Thèse n° 10 304

Data-Driven Control and Optimization under Noisy and Uncertain Conditions

Présentée le 1er décembre 2023

Faculté des sciences et techniques de l'ingénieur
Groupe SCI STI GFT
Programme doctoral en génie électrique

pour l'obtention du grade de Docteur ès Sciences

par

Baiwei GUO

Acceptée sur proposition du jury

Prof. C. N. Jones, président du jury
Prof. G. Ferrari Trecate, Prof. M. Kamgarpour, directeurs de thèse
Prof. S. Formentin, rapporteur
Prof. D. Krishnamoorthy, rapporteur
Prof. F. Dörfler, rapporteur

École
polytechnique
fédérale
de Lausanne

2023

To my parents.

Acknowledgements

I would like to express my deepest gratitude for my thesis supervisors Prof. Giancarlo Ferrari-Trecate and Prof. Maryam Kamgarpour. I feel so lucky to work with Gianni at DECODE group for the past four years. He is always caring and patient. His guidance helped me overcome dilemmas in research, tidy up my mind when it is messy and enhance my communication skills. From him, I came to understand the good qualities of a competent and responsible researcher. I was also fortunate to take Maryam's course and have her guidance during my master studies at ETH Zurich. Without her inspiration, I would not have the chance to pursue a career in control engineering. It was such a pleasure to still have her supervision at Lausanne. Her insights and rigour are invaluable treasures to me.

I would also like to thank my defence committee members, Prof. Colin Jones, Prof. Florian Dörfler, Prof. Simone Formentin and Prof. Dinesh Krishnamoorthy. They thoroughly scrutinised my thesis and raised many meaningful questions. I really enjoyed the fruitful discussion with them. Their recognition is such an honour for me.

My doctoral studies have been a collective effort, and I'm thankful for my collaborators: Orçun Karaca, Liang Xu, Mustafa Turan, Luca Furieri, Andrea Martin and Yuning Jiang. Orçun was my master thesis supervisor helping me establish the initial attitude and confidence towards research. I am glad to have him as an adviser and friend since then. Liang is also a tutor for me. When I get confused about research and career plan, he is always willing to share his experience unreservedly. The teamwork with Luca, Mustafa and Andrea was great memories. When working on projects, I felt empowered because of their support. Yuning is an inexhaustible source of brilliant ideas. Before settling down on a concrete plan to tackle a problem, I always tended to ask him for opinions. I would also like to thank my students Sepide Azhdari, Yang Wang and Younes Moussaif for the precious inspiration from their young mindsets. Besides, my research in EPFL would become much harder but for the help from Christophe Salzman and Nicole Bouendin. Their efforts make the working experience in Laboratoire d'Automatique so smooth. I also enjoyed the discussions with Ilnura Usmanova, Yao Xuan, Prof. Alireza Karimi, Prof. Chuchu Fan and Prof. Gabriela Hug, who helped me form some very important concepts of research. The experience of hosting the summer school with Clara, Wouter, Tobias, Matilde is invaluable. From them I learned a lot. I also appreciate Elise and other members of NCCR Automation. This lively community gave us many opportunities to exchange and find new inspirations.

Beyond the realm of research, I enjoy the leisure time spent with friends. Clara, Johannes (Waibel), Johannes (Schwarz), Philippe and Emilio are around since the very beginning of my time in EPFL. At that time, parties, hiking tours and gaming sessions were daily practices. The bond with them also got me through the quarantine days. In that work-from-home period, walking down to Ouchy and playing cards at Parc de Milan were typical afternoon breaks. After Covid, I was also very happy to hang out with Mahrokh, Vaibhav, Luca, Andrea, Yingzhao, Sohail, Paul and Roland. New members of the lab, Zak, Jean Sebastian, Daniele, Loris, Mert, Manuel, Anna, Tingting, Andreas, Giulio, Nathan, Leonardo and Danilo, make it a much more pleasant and vibrant environment. I am also grateful to have the company of Dongdong, Xiaoxue, Wenjie, Jicheng, Xiaozhe, Jiawei, Zhaoming and Shaohui. The memories of hotpot parties and relaxed chit-chats with them warmed my heart.

Outside of EPFL, I've been blessed with incredible connections. I cannot remember how many workday afternoons were spent on talking with Jiaqi (Tang). It is harder and harder to find someone like him responding to whatever stupid things I say. Old friends are a treasure for me. I am so fortunate to keep in touch with Fangyuan, Xuguang, Xiaobin, Yangyang, Hanlin and Longzhu. The friends back to Zurich time, Zhongda, Dan (Wu), Tanuj, Yunni, Daifei and Jonathan (Dietrich) have given me so much support on life and career. I always want to restart the basketball training with Herbert, Changsheng and Tinggui as we did one year ago, even if it means bridging thousands of kilometers. I also miss the parties with Pengbo, Shengzhao and Yixin where our conversations have no bounds. Most of my evenings were spent on games. The objectives of the games are so clear, unlike research. Without the teammates Zhaoming, Jiasheng and Chao (Liu), the experiences would not be so wonderful.

Lastly, I would like to thank my parents. Things may get messy at times but I am not very afraid of the uncertainties ahead, since no matter what happens I can hide back home and simply find the delight out of the plain days.

Lausanne, November 3, 2023 Baiwei Guo

Abstract

Control systems operating in real-world environments often face disturbances arising from measurement noise and model mismatch. These factors can significantly impact the performance and safety of the system. In this thesis, we aim to leverage data to derive near-optimal solutions to robust control and optimization problems despite the uncertainties.

The first part focuses on data-driven robust optimal control of linear systems for trajectory tracking under measurement noise. Existing data-driven methods based on behavioral system theory use historical system trajectories to formulate robust optimal control problems. These approaches employ regularization terms to avoid overfitting. However, the corresponding suboptimality bounds are conservative due to the influence of the regularization terms on prediction error analysis. To overcome this problem, we derive two prediction error bounds which can be embedded in regularization-free robust control methods. One is attained by using bootstrap methods when resampling is affordable and the other relies on perturbation analysis of the behavioral model. These bounds enable the design of open-loop control inputs and closed-loop controllers to minimize the upper bound for the worst-case cost, while ensuring robust constraint satisfaction and suboptimality bounds that decrease to zero as noise diminishes.

The second part of this thesis addresses constrained optimization problems with model uncertainties. We assume that the objective and constraint functions are not known but can be queried while all the samples have to be feasible. This setting, called safe zeroth-order optimization, can be applied to various control problems including optimal power flow and controller tuning where system models are only partially known and safety (sample feasibility) is essential. To derive a stationary point, we propose a novel method, SZO-QQ, that iteratively constructs convex subproblems through local approximations of the unknown functions. These subproblems are easier to solve than those arising in Bayesian Optimization. We show that the iterates of SZO-QQ converge to the neighborhood of a stationary point. We also analyze the sample complexity needed to achieve a certain level of accuracy and demonstrate that SZO-QQ is more sample-efficient than log-barrier-based zeroth-order methods. To further enhance sample and computation efficiency, we propose SZO-LP, a variant of SZO-QQ that solves linear programs in each iteration. Experiments on an optimal power flow problem in a 30-bus grid highlight the scalability of our algorithms.

Résumé

Les systèmes d'asservissement fonctionnant dans des environnements réels sont souvent confrontés à des perturbations dues au bruit de mesures et erreurs de modélisation. Ces facteurs peuvent avoir un impact significatif sur la performance ainsi que la sécurité. Dans cette thèse, nous visons à exploiter les données pour dériver des solutions quasi optimales à des problèmes d'asservissement et d'optimisation robustes, malgré la présence d'incertitudes.

La première partie se concentre sur la commande « data-driven » optimal robuste des systèmes linéaires pour le suivi de trajectoire en présence de bruit de mesure. Les méthodes existantes fondées sur la théorie des systèmes comportementaux (« behavioral system theory ») utilisent des trajectoires historiques du système pour formuler un problème de commande optimal robuste. Cette approche utilise des termes de régularisation pour éviter le surapprentissage. Cependant, en faisant l'analyse de l'erreur de prédiction, les garanties de sous-optimalité sont conservatives en raison de l'influence des termes de régularisation. Pour surmonter ce problème, nous dérivons deux bornes d'erreur de prédiction qui peuvent être intégrées dans les méthodes de commande robuste sans régularisation. L'une est obtenue grâce à des méthodes de « bootstrap » lorsque le rééchantillonnage est peu cher, et l'autre repose sur l'analyse des perturbations du modèle comportemental. Ces bornes permettent de concevoir des entrées de commande en boucle ouverte et des régulateurs en boucle fermée afin de minimiser la borne supérieure du coût le plus défavorable, tout en garantissant la satisfaction robuste des contraintes et des bornes de sous-optimalité qui diminuent au fur et à mesure jusqu'à zéro lorsque le bruit disparait.

La seconde partie de cette thèse traite les problèmes d'optimisation sous contrainte avec incertitudes de modèle. Nous supposons que les fonctions d'objectif et de contrainte ne sont pas connues, mais peuvent être évaluées alors que tous les échantillons évalués doivent être faisables. Ce cadre, appelé « safe zeroth-order optimization », peut être appliqué à divers problèmes d'asservissement, y compris le flux de puissance optimal et le design de régulateurs, où les modèles des systèmes ne sont que partiellement connus et où la sécurité (faisabilité de l'échantillon) est essentielle. Pour dériver un point stationnaire, nous proposons une nouvelle méthode, SZO-QQ, qui construit itérativement des sous-problèmes convexes, par le biais d'approximations locales des fonctions inconnues. Ces sous-problèmes sont plus faciles à résoudre que ceux issus de l'optimisation bayésienne. Nous montrons que les itérations de SZO-QQ convergent vers le voisinage d'un point stationnaire. Nous analysons également la complexité échantillonnaire nécessaire pour atteindre un certain niveau de précision et démontrons que SZO-QQ est plus efficace en termes d'échantillonnage que les méthodes d'ordre zéro utilisant une barrière logarithmique. Pour améliorer l'efficacité échantillonnaire et de calcul, nous proposons SZO-LP, une variante de SZO-QQ qui résout des programmes linéaires à chaque itération. Des expériences sur un problème de flux d'énergie optimal dans un réseau à 30 bus mettent en évidence l'extensibilité de nos algorithmes.

摘要

运行在实际环境中的控制系统常常面临来自测量噪声和模型误差等不确定性的干扰。这些因素可能会显著影响系统的性能和安全。在这篇论文中，我们的研究目标是利用数据来得出鲁棒控制和优化问题的近似最优解，以尽量克服不确定性的影响。

第一部分主要讨论基于数据驱动的线性系统鲁棒控制。目标是在有测量噪声的情况下实现最优的轨迹跟踪控制。现有的基于系统行为理论（behavioral system theory）的数据驱动方法使用历史系统轨迹来构造鲁棒控制问题。这些方法通过引入正则化项以避免过度拟合。然而，由于正则化项对预测误差分析的影响，相应的次优性上界是保守的。为了克服这个问题，我们提出两种无正则化的鲁棒控制方法，他们使用的预测误差分析有所不同。第一种方法通过自助抽样（bootstrap）来获得误差上限，前提是使用者可以付出重采样所需要的资源。第二种方法则依赖于对行为模型的扰动分析。用这两种办法我们可以分别设计闭环控制器和开环控制输入。这样可以确保鲁棒性约束得到满足以及证明次优性上上界在一定范围内正比于噪声上界。

论文的第二部分涉及带有模型不确定性的约束优化问题。我们假设目标和约束函数是未知的，但可以通过在可行域内的采样来获取函数值的信息。这种问题设定属于安全零阶优化（safe zeroth-order optimization）的范畴。相应的解法可以应用于各种控制问题，包括对最优功率流（optimal power flow）的求解和对反馈控制器的自动整定。在这类应用场景中系统模型仅部分已知，而相应的安全性（即要求样本处于可行域内）是至关重要的。为了找到优化问题的稳态点，我们提出了一种新方法，称为SZO-QQ。这一方法通过对未知函数的局部估计来构建一系列子问题。这些子问题满足凸性，故比贝叶斯优化中出现的问题更容易求解。我们证明了在SZO-QQ的迭代过程中，决策变量的值会收敛到稳态点的附近。我们还给出了实现一定精度所需的样本数量上限，并证明SZO-QQ比基于对数障碍函数的零阶方法更节约样本。为了进一步提高采样和计算效率，我们提出了SZO-LP。作为SZO-QQ的一个变体，SZO-LP仅需要在每次迭代过程中求解一个线性规划问题。我们在一个30节点电网模型上进行试验以解决最优功率流问题。实验结果突出了我们的算法在向高维问题扩展的性能上相较于已有的方法具有一定优势。

Contents

Introduction

In our modern society, safety-critical engineering systems play a crucial role in shaping the infrastructure and the functioning of various industries. As these systems become increasingly complex, conventional model-based control approaches struggle to capture the full dynamics and behaviors, limiting their effectiveness. In response, contemporary control approaches are shifting towards data-driven schemes, envisioning unknown black-box systems and learning control policies from a collection of system input-output data. In this thesis, we mainly look into data-driven methods for the following two settings.

(i) Optimal control design for linear systems where the system dynamics are unknown;

(ii) Solving optimization problems with unknown objective and constraint functions.

Data-driven methods for both settings face challenges due to uncertainties affecting system dynamics and the data collection process. This introduction outlines three key challenges:

(a) Quantifying the impact of uncertainties on the prediction of system behavior is difficult: Traditional model-based approaches often assume a parametric model to describe system dynamics, enabling the utilization of various techniques to bound model mismatch arising from measurement noise [1, Section 7.4]. Consequently, when employing such models for trajectory prediction, we can estimate prediction errors by leveraging the model mismatch bounds [2, Chapter 8]. However, data-driven schemes do not rely on parametric model, making it formidable to apply similar error quantification techniques.

(b) Uncertainties introduce complexities into the process of experimental design: Building an accurate system representation (in Setting (i)) or a proxy for an unknown function (in Setting (ii)) requires the data to exhibit sufficient informativeness, which can be hindered by uncertainties. For instance, in Settiing (i), measurement noise has the potential to obscure the true system response, necessitating the acquisition of large datasets to ensure sound decision-making [3]. In Setting (ii), the uncertainties lying in the function values out of the data set make it hard to determine where to take the new samples such that we can use them for building decent function proxies.

(c) Derivation and improvement of performance guarantees while ensuring safety under uncertainties are formidable: The errors in system representations for (i) and function proxies for (ii) significantly impact the optimality of solutions derived from data-driven problem formulations. These errors may even lead to violations of the true constraints governing the system. Should methods be available to address Challenges (a) and (b), they can be employed to formulate a robust optimization problem that guarantees constraint satisfaction within the solution. Nonetheless, the quantification of conservatism remains an open question. One can also consider reducing the conservativeness by collecting more samples. The main challenge of this procedure revolves around ensuring the feasibility of these additional samples.

In this thesis, we make some attempts to address these challenges by leveraging partial knowledge of the model structure (e.g., linearity or regularity) to design sampling strategies for uncertainty quantification and utilize them in safe control and optimization.

Outline of the thesis

Part I: Data-driven Robust Control

The first part focuses on finite-horizon constrained linear-quadratic problems for linear systems under measurement noise (Setting (i)). These are core problems in control engineering, finding applications, for example, in voltage control for power systems [4] and vehicle trajectory tracking [5]. We employ the behavioral model as the data-driven system representation, where any trajectory of a given linear system is expressed as a linear combination of a Hankel (or Page) matrix's columns, representing the collected system trajectories. This way, we can derive a continuous map from data to controller design, which enables a perturbation analysis (empirical or analytical) for prediction error quantification (see Challenge (a)). Then, we formulate robust control problems using the derived error bounds. The solutions are guaranteed to satisfy the ground-truth constraints and suboptimality upper bounds (see Challenge (c)).

Chapter 1

In this chapter, our goal is to design output-feedback control policies for linear systems given a set of trajectory data with bounded measurement noise. To achieve this, model-based methods work when all system states are directly measurable or when measurement noise follows a known i.i.d. normal distribution [6, 7, 2]. In [8], the authors address this problem using a data-driven system representation, but their safety constraints only ensure stability in model-reference control. In applications like racing car trajectory planning and room temperature control [9], ensuring safety in terms of input-output trajectory constraints is also crucial.

The gap between existing literature and our objective motivates us to propose the Behavioral Input-Output Parametrization (BIOP) method. BIOP allows us to synthesize robust controllers solely based on input-output data and measurement error bounds. This method is built upon IOP [10], which treats closed-loop transfer matrices from disturbances to input and output signals as design parameters and exploits their affine relationships. In BIOP, we only need certain system responses, which can be obtained using state-of-the-art behavioral estimators, such as the signal matrix model method [11]. We also determine model mismatch upper bounds, attributed to measurement noise, through a bootstrap procedure. With this information, we formulate a robust control problem using BIOP to minimize an upper bound to the worst-case scenario.

To solve this problem, we introduce a quasiconvex relaxation technique and employ a golden-section search to find the solution. When the uncertainty is sufficiently small, the synthesized controller is both safe and nearly optimal. In this context, the suboptimality gap increases linearly with the level of model mismatch. We demonstrate the effectiveness of our algorithm by testing the suboptimality scaling in a controller synthesis task with varying model mismatch levels.

The contents of this chapter are based on the following published articles.

- L. Furieri, B. Guo, A. Martin, and G. Ferrari-Trecate, "Near-optimal design of safe output-feedback controllers from noisy data," IEEE Transactions on Automatic Control, vol. 68, no. 5, pp. 2699–2714, 2022.
- L. Furieri, B. Guo, A. Martin, and G. Ferrari-Trecate, "A behavioral input-output parametrization of control policies with suboptimality guarantees," in 60th IEEE Conference on Decision and Control (CDC), pp. 2539–2544, IEEE, 2021.

Chapter 2

To avoid the intermediate steps of system response identification and model mismatch estimation (which require extensive resampling) needed by the control scheme in Chapter 1, we propose a method to design robust controller by directly using randomly generated system trajectories. Different from Chapter 1 where data set is given and fixed, here we allow active design of experiments for data collection. Several recent papers have explored this setting [14, 15, 16, 17, 18]. However, as far as the author is aware, only one work, [19], offers rigorous suboptimality and feasibility guarantees in the presence of general bounded noise. The method in this research relies on tuning of some penalty terms. The corresponding suboptimality upper bound holds only for certain penalty coefficients which are hard to characterize in an explicit form. Considering these issues, we would like our robust control scheme to be independent of penalty tuning and enjoy a conservative suboptimality guarantee that vanish when the noise decreases to zero (in contrast with that in [19]).

To this aim, we propose a method for experiment design and Page matrix construction to derive a data-driven representation of Multiple-Input Single-Output (MISO) systems, avoiding ill-conditioned operations (in response to Challenge (b)). We then conduct perturbation analysis and formulate an min-max optimization problem for the optimal control task. The solution to this problem comes with an upper bound on suboptimality concerning the ground-truth optimum. We also extend our approach to Multiple-Input Multiple-Output (MIMO) systems while ensuring constraint satisfaction. We provide detailed implementation insights and report numerical studies demonstrating the method's applicability, notably in room temperature control.

The contents of this chapter are based on the following paper:

- B. Guo, Y. Jiang, C. N. Jones, and G. Ferrari-Trecate. "Data-driven robust control using prediction error bounds based on perturbation analysis". arXiv preprint arXiv:2308.14178, 2023.

Part II: Safe Zeroth-Order Optimization

In Part II, we focus on safe zeroth-order methods that solve unmodelled optimization problems (Setting (ii)) through feasible sampling (i.e., evaluating the unknown functions at a set of chosen feasible points). The uncertainty lies in the function values outside the sample set. To deal with Challenge (a), we utilize the Lipschitz and smoothness constants (can be found empirically) of the unknown functions to derive a range where the unknown function values might lie. This uncertainty quantification enables design of sampling mechanisms (in response to Challenge (b)) for gradient estimation with accuracy guarantees. We address Challenge (c) by constructing local feasible sets and using new samples to further decrease the objective function.

Chapter 3

Most of existing research on safe zeroth-order optimization assumes known constraints [21, 22, 23]. To address problems with unknown constraint functions, [24] adopts log-barrier functions to transfer constrained optimization into the unconstrained counterpart while Safe Bayesian Optimization [25] relies on global representation of the unknown functions and formulate subproblems to locate the next sample. However, in [24] the iterate progress can be extremely small at the proximity of the feasible region boundary, resulting in high sampling complexity. The non-convexity of the subproblems in [25] leads to a computational complexity growing exponentially with the problem dimension.

For better computational and sampling efficiency, we propose a novel method, SZO-QQ, that iteratively computes quadratic approximations of the constraint functions, constructs local feasible sets, and optimizes over them. The subproblems are convex Quadratically Constrained Quadratic Programs (QCQP). We prove that this method returns a primal-dual pair being η-KKT (a criterion measuring how close a primal-dual pair is to being KKT) within $O(1/\eta^2)$ iterations. Moreover, we numerically show that our method can achieve fast convergence compared with some state-of-the-art safe zeroth-order approaches. The effectiveness of the proposed approach is also illustrated by applying it to nonconvex optimization problems, including open-loop optimal control and Optimal Power Flow (OPF). The papers listed next are closely related:

- B. Guo, Y. Jiang, M. Kamgarpour, and G. Ferrari-Trecate, “Safe zeroth-order convex optimization using quadratic local approximations,” in 21st European Control Conference (ECC), IEEE, 2023, **winner of the Best Student Paper Award**.
- B. Guo, Y. Jiang, G. Ferrari-Trecate, and M. Kamgarpour, “Safe zeroth-order optimization using quadratic local approximations,” arXiv preprint, arXiv:2303.16659, 2023.

Chapter 4

We pursue to improve SZO-QQ by formulating easier-to-solve subproblems and avoiding too small step sizes resulting from the iterates’ proximity to feasible region boundary. Therefore, we propose Safe Zeroth-order Optimization using Linear Programs (SZO-LP). The SZO-LP method solves a linear program in each iteration to find a descent direction that tends towards the feasible region interior, followed by a step length determination. We prove that, under mild conditions, the iterates of SZO-LP have an accumulation point that is also the primal of a KKT pair. We apply SZO-LP to the OPF problem for comparison with SZO-QQ and other state-of-the-art approaches in terms of computational and sampling complexity. The following paper is closely related:

- B. Guo, Y. Wang, Y. Jiang, M. Kamgarpour, and G. Ferrari-Trecate, “Safe zeroth-order optimization using linear programs,” to appear in 62th IEEE Conference on Decision and Control (CDC), IEEE, 2023.

Other contributions

The following papers were published by the author during his doctoral studies, but the related results are not treated in details in this thesis due to their being off-topic.

- B. Guo, O. Karaca, S. Azhdari, M. Kamgarpour, and G. Ferrari-Trecate, “Actuator placement for structural controllability beyond strong connectivity and towards robustness,” in 60th IEEE Conference on Decision and Control (CDC), pp. 5294–5299, IEEE, 2021.

- B. Guo, O. Karaca, T. Summers, and M. Kamgarpour, “Actuator placement under structural controllability using forward and reverse greedy algorithms,” IEEE Transactions on Automatic Control, vol. 66, no. 12, pp. 5845–5860, 2020.
- B. Guo, O. Karaca, T. Summers, and M. Kamgarpour. “Actuator placement for optimizing network performance under controllability constraints,” in 2019 IEEE 58th Conference on Decision and Control (CDC), pp. 7140-7147. IEEE, 2019.
- L. Xu, B. Guo, and G. Ferrari-Trecate, “Finite-sample-based spectral radius estimation and stabilizability test for networked control systems,” in 20th European Control Conference (ECC), pp. 2087–2092, IEEE, 2022.
- L. Xu, M. S. Turan, B. Guo, and G. Ferrari-Trecate, “Non-conservative design of robust tracking controllers based on input-output data,” in Learning for Dynamics and Control, pp. 138–149, PMLR, 2021.
- L. Xu, B. Guo, C. Galimberti, M. Farina, R. Carli, and G. Ferrari-Trecate, “Suboptimal distributed lqr design for physically coupled systems,” IFAC-PapersOnLine, vol. 53, no. 2, pp. 11032–11037, 2020.

Part I

Data-driven Robust Control

1 Near-Optimal Design of Safe Output-Feedback Controllers

1.1 Introduction

Controllers for unknown systems can be designed according to two paradigms. *Model-based* methods follow a two-step procedure: first, data are exploited to identify the system parameters, and then a suitable controller is computed for the estimated model. On the other hand, *model-free* methods aim at directly learning an optimal control policy, without explicitly reconstructing an internal representation of the dynamical system. For a description of advantages and limitations of both approaches, we refer to [35], among recent surveys.

Given the intricacy of deriving rigorous suboptimality and sample-complexity bounds, most recent model-based and model-free approaches have focused on basic Linear Quadratic Regulator (LQR) and Linear Quadratic Gaussian (LQG) control problems as suitable benchmarks to establish how machine learning can be interfaced to the continuous action spaces typical of control [3, 36, 37, 38, 39, 40, 41]. For complex control tasks, it is more challenging to perform a thorough probabilistic analysis. Recent advances include [6, 42] for constrained and distributed LQR control with direct state measurements, respectively, and [43] for distributed output-feedback LQG.

Model-based methods may pose a difficulty when it comes to accurately identifying the state-space model of a large-scale system; this is the case, for instance, for complex networked systems such as the power grid, brain and traffic networks [44]. A promising data-driven approach that aims at bypassing a parametric state-space description of the system dynamics, while still being conceptually simple to implement for the users, hinges on the *behavioral framework* [45]. This approacch has gained renewed interest with the introduction of Data-EnablEd Predictive Control (DeePC) [46, 47, 48], which established that constrained output reference tracking can be effectively tackled in a Model Predictive Control (MPC) fashion by plugging adequately generated data into a convex optimization problem. The work [49] introduces data-driven formulations for some controller design tasks, and [16] derives stability guarantees for closed-loop control.

In many scenarios, however, exact data are not available. For instance, data can be corrupted by measurement noise or even by malicious attacks intended at fatally compromising the safety [50], the quality, and the reliability of the synthesized control policies. It is therefore essential that data-driven controllers are endowed with robustness guarantees. While some approaches have been suggested in the behavioral framework, e.g. [47, 51, 11, 52, 18, 53], it remains fairly unexplored how much noise-corrupted data affect the performance and the safety of data-driven control systems. Recently, [54, 55] have derived suboptimality [55] and sample-complexity [54] bounds for LQR through direct behavioral formulations based on 1) Linear Matrix Inequalities (LMI) and 2) the System Level Synthesis (SLS) approach, respectively. A limitation is that the internal system states must be measured, which is unrealistic for several large-scale systems [44]. Furthermore, while [55] proves that for low-enough noise a high-performing and robustly stabilizing controller can be found, the corresponding suboptimality growth rate is not explicitly derived.

1.1.1 Contributions

We propose a method for designing safe and near-optimal output-feedback control policies for linear systems in finite-horizon. Our approach is solely based on noisy data, and we explicitly characterize the growth rate of the suboptimality as a function of the mismatch between the true and estimated system. First, we develop a new relaxed optimization problem that guarantees safety while robustly accounting for noise-corrupted data. Second, we show that the incurred level of suboptimality converges to zero approximately as a linear function of the model mismatch incurred during a preliminary identification phase. Hence, upon using a consistent system estimator, the proposed controller is near-optimal in the limit of available data growing to infinity. The corresponding analysis differs from that of [37], in that a feasible solution to the proposed optimization problem must be characterized analytically while taking the safety constraints into account. In addition to dealing with constraints in an output-feedback setup — which is the main novelty with respect to [37, 6, 56, 57, 58] — the effect of the *uncertain* initial condition x_0 must be explicitly tracked in the cost. Indeed, [37] assumed that $x_0 = 0$ thanks to the considered infinite-horizon setting. On a more general level, our analysis has been inspired by [6], which combined robust control tools with classical identification techniques to ensure safety of unknown systems with suboptimality guarantees when states are fully observed. As we only have access to noisy output measurements, we exploit an input-output representation of the plant and analyze four different closed-loop responses to understand how process and output measurement noises impact safety and performance. In particular, we show a linear growth rate of the suboptimality in terms of the model mismatch level as compared to the ground-truth constrained output-feedback controller.

1.1.2 Structure of this chapter

Assuming knowledge of the underlying dynamics, Section 1.2 reviews the optimal control problem of interest and its model-based solution. Section 1.3 treats the case where we only have access to noisy input and output data; we propose an optimal control problem that ensures safety against bounded model mismatches, and discuss its numerical implementation. Section 1.4 quantifies the suboptimality incurred by our synthesis procedure as a function of the model mismatch. We present numerical experiments in Section 1.5 and include the supporting material in Section 1.6.

1.1.3 Notation of this chapter

We use $\mathbb{R}$ and $\mathbb{N}$ to denote the sets of real numbers and non-negative integers, respectively. We use I_n to denote the identity matrix of size $n \times n$ and $0_{m\times n}$ to denote the zero matrix of size $m \times n$. We write $\mathbf{x} = \text{vec}(x_1, \dots, x_N) \in \mathbb{R}^{Nn}$ to denote the vector obtained by stacking together the vectors $x_1, \dots, x_N \in \mathbb{R}^n$, and $\mathbf{M} = \text{blkdiag}(M_1, \dots, M_N)$ to denote a block-diagonal matrix with $M_1, \dots, M_N \in \mathbb{R}^{m\times n}$ on its diagonal block entries. For $\mathbf{M} = \begin{bmatrix} M_1^\mathsf{T} & \dots & M_N^\mathsf{T} \end{bmatrix}^\mathsf{T}$ we define the block-Toeplitz matrix

$$\text{Toep}_{m\times n}(\mathbf{M}) = \begin{bmatrix} M_1 & 0_{m\times n} & \dots & 0_{m\times n} \\ M_2 & M_1 & \dots & 0_{m\times n} \\ \vdots & \vdots & \ddots & \vdots \\ M_N & M_{N-1} & \dots & M_1 \end{bmatrix}.$$

More concisely, we will write Toep$(\cdot)$ when the dimensions of the blocks are clear from the context. The Kronecker product between $M \in \mathbb{R}^{m\times n}$ and $P \in \mathbb{R}^{p\times q}$ is denoted as $M \otimes P \in \mathbb{R}^{mp\times nq}$. For a vector $v \in \mathbb{R}^n$ and a matrix $A \in \mathbb{R}^{m\times n}$ we denote as $\|v\|_p$, $\|A\|_p$, their standard p-norm and induced p-norms, respectively. For a row vector $x \in \mathbb{R}^{1\times n}$we define $\|x\|_1^\star = \sum_{i=1}^n |x_i|$. The Frobenius norm of a matrix $M \in \mathbb{R}^{m\times n}$ is denoted by $\|M\|_F = \sqrt{\text{Trace}(M^\mathsf{T} M)}$. For a symmetric matrix M, we write $M \succ 0$ or $M \succeq 0$ if it is positive definite or positive semidefinite, respectively. We say that $x \sim \mathcal{D}(\mu, \Sigma)$ if the random variable $x \in \mathbb{R}^n$ follows a distribution with mean $\mu \in \mathbb{R}^n$ and covariance matrix $\Sigma \in \mathbb{R}^{n\times n}, \Sigma \succeq 0$.

A finite-horizon trajectory of length T is a sequence $\omega(0), \omega(1), \dots, \omega(T-1)$ with $\omega(t) \in \mathbb{R}^n$ for every $t = 0, 1, \dots, T-1$, which can be compactly written as

$$\boldsymbol{\omega}_{[0,T-1]} = \begin{bmatrix} \omega^\mathsf{T}(0) & \omega^\mathsf{T}(1) & \dots & \omega^\mathsf{T}(T-1) \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{nT}.$$

When the value of T is clear from the context, we will omit the subscript $[0, T-1]$. For a finite-horizon trajectory $\boldsymbol{\omega}_{[0,T-1]}$ we also define the Hankel matrix of depth L as

$$\mathcal{H}_L(\boldsymbol{\omega}_{[0,T-1]}) = \begin{bmatrix} \omega(0) & \omega(1) & \dots & \omega(T-L) \\ \omega(1) & \omega(2) & \dots & \omega(T-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ \omega(L-1) & \omega(L) & \dots & \omega(T-1) \end{bmatrix} .$$

1.2 Problem Statement: the Model-Based Case

In this section, we review safe output-feedback controller synthesis when the system model is known. We consider a discrete-time linear system with output observations, whose state-space representation is given by

$$x(t+1) = Ax(t) + Bu(t), \quad y(t) = Cx(t) + v(t)\,, \tag{1.1}$$

where $x(t) \in \mathbb{R}^n$ is the state of the system and $x(0) = x_0$ for a predefined $x_0 \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ is the control input, $y(t) \in \mathbb{R}^p$ is the observed output, and $v(t) \in \mathbb{R}^p$ denotes measurement noise $v(t) \sim \mathcal{D}(0, \Sigma_v)$, with $\Sigma_v \succ 0$. The system is controlled through a time-varying, dynamic affine control policy

$$u(t) = \sum_{k=0}^{t} K_{t,k} y(k) + g_t + w(t)\,, \tag{1.2}$$

where $K_{t,k}$ and g_t are the linear and affine parts of the policy, respectively, and $w(t) \in \mathbb{R}^m$ denotes noise on the input $w(t) \sim \mathcal{D}(0, \Sigma_w)$ with $\Sigma_w \succeq 0$, which acts as process noise.[1] Furthermore, we assume that the noise is bounded with

$$\|w\|_\infty \leq w_\infty\,, \quad \|v\|_\infty \leq v_\infty\,,$$

where $w_\infty, v_\infty > 0$. We consider the problem of synthesizing a feedback control policy that minimizes the expected value with respect to the disturbances of a quadratic objective defined over future input-output trajectories of length $N \in \mathbb{N}$:

$$J^2 := \mathbb{E}_{w,v}\left[\sum_{t=0}^{N-1} \Big(y(t)^\mathsf{T} Q_t y(t) + u(t)^\mathsf{T} R_t u(t)\Big)\right]\,, \tag{1.3}$$

where $Q_t \succeq 0$ and $R_t \succ 0$ for every $t = 0, \dots, N-1$.

The problem is made more challenging by the requirement that inputs and outputs satisfy

[1] The more general model $x(t+1) = Ax(t) + Bu(t) + w(t)$ would make the cost function depend on a specific realization A, B explicitly [59, Chapter 3]. Instead, the adopted noise model ensures that the cost only depends on the covariance matrix Σ_w and the coordinate-free parameter $\mathbf{G}$, thus making our theoretical bounds meaningful in a data-driven input-output setting.

the safety constraints

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} \in \Gamma_t \subseteq \mathbb{R}^{p+m}\,, \quad \forall t = 0, \dots, N-1\,, \tag{1.4}$$

where Γ_t is a nonempty polytope for every $t = 0, \dots, N-1$ defined as

$$\Gamma_t = \left\{(y, u) \in (\mathbb{R}^p, \mathbb{R}^m) |\ F_y^t y \leq b_y^t, F_u^t u \leq b_u^t \right\}\,, \tag{1.5}$$

with $F_y^t \in \mathbb{R}^{s \times p}$, $F_u^t \in \mathbb{R}^{s \times m}$ and $b_y^t, b_u^t \in \mathbb{R}^s$ for every $t = 0, \dots, N-1$. Despite (1.3) being convex in the input and output trajectories and Γ_t being polytopic, we highlight that minimizing (1.3) subject to (1.1), (1.2) and (1.4) is a non-convex problem in the control policy parameters $K_{t,k}$ and g_t. We refer the interested reader to [60, 61, 62, 63, 10] for classical and recent methods to overcome the non-convexity problem. For the rest of the chapter, we assume that there exists a control input (1.2) that complies with (1.4) for all possible realizations of $w(t)$ and $v(t)$.

Remark 1.1. *In this chapter, we analyze a finite-horizon control problem, which represents one iteration of a receding-horizon MPC implementation. It is therefore appropriate to compare the proposed approach with a* single *iteration of open-loop prediction approaches, such as the DeePC [46, 16]. The main difference is that we perform* closed-loop predictions, *i.e., we optimize over feedback policies $\pi(\cdot)$ such that $u(t) = \pi(y(t), \dots, y(0))$, while the DeePC [46, 16] performs* open-loop predictions, *i.e., it directly optimizes over input sequences $u(0), u(1), u(N-1)$. It is well-known that closed-loop predictions are less conservative. Indeed, by setting $K_{t,k} = 0$ in* (1.2) *the closed-loop policy reduces to an open-loop one. Most notably, closed-loop policies may preserve feasibility for significantly longer prediction horizons [64]. Naturally, the price to pay is an increased computational burden due to the larger dimensionality of the problem.*

1.2.1 Convex design through the IOP

By leveraging tools offered by the framework of the Input-Output Parametrization (IOP)[2] [10], one can formulate a convex optimization problem that computes the optimal safe feedback control policy by searching over the input-output closed-loop responses. The state-space equations (1.1) provide the following relations between trajectories

$$\mathbf{x}_{[0,N-1]} = \mathbf{P}_A(:,0)x_0 + \mathbf{P}_B\mathbf{u}_{[0,N-1]}\,, \tag{1.6}$$

$$\mathbf{y}_{[0,N-1]} = \mathbf{C}\mathbf{x}_{[0,N-1]} + \mathbf{v}_{[0,N-1]}\,, \tag{1.7}$$

[2] Similar to [62, 63], the IOP [10] yields a convex representation of input-output closed-loop responses. It is also numerically stable for the case of infinite-horizon stable plants and for finite-horizon control problems [65].

where $\mathbf{P}_A(:,0)$ denotes the first block-column of $\mathbf{P}_A$ and

$$\begin{aligned} \mathbf{P}_A &= (I-\mathbf{ZA})^{-1}\,, & \mathbf{P}_B &= (I-\mathbf{ZA})^{-1}\mathbf{ZB}\,,\\ \mathbf{A} &= I_N \otimes A\,, & \mathbf{B} &= I_N \otimes B\,,\\ \mathbf{C} &= I_N \otimes C\,, & \mathbf{Z} &= \begin{bmatrix} 0_{n\times n(N-1)} & 0_{n\times n} \\ I_{n(N-1)} & 0_{n(N-1)\times n}\end{bmatrix}. \end{aligned}$$

A few comments on the used notation are in order. First, the matrix $\mathbf{Z}$ is the block-downshift operator. Second, from now on we denote $\mathbf{G} = \mathbf{CP}_B$ to highlight that $\mathbf{G}$ is a block-Toeplitz matrix containing the first N components of the impulse response of the plant $\mathbf{G}(z) = C(zI-A)^{-1}B$. Last, the matrix $\mathbf{CP}_A(:,0)$ contains the entries of the observability matrix CA^i for $i = 0,\dots,N-1$. We denote the model-based free response of the system as $\mathbf{y}_0 = \mathbf{CP}_A(:,0)x_0$. The control policy can be rewritten as:

$$\mathbf{u}_{[0,N-1]} = \mathbf{K}\mathbf{y}_{[0,N-1]} + \mathbf{g} + \mathbf{w}_{[0,N-1]}\,, \tag{1.8}$$

where $\mathbf{K}$ and $\mathbf{g}$ are defined as:

$$\mathbf{K} = \begin{bmatrix} K_{0,0} & 0_{m\times p} & \dots & 0_{m\times p} \\ K_{1,0} & K_{1,1} & \ddots & 0_{m\times p} \\ \vdots & \vdots & \ddots & \vdots \\ K_{N-1,0} & K_{N-1,1} & \dots & K_{N-1,N-1}\end{bmatrix}, \mathbf{g} = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{N-1}\end{bmatrix}. \tag{1.9}$$

The safety constraints (1.4)-(1.5) take the form

$$\max_{\|\mathbf{v}\|_\infty \le v_\infty,\ \|\mathbf{w}\|_\infty \le w_\infty} \mathbf{F}_y\mathbf{y} \le \mathbf{b}_y\,, \quad \max_{\|\mathbf{v}\|_\infty \le v_\infty, \|\mathbf{w}\|_\infty \le w_\infty} \mathbf{F}_u\mathbf{u} \le \mathbf{b}_u\,, \tag{1.10}$$

with $\mathbf{F}_y = \text{blkdiag}(F_y^0,\dots,F_y^{N-1})$, $\mathbf{b}_y = \text{vec}(b_y^0,\dots,b_y^{N-1})$, $\mathbf{F}_u = \text{blkdiag}(F_u^0,\dots,F_u^{N-1})$, $\mathbf{b}_u = \text{vec}(b_u^0,\dots,b_u^{N-1})$, and $\max(\cdot)$ to be intended row-wise. By plugging the controller (1.8) into (1.6)-(1.7), it is easy to derive the relationships

$$\begin{bmatrix}\mathbf{y}\\ \mathbf{u}\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu}\end{bmatrix}\begin{bmatrix}\mathbf{v}+\mathbf{y}_0 \\ \mathbf{w}\end{bmatrix} + \begin{bmatrix}\mathbf{Gq}\\ \mathbf{q}\end{bmatrix}, \tag{1.11}$$

where

$$\mathbf{\Phi} = \begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu}\end{bmatrix} = \begin{bmatrix}(I-\mathbf{GK})^{-1} & (I-\mathbf{GK})^{-1}\mathbf{G} \\ \mathbf{K}(I-\mathbf{GK})^{-1} & (I-\mathbf{KG})^{-1}\end{bmatrix}, \tag{1.12}$$

and $\mathbf{q} = (I-\mathbf{KG})^{-1}\mathbf{g} = \mathbf{\Phi}_{uu}\mathbf{g}$. The parameters $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$, where $\mathbf{\Phi}_{yy} \in \mathbb{R}^{Np\times Np}$, $\mathbf{\Phi}_{yu} \in \mathbb{R}^{Np\times Nm}$, $\mathbf{\Phi}_{uy} \in \mathbb{R}^{Nm\times Np}$ and $\mathbf{\Phi}_{uu} \in \mathbb{R}^{Nm\times Nm}$, represent the four closed-loop responses defining the relationship between disturbances and input-output signals, while $\mathbf{q} \in \mathbb{R}^{Nm}$ represents the affine part of the disturbance-feedback control policy [62, 66]. To achieve a convex reformulation of the control problem under consideration, it is not hard to extend the IOP from [10] to account for the safety constraints (1.10) in a convex way. The result is summarized in the next proposition, whose proof is reported in Section 1.6.2 for completeness.

Proposition 1.1. *Consider the LTI system* (1.1) *evolving under the control policy* (1.8) *within a horizon of length $N \in \mathbb{N}$. Then:*

i) For any control policy $(\mathbf{K}, \mathbf{g})$ *that complies with the safety constraints, there exist four matrices* $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$ *and a vector* $\mathbf{q}$ *such that* $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$, $\mathbf{g} = \mathbf{\Phi}_{uu}^{-1}\mathbf{q}$, *and for all* $j = 1, \dots, sN$,

$$\begin{bmatrix} I & -\mathbf{G} \end{bmatrix} \mathbf{\Phi} = \begin{bmatrix} I & 0 \end{bmatrix}, \quad \mathbf{\Phi} \begin{bmatrix} -\mathbf{G} \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \tag{1.13}$$

$$\left\| \begin{bmatrix} v_\infty(F_{y,j}\mathbf{\Phi}_{yy})^\mathsf{T} \\ w_\infty(F_{y,j}\mathbf{\Phi}_{yu})^\mathsf{T} \end{bmatrix}^\mathsf{T} \right\|_1^\star + F_{y,j}(\mathbf{G}\mathbf{q} + \mathbf{\Phi}_{yy}\mathbf{y}_0) \leq \mathbf{b}_{y,j}\,, \tag{1.14}$$

$$\left\| \begin{bmatrix} v_\infty(F_{u,j}\mathbf{\Phi}_{uy})^\mathsf{T} \\ w_\infty(F_{u,j}\mathbf{\Phi}_{uu})^\mathsf{T} \end{bmatrix}^\mathsf{T} \right\|_1^\star + F_{u,j}(\mathbf{q} + \mathbf{\Phi}_{uy}\mathbf{y}_0) \leq \mathbf{b}_{u,j}\,, \tag{1.15}$$

$$\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu} \text{ with causal sparsities }^{3}, \tag{1.16}$$

where $F_{y,j} \in \mathbb{R}^{1\times Np}$, $F_{u,j} \in \mathbb{R}^{1\times Nm}$ *and* $\mathbf{b}_{u,j}, \mathbf{b}_{y,j} \in \mathbb{R}$ *are the* j*-th row of* $\mathbf{F}_y$, $\mathbf{F}_u$ *and* $\mathbf{b}_u, \mathbf{b}_y$, *respectively.*

ii) For any four matrices $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$ *complying with* (1.13)-(1.16) *and any vector* $\mathbf{q} \in \mathbb{R}^{mN}$, *the matrix* $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$ *is causal as per* (1.9) *and it yields the closed-loop responses* $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$. *Moreover, the affine policy* $(\mathbf{K}, \mathbf{g})$ *with* $\mathbf{g} = \mathbf{\Phi}_{uu}^{-1}\mathbf{q}$ *complies with the safety constraints.*

We remark that the IOP is well-suited to a data-driven output-feedback setup, as all affine control policies are directly parametrized through the impulse response parameters $\mathbf{G}$, without requiring an internal state-space representation. This is useful for two reasons. First, when dealing with unknown systems, the state-space parameters (A, B, C, x_0) can only be estimated up to an unknown change of variable, which may be problematic for defining the cost and the noise statistics [67]. Second, several large-scale systems feature a very large number of states, but a comparably small number of inputs and outputs, that is $n >> \max(m, p)$. In such applications, it is advantageous to bypass a state-space representation and directly deal with $\mathbf{G}$, whose dimensions do not depend on n.

