(r_T^2 - \mathrm{tr}(\mathcal{I}_T)) + \mathrm{tr}(\mathcal{X}_T) \quad \text{s.t.} \quad \begin{bmatrix} \mathcal{X}_T & \gamma \mathcal{I}_T & 0 \\ \gamma \mathcal{I}_T & \gamma \mathcal{I}_T & (\mathcal{K}_T - \mathcal{K}_{\circ,T})^* \\ 0 & \mathcal{K}_T - \mathcal{K}_{\circ,T} & (\mathcal{I}_T + \mathcal{F}_T^* \mathcal{F}_T)^{-1} \end{bmatrix} \succcurlyeq 0.$$

Moreover, the worst-case disturbance $\mathbf{w}_T^\star$ can be identified from the nominal disturbances $\mathbf{w}_{\circ,T}$ as $\mathbf{w}_T^\star = (\mathcal{I}_T - \gamma_\star^{-1} \mathcal{T}_{\mathcal{K}_{\circ,T}}^* \mathcal{T}_{\mathcal{K}_{\circ,T}})^{-1} \mathbf{w}_{\circ,T}$ where $\gamma_\star > 0$ and $\mathcal{K}_T^\star$ are the optimal solutions.

Note that the scaling of the SDP in Theorem 9.G.2 with the time horizon is prohibitive for many time-critical real-world applications. Therefore, we compare our infinite-horizon controller to the finite-horizon one in the simulation sections 9.6 and 9.H.

9.H Additional Simulations

Note on Comparison with Other Methods in the Literature:

As our work is the first to explore infinite-horizon distributionally robust control, our comparative experiments are constrained by the existing literature on finite-horizon distributionally robust control. Since the closest work to ours is that of [219], our numerical experiments primarily compare with their finite-horizon version that utilizes an SDP formulation.

Unfortunately, the framework in [220] only allows for time-independent disturbances. While this approach is valuable for partially observed systems, it is widely acknowledged that the optimal distributionally robust controller for fully observed systems remains the same as the standard LQR controller as long as the disturbances are independent (though not necessarily identical) [101]. Therefore, in our setup, the results from [220] simply reduce to the optimal LQR controller. This observation has also been noted in [220].

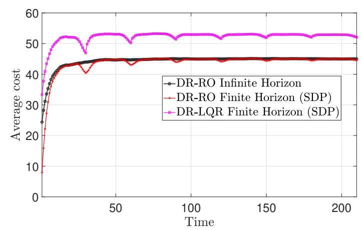
While in the main text we simulated under the worst-case distributions corresponding to each controller being compared, we include in this section of the appendix other systems under the worst-case distributions, and also under other disturbance realizations (namely sinusoidal and uniform distributions).

Additional Time Domain and Frequency Domain Simulations

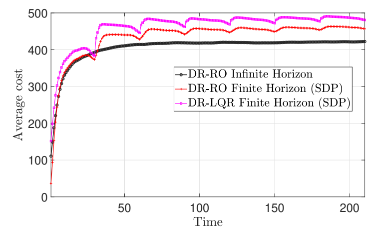
Time domain simulations: We repeat the same experiment of section 9.6 for 2 more systems, [REA4] and [HE3] [145]. [REA4] is a SISO system with 8 states and a stable A matrix, while [HE3] has 4 states and an unstable A matrix. The results are shown in figures 9.4,9.5. Similarly to our previous discussion, the infinite horizon DRRO controller achieves good performance across all systems, achieving the lowest cost under all considered noise scenarios.

In figures 9.6 and 9.7, we show the performance of the different DR controllers: (I) DR-RO in infinite horizon, (II) DR-RO in finite horizon and (III) DR-LQR in finite horizon under uniform noise and sinusoidal noise, respectively, for different systems. Note that the distributionally robust controller is guaranteed to perform better than other controllers under its own worst-case distribution, but has no guarantee of performance under other disturbances. Under uniform and sinusoidal noise, our infinite horizon DR-RO controller performs better than the finite horizon DR-LQR for systems [REA4] and [AC15], but worse than the finite horizon DR-LQR and on par with the finite horizon DR-RO for system [HE3].

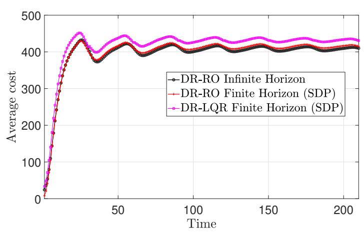
Frequency domain simulations We show in figure 9.8 the frequency domain representation of the square of the norm of the DR-RO controller and its approximation for [AC15] and [HE3], demonstrating that lower order approximations of $m(e^{j\omega})$ provide good estimates.



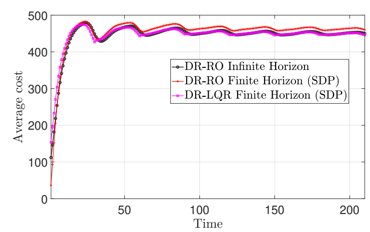
(a) White noise



(b) Worst disturbance for (I)

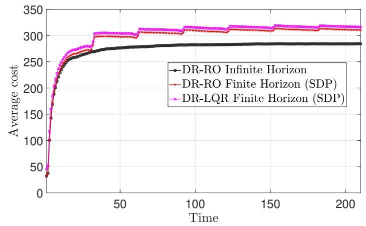


(c) Worst disturbance for (II)

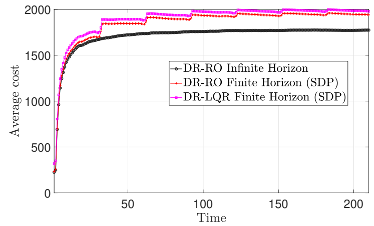


(d) Worst disturbance for (III)

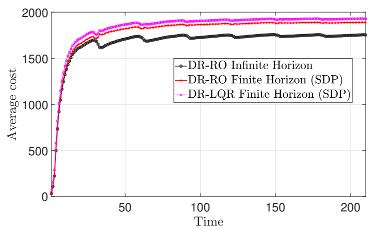
Figure 9.4: The control costs of the different DR controllers: (I) DR-RO in infinite horizon, (II) DR-RO in finite horizon and (III) DR-LQR in finite horizon under different disturbances for system [REA4] [145]. (a) is white noise, while (b), (c) and (d) are worst-case disturbances for each of the controllers, for $r = 1.5$. The finite-horizon controllers are re-applied every $s = 30$ steps. For all disturbances, the infinite horizon DRRO controller achieves lowest average cost, even in cases (c) and (d) where the finite horizon DR controllers are designed to minimize the cost.



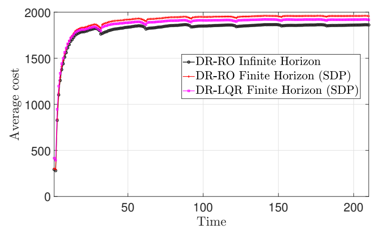
(a) White noise



(b) Worst disturbance for (I)

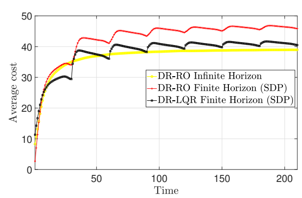


(c) Worst disturbance for (II)

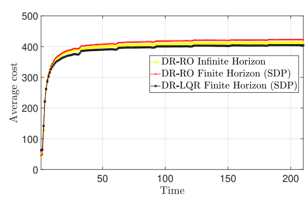


(d) Worst disturbance for (III)

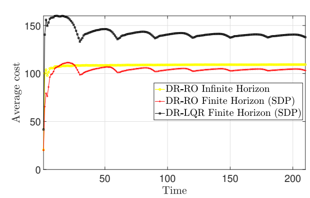
Figure 9.5: The control costs of the different DR controllers: (I) DR-RO in infinite horizon, (II) DR-RO in finite horizon and (III) DR-LQR in finite horizon under different disturbances for system [HE3] [145]. (a) is white noise, while (b), (c) and (d) are worst-case disturbances for each of the controllers, for $r = 1.5$. The finite-horizon controllers are re-applied every $s = 30$ steps. For all disturbances, the infinite horizon DRRO controller achieves lowest average cost, even in cases (c) and (d) where the finite horizon DR controllers are designed to minimize the cost.



(a) System [REA4]

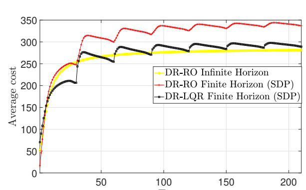


(b) System [HE3]

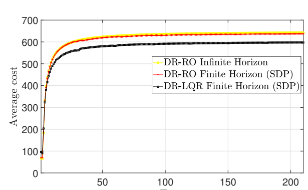


(c) System [AC15]

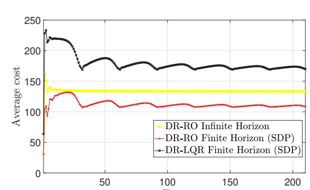
Figure 9.6: The control costs of the different DR controllers: DR-RO in infinite horizon, DR-RO in finite horizon and DR-LQR in finite horizon under uniform noise distributions (with amplitude=2) for different systems, for $r = 1.5$.



(a) System [REA4]

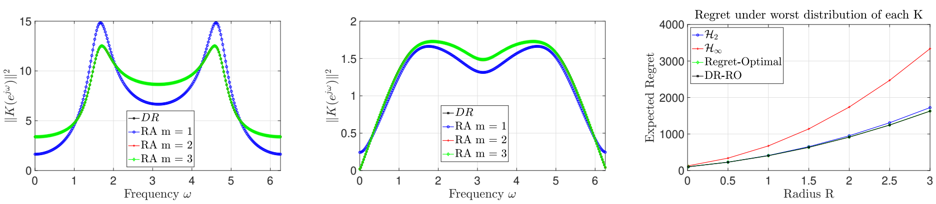


(b) System [HE3]



(c) System [AC15]

Figure 9.7: The control costs of the different DR controllers: DR-RO in infinite horizon, DR-RO in finite horizon and DR-LQR in finite horizon under sinusoidal noise distributions (frequency=1, phase=$\pi/4$, amplitude=2) for different systems, for $r = 1.5$.



(a) (b) (c)

Figure 9.8: The frequency domain representation of the square of the norm of the DR-RO controller $K(e^{jw})$ and its approximation for [AC15] 9.8a and [HE3] 9.8b. Figures 9.8a and 9.8b reaffirm our conclusions that lower order approximations of $m(e^{jw})$ still yield good estimates of the same. Figure 9.8c represents the worst case expected regret of $\mathcal{H}_2$, $\mathcal{H}_\infty$ and the RO controller.

Chapter 10

INFINITE-HORIZON DISTRIBUTIONALLY ROBUST KALMAN FILTERING

10.1 Introduction

The Kalman filter (KF), introduced by Rudolf Kalman in 1960 [121], is a fundamental tool for estimating dynamic signals generated by state-space models from noisy measurements. It has become indispensable across various fields, such as tracking [33], [252], navigation [85], [107], robotics [35], [119], [189], autonomous vehicles [65], [208], aerospace [11], [143], [181], earth sciences [15], [71], [98], [226], biomedicine [72], [185], [214], economics and finance [129], [200], [235]. Its efficacy hinges heavily on accurately modeling state-space parameters and noise statistics, which often deviate from the actual model due to statistical and approximation errors, inherent environmental uncertainties, and non-stationarities. These deviations can severely degrade performance [77], [86], [215], posing severe risks in safety-critical applications such as aircraft navigation and autonomous vehicles. Therefore, enhancing the robustness of the Kalman filter against inaccuracies and uncertainties is crucial for ensuring safe and reliable operation.

Traditionally, robustness in the Kalman filter has been addressed by treating uncertainties as adversarial, deterministic perturbations. In this context, the $\mathscr{H}_\infty$-filter [70], [87], [99], [101], [168], [204], [244] has garnered extensive research, driven by significant advances in robust control theory [14], [51], [101], [260]. The $\mathscr{H}_\infty$-filter enhances robustness by minimizing the worst-case mean-squared estimation error (MSE) attainable among all bounded energy (or power) disturbances. Although these uncertainties are presumed to arise from exogenous disturbances, the optimal $\mathscr{H}_\infty$-filter also ensures robustness against small modeling errors in state-space parameters [255]. More recently, regret-optimal filtering [82], [193] has been introduced to balance performance and robustness. Unlike the $\mathscr{H}_\infty$-filter, it minimizes the worst-case regret, defined as the excess error a causal estimator suffers compared to a clairvoyant estimator, among all bounded energy disturbances. While effective against large uncertainties, these filters neglect distributional information and may become overly conservative when faced with stochastic disturbances [179].

Distributionally robust (DR) estimation and filtering offers an alternative framework that addresses the limitations of traditional robust filtering. Pioneered by Kassam and Poor [127], [128] in the context of Wiener filtering [242], this approach enhances robustness against uncertainties through the use of ambiguity sets of plausible statistical models. The behavior of the resulting robust filter is intricately tied to the topology of the ambiguity set, which is often constructed as a ball induced by a statistical distance or divergence. Examples include the total variation (TV) distance [184], [230], the Kullback-Leibler (KL) divergence [146], [147], [264], [265], and the Wasserstein-2 (W_2) distance [29], [95], [153], [186], [203], [233], [234]. The filters derived from KL-ambiguity sets have been linked [28], [97], [146], [147] to risk-sensitive filters [99], [115], [216], [217], [237], which minimize the exponentiated squared estimation error. A significant drawback of KL-ambiguity is its limited expressivity, as it only includes distributions whose support matches the nominal distribution [109]. Due to its geometric interpretability as the optimal transportation metric [231], the W_2-distance has recently seen widespread adoption across various fields, including machine learning [9], computer vision [151], [177], control [7], [30], [89], [125], [219], [220], data compression [26], [144], [157], and robust optimization [24], [25], [74], [134], [166], [253], [257]. W_2-ambiguity sets offer richer expressivity, encompassing distributions with both discrete and continuous support. The W_2-distance also renders computationally tractable formulations for problems involving quadratic objectives, such as least mean-squared estimation [171], and linear-quadratic control [219].

Related Works

Recognizing these advantages, Shafieezadeh Abadeh, Nguyen, Kuhn, *et al.* [203] introduced a distributionally robust Kalman filter based on W_2-ambiguity sets confined to Gaussian distributions only. They derive state estimates at local time instances by minimizing the mean-squared error for the least favorable joint posterior distribution of the state-measurement vector, given past measurements. This iterative procedure, assuming iid Gaussian disturbances, incorporates the worst-case covariance of the previous state estimate into the nominal model for the subsequent time step. However, while this method inherently addresses state-space parameter mismatches, it lacks a global robustness guarantee over the entire time horizon and against non-iid or non-Gaussian disturbances. Similar temporally local approaches have also been studied in [95], [233], [234]. More recently, Lotidis, Bambos, Blanchet, *et al.* [153] took a different approach by imposing distributional uncertainty on the measurement noise process over the entire time horizon, assuming known iid process noise with known covariance. While the resulting filter demonstrates global robustness over the entire time horizon, the adversarial measurement noise is constrained by martingale conditions to prevent clairvoyance and dependence on future process noise realizations. Moreover, the assumption of known iid process noise is restrictive and does not provide robustness to modeling errors of the dynamics and the process noise.

Contributions

In this work, we consider the Wasserstein-2 distributionally robust Kalman filtering (W_2-DR-KF) of linear state-space models for both finite and infinite horizons. The probability distribution of the disturbances over the entire time horizon is assumed to lie in a W_2-ball of a specified radius centered at a given nominal distribution. We seek the optimal causal linear estimator of a target signal that minimizes the worst-case MSE within the W_2-ball. We cast this as a min-max optimization problem (**??**, Problem 10.2.2). Our approach differs drastically from the prior works [95], [153], [203], [233], [234] and possesses several advantages which can be listed as follows:

1. Global robustness to non-iid disturbances: In contrast to focusing on the worst-case MSE at local time instances under unknown iid disturbances [95], [203], [233], [234], our approach minimizes the cumulative MSE under the worst-case disturbance trajectory, thereby achieving global robustness for the entire horizon. Moreover, unlike [153], we impose no restrictions on the dependencies of the disturbances, accommodating arbitrarily correlated process and measurement noise sequences.

2. Bounded steady-state error: We derive the first infinite-horizon (aka steady-state) W_2-DR-KF, analogous to the steady-state Kalman and $\mathscr{H}_\infty$-filters [101], [120]. We show that the estimation error converges to a steady state (Corollary 10.3.5) with bounded covariance.

3. Efficient real-time implementation: The finite-horizon W_2-DR-KF requires solving an ill-scaled SDP (**??**), rendering it impractical for real-time implementation over long time horizons. However, our infinite-horizon W_2-DR-KF can be implemented efficiently, thanks to our novel rational approximation, thereby overcoming the scalability issues of SDP formulation.

Our contributions are summarized as follows:

1. Tractable convex formulation: We derive an SDP (**??**) formulation for the finite-horizon problem, and a concave-convex max-min optimization problem over positive-definite Toeplitz operators (Theorem 10.3.4) for the infinite horizon one.

2. Optimality of linear estimators for Gaussian nominal: We focus on linear estimators while allowing the distributions in the ambiguity set to be non-Gaussian. For Gaussian nominal distributions, we show the optimality of linear estimators (Theorem 10.3.1).

3. Characterization of the infinite-horizon DR-KF: We derive the infinite-horizon W_2-DR-KF via KKT conditions (Theorem 10.3.4) and show that the transfer function of the infinite-horizon W_2-DR-KF is non-rational, and thereby lacks a finite-order state-space realization. However, we also show that it can be uniquely characterized through a nonlinear finite-dimensional parametrization (Lemma 10.4.1).

4. An efficient algorithm to compute the optimal filter: Using frequency-domain techniques, we introduce an efficient algorithm, based on the Frank-Wolfe method, to compute the optimal infinite-horizon W_2-DR-KF (Algorithm 8). We construct the best rational approximation, in the $\mathscr{H}_\infty$-norm, of any given degree, for the non-rational optimal W_2-DR-KF via a novel convex program (Theorem 10.4.6).

Notations: Bare calligraphic letters ($\mathcal{K}$, $\mathcal{M}$, etc.) are reserved for operators, with the subscripted ones ($\mathcal{K}_T$, $\mathcal{M}_T$, etc.) being finite-dimensional. $\mathcal{I}$ is the identity operator with a suitable block size. Asterisk $\mathcal{M}^*$ denotes the adjoint of $\mathcal{M}$. $\succ$ is the usual positive-definite ordering. $\operatorname{tr}(\cdot)$ is the trace. $\|\cdot\|$ is the usual Euclidean norm. $\|\cdot\|_\infty$ and $\|\cdot\|_2$ are the $\mathscr{H}_\infty$ (operator) and $\mathscr{H}_2$ (Frobenius) norms, respectively. $\{\mathcal{M}\}_+$ and $\{\mathcal{M}\}_-$ denote the causal and strictly anti-causal parts. $\sqrt{\mathcal{M}}$ is the positive-definite symmetric square root. $\mathbb{S}_+^n$ is the set of psd matrices. $|z|$ is the magnitude and z^* is the conjugate of a complex number $z \in \mathbb{C}$. The complex unit circle is denoted by $\mathbb{T}$. $\mathbb{P}$ denotes a probability distribution and $\mathscr{P}_p$ is the set of distributions with finite p^{th} moment. $\mathbb{E}$ denotes the expectation. The Wasserstein-2 distance between distributions $\mathbb{P}_1, \mathbb{P}_2 \in \mathbb{R}^n$ is denoted by $\mathsf{W}_2(\mathbb{P}_1, \mathbb{P}_2)$ such that

$$\mathsf{W}_2(\mathbb{P}_1, \mathbb{P}_2) \triangleq \left(\inf \ \mathbb{E}\left[\|\mathbf{w}_1 - \mathbf{w}_2\|^2\right]\right)^{1/2}, \tag{10.1}$$

where the infimum is over all joint distributions of $(\mathbf{w}_1, \mathbf{w}_2)$ with marginals $\mathbf{w}_1 \sim \mathbb{P}_1$ and $\mathbf{w}_2 \sim \mathbb{P}_2$.

10.2 Problem Setup

In this section, we formulate the distributionally robust filtering problem for infinite horizon setting. To this end, consider the following state-space model:

$$\begin{aligned} x_{t+1} &= Ax_t + Bw_t, \\ y_t &= C_y x_t + v_t, \\ s_t &= C_s x_t, \end{aligned} \tag{10.2}$$

At time $t \in \mathbb{N}$, let $x_t \in \mathbb{R}^{\mathsf{d}_x}$ denote the unobserved *latent state*, $y_t \in \mathbb{R}^{\mathsf{d}_y}$ the *measurement*, $s_t \in \mathbb{R}^{\mathsf{d}_s}$ the unobserved *target signal* to be estimated, $w_t \in \mathbb{R}^{\mathsf{d}_w}$ the *process noise*, and $v_t \in \mathbb{R}^{\mathsf{d}_v}$ the *measurement noise*. The combined process-measurement noise sequence constitutes the *exogenous disturbance*. The setup presented above is quite general and widely adopted in the estimation and filtering literature [101], [120]. The usual state estimation problem is a specific instance of this setup with $C_s = I$. Moreover, we assume that (A, C_y) and (A, C_s) are detectable and (A, B) is controllable.

The Infinite-Horizon Distributionally Robust Filtering

Designing optimal filters for extended horizons can generally be impractical extended time horizons. To mitigate this, time-invariant steady-state filters are usually deployed for practical purposes. These filters can be characterized by their Markov parameters $\{\widehat{K}_t\}$, allowing the estimates $\{\widehat{s}_t\}$ to be computed as a convolution sum: $\widehat{s}_t = \sum_{s=0}^{t} \widehat{K}_{t-s} y_s$. This can be expressed compactly as $\widehat{\mathbf{s}} = \mathcal{K}\mathbf{y}$, where $\mathcal{K}$ is a bounded, causal, and doubly-infinite block Toeplitz operator constructed from the Markov parameters $\widehat{K}_t$. We denote the class of all such filtering policies by $\mathscr{K}$.

Here, $\mathbf{y}$ and $\widehat{\mathbf{s}}$ are the doubly infinite column vectors of measurements and estimates, respectively. Furthermore, letting by $\boldsymbol{\xi} = [\mathbf{w}; \mathbf{v}]$, and $\mathbf{s}$ be doubly-infinite disturbance and target signal vectors, respectively, the state-space dynamics (10.2) over an infinite-horizon can then be described as follows:

$$\begin{aligned} \mathbf{y} &= \mathcal{H}\mathbf{w} + \mathbf{v}, \\ \mathbf{s} &= \mathcal{L}\mathbf{w}, \end{aligned} \tag{10.3}$$

where $\mathcal{H}$ and $\mathcal{L}$ are strictly causal, doubly-infinite, block Toeplitz operators, completely described by the state-space parameters (A, B, C_y, C_s). The error transfer operator $\mathcal{T}_{\mathcal{K}} : \boldsymbol{\xi} \mapsto \mathbf{e} := \widehat{\mathbf{s}} - \mathbf{s}$ under a stationary causal filtering policy $\mathcal{K} \in \mathscr{K}$ is defined similarly as $\mathcal{T}_{\mathcal{K}} := \begin{bmatrix} \mathcal{K}\mathcal{H} - \mathcal{L} & \mathcal{K} \end{bmatrix}$. Note that these Toeplitz operators are equivalently identified by transfer function formalism. In particular, we have $\mathcal{H} \leftrightarrow H(z) := C_y(zI - A)^{-1}B$ and $\mathcal{L} \leftrightarrow L(z) := C_s(zI - A)^{-1}B$ for $z \in \mathbb{T}$.

Instead of focusing on a fixed horizon, we consider the time-averaged steady-state worst-case MSE as the horizon approaches infinity, *i.e.*,

$$\overline{\mathsf{E}}(\mathcal{K}, \rho) \triangleq \limsup_{T\to\infty} \frac{1}{T} \mathsf{E}_T(\mathcal{K}, \rho_T) = \limsup_{T\to\infty} \frac{1}{T} \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T} \left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \mathcal{K})\|^2\right]. \tag{10.4}$$

The limit above may generally be infinite without further specification of the asymptotics of the ambiguity set. To ensure the finiteness of the steady-state MSE, we make the following assumptions:

Assumption 10.2.1. The nominal disturbances $\{(w_t^\circ, v_t^\circ)\}$ form a zero-mean weakly stationary random process, *i.e.*, the cross covariance between (w_t°, v_t°) and (w_t°, v_t°) only depends on the difference $t-s$. Furthermore, the size of the ambiguity set for horizon $T>0$ scales as $\rho_T \sim \rho\sqrt{T}$ for a $\rho>0$.

The assumption on the radius ρ_T for varying T is justified, as the total energy of a random vector of length T from a weakly stationary process scales linearly with T. We state the distributionally robust filtering problem for the infinite horizon as follows:

Problem 10.2.2 (W_2-DR-KF over infinite-horizon)**.** Find a casual and time-invariant filter, $\mathcal{K} \in \mathscr{K}$, that minimizes the steady-state worst-case MSE defined in (10.4), *i.e.*,

$$\inf_{\mathcal{K}\in\mathscr{K}} \overline{\mathsf{E}}(\mathcal{K}, \rho) = \inf_{\mathcal{K}\in\mathscr{K}} \limsup_{T\to\infty} \frac{1}{T} \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T} \left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \mathcal{K})\|^2\right]. \tag{10.5}$$

10.3 Tractable Convex Formulations

In this section, we provide tractable formulations for the finite and infinite-horizon W_2-DR-KF problems. In **??**, we present an SDP formulation for the finite-horizon problem **??**. In Theorem 10.3.4, we reduce the infinite-horizon problem 10.2.2 to a tractable convex program via duality. We also characterize the optimal estimator and the worst-case distribution for both settings. The proofs of the theorems presented in this section are deferred to the Appendix.

Before proceeding with the main theorems, we present a minimax theorem establishing the optimality of linear filtering policies for Gaussian nominal distributions.

Theorem 10.3.1 (Minimax duality)**.** *Let $T > 0$ be a fixed horizon and Π_T be the class of non-linear causal estimators. Suppose that the nominal $\mathbb{P}_T^\circ$ is Gaussian.*

Then, the following holds:

$$\inf_{\pi_T \in \Pi_T} \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right] = \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \inf_{\pi_T \in \Pi_T} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right], \tag{10.6}$$

Moreover, (10.6) *admits a saddle point* $(\pi_T^\star, \mathbb{P}_T^\star)$ *such that the worst-case distribution* $\mathbb{P}_T^\star$ *is Gaussian and the optimal causal filter* $\pi_T^\star$ *is linear, i.e.,* $\pi_T^\star \in \mathscr{K}_T$*.*

For simplicity and clarity, we make the following assumption for the remainder of this paper.

Assumption 10.3.2. The nominal disturbances are uncorrelated, *i.e.*, $\mathbb{E}_{\mathbb{P}_T^\circ}[\boldsymbol{\xi}_T \boldsymbol{\xi}_T^*] = \mathcal{I}_T$ for any $T > 0$.

A Concave-Convex Optimization for the Infinite-Horizon Filtering

The scaling of the SDP in **??** with the time horizon is prohibitive for many time-critical real-world applications. Therefore, we shift our focus to the infinite-horizon W_2-DR-KF problem 10.2.2 to derive the optimal steady-state filtering policy.

Solving Problem 10.2.2 involves two major challenges. The first one is transforming the steady-state worst-case MSE for a fixed filtering policy $\mathcal{K} \in \mathscr{K}$, as defined in (10.4), to an equivalent convex optimization problem. We address this by leveraging the asymptotic convergence properties of Toeplitz matrices [125]. The second challenge is addressing the causality constraint on the estimator. To illustrate the triviality of non-causal estimation in the infinite-horizon setting, we present an analogous result as shown below:

Lemma 10.3.3. *Under the Assumptions 10.2.1 and 10.3.2,* $\mathcal{K}_\circ := \mathcal{L}\mathcal{H}^*(\mathcal{I} + \mathcal{H}\mathcal{H}^*)^{-1}$ *is the unique, optimal, non-causal estimator minimizing the steady-state worst-case MSE in* (10.4) *for any* $\rho > 0$*.*

We address the causality constraint by reformulating Problem 10.2.2 as a max-min optimization, where the inner minimization over the causal filtering policies is performed using the Wiener-Hopf technique [120], [239] (see Lemma 10.C.2). To this end, we introduce the *canonical spectral factorization*[1]

$$\Delta\Delta^* = \mathcal{I} + \mathcal{H}\mathcal{H}^*,$$

where both Δ and its inverse Δ^{-1} are causal operators. We state the equivalent formulation for the infinite-horizon W_2-DR-KF as follows.

[1] Essentially Cholesky factorization for Toeplitz operators.

Theorem 10.3.4 (Convex formulation of infinite-horizon W_2-DR-KF)**.** *Under the Assumptions 10.2.1 and 10.3.2, the Problem 10.2.2 is equivalent to the following feasible max-min problem:*

$$\sup_{\mathcal{M}\succ 0}\inf_{\mathcal{K}\in\mathscr{K}} \operatorname{Tr}(\mathcal{T}_{\mathcal{K}}\mathcal{T}_{\mathcal{K}}^{*}\mathcal{M}) \quad \text{s.t.} \quad \operatorname{Tr}(\mathcal{M}-2\sqrt{\mathcal{M}}+\mathcal{I}) \leq \rho^2. \tag{10.7}$$

Defining $\mathcal{K}_{\mathscr{H}_2} \coloneqq \{\mathcal{K}_{\circ}\Delta\}_{+}\Delta^{-1}$, the unique saddle point $(\mathcal{K}_{\star}, \mathcal{M}_{\star})$ of (10.7) *satisfies the following:*

$$\mathcal{K}_{\star} = \mathcal{K}_{\mathscr{H}_2} + \mathcal{U}_{\star}^{-1}\left\{\mathcal{U}_{\star}\{\mathcal{K}_{\circ}\Delta\}_{-}\right\}_{+}\Delta^{-1}, \tag{10.8a}$$

$$\mathcal{M}_{\star} = (\mathcal{I} - \gamma_{\star}^{-1}\mathcal{T}_{\mathcal{K}_{\star}}\mathcal{T}_{\mathcal{K}_{\star}}^{*})^{-2}, \tag{10.8b}$$

where $\mathcal{U}_{\star}^{*}\mathcal{U}_{\star} = \mathcal{M}_{\star}$ is the canonical spectral factorization with causal $\mathcal{U}_{\star}$ and $\mathcal{U}_{\star}^{-1}$, and $\gamma_{\star} > 0$ is the unique value satisfying the constraint with equality, i.e.,

$$\operatorname{Tr}\left[\left((\mathcal{I} - \gamma_{\star}^{-1}\mathcal{T}_{\mathcal{K}_{\star}}\mathcal{T}_{\mathcal{K}_{\star}}^{*})^{-1} - \mathcal{I}\right)^{2}\right] = \rho^2. \tag{10.9}$$

The optimal linear filter $\mathcal{K}_{\star}$, comprises the nominal Kalman (aka $\mathscr{H}_2$) filter, $\mathcal{K}_{\mathscr{H}_2}$ and an additive correction term that accounts for the correlations within the disturbance process. The correction term is derived directly from the optimal solution $\mathcal{M}_{\star}$ of (10.7) through spectral factorization.

As a result of devising infinite-horizon filters achieving finite optimal value in (10.7), we can deduce the boundedness of the steady-state error covariance.

Corollary 10.3.5. *The steady-state error has bounded covariance under the optimal $\mathcal{K}_{\star}$ in* (10.8).

10.4 An Efficient Algorithm

While the standard Kalman and $\mathscr{H}_\infty$-filters allow for finite-order spate-space realizations derived via algebraic methods, the optimal $\mathcal{K}_{\star}$ lacks such a realization since its transfer function $K_{\star}(z)$ is non-rational (Corollary 10.4.2) despite admitting a non-linear finite-dimensional parametrization (Lemma 10.4.1). Thus, we adopt a novel twofold approach to develop practical DR filters:

1. We introduce an efficient algorithm to compute the optimal positive-definite operator $\mathcal{M}_{\star}$ from (10.7). To address the challanges posed by its infinite-dimensional nature, we use the frequency-domain representation of $\mathcal{M}_{\star}$ as the power spectral density $M_{\star}(z)\succ 0$.

2. We develop a novel method to approximate the non-rational power spectral density $M_\star(z)$ in $\mathscr{H}_\infty$-norm using positive rational functions through convex optimization. This rational approximation is then used to derive an approximate rational filter with state-space realization via (10.8).

To this end, we adopt the transfer-function formalism for the rest of this paper with the correspondences: $\mathcal{M} \leftrightarrow M(z), \mathcal{U} \leftrightarrow U(z), \mathcal{K} \leftrightarrow K(z)$, and $\mathcal{T}_{\mathcal{K}} \leftrightarrow T_K(z)$ for $z \in \mathbb{T}$. The following lemma characterizes the optimal $M_\star(z)$, implying finite-dimensional parametrization.

Lemma 10.4.1. *Let $f : (\gamma, \Gamma) \mapsto \mathcal{M}$ return the unique solution of the implicit equation over $M(z)$,*

$$M(z) = \gamma^2[\gamma I - U(z)^{-1}\Gamma(I - z\overline{A})^{-1}\overline{B}\,\overline{B}^*(I - z\overline{A})^{-*}\Gamma^* U(z)^{-*} + T_{K_\circ}(z)T_{K_\circ}(z)^*]^{-2}, \forall z \in \mathbb{T} \tag{10.10}$$

where $U(z)^*U(z) = M(z)$ is the unique spectral factorization and $(\overline{A}, \overline{B}, \overline{C})$ are obtained from state-space parameters (see Section 10.D and (10.106)*). We have that $\mathcal{M}_\star = f(\gamma_\star, \Gamma_\star)$ where*

$$\Gamma_\star = \frac{1}{2\pi}\int_{-\pi}^{\pi} U_\star(\mathrm{e}^{\jmath\omega})\overline{C}(I - \mathrm{e}^{\jmath\omega}\overline{A})^{-1} d\omega, \tag{10.11}$$

and $\gamma_\star > 0$ is such that $\mathrm{Tr}(\mathcal{M}_\star - 2\sqrt{\mathcal{M}_\star} + \mathcal{I}) = \rho^2$,

As a consequence of Lemma 10.4.1, we deduce the non-rationally of the optimal W_2-DR-KF.

Corollary 10.4.2. *The spectral density $M_\star(z)$ and the transfer function $K_\star(z)$ are non-rational.*

Iterative Optimization Methods in the Frequency-Domain

Despite being a concave program, the infinite-dimensional nature of (10.7) hinders the direct application of standard optimization tools. To address this, we leverage frequency-domain analysis via transfer functions, enabling the use of standard tools with appropriate modifications. Specifically, we employ the modification of a Frank-Wolfe method [68], [116]. Our framework is versatile and can be extended to alternative approaches, including projected gradient descent [83], and the fixed-point iteration method used in [125]. A detailed pseudocode Algorithm 8 is provided in Section 10.D.

Frank-Wolfe: We define the following function and its (Gateaux) gradient [43]:

$$\Phi(\mathcal{M}) \triangleq \inf_{\mathcal{K}\in\mathscr{K}} \operatorname{Tr}\left(\mathcal{T}_{\mathcal{K}}\mathcal{T}_{\mathcal{K}}^{*}\mathcal{M}\right), \quad \text{and} \quad \nabla\Phi(\mathcal{M}) = \mathcal{U}^{-1}\left\{\mathcal{U}\mathcal{K}_{\circ}\Delta\right\}_{-}\left\{\mathcal{U}\mathcal{K}_{\circ}\Delta\right\}_{-}^{*}\mathcal{U}^{-*} + \mathcal{T}_{\mathcal{K}_{\circ}}\mathcal{T}_{\mathcal{K}_{\circ}}^{*}. \tag{10.12}$$

where $\mathcal{U}^{*}\mathcal{U} = \mathcal{M}$ is the spectral factorization. Rather than solving the optimization in (10.7) directly, the Frank-Wolfe method solves a linearized subproblem in consecutive steps. Namely, given the k^{th} iterate $\mathcal{M}_k$, the next iterate $\mathcal{M}_{k+1}$ is obtained via

$$\widetilde{\mathcal{M}}_k = \underset{\mathcal{M}\succcurlyeq I}{\arg\max}\ \operatorname{Tr}\left(\nabla\Phi(\mathcal{M}_k)\,\mathcal{M}\right) \quad \text{s.t.} \quad \operatorname{Tr}(\mathcal{M} - 2\sqrt{\mathcal{M}} + \mathcal{I}) \leq \rho^2, \tag{10.13a}$$

$$\mathcal{M}_{k+1} = (1-\eta_k)\mathcal{M}_k + \eta_k\widetilde{\mathcal{M}}_k, \tag{10.13b}$$

where $\eta_k \in [0,1]$ is a step-size, commonly set as $\eta_k = \frac{2}{k+2}$ [116]. Letting $\mathcal{G}_k \coloneqq \nabla\Phi(\mathcal{M}_k)$ be the gradient as in (10.12), Frank-Wolfe updates can be expressed equivalently using spectral densities as:

$$\widetilde{M}_k(z) = (I - \gamma_k^{-1}G_k(z))^{-2} \quad \text{and} \quad M_{k+1}(z) = (1-\eta_k)M_k(z) + \eta_k\widetilde{M}_k(z), \quad \forall z \in \mathbb{T} \tag{10.14}$$

where $\gamma_k > 0$ solves $\operatorname{Tr}\left[((I-\gamma_k^{-1}\mathcal{G}_k)^{-1} - I)^2\right] = \rho^2$. See Section 10.D for a closed-form $G_k(z)$.

Discretization: Instead of the continuous domain unit circle $\mathbb{T}$, we consider its uniform discretization by N points, $\mathbb{T}_N \coloneqq \{e^{j2\pi n/N} \mid n = 0,\ldots,N-1\}$. While the gradient update $G_k(z)$ for frequency z is applied to the next iterate $M_{k+1}(z)$ at that frequency, calculating $G_k(z)$ requires $M_k(z')$ at all other frequencies $z' \in \mathbb{T}$ due to spectral factorization involved. Thus, the full update for $M_{k+1}(z)$ needs $M_k(z')$ across the entire unit circle. This is overcome by finer discretization.

Spectral Factorization: Since the iterates $M_k(z)$ are non-rational spectral densities, the spectral factorization can only be performed approximately [199]. Specifically, we employ the algorithm proposed in [188] that uses discrete Fourier transform (DFT) and is based on Kolmogorov's method of factorization [132]. This method, tailored for scalar spectral densities (*i.e.*, for scalar target signals $\mathrm{d}_s = 1$), proves efficient as the associated error term, featuring a multiplicative phase factor, rapidly diminishes with finer discretization N. Matrix-valued spectral densities can also be tackled by various other algorithms [58], [243]. See Section 10.D for a pseudocode and details.

Bisection: We use bisection method to find the $\gamma_k > 0$ that solves $\operatorname{Tr}\left[((I-\gamma_k^{-1}\mathcal{G}_k)^{-1} - I)^2\right] = \rho^2$ in the Frank-Wolfe update (10.14). See Section 10.D for a pseudocode and further details.

Remark 10.4.3. The gradient $G_k(z)$ requires computation of the finite-dimensional parameter via (10.11), which can be performed using N-point trapezoidal integration. See Section 10.D for details.

We conclude this section with the following convergence result due to [116], [135].

Theorem 10.4.4 (Convergence of $\mathcal{M}_k$)**.** *There exists constants $\delta_N > 0$, depending on discretization N, and $\kappa > 0$, depending only on state-space parameters* (10.2) *and ρ, such that the iterates in* (10.13) *satisfy*

$$\Phi(\mathcal{M}_\star) - \Phi(\mathcal{M}_k) \leq \frac{2\kappa}{k+2}(1+\delta_N). \tag{10.15}$$

Rational Approximation using $\mathscr{H}_\infty$-norm

In the preceding section, we introduced a method to compute the optimal $M_\star(z)$ approximately on the unit circle. However, the resulting filtering policy is non-rational and cannot be realized as a state-space filter. In this section, we introduce a novel technique for obtaining approximate rational filtering policies. Instead of directly approximating the filter itself, our method involves an initial step of *approximating the power spectrum* $M_\star(z)$ by a ratio of positive fixed order polynomials, $P(z)/Q(z)$, to minimize the $\mathscr{H}_\infty$-norm of the approximation error. After finding a rational approximation $P(z)/Q(z)$ of $M_\star(z)$, we compute a state-space controller according to Equation (10.8). For simplicity, we focus on scalar target signals, namely, $\mathsf{d}_s = 1$.

Concretely, $P(z) = \sum_{k=-m}^{m} p_k z^{-k}$ and $Q(z) = \sum_{k=-m}^{m} q_k z^{-k}$ are Laurent polynomials of degree $m \in \mathbb{N}$ with symmetric coefficients $p_k = p_{-k} \in \mathbb{R}$ and $q_k = q_{-k} \in \mathbb{R}$. In other words, the polynomials $P(z)$ and $Q(z)$ are uniquely identified by $m+1$ real coefficients, $(p_0, \ldots, p_m)$ and $(q_0, \ldots, q_m)$. Given a positive spectral density $M(z) > 0$ for $z \in \mathbb{T}$, we seek positive polynomials $P(z), Q(z) > 0$ for $z \in \mathbb{T}$ of order at most $m \in \mathbb{N}$ that minimize the $\mathscr{H}_\infty$-norm of the rational approximation error, *i.e.*,

$$\min_{\substack{p_0,\ldots p_m \in \mathbb{R},\\ q_0,\ldots q_m \in \mathbb{R},\\ \varepsilon \geq 0}} \varepsilon \quad \text{s.t.} \quad \begin{array}{l} \textit{i)}\ P(z), Q(z) > 0 \text{ for all } z \in \mathbb{T}, \\ \textit{ii)}\ q_0 = 1, \\ \textit{iii)}\ \max\limits_{z \in \mathbb{T}} \left| \dfrac{P(z)}{Q(z)} - M(z) \right| \leq \varepsilon, \end{array} \tag{10.16}$$

where $\varepsilon \geq 0$ denotes an *upper bound on the approximation error*. The constraint $q_0 = 1$ eliminates redundancy in the problem since the fraction $P(z)/Q(z)$ is scale invariant. Unfortunately, the problem (10.16) is not convex in all the variables. Instead, Lemma 10.4.5 shows convexity for fixed $\varepsilon \geq 0$.

Lemma 10.4.5. *For a fixed $\varepsilon \geq 0$, the constraints (i-iii) define a jointly convex set for the coefficients.*

Proof. The constraints *(i-ii)* are affine inequalities, hence convex. Constraint *(iii)* is equivalent to

$$P(z) - (M(z)+\varepsilon)Q(z) \leq 0, \text{ and } P(z) - (M(z)-\varepsilon)Q(z) \geq 0, \text{ for all } z \in \mathbb{T}, \tag{10.17}$$

which are jointly affine inequalities in $(p_0, \ldots p_m)$ and $(q_0, \ldots q_m)$, hence convex. ■

This result enables us to obtain m^{th}-order rational approximations $P(z)/Q(z)$ of $M(z)$ with a fixed approximation precision ε, signifying our tolerance for deviations from $M(z)$, by solving a *convex feasibility problem*. Notice that the constraints *(i)* and *(iii)* (eqv. (10.17)) involve inequalities over the entire unit circle $\mathbb{T}$. Since the iterative method in Algorithm 8 only returns the values of $M(z)$ on the discretized unit circle $\mathbb{T}_N$, we can enforce these inequalities in the feasibility problem only for $\mathbb{T}_N$. While being an inexact approximation for *(iii)*, it is an exact characterization for *(i)* as long as $N > 2m$ by the Nyquist-Shannon sampling theorem [205]. See Section 10.D for a pseudocode.

Utilizing a convex feasibility oracle, our method can be used in two operational modes:

1. **Fixed order, best precision:** By iteratively reducing the precision ε we can revise the ε-feasible polynomials $P(z)$, $Q(z)$, effectively solving the non-convex problem (10.16) to obtain the best m^{th}-order rational approximation.

2. **Fixed precision, least order:** In contrast, we can seek the lowest degree rational approximation, which achieves a fixed precision ε.

Theorem 10.4.6. *The spectral factorization $U(z)^* U(z) = P(z)/Q(z)$ of a degree m rational approximation $P(z)/Q(z)$ admits a rational factor $U(z)$. Furthermore, the filter obtained from $U(z)$ using* (10.8)*, i.e., $K(z) = K_{\mathscr{H}_2}(z) + U(z)^{-1}\left\{U(z)\{K_\circ(z)\Delta(z)\}_-\right\}_+ \Delta(z)^{-1}$ is rational and can be realized as a state-space filter as highlighted below:*

$$\begin{aligned} \zeta_{t+1} &= \widetilde{F}\zeta_t + \widetilde{G}y_t, \\ \widehat{s}_t &= \widetilde{H}\zeta_t + \widetilde{L}y_t, \end{aligned} \tag{10.18}$$

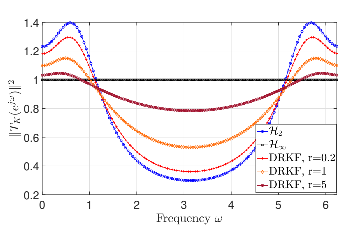
where $\zeta_t \in \mathbb{R}^{m+\mathsf{d}_x}$ is the filter state, and $(\widetilde{F}, \widetilde{G}, \widetilde{H}, \widetilde{L})$ are determined from (A, B, C_y, C_u) and $U(z)$.

10.5 Numerical Experiments

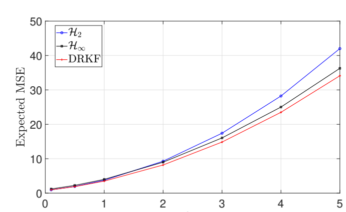
In this section, we compare the performance of finite and infinite horizon DR-KF filters with H_2, H_∞ filters, and other DRKFs [203], [153]. Our evaluation includes both frequency-domain and time-domain analyses, highlighting the effectiveness of the rational approximation method. The nominal distribution is assumed to be Gaussian with zero mean and identity covariance. Our results demonstrate that our DR-KF (in the finite and ∞ horizon) provides significant advantages over other DRKFs regarding stability, computational speed, and error reduction. The experiments were performed on a M1 Macbook Air with 8 GB of RAM.

Frequency Domain Evaluations

We study a typical tracking problem whose state-space model is $A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & \Delta t \end{bmatrix}^T$,$C_y = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $C_s = 1$ where the state corresponds to the position and velocity, the process noise is the exogenous acceleration, and Δt is the sampling time. We plot the frequency response of our DR-KF using the metric $|T_K(e^{j\omega})|^2 = \sigma(T_K^*(e^{j\omega})T_K(e^{j\omega}))$, where σ is the maximal singular value. We compare it to the classical H_2 (KF) and H_∞ (robust) filters. Figure 10.1a shows that the DR-KF interpolates well between the H_2 (KF) and H_∞ (robust) filters. Figure 10.1b illustrates the worst-case expected MSE. For smaller r, the DR-KF performs similarly to the $\mathcal{H}_2$ (KF) filter, while for larger r, its worst-case MSE approaches that of the robust filter. Overall, the DR-KF achieves the lowest worst-case expected MSE for any r. We investigate the behavior of the rational approximation across various values of the radius r. The results for degrees $m = 1, 2, 3$ are given in Table 10.1 . Approximations of order greater than 2 achieve an expected MSE closely matching the non-rational DR-KF for all values of r.



(a) Frequency response for the tracking problem.



(b) Worst-case expected MSE.

Figure 10.1: DR-KF versus the $\mathcal{H}_2, \mathcal{H}_\infty$ filters and the variation of the expected MSE with r.

	r=0.01	r=1	r=3	r=5
$DRKF$	0.7870	3.4948	14.842	34.110
RA(1)	0.7871	3.5818	15.954	38.327
RA(2)	0.7870	3.4948	14.844	34.124
RA(3)	0.7870	3.4948	14.834	34.024

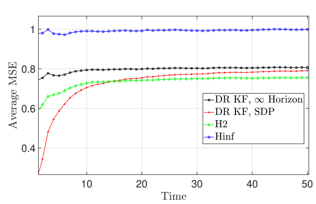
Table 10.1: The worst-case expected MSE of the non-rational DRKF, compared to the rational filters RA(1), RA(2), and RA(3), obtained from degree 1, 2, and 3 rational approximations to $U(e^{j\omega})$, for the system in Section 10.5.

	T=10	T=50	T=100	T=1000
DRMC	32.9 s	NAN	NAN	NAN
Our DRKF (finite)	0.65 s	7.3 s	194.9 s	NAN
Our DRKF (infinite)	6.6 s	6.6 s	6.6 s	6.6 s

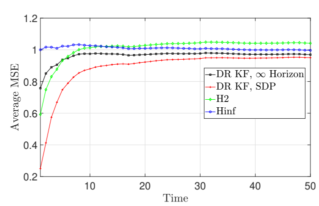
Table 10.2: The running time (in seconds) of different filters for the system in section 10.5. The DRMC is inefficient for T> 10, our DRKF (finite) is inefficient for T> 50 while our ∞ horizon DRKF can run for any horizon.

Time Domain Evaluations

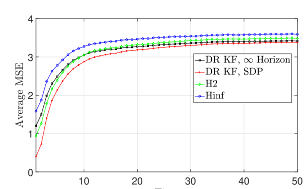
We assess the time-domain performance of both infinite and finite horizon DR-KF filters, comparing them with H_2 and H_∞ counterparts on the tracking problem introduced in Section 10.5. The average MSE over 50 time steps, aggregated across 1000 independent trials, is plotted. In Figure 10.2a, under white Gaussian noise, the H_2 (KF) filter outperforms others. Figures 10.2b and 10.2c correspond to correlated Gaussian noise and the worst-case noise for the finite horizon DRKF, respectively. In Figures 10.2b and 10.2c, the DRKF outperforms the classical filters, and the infinite horizon DRKF matches the finite horizon one. As we increase the time horizon, solving the finite horizon SDP becomes computationally infeasible, underscoring the advantage of the infinite horizon DRKF.



(a) White noise



(b) Correlated Gaussian noise

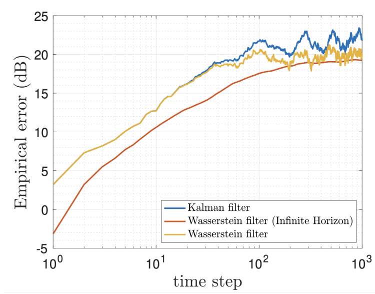


(c) Worst noise for DR-KF, SDP

Figure 10.2: The average MSE of the different filters for the tracking problem, under (a) white noise, (b) correlated Gaussian noise, and (c) worst-case noise for the finite horizon DR KF for the system in Section 10.5. While the H_2 filter (KF) performs best in (a), it behaves poorly in (b), (c). The DRKF achieves the lowest error in (b) and (c), and the finite and infinite horizon achieve similar average MSE at the end of the horizon.