From now on, to simplify the expressions appearing throughout the next sections and

[3] Specifically, they have the block lower-triangular sparsities resulting as per the expressions (1.12), the sparsity of $\mathbf{K}$ in (1.9) and that of $\mathbf{G}$.

without any loss of generality[4], we let $\mathbf{q} = \mathbf{g} = 0_{Nm\times 1}$, that is, we focus on *linear* control policies. We are ready to establish a convex formulation of the optimal control problem under study.

Proposition 1.2. *Consider the LTI system* (1.1). *The linear control policy that achieves the minimum of the cost functional* (1.3) *is given by* $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$*, where* $\mathbf{\Phi}_{uy}, \mathbf{\Phi}_{yy}$ *are optimal solutions to the following convex optimization problem:*

$$\begin{aligned}\min_{\mathbf{\Phi}} \quad & \left\| \begin{bmatrix} \mathbf{Q}^{\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_v^{\frac{1}{2}} & 0 & \mathbf{y}_0 \\ 0 & \mathbf{\Sigma}_w^{\frac{1}{2}} & 0 \end{bmatrix} \right\|_F^2 \\ & \text{subject to } \ (1.13) - (1.16)\,, \end{aligned} \tag{1.17}$$

where $\mathbf{Q} = $ *blkdiag*$(Q_0, \ldots, Q_{N-1})$*,* $\mathbf{R} = $ *blkdiag*$(R_0, \ldots, R_{N-1})$*,* $\mathbf{\Sigma}_v = I_N \otimes \Sigma_v$*,* $\mathbf{\Sigma}_w = I_N \otimes \Sigma_w$ *and where* (1.14)-(1.15) *are evaluated at* $\mathbf{q} = 0$.

Proof. We refer to Proposition 2 of [12] for a complete derivation of the cost function. To conclude the proof, it suffices to notice that the objective function and the safety constraints (1.14)-(1.15) are convex in $\mathbf{\Phi}$. □

When the system parameters (A, B, C, x_0) are known, a globally optimal solution ($\mathbf{\Phi}_{yy}^\star$, $\mathbf{\Phi}_{yu}^\star$, $\mathbf{\Phi}_{uy}^\star$, $\mathbf{\Phi}_{uu}^\star$) for problem (1.17) can be efficiently computed with off-the-shelf solvers. The corresponding globally optimal and safe control policy is then recovered as $\mathbf{K}^\star = \mathbf{\Phi}_{uy}^\star(\mathbf{\Phi}_{yy}^\star)^{-1}$.

The rest of this chapter contains our main contributions. Specifically, we address the following two questions:

Q1) How can we compute a safe control policy with performance close to that of $\mathbf{K}^\star$, solely based on libraries of *noisy* input-output trajectories?

Q2) How steeply does the suboptimality grow with respect to $\mathbf{K}^\star$ as the noise increases?

1.3 The Data-Driven Case: Robustly Safe Controller Synthesis From Noisy Data

We answer question Q1) by developing a method to synthesize near-optimal safe controllers from noisy data. The main result of this section is an optimization problem based on the IOP that tightly approximates the optimal and safe control policy, despite the fact

[4] One can redefine $\tilde{\mathbf{y}} = \begin{bmatrix}1 & \mathbf{y}^\mathsf{T}\end{bmatrix}^\mathsf{T}$, $\overline{\mathbf{v}} = \begin{bmatrix}1 & \mathbf{v}^\mathsf{T}\end{bmatrix}^\mathsf{T}$, $\overline{\mathbf{C}} = \begin{bmatrix}0_{1\times Nn} \\ \mathbf{C}\end{bmatrix}$, $\overline{\mathbf{K}} = \begin{bmatrix}\mathbf{g} & \mathbf{K}\end{bmatrix}$ and $\overline{\mathbf{\Phi}}$ as per (1.12) with $\overline{\mathbf{G}}$ and $\overline{\mathbf{K}}$ in place of $\mathbf{G}$ and $\mathbf{K}$, respectively. Minor modifications to (1.14)-(1.15) are needed as well.

that the noise-corrupted data only yield approximate estimates of the system impulse and free response. We conclude by offering novel insights on its properties and its numerical implementation based on convex optimization.

1.3.1 From noise-corrupted data to doubly-robust optimal control

From now on, the dynamics matrices (A, B, C) and the initial state x_0 are *unknown.* Instead, only the following data are available:

D1 A noisy system trajectory $\{y^h(t), u^h(t)\}_{t=-T}^{-1}$ recorded offline during an experiment.

D2 The cost matrices Q_t, R_t, the matrices Σ_v, Σ_w, the safety sets Γ_t, and the bounded sets $\boldsymbol{\mathcal{W}} = \{\mathbf{w}|\ \|\mathbf{w}\|_\infty \leq w_\infty\}$ and $\boldsymbol{\mathcal{V}} = \{\mathbf{v}|\ \|\mathbf{v}\|_\infty \leq v_\infty\}$ where disturbances live.

Our approach exploits the noisy data in **D1** to compute approximate system responses $\widehat{\mathbf{G}}$ and $\widehat{\mathbf{y}}_0$ in a preliminary identification step. We work under the following assumption.

Assumption 1.1. *Let* $\boldsymbol{\Delta} = \mathbf{G} - \widehat{\mathbf{G}}$ *and* $\boldsymbol{\delta}_0 = \mathbf{y}_0 - \widehat{\mathbf{y}}_0$*. There exist* $\epsilon_{2,G}, \epsilon_{\infty,G}, \epsilon_{2,y}, \epsilon_{\infty,y} > 0$ *such that,*

$$\begin{aligned} \|\boldsymbol{\Delta}\|_2 &\leq \epsilon_{2,G}, \quad \|\boldsymbol{\delta}_0\|_2 \leq \epsilon_{2,y}\,, \\ \|\boldsymbol{\Delta}\|_\infty &\leq \epsilon_{\infty,G}, \ \|\boldsymbol{\delta}_0\|_\infty \leq \epsilon_{\infty,y}\,. \end{aligned}$$

Note that, in practice, a meaningful bound on δ_0 is only available if A is stable or the time-horizon is sufficiently short. Let us define $\epsilon_2 = \max(\epsilon_{2,G}, \epsilon_{2,y})$ and $\epsilon_\infty = \max(\epsilon_{\infty,G}, \epsilon_{\infty,y})$. Assumption 1.1 can be fulfilled using different methods over the available data **D1**; for instance, one may utilize standard least-squares identification that comes with probabilistic and non-asymptotic error bounds [7, 68], or more sophisticated stochastic estimators based on behavioral theory such as maximum-likelihood predictors [11], which also come with quantifiable error bounds [69]. Our results are independent of the choice of the identification scheme. A discussion as to how recent behavioral approaches can be used for identification is reported in Section 1.6.1. These estimators will be used in the numerical examples in Section 1.5.

After condensing the effect of noise-corrupted data into model mismatch parameters $\boldsymbol{\Delta}, \boldsymbol{\delta}_0$, we formulate a *doubly-robust* control problem, that is, a problem where we enforce constraint satisfaction for 1) all possible model mismatches $(\boldsymbol{\Delta}, \boldsymbol{\delta}_0)$, and 2) all possible disturbances sequences $\mathbf{w} \in \boldsymbol{\mathcal{W}}$ and $\mathbf{v} \in \boldsymbol{\mathcal{V}}$. In particular, define $\boldsymbol{\theta} = (\boldsymbol{\Delta}, \boldsymbol{\delta}_0, \mathbf{w}, \mathbf{v})$ and let

$$\begin{aligned} \mathbf{y}(\mathbf{K}, \boldsymbol{\theta}) &= \widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0 + (\widehat{\mathbf{G}} + \boldsymbol{\Delta})\mathbf{u}(\mathbf{K}, \boldsymbol{\theta}) + \mathbf{v}\,, \\ \mathbf{u}(\mathbf{K}, \boldsymbol{\theta}) &= \mathbf{K}\mathbf{y}(\mathbf{K}, \boldsymbol{\theta}) + \mathbf{w}\,, \end{aligned}$$

be the closed-loop trajectories associated with a specific controller $\mathbf{K}$ and disturbance and mismatch realizations $\boldsymbol{\theta}$. Further, define the set of doubly-robust controllers as:

$$\boldsymbol{\mathcal{K}} = \{\mathbf{K} \text{ in (1.9)} |\ (\mathbf{y}(\mathbf{K},\boldsymbol{\theta}), \mathbf{u}(\mathbf{K},\boldsymbol{\theta})) \in \boldsymbol{\Gamma},\ \forall \boldsymbol{\theta} \in \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{W}} \times \boldsymbol{\mathcal{V}}\},$$

with $\boldsymbol{\mathcal{E}} = \{(\boldsymbol{\Delta}, \boldsymbol{\delta}_0)|\, \|\boldsymbol{\Delta}\|_p \leq \epsilon_p,\ \|\boldsymbol{\delta}_0\|_p \leq \epsilon_p,\ \forall p \in \{2,\infty\}\}$, $\boldsymbol{\Gamma} = \Gamma_0 \times \Gamma_1 \times \cdots \times \Gamma_{N-1}$, and assume that $\boldsymbol{\mathcal{K}}$ is not empty. Then, the doubly-robust problem of interest takes the form

$$\min_{\mathbf{K}\in\boldsymbol{\mathcal{K}}} \max_{(\boldsymbol{\Delta},\boldsymbol{\delta}_0)\in\boldsymbol{\mathcal{E}}} \sqrt{\mathop{\mathbb{E}}_{\mathbf{w},\mathbf{v}} \left[\mathbf{y}(\mathbf{K},\boldsymbol{\theta})^\mathsf{T}\mathbf{y}(\mathbf{K},\boldsymbol{\theta}) + \mathbf{u}(\mathbf{K},\boldsymbol{\theta})^\mathsf{T}\mathbf{u}(\mathbf{K},\boldsymbol{\theta})\right]}, \tag{1.18}$$

where we have selected the weights $\mathbf{Q}$, $\mathbf{R}$, $\boldsymbol{\Sigma}_w$, $\boldsymbol{\Sigma}_v$ to be identity matrices with appropriate dimensions. The same assumption is used in the rest of the chapter, in order to facilitate the derivations. However, we note that all our results can be easily adapted to non-identity weights. Next, we observe that the doubly-robust optimization problem admits an equivalent formulation in terms of the closed-loop response parameters.

Proposition 1.3. *Letting* $\boldsymbol{\Phi}_{yy} = \widehat{\boldsymbol{\Phi}}_{yy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}$, $\boldsymbol{\Phi}_{yu} = \boldsymbol{\Phi}_{yy}(\widehat{\mathbf{G}} + \boldsymbol{\Delta})$, $\boldsymbol{\Phi}_{uy} = \widehat{\boldsymbol{\Phi}}_{uy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}$, $\boldsymbol{\Phi}_{uu} = (I - \widehat{\boldsymbol{\Phi}}_{uy}\boldsymbol{\Delta})^{-1}\widehat{\boldsymbol{\Phi}}_{uu}$, *the optimization problem* (1.18) *is equivalent to*

$$\min_{\widehat{\boldsymbol{\Phi}}\in\boldsymbol{\Pi}} \max_{(\boldsymbol{\Delta},\boldsymbol{\delta}_0)\in\boldsymbol{\mathcal{E}}} \left\| \begin{bmatrix} \boldsymbol{\Phi}_{yy} & \boldsymbol{\Phi}_{yu} \\ \boldsymbol{\Phi}_{uy} & \boldsymbol{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} I & 0 & \widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0 \\ 0 & I & 0 \end{bmatrix} \right\|_F, \tag{1.19}$$

where the set of doubly-robust closed-loop responses $\boldsymbol{\Pi}$ *is*

$$\boldsymbol{\Pi} = \{\widehat{\boldsymbol{\Phi}}|\ (1.20) - (1.23),\ \forall j = 1, \ldots, sN,\ \forall (\boldsymbol{\Delta}, \boldsymbol{\delta}_0) \in \boldsymbol{\mathcal{E}}\},$$

with

$$\begin{bmatrix} I & -\widehat{\mathbf{G}} \end{bmatrix} \widehat{\boldsymbol{\Phi}} = \begin{bmatrix} I & 0 \end{bmatrix}, \quad \widehat{\boldsymbol{\Phi}} \begin{bmatrix} -\widehat{\mathbf{G}} \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \tag{1.20}$$

$$\left\| \begin{bmatrix} v_\infty \left(F_{y,j}\boldsymbol{\Phi}_{yy}\right)^\mathsf{T} \\ w_\infty \left(F_{y,j}\boldsymbol{\Phi}_{yu}\right)^\mathsf{T} \end{bmatrix} \right\|_1 + (F_{y,j}\boldsymbol{\Phi}_{yy})\left(\widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0\right) \leq \mathbf{b}_{y,j}, \tag{1.21}$$

$$\left\| \begin{bmatrix} v_\infty \left(F_{u,j}\boldsymbol{\Phi}_{uy}\right)^\mathsf{T} \\ w_\infty \left(F_{u,j}\boldsymbol{\Phi}_{uu}\right)^\mathsf{T} \end{bmatrix} \right\|_1 + (F_{u,j}\boldsymbol{\Phi}_{uy})\left(\widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0\right) \leq \mathbf{b}_{u,j}, \tag{1.22}$$

$$\widehat{\boldsymbol{\Phi}}_{yy}, \widehat{\boldsymbol{\Phi}}_{yu}, \widehat{\boldsymbol{\Phi}}_{uy}, \widehat{\boldsymbol{\Phi}}_{uu} \textit{ with causal sparsities.} \tag{1.23}$$

The proof of Proposition 1.3 can be found in Section 1.6.3. We remark that the closed-loop responses $\boldsymbol{\Phi}$ appearing in (1.19), (1.21) and (1.22) are associated with the *true* impulse response, whereas the closed-loop responses $\widehat{\boldsymbol{\Phi}}$ appearing in (1.20) and (1.23) are associated with the *estimated* impulse response. This is because, while we are interested in minimizing the cost and satisfying the safety constraints for the *real* system, we can only parametrize the closed-loop responses for the identified system.

The robust optimization problem (1.19) is non-convex in the cost and in the constraints because $\mathbf{\Phi}$ is a nonlinear function of the matrix variables $\widehat{\mathbf{\Phi}}$ and $\mathbf{\Delta}$. Therefore, it is challenging to find a feasible solution, let alone the optimal one. We note that, for the case of open-loop control policies, one may use constraint-tightening approaches such as those of [70, 71]. In this chapter, we propose an analysis that compares feedback control policies. Specifically, we derive suboptimality guarantees with respect to the optimal model-based linear feedback policy as a function of the model mismatch level.

1.3.2 Proposed relaxation for safe controller synthesis

Our first main result is to derive a relaxation of the intractable problem (1.19) that we can solve in practice. Our proposed approach is to 1) upper bound the cost function, and 2) tighten the safety constraints with more tractable expressions. In Section 1.4 we will explicitly quantify the suboptimality incurred by these approximations. At its core, this methodology is inspired by that developed in [6] for the state-feedback case without measurement noise. However, the addition of output-feedback and measurement noise leads to new terms both in the cost and the safety constraints that are more challenging to analyze.

The following two lemmas establish the basis for our relaxation. Let $J(\mathbf{G}, \mathbf{K}) = (\mathbb{E}_{\mathbf{w},\mathbf{v}}[\mathbf{y}^\mathsf{T}\mathbf{y} + \mathbf{u}^\mathsf{T}\mathbf{u}])^{\frac{1}{2}}$ denote the square root of the cost in (1.3). Lemma 1.1 provides the new expression which upper bounds $J(\mathbf{G}, \mathbf{K})$ and Lemma 1.2 provides a tightened form of the safety constraints. Their rather lengthy technical proof is reported in the Sections 1.6.4 and 1.6.5, respectively.

Lemma 1.1. *Let $\widehat{\mathbf{\Phi}}$ denote the closed-loop responses obtained by applying $\mathbf{K}$ to $\widehat{\mathbf{G}}$. Further assume that $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_2 \leq \gamma$, where $\gamma \in [0, \epsilon_2^{-1})$. Then, we have*

$$J(\mathbf{G}, \mathbf{K}) \leq \frac{J_{UB}}{1 - \epsilon_2 \gamma} \tag{1.24}$$

where

$$J_{UB} = \left\| \begin{bmatrix} \sqrt{1 + h(\epsilon_2, \gamma, \widehat{\mathbf{G}}) + h(\epsilon_2, \gamma, \widehat{\mathbf{y}}_0)}\,\widehat{\mathbf{\Phi}}_{yy} & \widehat{\mathbf{\Phi}}_{yu} & \widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \sqrt{1 + h(\epsilon_2, \gamma, \widehat{\mathbf{y}}_0)}\,\widehat{\mathbf{\Phi}}_{uy} & \widehat{\mathbf{\Phi}}_{uu} & \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F,$$

and $h(\epsilon, \gamma, \mathbf{Y}) = \epsilon^2(2 + \gamma\|\mathbf{Y}\|_2)^2 + 2\epsilon\,\|\mathbf{Y}\|_2\,(2 + \gamma\,\|\mathbf{Y}\|_2)$.

Lemma 1.1 exploits the upper bound $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_2 \leq \gamma$ to establish an explicit relationship between $J(\mathbf{G}, \mathbf{K})$, the cost obtained by applying a controller $\mathbf{K}$ to the real system $\mathbf{G}$, and $J(\widehat{\mathbf{G}}, \mathbf{K})$, the cost obtained by applying the same controller to the estimated system $\widehat{\mathbf{G}}$. To see this, notice that (1.24) can be equivalently rewritten as

$$J(\mathbf{G}, \mathbf{K}) \leq \frac{\left(J(\widehat{\mathbf{G}}, \mathbf{K})^2 + \|\widehat{\mathbf{\Phi}}_{yy}\|_F^2 (h(\epsilon_2, \gamma, \widehat{\mathbf{G}}) + \right.}{1 - \epsilon_2\gamma}$$

$$\frac{\left. +h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)) + \|\widehat{\mathbf{\Phi}}_{uy}\|_F^2 h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)\right)^{\frac{1}{2}}}{1-\epsilon_2\gamma}\,. \tag{1.25}$$

The expression (1.25) upper bounds the gap between $J(\mathbf{G},\mathbf{K})$ and $J(\widehat{\mathbf{G}},\mathbf{K})$ as a quantity that increases with ϵ_2 and with the norm of $\widehat{\mathbf{G}},\widehat{\mathbf{y}}_0,\widehat{\mathbf{\Phi}}$. We note that a similar result has appeared in [37, Proposition 3.2]. However, Lemma 1.1 additionally takes into account how an uncertain $\hat{x}(0)$ affects the cost through the free response $\widehat{\mathbf{y}}_0$. We now derive a tightened - yet more tractable - expression for the safety constraints (1.21)-(1.22).

Lemma 1.2. *Assume* $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty \leq \tau$, *where* $\tau \in [0,\epsilon_\infty^{-1})$. *Then, if for all* $j = 1,\dots,sN$ *the closed-loop responses* $\widehat{\mathbf{\Phi}}$ *satisfy the tightened safety constraints*

$$f_{1,j}(\widehat{\mathbf{\Phi}}) + f_{2,j}(\widehat{\mathbf{\Phi}}) + f_{3,j}(\widehat{\mathbf{\Phi}}) \leq \mathbf{b}_{y,j}\,, \tag{1.26}$$

$$f_{4,j}(\widehat{\mathbf{\Phi}}) + f_{5,j}(\widehat{\mathbf{\Phi}}) + f_{6,j}(\widehat{\mathbf{\Phi}}) \leq \mathbf{b}_{u,j}\,, \tag{1.27}$$

where

$$f_{1,j}(\widehat{\mathbf{\Phi}}) = \frac{v_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1^\star}{1-\epsilon_\infty\tau}\,,\quad f_{4,j}(\widehat{\mathbf{\Phi}}) = \frac{v_\infty\left\|F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\right\|_1^\star}{1-\epsilon_\infty\tau}\,,$$

$$f_{2,j}(\widehat{\mathbf{\Phi}}) = w_\infty\left\|\begin{bmatrix}\left(F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right)^\mathsf{T}\\ \epsilon_\infty\frac{1+\tau\|\widehat{\mathbf{G}}\|_\infty}{1-\epsilon_\infty\tau}\left(F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right)^\mathsf{T}\end{bmatrix}\right\|_1\,,$$

$$f_{5,j}(\widehat{\mathbf{\Phi}}) = w_\infty\left\|\begin{bmatrix}\left(F_{u,j}\widehat{\mathbf{\Phi}}_{uu}\right)^\mathsf{T}\\ \epsilon_\infty\frac{1+\tau\|\widehat{\mathbf{G}}\|_\infty}{1-\epsilon_\infty\tau}\left(F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\right)^\mathsf{T}\end{bmatrix}\right\|_1\,,$$

$$f_{3,j}(\widehat{\mathbf{\Phi}}) = F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\widehat{\boldsymbol{y}}_0 + \epsilon_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1^\star\left(\frac{1+\tau\|\widehat{\mathbf{y}}_0\|_\infty}{1-\epsilon_\infty\tau}\right)\,,$$

$$f_{6,j}(\widehat{\mathbf{\Phi}}) = F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\widehat{\boldsymbol{y}}_0 + \epsilon_\infty\left\|F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\right\|_1^\star\left(\frac{1+\tau\|\widehat{\mathbf{y}}_0\|_\infty}{1-\epsilon_\infty\tau}\right)\,,$$

then $\widehat{\mathbf{\Phi}}$ *satisfies the safety constraints* (1.21)-(1.22) *for all* $(\mathbf{\Delta},\boldsymbol{\delta}_0) \in \boldsymbol{\mathcal{E}}$.

Lemma 1.2 exploits the upper bound $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty \leq \tau$ to quantify the worst-case effect of the disturbances in increasing the values of the inputs and the outputs. In our setup, similar to [6], the feasible set shrinks in the presence of larger impulse and free response estimation error ϵ_∞. This is because (1.26)-(1.27) are more restrictive, and will eventually become infeasible for sufficiently large ϵ_∞. Instead, the effect of increasing the value of τ is less intuitive. Indeed, as τ increases, the constraint $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty \leq \tau$ softens while (1.26)-(1.27) tighten. It is therefore necessary to explicitly optimize over τ. We are now ready to establish a relaxation of problem (1.19).

Theorem 1.1. *Consider the following optimization problem:*

$$\begin{aligned} &\min_{\gamma \in [0,\epsilon_2^{-1}), \tau \in [0,\epsilon_\infty^{-1})} \frac{1}{1-\epsilon_2\gamma} \min_{\widehat{\mathbf{\Phi}}} && J_{UB} && (1.28)\\ &\text{subject to} && (1.20),\ (1.23)\,, \\ & && \left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_2 \leq \gamma\,, \quad \left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty \leq \tau\,, && (1.29)\\ & && (1.26)-(1.27), \quad \forall j = 1,\dots,sN\,, \end{aligned}$$

where J_{UB} is defined in Lemma 1.1. Then, (1.28) *has the following properties:*

i) upon fixing any specific values for $\gamma \in [0,\epsilon_2^{-1})$ and $\tau \in [0,\epsilon_\infty^{-1})$, the optimization problem is convex in $\widehat{\mathbf{\Phi}}$,

ii) all of its feasible solutions yield a controller $\widehat{\mathbf{K}} = \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1}$ complying with the safety constraints (1.14)-(1.15) *for the real system,*

iii) its minimal cost upper bounds that of (1.18).

Proof. Lemma 1.1 shows that the cost of (1.28) upper bounds $J(\mathbf{G},\mathbf{K}) = J(\mathbf{G},\widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1})$ for every feasible $\mathbf{K}$. Lemma 1.2 shows that (1.26)-(1.27) imply the doubly-robust constraints (1.21)-(1.22) for all $(\mathbf{\Delta}, \boldsymbol{\delta}_0) \in \boldsymbol{\mathcal{E}}$. Hence, $\widehat{\mathbf{K}} = \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1}$ complies with safety constraints (1.14)-(1.15) for the real system. When γ and τ are fixed, it remains to optimize over $\widehat{\mathbf{\Phi}}$. The cost function is convex in $\widehat{\mathbf{\Phi}}$ and so are the constraints of the inner optimization problem. □

Theorem 1.1 shows that problem (1.19), which is non-convex in its matrix variables, can be approximated as the problem of solving a convex optimization problem[5] for each choice of the scalar variables γ and τ. The ϵ-dependent suboptimality introduced by such an approximation will be quantified in the next section. The global optimum of (1.28) is thus determined by exhaustive search over the box $(\gamma,\tau) \in [0,\epsilon_2^{-1}) \times [0,\epsilon_\infty^{-1})$, for instance through gridding, random search [72] or bisection [73]. Gridding over (τ,γ) and solving a convex optimization problem each time may significantly increase the computational burden if we are interested in determining a near-optimal solution with very low tolerance. Similarly to [37], in the next proposition we show that the inner cost function in problem (1.28) can be made independent of γ by introducing a parameter $\alpha \in \mathbb{R}$ that acts as an upper bound to γ. As a result, the overall cost becomes quasiconvex[6] in γ, and the globally optimal $\gamma^\star(\tau)$ for each fixed τ can be found efficiently through golden-section search [75].

[5] Specifically, a semidefinite program (SDP) due to the presence of quadratic $\|\cdot\|_2$ constraints.

[6] A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is quasiconvex if and only if $f(\theta x_1 + (1-\theta)x_2) \leq \max(f(x_1), f(x_2))$ for every $x_1, x_2 \in \mathbb{R}^n$ and every $\theta \in [0,1]$. We refer to [74] for a comprehensive discussion.

Proposition 1.4. *Fix $\alpha \in [0, \epsilon_2^{-1})$ and consider the following optimization problem*

$$\begin{aligned} \min_{\gamma\in[0,\alpha],\tau\in[0,\epsilon_\infty^{-1})} \frac{1}{1-\epsilon_2\gamma} \min_{\widehat{\mathbf{\Phi}}} \quad & J^{\alpha}_{UB}(\widehat{\mathbf{\Phi}}) \\ \text{subject to} \quad & (1.20),\ (1.23),\ (1.26),\ (1.27),\ (1.29) \\ & \forall j = 1,\dots,sN\,, \end{aligned} \tag{1.30}$$

where $J^{\alpha}_{UB}(\widehat{\mathbf{\Phi}})$ is defined as

$$\left\| \begin{bmatrix} \sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{G}})+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widehat{\mathbf{\Phi}}_{yy} & \widehat{\mathbf{\Phi}}_{yu} & \widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widehat{\mathbf{\Phi}}_{uy} & \widehat{\mathbf{\Phi}}_{uu} & \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F .$$

Then, the statements i), ii) and iii) of Theorem 1.1 hold. Furthermore

iv) The cost function of problem (1.30) *is quasiconvex in γ.*

Proof. Since $\gamma \leq \alpha$, $\alpha < \epsilon_2^{-1}$, and $h(\epsilon, \gamma, \cdot)$ is a monotonically increasing function of γ, then the inequality (1.24) in Lemma 1.1 continues to hold when putting α in place of γ inside the $h(\cdot)$ functions. The constraints of (1.28) are unaffected. Hence, *i*), *ii*) and *iii*) of Theorem 1.1 continue to hold. It remains to prove *iv*). Let us fix any value for τ. First, notice that $J^{\alpha}_{UB}(\widehat{\mathbf{\Phi}})$ is a convex function of $\widehat{\mathbf{\Phi}}$ and does not depend on γ, and that the feasible set of the inner minimization in problem (1.30) is convex. Denote as $g(\gamma)$ the optimal value of the inner optimization problem. We are left with minimizing the functional $\frac{g(\gamma)}{1-\epsilon_2\gamma}$ over γ. We know that $g(\gamma)$ is convex in γ because it is obtained as the partial minimization of a convex functional over a convex set [76], and that $(1-\epsilon_2\gamma)$ is concave in γ. Since the ratio of a non-negative convex function and a positive concave function is quasiconvex, we conclude that the cost of problem (1.30) is quasiconvex in γ. □

In [37], the idea of using the parameter α was relying on a lemma from [77]. Here, we have derived an alternative self-contained proof that holds also for the case $x_0 \neq 0$. In summary, for a fixed $\alpha < \epsilon_2^{-1}$, for every τ gridding the interval $[0, \epsilon_\infty^{-1}]$ and for γ chosen according to golden-search, we solve the corresponding instance of the inner optimization problem in (1.30), which is convex in $\widehat{\mathbf{\Phi}}$. We also note that an infinite-horizon version of problem (1.30) can be established by adding a tail variable and adopting a finite-horizon approximation of stable transfer functions similar to [6].

Last, one may wonder whether the cost function of problem (1.30) is jointly quasiconvex in γ and τ, as conjectured in [6]. Here, we clarify that this may not be the case, even for the state-feedback framework of [6]. For instance, similar to the constraints (1.26)-(1.27) and those of [6], consider the function $s : \mathbb{R}^2 \to \mathbb{R}$ defined as $s(x, y) = \frac{|y|x}{(1-x)}$. Fixing

Table 1.1: Convexity properties for the proposed reformulations.

	QC in γ	QC in (τ,γ)	C for fixed (γ,τ)
(1.28)	X	X	✓
(1.30)	✓	X	✓

$x_1 = 1, x_2 = 0, y_1 = -\frac{1}{2}, y_2 = \frac{1}{2}$, one can verify that

$$s(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2) = |0.5 - \theta| \frac{\theta}{1-\theta} ,$$

is not quasiconvex for $\theta \in [0,1]$. Based on this reasoning, the cost of problem (1.30), and similarly the objective (2.3) in [6], may not be quasiconvex in τ. Hence, exhaustive search over τ remains the only solution in general. Table 1.1 summarizes the convexity properties of (1.28) and (1.30).

1.4 Suboptimality Analysis

In this section, we tackle question Q2) in Section 1.2 about performance degradation as a function of the level of model-mismatch due to noisy data. We denote as $\mathbf{K}^\star, \mathbf{\Phi}^\star$ the optimal controller for the real constrained problem (1.17) and corresponding closed-loop responses. Similarly, we denote as $\widehat{\mathbf{K}}^\star, \widehat{\mathbf{\Phi}}^\star$ the optimal controller for the optimization problem (1.30) and corresponding closed-loop responses. Further, we let $J^\star = J(\mathbf{G}, \mathbf{K}^\star)$ and $\hat{J} = J(\mathbf{G}, \widehat{\mathbf{K}}^\star)$. We aim to characterize the relative suboptimality gap $\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}}$, and specifically we will show that

$$\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}} \leq \mathcal{O}(\epsilon_2) + \tilde{S}(\epsilon_\infty, \epsilon_2) ,$$

where $\tilde{S}(\epsilon_\infty, \epsilon_2) = S(\epsilon_\infty)(1 + \mathcal{O}(\epsilon_2))$. Here, $S(\epsilon_\infty)$ quantifies the suboptimality incurred by tightening the constraints and is such that $S(0) = 0$. We prove that if ϵ_2 and ϵ_∞ are small enough and the optimal controller $\mathbf{K}^\star$ does not activate the safety constraints, then $S(\epsilon_\infty) = 0$ and the suboptimality shrinks to 0 linearly fast as ϵ_2 converges to 0. Otherwise, the gap may decrease according to $S(\epsilon_\infty)$, for which we provide a numerical plot in Section 1.5. In other words, for small estimation errors ϵ_2 and ϵ_∞, applying controller $\widehat{\mathbf{K}}^\star$ (which is solely computed from noisy data) to the *real* plant achieves almost optimal closed-loop performance while guaranteeing compliance with safety constraints. Surprisingly, despite the additional complexity of output-feedback and output noise, our bound matches the scaling with respect to $\epsilon = \max(\epsilon_2, \epsilon_\infty)$ that has been derived in [6] for the state-feedback case without measurement noise.

To prove the above statements, we first characterize a feasible solution to problem (1.30), which we later exploit to establish our suboptimality bound. The proof of Lemma 1.3 and Theorem 1.2 is reported in the Sections 1.6.6 and 1.6.7, respectively.

Lemma 1.3 (Feasible solution)**.** *Let* $\eta = \epsilon_2 \left\|\mathbf{\Phi}^\star_{uy}\right\|_2$ *and* $\zeta = \epsilon_\infty \left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty$*. Assume that the estimation errors are small enough to guarantee* $\eta < \frac{1}{5}$ *and* $\zeta < \frac{1}{2}$*, and select* $\alpha \in [\sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)}, \epsilon_2^{-1})$*. Consider the following optimization problem and its optimal solutions* $\mathbf{\Phi}^c$*:*

$$\mathbf{\Phi}^c \in \arg\min_{\mathbf{\Phi}} \quad \left\| \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} & \mathbf{\Phi}_{yy}\mathbf{y}_0 \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} & \mathbf{\Phi}_{uy}\mathbf{y}_0 \end{bmatrix} \right\|_F \tag{1.31}$$

$$\text{subject to} \begin{bmatrix} I & -\mathbf{G} \end{bmatrix} \mathbf{\Phi} = \begin{bmatrix} I & 0 \end{bmatrix}, \quad \mathbf{\Phi} \begin{bmatrix} -\mathbf{G} \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

$$\left\|\mathbf{\Phi}_{uy}\right\|_2 \leq \left\|\mathbf{\Phi}^\star_{uy}\right\|_2 , \left\|\mathbf{\Phi}_{uy}\right\|_\infty \leq \left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty ,$$

$$\phi_{1,j}(\mathbf{\Phi}) + \phi_{2,j}(\mathbf{\Phi}) + \phi_{3,j}(\mathbf{\Phi}) \leq \mathbf{b}_{y,j} , \tag{1.32}$$

$$\phi_{4,j}(\mathbf{\Phi}) + \phi_{5,j}(\mathbf{\Phi}) + \phi_{6,j}(\mathbf{\Phi}) \leq \mathbf{b}_{u,j} , \tag{1.33}$$

$$\forall j = 1, \ldots, sN ,$$

$$\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu} \text{ with causal sparsities} .$$

where

$$\begin{aligned}
\phi_{1,j}(\mathbf{\Phi}) &= \frac{v_\infty \left\|F_{y,j}\mathbf{\Phi}_{yy}\right\|_1^\star}{1-2\zeta}, \quad \phi_{4,j}(\mathbf{\Phi}) = \frac{v_\infty \left\|F_{u,j}\mathbf{\Phi}_{uy}\right\|_1^\star}{1-2\zeta}, \\
\phi_{2,j}(\mathbf{\Phi}) &= w_\infty \left\| \begin{bmatrix} (F_{y,j}\mathbf{\Phi}_{yu})^\mathsf{T} \\ 2\frac{\epsilon_\infty + \zeta \|\widehat{\mathbf{G}}\|_\infty}{1-2\zeta} (F_{y,j}\mathbf{\Phi}_{yy})^\mathsf{T} \end{bmatrix} \right\|_1 , \\
\phi_{5,j}(\mathbf{\Phi}) &= w_\infty \left\| \begin{bmatrix} (F_{u,j}\mathbf{\Phi}_{uu})^\mathsf{T} \\ 2\frac{\epsilon_\infty + \zeta \|\widehat{\mathbf{G}}\|_\infty}{1-2\zeta} (F_{u,j}\mathbf{\Phi}_{uy})^\mathsf{T} \end{bmatrix} \right\|_1 , \\
\phi_{3,j}(\mathbf{\Phi}) &= F_{y,j}\mathbf{\Phi}_{yy}\widehat{\mathbf{y}}_0 + 2\frac{\epsilon_\infty + \zeta \left\|\widehat{\mathbf{y}}_0\right\|_\infty}{1-2\zeta} \left\|F_{y,j}\mathbf{\Phi}_{yy}\right\|_1^\star , \\
\phi_{6,j}(\mathbf{\Phi}) &= F_{u,j}\mathbf{\Phi}_{uy}\widehat{\mathbf{y}}_0 + 2\frac{\epsilon_\infty + \zeta \left\|\widehat{\mathbf{y}}_0\right\|_\infty}{1-2\zeta} \left\|F_{u,j}\mathbf{\Phi}_{uy}\right\|_1^\star .
\end{aligned}$$

Then, the following expressions

$$\begin{aligned}
&\widetilde{\mathbf{\Phi}}_{yy} = \mathbf{\Phi}^c_{yy}(I+\mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1}, \ \widetilde{\mathbf{\Phi}}_{yu} = \mathbf{\Phi}^c_{yy}(I+\mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1}(\mathbf{G}-\mathbf{\Delta}), \\
&\widetilde{\mathbf{\Phi}}_{uy} = \mathbf{\Phi}^c_{uy}(I+\mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1}, \ \widetilde{\mathbf{\Phi}}_{uu} = (I+\mathbf{\Phi}^c_{uy}\mathbf{\Delta})^{-1}\mathbf{\Phi}^c_{uu}, \\
&\widetilde{\gamma} = \frac{\sqrt{2}\eta}{\epsilon_2(1-\eta)}, \qquad \widetilde{\tau} = \frac{\zeta}{\epsilon_\infty(1-\zeta)} ,
\end{aligned} \tag{1.34}$$

provide a feasible solution to problem (1.30)*.*

The main idea behind Lemma 1.3 is to construct a feasible solution to problem (1.30) from the set of closed-loop responses generated applying a cautious ground-truth optimal controller $\mathbf{K}^c = \mathbf{\Phi}^c_{uy}(\mathbf{\Phi}^c_{yy})^{-1}$ on the estimated system $\widehat{\mathbf{G}}$. In the absence of safety constraints, such a feasible solution could directly be established from the ground-truth optimal policy

$\mathbf{K}^\star$ similar to [37]. In the constrained case, however, one cannot expect the optimal solution $\mathbf{K}^\star$ to be feasible for $\widehat{\mathbf{G}}$ in (1.30) since (1.26)-(1.27) are more stringent than (1.14)-(1.15). Hence, in (1.31) we first compute $\mathbf{K}^c = \mathbf{\Phi}^c_{uy}(\mathbf{\Phi}^c_{yy})^{-1}$ as the optimal linear policy for the real system $\mathbf{G}$ under safety constraints that are more stringent than those of (1.30), and subsequently define $\widetilde{\mathbf{\Phi}}$ as the closed-loop responses generated applying $\mathbf{K}^c$ to $\widehat{\mathbf{G}}$. In this way, $\widetilde{\mathbf{\Phi}}$ is guaranteed to be feasible for (1.30), provided that the model mismatch is sufficiently small.

Clearly, the optimal solution $\mathbf{K}^c = \mathbf{\Phi}^c_{uy}(\mathbf{\Phi}^c_{yy})^{-1}$ to (1.31) will yield a suboptimal cost $J(\mathbf{G}, \mathbf{K}^c) \geq J(\mathbf{G}, \mathbf{K}^\star)$. We denote the corresponding suboptimality gap as

$$S(\epsilon_\infty) = \frac{J(\mathbf{G}, \mathbf{K}^c)^2 - J(\mathbf{G}, \mathbf{K}^\star)^2}{J(\mathbf{G}, \mathbf{K}^\star)^2}\,. \tag{1.35}$$

Note that, if the estimation error ϵ_∞ is too large, the optimization problem (1.31) may become infeasible. This is expected as the uncertainty level might be incompatible with the required safety. On the other hand, if the optimal solution to the non-noisy problem (1.17) does not activate the safety constraints, then the constraints of (1.31) remain inactive for small enough ϵ_∞. In such case we have that $S(\epsilon_\infty) = 0$.

We are now ready to state the main suboptimality result.

Theorem 1.2. *Let $\eta = \epsilon_2 \left\|\mathbf{\Phi}^\star_{uy}\right\|_2$ and $\zeta = \epsilon_\infty \left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty$. Assume that the estimation errors are small enough to guarantee $\eta < \frac{1}{5}$ and $\zeta < \frac{1}{2}$, and select $\alpha \in \left[\sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)}, 5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2\right]$. Moreover, assume that ϵ_∞ is small enough for the optimization problem* (1.31) *to be feasible. Then, when applying the controller $\widehat{\mathbf{K}}^\star$ optimizing* (1.28) *to the true plant $\mathbf{G}$, the relative error with respect to the true optimal cost is upper bounded as*

$$\begin{aligned}\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}} &\leq 20\eta + 4(M^c + V^c) + 4S(\epsilon_\infty)(1 + M^c + V^c)\\ &= \mathcal{O}\left(\epsilon_2\left(1 + \left\|\mathbf{\Phi}^\star_{uy}\right\|_2\right)(1 + \|\mathbf{G}\|_2 + \|\mathbf{y}_0\|_2)^2\right) + 4S(\epsilon_\infty)(1 + M^c + V^c)\,,\end{aligned} \tag{1.36}$$

where

$$\begin{aligned}M^c &= h(\epsilon_2, \alpha, \widehat{\mathbf{G}}) + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\|\mathbf{\Phi}^c_{uy}\right\|_2, \mathbf{G}) + h(\epsilon_2, \left\|\mathbf{\Phi}^c_{uy}\right\|_2, \mathbf{y}_0)\,,\\ V^c &= h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\|\mathbf{\Phi}^c_{uy}\right\|_2, \mathbf{y}_0)\,.\end{aligned}$$

We have expressed the suboptimality gap in the form (1.36) to highlight the presence of two main parts; the first addend scales as $\mathcal{O}\left(\epsilon_2\left(1 + \left\|\mathbf{\Phi}^\star_{uy}\right\|_2\right)(1 + \|\mathbf{G}\|_2 + \|\mathbf{y}_0\|_2)^2\right)$ and the second addend $S(\epsilon_\infty)(1 + M^c + V^c)$ is linked to the suboptimality of the tightened optimization program (1.31). The most important observation is that the suboptimality decreases at most *linearly* with ϵ_2 when $\max(\epsilon_2, \epsilon_\infty)$ is small enough. A linear suboptimality rate in the output-feedback case has first been observed for the unconstrained setup of

[37]. Recovering a similar suboptimality rate for the general case with constraints is one of the main novelties of this chapter. Indeed, despite recovering an upper bound that scales similarly to [37], the corresponding analysis in Section 1.6.7 is significantly complicated by the fact that the feasible solution used in [37] cannot be exploited anymore. Hence, one might expect that the suboptimality rate will worsen with respect to the unconstrained case of [37]. Theorem 1.2 shows, however, that the bound does not deteriorate for small-enough model mismatch levels. Turning our attention to the term $S(\epsilon_\infty)$, we observe through examples (cfr. Figure 1.2) that $S(\epsilon_\infty)$ sharply transitions from 0 to ∞ as ϵ_∞ increases. In practice, this example suggests that $S(\epsilon_\infty)$ might be interpreted as an indicator function; if $S(\epsilon_\infty) \approx 0$, then ϵ_∞ is small enough for the linear suboptimality rate to hold.

Our suboptimality bound (1.36) indicates features of the underlying unknown system that make it easier to be safely controlled based on noisy data. Notably, the suboptimality grows quadratically with the norm of the true impulse and free responses. This fact implies that an unknown unstable system will be more difficult to control for a long horizon. Last, we note that, surprisingly, our rate in terms of ϵ_2 matches that of [6] which was valid under the assumption of exact state measurements. In other words, our analysis shows that near-optimality can be ensured in complex data-driven control scenarios that combine hard safety requirements with noisy output measurements.

1.5 Numerical Experiments

In this section, we demonstrate numerically the effectiveness of the proposed framework in safely controlling unknown systems. In the experiments, we consider the single-input single-output unknown LTI system characterized by the matrices

$$A = \rho \begin{bmatrix} 1 & 0.25 \\ 0 & 1 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}, \; C = \begin{bmatrix} 1 & -1 \end{bmatrix}, \tag{1.37}$$

where $\rho > 0$ corresponds to the spectral radius of A. When $\rho < 1$, (1.37) is asymptotically stable, that is, its output converges to the origin at an exponential rate when the input is equal to 0. When $\rho = 1$, (1.37) is a marginally stable double-integrator system.

In all the following tests, the cost function is given by (1.3) for appropriate choices of the weights. The expectation in (1.3) is taken over future input/output disturbances with covariance matrices $\Sigma_w = I_m$ and $\Sigma_v = I_p$. We consider bounded disturbances between -1 and 1, that is, $w_\infty = v_\infty = 1$. Hence, each scalar disturbance is randomly chosen from $\{-1, 1\}$ with probability $\frac{1}{2}$. For solving optimization problems we use MOSEK [78], called through MATLAB via YALMIP [79].

1.5.1 Example: safe controller synthesis from noisy data

In our first test, we synthesize a safe output-feedback controller for system (1.37) with $\rho = 1$ from noisy data. We assume that $x_0 = x(1) = \begin{bmatrix} 6 & 0 \end{bmatrix}^\mathsf{T}$, where we set the initial time at $t = 1$ rather than $t = 0$ for compliance with MATLAB's indexing of vector entries.

The safety constraints are: $y(1) \in \mathbb{R}$ and

$$\begin{aligned} -5.5 \leq y(t) \leq 5.5, &\quad \forall t = 2, \dots, 12\,, \\ -100 \leq u(t) \leq 100\,, &\quad \forall t = 1, \dots, 11\,, \end{aligned}$$

for all realizations of noise $\|\mathbf{w}\|_\infty, \|\mathbf{v}\|_\infty \leq 1$, while minimizing the cost (1.3) with the weights $\mathbf{Q}$ and $\mathbf{R}$ in (1.17) set to the identity.

We first synthesize the optimal controller assuming that the available data are not affected by noise. To this end, we cast and solve the convex optimization problem (1.40). We verify that the optimal controller $\mathbf{K}^\star$ yields a cost $J(\mathbf{G}, \mathbf{K}^\star) = 69.88$. The green tubes in Figure 1.1 show the regions containing 50 realizations of the optimal closed-loop input and output trajectories. Due to the high level of noise, we can observe a significant variability in the trajectory values for different noise realizations. Nonetheless, all trajectories are safe.