Comparison to the DRKF in [203]

We first compare against [203] which assumes the states and measurements to be in a Wasserstein neighborhood around a nominal at each time step, robustifying immediately against model uncertainties. Authors in [203] don't consider noise correlations across time steps, their problem setup is in the finite-horizon, and they use the Frank-Wolfe algorithm to efficiently solve the problem. The system matrices that they consider is given by $A = \begin{bmatrix} 0.9802 & 0.0196 + 0.099\Delta \\ 0 & 0.9802 \end{bmatrix}$, $Q = \begin{bmatrix} 1.9608 & 0.0195 \\ 0.0195 & 1.9605 \end{bmatrix}$, $B = \begin{bmatrix} \sqrt{Q} & 0_{2\times 1} \end{bmatrix}$, $C_y = \begin{bmatrix} 1 & -1 \end{bmatrix}$, $C_s = I$, and Δ represents a scalar uncertainty (taken to be 1 as in [203]). We compare the performance of our infinite-horizon DRKF to [203] in Figure 10.3 under Gaussian noise. The plot shows that our DRKF outperforms [203] and has a more stable performance, even though we are disadvantaged in two ways: 1) our filter isn't explicitely designed for model uncertainties, 2) since we only consider estimations of linear combination of the state (C_s is a row vector), we get the total MSE from 2 different runs with $C_s = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $C_s = \begin{bmatrix} 0 & 1 \end{bmatrix}$, which is suboptimal.



(a) Average MSE in dB under Gaussian noise.

Figure 10.3: Average MSE for the KF, our DRKF, and the DRKF from [203], for system in section 10.5.

Comparison to the DR estimator of [153]

We contrast our approach with that of [153], termed linear quadratic estimator under martingale constraints (DRMC). Here are the key comparisons: 1) DRMC, akin to our approach, considers noise within a Wasserstein neighborhood around a baseline, allowing for correlations between process and measurement noise (achieved through a martingale sequence constraint). 2) DRMC assumes the process noise is sampled from the baseline and doesn't lie in the Wasserstein ball, a more restrictive assumption compared to ours. 3) DRMC's problem formulation is in the finite-horizon, claiming to have an efficient converging method to solve it. With a horizon of $T = 10$, they test their approach on a simple 1D system ($A = B = C_y = C_s = 1$) , which we also use for comparison. For $r = 0.2\sqrt{T}$ and under the worst-case noise for our finite-horizon

DRKF, the average MSE for DRMC is 0.86, closely matching our finite-horizon DRKF at 0.86 and our infinite-horizon DRKF at 0.88 at $T = 10$. For the same $r =$ under the worst-case noise of DRMC, the average MSE for DRMC is 0.78, close to our DRKF (0.81 for the finite-horizon and 0.83 for the infinite-horizon at $T = 10$). This shows that using the infinite-horizon controller for short horizons does not significantly compromise performance. Similar results are observed for other values of r. While the performances in this simple example are comparable, our filter is anticipated to excel for higher-diemsnional systems, due to its explicit consideration of robustness over process noise. However, our DRKFs outshine DRMC in efficiency. DRMC takes 32.9 seconds for $T = 10$, and becomes computationally infeasible beyond that. Our finite-horizon DRKF is faster and efficient up to $T = 50$, and our infinite-horizon DRKF remains unaffected by the time horizon. For details, see Table 10.2.

10.6 Conclusion

The main limitation in our work is that our H_∞-rational approximation method is limited to scalar target signals (*i.e.*, C_s is a row vector). Future work will address this limitation.

10.A Additional Discussion on the Problem Setup

Explicit Form of Finite-Horizon Model in (**??**)

The causal linear measurement model for the finite-horizon case in (**??**) can be stated explicitly as follows:

$$\underbrace{\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} C_y & 0 & 0 & \dots & 0 \\ C_yA & C_yB & 0 & \dots & 0 \\ C_yA^2 & C_yAB & C_yB & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_yA^T & C_yA^{T-1}B & C_yA^{T-2}B & \ddots & C_yB \end{bmatrix}}_{\mathcal{H}_T} \underbrace{\begin{bmatrix} x_0 \\ w_0 \\ w_1 \\ \vdots \\ w_{T-1} \end{bmatrix}}_{\mathbf{w}} + \underbrace{\begin{bmatrix} v_0 \\ v_0 \\ v_1 \\ \vdots \\ v_T \end{bmatrix}}_{\mathbf{v}} \tag{10.19a}$$

$$\underbrace{\begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_T \end{bmatrix}}_{\mathbf{s}} = \underbrace{\begin{bmatrix} C_s & 0 & 0 & \dots & 0 \\ C_sA & C_sB & 0 & \dots & 0 \\ C_sA^2 & C_sAB & C_sB & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_sA^T & C_sA^{T-1}B & C_sA^{T-2}B & \ddots & C_sB \end{bmatrix}}_{\mathcal{L}_T} \underbrace{\begin{bmatrix} x_0 \\ w_0 \\ w_1 \\ \vdots \\ w_{T-1} \end{bmatrix}}_{\mathbf{w}} \tag{10.19b}$$

A similar construction of $\mathcal{H}_T$ and $\mathcal{L}_T$ for time-varying systems can be performed by replacing the causal block elements of $\mathcal{H}_T$ and $\mathcal{L}_T$ with appropriate coefficients derived from the time-varying dynamics.

10.B Proofs of Theorems Related to Finite-Horizon Filtering

Proof of Theorem 10.3.1

Before we proceed with the proof, we first state the following useful deifnitions and results.

Definition 10.B.1 (Bures-Wasserstein distance [20])**.** For any two psd matrices $\Sigma_1, \Sigma_2 \in \mathbb{S}_+^d$, the Bures-Wasserstein distance between them is defined as follows:

$$\mathsf{BW}(\Sigma_1, \Sigma_2) \triangleq \sqrt{\mathrm{Tr}\left[\Sigma_1 + \Sigma_2 - 2\left(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2}\right)^{1/2}\right]}. \tag{10.20}$$

Definition 10.B.2 (Gelbrich distance [78])**.** For any two distributions $\mathbb{P}_1, \mathbb{P}_2 \in \mathscr{P}(\mathbb{R}^d)$ with means $\mu_1, \mu_2 \in \mathbb{R}^d$ and covariances $\Sigma_1, \Sigma_2 \in \mathbb{S}_+^d$, respectively, the Gelbrich distance between them is defined as follows:

$$\mathsf{G}(\mathbb{P}_1, \mathbb{P}_2) \triangleq \sqrt{\|\mu_1 - \mu_2\|^2 + \mathsf{BW}(\Sigma_1, \Sigma_2)^2}. \tag{10.21}$$

Lemma 10.B.3 (Gelbrich bound [78, Thm. 2.1])**.** *Consider two distributions $\mathbb{P}_1, \mathbb{P}_2 \in \mathscr{P}(\mathbb{R}^d)$ with means $\mu_1, \mu_2 \in \mathbb{R}^d$ and covariances $\Sigma_1, \Sigma_2 \in \mathbb{S}_+^d$, respectively. The W_2-distance between them satisfies*

$$\mathsf{W}_2(\mathbb{P}_1, \mathbb{P}_2) \geq \mathsf{G}(\mathbb{P}_1, \mathbb{P}_2), \tag{10.22}$$

where equality is attained if both $\mathbb{P}_1$ and $\mathbb{P}_2$ are Gaussian distributions.

Lemma 10.B.4 (Causal MMSE of Gaussian [120])**.** *Suppose the disturbances are distributed as Gaussian, i.e., $\boldsymbol{\xi}_T \sim \mathcal{N}(\boldsymbol{\mu}_T, \boldsymbol{\Sigma}_T)$ with mean $\boldsymbol{\mu}_T \in \Xi_T$ and covariance $\boldsymbol{\Sigma}_T$. Consider causal mean-square estimation of $\mathbf{s}_T$ from $\mathbf{y}_t$, i.e.,*

$$\inf_{\pi_T \in \Pi_T} \mathbb{E}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right]. \tag{10.23}$$

Then, there exists a causal (block lower-diagonal) matrix $\mathcal{K}_T^\star$ and a vector $\mathbf{b}_T^\star$, such that the optimal causal estimator $\pi_T^\star : \mathbf{y}_T \mapsto \widehat{\mathbf{s}}_T$ is affine with the following form:

$$\widehat{\mathbf{s}}_T = \mathcal{K}_T^\star \mathbf{y}_t + \mathbf{b}_T^\star. \tag{10.24}$$

Proof of Theorem 10.3.1: Clearly, we have the following weak duality,

$$\sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \inf_{\pi_T \in \Pi_T} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right] \leq \inf_{\pi_T \in \Pi_T} \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right]. \tag{10.25}$$

Let $\boldsymbol{\Sigma}_T^\circ \succ 0$ be the covariance of the nominal distribution $\mathbb{P}_T^\circ$. We start by bounding the lhs of (10.25) as follows:

$$\sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \inf_{\pi_T \in \Pi_T} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right] \overset{(a)}{\leq} \sup_{\mathsf{G}(\mathbb{P}_T, \mathbb{P}_T^\circ) \leq \rho_T} \inf_{\pi_T \in \Pi_T} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right], \tag{10.26}$$

$$\overset{(b)}{\leq} \sup_{\mathsf{G}(\mathbb{P}_T, \mathbb{P}_T^\circ) \leq \rho_T} \inf_{\mathcal{K}_T \in \mathscr{K}_T} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \mathcal{K}_T)\|^2\right], \tag{10.27}$$

$$= \sup_{\mathsf{G}(\mathbb{P}_T, \mathbb{P}_T^\circ) \leq \rho_T} \inf_{\mathcal{K}_T \in \mathscr{K}_T} \mathbb{E}_{\mathbb{P}_T}\left[\boldsymbol{\xi}_T^* \mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T} \boldsymbol{\xi}_T\right], \tag{10.28}$$

$$\overset{(c)}{=} \sup_{\mathsf{G}(\mathbb{P}_T, \mathbb{P}_T^\circ) \leq \rho_T} \inf_{\mathcal{K}_T \in \mathscr{K}_T} \operatorname{Tr}\left(\mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T} \mathbb{E}_{\mathbb{P}_T}\left[\boldsymbol{\xi}_T \boldsymbol{\xi}_T^*\right]\right), \tag{10.29}$$

$$\overset{(d)}{=} \sup_{\mathsf{BW}(\boldsymbol{\Sigma}_T, \boldsymbol{\Sigma}_T^\circ) \leq \rho_T} \inf_{\mathcal{K}_T \in \mathscr{K}_T} \operatorname{Tr}\left(\mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T} \boldsymbol{\Sigma}_T\right), \tag{10.30}$$

where (a) follows from the Gelbrich bound (Lemma 10.B.3), (b) follows from $\mathscr{K}_T \subset \Pi_T$, (c) follows from linearity of cyclic property of trace and the linearity of trace and the expectation, (d) follows from the definition of the Gelbrich distance (Definition 10.B.2). Note that, we can in general take the distributions involved to be zero-mean since any non-zero mean can be incorporated as an additive constant to the estimator, canceling the mean. Therefore, without loss of generality, we can restrict ourselves to zero-mean disturbances and linear estimators (instead of affine).

Following a similar reasoning, we obtain the following upper bound on the rhs of (10.25),

$$\inf_{\pi_T \in \Pi_T} \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \pi_T)\|^2\right] \leq \inf_{\mathcal{K}_T \in \mathscr{K}_T} \sup_{\mathsf{BW}(\boldsymbol{\Sigma}_T, \boldsymbol{\Sigma}_T^\circ) \leq \rho_T} \operatorname{Tr}\left(\mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T} \boldsymbol{\Sigma}_T\right). \tag{10.31}$$

Notice that the objective in the right-hand side of (10.31) is affine in $\boldsymbol{\Sigma}_T$ (hence concave) and quadratic in $\mathcal{K}_T$ (hence strictly convex whenever $\boldsymbol{\Sigma}_T \succ 0$). Furthermore, the constraint set $\mathscr{K}_T$ is affine, and the constraint $\mathsf{BW}(\boldsymbol{\Sigma}_T, \boldsymbol{\Sigma}_T^\circ)$ is convex [20]. Therefore, we have the following minimax duality.

$$\inf_{\mathcal{K}_T \in \mathscr{K}_T} \sup_{\mathsf{BW}(\boldsymbol{\Sigma}_T, \boldsymbol{\Sigma}_T^\circ) \leq \rho_T} \operatorname{Tr}\left(\mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T} \boldsymbol{\Sigma}_T\right) = \sup_{\mathsf{BW}(\boldsymbol{\Sigma}_T, \boldsymbol{\Sigma}_T^\circ) \leq \rho_T} \inf_{\mathcal{K}_T \in \mathscr{K}_T} \operatorname{Tr}\left(\mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T} \boldsymbol{\Sigma}_T\right). \tag{10.32}$$

We denote the saddle point of (10.32) by $(\mathcal{K}_T^\star, \mathbf{\Sigma}_T^\star)$. Notice that, when the nominal distribution is Gaussian $\mathbb{P}_T^\circ := \mathcal{N}(0, \mathbf{\Sigma}_T^\circ)$, the Gaussian distribution $\mathbb{P}_T^\star := \mathcal{N}(0, \mathbf{\Sigma}_T^\star)$ and the causal estimator $\pi_T^\star := \mathcal{K}_T^\star$ achieve the upper bound (10.30) with equality. Thus, from (10.25) and (10.31), we obtain the desired result. ■

Proof of ??

Before we proceed with the proof, we state the following result, which is the backbone for both **??** and Theorem 10.3.4.

Theorem 10.B.5 (Strong Duality in the Finite Horizon)**.** *Let the horizon $T > 0$ be fixed and $\mathcal{K}_T$ be a given estimator, which can be non-causal in general. Under the Assumption 10.3.2, the finite-horizon worst-case MSE* (**??**) *suffered by $\mathcal{K}_T$, i.e.,*

$$\mathsf{E}_T(\mathcal{K}_T, \rho_T) \;=\; \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T}\left[\boldsymbol{\xi}_T^* \mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T} \boldsymbol{\xi}_T\right], \tag{10.33}$$

attains a finite value and is equivalent to the following dual problem:

$$\mathsf{E}_T(\mathcal{K}_T, \rho_T) \;=\; \inf_{\gamma \geq 0} \gamma \rho_T^2 + \gamma \operatorname{Tr}\left[(\mathcal{I}_T - \gamma^{-1} \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*)^{-1} - \mathcal{I}_T\right] \quad \textit{s.t.} \quad \gamma \mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*. \tag{10.34}$$

Furthermore, the worst-case disturbance, $\boldsymbol{\xi}_T^\star$, can be identified from the nominal disturbance, $\boldsymbol{\xi}_T^\circ$, as

$$\boldsymbol{\xi}_T^\star = (\mathcal{I}_T - \gamma_\star^{-1} \mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T})^{-1} \boldsymbol{\xi}_T^\circ, \tag{10.35}$$

where $\gamma_\star$ is the optimal solution of (10.34) *and solves the following equation uniquely:*

$$\operatorname{Tr}\left[((\mathcal{I}_T - \gamma_\star^{-1} \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*)^{-1} - \mathcal{I}_T)^2\right] = \rho_T^2. \tag{10.36}$$

Proof. The proof follows closely from [219, Thm. 2 & 3] (and also from [89, Thm. IV.1]) by replacing the matrix C in Thm.2 of [219] with $\mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T}$. In that case, we get that

$$\mathsf{E}_T(\mathcal{K}_T, \rho_T) \;=\; \inf_{\gamma \geq 0} \gamma \rho_T^2 + \gamma \operatorname{Tr}\left[(\mathcal{I}_T - \gamma^{-1} \mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T})^{-1} - \mathcal{I}_T\right] \quad \text{s.t.} \quad \gamma \mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T}, \tag{10.37}$$

and the characterization of the optimal solution $\gamma_\star \geq 0$ as

$$\operatorname{Tr}\left[((\mathcal{I}_T - \gamma_\star^{-1} \mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T})^{-1} - \mathcal{I}_T)^2\right] = \rho_T^2. \tag{10.38}$$

Notice that (10.37) and (10.38) involve the term $\mathcal{T}_{\mathcal{K}_T}^* \mathcal{T}_{\mathcal{K}_T}$ whereas the desired formulations in (10.34) and (10.36) involve the term $\mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*$. To obtain the desired formulations, we appeal to matrix inversion identity, *i.e.*,

$$(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}^*\mathcal{T}_{\mathcal{K}_T})^{-1} = \mathcal{I}_T + \mathcal{T}_{\mathcal{K}_T}^*(\gamma\mathcal{I}_T - \mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}\mathcal{T}_{\mathcal{K}_T}, \tag{10.39}$$

$$= \mathcal{I}_T + \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}^*(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}\mathcal{T}_{\mathcal{K}_T}, \tag{10.40}$$

where the exact block dimensions of the identity operator $\mathcal{I}_T$ differ depending on where they appear and should be inferred from the context. We can evaluate the trace in (10.37) involving $\mathcal{T}_{\mathcal{K}_T}^*\mathcal{T}_{\mathcal{K}_T}$ as

$$\mathrm{Tr}\left[(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}^*\mathcal{T}_{\mathcal{K}_T})^{-1} - \mathcal{I}_T\right] = \mathrm{Tr}\left[\mathcal{I}_T + \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}^*(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}\mathcal{T}_{\mathcal{K}_T} - \mathcal{I}_T\right], \tag{10.41}$$

$$= \mathrm{Tr}\left[\gamma^{-1}\mathcal{T}_{\mathcal{K}_T}^*(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}\mathcal{T}_{\mathcal{K}_T}\right], \tag{10.42}$$

$$= \mathrm{Tr}\left[(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}\gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*\right], \tag{10.43}$$

where (10.41) is by (10.40), and (10.43) is by the cyclic property of trace. Noting that the condition $\gamma\mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T}^*\mathcal{T}_{\mathcal{K}_T}$ is equivalent to $\gamma\mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*$, we expand $(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}$ in (10.43) by the following Neumann series:

$$(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1} = \sum_{k=0}^{\infty}\left(\gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*\right)^k. \tag{10.44}$$

Thus, the expression in (10.43) can be written equivalently as

$$\mathrm{Tr}\left[(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}\gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*\right] = \mathrm{Tr}\left[\sum_{k=0}^{\infty}\left(\gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*\right)^{k+1}\right], \tag{10.45}$$

$$= \mathrm{Tr}\left[\sum_{k=1}^{\infty}\left(\gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*\right)^{k}\right], \tag{10.46}$$

$$= \mathrm{Tr}\left[(\mathcal{I}_T - \gamma^{-1}\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1} - \mathcal{I}_T\right], \tag{10.47}$$

giving the desired expression in (10.34). The desired expression in (10.36) can be obtained easily following similar algebraic manipulations. ■

Proof of ??: Let $\mathcal{K}_T$ be a given estimator, which can be non-causal. We have that

$$\mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^* = \begin{bmatrix}\mathcal{K}_T\mathcal{H}_T - \mathcal{L}_T & \mathcal{K}_T\end{bmatrix}\begin{bmatrix}\mathcal{H}_T^*\mathcal{K}_T^* - \mathcal{L}_T^* \\ \mathcal{K}_T^*\end{bmatrix}, \tag{10.48}$$

$$= (\mathcal{K}_T\mathcal{H}_T - \mathcal{L}_T)(\mathcal{K}_T\mathcal{H}_T - \mathcal{L}_T)^* + \mathcal{K}_T\mathcal{K}_T^*, \tag{10.49}$$

$$= \mathcal{K}_T(\mathcal{I}_T + \mathcal{H}_T\mathcal{H}_T^*)\mathcal{K}_T^* - \mathcal{K}_T\mathcal{H}_T\mathcal{L}_T^* - \mathcal{L}_T\mathcal{H}_T^*\mathcal{K}_T^* + \mathcal{L}_T\mathcal{L}_T^*, \tag{10.50}$$

$$\overset{(a)}{=} (\mathcal{K}_T - \mathcal{K}_T^\circ)(\mathcal{I}_T + \mathcal{H}_T\mathcal{H}_T^*)(\mathcal{K}_T - \mathcal{K}_T^\circ)^* + \mathcal{L}_T\mathcal{L}_T^* - \mathcal{K}_T^\circ(\mathcal{I}_T + \mathcal{H}_T\mathcal{H}_T^*)(\mathcal{K}_T^\circ)^*, \tag{10.51}$$

where $\mathcal{K}_T^\circ := \mathcal{L}_T \mathcal{H}_T^* (\mathcal{I}_T + \mathcal{H}_T \mathcal{H}_T^*)^{-1}$ and (a) is obtained from completion of squares.

Moreover, observe that

$$\begin{aligned}
\mathcal{T}_{\mathcal{K}_T^\circ} \mathcal{T}_{\mathcal{K}_T^\circ}^* &= \mathcal{L}_T \mathcal{L}_T^* - \mathcal{K}_T^\circ (\mathcal{I}_T + \mathcal{H}_T \mathcal{H}_T^*)(\mathcal{K}_T^\circ)^*, && (10.52)\\
&= \mathcal{L}_T \mathcal{L}_T^* - \mathcal{L}_T \mathcal{H}_T^* (\mathcal{I}_T + \mathcal{H}_T \mathcal{H}_T^*)^{-1} \mathcal{H}_T \mathcal{L}_T^*, && (10.53)\\
&= \mathcal{L}_T (\mathcal{I}_T - \mathcal{H}_T^* (\mathcal{I}_T + \mathcal{H}_T \mathcal{H}_T^*)^{-1} \mathcal{H}_T) \mathcal{L}_T^*, && (10.54)\\
&\overset{(b)}{=} \mathcal{L}_T (\mathcal{I}_T + \mathcal{H}_T^* \mathcal{H}_T)^{-1} \mathcal{L}_T^*, && (10.55)
\end{aligned}$$

where (b) follows from matrix inversion identity.

Thus, we have that

$$\mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^* = (\mathcal{K}_T - \mathcal{K}_T^\circ)(\mathcal{I}_T + \mathcal{H}_T \mathcal{H}_T^*)(\mathcal{K}_T - \mathcal{K}_T^\circ)^* + \mathcal{T}_{\mathcal{K}_T^\circ} \mathcal{T}_{\mathcal{K}_T^\circ}^* \succcurlyeq \mathcal{T}_{\mathcal{K}_T^\circ} \mathcal{T}_{\mathcal{K}_T^\circ}^* \quad (10.56)$$

Now, consider **??** without the causality constraint on the estimator. Using the strong duality result in Theorem 10.B.5, we can express **??** equivalently as

$$\inf_{\gamma \geq 0} \inf_{\mathcal{K}_T} \gamma \rho_T^2 + \gamma \operatorname{Tr}\left[(\mathcal{I}_T - \gamma^{-1} \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*)^{-1} - \mathcal{I}_T\right] \quad \text{s.t.} \quad \gamma \mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*. \quad (10.57)$$

Fixing $\gamma \geq 0$, we focus on the subproblem

$$\inf_{\mathcal{K}_T} \gamma \operatorname{Tr}\left[(\mathcal{I}_T - \gamma^{-1} \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*)^{-1}\right] \quad \text{s.t.} \quad \gamma \mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*. \quad (10.58)$$

Using the identity in (10.56), we can rewrite (10.58) in terms $\mathcal{K}_T^\circ$ as follows

$$\inf_{\mathcal{K}_T} \gamma^2 \operatorname{Tr}\left[(\gamma \mathcal{I}_T - (\mathcal{K}_T - \mathcal{K}_T^\circ)(\mathcal{I}_T + \mathcal{H}_T \mathcal{H}_T^*)(\mathcal{K}_T - \mathcal{K}_T^\circ)^* - \mathcal{T}_{\mathcal{K}_T^\circ} \mathcal{T}_{\mathcal{K}_T^\circ}^*)^{-1}\right] \text{ s.t. } \gamma \mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*.$$

Since the mapping $\mathcal{X} \mapsto (\mathcal{I}_T - \mathcal{X})^{-1}$ is operator monotone, the minimum over $\mathcal{K}_T$ is attained by $\mathcal{K}_T^\circ$ and the optimal value is given by the following optimization:

$$\inf_{\gamma \geq 0} \gamma \rho_T^2 + \gamma \operatorname{Tr}\left[(\mathcal{I}_T - \gamma^{-1} \mathcal{T}_{\mathcal{K}_T^\circ} \mathcal{T}_{\mathcal{K}_T^\circ}^*)^{-1} - \mathcal{I}_T\right] \quad \text{s.t.} \quad \gamma \mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T^\circ} \mathcal{T}_{\mathcal{K}_T^\circ}^*. \quad (10.59)$$

■

Proof of ??

Proof. Using the strong duality result in Theorem 10.B.5, we can express **??** equivalently as

$$\inf_{\substack{\gamma \geq 0 \\ \mathcal{K}_T \in \mathscr{K}}} \gamma(\rho_T^2 - \operatorname{Tr}(\mathcal{I}_T)) + \gamma^2 \operatorname{Tr}\left[(\gamma \mathcal{I}_T - \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*)^{-1}\right] \quad \text{s.t.} \quad \gamma \mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T} \mathcal{T}_{\mathcal{K}_T}^*.$$

(10.60)

Notice that we can express the rhs as

$$\gamma^2 \operatorname{Tr}\left[(\gamma\mathcal{I}_T - \mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}\right] = \inf_{\mathcal{X}_T \succ 0} \operatorname{Tr}(\mathcal{X}_T) \text{ s.t. } \mathcal{X}_T \succcurlyeq \gamma^2(\gamma\mathcal{I}_T - \mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}. \tag{10.61}$$

Using the Schur complement, we can rewrite the constraint $\mathcal{X}_T \succcurlyeq \gamma^2(\gamma\mathcal{I}_T - \mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*)^{-1}$ as

$$\begin{bmatrix} \mathcal{X}_T & \gamma\mathcal{I}_T \\ \gamma\mathcal{I}_T & \gamma\mathcal{I}_T - \mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^* \end{bmatrix} \succcurlyeq 0, \tag{10.62}$$

where we used the fact that $\gamma\mathcal{I}_T \succ \mathcal{T}_{\mathcal{K}_T}\mathcal{T}_{\mathcal{K}_T}^*$. Using the identity in (10.56), we can rewrite the matrix inequality (10.62) as

$$\begin{bmatrix} \mathcal{X}_T & \gamma\mathcal{I}_T \\ \gamma\mathcal{I}_T & \gamma\mathcal{I}_T - \mathcal{T}_{\mathcal{K}_T^\circ}\mathcal{T}_{\mathcal{K}_T^\circ}^* \end{bmatrix} - \begin{bmatrix} 0 \\ (\mathcal{K}_T - \mathcal{K}_T^\circ) \end{bmatrix} (\mathcal{I}_T + \mathcal{H}_T\mathcal{H}_T^*) \begin{bmatrix} 0 & (\mathcal{K}_T - \mathcal{K}_T^\circ)^* \end{bmatrix} \succcurlyeq 0. \tag{10.63}$$

As $(\mathcal{I}_T + \mathcal{H}_T\mathcal{H}_T^*) \succ 0$, by Schur complement theorem, we can reformulate the matrix inequality above as

$$\begin{bmatrix} \mathcal{X}_T & \gamma\mathcal{I}_T & 0 \\ \gamma\mathcal{I}_T & \gamma\mathcal{I}_T - \mathcal{T}_{\mathcal{K}_T^\circ}\mathcal{T}_{\mathcal{K}_T^\circ}^* & \mathcal{K}_T - \mathcal{K}_T^\circ \\ 0 & (\mathcal{K}_T - \mathcal{K}_T^\circ)^* & (\mathcal{I}_T + \mathcal{H}_T\mathcal{H}_T^*)^{-1} \end{bmatrix} \succcurlyeq 0. \tag{10.64}$$

■

10.C Proofs of Theorems Related to Infinite-Horizon Filtering

Proof of Lemma 10.3.3

Theorem 10.C.1 (Strong Duality in the Infinite-Horizon)**.** *Let $\mathcal{K}$ be a linear and time-invariant estimator (which can be non-causal in general) with bounded $\mathscr{H}_\infty$ norm. Under the Assumptions 10.2.1 and 10.3.2, the infinite-horizon worst-case MSE* (10.4) *suffered by $\mathcal{K}$, i.e.,*

$$\overline{\mathsf{E}}(\mathcal{K}, \rho) = \limsup_{T\to\infty} \frac{1}{T} \sup_{\mathbb{P}_T \in \mathscr{W}_T(\mathbb{P}_T^\circ, \rho_T)} \mathbb{E}_{\mathbb{P}_T}\left[\|\mathbf{e}_T(\boldsymbol{\xi}_T, \mathcal{K})\|^2\right], \tag{10.65}$$

attains a finite value and is equivalent to the following dual problem:

$$\overline{\mathsf{E}}(\mathcal{K}, \rho) = \inf_{\gamma \geq 0} \gamma\rho^2 + \gamma \operatorname{Tr}\left[(\mathcal{I} - \gamma^{-1}\mathcal{T}_{\mathcal{K}}\mathcal{T}_{\mathcal{K}}^*)^{-1} - \mathcal{I}\right] \quad \text{s.t.} \quad \gamma\mathcal{I} \succ \mathcal{T}_{\mathcal{K}}\mathcal{T}_{\mathcal{K}}^*. \tag{10.66}$$

Furthermore, the worst-case disturbance, $\boldsymbol{\xi}_\star$, can be identified from the nominal disturbance, $\boldsymbol{\xi}_\circ$, as

$$\boldsymbol{\xi}_\star = (\mathcal{I} - \gamma_\star^{-1}\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}})^{-1}\boldsymbol{\xi}_\circ, \tag{10.67}$$

where $\gamma_\star$ is the optimal solution of (10.66) *and solves the following equation uniquely:*

$$\operatorname{Tr}\left[((\mathcal{I}-\gamma_\star^{-1}\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*)^{-1}-\mathcal{I})^2\right]=\rho^2. \tag{10.68}$$

Proof. The proof of this result closely tracks the proof of Thm. 5 in [125] By replacing $\mathcal{C}_\mathcal{K}$ in the proof of Thm. 5 in [125] with $\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*$. ■

Proof of Lemma 10.3.3 : The proof follows closely from the proof of **??** in Section 10.B. Let $\mathcal{K}$ be linear time-invariant estimator, which can be non-causal, with bounded $\mathscr{H}_\infty$ norm. We have that

$$\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*=(\mathcal{K}-\mathcal{K}_\circ)(\mathcal{I}+\mathcal{H}\mathcal{H}^*)(\mathcal{K}-\mathcal{K}_\circ)^*+\mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*\succcurlyeq\mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^* \tag{10.69}$$

where $\mathcal{K}^\circ:=\mathcal{L}\mathcal{H}^*(\mathcal{I}+\mathcal{H}\mathcal{H}^*)^{-1}$ and $\mathcal{T}_{\mathcal{K}_T^\circ}\mathcal{T}_{\mathcal{K}_T^\circ}^*=\mathcal{L}(\mathcal{I}+\mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^*$.

Now, consider Problem 10.2.2 without the causality constraint on the estimator. Using the strong duality result in Theorem 10.C.1, we can express Equation (10.5) equivalently as

$$\inf_{\gamma\geq 0}\inf_{\mathcal{K}}\gamma\rho^2+\gamma\operatorname{Tr}\left[(\mathcal{I}-\gamma^{-1}\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*)^{-1}-\mathcal{I}\right]\quad\text{s.t.}\quad\gamma\mathcal{I}\succ\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*. \tag{10.70}$$

Fixing $\gamma\geq 0$, we focus on the subproblem

$$\inf_{\mathcal{K}}\gamma\operatorname{Tr}\left[(\mathcal{I}-\gamma^{-1}\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*)^{-1}\right]\quad\text{s.t.}\quad\gamma\mathcal{I}\succ\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*. \tag{10.71}$$

Using the identity in (10.69), we can rewrite (10.71) in terms $\mathcal{K}_\circ$ as follows

$$\inf_{\mathcal{K}}\gamma^2\operatorname{Tr}\left[(\gamma\mathcal{I}-(\mathcal{K}-\mathcal{K}_\circ)(\mathcal{I}+\mathcal{H}\mathcal{H}^*)(\mathcal{K}-\mathcal{K}_\circ)^*-\mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*)^{-1}\right]\text{ s.t. }\gamma\mathcal{I}\succ\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*.$$

Since the mapping $\mathcal{X}\mapsto(\mathcal{I}-\mathcal{X})^{-1}$ is operator monotone, the minimum over $\mathcal{K}$ is attained by $\mathcal{K}_\circ$ and the optimal value is given by the following optimization:

$$\inf_{\gamma\geq 0}\gamma\rho^2+\gamma\operatorname{Tr}\left[(\mathcal{I}-\gamma^{-1}\mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*)^{-1}-\mathcal{I}\right]\quad\text{s.t.}\quad\gamma\mathcal{I}\succ\mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*. \tag{10.72}$$

■

Proofs of Theorem 10.3.4 and Corollary 10.3.5

Lemma 10.C.2 (Wiener-Hopf Method [120]). *For a bounded and positive definite Toeplitz operator $\mathcal{M}\succ 0$, let $\mathcal{M}\mapsto\Phi(\mathcal{M})$ be a mapping defined as*

$$\Phi(\mathcal{M})\triangleq\inf_{\mathcal{K}\in\mathscr{K}}\operatorname{Tr}\left(\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*\mathcal{M}\right). \tag{10.73}$$

Denote by $\mathcal{M} = \mathcal{U}^*\mathcal{U}$ and $\Delta\Delta^* = \mathcal{I} + \mathcal{H}\mathcal{H}^*$ the canonical spectral factorizations[2] *where $\mathcal{U}$, Δ as well as their inverses $\mathcal{U}^{-1}$, Δ^{-1} are causal operators. The following statements hold:*

i. The optimal causal solution to (10.12) *is given by*

$$\mathcal{K} = \mathcal{U}^{-1}\left\{\mathcal{U}\mathcal{K}_\circ\Delta\right\}_+ \Delta^{-1} = \mathcal{K}_{\mathscr{H}_2} + \mathcal{U}^{-1}\left\{\mathcal{U}\{\mathcal{K}_\circ\Delta\}_-\right\}_+ \Delta^{-1}, \tag{10.74}$$

where $\mathcal{K}_{\mathscr{H}_2} := \left\{\mathcal{K}_\circ\Delta\right\}_+ \Delta^{-1}$ is the Kalman filter.

ii. The function Φ can be written in closed form as

$$\Phi(\mathcal{M}) = \mathrm{Tr}\left[\{\mathcal{U}\mathcal{K}_\circ\Delta\}_- \{\mathcal{U}\mathcal{K}_\circ\Delta\}_-^*\right] + \mathrm{Tr}\left(\mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*\mathcal{M}\right), \tag{10.75}$$

where $\mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^* = \mathcal{L}(\mathcal{I} + \mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^*$.

iii. The gradient of Φ has the following closed form

$$\nabla\Phi(\mathcal{M}) = \mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^* = \mathcal{U}^{-1}\left\{\mathcal{U}\mathcal{K}_\circ\Delta\right\}_- \left\{\mathcal{U}\mathcal{K}_\circ\Delta\right\}_-^* \mathcal{U}^{-*} + \mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*. \tag{10.76}$$

Proof. Using the identity (10.69) and the cyclic property of Tr, the objective can be written as,

$$\inf_{\mathcal{K}\in\mathscr{K}} \mathrm{Tr}\left[(\mathcal{K}-\mathcal{K}_\circ)\Delta\Delta^*(\mathcal{K}-\mathcal{K}_\circ)^*\mathcal{M}\right] = \inf_{\mathcal{K}\in\mathscr{K}} \mathrm{Tr}\left[(\mathcal{K}\Delta-\mathcal{K}_\circ\Delta)(\mathcal{K}\Delta-\mathcal{K}_\circ\Delta)^*\mathcal{U}^*\mathcal{U}\right] \tag{10.77}$$

$$= \inf_{\mathcal{K}\in\mathscr{K}}\ \inf_{\mathcal{K}\in\mathscr{K}} \mathrm{Tr}\left[(\mathcal{U}\mathcal{K}\Delta-\mathcal{U}\mathcal{K}_\circ\Delta)(\mathcal{U}\mathcal{K}\Delta-\mathcal{U}\mathcal{K}_\circ\Delta)^*\right] \tag{10.78}$$

$$= \inf_{\mathcal{K}\in\mathscr{K}} \left\|\mathcal{U}\mathcal{K}\Delta-\mathcal{U}\mathcal{K}_\circ\Delta\right\|_{\mathscr{H}_2}^2, \tag{10.79}$$

where $\|\cdot\|_2$ represents the $\mathscr{H}_2$ norm. Since $\Delta, \mathcal{K}$ and $\mathcal{U}$ are causal, and $\mathcal{U}\mathcal{K}_\circ\Delta$ can be broken into causal and non-causal parts, it is evident that the (causal) filter that minimises the objective is the one that makes the term $\mathcal{U}\mathcal{K}\Delta - \mathcal{U}\mathcal{K}_\circ\Delta$ strictly anti-causal, cancelling off the causal part of $\Delta\mathcal{K}_\circ\mathcal{L}$. This means that the optimal filter satisfies,

$$\mathcal{U}\mathcal{K}\Delta = \left\{\mathcal{U}\mathcal{K}_\circ\Delta\right\}_+. \tag{10.80}$$

[2]The canonical spectral factorization is essentially the Toeplitz operator counterpart of Cholesky decomposition of finite-dimensional matrices.

Also, since $\mathcal{U}^{-1}$ and Δ^{-1} are causal, the optimal causal filter is given by

$$\mathcal{K} = \mathcal{U}^{-1}\left\{\mathcal{U}\mathcal{K}_\circ\Delta\right\}_+ \Delta^{-1}. \tag{10.81}$$

Furthermore, using the identity $\mathcal{K}_\circ\Delta = \{\mathcal{K}_\circ\Delta\}_+ + \{\mathcal{K}_\circ\Delta\}_-$, we get

$$\begin{aligned}
\mathcal{K} &= \mathcal{U}^{-1}\left\{\mathcal{U}\{\mathcal{K}_\circ\Delta\}_+\right\}_+ \Delta^{-1} + \mathcal{U}^{-1}\left\{\mathcal{U}\{\mathcal{K}_\circ\Delta\}_-\right\}_+ \Delta^{-1}, && (10.82)\\
&= \mathcal{U}^{-1}\mathcal{U}\{\mathcal{K}_\circ\Delta\}_+\Delta^{-1} + \mathcal{U}^{-1}\left\{\mathcal{U}\{\mathcal{K}_\circ\Delta\}_-\right\}_+ \Delta^{-1}, && (10.83)\\
&= \{\mathcal{K}_\circ\Delta\}_+\Delta^{-1} + \mathcal{U}^{-1}\left\{\mathcal{U}\{\mathcal{K}_\circ\Delta\}_-\right\}_+ \Delta^{-1}. && (10.84)
\end{aligned}$$

Plugging this solution to $\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*$, we get

$$\begin{aligned}
\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^* &= \mathcal{U}^{-1}(\mathcal{U}\mathcal{K}\Delta - \mathcal{U}\mathcal{K}_\circ\Delta)(\mathcal{U}\mathcal{K}\Delta - \mathcal{U}\mathcal{K}_\circ\Delta)^*\mathcal{U}^{-*} + \mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*, && (10.85)\\
&= \mathcal{U}^{-1}(\{\mathcal{U}\mathcal{K}_\circ\Delta\}_+ - \mathcal{U}\mathcal{K}_\circ\Delta)(\{\mathcal{U}\mathcal{K}_\circ\Delta\}_+ - \mathcal{U}\mathcal{K}_\circ\Delta)^*\mathcal{U}^{-*} + \mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*, && (10.86)\\
&= \mathcal{U}^{-1}\left\{\mathcal{U}\mathcal{K}_\circ\Delta\right\}_-\left\{\mathcal{U}\mathcal{K}_\circ\Delta\right\}_-^*\mathcal{U}^{-*} + \mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*. && (10.87)
\end{aligned}$$

Then, the objective becomes

$$\begin{aligned}
\operatorname{Tr}(\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*\mathcal{M}) &= \operatorname{Tr}(\mathcal{U}^{-1}\left\{\mathcal{U}\mathcal{K}_\circ\Delta\right\}_-\left\{\mathcal{U}\mathcal{K}_\circ\Delta\right\}_-^*\mathcal{U}^{-*}\mathcal{M}) + \operatorname{Tr}(\mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*\mathcal{M}), && (10.88)\\
&= \operatorname{Tr}(\{\mathcal{U}\mathcal{K}_\circ\Delta\}_-\{\mathcal{U}\mathcal{K}_\circ\Delta\}_-^*) + \operatorname{Tr}(\mathcal{T}_{\mathcal{K}_\circ}\mathcal{T}_{\mathcal{K}_\circ}^*\mathcal{M}). && (10.89)
\end{aligned}$$

Finally, by Danskin theorem [43], the gradient of Φ is simply $\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*$ evaluated at the optimal $\mathcal{K}$ as given in (10.87). ■

Lemma 10.C.3. *Let $\gamma > \inf_{\mathcal{K}\in\mathscr{K}} \|\mathcal{T}_\mathcal{K}\|_\infty^2$ be fixed. Then, we have the following duality*

$$\inf_{\substack{\mathcal{K}\in\mathscr{K},\\ \gamma\mathcal{I}\succ\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*}} \gamma\operatorname{Tr}\left[(\mathcal{I}-\gamma^{-1}\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*)^{-1}-\mathcal{I}\right] = \sup_{\mathcal{M}\succ 0}\inf_{\mathcal{K}\in\mathscr{K}} \operatorname{Tr}(\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*\mathcal{M}) - \gamma\operatorname{Tr}\left(\mathcal{M}-2\sqrt{\mathcal{M}}+\mathcal{I}\right). \tag{10.90}$$

Proof. The convex mapping $\mathcal{X}\mapsto\operatorname{Tr}\mathcal{X}^{-1}$ for $\mathcal{X}\succ 0$ can be expressed via Fenchel duality as

$$\sup_{\mathcal{M}\succ 0} -\operatorname{Tr}(\mathcal{X}\mathcal{M}) + 2\operatorname{Tr}(\sqrt{\mathcal{M}}) = \begin{cases}\operatorname{Tr}(\mathcal{X}^{-1}), & \text{if } \mathcal{X}\succ 0\\ +\infty, & \text{o.w.}\end{cases} \tag{10.91}$$

Using the identity (10.91), we rewrite the original problem as,

$$\inf_{\mathcal{K}\in\mathscr{K}}\sup_{\mathcal{M}\succ 0} \operatorname{Tr}(\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*\mathcal{M}) - \gamma\operatorname{Tr}\left(\mathcal{M}-2\sqrt{\mathcal{M}}+\mathcal{I}\right). \tag{10.92}$$

Notice that, the objective above is strictly convex in $\mathcal{K}$ and strictly concave in $\mathcal{M}$. Furthermore, the primal and dual problems are feasible since $\gamma > \inf_{\mathcal{K}\in\mathscr{K}}\|\mathcal{T}_\mathcal{K}\|_\infty^2$. Thus, the proof follows from the minimax theorem. ■

***Proof of Theorem 10.3.4*:** Consider Problem 10.2.2. Using the strong duality result in Theorem 10.C.1, we can express Equation (10.5) equivalently as

$$\inf_{\gamma \geq 0} \inf_{\mathcal{K} \in \mathscr{K}} \gamma \rho^2 + \gamma \operatorname{Tr}\left[(\mathcal{I} - \gamma^{-1} \mathcal{T}_{\mathcal{K}} \mathcal{T}_{\mathcal{K}}^*)^{-1} - \mathcal{I} \right] \quad \text{s.t.} \quad \gamma \mathcal{I} \succ \mathcal{T}_{\mathcal{K}} \mathcal{T}_{\mathcal{K}}^*. \tag{10.93}$$

Fixing $\gamma \geq 0$, we focus on the subproblem

$$\inf_{\mathcal{K}} \gamma \operatorname{Tr}\left[(\mathcal{I} - \gamma^{-1} \mathcal{T}_{\mathcal{K}} \mathcal{T}_{\mathcal{K}}^*)^{-1} - \mathcal{I} \right] \quad \text{s.t.} \quad \gamma \mathcal{I} \succ \mathcal{T}_{\mathcal{K}} \mathcal{T}_{\mathcal{K}}^*. \tag{10.94}$$

Using Lemma 10.C.3, we can reformulate (10.94) as

$$\sup_{\mathcal{M} \succ 0} \inf_{\mathcal{K} \in \mathscr{K}} \operatorname{Tr}(\mathcal{T}_{\mathcal{K}} \mathcal{T}_{\mathcal{K}}^* \mathcal{M}) - \gamma \operatorname{Tr}\left(\mathcal{M} - 2\sqrt{\mathcal{M}} + \mathcal{I} \right). \tag{10.95}$$

Thus, the original formulation in (10.93) can be expressed as

$$\inf_{\gamma \geq 0} \sup_{\mathcal{M} \succ 0} \inf_{\mathcal{K} \in \mathscr{K}} \operatorname{Tr}(\mathcal{T}_{\mathcal{K}} \mathcal{T}_{\mathcal{K}}^* \mathcal{M}) + \gamma \left(\rho^2 - \operatorname{Tr}(\mathcal{M} - 2\sqrt{\mathcal{M}} + \mathcal{I}) \right). \tag{10.96}$$

Note that the objective above is affine in $\gamma \geq 0$ and strictly concave in $\mathcal{M}$. Moreover, primal and dual feasibility hold, enabling the exchange of $\inf_{\gamma \geq 0} \sup_{\mathcal{M} \succ 0}$ resulting in

$$\sup_{\mathcal{M} \succ 0} \inf_{\mathcal{K} \in \mathscr{K}} \inf_{\gamma \geq 0} \operatorname{Tr}(\mathcal{T}_{\mathcal{K}} \mathcal{T}_{\mathcal{K}}^* \mathcal{M}) + \gamma \left(\rho^2 - \operatorname{Tr}(\mathcal{M} - 2\sqrt{\mathcal{M}} + \mathcal{I}) \right), \tag{10.97}$$

where the inner minimization over γ reduces the problem to its constrained version in Equation (10.7).

Finally, the form of the optimal $\mathcal{K}_\star$ follows from the Wiener-Hopf technique in Lemma 10.C.2 and the optimal $\gamma_\star$ and $\mathcal{M}_\star$ can be obtained using the strong duality theorem in (10.C.1).

***Proof of Corollary 10.3.5*:** This result follows immediately from the finiteness of the time-averaged infinite-horizon MSE.

10.D Additional Discussion on Frequency-domain Optimization Method

Pseudocode for Frequency-domain Iterative Optimization Method Solving Equation (10.7)

Frequency-Domain Characterization of the Optimal Solution of Equation (10.5)

We present the frequency-domain formulation of the saddle point $(\mathcal{K}_\star, \mathcal{M}_\star)$ derived in Theorem 10.3.4 to reveal the structure of the solution. We first introduce the following useful results:

Algorithm 8 Frequency-domain iterative optimization method solving Equation (10.7)

Input: Radius $\rho > 0$, state-space model (A, B, C_y, C_s), discretization $N > 0$, tolerance $\epsilon > 0$
Compute $(\overline{A}, \overline{B}, \overline{C})$ from (A, B, C_y, C_s) using (10.106)
Generate frequency samples $\mathbb{T}_N \coloneqq \{\mathrm{e}^{j2\pi n/N} \mid n = 0, \dots, N-1\}$
Initialize $M_0(z) \leftarrow I$ for $z \in \mathbb{T}_N$, and $k \leftarrow 0$
repeat
    Set the step size $\eta_k \leftarrow \frac{2}{k+2}$
    Compute the spectral factor $U_k(z) \leftarrow \texttt{SpectralFactor}(M_k)$ (see Section 10.D)
    Compute the parameter $\Gamma_k \leftarrow \dfrac{1}{N} \sum_{z \in \mathbb{T}_N} U_k(z)\overline{C}(I - z\overline{A})^{-1}$ (see Section 10.D)
    Compute the gradient for $z \in \mathbb{T}_N$ (see Section 10.D)
        $G_k(z) \leftarrow U_k(z)^{-1}\Gamma_k(I - z\overline{A})^{-1}\overline{B}\,\overline{B}^*(I - z\overline{A})^{-*}\Gamma_k^* U_k(z)^{-*} + T_{K_\circ}(z)T_{K_\circ}(z)^*$
    Solve the linear subproblem (10.13a) via bisection (see Section 10.D)
        $\widetilde{M}_k(z) \leftarrow \texttt{Bisection}(G_k, \rho, \epsilon)$ for $z \in \mathbb{T}_N$.
    Set $M_{k+1}(z) \leftarrow (1 - \eta_k)M_k(z) + \eta_k \widetilde{M}_k(z)$ for $z \in \mathbb{T}_N$.
    Increment $k \leftarrow k + 1$
until $\|M_{k+1} - M_k\| / \|M_k\| \leq \epsilon$
Compute $K(z) \leftarrow \texttt{RationalApproximate}(M_{k+1})$ (see Section 10.D)

Lemma 10.D.1 ([101, pg. 261])**.** *Given $H(z) \coloneqq C_y(zI - A)^{-1}B$, consider the canonical spectral factorization $\Delta(z)\Delta(z)^* I + H(z)H(z)^*$ for $z \in \mathbb{T}$. We have that*

$$\Delta(z) = (I + C_y(zI - A)^{-1}F_P)R_e^{1/2}, \tag{10.98}$$

$$\Delta(z)^{-1} = R_e^{-1/2}(I - C_y(zI - A_P)^{-1}F_P), \tag{10.99}$$

where $R_e \coloneqq I + C_y P C_y^*$, $F_P \coloneqq (APC_y^*)R_e^{-1}$, $A_P \coloneqq A - F_P C_y$, and P is the unique positive semidefinite solution to the following discrete algebraic Riccati equation (DARE)

$$P = APA^* + BB^* - F_P R_e F_P^*. \tag{10.100}$$

Denoting by $M_\star(z)$ and $T_{K_\star}(z)$ the transfer functions corresponding to the optimal $\mathcal{M}_\star$ and $\mathcal{T}_{\mathcal{K}_\star}$, respectively, the optimality condition in (10.8) takes the equivalent form:

$$\textit{i. } M_\star(z) = \left(I - \gamma_\star^{-1} T_{K_\star}(z) T_{K_\star}(z)^*\right)^{-2}, \tag{10.101a}$$

$$\textit{ii. } T_{K_\star}(z) T_{K_\star}(z)^* = U_\star(z)^{-1} \{\mathcal{U}_\star \mathcal{S}\}_-(z) \{\mathcal{U}_\star \mathcal{S}\}_-(z)^* U_\star(z)^{-*} + T_{K_\circ}(z) T_{K_\circ}(z)^*, \tag{10.101b}$$

$$\textit{iii. } \operatorname{Tr}\left[\left((I - \gamma_\star^{-1} T_{K_\star}(z) T_{K_\star}(z)^*)^{-1} - I\right)^2\right] = \rho^2, \tag{10.101c}$$

where $\mathcal{S} := \{\mathcal{K}_\circ \Delta\}_-$ is a strictly anticausal operator and $U_\star(z)$ is the transfer function corresponding to the causal canonical factor $\mathcal{U}_\star$

The transfer function corresponding to the operator $\mathcal{S}$ takes a rational form as

$$S(z) := \overline{C}(z^{-1} I - \overline{A})^{-1} \overline{B}, \tag{10.102}$$

where $(\overline{A}, \overline{B}, \overline{C})$ are determined by the original state-space parameters (A, B, C_y, C_s). The following lemma explicitly states this result.