We then discuss the case where the available data are affected by noise. In order for the tightened constraints of (1.30) to be feasible, we consider noisy estimates $(\widehat{\mathbf{G}}, \widehat{\mathbf{y}}_0)$ with $\epsilon = 0.01$ and compute a near-optimal solution to the proposed optimization problem (1.28). As discussed in Section 1.3.2, this can be achieved by 1) extensive or random search over γ and τ, or 2) extensive search over τ and golden-section search over γ. Even if the first solution comes without strong theoretical guarantees, extensive search over γ and τ may be simpler to implement as it avoids the delicate task of tuning the parameter α. Specifically, for this example we have searched over 100 randomly extracted values of γ and τ in the interval $[0, \epsilon^{-1})$. A potential improvement to this heuristic could be to use a bisection algorithm, as proposed in [73] for example.

Proceeding as above, we synthesize a robustly safe controller $\widehat{\mathbf{K}}^\star$ yielding a cost of $J(\mathbf{G}, \widehat{\mathbf{K}}^\star) = 140.54$. The corresponding suboptimality gap is $\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}} = 3.049$. In Figure 1.1, the trajectories and variability levels resulting from $\widehat{\mathbf{K}}^\star$ for 50 noise realizations are plotted in blue. We observe that, since $\widehat{\mathbf{K}}^\star$ is synthesized using noise-corrupted data, it leads to safer, but more conservative trajectories. Indeed, due to uncertainty, higher control effort is spent to keep the output further from the constraints.

It is informative to inspect the robust suboptimality gap $S(\epsilon_\infty)$ incurred by the tightened optimization problem (1.31) that we have used in the analysis to characterize a feasible solution to (1.30). In Figure 1.2, we plot $S(\epsilon_\infty)$ assuming $x_0 = x(1) = \begin{bmatrix} 1 & 0 \end{bmatrix}^\mathsf{T}$ and requiring $-3 \leq y(t) \leq 3$ for $t = 1, \dots, 7$. The example exhibits a fast transition from infeasibility for $\epsilon_\infty > 0.118$ to near-optimality for $\epsilon_\infty < 0.115$. This fact leads to the following observation: high-performing safe controllers can be synthesized by solving (1.30) even when the optimization problem (1.31) is infeasible, i.e. $S(\epsilon_\infty) = \infty$. In such cases the suboptimality bound (1.36) is not applicable, but a robustly safe controller has been synthesized nonetheless. This phenomenon is consistent with the numerical examples of [6] for the state-feedback case.

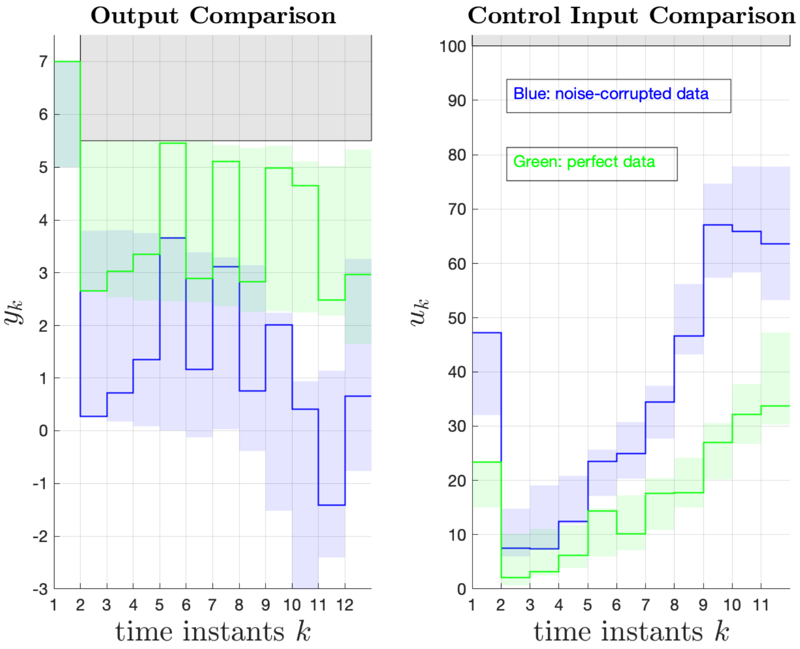

Figure 1.1: Closed-loop trajectories. The grey region indicates unsafe input and output values. The green and blue regions contain trajectories for 50 noise realizations obtained through $\mathbf{K}^\star$ and $\widehat{\mathbf{K}}^\star$, respectively. Green and blue lines represent a specific trajectory in both settings.

1.5.2 Example: suboptimality scaling beyond least-squares estimation

The bound (1.36) in Theorem 1.2 states that a low estimation error level ϵ is crucial in ensuring safety and near-optimality when controlling unknown systems based on noisy data. One advantage of the proposed formulation is that it is directly compatible with behavioral estimation approaches beyond LS identification for the reconstruction of the impulse and free responses, such as data-enabled Kalman filtering [51] and Signal Matrix Model (SMM) [11, 80] . In our last test, we drop the constraints for both the input and the outputs thus putting our focus on 1) validating the linear scaling of the suboptimality gap (1.36), and 2) showcasing that, for instance, SMM-based estimation [80] may lead to significantly lower error levels given the same amount of data. We consider the system (1.37) with different values of $\rho \in [0.9, 0.93, 0.96, 0.99, 1]$ and $x_0 = x(1) = \begin{bmatrix} 6 & 0 \end{bmatrix}^\mathsf{T}$, over a time-horizon of length $N = 11$. The cost function weights in (1.17) are selected as $Q(t) = I_p$ for every $t = 1, \dots, 11$,

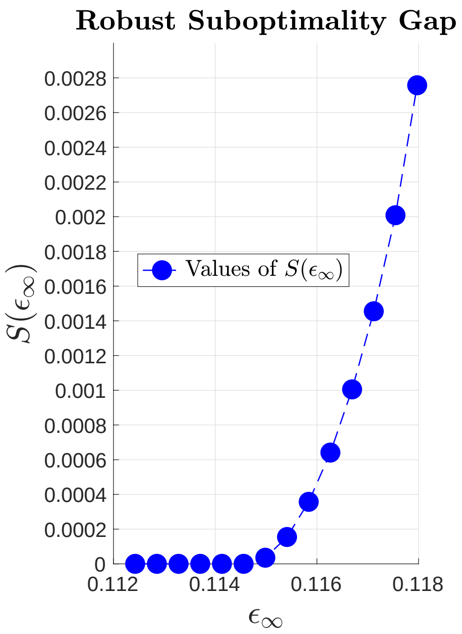

Figure 1.2: Robust suboptimality gap $S(\epsilon_\infty)$. This quantity can be interpreted as an indicator as to whether the guarantee (1.36) holds for a given ϵ_∞.

$Q(12) = 20I_p$ and $R(t) = 0.05$ for every $t = 1, \dots, 11$.

Behavioral estimation: LS vs SMM

For a fixed value of ρ, we gather system trajectories of length 200 time-steps which are corrupted by input and output Gaussian noise with covariance matrices equal to σI. For each experiment, we fix the variance $\sigma \geq 0$ and select a random exploration control input $\mathbf{u}$. We collect 1000 different trajectories for different realizations of the corrupting noise. For each realization of the trajectories, we compute 1) the LS solution (G_{LS}, g_{LS}) using (1.41) and the corresponding impulse and free responses $\widetilde{\mathbf{G}}_{LS}, \widetilde{\mathbf{y}}_{0,LS}$, and 2) the ML solution (G_{ML}, g_{ML}) using (1.42)-(1.43) and the corresponding impulse and free responses $\widetilde{\mathbf{G}}_{ML}, \widetilde{\mathbf{y}}_{0,ML}$. For each estimation, we determine the incurred error levels $\epsilon_{2,G}$, $\epsilon_{\infty,G}$, $\epsilon_{2,y}$ and $\epsilon_{\infty,y}$[7]. Last, we record the 90-th percentile of these values, both for SMM and LS estimation.

In Figure 1.3 we compare the values of ϵ_2 and ϵ_∞ incurred by both estimation techniques. We observe that SMM may yield significantly smaller estimation errors than LS identification. While a full sample-complexity analysis is still unavailable beyond least-squares [3, 7, 54], these examples showcase an advantage in using more sophisticated estimation techniques for safe data-driven control.

[7] Since the real system is unavailable, in practice this can be done using a bootstrap procedure.

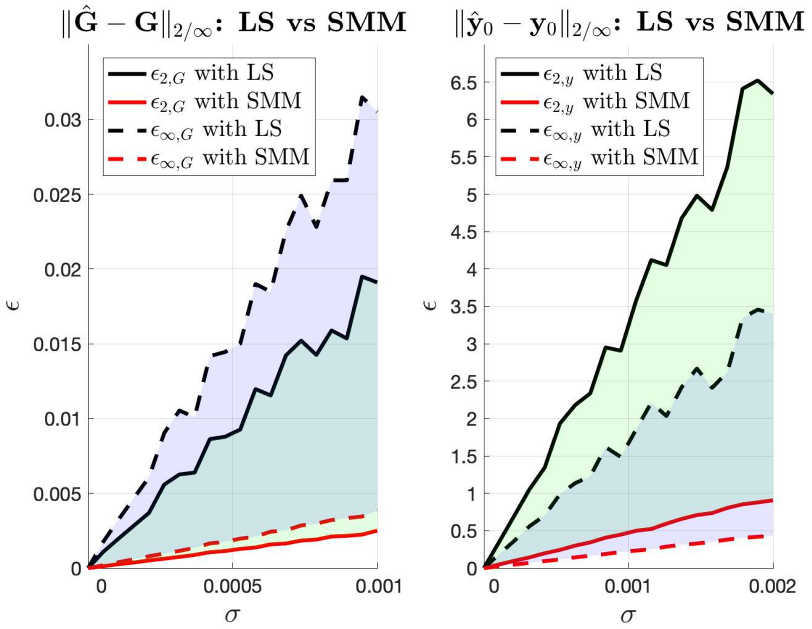

Figure 1.3: Estimation error in function of the corrupting noise. ML estimation through the SMM yields significantly smaller errors than LS. The green and blue regions indicate the gap for the 2-norm and ∞-norm, respectively.

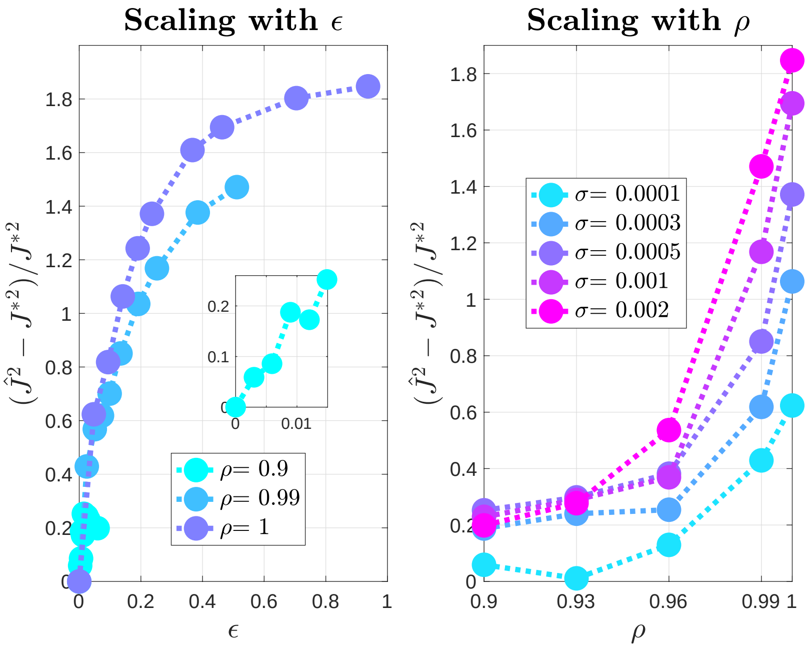

Figure 1.4: Suboptimality gap as a function of ϵ_2 (obtained through SMM estimation) for increasing values of the spectral radius ρ of matrix A (on the left). Suboptimality gap as a function of ρ for increasing values of σ (on the right).

Suboptimality scaling

Having exploited ML estimation to construct approximate impulse and free responses and the corresponding error-levels, we are ready to solve the optimization problem (1.30). Since constraints are not present in this example, (1.30) can be simplified to the quasiconvex formulation we have proposed in [12], where the optimization variable τ is not present. The parameter α is tuned empirically in the interval $\alpha \in [\sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)}, \epsilon_2^{-1})$.[8]

Figure 1.4 shows the suboptimality gap one incurs by applying the controller $\widehat{\mathbf{K}}^\star$ obtained through the proposed approach. On the left, we consider increasing levels of the estimation error level ϵ_2 for each choice of the spectral radius $\rho = 0.9, 0.99, 1$. On the right, we conversely consider increasing levels of the spectral radius for each choice of the estimation error level ϵ_2. In both cases, we plot the suboptimality gap $\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}}$. It can be observed that, consistently with Theorem 1.2, 1) the gap linearly converges to 0 as ϵ_2 converges to 0, and 2) the gap may grow faster than linearly with the spectral radius ρ as a larger ρ generally leads to larger $\|\mathbf{G}\|_2$. We also observe that larger ρ may lead to higher model mismatch values ϵ. Finally, we remark that, in finite-horizon, our formulations are valid for unstable systems with $\rho > 1$. However, it is inherently challenging to collect trajectories of an unstable system, as the values to be plugged into the corresponding optimization problems will become too large to be handled by numerical solvers. For unstable systems in a data-driven scenario, it is common to assume knowledge of a pre-stabilizing controller [38, 37].

1.6 Appendices

1.6.1 Willems' lemma and behavioral theory for synthesizing safe controllers

We recall the definition of persistency of excitation and the result known as the *Fundamental Lemma* for LTI systems [81].

Definition 1.1. *We say that* $\mathbf{u}^h_{[0,T-1]}$ *is* persistently exciting *(PE) of order L if the Hankel matrix* $\mathcal{H}_L(\mathbf{u}^h_{[0,T-1]})$ *is full row-rank.*

A necessary condition for the matrix $\mathcal{H}_L(\mathbf{u}^h_{[0,T-1]})$ to be full row-rank is that it has at least as many columns as rows. It follows that the input trajectory $\mathbf{u}^h_{[0,T-1]}$ must be long enough to satisfy $T \geq (m+1)L - 1$.

Lemma 1.4 (Theorem 3.7, [81]). *Consider system* (1.1). *Assume that* (A, B) *is controllable and that there is no noise. Let* $\{\mathbf{y}^h_{[0,T-1]}, \mathbf{u}^h_{[0,T-1]}\}$ *be a system trajectory of length T that has been recorded during a past experiment. Then, if* $\mathbf{u}^h_{[0,T-1]}$ *is PE of order* $n + L$*, the*

[8] The value $\eta = \epsilon_2 \left\|\mathbf{\Phi}^\star_{uy}\right\|_2$ is unknown in practice because $\mathbf{\Phi}^\star_{uy}$ is unavailable. One can then tune α according to $\alpha < \epsilon_2^{-1}$.

signals $\mathbf{y}^{\star}_{[0,L-1]} \in \mathbb{R}^{pL}$ *and* $\mathbf{u}^{\star}_{[0,L-1]} \in \mathbb{R}^{mL}$ *are trajectories of* (1.1) *if and only if there exists* $g \in \mathbb{R}^{T-L+1}$ *such that*

$$\begin{bmatrix} \mathcal{H}_L(\mathbf{y}^h_{[0,T-1]}) \\ \mathcal{H}_L(\mathbf{u}^h_{[0,T-1]}) \end{bmatrix} g = \begin{bmatrix} \mathbf{y}^{\star}_{[0,L-1]} \\ \mathbf{u}^{\star}_{[0,L-1]} \end{bmatrix} . \tag{1.38}$$

We proceed by showing how Lemma 1.4 allows one to derive a data-driven formulation of (1.17) when the data are not noisy. We work under the following assumptions that are standard in the behavioral framework.

Assumption 1.2. *The data-generating LTI system* (1.1) *is such that* (A, B) *is controllable and* (A, C) *is observable.*

Assumption 1.3. *The historical input trajectory* $\mathbf{u}^h_{[0,\tilde{T}-1]}$ *is PE of order* $n + T_{ini} + N$*, where* $T_{ini} \geq l$ *and* l *is the smallest integer such that the matrix*

$$\begin{bmatrix} C^{\mathsf{T}} & (CA)^{\mathsf{T}} & \dots & (CA^{l-1})^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} ,$$

has full row-rank. Note that if Assumption 1.2 holds, then $l \leq n$.

Further, we give the following definition.

Definition 1.2. *The available data in* **D1** *are further split as follows:*

i) a recent *system trajectory of length* T_{ini}*:* $\left\{\mathbf{y}^r_{[0,T_{ini}-1]}, \mathbf{u}^r_{[0,T_{ini}-1]}\right\}$*, with* $\mathbf{y}^r_{[0,T_{ini}-1]} = \mathbf{y}_{[-T_{ini},-1]}$ *and* $\mathbf{u}^r_{[0,T_{ini}-1]} = \mathbf{u}_{[-T_{ini},-1]}$*,*

ii) a historical *system trajectory of length* $\tilde{T}$*:* $\left\{\mathbf{y}^h_{[0,\tilde{T}-1]}, \mathbf{u}^h_{[0,\tilde{T}-1]}\right\}$*, with* $\mathbf{y}^h_{[0,\tilde{T}-1]} = \mathbf{y}_{[-T_h,-T_h+\tilde{T}-1]}$ *and* $\mathbf{u}^h_{[0,\tilde{T}-1]} = \mathbf{u}_{[-T_h,-T_h+\tilde{T}-1]}$ *for* $T_h \in \mathbb{N}$ *such that* $T_h \geq \tilde{T}$ *and* $\tilde{T} \leq T$.

The *historical* data are to be used in substitution of the system model, while the *recent* data reflect the system initial state $x_0 \in \mathbb{R}^n$ [82]. By exploiting (1.38), one can derive a constrained version of the BIOP derived in [12] as follows:

Proposition 1.5 (Safe Behavioral IOP). *Consider the LTI system* (1.1)*, whose parameters* (A, B, C, x_0) *are* unknown*, and let Assumptions 1.2-1.3 hold. Further assume that the historical and recent trajectories are* not *affected by noise. Let* (G, g) *be any solutions to the linear system of equations*

$$\begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix} \begin{bmatrix} G & g \end{bmatrix} = \begin{bmatrix} 0_{mT_{ini}\times m} & \mathbf{u}^r_{[0,T_{ini}-1]} \\ 0_{pT_{ini}\times m} & \mathbf{y}^r_{[0,T_{ini}-1]} \\ \begin{bmatrix} I_m & 0_{m\times m(N-1)} \end{bmatrix}^{\mathsf{T}} & 0_{mN\times 1} \end{bmatrix} , \tag{1.39}$$

where $\begin{bmatrix} U_p \\ U_f \end{bmatrix} = \mathcal{H}_{T_{ini}+N}(\mathbf{u}^h_{[0,\tilde{T}-1]})$ *and* $\begin{bmatrix} Y_p \\ Y_f \end{bmatrix} = \mathcal{H}_{T_{ini}+N}(\mathbf{y}^h_{[0,\tilde{T}-1]})$. *Then, the optimization problem* (1.17) *is equivalent to*

$$
\begin{aligned}
\min_{\mathbf{\Phi}} \ & \left\| \begin{bmatrix} \mathbf{Q}^{\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_v^{\frac{1}{2}} & 0 & Y_f g \\ 0 & \mathbf{\Sigma}_w^{\frac{1}{2}} & 0 \end{bmatrix} \right\|_F^2 \qquad (1.40)\\
\text{subject to} \ & \begin{bmatrix} I & -\operatorname{Toep}(Y_f G) \end{bmatrix} \mathbf{\Phi} = \begin{bmatrix} I & 0 \end{bmatrix} , \\
& \mathbf{\Phi} \begin{bmatrix} -\operatorname{Toep}(Y_f G) \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix} , \\
& \left\| \begin{bmatrix} v_\infty \left(F_{y,j}\mathbf{\Phi}_{yy}\right)^\mathsf{T} \\ w_\infty \left(F_{y,j}\mathbf{\Phi}_{yu}\right)^\mathsf{T} \end{bmatrix}^\mathsf{T} \right\|_1 + \left(F_{y,j}\mathbf{\Phi}_{yy}\right) Y_f g \leq \mathbf{b}_{y,j} , \\
& \left\| \begin{bmatrix} v_\infty \left(F_{u,j}\mathbf{\Phi}_{uy}\right)^\mathsf{T} \\ w_\infty \left(F_{u,j}\mathbf{\Phi}_{uu}\right)^\mathsf{T} \end{bmatrix}^\mathsf{T} \right\|_1 + \left(F_{u,j}\mathbf{\Phi}_{uy}\right) Y_f g \leq \mathbf{b}_{u,j} , \\
& \forall j = 1, \ldots, sN , \\
& \mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu} \ \textit{with causal sparsities.}
\end{aligned}
$$

The proof of Proposition 1.5 is analogous to that of Theorem 1 in [12], with the addition of the safety constraints as per Proposition 1.1. Since the historical and recent data are not noisy, $Y_f G$ and $Y_f g$ yield the *true* impulse response matrix $\mathbf{G}$ and free response $\mathbf{y}_0$ and the optimal solution of (1.40) recovers the optimal safe controller $\mathbf{K}^\star$ for the real system.

In practice, exact historical and recent data are not available. As per the noise model in the dynamics (1.1)-(1.2), one may assume that historical and recent trajectories are affected by additive noise $w^h(t), w^r(t), v^h(t), v^r(t)$[9] at all time instants, with zero expected values and variances $\mathbf{\Sigma}^h_w, \mathbf{\Sigma}^r_w, \mathbf{\Sigma}^h_v, \mathbf{\Sigma}^r_v$ respectively. Hence, the matrix on the left-hand-side of (1.39) becomes full row-rank almost surely, and any solution to (1.39) leads to potentially different estimates of the system free and impulse responses, which do not necessarily match the exact ones. This issue is well-known in the behavioral theory literature, and several mitigation strategies have recently been proposed [46, 47, 49, 11, 51, 80]. For instance, a behavioral LS estimator akin to the impulse-response identification of [7, 37] is given by

$$
\begin{bmatrix} G_{LS} & g_{LS} \end{bmatrix} = \begin{bmatrix} \hat{U}_p \\ \hat{Y}_p \\ \hat{U}_f \end{bmatrix}^+ \begin{bmatrix} 0_{mT_{ini}\times m} & \mathbf{u}^r_{[0,T_{ini}-1]} \\ 0_{pT_{ini}\times m} & \mathbf{y}^r_{[0,T_{ini}-1]} \\ \begin{bmatrix} I_m & 0_{m\times m(N-1)} \end{bmatrix}^\mathsf{T} & 0_{mN\times 1} \end{bmatrix} , \qquad (1.41)
$$

[9] where "w" and "v" denote input and output noise, respectively, and the apices r and h denote recent data and historical data, respectively.

while the ML estimator [11] is computed as

$$\begin{aligned} G_{ML} = \arg\min_{G} \quad & -\,log\left[p\left(\begin{bmatrix}\Xi_y \\ Y_f G\end{bmatrix} \mid G, Y_f\right)\right] && (1.42)\\ \text{subject to} \quad & \begin{bmatrix}\hat{U}_p \\ \hat{U}_f\end{bmatrix} G = \begin{bmatrix} 0_{mT_{ini}\times m} \\ \begin{bmatrix} I_m & 0_{m\times m(N-1)}\end{bmatrix}^{\mathsf{T}}\end{bmatrix}, \\ g_{ML} = \arg\min_{g} \quad & -\,log\left[p\left(\begin{bmatrix}\xi_y \\ Y_f g\end{bmatrix} \mid g, Y_f\right)\right] && (1.43)\\ \text{subject to} \quad & \begin{bmatrix}\hat{U}_p \\ \hat{U}_f\end{bmatrix} g = \begin{bmatrix}\mathbf{u}^r_{[0,T_{ini}-1]} \\ 0_{mN\times 1}\end{bmatrix}, \end{aligned}$$

where the residuals $\Xi_y = (Y_p - \hat{Y}_p)G$ and $\xi_y = (Y_p - \hat{Y}_p)g$ denote the fitting deviation from the most recent output measurements, and $p(a|b)$ indicates the probability of event a conditioned to b.

1.6.2 Proof of Proposition 1.1

For the first statement, notice that the controller $\mathbf{K}$ achieves the closed-loop responses (1.12). Now select $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$ as

$$\begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu}\end{bmatrix} = \begin{bmatrix}(I-\mathbf{GK})^{-1} & (I-\mathbf{GK})^{-1}\mathbf{G} \\ \mathbf{K}(I-\mathbf{GK})^{-1} & (I-\mathbf{KG})^{-1}\end{bmatrix}, \tag{1.44}$$

and $\mathbf{q} = \mathbf{\Phi}_{uu}\mathbf{g}$. Clearly, $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$ and $\mathbf{g} = \mathbf{\Phi}_{uu}^{-1}\mathbf{q}$, and by plugging the corresponding expressions, we verify that (1.13) and (1.16) are satisfied. It remains to prove that (1.14)-(1.15) are satisfied. In (1.10), substitute $\mathbf{y}$ and $\mathbf{u}$ with their closed-loop expressions (1.11). It follows that the addends separately depend on $\mathbf{w}$ or $\mathbf{v}$. Hence, (1.10) can be rewritten as

$$\begin{aligned} &\max_{\|\mathbf{v}\|_\infty \le v_\infty} (\mathbf{F}_y\mathbf{\Phi}_{yy})\,\mathbf{v} + \max_{\|\mathbf{w}\|_\infty \le w_\infty} (\mathbf{F}_y\mathbf{\Phi}_{yu})\,\mathbf{w} + \\ &+ \mathbf{F}_y\mathbf{G}\mathbf{q} + (\mathbf{F}_y\mathbf{\Phi}_{yy})\,\mathbf{C}\mathbf{P}_A(:,0)x_0 \le \mathbf{b}_y\,, \end{aligned} \tag{1.45}$$

$$\begin{aligned} &\max_{\|\mathbf{v}\|_\infty \le v_\infty} (\mathbf{F}_u\mathbf{\Phi}_{uy})\,\mathbf{v} + \max_{\|\mathbf{w}\|_\infty \le w_\infty} (\mathbf{F}_u\mathbf{\Phi}_{uu})\,\mathbf{w} + \\ &+ \mathbf{F}_u\mathbf{q} + (\mathbf{F}_y\mathbf{\Phi}_{uy})\,\mathbf{C}\mathbf{P}_A(:,0)x_0 \le \mathbf{b}_u\,, \end{aligned} \tag{1.46}$$

where the $\max(\cdot)$ is to be intended row-wise. The expressions (1.45)-(1.46) are already convex in $\mathbf{\Phi}, \mathbf{q}$. To have a more explicit expression, similar to [6] we utilize the well-known property that the $\|\cdot\|_1$ and the $\|\cdot\|_\infty$ vector norms are dual of each other [76], that is $k\,\|x\|_1 = \max_{\|w\|_\infty \le k} x^{\mathsf{T}}w$. The result follows immediately by inspecting (1.45)-(1.46) and letting x^{T} be equal to either $F_{y,j}\mathbf{\Phi}_{yy}$, $F_{y,j}\mathbf{\Phi}_{uy}$, $F_{y,j}\mathbf{\Phi}_{yu}$ or $F_{y,j}\mathbf{\Phi}_{uu}$, and letting k be equal to either v_∞ or w_∞.

For the second statement, it is easy to notice that $\mathbf{K}$ is causal by construction because $\mathbf{\Phi}_{uy}$ and $\mathbf{\Phi}_{yy}$ are block lower-triangular. By selecting the controller $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$ one has

$$\begin{aligned}(I - \mathbf{G}\mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1})^{-1} &= (I - \mathbf{G}\mathbf{\Phi}_{uy}(I + \mathbf{G}\mathbf{\Phi}_{uy})^{-1})^{-1}\\ &= ((I + \mathbf{G}\mathbf{\Phi}_{uy} - \mathbf{G}\mathbf{\Phi}_{uy})(I + \mathbf{G}\mathbf{\Phi}_{uy})^{-1})^{-1}\\ &= I + \mathbf{G}\mathbf{\Phi}_{uy} = \mathbf{\Phi}_{yy}\,,\end{aligned}$$

which shows that $\mathbf{\Phi}_{yy}$ is the closed-loop response from $\mathbf{v}_{[0,N-1]} + \mathbf{CP}_A(:,0)x_0$ to $\mathbf{y}_{[0,N-1]}$ as per (1.12). Similar computations hold for the remaining closed-loop responses. For the safety constraints, select any $\mathbf{\Phi}$ and $\mathbf{q}$ complying with (1.14)-(1.15). It is easy to verify by direct computation that, for any $\mathbf{w}$ and $\mathbf{v}$, the same input and output trajectories defined at (1.11) are obtained by letting $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$ and $\mathbf{g} = \mathbf{\Phi}_{uu}^{-1}\mathbf{q}$ in (1.6), (1.7), (1.8). Hence, the safety constraints are satisfied for any disturbance realization.

1.6.3 Proof of Proposition 1.3

We first prove that $\mathbf{K} \in \boldsymbol{\mathcal{K}} \implies \widehat{\mathbf{\Phi}} \in \mathbf{\Pi}$, where

$$\widehat{\mathbf{\Phi}} := \begin{bmatrix}(I - \widehat{\mathbf{G}}\mathbf{K})^{-1} & (I - \widehat{\mathbf{G}}\mathbf{K})^{-1}\widehat{\mathbf{G}}\\ \mathbf{K}(I - \widehat{\mathbf{G}}\mathbf{K})^{-1} & (I - \mathbf{K}\widehat{\mathbf{G}})^{-1}\end{bmatrix}. \tag{1.47}$$

Let us fix $(\mathbf{\Delta}, \boldsymbol{\delta}_0) \in \mathcal{E}$. By substitution of $\widehat{\mathbf{\Phi}}$ inside the blocks of $\mathbf{\Phi}$ defined in the proposition statement, one has $\mathbf{\Phi} = \begin{bmatrix}(I - (\widehat{\mathbf{G}} + \mathbf{\Delta})\mathbf{K})^{-1} & (I - (\widehat{\mathbf{G}} + \mathbf{\Delta})\mathbf{K})^{-1}(\widehat{\mathbf{G}} + \mathbf{\Delta})\\ \mathbf{K}(I - (\widehat{\mathbf{G}} + \mathbf{\Delta})\mathbf{K})^{-1} & (I - \mathbf{K}(\widehat{\mathbf{G}} + \mathbf{\Delta}))^{-1}\end{bmatrix}$. From (1.11)-(1.12) the closed-loop trajectories obtained by applying $\mathbf{K}$ to the system $\widehat{\mathbf{G}} + \mathbf{\Delta}$ are given by

$$\begin{bmatrix}\mathbf{y}(\mathbf{K}, \boldsymbol{\theta})\\ \mathbf{u}(\mathbf{K}, \boldsymbol{\theta})\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu}\\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu}\end{bmatrix}\begin{bmatrix}\mathbf{v} + \widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0\\ \mathbf{w}\end{bmatrix}.$$

Proceeding as in the proof of Proposition 1.1, one can show that "$(\mathbf{y}(\mathbf{K}, \boldsymbol{\theta}), \mathbf{u}(\mathbf{K}, \boldsymbol{\theta})) \in \mathbf{\Gamma}$" for every $\boldsymbol{\theta} \in \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{W}} \times \boldsymbol{\mathcal{V}}$ is the same as "(1.21)-(1.22)" for every $(\mathbf{\Delta}, \boldsymbol{\delta}_0) \in \boldsymbol{\mathcal{E}}$. Since (1.20) and (1.23) are verified by construction, the proof is concluded. Further, for any $(\mathbf{\Delta}, \boldsymbol{\delta}_0)$, the cost of (1.18) achieved by $\mathbf{K}$ is identical to the cost of (1.19) achieved by $\widehat{\mathbf{\Phi}}$ as proven in Proposition 1.2.

Next, we show $\widehat{\mathbf{\Phi}} \in \mathbf{\Pi} \implies \widehat{\mathbf{K}} \in \boldsymbol{\mathcal{K}}$, where $\widehat{\mathbf{K}} := \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1}$. Using (1.20), one can verify that

$$\mathbf{\Phi}_{yy} := \widehat{\mathbf{\Phi}}_{yy}(I - \mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1} = (I - (\widehat{\mathbf{G}} + \mathbf{\Delta})\widehat{\mathbf{K}})^{-1}\,,$$

and similarly, that all other equalities in (1.47) hold by substituting $\widehat{\mathbf{\Phi}}$ with $\mathbf{\Phi}$ and $\mathbf{K}$ with $\widehat{\mathbf{K}}$. Then

$$\begin{bmatrix}\mathbf{y}(\widehat{\mathbf{K}}, \boldsymbol{\theta})\\ \mathbf{u}(\widehat{\mathbf{K}}, \boldsymbol{\theta})\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu}\\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu}\end{bmatrix}\begin{bmatrix}\mathbf{v} + \widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0\\ \mathbf{w}\end{bmatrix}.$$

But "(1.21)-(1.22)" for every $(\mathbf{\Delta}, \boldsymbol{\delta}_0) \in \boldsymbol{\mathcal{E}}$, which hold by definition, imply that $(\mathbf{y}(\widehat{\mathbf{K}}, \boldsymbol{\theta}), \mathbf{u}(\widehat{\mathbf{K}}, \boldsymbol{\theta})) \in \mathbf{\Gamma}$ for every $\boldsymbol{\theta} \in \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{W}} \times \boldsymbol{\mathcal{V}}$ (see the proof of Proposition 1.1). Further, for any $(\mathbf{\Delta}, \boldsymbol{\delta}_0)$, the cost of (1.19) achieved by $\widehat{\mathbf{\Phi}}$ is identical to the cost of (1.18) achieved by $\widehat{\mathbf{K}}$ as proven in Proposition 1.2.

1.6.4 Proof of Lemma 1.1

The objective function in Proposition 1.3 can be written as the square-root of the sum of the square of the Frobenius norms of each of its six blocks. For the upper-left block, since $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_2 \leq \gamma < \epsilon_2^{-1}$ by assumption, we have

$$\begin{aligned}\|\widehat{\mathbf{\Phi}}_{yy}(I - \mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}\|_F &\leq \|\widehat{\mathbf{\Phi}}_{yy}\|_F \left\|\sum_{k=0}^{\infty}(\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^k\right\|_2 \\ &\leq \|\widehat{\mathbf{\Phi}}_{yy}\|_F \sum_{k=0}^{\infty} \left\|\epsilon_2\widehat{\mathbf{\Phi}}_{uy}\right\|_2^k = \|\widehat{\mathbf{\Phi}}_{yy}\|_F \left(1 - \epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2\right)^{-1} ,\end{aligned}$$

where the convergence of the series follows from $\mathbf{\Delta}$ and $\widehat{\mathbf{\Phi}}_{uy}$ having zero-entries diagonal blocks by construction. Similarly,

$$\begin{aligned}\|\widehat{\mathbf{\Phi}}_{uy}(I - \mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}\|_F &\leq \|\widehat{\mathbf{\Phi}}_{uy}\|_F \left(1 - \epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2\right)^{-1} , \\ \|(I - \widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}\widehat{\mathbf{\Phi}}_{uu}\|_F &\leq \|\widehat{\mathbf{\Phi}}_{uu}\|_F \left(1 - \epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2\right)^{-1} .\end{aligned}$$

Next, we have

$$\begin{aligned}&\|\widehat{\mathbf{\Phi}}_{yy}(I - \mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}} + \mathbf{\Delta})\|_F \\ \leq&\|\widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{G}}\|_F + \|\widehat{\mathbf{\Phi}}_{yy}\mathbf{\Delta}\|_F + \left\|\widehat{\mathbf{\Phi}}_{yy}\left(\sum_{k=1}^{\infty}(\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^k\right)(\widehat{\mathbf{G}} + \mathbf{\Delta})\right\|_F \\ \leq&\|\widehat{\mathbf{\Phi}}_{yu}\|_F + \epsilon_2\|\widehat{\mathbf{\Phi}}_{yy}\|_F + \|\widehat{\mathbf{\Phi}}_{yy}\|_F \frac{\epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2(\|\widehat{\mathbf{G}}\|_2 + \epsilon_2)}{1 - \epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2} \\ \leq&\frac{\|\widehat{\mathbf{\Phi}}_{yu}\|_F + \epsilon_2\|\widehat{\mathbf{\Phi}}_{yy}\|_F(2 + \|\widehat{\mathbf{\Phi}}_{uy}\|_2\|\widehat{\mathbf{G}}\|_2)}{1 - \epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2} ,\end{aligned}$$

and therefore, by developing the squares and using that $\left\|\widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{G}}\right\|_F \leq \|\widehat{\mathbf{\Phi}}_{yy}\|_F\|\widehat{\mathbf{G}}\|_2$ we obtain

$$\|\widehat{\mathbf{\Phi}}_{yy}(I - \mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}} + \mathbf{\Delta})\|_F^2 \leq \frac{\left(\|\widehat{\mathbf{\Phi}}_{yu}\|_F^2 + \|\widehat{\mathbf{\Phi}}_{yy}\|_F^2 h(\epsilon_2, \gamma, \widehat{\mathbf{G}})\right)}{(1 - \epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2)^2} .$$

Proceeding analogously, one can also prove that

$$\|\widehat{\mathbf{\Phi}}_{yy}(I - \mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\boldsymbol{y}}_0 + \boldsymbol{\delta}_0)\|_F^2 \leq \frac{1}{(1 - \epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2)^2}\left(\|\widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0\|_F^2 + \|\widehat{\mathbf{\Phi}}_{yy}\|_F^2 h(\epsilon_2, \gamma, \widehat{\mathbf{y}}_0)\right) ,$$

$$\|\widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\boldsymbol{y}}_0+\boldsymbol{\delta}_0)\|_F^2 \leq \frac{1}{(1-\epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2)^2}\left(\|\widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0\|_F^2+\|\widehat{\mathbf{\Phi}}_{uy}\|_F^2 h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)\right).$$

Therefore, combining the above inequalities we finally conclude that

$$J(\mathbf{G},\mathbf{K}) \leq \frac{\left(\left\|\begin{bmatrix}\widehat{\mathbf{\Phi}}_{yy} & \widehat{\mathbf{\Phi}}_{yu} & \widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0\\ \widehat{\mathbf{\Phi}}_{uy} & \widehat{\mathbf{\Phi}}_{uu} & \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0\end{bmatrix}\right\|_F^2 + \|\widehat{\mathbf{\Phi}}_{yy}\|_F^2(h(\epsilon_2,\gamma,\widehat{\mathbf{G}})+\right.}{1-\epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2}$$
$$\frac{\left.+h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0))+\|\widehat{\mathbf{\Phi}}_{uy}\|_F^2 h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)\right)^{\frac{1}{2}}}{1-\epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2}.$$

1.6.5 Proof of Lemma 1.2

By using the fact that for $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ we have that $\left\|\begin{bmatrix}x^\mathsf{T} & y^\mathsf{T}\end{bmatrix}\right\|_1 = \|x^\mathsf{T}\|_1 + \|y^\mathsf{T}\|_1$, the left-hand-sides of (1.21)-(1.22) are each made of three addends. The proof hinges on upper bounding each one of them for a generic $(\mathbf{\Delta},\boldsymbol{\delta_0}) \in \boldsymbol{\mathcal{E}}$. We report the full derivations for the most informative of them. Exploiting Holder's inequality and the relation $\left\|I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty \leq \frac{1}{1-\epsilon_\infty\tau}$, which can be derived by proceeding as in the proof of Lemma 1.1, we have

$$v_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}\right\|_1^\star \leq v_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1^\star\left\|(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}\right\|_\infty \leq \frac{v_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1^\star}{1-\epsilon_\infty\tau},$$

which is equal to $f_{1,j}(\widehat{\mathbf{\Phi}})$. Next, recalling $\widehat{\mathbf{\Phi}}_{yu}=\widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{G}}$,

$$\begin{aligned}
&w_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}}+\mathbf{\Delta})\right\|_1^\star\\
&\leq w_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right\|_1^\star + \max_{\|\mathbf{w}\|_\infty\leq w_\infty}|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\mathbf{\Delta}\mathbf{w}|+\\
&\quad+\max_{\|\mathbf{w}\|_\infty\leq w_\infty}|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}}+\mathbf{\Delta})\mathbf{w}|\\
&\leq w_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right\|_1^\star + w_\infty\epsilon_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1^\star+\\
&\quad+w_\infty\epsilon_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1^\star\left\|\widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}}+\mathbf{\Delta})\right\|_\infty\\
&\leq w_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right\|_1^\star + w_\infty\epsilon_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1^\star\left(1+\tau\frac{\left\|\widehat{\mathbf{G}}\right\|_\infty+\epsilon_\infty}{1-\epsilon_\infty\tau}\right)\\
&= w_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right\|_1^\star + w_\infty\epsilon_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1^\star\left(\frac{1+\tau\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\epsilon_\infty\tau}\right)\\
&= f_{2,j}(\widehat{\mathbf{\Phi}}).
\end{aligned}$$

Lastly, remembering that $\widehat{\mathbf{\Phi}}_{uu} = I + \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{G}}$ and noticing that

$$(I - \widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}\widehat{\mathbf{\Phi}}_{uu} = \widehat{\mathbf{\Phi}}_{uu} + \widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta}(I - \widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}\widehat{\mathbf{\Phi}}_{uu}\,,$$

we have

$$\begin{aligned}
& w_\infty \left\| F_{u,j}(I - \widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}\widehat{\mathbf{\Phi}}_{uu} \right\|_1^\star \\
& \leq w_\infty \left\| F_{u,j}\widehat{\mathbf{\Phi}}_{uu} \right\|_1^\star + w_\infty \epsilon_\infty \frac{\left\| F_{u,j}\widehat{\mathbf{\Phi}}_{uy} \right\|_1^\star}{1 - \epsilon_\infty \left\| \widehat{\mathbf{\Phi}}_{uy} \right\|_\infty} \left(\left\| \widehat{\mathbf{G}} \right\|_\infty \left\| \widehat{\mathbf{\Phi}}_{uy} \right\|_\infty + 1 \right) \\
& \leq w_\infty \left\| F_{u,j}\widehat{\mathbf{\Phi}}_{uu} \right\|_1^\star + w_\infty \epsilon_\infty \left\| F_{u,j}\widehat{\mathbf{\Phi}}_{uy} \right\|_1^\star \frac{1 + \tau \left\| \widehat{\mathbf{G}} \right\|_\infty}{1 - \epsilon_\infty \tau} \\
& = f_{5,j}(\widehat{\mathbf{\Phi}})\,.
\end{aligned}$$

Similar computations allow one to derive the upper bounds for the remaining terms.