Lemma 10.D.2 ([101, pg. 261] and [193, Lem. 6])**.** *We have that*

$$\mathcal{K}_\circ \Delta = \mathcal{K}_{\mathscr{H}_2} \Delta + \mathcal{S}, \tag{10.103}$$

where $\mathcal{K}_{\mathscr{H}_2}$ is the nominal causal $\mathscr{H}_2$ (aka Kalman) filter and $\mathcal{S} := \{\mathcal{K}_\circ \Delta\}_-$ is strictly anti-causal. Furthermore, the corresponding transfer functions take an explicit form as highlighted below

$$K_{\mathscr{H}_2}(z) := C_s P C_y^\star R_e^{-1} + C_s(I - P C_y^* R_e^{-1} C_y)(zI - A_P)^{-1} F_P, \tag{10.104}$$

$$S(z) := C_s P A_P^* (z^{-1} I - A_P^*)^{-1} C_y^* R_e^{-1/2}, \tag{10.105}$$

where P, R_e, F_P, and A_P are defined as in Lemma 10.D.1.

Thus, we have that

$$\overline{A} \triangleq A_P^*, \quad \overline{B} \triangleq C_y^\star R_e^{-1/2}, \quad \overline{C} \triangleq C_s P A_P^*. \tag{10.106}$$

Notice that for a causal $U(z)$ and strictly anti-causal $S(z)$, the strictly anti-causal part $\{U(z)S(z)\}_-$ may not have any poles from $U(z)$, and all of its poles must be from the strictly anti-causal $S(z)$. This observation is formally expressed in the following lemma.

Lemma 10.D.3. *Let $\mathcal{U}$ be a causal and causally invertible operator, which can be non-rational in general. Then, the strictly anti-causal operator $\{\mathcal{U}\mathcal{S}\}_-$ admits a rational transfer function, i.e.,*

$$\{\mathcal{U}\mathcal{S}\}_-(z) = \Gamma(z^{-1}I - \overline{A})^{-1}\overline{B}, \tag{10.107}$$

where

$$\Gamma \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} U(\mathrm{e}^{\jmath\omega})\overline{C}(I - \mathrm{e}^{\jmath\omega}\overline{A})^{-1}d\omega. \tag{10.108}$$

Proof. Consider the z-transform expansions of $U(z)$ and $S(z)$:

$$U(z) = \sum_{k=0}^{\infty}\widehat{U}_k z^{-k}, \quad \text{and} \quad S(z) = \sum_{l=0}^{\infty}\overline{C}\,\overline{A}^l\,\overline{B}z^{l+1}, \tag{10.109}$$

where the time-domain coefficients $\widehat{U}_k$ can be derived from the Fourier series integrals as

$$\widehat{U}_k \coloneqq \frac{1}{2\pi}\int_{-\pi}^{\pi} U(\mathrm{e}^{\jmath\omega})\mathrm{e}^{j\omega k}d\omega. \tag{10.110}$$

Multiplying $U(z)$ and $S(z)$ and taking the strictly anti-causal parts, *i.e.*, terms with positive powers of z, we get

$$\begin{aligned}
\{U(z)S(z)\}_- &= \left\{\left(\sum_{k=0}^{\infty}\widehat{U}_k z^{-k}\right)\left(\sum_{l=0}^{\infty}\overline{C}\,\overline{A}^l\,\overline{B}z^{l+1}\right)\right\}_-, &(10.111)\\
&= \left(\sum_{k=0}^{\infty}\widehat{U}_k\overline{C}\,\overline{A}^k\right)\left(\sum_{l=0}^{\infty}\overline{A}^l\,\overline{B}z^{l+1}\right), &(10.112)\\
&= \Gamma(z^{-1}I - \overline{A})^{-1}\overline{B}, &(10.113)
\end{aligned}$$

where $\Gamma = \sum_{k=0}^{\infty}\widehat{U}_k\overline{C}\,\overline{A}^k$ which can be expressed as an integral

$$\Gamma = \frac{1}{2\pi}\int_{-\pi}^{\pi} U(\mathrm{e}^{\jmath\omega})\overline{C}(I - \mathrm{e}^{\jmath\omega}\overline{A})^{-1}d\omega. \tag{10.114}$$

using Parseval's theorem. ∎

Proofs of Lemma 10.4.1 and Corollary 10.4.2

Proof of Lemma 10.4.1: Using Lemma 10.D.3, the frequency-domain optimality equations (10.101) can be reformulated explicitly as follows

i. $M_\star(z) = \left(I - \gamma_\star^{-1}T_{K_\star}(z)T_{K_\star}(z)^*\right)^{-2}$, (10.115a)

ii. $T_{K_\star}(z)T_{K_\star}(z)^* = U_\star(z)^{-1}\Gamma_\star(I - z\overline{A})^{-1}\overline{B}\,\overline{B}^*(I - z\overline{A})^{-*}\Gamma_\star^* U_\star(z)^{-*} + T_{K_\circ}(z)T_{K_\circ}(z)^*$, (10.115b)

iii. $\operatorname{Tr}\left[\left((I - \gamma_\star^{-1}T_{K_\star}(z)T_{K_\star}(z)^*)^{-1} - I\right)^2\right] = \rho^2$, (10.115c) where

$$\Gamma_\star \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} U_\star(e^{\jmath\omega})\overline{C}(I-e^{\jmath\omega}\overline{A})^{-1}d\omega, \tag{10.116}$$

and $(\overline{A},\overline{B},\overline{C})$ are as in (10.106). This gives us the desired result. ■

Proof of Corollary 10.4.2: Define $S_\star(z) := \Gamma_\star(I-z\overline{A})^{-1}\overline{B}$ for notational convenience. We rewrite the optimality conditions in (10.115) as

$$\textit{i. } (U_\star(z)^*U_\star(z))^{-1/2} = I - \gamma_\star^{-1}T_{K_\star}(z)T_{K_\star}(z)^* \tag{10.117}$$

$$\textit{ii. } T_{K_\star}(z)T_{K_\star}(z)^* = U_\star(z)^{-1}S_\star(z)S_\star(z)^*U_\star(z)^{-*} + T_{K_\circ}(z)T_{K_\circ}(z)^* \tag{10.118}$$

By plugging *ii.* into *i.*, we get

$$0 = I - (U_\star(z)^*U_\star(z))^{-1/2} - \gamma_\star^{-1}\left(U_\star(z)^{-1}S_\star(z)S_\star(z)^*U_\star(z)^{-*} + T_{K_\circ}(z)T_{K_\circ}(z)^*\right) = 0, \tag{10.119}$$

Multiplying by $U_\star(z)$ from the left and by $U_\star(z)^*$ from the right, we get

$$0 = U_\star(z)U_\star(z)^* - (U_\star(z)U_\star(z)^*)^{1/2} - \gamma_\star^{-1}\left(S_\star(z)S_\star(z)^* + U_\star(z)T_{K_\circ}(z)T_{K_\circ}(z)^*\right)U_\star(z)^*,$$

which can be written further as

$$U_\star(z)U_\star(z)^* = \frac{1}{4}\left(I + \sqrt{I + 4\gamma^{-1}\left(S_\star(z)S_\star(z)^* + U_\star(z)T_{K_\circ}(z)T_{K_\circ}(z)^*\right)U_\star(z)^*}\right)^2. \tag{10.120}$$

Notice that while $S_\star(z)S_\star(z)^*$ is rational, the expression above involves its positive definite square root, which does not generally preserve rationality, implying the desired result. ■

Additional Discussion on the Computation of Gradients

By the Wiener-Hopf technique discussed in Lemma 10.C.2, the gradient $\mathcal{G}_k = \nabla\Phi(\mathcal{M}_k)$ can be obtained as

$$G_k(z) = U_k(z)^{-1}\left\{\mathcal{U}_k\mathcal{K}_\circ\Delta\right\}_-(z)\left\{\mathcal{U}_k\mathcal{K}_\circ\Delta\right\}_-(z)^*U_k^{-*} + T_{K_\circ}T_{K_\circ}^*, \tag{10.121}$$

where $\mathcal{U}_k^*\mathcal{U}_k = \mathcal{M}_k$ is the unique spectral factorization. Furthermore, by Lemma 10.D.3, we can reformulate the gradient $G_k(z)$ more explicitly as

$$G_k(z) = U_k(z)^{-1}\Gamma_k(I-z\overline{A})^{-1}\overline{B}\,\overline{B}^*(I-z\overline{A})^{-*}\Gamma_k^*U_k^{-*} + T_{K_\circ}T_{K_\circ}^*, \tag{10.122}$$

where

$$\Gamma_k \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} U_k(e^{\jmath\omega})\overline{C}(I-e^{\jmath\omega}\overline{A})^{-1}d\omega. \tag{10.123}$$

Here, the spectral factor $U_k(z)$ is obtained for $z \in \mathbb{T}_N$ by Section 10.D. Similarly, the parameter Γ_k can be computed numerically using the trapezoid rule over the discrete domain $\mathbb{T}_N$, *i.e.*,

$$\Gamma_k \leftarrow \frac{1}{N} \sum_{z \in \mathbb{T}_N} U_k(z) \overline{C}(I - z\overline{A})^{-1}. \tag{10.124}$$

Noting that $T_{K_\circ} T_{K_\circ}^*$ is rational and depends only on the system, the gradient $G_k(z)$ can be efficiently computed for $z \in \mathbb{T}_N$.

Implementation of Spectral Factorization

To perform the spectral factorization of an irrational function $M(z)$, we use a spectral factorization method via discrete Fourier transform, which returns samples of the spectral factor on the unit circle. First, we compute $\Lambda(z)$ for $z \in \mathbb{T}_N$, which is defined to be the logarithm of $M(z)$, then we take the inverse discrete Fourier transform λ_k for $k = 0, \ldots, N-1$ of $\Lambda(z)$ which we use to compute the spectral factorization as

$$U(z_n) \leftarrow \exp\left(\frac{1}{2}\lambda_0 + \sum_{k=1}^{N/2-1} \lambda_k z_n^{-k} + \frac{1}{2}(-1)^n \lambda_{N/2}\right)$$

for $k = 0, \ldots, N-1$ where $z_n = \mathrm{e}^{j2\pi n/N}$.

The method is efficient without requiring rational spectra, and the associated error term, featuring a purely imaginary logarithm, rapidly diminishes with an increased number of samples. It is worth noting that this method is explicitly designed for scalar functions.

Algorithm 9 `SpectralFactor`: Spectral Factorization via DFT

1: **Input:** Scalar positive spectrum $M(z) > 0$ on $\mathbb{T}_N \coloneqq \{\mathrm{e}^{j2\pi n/N} \mid n{=}0, \ldots, N{-}1\}$
2: **Output:** Causal spectral factor $U(z)$ of $M(z) > 0$ on $\mathbb{T}_N$
3: Compute the cepstrum $\Lambda(z) \leftarrow \log(M(z))$ on $z \in \mathbb{T}_N$.
4: Compute the inverse DFT
$\lambda_k \leftarrow \mathrm{IDFT}(\Lambda(z))$ for $k = 0, \ldots, N{-}1$
5: Compute the spectral factor for $z_n = \mathrm{e}^{j2\pi n/N}$

$$U(z_n) \leftarrow \exp\left(\frac{1}{2}\lambda_0 + \sum_{k=1}^{N/2-1} \lambda_k z_n^{-k} + \frac{1}{2}(-1)^n \lambda_{N/2}\right), \quad n = 0, \ldots, N{-}1$$

Implementation of Bisection Method

To find the optimal parameter γ_k that solves $\mathrm{Tr}\left[((I-\gamma_k^{-1}\mathcal{G}_k)^{-1}-I)^2\right]=\rho^2$ in the Frank-Wolfe update (10.14), we use a bisection algorithm. The pseudo code for the bisection algorithm can be found in Algorithm 10. We start off with two guesses of γ *i.e.*$(\gamma_{left}, \gamma_{right})$ with the assumption that the optimal γ lies between the two values (without loss of generality).

Algorithm 10 `Bisection Algorithm`

Input: $\gamma_{right}, \gamma_{left}$
Compute the gradient at γ_{right}: $grad\_\gamma_{right}$
while $\mid \gamma_{right} - \gamma_{left} \mid > \epsilon$ **do**
    Calculate the midpoint γ_{mid} between γ_{left} and γ_{right}
    Compute the gradient at γ_{mid}
    if the gradient at γ_{mid} is zero **then**
        return γ_{mid} {Root found}
    else if the gradient at γ_{mid} is positive **then**
        Update γ_{right} to γ_{mid}
    else
        Update γ_{left} to γ_{mid}
    end if
end while
return the average of γ_{left} and γ_{right} {Approximate root}

Proof of Theorem 10.4.4

Our proof of convergence follows closely from the proof technique used in [116]. In particular, since the unit circle is discretized and the computation of the gradients $G_k(z)$ are approximate, the linear suboptimal problem is solved up to an approximation, δ_N, which depends on the problem parameters, and the discretization level N. Namely,

$$\mathrm{Tr}(\nabla\Phi(\mathcal{M}_k)\widetilde{\mathcal{M}}_{k+1}) \geq \sup_{\mathcal{M}\in\Omega_\rho} \mathrm{Tr}(\nabla\Phi(\mathcal{M}_k)\mathcal{M}) - \delta_N \tag{10.125}$$

where

$$\Omega_\rho \coloneqq \{\mathcal{M} \succ 0 \mid \mathrm{Tr}(\mathcal{M} - 2\sqrt{\mathcal{M}} + \mathcal{I}) \leq \rho^2\}, \tag{10.126}$$

Therefore, using Theorem 1 of [116], we obtain

$$\Phi(\mathcal{M}_\star) - \Phi(\mathcal{M}_k) \leq \frac{2\kappa}{k+2}(1+\delta_N). \tag{10.127}$$

where

$$\kappa := \sup_{\substack{\mathcal{M},\widetilde{\mathcal{M}}\in\Omega_\rho \\ \eta\in[0,1] \\ \mathcal{M}'=\mathcal{M}+\eta(\widetilde{\mathcal{M}}-\mathcal{M})}} \frac{2}{\eta^2}\left(\mathrm{Tr}(\mathcal{M}'\nabla\Phi(\mathcal{M}))-\Phi(\mathcal{M}')\right). \tag{10.128}$$

Implementation of Rational Approximation

We present the pseudocode of `RationalApproximation`.

Algorithm 11 `RationalApproximation`

1: **Input:** Scalar positive spectrum $M(z) > 0$ on $\mathbb{T}_N := \{e^{j2\pi n/N} \mid n=0,\ldots,N-1\}$, and a small positive scalar ϵ
2: **Output:** Causal rational filter $K(z)$ on $\mathbb{T}_N$
3: Get $P(z), Q(z)$ by solving the convex optimization in (10.16), for fixed ϵ, given $M(z)$
4: Get the rational spectral factors of $P(z), Q(z)$, which are $S_P(z), S_Q(z)$ using the canonical Factorization method in [199]
5: Get $U^r(z)$,the rational spectral factor of$M(z)$, as $S_P(z)/S_Q(z)$
6: Get $K(z)$ from the formulation in (10.18), (10.146)

Proof of Theorem 10.4.6

We write the DR estimator, $K(e^{j\omega})$, as a sum of causal functions:

$$\begin{aligned} K(e^{j\omega}) &= U^{-1}\{UK_0\Delta\}_+\Delta^{-1} && (10.129)\\ &= U^{-1}(U\{K_0\Delta\}_+ + \{U\{K_0\Delta\}_-\}_+)\Delta^{-1} && (10.130)\\ &= \{K_0\Delta\}_+\Delta^{-1} + U^{-1}\{U\{K_0\Delta\}_-\}_+\Delta^{-1} && (10.131)\end{aligned}$$

where we drop the dependence of Δ, K_0 and U on $e^{j\omega}$.

Given the spectral factor $U(e^{j\omega})$ in rational form as $U(e^{j\omega}) = \tilde{D}^{1/2}(I + \tilde{C}(e^{j\omega}I - \tilde{A})^{-1}\tilde{B})$, its inverse is given by:

$$U^{-1}(e^{j\omega}) = (I - \tilde{C}(e^{j\omega}I - (\tilde{A} - \tilde{B}\tilde{C}))^{-1}\tilde{B})\tilde{D}^{-1/2} \tag{10.132}$$

From the above, we have:

$$\{K_0\Delta\}_- = T(z) = C_sPA_P^*(z^{-1}I - A_P^*)^{-1}C_y^*(I + C_yPC_y^*)^{-*/2} \tag{10.133}$$

Multiplying the above equation with U, and taking its causal part, we get:

$$\begin{aligned}\{U\{\Delta K_0\}_-\}_+ =&\{\tilde{D}^{1/2}C_sPA_P^*(z^{-1}I - A_P^*)^{-1}C_y^*(I + C_yPC_y^*)^{-*/2} + \\ &\tilde{D}^{1/2}\tilde{C}(e^{j\omega}I - \tilde{A})^{-1}\tilde{B}C_sPA_P^*(z^{-1}I - A_P^*)^{-1}C_y^*(I + C_yPC_y^*)^{-*/2}\}_+\end{aligned} \tag{10.134}$$

Given that the term $\tilde{D}^{1/2}C_sPA_P^*(z^{-1}I - A_P^*)^{-1}C_y^*(I + C_yPC_y^*)^{-*/2}$ is strictly anticausal, and considering the matrix U_{ly} which solves the lyapunov equation: $\tilde{A}U_{ly}A_P^* + \tilde{B}C_sPA_P^* = U_{ly}$, we get $\{U\{K_0\Delta\}_-\}_+$ as:

$$
\begin{aligned}
&\{U\{\Delta K_0\}_-\}_+ \\
&= \{\tilde{D}^{1/2}\tilde{C}\left((zI - \tilde{A})^{-1}\tilde{A}U_{ly} + U_{ly}A_P^*(z^{-1}I - A_P^*)^{-1} + U_{ly}\right)C_y^*(I + C_yPC_y^*)^{-*/2}\}_+ \\
&&& (10.135) \\
&= \tilde{D}^{1/2}\tilde{C}\left((zI - \tilde{A})^{-1}\tilde{A} + I\right)U_{ly}C_y^*(I + C_yPC_y^*)^{-*/2} && (10.136) \\
&= z\tilde{D}^{1/2}\tilde{C}(zI - \tilde{A})^{-1}U_{ly}C_y^*(I + C_yPC_y^*)^{-*/2} && (10.137)
\end{aligned}
$$

Now, multiplying equation (10.137) by the inverse of U (10.132), we get:

$$
\begin{aligned}
U^{-1}\{U\{K_0\Delta\}_-\}_+ &= z(I + \tilde{C}(e^{\jmath\omega}I - \tilde{A})^{-1}\tilde{B})^{-1}\tilde{C}(zI - \tilde{A})^{-1}U_{ly}C_y^*(I + C_yPC_y^*)^{-*/2} \\
&&& (10.138) \\
&= z\tilde{C}(I + (zI - \tilde{A})^{-1}\tilde{B}\tilde{C})^{-1}(zI - \tilde{A})^{-1}U_{ly}C_y^*(I + C_yPC_y^*)^{-*/2} \\
&&& (10.139) \\
&= z\tilde{C}(zI - \tilde{A}_P)^{-1}U_{ly}C_y^*(I + C_yPC_y^*)^{-*/2} && (10.140) \\
&= \tilde{C}(I + \tilde{A}_P(zI - \tilde{A}_P)^{-1})U_{ly}C_y^*(I + C_yPC_y^*)^{-*/2} && (10.141) \\
&&& (10.142)
\end{aligned}
$$

where $\tilde{A}_P = \tilde{A} - \tilde{B}\tilde{C}$.

The inverse of Δ is given by $\Delta^{-1}(z) = (I + C_yPC_y^*)^{-1/2}(I - C_y(zI - A_P)^{-1}K_P)$, and we already showed that $\{K_0\Delta\}_+ = C_s(zI - A)^{-1}APC_y^*(I + C_yPC_y^*)^{-*/2} + C_sPC_y^*(I + C_yPC_y^*)^{-*/2}$.

Then we can get the 2 terms of equation (10.131):

$$
\{K_0\Delta\}_+\Delta^{-1} = C_sPC_y^*(I+C_yPC_y^*)^{-1} + C_s\left(I - PC_y^*(I + C_yPC_y^*)^{-1}C_y\right)(zI - A_P)^{-1}K_P \tag{10.143}
$$

and

$$
\begin{aligned}
&U^{-1}\{U\{K_0\Delta\}_-\}_+\Delta^{-1} \\
&= \left(\tilde{C}U_{ly}C_y^*(I+C_yPC_y^*)^{-*/2} + \tilde{C}\tilde{A}_P(zI-\tilde{A}_P)^{-1}U_{ly}C_y^*(I+C_yPC_y^*)^{-*/2}\right) \\
&\times \left((I+C_yPC_y^*)^{-1/2} - (I+C_yPC_y^*)^{-1/2}C_y(zI-A_P)^{-1}K_P\right) \qquad (10.144) \\
&= \tilde{C}\tilde{A}_P(zI-\tilde{A}_P)^{-1}U_{ly}C_y^*(I+C_yPC_y^*)^{-1}\left(I - C_y(zI-A_P)^{-1}K_P)\right) \\
&- \tilde{C}U_{ly}C_y^*(I+C_yPC_y^*)^{-1}C_y(zI-A_P)^{-1}K_P \\
&+ \tilde{C}U_{ly}C_y^*(I+C_yPC_y^*)^{-1} \qquad (10.145)
\end{aligned}
$$

Finally, summing equations (10.143) and (10.144), we get the controller $K(\mathrm{e}^{j\omega})$ in its rational form:

$$
K(\mathrm{e}^{j\omega}) = \begin{bmatrix}\tilde{C}\tilde{A}_P & -C_s + C_sPC_y^*(I+C_yPC_y^*)^{-1}C_y + \tilde{C}U_{ly}C_y^*(I+C_yPC_y^*)^{-1}C_y\end{bmatrix} \qquad (10.146)
$$

$$
\times \left(zI - \begin{bmatrix}\tilde{A}_P & U_{ly}C_y^*(I+C_yPC_y^*)^{-1}C_y \\ 0 & A_P\end{bmatrix}\right)^{-1} \begin{bmatrix}U_{ly}C_y^*(I+C_yPC_y^*)^{-1} \\ -K_P\end{bmatrix} \qquad (10.147)
$$

$$
+ \tilde{C}U_{ly}C_y^*(I+C_yPC_y^*)^{-1} + C_sPC_y^*(I+C_yPC_y^*)^{-1} \qquad (10.148)
$$

which can be explicitly rewritten as in equation (10.18), where $\widetilde{F}$,$\widetilde{G}$,$\widetilde{H}$ and $\widetilde{L}$ are defined accordingly.

10.E Simulation Results

Another tracking problem

We study another tracking problem, standard in the filtering community, whose state-space model is

$F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \Delta t & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \Delta t & 1 \end{bmatrix}$, $G = \begin{bmatrix} \Delta t & 0 \\ 0.5(\Delta t)^2 & 0 \\ 0 & \Delta t \\ 0 & 0.5(\Delta t)^2 \end{bmatrix}$,$H = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $L = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$,

with $\Delta t = 1$. The results are shown in the plots below.

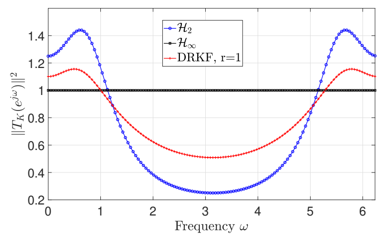
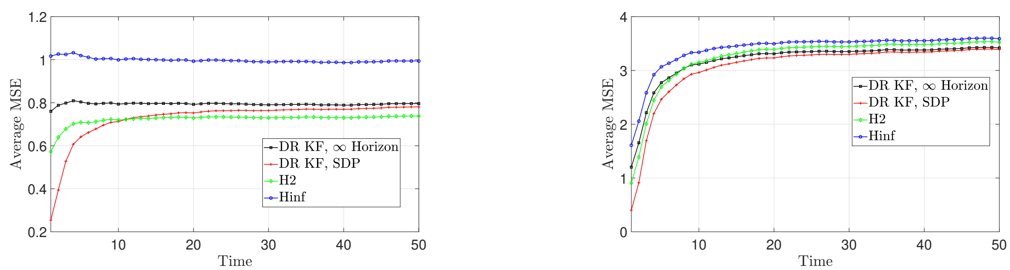


Figure 10.4: The frequency response of different filters ($\mathcal{H}_2, \mathcal{H}_\infty$ and DRKF) for the tracking problem in section 10.E. The worst-case expected MSE is 3.99 for H_∞ , 3.77 for H_2 and 3.47 (lowest) for DRKF.



(a) White noise

(b) Worst noise for DR-KF, SDP

Figure 10.5: The average MSE of the different filters horizon under different disturbances for the tracking problem in section 10.E. (a) is white noise, while (b) is the worst-case noise for the finite horizon DR KF (SDP). While the KF performs best under gaussian noise, the DRKF achieves the lowest error in most of other scenarios (including the more realistic case of correlated noise), and the finite and infinite horizon achieve similar avergae MSE at the end of the horizon.

Chapter 11

MIXED $\mathscr{H}_2/\mathscr{H}_\infty$ CONTROL

11.1 Introduction

Performance and robustness are the two most desired characteristics of a controller. Tackling uncertainty plays a pivotal role in the realm of control, especially within control systems that are frequently exposed to a range of uncertainties including external disruptions, inaccuracies in measurements, deviations in models, and time-varying system dynamics. Ignoring such uncertainties during the development of control policies can significantly undermine performance. The traditional $\mathscr{H}_2$ [121] and $\mathscr{H}_\infty$ [51] control are the two main approaches to address the dichotomy of robustness and performance. While $\mathscr{H}_2$ control aims to achieve optimal average case performance for stochastic white disturbances, $\mathscr{H}_\infty$ control guarantees robustness to worst-case deterministic disturbances with bounded energy. They both result in rational controllers which can be efficiently synthesized by solving a set of algebraic Riccati equations.

When it comes to designing a controller that is both robust to worst-case disturbances and optimal in the average case performance, a natural idea is to find the best $\mathscr{H}_2$-optimal controller among a family of suboptimal $\mathscr{H}_\infty$ controllers. Also known as the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control, these controllers provide a trade-off between the performance of $\mathscr{H}_2$ and the robustness of the $\mathscr{H}_\infty$ controllers. Although both $\mathscr{H}_2$ and central $\mathscr{H}_\infty$ controllers are rational and admit finite-order state-space realizations [51], Megretski [165] showed that the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller has infinite order whenever the $\mathscr{H}_\infty$-norm constraint is not redundant.

Prior Works

The mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control problem and its several variations have been extensively studied in the past. [167] is one of the earliest works establishing a connection between the maximum entropy $\mathscr{H}_\infty$ solution and the $\mathscr{H}_2$ performance. [18] proposed the first conceptualization and formulation of mixed $\mathscr{H}_2$ control with $\mathscr{H}_\infty$ constraints, and obtained a fixed order dynamic output feedback controller for an auxiliary objective upper bounding the $\mathscr{H}_2$-norm. [130], [259] approached this problem similarly by considering auxiliary systems or objectives to bound the desired mixed $\mathscr{H}_2/\mathscr{H}_\infty$ objective.

Unlike the pure $\mathscr{H}_2$ and pure $\mathscr{H}_\infty$ controllers, it was proved by [165] that the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller is non-rational whenever the $\mathscr{H}_\infty$ constraint is active. Since non-rational functions do not admit finite-dimensional state-space realizations, future works focused either on finite-dimensional approximations or more tractable auxiliary objectives.

Among these, [94] addressed this problem by converting it into a series of convex subproblems. [238] proposed a homotopy algorithm for fixed-order controller synthesis whereas [201] considered a gradient-based method. [104] proposed an approximate finite-dimensional parametrization by choosing a finite basis for the Youla parameter to compute the suboptimal solutions. The performance characteristics of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ objective have also been studied in [100] where the authors provide a relation between the $\mathscr{H}_2$-norms of the pure $\mathscr{H}_2$-optimal controller and the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller.

Contributions

In this work, we study the infinite-horizon mixed $\mathscr{H}_2$-optimal control with $\mathscr{H}_\infty$-norm constraints of finite-order, discrete-time, linear time-invariant (LTI) systems. While several past works [52], [100] assumed separation of the disturbances into stochastic and deterministic components, we make no such separation assumption. In particular, we consider the full-information setting where the controller can access current and past disturbances. Our major contributions are as follows:

i. *The Exact Stabilizing Optimal Controller:* Moving away from earlier approaches that utilized approximation strategies and auxiliary problems to obtain a finite-dimensional optimization problem, we find the *exact stabilizing optimal controller* in the frequency-domain for the infinite-horizon mixed $\mathscr{H}_2/\mathscr{H}_\infty$ problem.

ii. *A Finite-Dimensional Characterization:* While we confirm that the optimal controller is irrational as shown in [165], we show that a finite-dimensional parameter completely characterizes it.

iii. *An Efficient Numerical Method:* Exploiting the finite-dimensional characterization, we propose an iterative fixed point method to compute the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller in the frequency domain to arbitrary fidelity.

iv. *Fixed-order Rational Approximations:* Given a finite order, we find the best rational approximation (in H_∞ norm) to the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller and obtain a state space structure for the approximate controller. We provide numerical simulations to analyze the performance of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller and its rational approximations for different orders.

11.2 Preliminaries

Notations: Calligraphic symbols ($\mathcal{K}$, $\mathcal{M}$, $\mathcal{L}$, etc.) represent operators, with $\mathcal{I}$ symbolizing the identity operator. Letters in boldface ($\mathbf{x}$, $\mathbf{u}$, $\mathbf{w}$, etc.) refer to infinite sequences. The notation $\mathcal{M}^*$ indicates the adjoint of the operator $\mathcal{M}$, while $\succcurlyeq$ signifies the Loewner order. tr denotes the normalized trace over block Laurent operators, and Tr denotes the usual trace over finite dimensional matrices. For $p \in [1, \infty]$, we denote by $\mathscr{L}_p^{n\times m}$ the Banach space of $n \times m$-block Laurent operators whose transfer matrix has finite L_p norm on the unit circle. Similarly, $\mathscr{H}_p^{n\times m}$ is the Banach space of causal $n \times m$-block Laurent operators whose transfer matrix has finite L_p norm on the unit circle. Norms $\|\cdot\|_\infty$ and $\|\cdot\|_2$ represent the $\mathscr{H}_\infty/L_\infty$ and $\mathscr{H}_2/L_\infty$ norms, respectively. $\|\cdot\|$ is reserved for Euclidean norm for vectors and ℓ_2 norm of sequences. The expressions $\{\mathcal{M}\}_+$ and $\{\mathcal{M}\}_-$ delineate the causal and strictly anti-causal portions of an operator $\mathcal{M}$. The absolute value of an operator is defined as $|\mathcal{M}| \coloneqq \sqrt{\mathcal{M}^*\mathcal{M}}$. $\mathscr{T}_+^m$ denotes the set of positive symmetric trigonometric polynomials of degree m. $\mathbb{S}_+^m$ denotes the set of symmetric positive semidefinite matrices of size $m \times m$. $\overline{\sigma}(M)$ denotes the maximum singular value of a matrix M.

Linear-Quadratic Control

We consider a discrete-time linear time-invariant (LTI) dynamical system described by its state-space representation as:

$$x_{t+1} = Ax_t + B_u u_t + B_w w_t, \tag{11.1}$$

Here, the vectors $x_t \in \mathbb{R}^{d_x}$, $u_t \in \mathbb{R}^{d_u}$, and $w_t \in \mathbb{R}^{d_w}$ denote the state, control input, and exogenous disturbance at time $t \in \mathbb{Z}$, respectively. At a given time instance $t \in \mathbb{Z}$, the controller suffers a per stage cost $c_t \coloneqq x_t^\intercal Q x_t + u_t^\intercal R u_t$, where $Q, R \succ 0$. We assume that (A, B_u) and (A, B_w) are stabilizable. For notational convenience, we take $Q = I_{d_x}$ and $R = I_{d_u}$ without loss of generality by change of variables $x_t \mapsto Q_t^{-1/2} x_t$ and $u_t \mapsto R_t^{-1/2} u_t$ so that $c_t = \|x_t\|^2 + \|u_t\|^2$.

Input-Output Formalism Throughout this paper, we employ an operator-theoretic framework to represent the state-space dynamics in Equation (11.1) as input-output maps. We use $\mathbf{x} \coloneqq \{x_t\}_{t\in\mathbb{Z}}$, $\mathbf{u} \coloneqq \{u_t\}_{t\in\mathbb{Z}}$, and $\mathbf{w} \coloneqq \{w_t\}_{t\in\mathbb{Z}}$ to represent the state, control input, and disturbance sequences, respectively. The dynamical relations among these variables dictated by the state-space structure in (11.1) can be captured equivalently and succinctly using input/output transfer operators as:

$$\mathbf{x} = \mathcal{F}\mathbf{u} + \mathcal{G}\mathbf{w}, \tag{11.2}$$

where $\mathcal{F}$ and $\mathcal{G}$ symbolize strictly causal $d_x \times d_u$ and $d_x \times d_w$-block Laurent operators mapping control inputs $\mathbf{u}$ and the disturbances $\mathbf{w}$ to the states $\mathbf{x}$. The corresponding transfer matrices are given by

$$F(z) = C(zI - A)^{-1} B_u, \quad G(z) = C(zI - A)^{-1} B_w. \tag{11.3}$$

Controller We consider the full-information setting where the control input u_t at any time $t \in \mathbb{Z}$ has causal access to past disturbances, $\{w_\tau\}_{\tau \leq t}$. In particular, we restrict our attention to causal LTI controllers such that $u_t = \sum_{s \leq t} \widehat{K}_{t-s} w_s$ where $\{\widehat{K}_t\}_{t \geq 0}$ is the Markov parameters of the controller. We succinctly capture this relationship via a linear mapping

$$\mathcal{K} : \mathbf{w} \mapsto \mathbf{u} := \mathcal{K}\mathbf{w}$$

where $\mathcal{K}$ is a causal $d_u \times d_w$-block Laurent operator. Furthermore, we seek controllers that map ℓ_2 disturbances to ℓ_2 control inputs, which makes them bounded. Therefore, $\mathcal{K}$ must be a member of $\mathscr{H}_\infty^{d_u \times d_w}$.

When the underlying system in (11.2) is in a feedback loop with a fixed controller $\mathcal{K}$, the states and the control inputs are determined completely by the disturbances through the closed-loop transfer operator given by:

$$\mathcal{T}_{\mathcal{K}} : \mathbf{w} \mapsto \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} := \begin{bmatrix} \mathcal{F}\mathcal{K} + \mathcal{G} \\ \mathcal{K} \end{bmatrix} \mathbf{w}. \tag{11.4}$$

We use $K(z)$ and $T_K(z)$ to represent the z-domain transfer matrices corresponding to the operators $\mathcal{K}$ and $\mathcal{T}_{\mathcal{K}}$, respectively.

In the remainder of this paper, our focus will primarily be on the quadratic expression $\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}} \succcurlyeq 0$, which can be expressed in terms of $\mathcal{K}$ after a completion-of-squares as

$$\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}} = (\mathcal{K} - \mathcal{K}_{\text{nc}})^* (\mathcal{I} + \mathcal{F}^* \mathcal{F})(\mathcal{K} - \mathcal{K}_{\text{nc}}) + \mathcal{T}_{\mathcal{K}_{\text{nc}}}^* \mathcal{T}_{\mathcal{K}_{\text{nc}}}. \tag{11.5}$$

Here, $\mathcal{K}_{\text{nc}} := -(\mathcal{I} + \mathcal{F}^* \mathcal{F})^{-1} \mathcal{F}^* \mathcal{G}$ denotes the unique non-causal controller such that $\mathcal{T}_{\mathcal{K}_{\text{nc}}}^* \mathcal{T}_{\mathcal{K}_{\text{nc}}} \preccurlyeq \mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}$ for all $\mathcal{K}$, [101].

Mixed $\mathscr{H}_2/\mathscr{H}_\infty$ Control

This paper explores the problem of mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control, aiming to design a causal controller minimizing the $\mathscr{H}_2$ norm of the closed-loop transfer operator, $\mathcal{T}_\mathcal{K}$, while ensuring that its $\mathscr{H}_\infty$ norm is bounded above by a fixed constant $\gamma > 0$. This can be stated formally as a constrained optimization problem as follows:

Problem 11.2.1 (Mixed $\mathscr{H}_2/\mathscr{H}_\infty$ Control)**.** Given an achievable $\mathscr{H}_\infty$ norm bound $\gamma > 0$, find a causal and bounded controller $\mathcal{K} \in \mathscr{H}_\infty^{d_u \times d_w}$ that achieves the minimum closed-loop $\mathscr{H}_2$ norm among all γ-suboptimal $\mathscr{H}_\infty$ controllers, *i.e.*,

$$\inf_{\mathcal{K} \in \mathscr{H}_\infty^{d_u \times d_w}} \|\mathcal{T}_\mathcal{K}\|_2^2 \quad \text{s.t.} \quad \|\mathcal{T}_\mathcal{K}\|_\infty \leq \gamma. \tag{11.6}$$

We will drop the superscript over $\mathscr{L}$ and $\mathscr{H}$ spaces whenever the block dimensions are clear. Here, the $\mathscr{H}_2$ norm of $\mathcal{T}_\mathcal{K}$ is defined as

$$\|\mathcal{T}_\mathcal{K}\|_2 \coloneqq \sqrt{\operatorname{tr}(\mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K})} = \sqrt{\int_0^{2\pi} \frac{d\omega}{2\pi} \operatorname{Tr}(T_K(\mathrm{e}^{\jmath\omega})^* T_K(\mathrm{e}^{\jmath\omega}))} \tag{11.7}$$

The $\mathscr{H}_2$ criterion can be derived from the infinite-horizon average expected cost when the disturbances constitute a white random process with $\mathbb{E}[w_s w_t^\intercal] = I_{d_w} \delta_{s-t}$, namely,

$$\limsup_{T\to\infty} \frac{1}{2T+1}\, \mathbb{E}\left[\sum\nolimits_{t=-T}^{T} \|x_t\|^2 + \|u_t\|^2\right] = \|\mathcal{T}_\mathcal{K}\|_2^2. \tag{11.8}$$

We denote by $\mathcal{K}_{\mathscr{H}_2} \coloneqq \arg\min\{\|\mathcal{T}_\mathcal{K}\|_2 \mid \mathcal{K} \in \mathscr{H}_\infty^{d_u \times d_w}\}$ the optimal $\mathscr{H}_2$ controller.

Similarly, the $\mathscr{H}_\infty$ norm of $\mathcal{T}_\mathcal{K}$ corresponds to its operator norm as a mapping from $\ell_2(\mathbb{Z})$ to $\ell_2(\mathbb{Z})$, *i.e.*,

$$\|\mathcal{T}_\mathcal{K}\|_\infty \coloneqq \sup_{\mathbf{w} \in \ell_2/\{0\}} \frac{\|\mathcal{T}_\mathcal{K}\mathbf{w}\|}{\|\mathbf{w}\|} = \max_{\omega \in [0,2\pi)} \overline{\sigma}(T_K(\mathrm{e}^{\jmath\omega})) \tag{11.9}$$

Since $\mathcal{T}_\mathcal{K}$ consists of 2×1 block of causal Laurent operators, the corresponding transfer matrix $T_K(z)$ is analytic outside the unit circle whenever it is bounded on the unit circle. The $\mathscr{H}_\infty$ criterion can be derived from the worst-case infinite-horizon cost incurred by $\mathcal{K}$ among all bounded energy (or bounded power) disturbances, namely,

$$\sup_{\|\mathbf{w}\| \leq 1} \sum_{t=-\infty}^{\infty} \|x_t\|^2 + \|u_t\|^2 = \sup_{\|\mathbf{w}\| \leq 1} \|\mathbf{x}\|^2 + \|\mathbf{u}\|^2 = \|\mathcal{T}_\mathcal{K}\|_\infty^2.$$

We denote by $\mathcal{K}_{\mathscr{H}_\infty} \coloneqq \arg\min\{\|\mathcal{T}_\mathcal{K}\|_\infty \mid \mathcal{K} \in \mathscr{H}_\infty^{d_u \times d_w}\}$ the optimal $\mathscr{H}_\infty$ controller.

If it exists, the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller achieves the minimum expected cost for stochastic disturbances while guaranteeing a certain degree of robustness against adversarial bounded energy (or bounded power) disturbances.

Let $\gamma_2 := \|\mathcal{T}_{\mathcal{K}_{\mathscr{H}_2}}\|_\infty$ and $\gamma_\infty := \|\mathcal{T}_{\mathcal{K}_{\mathscr{H}_\infty}}\|_\infty$ be the $\mathscr{H}_\infty$ norms of closed-loop transfer operators under the optimal $\mathscr{H}_2$ and $\mathscr{H}_\infty$ controllers, respectively. Note that for $\gamma \geq \gamma_2$, the $\mathscr{H}_\infty$-norm constraint in Problem 11.2.1 is redundant, and the optimal solution coincides with $\mathcal{K}_{\mathscr{H}_2}$. Moreover, for $\gamma < \gamma_\infty$, Problem 11.2.1 is not feasible since no causal controller can achieve $\mathscr{H}_\infty$ norm less than that of the $\mathscr{H}_\infty$ controller, $\mathcal{K}_{\mathscr{H}_\infty}$. For $\gamma = \gamma_\infty$, the solution coincides with $\mathcal{K}_{\mathscr{H}_\infty}$. Thus, we are interested in non-trivial solutions for $\gamma \in (\gamma_\infty, \gamma_2)$, which interpolate between the optimal $\mathscr{H}_2$ and $\mathscr{H}_\infty$ controllers.

11.3 Main Results

In this section, we present the main theoretical results of our paper. In Theorem 11.3.1, we formulate the Lagrange dual of Problem 11.2.1 and establish strong duality. In Theorem 11.3.4, we state the optimal controller and argue that it is irrational.

First, we form the Lagrange dual of Problem 11.2.1 in the following theorem.

Theorem 11.3.1 (Strong Duality). *Let $\gamma \in (\gamma_\infty, \gamma_2)$ be an admissible $\mathscr{H}_\infty$ norm bound. The infinite-horizon mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control problem (Problem 11.2.1) is equivalent to the following max-min problem:*

$$\max_{\substack{\Lambda \in \mathscr{L}_1,\\ \Lambda \succcurlyeq 0}} \min_{\mathcal{K} \in \mathscr{H}_\infty} \operatorname{tr}(\mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K}(\mathcal{I} + \Lambda)) - \gamma^2 \operatorname{tr}(\Lambda), \tag{11.10}$$

where the dual variable $\Lambda \in \mathscr{L}_1^{d_w \times d_w}$ is a positive definite, self-adjoint, $d_w \times d_w$-block Laurent operator.

Proof. The proof of this theorem relies on the Lagrange duality theory for infinite-dimensional optimization on Banach spaces. Consider Problem 11.2.1. The $\mathscr{H}_\infty$ norm constraint $\|\mathcal{T}_\mathcal{K}\|_\infty \leq \gamma$ is equivalent to the L_∞ norm constraint $\|\mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K}\|_\infty \leq \gamma^2$, which in turn is equivalent to the following operator inequality constraint:

$$\mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K} - \gamma^2 \mathcal{I} \preccurlyeq 0. \tag{11.11}$$

Since $\mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K} \in \mathscr{L}_\infty^{d_w \times d_w}$ is self-adjoint, this operator inequality constraint is, by definition, equivalent to

$$\operatorname{tr}(\Lambda(\mathcal{T}_\mathcal{K}^* \mathcal{T}_\mathcal{K} - \gamma^2 \mathcal{I})) \leq 0, \quad \forall \Lambda \succcurlyeq 0, \tag{11.12}$$

where $\Lambda \in \mathscr{L}_1^{d_w \times d_w}$ is a positive-definite, self-adjoint, block Laurent operator with bounded absolute trace, *i.e.*, $\operatorname{tr}(|\Lambda|) < +\infty$. This constraint can be reincorporated into the primal problem in (11.6) via Lagrangian, which yields an equivalent min-max problem:

$$\inf_{\mathcal{K} \in \mathscr{K}} \sup_{\Lambda \succcurlyeq 0} \operatorname{tr}(\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}) + \operatorname{tr}(\Lambda(\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}} - \gamma^2 \mathcal{I})). \tag{11.13}$$

Notice that the Lagrangian objective function is strictly convex (indeed quadratic) in $\mathcal{K}$, and affine in the dual variable $\Lambda \succcurlyeq 0$. Furthermore, *Slater's conditions* are satisfied since $\mathcal{K}_{\mathscr{H}_\infty} \in \mathscr{H}_\infty$ and $\|\mathcal{T}_{\mathcal{K}_{\mathscr{H}_\infty}}\|_\infty =: \gamma_\infty < \gamma$. Thus, by [254, Thm. 2.9.2], strong duality holds, and the supremum is attained. Moreover, the inner infimum is attained for a fixed $\Lambda \succcurlyeq 0$ due to strict convexity wrt $\mathcal{K}$. Rearranging terms, we get (11.10). ■

As opposed to the $\mathscr{H}_\infty$-norm constrained primal problem in (11.6), the max-min problem in (11.10) is more manageable as it is not norm constrained and strictly convex in $\mathcal{K}$. This comes at the expense of an additional maximization step.

Notice that the inner minimization over causal $\mathcal{K}$ in (11.10) is nothing but $(\mathcal{I} + \Lambda)$-weighted squared $\mathscr{H}_2$-norm objective. Therefore, the inner minimization can be considered a stochastic optimal control problem under weakly stationary disturbances with autocovariance operator $\mathcal{I} + \Lambda$. This can be carried out tractably using the Wiener-Hopf method [251] and canonical spectral factorization of $\mathcal{I} + \Lambda$ as stated in Lemma 11.3.2.

Lemma 11.3.2 (Wiener-Hopf Method)**.** *Let $\Lambda \in \mathscr{L}_1$ be a positive-definite and self-adjoint block Laurent operator. Consider the problem of finding an optimal causal controller, $\mathcal{K}$, minimizing $(\mathcal{I} + \Lambda)$-weighted $\mathscr{H}_2$ norm, i.e.,*

$$\min_{\mathcal{K} \in \mathscr{H}_\infty} \operatorname{tr}(\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}(\mathcal{I} + \Lambda)) \tag{11.14}$$

The unique solution to this problem is given by,

$$\mathcal{K}_\Lambda = \mathcal{K}_{\mathscr{H}_2} + \Delta^{-1} \left\{ \{\Delta \mathcal{K}_{nc}\}_{-} \mathcal{L} \right\}_{+} \mathcal{L}^{-1}, \tag{11.15}$$

where Δ and $\mathcal{L}$ are the unique causal and causally invertible canonical spectral factors such that $\Delta^* \Delta = \mathcal{I} + \mathcal{F}^* \mathcal{F}$ and $\mathcal{L}\mathcal{L}^* = \mathcal{I} + \Lambda$.

Proof. Inserting the equation (11.5) in place of $\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}$ and replacing $\mathcal{I} + \mathcal{F}^* \mathcal{F}$ and $\mathcal{I} + \Lambda$ by their corresponding canonical spectral factors, we rewrite (11.15) as

$$\min_{\mathcal{K} \in \mathscr{H}_\infty} \|\Delta \mathcal{K} \mathcal{L} - \Delta \mathcal{K}_{\mathrm{nc}} \mathcal{L}\|_2^2 + \operatorname{tr}(\mathcal{T}_{\mathcal{K}_{\mathrm{nc}}}^* \mathcal{T}_{\mathcal{K}_{\mathrm{nc}}}(\mathcal{I} + \Lambda)). \tag{11.16}$$

Notice that the term $\mathrm{tr}(\mathcal{T}_{\mathcal{K}_{\mathrm{nc}}}^{*}\mathcal{T}_{\mathcal{K}_{\mathrm{nc}}}(\mathcal{I}+\Lambda))$ does not depend on $\mathcal{K}$ and, therefore, can be omitted. Inside the $\mathscr{H}_2$ norm, $\Delta\mathcal{K}\mathcal{L}$ is causal for any $\mathcal{K}$ whereas $\Delta\mathcal{K}_{\mathrm{nc}}\mathcal{L}$ is mixed (non-causal). Therefore, $\Delta\mathcal{K}\mathcal{L}$ can only minimize the objective by matching itself to the causal part of $\Delta\mathcal{K}_{\mathrm{nc}}\mathcal{L}$, [101], [251]. The optimal solution satisfies $\Delta\mathcal{K}_{\Lambda}\mathcal{L} = \{\Delta\mathcal{K}_{\mathrm{nc}}\mathcal{L}\}_{+}$ which yields $\mathcal{K}_{\Lambda} = \Delta^{-1}\{\Delta\mathcal{K}_{\mathrm{nc}}\mathcal{L}\}_{+}\mathcal{L}^{-1}$. By splitting $\Delta\mathcal{K}_{\mathrm{nc}} = \{\Delta\mathcal{K}_{\mathrm{nc}}\}_{+} + \{\Delta\mathcal{K}_{\mathrm{nc}}\}_{-}$ and noting that $\Delta^{-1}\{\Delta\mathcal{K}_{\mathrm{nc}}\}_{+}$ is the $\mathscr{H}_2$-controller [101], we get the desired result. ■

Remark 11.3.3. Notice that $\Lambda = 0$ yields the $\mathscr{H}_2$-optimal solution since $\mathcal{L}\mathcal{L}^{*} = \mathcal{I}$ while any other choice of $\Lambda \succcurlyeq 0$ yields a controller which is a combination of the $\mathscr{H}_2$-controller, $\mathcal{K}_{\mathscr{H}_2}$, and a compensation term $\Delta^{-1}\{\{\Delta\mathcal{K}_{\mathrm{nc}}\}_{-}\mathcal{L}\}_{+}\mathcal{L}^{-1}$. This term accounts for the correlations in the disturbance as dictated by the strictly positive autocovariance operator $\mathcal{I}+\Lambda$.

Given any $\Lambda \succcurlyeq 0$, the inner minimization solution of the Lagrange dual max-min problem in (11.10) can be derived in closed form using (11.15). This allows us to simplify the primal problem over causal operators $\mathcal{K}$ in (11.6), into a dual problem over positive operators $\Lambda \succcurlyeq 0$. In Theorem 11.3.4, we state the optimality conditions for the dual variable $\Lambda \succcurlyeq 0$.

Theorem 11.3.4 (Saddle Point)**.** *Let $\gamma \in (\gamma_2, \gamma_\infty)$ be a fixed admissible $\mathscr{H}_\infty$ norm bound and $(\mathcal{K}_\gamma, \Lambda_\gamma)$ be a saddle point of the max-min problem in* (11.10)*. Moreover, let $\mathcal{L}_\gamma$ and Δ be the unique causal and causally invertible spectral factors of $\mathcal{I}+\Lambda_\gamma = \mathcal{L}_\gamma\mathcal{L}_\gamma^{*}$ and $\mathcal{I}+\mathcal{F}^{*}\mathcal{F} = \Delta^{*}\Delta$, respectively. Then, $(\mathcal{K}_\gamma, \Lambda_\gamma)$ satisfies the following necessary and sufficient conditions*[1]*:*

$$\mathcal{K}_\gamma = \mathcal{K}_2 + \Delta^{-1}\left\{\{\Delta\mathcal{K}_{nc}\}_{-}\mathcal{L}_\gamma\right\}_{+}\mathcal{L}_\gamma^{-1}, \quad \textit{and} \tag{11.17}$$

$$\mathcal{L}_\gamma^{*}\mathcal{L}_\gamma = \max\left\{\gamma^{-2}\left(\mathcal{S}^{*}\mathcal{S} + \mathcal{L}_\gamma^{*}\mathcal{T}_{\mathcal{K}_{nc}}^{*}\mathcal{T}_{\mathcal{K}_{nc}}\mathcal{L}_\gamma\right),\ \mathcal{I}\right\}, \tag{11.18}$$

where $\mathcal{S}_\gamma \coloneqq \{\Delta\mathcal{K}_{nc}\mathcal{L}_\gamma\}_{-}$.

Corollary 11.3.5. *When the disturbances are scalar ($d_w = 1$), we have,*

$$\mathcal{L}_\gamma^{*}\mathcal{L}_\gamma = \max\left\{\frac{\{\Delta\mathcal{K}_{nc}\mathcal{L}_\gamma\}_{-}^{*}\{\Delta\mathcal{K}_{nc}\mathcal{L}_\gamma\}_{-}}{\gamma^2 - \mathcal{T}_{\mathcal{K}_{nc}}^{*}\mathcal{T}_{\mathcal{K}_{nc}}},\ \mathcal{I}\right\}. \tag{11.19}$$

Proof. Existence of a saddle point $(\mathcal{K}_\gamma, \Lambda_\gamma)$ is ensured by Theorem 11.3.1. We can characterize the saddle point by a Banach space analog of Karush-Kuhn-Tucker (KKT) conditions [254, Thm. 2.9.2]:

[1] $\max\{\mathcal{X},\mathcal{Y}\} \triangleq \frac{1}{2}(\mathcal{X}+\mathcal{Y}+|\mathcal{X}-\mathcal{Y}|)$

i. Stationarity: $\mathcal{K}_\gamma \in \arg\min\{\mathrm{tr}(\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}(\mathcal{I}+\Lambda_\gamma)) \mid \mathcal{K} \in \mathscr{H}_\infty\}$,

ii. Primal feasibility: $\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma} - \gamma^2\mathcal{I} \preccurlyeq 0$,

iii. Dual feasibility: $\Lambda_\gamma \succcurlyeq 0$,

iv. Complementary slackness: $\mathrm{tr}(\Lambda_\gamma(\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma} - \gamma^2\mathcal{I})) = 0$.

By the Wiener-Hopf method in Lemma 11.3.2, the stationarity condition (i) immediately implies Eq. (11.17).

To characterize the dual variable Λ_γ, we note that the complementary slackness condition (iv) can be equivalently expressed as

$$\Lambda_\gamma \left(\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma} - \gamma^2\mathcal{I}\right) = 0. \tag{11.20}$$

since $\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma} - \gamma^2\mathcal{I} \preccurlyeq 0$ and $\Lambda_\gamma \succcurlyeq 0$ by primal (ii) and dual (iii) feasibility. Furthermore, this condition can be written alternatively as $(\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma} - \gamma^2\mathcal{I})\Lambda_\gamma = 0$. Therefore, the operators $\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma} - \gamma^2\mathcal{I}$ and Λ_γ commute. This implies that they share the same eigenspace and can be simultaneously unitarily diagonalized, so do the operators $\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma}$ and Λ_γ.

For the optimal controller $\mathcal{K}_\gamma$ in (11.17), we have

$$\Delta\mathcal{K}_\gamma - \Delta\mathcal{K}_{\text{nc}} = \{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_\gamma\}_+\mathcal{L}_\gamma^{-1} - \Delta\mathcal{K}_{\text{nc}}\mathcal{L}_\gamma\mathcal{L}_\gamma^{-1}, \tag{11.21}$$

$$= -\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_\gamma\}_- =: -\mathcal{S}_\gamma. \tag{11.22}$$

Therefore, the quadratic expression $\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma}$ becomes

$$\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma} = (\Delta\mathcal{K}_\gamma - \Delta\mathcal{K}_{\text{nc}})^*(\Delta\mathcal{K}_\gamma - \Delta\mathcal{K}_{\text{nc}}) + \mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}, \tag{11.23}$$

$$= \mathcal{L}_\gamma^{-*}\mathcal{S}^*\mathcal{S}\mathcal{L}_\gamma^{-1} + \mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}. \tag{11.24}$$

Substituting the optimal value of $\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma}$ in (11.20), we get,

$$\Lambda_\gamma \left(\mathcal{L}_\gamma^{-*}\mathcal{S}^*\mathcal{S}\mathcal{L}_\gamma^{-1} + \mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}} - \gamma^2\mathcal{I}\right) = 0. \tag{11.25}$$

Using the identity $\Lambda_\gamma = \mathcal{L}_\gamma\mathcal{L}_\gamma^* - \mathcal{I}$, we get,

$$\mathcal{L}_\gamma\mathcal{S}^*\mathcal{S}\mathcal{L}_\gamma^{-1} + \left(\mathcal{L}_\gamma\mathcal{L}_\gamma^* - \mathcal{I}\right)\left(\mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}} - \gamma^2\mathcal{I}\right) - \mathcal{L}_\gamma^{-*}\mathcal{S}^*\mathcal{S}\mathcal{L}_\gamma^{-1} = 0.$$

Multiplying the identity above by $\mathcal{L}_\gamma$ on the right and by $\mathcal{L}_\gamma^*$ on the left, we get,

$$\mathcal{L}_\gamma^*\mathcal{L}_\gamma\mathcal{S}^*\mathcal{S} + \mathcal{L}_\gamma^*\left(\mathcal{L}_\gamma\mathcal{L}_\gamma^* - \mathcal{I}\right)\left(\mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}} - \gamma^2\mathcal{I}\right)\mathcal{L}_\gamma - \mathcal{S}^*\mathcal{S} = 0.$$

The identity above can be simplified further by factorizing $\mathcal{L}_\gamma^*\mathcal{L}_\gamma - \mathcal{I}$ as follows:

$$\begin{aligned}
0 &= (\mathcal{L}_\gamma^*\mathcal{L}_\gamma - \mathcal{I})\mathcal{S}^*\mathcal{S} + \left(\mathcal{L}_\gamma^*\mathcal{L}_\gamma\mathcal{L}_\gamma^* - \mathcal{L}_\gamma^*\right)\left(\mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}} - \gamma^2\mathcal{I}\right)\mathcal{L}_\gamma, \\
&= (\mathcal{L}_\gamma^*\mathcal{L}_\gamma - \mathcal{I})\mathcal{S}^*\mathcal{S} + \left(\mathcal{L}_\gamma^*\mathcal{L}_\gamma - \mathcal{I}\right)\mathcal{L}_\gamma^*\left(\mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}} - \gamma^2\mathcal{I}\right)\mathcal{L}_\gamma, \\
&= (\mathcal{L}_\gamma^*\mathcal{L}_\gamma - \mathcal{I})\left(\mathcal{S}^*\mathcal{S} + \mathcal{L}_\gamma^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}\mathcal{L}_\gamma - \gamma^2\mathcal{L}_\gamma^*\mathcal{L}_\gamma\right).
\end{aligned} \tag{11.26}$$

Starting with $\mathcal{T}_{\mathcal{K}_\gamma}^*\mathcal{T}_{\mathcal{K}_\gamma} - \gamma^2\mathcal{I} \preccurlyeq 0$ instead, a similar identity can be derived:

$$0 = \left(\mathcal{S}^*\mathcal{S} + \mathcal{L}_\gamma^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}\mathcal{L}_\gamma - \gamma^2\mathcal{L}_\gamma^*\mathcal{L}_\gamma\right)(\mathcal{L}_\gamma^*\mathcal{L}_\gamma - \mathcal{I}). \tag{11.27}$$

Therefore, by (11.26) and (11.27), we have that $\mathcal{S}^*\mathcal{S} + \mathcal{L}_\gamma^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}\mathcal{L}_\gamma$ and $\mathcal{L}_\gamma^*\mathcal{L}_\gamma$ commute, and thus are simultaneously diagonalized.

Defining $\mathcal{X}_\gamma := \gamma^{-2}(\mathcal{S}^*\mathcal{S} + \mathcal{L}_\gamma^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}}\mathcal{L}_\gamma)$ and $\mathcal{N}_\gamma := \mathcal{L}_\gamma^*\mathcal{L}_\gamma$, we can rewrite (11.26) and (11.27) as

$$(\mathcal{N}_\gamma - \mathcal{I})(\mathcal{N}_\gamma - \mathcal{X}_\gamma) = 0, \tag{11.28}$$

with $\mathcal{N}_\gamma - \mathcal{I} \succcurlyeq 0$ by dual feasibility, $\mathcal{N}_\gamma - \mathcal{X}_\gamma \succcurlyeq 0$ by primal feasibility, and commuting operators $\mathcal{N}_\gamma$ and $\mathcal{X}_\gamma$. Since these operators commute, we can solve the quadratic equation (11.28) for $\mathcal{N}_\gamma$ in terms of $\mathcal{X}_\gamma$. This essentially gives,

$$\mathcal{N}_\gamma = \frac{1}{2}\left(\mathcal{X}_\gamma + \mathcal{I} + |\mathcal{X}_\gamma - \mathcal{I}|\right) = \max\left\{\mathcal{X}_\gamma, \mathcal{I}\right\} \tag{11.29}$$

which is the only possible solution of (11.28) that satisfies the primal and dual feasibility constraints. This expression is essentially an operator-theoretic analog of taking the maximum between two elements. Thus, we get (11.18).

When the disturbances are scalar, *i.e.*, $d_w = 1$, all $d_w \times d_w$-block Laurent operators commute with each other. Thus, the expression (11.27) can be simplified further as

$$0 = \left(\mathcal{S}^*\mathcal{S} + \mathcal{L}_\gamma^*\mathcal{L}_\gamma(\mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}} - \gamma^2)\right)(\mathcal{L}_\gamma^*\mathcal{L}_\gamma - \mathcal{I}). \tag{11.30}$$

Upon solving it for $\mathcal{L}_\gamma^*\mathcal{L}_\gamma$ in the same fashion, we immediately get (11.19)

■

11.4 A Fixed-Point Iteration Algorithm

Henceforth in this paper, we focus our analysis on the case when $d_w = 1$, *i.e.*, scalar disturbances. In subsection 11.4, we first demonstrate that the Karush-Kuhn-Tucker (KKT) conditions outlined in equation (11.19) can be exclusively defined by a finite-dimensional parameter, $\overline{B}_\gamma$, within the frequency domain. However, the optimal controller lacks a rational nature, preventing it from being represented through a finite-dimensional state-space model. Following this, in subsection 11.4, we introduce a fixed-point iteration method for any given $\gamma \in (\gamma_\infty, \gamma_2)$. This method is designed to determine $\overline{B}_\gamma$, enabling the calculation of the optimal controller, $K_\gamma(e^{\jmath\omega})$ in the frequency domain.

Finite-Dimensional Parameterization of the Optimal Controller

In the following Theorem 11.4.1, we establish that the strictly anticausal transfer function $S_\gamma(e^{\jmath\omega})$ defined below can be expressed using a finite-dimensional state-space model. This theorem indicates that the right-hand side of (11.34) for $N_\gamma(e^{\jmath\omega})$, involving the square root of the rational term $S_\gamma(e^{\jmath\omega})^* S_\gamma(e^{\jmath\omega})$, is no longer rational due to the square root operation. This observation leads us to Corollary 11.4.2.

Theorem 11.4.1 (Frequency-domain solution)**.** *Define* $S_\gamma(e^{\jmath\omega}) \coloneqq \{\Delta K_\circ L_\gamma\}_{-}(e^{\jmath\omega})$. *Then,*

$$S_\gamma(e^{\jmath\omega}) = \overline{C}(e^{-\jmath\omega} I - \overline{A})^{-1}\overline{B}_\gamma, \tag{11.31}$$

$$\textit{with} \quad \overline{B}_\gamma = \frac{1}{2\pi}\int_0^{2\pi} (I - e^{\jmath\omega}\overline{A})^{-1}\overline{D} L_\gamma(e^{\jmath\omega}) d\omega. \tag{11.32}$$

The optimal controller in the frequency domain is given by,

$$K_\gamma(e^{\jmath\omega}) = K_{nc}(e^{\jmath\omega}) - \Delta^{-1}(e^{\jmath\omega}) S_\gamma(e^{\jmath\omega}) L_\gamma^{-1}(e^{\jmath\omega}), \tag{11.33}$$

$$N_\gamma(e^{\jmath\omega}) = \max\left\{ \frac{\{\Delta\mathcal{K}_{nc}\mathcal{L}_\gamma\}_{-}^{*}(e^{\jmath\omega})\{\Delta\mathcal{K}_{nc}\mathcal{L}_\gamma\}_{-}(e^{\jmath\omega})}{\gamma^2 - T_{\mathcal{K}_{nc}}(e^{\jmath\omega})^* T_{\mathcal{K}_{nc}}(e^{\jmath\omega})}, 1 \right\}, \tag{11.34}$$

where,

$$K_{lqr} \coloneqq (R + B_u^\top P B_u)^{-1} B_u^\top P A \tag{11.35}$$

$$P = Q + A^\top P A - A^\top P B_u (R + B_u^\top P B_u)^{-1} B_u^\top P A \tag{11.36}$$

$$\overline{A} \coloneqq \left(A - B_u K_{lqr}\right)^\top \tag{11.37}$$

$$\overline{C} \coloneqq -(R + B_u^\top P B_u)^{-\top/2} B_u^\top \tag{11.38}$$

$$\overline{D} \coloneqq \left(A - B_u K_{lqr}\right)^\top P B_w \tag{11.39}$$

$$N_\gamma(e^{\jmath\omega}) \coloneqq L_\gamma(e^{\jmath\omega})^* L_\gamma(e^{\jmath\omega}). \tag{11.40}$$

Proof. Using the identity $\{\mathcal{X}\}_{+} = \mathcal{X} - \{\mathcal{X}\}_{-}$, we restate the KKT equations (11.19) in the frequency domain as in (11.33), (11.34).

We introduce the Linear Quadratic Regulator (LQR) controller K_{lqr}, the corresponding closed-loop matrix $A - B_u K_{\text{lqr}}$, and the unique solution to the LQR Riccati equation that stabilizes the system, $P \succ 0$ to write $S_\gamma(e^{\jmath\omega})$ in (11.31). The Riccati equation emerges from the spectral factorization of $\Delta^*\Delta = \mathcal{I} + \mathcal{F}^*\mathcal{F}$. See [125, Lemma 11] for example. ■

Corollary 11.4.2. *For any given $\gamma \in (\gamma_\infty, \gamma_2)$, both $N_\gamma(e^{\jmath\omega})$ and the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller $K_\gamma(e^{\jmath\omega})$, are characterized as irrational. Consequently, $K_\gamma(e^{\jmath\omega})$ cannot be realized by a finite-dimensional state-space model.*

Although $K_\gamma(e^{\jmath\omega})$ cannot be modeled in a finite-dimensional state-space, Lemma 11.4.1 introduces a finite-dimensional parameterization for $N_\gamma(e^{\jmath\omega})$ via $\overline{B}_\gamma$. Theorem 11.4.3 further verifies that $\overline{B}_\gamma$ directly determines $N_\gamma(e^{\jmath\omega})$, and by extension, the suboptimal controller $K_\gamma(e^{\jmath\omega})$.

Theorem 11.4.3 (Fixed-Point Solution)**.** *Assuming $d_w = 1$, $\gamma \in (\gamma_\infty, \gamma_2)$, we consider a sequence of mappings:*

$$F_1 : \overline{B} \mapsto S(e^{\jmath\omega}) = \overline{C}(e^{-\jmath\omega} I - \overline{A})^{-1}\overline{B} \tag{11.41}$$

$$F_{2,\gamma} : S(e^{\jmath\omega}) \mapsto N_\gamma(e^{\jmath\omega})$$
$$= \max\left\{ \frac{S(e^{\jmath\omega})^* S(e^{\jmath\omega})}{\gamma^2 - T_{\mathcal{K}_{nc}}(e^{\jmath\omega})^* T_{\mathcal{K}_{nc}}(e^{\jmath\omega})}, 1 \right\}. \tag{11.42}$$

$$F_3 : N(e^{\jmath\omega}) \mapsto L(e^{\jmath\omega}) \tag{11.43}$$

$$F_4 : L(e^{\jmath\omega}) \mapsto \overline{B} = \frac{1}{2\pi}\int_0^{2\pi} (I - e^{\jmath\omega}\overline{A})^{-1}\overline{D}L(e^{\jmath\omega})d\omega. \tag{11.44}$$

where F_3 produces a unique spectral factor of $N(e^{\jmath\omega}) > 0$. The composite function $F_4 \circ F_3 \circ F_{2,\gamma} \circ F_1 : \overline{B} \mapsto \overline{B}$ has a unique fixed point $\overline{B}_\gamma$, with $N_\gamma(e^{\jmath\omega}) \equiv F_{2,\gamma} \circ F_1(\overline{B}_\gamma)$ fulfilling the KKT conditions (11.19).

Proof. Consider the optimality condition (11.34). Note that for $\gamma > \gamma_\infty$, $N_\gamma(e^{\jmath\omega})$ in (11.34) is well-defined. Thus, the map $F_4 \circ F_3 \circ F_{2_\gamma} \circ F_1 : \overline{B} \mapsto \overline{B}$ described earlier admits a fixed point $\overline{B}_\gamma$ for a fixed γ. Since (11.13) is concave in Λ (or $\mathcal{L}\mathcal{L}^* = \Lambda + \mathcal{I}$), the optimal solution $\mathcal{N}_\gamma$ is unique. Given that $M_\gamma(e^{\jmath\omega}) = L_\gamma(e^{\jmath\omega})L_\gamma^*(e^{\jmath\omega})$, where $L_\gamma(e^{\jmath\omega})$ represents a spectral factor of $M_\gamma(e^{\jmath\omega})$ that is both causal and causally invertible, it follows that $L_\gamma(e^{\jmath\omega})$ is uniquely determined apart from a unitary transformation. By establishing a specific choice for the unitary transformation during the spectral factorization process, for example, opting for positive-definite factors at infinity as outlined by [59], we ensure the uniqueness of $L_\gamma(e^{\jmath\omega})$. Consequently, with $\overline{A}$ and $\overline{D}$ being constants, the expression for $\overline{B}_\gamma = \frac{1}{2\pi}\int_0^{2\pi}(I - e^{\jmath\omega}\overline{A})^{-1}\overline{D}L_\gamma(e^{\jmath\omega})d\omega$ is also uniquely determined. ■

Algorithm

Now that we know a fixed point solution exists to our problem, we can use Theorem 11.4.3 to obtain the following iterative fixed-point algorithm. Once we have an optimal $N_\gamma(e^{j\omega})$, we can find the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller $K_\gamma(\mathrm{e}^{j\omega})$ using (11.33).

Algorithm 12 FixedMHH: Dual Fixed-Point Iteration via Spectral Factorization

---

Input: $\gamma > \gamma_{H_\infty}$, initialise $\overline{B}$, system $(\overline{A}, \overline{D}, \overline{C})$
repeat
  $N_\gamma^{(n)}(\mathrm{e}^{j\omega}) \leftarrow F_{2,\gamma} \circ F_1(\overline{B}^{(n)})$
  $L_\gamma^{(n+1)}(\mathrm{e}^{j\omega}) \leftarrow \texttt{SpectralFactor}(N_\gamma^{(n)}(\mathrm{e}^{j\omega}))$
  $\overline{B}^{n+1} = F_4(L_\gamma^{(n+1)}(\mathrm{e}^{j\omega}))$
  $n \leftarrow n+1$
until convergence of $N_\gamma^{(n)}(\mathrm{e}^{j\omega})$

11.5 Convergence Analysis

In this section, we provide a proof of convergence of FixedMHH for the particular case when $d_x = d_w = 1$. We first show the fixed point, that Algorithm 12 converges to, is unique when $\gamma > \gamma_\infty$. Then, we show that the iterates produce a sequence of monotonically increasing spectra $\mathcal{N}$. Due to this, the algorithm must converge to the unique fixed point. Formally, we state this in the following lemma.

Lemma 11.5.1 (Monotonicity). *For $d_x = d_w = 1$, the composite mapping $F_{2,\gamma} \circ F_1 \circ F_4 \circ F_3 : N(\mathrm{e}^{j\omega}) \mapsto N_\gamma(\mathrm{e}^{j\omega})$ is monotonic, i.e., for any two positive power spectral densities such that $0 < N_1(e^{j\omega}) \leq N_2(e^{j\omega})$ for all $\omega \in [0, 2\pi)$, the mapping $F_{2,\gamma} \circ F_1 \circ F_4 \circ F_3$ preserves their order.*

Proof. Let $\gamma > \gamma_\infty$. Consider now two spectra $0 \prec \mathcal{N}_1 \preccurlyeq \mathcal{N}_2$ that are represented in the frequency domain as $N_1(e^{j\omega}) \leq N_2(e^{j\omega})$ for all $\omega \in [0, 2\pi)$. Now the spectra $\mathcal{N}_1, \mathcal{N}_2$ are passed through one iteration of Algorithm 12, *i.e.*, $F_3 \circ F_4 \circ F_1 \circ F_{2_\gamma} \circ : \mathcal{N} \mapsto \overline{\mathcal{N}}$ to get $\overline{\mathcal{N}_1}, \overline{\mathcal{N}_2}$ respectively. We want to show that $\overline{n}_1(e^{j\omega}) \leq \overline{n}_2(e^{j\omega}) \forall \omega$ *i.e.*, Algorithm 12 preserves the order of $\mathcal{N}_1 \preccurlyeq \mathcal{N}_2$. We have that

$$\overline{N_1}(\mathrm{e}^{j\omega}) = \max\left\{1, \frac{\{\Delta\mathcal{K}_{\mathrm{nc}}\mathcal{L}_1\}_-^*(\mathrm{e}^{j\omega})\{\Delta\mathcal{K}_{\mathrm{nc}}\mathcal{L}_1\}_-(\mathrm{e}^{j\omega})}{\gamma^2 - T_{\mathcal{K}_{\mathrm{nc}}}(\mathrm{e}^{j\omega})^* T_{\mathcal{K}_{\mathrm{nc}}}(\mathrm{e}^{j\omega})}\right\} \tag{11.45}$$

$$\overline{N_2}(\mathrm{e}^{j\omega}) = \max\left\{1, \frac{\{\Delta\mathcal{K}_{\mathrm{nc}}\mathcal{L}_2\}_-^*(\mathrm{e}^{j\omega})\{\Delta\mathcal{K}_{\mathrm{nc}}\mathcal{L}_2\}_-(\mathrm{e}^{j\omega})}{\gamma^2 - T_{\mathcal{K}_{\mathrm{nc}}}(\mathrm{e}^{j\omega})^* T_{\mathcal{K}_{\mathrm{nc}}}(\mathrm{e}^{j\omega})}\right\}. \tag{11.46}$$

Moreover, denoting by $\mathcal{N}_1 = \mathcal{L}_1^*\mathcal{L}_2$ and $\mathcal{N}_2 = \mathcal{L}_2^*\mathcal{L}_2$ the unique spectral factorizations, we have that

$$\begin{aligned}
\operatorname{tr}\left(\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_1\}_-^*\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_1\}_-\right) &\overset{(a)}{=} \inf_{\mathcal{K}\in\mathscr{H}_\infty} \operatorname{tr}((\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} - \mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}})\mathcal{N}_1) \\
&\overset{(b)}{\leq} \inf_{\mathcal{K}\in\mathscr{H}_\infty} \operatorname{tr}((\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} - \mathcal{T}_{\mathcal{K}_{\text{nc}}}^*\mathcal{T}_{\mathcal{K}_{\text{nc}}})\mathcal{N}_2) \qquad (11.47) \\
&\overset{(c)}{=} \operatorname{tr}\left(\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_2\}_-^*\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_2\}_-\right),
\end{aligned}$$

where the identities (a), (c) are due to Wiener-Hopf technique in Lemma 11.3.2 and (b) is due to $\mathcal{N}_1 \preccurlyeq \mathcal{N}_2$.

Notice that, for $d_x = d_w = 1$, we have that

$$\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_1\}_-^*\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_1\}_- = |\overline{C}|^2|\overline{B}_1|^2|\mathrm{e}^{-j\omega}I - \overline{A}|^{-2}, \qquad (11.48)$$
$$\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_2\}_-^*\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_2\}_- = |\overline{C}|^2|\overline{B}_2|^2|\mathrm{e}^{-j\omega}I - \overline{A}|^{-2}, \qquad (11.49)$$

and therefore, their traces take the form

$$\begin{aligned}
&\operatorname{tr}\left(\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_1\}_-^*\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_1\}_-\right) \\
&\qquad = |\overline{C}|^2|\overline{B}_1|^2 \int_0^{2\pi} |\mathrm{e}^{-j\omega}I - \overline{A}|^{-2}\frac{d\omega}{2\pi}, \qquad (11.50) \\
&\operatorname{tr}\left(\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_2\}_-^*\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_2\}_-\right) \\
&\qquad = |\overline{C}|^2|\overline{B}_2|^2 \int_0^{2\pi} |\mathrm{e}^{-j\omega}I - \overline{A}|^{-2}\frac{d\omega}{2\pi}. \qquad (11.51)
\end{aligned}$$

By this observation and using the inequality (11.47), we have that $|\overline{B}_1|^2 \leq |\overline{B}_1|^2$ and thus the monotonicity of the traces implies the monotonicity of the operators, *i.e.*, for all $\omega \in [0, 2\pi)$, the following holds

$$\begin{aligned}
&\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_1\}_-^*(\mathrm{e}^{\jmath\omega})\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_1\}_-(\mathrm{e}^{\jmath\omega}) \\
&\qquad = |\overline{C}|^2|\overline{B}_1|^2|\mathrm{e}^{-j\omega}I - \overline{A}|^{-2}, \qquad (11.52) \\
&\qquad \leq |\overline{C}|^2|\overline{B}_2|^2|\mathrm{e}^{-j\omega}I - \overline{A}|^{-2}, \qquad (11.53) \\
&\qquad = \{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_2\}_-^*(\mathrm{e}^{\jmath\omega})\{\Delta\mathcal{K}_{\text{nc}}\mathcal{L}_2\}_-(\mathrm{e}^{\jmath\omega}). \qquad (11.54)
\end{aligned}$$

Hence, we have that $\overline{N_1}(\mathrm{e}^{j\omega}) \leq \overline{N_2}(\mathrm{e}^{j\omega})$ for all $\omega \in [0, 2\pi)$ using (11.45) and (11.46). ■

We can now initialise the algorithm with $\overline{B}^{(0)} = 0$ or $\overline{N}^{(0)}(\mathrm{e}^{j\omega}) = 1$. Since after each iteration, $\overline{N}^{(n)}(\mathrm{e}^{j\omega}) \geq 1$ for all $\omega \in [0, 2\pi)$, FixedMHH generates a monotonically increasing sequence of $\{\overline{N}^{(n)}(\mathrm{e}^{j\omega})\}$ for all $\omega \in [0, 2\pi)$, which converges to the unique fixed point. We state this formally in the following theorem.

Theorem 11.5.2. *For $d_x = d_w = 1$ and $\overline{B}^{(0)} = 0$, the sequence of iterates $\{\overline{N}^{(n)}(\mathrm{e}^{j\omega})\}$ generated by FixedMHH is monotonically increasing and converges to the optimal solution in* (11.19).

Proof. The proof follows directly from the repeated application of monotonicity result on Lemma 11.5.1 and [249, Thm. 2] ■

Remark 11.5.3. Although we present a proof of convergence for the particular case of scalar systems $d_x = 1$, empirical evidence suggests (see Section 11.7) that the algorithm is exponentially convergent with a faster rate of convergence for larger $\gamma > 0$.

11.6 Rational Approximation

This section describes a practical approach to devising state-space controllers that serve as approximations to our irrational γ-optimal controller (11.33). Rather than attempting to approximate the controller directly, we *approximate the power spectrum* $N(\mathrm{e}^{\jmath\omega})$, aiming to reduce the $\mathscr{H}_\infty$-norm of the approximation error through the use of positive rational functions. While the problem of approximating with rational functions typically does not lead to convex formulations, we demonstrate in Theorem 11.6.3 that the process of approximating positive power spectra through the use of ratios of positive fixed-order polynomials can indeed be framed as a convex feasibility problem. The problem is stated as follows:

Problem 11.6.1 (Rational Approximation via $\mathscr{H}_\infty$-norm)**.** For a given positive spectrum $\mathcal{N}$, identify the optimal rational approximation utilizing the $\mathscr{H}_\infty$ norm, with an order not exceeding $m \in \mathbb{N}$. Specifically,

$$\inf_{\mathcal{P},\mathcal{Q} \in \mathscr{T}_+^m} \|\mathcal{P}/\mathcal{Q} - \mathcal{N}\|_\infty \text{ subject to } \operatorname{tr}(\mathcal{Q}) = 1, \tag{11.55}$$

where $\mathscr{T}_{m,+}$ is the set of positive symmetric polynomials of order less than or equal to m and the constraint $\operatorname{tr}(\mathcal{Q}) = 1$ is to avoid redundancy in solutions.

Definition 11.6.2. Given an $\epsilon > 0$ approximation bound, the ϵ-sublevel set of Problem 11.6.1 is defined as

$$\mathscr{S}_\epsilon := \{(\mathcal{P}, \mathcal{Q}) \mid \|\mathcal{P}/\mathcal{Q} - \mathcal{N}\|_\infty \leq \epsilon,\ \operatorname{tr}(\mathcal{Q}) = 1\}$$

Theorem 11.6.3 (Feasibility of $\mathscr{S}_\epsilon$, [124, Thm. 5.5])**.** *Given an accuracy level $\epsilon > 0$ and $m \in \mathbb{N}$ is a fixed order, the polynomials $\mathcal{P}$ and $\mathcal{Q}$ of order m belong to the ϵ-sub-level set, i.e.$(\mathcal{P}, \mathcal{Q}) \in \mathscr{S}_\epsilon$ if and only if there exists $\mathbf{P}, \mathbf{Q} \in \mathbb{S}_+^{m+1}$ such that* $\operatorname{tr}(\mathbf{Q}) = 1$ *and for all $\omega \in [0, 2\pi)$, .*

$$1)\ \operatorname{Tr}(\mathbf{P}\boldsymbol{\Theta}(\mathrm{e}^{\jmath\omega})) - (N(\mathrm{e}^{\jmath\omega}) + \epsilon)\operatorname{Tr}(\mathbf{Q}\boldsymbol{\Theta}(\mathrm{e}^{\jmath\omega})) \leq 0, \tag{11.56}$$

$$2)\ \operatorname{Tr}(\mathbf{P}\boldsymbol{\Theta}(\mathrm{e}^{\jmath\omega})) - (N(\mathrm{e}^{\jmath\omega}) - \epsilon)\operatorname{Tr}(\mathbf{Q}\boldsymbol{\Theta}(\mathrm{e}^{\jmath\omega})) \geq 0, \tag{11.57}$$

Although the above equations hold for all frequencies, practical implementation necessitates focusing on a finite selection of frequency samples. To sidestep this limitation, the analysis can be limited to a select set of frequencies, defined as $\Omega_N = \{\omega = 2\pi k/N | k = 1, \ldots, N\}$, where N, chosen to be much larger than m, provides a dense sampling of the frequency domain. While this approach inherently approximates the full spectrum of frequencies, increasing the number of sampled frequencies N allows for an arbitrary improvement in the approximation's accuracy. Consequently, this method facilitates the transformation of the rational function approximation problem into a convex feasibility problem, addressable through the application of Linear Matrix Inequalities (LMIs) alongside a finite collection of affine (in)equality constraints. Upon achieving a rational approximation of $N(\mathrm{e}^{\jmath\omega})$, we then derive a state-space controller as outlined in (11.17). Note that a canonical factor of the rational approximation of $N(\mathrm{e}^{\jmath\omega})$ can always be found due to the following lemma.

Lemma 11.6.4 (Canonical Factorization [199, Lem. 1])**.** *Given a Laurent polynomial of order m, $P(z) = \sum_{k=-m}^{m} p_k z^{-k}$, where $p_k = p_{-k} \in \mathbb{R}$, and $P(\mathrm{e}^{\jmath\omega}) > 0$, it can be shown that a canonical factor $L(z) = \ell_0 + \ell_1 z^{-1} + \ldots + \ell_m z^{-m}$ exists. This factor satisfies $P(\mathrm{e}^{\jmath\omega}) = |L(\mathrm{e}^{\jmath\omega})|^2$, with all roots of $L(z)$ lying inside the unit circle.*

11.7 Numerical Results

In this section, we analyze the performance of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller and the rational approximation method for 2 systems. We use benchmark models from [145]. In particular, we test the aircraft system [AC17] and a chemical reactor system [REA4]. First, we present frequency domain plots for each system, highlighting their performances. The plots show the performance of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ and the rational approximations. Additional data is presented in Tables 11.1 and 11.2. Next, we provide numerical evidence supporting the exponential convergence of the proposed algorithm Algorithm 12. Note that in this section, for brevity, a rational approximation of order m is denoted $RA(m)$.

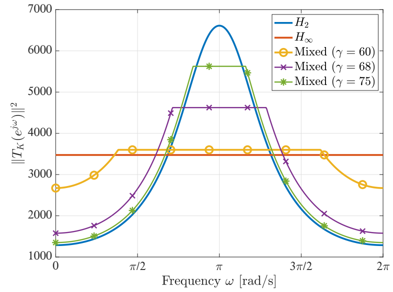


Figure 11.1: The spectral norm, $\overline{\sigma}(T_K^*(\mathrm{e}^{\jmath\omega})T_K(\mathrm{e}^{\jmath\omega}))$ of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller for $\gamma \in \{60, 68, 75\}$ at different frequency values, for the system [AC17]. The cost of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller follows H_2 and clips at the threshold γ for some frequencies.

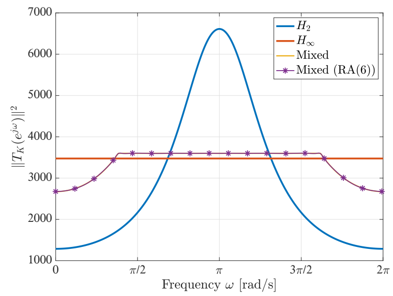


Figure 11.2: The spectral norm, $\overline{\sigma}(T_K^*(\mathrm{e}^{\jmath\omega})T_K(\mathrm{e}^{\jmath\omega}))$ of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller ($\gamma = 60$) and a 6^{th} order rational approximation at different frequency values, for the system [AC17]. The cost of the rational controller closely follows the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller.

We first present the performance metrics for the [AC17] system. For this system, $\gamma_\infty = 58.94$ and $\gamma_2 = 81.309$. Thus, the $\mathscr{H}_\infty$ norm of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller $\gamma \in (58.94, 81.309]$. The system is a 4^{th} order system. The system matrices are,

$$A = \begin{bmatrix} -2.98 & .93 & 0 & -.034 \\ -.99 & -.21 & .035 & -.001 \\ 0 & 0 & 0 & 1 \\ .39 & -5.55 & 0 & -1.89 \end{bmatrix} B_u = \begin{bmatrix} -.032 \\ 0 \\ 0 \\ -1.6 \end{bmatrix}.$$

We present the frequency domain plots in Figures 11.1 and 11.2. Note that $\|\mathcal{T}_\mathcal{K}\|_\infty^2 = \max_{0\le\omega\le 2\pi} \overline{\sigma}(T_K^*(\mathrm{e}^{\jmath\omega})T_K(\mathrm{e}^{\jmath\omega}))$. This metric is visualised in Figure 11.1 which highlights the mixed nature of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller. For γ close to γ_∞, the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller closely follows the H_∞ controller for most of the frequencies but still has less area under the curve ($\|\mathcal{T}_\mathcal{K}\|_2$) and seems to follow the H_2 controller. As we increase γ, we see that the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller clips the value of $\|\mathcal{T}_\mathcal{K}\|_\infty$ at γ for some frequencies, and for the other frequencies, it follows

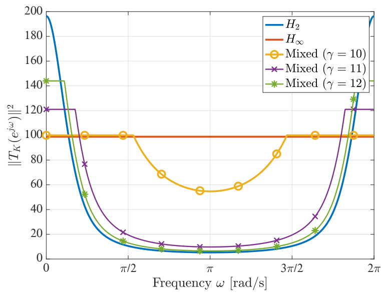


Figure 11.3: The spectral norm, $\overline{\sigma}(T_K^*(\mathrm{e}^{\jmath\omega})T_K(\mathrm{e}^{\jmath\omega}))$ of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller for $\gamma \in \{10, 11, 12\}$ at different frequency values, for system [REA4]. The cost of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller follows H_2 and clips at the threshold γ for some frequencies.

the H_2 controller which is to minimise $\|\mathcal{T}_\mathcal{K}\|_2$. We now focus on the performance of the rational approximations of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller. Since we approximate the spectrum $\mathcal{N}$, the order of controller is given by the order of the system plus the order of the rational approximation. Table 11.1 highlights the performance metrics of the rational approximations. Moreover, Figure 11.2 showcases how a rational approximation looks in the frequency domain. One can observe from Figure 11.2 that $\|\mathcal{T}_\mathcal{K}\|_\infty$ for the rational approximation can be slightly higher than that of the actual mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller. As can be seen from Table 11.1, a higher order approximation results in a controller with performance metrics close the the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller. For the [AC17] system, a 6^{th} order approximation of the spectrum $\mathcal{N}$ provides seems to be a good choice for the controller.

	$\gamma = 60$		$\gamma = 68$	
	$\|\mathcal{T}_\mathcal{K}\|_2$	$\|\mathcal{T}_\mathcal{K}\|_\infty$	$\|\mathcal{T}_\mathcal{K}\|_2$	$\|\mathcal{T}_\mathcal{K}\|_\infty$
H_∞	58.94	58.94	58.94	58.94
Mixed $\mathscr{H}_2/\mathscr{H}_\infty$	57.92	60	54.94	68
RA(1)	58.14	60.36	54.94	69.46
RA(3)	58.04	60.42	54.95	68.31
RA(6)	57.92	60.07	54.94	68.07
H_2	54.28	81.309	54.28	81.309

Table 11.1: The performance characteristics of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller obtained from degree 1, 2, and 3 rational approximations to $N(\mathrm{e}^{\jmath\omega})$.

We now present the numerical results for the system [REA4]. This is an 8^{th} order system which is a chemical reactor system. The system matrices can be found in [145]. As in the case of [AC17], we see in Figure 11.3, that the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller closely follows the H_∞ controller for most of the frequencies but still has

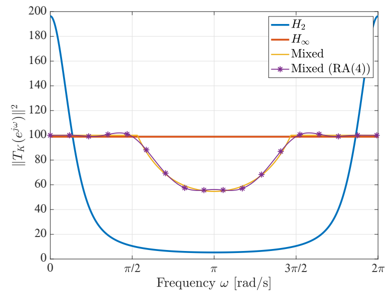


Figure 11.4: The spectral norm, $\overline{\sigma}(T_K^*(e^{\jmath\omega})T_K(e^{\jmath\omega}))$ of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller ($\gamma = 10$) and a 4^{th} order rational approximation at different frequency values, for system [REA4]. The cost of the rational controller closely follows the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller.

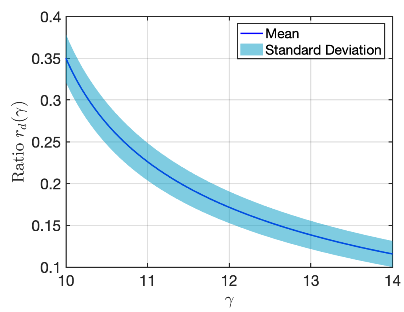


Figure 11.5: The variation of $r_d(\gamma)$ (defined in (11.58)) with γ for the [REA4] system. The plot indicates different $\mathcal{N}_1, \mathcal{N}_2$ in (11.58) chosen at random. Note that the contraction ratio is always less than 1 and decreases with an increase in γ.

less area under the curve ($\|\mathcal{T}_\mathcal{K}\|_2$) and seems to follow the H_2 controller for some part of the frequencies. The performance of a controller obtained via a 4^{th} order rational approximation of the spectrum $\mathcal{N}$ is shown in Figure 11.4. As expected, a higher order approximation results in a controller that better approximates the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller. For the [REA4] system, a controller obtained via an 8^{th} order approximation well approximates the optimal mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller, as shown in Table 11.2.

We now present numerical evidence that suggests an exponential convergence of our iterative method Algorithm 12. In Figure 11.5, we consider the ratio,

$$r_d(\gamma) = \frac{d(\mathcal{N}_1^{(1)}, \mathcal{N}_2^{(1)})}{d(\mathcal{N}_1^{(0)}, \mathcal{N}_2^{(0)})}. \tag{11.58}$$

Here, $\mathcal{N}_1(e^{\jmath\omega})$ and $\mathcal{N}_2(e^{\jmath\omega})$ are the frequency domain representation of the power spectra $\mathcal{N}_1$ and $\mathcal{N}_2$. $N^{(1)}(e^{\jmath\omega})$ is the spectrum obtained after one iteration of Algorithm 12 initialized with $N(e^{\jmath\omega})$, and $d(\cdot)$ is a distance metric. For our results,

	$\gamma = 10$		$\gamma = 12$	
	$\|\mathcal{T}_{\mathcal{K}}\|_2$	$\|\mathcal{T}_{\mathcal{K}}\|_\infty$	$\|\mathcal{T}_{\mathcal{K}}\|_2$	$\|\mathcal{T}_{\mathcal{K}}\|_\infty$
H_∞	9.94	9.94	9.94	9.94
Mixed $\mathscr{H}_2/\mathscr{H}_\infty$	9.2	10	6.17	12
RA(2)	9.3	10.05	6.21	12.09
RA(4)	9.21	10.1	6.177	12.02
RA(8)	9.2	10.008	6.17	12.01
H_2	6.06	14.01	6.06	14.01

Table 11.2: The performance characteristics of the mixed $\mathscr{H}_2/\mathscr{H}_\infty$ controller obtained from degree 1, 2, and 3 rational approximations to $N(\mathrm{e}^{\jmath\omega})$.

we consider the H_∞ norm given by,

$$d(\mathcal{N}_1, \mathcal{N}_2) = \max_{\omega\in[0,2\pi]} \overline{\sigma}(N_1(\mathrm{e}^{\jmath\omega}) - N_2(\mathrm{e}^{\jmath\omega})). \tag{11.59}$$

We consider various random initializations of the $\mathcal{N}_1$ and $\mathcal{N}_2$. What we would like to observe is $r_d(\gamma) < 1$ for all $\gamma > \gamma_\infty$. Since this would mean that after each iteration, the spectrum is close to the optimal solution. As can be seen in Figure 11.5, the contraction ratio is indeed a decreasing function of γ and is always less than 1. This suggests that Algorithm 12 is exponentially convergent.

11.8 Conclusion

In this paper, we studied the problem of mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control in the infinite-horizon setting. We provide the exact closed-form solution to the infinite-horizon mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control in the frequency domain. Despite being non-rational, we show that the optimal controller admits a finite-dimensional parameterization. Leveraging this fact, we introduce an efficient iterative algorithm that finds the optimal causal controller in the frequency domain. We show that this algorithm is convergent when the system is scalar and present numerical evidence for exponential convergence of the proposed algorithm. To obtain a finite order controller, we use a rational approximation method (based on the H_∞ norm) and present its performance. In future works, we will extend our results to cases when $d_w > 1$. As mentioned in the paper, numerical results hint that the algorithm is exponentially convergent. Future works will involve analyzing the convergence properties of the proposed iterative algorithm.

Chapter 12

STRONG DUALITY

Our primary goal in this section is to construct a dual problem for such a general class of primal control problems by leveraging the Fenchel conjugate of convex functions. The main advantage of duality is that the complicated optimization over causal controllers of an arbitrary convex objective can be transformed into a minimization over a simpler quadratic objective at the expense of an additional maximization over a dual variable.

While Assumption 8.2.4 makes (P) a convex program, it is still an infinite-dimensional problem. However, we are not yet concerned with infinite-dimensionality of (P) and we will address effective numerical computation of the optimal solution in **??**. Until then, our primary focus is to derive a tractable set of necessary and sufficient conditions for the optimal solution of the infinite-dimensional convex program (P).

In its current form, (P) is not amenable for deriving the optimality conditions for the controller $\mathcal{K}$. The major roadblock sitting to achieving this is the non-triviality of the causality constraint on the controller, even for very simple objective functions. In fact, in the full-information setting, there exists a universally optimal non-causal controller minimizing any proper, convex, and monotonically decreasing function f:

Example 12.0.1 (Causality is non-trivial)**.** Consider the full-information setting where $\mathcal{T}_{\mathcal{K}} = \begin{bmatrix} \mathcal{P}_{xu}\mathcal{K} + \mathcal{P}_{xw} \\ \mathcal{K} \end{bmatrix}$. Suppose $f : \mathscr{L}_1 \to \overline{\mathbb{R}}$ satisfy Assumption 8.2.4 and there is at least one $\mathcal{K}' \in \mathscr{L}_\infty$, which can be non-causal, such that $f(\mathcal{T}_{\mathcal{K}'}^* \mathcal{T}_{\mathcal{K}'}) < \infty$. By completion-of-squares, we get

$$\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}} = (\Delta\mathcal{K} - \Delta\mathcal{K}_{\text{nc}})^* (\Delta\mathcal{K} - \Delta\mathcal{K}_{\text{nc}}) + \mathcal{T}_{\mathcal{K}_{\text{nc}}}^* \mathcal{T}_{\mathcal{K}_{\text{nc}}} \succcurlyeq \mathcal{T}_{\mathcal{K}_{\text{nc}}}^* \mathcal{T}_{\mathcal{K}_{\text{nc}}}, \quad \forall \mathcal{K} \in \mathscr{L}_\infty \tag{12.1}$$

where $\mathcal{K}_{\text{nc}} \coloneqq -(\mathcal{I} + \mathcal{P}_{xu}^* \mathcal{P}_{xu})^{-1} \mathcal{P}_{xu}^* \mathcal{P}_{xw}$ is the unique $\mathscr{H}_2$-optimal non-causal controller and $\Delta^*\Delta = \mathcal{I} + \mathcal{P}_{xu}^* \mathcal{P}_{xu}$ is the canonical spectral factorization. Since $\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}$ is dominated by $\mathcal{T}_{\mathcal{K}_{\text{nc}}}^* \mathcal{T}_{\mathcal{K}_{\text{nc}}}$, we have that $f(\mathcal{T}_{\mathcal{K}_{\text{nc}}}^* \mathcal{T}_{\mathcal{K}_{\text{nc}}}) \leq f(\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}})$ for all $K \in \mathscr{L}_\infty$ and $f(\mathcal{T}_{\mathcal{K}_{\text{nc}}}^* \mathcal{T}_{\mathcal{K}_{\text{nc}}})$ is finite. Thus, $\mathcal{K}_{\text{nc}}$ is the universally optimal non-causal controller.

This example highlights the inherent complexity of addressing causality constraints. Causality, combined with the infinite-dimensional nature of resulting optimization problems, is arguably what sets optimal control apart as significantly more challenging than static optimization. Indeed, beyond classical problems like $\mathscr{H}_2$ and $\mathscr{H}_\infty$ control, and the more recent regret-optimal control framework, there remains a substantial gap in our ability to systematically solve generalized control problems of the form presented in Problem 8.2.3. This profound challenge is exemplified by the absence of an analytic closed-form solution for the renowned mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control problem, despite the fact that the pure $\mathscr{H}_2$ and $\mathscr{H}_\infty$ problems have been solved and thoroughly understood for decades.

The primary objective of this section is to systematically derive a tractable dual reformulation of Problem 8.2.3, thereby addressing the challenges associated with solving the causality-constrained primal problem (P). The cornerstone of this dual reformulation is the Fenchel conjugate of the convex function $f : \mathscr{L}_\infty(\mathbb{S}^{d_w}) \to \overline{\mathbb{R}}$. To facilitate this development, we begin by introducing the following technical assumption on the convex function f.

Assumption 12.0.2. The convex function $f : \mathscr{L}_\infty(\mathbb{S}_+^{d_w}) \to \overline{\mathbb{R}}$ is

i. proper, *i.e.*, its domain is non-empty, $\mathbf{dom}(f) \neq \emptyset$, and $f > -\infty$,

ii. weak*-lower semicontinous (w*-l.s.c.), *i.e.*, its epigraph $\mathbf{epi}(f)$ is convex and closed in $\mathscr{L}_\infty \times \mathbb{R}$ with respect to the *weak* topology* induced by its predual, $\mathscr{L}_1 \times \mathbb{R}$.

Remark 12.0.3. The technical assumption that f is weak*-lower semicontinous ensures that the supporting hyperplanes of the epigraph $\mathbf{epi}(f)$ correspond to affine functions of the form $\mathscr{L}_\infty \ni \mathcal{C} \mapsto \langle \mathcal{C}, \mathcal{M} \rangle + b$ where $b \in \mathbb{R}$ and $\mathcal{M} \in \mathscr{L}_1$, the predual of $\mathscr{L}_\infty$. This is in contrast to a broader class of affine functions induced by the dual space, where $\mathcal{M} \in \mathscr{L}_\infty^*$. The distinction stems from the non-reflexivity of the Banach spaces $\mathscr{L}_1$ and $\mathscr{L}_\infty$, as reflected in the chain of inclusions $(\mathscr{L}_1^* = \mathscr{L}_\infty) \subsetneq \mathscr{L}_1 \subsetneq (\mathscr{L}_1^{**} = \mathscr{L}_\infty^*)$. While this assumption rules out certain edge cases, it encompasses all the interesting examples considered in this paper.

Equipped with the Assumption 12.0.2, the Fenchel conjugate $f^* : \mathscr{L}_1(\mathbb{S}^{d_w}) \to \overline{\mathbb{R}}$ of the convex function $f : \mathscr{L}_\infty(\mathbb{S}^{d_w}) \to \overline{\mathbb{R}}$ is formally defined as

$$f^*(\mathcal{M}) \coloneqq \sup_{\mathcal{C} \in \mathscr{L}_\infty} \langle \mathcal{C}, \mathcal{M} \rangle - f(\mathcal{C}), \tag{12.2}$$

where $\mathcal{M} \in \mathscr{L}_1$ is the dual variable. Moreover, Fenchel-Moreau theorem establishes a duality pairing between f and its conjugate f^* as $f^{**} := (f^*)^* = f$ [56, Prop. 4.1], namely,

$$f(\mathcal{C}) = \sup_{\mathcal{M}\in\mathscr{L}_1} \langle \mathcal{C},\, \mathcal{M}\rangle - f^*(\mathcal{M}). \tag{12.3}$$

Moreover, it turns out that such a function f is monotonic if and only if the domain of the conjugate functions f^* only contains positive operators.