1.6.6 Proof of Lemma 1.3

First, it is easy to verify that $\widetilde{\mathbf{\Phi}}$ satisfies the constraints in (1.30); indeed, $\widetilde{\mathbf{\Phi}}$ comprises the closed-loop responses when we apply $\mathbf{K}^c$ to the estimated plant $\widehat{\mathbf{G}}$. Next, we have

$$\begin{aligned}
\left\| \widetilde{\mathbf{\Phi}}_{uy} \right\|_2 &= \left\| \mathbf{\Phi}^c_{uy}(I + \mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1} \right\|_2 \\
&\leq \frac{\left\| \mathbf{\Phi}^c_{uy} \right\|_2}{1 - \epsilon_2 \left\| \mathbf{\Phi}^c_{uy} \right\|_2} \leq \sqrt{2} \frac{\left\| \mathbf{\Phi}^c_{uy} \right\|_2}{1 - \epsilon_2 \left\| \mathbf{\Phi}^c_{uy} \right\|_2} \\
&\leq \sqrt{2} \frac{\left\| \mathbf{\Phi}^\star_{uy} \right\|_2}{1 - \epsilon_2 \left\| \mathbf{\Phi}^\star_{uy} \right\|_2} = \sqrt{2} \frac{\eta}{\epsilon_2(1 - \eta)} = \widetilde{\gamma}\,.
\end{aligned}$$

Since $\alpha \in [\sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)}, \epsilon_2^{-1})$ and $\eta < \frac{1}{5}$, then $\tilde{\gamma} \leq \alpha < \epsilon_2^{-1}$. Hence $\widetilde{\gamma}$ is feasible. Similarly,

$$\begin{aligned}
\left\| \widetilde{\mathbf{\Phi}}_{uy} \right\|_\infty &= \left\| \mathbf{\Phi}^c_{uy}(I + \mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1} \right\|_\infty \\
&\leq \frac{\left\| \mathbf{\Phi}^c_{uy} \right\|_\infty}{1 - \epsilon_\infty \left\| \mathbf{\Phi}^c_{uy} \right\|_\infty} \leq \frac{\left\| \mathbf{\Phi}^\star_{uy} \right\|_\infty}{1 - \epsilon_\infty \left\| \mathbf{\Phi}^\star_{uy} \right\|_\infty} = \frac{\zeta}{\epsilon_\infty(1 - \zeta)} = \widetilde{\tau}\,.
\end{aligned}$$

Since $\zeta < \frac{1}{2}$, then $\tilde{\tau} < \epsilon_\infty^{-1}$ and hence it is a feasible value for τ. It remains to show that $\widetilde{\mathbf{\Phi}}$ satisfies the safety constraints (1.26)-(1.27). We know that $\mathbf{\Phi}^c$ is feasible for (1.31), and hence $\phi_{1,j}(\mathbf{\Phi}^c) + \phi_{2,j}(\mathbf{\Phi}^c) + \phi_{3,j}(\mathbf{\Phi}^c) \leq \mathbf{b}_{y,j}$ and $\phi_{4,j}(\mathbf{\Phi}^c) + \phi_{5,j}(\mathbf{\Phi}^c) + \phi_{6,j}(\mathbf{\Phi}^c) \leq \mathbf{b}_{u,j}$. We conclude the proof by showing that $f_{i,j}(\widetilde{\mathbf{\Phi}}) \leq \phi_{i,j}(\mathbf{\Phi}^c)$ for every $i = 1, \ldots, 6$. We report the full derivations for the most informative terms.

$$f_{1,j}(\widetilde{\mathbf{\Phi}}) = \frac{v_\infty \left\| F_{y,j} \left(\mathbf{\Phi}^c_{yy} - \mathbf{\Phi}^c_{yy}\mathbf{\Delta}\mathbf{\Phi}^c_{uy} \left(I + \mathbf{\Delta}\mathbf{\Phi}^c_{uy} \right)^{-1} \right) \right\|_1^\star}{1 - \epsilon_\infty \widetilde{\tau}}$$

$$
\begin{aligned}
&\leq \frac{v_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star + \frac{v_\infty \epsilon_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star \left\|\mathbf{\Phi}^c_{uy}\right\|_\infty}{1-\epsilon_\infty \left\|\mathbf{\Phi}^c_{uy}\right\|_\infty}}{1-\epsilon_\infty \widetilde{\tau}} \\
&\leq \frac{v_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star + \frac{v_\infty \epsilon_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star \left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty}{1-\epsilon_\infty \left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty}}{1-\epsilon_\infty \widetilde{\tau}} \\
&\leq \frac{v_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star}{1-2\zeta} = \phi_{1,j}(\mathbf{\Phi}^c)\,.
\end{aligned}
$$

Similarly, it is easy to show that $f_{4,j}(\widetilde{\mathbf{\Phi}}) \leq \phi_{4,j}(\mathbf{\Phi}^c)$. Next, recalling (1.34) and observing that

$$
\begin{aligned}
\widetilde{\mathbf{\Phi}}_{yu} &= \mathbf{\Phi}^c_{yu} - \mathbf{\Phi}^c_{yy}\mathbf{\Delta} - \mathbf{\Phi}^c_{yy}\mathbf{\Delta}\mathbf{\Phi}^c_{uy}(I+\mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1}\widehat{\mathbf{G}}\,, \\
\widetilde{\mathbf{\Phi}}_{yy} &= \mathbf{\Phi}^c_{yy} - \mathbf{\Phi}^c_{yy}\mathbf{\Delta}\mathbf{\Phi}^c_{uy}\left(I+\mathbf{\Delta}\mathbf{\Phi}^c_{uy}\right)^{-1}\,,
\end{aligned}
$$

we have

$$
\begin{aligned}
&f_{2,j}(\widetilde{\mathbf{\Phi}}) \\
&\leq w_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yy}(I+\mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1}(\mathbf{G}-\mathbf{\Delta})\right\|_1^\star + \\
&\quad + w_\infty \epsilon_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yy}(I+\mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1}\right\|_1^\star \left(\frac{1+\widetilde{\tau}\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\epsilon_\infty\widetilde{\tau}}\right) \\
&\leq w_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yu}\right\|_1^\star + w_\infty\epsilon_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star\left(1+\frac{\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\epsilon_\infty\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty}\right) + \\
&\quad + w_\infty\epsilon_\infty \frac{\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star\left(\frac{1+\widetilde{\tau}\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\epsilon_\infty\widetilde{\tau}}\right)}{1-\epsilon_\infty\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty} \leq w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yu}\right\|_1^\star + \\
&\quad + w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star\left(\epsilon_\infty + \frac{\zeta\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\zeta} + \frac{\epsilon_\infty + \frac{\zeta\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\zeta}}{\left(1-\frac{\zeta}{1-\zeta}\right)(1-\zeta)}\right) \\
&= w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yu}\right\|_1^\star + 2w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star\frac{(1-\zeta)\left(\epsilon_\infty+\zeta\left\|\widehat{\mathbf{G}}\right\|_\infty\right)}{1-2\zeta} \\
&\leq w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yu}\right\|_1^\star + 2w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1^\star\frac{\left(\epsilon_\infty+\zeta\left\|\widehat{\mathbf{G}}\right\|_\infty\right)}{1-2\zeta} \\
&\leq \phi_{2,j}(\mathbf{\Phi}^c)\,.
\end{aligned}
$$

Similarly, $f_{3,j}(\widetilde{\mathbf{\Phi}}) \leq \phi_{3,j}(\mathbf{\Phi}^c)$ and $f_{6,j}(\widetilde{\mathbf{\Phi}}) \leq \phi_{6,j}(\mathbf{\Phi}^c)$. By only noticing that $\|\mathbf{\Phi}^c_{uu}\|_\infty \leq 1 + \left\|\mathbf{\Phi}^c_{uy}\right\|_\infty\left(\left\|\widehat{\mathbf{G}}\right\|_\infty + \epsilon_\infty\right)$ and that $(1+\zeta)(1-2\zeta) \leq 1-\zeta$ for every $\zeta > 0$, analogous computations lead to $f_{5,j}(\widetilde{\mathbf{\Phi}}) \leq \phi_{5,j}(\mathbf{\Phi}^c)$.

1.6.7 Proof of Theorem 1.2

By denoting as $\widehat{\mathbf{\Phi}}^\star$ the closed-loop responses obtained by applying $\widehat{\mathbf{K}}^\star$ to $\widehat{\mathbf{G}}$, we have by Lemma 1.1 and by $\gamma \leq \alpha$

$$J(\mathbf{G},\widehat{\mathbf{K}}^\star) \leq \frac{1}{1-\epsilon_2\gamma^\star}\left\|\begin{bmatrix}\sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{G}})+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widehat{\mathbf{\Phi}}^\star_{yy} & \widehat{\mathbf{\Phi}}^\star_{yu} & \widehat{\mathbf{\Phi}}^\star_{yy}\widehat{\mathbf{y}}_0\\ \sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widehat{\mathbf{\Phi}}^\star_{uy} & \widehat{\mathbf{\Phi}}^\star_{uu} & \widehat{\mathbf{\Phi}}^\star_{uy}\widehat{\mathbf{y}}_0\end{bmatrix}\right\|_F,$$

where $\gamma^\star$ is optimal for (1.30). By Lemma 1.3, under the assumptions on η,ζ,α we have that $(\widetilde{\gamma},\widetilde{\tau},\widetilde{\mathbf{\Phi}})$ belongs to the feasible set of (1.30). Hence, by suboptimality of any feasible solution:

$$J(\mathbf{G},\widehat{\mathbf{K}}^\star) \leq \frac{1}{1-\epsilon_2\widetilde{\gamma}}\left\|\begin{bmatrix}\sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{G}})+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widetilde{\mathbf{\Phi}}_{yy} & \widetilde{\mathbf{\Phi}}_{yu} & \widetilde{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0\\ \sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widetilde{\mathbf{\Phi}}_{uy} & \widetilde{\mathbf{\Phi}}_{uu} & \widetilde{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0\end{bmatrix}\right\|_F.$$

Using the definition of $\widetilde{\mathbf{\Phi}}$ from Lemma 1.3, we now relate

$$\widetilde{C}=\left\|\begin{bmatrix}\sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{G}})+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widetilde{\mathbf{\Phi}}_{yy} & \widetilde{\mathbf{\Phi}}_{yu} & \widetilde{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0\\ \sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widetilde{\mathbf{\Phi}}_{uy} & \widetilde{\mathbf{\Phi}}_{uu} & \widetilde{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0\end{bmatrix}\right\|_F,$$

to the optimal cost of problem (1.31). Recalling the expressions of M^c and V^c, and similarly to Lemma 1.3,

$$\begin{aligned}\widetilde{C} = \Bigg(&\left\|\begin{bmatrix}\widetilde{\mathbf{\Phi}}_{yy} & \widetilde{\mathbf{\Phi}}_{yu} & \widetilde{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0\\ \widetilde{\mathbf{\Phi}}_{uy} & \widetilde{\mathbf{\Phi}}_{uu} & \widetilde{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0\end{bmatrix}\right\|_F^2 + \\ &+\Big(h(\epsilon_2,\alpha,\widehat{\mathbf{G}})+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)\Big)\left\|\widetilde{\mathbf{\Phi}}_{yy}\right\|_F^2+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)\left\|\widetilde{\mathbf{\Phi}}_{uy}\right\|_F^2\Bigg)^{\frac{1}{2}}\\ \leq\ & \frac{\sqrt{J(\mathbf{G},\mathbf{K}^c)^2+M^c\left\|\mathbf{\Phi}^c_{yy}\right\|_F^2+V^c\left\|\mathbf{\Phi}^c_{uy}\right\|_F^2}}{1-\epsilon_2\left\|\mathbf{\Phi}^c_{uy}\right\|_2}.\end{aligned}$$

Thus, we have established the chain of inequalities

$$J(\mathbf{G},\widehat{\mathbf{K}}^\star)^2 \leq \frac{\widetilde{C}^2}{(1-\epsilon_2\widetilde{\gamma})^2} \leq \frac{\left(J(\mathbf{G},\mathbf{K}^c)^2+M^c\left\|\mathbf{\Phi}^c_{yy}\right\|_F^2+V^c\left\|\mathbf{\Phi}^c_{uy}\right\|_F^2\right)}{(1-\epsilon_2\widetilde{\gamma})^2(1-\epsilon_2\left\|\mathbf{\Phi}^c_{uy}\right\|_2)^2}.$$

Next, notice that, by definition, we have $J(\mathbf{G},\mathbf{K}^c)^2=(S(\epsilon_\infty)+1)J(\mathbf{G},\mathbf{K}^\star)^2$. Recalling that $\left\|\mathbf{\Phi}^c_{uy}\right\|_2 \leq \left\|\mathbf{\Phi}^\star_{uy}\right\|_2$ and $\left\|\mathbf{\Phi}^c_{yy}\right\|_2 \leq \left\|\mathbf{\Phi}^\star_{yy}\right\|_2$, observing that $\eta < \frac{1}{5}$ implies $1-(1+\sqrt{2})\eta \leq 2$, and further noticing that if $M,V>0$, then $Ma^2+Vb^2 \leq (M+V)(a^2+b^2)$, we can establish:

$$\frac{J(\mathbf{G},\widehat{\mathbf{K}}^\star)^2-J(\mathbf{G},\mathbf{K}^\star)^2}{J(\mathbf{G},\mathbf{K}^\star)^2} \leq \left(\frac{1}{(1-\epsilon_2\left\|\mathbf{\Phi}^c_{uy}\right\|_2)^2(1-\epsilon_2\widetilde{\gamma})^2}\right)\times$$

$$
\begin{aligned}
&\times \left(S(\epsilon_\infty) + 1 + \frac{M^c \left\|\mathbf{\Phi}^c_{yy}\right\|_F^2 + V^c \left\|\mathbf{\Phi}^c_{uy}\right\|_F^2}{J(\mathbf{G},\mathbf{K}^\star)^2} \right) - 1 \\
&\leq \left(\frac{1}{(1-\eta)^2(1-\sqrt{2}\frac{\eta}{1-\eta})^2} - 1 + \frac{S(\epsilon_\infty)}{(1-\eta)^2(1-\sqrt{2}\frac{\eta}{1-\eta})^2} \right) + \\
&+ \frac{M^c \left\|\mathbf{\Phi}^c_{yy}\right\|_F^2 + V^c \left\|\mathbf{\Phi}^c_{uy}\right\|_F^2}{(1-\eta)^2(1-\sqrt{2}\frac{\eta}{1-\eta})^2 J(\mathbf{G},\mathbf{K}^\star)^2} \\
&\leq \eta \left(\frac{2(1+\sqrt{2}) - (1+\sqrt{2})^2\eta}{(1-(1+\sqrt{2})\eta)^2} \right) + \frac{S(\epsilon_\infty)}{(1-(1+\sqrt{2})\eta)^2} + \\
&+ \frac{(M^c + V^c) J(\mathbf{G},\mathbf{K}^c)^2}{(1-(1+\sqrt{2})\eta)^2 J(\mathbf{G},\mathbf{K}^\star)^2} \\
&\leq 20\eta + 4(M^c + V^c) + 4S(\epsilon_\infty)(1 + M^c + V^c)\,.
\end{aligned}
$$

Last, we prove that $20\eta + 4(M^c + V^c) = \mathcal{O}\left(\epsilon_2 \left(1 + \left\|\mathbf{\Phi}^\star_{uy}\right\|_2\right)(1 + \|\mathbf{G}\|_2 + \|\mathbf{y}_0\|_2)^2\right)$. First, notice that $M^c + V^c \leq M^\star + V^\star$, where

$$
\begin{aligned}
M^\star &= h(\epsilon_2, \alpha, \widehat{\mathbf{G}}) + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\|\mathbf{\Phi}^\star_{uy}\right\|_2, \mathbf{G}) + h(\epsilon_2, \left\|\mathbf{\Phi}^\star_{uy}\right\|_2, \mathbf{y}_0)\,, \\
V^\star &= h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\|\mathbf{\Phi}^\star_{uy}\right\|_2, \mathbf{y}_0)\,.
\end{aligned}
$$

Using $\alpha \leq 5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2$, $\eta < \frac{1}{5}$, $\left\|\widehat{\mathbf{G}}\right\|_2 \leq \|\mathbf{G}\|_2 + \epsilon_2$ and $\|\widehat{\mathbf{y}}_0\|_2 \leq \|\mathbf{y}_0\|_2 + \epsilon_2$, we deduce that

$$
\begin{aligned}
M^\star &\leq 2\Big[\epsilon_2^2(2 + 5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2 \|\mathbf{G}\|_2)^2 + 2\epsilon_2 \|\mathbf{G}\|_2 (2 + 5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2 \|\mathbf{G}\|_2) \\
&+ \epsilon_2^2 (2 + 5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2 \|\mathbf{y}_0\|_2)^2 \\
&+ 2\epsilon_2 \|\mathbf{y}_0\|_2 (2 + 5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2 \|\mathbf{y}_0\|_2)\Big] + \mathcal{O}(\epsilon_2^2) \\
&= \mathcal{O}\left(\epsilon_2 \left(1 + \left\|\mathbf{\Phi}^\star_{uy}\right\|_2\right)(1 + \|\mathbf{G}\|_2 + \|\mathbf{y}_0\|_2)^2\right)\,,
\end{aligned}
$$

and, similarly, $V^\star = \mathcal{O}\left(\epsilon_2 \left(1 + \left\|\mathbf{\Phi}^\star_{uy}\right\|_2\right)(1 + \|\mathbf{y}_0\|_2)^2\right)$.

1.7 Final remarks

In this chapter, we have analyzed how much the model-mismatch due to noisy data can impact the safety and performance of output-feedback control systems with constraints. By deriving a suitable problem relaxation, we have proven that, despite the presence of constraints, the suboptimality of our proposed problem relaxation increases at most linearly for small model mismatches incurred during system identification.

We notice that our analysis towards the influence of noise is mainly based on the error upper bounds of the system's impulse and free response estimation (i.e., Assumption 1.1). Although we can use highly effective methods (e.g., SMM [11]) for response identification, a rigorous and efficient scheme for error analysis is still lacking. The bootstrap approach we used in Section 1.5.2 usually requires extensive resampling. For now, there does not exist any guarantee on the number of resamples needed such that the error analysis achieves a certain accuracy level. Therefore, it is interesting if we can find some data-driven approach that gives an upper bound for response identification errors.

2 Robust Data-Driven Control Based on Perturbation Analysis

2.1 Introduction

The method proposed in Chapter 1 requires impulse and free response identification based on a given trajectory data set. In this procedure, resampling, which is needed for identification error analysis, might be time-consuming. Moreover, there exists no finite-sample guarantees on the bootstrap results. In this chapter, we study active experiment design for data collection and provide an upper bound to the trajectory prediction error given by a behavioral-model-based method. By using this result, we can develop an open-loop robust control approach for the regulation of linear systems under measurement noise.

Several data-driven works approximately solve this problem with performance guarantees from different perspectives and under different assumptions. The authors of [14] minimize the worst-case trajectory tracking cost by reformulating the minmax problem into a semidefinite program through the S-lemma. However, the optimality results in [14] rely on the assumption that the noise sequence satisfies a cumulative quadratic constraint. Therefore, they do not hold when, e.g., the entries of the noise sequence are only known to satisfy box constraints, which is a common scenario in practice (many sensors are subject to bounded noise [83]). To address this issue, [19] adopts a different reformulation where model mismatch (the difference between measured recent output and that provided by the behavioral model) penalty is added to the tracking cost. The involved minmax problem is solvable by robust optimization techniques and the authors derive a suboptimality gap bound. However, this bound is conservative in the sense that it does not vanish when the noise decreases to zero.

To develop a control approach that enjoys both a less conservative suboptimality gap and the guarantee of constraint satisfaction, we leverage perturbation analysis for assessing the influence of measurement noise on the behavioral-model predictions. Research on perturbation analysis of data-driven prediction includes [16], [84] and [85]. In [16], the authors use a cost function similar with the one in [19], aiming to minimize the sum of the mismatch penalty and the tracking cost. Therefore, the prediction error upper bound relies on a variable related to the optimal control task and thus can only be evaluated after the optimal control problem has been solved. In contrast, the robust fundamental lemma proposed in [84] gives a prediction error analysis independent of control tasks, but it is limited to input-state systems and one-step-ahead prediction. The work in [85] extends [84] by deriving lower bounds for the singular values of the Hankel matrix in the behavioral model for input-output systems. However, the lower bounds depend on the unknown ground truth system model.

Targeting the limitation of existing methods, our robust control approach, relying solely on system input-output data, enjoys robust constraint satisfaction and a suboptimality bound that vanishes when noise decreases to zero.

2.1.1 Contributions

Our main contributions are the following:

(a) We propose an input generation strategy to collect historical data (in contrast with given and fixed historical data in Chapter 1) under bounded measurement noise for the construction of a Page matrix (a variant of Hankel matrix, see [86]). Then, we derive an error bound of the data-driven predictions, which only relies on noisy data. This bound is valid when the historical inputs achieve sufficient “persistent-excitation-to-noise ratio” (rigorously stated in Assumption 2.3) and the observability index (see Definition 2.2) is identified correctly. The first condition can be satisfied if collecting multiple historical data sets for averaging the noisy measurements or enlarging the input signals is allowed. We achieve the second through a data-driven method with correctness guarantee.

(b) For unconstrained regulation of MISO systems, in order to minimize the worst-case cost, we utilize the new prediction error bound to formulate a minmax problem and bound the suboptimality gap. The derived bound decreases to zero as the measurement noise converges to zero. This scheme can be extended to regulation of MIMO systems and robust constraint satisfaction.

2.1.2 Structure of this chapter

The rest of this chapter is organized as follows: in Section 2.2, we introduce the basics of data-driven control and formulate the robust control problem. Section 2.3 focuses on Multiple-Input Single-Output (MISO) systems with no input/output constraints and conducts perturbation analysis on the noisy behavioral model used for trajectory prediction. The novel robust control scheme is proposed in Section 2.4 where the associated upper bound to the worst-case cost derived by using perturbation analysis is minimized and the suboptimality gap is bounded. We extend our method to the regulation of Multiple-Input Multiple-Output (MIMO) systems and to robust constraint satisfaction in Section 2.5. Experiments illustrating the performance of the proposed approach are shown in Section 2.6.

2.1.3 Notations of this chapter

Given a time-varying vector variable v, we use v_t to denote its value at time instant t, let $[t_1, t_2] = \{t_1, t_1 + 1, ..., t_2\}$ and set $v_{[t_1,t_2]} := \{v_{t_1}, v_{t_1+1}, \dots, v_{t_2}\}$, $\operatorname{col}(v_{[t_1,t_2]}) := \begin{bmatrix} v_{t_1}^\top & v_{t_1+1}^\top & \dots & v_{t_2}^\top \end{bmatrix}^\top$. We use $\|x\|_i$ to denote the ℓ_i norm of x. Moreover, $||x||$ is the ℓ_2 norm. Given positive semidefinite matrix Q, the term $\|x\|_Q^2$ denotes $x^\top Qx$. For a matrix M, $||M||_i$ denotes the matrix i-norm while $||M||_{\max} := \max_{i,j} |m_{ij}|$ and $||M|| := ||M||_2$. We also denote $\sigma_{\min}(M)$ as the smallest singular value of M and $\sigma_{\max}(M)$ as the largest. We use $M_{i,\cdot}$ to denote the i-th row, $M_{\cdot,i}$ the i-th column and $M_{i:j,\cdot}$ the submatrix consisting of the rows of M from the i-th to the j-th. For a given $x \in \mathbb{R}$, we use notation $\lfloor x \rfloor$ to denote the floor function, i.e., $\lfloor x \rfloor = \max\{z \in \mathbb{Z} \mid z \leq x\}$. For $y \in \mathbb{R}^n$ and $r > 0$, we let $\mathcal{B}(y, r) := \{x : \|x - y\| \leq r\}$. The identity matrix with n rows is denoted as I_n. The pseudo-inverse of a matrix H is written as $H^\dagger$.

2.2 Preliminaries and Problem Formulation

2.2.1 Preliminaries: data-driven description of linear systems

This chapter considers the regulation of a discrete-time linear time-invariant (LTI) system with the following controllable and observable minimal realization,

$$x_{t+1} = Ax_t + Bu_t\,, \quad y_t = Cx_t + Du_t, \tag{2.1}$$

where $u_t \in \mathbb{R}^m$, $y_t \in \mathbb{R}^p$ and $x_t \in \mathbb{R}^{n_x}$. We assume that n_x and the system matrices A, B, C, D are unknown. Instead of identifying the system matrices, we utilize a behavioral model to characterize the possible trajectories in a horizon of length L. To build this system representation, we excite the linear system with an input sequence $u_{[1,T]}$ and collect the output data $y_{[1,T]}$, where $T > L$. The L-Page matrix of $u_{[1,T]}$ is given by

$$\mathcal{P}_L(u_{[1,T]}) := \begin{bmatrix} u_1 & u_{L+1} & \cdots & u_{l_\mathrm{h}L-L+1} \\ u_2 & u_{L+2} & \cdots & u_{l_\mathrm{h}L-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ u_L & u_{2L} & \cdots & u_{l_\mathrm{h}L}, \end{bmatrix}$$

where $l_\mathrm{h} = \lfloor \frac{T}{L} \rfloor$ denotes the number of columns [86]. To evaluate whether the input sequence $u_{[1,T]}$ along with the resulting outputs is sufficiently informative to uniquely determine system (2.1), we introduce the following definition analogous to persistent excitation in system identification.

Definition 2.1 ([86])**.** *For $L, T, d \in \mathbb{Z}^+$, we say the input sequence $u_{[1,T]}$ is L-Page exciting* *of order d if the following matrix has full row rank,*

$$\mathcal{P}_{L,d}\left(u_{[1,T]}\right) := \begin{pmatrix} \mathcal{P}_L\left(u_{[1,T-(d-1)L]}\right) \\ \mathcal{P}_L\left(u_{[L+1,T-(d-2)L]}\right) \\ \vdots \\ \mathcal{P}_L\left(u_{[L(d-1)+1,T]}\right) \end{pmatrix}.$$

By using the collected input output data $(u_{[1,T]}, y_{[1,T]})$, called *historical*, one might be able to determine whether another trajectory $(u^r_{[1,L]}, y^r_{[1,L]})$, called *recent*, is generated by system (2.1). Rigorously, we have the following result, which is a variant of the well-known Willem's Fundamental Lemma [81].

Lemma 2.1 ([86, Theorem 2.1])**.** *For the LTI system described in* (2.1)*, given a T-length historical trajectory $(u_{[1,T]}, y_{[1,T]})$ where $u_{[1,T]}$ is L-Page exciting of order $n_x + 1$ and the L-length recent trajectory $(u^r_{[1,L]}, y^r_{[1,L]})$, there exists $x^r_{[1,L]}$ with $x_i \in \mathbb{R}^{n_x}$, $1 \le i \le L$, such that $(x^r_{[1,L]}, u^r_{[1,L]}, y^r_{[1,L]})$ satisfies* (2.1) *if and only if there exists a vector $g \in \mathbb{R}^{l_\mathrm{h}}$ such that*

$$\begin{bmatrix} \mathcal{P}_L\left(u_{[1,T]}\right) \\ \mathcal{P}_L\left(y_{[1,T]}\right) \end{bmatrix} g = \begin{bmatrix} \mathrm{col}(u^r_{[1,L]}) \\ \mathrm{col}(y^r_{[1,L]}) \end{bmatrix}. \tag{2.2}$$

2.2.2 Problem formulation: robust control under measurement noise

Given the historical data $(u_{[1,T]}, y_{[1,T]})$ where $u_{[1,T]}$ is L-Page exciting of order $n_x + 1$ and an l_p-long initial trajectory $(u^r_{[1,l_\mathrm{p}]}, y^r_{[1,l_\mathrm{p}]})$ with $l_\mathrm{p} < L$, we consider the following regulation problem for the trajectory $(u^r_{[1,T]}, y^r_{[1,T]})$ from time $l_\mathrm{p}+1$ to time L:

$$\begin{aligned} \min_{u^r_{[l_\mathrm{p}+1,L]}, y^r_{[l_\mathrm{p}+1,L]}} \quad & \sum_{i=1}^{L-l_\mathrm{p}} (||u^r_{l_\mathrm{p}+i}||^2 + ||y^r_{l_\mathrm{p}+i}||^2) \\ \text{s.t.} \quad & \text{there exists } g \text{ such that} \\ & u_{[1,T]}, y_{[1,T]}, u^r_{[1,L]}, y^r_{[1,L]} \text{ satisfy (2.2)}. \end{aligned} \tag{2.3}$$

For convenience, we let $l_\mathrm{f} = L - l_\mathrm{p}$ and use the following notations for historical data,

$$\begin{aligned} U_\mathrm{p} &= \begin{bmatrix} I_{ml_\mathrm{p}} & 0 \end{bmatrix} \mathcal{P}_L\left(u_{[1,T]}\right), U_\mathrm{f} = \begin{bmatrix} 0 & I_{ml_\mathrm{f}} \end{bmatrix} \mathcal{P}_L\left(u_{[1,T]}\right), \\ Y_\mathrm{p} &= \begin{bmatrix} I_{pl_\mathrm{p}} & 0 \end{bmatrix} \mathcal{P}_L\left(y_{[1,T]}\right), Y_\mathrm{f} = \begin{bmatrix} 0 & I_{pl_\mathrm{f}} \end{bmatrix} \mathcal{P}_L\left(y_{[1,T]}\right) \end{aligned} \tag{2.4}$$

and for recent data,

$$\begin{aligned} u_\mathrm{p} &= \mathrm{col}(u^r_{[1,l_\mathrm{p}]}), u_\mathrm{f} = \mathrm{col}(u^r_{[l_\mathrm{p}+1,L]}), \\ y_\mathrm{p} &= \mathrm{col}(y^r_{[1,l_\mathrm{p}]}), y_\mathrm{f} = \mathrm{col}(y^r_{[l_\mathrm{p}+1,L]}). \end{aligned} \tag{2.5}$$

In practice, output measurements are subject to noise. Specifically, for any $i \in [1, \dots, T]$, $j \in [1, \dots, l_\mathrm{p}]$, there exist noise vectors $w_i, w_j^r \in \mathbb{R}^p$ such that the measurements are $\hat{y}_i = y_i + w_i$ and $\hat{y}_j^r = y_j^r + w_j^r$. We build $\widehat{Y}_\mathrm{p}, \widehat{Y}_\mathrm{f}$ and $\hat{y}_\mathrm{p}$ from $\hat{y}_{[1,T]}$ and $\hat{y}^r_{[1,l_\mathrm{p}]}$ as noisy counterparts of $Y_\mathrm{p}, Y_\mathrm{f}$ and y_p. In this chapter, we only consider bounded noise, as stated in the following assumption.

Assumption 2.1. *For the noisy measurements* $\widehat{Y}_\mathrm{p}, \widehat{Y}_\mathrm{f}, \hat{y}_\mathrm{p}$, *corresponding to* $Y_\mathrm{p}, Y_\mathrm{f}, y_\mathrm{p}$, *the following holds,*

$$\max\{||\widehat{Y}_\mathrm{p} - Y_\mathrm{p}||_{\max}, ||\widehat{Y}_\mathrm{f} - Y_\mathrm{f}||_{\max}, ||\hat{y}_\mathrm{p} - y_\mathrm{p}||_\infty\} \leq \delta,$$

where $\delta > 0$ *is a known constant.*

Given a fixed control strategy, different noise realizations lead to different performances. To attenuate the influence of the uncertainties, we aim to **design a robust control scheme where the worst-case regulation cost is minimized**. In general, the worst-case cost is hard to compute and we will provide an upper bound $c_\mathrm{worst}(u_\mathrm{f})$ to it given the historical data $u_{[1,T]}, \hat{y}_{[1,T]}$ and the recent data u_p and $\hat{y}_\mathrm{p}$. We formulate the robust control problem as

$$u_\mathrm{f}^* = \mathrm{argmin}_{u_\mathrm{f}} c_\mathrm{worst}(u_\mathrm{f}).$$

This way, we ensure that the true cost induced by u_f^* is less than $c_\mathrm{worst}(u_\mathrm{f}^*)$.

In Section 2.4 we propose a formulation of $c_\mathrm{worst}(u_\mathrm{f})$. But before that, in Section 2.3, by assuming u_f is given, we introduce tools of data-driven prediction based on behavioral models and conduct perturbation analysis, which allows us to estimate the worst-case cost by using the noisy data. To avoid bulky statements, we only consider the MISO case in Sections 2.3 and 2.4, while the extension to the MIMO case is presented in Section 2.5.

2.3 Data-Driven Prediction with Perturbation Analysis for MISO Systems

In this section, we investigate, for MISO systems, a data-driven prediction method that generates an estimate $\hat{y}_\mathrm{f}$ for y_f based on historical data along with u_p, $\hat{y}_\mathrm{p}$ and u_f. Specifically, we propose a method for generating historical data enabling the derivation of an upper bound for the prediction error. This ingredient is essential for the computation of upper bounds to worst-case costs in Section 2.4.

2.3.1 The proposed data-driven prediction scheme

For MISO systems, we adopt the data-driven prediction method used in the framework of **PEM-MPC**[1] proposed in [87]. Specifically, we look into the following prediction scheme

$$\hat{g}(u_\mathrm{f}) = \widehat{H}^\dagger \hat{b}(u_\mathrm{f}),\ \hat{y}_\mathrm{f}(u_\mathrm{f}) = \widehat{Y}_\mathrm{f}\hat{g}(u_\mathrm{f}),\ \text{where } \widehat{H} = \begin{bmatrix} U_\mathrm{p}^\top & \widehat{Y}_\mathrm{p}^\top & U_\mathrm{f}^\top \end{bmatrix}^\top,\ \hat{b}(u_\mathrm{f}) = \begin{bmatrix} u_\mathrm{p}^\top & \hat{y}_\mathrm{p}^\top & u_\mathrm{f}^\top \end{bmatrix}^\top. \tag{2.6}$$

We also define the noiseless counterparts of $\widehat{H}, \hat{b}(u_\mathrm{f}), \hat{g}(u_\mathrm{f}), \hat{y}_\mathrm{f}(u_\mathrm{f})$ in (2.6) as $\overline{H}, \bar{b}(u_\mathrm{f}), \bar{g}(u_\mathrm{f}), \bar{y}_\mathrm{f}(u_\mathrm{f})$. One can easily verify that $\|\bar{g}(u_\mathrm{f})\| = \min\{\|g\| : \bar{H}g = \bar{b}(u_\mathrm{f}) \text{ and } Y_\mathrm{f}g = \bar{y}_\mathrm{f}(u_\mathrm{f})\}$. Before elaborating on the historical data collected and the construction of $\widehat{H}$ in Assumption 2.2, we introduce the concept of observability index.

Definition 2.2. *We say l_o is the observability index of the system* (2.1) *if the l_o is the smallest positive integer l such that the observability matrix $\mathcal{O}(l) := \begin{bmatrix} C^\top & (CA)^\top & \cdots & (CA^{l-1})^\top \end{bmatrix}^\top$ is full column rank.*

Assumption 2.2. *Given horizon length $L > l_o$, the historical inputs of the MISO system* (2.1) *are generated using the following setting:*

$$x_1 = 0, \tag{2.7a}$$

$$u_{[1,L]}, u_{[L+1,2L]}, \ldots, u_{[4L^2+1,4L^2+L]} \overset{\text{i.i.d.}}{\sim} \mathcal{Q}, \tag{2.7b}$$

$$T \geq 4L^2 + L \text{ and } u_{[1,T]} \text{ is } L\text{-Page exciting of order } n_x + 1, \tag{2.7c}$$

where $\mathcal{Q}$ is a probability distribution such that for $x \in \mathbb{R}^{mL}$

$$\text{if } x \sim \mathcal{Q} \text{ then for any subspace } V \subsetneqq \mathbb{R}^{mL},\ \mathbb{P}(\{x \in V\}) = 0. \tag{2.8}$$

The Page matrices $U = \mathcal{P}_L\left(u_{[1,T]}\right)$ and $\widehat{Y} = \mathcal{P}_L\left(\hat{y}_{[1,T]}\right)$ are split into $U_\mathrm{p}, U_\mathrm{f}, \widehat{Y}_\mathrm{p}, \widehat{Y}_\mathrm{f}$ using (2.4) *with*

$$l_\mathrm{p} = l_\mathrm{o}, \tag{2.9}$$

and $\widehat{H}$ is built using (2.6).

For identification of l_o using data, we refer to Section 2.3.3 and Algorithm 1. Under Assumption 2.2, we will see in Section 2.3.2 that the noiseless matrix $\overline{H}$ is almost surely full row rank. This property is essential for upperbounding the prediction error resulting from (2.6). The reason is that, if $\overline{H}$ has a zero singular value, even infinitesimal measurement noise can result in a huge prediction error due to the use of the pseudoinverse in (2.6).

2.3.2 Perturbation analysis

We show in Theorem 2.1 and Theorem 2.2 that the prediction error resulting from the scheme (2.6) can be bounded.

[1] PEM stands for *Prediction Error Method* .

Theorem 2.1. *Under Assumption 2.2, we have* $\mathbb{P}(\{\overline{H}$ *is full row rank*$\}) = 1$.

The proof of Theorem 2.1 is given in Section 2.7.2.

Remark 2.1. *Theorem 2.1 does not hold for MIMO systems. As a simple example, when there are two outputs measuring the same quantity, almost surely the noiseless* $\overline{H}$ *is not full row rank.*

By utilizing Theorem 2.1, we show in Theorem 2.2 that, when the measurement noise is sufficiently small (as indicated below), the prediction errors given by the least-square solution (2.6) can be bounded.

Assumption 2.3. *With* $\widehat{H}$ *constructed according to Assumption 2.2, we have*

$$\delta < \frac{\sigma_{\min}(\widehat{H})}{2l_{\mathrm{h}}}. \tag{2.10}$$

Remark 2.2. *As will be shown later in Theorem 2.2, under Assumption 2.3, an upper bound for the prediction error is roughly proportional to* $\sigma_{\min}(\widehat{H})^{-1}$*. Therefore, we can regard* $\sigma_{\min}(\widehat{H})$ *as a measure of "persistent excitation" and thus* (2.10) *requires the "persistent-excitation-to-noise ratio" to be sufficiently large. In [16, Lemma 1], a similar assumption is made. To satisfy Assumption 2.3, one can decrease the noise magnitude by collecting multiple historical data sets and using averaging techniques if the entries of the noise sequence are independent and identically distributed (for details see Remark 2.4). Since from [88, Theorem 4.3] we have*

$$\sigma_{\min}(\widehat{H}) \geq \sigma_{\min}(\overline{H}) - \|\widehat{H} - \overline{H}\|, \tag{2.11}$$

we can also amplify the historical input signals in (2.7) *for obtaining a larger* $\sigma_{\min}(\overline{H})$ *and hence increasing* $\sigma_{\min}(\widehat{H})$*, since* $\|\widehat{H} - \overline{H}\|$ *only depends on the noise sequence (see* (2.12)*).*

Theorem 2.2. *We denote* $\overline{\Delta}_g(u_{\mathrm{f}}) := \bar{g}(u_{\mathrm{f}}) - \hat{g}(u_{\mathrm{f}})$ *and* $\overline{\Delta}_{y_{\mathrm{f}}}(u_{\mathrm{f}}) := \bar{y}_{\mathrm{f}}(u_{\mathrm{f}}) - \hat{y}_{\mathrm{f}}(u_{\mathrm{f}})$*. If Assumption 2.3 holds, we have for MISO systems*

$$||\overline{\Delta}_g(u_{\mathrm{f}})|| \leq \mathcal{C}(u_{\mathrm{f}})\delta, \text{ where } \mathcal{C}(u_{\mathrm{f}}) = 2\sigma_{\min}(\widehat{H})^{-1}(\sqrt{l_{\mathrm{p}}} + l_{\mathrm{h}}||\hat{g}(u_{\mathrm{f}})||), \text{ and}$$

$$||\overline{\Delta}_{y_{\mathrm{f}}}(u_{\mathrm{f}})|| \leq \mathcal{C}(u_{\mathrm{f}})||\widehat{Y}_{\mathrm{f}}||\delta + l_{\mathrm{h}}(\|\hat{g}(u_{\mathrm{f}})\| + \mathcal{C}(u_{\mathrm{f}}))\delta.$$

Proof. We let $E := \widehat{H} - \overline{H}$. Due to Assumption 2.1, we have

$$||E|| \leq l_{\mathrm{h}}\delta. \tag{2.12}$$

Then the following inequalities hold,

$$\begin{aligned}
\frac{\sigma_{\min}(\widehat{H})}{2}||\hat{g}(u_{\mathrm{f}}) - \bar{g}(u_{\mathrm{f}})|| &\leq (\sigma_{\min}(\widehat{H}) - ||E||)||\hat{g}(u_{\mathrm{f}}) - \bar{g}(u_{\mathrm{f}})|| \\
&\leq \sigma_{\min}(\widehat{H} - E)||\hat{g}(u_{\mathrm{f}}) - \bar{g}(u_{\mathrm{f}})|| \\
&\leq ||\overline{H}(\hat{g}(u_{\mathrm{f}}) - \bar{g}(u_{\mathrm{f}}))|| \\
&\leq ||\widehat{H}\hat{g}(u_{\mathrm{f}}) - \overline{H}\bar{g}(u_{\mathrm{f}}) + (\overline{H} - \widehat{H})\hat{g}(u_{\mathrm{f}})|| \\
&\leq ||\hat{b}(u_{\mathrm{f}}) - \bar{b}(u_{\mathrm{f}})|| + ||E|| \cdot ||\hat{g}(u_{\mathrm{f}})|| \\
&= (\sqrt{l_{\mathrm{p}}} + l_{\mathrm{h}}||\hat{g}(u_{\mathrm{f}})||)\delta,
\end{aligned} \tag{2.13}$$

where the first inequality is due to (2.10) in Assumption 2.3, the second results from (2.11) and $\hat{b}(u_{\mathrm{f}})$ is defined in (2.6). By simplifying (2.13), we have

$$||\overline{\Delta}_g(u_{\mathrm{f}})|| \leq \mathcal{C}(u_{\mathrm{f}})\delta.$$

Based on this inequality, we have

$$\begin{aligned}
||\hat{y}_{\mathrm{f}}(u_{\mathrm{f}}) - \bar{y}_{\mathrm{f}}(u_{\mathrm{f}})|| &= ||\widehat{Y}_{\mathrm{f}}(\hat{g}(u_{\mathrm{f}}) - \bar{g}(u_{\mathrm{f}})) + (\widehat{Y}_{\mathrm{f}} - Y_{\mathrm{f}})\bar{g}(u_{\mathrm{f}})|| \\
&\leq ||\widehat{Y}_{\mathrm{f}}|| \cdot ||\hat{g}(u_{\mathrm{f}}) - \bar{g}(u_{\mathrm{f}})|| + ||\widehat{Y}_{\mathrm{f}} - Y_{\mathrm{f}}|| \cdot ||\bar{g}(u_{\mathrm{f}})|| \\
&\leq ||\widehat{Y}_{\mathrm{f}}|| \cdot ||\hat{g}(u_{\mathrm{f}}) - \bar{g}(u_{\mathrm{f}})|| + \\
&\quad l_{\mathrm{h}}\delta(||\hat{g}(u_{\mathrm{f}})|| + ||\bar{g}(u_{\mathrm{f}}) - \hat{g}(u_{\mathrm{f}})||) \\
&\leq \mathcal{C}(u_{\mathrm{f}})||\widehat{Y}_{\mathrm{f}}||\delta + l_{\mathrm{h}}(\|\hat{g}(u_{\mathrm{f}})\| + \mathcal{C}(u_{\mathrm{f}}))\delta.
\end{aligned} \tag{2.14}$$

□

In the literature, similar results on perturbation analysis accounting for measurement noise can be found in [7] and [89]. However, the associated bounds utilize the unknown true system model. In the behavioral model framework, [16] also derives a prediction error bound after solving an optimal control problem. Therefore, the bound is only valid for the optimal control sequence u_{f}^* determined by the specific problem formulation. In contrast, our upper bound in Theorem 2.2 can be calculated directly from the noisy data for any given u_{f}. This feature allows us to formulate in Sections 2.4 and 2.5 a min-max regulation problem where robust constraint satisfaction can be enforced.

2.3.3 Data-driven identification of l_{o}

We discuss the identification of l_{o}, which is needed in the construction of $\widehat{H}$ for satisfying Assumption 2.2. For this aim, we discuss in Proposition 2.1 whether $\overline{H}$ almost surely has full rank when Assumption 2.2 does not hold. Based on this result, we propose a method to identify l_{o}.

Proposition 2.1. *If Assumption 2.2 does not hold because $l_\mathrm{p} < l_\mathrm{o}$, then*

$$\mathbb{P}(\{\overline{H} \textit{ is full row rank}\}) = 1.$$

Furthermore, if $l_\mathrm{p} > l_\mathrm{o}$, $\mathbb{P}(\{\overline{H} \textit{ is full row rank}\}) = 0$.

The proof is reported in Section 2.7.3. Theorem 2.1 and Proposition 2.1 say that l_o is the largest value for l_p such that the constructed $\overline{H}$ is full row rank. We notice that the noise matrix $E = \widehat{H} - \overline{H}$ satisfies that $\|E\| \leq l_\mathrm{h}\delta$ [2]. Therefore, if $\overline{H}$ is not full row rank, i.e., $\sigma_{\min}(\overline{H}) = 0$, we have

$$\sigma_{\min}(\widehat{H}) \leq \sigma_{\min}(\overline{H}) + \sigma_{\max}(E) \leq l_\mathrm{h}\delta. \tag{2.15}$$

By utilizing (2.15), we propose Algorithm 1 for identification of l_o and show the correctness of the derived result in Proposition 2.2.

Proposition 2.2. *Under Assumption 2.3, if $L > l_\mathrm{o}$, Algorithm 1 returns l_o, the true observability index, almost surely.*

Proof. From Assumption 2.3 and (2.15), we know that $\sigma_{\min}(\widehat{H}) \geq 2l_\mathrm{h}\delta$ if $l_\mathrm{p} = l_\mathrm{o}$ and $\sigma_{\min}(\widehat{H}) \leq l_\mathrm{h}\delta$ if $l_\mathrm{p} > l_\mathrm{o}$. Since Algorithm 1 terminates when $\sigma_{\min}(\widehat{H}) \leq l_\mathrm{h}\delta$ is verified, we only need to show that $\sigma_{\min}(\widehat{H}) > l_\mathrm{h}\delta$ almost surely if $l_\mathrm{p} < l_\mathrm{o}$.

In the remainder of this proof, we denote with $[U_{\mathrm{p},k}^\top \; U_{\mathrm{f},k}^\top]^\top$ and $[\widehat{Y}_{\mathrm{p},k}^\top \; \widehat{Y}_{\mathrm{f},k}^\top]^\top$ the partitions of the Page matrices U and Y, respectively, where $U_{\mathrm{p},k}$ has km rows and $Y_{\mathrm{p},k}$ has kp rows. Correspondingly, we write $\widehat{H}_k := [U_{\mathrm{p},k}^\top \; \widehat{Y}_{\mathrm{p},k}^\top \; U_{\mathrm{f},k}^\top]^\top$. We notice that $\widehat{H}_k$ is a submatrix consisting of a fraction of $\widehat{H}_{l_\mathrm{o}}$'s rows and $\widehat{H}_{l_o}$ is full row rank almost surely (Theorem 2.1). According to Lemma 2.4 in Section 2.7.4, we have that almost surely $\sigma_{\min}(\widehat{H}_k) \geq \sigma_{\min}(\widehat{H}_{l_\mathrm{o}}) \geq 2l_\mathrm{h}\delta$ for any $k < l_\mathrm{o}$. □

Remark 2.3. *Hereafter, we describe a heuristic method which can be used as supplement to Algorithm 1 when Assumption 2.3 is not satisfied. Specifically, in Line 1 of Algorighm 1, after deriving $(u_{[1,T]}, \hat{y}_{[1,T]})$, we can generate several extra historical trajectories $(u(\alpha)_{[1,T]}, y(\alpha)_{[1,T]})$ using the initial state $x_1 = 0$ and the input sequence $u(\alpha)_{[1,T]} = \alpha u_{[1,T]}$ for different values of α. Based on the data, we can construct Page matrices $U(\alpha), Y(\alpha)$. In Lines 4 and 5, we obtain the submatrices $U_\mathrm{p}(\alpha), U_\mathrm{f}(\alpha), \widehat{Y}_\mathrm{p}(\alpha), \widehat{Y}_\mathrm{f}(\alpha)$ and $\widehat{H}(\alpha)$. Due to the linearity of the system and the zero initial state, we have $\overline{H}(\alpha) = \alpha\overline{H}(1)$ and thus $\sigma_{\min}(\overline{H}(\alpha))$ is proportional to α if $\sigma_{\min}(\overline{H}(1)) \neq 0$. Therefore, when observing that $\sigma_{\min}(\widehat{H}(\alpha))$ increases approximately proportionally with α, we claim that $\sigma_{\min}(\overline{H}(1)) \neq 0$.*

[2] To show this result, one only needs Cauchy–Schwarz inequality and Assumption 2.1.