Theorem 12.0.4. *Let $f : \mathscr{L}_\infty \to \overline{\mathbb{R}}$ satisfy Assumption 12.0.2. Then, f is monotonic if and only if $\mathbf{dom}(f^*) \subset \mathscr{L}_1^+$.*

Plugging in the identity (12.3), we rewrite the primal optimization problem (P) as a minimax problem as follows:

$$p_\star = \inf_{\mathcal{K}\in\mathscr{K}} \sup_{\mathcal{M}\in\mathscr{L}_1} \langle \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K},\, \mathcal{M}\rangle - f^*(\mathcal{M}). \tag{P$'$}$$

The minimax problem (P$'$) naturally leads to a dual problem, which can be derived by interchanging the $\inf$ and $\sup$ operations:

$$\begin{aligned} d_\star &:= \sup_{\mathcal{M}\in\mathscr{L}_1} \inf_{\mathcal{K}\in\mathscr{K}} \langle \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K},\, \mathcal{M}\rangle - f^*(\mathcal{M}), \\ &= \sup_{\mathcal{M}\succcurlyeq 0} \inf_{\mathcal{K}\in\mathscr{K}} \langle \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K},\, \mathcal{M}\rangle - f^*(\mathcal{M}), \end{aligned} \tag{12.4}$$

where $d_\star \in \overline{\mathbb{R}}$ is the dual value. The last equality follows from the fact that $\inf\{\langle \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K},\, \mathcal{M}\rangle \mid \mathcal{K} \in \mathscr{K}\} = -\infty$ whenever $\mathcal{M} \in \mathscr{L}_1$ has strictly negative eigenvalues[1].

Problem 12.0.5 (Dual Control Problem)**.** Given a convex and monotonic function $f : \mathscr{L}_\infty \to \overline{R}$ satisfying Assumption 12.0.2 and its Fenchel conjugate $f^* : \mathscr{L}_1 \to \overline{R}$, find a positive-definite operator $\mathcal{M} \in \mathscr{L}_1$ that minimizes the following objective:

$$\boxed{d_\star = \sup_{\substack{\mathcal{M}\in\mathscr{L}_1,\\ \mathcal{M}\succcurlyeq 0}} \inf_{\mathcal{K}\in\mathscr{K}} \langle \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K},\, \mathcal{M}\rangle - f^*(\mathcal{M}),} \tag{D}$$

where $d_\star \in \overline{\mathbb{R}}$ is the dual value.

It is evident from the minimax inequality that weak duality holds, *i.e.*, $d_\star \leq p_\star$. The main result of this section, as established in Theorem 12.0.6, demonstrates that strong minimax duality also holds, *i.e.*, $d_\star = p_\star$.

[1] More precisely, whenever $\mathcal{T}_{yw}\mathcal{M}\mathcal{T}_{yw}^* \not\succcurlyeq 0$.

Theorem 12.0.6 (Strong Minimax Duality)**.** *Let Assumptions 8.2.4 and 12.0.2 hold. If there exists a causal and stabilizing controller $\mathcal{K}_0 \in \mathscr{K}$ and a bounded operator $\mathcal{C}_0 \in \mathscr{L}_\infty$ such that $\mathcal{C}_0 \succ \mathcal{T}_{\mathcal{K}_0}^* \mathcal{T}_{\mathcal{K}_0}$ and $f(\mathcal{C}_0) < +\infty$, then the primal problem* (P) *admits strong duality in*

$$\inf_{\mathcal{K} \in \mathscr{K}} f(\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}) = \max_{\substack{\mathcal{M} \in \mathrm{ba}, \\ \mathcal{M} \succcurlyeq 0}} \inf_{\mathcal{K} \in \mathscr{K}} \langle \mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}, \, \mathcal{M} \rangle - f^*(\mathcal{M}), \tag{12.5}$$

and the maximum is attained at a positive definite point $\mathcal{M}_\star \in \mathrm{ba}$.

Theorem 12.0.7 (Strong duality in $\mathscr{L}_1$)**.** *Define the perturbation functions $h : \mathscr{L}_\infty \to \overline{\mathbb{R}}$ and $g : \mathscr{L}_1 \to \overline{\mathbb{R}}$ as*

$$h(\Delta) = \inf_{\mathcal{K} \in \mathscr{K}} f(\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}} + \Delta), \tag{12.6}$$

$$g(\nabla) = \sup_{\mathcal{M} \in \mathscr{L}_1} V(\mathcal{M}) - f^*(\mathcal{M} + \nabla), \tag{12.7}$$

where h is convex, whereas g is concave. The following statements are equivalent:

i. Strong duality between problems Equation (P) *and Equation* (D) *hold.*

ii. h is weak-lower semicontinous at 0. Namely, for any uniformly norm bounded sequence $\{\Delta_n\}_{n \in \mathbb{N}} \subset \mathscr{L}_\infty \cap \mathbb{B}_r$ that weak* converges to 0, i.e., $\mathrm{w}^*\text{-}\lim_{n\to\infty} \Delta_n = 0$, it holds that $h(0) \leq \liminf_{n\to\infty} h(\Delta_n)$.*

iii. g is upper semicontinous at 0. Namely, for any sequence $\{\nabla_n\}_{n \in \mathbb{N}} \subset \mathscr{L}_1$ that converges to 0 in norm, i.e., $\lim_{n\to\infty} \|\nabla_n\|_1$=0, it holds that $g(0) \geq \limsup_{n\to\infty} g(\nabla_n)$.

Theorem 12.0.8 (Sufficient Conditions for Strong Duality)**.** *The following conditions are sufficient for strong duality:*

i. There exists $\mathcal{K}_0 \in \mathscr{H}_\infty$ and $\mathcal{C}_0 \in \mathbf{dom}(f)$ such that $\mathcal{C}_0 \succcurlyeq \mathcal{T}_{\mathcal{K}_0}^* \mathcal{T}_{\mathcal{K}_0}$ and f is weak-continuous at $\mathcal{C}_0$.*

- *Canonical example of weak*-continuous convex functions are simply affine functionals $\ell(\mathcal{C}) = \langle \mathcal{C}, \, \mathcal{M} \rangle + b$ where $\mathcal{M} \in \mathscr{L}_1^+$ and $b \in \mathbb{R}$. Suppose, there exists a finite number of positive-definite dual variables $\mathcal{M}_1, \ldots, \mathcal{M}_k \in \mathscr{L}_1^+$, a vector $\boldsymbol{b} \in \mathbb{R}^k$, and a proper, convex, lower semicontinuous function $h : \mathbb{R}^k \to \overline{\mathbb{R}}$ such that $f(\mathcal{C}) = h(\boldsymbol{\ell}(\mathcal{C}))$ and*

$\boldsymbol{\ell}(\mathcal{C}_0) \in \mathbf{ri}(h)$ *where* $\boldsymbol{\ell} : \mathscr{L}_\infty \to \mathbb{R}^k$ *is a weak*-continuous affine transformation defined as*

$$\boldsymbol{\ell}(\mathcal{C}) = \begin{bmatrix} \ell_1(\mathcal{C}) & \dots & \ell_k(\mathcal{C}) \end{bmatrix}^\intercal = \begin{bmatrix} \langle \mathcal{C}, \mathcal{M}_1 \rangle & \dots & \langle \mathcal{C}, \mathcal{M}_k \rangle \end{bmatrix}^\intercal + \boldsymbol{b}$$

.

ii. $\mathbf{cone}(\mathscr{L}_1^+ - \mathbf{dom}(f^*))$ *is a closed linear subspace of* $\mathscr{L}_1$ *(in other words,* $0 \in \mathbf{sqri}(\mathscr{L}_1^+ - \mathbf{dom}(f^*))$ *),*

iii. $\mathbf{dom}(f)$ *is weak*-compact.*

iv. The function f *norm coercive, i.e.,* $f(\mathcal{C}) \to \infty$ *as* $\|\mathcal{C}\|_\infty \to \infty$*.*

Remark 12.0.9. The condition on the existence of $\mathcal{K}_0 \in \mathscr{K}$ and $\mathcal{C}_0 \in \mathscr{L}_\infty$ such that $\mathcal{C}_0 \succ \mathcal{T}_{\mathcal{K}_0}^* \mathcal{T}_{\mathcal{K}_0}$ and $f(\mathcal{C}_0) < +\infty$ is essentially a Slater's condition requiring existence of an interior feasible point.

A discerning reader might question the utility of transforming the primal problem (P) into its dual formulation (D), given the introduction of an additional maximization step. The principal advantage lies in the linearization of the objective function with respect to the quadratic term $\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}$ through the inclusion of the dual variable $\mathcal{M} \succcurlyeq 0$ and the Fenchel conjugate $f^*(\mathcal{M})$. This reformulation facilitates a more tractable approach to the minimization problem over causally constrained controllers $\mathcal{K} \in \mathscr{K}$, as the inner minimization reduces to a quadratic objective in $\mathcal{K}$. While the additional maximization over the dual variable $\mathcal{M} \succcurlyeq 0$ may still pose challenges, addressing the positivity constraint on $\mathcal{M}$ turns out to be more manageable generally than enforcing the causality constraint on $\mathcal{K}$ within the same framework.

Before proceeding to the proof of this result in Section 12.3, we first illustrate its applicability across a range of control problems, encompassing both classical examples and more contemporary cases of interest.

12.1 Examples: Norm-optimal Control

Example 12.1.1 (Standard $\mathscr{H}_2$ optimal control). The $\mathscr{H}_2$ optimal controller aims to minimize the $\mathscr{H}_2$ norm of the closed-loop transfer function:

$$\eta_{\mathscr{H}_2}^2 \coloneqq \inf_{\mathcal{K} \in \mathscr{K}} \|\mathcal{T}_{\mathcal{K}}\|_2^2 = \inf_{\mathcal{K} \in \mathscr{K}} \operatorname{tr}(\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}}), \tag{12.8}$$

where $\eta_{\mathscr{H}_2} > 0$ is the minimum $\mathscr{H}_2$ norm. The $\mathscr{H}_2$ optimal controller, whenever it exists and is unique, is denoted by $\mathcal{K}_{\mathscr{H}_2} \coloneqq \arg\min\{\|\mathcal{T}_{\mathcal{K}}\|_2^2 \mid \mathcal{K} \in \mathscr{K}\}$.

It is clear to see that the $\mathscr{H}_2$ optimal control problem can be cast in the form of Problem 8.2.3 by picking

$$f_{\mathscr{H}_2}(\mathcal{C}) \coloneqq \operatorname{tr}(\mathcal{C}) \text{ with } \mathbf{dom}(f_{\mathscr{H}_2}) = \mathscr{L}_1. \tag{12.9}$$

The $\mathscr{H}_2$ norm of the closed-loop transfer operator $\mathcal{T}_{\mathcal{K}}$ has several operational interpretations. The first interpretation establishes a connection to stochastic control. Specifically, the $\mathscr{H}_2$ norm $\|\mathcal{T}_{\mathcal{K}}\|_2^2$ is equal to the steady-state expected squared-norm of the regulated output state-space model (8.42) driven by the feedback controller $\mathcal{K}$ when the disturbance process $\{w_t\}$ is a white noise with $\mathbb{E}[w_t w_s^\intercal] = \delta_{t-s} I_{d_w}$, *i.e.*,

$$\|\mathcal{T}_{\mathcal{K}}\|_2^2 = \limsup_{T\to\infty} \frac{1}{T}\, \mathbb{E}\left[\sum\nolimits_{t=0}^{T-1} \|z_t\|^2\right]. \tag{12.10}$$

Furthermore, when the disturbances are white Gaussian noise, the $\mathscr{H}_2$ optimal controller coincides with the optimal LQG controller .

The second operational interpretation established a connection as a system gain. Namely, the $\|\mathcal{T}_{\mathcal{K}}\|_2^2$ is the worst-case system gain of the closed-loop transfer operator $\mathcal{T}_{\mathcal{K}} : \mathbf{w} \mapsto \mathbf{z}$ from bounded energy disturbances $\mathbf{w} \in \ell_2$ to bounded regulated output $\mathcal{T}_{\mathcal{K}}\mathbf{w} \in \ell_\infty$, *i.e.*,

$$\|\mathcal{T}_{\mathcal{K}}\|_2 = \sup_{\substack{\mathbf{w}\in\ell_2,\\ \mathbf{w}\neq 0}} \frac{\|\mathcal{T}_{\mathcal{K}}\mathbf{w}\|_{\ell_\infty}}{\|\mathbf{w}\|_{\ell_2}} = \sup_{\|\mathbf{w}\|_{\ell_2}\leq 1} \|\mathcal{T}_{\mathcal{K}}\mathbf{w}\|_{\ell_\infty}. \tag{12.11}$$

This is essentially the operator norm of $\mathcal{T}$ as a mapping from ℓ_2 to ℓ_∞ .

Example 12.1.2 (Weighted $\mathscr{H}_2$ optimal control). Let $\mathcal{M} \succcurlyeq 0$ be a positive-definite $\mathscr{L}_1$ operator. The $\mathcal{M}$-weighted $\mathscr{H}_2$ optimal control problem is formulated as

$$\inf_{\mathcal{K}\in\mathscr{K}} \operatorname{tr}(\mathcal{T}_{\mathcal{K}}^* \mathcal{T}_{\mathcal{K}} \mathcal{M}). \tag{12.12}$$

The corresponding convex function here is simply

$$f_{\mathcal{M},\mathscr{H}_2}(\mathcal{C}) \coloneqq \operatorname{tr}(\mathcal{C}\mathcal{M}) = \langle \mathcal{C},\, \mathcal{M}\rangle \text{ with } \mathbf{dom}(f_{\mathcal{M},\mathscr{H}_2}) = \mathscr{L}_\infty. \tag{12.13}$$

Clearly, the usual $\mathscr{H}_2$ optimal control corresponds to the special case $\mathcal{M} = \mathcal{I}$. When the weight operator $\mathcal{M}$ is strictly positive definite and satisfies the Paley-Wiener-Szegő condition, *i.e.*, $\log(\mathcal{M}) \in \mathscr{L}_1$, , then $\mathcal{M}$ can be realized as the auto-covariance operator of a weakly-stationary stochastic process $\{w_t\}$ such that $\mathbb{E}[w_t w_s^\intercal] = \widehat{M}_{t-s}$. In that case, the corresponding transfer function $M(z)$ is the power spectral density of the process $\{w_t\}$. Under these conditions, one can interpret $\mathrm{tr}(\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}\mathcal{M})$ as the steady-state expected squared-norm of the regulated output z_t as in (**??**) for a weakly-stationary and colored disturbance process $\{w_t\}$ with auto-covariance operator $\mathcal{M}$.

Example 12.1.3 (**$\mathscr{H}_\infty$ optimal control**)**.** The standard $\mathscr{H}_\infty$ control problem is formulated as

$$\gamma_{\mathscr{H}_\infty}^2 := \inf_{\mathcal{K}\in\mathscr{K}} \|\mathcal{T}_{\mathcal{K}}\|_\infty^2 = \inf_{\mathcal{K}\in\mathscr{K}} \|\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}\|_\infty, \tag{12.14}$$

where $\gamma_{\mathscr{H}_\infty} > 0$ is the minimum $\mathscr{H}_\infty$ norm. The $\mathscr{H}_\infty$ optimal controller, whenever it exists and is unique, is denoted by $\mathcal{K}_{\mathscr{H}_\infty} := \arg\min\{\|\mathcal{T}_{\mathcal{K}}\|_\infty^2 \mid \mathcal{K} \in \mathscr{K}\}$.

The corresponding convex function is

$$f_{\mathscr{H}_\infty}(\mathcal{C}) := \|\mathcal{C}\|_\infty \text{ with } \mathbf{dom}(f_{\mathscr{H}_\infty}) = \mathscr{L}_\infty. \tag{12.15}$$

The $\mathscr{H}_\infty$ norm $\|\mathcal{T}_{\mathcal{K}}\|_\infty$ can be interpreted as the worst-case system gain of the closed-loop transfer operator $\mathcal{T}_{\mathcal{K}} : \mathbf{w} \mapsto \mathbf{z}$ as a mapping from ℓ_2 to ℓ_2, *i.e.*,

$$\|\mathcal{T}_{\mathcal{K}}\|_\infty = \sup_{\substack{\mathbf{w}\in\ell_2,\\ \mathbf{w}\neq 0}} \frac{\|\mathcal{T}_{\mathcal{K}}\mathbf{w}\|_{\ell_2}}{\|\mathbf{w}\|_{\ell_2}} = \sup_{\|\mathbf{w}\|_{\ell_2}\leq 1} \|\mathcal{T}_{\mathcal{K}}\mathbf{w}\|_{\ell_2}. \tag{12.16}$$

Example 12.1.4 (H_{2p} **optimal control**)**.** For $p \in [1, \infty]$

$$\inf_{\mathcal{K}\in\mathscr{K}} \|\mathcal{T}_{\mathcal{K}}\|_{2p}^2 = \inf_{\mathcal{K}\in\mathscr{K}} \|\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}\|_p \tag{12.17}$$

$$f_{H_{2p}}(\mathcal{C}) := \|\mathcal{C}\|_p \text{ with } \mathbf{dom}(f_{H_{2p}}) = \mathscr{L}_p. \tag{12.18}$$

Example 12.1.5 (**Mixed $\mathscr{H}_2/\mathscr{H}_\infty$ control**)**.**

$$\begin{array}{l} \inf_{\mathcal{K}\in\mathscr{K}} \|\mathcal{T}_{\mathcal{K}}\|_2^2, \\ \text{s.t. } \|\mathcal{T}_{\mathcal{K}}\|_\infty \leq \gamma, \end{array} = \inf_{\mathcal{K}\in\mathscr{K}} \mathrm{tr}(\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}) + \begin{cases} 0, & \|\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}\|_\infty \leq \gamma^2, \\ +\infty, & \text{o.w.} \end{cases} \tag{12.19}$$

$$f_{\gamma,\mathscr{H}_2/\mathscr{H}_\infty}(\mathcal{C}) := \mathrm{tr}(\mathcal{C}) + \begin{cases} 0, & \|\mathcal{C}\|_\infty \leq \gamma^2, \\ +\infty, & \text{o.w.}, \end{cases} \tag{12.20}$$

with $\mathbf{dom}(f_{\gamma,\mathscr{H}_2/\mathscr{H}_\infty}) = \{\mathcal{C} \in \mathscr{L}_\infty \mid \|\mathcal{C}\|_\infty \leq \gamma^2\}$

12.2 Examples: Distributionally Robust Control

Example 12.2.1 (Risk-sensitive control). The risk-sensitive control objective aims to minimize an exponential cost, formulated below

$$\inf_{\mathcal{K}\in\mathscr{K}} \gamma \log\left(\mathbb{E}_{\mathbf{w}\sim\mathbb{P}_\circ}\left[\mathrm{e}^{\gamma^{-1}\mathbf{w}^*\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}\mathbf{w}}\right]\right), \tag{12.21}$$

where $\gamma > 0$ is the risk parameter and $\mathbb{P}_\circ$ is a nominal probability distribution of the disturbances. The expectation above should be understood formally as the time-averaged limit of finite-horizon risk-sensitive costs. The convex function corresponding to this problem is given by

$$f_{\gamma,\mathrm{RS}}(\mathcal{C}) \coloneqq \gamma \log\left(\mathbb{E}_{\mathbf{w}\sim\mathbb{P}_\circ}\left[\mathrm{e}^{\gamma^{-1}\mathbf{w}^*\mathcal{C}\mathbf{w}}\right]\right) \text{ with } \mathbf{dom}(f_{\gamma,\mathrm{RS}}) = \mathscr{L}_\infty. \tag{12.22}$$

With the decreasing value of γ, The risk-sensitive objective resolves the gap between smaller and larger cost values. It penalizes higher cost levels relatively more than the smaller values as γ decreases. This essentially incentivizes the controller to be more risk-averse to reduce the chances of yielding higher costs.

In the special case of the nominal distribution $\mathbb{P}_\circ$ of disturbances forming a stationary Gaussian process with auto-covariance operator $\mathcal{M}_\circ \succ 0$, the risk-sensitive objective simplifies further to

$$\inf_{\mathcal{K}\in\mathscr{K}} -\frac{\gamma}{2}\operatorname{logdet}(\mathcal{I} - 2\gamma^{-1}\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}\mathcal{M}_\circ), \tag{12.23}$$

where $\operatorname{logdet}(\cdot)$ should be understood as $\operatorname{tr}(\log(\cdot))$.

The corresponding convex function then becomes

$$f_{\gamma,\mathrm{RS}}(\mathcal{C}) \coloneqq -\frac{\gamma}{2}\operatorname{logdet}(\mathcal{I} - 2\gamma^{-1}\mathcal{C}\mathcal{M}_\circ), \tag{12.24}$$

with $\mathbf{dom}(f_{\gamma,\mathrm{RS}}) = \left\{\mathcal{C} \in \mathscr{L}_\infty \mid \mathcal{C} \prec \frac{\gamma}{2}\mathcal{M}_\circ^{-1}\right\}$.

Example 12.2.2 (KL **distributionally robust control).**

$$\inf_{\substack{\mathcal{K}\in\mathscr{K},\\ \gamma>2\|\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}\mathcal{M}_\circ\|_\infty}} \gamma r - \frac{\gamma}{2}\operatorname{logdet}(\mathcal{I} - 2\gamma^{-1}\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}\mathcal{M}_\circ) \tag{12.25}$$

$$f_{\mathrm{KL}}(\mathcal{C}) \coloneqq \inf_{\gamma>2\|\mathcal{C}\mathcal{M}_\circ\|_\infty} \gamma r - \frac{\gamma}{2}\operatorname{logdet}(\mathcal{I} - 2\gamma^{-1}\mathcal{C}\mathcal{M}_\circ), \tag{12.26}$$

with $\mathbf{dom}(f_{\mathrm{KL}}) = \mathscr{L}_\infty$.

$$\inf_{\mathcal{K}\in\mathscr{K}} -\gamma \operatorname{logdet}(\mathcal{I} - \gamma^{-1}\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}) \tag{12.27}$$

Example 12.2.3 (Wasserstein distributionally robust control). When the ambiguity set of plausible probability distributions of disturbances is constructed as a Wasserstein-2 ball, the distributionally robust controller can be obtained by solving the following primal optimization problem:

$$\inf_{\substack{\mathcal{K}\in\mathscr{K},\\ \gamma>\|\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}\|_\infty}} \gamma r^2 + \gamma \operatorname{tr}\left[\left((\mathcal{I}-\gamma^{-1}\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}})^{-1}-\mathcal{I}\right)\mathcal{M}_\circ\right], \tag{12.28}$$

where $\mathcal{M}_\circ \in \mathscr{L}_1^+$ is the auto-covariance operator of the nominal disturbance process, which is assumed to be weakly stationary, and γ is a Lagrange multiplier determined by the desired radius of the Wasserstein-2 ball, $r > 0$. The corresponding convex function for this optimization problem then becomes:

$$f_{\mathsf{W}_2}(\mathcal{C}) \coloneqq \inf_{\gamma>\|\mathcal{C}\|_\infty} \gamma r^2 + \gamma \operatorname{tr}\left[\left((\mathcal{I}-\gamma^{-1}\mathcal{C})^{-1}-\mathcal{I}\right)\mathcal{M}_\circ\right], \tag{12.29}$$

with $\mathbf{dom}(f_{\mathsf{W}_2}) = \mathscr{L}_\infty^+$. The suboptimal problem

$$\inf_{\mathcal{K}\in\mathscr{K}} \gamma \operatorname{tr}\left[(\mathcal{I}-\gamma^{-1}\mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}})^{-1}\right] \tag{12.30}$$

The corresponding convex function

$$f_{\gamma,\mathsf{W}_2} \coloneqq \gamma \operatorname{tr}\left[(\mathcal{I}-\gamma^{-1}\mathcal{C})^{-1}\right] \text{ with } \mathbf{dom}(f_{\gamma,\mathsf{W}_2}) = \{\mathcal{C}\in\mathscr{L}_\infty \mid \mathcal{C} \prec \gamma\mathcal{I}\}. \tag{12.31}$$

Remark 12.2.4. Distributionally robust interpretation of problems of the type

$$\sup_{\mathcal{M}\in\mathscr{M}} V(\mathcal{M}) \tag{12.32}$$

where $\mathscr{M} \subseteq \mathscr{L}_1(\mathbb{S}_+)$ is a convex subset. This is analogous to coherent risk measures. In fact, one can view this problem as a non-commutative generalization of coherent risk measures.

Furthermore, the general setting

$$\sup_{\mathcal{M}\succcurlyeq 0} V(\mathcal{M}) - f_*(\mathcal{M}) \tag{12.33}$$

is related to convex risk measures.

Remark 12.2.5 (Norm interpretation).

$$\sup_{\mathcal{M}\in\mathscr{M}} \sqrt{\langle \mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}},\, \mathcal{M}\rangle} \tag{12.34}$$

is a norm for $\mathcal{T}_{\mathcal{K}}$ whenever $\mathscr{M} \cap \operatorname{int}(\mathscr{L}_1(\mathbb{S}_+)) \neq \varnothing$, that is, there exists a strictly positive definite element $\mathcal{M} \succ 0$ of $\mathscr{M}$.

12.3 Proof of Strong Duality in Theorem 12.0.6

We need some technical results in order to prove the strong duality.

Lemma 12.3.1. *Let $f : \mathscr{L}_\infty \to \overline{\mathbb{R}}$ satisfy Assumption 8.2.4. Then, we have that*

$$\inf_{\mathcal{K}\in\mathscr{K}} f(\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}) = \inf_{\mathcal{C}\in\mathscr{L}_\infty} f(\mathcal{C}) + \inf_{\mathcal{K}\in\mathscr{K}} \sup_{\substack{\mathcal{M}\in\mathscr{L}_1,\\ \mathcal{M}\succcurlyeq 0}} \langle \mathcal{M},\, \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} - \mathcal{C}\rangle. \tag{12.35}$$

Proof. First, we invoke the monotonicity of f to express $f(\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K})$ as

$$\begin{aligned} f(\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}) &= \inf_{\mathcal{C}\in\mathscr{L}_\infty} f(\mathcal{C}) \ \text{ s.t. } \mathcal{C} \succcurlyeq \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}, & (12.36)\\ &= \inf_{\mathcal{C}\in\mathscr{L}_\infty} f(\mathcal{C}) + \sup_{\substack{\mathcal{M}\in\mathscr{L}_1,\\ \mathcal{M}\succcurlyeq 0}} \langle \mathcal{M},\, \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} - \mathcal{C}\rangle, & (12.37)\end{aligned}$$

where $\mathcal{M} \succcurlyeq 0$ is the Lagrange multiplier for the constraint $\mathcal{C} \succcurlyeq \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}$. ■

Lemma 12.3.2 (LDU Decomposition of Block Operators)**.** *Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be Hilbert spaces and $\mathcal{A}_{ij} : \mathscr{H}_i \to \mathscr{H}_j$ for $i, j \in \{1, 2\}$ be bounded linear operators. Consider the block operator*

$$\mathcal{A} := \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} : \mathscr{H}_1 \oplus \mathscr{H}_2 \to \mathscr{H}_1 \oplus \mathscr{H}_2 \tag{12.38}$$

i. $\mathcal{A}$ is a bounded operator.

ii. If $\mathcal{A}_{11}$ has a bounded inverse, then $\mathcal{A}$ decomposes into three bounded block operators $\mathcal{L}$, $\mathcal{D}$, and $\mathcal{U}$ as

$$\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathcal{I}_1 & 0 \\ \mathcal{A}_{21}\mathcal{A}_{11}^{-1} & \mathcal{I}_2 \end{bmatrix}}_{\mathcal{L}} \underbrace{\begin{bmatrix} \mathcal{A}_{11} & 0 \\ 0 & \mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12} \end{bmatrix}}_{\mathcal{D}} \underbrace{\begin{bmatrix} \mathcal{I}_1 & \mathcal{A}_{11}^{-1}\mathcal{A}_{12} \\ 0 & \mathcal{I}_2 \end{bmatrix}}_{\mathcal{U}}. \tag{12.39}$$

iii. Suppose $\mathcal{A}_{21} = \mathcal{A}_{12}^*$, $\mathcal{A}_{11}$ and $\mathcal{A}_{22}$ are self-adjoint, and $\mathcal{A}_{11} \succ 0$ with bounded inverse. Then, $\mathcal{A} \succcurlyeq 0$ (resp. $\mathcal{A} \succ 0$) if and only if the Schur complement is positive-definite, i.e., $\mathcal{A}_{22} - \mathcal{A}_{12}^*\mathcal{A}_{11}^{-1}\mathcal{A}_{12} \succcurlyeq 0$ (resp. $\mathcal{A}_{22} - \mathcal{A}_{12}^*\mathcal{A}_{11}^{-1}\mathcal{A}_{12} \succ 0$.)

Lemma 12.3.3. *Consider the mapping $F : \mathscr{H}_\infty \to \mathscr{L}_\infty^+$ defined as $F(\mathcal{K}) = \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}$. We can define its epigraph using the partial order over the operator algebra as follows*

$$\mathbf{epi}(F) := \{(\mathcal{K}, \mathcal{C}) \in \mathscr{H}_\infty \times \mathscr{L}_\infty \mid F(\mathcal{K}) = \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} \preccurlyeq \mathcal{C}\} \tag{12.40}$$

We have the following properties:

i. $(\mathcal{K},\mathcal{C}) \in \mathbf{epi}(F)$ *if and only if* $\mathcal{K} \in \mathscr{H}_\infty$, $\mathcal{C} \in \mathscr{L}_\infty$ *and* $\begin{bmatrix} \mathcal{I} & \mathcal{T}_\mathcal{K} \\ \mathcal{T}_\mathcal{K}^* & \mathcal{C} \end{bmatrix} \succcurlyeq 0.$

ii. $\mathbf{epi}(F)$ *is convex in* $\mathscr{H}_\infty \times \mathscr{L}_\infty$.

iii. $\mathbf{epi}(F)$ *is closed in* $\mathscr{H}_\infty \times \mathscr{L}_\infty$ *when endowed with the product weak*-topology* $\sigma(\mathscr{L}_\infty, \mathscr{L}_1) \times \sigma(\mathscr{H}_\infty, \mathscr{L}_1/\mathscr{H}_1^0)$.

Proof. For closedness, first use the fact that $\mathbf{epi}(F)$ is convex set. Then, by the Krein-Smulian theorem, the closedness of $\mathbf{epi}(F)$ is equivalent to the closedness of $\mathbf{epi}(F)$ intersected with bounded sets $\mathbb{B}_r$ with $r > 0$. Then, by Banach-Alaoglu, $\mathbf{epi}(F) \cap \mathbb{B}_r$ is weak*-compact. Since its predual is separable, the weak*-compactness of bounded sets is equivalent to weak*-sequential compactness. Then, show that for any weak*-convergent sequence in $\mathbf{epi}(F) \cap \mathbb{B}_r$, the converged point remains in $\mathbf{epi}(F) \cap \mathbb{B}_r$. ■

Definition 12.3.4. Let $\chi : \mathscr{L}_\infty \to \overline{\mathbb{R}}$ and $V : \mathscr{L}_1 \to \overline{\mathbb{R}}$ be two functions defined as

$$\chi(\mathcal{C}) := \inf_{\mathcal{K}\in\mathscr{K}} \sup_{\substack{\mathcal{M}\in\mathscr{L}_1,\\ \mathcal{M}\succcurlyeq 0}} \langle \mathcal{M},\, \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} - \mathcal{C}\rangle, \ \text{ and } V(\mathcal{M}) := \inf_{\mathcal{K}\in\mathscr{K}} \langle \mathcal{M},\, \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}\rangle. \tag{12.41}$$

We call V as the *value function*.

The following lemma asserts that χ is an indicator function of a convex subset.

Lemma 12.3.5. *The following statements hold:*

i. $\chi : \mathscr{L}_\infty \to \overline{\mathbb{R}}$ *is the indicator function of a set* $\mathscr{C}_{\mathscr{H}_\infty} \subset \mathscr{L}_\infty$ *defined as*

$$\mathscr{C}_{\mathscr{H}_\infty} := \{\mathcal{C} \in \mathscr{L}_\infty \mid \exists \mathcal{K} \in \mathscr{H}_\infty,\ \mathcal{C} \succcurlyeq \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}\}, \tag{12.42}$$

namely, $\mathbf{dom}(\chi) = \mathscr{C}_{\mathscr{H}_\infty}$ *and it is exactly zero on its domain.*

ii. The set $\mathscr{C}_{\mathscr{H}_\infty} \subset \mathscr{L}_\infty^+$ *is non-empty, convex, w*-closed. In other words,* χ *is proper, convex, w*-lower-semicontinuous, and monotonically decreasing.*

iii. The Fenchel conjugate $\chi^* : \mathscr{L}_1 \to \overline{\mathbb{R}}$ *satisfies* $\chi^*(\mathcal{M}) = \begin{cases} -V(-\mathcal{M}), & \text{if } \mathcal{M} \preccurlyeq 0, \\ +\infty, & \text{o.w.} \end{cases}$

Proof. i. By the definition of the positive definite order, for any $\mathcal{C} \in \mathscr{L}_\infty$ and $\mathcal{K} \in \mathscr{H}_\infty$, we have that

$$\sup_{\substack{\mathcal{M}\in\mathscr{L}_1,\\ \mathcal{M}\succcurlyeq 0}} \langle \mathcal{M},\, \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} - \mathcal{C}\rangle = \begin{cases} 0, & \text{if } \mathcal{C} \succcurlyeq \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}, \\ +\infty, & \text{o.w.} \end{cases} \tag{12.43}$$

Therefore, $\chi(\mathcal{C}) = 0$ if there exists a $\mathcal{K} \in \mathscr{K}$ such that $\mathcal{C} \succcurlyeq \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}$ and $\chi(\mathcal{C}) = +\infty$ if no such $\mathcal{K} \in \mathscr{K}$ exists. This makes χ an indicator function of its domain $\mathbf{dom}(\chi) = \{\mathcal{C} \in \mathscr{L}_\infty \mid \exists \mathcal{K} \in \mathscr{H}_\infty,\, \mathcal{C} \succcurlyeq \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}\}$.

ii. (Convexity of $\mathscr{C}_{\mathscr{H}_\infty}$) Non-emptiness follows trivially since $\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} \in \mathscr{C}_{\mathscr{H}_\infty}$ for any $\mathcal{K} \in \mathscr{H}_\infty$. Let $\lambda \in [0,1]$ and $\mathcal{C}_0, \mathcal{C}_1 \in \mathscr{C}_{\mathscr{H}_\infty}$. This means there exists two causal and stable controllers $\mathcal{K}_0, \mathcal{K}_1 \in \mathscr{H}_\infty$ such that $\mathcal{T}_{\mathcal{K}_0}^*\mathcal{T}_{\mathcal{K}_0} \preccurlyeq \mathcal{C}_0$ and $\mathcal{T}_{\mathcal{K}_1}^*\mathcal{T}_{\mathcal{K}_1} \preccurlyeq \mathcal{C}_1$. Now, define $\mathcal{K}_\lambda \coloneqq \lambda\mathcal{K}_0 + (1-\lambda)\mathcal{K}_1$, then

$$\lambda\mathcal{C}_0 + (1-\lambda)\mathcal{C}_1 \succcurlyeq \lambda\mathcal{T}_{\mathcal{K}_0}^*\mathcal{T}_{\mathcal{K}_0} + (1-\lambda)\mathcal{T}_{\mathcal{K}_1}^*\mathcal{T}_{\mathcal{K}_1} \succcurlyeq \mathcal{T}_{\mathcal{K}_\lambda}^*\mathcal{T}_{\mathcal{K}_\lambda}, \tag{12.44}$$

where the last inequality follows from the operator convexity of the mapping $\mathcal{K} \mapsto \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}$ as shown in (**??**) for the proof of **??**. Thus, $\lambda\mathcal{C}_0 + (1-\lambda)\mathcal{C}_1 \in \mathscr{C}_{\mathscr{H}_\infty}$.

(w*-closedness of $\mathscr{C}_{\mathscr{H}_\infty}$) *Step 1: Restricting to norm bounded sets:* Since $\mathscr{C}_{\mathscr{H}_\infty} \subset \mathscr{L}_\infty^+$ is convex, it is $\sigma(\mathscr{L}_\infty, \mathscr{L}_1)$-closed if and only if the intersection $\mathscr{C}_{\mathscr{H}_\infty} \cap \{\mathcal{C} \in \mathscr{L}_\infty \mid \|\mathcal{C}\|_\infty \leq \gamma\}$ is $\sigma(\mathscr{L}_\infty, \mathscr{L}_1)$-closed for each $\gamma > 0$ by Krein–Šmulian Theorem (**??**).

Step 2: Sequential Closure of Bounded Sets: Fix any $\gamma > 0$. The $\|.\|_\infty$-ball of size $\gamma > 0$ is $\sigma(\mathscr{L}_\infty, \mathscr{L}_1)$-compact by Banach-Alaoglu Theorem (**??**). Moreover, since the predual $\mathscr{L}_1$ is a separable space, the norm bounded ball is sequentially compact, meaning every sequence in it has a convergent subsequence. Therefore, the intersection $\mathscr{C}_{\mathscr{H}_\infty} \cap \{\mathcal{C} \in \mathscr{L}_\infty \mid \|\mathcal{C}\|_\infty \leq \gamma\}$ is $\sigma(\mathscr{L}_\infty, \mathscr{L}_1)$-closed if and only if every $\sigma(\mathscr{L}_\infty, \mathscr{L}_1)$-convergent sequence in it converges to a point in it. Concretely, let $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathscr{C}_{\mathscr{H}_\infty} \cap \{\mathcal{C} \in \mathscr{L}_\infty \mid \|\mathcal{C}\|_\infty \leq \gamma\}$, converging to a point $\mathcal{C}_\star \in \mathscr{L}_\infty$ with $\|\mathcal{C}_\star\|_\infty \leq \gamma$ in the $\sigma(\mathscr{L}_\infty, \mathscr{L}_1)$ sense. This means there exists a sequence of causal and bounded controllers $\{\mathcal{K}_n\}_{n\in\mathbb{N}} \subset \mathscr{H}_\infty$ such that $\mathcal{T}_{\mathcal{K}_n}^*\mathcal{T}_{\mathcal{K}_n} \preccurlyeq \mathcal{C}_n \preccurlyeq \gamma\mathcal{I}$ and $\lim_{n\to\infty}\langle \mathcal{M},\, \mathcal{C}_n\rangle = \langle \mathcal{M},\, \mathcal{C}_\star\rangle$ for all $\mathcal{M} \in \mathscr{L}_1$.

Step 3: Uniform Boundedness of the Controllers: Notice that one can decompose the closed-loop transfer operator as $\mathcal{T}_\mathcal{K} = \mathcal{T}_1\mathcal{K}\mathcal{T}_2 + \mathcal{T}_3 = \mathcal{U}\Delta\mathcal{K}\nabla\mathcal{V}^* + \mathcal{T}_3$

where $\mathcal{T}_1 = \mathcal{U}\Delta$ and $\mathcal{T}_2 = \nabla\mathcal{V}^*$ are inner-outer factorizations such that $\Delta, \Delta^{-1}, \nabla, \nabla^{-1} \in \mathscr{RH}_\infty$ outer operators and $\mathcal{U}, \mathcal{V}^* \in \mathscr{RH}_\infty$ are inner operators such that $\mathcal{U}^*\mathcal{U} = \mathcal{I}$ and $\mathcal{V}^*\mathcal{V} = \mathcal{I}$. We can bound the $\mathscr{H}_\infty$-norm of $\mathcal{K}$ above by the $\mathscr{H}_\infty$-norm of $\mathcal{T}_\mathcal{K}$ as follows:

$$\begin{aligned}
\|\mathcal{K}\|_\infty &= \|\Delta^{-1}\mathcal{U}^*\mathcal{U}\Delta\mathcal{K}\nabla\mathcal{V}^*\mathcal{V}\nabla^{-1}\|_\infty, && (12.45)\\
&\leq \|\Delta^{-1}\mathcal{U}^*\|_\infty\|\mathcal{V}\nabla^{-1}\|_\infty\|\mathcal{U}\Delta\mathcal{K}\nabla\mathcal{V}^*\|_\infty, && (12.46)\\
&= \|\Delta^{-1}\|_\infty\|\nabla^{-1}\|_\infty\|\mathcal{T}_\mathcal{K} - \mathcal{T}_3\|_\infty, && (12.47)\\
&\leq \|\Delta^{-1}\|_\infty\|\nabla^{-1}\|_\infty(\|\mathcal{T}_\mathcal{K}\|_\infty + \|\mathcal{T}_3\|_\infty) && (12.48)
\end{aligned}$$

Since $\sup_{n\in\mathbb{N}} \|\mathcal{T}_{\mathcal{K}_n}\|_\infty \leq \sqrt{\gamma}$, the sequence $\{\mathcal{K}_n\}_{n\in\mathbb{N}} \subset \mathscr{H}_\infty$ is uniformly norm bounded.

Step 4: Existence of a Viable Controller: By the Banach-Alaoglu Theorem (**??**), there exists a subsequence $\{\mathcal{K}_{n_k}\}_{k\in\mathbb{N}} \subset \mathscr{H}_\infty$ which converges to a point $\mathcal{K}_\star \in \mathscr{H}_\infty$ under the $\sigma(\mathscr{H}_\infty, \mathscr{L}_1/\mathscr{H}_1^0)$ topology, *i.e.*, $\lim_{k\to\infty}\langle\mathcal{K}_{n_k}, \Lambda\rangle = \langle\mathcal{K}_\star, \Lambda\rangle$ for all $\Lambda \in \mathscr{L}_1$.

The proof is completed if $\mathcal{T}_{\mathcal{K}_\star}^*\mathcal{T}_{\mathcal{K}_\star} \preccurlyeq \mathcal{C}_\star$. By the Schur complement lemma (Lemma 12.3.2), the condition $\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} \preccurlyeq \mathcal{C}$ is equivalent to the positive-definiteness of the following block operator:

$$\begin{bmatrix} \mathcal{I} & \mathcal{T}_\mathcal{K} \\ \mathcal{T}_\mathcal{K}^* & \mathcal{C} \end{bmatrix} \succcurlyeq 0. \tag{12.49}$$

Therefore, we have that

$$\sup_{\substack{\Lambda\in\mathscr{L}_1\\ \Lambda\succcurlyeq 0}} -\left\langle \begin{bmatrix} \mathcal{I} & \mathcal{T}_\mathcal{K} \\ \mathcal{T}_\mathcal{K}^* & \mathcal{C} \end{bmatrix}, \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}^* & \Lambda_{22} \end{bmatrix} \right\rangle = \begin{cases} 0, & \text{if } \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K} \preccurlyeq \mathcal{C}, \\ +\infty, & \text{o.w.} \end{cases} \tag{12.50}$$

Using this fact, we obtain that

$$\sup_{\substack{\Lambda\in\mathscr{L}_1\\ \Lambda\succcurlyeq 0}} -\left\langle \begin{bmatrix} \mathcal{I} & \mathcal{T}_{\mathcal{K}_\star} \\ \mathcal{T}_{\mathcal{K}_\star}^* & \mathcal{C}_\star \end{bmatrix}, \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}^* & \Lambda_{22} \end{bmatrix} \right\rangle \tag{12.51}$$

$$= \sup_{\substack{\Lambda\in\mathscr{L}_1\\ \Lambda\succcurlyeq 0}} -\operatorname{tr}(\Lambda_{11}) - 2\Re\left\{\langle \mathcal{T}_{\mathcal{K}_\star},\, \Lambda_{12}\rangle\right\} - \langle \mathcal{C}_\star,\, \Lambda_{22}\rangle, \tag{12.52}$$

$$\overset{(a)}{=} \sup_{\substack{\Lambda\in\mathscr{L}_1\\ \Lambda\succcurlyeq 0}} \lim_{k\to\infty} -\operatorname{tr}(\Lambda_{11}) - 2\Re\left\{\langle \mathcal{T}_{\mathcal{K}_{n_k}},\, \Lambda_{12}\rangle\right\} - \langle \mathcal{C}_{n_k},\, \Lambda_{22}\rangle, \tag{12.53}$$

$$\overset{(b)}{\leq} \liminf_{k\to\infty} \sup_{\substack{\Lambda\in\mathscr{L}_1\\ \Lambda\succcurlyeq 0}} -\operatorname{tr}(\Lambda_{11}) - 2\Re\left\{\langle \mathcal{T}_{\mathcal{K}_{n_k}},\, \Lambda_{12}\rangle\right\} - \langle \mathcal{C}_{n_k},\, \Lambda_{22}\rangle, \tag{12.54}$$

$$\overset{(c)}{=} \liminf_{k\to\infty} \sup_{\substack{\Lambda\in\mathscr{L}_1\\ \Lambda\succcurlyeq 0}} -\left\langle \begin{bmatrix} \mathcal{I} & \mathcal{T}_{\mathcal{K}_{n_k}} \\ \mathcal{T}_{\mathcal{K}_{n_k}}^* & \mathcal{C}_{n_k} \end{bmatrix}, \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}^* & \Lambda_{22} \end{bmatrix} \right\rangle, \tag{12.55}$$

$$\overset{(d)}{=} 0, \tag{12.56}$$

where equality (a) follows from the weak*-convergence of sequences $\{\mathcal{K}_{n_k}\}_{k\in\mathbb{N}}$ and $\{\mathcal{C}_{n_k}\}_{k\in\mathbb{N}}$, inequality (b) follows from the $\lim/\sup$ inequality, equality (c) follows from the fact that $\mathcal{T}_{\mathcal{K}_{n_k}}^* \mathcal{T}_{\mathcal{K}_{n_k}} \preccurlyeq \mathcal{C}_{n_k}$, and the equality (d) follows from (12.50).

Thus, we conclude that $\mathscr{C}_{\mathscr{H}_\infty}$ is weak*-closed. Hence, the indicator function χ is proper, convex, and weak* lower-semicontinuous.

iii. By the definition of Fenchel conjugate, we have

$$\chi^*(\mathcal{M}) = \sup_{\mathcal{C}\in\mathscr{L}_\infty} \langle \mathcal{M},\, \mathcal{C}\rangle - \chi(\mathcal{C}), \tag{12.57}$$

$$= \sup_{\mathcal{C}\in\mathscr{L}_\infty} \langle \mathcal{M},\, \mathcal{C}\rangle \quad \text{s.t.} \quad \exists \mathcal{K}\in\mathscr{K}, \mathcal{C}\succcurlyeq \mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}, \tag{12.58}$$

$$= \sup_{\substack{\mathcal{C}\in\mathscr{L}_\infty,\\ \mathcal{K}\in\mathscr{K}}} \langle \mathcal{M},\, \mathcal{C}\rangle \quad \text{s.t.} \quad \mathcal{C}\succcurlyeq \mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}. \tag{12.59}$$

When $\mathcal{M}\preccurlyeq 0$, we have

$$\chi^*(\mathcal{M}) = \sup_{\substack{\mathcal{C}\in\mathscr{L}_\infty,\\ \mathcal{K}\in\mathscr{K}}} -\langle -\mathcal{M},\, \mathcal{C}\rangle \quad \text{s.t.} \quad \mathcal{C}\succcurlyeq \mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}, \tag{12.60}$$

$$= -\inf_{\substack{\mathcal{C}\in\mathscr{L}_\infty,\\ \mathcal{K}\in\mathscr{K}}} \langle -\mathcal{M},\, \mathcal{C}\rangle \quad \text{s.t.} \quad \mathcal{C}\succcurlyeq \mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}, \tag{12.61}$$

$$= -\inf_{\mathcal{K}\in\mathscr{K}} \langle -\mathcal{M},\, \mathcal{T}_{\mathcal{K}}^*\mathcal{T}_{\mathcal{K}}\rangle = -V(-\mathcal{M}). \tag{12.62}$$

When the spectrum of $\mathcal{M}$ has otherwise strictly positive elements, then $\mathcal{C}$ can get arbitrarily large, yielding $\chi(\mathcal{M}) = +\infty$.

■

We are now ready to prove Theorem 12.0.6.

Proof of Strong Duality in Theorem 12.0.6: By Lemma 12.3.1 and Lemma 12.3.5, we can write the primal objective (P) as a minimization of the sum of two convex functions over the cone of positive definite operators in $\mathscr{L}_\infty$:

$$\inf_{\mathcal{K}\in\mathscr{K}} f(\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}) = \inf_{\mathcal{C}\in\mathscr{L}_\infty} f(\mathcal{C}) + \chi(\mathcal{C}). \tag{12.63}$$

Since there exists $\mathcal{K}_0 \in \mathscr{K}$ and $\mathcal{C}_0 \in \mathscr{L}_\infty$ such that $\mathcal{C}_0 \succ \mathcal{T}_{\mathcal{K}_0}^*\mathcal{T}_{\mathcal{K}_0}$, then $\chi(\mathcal{C}_0) = 0$. Moreover, there is a neighborhood $\mathscr{N}_0 \subset \mathscr{L}_\infty$ around $\mathcal{C}_0$ such that $\chi(\mathcal{C}) = 0$ for all $\mathcal{C} \in \mathscr{N}_0$, hence χ is continuous at $\mathcal{C}_0$. Since $\mathcal{C}_0$ is in the domains of both f and χ, and χ is continuous at $\mathcal{C}_0$, *Fenchel-Rockafeller duality theorem* [56, Thm. 4.1], [254, Cor. 2.8.5] implies that the following strong duality holds:

$$\inf_{\mathcal{K}\in\mathscr{K}} f(\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}) = \inf_{\mathcal{C}\in\mathscr{L}_\infty} f(\mathcal{C}) + \chi(\mathcal{C}) = \max_{\mathcal{M}\in\mathscr{L}_1} -\chi^*(\mathcal{M}) - f^*(-\mathcal{M}), \tag{12.64}$$

and the maximum is attained at a point $\mathcal{M}_\star \in \mathscr{L}_1$. Plugging in the conjugate of χ derived in Lemma 12.3.5, we get

$$\begin{aligned}
\inf_{\mathcal{K}\in\mathscr{K}} f(\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}) &= \max_{\substack{\mathcal{M}\in\mathscr{L}_1,\\ \mathcal{M}\preccurlyeq 0}} V(-\mathcal{M}) - f^*(-\mathcal{M}), && (12.65)\\
&= \max_{\substack{\mathcal{M}\in\mathscr{L}_1,\\ \mathcal{M}\succcurlyeq 0}} V(\mathcal{M}) - f^*(\mathcal{M}), && (12.66)\\
&= \max_{\substack{\mathcal{M}\in\mathscr{L}_1,\\ \mathcal{M}\succcurlyeq 0}} \inf_{\mathcal{K}\in\mathscr{K}} \langle \mathcal{M},\, \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}\rangle - f^*(\mathcal{M}). && (12.67)
\end{aligned}$$

■

Chapter 13

PRIMAL-DUAL OPTIMALITY CONDITIONS

In the preceding section, we derived the dual reformulation (D) of the primal optimal control problem (P) and showed in Theorem 12.0.6 that the strong duality holds under some regularity conditions. We also presented various example control problems, both historically well-known and more recently studied, deriving their dual reformulations.

Towards the ultimate goal of devising a systematic and practical control synthesis paradigm for generalized control objectives, we aim to establish necessary and sufficient conditions for the optimality of a controller, if it exists, in the primal program (P) in this section. The primary utility of deriving the dual program lies in enabling explicit was to perform the optimization over the controller $\mathcal{K}$ by instead solving a simple $\mathcal{M}$-weighted $\mathscr{H}_2$ control problem at the expense of solving an additional optimization problem over the positive-definite dual variable $\mathcal{M} \succcurlyeq 0$. In particular, viewing the dual program (D) as a concave-convex zero-sum game between a positive-definite dual variable $\mathcal{M} \succcurlyeq 0$ and a causal controller $\mathcal{K} \in \mathscr{K}$, we show in Theorem 13.0.1 that the optimal solutions to the primal and dual control problems can readily be characterized as the saddle point of this game. The derivations rely on two main technical tools: (i) the Wiener-Hopf method to derive the optimal controller $\mathcal{K}$ in terms of the dual variable $\mathcal{M}$, and (ii) Karush-Kuhn-Tucker (KKT) conditions to derive the optimality conditions for the dual variable $\mathcal{M}$ in terms of the controller $\mathcal{K}$.