Algorithm 1 Data-driven observability index identification for MISO systems

```
Input: horizon length $L$ for the Page matrices
Output: the system order $l_\mathrm{o}$
Use (2.7) to generate historical data $(u_{[1,T]}, \hat{y}_{[1,T]})$ to construct Page matrices $U = \mathcal{P}_L(u_{[1,T]})$ and $\widehat{Y} = \mathcal{P}_L(\hat{y}_{[1,T]})$.
$k \leftarrow 1, \mathrm{TER} = 0$
while $\mathrm{TER} = 0$ do
    Partition $U = [U_\mathrm{p}^\top \; U_\mathrm{f}^\top]^\top$, $\widehat{Y} = [\widehat{Y}_\mathrm{p}^\top \; \widehat{Y}_\mathrm{f}^\top]^\top$ such that $U_\mathrm{p}$ has $km$ rows (i.e., $l_\mathrm{p} = k$)
    Build $\widehat{H} = [U_\mathrm{p}^\top \; \widehat{Y}_\mathrm{p}^\top \; U_\mathrm{f}^\top]^\top$
    if $\sigma_{\min}(\widehat{H}) \leq l_\mathrm{h}\delta$ then
        $l_\mathrm{o} \leftarrow k - 1$, TER$\leftarrow 1$
    end if
    $k \leftarrow k + 1$
end while
```

2.4 Robust Control for Regulation of MISO Systems with Suboptimality Guarantees

In this section, we propose a data-driven robust control method for MISO systems, where we use the prediction error bounds in Theorem 2.2 to calculate $c_\mathrm{worst}(u_\mathrm{f})$, an upperbound to the regulation cost. To justify that this upperbound is not too conservative, we study the suboptimality of the derived input sequence and compare it with the optimal input sequence. The extension to multiple-output systems is provided in Section 2.5.

With the notation $\Delta := \{\Delta_{Y_\mathrm{p}}, \Delta_{Y_\mathrm{f}}, \Delta_{y_\mathrm{p}}\}$, the robust regulation problem is formulated as the following bilevel program where the inner problem calculates an upperbound for the worst-case cost and the outer problem optimizes the input such that the cost upperbound is minimized,

$$\min_{u_\mathrm{f}} \max_{\Delta, g, y_\mathrm{f}} \quad y_\mathrm{f}^\top y_\mathrm{f} + u_\mathrm{f}^\top u_\mathrm{f} \tag{2.16a}$$

$$\text{s.t.} \quad \begin{bmatrix} U_\mathrm{p} \\ \widehat{Y}_\mathrm{p} \\ U_\mathrm{f} \\ \widehat{Y}_\mathrm{f} \end{bmatrix} g + \begin{bmatrix} 0 \\ \Delta_{Y_\mathrm{p}} \\ 0 \\ \Delta_{Y_\mathrm{f}} \end{bmatrix} g = \begin{bmatrix} u_\mathrm{p} \\ \hat{y}_\mathrm{p} \\ u_\mathrm{f} \\ y_\mathrm{f} \end{bmatrix} + \begin{bmatrix} 0 \\ \Delta_{y_\mathrm{p}} \\ 0 \\ 0 \end{bmatrix} \tag{2.16b}$$

$$\max\left(||\Delta_{Y_\mathrm{p}}||_{\max}, ||\Delta_{Y_\mathrm{f}}||_{\max}, ||\Delta_{y_\mathrm{p}}||_\infty\right) \leq \delta \tag{2.16c}$$

$$||g - \hat{g}(u_\mathrm{f})||^2 \leq \mathcal{C}^2(u_\mathrm{f})\delta^2. \tag{2.16d}$$

where $\mathcal{C}(u_\mathrm{f})$ is defined in Theorem 2.2. We use the alternating optimization method in [90] to solve problem (2.16)(for details see Section 2.6.1). We denote the solution to the outer problem as $\check{u}_\mathrm{f}$. Given any input sequence u_f, we let $\bar{c}(u_\mathrm{f})$ be the resulting true regulation cost and $c_\mathrm{worst}(u_\mathrm{f})$ be the optimal objective value of the inner problem of (2.16). In the following theorem, we show that $c_\mathrm{worst}(u_\mathrm{f})$ is indeed an upperbound to the true regulation cost.

Theorem 2.3. *If Assumptions 2.1, 2.2 and 2.3 hold, for any u_f we have $c_\mathrm{worst}(u_\mathrm{f}) \geq \bar{c}(u_\mathrm{f})$.*

Proof. The noise realization $(\overline{\Delta}_{Y_\mathrm{p}}, \overline{\Delta}_{Y_\mathrm{f}}, \overline{\Delta}_{y_\mathrm{p}})$ satisfies that

$$\max\left(||\overline{\Delta}_{Y_\mathrm{p}}||_{\max}, ||\overline{\Delta}_{Y_\mathrm{f}}||_{\max}, ||\overline{\Delta}_{y_\mathrm{p}}||_{\infty}\right) \leq \delta.$$

With the vector $\bar{g}(u_\mathrm{f})$ we can reconstruct the noiseless system output, i.e., $(\widehat{Y}_\mathrm{p} + \overline{\Delta}_{y_\mathrm{p}})\bar{g}(u_\mathrm{f}) = y_\mathrm{p}$, $(\widehat{Y}_\mathrm{f} + \overline{\Delta}_{y_\mathrm{f}})\bar{g}(u_\mathrm{f}) = \bar{y}_\mathrm{f}(u_\mathrm{f})$. Now we see that $u_\mathrm{f}, \overline{\Delta}_{Y_\mathrm{p}}, \overline{\Delta}_{Y_\mathrm{f}}$, $\overline{\Delta}_{y_\mathrm{p}}, \bar{g}(u_\mathrm{f})$, and $\bar{y}_\mathrm{f}(u_\mathrm{f})$ satisfy (2.16b), (2.16c) and the corresponding regulation cost is $\bar{c}(u_\mathrm{f})$. Meanwhile, considering the error bounds in Theorem 2.2, the constraint (2.16d) is also satisfied. Since $c_{\mathrm{worst}}(u_\mathrm{f})$ is the maximum cost in the inner problem of (2.16), we have $c_{\mathrm{worst}}(u_\mathrm{f}) \geq \bar{c}(u_\mathrm{f})$. □

In (2.16), the outer problem seeks a u_f that minimizes this upperbound. Similar ideas that minimize a cost upperbound can be found in [89, 13]. To discuss the conservativeness of this approach, we need to bound the suboptimality of the solution to (2.16) when compared to the noiseless case. To this aim, we define the tuple $\widehat{\mathcal{Y}} := (\widehat{Y}_\mathrm{p}, \widehat{Y}_\mathrm{f}, \widehat{y}_\mathrm{p})$, consider the control input sequences

$$\begin{aligned} \hat{u}_\mathrm{f}^* &: = \mathrm{argmin}_{u_\mathrm{f}} \|\hat{y}_\mathrm{f}(u_\mathrm{f})\|^2 + \|u_\mathrm{f}\|^2, \\ u_\mathrm{f}^* &: = \mathrm{argmin}_{u_\mathrm{f}} \|\bar{y}_\mathrm{f}(u_\mathrm{f})\|^2 + \|u_\mathrm{f}\|^2, \end{aligned} \tag{2.17}$$

and bound the error $\|\hat{u}_\mathrm{f}^* - u_\mathrm{f}^*\|$ in the following lemma whose proof is given in Section 2.7.5.

Lemma 2.2. *Under Assumption 2.1, 2.2 and 2.3, we define the following polynomials in δ,*

$$\begin{aligned} \mathcal{F}_1(\delta, \widehat{\mathcal{Y}}) &: = 2l_\mathrm{h}||\widehat{H}^\dagger||(1 + 4||\widehat{Y}_\mathrm{p}|| \cdot ||\widehat{H}^\dagger||)\delta \\ \mathcal{F}_2(\delta, \widehat{\mathcal{Y}}) &: = (2||\widehat{K}_1|| + \mathcal{F}_1(\delta, \widehat{\mathcal{Y}}))\mathcal{F}_1(\delta, \widehat{\mathcal{Y}}) \\ \mathcal{F}_3(\delta, \widehat{\mathcal{Y}}) &: = \sqrt{l_\mathrm{p}}\left\|\widehat{K}_2^\top \widehat{K}_1\right\|\delta + \left(\left\|\hat{b}(0)\right\| + \sqrt{l_\mathrm{p}}\delta\right)\mathcal{F}_2(\delta, \widehat{\mathcal{Y}}) \\ \mathcal{F}(\delta, \widehat{\mathcal{Y}}) &: = \left\|(\widehat{K}_2^\top \widehat{K}_2 + I)^{-1}\right\|\mathcal{F}_3(\delta, \widehat{\mathcal{Y}}) + \left(\left\|\widehat{K}_2^\top \widehat{K}_1 \hat{b}(0)\right\| + \mathcal{F}_3(\delta, \widehat{\mathcal{Y}})\right)\mathcal{F}_2(\delta, \widehat{\mathcal{Y}}), \end{aligned} \tag{2.18}$$

where $\widehat{K}_1 := \widehat{Y}_\mathrm{f}\widehat{H}^\dagger$, $\widehat{K}_2 := \widehat{K}_1[0\ \ 0\ \ I]^\top$ and $\hat{b}(u_\mathrm{f})$ is defined in (2.6). *Then, by letting $\eta(\delta, \widehat{\mathcal{Y}}) = 1 + \frac{||\widehat{H}^\dagger||\mathcal{F}(\delta, \widehat{\mathcal{Y}})}{||\hat{g}(\hat{u}_\mathrm{f}^*)||}$, we have $||\hat{u}_\mathrm{f}^* - u_\mathrm{f}^*|| \leq \mathcal{F}(\delta, \widehat{\mathcal{Y}})$ and*

$$||\hat{g}(u_\mathrm{f}^*)|| \leq \eta(\delta, \widehat{\mathcal{Y}})||\hat{g}(\hat{u}_\mathrm{f}^*)||. \tag{2.19}$$

Recall that $c_{\mathrm{worst}}(u_\mathrm{f})$ and $\bar{c}(u_\mathrm{f})$ are, respectively, the worst-case cost derived by the inner problem of (2.16) and the resulting true regulation cost when u_f is applied, u_f^* is the optimal input for the noiseless case (see (2.17)) and $\check{u}$ is the optimal solution to (2.16). In the following, we compare $\bar{c}(\check{u}_\mathrm{f})$ with $\bar{c}(u_\mathrm{f}^*)$ to see how much suboptimality is introduced by solving (2.16) and applying $\check{u}_\mathrm{f}$.

Theorem 2.4. *Let Assumption 2.1, 2.2 and 2.3 hold and define*

$$
\begin{aligned}
\mathcal{C}_1(\delta, \widehat{\mathcal{Y}}) &: = 2\sigma_{\min}^{-1}(\widehat{H})(\sqrt{l_{\mathrm{p}}} + \eta(\delta, \widehat{\mathcal{Y}})l_{\mathrm{h}}||\hat{g}(\hat{u}_{\mathrm{f}}^*)||)\delta \\
\mathcal{C}_2(\delta, \widehat{\mathcal{Y}}) &: = (||\widehat{Y}_{\mathrm{f}}|| + l_{\mathrm{h}}\delta)\mathcal{C}_1(\delta, \widehat{\mathcal{Y}}) + \eta(\delta, \widehat{\mathcal{Y}})l_{\mathrm{h}}||\hat{g}(\hat{u}_{\mathrm{f}}^*)||\delta \\
\mathcal{C}_3(\delta, \widehat{\mathcal{Y}}) &: = 8(\mathcal{C}_2(\delta, \widehat{\mathcal{Y}}))^2 + 4\left\|\widehat{Y}_{\mathrm{f}}\right\| \cdot ||\hat{g}(\hat{u}_{\mathrm{f}}^*)||\eta(\delta, \widehat{\mathcal{Y}})\mathcal{C}_2(\delta, \widehat{\mathcal{Y}}).
\end{aligned} \tag{2.20}
$$

One has that $\bar{c}(\check{u}_{\mathrm{f}}) - \bar{c}(u_{\mathrm{f}}^*) \leq \mathcal{C}_3(\delta, \widehat{\mathcal{Y}})$. Morever, the upperbound $\mathcal{C}_3(\delta, \widehat{\mathcal{Y}})$, computable by using only the noisy measurements, converges in probability to 0 as δ converges to 0.

Proof. We denote $(\widetilde{\Delta}_{Y_{\mathrm{p}}}, \widetilde{\Delta}_{Y_{\mathrm{f}}}, \widetilde{\Delta}_{y_{\mathrm{p}}}, \tilde{g}(u_{\mathrm{f}}))$ as the solution to the following optimization problem

$$
\begin{aligned}
\max_{\Delta_{Y_{\mathrm{p}}}, \Delta_{Y_{\mathrm{f}}}, \Delta_{y_{\mathrm{p}}}, g} \quad & g^\top(\widehat{Y}_{\mathrm{f}} + \Delta_{Y_{\mathrm{f}}})^\top(\widehat{Y}_{\mathrm{f}} + \Delta_{Y_{\mathrm{f}}})g + u_{\mathrm{f}}^\top u_{\mathrm{f}} \\
\text{s.t.} \quad & \begin{bmatrix} U_{\mathrm{p}} \\ \widehat{Y}_{\mathrm{p}} \\ U_{\mathrm{f}} \end{bmatrix} g + \begin{bmatrix} 0 \\ \Delta_{Y_{\mathrm{p}}} \\ 0 \end{bmatrix} g = \begin{bmatrix} u_{\mathrm{p}} \\ \hat{y}_{\mathrm{p}} \\ u_{\mathrm{f}}^* \end{bmatrix} + \begin{bmatrix} 0 \\ \Delta_{y_{\mathrm{p}}} \\ 0 \end{bmatrix} \\
& \max\left(||\Delta_{Y_{\mathrm{p}}}||_{\max}, ||\Delta_{Y_{\mathrm{f}}}||_{\max}, ||\Delta_{y_{\mathrm{p}}}||_\infty\right) \leq \delta \\
& ||g - \hat{g}(u_{\mathrm{f}})||^2 \leq \mathcal{C}^2(u_{\mathrm{f}})\delta^2
\end{aligned} \tag{2.21}
$$

and also define $\tilde{y}_{\mathrm{f}}(u_{\mathrm{f}}) := (\widehat{Y}_{\mathrm{f}} + \widetilde{\Delta}_{y_{\mathrm{f}}})\tilde{g}(u_{\mathrm{f}})$. Since $||\hat{g}(u_{\mathrm{f}}^*) - \bar{g}(u_{\mathrm{f}}^*)|| \leq \mathcal{C}(u_{\mathrm{f}}^*)\delta$ according to Theorem 2.2, we have

$$
\begin{aligned}
\frac{1}{2}\sigma_{\min}(\widehat{H})||\hat{g}(u_{\mathrm{f}}^*) - \bar{g}(u_{\mathrm{f}}^*)|| &\leq (\sqrt{l_{\mathrm{p}}} + l_{\mathrm{h}}||\bar{g}(u_{\mathrm{f}}^*)||)\delta \\
&\leq (\sqrt{l_{\mathrm{p}}} + \eta(\delta, \widehat{\mathcal{Y}})l_{\mathrm{h}}||\hat{g}(\hat{u}_{\mathrm{f}}^*)||)\delta,
\end{aligned} \tag{2.22}
$$

where $\eta(\delta, \widehat{\mathcal{Y}})$ is derived in Lemma 2.2. The inequality (2.22) implies $||\hat{g}(u_{\mathrm{f}}^*) - \bar{g}(u_{\mathrm{f}}^*)|| \leq \mathcal{C}_1(\delta, \widehat{\mathcal{Y}})$. Similarly, we have $||\hat{g}(u_{\mathrm{f}}^*) - \tilde{g}(u_{\mathrm{f}}^*)|| \leq \mathcal{C}_1(\delta, \widehat{\mathcal{Y}})$. Consequently, we can derive

$$
\begin{aligned}
||\hat{y}_{\mathrm{f}}(u_{\mathrm{f}}^*) - \bar{y}_{\mathrm{f}}(u_{\mathrm{f}}^*)|| &= ||\widehat{Y}_{\mathrm{f}}(\hat{g}(u_{\mathrm{f}}^*) - \bar{g}(u_{\mathrm{f}}^*)) - \overline{\Delta}_{Y_{\mathrm{f}}}\bar{g}(u_{\mathrm{f}}^*)|| \\
&\leq ||\widehat{Y}_{\mathrm{f}}|| \cdot ||\hat{g}(u_{\mathrm{f}}^*) - \bar{g}(u_{\mathrm{f}}^*)|| + ||\overline{\Delta}_{Y_{\mathrm{f}}}|| \cdot ||\bar{g}(u_{\mathrm{f}}^*)|| \\
&\leq ||\widehat{Y}_{\mathrm{f}}|| \cdot ||\hat{g}(u_{\mathrm{f}}^*) - \bar{g}(u_{\mathrm{f}}^*)|| + \\
&\qquad l_{\mathrm{h}}\delta(||\hat{g}(u_{\mathrm{f}}^*)|| + ||\bar{g}(u_{\mathrm{f}}^*) - \hat{g}(u_{\mathrm{f}}^*)||) \\
&\leq \underbrace{(||\widehat{Y}_{\mathrm{f}}|| + l_{\mathrm{h}}\delta)\mathcal{C}_1(\delta, \widehat{\mathcal{Y}}) + \eta(\delta, \widehat{\mathcal{Y}})l_{\mathrm{h}}||\hat{g}(\hat{u}_{\mathrm{f}}^*)||\delta}_{\mathcal{C}_2(\delta, \widehat{\mathcal{Y}})},
\end{aligned} \tag{2.23}
$$

$||\hat{y}(u_{\mathrm{f}}^*) - \tilde{y}(u_{\mathrm{f}}^*)|| \leq \mathcal{C}_2(\delta, \widehat{\mathcal{Y}})$ and $||\bar{y}(u_{\mathrm{f}}^*) - \tilde{y}(u_{\mathrm{f}}^*)|| \leq 2\mathcal{C}_2(\delta, \widehat{\mathcal{Y}})$. Finally, since $\check{u}_{\mathrm{f}}$ is the solution to the outer problem of (2.16) while u_{f}^* is feasible, $c_{\text{worst}}(\check{u}_{\mathrm{f}}) \leq c_{\text{worst}}(u_{\mathrm{f}}^*) = ||\tilde{y}_{\mathrm{f}}(u_{\mathrm{f}}^*)||^2 + ||u_{\mathrm{f}}^*||^2$, based on which we have

$$\begin{aligned}\bar{c}(\check{u}_{\mathrm{f}}) - \bar{c}(u_{\mathrm{f}}^*) &\le c_{\mathrm{worst}}(\check{u}_{\mathrm{f}}) - \bar{c}(u_{\mathrm{f}}^*)\\ &\le c_{\mathrm{worst}}(u_{\mathrm{f}}^*) - \bar{c}(u_{\mathrm{f}}^*)\\ &= \|\tilde{y}_{\mathrm{f}}(u_{\mathrm{f}}^*)\|^2 + ||u_{\mathrm{f}}^*||^2 - \bar{c}(u_{\mathrm{f}}^*)\\ &\le ||\bar{y}_{\mathrm{f}}(u_{\mathrm{f}}^*)||^2 + ||u_{\mathrm{f}}^*||^2 + ||\tilde{y}_{\mathrm{f}}(u_{\mathrm{f}}^*) - \bar{y}_{\mathrm{f}}(u_{\mathrm{f}}^*)||^2 - \bar{c}(u_{\mathrm{f}}^*) +\\ &\quad 2(||\hat{y}_{\mathrm{f}}(u_{\mathrm{f}}^*)|| + ||\bar{y}_{\mathrm{f}}(u_{\mathrm{f}}^*) - \hat{y}_{\mathrm{f}}(u_{\mathrm{f}}^*)||) \cdot ||\tilde{y}_{\mathrm{f}}(u_{\mathrm{f}}^*) - \bar{y}_{\mathrm{f}}(u_{\mathrm{f}}^*)|| \qquad (2.24)\\ &\le 2(\eta(\delta, \widehat{\mathcal{Y}}) \left\|\widehat{Y}_{\mathrm{f}}\right\| \cdot ||\hat{g}(\hat{u}_{\mathrm{f}}^*)|| + \mathcal{C}_2(\delta, \widehat{\mathcal{Y}})) \cdot 2\mathcal{C}_2(\delta, \widehat{\mathcal{Y}}) + 4(\mathcal{C}_2(\delta, \widehat{\mathcal{Y}}))^2\\ &= \underbrace{8(\mathcal{C}_2(\delta, \widehat{\mathcal{Y}}))^2 + 4\left\|\widehat{Y}_{\mathrm{f}}\right\| \cdot ||\hat{g}(\hat{u}_{\mathrm{f}}^*)||\eta(\delta, \widehat{\mathcal{Y}})\mathcal{C}_2(\delta, \widehat{\mathcal{Y}})}_{\mathcal{C}_3(\delta, \widehat{\mathcal{Y}})}.\end{aligned}$$

As δ goes to 0, the noisy measurement $\widehat{\mathcal{Y}}$ converges in probability to $\overline{\mathcal{Y}} := (Y_{\mathrm{p}}, Y_{\mathrm{f}}, y_{\mathrm{p}})$. Therefore, we see that $\sigma_{\min}^{-1}(\widehat{H})$ converges to in probability to $\sigma_{\min}^{-1}(\overline{H})$, $\eta(\delta, \widehat{\mathcal{Y}})$ to 1 and thus $\mathcal{C}_3(\delta, \widehat{\mathcal{Y}})$ to 0. □

Now, we see that as the measurement noise diminishes, the regulation cost resulting from the solution to (2.16) decreases to the minimal value. This is different from the suboptimality bound in [19] where the achieved regulation cost is only shown to be less than twice the minimal value.

Remark 2.4. *If the measurement noise sequence is i.i.d., by sampling the same historical trajectories for N times to construct $\widehat{Y}_{\mathrm{p},i}, \widehat{Y}_{\mathrm{f},i}, \hat{y}_{\mathrm{p},i}$ for $i = 1, \ldots, N$ (representing instances of $Y_{\mathrm{p}}, Y_{\mathrm{f}}, y_{\mathrm{p}}$ with independent noise realizations) and calculating the average values $\widehat{Y}_{\mathrm{p}}^{\mathsf{avg}} = (1/N)\sum_{i=1}^N \widehat{Y}_{\mathrm{p},i}, \widehat{Y}_{\mathrm{f}}^{\mathsf{avg}} = (1/N)\sum_{i=1}^N \widehat{Y}_{\mathrm{f},i}, \hat{y}_{\mathrm{p}}^{\mathsf{avg}} = (1/N)\sum_{i=1}^N \hat{y}_{\mathrm{f},i}$, we can attenuate the influence of noise. Specifically, given any $\epsilon > 0$, there exists $0 < \delta^{\mathsf{new}}(N, \epsilon) < \delta$ such that $\delta^{\mathsf{new}}(N, \epsilon)$ converges to 0 as N goes to infinity and*

$$\max\{||\widehat{Y}_{\mathrm{p}}^{\mathsf{avg}} - Y_{\mathrm{p}}||_{\max}, ||\widehat{Y}_{\mathrm{f}}^{\mathsf{avg}} - Y_{\mathrm{f}}||_{\max}, ||\hat{y}_{\mathrm{p}}^{\mathsf{avg}} - y_{\mathrm{p}}||_{\max}\} \le \delta^{\mathsf{new}}(N, \epsilon)$$

holds with a probability of $1 - \epsilon$. The averaging technique can be used to satisfy Assumption 2.3. Moreover, it allows one to conduct a sampling complexity analysis for the robust control scheme (2.16) *(i.e., upperbounding the number of samples required to achieve a given suboptimality level), similar with the ones conducted in [89, 32] for model-based schemes.*

Remark 2.5. *If the formulation* (2.16) *is extended to solve a trajectory tracking problem with an objective function $\|y_{\mathrm{f}} - y_{\mathrm{ref}}\|_Q + \|u_{\mathrm{f}}\|_R$ with positive semidefinite matrices $Q \in \mathbb{R}^{p\times p}$ and $R \in \mathbb{R}^{m\times m}$, it is easy to modify Theorem 2.4 for upperbounding the suboptimality gap.*

2.5 Extensions

2.5.1 MIMO systems

According to Remark 2.1, for MIMO systems we cannot bound the prediction error $||\hat{y}(u_\mathrm{f}) - \bar{y}(u_\mathrm{f})||$ resulting from the scheme (2.6) and thus cannot analyse the suboptimality of our robust control framework (2.16). Here, we propose a method where Page matrices are built seperately for each output.

Suppose we have outputs $y^1, y^2, \dots, y^p$. For each output y^i, we have a MISO sub-system with minimal realization

$$\begin{aligned} x^i_{t+1} &= A^i x^i_t + B^i u_t, \\ y^i_t &= C^i x^i_t + D^i u_t, \end{aligned} \tag{2.25}$$

where the matrices A^i, B^i, C^i, D^i are unknown. We build, according to (2.7), the Page matrices U^i_p, U^i_f, $\widehat{Y}^i_\mathrm{p}$, $\widehat{Y}^i_\mathrm{f}$, $\widehat{H}^i$ along with the recent vectors u^i_p, u^i_f, $\hat{y}^i_\mathrm{p}$ for each i. Similar with (2.16), we can write the robust control problem as

$$\min_{u_\mathrm{f}} \max_{\Delta, g, y_\mathrm{f}} \quad ||u_\mathrm{f}||^2 + \sum_{i=1}^{n} ||y^i_\mathrm{f}||^2 \tag{2.26a}$$

$$\text{s.t.} \quad \forall i, \quad \begin{bmatrix} U^i_\mathrm{p} \\ \widehat{Y}^i_\mathrm{p} \\ U^i_\mathrm{f} \\ \widehat{Y}^i_\mathrm{f} \end{bmatrix} g^i + \begin{bmatrix} 0 \\ \Delta^i_{Y_\mathrm{p}} \\ 0 \\ \Delta^i_{Y_\mathrm{f}} \end{bmatrix} g^i = \begin{bmatrix} u^i_\mathrm{p} \\ \hat{y}^i_\mathrm{p} \\ u^i_\mathrm{f} \\ y^i_\mathrm{f} \end{bmatrix} + \begin{bmatrix} 0 \\ \Delta^i_{y_\mathrm{p}} \\ 0 \\ 0 \end{bmatrix} \tag{2.26b}$$

$$\max\Big(||\Delta^i_{Y_\mathrm{p}}||_{\max}, ||\Delta^i_{Y_\mathrm{f}}||_{\max}, ||\Delta^i_{y_\mathrm{p}}||_\infty\Big) \le \delta \tag{2.26c}$$

$$||g^i - \hat{g}^i(u_\mathrm{f})||^2 \le (\mathcal{C}^i(u_\mathrm{f})\delta)^2, \tag{2.26d}$$

where $\mathcal{C}^i(u_\mathrm{f}) = 2\sigma^{-1}_{\min}(\widehat{H}^i)(\sqrt{l^i_\mathrm{p}} + l^i_\mathrm{h}||\hat{g}^i(u_\mathrm{f})||)$.

To justify this formulation, we notice that, Theorem 2.3 and 2.4 can be applied to every sub-system in (2.25). Summing up all the suboptimality bounds, we can get the suboptimality bound for the whole system. The alternating method in [90] is also applicable to solve (2.26).

2.5.2 Input and output constraints

Suppose now there are an input constraint $u_\mathrm{f} \in \mathcal{U}$ and an output constraint $(\bar{y}^1_\mathrm{f}(u_\mathrm{f}), \dots, \bar{y}^p_\mathrm{f}(u_\mathrm{f})) \in \mathcal{Y}$. The input constraint $u_\mathrm{f} \in \mathcal{U}$ can be easily embedded into the outer optimization problem in (2.16). To ensure satisfaction of the output constraint, we have to consider the prediction error and regard as infeasible the sequence u_f that has the slightest chance of violating the output constraint. By applying the inequalities in Theorem 2.2 to every MISO subsystem, we derive the prediction error bound $\mathcal{E}^i(u_\mathrm{f}, \delta)$ for the i-th output, i.e.,

$\|\hat{y}_{\mathrm{f}}^i(u_{\mathrm{f}}) - \bar{y}_{\mathrm{f}}^i(u_{\mathrm{f}})\| \leq \mathcal{E}^i(u_{\mathrm{f}}, \delta)$. Therefore, we can guarantee that $\bar{y}_{\mathrm{f}}^i(u_{\mathrm{f}}) \in \mathcal{B}(\hat{y}_{\mathrm{f}}^i(u_{\mathrm{f}}), \mathcal{E}^i(u_{\mathrm{f}}, \delta))$. Let

$$\mathcal{Y}_{\mathrm{f}}(u_{\mathrm{f}}) := \{(y_{\mathrm{f}}^1, \ldots, y_{\mathrm{f}}^p) | y_{\mathrm{f}}^i \in \mathcal{B}(\hat{y}_{\mathrm{f}}^i(u_{\mathrm{f}}), \mathcal{E}^i(u_{\mathrm{f}}, \delta)), \forall i\}$$

be the region where the output sequence might lie. Then, we can ensure the satisfaction of the output constraints by enforcing the input in (2.26) to additionally verify that

$$\mathcal{Y}_{\mathrm{f}}(u_{\mathrm{f}}) \subset \mathcal{Y}. \tag{2.27}$$

We call the optimization problem (2.26) under (2.27) Safe Data-Driven Minmax Control (SDDMC) . Enforcing (2.27) for general $\mathcal{Y}$ can be challenging [91]. However, if

$$\mathcal{Y} = \{(\bar{y}_{\mathrm{f}}^1, \ldots, \bar{y}_{\mathrm{f}}^p) | y_{-}^i \leq \bar{y}_{\mathrm{f}}^i \leq y_{+}^i, \forall i\}$$

is a box constraint, the constraint (2.27) translates to another box constraint, which is for any i,

$$y_{-}^i + \mathcal{E}^i(u_{\mathrm{f}}, \delta) \leq \hat{y}_{\mathrm{f}}^i(u_{\mathrm{f}}) \leq y_{+}^i - \mathcal{E}^i(u_{\mathrm{f}}, \delta). \tag{2.28}$$

2.6 Numerical Studies

In this section, through numerical experiments, we elaborate on the implementation details, verify the theoretical results and test the performance of SDDMC.

2.6.1 Trajectory regulation for a SISO system

To begin with, we consider a SISO system with the following system matrices

$$A = 0.99 * \begin{bmatrix} 0.7 & 0.2 & 0 \\ 0.3 & 0.7 & -0.1 \\ 0 & -0.2 & 0.8 \end{bmatrix},\ B = \begin{bmatrix} 1 \\ 2 \\ 1.5 \end{bmatrix},\ C = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix},\ D = 0. \tag{2.29}$$

We assume these matrices along with the observability index $l_{\mathrm{o}} = 3$ are unknown to us. We aim to identify the observability index l_{o}, construct the relevant Page matrices, observe the trajectory prediction error and then solve a robust control problem.

We collect historical data by exciting the system according to (2.7) with $\mathcal{Q} = \mathcal{N}(0, 4)$, $L = 8$ and $T = 160$. The i.i.d. measurement noise in the simulation is sampled from the uniform distribution on $[-\delta, \delta]$ with $\delta = 10^{-3}$. Algorithm 1 returns the correct observability index $l_{\mathrm{o}} = 3$. Now we can derive the matrices $U_{\mathrm{p}}, U_{\mathrm{f}}, Y_{\mathrm{p}}, Y_{\mathrm{f}}$ according to Assumption 2.2. We use (2.6) to predict the future output with a horizon of length 3 following a fixed recent trajectory. We notice that if $\widehat{H}$ is constructed using (2.6) and (2.5) with $l_{\mathrm{p}} = l_{\mathrm{o}} = 3$, the prediction error is 1.8×10^{-2}. If we set $l_{\mathrm{p}} = 4$, the error becomes 1.4×10^{4}. Through this comparison, we see the important role of observability index identification for the prediction scheme (2.6).

We then test our robust control scheme (2.16) on the SISO system (2.29). With $l_\mathrm{p} = l_\mathrm{f} = 3$ and the recent trajectory $(u_\mathrm{p}, y_\mathrm{p})$ where $u_\mathrm{p} = [-5.2254, 7.2684, -22.5535]^\top$ and $y_\mathrm{p} = [-1.1242, -23.7291, 13.3406]^\top$, we aim to achieve a minimum regulation cost $\bar{c}(u_\mathrm{f}) = 10u_\mathrm{f}^\top * u_\mathrm{f} + \bar{y}_\mathrm{f}(u_\mathrm{f})^\top \bar{y}_\mathrm{f}(u_\mathrm{f})$. We use the scheme (2.16) to address this regulation problem.

We elaborate on how to solve (2.16). To use the alternating method, we initialize the input sequence by $u_\mathrm{f} = \hat{u}_\mathrm{f}^*$ (defined in (2.17)). When the input is fixed, the inner maximization problem of (2.16) is solved through IPOPT [92] using the primal-dual barrier approach, where the feasible solution $(\Delta_{Y_\mathrm{p}}, \Delta_{Y_\mathrm{f}}, \Delta_{y_\mathrm{p}}, g) = (0, 0, \widehat{Y}_\mathrm{p}\hat{g}(u_\mathrm{f}) - \hat{y}_\mathrm{p}, \hat{g}(u_\mathrm{f}))$ is set as the initial point. After deriving $(\widetilde{\Delta}_{Y_\mathrm{p}}, \widetilde{\Delta}_{Y_\mathrm{f}}, \widetilde{\Delta}_{y_\mathrm{p}}, \tilde{g})$, the solution to the inner problem, the outer problem of (2.16) is simply

$$\min_{u_\mathrm{f}} \widetilde{G}(\widetilde{\Delta}_{Y_\mathrm{p}}, \widetilde{\Delta}_{Y_\mathrm{f}}, \widetilde{\Delta}_{y_\mathrm{p}}, \tilde{g}, u_\mathrm{f}) := \left\| (\widetilde{Y}_\mathrm{f} + \widetilde{\Delta}_{Y_\mathrm{f}}) \begin{bmatrix} U_\mathrm{p} \\ \widehat{Y}_\mathrm{p} + \widetilde{\Delta}_{Y_\mathrm{p}} \\ U_\mathrm{f} \end{bmatrix}^\dagger \begin{bmatrix} u_\mathrm{p} \\ y_\mathrm{p} + \widetilde{\Delta}_{y_\mathrm{p}} \\ u_\mathrm{f} \end{bmatrix} \right\| + \|u_\mathrm{f}\|^2,$$

where the objective is a quadratic function of u_f and the explicit expression of the solution can be readily computed. With the new input sequence at hand, we can start the next iteration. We terminate this algorithm when the difference of input sequences derived in two iterations has a 2-norm less than 10^{-4}. Currently, we do not have any results on the convergence for this iterative scheme. Empirically, we observe in solving the regulation problem defined above that at most three iterations are needed before the termination even when δ increases to 1.

Then, we solve 50 instances of the regulation problem with independent noise realizations. To evaluate the performance of the control scheme (2.16), we calculate for each instance the relative suboptimality, defined as $\frac{\bar{c}(\breve{u}_\mathrm{f}) - \bar{c}(u_\mathrm{f}^*)}{\bar{c}(\breve{u}_\mathrm{f})}$. In Fig. 2.1, we show the mean relative suboptimality for different δ. We see that the input sequence derived based on (2.16) and the alternating method achieves a suboptimality gap that decreases to 0 as the noise diminishes, which coincides with Theorem 2.4. Although we find that with $\delta > 10^{-2}$ Assumption 2.3 is not satisfied and thus Theorem 2.4 is not valid, the relative suboptimality resulting from the robust control scheme (2.27) is empirically small even when $\delta = 1$. When the measurement noise is i.i.d., we can weaken Assumption 2.3 such that Theorem 2.1 and Theorem 2.4 are valid for larger noise level for a high probability. However, to achieve this, we need to use random matrix analysis to tighten (2.12), which is out of the scope of this chapter.

2.6.2 Application to a MIMO system with constraints: room temperature control

We apply SDDMC proposed in Section 2.5.2 to room temperature control. We consider a small building model taken from [93], where the temperature dynamic is normalized and

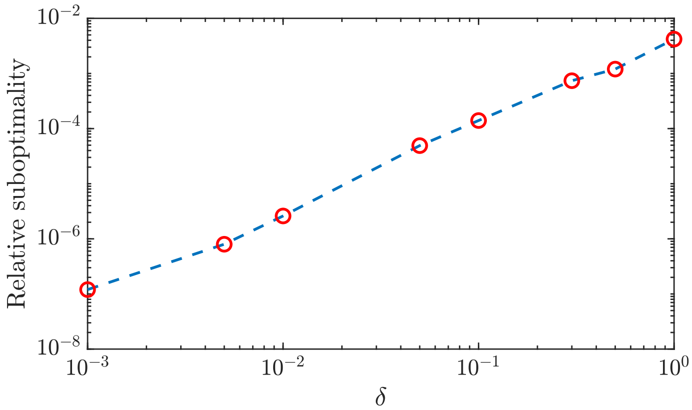

Figure 2.1: Mean relative suboptimality for different δ

linearized at the equilibrium temperature $T = 15°\mathrm{C}$ (which coincides with the outdoor temperature), with a sampling time of 0.5 hour. The model is described by (2.1), with

$$A = \begin{bmatrix} 0.8511 & 0.0541 & 0.0707 \\ 0.1293 & 0.8635 & 0.0055 \\ 0.0989 & 0.0032 & 0.7541 \end{bmatrix}, \quad B = \begin{bmatrix} 0.07 \\ 0.006 \\ 0.004 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, D = 0.$$

The three states represent the temperature at different spots while only two of them can be measured as indicated in the matrix C. Notice that, due to normalization, a state with a value x translates to $(x+15)°\mathrm{C}$ at the corresponding spot. We assume that, at time $t = 0$, the system is at equilibrium. Therefore, the indoor and outdoor temperature is 15°C. To increase the indoor temperature to 25°C and balance between user comfort and energy consumption, we let the cost function be $\bar{c}(u_\mathrm{f}) = 10u_\mathrm{f}^\top * u_\mathrm{f} + (\bar{y}_\mathrm{f}(u_\mathrm{f}) - 10 * [1\ 1]^\top)^\top (\bar{y}_\mathrm{f}(u_\mathrm{f}) - 10 * [1\ 1]^\top)$ covering the horizon $0 \le t \le 4$, while enforcing the the first output to be no less than 5 for $t \ge 1$, i.e., the first spot to be at least 20°C in half an hour after the experiment starts.

We run SDDMC in 50 experiments with independent noise realizations ($\delta = 0.01$). We show in Fig. 2.2 the trajectory of the first output. We also compare with PEM in [87] where the prediction given by (2.6) is regarded as the true output. From Fig. 2.2, we observe that the lower bound 20°C for the first output is always satisfied if SDDMC is applied while PEM generates several trajectories with constraint violations. Meanwhile, we also plot the mean relative suboptimality in Fig. 2.3 for different δ. The suboptimality decreases to 0 when δ diminishes. A source of suboptimality, when δ gets larger, is the conservative estimate of the prediction error bound. This is also the main reason why in Fig. 2.2 the output trajectory is substantially far away from the constraint boundaries. If one can shrink the uncertainty set, the output trajectory can get closer to the boundary, therefore leading to better optimality.

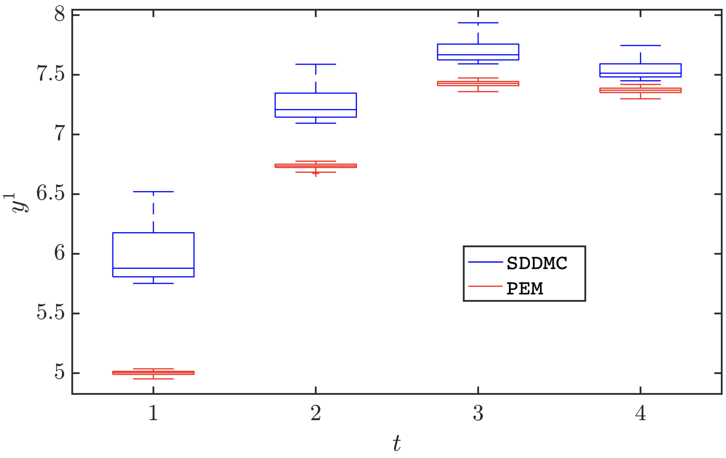

Figure 2.2: Trajectory of the first output: SDDMC v.s. PEM ($\delta = 0.01$)

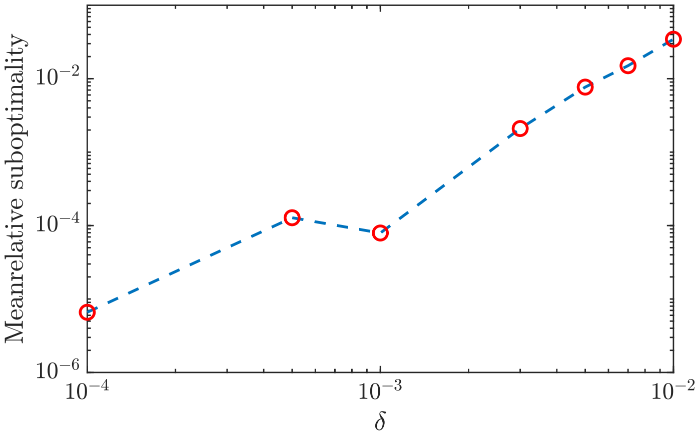

Figure 2.3: Room temperature control: mean relative suboptimality for different δ

2.7 Appendices

2.7.1 A preliminary lemma for the proof of Theorem 2.1

Lemma 2.3. *If the system* (2.1) *is SISO and Assumption 2.2 holds, we let*

$$\mathcal{W} := \begin{bmatrix} u_{L+1} & \cdots & u_{2L} \\ u_{3L+1} & \cdots & u_{4L} \\ \vdots & \vdots & \vdots \\ u_{2Ll_L-L+1} & \cdots & u_{2Ll_L} \end{bmatrix},$$

$$\mathcal{X} := \begin{bmatrix} Cx_{L+1} & \cdots & CA^{l_{\mathrm{p}}-1}x_{L+1} \\ Cx_{3L+1} & \cdots & CA^{l_{\mathrm{p}}-1}x_{3L+1} \\ \vdots & \vdots & \vdots \\ Cx_{2Ll_L-L+1} & \cdots & CA^{l_{\mathrm{p}}-1}x_{2Ll_L-L+1} \end{bmatrix},$$

$$\mathcal{V} := \begin{bmatrix} \mathcal{W} & \mathcal{X} \end{bmatrix}. \tag{2.30}$$

Then

$$\mathbb{P}(\{\mathcal{V} \text{ is full rank}\}) = 1.$$

Proof. We first notice that $l_{\mathrm{p}} = l_{\mathrm{o}} \geq n_x$, otherwise, for a single-output system, $\mathcal{O}(l_{\mathrm{p}})$ cannot be column full rank. By the Cayley-Hamilton theorem, we know $\mathcal{O}(l_{\mathrm{p}})$ has the same row rank as $\mathcal{O}(n_x)$. Since $\mathcal{O}(l_{\mathrm{p}})$ is full column rank, $\mathcal{O}(n_x)$ is also. According to Definition 2.2, $l = l_{\mathrm{p}}$ is the smallest such that $\mathcal{O}(n_x)$ is full row rank and thus $l_{\mathrm{p}} \leq n_x$. Considering these facts, we conclude that $l_{\mathrm{p}} = n_x$. Now we proceed by proving that the event $F_{\mathcal{X}} = \{\mathcal{X}_{1:l_{\mathrm{p}},\cdot}$ is full row rank$\}$ takes place almost surely.