We devote the rest of this section to a detailed discussion of our approach and revisiting the previously introduced example control problems to derive the optimal controllers. It is worthwhile to note that although some of these controllers are already well-known and have been derived several times using various meticulously involved and specific techniques, our approach not only brings forth a novel and systematic methodology that collectively subsumes all these problems and the specific techniques in a single, unified framework, but it also enables deriving explicit forms of optimal controllers of previously open problems, such as the mixed $\mathscr{H}_2/\mathscr{H}_\infty$, as well as novel control objectives, such as $\mathscr{H}_p$-optimal control.

Theorem 13.0.1 (Optimality and Saddle Point Conditions)**.**

13.1 From Value Function to Wiener-Hopf Technique

Recall the strong duality (12.5) derived in Theorem 12.0.6:

$$\inf_{\mathcal{K}\in\mathscr{K}} f(\mathcal{T}_{\mathcal{K}}^{*}\mathcal{T}_{\mathcal{K}}) = \max_{\substack{\mathcal{M}\in\mathscr{L}_1,\\ \mathcal{M}\succcurlyeq 0}} \underbrace{\inf_{\mathcal{K}\in\mathscr{K}} \langle \mathcal{T}_{\mathcal{K}}^{*}\mathcal{T}_{\mathcal{K}},\, \mathcal{M}\rangle}_{=:V(\mathcal{M})} - f^{*}(\mathcal{M}), \tag{13.1}$$

where $V : \mathscr{L}_1 \to \overline{\mathbb{R}}$ is called the *value function* as defined in Definition 12.3.4. While a positive-definite solution $\mathcal{M}_{\star} \in \mathscr{L}_1$ to the dual objective is guaranteed to exist under the assumptions of Theorem 12.0.6, the same theorem does not explicitly provide such existential guarantees for an optimal causal and stabilizing LTI controller $\mathcal{K}_{\star} \in \mathscr{K}$ solving the primal control problem (P). Furthermore, we ultimately wish to obtain a computationally feasible set of necessary and sufficient conditions for its optimality in order to synthesize and deploy such a controller for real-time implementation.

As a first step towards characterizing the optimality conditions, we study the inner minimization in the dual control problem (13.1), namely the *value function*, $V(\mathcal{M})$, defined in Definition 12.3.4. As pointed out in **??**, this objective is simply a weighted $\mathscr{H}_2$ problem where the positive-definite weighting $\mathcal{M} \succcurlyeq 0$ shapes the spectrum of the closed-loop transfer operator $\mathcal{T}_{\mathcal{K}} : \mathbf{w} \mapsto \mathbf{z}$. Moreover, the weight $\mathcal{M} \succcurlyeq 0$ can be interpreted as an auto-covariance operator of a weakly (aka. wide-sense) stationary random process under a mild regularity condition (*i.e.*, *Paley-Wiener-Szegő criterion*), hence, elucidating and reinforcing the distributionally robust and risk-aversive interpretations of the dual control problem in Remark 12.2.4 as well as the game-theoretic view presented earlier.

In the rest of this subsection, we first state some useful properties of the value function $V(\mathcal{M})$ in Lemma 13.1.1. Then, in Theorem 13.3.5, we obtain a closed-form expression for a minimizing controller $\mathcal{K} \in \mathscr{K}$ in terms of dual variable $\mathcal{M}$ using the infamous *Wiener-Hopf technique*.

Lemma 13.1.1 (Properties of V)**.** *The following holds: the function $V : \mathscr{L}_1 \to \overline{\mathbb{R}}$ is*

i. proper, concave, upper-semicontinuous with the domain

$$\mathbf{dom}(V) = \{\mathcal{M} \in \mathscr{L}_1 \mid \mathcal{T}_2\mathcal{M}\mathcal{T}_2^{*} \succcurlyeq 0\}, \tag{13.2}$$

ii. positively homogeneous, i.e., $V(\lambda\mathcal{M}) = \lambda V(\mathcal{M})$ for any $\mathcal{M} \in \mathscr{L}_1$ and $\lambda \geq 0$, *iii.* *strictly monotonically increasing, i.e.,* $V(\mathcal{M}_1) < V(\mathcal{M}_2)$ *whenever* $\mathcal{M}_1 \prec \mathcal{M}_2$*,*

iv. *superadditive, i.e.,* $V(\mathcal{M}_1+\mathcal{M}_2) \geq V(\mathcal{M}_1)+V(\mathcal{M}_2)$ *for any* $\mathcal{M}_1, \mathcal{M}_2 \in \mathscr{L}_1$.

v. *bounded above as* $V(\mathcal{M}) \leq \|M\|_1 \inf_{\mathcal{K}\in\mathscr{K}} \|\mathcal{T}_\mathcal{K}\|_\infty^2$ *for any* $\mathcal{M} \in \mathscr{L}_1$*, and bounded below as* $V(\mathcal{M}) \geq \|\mathcal{M}^{-1}\|_\infty^{-1} \inf_{\mathcal{K}\in\mathscr{H}_\infty} \|\mathcal{T}_\mathcal{K}\|_2^2$ *whenever* $\mathcal{M} \succcurlyeq 0$.

Lemma 13.1.2 (Superdifferential of $V(\mathcal{M})$)**.** *The following properties about the superdifferential of* V *hold:*

i. *has non-empty, weak*-closed, and convex superdifferential* $\overline{\partial}V(\mathcal{M}) \subset \mathscr{L}_\infty^+$ *at every interior point* $\mathcal{M} \in \mathbf{ri}_{\|\cdot\|_1}(\mathbf{dom}(V))$ *(wrt. norm topology).*

ii. *A bounded and positive definite operator* $\mathcal{G}_0 \in \mathscr{L}_\infty$ *is a supergradient* $\mathcal{G}_0 \in \overline{\partial}V(\mathcal{M}_0)$ *at* $\mathcal{M}_0 \in \mathbf{dom}(V)$ *if and only if* $\mathcal{G}_0 \in \mathscr{C}_{\mathscr{H}_\infty}$ *and* $\langle \mathcal{G}_0,\, \mathcal{M}_0\rangle = V(\mathcal{M}_0)$ *(equivalently,* $\langle \mathcal{G}_0,\, \mathcal{M}\rangle \geq V(\mathcal{M})$ *for all positive definite* $\mathcal{M} \in \mathscr{L}_1$*, with equality at* $\mathcal{M}_0$*).*

iii. *If* $\mathcal{M}_0 \in \mathbf{dom}(V)$*, then* $\overline{\partial}V(\lambda\mathcal{M}_0) = \overline{\partial}V(\mathcal{M}_0)$ *for any* $\lambda > 0$.

Lemma 13.1.3 ($V(\mathcal{M})$ is monotonic under transformation)**.** *Let* $\mathcal{M} \in \mathscr{L}_1^{n\times n}$ *be positive definite and not exactly zero. Suppose* $\mathcal{M}$ *admits a canonical spectral factorization* $\mathcal{M} = \mathcal{L}\mathcal{L}^*$ *with* $\mathcal{L} \in \mathscr{H}_2^{n\times n}$ *with* $\mathbf{rank}(\mathcal{L}) \leq n$ *being left-outer and Let* $\mathcal{G} \in \overline{\partial}V(\mathcal{M})$ *be a supergradient. Define the transformation* $\tau : \mathcal{M} \to \mathcal{L}^*\mathcal{G}\mathcal{L}$. *Then,*

$$\frac{V(\tau(\mathcal{M}))}{\mathrm{tr}(\mathcal{M})} \geq \left(\frac{V(\mathcal{M})}{\mathrm{tr}(M)}\right)^2. \tag{13.3}$$

13.2 Inner-Outer and Spectral Factorizations

Definition 13.2.1 (Inner and Outer Operators [67, Ch. 7])**.** A bounded, causal operator $\mathcal{U} \in \mathscr{H}_\infty^{n\times m}$ (resp. anti-causal operator $\mathcal{V} \in {}^*\mathscr{H}_\infty^{n\times m}$) with $n \geq m$ is called *inner* (resp. *co-inner*) if $\mathcal{U}^*\mathcal{U} = \mathcal{I}_m$ (resp. $\mathcal{V}^*\mathcal{V} = \mathcal{I}_m$).

A causal operator $\mathcal{L} \in \mathscr{H}_p^{n\times m}$ with $n \geq m$ (resp. $n \leq m$) is called *left-outer* (resp. *right-outer*) if there exists a causal and (marginally) stable operator $\mathcal{G} \in \mathscr{H}^{m\times n}$ (not necessarily in $\mathscr{H}_p$) such that $\mathcal{G}\mathcal{L} = \mathcal{I}_m$ (resp. $\mathcal{L}\mathcal{G} = \mathcal{I}_n$). Such an operator $\mathcal{G}$ is called the left-inverse (resp. right-inverse) of $\mathcal{L}$ and denoted as $\mathcal{L}^\dagger := \mathcal{G}$. $\mathcal{L} \in \mathscr{H}_p^{n\times m}$ is simply called *outer* if it is both left- and right-outer with the same inverse and the inverse is denoted as $\mathcal{L}^{-1}$.

Lemma 13.2.2. *Given $\mathcal{L} \in \mathscr{H}_p^{n\times m}$ with $n \leq m$, the following statements hold:*

i. $\mathcal{L}$ is right-outer if and only if the set $\mathcal{L}\mathscr{T}_+^{n\times 1}$ is dense in $\mathscr{H}_p^{n\times 1}$ [103, Thm. 10].

ii. If $n = m$, $\mathcal{L}$ is outer if and only if there exists $\mathcal{H} \in \mathscr{H}_1^{n\times n}$ such that $\mathcal{L} = \exp(\mathcal{H})$.

Theorem 13.2.3 (Inner-Outer Factorization)**.** *For $p \in [1, \infty]$, let $\mathcal{T} \in \mathscr{H}_p^{n\times m}$ be a causal operator.*

i. If the left-absolute value $|\mathcal{T}|_l := (\mathcal{T}^*\mathcal{T})^{\frac{1}{2}} \in \mathscr{L}_p^{m\times m}$ is non-singular (i.e., $T(z) \in \mathbb{C}^{n\times m}$ is full column rank a.e. on $z \in \mathbb{T}$ with $n \geq m$), then there exists an inner operator $\mathcal{U} \in \mathscr{H}_\infty^{n\times m}$ and an outer operator $\Delta \in \mathscr{H}_p^{m\times m}$ such that $\mathcal{T} = \mathcal{U}\Delta$.

ii. If the right-absolute value $|\mathcal{T}|_r := (\mathcal{T}\mathcal{T}^*)^{\frac{1}{2}} \in \mathscr{L}_p^{n\times n}$ is non-singular (i.e., $T(z) \in \mathbb{C}^{n\times m}$ is full row rank a.e. on $z \in \mathbb{T}$ with $n \leq m$), then there exists a co-inner operator $\mathcal{V} \in {}^*\mathscr{H}_\infty^{m\times n}$ and an outer operator $\nabla \in \mathscr{H}_p^{n\times n}$ such that $\mathcal{T} = \nabla\mathcal{V}^*$.

Definition 13.2.4 (Paley-Wiener-Szegő Criterion)**.** A strictly positive-definite Laurent operator $\mathcal{M} \succ 0$ in $\mathscr{L}_1^{n\times n}$ with the transfer matrix $M(\mathrm{e}^{\jmath\omega}) \succ 0$ a.s. is an *auto-covariance operator* of a weakly stationary stochastic process if it satisfies the *Paley-Wiener-Szegő criterion*: The logarithm $\log(\mathcal{M})$ is in $\mathscr{L}_1$, *i.e.*,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{logdet}(M(\mathrm{e}^{\jmath\omega}))\,\mathrm{d}\omega > -\infty. \tag{13.4}$$

The transfer matrix $M(\mathrm{e}^{\jmath\omega}) \in \mathbb{S}_{++}^n$ of such an auto-covariance operator $\mathcal{M} \in \mathscr{L}_{1,++}^{n\times n}$ is called a *power spectral density*.

Theorem 13.2.5 (Spectral Factorization[240, Thm. 7.13], [102, Thm. 9])**.** *A positive-definite operator $\mathcal{M} \succcurlyeq 0$ in $\mathscr{L}_1^{n\times n}$ is an auto-covariance operator if and only if it admits a left (or right)* spectral factorization *$\mathcal{M} = \mathcal{L}\mathcal{L}^*$ (resp. $\mathcal{M} = \mathcal{G}^*\mathcal{G}$) where*

i. $\mathcal{L} \in \mathscr{H}_2^{n\times n}$ (resp. $\mathcal{G} \in \mathscr{H}_2^{n\times n}$) is a causal left (resp. right) spectral factor *with stable transfer matrix $L(\mathrm{e}^{\jmath\omega})$ (resp. $G(\mathrm{e}^{\jmath\omega})$), and* *ii.* $\det(L(\infty)) \neq 0$ *(resp.* $\det(G(\infty)) \neq 0$*) so that* $\mathcal{L}$ *(resp.* $\mathcal{G}$*) is invertible in* $\mathscr{H}_2^{n\times n}$*, i.e.,* $\mathcal{L}^{-1} \in \mathscr{H}_2^{n\times n}$ *(resp.* $\mathcal{G}^{-1} \in \mathscr{H}_2^{n\times n}$*).*

In that case, the Paley-Wiener-Szegő Criterion is equivalent to

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{logdet}(M(\mathrm{e}^{\jmath\omega}))\,\mathrm{d}\omega = 2\log|\det(L(\infty))| = 2\log|\det(G(\infty))| > -\infty. \tag{13.5}$$

The unique left (resp. right) spectral factor $\mathcal{L}$ *(resp.* $\mathcal{G}$*) with symmetric and strictly positive-definite* $L(\infty) \succ 0$ *(resp.* $G(\infty) \succ 0$*) is called the left (resp. right)* canonical spectral factor.

Notice that the Paley-Wiener-Szegő Criterion is trivially satisfied under the sufficient condition that $\mathcal{M} \succcurlyeq \alpha\mathcal{I}$ for a positive scalar $\alpha > 0$.

$$\mathcal{M} = \mathcal{L}\mathcal{R}\mathcal{L}^* \tag{13.6}$$

where $R(z) = R \in \mathbb{S}_{++}^{n\times n}$ for all $z \in \mathbb{C}$ and $\mathcal{L}, \mathcal{L}^{-1} \in \mathscr{H}_2^{n\times n}$ such that $L(\infty) = L^{-1}(\infty) = I_n$ and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{logdet}(M(\mathrm{e}^{\jmath\omega}))\,\mathrm{d}\omega = \operatorname{logdet}(R). \tag{13.7}$$

13.3 Weighted-$\mathscr{L}_2$ Spaces and Wiener-Hopf Projection

Definition 13.3.1 (Weighted $\mathscr{L}_2$ and $\mathscr{H}_2$ Spaces)**.** Let $\mathcal{M} \in \mathscr{L}_1$ and $\mathcal{M} \succcurlyeq 0$. We call $\mathcal{T} \in \mathscr{L}_{2,\mathcal{M}}$ if and only if $\mathcal{T}\mathcal{M}\mathcal{T}^* \in \mathscr{L}_1$. Furthermore, $\mathcal{T} \in \mathscr{H}_{2,\mathcal{M}}$ iff it is causal and $\mathcal{T} \in \mathscr{L}_{2,\mathcal{M}}$.

Lemma 13.3.2 (Properties of $\mathscr{L}_{2,\mathcal{M}}$ and $\mathscr{H}_{2,\mathcal{M}}$ [241, Lem. 4.1-4.2])**.** *Let* $\mathcal{M} \in \mathscr{L}_1$ *satisfy the Paley-Wiener-Szegő conditions and let* $\mathcal{M} = \mathcal{L}\mathcal{L}^*$ *be its canonical spectral factorization.*

i. The space $\mathscr{L}_{2,\mathcal{M}}$ *is a Hilbert space with the inner product and norms defined as*

$$\langle \mathcal{T}_1, \mathcal{T}_2\rangle_{\mathcal{M}} := \operatorname{tr}(\mathcal{T}_1\mathcal{M}\mathcal{T}_2^*), \quad \|\mathcal{T}\|_{\mathcal{M}} := \sqrt{\operatorname{tr}(\mathcal{T}\mathcal{M}\mathcal{T}^*)}. \tag{13.8}$$

ii. The space $\mathscr{H}_{2,\mathcal{M}} \subset \mathscr{L}_{2,\mathcal{M}}$ *is a sub-Hilbert space with the same inner product and norm.*

iii. $\mathcal{T} \in \mathscr{L}_{2,\mathcal{M}}$ *(resp.* $\mathcal{T} \in \mathscr{H}_{2,\mathcal{M}}$*) if and only if* $\mathcal{T}\sqrt{\mathcal{M}} \in \mathscr{L}_2$ *(resp.* $\mathcal{T}\mathcal{L} \in \mathscr{H}_2$*).*

iv. $\mathscr{L}_\infty \subset \mathscr{L}_{2,\mathcal{M}}$ *and* $\mathscr{H}_\infty \subset \mathscr{H}_{2,\mathcal{M}}$.

v. Rational functions $\mathscr{R}\mathscr{L}_\infty$ *(resp.* $\mathscr{R}\mathscr{H}_\infty$*) are dense in* $\mathscr{L}_{2,\mathcal{M}}$ *(resp.* $\mathscr{H}_{2,\mathcal{M}}$*) wrt the* $\|\cdot\|_\mathcal{M}$ *norm topology.*

vi. $\|\mathcal{T}\|_\mathcal{M} \leq \|\mathcal{T}\|_\infty \operatorname{tr}(\mathcal{M})$

Lemma 13.3.3 (Hilbert Projection Theorem)**.** *Let* $(H, \|\cdot\|)$ *be a Hilbert space. For any vector* $x \in H$ *and a closed convex subset* $C \subseteq H$*, there exists a unique point* $z_\star \in C$ *such that* $\|x - z_\star\| \leq \|x - z\|$ *for all* $z \in C$*. In other words, the mapping* $P_C : H \to C$ *defined as*

$$P_C(x) = \arg\min_{z\in C} \|x - z\|, \tag{13.9}$$

is well-defined and unique.

Lemma 13.3.4 (Projection)**.** *Let* $\mathcal{M} \in \mathscr{L}_1$ *be an auto-covariance operator with canonical spectral factorization* $\mathcal{M} = \mathcal{L}\mathcal{L}^*$*. Given a non-causal bounded operator* $\mathcal{K}_\circ \in \mathscr{L}_\infty$*, consider its best bounded and causal* $\mathscr{H}_\infty$ *approximation in* $\mathcal{M}$*-weighted* $\mathscr{L}_{2,\mathcal{M}}$ *norm:*

$$\mathbf{dist}^2_{\mathscr{L}_{2,\mathcal{M}}}(\mathcal{K}_\circ, \mathscr{H}_\infty) \coloneqq \inf_{\mathcal{K}\in\mathscr{K}} \left\{\|\mathcal{K} - \mathcal{K}_\circ\|^2_\mathcal{M} = \operatorname{tr}\left((\mathcal{K} - \mathcal{K}_\circ)\mathcal{M}(\mathcal{K} - \mathcal{K}_\circ)^*\right)\right\} \tag{13.10}$$

The following statements hold:

i. $\mathbf{dist}^2_{\mathscr{L}_{2,\mathcal{M}}}(\mathcal{K}_\circ, \mathscr{H}_\infty) = \mathbf{dist}^2_{\mathscr{L}_{2,\mathcal{M}}}(\mathcal{K}_\circ, \mathscr{H}_{2,\mathcal{M}}) \coloneqq \min_{\mathcal{K}\in\mathscr{H}_{2,\mathcal{M}}} \|\mathcal{K} - \mathcal{K}_\circ\|^2_\mathcal{M}$.

ii. The optimal distance is equal to $\mathbf{dist}^2_{\mathscr{L}_{2,\mathcal{M}}}(\mathcal{K}_\circ, \mathscr{H}_{2,\mathcal{M}}) = \|\{\mathcal{K}_\circ\mathcal{L}\}_-\|^2_2$*, which is achieved by* $\mathcal{K}_\star \coloneqq \{\mathcal{K}_\circ\mathcal{L}\}_+ \mathcal{L}^{-1} \in \mathscr{H}_{2,\mathcal{M}}$.

iii. There exists a sequence $\{\mathcal{K}_n\} \subset \mathscr{R}\mathscr{H}_\infty$ *such that* $\mathcal{K}_n \xrightarrow{\mathscr{H}_{2,\mathcal{M}}} \mathcal{K}_\star$*, namely,* $\lim_{n\to\infty} \|\mathcal{K}_n\mathcal{L} - \mathcal{K}_\star\mathcal{L}\|_2 = 0$ *and* $\lim_{n\to\infty} \|\mathcal{K}_n - \mathcal{K}_\circ\|^2_\mathcal{M} = \mathbf{dist}^2_{\mathscr{L}_{2,\mathcal{M}}}(\mathcal{K}_\circ, \mathscr{H}_\infty)$.

iv. If $\mathcal{M} \succ 0$ *and* $\mathcal{K}_\circ \in \mathscr{R}\mathscr{H}_\infty$*, then the optimal solution is bounded, i.e.,* $\mathcal{K}_\star \in \mathscr{H}_\infty$.

Proof.

$$\begin{aligned}
\|\mathcal{K} - \mathcal{K}_\circ\|^2_\mathcal{M} &= \|\mathcal{K} - \mathcal{K}_\star + \mathcal{K}_\star - \mathcal{K}_\circ\|^2_\mathcal{M} && (13.11)\\
&= \|\mathcal{K} - \mathcal{K}_\star\|^2_\mathcal{M} + \|\mathcal{K}_\star - \mathcal{K}_\circ\|^2_\mathcal{M}, && (13.12)\\
&= \|\mathcal{K}\mathcal{L} - \{\mathcal{K}_\circ\mathcal{L}\}_+\|^2_2 + V(\mathcal{M}) && (13.13)
\end{aligned}$$

<u>*Step 1: Density of $\mathscr{H}_\infty$.*</u> Since $\mathcal{L} \in \mathscr{H}_2$ is outer, and $\{\mathcal{K}_\circ\mathcal{L}\} \in \mathscr{H}_2$, there exists a sequence $\{\mathcal{K}_n\}_{n\in\mathbb{N}} \subset \mathscr{H}_\infty$ such that

$$\|\mathcal{K}_n\mathcal{L} - \{\mathcal{K}_\circ\mathcal{L}\}_+\|_2 \to 0. \tag{13.14}$$

<u>*Convergence to optimal value.*</u> This implies

$$\|\mathcal{K}_n - \mathcal{K}_\circ\|_{\mathcal{M}}^2 = \|\mathcal{K}_n\mathcal{L} - \{\mathcal{K}_\circ\mathcal{L}\}_+\|_2^2 + V(\mathcal{M}) \to V(\mathcal{M}) \tag{13.15}$$

<u>*Boundedness.*</u> As $\mathcal{M} \succ 0$, $0 \preccurlyeq \mathcal{M}^{-1} \in \mathscr{L}_\infty$ and $\mathcal{L}^{-1} \in \mathscr{L}_\infty$ is outer. Therefore, for any $\mathcal{K} \in \mathscr{H}_\infty$

$$\|\mathcal{K} - \mathcal{K}_\circ\|_\infty^2 \leq \|\mathcal{K} - \mathcal{K}_\circ\|_2^2 = \|(\mathcal{K}\mathcal{L} - \mathcal{K}_\circ\mathcal{L})\mathcal{L}^{-1}\|_2^2, \tag{13.16}$$

$$\leq \|\mathcal{K}\mathcal{L} - \mathcal{K}_\circ\mathcal{L}\|_2^2\,\|\mathcal{L}^{-1}\|_\infty^2 = \|\mathcal{K} - \mathcal{K}_\circ\|_{\mathcal{M}}^2\,\|\mathcal{M}^{-1}\|_\infty. \tag{13.17}$$

Thus,

$$\limsup_{n\to\infty} \|\mathcal{K}_n - \mathcal{K}_\circ\|_\infty^2 \leq \|\mathcal{M}^{-1}\|_\infty \limsup_{n\to\infty} \|\mathcal{K}_n - \mathcal{K}_\circ\|_{\mathcal{M}}^2 = \|\mathcal{M}^{-1}\|_\infty\, V(\mathcal{M}) \tag{13.18}$$

Moreover, for any $\mathcal{K} \in \mathscr{H}_\infty$, we have

$$\|\mathcal{K}\|_\infty = \|\mathcal{K} - \mathcal{K}_\circ + \mathcal{K}_\circ\|_\infty \leq \|\mathcal{K} - \mathcal{K}_\circ\|_\infty + \|\mathcal{K}_\circ\|_\infty, \tag{13.19}$$

by triangle inequality. Thus,

$$\limsup_{n\to\infty} \|\mathcal{K}_n\|_\infty \leq \sqrt{\|\mathcal{M}^{-1}\|_\infty\, V(\mathcal{M})} + \|\mathcal{K}_\circ\|_\infty. \tag{13.20}$$

By definition of $\limsup$, for any small enough $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that

$$\sup_{n\geq N_\epsilon} \|\mathcal{K}_n\|_\infty \leq \sqrt{\|\mathcal{M}^{-1}\|_\infty\, V(\mathcal{M})} + \|\mathcal{K}_\circ\|_\infty + \epsilon. \tag{13.21}$$

In other words, the sequence $\{\mathcal{K}_n\}_{n\geq N_\epsilon}$ is uniformly bounded. Since closed and bounded set in $\mathscr{H}_\infty$ are weak*-compact by Banach-Alaoglu theorem, there exists a uniformly bounded subsequence $\{\mathcal{K}_{n_k}\}_{k\in\mathbb{N}}$ and $\mathcal{K}_\infty \in \mathscr{H}_\infty$ such that $\mathcal{K}_{n_k} \xrightarrow{w^*} \mathcal{K}_\infty$.

■

Theorem 13.3.5 (Wiener-Hopf Technique [120], [250], [251])**.** *Without loss of generality, let $d_z \geq d_u$ and $d_w \geq d_y$. Suppose $\mathcal{T}_2\mathcal{M}\mathcal{T}_2^*$ for $\mathcal{M} \in \mathscr{L}_1^+$ satisfies Paley-Wiener-Szegő criterion in Definition 13.2.4. The* value function $V : \mathscr{L}_1 \to \overline{\mathbb{R}}$ *defined as*

$$V(\mathcal{M}) \triangleq \inf_{\mathcal{K}\in\mathscr{K}} \langle \mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K},\, \mathcal{M}\rangle = \operatorname{tr}(\mathcal{T}_\mathcal{K}^*\mathcal{T}_\mathcal{K}\mathcal{M}), \tag{13.22}$$

admits an optimal solution $\mathcal{K}_\star \in \mathscr{H}_{2,\mathcal{T}_2\mathcal{M}\mathcal{T}_2^*}^{d_u\times d_y}$ *given by*

$$\mathcal{K}_\star := \Delta^{-1}\left\{\Delta\mathcal{K}_{nc}\nabla\mathcal{L}\right\}_+\mathcal{L}^{-1}\nabla^{-1}, \tag{13.23}$$

where $\mathcal{K}_{nc} \in \mathscr{L}_{2,\mathcal{T}_2\mathcal{M}\mathcal{T}_2^*}^{d_u\times d_y}$ *is the optimal non-causal controller as*

$$\mathcal{K}_{nc} := -(\mathcal{T}_1^*\mathcal{T}_1)^{-1}\mathcal{T}_1^*\,\mathcal{T}_3\,\mathcal{M}\mathcal{T}_2^*(\mathcal{T}_2\mathcal{M}\mathcal{T}_2^*)^{-1}, \tag{13.24}$$

$\mathcal{T}_1 = \mathcal{U}\Delta$ *(resp.* $\mathcal{T}_2 = \nabla\mathcal{V}^*$*) is inner (resp. co-inner) factorization such that* $\mathcal{T}_1^*\mathcal{T}_1 = \Delta^*\Delta$ *(resp.* $\mathcal{T}_2\mathcal{T}_2^* = \nabla\nabla^*$*) is left (resp. right) canonical spectral factorization with* $\Delta, \Delta^{-1} \in \mathscr{RH}_\infty^{d_u\times d_u}$ *(resp.* $\nabla, \nabla^{-1} \in \mathscr{RH}_\infty^{d_y\times d_y}$*) outer and* $\mathcal{U} = \mathcal{T}_1\Delta^{-1} \in \mathscr{RH}_\infty^{d_z\times d_u}$ *(resp.* $\mathcal{V}^* = \nabla^{-1}\mathcal{T}_2 \in \mathscr{RH}_\infty^{d_y\times d_w}$*) is inner (resp. co-inner) with* $\mathcal{U}^*\mathcal{U} = \mathcal{I}_{d_u}$ *(resp.* $\mathcal{V}^*\mathcal{V} = \mathcal{I}_{d_y}$*), and* $\mathcal{L}\mathcal{L}^* := \mathcal{U}^*\mathcal{M}\mathcal{U}$ *is right-canonical spectral factorization with* $\mathcal{L} \in \mathscr{H}_2^{d_y\times d_y}$ *outer.*

Moreover, if $\mathcal{T}_2\mathcal{M}\mathcal{T}_2^* \succ 0$ *strictly positive, then* $\mathcal{K}_\star \in \mathscr{H}_2^{d_u\times d_y}$*, and if* $\mathcal{M} \in \mathscr{L}_\infty$*, then* $\mathcal{K}_\star \in \mathrm{BMOA} \subsetneq \bigcap_{p\geq 1}\mathscr{H}_{2p}^{d_u\times d_y}$.

Proof. We recast the problem with unitary extensions of the inner and co-inner factors so that only one block of a 2×2 matrix depends on the controller. The minimization then reduces to an orthogonal projection in $\mathscr{H}_2$.

1. Inner / co-inner completions. From the left-canonical factorization $\mathcal{T}_1 = \mathcal{U}\Delta$ we already have $\mathcal{U} \in \mathscr{RH}_\infty^{d_z\times d_u}$ inner ($\mathcal{U}^*\mathcal{U} = I_{d_u}$). Because $d_z \geq d_u$ there exists an *inner complement*

$$\widetilde{\mathcal{U}} \in \mathscr{RH}_\infty^{d_z\times(d_z-d_u)} \quad \text{with} \quad \overline{\mathcal{U}} := \begin{bmatrix}\mathcal{U} & \widetilde{\mathcal{U}}\end{bmatrix} \quad \Longrightarrow \quad \overline{\mathcal{U}}^*\overline{\mathcal{U}} = \overline{\mathcal{U}}\,\overline{\mathcal{U}}^* = I_{d_z}.$$

Likewise, the right-canonical factorization $\mathcal{T}_2 = \nabla\mathcal{V}^*$ gives the *co-inner* $\mathcal{V}^* \in \mathscr{RH}_\infty^{d_y\times d_w}$ ($\mathcal{V}^*\mathcal{V} = I_{d_y}$). Because $d_w \geq d_y$ we can choose a *co-inner complement* $\widetilde{\mathcal{V}}^*$ so that

$$\overline{\mathcal{V}} := \begin{bmatrix}\mathcal{V} & \widetilde{\mathcal{V}}\end{bmatrix} \quad \text{satisfies} \quad \overline{\mathcal{V}}^*\overline{\mathcal{V}} = \overline{\mathcal{V}}\,\overline{\mathcal{V}}^* = I_{d_w},$$

hence both $\overline{\mathcal{U}}$ and $\overline{\mathcal{V}}$ are *unitary* transfer matrices.

2. Four-block lifting of the closed loop. For a causal controller $\mathcal{K}$ set $\mathcal{Q} = \Delta\mathcal{K}\nabla$. The standard Youla parameterization gives

$$\mathcal{T}_{\mathcal{K}} = \mathcal{V}\,\mathcal{Q}\,\mathcal{U}^* \;+\; \mathcal{T}_3.$$

Introduce the lifted map

$$S_{\mathcal{K}} := \overline{\mathcal{V}}^{*}\,\mathcal{T}_{\mathcal{K}}\overline{\mathcal{U}} = \begin{bmatrix} \mathcal{Q}+\mathcal{V}^{*}\mathcal{T}_{3}\mathcal{U} & \mathcal{V}^{*}\mathcal{T}_{3}\widetilde{\mathcal{U}} \\ \widetilde{\mathcal{V}}^{*}\mathcal{T}_{3}\mathcal{U} & \widetilde{\mathcal{V}}^{*}\mathcal{T}_{3}\widetilde{\mathcal{U}} \end{bmatrix}.$$

Because $\overline{\mathcal{V}}$ and $\overline{\mathcal{U}}$ are unitary,

$$V(\mathcal{M}) = \mathrm{tr}(\mathcal{T}_{\mathcal{K}}^{*}\mathcal{T}_{\mathcal{K}}\mathcal{M}) = \mathrm{tr}\big(S_{\mathcal{K}}^{*}S_{\mathcal{K}}\,\overline{\mathcal{U}}^{*}\mathcal{M}\overline{\mathcal{U}}\big).$$

3. Block-dispensational of the weight. Let $\mathcal{M} = \mathcal{G}\mathcal{G}^{*}$ be the canonical $\mathscr{H}_2$ spectral factorization and factor $\mathcal{U}^{*}\mathcal{M}\mathcal{U} = \mathcal{L}\mathcal{L}^{*}$ with $\mathcal{L}, \mathcal{L}^{-1} \in \mathscr{H}_2$ outer. A direct algebra (Schur complement) shows

$$\overline{\mathcal{U}}^{*}\mathcal{M}\overline{\mathcal{U}} = \begin{bmatrix} \mathcal{L}\mathcal{L}^{*} & 0 \\ 0 & \mathcal{M}_{\perp} \end{bmatrix}, \qquad \mathcal{M}_{\perp} \succeq 0, \qquad \mathcal{M}_{\perp} \text{ independent of } \mathcal{Q}.$$

Consequently

$$V(\mathcal{M}) = \big\|\,(\mathcal{Q}+\mathcal{V}^{*}\mathcal{T}_{3}\mathcal{U})\,\mathcal{L}\big\|_{\mathscr{H}_2}^{2} \;+\; \mathrm{tr}\big(\widetilde{\mathcal{V}}^{*}\mathcal{T}_{3}\widetilde{\mathcal{U}}\;\mathcal{M}_{\perp}\,\widetilde{\mathcal{U}}^{*}\mathcal{T}_{3}^{*}\widetilde{\mathcal{V}}\big),$$

and the second term no longer depends on $\mathcal{Q}$.

4. Orthogonal projection in $\mathscr{H}_2$. Define

$$\mathcal{X} := \mathcal{V}^{*}\mathcal{T}_{3}\mathcal{M}\mathcal{U}\mathcal{L}^{-*} \in \mathscr{L}_{2}^{d_u\times d_y}.$$

The same norm estimate as in the original proof yields $\mathcal{X} \in \mathscr{L}_2$, whence the orthogonal decomposition $\mathcal{X} = \{\mathcal{X}\}_{+} + \{\mathcal{X}\}_{-}$ (with $\{\cdot\}_{+}$ the orthogonal projector onto $\mathscr{H}_2$). Choosing

$$\boxed{\mathcal{Q}_{\star} = -\{\mathcal{X}\}_{+}\,\mathcal{L}^{-1} \in \mathscr{H}_{2}^{d_u\times d_y}}$$

forces the analytic part of $(\mathcal{Q}+\mathcal{V}^{*}\mathcal{T}_{3}\mathcal{U})\mathcal{L}$ to vanish and attains the global minimum of $V(\mathcal{M})$.

5. Optimal controller and regularity. Since $\mathcal{K} = \Delta^{-1}\mathcal{Q}\nabla^{-1}$ we finally obtain

$$\boxed{\mathcal{K}_{\star} = \Delta^{-1}\{\,\Delta\mathcal{K}_{\mathrm{nc}}\nabla\mathcal{L}\,\}_{+}\,\mathcal{L}^{-1}\nabla^{-1} = \Delta^{-1}\big\{\,\mathcal{V}^{*}\mathcal{T}_{3}\mathcal{M}\mathcal{U}\mathcal{L}^{-*}\,\big\}_{+}\,\mathcal{L}^{-1}\nabla^{-1}} \in \mathscr{H}_{2,\mathcal{T}_2\mathcal{M}\mathcal{T}_2^{*}}^{d_u\times d_y}.$$

If $\mathcal{T}_{2}\mathcal{M}\mathcal{T}_{2}^{*} \succ 0$ then every factor is analytic and square integrable, hence $\mathcal{K}_{\star} \in \mathscr{H}_2$. When $\mathcal{M} \in \mathscr{L}_{\infty}$, $\mathcal{X} \in \mathrm{BMOA}$, so $\mathcal{K}_{\star} \in \mathrm{BMOA} \subsetneq \bigcap_{p\geq 1}\mathscr{H}_{2p}$.

This completes the proof. The four-block formulation shows that the inner complements $\widetilde{\mathcal{U}}, \widetilde{\mathcal{V}}$ merely embed the problem into a unitary environment; only the *upper-left* block is optimization-relevant.

■

Theorem 13.3.6 (Sub-Nehari Problem)**.** *Let $\mathcal{K}_\circ \in \mathscr{L}_\infty$ and $\mathcal{C} \in \mathscr{L}_\infty$ be self-adjoint. Then, there exists a bounded and causal operator $\mathcal{K} \in \mathscr{H}_\infty$ such that*

$$(\mathcal{K} - \mathcal{K}_\circ)^*(\mathcal{K} - \mathcal{K}_\circ) \prec \mathcal{C} \tag{13.25}$$

if and only if there exists a J-canonical factorization

$$\begin{bmatrix} \mathcal{I} & -\mathcal{K}_\circ \\ -\mathcal{K}_\circ^* & -\mathcal{C} + \mathcal{K}_\circ^*\mathcal{K}_\circ \end{bmatrix} = \begin{bmatrix} \mathcal{L}_{11}^* & \mathcal{L}_{21}^* \\ \mathcal{L}_{12}^* & \mathcal{L}_{22}^* \end{bmatrix} \begin{bmatrix} \mathcal{I} & \mathbf{0} \\ \mathbf{0} & -\mathcal{I} \end{bmatrix} \begin{bmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{bmatrix}, \tag{13.26}$$

where

$$\begin{bmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{bmatrix}, \begin{bmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{bmatrix}^{-1}, \text{ and } \mathcal{L}_{11}^{-1} \tag{13.27}$$

are causal and bounded, and $\mathcal{L}_{21}$ is strictly causal. All such causal and bounded operators are parameterized asver

$$\mathcal{K} = (\mathcal{L}_{11} - \mathcal{S}\mathcal{L}_{21})^{-1}(\mathcal{S}\mathcal{L}_{22} - \mathcal{L}_{12}) \tag{13.28}$$

for any causal and strictly contractive operator $\mathcal{S}$.

13.4 Finite-Dimensional Parametrization of $\mathcal{M}_\star$

We first obtain an equivalent condition of optimality of $\mathcal{M}_\star$.

Lemma 13.4.1. *Define the anti-causal operator $\mathcal{T} \coloneqq \{\Delta\mathcal{K}_\circ\}_{-}$. The optimality condition in* (9.16) *takes the equivalent form:*

$$\mathcal{L}_\star^*\mathcal{L}_\star = \frac{1}{4}\left(\mathcal{I} + \sqrt{\mathcal{I} + 4\gamma_\star^{-1}\{\mathcal{T}\mathcal{L}_\star\}_{-}^*\{\mathcal{T}\mathcal{L}_\star\}_{-}}\right)^2 \tag{13.29}$$

where $\gamma_\star > 0$ is such that $\mathsf{BW}(\mathcal{L}_\star\mathcal{L}_\star^*, \mathcal{I}) = r$.

Denoting $\mathcal{N}_\star \coloneqq \mathcal{L}_\star^*\mathcal{L}_\star$, there exists a one-to-one mapping between $\mathcal{M}_\star = \mathcal{L}_\star\mathcal{L}_\star^*$ and $\mathcal{N}_\star$ due to the uniqueness of the spectral factorization. Consequently, we interchangeably refer to both $\mathcal{N}_\star$ and $\mathcal{M}_\star$ as the optimal solution. The following theorem characterizes the optimal $\mathcal{N}_\star$ in the frequency domain, implying a finite-dimensional parametrization.

Theorem 13.4.2. *Denoting by* $T(z)=\overline{C}(z^{-1}I-\overline{A})^{-1}\overline{B}$ *the transfer function of the anti-causal operator* $\mathcal{T}=\{\Delta\mathcal{K}_\circ\}_-$*, let* $f:(\gamma,\Gamma)\in\mathbb{R}\times\mathbb{R}^{d_x\times d_w}\mapsto\mathcal{N}$ *return a pd operator with a transfer function* $z\in\mathbb{T}\mapsto N(z)$ *taking the form*

$$\frac{1}{4}\left(I+\sqrt{I+4\gamma^{-1}\Gamma^*(z^{-1}I-\overline{A})^{-*}\overline{C}^*\overline{C}(z^{-1}I-\overline{A})^{-1}\Gamma}\right)^2,$$

where $(\overline{A},\overline{B},\overline{C})$ *are obtained from the state-space parameters of the system in* (9.2) *(see Section 9.D). Then, the optimal solution* $\mathcal{N}_\star=\mathcal{L}_\star^*\mathcal{L}_\star$ *in* (13.29) *satisfies* $\mathcal{N}_\star=f(\gamma_\star,\Gamma_\star)$ *where*

$$\Gamma_\star\coloneqq\frac{1}{2\pi}\int_0^{2\pi}(I-\mathrm{e}^{\jmath\omega}\overline{A})^{-1}\overline{B}L_\star(\mathrm{e}^{\jmath\omega})d\omega, \tag{13.30}$$

and $\gamma_\star>0$ *is such that* $\mathsf{BW}(\mathcal{L}_\star\mathcal{L}_\star^*,\mathcal{I})=r$*.*

Notice in Theorem 13.4.2 that $N_\star(z)$ involves the square root of a rational term. In general, the square root does not preserve rationality. We thus get Corollary 13.4.3.

Corollary 13.4.3. *The optimal* W_2*-DR-RO controller,* $K_\star(z)$*, and* $N_\star(z)$ *are non-rational. Thus,* $K_\star(z)$ *does not admit a finite-dimensional state-space form.*

Given the non-rationality of the controller $K_\star(z)$, Kargin*, Hajar*, Malik*, *et al.* [123] proposes a fixed-point algorithm exploiting the finite-parametrization of the controller. In the next section, we propose an alternative efficient optimization algorithm, which, in contrast to the fixed-point algorithm, has *provable convergence guarantee* to the saddle point $(\mathcal{K}_\star,\mathcal{N}_\star)$.

Chapter 14

FREQUENCY-DOMAIN OPTIMIZATION TECHNIQUES

State-space controllers form the backbone of modern control engineering. Their compact structure, interpretability, and ease of implementation make them particularly well-suited for real-time and embedded applications. Moreover, state-space realizations admit highly efficient and scalable implementations, allowing them to be deployed reliably in high-dimensional systems. As such, it is a natural and desirable goal to design finite-dimensional state-space controllers for any real-world setting.

Nonetheless, as demonstrated in the preceding chapter, the optimal controllers arising from many advanced control problems—particularly those involving distributional robustness or multi-objective criteria—are often *non-rational*. That is, they do not admit realizations as finite-dimensional state-space systems and are fundamentally non-implementable in exact form with finite-memory. Consequently, the most viable alternative is to seek a suboptimal controller of finite degree that offers provable performance guarantees.

Concretely, we seek to synthesize a causal, finite-dimensional state-space controller $\widehat{\mathcal{K}} \in \mathscr{K}$ of finite McMillan degree $\deg(\widehat{\mathcal{K}}) = d_K \in \mathbb{N}$ specified by the realization

$$\begin{aligned} s_{t+1} &= A_{\widehat{\mathcal{K}}} s_t + B_{\widehat{\mathcal{K}}} y_t, \\ u_t &= C_{\widehat{\mathcal{K}}} s_t + D_{\widehat{\mathcal{K}}} y_t, \end{aligned} \tag{14.1}$$

where $s_t \in \mathbb{R}^{d_K}$ denotes the internal state of the controller, and $(A_{\widehat{\mathcal{K}}}, B_{\widehat{\mathcal{K}}}, C_{\widehat{\mathcal{K}}}, D_{\widehat{\mathcal{K}}})$ are its state-space parameters. Let $d_{\mathcal{P}} := \max(d_x, d_u, d_w, d_y, d_z)$ denote the effective dimension of the plant model and $J_\star := \inf_{\mathcal{K} \in \mathscr{K}} J(\mathcal{K})$ be the optimal achievable performance.

A controller is said to be *practically implementable* if it satisfies a prescribed degree bound d_K, achieves a specified relative performance loss $\varepsilon > 0$, and can be synthesized efficiently. We formalize this synthesis problem as follows:

Problem 14.0.1 (Practical Controller Design). Given a degree bound $d_K \in \mathbb{N}$ and a target relative performance gap $\varepsilon_P > 0$, develop an efficient algorithm to synthesize a controller

$$\widehat{\mathcal{K}} \equiv \left[\begin{array}{c|c} A_{\widehat{\mathcal{K}}} & B_{\widehat{\mathcal{K}}} \\ \hline C_{\widehat{\mathcal{K}}} & D_{\widehat{\mathcal{K}}} \end{array}\right] \in \mathscr{K} \cap \mathscr{RH}_\infty \tag{14.2}$$

with $\deg(\widehat{\mathcal{K}}) \leq d_K$ and computable in time $\mathrm{O}(\mathsf{poly}(d_{\mathcal{P}}, d_K, \varepsilon^{-1}))$, such that the primal relative suboptimality gap satisfies

$$\mathsf{gap}_P(\widehat{\mathcal{K}}) \coloneqq \frac{J(\widehat{\mathcal{K}}) - J_\star}{J_\star} \leq \varepsilon_P. \tag{14.3}$$

This design problem inherently reflects a three-way trade-off between performance accuracy, controller complexity (*i.e.*, degree), and computational complexity. To make this trade-off explicit, we distinguish between two complementary problem formulations:

- **Fixed-Degree Design with Optimal Accuracy:** Given a prescribed maximum degree $d_K \in \mathbb{N}$, develop an algorithm that synthesizes the best-performing rational controller of degree at most d_K, *i.e.*, solve:

$$\inf_{\widehat{\mathcal{K}} \in \mathscr{K} \cap \mathscr{R}\mathscr{H}_\infty} \quad J(\widehat{\mathcal{K}}) \quad \text{s.t.} \quad \deg(\widehat{\mathcal{K}}) \leq d_K. \tag{14.4}$$

- **Fixed-Accuracy Design with Minimal Degree:** Given a prescribed performance accuracy $\varepsilon_P > 0$, develop an algorithm that synthesizes a rational controller of minimal degree achieving the desired accuracy, *i.e.*, solve:

$$\inf_{\widehat{\mathcal{K}} \in \mathscr{K} \cap \mathscr{R}\mathscr{H}_\infty} \quad \deg(\widehat{\mathcal{K}}) \quad \text{s.t.} \quad \frac{J(\widehat{\mathcal{K}}) - J_\star}{J_\star} \leq \varepsilon_P. \tag{14.5}$$

While it is desirable to directly solve either of these *finite-dimensional* optimization problems rather than confronting the original infinite-dimensional optimal control problem, both formulations (14.4) and (14.5) are computationally intractable as the set of causal systems with bounded McMillan degree is highly non-convex, thereby rendering the resulting optimization problems non-convex and NP-hard[1].

Dual Perspective

A promising alternative to synthesizing practically implementable finite-degree controllers is to approach it from the perspective of the dual optimization problem, as developed in detail in the preceding chapters. Within the Wiener–Hopf framework, it is known that a finite-degree dual solution $\widehat{\mathcal{M}}$ induces a corresponding finite-degree controller $\widehat{\mathcal{K}}$. This observation motivates the strategy of searching for a finite-degree suboptimal solution $\widehat{\mathcal{M}}$ to the dual control problem, which comes similarly in two flavors:

[1]The same prohibitive complexity issue persist even if one seeks to synthesize the closed-loop system directly, as in System Level Synthesis, rather than designing the controller explicitly.

- **Fixed-Degree Dual Optimization:**

$$\sup_{\widehat{\mathcal{M}} \in \mathscr{RL}_1^+} V(\widehat{\mathcal{M}}) - f^*(\widehat{\mathcal{M}}) \quad \text{s.t.} \quad \deg(\widehat{\mathcal{M}}) \le d_M, \tag{14.6}$$

- **Fixed-Accuracy Dual Optimization:**

$$\inf_{\widehat{\mathcal{M}} \in \mathscr{RL}_1^+} \deg(\widehat{\mathcal{M}}) \quad \text{s.t.} \quad \frac{J_\star - (V(\widehat{\mathcal{M}}) - f^*(\widehat{\mathcal{M}}))}{J_\star} \le \varepsilon_D, \tag{14.7}$$

where $d_M \in \mathbb{N}$ is a prescribed upper bound on the McMillan degree of the rational dual variable $\widehat{\mathcal{M}} \in \mathscr{RL}_1^+$ and $\varepsilon_D > 0$ is a target bound on the dual suboptimality gap.

Despite their appeal, these formulations remain non-convex and computationally intractable for the same fundamental reason: the inherent non-convexity of the degree constraint. As a result, directly searching over finite-degree rational power spectral densities in the dual domain is no more tractable than in the primal setting.

Moreover, even if a viable solution $\widehat{\mathcal{M}} \in \mathscr{RL}_1^+$ with dual suboptimality certificate ε_D is successfully obtained, and a finite-degree controller $\widehat{\mathcal{K}} \in \mathscr{K} \cap \mathscr{RH}_\infty$ is constructed from $\widehat{\mathcal{M}}$ via the Wiener-Hopf technique, one must still establish a corresponding primal performance guarantee. In other words, it is necessary to convert the dual suboptimality certificate ε_D into a primal suboptimality bound ε_P for the synthesized controller $\widehat{\mathcal{M}}$. This conversion is generally nontrivial, and no universal theory guarantees its validity for arbitrary convex problems. Explicit bounds on the primal gap can be derived in certain structured problems.

This begs the question:

> *How can we construct feasible and practically implementable finite-degree state-space controllers $\widehat{\mathcal{K}}$, as in* (14.1)*, that satisfy a prescribed suboptimality gap and a maximum allowable McMillan degree, in a computationally tractable manner, for a given convex infinite-horizon control problem with generalized performance objectives?*

Our Approach: Optimize-then-Approximate

Although the inherent non-convexity and computational intractability of finite-degree controller synthesis may initially appear insurmountable, the central thesis of this dissertation is that the underlying infinite-dimensional optimization problem can, in fact, be approached in a computationally tractable manner. This tractability stems from the convexity of the formulation and the finite-dimensional structure of the plant dynamics. As detailed in the preceding chapters, the optimal solutions to both the primal and dual formulations of the control problem can often be characterized as a saddle point $(\mathcal{K}_\star, \mathcal{M}_\star)$ through the Karush–Kuhn–Tucker (KKT) conditions. Moreover, in a broad class of relevant problems, the dual optimal solution $M_\star(z)$ admits a closed-form, non-rational transfer function representation that is parametrized by a finite-dimensional object uniquely identifying the solution.