In order to show $\mathbb{P}(F_{\mathcal{X}}) = 1$, we use an induction argument. We first look into the first row of $\mathcal{X}$. According to the system dynamic (2.1), we have

$$x_{L+1} = \begin{bmatrix} B & AB & \cdots & A^{L-1}B \end{bmatrix} \begin{bmatrix} u_L & u_{L-1} & \cdots & u_1 \end{bmatrix}^\top, \quad \mathcal{X}_{1,\cdot} = x_{L+1}^\top \mathcal{O}(l_{\mathrm{p}})^\top. \tag{2.31}$$

Since $x_{L+1} = 0$ defines a subspace in $\mathbb{R}^{mL}$ where $u_{[1,L]}$ resides, by using (2.8) we have $\mathbb{P}\{\mathcal{X}_{1,\cdot} = 0\} = \mathbb{P}\{x_{L+1} = 0\} = 0$. Then, in the following, we show that, for any $l < l_{\mathrm{p}}$,

$$\mathbb{P}\{\mathcal{X}_{1:(l+1),\cdot} \text{ has row full rank} | \mathcal{X}_{1:l,\cdot} \text{ has row full rank}\} = 1. \tag{2.32}$$

To this end, we let Θ be the event where the inputs are fixed from $t = 1$ to $t = (2l-1)L$ and suppose $\Theta \in \{\mathcal{X}_{1:l,\cdot}$ has row full rank$\}$, then there exists a vector $x_{\mathrm{n}} \in \mathbb{R}^{l_{\mathrm{p}}}$ and $x_{\mathrm{n}} \neq 0$ such that $\mathcal{X}_{i,\cdot}x_{\mathrm{n}} = 0$ holds for any $i \leq l$. If $\mathcal{X}_{1:(l+1),\cdot}$ does not have full rank, we have

$$\mathcal{X}_{l+1,\cdot}x_{\mathrm{n}} = 0. \tag{2.33}$$

By noticing $\mathcal{X}_{l+1,\cdot} = x_{(2l+1)L+1}^\top \mathcal{O}(l_\mathrm{p})^\top$, we can rewrite (2.33) into

$$x_\mathrm{n}^\top \mathcal{O}(l_\mathrm{p}) x_{(2l-1)L+1} + x_\mathrm{n}^\top \mathcal{O}(l_\mathrm{p}) \begin{bmatrix} B & \cdots & A^{2L-1}B \end{bmatrix} \begin{bmatrix} u_{(2l+1)L} & \cdots & u_{(2l-1)L+1} \end{bmatrix}^\top = 0. \quad (2.34)$$

By noticing that $\mathcal{O}(l_\mathrm{p})$ is full rank and $x_\mathrm{n}^\top \mathcal{O}(l_\mathrm{p}) \neq 0$, we see $x_\mathrm{n}^\top \mathcal{O}(l_\mathrm{p}) \begin{bmatrix} B & \cdots & A^{2L-1}B \end{bmatrix} \neq 0$ since $\begin{bmatrix} B & \cdots & A^{2L-1}B \end{bmatrix}$ is full row rank. Then, the realizations of $u_{(2l-1)L+1}, \ldots, u_{(2l+1)L}$ avoids (2.34) almost surely due to (2.8). Therefore, the claim (2.32) holds because

$$\mathbb{P}\{\mathcal{X}_{1:(l+1),\cdot} \text{ has row full rank} | \Theta\} = 1.$$

Due to (2.32), if $l < l_\mathrm{p}$ and $\mathbb{P}\{\mathcal{X}_{1:l,\cdot} \text{ has row full rank}\} = 1$, then $\mathbb{P}\{\mathcal{X}_{1:(l+1),\cdot} \text{ has row full rank}\} = 1$. By induction, we know $\mathbb{P}\{\mathcal{X}_{1:l_\mathrm{p},\cdot} \text{ has row full rank}\} = 1$.

Now, almost surely, the first l_p rows of $\mathcal{X}$ are linear independent while the last L rows of $\mathcal{W}$ are linear independent. Since the inputs in the $(l_\mathrm{p}+i)$-th row of $\mathcal{W}$ are independent of the elements in the $(l_\mathrm{p}+i)$-th row of $\mathcal{X}$ for $1 \leq i \leq L$, we can again use the induction technique to show that the first $(l_\mathrm{p}+i)$ rows of $\mathcal{V}$ are linearly independent almost surely for $1 \leq i \leq L$. Thus, we have that $\mathbb{P}(\{\mathcal{V} \text{ is full rank}\}) = 1$. □

2.7.2 Proof of Theorem 2.1

In this proof, we **only** consider single-input cases for simplicity since it can be easily adapted to multiple-input cases. The main idea is to exploit the relationship between $\overline{H}$ and $\mathcal{V}$ in Lemma 2.3.

We let $l_L = L + l_p$ and denote $\overline{H}^\circ$ as the submatrix consisting of the $2, 4, \ldots, 2l_L$-th columns of $\overline{H}$. We notice that for any i, j,

$$y_{L(2i-1)+j} = CA^{j-1} x_{L(2i-1)+1} + \sum_{k=1}^{j-1} CA^k B u_{L(2i-1)+k}. \quad (2.35)$$

By substituting (2.35) into $\overline{H}^\circ$ and using elementary row transformation to simplify $\overline{H}^\circ$, we see that $\overline{H}^\circ$ is full row rank if $\mathcal{V}$, defined in (2.30), has full rank. From Lemma 2.3, we have that $\mathbb{P}(\{\overline{H} \text{ is full row rank}\}) \geq \mathbb{P}(\{\overline{H}^\circ \text{ is full row rank}\}) \geq \mathbb{P}(\{\mathcal{V} \text{ is full rank}\}) = 1$.

2.7.3 Proof of Proposition 2.1

The proof for the case $l_\mathrm{p} < l_\mathrm{o}$ is an easy modification of that for Theorem 2.1 and thus omitted. If $l_\mathrm{p} > l_\mathrm{o}$, we notice that there exist matrices $\mathcal{A}_i \in \mathbb{R}^{p \times p}$, $\mathcal{F}_i \in \mathbb{R}^{p \times m}$ for $1 \leq i \leq l_\mathrm{o}$

such that the l_{o}th-order ARX model

$$y_t = \sum_{i=1}^{l_{\mathrm{o}}} (\mathcal{A}_i y_{t-i} + \mathcal{F}_i u_{t-i}) \tag{2.36}$$

describes exactly the system (2.1). Therefore, we almost surely have that the $(l_{\mathrm{o}}p+1)$-th row of Y_{p} is a linear combination of the first $l_{\mathrm{o}}p$ rows of U_{p} and the first $l_{\mathrm{o}}p$ rows of Y_{p}, which concludes the proof.

2.7.4 A supporting lemma for Proposition 2.2

Lemma 2.4. *Given a full-row-rank matrix $A \in \mathbb{R}^{m\times n}$ with $m < n$ and a row vector $z \in \mathbb{R}^{m\times n}$, we have that*

$$\sigma_{\min}([A^\top z^\top]^\top) \le \sigma_{\min}(A).$$

Proof. According to the definition of singular values, we have

$$\begin{aligned}\sigma_{\min}([A^\top z^\top]^\top) &= \min_{v\in\mathbb{R}^{m+1},\|v\|=1} \|[A^\top z^\top]v\| \\ &\le \min_{v\in\mathbb{R}^{m+1},\|v\|=1,v_{m+1}=0} \|[A^\top z^\top]v\| \\ &\le \min_{w\in\mathbb{R}^{m},\|w\|=1} \|A^\top v\| = \sigma_{\min}(A)\end{aligned} \tag{2.37}$$

□

2.7.5 Proof of Lemma 2.2

By substituting (2.6) to (2.17), we have

$$\hat{u}_{\mathrm{f}}^* = \operatorname*{argmin}_{u_{\mathrm{f}}} \, (\widehat{K}_1\hat{b}(0) + \widehat{K}_2 u_{\mathrm{f}})^\top (\widehat{K}_1\hat{b}(0) + \widehat{K}_2 u_{\mathrm{f}}) + u_{\mathrm{f}}^\top u_{\mathrm{f}}. \tag{2.38}$$

The solution is $\hat{u}_{\mathrm{f}}^* = -(\widehat{K}_2^\top \widehat{K}_2 + I)^{-1}\widehat{K}_2^\top \widehat{K}_1\hat{b}(0)$. With $\overline{K}_1$, $\overline{K}_2$ being the noiseless counterparts of $\widehat{K}_1$, $\widehat{K}_2$ respectively, we have $u_{\mathrm{f}}^* = -(\overline{K}_2^\top \overline{K}_2 + I)^{-1}\overline{K}_2^\top \overline{K}_1\bar{b}(0)$. In the following, we aim to bound $||\hat{u}_{\mathrm{f}}^* - u_{\mathrm{f}}^*||$.

Firstly, we analyse the influence of noise on $\widehat{K}_1$ and $\widehat{K}_2$. Due to Assumption 2.3,

$$\sigma_{\min}(\overline{H}) \ge \sigma_{\min}(\widehat{H}) - ||E|| \ge \frac{1}{2}\sigma_{\min}(\widehat{H}).$$

Thus, $||\overline{H}^\dagger|| = \sigma_{\min}^{-1}(\overline{H}) \le 2\sigma_{\min}^{-1}(\widehat{H})$. Considering the following conclusion in perturbation analysis [94],

$$||\overline{H}^\dagger - \widehat{H}^\dagger|| \le 2\max\{||\overline{H}^\dagger||^2, ||\widehat{H}^\dagger||^2\}||E|| \tag{2.39}$$

we have $||\overline{H}^{\dagger} - \widehat{H}^{\dagger}|| \leq 8l_{\mathrm{h}}||\widehat{H}^{\dagger}||^2\delta$ and

$$\begin{aligned}||\widehat{K}_1 - \overline{K}_1|| &\leq ||\widehat{Y}_{\mathrm{p}}|| \cdot ||\widehat{H}^{\dagger} - H^{\dagger}|| + ||\widehat{Y}_{\mathrm{p}} - Y_{\mathrm{p}}|| \cdot ||\overline{H}^{\dagger}|| \\ &\leq 8l_{\mathrm{h}}||\widehat{Y}_{\mathrm{p}}|| \cdot ||\widehat{H}^{\dagger}||^2\delta + 2l_{\mathrm{h}}||\widehat{H}^{\dagger}||\delta \\ &= \underbrace{2l_{\mathrm{h}}||\widehat{H}^{\dagger}||(1 + 4||\widehat{Y}_{\mathrm{p}}|| \cdot ||\widehat{H}^{\dagger}||)\delta}_{\mathcal{F}_1(\delta,\widehat{\mathcal{Y}})}.\end{aligned} \tag{2.40}$$

Therefore, $||\widehat{K}_2 - \overline{K}_2|| \leq \mathcal{F}_1\delta$. Similarly we have

$$\max_{i=1,2}\left\{||\widehat{K}_2^{\top}\widehat{K}_i - \overline{K}_2^{\top}\overline{K}_i||\right\} \leq \underbrace{(2||\widehat{K}_1|| + \mathcal{F}_1(\delta,\widehat{\mathcal{Y}}))\mathcal{F}_1(\delta,\widehat{\mathcal{Y}})}_{\mathcal{F}_2(\delta,\widehat{\mathcal{Y}})}.$$

Secondly, again following (2.39), we have

$$(\widehat{K}_2^{\top}\widehat{K}_2 + I)^{-1} - (\overline{K}_2^{\top}\overline{K}_2 + I)^{-1} \leq 2||\widehat{K}_2^{\top}\widehat{K}_2 - \overline{K}_2^{\top}\overline{K}_2||. \tag{2.41}$$

Through the same technique for deducing (2.40) based on the upperbound of $||\overline{H}^{\dagger} - \widehat{H}^{\dagger}||$, we utilize (2.41) to conclude that

$$\left\|\widehat{K}_2^{\top}\widehat{K}_1\hat{b}(0) - \overline{K}_2^{\top}\overline{K}_1\bar{b}(0)\right\| \leq \underbrace{\sqrt{l_{\mathrm{p}}}\left\|\widehat{K}_2^{\top}\widehat{K}_1\right\|\delta + \left(\left\|\hat{b}(0)\right\| + \sqrt{l_{\mathrm{p}}}\delta\right)\mathcal{F}_2(\delta,\widehat{\mathcal{Y}})}_{\mathcal{F}_3(\delta,\widehat{\mathcal{Y}})},$$

$$||\hat{u}_{\mathrm{f}}^* - u_{\mathrm{f}}^*|| \leq \underbrace{\left\|(\widehat{K}_2^{\top}\widehat{K}_2 + I)^{-1}\right\|\mathcal{F}_3(\delta,\widehat{\mathcal{Y}}) + \left(\left\|\widehat{K}_2^{\top}\widehat{K}_1\hat{b}(0)\right\| + \mathcal{F}_3(\delta,\widehat{\mathcal{Y}})\right)\mathcal{F}_2(\delta,\widehat{\mathcal{Y}})}_{\mathcal{F}(\delta,\widehat{\mathcal{Y}})}$$

and $\mathcal{F}(\delta,\widehat{\mathcal{Y}})$ converges to 0 as δ goes to 0.

Finally, we have $||\hat{g}(u_{\mathrm{f}}^*) - \hat{g}(\hat{u}_{\mathrm{f}}^*)|| \leq ||\widehat{H}^{\dagger}|| \cdot ||\hat{u}_{\mathrm{f}}^* - u_{\mathrm{f}}^*||$ and then

$$||\hat{g}(u_{\mathrm{f}}^*)|| \leq ||\hat{g}(\hat{u}_{\mathrm{f}}^*)|| + ||\widehat{H}^{\dagger}||\mathcal{F}(\delta,\widehat{\mathcal{Y}}) = ||\hat{g}(\hat{u}_{\mathrm{f}}^*)||(1 + \frac{||\widehat{H}^{\dagger}||\mathcal{F}(\delta,\widehat{\mathcal{Y}})}{||\hat{g}(\hat{u}_{\mathrm{f}}^*)||}). \tag{2.42}$$

2.8 Final Remarks

In this chapter, we proposed a method to construct the Page matrices for linear system trajectory prediction such that the error due to the measurement noise can be bounded using solely collected data. Based on this error bound, we designed the minmax robust control scheme such that the suboptimality gap is bounded. This scheme is also extended to solve regulation problems for MIMO systems with input/output constraints. The experiments illustrated that with our method of Page matrix construction we can achieve small prediction error and, by using the proposed robust control method, the suboptimality gap can be small for unconstrained optimal control problems. For constrained problems, we saw that the constraints were respected for different noise realizations.

We observe the conservativeness of our method in the experiment on room temperature control, which not only leads to suboptimality but also might result in infeasibility. The conservativeness mainly stems from the small value of $\sigma_{\min}(\widehat{H})$, i.e., "insufficient persistent excitation". Currently in this chapter, we used randomly generated input sequences. A possible direction for larger $\sigma_{\min}(\widehat{H})$ is experiment design. To the best of the author's knowledge, an efficient data-driven method to design system excitations is still open in the literature. This has to be done through constructing intermediate data-driven system representations as a guidance for designing new excitations.

Safe Zeroth-Order Optimization Part II

3 Safe Zeroth-Order Optimization Using Quadratic Local Approximations

3.1 Introduction

Classical techniques for zeroth-order optimization can be classified as direct-search-based (where a set of points around the current point is searched for a lower value of the objective function), gradient-descent-based (where the gradients are estimated based on samples), and proxy-based (where a local model of the objective function around the current point is built and used for local optimization) [95, Chapter 9]. Examples of these three categories for unconstrained optimization are, respectively, pattern search methods [21], randomized stochastic gradient-free methods [96], and trust region methods [97].

In case the explicit formulations of both objective and constraint functions are not available, the work [98] proposes a proxy-based variant of the Frank-Wolfe algorithm, which enjoys sample feasibility and convergence towards a neighborhood of the optimal point with high probability. However, this method only addresses convex objectives and polytopic constraints. When the unmodelled constraints are nonlinear, one can use two-phase proxi-based methods [99, 100] where an optimization phase reduces the objective function subject to relaxed constraints and a restoration phase modifies the result of the first phase to regain feasibility. A drawback of these approaches is the lack of a guarantee for sample feasibility.

For ensuring sample feasibility, the zeroth-order methods of [25, 101, 102], including proxy-based SafeOPT and its variants, assume the knowledge of a Lipschitz constant of the objective and constraint functions, while [103] utilizes the Lipschitz constants of the gradients of these functions (the so called smoothness constants). With these quantities, one can build local proxies for the constraint functions and provide a confidence interval for the true function values. By starting from a feasible point, [25, 102, 103] utilize the proxies to search for potential minimizers. However, for each search, one may have to use a global optimization method to solve a non-convex problem, which makes the algorithm computationally intractable if there are many decision variables.

Another research direction aimed at the feasibility of the samples is to include barrier functions in the objective to penalize the proximity to the boundary of the feasible set [104, 105]. In this category, extremum seeking methods estimate the gradient of the new objective function by adding sinusoidal perturbations to the decision variables [106]. However, due to the perturbations, these methods have to adopt a sufficiently large penalty coefficient to ensure all the samples fall in the feasible region. This strategy sacrifices optimality since deriving a near-optimal solution requires a small penalty coefficient. In contrast, another gradient-descent-based algorithm, LB-SGD , proposed in [24] uses log-barrier functions and ensures the feasibility of the samples despite a small penalty coefficient. After calculating a descent direction for the cost function with log-barrier penalties, this method exploits the Lipshcitz and smoothness constants of the constraint functions to build local safe sets for selecting the step size of the descent. Although LB-SGD comes with a polynomial worst-case complexity in problem dimension, it might converge slowly, even for convex problems. The reason is that as the iterates approach the boundary of the feasible set, the log-barrier function and its derivative become very large, leading to very conservative local feasible sets and slow progress of the iterates.

Safe zeroth-order optimization has been an increasingly important topic in the learning-based control community. One application is constrained optimal control problems with unknown system dynamics. To solve these problems, traditional model-based methods rely on system identification which might be challenging, for example, when the order of the ground-truth model is unknown or sufficient excitation required for modeling may lead to infeasibility. Data-driven methods based on Willems' lemma [81], such as [46, 11, 13], provide near-optimal controllers that satisfy safety constraints without requiring system identification. However, ensuring the feasibility during data collection to implement these methods remains an open question. An alternative approach to finding feasible solutions to optimal control problems is to design a safety filter [107]. This filter is activated when the reachability analysis indicates the possibility of constraint violation. While successful in practice, it is hard to obtain a convergence guarantee and sample complexity of this approach for learning an optimal controller. In reinforcement learning, Constrained Policy Optimization [108] and Learning Control Barrier Functions [109] (model-free) are used to find the optimal safe controller, but feasibility during training cannot be ensured. Bayesian Optimization can also be applied to optimal control in a zeroth-order manner. For example, [110] proposes Violation-Aware Bayesian Optimization to optimize three set points of a vapor compression system, [111] utilizes SafeOPT to tune a linear control law with two parameters for quadrotors, and [112] implements the Goal Oriented Safe Exploration algorithm in [101] to optimize a PID controller with three parameters for a rotational axis drive. Although these variants of Bayesian Optimization offer guarantees of sample feasibility, they scale poorly to high-dimensional systems due to the non-convexity of the subproblems and the need for numerous samples.

Optimal Power Flow (OPF) is an example of large-scale optimization problems that can benefit from zeroth-order optimization. Its objective is to allocate the active and reactive power generation, transmission line flows and voltage levels to minimize costs while satisfying operational and security constraints such as transmission line capacity and voltage level limits. In recent years, OPF has gained considerable attention due to the rising demand for efficient and reliable operation of power systems, as well as the integration of renewable energy sources and energy storage systems [113]. However, the application of OPF to power system operation is a significant challenge due to the difficulties in accurately deriving a system model. Therefore, we consider applying our model-free method to solve OPF problems.

3.1.1 Contributions

We develop a zeroth-order method for smooth optimization problems with guaranteed sample feasibility and convergence. The approach is based on designing quadratic local proxies of the objective and constraint functions and iteratively solving the Quadractially Constrained Quadratic Programming (QCQP) subproblems. On this algorithm, we have the following results:

(a) We show in Section 3.4.1 that, under mild assumptions, our safe zeroth-order algorithm has iterates whose accumulation points are KKT pairs even for **non-convex** problems.

(b) Besides the asymptotic results in (a), given $\eta > 0$, we add **termination conditions** to the zeroth-order algorithm and guarantee in Section 3.4.2 that the returned primal-dual pair is an **η-KKT** pair (see Definition 3.1) of the optimization problem. We further show in Section 3.5 that under mild assumptions our algorithm terminates in $O(\frac{1}{\eta^2})$ iterations and requires $O(\frac{1}{\eta^2})$ samples.

(c) We present in Section 3.6 an example illustrating that our algorithm achieves faster convergence than state-of-the-art zeroth-order methods that guarantee sample feasibility. We further apply the algorithm to optimal control and optimal power flow problems, showing that the results returned by our algorithm are almost identical to those provided by commercial solvers utilizing the true model.

3.1.2 Structure of this chapter

We introduce the problem formulation in Section 3.2 and propose the algorithm SZO-QQ in Section 3.3. On the properties of SZO-QQ's iterates and output, we show asymptotic results in Section 3.4 and provide complexity analysis in Section 3.5. In Section 3.6, we compare SZO-QQ with other state-of-the-art zeroth-order methods and apply it to optimal control and optimal power flow.

3.1.3 Notations of this chapter

We use $e_i \in \mathbb{R}^d$ to define the i-th standard basis of vector space $\mathbb{R}^d$ and $\|\cdot\|$ to denote the 2-norm throughout the chapter. Given a vector $x \in \mathbb{R}^d$ and a scalar $\epsilon > 0$, we write $x = [x^{(1)}, \dots, x^{(d)}]^\top$, $\mathcal{B}(x, \epsilon) = \{y : \|y - x\| \leq \epsilon\}$ and $\mathcal{SP}(x, \epsilon) = \{y : \|y - x\| = \epsilon\}$. We use $\mathbb{Z}_i^j = \{i, i+1, \dots, j\}$ to define the set of integers ranging from i to j with $i < j$. For two vectors $x, y \in \mathbb{R}^d$, we use $\langle x, y\rangle := x^\top y$ to define the inner product.

3.2 Problem Formulation

We address the constrained optimization problem

$$\min_{x \in \mathbb{R}^d} \; f_0(x) \quad \text{subject to} \;\; x \in \Omega \tag{3.1}$$

with feasible set $\Omega := \{x \in \mathbb{R}^d : f_i(x) \leq 0, i \in \mathbb{Z}_1^m\}$. We consider the setting where the continuously differentiable functions $f_i : \mathbb{R}^d \to \mathbb{R}$ are not explicitly known but can be queried. In this chapter, we aim to derive a local optimization algorithm that for any given $\eta > 0$ returns an η-approximate KKT pair of (3.1) defined as follows.

Definition 3.1. *For $\eta > 0$, a pair (x, λ) with $x \in \Omega$ and $\lambda \in \mathbb{R}^m_{\geq 0}$ is an η-approximate Karush–Kuhn–Tucker (η-KKT for short) pair of the problem (3.1) if*

$$\|\nabla f_0(x) + \sum_{i=1}^m \lambda^{(i)} \nabla f_i(x)\| \leq \eta, \tag{3.2a}$$

$$|\lambda^{(i)} f_i(x)| \leq \eta, \quad i \in \mathbb{Z}_1^m. \tag{3.2b}$$

If $(x^*, \lambda^*) \in \Omega \times \mathbb{R}^m_{\geq 0}$ fulfills (3.2) with $\eta = 0$, we say that it is a KKT pair.

For any optimization problem with differentiable objective and constraint functions for which strong duality holds, any pair of primal and dual optimal points must be a KKT pair. If the optimization problem is furthermore convex, any KKT pair satisfies primal and dual optimality [76]. The optimization methods that aim to obtain a KKT pair, such as Newton-Raphson and interior point methods, might converge to a local optimum. Despite this drawback, these local methods are extensively applied, because local algorithms are more efficient to implement and KKT pairs are good enough for some applications, such as machine learning [114], optimal control [115] and optimal power flow [116]. Considering that, in general, numerical solvers cannot return an exact KKT pair, the concept of η-KKT pair indicates how close primal and dual solutions are to a KKT pair [117]. According to [117, Theorem 3.6], under mild assumptions, one can make the η-KKT pair arbitrarily close (in Euclidean distance) to an KKT pair of (3.1) by decreasing η. In many numerical optimization methods [118, 119], one can trade off accuracy against efficiency by tuning η.

We assume, without loss of generality, the objective function $f_0(x)$ is explicitly known and linear. Indeed, when the function $f_0(x)$ in (3.1) is not known but can be queried, the problem in (3.1) can be written as

$$\begin{aligned} \min_{(x,\gamma)\in\mathbb{R}^{d+1}} \quad & \gamma \\ \text{subject to} \quad & f_0(x) - \gamma \leq 0, \\ & f_i(x) \leq 0, \quad i \in \mathbb{Z}_1^m, \end{aligned}$$

where the objective function is now known and linear. Throughout this chapter, we make the following assumptions on the smoothness of the objective and constraint functions and the availability of a strictly feasible point.

Assumption 3.1. *The functions $f_i(x)$, $i \in \mathbb{Z}_0^m$ are continuously differentiable and we know constants $L_i, M_i > 0$ such that for any x_1, $x_2 \in \mathbb{R}^d$,*

$$|f_i(x_1) - f_i(x_2)| \leq L_i\|x_1 - x_2\|, \tag{3.3a}$$

$$\|\nabla f_i(x_1) - \nabla f_i(x_2)\| \leq M_i\|x_1 - x_2\|. \tag{3.3b}$$

We also assume that the known Lipschitz and smoothness constants L_i and M_i verify that

$$L_i > L_{i,\text{inf}} \text{ and } M_i > M_{i,\text{inf}}, \tag{3.4}$$

where

$$\begin{aligned} L_{i,\text{inf}} &:= \inf\{L_i : \text{ (3.3a) } \textit{holds}, \forall x_1, x_2 \in \Omega\}, \\ M_{i,\text{inf}} &:= \inf\{M_i : \text{ (3.3b) } \textit{holds}, \forall x_1, x_2 \in \Omega\}. \end{aligned}$$

In the remainder of this chapter, we also define $L_{\max} = \max_{i\geq 1} L_i$ and $M_{\max} = \max_{i\geq 1} M_i$.

Remark 3.1. *The bounds in* (3.3) *are utilized in several works on zeroth-order optimization, e.g., [120, 121]. As it will be clear in the sequel, these bounds allow one to estimate the error of local approximations of the unknown functions and their derivatives. In practice, it is usually impossible to obtain $L_{i,\text{inf}}$ and $M_{i,\text{inf}}$, thus we only assume to know the upperbounds $L_i > L_{i,\text{inf}}$ and $M_i > M_{i,\text{inf}}$. In case L_i and M_i are not known, we regard them as hyperparameters and describe how to tune them in Remark 3.3.*

Assumption 3.2. *There exists a known strictly feasible point x_0, i.e., $f_i(x_0) < 0$ for all $i \in \mathbb{Z}_1^m$.*

Remark 3.2. *The existence of a strictly feasible point is called Slater's Condition and is commonly assumed in several optimization methods [76]. Moreover, several works on safe learning [25, 24] assume a strictly feasible point used for initializing the algorithm. Assumption 3.2 is necessary for designing an algorithm whose iterates remain feasible since the constraint functions are unknown a priori. Practically, it holds in several applications. For example, in any robot mission planning, the robot is placed initially at a safe point* *and needs to gradually explore the neighboring regions while ensuring the feasibility of its trajectory. Similarly, in the optimization of manufacturing processes, often an initial set of (suboptimal) design parameters satisfying the problem constraints are known [122]. Another example is frequency control of power grids, where the initial frequency is guaranteed to lie within certain bounds by suitably regulating the power reserves and loads [116].*

Assumption 3.3. *There exists $\beta \in \mathbb{R}$ such that the sublevel set $\mathcal{P}_\beta = \{x \in \Omega : f_0(x) \leq \beta\}$ is bounded and includes the initial feasible point x_0.*

Under Assumption 3.3, for any iterative algorithm producing non-increasing objective function values $\{f_0(x_k)\}_{k\geq 0}$, we ensure the iterates $\{x_k\}_{k\geq 0}$ to be within the bounded set $\mathcal{P}_\beta$.

We highlight that Assumptions 1-3 stand *throughout this chapter*. By exploiting them, we design in the following section an algorithm that iteratively optimizes $f_0(x)$.

3.3 Safe Zeroth-Order Algorithm

Before introducing the iterative algorithm, this section proposes an approach to construct local feasible sets by using samples around a given strictly feasible point. To do so, we recall the properties of a gradient estimator constructed through finite differences.

The gradients of the unknown functions $\{f_i\}_{i=1}^m$ can be approximated as

$$\nabla^\nu f_i(x) := \sum_{j=1}^{d} \frac{f_i(x+\nu e_j) - f_i(x)}{\nu} e_j \tag{3.5}$$

where $\nu > 0$. From Assumption 3.1, we have the following result about the estimation error

$$\Delta_i^\nu(x) := \nabla^\nu f_i(x) - \nabla f_i(x).$$

Lemma 3.1 ([123], Theorem 3.2)**.** *Under Assumption 3.1, we have*

$$\|\Delta_i^\nu(x)\|_2 \leq \alpha_i \nu, \text{ with } \alpha_i = \frac{\sqrt{d}M_i}{2}. \tag{3.6}$$

3.3.1 Local feasible set construction

Based on (3.5) and (3.6), one can build a local feasible set around a strictly feasible point x_0 as follows.

Theorem 3.1. *For any strictly feasible point x_0, let*

$$l_0^* = \min_{i\in\{1,\ldots,m\}} -f_i(x_0)/L_{\max}, \tag{3.7}$$

and $\nu_0^* = l_0^*/\sqrt{d}$*, where* $L_{\max} = \max_{i\geq 1} L_i$*. Define*

$$\mathcal{S}_i^{(0)}(x_0) := \Big\{x : f_i(x_0) + \nabla^{\nu_0^*} f_i\,(x_0)^\top\,(x - x_0) + 2M_i\|x - x_0\|^2 \leq 0\Big\}. \tag{3.8}$$

Under Assumption 3.1, all the samples needed for computing $\nabla^{\nu_0^*} f_i\,(x_0)$ *are feasible and the set* $\mathcal{S}^{(0)}(x_0) := \bigcap_{i=1}^m \mathcal{S}_i^{(0)}(x_0)$ *satisfies* $\mathcal{S}^{(0)}(x_0) \subset \Omega$*.*

For completeness, we include the proof of Theorem 3.1 in Section 3.7.1. By construction, we see that if x_0 is strictly feasible, then x_0 belongs to the interior of $\mathcal{S}^{(0)}(x_0)$ and thus $\mathcal{S}^{(0)}(x_0) \neq \emptyset$. Moreover, the set $\mathcal{S}^{(0)}(x_0)$ is convex since $\mathcal{S}^{(0)}(x_0) = \cap_{i=1}^m \mathcal{S}_i^{(0)}(x_0)$ and $\mathcal{S}_i^{(0)}(x_0)$ is a d-dimensional ball for any i. We call $\mathcal{S}^{(0)}(x_0)$ a *local feasible set around* x_0.

Remark 3.3. *The feasibility of* $\mathcal{S}^{(0)}(x_0)$ *is a consequence of Assumption 3.1. Next, we comment on the missing knowledge of* L_i *and* M_i *verifying* (3.4)*. In this case, the set* $\mathcal{S}^{(0)}(x_0)$ *built based on the initial guesses,* L_i *and* M_i*, might not be feasible. When infeasible samples are generated, one can multiply* L_i *and* M_i *for* $i \in \mathbb{Z}_1^m$ *by* $\beta > 1$*. This way, at most* $m + \sum_{i=1}^m \max\{\log_\beta(L_{i,\inf}/L_i), \log_\beta(M_{i,\inf}/M_i)\}$ *infeasible samples are encountered, where* L_i *and* M_i *are the initial guesses. At the same time, one should avoid using a too large value for* M_i*, since if* $M_i \gg M_{i,\inf}$*, the approximation used to construct* $\mathcal{S}^{(0)}(x_0)$ *can be very conservative. We refer the readers to Theorem 3.4, for a discussion on the growth of the complexity of the proposed method with* $L_{\max} + M_{\max}$*, and Section 3.6, for an example illustrating the impact of* $M_{\max}$ *on the convergence.*

One can find in [24] and [25] a different formulation of local feasible sets. In Section 3.7.2, we compare the two formulations and explain why $\mathcal{S}^{(0)}(x_0)$ is the less conservative.

3.3.2 The proposed algorithm

The proposed method to solve problem (3.1), called Safe Zeroth-Order Sequential QCQP (SZO-QQ) , is summarized in Algorithm 2. The idea is to start from a strictly feasible initial point x_0 and iteratively solve (SP1) in Algorithm 2 until two termination conditions are satisfied. Below, we expand on the main steps of the algorithm.

Providing input data. The input to the Algorithm 2 includes an initial feasible point x_0 (see Assumption 3.2) and three parameters μ, ξ, Λ. We will describe in Section 3.4 the selection of ξ and Λ to ensure that Algorithm 2 returns an η-KKT pair of (3.1). The impact of μ on the convergence will be shown in Theorem 3.4.

Building local feasible sets (Line 4). For a strictly feasible x_k, we use (3.7) to define l_k^* and let the step size of the finite differences for gradient estimation be

$$\nu_k^* = \min\{\frac{l_k^*}{\sqrt{d}}, \frac{1}{k}, \frac{\eta}{12\alpha_{\max} m\Lambda}\}. \tag{3.9}$$

Algorithm 2 Safe Zeroth-Order Sequential QCQP (SZO-QQ)

```
Input: $\mu, \xi, \Lambda > 0$, initial feasible point $x_0 \in \Omega$
Output: $\tilde{x}$, $\tilde{\lambda}$, $\tilde{k}$
 1: Choose $M_i > M_{i,\inf}$, for $i \in \mathbb{Z}_1^m$
 2: $k \leftarrow 0, \mathrm{TER} = 0$
 3: while $\mathrm{TER} = 0$ do
 4:     Compute $\mathcal{S}^{(k)}(x_k)$ based on (4.1) and (3.9).
 5:     Compute the optimal primal and dual solutions $(x_{k+1}, \lambda_{k+1}^{\circ})$ of
            $\min_{x \in \mathcal{S}^{(k)}(x_k)} f_0(x) + \mu\|x - x_k\|^2$        (SP1)
 6:     if $\|x_{k+1} - x_k\| \leq \xi$ then
 7:
            $\lambda_{k+1} \leftarrow \operatorname{argmin}_{\lambda_{k+1} \in \mathbb{R}_+^m} \|\lambda_{k+1}\|_\infty$  s.t. (3.10)        (SP2)
 8:         if $\|\lambda_{k+1}\|_\infty \leq 2\Lambda$ then
 9:             $\tilde{x} \leftarrow x_{k+1}, \tilde{\lambda} \leftarrow \lambda_{k+1}$, $\tilde{k} \leftarrow k + 1$, $\mathrm{TER} \leftarrow 1$
10:         end if
11:     end if
12:     $k \leftarrow k + 1$
13: end while
```

Moreover, we use (4.1) to define $\mathcal{S}^{(k)}(x_k)$ in (SP1). The bounds $\nu_k^* \leq 1/k$ and $\nu_k^* \leq \eta/(12\alpha_{\max} m\Lambda)$ in (3.9) are utilized to verify the approximate KKT conditions (3.2) (see Theorem 3.3).

Solving a subproblem (Line 5). Based on the local feasible set, we formulate the subproblem (SP1) of Algorithm 2. The regulation term $\mu\|x - x_k\|^2$ along with $2M_i\|x - x_k\|$ in $\mathcal{S}_i^{(k)}(x_k)$ prevents too large step sizes. If $\|x - x_k\|$ is large, the proxies used in (SP1) are not accurate. The regulation term can also be found in the proximal trust-region method in [124]. With it, we can ensure that $\|x_{k+1} - x_k\|$ converges to 0 (see Proposition 3.1) and conduct complexity analysis (see Theorem 3.4). Since f_0 is assumed, without loss of generality, to be known and linear (see Section 3.2), (SP1) is a known convex QCQP . We let $(x_{k+1}, \lambda_{k+1}^{\circ})$ be the optimal primal and dual solutions to (SP1). As shown in the proof of Theorem 3.1, the bound $M_i > M_{i,\inf}$ from Assumption 3.1 implies that $x_{k+1} \in \Omega$ is strictly feasible. Although x_k is strictly feasible for any k, it is possible that the sequence $\{x_k\}_{k \geq 1}$ converges to a point on the boundary of Ω.

Checking termination conditions (Line 6-11). We introduce two termination conditions guaranteeing that the pair $(\tilde{x}, \tilde{\lambda})$ returned by Algorithm 2 is an η-KKT pair. The first one (Line 6) requires that $\|x_{k+1} - x_k\|$ is smaller than a given threshold ξ while the second requires that the solution to the optimization problem (SP2) is small enough (Line 8). The constraint of (SP2) is

$$\max\left\{\delta_1(k,\lambda_{k+1}),\max\{\delta_2^{(i)}(k,\lambda_{k+1}):i\geq 1\}\right\}\leq\frac{\eta}{2},\tag{3.10}$$

where

$$\begin{aligned}\delta_1(k,\lambda_{k+1})&:=\left\|\nabla f_0\left(x_{k+1}\right)+2\mu(x_{k+1}-x_k)+\sum_{i=1}^{m}\lambda_{k+1}^{(i)}\left(\nabla^{\nu_k^*}f_i\left(x_k\right)+4M_i(x_{k+1}-x_k)\right)\right\|,\\ \delta_2^{(i)}(k,\lambda_{k+1})&:=\left|\lambda_{k+1}^{(i)}\left(f_i(x_k)+\nabla^{\nu_k^*}f_i\left(x_k\right)\left(x_{k+1}-x_k\right)+2M_i\|x_{k+1}-x_k\|^2\right)\right|.\end{aligned}\tag{3.11}$$

Observe that $\delta_1(k,\lambda_{k+1})$ and $\delta_2^{(i)}(k,\lambda_{k+1})$ in (3.11) originate from the KKT conditions for (SP1). Therefore, by solving (SP2) we obtain the smallest-norm vector λ_{k+1} such that (x_{k+1},λ_{k+1}) is a $\eta/2$-KKT pair of (SP1). To solve (SP2), we reformulate it as a convex QCQP and use λ_{k+1}° as an initial feasible solution. If the two conditions are satisfied at the $(k+1)$-th iteration, then the algorithm outputs in Line 9 are $\tilde{x}=x_{k+1}$, $\tilde{\lambda}=\lambda_{k+1}$ and $\tilde{k}=k+1$.

Algorithm 2 is similar to Sequential QCQP (SQCQP) [125]. In SQCQP, at each iteration, quadratic proxies for constraint functions are built based on the local gradient vectors and Hessian matrices. The application of SQCQP to optimal control has received increasing attention [126, 127], due to the development of efficient solvers for QCQP subproblems [128]. Different from SQCQP [125], Algorithm 2 can guarantee sample feasibility and does not require the knowledge of Hessian matrices, which are costly to obtain for zeroth-order methods. As Hessian matrices are essential for proving the convergence of SQCQP in [125], we cannot use the same arguments in [125] to show the properties of SZO-QQ's iterates. In the following two sections, we state the properties of $(\tilde{x},\tilde{\lambda})$ and analyze the efficiency of the algorithm.

3.4 Properties of SZO-QQ's Iterates and Output

In this section, we aim to show that, for a suitable ξ, the pair $(\tilde{x},\tilde{\lambda})$ derived in Algorithm 2 is η-KKT. We start by considering the infinite sequence of Algorithm 2's iterates $\{x_k\}_{k\geq 1}$ when the termination conditions in Line 6 and 8 of Algorithm 2 are removed. We show that the sequence $\{x_k\}_{k\geq 1}$ has accumulation points and, for any accumulation point x_c, under mild assumptions, there exists $\lambda_c\in\mathbb{R}^m_{\geq 0}$ such that (x_c,λ_c) is a KKT pair of (3.1). Based on this result, we then study the activation of the two termination conditions and prove that they are satisfied within a finite number of iterations. Finally, we show that if ξ is carefully chosen, the derived pair $(\tilde{x},\tilde{\lambda})$ is η-KKT.

3.4.1 On the accumulation points of $\{x_k\}_{k\geq 1}$

Proposition 3.1. *If the termination conditions are removed, the sequence $\{x_k\}_{k\geq 1}$ in Algorithm 2 has the following properties:*

1. *the sequence $\{f_0(x_k)\}_{k\geq 1}$ is non-increasing;*
2. *$\{x_k\}_{k\geq 1}$ has at least one accumulation point x_c and $\{\|x_{k+1} - x_k\|\}_{k\geq 1}$ converges to 0;*
3. *$\lim_{k\to\infty} f_0(x_k) = f_0(x_c) > -\infty$.*

The proof is provided in Section 3.7.3. It mainly exploits the following inequality,

$$f_0(x_{k+1}) + \mu\|x_{k+1} - x_k\|^2 \leq f_0(x_k), \tag{3.12}$$

which is due to the optimality of x_{k+1} for (SP1) in Algorithm 2. The monotonicity of $f_0(x_k)$ and the convergence of $\|x_{k+1} - x_k\|$ are direct consequences of (3.12). By utilizing the monotonicity, we have that, for any $k \geq 1$, x_k belongs to the bounded set $\mathcal{P}_\beta$ (see Assumption 3.3 for the definition of $\mathcal{P}_\beta$). Due to Bolzano–Weierstrass theorem, there exists an accumulation point of $\{x_k\}_{k\geq 1}$. The continuity of $f_0(x)$ gives us Point 3 of Proposition 3.1.

Based on Proposition 3.1, we can show that under Assumption 3.4 below, there exists an accumulation point of $\{x_k\}_{k\geq 1}$ that allows one to build a KKT pair.

Assumption 3.4. *There exists an accumulation point x_c of $\{x_k\}_{k\geq 1}$ such that the Linear Independent Constraint Qualification (LICQ) holds at x_c for* (3.1)*, which is to say the gradients $\nabla f_i(x_c)$ with $i \in \mathcal{A}(x_c) := \{i : f_i(x_c) = 0\}$ are linearly independent.*

Assumption 3.4 is widely used in optimization [129]. For example, it is used to prove the properties of the limit point of the Interior Point Method [95]. With this assumption, if x_c is a local minimizer, then there exists $\lambda_c \in \mathbb{R}^m_{\geq 0}$ such that (x_c, λ_c) is a KKT pair [95, Theorem 12.1], which will be used in the proof of Theorem 3.2.

Theorem 3.2. *Let Assumption 3.4 hold, and let x_c be an accumulation point of $\{x_k\}_{k\geq 1}$ where LICQ is verified. Then, there exists a unique $\lambda_c \in \mathbb{R}^m_{\geq 0}$ such that (x_c, λ_c) is a KKT pair of the problem* (3.1)*.*

We only consider the case where $\mathcal{A}(x_c)$ is not empty in the proof of Theorem 3.2, provided in Section 3.7.5. The proof can be easily extended to account for the case where x_c is in the interior of Ω. To show Theorem 3.2, we exploit a preliminary result (Lemma 3.5, stated and proved in Section 3.7.4) where we construct an auxiliary problem (3.24) and show that x_c is an optimizer to (3.24). We notice that the KKT conditions of (3.24) evaluated at x_c coincide with those of (3.1) evaluated at the same point. Due to LICQ, there exists a unique $\lambda_c \in \mathbb{R}_{\geq 0}$ such that (x_c, λ_c) is a common KKT pair of (3.24) and (3.1) [95, Section 12.3].

3.4.2 The output of Algorithm 2 is an η-KKT pair

The result in Theorem 3.2 is asymptotic, but in practice, only finitely many iterations can be computed. From now on, we take the termination conditions of Algorithm 2 into account and show that, given any $\eta > 0$, by suitably tuning $\xi > 0$, Algorithm 2 returns an η-KKT pair. First, we make the following assumption, which allows us to conclude in Proposition 3.2 the finite termination of Algorithm 2.

Assumption 3.5. *The KKT pair (x_c, λ_c) in Theorem 3.2 satisfies $\|\lambda_c\|_\infty < \Lambda$, where $\Lambda > 0$ is the input in Algorithm 2.*

Assumptions on the bound of the dual variable are adopted in the literature of primal-dual methods, including [130, 131]. We illustrate Assumption 3.5 in Section 3.7.6 where we show in an example that Λ is related to the geometric properties of the feasible region.

Remark 3.4. *In case it is hard to choose a value of Λ fulfilling Assumption 3.5, we can replace Λ with $\gamma\|\lambda_{k+1}\|_\infty$, where $\gamma > 1$ and λ_{k+1} is the solution to the problem* (SP2) *in Algorithm 2, every time the second termination condition (Line 8 in Algorithm 2) is violated. Note that every time Λ gets updated, it becomes at least $2\gamma - 1$ times larger. Similar updating rules can also be found in [130]. In this way, we are guaranteed to find Λ that satisfies Assumption 3.5 after a finite number of iterations. However, we also notice that this updating mechanism generates a conservative guess for Λ if $\|\lambda_k\|_\infty \gg \|\lambda_c\|_\infty$ for some k. In Theorem 3.3, we will set ξ in Algorithm 2 to be proportional to Λ^{-1} so that the returned pair is an η-KKT pair. Consequently, a conservative Λ can increase the number of iterations required by Algorithm 2.*

Proposition 3.2. *Let Assumptions 3.4 and 3.5 hold, Algorithm 2 terminates in a finite number of iterations.*

According to Proposition 3.1, the first termination condition is satisfied in Algorithm 2 whenever k is sufficiently large. In the proof of Proposition 3.2 (provided in Section 3.7.7), we show that $\lambda_{k+1} = \lambda_c$ is a feasible solution to (SP2) when x_{k+1} is close enough to x_c. Thus, for sufficiently large k, the second termination is satisfied since $\|\lambda_c\|_\infty < \Lambda$.

Recall that Algorithm 2 returns $\tilde{x}, \tilde{\lambda}$ and $\tilde{k}$, which are dependent on the chosen value for ξ. For a given accuracy indicator $\eta > 0$, in the following, we show how to select ξ such that $(\tilde{x}, \tilde{\lambda})$ is an η-KKT pair.

Theorem 3.3. *Let Assumptions 3.4 and 3.5 hold, and let*

$$\xi = h(\eta) := \min\left\{\frac{\eta}{60\Lambda\sum_{i=1}^m M_i}, \frac{\eta}{12\mu}, 1, \frac{\eta}{4\Lambda(\alpha_{\max} + 2L_{\max} + 2M_{\max})}\right\} \tag{3.13}$$

be satisfied, where μ is a parameter of (SP1), *$\alpha_{\max} = \max_{1\leq i\leq m} \alpha_i$ and α_i is defined in* (3.6). *Then the output $(\tilde{x}, \tilde{\lambda})$ of Algorithm 2 is an η-KKT pair of* (3.1) .

The proof of Theorem 3.3 can be found in Section 3.7.8. In summary, to ensure that the pair $(\tilde{x}, \tilde{\lambda})$ is an η-KKT pair, we need to set ξ in Algoirthm 2 to be $h(\eta)$ in (3.13) while selecting L_i, M_i and Λ to satisfy Assumptions 3.3 and 3.5 (see Remarks 3.3 and 3.4).

3.5 Complexity Analysis

In this section, we aim to give an upper bound, dependent on η, for the number of Algorithm 2's iterates. To this purpose, we consider the following assumption, which allows us to show the convergence of $\{x_k\}_{k\geq 1}$ in Lemma 3.2.

Assumption 3.6. *The accumulation point x_c in Assumption 3.4, which is already known to be the primal of some KKT pair, is a strict local minimizer, i.e., there exists a neighborhood $\mathcal{N}$ of x_c such that $f_0(x) > f_0(x_c)$ for any $x \in \mathcal{N} \cap \Omega \setminus x_c$.*

If Second-Order Sufficient Condition (SOSC) for optimality [95, Chapter 12.5] is satisfied, Assumption 3.6 holds. Since this assumption does not rely on the twice differentiability of the objective and constraint functions, it is more general than SOSC, which is commonly assumed in the optimization literature [132, 133].

Lemma 3.2. *Let Assumption 3.6 holds, $\{x_k\}_{k\geq 1}$ converges.*

The proof of Lemma 3.2 is in Section 3.7.9. In the rest of this section, we consider Assumption 3.6. Let x_c be the limit point of $\{x_k\}_{k\geq 1}$ and note that there exists λ_c such that (x_c, λ_c) is a KKT pair. Then we show in Lemma 3.3, with the proof in Section 3.7.10, that λ_k°, the optimal dual solution to (SP1), converges to λ_c.

Lemma 3.3. *Let Assumptions 3.4, 3.5 and 3.6 hold, λ_k° converges to λ_c.*

With Lemma 3.3 and Assumption 3.5, we know that for any η there exists K, independent of η, such that $\|\lambda_k\|_\infty \leq 2\Lambda$ for any $k \geq K$. Therefore, for sufficiently small η, Algorithm 2 terminates whenever the first termination condition in Line 6 is satisfied. We can now conclude in Theorem 3.4 on the complexity of Algorithm 2 by analyzing only the first termination condition. The proof of Theorem 3.4 is in Section 3.7.11.

Theorem 3.4. *Let Assumptions 3.4, 3.5 and 3.6 hold, there exists $\bar{\eta} > 0$ such that, for any $\eta < \bar{\eta}$, Algorithm 2 terminates within $\overline{\mathcal{K}}(\eta) + 1$ iterations, where*

$$\overline{\mathcal{K}}(\eta) = \frac{f_0(x_0) - \inf\{f_0(x) : x \in \Omega\}}{\mu(h(\eta))^2},$$

and μ is the coefficient of the quadratic penalty term in (SP1)*, and $h(\eta)$ is defined in* (3.13)*. Thus, for any $0 < \eta \leq 1$, Algorithm 2 takes at most $O((M_{\max} + L_{\max})^2/\eta^2)$ iterations to return $(\tilde{x}, \tilde{\lambda})$, an η-KKT pair of the problem* (3.1).

Discussion: We compare the sample and computation complexity of SZO-QQ with two other existing safe zeroth-order methods, namely, LB-SGD in [24] and SafeOPT in [111]. We remind the readers that these methods have different assumptions. Specifically, given the black-box optimization problem (3.1), SafeOPT assumes that $f_i(x), i \in \mathbb{Z}_{i=0}^m$, has bounded norm in a suitable Reproducing Kernel Hilbert Space while both SZO-QQ and LB-SGD assume the knowledge of the smoothness constants. Regarding sample complexity, SZO-QQ needs $O(\frac{1}{\eta^2})$ samples under Assumption 3.6 to generate an η-KKT pair while LB-SGD and SafeOPT require $O(\frac{1}{\eta^3})$ and at least $O(\frac{1}{\eta^2})$ samples[1] respectively. We also highlight that the computational complexity of each iteration of both LB-SGD and SZO-QQ stays fixed while the computation time required for the Gaussian Process regression involved in each iteration of SafeOPT increases as the data set gets larger. The high computational cost is one of the main reasons why SafeOPT scales poorly to high-dimensional problems. Numerical results comparing the computation time and the number of samples required by these methods are provided in Section 3.6.1.

In contrast, the Interior Point Method, based on the assumption of $f_i(x)$ being twice continuously differentiable, achieves superlinear convergence [95] by utilizing the true model of the optimization problem, which translates into at most $O(\log\frac{1}{\eta})$ iterations. The gap between $O(\frac{1}{\eta^2})$ of SZO-QQ and $O(\log\frac{1}{\eta})$ may be either the price we pay for the lack of the first-order and second-order information of the objective and constraint functions or due to the conservative analysis in Theorem 3.4. To see whether there exists a tighter complexity bound than $O(\frac{1}{\eta^2})$ for Algorithm 2, an analysis on the convergence rate is needed, which is left as future work.

3.6 Numerical Results

In this section, we present three numerical experiments to test the performance of Algorithm 2. The first is a two-dimensional problem where we compare SZO-QQ with other existing zeroth-order methods and discuss the impact of parameters. In the remaining two examples, we apply our method to solve optimal control and optimal power flow problems, which have more dimensions and constraints. All the numerical experiments have been executed on a PC with an Intel Core i9 processor.

[1] In [24, Theorem 8], the authors claim the iterations of LB-SGD needed for obtaining an η-KKT to be $O(\frac{1}{\eta^3})$. Since the number of samples stays fixed and does not rely on η in each iteration, the number of samples required is also $O(\frac{1}{\eta^3})$. In [25, Theorem 1], $O(\frac{1}{\eta^2})$ samples are needed in SafeOPT to obtain an η-suboptimal point. For unconstrained optimization with a strongly convex objective function, there exists $\alpha > 0$ such that any η-KKT pair has a $\alpha\eta$-suboptimal primal. Therefore, SafeOPT requires at least $O(\frac{1}{\eta^2})$ samples to obtain an η-KKT pair of (3.1).

3.6.1 Solving an unknown non-convex QCQP

We evaluate SZO-QQ and compare it with alternative safe zeroth-order methods in the following non-convex example,

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & f_0(x) = 0.1 \times (x^{(1)})^2 + x^{(2)} \\ \text{subject to} \quad & f_1(x) = 0.5 - \|x + [0.5 \ \ -0.5]^\top\|^2 \leq 0, \\ & f_2(x) = x^{(2)} - 1 \leq 0, \\ & f_3(x) = (x^{(1)})^2 - x^{(2)} \leq 0. \end{aligned}$$

We assume that the functions $f_i(x), i = 1, 2, 3$, are unknown but can be queried. A strictly feasible initial point $x_0 = [0.9 \ \ 0.9]^\top$ is given. The unique optimum $x_* = [0 \ \ 0]^\top$ is not strictly feasible. According to Theorem 3.2, the iterates of SZO-QQ will get close to x_*, which allows us to see whether SZO-QQ stays safe and whether the convergence is fast when the iterates are close to the feasible region boundary. The experiment results allow us to discuss the following three aspects, respectively on the derivation of a 10^{-2}-KKT pair, performance evaluation, and parameter tuning.

Selection of ξ for deriving a 10^{-2}-KKT pair

To begin, we fix $\eta = 10^{-2}$ and aim to derive an η-KKT pair. By setting $\Lambda = 1.5$ and $L_i = 5$, $M_i = 3$ for any $i \geq 1$, we calculate $\xi = 1.51 \times 10^{-5}$ according to (3.13). With these values, SZO-QQ returns in 3.3 seconds an η'-KKT pair with $\eta' = 9.21 \times 10^{-4} < \eta$. We also observe that $\|\lambda_k\|_\infty$ converges to 1 and thus the original guess $\Lambda = 1.5$ satisfies Assumption 3.5. Now we see that we indeed derive an η-KKT pair, which coincides with Theorem 3.3. To further evaluate the performance of SZO-QQ in terms of how fast the objective function value decreases, we eliminate the termination conditions in the remainders of Section 3.6.1.

Performance comparison with other methods

We run SZO-QQ with $\xi = 0$ and compare with LB-SGD [24], Extremum Seeking [106] and SafeOptSwarm[2] [111]. Among these methods, SZO-QQ, LB-SGD, and SafeOpt-Swarm have theoretical guarantees for sample feasibility (at least with a high probability). Only SZO-QQ and LB-SGD require Assumption 3.1 on Lipschitz and smoothness constants. For these two approaches, by trial and error (see Remark 3.3), we choose $L_i = 5$ and $M_i = 3$ for any $i \geq 1$. The penalty coefficient μ of Algorithm 2 in (SP1) is set to be 0.001. For both LB-SGD and Extremum Seeking are barrier-function-based, we use the reformulated unconstrained problem $\min_x f_{\log}(x, \mu_{\log})$, where $f_{\log}(x, \mu_{\log}) := f_0(x) - \mu_{\log} \sum_{i=1}^{4} \log(-f_i(x))$, and $\mu_{\log} = 0.001$.

[2] SafeOptSwarm is a variant of SafeOpt (recall Section 3.1). The former add heuristics to make SafeOpt in [25] more tractable for higher dimensions.

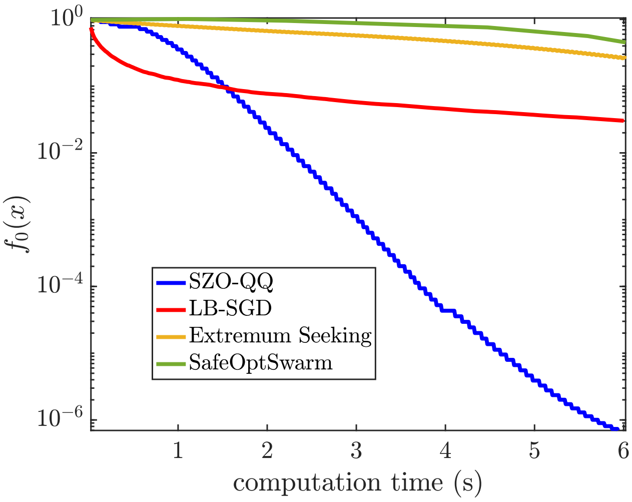

Figure 3.1: Objective value as a function of computation time.

In Fig. 3.1, we show the objective function values versus the computation time. During the experiments, none of the methods violates the constraints. Regarding the convergence to the minimum, we see that LB-SGD has the most decrease in the objective function value in the first 1.5 seconds due to the low complexity of each iteration. In these 1.5 seconds, LB-SGD utilizes 67856 function samples while SZO-QQ only 252. Afterward, SZO-QQ achieves a better solution. In the first 6 seconds, SZO-QQ shows a clear convergence trend to the optimum, consistent with Theorem 3.2, while SafeOptSwarm only finishes 6 iterations and 28 function samples.

LB-SGD slows down when the iterates are close to the boundary of the feasible set (see Section 3.7.2 for the explanation for this phenomenon). Meanwhile, the slow convergence of Extremum Seeking is due to its small learning rate. If the learning rate is large, the iterates might be brought too close to the boundary of the feasible set, and then the perturbation added by this method would lead to constraint violation. These considerations constitute the main dilemma in parameter tuning for Extremum Seeking. Meanwhile, exploring the unknown functions in SafeOptSwarm is based on Gaussian Process (GP) regression models instead of local perturbations. Since SafeOptSwarm does not exploit the convexity of the problem, it maintains a safe set and tries to expand it to find the global minimum. Empirically, this method samples many points close to the boundary of the feasible region, which is also observed in [101]. These samples along with the computational complexity of GP regression, are the main reason for the slow convergence of SafeOptSwarm. We also run LB-SGD and Extremum Seeking with different penalty coefficients $\mu_{\log}$ to check whether the slow convergence is due to improper parameter tuning. We see that with larger $\mu_{\log}$ the performance of the log-barrier-based methods deteriorates. This is probably because the optimum of the unconstrained problem $f_{\log}(x, \mu_{\log})$ deviates more from the optimum as $\mu_{\log}$ increases. With smaller $\mu_{\log}$, the Extremum Seeking method leads to constraint violation while the performance of LB-SGD barely changes.

Impact of the parameters L_i and M_i

To show the impact of conservative guesses of L_i and M_i, we consider 9 test cases of different values for the pair (L, M). We use L as the Lipschitz constant for all the objective and constraint functions and M as the smoothness constant. We illustrate in Figures 3.2 and 3.3 the decrease of the objective function values when SZO-QQ and LB-SGD are applied to solve the 9 test cases. From the figures, we see that the time required by SZO-QQ to achieve an objective function value less than 10^{-2} grows with M. Despite this, across all the cases SZO-QQ is the first to achieve an objective function value of 10^{-2}. Another observation is that the performance of SZO-QQ is more sensitive to varying M while LB-SGD is more sensitive to varying L. This is due to the differences in the local feasible set formulations in both methods. Indeed, in SZO-QQ the constant L_i is only related to the gradient estimation, and the size of the local feasible set $\mathcal{S}^{(k)}(\cdot)$ in (4.1) is mainly decided by M_i, while in LB-SGD the size of $\mathcal{T}^{(k)}(\cdot)$, for any k, is mainly dependent on the Lipschitz constants L_i for $i \geq 1$.

We also study the case where the initial guesses of Lipschitz and smoothness constants are wrong, i.e., (3.3) in Assumption 3.1 is violated. With $L = 0.2$ and $M = 0.2$, we encounter an infeasible sample. Then we follow the method in Remark 3.3 to multiply the constants by 2 every time an infeasible sample is generated. With $L = M = 0.8$, every sample is feasible and we derive in 2 seconds an objective function value of 4×10^{-7}. In total, we generate two infeasible samples. Although the setting $L = M = 0.8$ still fails to satisfy Assumption 3.1, with these constants, SZO-QQ is able to generate iterates that have a subsequence converging to a KKT pair. The readers can check that Theorem 3.2 holds even if the guesses for Lipschitz and smoothness constants do not verify (3.3) in Assumption 3.1 (see the proof of Theorem 3.2 in Section 3.7.5).

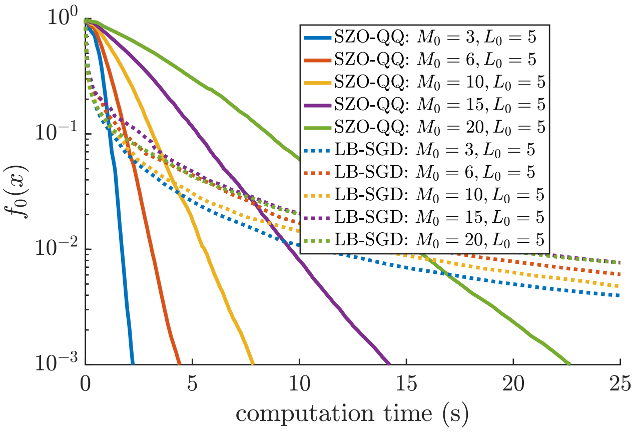

Figure 3.2: Objective value as a function of computation time: $L_0 = 5$ fixed and M_0 varied

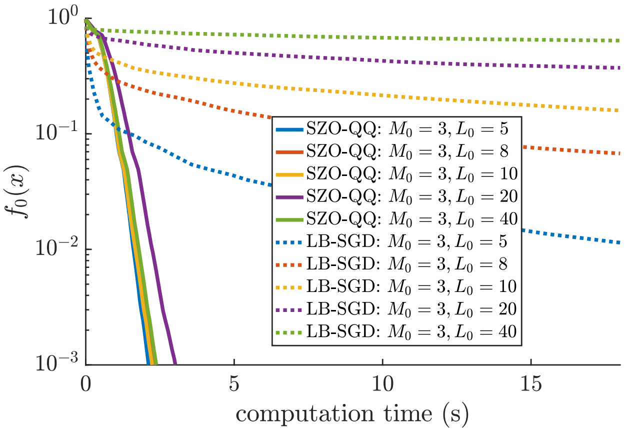

Figure 3.3: Objective value as a function of computation time: $M_0 = 3$ fixed and L_0 varied

3.6.2 Open-loop optimal control with unmodelled disturbance

SZO-QQ can be applied to deterministic optimal control problems with unknown nonlinear dynamics by using only feasible samples. To illustrate, we consider a nonlinear system with dynamics $x_{k+1} = Ax_k + Bu_k + \delta(x_k)[1\ 0]^\top$, where $x_k \in \mathbb{R}^2$ for $k \geq 0$ and $x_0 = [1\ 1]^\top$. The matrices

$$A = \begin{bmatrix} 1.1 & 1 \\ -0.5 & 1.1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and the expression of the disturbance $\delta(x) := 0.1 * (x^{(2)})^2$ are unknown. We aim to design the input u_k for $i \in \mathbb{Z}_0^5$ to minimize the cost $\sum_{k=0}^{5} \left(x_{k+1}^\top Q x_{k+1} + u_k^\top R u_k\right)$ where $Q = 0.5\,\mathbf{I}_2$ and $R = 2\,\mathbf{I}_2$ with identity matrix $\mathbf{I}_2 \in \mathbb{R}^{2\times 2}$ while enforcing $\|x_{k+1}\|_\infty \leq 0.7$ and $\|u_k\|_\infty \leq 1.5$ for $0 \leq k \leq 5$. Since we assume all the states are measured, we can evaluate the objective and constraints. In this example, we assume to have a feasible sequence of inputs $\{u_k\}_{0\leq k\leq 5}$ that leads to a safe trajectory and results in a cost of 6.81 (as in Assumption 3.2). Different from the settings in the model-based safe learning control methods [134, 135], we do not assume that this safe trajectory is sufficient for identifying the system dynamics with small error bounds. If the error bounds are huge, the robust control problems formulated in [134, 135] may become infeasible.

We run SZO-QQ to further decrease the cost resulting from the initial safe trajectory and derive within 146 seconds of computation an input sequence that satisfies all the constraints and achieves a cost of 5.96. This cost is the same as the one obtained when assuming the dynamics are known and applying the solver IPOPT [92]. This observation is consistent with Theorem 3.2 on the convergence to a KKT pair. In this experiment, we set $L_i = M_i = 20$ for $i \geq 1$, $\mu = 10^{-4}$ and $\eta = 10^{-1}$. Thus, the parameter ξ adopted is 2×10^{-5} according to (3.13).

3.6.3 Optimal power flow for an unmodelled electric network

In this section, we apply SZO-QQ to solve the AC Optimal Power Flow (OPF) for the IEEE 30-bus system in [136].

Formulation of the OPF problem

To formulate an OPF problem, we introduce the following notations and assumptions:

- Let $B = \{b_1, b_2, \dots, b_n\}$ be the bus set and let $T = \{(b_i, b_j) :$ there is a transmission line between b_i and $b_j\}$ be a set of undirected edges representing the transmission lines;
- We denote P_{G_i}, P_{L_i}, Q_{L_i}, U_i and θ_i as the active power generation, active power consumption, reactive power consumption, voltage and voltage angle at b_i;
- From the b_i to b_j, the active power and the reactive power transferred are written respectively as $P_{ij}(U_i, U_j, \theta_i, \theta_j)$ and $Q_{ij}(U_i, U_j, \theta_i, \theta_j)$, while the current is denoted as $I_{ij}(U_i, U_j, \theta_i, \theta_j)$. We refer the readers to [137] for the explicit expressions of these functions;
- We also assume that there are n_G generators at the buses b_i, $i \in \mathbb{Z}_1^{n_G}$ and b_1 is a slack bus providing active power to maintain the power balance within the network and has a voltage angle of 0.

Then the OPF problem is formulated [137] as

$$\min_{P_{G_i}, U_i, \theta_i} \sum_{i=1}^{n_G} C_i(P_{G_i}) \tag{3.14a}$$

subject to

$$P_{G_i} = P_{L_i} + \sum_{(i,j)\in T} P_{ij}(U_i, U_j, \theta_i, \theta_j), \ \ \forall i \tag{3.14b}$$

$$-Q_{L_i} = \sum_{(i,j)\in T} Q_{ij}(U_i, U_j, \theta_i, \theta_j), \ \ i > n_G \tag{3.14c}$$

$$P_{G_i} = 0, \ \ \text{for } i > n_G, \ \ \theta_1 = 0, \tag{3.14d}$$

$$P_{G,\min} \leq P_{G_i} \leq P_{G,\max}, \text{for } \ i \leq n_G, \tag{3.14e}$$

$$I_{ij,\min} \leq I_{ij}(U_i, U_j, \theta_i, \theta_j) \leq I_{ij,\max}, \ \ \forall (i,j) \in T, \tag{3.14f}$$

$$U_{\min} \leq U_i \leq U_{\max}, \forall i. \tag{3.14g}$$

where $C_i(\cdot)$ is a quadratic function accounting for the generation cost and the equations (3.14e)-(3.14g) give the safe intervals for the corresponding variables.

The main challenges of OPF applications lie in modelling the system and deriving the accurate expressions of (3.14). The difficulties include the nonlinearity of device dynamics, slowly changing physical parameters and disturbances [138]. Inaccurate models can result in suboptimal OPF solutions (leading to more generation cost) or violate the true hard constraints (causing damages to devices) [139]. Therefore, we consider the black-box setting and use SZO-QQ.

To this aim, we reformulate (3.14) as optimization with only inequality constraint to fit (3.1) used by SZO-LP. Let $\{P_{G_i}\}_{i=2}^{n_G}$ and $\{U_i\}_{i=1}^{n_G}$ be the main decision variables. Then by assigning values to $\{P_{G_i}\}_{i=2}^{n_G}$ and $\{U_i\}_{i=1}^{n_G}$, one can solve the power flow equations (3.14b)-(3.14d) to derive the values for all the other decision variables in (3.14). Therefore, (3.14b)-(3.14d) give us the functions

$$\begin{aligned} U_i &= U_i(\{P_{G_j}\}_{j=2}^{n_G}, \{U_j\}_{j=1}^{n_G}), \quad i = n_G + 1, \dots, n, \\ \theta_i &= \theta_i(\{P_{G_j}\}_{j=2}^{n_G}, \{U_j\}_{j=1}^{n_G}), \quad i = 1, \dots, n. \end{aligned} \tag{3.15}$$

By substituting (3.15) to (3.14), we obtain a reformulation where $\{\{P_{G_j}\}_{j=2}^{n_G}, \{U_j\}_{j=1}^{n_G}\}$ are the only decision variables and there are not equality constraints.

Experiment results

Given a set of values for all the 11 decision variables, we can utilize a black-box simulation model in `Matpower` [140] to sample the voltages of all the 30 buses and the power through all the transmission lines in this network. We assume to have initial values for all the decision variables such that the constraints are satisfied. In practice, initial values of the decision variables verifying the safety constraints in power systems are not hard to find due to various mechanisms for robust operation, e.g., droop control for power generation, shunt capacitor control and load shedding.

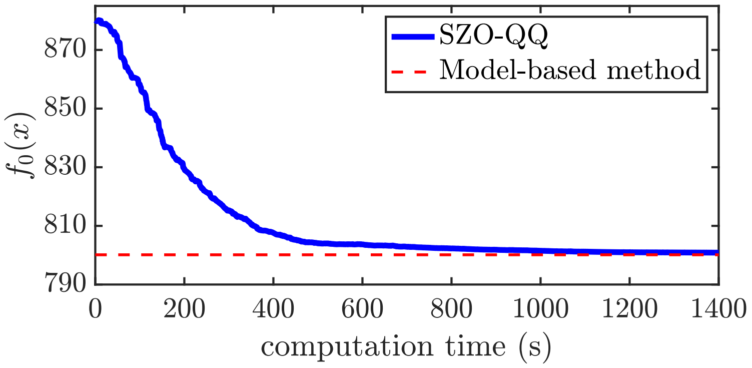

Figure 3.4: Objective value as a function of computation time

The numerical experiment is executed on a PC with an Intel Core i9 processor. The QCQP subproblems are SZO-QQ is solved using Gurobi [141] with Yalmip [142] as the interface. Since the structure of the subproblems is fixed, we use OPTIMIZER function in Yalmip to speed up the computation. In this experiment, we set $\mu = 0.001$, $\xi = 0.002$, $\Lambda = 2$,

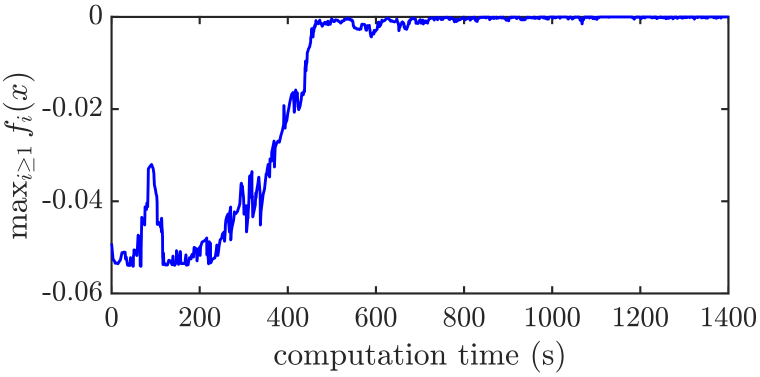

Figure 3.5: Largest constraint value as a function of computation time

$M_i = 0.2$ and $L_i = 1$ for $i \geq 1$. In Figure 3.4, we illustrate the decrease of the cost and compare it with the cost derived by using Gurobi to solve the model-based optimal power flow. We see that the achieved cost within 1400 seconds is close to what the model-based method derives, which is again consistent with Theorem 3.2. Meanwhile, from Figure 3.5, which depicts the largest constraint function value with respect to the computation time, we see that, even though the decision variables can get very close to the boundary of the feasible set, the constraints are never violated.

Compared with only 1 second used by the model-based method to derive the solution, SZO-QQ is slow. One reason is that solving the QCQP subproblems of SZO-QQ takes too much time for this experiment, since the size of subproblems is the same as the original problem. Another factor is that, as shown in Figs. 3.4 and 3.5, the decrease of the objective functions slows down as the iterates get close to the feasible region boundary.

3.7 Appendices

3.7.1 Proof of Theorem 3.1

We notice that $\nu_0^* \leq l_0^*$ and, from Assumption 3.1, $f_i(x_0 + \nu_0^* e_j) < f_i(x_0) + l_0^* L_{\max} = 0$ for any $i \in \mathbb{Z}_1^m$, $j \in \mathbb{Z}_1^d$. which shows the samples' feasibility. To show the feasibility of $\mathcal{S}^{(0)}(x_0)$, we first partition $\mathcal{S}_i^{(0)}(x_0)$ as

$$\mathcal{S}_i^{(0)}(x_0) = \Big(\mathcal{S}_i^{(0)}(x_0) \bigcap \mathcal{B}(x_0, l_0^*)\Big) \bigcup \Big(\mathcal{S}_i^{(0)}(x_0) \setminus \mathcal{B}(x_0, l_0^*)\Big)$$

and notice that $\mathcal{S}_i^{(0)}(x_0) \bigcap \mathcal{B}(x_0, l_0^*) \subseteq \Omega$. Then, it only remains to show $\mathcal{S}_i^{(0)}(x_0) \setminus \mathcal{B}(x_0, l_0^*) \subseteq \Omega$.

For $x \in \mathcal{S}_i^{(0)}(x_0) \setminus B_{l_0^*}(x_0)$, we have $\sqrt{d}\nu_0^* = l_0^* \leq \|x - x_0\|$. By the mean value theorem, for any i there exists $\theta_i \in [0,1]$ such that

$$\begin{aligned}f_i(x) =& f_i(x_0) + \nabla f_i\left(x_0 + \theta_i(x-x_0)\right)^\top (x-x_0) \qquad (3.16)\\ =& f_i(x_0) + \nabla f_i\left(x_0\right)^\top (x-x_0) + \left(\nabla f_i\left(x_0 + \theta_i(x-x_0)\right) - \nabla f_i\left(x_0\right)\right)^\top (x-x_0)\\ \leq& f_i(x_0) + \nabla^{\nu_0^*} f_i\left(x_0\right)^\top (x-x_0) + \frac{\sqrt{d}\nu_0^* M_i}{2}\|x-x_0\| + M_i\|x-x_0\|^2\\ \leq& f_i(x_0) + \nabla^{\nu_0^*} f_i\left(x_0\right)^\top (x-x_0) + 2M_i\|x-x_0\|^2\\ \leq& 0,\end{aligned}$$

where the first inequality is due to Assumption 3.1 while the second one can be derived from the definition of ν_0^* in Theorem 3.1. Hence, $\mathcal{S}_i^{(0)}(x_0) \setminus \mathcal{B}(x_0, l_0^*) \subseteq \Omega$. Since $\mathcal{S}_i^{(0)}(x_0) \subseteq \Omega, \forall i$, then $\mathcal{S}^{(0)}(x_0) \subseteq \Omega$.

3.7.2 Comparison between two different formulations of local safe sets

The works [24, 25] adopt an alternative approximation of the constraints and, in particular, form a local feasible set

$$\mathcal{T}^{(0)}(x_0) := \bigcap_{i=1}^{m} \left\{ x : \|x - x_0\| \leq -\frac{f_i(x_0)}{L_i} \right\}.$$

We see that $\mathcal{T}^{(0)}(x_0) = \{x : F_i^L(x) \leq 0, \forall i\}$ where $F_i^L(x) := f_i(x_0) + L_i\|x-x_0\|$ is linear in $\|x - x_0\|$. In contrast, $\mathcal{S}^{(0)}(x_0) = \{x : F_i^M(x) \leq 0, \forall i\}$ where $F_i^M(x) := f_i(x_0) + \nabla^{\nu_0^*} f_i\left(x_0\right)^\top (x-x_0) + 2M_i\|x-x_0\|^2$ is a quadratic approximation of $f_i(x)$. In the following proposition, we show that if x_0 is sufficiently close to the boundary of the feasible set, $\mathcal{T}^{(0)}(x_0) \subset \mathcal{S}^{(0)}(x_0)$, which means that $\mathcal{S}^{(0)}(x_0)$ is less conservative.

Proposition 3.3. *Let $\ell_{\min} = \min_{i\geq 1}(L_i - L_{i,\inf})$. For x_0, if*

$$\min_{i\geq 1} - f_i(x_0) \leq \frac{L_{\max}\ell_{\min}}{4M_{\max}}, \tag{3.17}$$

then $\mathcal{T}^{(0)}(x_0) \subset \mathcal{S}^{(0)}(x_0)$.

Proof. For any $x \in \mathcal{T}^{(0)}(x_0)$, we have

$$\|x - x_0\| \leq \frac{\min -f_i(x_0)}{L_{\max}} \leq \frac{\ell_{\min}}{4M_{\max}}$$

and thus

$$2M_{\max}\|x-x_0\|^2 \leq \frac{\ell_{\min}}{2}\|x-x_0\|. \tag{3.18}$$

Considering $\nu_0^* = \frac{\min -f_i(x_0)}{\sqrt{d}L_{\max}}$ and (3.6), we have

$$\left\|\Delta_i^{\nu_0^*}(x)\right\|_2 \leq \frac{\ell_{\min}}{2}$$

and thus

$$\begin{aligned}\nabla^{\nu_0^*} f_i(x_0)^\top (x - x_0) =& (\nabla f_i(x_0) + \Delta_i^{\nu_0^*}(x))^\top (x - x_0)\\ \leq& L_{i,\inf}\|x - x_0\| + \frac{\ell_{\min}}{2}\|x - x_0\|\end{aligned} \tag{3.19}$$

By summing up (3.18) and (3.19), we have for any $x \in \mathcal{T}^{(0)}(x_0)$

$$F_i^M(x) \leq f_i(x_0) + L_i\|x - x_0\| = F_i^L(x) \leq 0,$$

and thus $x \in \mathcal{S}^{(0)}(x_0)$. ■

3.7.3 Proof of Proposition 3.1

Proof of Point 1. Given any k, we have $x_k \in \mathcal{S}^{(k)}(x_k)$ and $x_{k+1} = \arg\min_{x\in\mathcal{S}^{(k)}(x_k)} f_0(x) + \mu\|x - x_k\|^2$. Thus,

$$f_0(x_{k+1}) + \mu\|x_{k+1} - x_k\|^2 \leq f_0(x_k) + \mu\|x_k - x_k\|^2 = f_0(x_k). \tag{3.20}$$

Proof of Point 2. For $k \geq 0$, one has $f_0(x_k) \leq f_0(x_0) < \beta$ according to Assumption 3.3. Now we know $\{x_k\}_{k\geq 1}$ is within the set $\mathcal{P}_\beta$. Due to the boundedness of the set $\mathcal{P}_\beta$, we can use Bolzano–Weierstrass theorem to conclude that there exists a subsequence of $\{x_k\}_{k\geq 1}$ that converges. Hence, $\{x_k\}_{k\geq 1}$ has at least one accumulation point x_c. According to (3.20), $f_0(x_{k+1}) \leq f_0(x_1) - \mu\sum_{i=1}^k \|x_{i+1} - x_i\|^2$. Since $f_0(x)$ is a continuous function on the compact set $\mathcal{P}_\beta$, $\inf_{x\in\mathcal{P}_\beta} f_0(x) > -\infty$. Therefore, $\sum_{i=1}^\infty \|x_{i+1} - x_i\|^2 < \infty$ and $\|x_{k+1} - x_k\|$ converges to 0.

Proof of Point 3. The sequence $\{f_0(x_k)\}_{k\geq 1}$ converging to $f_0(x_c)$ is a direct consequence of Point 2 in Proposition 3.1 and the continuity of $f_0(x)$.

3.7.4 Preliminary results towards the proof of Theorem 3.2

In this section, we only consider the case where $\mathcal{A}(x_c) \neq \emptyset$. Before stating the preliminary results, we have the following notations on the local feasible set $\mathcal{S}^{(k)}(x_k)$ of (SP1) in Algorithm 2. Our preliminary results are on the properties of the "limit" of these feasible sets as k goes to infinity. We define for strictly feasible $x \in \Omega$,

$$\begin{aligned} O_i^{(k)}(x) :=& x - \frac{\nabla^{\nu_k^*} f_i(x)}{2M_i}, \\ \left(r_i^{(k)}(x)\right)^2 :=& -\frac{f_i(x)}{M_i} + \frac{\|\nabla^{\nu_k^*} f_i(x)\|^2}{4M_i^2}, \end{aligned} \tag{3.21}$$

which allows us to write

$$\mathcal{S}_i^{(k)}(x_k) = \mathcal{B}(O_i^{(k)}(x_k), r_i^{(k)}(x_k)).$$

We let $\{x_{k_p}\}_{p\geq 1}$ be a subsequence converging to x_c. Since $\{\nu_k^*\}_{k\geq 1}$ converges to 0 (see (3.9)), we have

$$O_i^{(k_p)}(x_{k_p}) \to O_i(x_c),\ r_i^{(k_p)}(x_{k_p}) \to r_i(x_c) \text{ as } p \to \infty, \tag{3.22}$$

where

$$O_i(x_c) := x_c - \frac{\nabla f_i(x_c)}{2M_i},\ (r_i(x_c))^2 := -\frac{f_i(x)}{M_i} + \frac{\|\nabla f_i(x_c)\|^2}{4M_i^2}.$$

Then, we write

$$\mathcal{S}_i(x_c) := \mathcal{B}(O_i(x_c), r_i(x_c)), \mathcal{S}(x_c) := \bigcap_{i=1}^{m} \mathcal{S}_i(x_c).$$

With these notations, we can state and prove the following lemmas on the properties of $\mathcal{S}(x_c)$.

Lemma 3.4. *Let Assumption 3.4 holds and $\mathcal{A}(x_c) \neq \emptyset$ hold, where $\mathcal{A}(x_c) := \{i : f_i(x_c) = 0\}$, then*

1. *there exists $x \in \Omega$ that is strictly feasible with respect to $\mathcal{S}(x_c)$, i.e. for any $i \geq 1$, $f_i^c(x) < 0$ where*

   $$f_i^c(x) := f_i(x_c) + \nabla f_i(x_c)^\top (x - x_c) + 2M_i \|x - x_c\|^2.$$

   For any $x \in \{x : f_i^c(x) < 0, \forall i\}$, there exists $k \in \mathbb{N}$ such that x belongs to $\mathcal{S}^{(k)}(x_k)$;

2. *there exists $x_s \in \mathcal{S}(x_c)$ such that $x_s \neq x_c$. For any such x_s and any $0 < t < 1$, we let $x(t) = tx_c + (1-t)x_s$ and have that $x(t)$ is strictly feasible with respect to $\mathcal{S}(x_c)$.*

Proof of Point 1. We let $\mathcal{A}(x_c) = \{i_1, \ldots, i_l\}$. There exists $y \in \mathbb{R}^d$ such that

$$Jy = \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix}, \text{ where } J = \begin{bmatrix} \nabla f_{i_1}(x_c)^\top \\ \vdots \\ \nabla f_{i_l}(x_c)^\top \end{bmatrix}, \tag{3.23}$$

because J is full row rank due to LICQ. For any y satisfying (4.13), if $\epsilon_0 = 1/(4M_{\max}\|y\|)$, then, for any $\epsilon < \epsilon_0$, $x = x_c + \epsilon y/\|y\|$ and $i \in \mathcal{A}(x_c)$,

$$f_i^c(x) = 0 - \epsilon/\|y\| + 2M_i\epsilon^2 < 0.$$

Since $f_i(x_c) < 0$ for any $i \notin \mathcal{A}(x_c)$, there exists $\epsilon_c > 0$ such that, for $\epsilon < \epsilon_c$ and $x = x_c + \epsilon_c y/\|y\|$, $f_i^c(x) < 0$ for any $i \notin \mathcal{A}(x_c)$. Thus, with $\epsilon = \min\{\epsilon_0, \epsilon_c\}$ and $x = x_c + \epsilon y/\|y\|$, we have $f_i^c(x) < 0$ for any i. Since x_c is an accumulation point, there exists $k \in \mathbb{N}$ such that x belongs to $\mathcal{S}^{(k)}(x_k)$.

Proof of Point 2. We utilize the first point and the fact that x_c is not strictly feasible with respect to Ω to conclude that x_c is not strictly feasible either with respect to $\mathcal{S}(x_c)$ and thus there exists $x_s \in \mathcal{S}(x_c)$ verifying $x_s \neq x_c$. Considering that $f_i^c(x)$ is strongly convex, we have, for any i and any $0 < t < 1$, $f_i^c(x(t)) < \max\{f_i^c(x_c), f_i^c(x_s)\} \leq 0$.

Lemma 3.5. *Let Assumption 3.4 and $\mathcal{A}(x_c) \neq \emptyset$ hold, then x_c is the unique optimum of the convex optimization*

$$\min_{x\in\mathcal{S}(x_c)} f_0(x) + \mu\|x - x_c\|^2. \tag{3.24}$$

Moreover, the optimizer λ_c for the dual variable of (3.24) *is also unique.*

Proof. We prove the optimality of x_c by contradiction. Assume x_c is not the optimum of (3.24), then there exists $x_s \in \mathcal{S}(x_c)$ verifying $f_0(x_s) + \mu\|x_s - x_c\|^2 < f_0(x_c)$. According to the second point of Lemma 3.4 in Section 3.7.4 and the continuity of $f_0(x)$, there exists $0 < t < 1$ such that with $x(t) = tx_c + (1-t)x_s$ we have $f_0(x(t)) + \mu\|x(t) - x_c\|^2 < f_0(x_c)$ and $x(t)$ is strictly feasible with respect to $\mathcal{S}(x_c)$. We let $\{x_{k_p}\}_{p\geq 1}$ be a subsequence of $\{x_k\}_{k\geq 1}$ that converges to x_c. Considering the first point of Lemma 3.4, there exists p such that $x(t) \in \mathcal{S}^{(k_p)}(x_{k_p})$. Because of the convergence of the subsequence, we can assume without loss of generality that p is sufficiently large so that $f_0(x(t)) + \mu\|x(t) - x_{k_p}\|^2 < f_0(x_c)$. Due to the optimality of x_{k_p+1} for the problem (SP1) in Algorithm 2 when $k = k_p$, we have $f_0(x_{k_p+1}) + \mu\|x_{k_p+1} - x_{k_p}\|^2 < f_0(x_c)$, which contradicts the monotonicity of $\{f_0(x_k)\}_{k\geq 1}$ in Proposition 3.1. Due to optimality of x_c and LICQ, there exists $\lambda_c \in \mathbb{R}^m$ such that (x_c, λ_c) is a KKT pair of (3.24).

We prove the uniqueness of x_c and λ_c also by contradiction. Assume there exists $x_o \in \mathcal{S}(x_c) \setminus \{x_c\}$ such that $f_0(x_o) + \mu\|x_o - x_c\|^2 = f_0(x_c)$. Due to the strong convexity of the function $f_0(x) + \mu\|x - x_c\|^2$, we know

$$f_0(\frac{x_o + x_c}{2}) + \mu\|\frac{x_o + x_c}{2} - x_c\|^2 < f_0(x_c),$$

which contradicts the optimality of x_c for (3.24). Assume $(x_c, \lambda_{c,1})$ and $(x_c, \lambda_{c,2})$ are two KKT pairs of (3.24) with $\lambda_{c,1} \neq \lambda_{c,2}$, then for $j = 1, 2$, $\lambda_{c,j}^{(i)} = 0$ for any $i \neq \mathcal{A}(x_c)$,

$$\sum_{i=1}^{m} \lambda_{c,j}^{(i)} \nabla f_i(x_c) = -\nabla f_0(x_c)$$

and thus

$$\sum_{i=1}^{m}(\lambda_{c,1}^{(i)} - \lambda_{c,2}^{(i)})\nabla f_i(x_c) = 0.$$

This contradicts LICQ at x_c since $\lambda_{c,1} - \lambda_{c,2} \neq 0$. ■

3.7.5 Proof of Theorem 3.2

This proof only considers the case where $\mathcal{A}(x_c) \neq \emptyset$ and can be easily extended to "$\mathcal{A}(x_c) = \emptyset$". According to Lemma 3.5, there exists a $\lambda_c \in \mathbb{R}^m$ such that (x_c, λ_c) is a KKT pair of (3.24), i.e.,

$$\nabla f_0(x_c) + \sum_{i \in \mathcal{A}(x_c)} \lambda_c^{(i)} \nabla f_i(x_c) = 0$$
$$\text{and } \ \lambda_c^{(i)} = 0 \text{ for } i \notin \mathcal{A}(x_c),$$

which coincides with KKT conditions of (3.1). Thus, (x_c, λ_c) is also a KKT pair of of (3.1). Following the same arguments used in the proof of Lemma 3.5, one can exploit LICQ to show that there does not exist $\lambda_{c,2} \neq \lambda_c$ such that $(x_c, \lambda_{c,2})$ is a KKT pair of (3.1).

3.7.6 Geometric illustration of an upperbound to $\|\lambda_c\|_\infty$

With the following example, we aim to illustrate that $\|\lambda_c\|_\infty$ is related to the geometric properties of the feasible region. We consider an instance of the optimization problem (3.1) where $d = 2$, the feasible region is convex, and (x_c, λ_c) is a KKT pair. We only consider the non-degenerate case where $\mathcal{A}(x_c) = \{1, 2\}$ and assume LICQ holds at x_c, i.e., $\nabla f_i(x_c)$ are linearly independent for $i = 1, 2$. The objective and constraint functions are normalized at x_c, i.e., $\|\nabla f_i(x_c)\| = 1$ for $i \in \mathbb{Z}_0^2$. Then, we use coordinate transformation such that $\nabla f_0(x_c) = \begin{bmatrix} 0 & -1 \end{bmatrix}^\top$. Since $\mathcal{A}(x_c) = \{1, 2\}$, the KKT pair (x_c, λ_c) satisfies that $\lambda_c^{(1)}, \lambda_c^{(2)} \geq 0$ and

$$\begin{aligned} f_1(x_c) \leq 0, \ f_2(x_c) \leq 0 \\ \nabla f_0(x_c) + \lambda_c^{(1)} \nabla f_1(x_c) + \lambda_c^{(2)} \nabla f_2(x_c) = 0 \\ \lambda_c^{(1)} f_1(x_c) = 0, \ \lambda_c^{(2)} f_2(x_c) = 0. \end{aligned} \tag{3.25}$$

Let θ_i be the angle between $-\nabla f_0(x_c)$ and $\nabla f_i(x_c)$ for $i = 1, 2$. Due to the convexity of the feasible region, $0 < \theta_1 + \theta_2 < \pi$. By solving (3.25), we have that

$$\lambda_c^{(1)} = \frac{|\sin\theta_2|}{\sin(\theta_1 + \theta_2)}, \ \lambda_c^{(2)} = \frac{|\sin\theta_1|}{\sin(\theta_1 + \theta_2)}.$$

We illustrate in Fig. 3.6 how to construct θ_1 and θ_2.

We notice that

$$\|\lambda_c\|_\infty < (\sin\theta)^{-1},$$

where $\theta = \pi - \theta_1 - \theta_2$ is the angle between the two lines $l_1 := \{x : (x - x_c)^\top \nabla f_1(x_c) = 0\}$ and $l_2 := \{x : (x - x_c)^\top \nabla f_2(x_c) = 0\}$. These two lines are actually the boundaries formed by the constraint functions $f_1(x)$ and $f_2(x)$ linearized at $x = x_c$. From the above conclusions, we see that for 2-dimensional optimization problems, we need a large Λ to satisfy Assumption 3.5 only when the angle θ is small.