This observation motivates a two-stage approach:

i. **Numerically solve the infinite-dimensional optimization:** Rather than attempting to solve the finite-dimensional but inherently non-convex and computationally intractable formulations in (14.4),(14.5),(14.6), and (14.7), we instead pursue the solution of the original infinite-dimensional dual control problem directly. While both the primal and dual decision variables $(\mathcal{K}, \mathcal{M})$ are intrinsically infinite-dimensional operators, we exploit their transfer function representations in the Fourier domain to develop high-fidelity and scalable numerical optimization algorithms.

ii. **Approximate the solution with a rational controller:** Upon obtaining a high-fidelity solution to the infinite-dimensional control problem, direct implementation may still not be viable if the associated transfer function is non-rational. In such cases, we proceed to construct a rational controller of bounded McMillan degree that approximates the optimal infinite-dimensional solution. This approximation must be carried out in a manner that ensures theoretical guarantees on the resulting suboptimality, thereby enabling reliable implementation without significant degradation in performance.

In this chapter, we focus exclusively on the first stage of the proposed framework: the efficient computation of approximate solutions to infinite-dimensional optimization problems. The second stage, namely, the approximation and realization of these solutions via finite-dimensional rational controllers, is deferred to the subsequent chapter.

Transfer Function Representation and Computational Amenability

The primary mathematical objects considered in this study are infinite-dimensional operators that either represent linear time-invariant (LTI) systems—such as the controller $\mathcal{K}$, the closed-loop operator $\mathcal{T}_{\mathcal{K}}$, and related constructs—or encode statistical correlations, such as the autocovariance operator $\mathcal{M}$ of stochastic disturbances. This representation offers a compact and powerful means of describing the global behavior of dynamical systems and stochastic processes across the entire time horizon. Their algebraic manipulation follows rules closely resembling those of finite-dimensional matrix algebra, making them analytically tractable and conceptually familiar.

Crucially, these operators admit several equivalent representations, among which the most computationally tractable is the transfer function (or Fourier) representation. Specifically, any Laurent operator $\mathscr{X} \in \mathscr{L}_p(\mathbb{V}, \mathbb{W})$ admits an associated transfer matrix function $X \in L_p(\mathbb{T}, \mathscr{B}(\mathbb{V}, \mathbb{W}))$ defined by its Fourier series:

$$X(\mathrm{e}^{\jmath\omega}) = \sum_{t\in\mathbb{Z}} \widehat{X}_t \mathrm{e}^{-\jmath\omega t}, \tag{14.8}$$

where $\omega \in [-\pi, \pi)$. This representation provides a concrete object—namely, a function defined on the unit circle $\mathbb{T}$—that is well-suited for digital computation and numerical optimization.

Although the time-domain (via Markov parameters $(\widehat{X}_t)_{t\in\mathbb{N}}$), operator-theoretic (via $\mathscr{X}$), and frequency-domain (via $X(z)$) formulations are all mathematically equivalent, the transfer function representation is especially advantageous from a computational standpoint. Many systems of practical interest exhibit spectra composed of simple or structured components, such as those arising from first-order IIR or AR(1) processes, which allows for an effective finite-dimensional representation, such as rational functions.

Even in instances where the optimal transfer functions are inherently non-rational, as in Wasserstein distributionally robust control and mixed $\mathscr{H}_2/\mathscr{H}_\infty$ formulations, it is often possible to express these functions through structured parametric forms defined over finite-dimensional parameter spaces. These representations, while not rational in the classical sense, nonetheless permit tractable evaluation, manipulation, and optimization. Moreover, even in the absence of an explicit finite-dimensional parametrization, a transfer function can still be effectively approximated by discretizing the frequency domain and storing a finite (albeit potentially large) set of spectral samples on the unit circle. This discretized spectral representation retains sufficient fidelity for many practical purposes and facilitates algorithmic implementation. As such, the transfer function framework offers a particularly suitable foundation for digital computation. In what follows, we shall develop efficient numerical algorithms that exploit the structural advantages of the transfer function representation to enable tractable optimization in infinite-dimensional controller synthesis problems.

Chapter Overview

Whereas the preceding chapters established duality and optimality results in a general abstract framework and subsequently applied these foundational master theorems to derive explicit solutions for specific problem instances, the present chapter adopts a complementary, bottom-up approach. This reversal is motivated by the intrinsic difficulty of designing a universal algorithm that simultaneously exhibits all desirable properties, such as efficiency, scalability, and robustness, across a broad class of problems. In contrast, leveraging structural features unique to particular problem settings often enables the design of more specialized and computationally efficient algorithms. Accordingly, we begin by examining algorithmic strategies tailored to specific subclasses of problems and culminate with a general-purpose framework that integrates insights from these specialized cases.

In particular, Section 14.1 focuses on problems for which the optimal dual solution $M_\star(z)$ admits a finite-dimensional parametric transfer function representation. The most favorable and practically relevant scenario occurs when this parametrization permits the reformulation of the original infinite-dimensional convex program into an equivalent finite-dimensional convex problem. Such reformulations significantly reduce computational burden and allow for the direct application of standard convex optimization techniques. Even when an exact finite-dimensional convex reformulation is unavailable, the underlying parametric structure of $M_\star(z)$ can still be exploited to construct efficient algorithms, such as those based on fixed-point iterations. These approaches reduce the computation of $M_\star(z)$ to a finite-dimensional parameter search constrained by optimality conditions.

The most technically challenging regime arises when the dual solution lacks a tractable closed-form or parametric representation. In such instances, the discretization of the transfer function spectrum becomes indispensable for practical computation. Section 14.2 and Section 14.3 address these cases in the context of constrained and unconstrained dual formulations, respectively. Beyond the inherent difficulties posed by spectral discretization, an additional obstacle is the inapplicability of standard first-order methods. The underlying optimization problems are typically posed over Banach spaces, often endowed with weak topologies that do not admit a natural metric structure. Consequently, classical convergence analyses grounded in notions such as strong convexity or Lipschitz continuity defined with respect to a metric fail to apply. To address this, we introduce a class of provably convergent and computationally tractable algorithms based on Bregman divergences. These divergences serve as a more appropriate surrogate for metric-based notions, thereby facilitating the development of convergence theory in this more general setting.

14.1 Algorithms Utilizing Finite-Dimensional Parametrization

In this section, we develop provably efficient algorithms for computing the optimal dual solution $\mathcal{M}_\star$, in instances where it admits a closed-form and tractable parametric representation in the Fourier domain. As a starting point, we recall the dual formulation of the full-information Wasserstein-2 regret-optimal distributionally robust control problem. Specifically, consider the setting with the nominal autocovariance $\mathcal{M}_\circ = \mathcal{I}$ and the regularized γ-suboptimal form of this problem for $\gamma > \gamma_{\mathrm{RO}}$:

$$\max_{\substack{\mathcal{M}\in\mathscr{L}_1^+, \\ \mathcal{M}\succcurlyeq\mathcal{I}}} V(\mathcal{M}) - \gamma \operatorname{tr}\left((\sqrt{\mathcal{M}} - \mathcal{I})^2\right). \tag{14.9}$$

As established in the preceding chapter, the optimal solution $\mathcal{M}_\star$ admits a canonical spectral factorization $\mathcal{M}_\star = \mathcal{L}_\star\mathcal{L}_\star^* \in \mathscr{L}_\infty^+$ with a canonical factor $\mathcal{L}_\star \in \mathscr{H}_\infty$ whose Fourier-domain representation satisfies the parametric form:

$$L_\star(\mathrm{e}^{\jmath\omega})^* L_\star(\mathrm{e}^{\jmath\omega}) = \left(\frac{I + \sqrt{I + 4\gamma^{-1}\Gamma_\star^\intercal(\mathrm{e}^{\jmath\omega} I - A_{\mathsf{lqr}})^{-1}\overline{C}^\intercal\overline{C}(\mathrm{e}^{-\jmath\omega} I - A_{\mathsf{lqr}}^\intercal)^{-1}\Gamma_\star}}{2}\right)^2,$$

for $\omega \in [-\pi, \pi)$, where $A_{\mathsf{lqr}} := A - B_u K_{\mathsf{lqr}}$ denotes the closed-loop transition matrix associated with the linear-quadratic regulator (LQR), and $K_{\mathsf{lqr}} := (R + B_u^\intercal P B_u)^{-1} B_u^\intercal P A$ is the optimal LQR feedback gain. The matrix $\overline{C} := -(R + B_u^\intercal P B_u)^{-1/2} B_u^\intercal$ depends solely on the plant's state-space data, and $P = \mathtt{dare}(A, B_u, Q, R)$ is the unique positive semidefinite solution of the discrete-time algebraic Riccati equation (DARE) associated with the LQR problem.

The optimal finite-dimensional parameter $\Gamma_\star \in \mathbb{R}^{d_x \times d_w}$, which completely characterizes the dual solution, is determined via the following integral involving the canonical factor $L_\star(\mathrm{e}^{\jmath\omega})$:

$$\Gamma_\star := \frac{1}{2\pi}\int_{-\pi}^{\pi} (I - \mathrm{e}^{\jmath\omega} A_{\mathsf{lqr}}^\intercal)^{-1} A_{\mathsf{lqr}}^\intercal P B_w L_\star(\mathrm{e}^{\jmath\omega})\,\mathrm{d}\omega. \tag{14.10}$$

As the spectral factorization $\mathcal{M}_\star = \mathcal{L}_\star\mathcal{L}_\star^*$ is unique up to a constant unitary transformation on the right, *i.e.*, any alternative spectral factor takes the form

$L_\star(\mathrm{e}^{\jmath\omega}) \mapsto L_\star(\mathrm{e}^{\jmath\omega})U$ for a constant unitary matrix U, the corresponding optimal parameter $\Gamma_\star$ is likewise unique up to the same unitary transformation on the left, *i.e.*, $\Gamma_\star \mapsto \Gamma_\star U$.

Since the optimal solution necessarily admits the aforementioned parametric form, it becomes viable to reformulate the dual control problem by restricting the dual decision variable $\mathcal{M} \in \mathscr{L}_1^+$ to the subclass of operators whose canonical spectral factor $L(\mathrm{e}^{\jmath\omega})$ conforms to this structure. Specifically, we consider the reformulated problem:

$$\max_{\substack{\mathcal{M}=\mathcal{L}\mathcal{L}^*\in\mathscr{L}_1^+,\\ \mathcal{M}\succcurlyeq\mathcal{I},\\ \Gamma\in\mathbb{R}^{d_x\times d_w}}} V(\mathcal{M}) - \gamma\,\mathrm{tr}\left((\sqrt{\mathcal{M}}-\mathcal{I})^2\right),$$

$$\text{subject to,} \tag{14.11}$$

$$L(\mathrm{e}^{\jmath\omega})^* L(\mathrm{e}^{\jmath\omega}) = \left(\frac{I+\sqrt{I+4\gamma^{-1}\Gamma^\mathsf{T}\Psi(\mathrm{e}^{\jmath\omega})^*\Psi(\mathrm{e}^{\jmath\omega})\Gamma}}{2}\right)^2, \ \forall\omega\in[-\pi,\pi)$$

where the strictly anti-causal transfer matrix $\Psi(\mathrm{e}^{\jmath\omega}) := \overline{C}(\mathrm{e}^{-\jmath\omega}I - A_{\mathsf{lqr}}^\mathsf{T})^{-1}$ determined entirely by the plant's state-space matrices. Crucially, this reformulation renders the infinite-dimensional problem effectively finite-dimensional, as the dual variable $\mathcal{M} \in \mathscr{L}_1^+$ is fully characterized, up to a constant unitary transformation, by the finite-dimensional parameter $\Gamma \in \mathbb{R}^{d_x\times d_w}$. What remains to be determined is whether this equivalent finite-dimensional reformulation can be solved efficiently.

14.2 Algorithms for Dual Problems in Constrained Form

14.3 Algorithms for General Unconstrained Setting

Chapter 15

FINITE-DIMENSIONAL REALIZATION

The preceding section determined that the optimal solution, denoted as $\mathcal{N}_\star$, is non-rational and lacks a state-space representation. Nevertheless, Algorithm 4 introduced in Section 9.4 can effectively approximate it in the frequency domain. Indeed, after convergence, the algorithm returns the optimal finite parameter, $\Gamma_\star$, which can be used to compute $N_\star(z)$ at *any arbitrary frequency* using Theorem 13.4.2, and thus $K_\star(z)$ (see Algorithm 4 in Section 9.E). However, a state-space controller must be devised for any practical real-time implementation.

This section introduces an efficient method to obtain state-space controllers approximating the non-rational optimal controller. Instead of directly approximating the controller itself, our method involves an initial step of *approximating the power spectrum* $N_\star(z)$ *of the worst-case disturbance* to minimize the $\mathscr{H}_\infty$-norm of the approximation error using positive rational functions. While problems involving rational function approximation generally do not admit a convex formulation, we show in Theorem 15.2.4 that approximating positive power spectra by a ratio of positive fixed order polynomials can be cast as a convex feasibility problem. After finding a rational approximation of $N_\star(z)$, we compute a state-space controller according to (9.16a). For the sake of simplicity, we focus on scalar disturbances, *i.e.*, $d_w = 1$.

15.1 State-Space Models from Rational Power Spectra

As established in Theorem 9.3.2, the derivation of a optimal controller $K_\star$ is achieved through the positive operator $\mathcal{N}_\star = \mathcal{L}_\star^* \mathcal{L}_\star$ using the Wiener-Hopf technique. Specifically, we have $\mathcal{K}_\star = \mathcal{K}_{\mathscr{H}_2} + \Delta^{-1} \left\{ \{\Delta \mathcal{K}_\circ\}_- \mathcal{L}_\star \right\}_+ \mathcal{L}_\star^{-1} \mathcal{L}_\star^{-1}$. Since other controllers of interest, including $\mathscr{H}_2$, $\mathscr{H}_\infty$, and RO, can all be formulated this way, we focus on obtaining approximations to positive power spectra.

It is worth noting that a positive and symmetric rational approximation $\widehat{N}(z)$ of order $m \in \mathbb{N}$ can be represented as a ratio $\widehat{N}(z) = P(z)/Q(z)$ of two positive symmetric polynomials $P(z) = p_0 + \sum_{k=1}^m p_k (z^k + z^{-k})$, and $Q(z) = q_0 + \sum_{k=1}^m q_k (z^k + z^{-k})$. When such $P(z), Q(z)$ exist, we can obtain a rational spectral factorization of $\widehat{N}(z)$ by obtaining spectral factorization for $P(z)$, and $Q(z)$.

Finally, we end this section by stating an exact characterization of positive trig. polynomials. While verifying the positivity condition for general functions might pose challenges, the convex cone of positive symmetric trigonometric polynomials, $\mathscr{T}_{m,+}$, possess a characterization through a linear matrix inequality (LMI), as outlined below:

Lemma 15.1.1 (Trace parametrization of $\mathscr{T}_{m,+}$ [55, Thm. 2.3])**.** *For $k = [-m, m]$, let $\mathbf{\Theta}_k \in \mathbb{R}^{(m+1)\times(m+1)}$ be the primitive Toeplitz matrix with ones on the k^{th} diagonal and zeros everywhere else. Then, $P(z) = p_0 + \sum_{k=1}^{m} p_k(z^k + z^{-k}) > 0$ if and only if there exists a real positive definite matrix $\mathbf{P} \in \mathbb{S}_+^{m+1}$ such that*

$$p_k = \operatorname{tr}(\mathbf{P}\mathbf{\Theta}_k), \;\; k = 0, \ldots, m. \tag{15.1}$$

According to Lemma 15.1.1, any positive trig. polynomial of order at most m can be expressed (non-uniquely) as $P(z) = \sum_{k=-r}^{r} \operatorname{tr}(\mathbf{P}\mathbf{\Theta}_k) z^{-1} = \operatorname{tr}(\mathbf{P}\mathbf{\Theta}(z))$. Here, $\mathbf{\Theta}(z) \coloneqq \sum_{k=-r}^{r} \mathbf{\Theta}_k z^{-1}$.

15.2 Rational Approximation using $\mathscr{H}_\infty$-norm

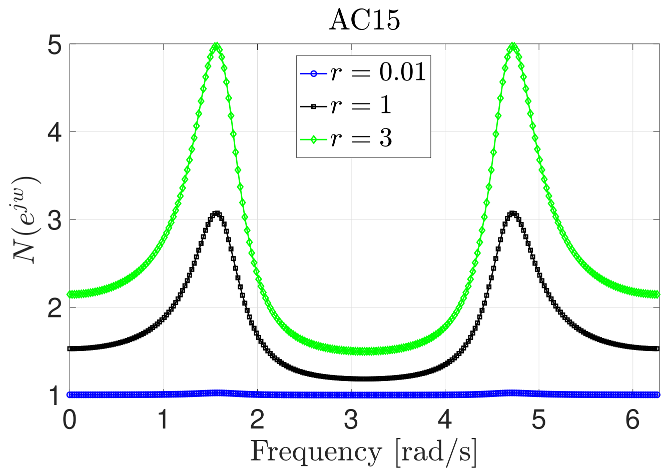
In this context, we present a novel and efficient approach for deriving rational approximations of non-rational power spectra. Our method bears similarities to the flexible uniform rational approximation approach described in [207], which approximates a function with a rational form while imposing the positivity of the denominator of the rational form as a constraint. Our method uses $\mathscr{H}_\infty$-norm as criteria to address the approximation error effectively. First, consider the following problem:

Problem 15.2.1 (Rational approximation via H_∞-norm minimization)**.** Given a positive spectrum $\mathcal{N}$, find the best rational approximation of order at most $m \in \mathbb{N}$ with respect to H_∞ norm, *i.e.*,

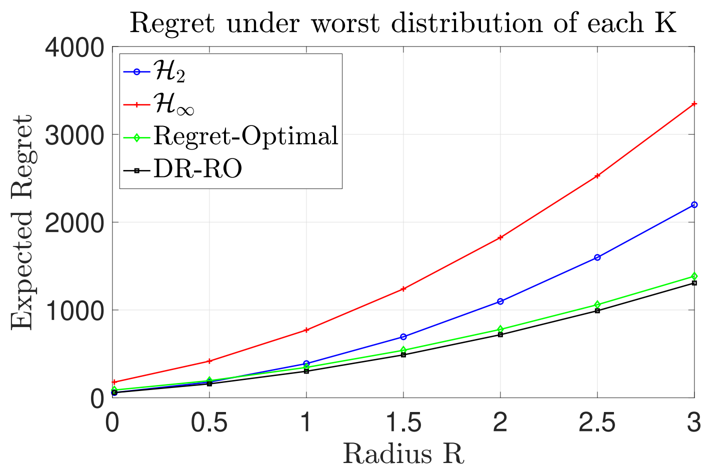
$$\inf_{\mathcal{P},\mathcal{Q} \in \mathscr{T}_{m,+}} \|\mathcal{P}/\mathcal{Q} - \mathcal{N}\|_\infty \text{ s.t. } \operatorname{tr}(\mathcal{Q}) = 1 \tag{15.2}$$

Note that the constraint $\operatorname{tr}(\mathcal{Q}) = 1$, equivalent to $q_0 = 1$, eliminates redundancy in the problem since the fraction $\mathcal{P}/\mathcal{Q}$ is scale invariant.

While the objective function in Equation (15.2) is convex with respect to $\mathcal{P}$ and $\mathcal{Q}$ individually, *it is not jointly convex in* $(\mathcal{P}, \mathcal{Q})$. In this form, Problem 15.2.1 is not amenable to standard convex optimization tools.



(a) The frequency domain representation of $\mathcal{N}$ for $r = 0.01, 1, 3$ for system [AC15].



(b) The worst-case expected regret of different controllers for the system [AC15].

Figure 15.1: Variation of $\mathcal{N}$ with r and the performance of the W$_2$-DR-RO controller versus the H_2, H_∞, and RO controller.

To circumvent this issue, we instead consider the sublevel sets of the objective function in Equation (15.2).

Definition 15.2.2. For a given $\epsilon > 0$ approximation bound, the ϵ-sublevel set of Problem 15.2.1 is defined as

$$\mathscr{S}_\epsilon := \{(\mathcal{P}, \mathcal{Q}) \mid \|\mathcal{P}/\mathcal{Q} - \mathcal{N}\|_\infty \leq \epsilon,\ \operatorname{tr}(\mathcal{Q}) = 1\}.$$

By applying the definition of $\mathscr{H}_\infty$-norm, we have that

$$\begin{aligned} \|\mathcal{P}/\mathcal{Q} - \mathcal{N}\|_\infty &= \max_{z \in \mathbb{T}} |P(z)/Q(z) - N(z)| \leq \epsilon \\ \iff &\begin{cases} P(z) - (N(z) + \epsilon)\, Q(z) \leq 0, \\ P(z) - (N(z) - \epsilon)\, Q(z) \geq 0, \end{cases} \end{aligned} \tag{15.3}$$

where the last set of inequalities hold for all $z \in \mathbb{T}$. Notice that the inequalities in Equation (15.3) and the positivity constraints on $\mathcal{P}$, $\mathcal{Q}$ are jointly affine in $(\mathcal{P}, \mathcal{Q})$. Moreover, the equation $\mathrm{tr}(\mathcal{Q}) = 1$ is an affine equality constraint. Therefore, we have the following claim.

Lemma 15.2.3. *The set $\mathscr{S}_\epsilon$ is jointly convex in $(\mathcal{P}, \mathcal{Q})$.*

Unlike its non-convex optimization counterpart in Problem 15.2.1, a membership oracle for the convex set $\mathscr{S}_\epsilon$ offers a means to obtain accurate rational approximations for non-rational functions. According to Lemma 15.1.1, the positive trig. polynomials $(\mathcal{P}, \mathcal{Q}) \in \mathscr{S}_\epsilon$ can be parameterized by psd matrices $\mathbf{P}$ and $\mathbf{Q}$. This allows the equality constraint $\mathrm{tr}(\mathcal{Q})$ and the affine inequalities in (15.3) to be expressed as Linear Matrix Inequalities (LMIs) in terms of $\mathbf{P}$ and $\mathbf{Q}$. The resulting theorem characterizes the ϵ-sublevel sets.

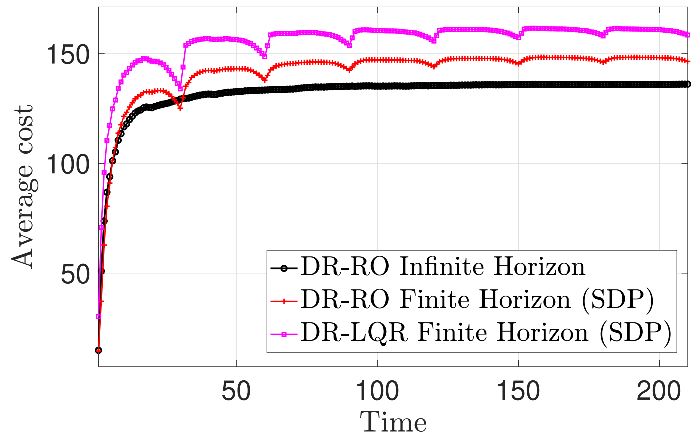
Theorem 15.2.4 (Feasibility of $\mathscr{S}_\epsilon$)**.** *Let $\epsilon > 0$ be a given accuracy level, and $m \in \mathbb{N}$ is a fixed order. The trig. polynomials $\mathcal{P}$ and $\mathcal{Q}$ of order m belong to the ϵ-sublevel set, $(\mathcal{P}, \mathcal{Q}) \in \mathscr{S}_\epsilon$ if and only if there exists $\mathbf{P}, \mathbf{Q} \in \mathbb{S}_+^{m+1}$ such that $\mathrm{tr}\,(\mathbf{Q}) = 1$ and for all $z \in \mathbb{T}$,*

$$1)\ \mathrm{tr}\,(\mathbf{P}\boldsymbol{\Theta}(z)) - (N(z)+\epsilon)\,\mathrm{tr}\,(\mathbf{Q}\boldsymbol{\Theta}(z)) \leq 0, \tag{15.4}$$

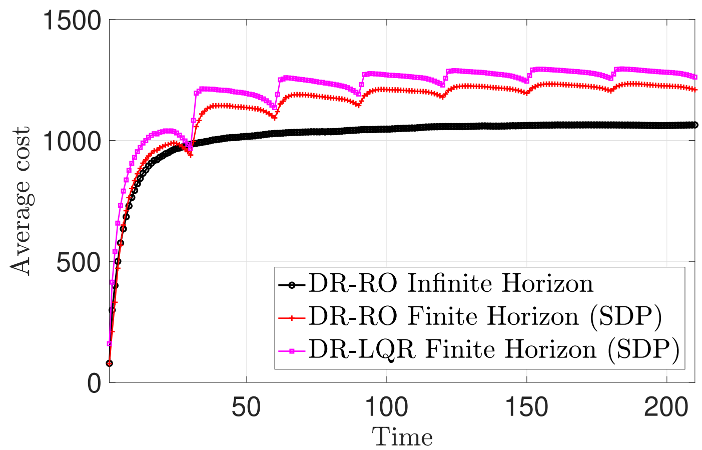
$$2)\ \mathrm{tr}\,(\mathbf{P}\boldsymbol{\Theta}(z)) - (N(z)-\epsilon)\,\mathrm{tr}\,(\mathbf{Q}\boldsymbol{\Theta}(z)) \geq 0. \tag{15.5}$$

The sole limitation in this approach arises from the fact that for a non-rational $N(z)$, the set of infinitely many inequalities in (15.3) cannot be precisely characterized by a finite number of constraints, as seen in the trace parametrization of positive polynomials. To overcome this challenge, one can address the inequalities in (15.3) solely for a finite set of frequencies, such as $\mathbb{T}_N = \{\mathrm{e}^{j2\pi n/N} \mid n = 0, \ldots, N-1\}$ for $N \gg m$. While this introduces an approximation, the method's accuracy can be enhanced arbitrarily by increasing the frequency samples. By taking this approach, the problem of rational function approximation can be reformulated as a convex feasibility problem involving LMIs and a finite number of affine (in)equality constraints.

It is crucial to note that our algorithm can be used in the following two operational modes. These modes highlight the algorithm's adaptability for the given two use cases.



(a) White noise

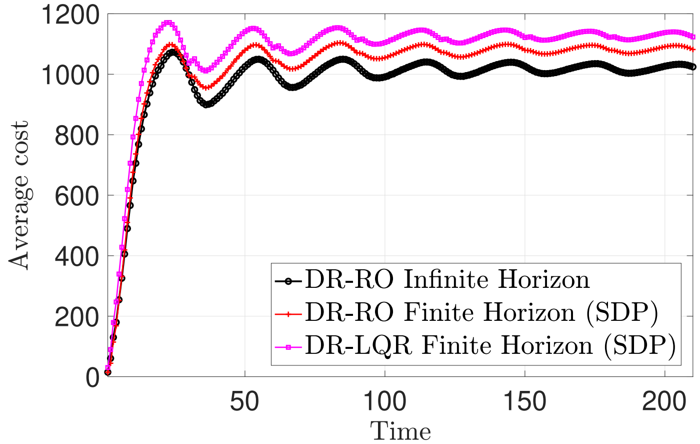


(b) Worst disturbance for W_2-DR-RO, infinite horizon

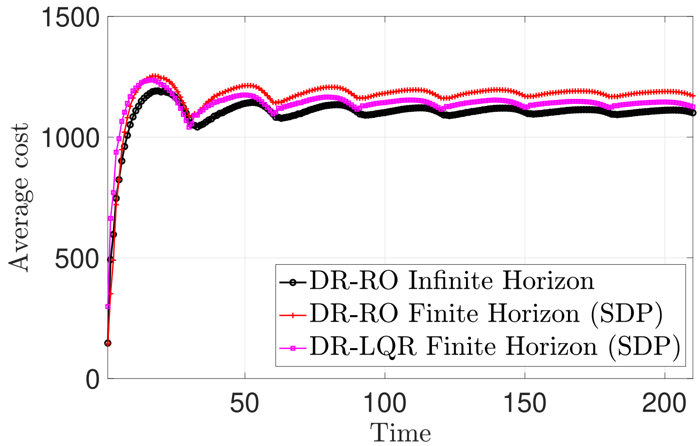
Figure 15.2: The control costs of different DR controllers under (a) white noise and (b) worst disturbance for W_2-DR-RO in infinite horizon, for system [AC15]. The finite-horizon controllers are re-applied every $s = 30$ steps. The infinite horizon W_2-DR-RO controller achieves the lowest average cost compared to the finite-horizon controllers.

1. **Best Precision for a given degree** By adjusting the parameter ϵ, which signifies our tolerance for deviations from $M(e^{jw})$, we can refine the approximation's accuracy. This method is particularly valuable when finding the best possible polynomial representation of $M(e^{jw})$ for a given degree.

2. **Lowest Degree for a given precision** In contrast, we can ask for the lowest degree polynomial, which achieves a certain precision level ϵ. This mode is advantageous when the priority is to minimize computational overhead or when we need a simpler polynomial approximation, as long as the approximation



(a) Worst disturbance for W_2-DR-RO, finite horizon



(b) Worst disturbance for DR-LQR, finite horizon

Figure 15.3: The control costs of different DR controllers under (a) worst disturbances for W_2-DR-RO in finite horizon and (b) worst disturbances for DR-LQR in finite horizon, for system [AC15]. The finite-horizon controllers are re-applied every $s = 30$ steps. Despite being designed to minimize the cost under specific disturbances, the finite horizon DR controllers are outperformed by the infinite horizon W_2-DR-RO controller.

remains within acceptable accuracy bounds

15.3 Obtaining State-Space Controllers

Note that given the polynomial z-spectra, we require its spectral factorization to obtain the state-space controller that approximates the W_2-DR-RO controller. The following Lemma introduces a simple way to obtain such an approximation

Lemma 15.3.1 (Canonical factor of polynomial z-spectra [199, Lem. 1])**.** *Consider* *a Laurent polynomial of degree m, $P(z) = \sum_{k=-m}^{m} p_k z^{-k}$, with $p_k = p_{-k} \in \mathbb{R}$, such that $P(z) > 0$. Then, there exists a canonical factor $L(z) = \sum_{k=0}^{m} \ell_k z^{-k}$ such that $P(z) = |L(z)|^2$ and $L(z)$ has all of its root in $\mathbb{T}$.*

Using Lemma 15.3.1, we can compute spectral factors by factorizing the symmetric positive polynomials and multiplying all the factors with stable roots together. Consequently, this rational spectral factor enables the derivation of a rational controller, denoted as $K(z)$ (refer to Section 15.3).

Now we present the W_2-DR-RO controller in state-space form.

Lemma 15.3.2. *Let $\tilde{L}(z)$ be the rational factor of the spectral factorization $\tilde{N}(z) = \tilde{L}(z)^* \tilde{L}(z) = P(z)/Q(z)$ of a degree m rational approximation $P(z)/Q(z)$. The controller obtained from $\tilde{L}(z)$ using* (9.16), *i.e., $K(z) = K_{\mathscr{H}_2}(z) + \Delta(z)^{-1} \left\{ \{\Delta(z) K_\circ(z)\}_- \tilde{L}(z) \right\}_+ \tilde{L}(z)^{-1}$ is rational and can be realized as a state-space controller as follows:*

$$\begin{aligned} e(t+1) &= \widetilde{F} e(t) + \widetilde{G} w(t), \\ u(t) &= \widetilde{H} e(t) + \widetilde{J} w(t)) \end{aligned} \tag{15.6}$$

where e_t is the controller state, and $(\widetilde{F}, \widetilde{G}, \widetilde{H}, \widetilde{J})$ are determined from (A, B_u, B_w) and $\tilde{L}(z)$.

BIBLIOGRAPHY

[1] Y. Abbasi-Yadkori, D. Pál, and C. Szepesvári, "Improved algorithms for linear stochastic bandits," in *Advances in Neural Information Processing Systems*, 2011, pp. 2312–2320.

[2] 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*, 2011, pp. 1–26.

[3] M. Abeille and A. Lazaric, "Thompson sampling for linear-quadratic control problems," *arXiv preprint arXiv:1703.08972*, 2017.

[4] M. Abeille and A. Lazaric, "Improved regret bounds for thompson sampling in linear quadratic control problems," in *International Conference on Machine Learning*, 2018, pp. 1–9.

[5] M. Abeille and A. Lazaric, "Efficient optimistic exploration in linear-quadratic regulators via lagrangian relaxation," *arXiv preprint arXiv:2007.06482*, 2020.

[6] S. Agrawal, "Recent advances in multiarmed bandits for sequential decision making," *Operations Research & Management Science in the Age of Analytics*, pp. 167–188, 2019.

[7] L. Aolaritei, M. Fochesato, J. Lygeros, and F. Dörfler, "Wasserstein tube mpc with exact uncertainty propagation," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, 2023, pp. 2036–2041. DOI: `10.1109/CDC49753.2023.10383526`.

[8] L. Aolaritei, N. Lanzetti, and F. Dörfler, "Capture, propagate, and control distributional uncertainty," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, 2023, pp. 3081–3086. DOI: `10.1109/CDC49753.2023.10383424`.

[9] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proceedings of the 34th International Conference on Machine Learning*, D. Precup and Y. W. Teh, Eds., ser. Proceedings of Machine Learning Research, vol. 70, PMLR, Jun. 2017, pp. 214–223. [Online]. Available: `https://proceedings.mlr.press/v70/arjovsky17a.html`.

[10] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA '07, New Orleans, Louisiana: Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035, ISBN: 9780898716245.

[11] Arthur E. Bryson, "Kalman Filter Divergence and Aircraft Motion Estimators," en, *Journal of Guidance and Control*, vol. 1, no. 1, pp. 71–79, Jan. 1978, ISSN: 0162-3192. DOI: 10.2514/3.55745. [Online]. Available: https://arc.aiaa.org/doi/10.2514/3.55745 (visited on 05/18/2024).

[12] K. Azizzadenesheli, E. Brunskill, and A. Anandkumar, "Efficient exploration through bayesian deep q-networks," in *2018 Information Theory and Applications Workshop (ITA)*, 2018, pp. 1–9. DOI: 10.1109/ITA.2018.8503252.

[13] D. Baron and T. Weissman, "An MCMC approach to lossy compression of continuous sources," in *2010 Data Compression Conference*, IEEE, 2010, pp. 40–48.

[14] T. Başar and P. Bernhard, *H_∞-Optimal Control and Related Minimax Design Problems*. Boston, MA: Birkhäuser Boston, 2008, ISBN: 978-0-8176-4756-8 978-0-8176-4757-5. DOI: 10.1007/978-0-8176-4757-5. [Online]. Available: http://link.springer.com/10.1007/978-0-8176-4757-5 (visited on 05/19/2024).

[15] E. Baziw and I. Weir-Jones, "Application of Kalman Filtering Techniques for Microseismic Event Detection," en, in *The Mechanism of Induced Seismicity*, C. I. Trifu, Ed., Basel: Birkhäuser Basel, 2002, pp. 449–471, ISBN: 978-3-7643-6653-7 978-3-0348-8179-1. DOI: 10.1007/978-3-0348-8179-1_19. [Online]. Available: http://link.springer.com/10.1007/978-3-0348-8179-1_19 (visited on 05/18/2024).

[16] R. BELLMAN, "A markovian decision process," *Journal of Mathematics and Mechanics*, vol. 6, no. 5, pp. 679–684, 1957, ISSN: 00959057, 19435274. [Online]. Available: http://www.jstor.org/stable/24900506 (visited on 05/12/2025).

[17] R. E. Bellman, *Dynamic Programming and Multistage Decision Processes of Stochastic Type*. Santa Monica, CA: RAND Corporation, 1954.

[18] D. Bernstein and W. Haddad, "Lqg control with an h/sup infinity / performance bound: A riccati equation approach," *IEEE Transactions on Automatic Control*, vol. 34, no. 3, pp. 293–305, 1989. DOI: 10.1109/9.16419.

[19] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena scientific Belmont, MA, 1995, vol. 2.

[20] R. Bhatia, T. Jain, and Y. Lim, "On the bures–wasserstein distance between positive definite matrices," *Expositiones Mathematicae*, vol. 37, no. 2, pp. 165–191, 2019, ISSN: 0723-0869. DOI: https://doi.org/10.1016/j.exmath.2018.01.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0723086918300021.

[21] S. Bittanti, M. C. Campi, *et al.*, "Adaptive control of linear time invariant systems: The "bet on the best" principle," *Communications in Information & Systems*, vol. 6, no. 4, pp. 299–320, 2006.

[22] D. Blackwell, L. Breiman, and A. Thomasian, "The capacity of a class of channels," *The Annals of Mathematical Statistics*, pp. 1229–1241, 1959.

[23] D. Blackwell, L. Breiman, and A. Thomasian, "The capacities of certain channel classes under random coding," *The Annals of Mathematical Statistics*, vol. 31, no. 3, pp. 558–567, 1960.

[24] J. Blanchet, D. Kuhn, J. Li, and B. Taskesen, *Unifying Distributionally Robust Optimization via Optimal Transport Theory*, arXiv:2308.05414 [math, stat], Aug. 2023. [Online]. Available: http://arxiv.org/abs/2308.05414 (visited on 08/15/2023).

[25] J. Blanchet, J. Li, S. Lin, and X. Zhang, *Distributionally Robust Optimization and Robust Statistics*, arXiv:2401.14655 [stat], Jan. 2024. [Online]. Available: http://arxiv.org/abs/2401.14655 (visited on 05/11/2024).

[26] Y. Blau and T. Michaeli, "Rethinking lossy compression: The rate-distortion-perception tradeoff," in *International Conference on Machine Learning*, PMLR, 2019, pp. 675–685.

[27] N. Blits and M. Feder, "Universal rateless coding with finite message set," in *2012 IEEE International Symposium on Information Theory Proceedings*, IEEE, 2012, pp. 1772–1776.

[28] R. Boel, M. James, and I. Petersen, "Robustness and risk-sensitive filtering," in *Proceedings of the 36th IEEE Conference on Decision and Control*, vol. 3, San Diego, CA, USA: IEEE, 1997, pp. 2273–2278, ISBN: 978-0-7803-4187-6. DOI: 10.1109/CDC.1997.657112. [Online]. Available: http://ieeexplore.ieee.org/document/657112/ (visited on 05/17/2024).

[29] J.-S. Brouillon, F. Dörfler, and G. Ferrari-Trecate, "Regularization for Distributionally Robust State Estimation and Prediction," *IEEE Control Systems Letters*, pp. 1–1, 2023, ISSN: 2475-1456. DOI: 10.1109/LCSYS.2023.3287803. [Online]. Available: https://ieeexplore.ieee.org/document/10158348/ (visited on 05/11/2024).

[30] J.-S. Brouillon, A. Martin, J. Lygeros, F. Dörfler, and G. F. Trecate, *Distributionally robust infinite-horizon control: From a pool of samples to the design of dependable controllers*, 2023. arXiv: 2312.07324 [math.OC].

[31] M. C. Campi and P. Kumar, "Adaptive linear quadratic gaussian control: The cost-biased approach revisited," *SIAM Journal on Control and Optimization*, vol. 36, no. 6, pp. 1890–1907, 1998.

[32] A. Cassel, A. Cohen, and T. Koren, "Logarithmic regret for learning linear quadratic regulators efficiently," *arXiv preprint arXiv:2002.08095*, 2020.

[33] Y. Chan, A. Hu, and J. Plant, "A Kalman Filter Based Tracking Scheme with Input Estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-15, no. 2, pp. 237–244, Mar. 1979, ISSN: 0018-9251. DOI: 10.1109/TAES.1979.308710. [Online]. Available: http://ieeexplore.ieee.org/document/4102140/ (visited on 05/18/2024).

[34] O. Chapelle and L. Li, "An empirical evaluation of thompson sampling," in *Advances in neural information processing systems*, 2011, pp. 2249–2257.

[35] S. Y. Chen, "Kalman Filter for Robot Vision: A Survey," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 11, pp. 4409–4420, Nov. 2012, ISSN: 0278-0046, 1557-9948. DOI: 10.1109/TIE.2011.2162714. [Online]. Available: http://ieeexplore.ieee.org/document/5985520/ (visited on 05/18/2024).

[36] X. Chen and E. Hazan, "Black-box control for linear dynamical systems," 2020.

[37] A. Coates and A. Y. Ng, "Learning feature representations with k-means," in *Neural Networks: Tricks of the Trade: Second Edition*, Springer, 2012, pp. 561–580.

[38] A. Coates and A. Y. Ng, "Learning Feature Representations with K-Means," in *Neural Networks: Tricks of the Trade*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds., vol. 7700, Series Title: Lecture Notes in Computer Science, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 561–580, ISBN: 978-3-642-35288-1 978-3-642-35289-8. DOI: 10.1007/978-3-642-35289-8_30. [Online]. Available: http://link.springer.com/10.1007/978-3-642-35289-8_30 (visited on 02/06/2025).

[39] A. Cohen, A. Hassidim, T. Koren, N. Lazic, Y. Mansour, and K. Talwar, "Online linear quadratic control," *arXiv preprint arXiv:1806.07104*, 2018.

[40] A. Cohen, T. Koren, and Y. Mansour, "Learning linear-quadratic regulators efficiently with only $\sqrt{T}$ regret," *arXiv preprint arXiv:1902.06223*, 2019.

[41] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, en, 1st ed. Wiley, Sep. 2005, ISBN: 978-0-471-24195-9 978-0-471-74882-3. DOI: 10.1002/047174882X. [Online]. Available: https://onlinelibrary.wiley.com/doi/book/10.1002/047174882X (visited on 05/05/2025).

[42] I. Csiszar, "Arbitrarily varying channels with general alphabets and states," *IEEE Transactions on Information Theory*, vol. 38, no. 6, pp. 1725–1742, 1992.

[43] J. M. Danskin, "The theory of max-min, with applications," *SIAM Journal on Applied Mathematics*, vol. 14, no. 4, pp. 641–664, 1966, ISSN: 00361399. [Online]. Available: http://www.jstor.org/stable/2946123 (visited on 12/11/2023).

[44] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "Regret bounds for robust adaptive control of the linear quadratic regulator," in *Advances in Neural Information Processing Systems*, 2018, pp. 4188–4197.

[45] A. Deshpande, P. Kacham, and R. Pratap, "Robust k-means++," in *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, J. Peters and D. Sontag, Eds., ser. Proceedings of Machine Learning Research, vol. 124, PMLR, Aug. 2020, pp. 799–808. [Online]. Available: `https://proceedings.mlr.press/v124/deshpande20a.html`.

[46] A. Deshpande and R. Pratap, *Improved Outlier Robust Seeding for k-means*, arXiv:2309.02710 [cs], Sep. 2023. DOI: `10.48550/arXiv.2309.02710`. [Online]. Available: `http://arxiv.org/abs/2309.02710` (visited on 02/06/2025).

[47] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using k-means clustering algorithm and subtractive clustering algorithm," *Procedia Computer Science*, vol. 54, pp. 764–771, 2015.

[48] A. Didier, J. Sieber, and M. N. Zeilinger, "A system-level approach to regret optimal control," *IEEE Control Systems Letters*, vol. 6, pp. 2792–2797, 2022.

[49] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 3072–3081, Nov. 2001.

[50] O. Dorabiala, J. N. Kutz, and A. Y. Aravkin, "Robust trimmed k -means," en, *Pattern Recognition Letters*, vol. 161, pp. 9–16, Sep. 2022, ISSN: 01678655. DOI: `10.1016/j.patrec.2022.07.007`. [Online]. Available: `https://linkinghub.elsevier.com/retrieve/pii/S0167865522002185` (visited on 02/06/2025).

[51] J. Doyle, K. Glover, P. Khargonekar, and B. Francis, "State-space solutions to standard H_2 and H_∞ control problems," in *1988 American Control Conference*, Atlanta, GA, USA: IEEE, Jun. 1988, pp. 1691–1696. DOI: `10.23919/ACC.1988.4789992`. [Online]. Available: `https://ieeexplore.ieee.org/document/4789992/` (visited on 12/11/2023).

[52] J. Doyle, K. Zhou, and B. Bodenheimer, "Optimal control with mixed H_2 and H_∞ performance objectives," in *1989 American Control Conference*, IEEE, 1989, pp. 2065–2070.

[53] J. C. Doyle, "Structured uncertainty in control system design," in *1985 24th IEEE Conference on Decision and Control*, 1985, pp. 260–265. DOI: `10.1109/CDC.1985.268842`.

[54] S. C. Draper, B. J. Frey, and F. R. Kschischang, "Efficient variable length channel coding for unknown dmcs," in *International Symposium onInformation Theory, 2004. ISIT 2004. Proceedings.*, IEEE, 2004, pp. 379–379.

[55] B. Dumitrescu, *Positive Trigonometric Polynomials and Signal Processing Applications* (Signals and Communication Technology). Cham: Springer International Publishing, 2017, ISBN: 978-3-319-53687-3 978-3-319-53688-0. DOI: `10.1007/978-3-319-53688-0`. [Online]. Available: `http://link.springer.com/10.1007/978-3-319-53688-0` (visited on 01/15/2024).

[56] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, en. Society for Industrial and Applied Mathematics, Jan. 1999, ISBN: 978-0-89871-450-0 978-1-61197-108-8. DOI: `10.1137/1.9781611971088`. [Online]. Available: `http://epubs.siam.org/doi/book/10.1137/1.9781611971088` (visited on 03/08/2021).

[57] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge university press, 2011.

[58] L. Ephremidze, *An Elementary Proof of the Polynomial Matrix Spectral Factorization Theorem*, arXiv:1011.3777 [math], Nov. 2010. [Online]. Available: `http://arxiv.org/abs/1011.3777` (visited on 11/15/2023).

[59] L. Ephremidze, F. Saied, and I. M. Spitkovsky, “On the Algorithmization of Janashia-Lagvilava Matrix Spectral Factorization Method,” *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 728–737, Feb. 2018, ISSN: 0018-9448, 1557-9654. DOI: `10.1109/TIT.2017.2772877`. [Online]. Available: `http://ieeexplore.ieee.org/document/8105834/` (visited on 11/15/2023).

[60] L. Falconi, A. Ferrante, and M. Zorzi, “A new perspective on robust performance for LQG control problems,” in *2022 IEEE 61st Conference on Decision and Control (CDC)*, Cancun, Mexico: IEEE, Dec. 2022, pp. 3003–3008, ISBN: 978-1-66546-761-2. DOI: `10.1109/CDC51059.2022.9992567`. [Online]. Available: `https://ieeexplore.ieee.org/document/9992567/` (visited on 06/10/2024).

[61] M. K. S. Faradonbeh, M. S. S. Faradonbeh, and M. Bayati, “Thompson sampling efficiently learns to control diffusion processes,” in *Proceedings of the 36th International Conference on Neural Information Processing Systems*, ser. NIPS ’22, New Orleans, LA, USA: Curran Associates Inc., 2022, ISBN: 9781713871088.

[62] M. K. S. Faradonbeh, A. Tewari, and G. Michailidis, “Optimism-based adaptive regulation of linear-quadratic systems,” *arXiv preprint arXiv:1711.07230*, 2017.

[63] M. K. S. Faradonbeh, A. Tewari, and G. Michailidis, “On adaptive linear–quadratic regulators,” *Automatica*, vol. 117, p. 108 982, 2020.

[64] 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.

[65] W. Farag, “Kalman-filter-based sensor fusion applied to road-objects detection and tracking for autonomous vehicles,” en, *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 235, no. 7, pp. 1125–1138, Aug. 2021, ISSN: 0959-6518, 2041-3041. DOI: 10.1177/0959651820975523. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0959651820975523 (visited on 05/18/2024).

[66] N. Fournier and A. Guillin, “On the rate of convergence in Wasserstein distance of the empirical measure,” en, *Probability Theory and Related Fields*, vol. 162, no. 3-4, pp. 707–738, Aug. 2015, ISSN: 0178-8051, 1432-2064. DOI: 10.1007/s00440-014-0583-7. [Online]. Available: https://link.springer.com/10.1007/s00440-014-0583-7 (visited on 02/06/2025).

[67] B. A. Francis, Ed., *A Course in H_∞ Control Theory* (Lecture Notes in Control and Information Sciences), en. Berlin/Heidelberg: Springer-Verlag, 1987, vol. 88, ISBN: 978-3-540-17069-3. DOI: 10.1007/BFb0007371. [Online]. Available: http://link.springer.com/10.1007/BFb0007371 (visited on 03/04/2025).

[68] M. Frank and P. Wolfe, “An algorithm for quadratic programming,” en, *Naval Research Logistics Quarterly*, vol. 3, no. 1-2, pp. 95–110, Mar. 1956, ISSN: 0028-1441, 1931-9193. DOI: 10.1002/nav.3800030109. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/nav.3800030109 (visited on 05/11/2024).

[69] B. Friedland, *Control system design: an introduction to state-space methods*. Courier Corporation, 2012.

[70] M. Fu, C. E. De Souza, and L. Xie, “H_∞ estimation for uncertain systems,” en, *International Journal of Robust and Nonlinear Control*, vol. 2, no. 2, pp. 87–105, Aug. 1992, ISSN: 1049-8923, 1099-1239. DOI: 10.1002/rnc.4590020202. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/rnc.4590020202 (visited on 05/17/2024).

[71] G. Galanis, P. Louka, P. Katsafados, I. Pytharoulis, and G. Kallos, “Applications of Kalman filters based on non-linear functions to numerical weather predictions,” en, *Annales Geophysicae*, vol. 24, no. 10, pp. 2451–2460, Oct. 2006, ISSN: 1432-0576. DOI: 10.5194/angeo-24-2451-2006. [Online]. Available: https://angeo.copernicus.org/articles/24/2451/2006/ (visited on 05/18/2024).

[72] A. Galka, O. Yamashita, T. Ozaki, R. Biscay, and P. Valdés-Sosa, “A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering,” en, *NeuroImage*, vol. 23, no. 2, pp. 435–453, Oct. 2004, ISSN: 10538119. DOI: 10.1016/j.neuroimage.2004.02.022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1053811904001144 (visited on 05/18/2024).

[73] A. Gamal, "The capacity of a class of broadcast channels," *IEEE Transactions on Information Theory*, vol. 25, no. 2, pp. 166–169, 1979.

[74] R. Gao and A. Kleywegt, "Distributionally robust stochastic optimization with wasserstein distance," *Mathematics of Operations Research*, vol. 48, no. 2, pp. 603–655, 2023. DOI: 10.1287/moor.2022.1275.

[75] L. A. García-Escudero, A. Gordaliza, C. Matrán, and A. Mayo-Iscar, "A review of robust clustering methods," en, *Advances in Data Analysis and Classification*, vol. 4, no. 2-3, pp. 89–109, Sep. 2010, ISSN: 1862-5347, 1862-5355. DOI: 10.1007/s11634-010-0064-5. [Online]. Available: http://link.springer.com/10.1007/s11634-010-0064-5 (visited on 02/06/2025).

[76] M. Garey, D. Johnson, and H. Witsenhausen, "The complexity of the generalized lloyd-max problem (corresp.)," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 255–256, Mar. 1982.

[77] A. Gelb and Analytic Sciences Corporation, Eds., *Applied optimal estimation*, eng, 19. printing. Cambridge, Mass.: M.I.T. Press, 2006, ISBN: 978-0-262-57048-0 978-0-262-20027-1.

[78] M. Gelbrich, "On a Formula for the L ^{2} Wasserstein Metric between Measures on Euclidean and Hilbert Spaces," en, *Mathematische Nachrichten*, vol. 147, no. 1, pp. 185–203, Jan. 1990, ISSN: 0025-584X, 1522-2616. DOI: 10.1002/mana.19901470121. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/mana.19901470121 (visited on 01/15/2024).

[79] A. Georgogiannis, "Robust k-means: A theoretical revisit," in *Proceedings of the 30th International Conference on Neural Information Processing Systems*, ser. NIPS'16, event-place: Barcelona, Spain, Red Hook, NY, USA: Curran Associates Inc., 2016, pp. 2891–2899, ISBN: 978-1-5108-3881-9.

[80] G. Goel and B. Hassibi, *Regret-optimal control in dynamic environments*, arXiv:2010.10473 [cs, eess, math], Feb. 2021. [Online]. Available: http://arxiv.org/abs/2010.10473 (visited on 12/10/2023).

[81] G. Goel and B. Hassibi, "Regret-optimal measurement-feedback control," in *Learning for Dynamics and Control*, PMLR, 2021, pp. 1270–1280.

[82] G. Goel and B. Hassibi, "Regret-optimal estimation and control," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 3041–3053, 2023.

[83] A. A. Goldstein, "Convex programming in Hilbert space," *Bulletin of the American Mathematical Society*, vol. 70, no. 5, pp. 709–710, 1964.

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

[85] M. Grewal, V. Henderson, and R. Miyasako, "Application of Kalman filtering to the calibration and alignment of inertial navigation systems," in *29th IEEE Conference on Decision and Control*, Honolulu, HI, USA: IEEE, 1990, 3325–3334 vol.6. DOI: 10.1109/CDC.1990.203410. [Online]. Available: http://ieeexplore.ieee.org/document/203410/ (visited on 05/18/2024).

[86] M. S. Grewal and A. P. Andrews, *Kalman filtering: theory and practice* (Prentice-Hall information and system sciences series), eng. Englewood Cliffs, N.J: Prentice-Hall, 1993, ISBN: 978-0-13-211335-9.

[87] M. J. Grimble, " H_∞ Design of Optimal Linear Filters," *Linear Circuit Systems and Signal Processing : Theory and Application (Proc. MTNS '87)*, pp. 533–540, 1987. [Online]. Available: https://cir.nii.ac.jp/crid/1573105974308524416.

[88] P. M. E. M. van der Grinten, "Uncertainty in measurement and control," *Statistica Neerlandica*, vol. 22, no. 1, pp. 43–63, 1968. DOI: https://doi.org/10.1111/j.1467-9574.1960.tb00617.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9574.1960.tb00617.x.

[89] J. Hajar, T. Kargin, and B. Hassibi, "Wasserstein distributionally robust regret-optimal control under partial observability," in *2023 59th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2023, pp. 1–6. DOI: 10.1109/Allerton58177.2023.10313386,

[90] J. Hajar, T. Kargin, V. Malik, and B. Hassibi, "The distributionally robust infinite-horizon lqr," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, 2024, pp. 6539–6546. DOI: 10.1109/CDC56724.2024.10886462,

[91] J. Hajar, O. Sabag, and B. Hassibi, *Regret-Optimal Control under Partial Observability*, arXiv:2311.06433 [cs, eess, math], Nov. 2023. [Online]. Available: http://arxiv.org/abs/2311.06433 (visited on 11/20/2023).

[92] J. Hajar*, T. Kargin*, V. Malik, and B. Hassibi, "Distributionally robust kalman filtering over an infinite-horizon," in *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025, pp. 1–5. DOI: 10.1109/ICASSP49660.2025.10889742,

[93] A. Hakobyan and I. Yang, "Wasserstein distributionally robust control of partially observable linear systems: Tractable approximation and performance guarantee," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, IEEE, 2022, pp. 4800–4807.

[94] B. Halder, B. Hassibi, and T. Kailath, "Design of optimal mixed h/sub 2//h/sub/spl infin//static state feedback controllers," in *Proceedings of the*

1998 American Control Conference. ACC (IEEE Cat. No. 98CH36207), IEEE, vol. 5, 1998, pp. 3239–3243.

[95] B. Han, *Distributionally robust Kalman filtering with volatility uncertainty*, arXiv:2302.05993 [math], Feb. 2023. [Online]. Available: `http://arxiv.org/abs/2302.05993` (visited on 05/11/2024).

[96] L. Hansen and T. Sargent, *Robustness*. Princeton University Press, 2008, ISBN: 978-0-691-11442-2. [Online]. Available: `https://books.google.com/books?id=7wwV27TvR8EC`.

[97] L. P. Hansen and T. J. Sargent, "Robust estimation and control under commitment," en, *Journal of Economic Theory*, vol. 124, no. 2, pp. 258–301, Oct. 2005, ISSN: 00220531. DOI: `10.1016/j.jet.2005.06.006`. [Online]. Available: `https://linkinghub.elsevier.com/retrieve/pii/S0022053105001675` (visited on 05/17/2024).

[98] J. C. Hargreaves, J. D. Annan, N. R. Edwards, and R. Marsh, "An efficient climate forecasting method using an intermediate complexity Earth System Model and the ensemble Kalman filter," en, *Climate Dynamics*, vol. 23, no. 7-8, pp. 745–760, Dec. 2004, ISSN: 0930-7575, 1432-0894. DOI: `10.1007/s00382-004-0471-4`. [Online]. Available: `http://link.springer.com/10.1007/s00382-004-0471-4` (visited on 05/18/2024).

[99] B. Hassibi, A. Sayed, and T. Kailath, "Linear estimation in Krein spaces. II. Applications," *IEEE Transactions on Automatic Control*, vol. 41, no. 1, pp. 34–49, Jan. 1996, ISSN: 00189286. DOI: `10.1109/9.481606`. [Online]. Available: `http://ieeexplore.ieee.org/document/481606/` (visited on 05/17/2024).

[100] B. Hassibi and T. Kailath, "Upper bounds for mixed h/sup 2//h/sup/spl infin//control," in *Proceedings of the 37th IEEE Conference on Decision and Control (Cat. No. 98CH36171)*, IEEE, vol. 1, 1998, pp. 652–657.

[101] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic Estimation and Control*. Society for Industrial and Applied Mathematics, 1999. DOI: `10.1137/1.9781611970760`. eprint: `https://epubs.siam.org/doi/pdf/10.1137/1.9781611970760`. [Online]. Available: `https://epubs.siam.org/doi/abs/10.1137/1.9781611970760`.

[102] H. Helson and D. Lowdenslager, "Prediction theory and Fourier Series in several variables," en, *Acta Mathematica*, vol. 99, no. 0, pp. 165–202, 1958, ISSN: 0001-5962. DOI: `10.1007/BF02392425`. [Online]. Available: `http://projecteuclid.org/euclid.acta/1485892263` (visited on 02/28/2025).

[103] H. Helson and D. Lowdenslager, "Prediction theory and fourier series in several variables. II," en, *Acta Mathematica*, vol. 106, no. 3-4, pp. 175–213, 1961, ISSN: 0001-5962. DOI: 10.1007/BF02545786. [Online]. Available: http://projecteuclid.org/euclid.acta/1485889232 (visited on 03/09/2025).

[104] H. A. Hindi, B. Hassibi, and S. P. Boyd, "Multiobjective h/sub 2//h/sub/spl infin//-optimal control via finite dimensional q-parametrization and linear matrix inequalities," in *Proceedings of the 1998 American Control Conference. ACC (IEEE Cat. No. 98CH36207)*, IEEE, vol. 5, 1998, pp. 3244–3249.

[105] B. Ho and R. E. Kálmán, "Effective construction of linear state-variable models from input/output functions," *at-Automatisierungstechnik*, vol. 14, no. 1-12, pp. 545–548, 1966.

[106] K. Honda, A. Notsu, and H. Ichihashi, "Fuzzy pca-guided robust k-means clustering," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 1, pp. 67–79, 2010. DOI: 10.1109/TFUZZ.2009.2036603.

[107] Honghui Qi and J. Moore, "Direct Kalman filtering approach for GPS/INS integration," en, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 2, pp. 687–693, Apr. 2002, ISSN: 0018-9251. DOI: 10.1109/TAES.2002.1008998. [Online]. Available: http://ieeexplore.ieee.org/document/1008998/ (visited on 05/18/2024).

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

[109] Z. Hu and L. J. Hong, *Kullback-Leibler divergence constrained distributionally robust optimization*, Nov. 2012. [Online]. Available: https://optimization-online.org/2012/11/3677/ (visited on 04/03/2024).

[110] IFR International Federation of Robotics, *Record of 4 Million Robots in Factories Worldwide — ifr.org*, Sep. 2024. [Online]. Available: https://ifr.org/ifr-press-releases/news/record-of-4-million-robots-working-in-factories-worldwide (visited on 05/11/2025).

[111] S. Ihara, "On the capacity of channels with additive non-Gaussian noise," *Information and Control*, vol. 37, no. 1, pp. 34–39, Apr. 1978.

[112] IoT Analytics, *Smart electricity meter market 2024: Global adoption landscape — iot-analytics.com*, Feb. 2024. [Online]. Available: https://iot-analytics.com/smart-meter-adoption/ (visited on 05/11/2025).

[113] T. Ishihara, H.-J. Guo, and H. Takeda, "A design of discrete-time integral controllers with computation delays via loop transfer recovery," *Automatica*, vol. 28, no. 3, pp. 599–603, 1992.

[114] T. Ishihara, H.-J. Guo, and H. Takeda, "A design of discrete-time integral controllers with computation delays via loop transfer recovery," *Automatica*, vol. 28, no. 3, pp. 599–603, 1992.

[115] D. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," en, *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 124–131, Apr. 1973, ISSN: 0018-9286. DOI: `10.1109/TAC.1973.1100265`. [Online]. Available: `http://ieeexplore.ieee.org/document/1100265/` (visited on 05/17/2024).

[116] M. Jaggi, "Revisiting Frank-Wolfe: Projection-Free Sparse Convex Optimization," in *Proceedings of the 30th International Conference on Machine Learning*, S. Dasgupta and D. McAllester, Eds., ser. Proceedings of Machine Learning Research, Issue: 1, vol. 28, Atlanta, Georgia, USA: PMLR, Jun. 2013, pp. 427–435. [Online]. Available: `https://proceedings.mlr.press/v28/jaggi13.html`.

[117] T. Jaksch, R. Ortner, and P. Auer, "Near-optimal regret bounds for reinforcement learning," *Journal of Machine Learning Research*, vol. 11, no. Apr, pp. 1563–1600, 2010.

[118] S. Jalali and T. Weissman, "Block and sliding-block lossy compression via mcmc," *IEEE transactions on communications*, vol. 60, no. 8, pp. 2187–2198, 2012.

[119] F. Janabi-Sharifi and M. Marey, "A Kalman-Filter-Based Method for Pose Estimation in Visual Servoing," *IEEE Transactions on Robotics*, vol. 26, no. 5, pp. 939–947, Oct. 2010, ISSN: 1552-3098, 1941-0468. DOI: `10.1109/TRO.2010.2061290`. [Online]. Available: `http://ieeexplore.ieee.org/document/5560877/` (visited on 05/18/2024).

[120] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation* (Prentice Hall information and system sciences series). Upper Saddle River, N.J: Prentice Hall, 2000, ISBN: 978-0-13-022464-4.

[121] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," en, *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, Mar. 1960, ISSN: 0021-9223. DOI: `10.1115/1.3662552`. (visited on 09/26/2022).

[122] T. Kargin*, J. Hajar*, V. Malik, and B. Hassibi, *Distributionally Robust Kalman Filtering over Finite and Infinite Horizon*, arXiv:2407.18837, Jul. 2024. [Online]. Available: `http://arxiv.org/abs/2407.18837`,

[123] T. Kargin*, J. Hajar*, V. Malik*, and B. Hassibi, *Wasserstein distributionally robust regret-optimal control in the infinite-horizon*, 2023. arXiv: `2312.17376 [eess.SY]`. [Online]. Available: `https://arxiv.org/abs/2312.17376`,

[124] T. Kargin*, J. Hajar*, V. Malik*, and B. Hassibi, "Infinite-Horizon Distributionally Robust Regret-Optimal Control," in *Proceedings of the 41st International Conference on Machine Learning*, R. Salakhutdinov, Z. Kolter, K. Heller, *et al.*, Eds., ser. Proceedings of Machine Learning Research, vol. 235, PMLR, Jul. 2024, pp. 23 187–23 214. [Online]. Available: `https://proceedings.mlr.press/v235/kargin24a.html`,

[125] T. Kargin*, J. Hajar*, V. Malik*, and B. Hassibi, "Wasserstein distributionally robust regret-optimal control over infinite-horizon," in *Proceedings of the 6th Annual Learning for Dynamics & Control Conference*, A. Abate, M. Cannon, K. Margellos, and A. Papachristodoulou, Eds., ser. Proceedings of Machine Learning Research, vol. 242, PMLR, Jul. 2024, pp. 1688–1701. [Online]. Available: `https://proceedings.mlr.press/v242/kargin24a.html`,

[126] T. Kargin*, S. Lale*, K. Azizzadenesheli, A. Anandkumar, and B. Hassibi, "Thompson sampling achieves $\tilde{O}(\sqrt{T})$ regret in linear quadratic control," in *Proceedings of Thirty Fifth Conference on Learning Theory*, P.-L. Loh and M. Raginsky, Eds., ser. Proceedings of Machine Learning Research, vol. 178, PMLR, Jul. 2022, pp. 3235–3284. [Online]. Available: `https://proceedings.mlr.press/v178/kargin22a.html`,

[127] S. Kassam and H. Poor, "Robust techniques for signal processing: A survey," *Proceedings of the IEEE*, vol. 73, no. 3, pp. 433–481, 1985, ISSN: 0018-9219. DOI: `10.1109/PROC.1985.13167`. [Online]. Available: `http://ieeexplore.ieee.org/document/1457435/` (visited on 05/17/2024).

[128] S. A. Kassam and T. L. Lim, "Robust Wiener filters," en, *Journal of the Franklin Institute*, vol. 304, no. 4-5, pp. 171–185, Oct. 1977, ISSN: 00160032. DOI: `10.1016/0016-0032(77)90011-4`. [Online]. Available: `https://linkinghub.elsevier.com/retrieve/pii/0016003277900114` (visited on 05/17/2024).

[129] B. P. Kellerhals, *Financial Pricing Models in Continuous Time and Kalman Filtering* (Lecture Notes in Economics and Mathematical Systems), G. Fandel, W. Trockel, C. D. Aliprantis, and D. Kovenock, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, vol. 506, ISBN: 978-3-540-42364-5 978-3-662-21901-0. DOI: `10.1007/978-3-662-21901-0`. [Online]. Available: `http://link.springer.com/10.1007/978-3-662-21901-0` (visited on 05/18/2024).

[130] P. P. Khargonekar and M. A. Rotea, "Mixed h/sub 2//h/sub infinity/control: A convex optimization approach," *IEEE Transactions on Automatic Control*, vol. 36, no. 7, pp. 824–837, 1991.

[131] K. Kim and I. Yang, "Distributional robustness in minimax linear quadratic control with wasserstein distance," *SIAM Journal on Control and Optimization*, vol. 61, no. 2, pp. 458–483, 2023. DOI: `10.1137/22M1494105`.

eprint: https://doi.org/10.1137/22M1494105. [Online]. Available: https://doi.org/10.1137/22M1494105.

[132] A. N. Kolmogorov, “Sur l’interpolation et extrapolation des suites stationnaires,” *CR Acad. Sci*, vol. 208, no. S 2043, 1939.

[133] A. Kolmogorov, “On the Shannon theory of information transmission in the case of continuous signals,” *IRE Transactions on Information Theory*, vol. 2, no. 4, pp. 102–108, Dec. 1956.

[134] D. Kuhn, P. M. Esfahani, V. A. Nguyen, and S. Shafieezadeh-Abadeh, “Wasserstein distributionally robust optimization: Theory and applications in machine learning,” in *Operations research & management science in the age of analytics*, Informs, 2019, pp. 130–166.

[135] S. Lacoste-Julien and M. Jaggi, *An Affine Invariant Linear Convergence Analysis for Frank-Wolfe Algorithms*, arXiv:1312.7864 [math], Jan. 2014. [Online]. Available: http://arxiv.org/abs/1312.7864 (visited on 05/05/2024).

[136] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, “Logarithmic regret bound in partially observable linear dynamical systems,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 20 876–20 888, 2020.

[137] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, “Regret minimization in partially observable linear quadratic control,” *arXiv preprint arXiv:2002.00082*, 2020.

[138] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, “Adaptive control and regret minimization in linear quadratic gaussian (lqg) setting,” in *2021 American Control Conference (ACC)*, IEEE, 2021, pp. 2517–2522.

[139] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, “Finite-time system identification and adaptive control in autoregressive exogenous systems,” in *Learning for Dynamics and Control*, PMLR, 2021, pp. 967–979.

[140] 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.

[141] A. Lapidoth, “Nearest neighbor decoding for additive non-gaussian noise channels,” *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 1520–1529, 1996.

[142] A. Lapidoth and P. Narayan, “Reliable communication under channel uncertainty,” *IEEE transactions on Information Theory*, vol. 44, no. 6, pp. 2148–2177, 1998.

[143] E. Lefferts, F. Markley, and M. Shuster, "Kalman Filtering for Spacecraft Attitude Estimation," en, *Journal of Guidance, Control, and Dynamics*, vol. 5, no. 5, pp. 417–429, Sep. 1982, ISSN: 0731-5090, 1533-3884. DOI: 10.2514/3.56190. [Online]. Available: https://arc.aiaa.org/doi/10.2514/3.56190 (visited on 05/18/2024).

[144] E. Lei, H. Hassani, and S. S. Bidokhti, "Out-of-distribution robustness in deep learning compression," *arXiv preprint arXiv:2110.07007*, 2021.

[145] F. Leibfritz and W. Lipinski, "Description of the benchmark examples in compleib 1.0," *Dept. Math, Univ. Trier, Germany*, vol. 32, 2003.

[146] B. Levy and R. Nikoukhah, "Robust Least-Squares Estimation With a Relative Entropy Constraint," en, *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 89–104, Jan. 2004, ISSN: 0018-9448. DOI: 10.1109/TIT.2003.821992. [Online]. Available: http://ieeexplore.ieee.org/document/1262619/ (visited on 05/17/2024).

[147] B. C. Levy and R. Nikoukhah, "Robust State Space Filtering Under Incremental Model Perturbations Subject to a Relative Entropy Tolerance," *IEEE Transactions on Automatic Control*, vol. 58, no. 3, pp. 682–695, Mar. 2013, ISSN: 0018-9286, 1558-2523. DOI: 10.1109/TAC.2012.2219952. [Online]. Available: http://ieeexplore.ieee.org/document/6308696/ (visited on 05/10/2024).

[148] H. LI, S. Sugasawa, and S. Katayama, "Adaptively robust and sparse \$k\$-means clustering," *Transactions on Machine Learning Research*, 2024, ISSN: 2835-8856. [Online]. Available: https://openreview.net/forum?id=EhC84fT2yA.

[149] R. Liu, G. Shi, and P. Tokekar, "Data-driven distributionally robust optimal control with state-dependent noise," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 9986–9991. DOI: 10.1109/IROS55552.2023.10342392.

[150] T. Liu, Y. Lu, B. Zhu, and H. Zhao, "Clustering High-Dimensional Data via Feature Selection," en, *Biometrics*, vol. 79, no. 2, pp. 940–950, Jun. 2023, ISSN: 0006-341X, 1541-0420. DOI: 10.1111/biom.13665. [Online]. Available: https://academic.oup.com/biometrics/article/79/2/940-950/7513996 (visited on 02/06/2025).

[151] Y. Liu, L. Zhu, M. Yamada, and Y. Yang, "Semantic Correspondence as an Optimal Transport Problem," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA: IEEE, Jun. 2020, pp. 4462–4471, ISBN: 978-1-72817-168-5. DOI: 10.1109/CVPR42600.2020.00452. [Online]. Available: https://ieeexplore.ieee.org/document/9156442/ (visited on 05/19/2024).

[152] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[153] K. Lotidis, N. Bambos, J. Blanchet, and J. Li, "Wasserstein Distributionally Robust Linear-Quadratic Estimation under Martingale Constraints," in *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, F. Ruiz, J. Dy, and J.-W. van de Meent, Eds., ser. Proceedings of Machine Learning Research, vol. 206, PMLR, Apr. 2023, pp. 8629–8644. [Online]. Available: `https://proceedings.mlr.press/v206/lotidis23a.html`.

[154] S. Loyka and C. D. Charalambous, "A general formula for compound channel capacity," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3971–3991, 2016.

[155] D. J. MacKay, "Fountain codes," *IEE Proceedings-Communications*, vol. 152, no. 6, pp. 1062–1068, 2005.

[156] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Oakland, CA, USA, vol. 1-14, 1967, pp. 281–297.

[157] V. Malik, T. Kargin, V. Kostina, and B. Hassibi, "A Distributionally Robust Approach to Shannon Limits using the Wasserstein Distance," in *2024 IEEE International Symposium on Information Theory (ISIT)*, Athens, Greece: IEEE, Jul. 2024, pp. 861–866, ISBN: 9798350382846. DOI: `10.1109/ISIT57864.2024.10619597`. [Online]. Available: `https://ieeexplore.ieee.org/document/10619597/`,

[158] V. Malik*, T. Kargin*, J. Hajar, and B. Hassibi, "Optimal infinite-horizon mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control," in *2024 60th Annual Allerton Conference on Communication, Control, and Computing*, 2024, pp. 1–8. DOI: `10.1109/Allerton63246.2024.10735276`,

[159] H. Mania, S. Tu, and B. Recht, "Certainty equivalent control of lqr is efficient," *arXiv preprint arXiv:1902.07826*, 2019.

[160] K. N. Markelle Kelly Rachel Longjohn, *The UCI machine learning repository*. [Online]. Available: `https://archive.ics.uci.edu`.

[161] A. Martin, L. Furieri, F. Dorfler, J. Lygeros, and G. Ferrari-Trecate, "Safe control with minimal regret," in *Learning for Dynamics and Control Conference*, PMLR, 2022, pp. 726–738.

[162] D. Martinelli, A. Martin, G. Ferrari-Trecate, and L. Furieri, *Closing the Gap to Quadratic Invariance: A Regret Minimization Approach to Optimal Distributed Control*, arXiv:2311.02068 [cs, eess], Nov. 2023. [Online]. Available: `http://arxiv.org/abs/2311.02068` (visited on 11/15/2023).

[163] J. Max, "Quantizing for minimum distortion," *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.

[164] McKinsey Center for Future Mobility, “Autonomous driving’s future: Convenient and connected,” McKinsey & Company, Tech. Rep., Jan. 2023. [Online]. Available: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/autonomous-drivings-future-convenient-and-connected.

[165] A. Megretski, “On the order of optimal controllers in the mixed H-2/H-infinity control,” in *Proceedings of 1994 33rd IEEE Conference on Decision and Control*, IEEE, vol. 4, 1994, pp. 3173–3174.

[166] P. Mohajerin Esfahani and D. Kuhn, “Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations,” en, *Mathematical Programming*, vol. 171, no. 1-2, pp. 115–166, Sep. 2018, ISSN: 0025-5610, 1436-4646. DOI: 10.1007/s10107-017-1172-1. [Online]. Available: http://link.springer.com/10.1007/s10107-017-1172-1 (visited on 12/11/2023).

[167] D. Mustafa and K. Glover, “Controllers which satisfy a closed-loop h/sup infinity /-norm bound and maximize an entropy integral,” in *Proceedings of the 27th IEEE Conference on Decision and Control*, 1988, 959–964 vol.2. DOI: 10.1109/CDC.1988.194456.

[168] K. Nagpal and P. Khargonekar, “Filtering and smoothing in an H/sup infinity / setting,” *IEEE Transactions on Automatic Control*, vol. 36, no. 2, pp. 152–166, Feb. 1991, ISSN: 00189286. DOI: 10.1109/9.67291. [Online]. Available: http://ieeexplore.ieee.org/document/67291/ (visited on 05/17/2024).

[169] National Highway Traffic Safety Administration, “Incident Reporting for Automated Driving Systems (ADS) and Level 2 Advanced Driver Assistance Systems (ADAS),” United States Department of Transportation, Washington, DC, Standing General Order, Jan. 2021. [Online]. Available: https://www.nhtsa.gov/laws-regulations/standing-general-order-crash-reporting.

[170] National Transportation Safety Board, “Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian, Tempe, Arizona, March 18, 2018,” Washington, DC, Highway Accident Report NTSB/HAS-19/03, Nov. 2019. [Online]. Available: https://www.ntsb.gov/investigations/Pages/HWY18MH010.aspx.

[171] V. A. Nguyen, S. Shafieezadeh-Abadeh, D. Kuhn, and P. M. Esfahani, *Bridging Bayesian and Minimax Mean Square Error Estimation via Wasserstein Distributionally Robust Optimization*, arXiv:1911.03539 [cs, math, stat], Jan. 2021. [Online]. Available: http://arxiv.org/abs/1911.03539 (visited on 09/08/2023).

[172] S. Orfanidis, "An exact solution of the time-invariant discrete kalman filter," *IEEE Transactions on Automatic Control*, vol. 27, no. 1, pp. 240–242, 1982. DOI: 10.1109/TAC.1982.1102836.

[173] I. Osband and B. Van Roy, "Why is posterior sampling better than optimism for reinforcement learning?" In *Proceedings of the 34th International Conference on Machine Learning*, D. Precup and Y. W. Teh, Eds., ser. Proceedings of Machine Learning Research, vol. 70, PMLR, Jun. 2017, pp. 2701–2710. [Online]. Available: https://proceedings.mlr.press/v70/osband17a.html.

[174] Y. Ouyang, M. Gagrani, and R. Jain, "Control of unknown linear systems with thompson sampling," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2017, pp. 1198–1205. DOI: 10.1109/ALLERTON.2017.8262873.

[175] Y. Ouyang, M. Gagrani, and R. Jain, "Learning-based control of unknown linear systems with thompson sampling," *arXiv preprint arXiv:1709.04047*, 2017.

[176] S. Oymak and N. Ozay, "Non-asymptotic identification of lti systems from a single trajectory," in *2019 American Control Conference (ACC)*, 2019, pp. 5655–5661. DOI: 10.23919/ACC.2019.8814438.

[177] B. Ozaydin, T. Zhang, D. Bhattacharjee, S. Süsstrunk, and M. Salzmann, *OMH: Structured Sparsity via Optimally Matched Hierarchy for Unsupervised Semantic Segmentation*, arXiv:2403.06546 [cs], Apr. 2024. [Online]. Available: http://arxiv.org/abs/2403.06546 (visited on 05/19/2024).

[178] I. Petersen, M. James, and P. Dupuis, "Minimax optimal control of stochastic uncertain systems with relative entropy constraints," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 398–412, Mar. 2000, ISSN: 00189286. DOI: 10.1109/9.847720. [Online]. Available: http://ieeexplore.ieee.org/document/847720/ (visited on 06/10/2024).

[179] I. R. Petersen and A. V. Savkin, *Robust Kalman filtering for signals and systems with large uncertainties* (Control engineering). Boston: Birkhäuser, 1999, ISBN: 978-0-8176-4089-7.

[180] S. Pichorides, "On the best values of the constants in the theorem of M. Riesz, Zygmund and Kolmogorov," eng, *Studia Mathematica*, vol. 44, no. 2, pp. 165–179, 1972. DOI: 10.4064/sm-44-2-165-179. [Online]. Available: http://eudml.org/doc/217677.

[181] M. E. Pittelkau, "Kalman Filtering for Spacecraft System Alignment Calibration," en, *Journal of Guidance, Control, and Dynamics*, vol. 24, no. 6, pp. 1187–1195, Nov. 2001, ISSN: 0731-5090, 1533-3884. DOI: 10.2514/2.4834. [Online]. Available: https://arc.aiaa.org/doi/10.2514/2.4834 (visited on 05/18/2024).

[182] Y. Polyanskiy and Y. Wu, "Information theory: From coding to learning," *Book draft*, 2022.

[183] H. Poor, "The rate-distortion function on classes of sources determined by spectral capacities," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 19–26, 1982.

[184] H. Poor, "On robust wiener filtering," en, *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 531–536, Jun. 1980, ISSN: 0018-9286. DOI: 10.1109/TAC.1980.1102349. [Online]. Available: http://ieeexplore.ieee.org/document/1102349/ (visited on 05/17/2024).

[185] C. Poupon, A. Roche, J. Dubois, J.-F. Mangin, and F. Poupon, "Real-time MR diffusion tensor and Q-ball imaging using Kalman filtering," en, *Medical Image Analysis*, vol. 12, no. 5, pp. 527–534, Oct. 2008, ISSN: 13618415. DOI: 10.1016/j.media.2008.06.004. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1361841508000571 (visited on 05/18/2024).

[186] H. Prabhat and R. Bhattacharya, *Optimal State Estimation in the Presence of Non-Gaussian Uncertainty via Wasserstein Distance Minimization*, arXiv:2403.13828 [cs, eess, math, stat], Mar. 2024. [Online]. Available: http://arxiv.org/abs/2403.13828 (visited on 05/11/2024).

[187] 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.

[188] C. Rino, "Factorization of spectra by discrete Fourier transforms (Corresp.)," en, *IEEE Transactions on Information Theory*, vol. 16, no. 4, pp. 484–485, Jul. 1970, ISSN: 0018-9448. DOI: 10.1109/TIT.1970.1054502. [Online]. Available: http://ieeexplore.ieee.org/document/1054502/ (visited on 11/15/2023).

[189] G. Rodriguez, "Kalman filtering, smoothing, and recursive robot arm forward and inverse dynamics," *IEEE Journal on Robotics and Automation*, vol. 3, no. 6, pp. 624–639, Dec. 1987, ISSN: 0882-4967. DOI: 10.1109/JRA.1987.1087147. [Online]. Available: http://ieeexplore.ieee.org/document/1087147/ (visited on 05/18/2024).

[190] W. Root and P. Varaiya, "Capacity of classes of gaussian channels," *SIAM Journal on Applied Mathematics*, vol. 16, no. 6, pp. 1350–1393, 1968.

[191] O. Sabag, G. Goel, S. Lale, and B. Hassibi, "Regret-optimal controller for the full-information problem," in *2021 American Control Conference (ACC)*, IEEE, 2021, pp. 4777–4782.

[192] O. Sabag, G. Goel, S. Lale, and B. Hassibi, *Regret-optimal lqr control*, 2023. arXiv: 2105.01244 [math.OC].

[193] O. Sabag and B. Hassibi, "Regret-Optimal Filtering for Prediction and Estimation," *IEEE Transactions on Signal Processing*, vol. 70, pp. 5012–5024, 2022, ISSN: 1053-587X, 1941-0476. DOI: `10.1109/TSP.2022.3212153`. [Online]. Available: `https://ieeexplore.ieee.org/document/9911672/` (visited on 11/20/2023).

[194] D. Sakrison, "The rate of a class of random processes," *IEEE Transactions on Information Theory*, vol. 16, no. 1, pp. 10–16, 1970.

[195] D. J. Sakrison, "The rate distortion function for a class of sources," *Information and Control*, vol. 15, no. 2, pp. 165–195, 1969.

[196] S. Samuelson and I. Yang, "Data-driven distributionally robust control of energy storage to manage wind power fluctuations," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*, 2017, pp. 199–204. DOI: `10.1109/CCTA.2017.8062463`.

[197] F. Santambrogio, *Optimal transport for applied mathematicians: calculus of variations, PDEs, and modeling* (Progress in nonlinear differential equations and their applications 87), eng. Cham: Birkhauser, 2015, ISBN: 978-3-319-20827-5.

[198] T. Sarkar, A. Rakhlin, and M. A. Dahleh, "Finite time lti system identification," *Journal of Machine Learning Research*, vol. 22, no. 26, pp. 1–61, 2021. [Online]. Available: `http://jmlr.org/papers/v22/19-725.html`.

[199] A. H. Sayed and T. Kailath, "A survey of spectral factorization methods," en, *Numerical Linear Algebra with Applications*, vol. 8, no. 6-7, pp. 467–496, Sep. 2001, ISSN: 1070-5325, 1099-1506. DOI: `10.1002/nla.250`. [Online]. Available: `https://onlinelibrary.wiley.com/doi/10.1002/nla.250` (visited on 11/15/2023).

[200] W. Schneider, "Analytical uses of Kalman filtering in econometrics — A survey," en, *Statistical Papers*, vol. 29, no. 1, pp. 3–33, Dec. 1988, ISSN: 0932-5026, 1613-9798. DOI: `10.1007/BF02924508`. [Online]. Available: `http://link.springer.com/10.1007/BF02924508` (visited on 05/18/2024).

[201] E. Schomig, M. Sznaier, and U.-L. Ly, "Mixed h2/h-infinity control of multimodel plants," *Journal of Guidance, Control, and Dynamics*, vol. 18, no. 3, pp. 525–531, 1995.

[202] G. Serra, P. A. Stavrou, and M. Kountouris, "On the computation of the gaussian rate-distortion-perception function," *arXiv preprint arXiv:2311.09190*, 2023.

[203] S. Shafieezadeh Abadeh, V. A. Nguyen, D. Kuhn, and P. M. Mohajerin Esfahani, "Wasserstein distributionally robust kalman filtering," in *Advances in Neural Information Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds.,

vol. 31, Curran Associates, Inc., 2018. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2018/file/15212f24321aa2c3dc8e9acf820f3c15-Paper.pdf.

[204] U. Shaked and Y. Theodor, "H/sub infinity /-optimal estimation: A tutorial," in *[1992] Proceedings of the 31st IEEE Conference on Decision and Control*, Tucson, AZ, USA: IEEE, 1992, pp. 2278–2286, ISBN: 978-0-7803-0872-5. DOI: 10.1109/CDC.1992.371384. [Online]. Available: http://ieeexplore.ieee.org/document/371384/ (visited on 05/17/2024).

[205] C. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949. DOI: 10.1109/JRPROC.1949.232969.

[206] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[207] N. Sharon, V. Peiris, N. Sukhorukova, and J. Ugon, *Flexible rational approximation and its application for matrix functions*, 2025. arXiv: 2108.09357 [math.NA]. [Online]. Available: https://arxiv.org/abs/2108.09357.

[208] S. Shoval, I. Zeitoun, and E. Lenz, "Implementation of a Kalman filter in positioning for autonomous vehicles, and its sensitivity to the process parameters," en, *The International Journal of Advanced Manufacturing Technology*, vol. 13, no. 10, pp. 738–746, Oct. 1997, ISSN: 0268-3768, 1433-3015. DOI: 10.1007/BF01179074. [Online]. Available: http://link.springer.com/10.1007/BF01179074 (visited on 05/18/2024).

[209] N. Shulman, "Communication over an unknown channel via common broadcasting," Ph.D. dissertation, Citeseer, 2003.

[210] M. Simchowitz and D. J. Foster, "Naive exploration is optimal for online lqr," *arXiv preprint arXiv:2001.09576*, 2020.

[211] M. Simchowitz, K. Singh, and E. Hazan, "Improper learning for non-stochastic control," in *Proceedings of Thirty Third Conference on Learning Theory*, J. Abernethy and S. Agarwal, Eds., ser. Proceedings of Machine Learning Research, vol. 125, PMLR, Sep. 2020, pp. 3320–3436. [Online]. Available: https://proceedings.mlr.press/v125/simchowitz20a.html.

[212] M. Sion, "On general minimax theorems.," *Pacific Journal of Mathematics*, vol. 8, no. 1, pp. 171–176, 1958, Publisher: Pacific Journal of Mathematics, A Non-profit Corporation. DOI: 10.2140/pjm.1958.8.171.

[213] M. Smith, "On stabilization and the existence of coprime factorizations," *IEEE Transactions on Automatic Control*, vol. 34, no. 9, pp. 1005–1007, Sep. 1989, ISSN: 00189286. DOI: 10.1109/9.35819. [Online]. Available:

http://ieeexplore.ieee.org/document/35819/ (visited on 09/26/2024).

[214] R. L. Smith, A. A. A. Rahni, J. Jones, and K. Wells, "A Kalman-Based Approach With EM Optimization for Respiratory Motion Modeling in Medical Imaging," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 4, pp. 410–420, Jul. 2019, ISSN: 2469-7311, 2469-7303. DOI: 10.1109/TRPMS.2018.2879441. [Online]. Available: https://ieeexplore.ieee.org/document/8520867/ (visited on 05/18/2024).

[215] H. W. Sorenson and Institute of Electrical and Electronics Engineers, Eds., *Kalman filtering, theory and application* (IEEE Press selected reprint series), eng. New York: IEEE Press, 1985, ISBN: 978-0-7803-0421-5 978-0-87942-191-5.

[216] J. Speyer, J. Deyst, and D. Jacobson, "Optimization of stochastic linear systems with additive measurement and process noise using exponential performance criteria," en, *IEEE Transactions on Automatic Control*, vol. 19, no. 4, pp. 358–366, Aug. 1974, ISSN: 0018-9286. DOI: 10.1109/TAC.1974.1100606. [Online]. Available: http://ieeexplore.ieee.org/document/1100606/ (visited on 05/17/2024).

[217] J. Speyer, C.-H. Fan, and R. Banavar, "Optimal stochastic estimation with exponential cost criteria," in *[1992] Proceedings of the 31st IEEE Conference on Decision and Control*, Tucson, AZ, USA: IEEE, 1992, pp. 2293–2299, ISBN: 978-0-7803-0872-5. DOI: 10.1109/CDC.1992.371382. [Online]. Available: http://ieeexplore.ieee.org/document/371382/ (visited on 05/11/2024).

[218] M. J. A. Strens, "A bayesian framework for reinforcement learning," in *Proceedings of the Seventeenth International Conference on Machine Learning*, ser. ICML '00, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000, pp. 943–950, ISBN: 1558607072.

[219] F. A. Taha, S. Yan, and E. Bitar, "A Distributionally Robust Approach to Regret Optimal Control using the Wasserstein Distance," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, Singapore, Singapore: IEEE, Dec. 2023, pp. 2768–2775, ISBN: 9798350301243. DOI: 10.1109/CDC49753.2023.10384311. [Online]. Available: https://ieeexplore.ieee.org/document/10384311/ (visited on 01/31/2024).

[220] B. Taskesen, D. Iancu, C. Kocyigit, and D. Kuhn, "Distributionally robust linear quadratic control," in *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., vol. 36, Curran Associates, Inc., 2023, pp. 18 613–18 632. [Online]. Available: https://proceedings.neurips.cc/paper_files/

paper/2023/file/3b7a66b2d1258e892c89f485b8f896e0-Paper-Conference.pdf.

[221] A. Tchamkerten and E. Telatar, "A feedback strategy for binary symmetric channels," in *Proceedings of the IEEE International Symposium on Information Theory*, 2002, p. 362.

[222] W. R. Thompson, "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples," *Biometrika*, vol. 25, no. 3/4, pp. 285–294, 1933.

[223] A. Tsiamis, N. Matni, and G. Pappas, "Sample complexity of kalman filtering for unknown systems," in *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, A. M. Bayen, A. Jadbabaie, G. Pappas, *et al.*, Eds., ser. Proceedings of Machine Learning Research, vol. 120, PMLR, Oct. 2020, pp. 435–444. [Online]. Available: https://proceedings.mlr.press/v120/tsiamis20a.html.

[224] A. Tsiamis and G. J. Pappas, "Finite sample analysis of stochastic system identification," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 3648–3654. DOI: 10.1109/CDC40024.2019.9029499.

[225] S. Tu, "Sample Complexity Bounds for the Linear Quadratic Regulator," Issue: UCB/EECS-2019-42, PhD Thesis, EECS Department, University of California, Berkeley, May 2019. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2019/EECS-2019-42.html.

[226] D. Tuan Pham, J. Verron, and M. Christine Roubaud, "A singular evolutive extended Kalman filter for data assimilation in oceanography," en, *Journal of Marine Systems*, vol. 16, no. 3-4, pp. 323–340, Oct. 1998, ISSN: 09247963. DOI: 10.1016/S0924-7963(97)00109-7. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0924796397001097 (visited on 05/18/2024).

[227] I. Tzortzis, C. D. Charalambous, and T. Charalambous, "Dynamic programming subject to total variation distance ambiguity," *SIAM Journal on Control and Optimization*, vol. 53, no. 4, pp. 2040–2075, 2015. DOI: 10.1137/140955707. eprint: https://doi.org/10.1137/140955707. [Online]. Available: https://doi.org/10.1137/140955707.

[228] I. Tzortzis, C. D. Charalambous, T. Charalambous, C. K. Kourtellaris, and C. N. Hadjicostis, "Robust Linear Quadratic Regulator for uncertain systems," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Las Vegas, NV, USA: IEEE, Dec. 2016, pp. 1515–1520, ISBN: 978-1-5090-1837-6. DOI: 10.1109/CDC.2016.7798481. [Online]. Available: http://ieeexplore.ieee.org/document/7798481/ (visited on 07/16/2023).

[229] U.S.-Canada Power System Outage Task Force, "Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations," Tech. Rep., Apr. 2004. [Online]. Available: https://www.energy.gov/sites/prod/files/oeprod/DocumentsandMedia/BlackoutFinal-Web.pdf.

[230] K. Vastola and H. Poor, "Robust Wiener- Kolmogorov theory," en, *IEEE Transactions on Information Theory*, vol. 30, no. 2, pp. 316–327, Mar. 1984, ISSN: 0018-9448. DOI: 10.1109/TIT.1984.1056875. [Online]. Available: http://ieeexplore.ieee.org/document/1056875/ (visited on 05/17/2024).

[231] C. Villani, *Optimal transport: old and new* (Grundlehren der mathematischen Wissenschaften 338). Berlin: Springer, 2009, ISBN: 978-3-540-71049-3.

[232] X. Wan, "Application of k-means algorithm in image compression," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 563, 2019, p. 052 042.

[233] S. Wang, Z. Wu, and A. Lim, "Robust State Estimation for Linear Systems Under Distributional Uncertainty," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5963–5978, 2021, ISSN: 1053-587X, 1941-0476. DOI: 10.1109/TSP.2021.3118540. [Online]. Available: https://ieeexplore.ieee.org/document/9563203/ (visited on 04/06/2024).

[234] S. Wang and Z.-S. Ye, "Distributionally Robust State Estimation for Linear Systems Subject to Uncertainty and Outlier," *IEEE Transactions on Signal Processing*, vol. 70, pp. 452–467, 2022, ISSN: 1053-587X, 1941-0476. DOI: 10.1109/TSP.2021.3136804. [Online]. Available: https://ieeexplore.ieee.org/document/9656678/ (visited on 04/06/2024).

[235] M. W. Watson, "Applications of Kalman Filter Models in Econometrics," English, Ph.D. dissertation, University of California, San Deigo, San Deigo, 1980. [Online]. Available: https://www.proquest.com/openview/9e5eef36e1edebca2bf3225edf9538f4/1?pq-origsite=gscholar&cbl=18750&diss=y.

[236] T. A. Welch, "A technique for high-performance data compression," *Computer*, vol. 17, no. 06, pp. 8–19, 1984.

[237] P. Whittle, "Risk-sensitive linear/quadratic/gaussian control," en, *Advances in Applied Probability*, vol. 13, no. 4, pp. 764–777, Dec. 1981, ISSN: 0001-8678, 1475-6064. DOI: 10.2307/1426972. [Online]. Available: https://www.cambridge.org/core/product/identifier/S0001867800036508/type/journal_article (visited on 04/06/2024).

[238] M. Whorton, H. Buschek, and A. J. Calise, "Homotopy algorithm for fixed order mixed h2/h-infinity design," *Journal of guidance, control, and dynamics*, vol. 19, no. 6, pp. 1262–1269, 1996.

[239] N. Wiener and E. Hopf, *Über eine Klasse singulärer Integralgleichungen* (Sitzungsberichte der Preussischen Akademie der Wissenschaften. Physikalisch-mathematische Klasse). Akad. d. Wiss., 1931.

[240] N. Wiener and P. Masani, "The prediction theory of multivariate stochastic processes: I. The regularity condition," en, *Acta Mathematica*, vol. 98, no. 0, pp. 111–150, 1957, ISSN: 0001-5962. DOI: 10.1007/BF02404472. [Online]. Available: http://projecteuclid.org/euclid.acta/1485892246 (visited on 03/03/2025).

[241] N. Wiener and P. Masani, "The prediction theory of multivariate stochastic processes, II: The linear predictor," en, *Acta Mathematica*, vol. 99, no. 0, pp. 93–137, 1958, ISSN: 0001-5962. DOI: 10.1007/BF02392423. [Online]. Available: http://projecteuclid.org/euclid.acta/1485892261 (visited on 03/03/2025).

[242] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications*, en. The MIT Press, Aug. 1949, ISBN: 978-0-262-25719-0. DOI: 10.7551/mitpress/2946.001.0001. [Online]. Available: https://direct.mit.edu/books/book/4361/Extrapolation-Interpolation-and-Smoothing-of (visited on 05/18/2024).

[243] G. T. Wilson, "The Factorization of Matricial Spectral Densities," *SIAM Journal on Applied Mathematics*, vol. 23, no. 4, pp. 420–426, 1972, Publisher: Society for Industrial and Applied Mathematics, ISSN: 00361399. DOI: 10.1137/0123044. [Online]. Available: http://www.jstor.org/stable/2100089 (visited on 11/20/2023).

[244] L. Xie, C. E. De Souza, and M. Fu, "H_∞ estimation for discrete-time linear uncertain systems," en, *International Journal of Robust and Nonlinear Control*, vol. 1, no. 2, pp. 111–123, Apr. 1991, ISSN: 1049-8923, 1099-1239. DOI: 10.1002/rnc.4590010206. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/rnc.4590010206 (visited on 05/17/2024).

[245] E.-H. Yang and J. C. Kieffer, "Simple universal lossy data compression schemes derived from the lempel-ziv algorithm," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 239–245, 1996.

[246] E.-h. Yang, Z. Zhang, and T. Berger, "Fixed-slope universal lossy data compression," *IEEE Transactions on Information theory*, vol. 43, no. 5, pp. 1465–1476, 1997.

[247] I. Yang, "Wasserstein distributionally robust stochastic control: A data-driven approach," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3863–3870, 2020.

[248] I. Yang, "Wasserstein distributionally robust stochastic control: A data-driven approach," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3863–3870, 2021. DOI: 10.1109/TAC.2020.3030884.

[249] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995, ISSN: 07338716. DOI: 10.1109/49.414651. [Online]. Available: http://ieeexplore.ieee.org/document/414651/ (visited on 03/20/2024).

[250] D. Youla, J. Bongiorno, and H. Jabr, "Modern Wiener–Hopf design of optimal controllers Part I: The single-input-output case," en, *IEEE Transactions on Automatic Control*, vol. 21, no. 1, pp. 3–13, Feb. 1976, ISSN: 0018-9286. DOI: 10.1109/TAC.1976.1101139. [Online]. Available: http://ieeexplore.ieee.org/document/1101139/ (visited on 09/26/2024).

[251] D. Youla, H. Jabr, and J. Bongiorno, "Modern Wiener-Hopf design of optimal controllers–Part II: The multivariable case," en, *IEEE Transactions on Automatic Control*, vol. 21, no. 3, pp. 319–338, Jun. 1976, ISSN: 0018-9286. DOI: 10.1109/TAC.1976.1101223. [Online]. Available: http://ieeexplore.ieee.org/document/1101223/ (visited on 12/09/2023).

[252] Youngrock Yoon, A. Kosaka, and A. Kak, "A New Kalman-Filter-Based Framework for Fast and Accurate Visual Tracking of Rigid Objects," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1238–1251, Oct. 2008, ISSN: 1552-3098, 1941-0468. DOI: 10.1109/TRO.2008.2003281. [Online]. Available: http://ieeexplore.ieee.org/document/4631506/ (visited on 05/18/2024).

[253] M.-C. Yue, D. Kuhn, and W. Wiesemann, "On linear optimization over Wasserstein balls," en, *Mathematical Programming*, vol. 195, no. 1-2, pp. 1107–1122, Sep. 2022, ISSN: 0025-5610, 1436-4646. DOI: 10.1007/s10107-021-01673-8. [Online]. Available: https://link.springer.com/10.1007/s10107-021-01673-8 (visited on 05/18/2024).

[254] C. Zalinescu, *Convex Analysis in General Vector Spaces*, en. WORLD SCIENTIFIC, Jul. 2002, ISBN: 978-981-238-067-8 978-981-277-709-6. DOI: 10.1142/5021. [Online]. Available: https://www.worldscientific.com/worldscibooks/10.1142/5021 (visited on 09/27/2024).

[255] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses," en, *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 301–320, Apr. 1981, ISSN: 0018-9286. DOI: 10.1109/TAC.1981.1102603. [Online]. Available: http://ieeexplore.ieee.org/document/1102603/ (visited on 12/11/2023).

[256] A. Zelig, H. Kariti, and N. Kaplan, "KMD clustering: Robust general-purpose clustering of biological data," en, *Communications Biology*, vol. 6, no. 1, p. 1110, Nov. 2023, ISSN: 2399-3642. DOI: 10.1038/s42003-023-05480-z. [Online]. Available: https://www.nature.com/articles/s42003-023-05480-z (visited on 02/06/2025).

[257] C. Zhao and Y. Guan, "Data-driven risk-averse stochastic optimization with Wasserstein metric," en, *Operations Research Letters*, vol. 46, no. 2, pp. 262–267, Mar. 2018, ISSN: 01676377. DOI: 10.1016/j.orl.2018.01.011. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167637718300506 (visited on 11/30/2023).

[258] Z. Zhong and J.-J. Zhu, "Nonlinear wasserstein distributionally robust optimal control," in *ICML Workshop on New Frontiers in Learning, Control, and Dynamical Systems*, 2023. [Online]. Available: https://openreview.net/forum?id=4ZhPzV5boz.

[259] K. Zhou, J. Doyle, K. Glover, and B. Bodenheimer, "Mixed H_2 and H_∞ control," in *1990 American Control Conference*, IEEE, 1990, pp. 2502–2507.

[260] K. Zhou, J. C. Doyle, and K. Glover, *Robust and optimal control*. Upper Saddle River, N.J: Prentice Hall, 1996, ISBN: 978-0-13-456567-5.

[261] I. Ziemann and H. Sandberg, "Regret lower bounds for learning linear quadratic gaussian systems," *IEEE Transactions on Automatic Control*, vol. 70, no. 1, pp. 159–173, 2025. DOI: 10.1109/TAC.2024.3439132.

[262] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Transactions on information theory*, vol. 23, no. 3, pp. 337–343, 1977.

[263] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE transactions on Information Theory*, vol. 24, no. 5, pp. 530–536, 1978.

[264] M. Zorzi, *On the Robustness of the Bayes and Wiener Estimators under Model Uncertainty*, arXiv:1508.01904 [math], May 2017. [Online]. Available: http://arxiv.org/abs/1508.01904 (visited on 05/17/2024).

[265] M. Zorzi, "Robust Kalman Filtering Under Model Perturbations," *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 2902–2907, Jun. 2017, ISSN: 0018-9286, 1558-2523. DOI: 10.1109/TAC.2016.2601879. [Online]. Available: http://ieeexplore.ieee.org/document/7549014/ (visited on 05/11/2024).