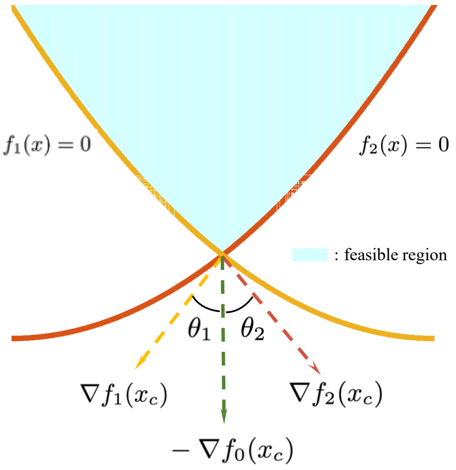

Figure 3.6: The illustration of angles θ_1 and θ_2

3.7.7 Proof of Proposition 3.2

Since (x_c, λ_c) is a KKT pair where $\|\lambda_c\|_\infty < \Lambda$, by using triangular inequalities on the norm terms defining $\delta_1(k, \lambda_c)$, we obtain

$$\begin{aligned}
&\delta_1(k, \lambda_c) \\
\leq &\|\nabla f_0(x_c) + \sum_{i=1}^{m} \lambda_c^{(i)} \nabla f_i\left(x_c\right)\| + \|\nabla f_0(x_k) - \nabla f_0(x_c)\| \\
&+ 4\mu \|x_{k+1} - x_k\| + \sum_{i=1}^{m} \Lambda \Big(\|\nabla^{\nu_k^*} f_i\left(x_k\right) - \nabla f_i\left(x_k\right)\| \\
&+ \|\nabla f_i\left(x_k\right) - \nabla f_i\left(x_c\right)\| \Big) + 4mM_{\max}\Lambda \|x_{k+1} - x_k\|.
\end{aligned}$$

Similar computations for $\delta_2^{(i)}(k, \lambda_c)$ give

$$\begin{aligned}
\delta_2^{(i)}(k, \lambda_c) \leq &|\lambda_c^{(i)} f_i(x_c)| + \Lambda |f_i(x_k) - f_i(x_c)| \\
&+ \Lambda \Big(L_i |x_{k+1} - x_k| + 2M_i \|x_{k+1} - x_k\|^2 \Big).
\end{aligned}$$

We let $\{x_{k_p}\}_{p \geq 1}$ be an subsequence that converges to x_c. Considering that the gradient estimation error converges to 0 (see (3.9) and Lemma 3.1), we know the term $\|\nabla^{\nu_{k_p}^*} f_i\left(x_{k_p}\right) -$

$\nabla f_i\left(x_{k_p}\right) \|$ converges to 0 as p goes to infinity. Therefore, we have

$$\lim_{p\to\infty} \delta_1(k_p, \lambda_c) = 0 \text{ and } \lim_{p\to\infty} \max_{1\le i\le m} |\delta_2^{(i)}(k_p, \lambda_c)| = 0.$$

Thus, for any k_0 and $\eta > 0$, one can find $k_\Lambda > k_0$ such that $\max\{\delta_1(k_\Lambda, \lambda_c), \max_{1\le i\le m} |\delta_2^{(i)}(k_\Lambda, \lambda_c)|\} < \eta/2$. For $k = k_\Lambda$ in SZO-QQ, we have that $\lambda_{k_\Lambda+1}$, the solution to (SP2), has an infinite norm less than 2Λ because λ_c is a feasible solution to (SP2) and $\|\lambda_c\|_\infty < 2\Lambda$, which is to say that the second termination condition of Algorithm 2 is satisfied when $k = k_\Lambda$.

Since $\|x_{k+1} - x_k\|$ converges to 0 as k goes to infinity (see Proposition 3.1), we can choose k_0 to be sufficiently large so that, for any $k > k_0$, $\|x_{k+1} - x_k\| \le \xi$. Thus, the two termination conditions are satisfied when $k = k_\Lambda$.

3.7.8 Proof of Theorem 3.3

The pair $(\tilde{x}, \tilde{\lambda})$ and the index $\tilde{k}$ returned by Algorithm 2 satisfy $\tilde{x} = x_{\tilde{k}}$ and

$$\max\left\{\delta_1(\tilde{k}, \tilde{\lambda}), \max\{\delta_2^{(i)}(\tilde{k}, \tilde{\lambda}) : i \ge 1\}\right\} \le \frac{\eta}{2}. \tag{3.26}$$

By using triangular inequalities we have for any i,

$$\begin{aligned} &\|\nabla^{\nu^*_{\tilde{k}-1}} f_i\left(x_{\tilde{k}-1}\right) - \nabla f_i\left(x_{\tilde{k}}\right)\| \\ \le&\|\nabla^{\nu^*_{\tilde{k}-1}} f_i\left(x_{\tilde{k}-1}\right) - \nabla f_i\left(x_{\tilde{k}-1}\right)\| + \|\nabla f_i\left(x_{\tilde{k}}\right) - \nabla f_i\left(x_{\tilde{k}-1}\right)\| \\ \le&\alpha\xi + L_i\xi. \end{aligned} \tag{3.27}$$

Then, based on (3.13), (3.26) and (3.27), we have

$$\begin{aligned} &||\nabla f_0\left(x_{\tilde{k}}\right) + \sum_{i=1}^m \tilde{\lambda}^{(i)} \nabla f_i\left(x_{\tilde{k}}\right)\| \\ \le&\left\|\nabla f_0\left(x_{\tilde{k}}\right) + \sum_{i=1}^m \tilde{\lambda}^{(i)}\left(\nabla^{\nu^*_{\tilde{k}-1}} f_i\left(x_{\tilde{k}-1}\right) + 4M_i(x_{\tilde{k}} - x_{\tilde{k}-1})\right) + 2\mu(x_{\tilde{k}} - x_{\tilde{k}-1})\right\| \\ &+ \|2\mu(x_{\tilde{k}} - x_{\tilde{k}-1})\| + 2\Lambda\sum_{i=1}^m\left(4\left\|M_i(x_{\tilde{k}} - x_{\tilde{k}-1})\right\|\right) \\ &+ 2\Lambda\sum_{i=1}^m\left(\left\|\nabla^{\nu^*_{\tilde{k}-1}} f_i\left(x_{\tilde{k}-1}\right) - \nabla f_i\left(x_{\tilde{k}}\right)\right\|\right) \\ \le&\eta/2 + 2\Lambda\sum_{i=1}^m\left(5M_i\xi + \alpha_i\nu^*_{\tilde{k}-1}\right) + 2\mu\xi \le \eta, \end{aligned}$$

and

$$\begin{aligned}
&\left\|\tilde{\lambda}^{(i)} f_i(x_{\tilde{k}})\right\| \\
\leq & \left\|\tilde{\lambda}^{(i)}\left(f_i(x_{\tilde{k}-1}) + \nabla^{\nu^*_{\tilde{k}-1}} f_i\left(x_{\tilde{k}-1}\right)(x_{\tilde{k}} - x_{\tilde{k}-1}) + 2M_i\|x_{\tilde{k}} - x_{\tilde{k}-1}\|^2\right)\right\| \\
&+ 2\Lambda\Bigg(\|f_i(x_{\tilde{k}}) - f_i(x_{\tilde{k}-1})\| + \|\nabla^{\nu^*_{\tilde{k}-1}} f_i\left(x_{\tilde{k}-1}\right) - \nabla f_i\left(x_{\tilde{k}-1}\right)\| \cdot \|x_{\tilde{k}} - x_{\tilde{k}-1}\| \\
&+ \|\nabla f_i(x_{\tilde{k}-1})\| \cdot \|x_{\tilde{k}} - x_{\tilde{k}-1}\| + 2M_i\|x_{\tilde{k}} - x_{\tilde{k}-1}\|^2\Bigg) \\
\leq & \eta/2 + 2\Lambda(2L_i\xi + \alpha_i\xi^2 + 2M_i\xi^2) \\
\leq & \eta/2 + 2\Lambda(2L_i + \alpha_i + 2M_i)\xi \leq \eta,
\end{aligned}$$

which concludes the proof.

3.7.9 Proof of Lemma 3.2

We assume x_τ is an accumulation point of $\{x_k\}_{k\geq 1}$ and is also a strict local minimizer. We show the convergence of $\{x_k\}_{k\geq 1}$ by contradiction. We assume that $\mathcal{C} \setminus \{x_\tau\} \neq \emptyset$, where $\mathcal{C}$ is the set of accumulation points of $\{x_k\}_{k\geq 1}$. Then, there exists $\epsilon > 0$ such that $\mathcal{C} \cap (\Omega \setminus \mathcal{B}(x_\tau, \epsilon)) \neq \emptyset$ and any $x \in \mathcal{B}(x_\tau, \epsilon) \setminus \{x_\tau\}$ verifies $f_0(x_\tau) < f_0(x)$. Since the sphere $\mathcal{SP}(x_\tau, \epsilon) = \{x : \|x - x_\tau\| = \epsilon\}$ is compact, we let $\sigma = \inf_{x \in \mathcal{SP}(x_\tau, \epsilon)} f_0(x)$ and have $\sigma > f_0(x_\tau)$. Therefore, there exists $k_\alpha > 0$ such that $f_0(x_{k_\alpha}) < (\sigma + f_0(x_\tau))/2$.

Since there exists an accumulation point outside $\mathcal{B}(x_\tau, \epsilon)$ and $\{x_{k+1} - x_k\}_{k\geq 1}$ converges to 0, we can find $k_\beta > k_\alpha$ such that $x_{k_\beta} \in \mathcal{B}(x_\tau, \epsilon)$, $x_{k_\beta+1} \notin \mathcal{B}(x_\tau, \epsilon)$ and $\|x_{k_\beta+1} - x_{k_\beta}\| \leq (\sigma - f_0(x_\tau))/(4L_{\max})$. Let $\tilde{x} = \{x : \text{there exists } t \in [0,1] \text{ such that } x = tx_{k_\beta} + (1-t)x_{k_\beta+1}\} \cap \mathcal{SP}(x_\tau, \epsilon)$, i.e., $\tilde{x}$ is the intersection of $\mathcal{SP}(x_\tau, \epsilon)$ and the line segment between x_{k_β} and $x_{k_\beta+1}$. Then, $\|x_{k_\beta} - \tilde{x}\| \leq (\sigma - f_0(x_\tau))/(4L_{\max})$ and thus

$$\begin{aligned}
f_0(x_{k_\beta}) &\geq f_0(\tilde{x}) - \frac{\sigma - f_0(x_\tau)}{4} \\
&\geq \sigma - \frac{\sigma - f_0(x_\tau)}{4} \\
&> (\sigma + f_0(x_\tau))/2 > f_0(x_{k_\alpha}),
\end{aligned}$$

which contradicts with the monotonicity of $\{f_0(x_k)\}_{k\geq 1}$ shown in Proposition 3.1. ■

3.7.10 Proof of Lemma 3.3

To begin with, we let $\mathcal{D}_\lambda(y,\nu) \subset \mathbb{R}^m$ be the optimal solution set of the dual of the following convex problem:

$$\begin{aligned}\mathsf{P}(y,\nu) : \min_{x\in\mathbb{R}^d} \quad & f_0(x) + \mu\|x-y\|^2 \\ \text{subject to} \quad & f_i(y) + \Bigg(\underbrace{\Delta_i^\nu(y) + \nabla f_i(y)}_{\nabla_i^\nu f_i(y)}\Bigg)^\top (x-y) + 2M_i\|x-y\|^2 \le 0.\end{aligned} \tag{3.28}$$

Notice that the problem $\mathsf{P}(x_k, \nu_k^*)$ coincides with (SP1) in Algorithm 2. Then, we have

$$\begin{aligned}\mathcal{D}_\lambda(y,\nu) := \arg\max_{\lambda\ge 0}\min_x f_0(x) + \mu\|x-y\|^2 + \sum_{i=1}^m \lambda^{(i)}\Bigg(f_i(y) \\ \Big(\Delta_i^\nu(y) + \nabla f_i(y)\Big)^\top (x-y) + 2M_i\|x-y\|^2\Bigg).\end{aligned}$$

By solving the inner minimization problem which is an unconstrained convex quadratic programming, we know that there exist $p \in \mathbb{R}_{\ge 0}^{m\times 1}$ and $a \in \mathbb{R}_{>0}$, independent of y and ν, such that

$$\mathcal{D}_\lambda(y,\nu) = \arg\max_{\lambda\ge 0} G(\lambda, y, \nu),$$

where for $\lambda \in \mathbb{R}_{\ge 0}^m$,

$$G(\lambda, y, \nu) := \frac{\lambda^\top Q(y,\nu)\lambda + q^\top(y,\nu)\lambda + b(y,\nu)}{p^\top\lambda + a},$$

the functions $Q(y,\nu) \in \mathbb{R}^{m\times m}, q(y,\nu) \in \mathbb{R}^{m\times 1}, b(y,\nu) \in \mathbb{R}$ are continuous in (y,ν) and $Q(y,\nu)$ is negative definite.

From the continuity of $Q(y,\nu), q(y,\nu)$ and $b(y,\nu)$, the function $G(\lambda,y,\nu)$ is continuous in all arguments for $\lambda \ge 0$. Due to the continuity and the uniqueness of the optimal dual solution $\mathcal{D}_\lambda(x_c, 0) = \{\lambda_c\}$ (see Lemma 3.5), we can use perturbation theory [143, Proposition 4.4] to conclude that $\mathcal{D}_\lambda(y,\nu)$ is upper semicontinuous at $(y,\nu) = (x_c, 0)$.

Definition 3.2. *Let W and V be two vector spaces. A multifunction $F : W \to \mathcal{P}(V)$, where $\mathcal{P} := \{P : P \subset V\}$, is said to be upper semicontinuous at w_0 if for any neighborhood $\mathcal{N}_V$ of $F(w_0)$, there exists a neighborhood $\mathcal{N}_W$ of w_0 such that the inclusion $F(w) \subset \mathcal{N}_V$ holds for any $w \in \mathcal{N}_W$.*

Considering the convergence of (x_k, ν_k^*) to $(x_c, 0)$ and the upper semicontinuity of $\mathcal{D}_\lambda(y,\nu)$ at $(y,\nu) = (x_c, 0)$, for any $\delta > 0$, there exists $k_\delta > 0$ such that $\mathcal{D}_\lambda(x_k, \nu_k^*) \subset \mathcal{B}(\lambda_c, \delta)$ for any $k > k_\delta$. Since $\lambda_{k+1}^\circ \in \mathcal{D}_\lambda(x_k, \nu_k^*) \subset \mathcal{B}(\lambda_c, \delta)$, we have $\lambda_{k+1}^\circ \in \mathcal{B}(\lambda_c, \delta)$ for any $k > k_\delta$.

3.7.11 Proof of Theorem 3.4

According to Lemma 3.3, there exists $\bar{k} > 0$ such that $\|\lambda^{\circ}_{k+1}\|_\infty \leq 2\Lambda$ for any $k \geq \bar{k}$. Since λ°_{k+1} is a feasible solution of (SP2) in Algorithm 2, λ_{k+1}, the optimal solution of (SP2), also satisfies $\|\lambda_{k+1}\|_\infty \leq 2\Lambda$.

Recall the definition of $h(\eta)$ in (3.13). We let $\bar{\xi} := \inf_{k \leq \bar{k}} \|x_{k+1} - x_k\|$, $\bar{\eta} := \inf\{\eta : h(\eta) \geq \bar{\xi}/2\}$ and consider the case where $\eta < \bar{\eta}$. We first notice that if $\|x_{k+1} - x_k\| \leq h(\eta)$, we have $\|x_{k+1} - x_k\| < \bar{\xi}$ and thus $k > \bar{k}$. We then let $\mathcal{K}(\eta) := \max\{k : \|x_{k+1} - x_k\| > h(\eta)\} + 1$. Since $\|x_{\mathcal{K}(\eta)+1} - x_{\mathcal{K}(\eta)}\| \leq h(\eta)$, we have that $\mathcal{K}(\eta) > \bar{k}$ and thus $\|\lambda_{k+1}\|_\infty \leq 2\Lambda$, which is equivalent to say that with $k = \mathcal{K}(\eta)$, the two termination conditions in Algoirthm 2 are satisfied. Then $\tilde{k}$, the iteration number returned by Algorithm 2, verifies that $\tilde{k} \leq \mathcal{K}(\eta) + 1$. According to (3.12),

$$f_0(x_0) - \inf\{f_0(x) : x \in \Omega\} \geq f_0(x_0) - f_0(x_{\mathcal{K}(\eta)}) \geq \mu\mathcal{K}(\eta)(h(\eta))^2 \tag{3.29}$$

and thus $\tilde{k} \leq \mathcal{K}(\eta) + 1 \leq \overline{\mathcal{K}}(\eta) + 1$. Therefore, according to the definition of $h(\eta)$ in (3.13) there exists $A_1 > 0$ such that $\mathcal{K}(\eta) + 1 \leq A_1((L_{\max} + M_{\max})/\eta)^2$.

For $\bar{\eta} \leq \eta \leq 1$, we let $A_2 = \sup_{\bar{\eta} \leq \eta \leq 1} \frac{\mathcal{K}(\eta)+1}{(L_{\max}+M_{\max})^2}$. According to the definition of $\mathcal{K}(\eta)$, we have that $\mathcal{K}(\eta)$ is monotonously decreasing with respect to η. Therefore, $\mathcal{K}(\eta) \leq \mathcal{K}(\bar{\eta}) \leq \overline{\mathcal{K}}(\bar{\eta})$ and thus A_2 is finite. Since $\mathcal{K}(\eta) + 1 \leq A_2((L_{\max} + M_{\max})/\eta)^2$, by letting $A = \max\{A_1, A_2\}$, we have for any $0 < \eta \leq 1$, $\mathcal{K}(\eta) + 1 \leq A((L_{\max} + M_{\max})/\eta)^2$.

3.8 Final Remarks

The most notable feature of safe zeroth-order optimization methods is the feasibility of samples. In this chapter, we focused on achieving sample feasibility by constructing local feasible sets based on proxies of the constraint functions and their corresponding error upper bounds. As the iterates approach the boundary of the feasible region, the local feasible sets become smaller, resulting in slower progress of the iterates (see Section 3.7.2). To ensure a rapid decrease in the objective function value, it is essential to maximize the size of the local feasible sets. This can be achieved through two potential directions. Firstly, we can use less conservative proxies and error upper bounds, for instance, by incorporating local smoothness constants. Secondly, we can avoid getting too close to the boundary by keeping the iterates away from it. In next chapter, our primary focus will be on exploring the latter direction.

4 Safe Zeroth-Order Optimization Using Linear Programs

4.1 Introduction

In this chapter, we look into the same black-box optimization problem as in Chapter 3. Therefore, we inherit the problem formulation and the notations from Sections 3.2 and 3.1.3. As we have seen in Section 3.6.3, regarding the application of SZO-QQ to large-scale problems, most computation time is spent on solving QCQP subproblems and the iterate progress slows down when the iterates get close to the feasible region boundary. To further decrease the computation and sampling complexities, in this chapter, we propose an alternative method where the subproblems are easier to solve than QCQPs in SZO-QQ and the iterates have the tendency to stay away from the feasible region boundary.

4.1.1 Contributions

The main contributions of this chapter are summarized as follows:

(a) We present a novel approach called Safe Zeroth-Order optimization using Linear Programs (SZO-LP). This method iteratively solves linear programming subproblems to derive descent directions and then decides the step length by sampling;

(b) We show that, under mild assumptions, a subsequence of SZO-LP's iterates converges to the primal of a KKT pair (see Definition 3.1);

(c) By application to an IEEE 30-bus benchmark problem, we show that SZO-LP can efficiently solve an OPF problem with 11 decision variables and 158 constraints. We compare SZO-LP with state-of-the-art approaches and demonstrate its advantages in terms of computation time and the number of samples required.

4.1.2 Structure of this chapter

We present SZO-LP in Section 4.2 and show the convergence properties of this method in Section 4.3. In Section 4.4, we demonstrate the performance of SZO-LP when used to solve the OPF problem.

4.2 Algorithm: SZO-LP

In this section, we present the necessary tools required by SZO-LP and the steps of its implementation. Unlike SZO-QQ, where the gradient estimation step size is explicitly expressed by the rule (3.9), in SZO-LP, we adjust the step size based on the requirement for gradient estimation accuracy. For this aim, we let

$$\nu(\epsilon) := \frac{2\epsilon}{\sqrt{d}M_{\max}}.$$

This function describes the step size for gradient estimation in (3.5) needed to achieve an gradient estimation error less than ϵ, as indicated in the following proposition.

Proposition 4.1. *Recall from Section 3.3 that* $\Delta_i^{\nu}(x) := \nabla^{\nu} f_i\,(x) - \nabla f_i(x)$. *Then,*

$$\|\Delta_i^{\nu(\epsilon)}(x)\| \leq \epsilon.$$

This proposition is a direct consequence of Lemma 3.1. Recall that for the initial point x_0 we can build a local feasible set $\mathcal{S}^{(0)}(x_0)$ as in (4.1). In this chapter, we make a slight modification to the local feasible set construction. Specifically, we let $l_0^* = \min_{i\in\{1,\ldots,m\}} \; -f_i(x_0)/L_{\max}$ and

$$\nu_0^*(\epsilon) := \min\{l_0^*/\sqrt{d}, \nu(\epsilon)\}.$$

Since $\nu_0^*(\epsilon) \leq \nu(\epsilon)$, we have $\|\Delta_i^{\nu_0^*(\epsilon)}(x_0)\| \leq \epsilon$ for any $\epsilon > 0$ and $i \geq 1$. Then given any $\epsilon_0 > 0$, the set

$$\begin{aligned} &\mathcal{S}^{(0)}(x_0) := \cap_{i=1}^m \mathcal{S}_i^{(0)}(x_0), \text{ where} \\ &\mathcal{S}_i^{(0)}(x_0) := \Big\{x : f_i(x_0) + \nabla^{\nu_0^*(\epsilon_0)} f_i\,(x_0)^\top\,(x - x_0) + 2M_i\|x - x_0\|^2 \leq 0\Big\}, \end{aligned} \tag{4.1}$$

is feasible as shown in the following proposition, which is based on Theorem 3.1.

Proposition 4.2. *All the samples used to construct* $\mathcal{S}^0(x_0)$ *are feasible. Moreover, the set* $\mathcal{S}^{(0)}(x_0)$ *is convex and any* $x \in \mathcal{S}^{(0)}(x_0)$ *is strictly feasible.*

In the lack of explicit constraint functions, a local feasible set is a common tool of several zeroth-order methods [25, 24, 27] to ensure the feasibility of the iterates, though the specific formulations are different. In the following, we propose our method where the local feasible sets are used to select the step length for the derived descent direction.

4.2.1 Algorithm: Safe Zeroth-Order optimization using Linear Programs (SZO-LP)

The main idea of the SZO-LP method, shown in Algorithm 3, is to iteratively select a descent direction by executing in Line 7 $\texttt{LP}(x_k, \epsilon_k)$ defined in (4.4). Thanks to the tightening contant ϵ_k in the linear program involved in $\texttt{LP}(x_k, \epsilon_k)$, the descent direction we obtain points into the iterior of the feasible set. Along this direction, we select the step length (Line 9-14) based on local feasible sets and the pre-defined length

$$\gamma(\epsilon_k) := \frac{\epsilon_k}{4(M_{\max} + L_{\max})}.$$

Algorithm 3 Safe Zeroth-Order optimization using Linear Programs (SZO-LP)

Input: ϵ_0, $\epsilon_{\min}$, K_{switch}, initial feasible point $x_0 \in \Omega$

Output: $\tilde{x}$

```
1:  $k \leftarrow 0, \text{TER} = 0$
2:  while $\epsilon_k > \epsilon_{\min}$ do
3:      $s_{\text{tmp}} \leftarrow \texttt{LP}(x_k, 2\epsilon_k)$
4:      if $\nabla^{\nu_k^*(2\epsilon_k)} f_0(x_k)^\top s_{\text{tmp}} \le -4\epsilon_k$ then
5:          $\epsilon_{k+1} \leftarrow 2\epsilon_k$, $x_{k+1} \leftarrow x_k$
6:      else
7:          $s_k^* = \texttt{LP}(x_k, \epsilon_k)$
8:          if $\nabla^{\nu_k^*(\epsilon_k)} f_0(x_k)^\top s_k^* \le -2\epsilon_k$ then
9:              if $k < K_{\text{switch}}$ then
                    $\beta_k = \arg\max_{\beta \ge 0} \beta \text{ s.t. } x_k + \beta s_k^* \in \mathcal{S}^{(k)}(x_k),$          (4.2)
                    $\alpha_k = \arg\min_{\alpha \in \{\beta_k, \gamma(\epsilon_k)\}} f_0(x_k + \alpha s_k^*)$          (4.3)
10:                 $x_{k+1} \leftarrow x_k + \alpha_k s_k^*$, $\epsilon_{k+1} \leftarrow \epsilon_k$
11:             else
12:                 $x_{k+1} \leftarrow x_k + \gamma(\epsilon_k) s_k^*$, $\epsilon_{k+1} \leftarrow \epsilon_k$
13:             end if
14:         else
15:             $\epsilon_{k+1} \leftarrow \epsilon_k/2$, $x_{k+1} \leftarrow x_k$
16:         end if
17:     end if
18:     $k \leftarrow k + 1$
19: end while
```

The essential steps are as follows:

Providing the input data

The input includes an initial strictly feasible point x_0 (see Assumption 3.2) and a tightening constant ϵ_0. Each iteration of the algorithm generates a new tightening constant ϵ_k, which can be equal to ϵ_{k-1}, $2\epsilon_{k-1}$ or $\epsilon_{k-1}/2$. Since ϵ_k converges to 0 (see Theorem 4.1), the user can control the termination by providing a lower bound $\epsilon_{\min}$ for ϵ_k. The parameter K_{switch} marks the boundary of two methods for selecting step length, see the last bullet point.

Building local feasible sets

For a strictly feasible x_k, we use (3.7) to define l_k^* and

$$\nu_k^*(\epsilon_k) := \min\{l_k^*/\sqrt{d}, \nu(\epsilon_k)\}.$$

We then use $\nu_k^*(\epsilon_k)$ and (4.1) to define $\mathcal{S}^{(k)}(x_k)$, a local feasible set around x_k. From Theorem 4.2 we know that if $x_{k+1} \in \mathcal{S}^{(k)}(x_k)$ then x_{k+1} is also strictly feasible.

Solving subproblems for the descent diretion

In each iteration, we execute in Line 7 $\texttt{LP}(x_k, \epsilon_k)$ to derive a search direction, which returns

$$\begin{aligned} \operatorname*{arg\,min}_{\|s\|_1 \leq 1} \quad & (\nabla^{\nu_k^*(\epsilon_k)} f_0(x_k))^\top s \\ \text{s.t.} \quad & (\nabla^{\nu_k^*(\epsilon_k)} f_i(x_k))^\top s + 2\epsilon_k \leq 0, \\ & \forall i \in \mathcal{A}(x_k, \epsilon_k), \end{aligned} \tag{4.4}$$

or NaN if (4.4) is not feasible. Here, $\mathcal{A}(x, \epsilon) := \{i : f_i(x) \geq -2\epsilon\}$ is the near-active constraint index set. The solution to (4.4) is a direction that not only gives a fast descent but also points into the interior of the feasible region Ω (away from the boundary). In (4.4), due to the tightening constant ϵ_k, along the direction s_k^* in Line 7, the constraint function values decrease. Therefore, moving along the direction s_k^* we indeed stay away from the boundary of Ω. This direction helps to avoid small values of $-f_i(x_k)$, which lead to conservative local feasible sets $\mathcal{S}^{(k)}(x_k)$. Moreover, the inclusion of only near-active constraints makes (4.4) small-size and easy to solve. We will later see in Theorem 4.1 that ϵ_k converges to 0. Therefore, it is still possible that a subsequence of the iterates converges to a point on the feasible set boundary.

We also let $s_{\text{tmp}} = \texttt{LP}(x_k, 2\epsilon_k)$ and check in Line 4 whether $\nabla^{\nu_k^*(2\epsilon_k)} f_0(x_k)^\top s_{\text{tmp}} \leq -4\epsilon_k$, which allows us to have Proposition 4.3, the proof of which is in Section 4.5.1. This proposition will be later used to show in Theorem 4.2 the properties of the $\{x_k\}_{k\geq 1}$ as k goes to infinity.

Proposition 4.3. *Any ϵ_k entering Line 7 satisfies*

$$\begin{aligned}\epsilon_k \geq \frac{1}{8}\sup\{\epsilon : & s = \mathtt{LP}(x_k, \epsilon) \textit{ verifies} \\ & \nabla^{\nu_k^*(\epsilon)} f_0(x_k)^\top s \leq -2\epsilon\}.\end{aligned} \tag{4.5}$$

Deciding the step length

When a direction s_k^* derived in Line 7 gives sufficient descent (i.e., $\nabla^{\nu_k^*(\epsilon_k)} f_0(x_k)^\top s_k^* \leq -2\epsilon_k$), we move along the tentative direction s_k^*. To decide the step length, we consider the local feasible set and the pre-defined step length $\gamma(\epsilon_k)$ that is guaranteed to achieve a non-trivial descent (see Lemma 4.1). In (4.2), we calculate by bisection the largest step length within the local feasible set to derive α_k in (4.3). The use of local feasible sets in Line 10 allows us to obtain a larger step length than $\gamma(\epsilon)$, when x_k is not close to the boundary of the feasible set. This is because, from the formulation (4.1), smaller values of $f_i(x_k)$ lead to larger sizes of $\mathcal{S}_i^{(k)}(x_k)$ while $\gamma(\epsilon_k)$ is independent of how far the iterates are from the feasible set boundary. When $k > K_{\text{switch}}$, we let the step length be $\gamma(\epsilon_k)$ as in Line 12, which is useful for the proof of the iterates' properties as k goes to infinity (see Theorem 4.2). The selection of K_{switch} is not critical since we use the step length in Line 10 for $k < K_{\text{switch}}$ instead of that defined in Line 12 only to accelerate the descent in the early iterations of the algorithm.

On the other hand, if the direction s_k^* cannot give sufficient descent, we let $\epsilon_{k+1} = \epsilon_k/2$ in Line 15 to relax the tightened constraints in (4.4). This relaxation makes it easier for s_{k+1}^* to give sufficient descent, i.e., to satisfy $\nabla^{\nu_k^*(\epsilon_{k+1})} f_0(x_{k+1})^\top s_{k+1}^* \leq -2\epsilon_{k+1}$. Only when s_{k+1}^* gives sufficient descent will we move along s_{k+1}^* to a new point.

We refer the readers to Remark 4.1 for how SZO-LP is compared with some state-of-the-art methods.

4.3 Convergence Properties of the Approach

In this section, we aim to show that, under mild conditions and by letting $\epsilon_{\min} = 0$, the sequence $\{x_k\}_{k\geq 1}$ produced in Algorithm 3 has an accumulation point x_c that is also the primal of a KKT pair of (3.1). To start with, we show in Lemma 4.1 that, whenever $x_{k+1} \neq x_k$, the new iterate x_{k+1} is strictly feasible and the objective function value gets a non-trivial decrease.

Lemma 4.1. *Suppose s_k^* derived in Line 7 of Algorithm 3, satisfies*

$$\nabla^{\nu_k^*(\epsilon_k)} f_0(x_k)^\top s_k^* \leq -2\epsilon_k.$$

We have that $x_k + \gamma(\epsilon_k)s_k^*$ is strictly feasible. Furthermore $x_k + \gamma(\epsilon_k)s_k^*$ satisfies

$$f_0(x_k + \gamma(\epsilon_k)s_k^*) - f_0(x_k) < -\epsilon_k^2/(8(M_{\max} + L_{\max})). \tag{4.6}$$

The proof of Lemma 4.1 is in Section 4.5.2. The main idea is to utilize the smoothness constants in Assumption 3 to upper-bound $f_i(x_k + \gamma(\epsilon_k))$ for $i \in \mathbb{Z}_0^m$. Based on this lemma, we have the following theorem on the sequences $\{x_k\}_{k\geq 1}$ and $\{\epsilon_k\}_{k\geq 1}$ as k goes to infinity.

Theorem 4.1. *The following arguments hold:*

1. *The sequence $\{f_0(x_k)\}_{k\geq 1}$ is non-increasing;*
2. *There exists at least one accumulation point of the sequence $\{x_k\}_{k\geq 1}$. For any accumulation point x_c,*
$$\lim_{k\to\infty} f_0(x_k) = f_0(x_c) > -\infty.$$
3. *The sequence $\{\epsilon_k\}_{k\geq 1}$ converges to 0.*

Proof. The first point is a direct consequence of Lemma 4.1, which implies that whenever the iterate moves to a new point the objective function value decreases.

Proof of Point 2. Since $\{f_0(x_k)\}_{k\geq 1}$ is non-increasing, $f_0(x_k) \leq f_0(x_0)$, for any $k \geq 1$, and thus $x_k \in \mathcal{P}_\beta$. Due to the boundedness of $\mathcal{P}_\beta$, by the Bolzano–Weierstrass theorem, we know that there exists at least one accumulation point of $\{x_k\}_{k\geq 1}$. For any accumulation point x_c, there exists a subsequence $\{x_{k_p}\}_{p\geq 1}$ converging to x_c. Due to the continuity of $f_0(x)$,

$$\lim_{p\to\infty} f_0(x_{k_p}) = f_0(x_c) > -\infty.$$

By utilizing again the monotonicity of $\{f_0(x_k)\}_{k\geq 1}$, we have $\lim_{k\to\infty} f_0(x_k) = f_0(x_c) > -\infty$.

Proof of Point 3. We show this result through contradiction by assuming that $\{\epsilon_k\}_{k\geq 1}$ does not diminish as k goes to infinity. Based on this assumption, we can show that $\{\epsilon_k\}_{k\geq 1}$ does not converge to any non-zero values. If $\{\epsilon_k\}_{k\geq 1}$ converges to a non-zero value, by noticing that $\epsilon_k \in \{\epsilon_0 * 2^i : i \in \mathbb{Z}\}$, we have that there exists $K > 0$ such that any $k > K$ verifies $\epsilon_k = \epsilon_{k-1}$. Then for any $k > K$ the new iterate x_{k+1} is derived in Line 10 or 12 in Algorithm 3. According to Lemma 4.1, $f_0(x_{k+1}) \leq f_0(x_k) - \epsilon_k^2/(8(M_{\max} + L_{\max}))$ and thus $f_0(x_k)$ goes to $-\infty$ as k goes to $+\infty$, which contradicts Point 2. Therefore, ϵ_k does not converge.

Since from Algorithm 3 $\{\epsilon_k\}_{k\geq 1}$ is bounded, we can conclude that $\{\epsilon_k\}_{k\geq 1}$ has multiple accumulation points. Then there are $\epsilon > 0$ and infinitely many k such that $\epsilon_k = \epsilon$ and $\epsilon_{k-1} = \epsilon/2$. For any k of this kind, there exists $k' \geq k$ verifying $f_0(x_{k'+1}) \leq f_0(x_{k'}) - \epsilon^2/(8(M_{\max} + L_{\max}))$. Consequently there are infinitely many k' verifying $f_0(x_{k'+1}) \leq f_0(x_{k'}) - \epsilon^2/(8(M_{\max} + L_{\max}))$, which again contradicts Point 2. ■

Theorem 4.1 offers us the essential tools to show in Theorem 4.2 the properties of an accumulation point of $\{x_k\}_{k\geq 1}$ under Assumption 3.4.

Theorem 4.2. *Suppose the iterates $\{x_k\}_{k\geq 1}$ with an accumulation point x_c satisfy Assumption 3.4, then there exists $\lambda_c \in \mathbb{R}^m_{\geq 0}$ such that (x_c, λ_c) is a KKT pair of* (3.1).

The proof, in Section 4.5.3, is based on contradiction. If x_c is not the primal of a KKT pair, we can find $r > 0$, $\epsilon > 0$ and $s_\epsilon \in \mathbb{R}^d$ such that for any $x_k \in \mathcal{B}_r(x_c)$ the solution $s = \texttt{LP}(x_k, \epsilon)$ verifies $\nabla^{\nu_k^*(\epsilon)} f_0(x_k)^\top s \leq -2\epsilon$. There are infinitely many k such that $x_k \in \mathcal{B}_r(x_c)$ and s_k^* is derived through Line 7 in Algorithm 3. For any of these ks, according to (4.5), $\epsilon_k \geq \epsilon/8$, which contradicts Point 3 of Theorem 4.1.

Remark 4.1. *Like SZO-QQ [27] and LB-SGD [24], the samples in SZO-LP are all feasible and the iterates, under mild assumptions, have an accumulation point that is also the primal of a KKT pair. In contrast, the tightening constant ϵ_k of SZO-LP keeps the iterates away from the boundary of the feasible set and leads to less conservative local feasible sets than those used in SZO-QQ and LB-SGD. Moreover, due to the use of the near-active set $\mathcal{A}(x_k, \epsilon_k)$ the subproblems* (4.4) *are smaller-size and easier to solve than the QCQPs in SZO-QQ and nonconvex subproblems in Safe Bayesian Optimization methods [25, 144]. However, to rigorously show these advantages, we need to upper bound the number of iterations needed by SZO-LP given certain accuracy requirements, which is left as future work.*

4.4 Experiment on the OPF Problem

To illustrate the performance of SZO-LP, we apply it to the OPF problem formulated in Section 3.6.3. In total, there are 11 decision variables and 158 constraints. We do not assume knowledge of the system model for the optimization task. However, given a set of values for all 11 decision variables, we can use a black-box simulation model in `Matpower` [140] to sample the voltages of all the 30 buses and the current through all the transmission lines in the network. Additionally, we assume the availability of initial values for all the decision variables to start the SZO-LP algorithm from a feasible point.

We employ SZO-LP to reduce the quadratic cost induced by the initial decision values. The numerical experiments are executed on a PC with an Intel Core i9 processor. The solver we adopt for subproblems (4.4) is LINPROG in Matlab. We let $M_i = M_{\max} = 0.13$ and $L_i = L_{\max} = 0.5$. The tuning of these two parameters is described in Remark (3.3). Moreover, we set $\epsilon_0 = 0.05$, $\epsilon_{\min} = 10^{-6}$ and $K_{\text{switch}} = 200$.

In Figures 4.1 and 4.2 , we present the results of our numerical experiments, where we compare the performance of SZO-LP with SZO-QQ [27], LB-SGD [24] and Extremum Seeking [106]. The QCQP subproblems in SZO-QQ are solved using MOSEK. The reference solution of the OPF problem is returned by the optimization based on the true model and utilizing

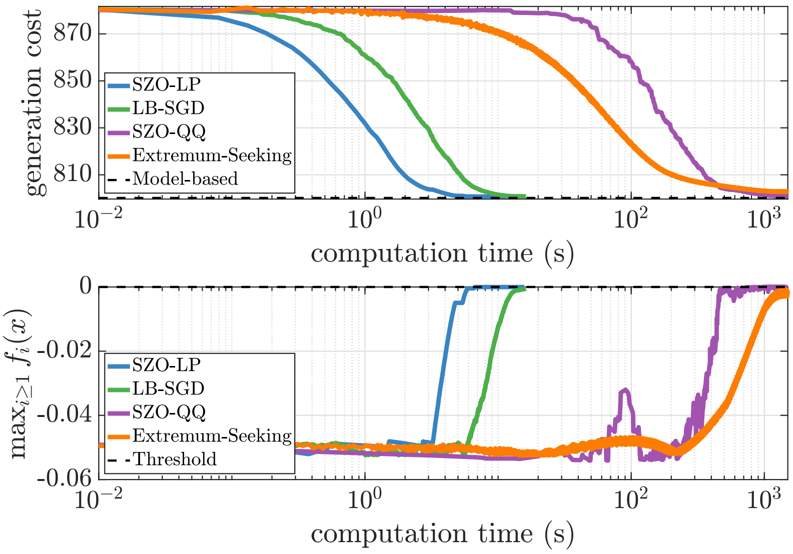

Figure 4.1: Decrease of cost and growth of the largest constraint function values with respect to computation time

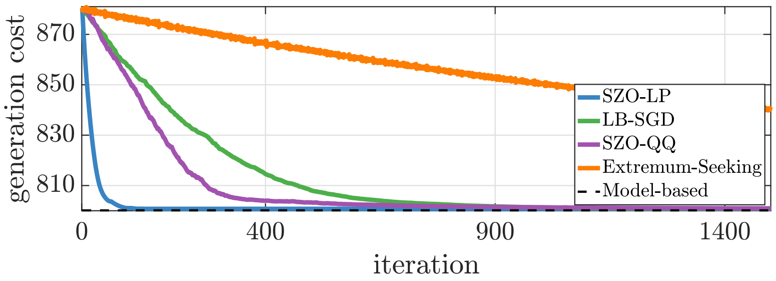

Figure 4.2: Decrease of cost with respect to the number of iterations

Gurobi [141] as the solver. The computation time in Figure 4.1 includes that consumed by power grid simulation (through Matpower) when we query the objective and constraint functions. We observe that all four methods keep the iterates feasible and eventually achieve a generation cost very close to that (800.14) derived based on the true model. However, SZO-LP achieves a faster decrease in the generation cost than the other methods.

One main reason for the superior performance of SZO-LP over SZO-QQ with respect to computation time shown in Figure 4.1, is that the linear programming subproblems can be solved faster. We notice that to finish the first 60 subproblems, SZO-LP takes 5.63 seconds while SZO-QQ takes 72.06 seconds. Firstly, the subproblem in SZO-LP only takes into account the near-active constraints while the subproblem in SZO-QQ involves all constraints. Among the iterations of SZO-LP, the largest number of constraints is 2. Secondly, although the big gap in efficiency shown in Figure 4.1 may be due to the specific solvers we select, linear programs, in general, are open to a wider selection of solvers and thus allow for more efficient implementations.

Unlike SZO-LP and SZO-QQ, LB-SGD and Extremum Seeking do not require solving any subproblems, thus allowing for more iterations within a certain time length. This is why LB-SGD can also achieve a low generation cost in a short time. However, considering the four methods take the same number of samples every iteration, LB-SGD and Extremum Seeking are less sample-efficient than SZO-LP and SZO-QQ since they require more iterations as shown in Figure 4.2. Moreover, since LB-SGD and Extremum Seeking are based on log barriers, these two methods require tuning of the barrier function coefficients. Improper tuning might lead to suboptimality in LB-SGD or even infeasibility in Extremum Seeking.

SZO-LP has another advantage over SZO-QQ, which is the feature of SZO-LP keeping the iterates away from the feasible set boundary before getting close to the primal of a KKT pair. Iterates getting too close to the feasible set boundary might impede the decrease of the cost. To see this point, we notice from Figure 4.1 that in SZO-QQ the decrease of the generation cost slows down when the largest constraint function value is larger than -0.005. The reason is that, when the largest constraint function value is close to 0, the local feasible set constructed in SZO-QQ gets conservative, and thus the step length becomes small. When the largest constraint function value gets larger than -0.005 for the first time, the generation cost in SZO-QQ is 805.27 while the corresponding cost in SZO-LP is 801.77, which is much closer to 800.14 (derived by optimization based on the true model). Therefore, we see that in SZO-QQ the decrease of the objective function value can slow down at a much earlier stage.

In conclusion, from the experiment results, we see that SZO-LP is the most computation-efficient and sample-efficient method, among the four approaches.

4.5 Appendices

4.5.1 The proof of Proposition 4.3

We first show that if for some $k > 0$ and $\epsilon_\alpha > 0$

$$s_1 = \texttt{LP}(x_k, \epsilon_\alpha) \text{ verifies } \nabla^{\nu_k^*(\epsilon_\alpha)} f_0(x_k)^\top s_2 \leq -2\epsilon_\alpha \tag{4.7}$$

then for any $\epsilon_\beta \leq \epsilon_\alpha/4$

$$s_2 = \texttt{LP}(x_k, \epsilon_\beta) \text{ verifies } \nabla^{\nu_k^*(\epsilon_\beta)} f_0(x_k)^\top s_2 \leq -2\epsilon_\beta. \tag{4.8}$$

With (4.7), we notice that s_1 with $\|s_1\|_1 \leq 1$ is a feasible solution to the linear program involved in $\texttt{LP}(x_k, \epsilon_\beta)$. This is because for any $i \in \mathcal{A}(x_k, \epsilon_\beta) \subset \mathcal{A}(x_k, \epsilon_\alpha)$, we have

$$\begin{aligned}
&\langle \nabla^{\nu_k^*(\epsilon_\beta)} f_i(x_k), s_1 \rangle \\
\leq &\langle \nabla f_i(x_k), s_1 \rangle + |\langle \nabla f_i(x_k) - \nabla^{\nu_k^*(\epsilon_\beta)} f_i(x_k), s_1 \rangle| \\
\leq &\langle \nabla^{\nu_k^*(\epsilon_\alpha)} f_i(x_k), s_1 \rangle + |\Delta_i^{\nu_k^*(\epsilon_\alpha)}(x_k)| + |\Delta_i^{\nu_k^*(\epsilon_\beta)}(x_k)| \\
\leq &-2\epsilon_\alpha + \epsilon_\alpha + \epsilon_\alpha/4 < -\epsilon_\alpha/2 \leq -2\epsilon_\beta.
\end{aligned}$$

Similarly, we can show that

$$\langle \nabla^{\nu_k^*(\epsilon_\beta)} f_0(x_k), s_1 \rangle \leq -2\epsilon_\beta.$$

Considering that s_2 is the optimum of the linear program involved in $\texttt{LP}(x_k, \epsilon_\beta)$, (4.8) holds.

Then if ϵ_k enters Line 7 of Algorithm 3, the condition in Line 4 "$s_{\text{tmp}} = \texttt{LP}(x_k, 2\epsilon_k)$ verifying $\nabla^{\nu_k^*(2\epsilon_k)} f_0(x_k)^\top s_{\text{tmp}} \leq -4\epsilon_k$" does not hold. By letting

$$\epsilon_\alpha = \sup\{\epsilon : s = \texttt{LP}(x_k, \epsilon) \text{ verifies } \nabla^{\nu_k^*(\epsilon)} f_0(x_k)^\top s \leq -2\epsilon\},$$

for any $\epsilon_\beta \leq \epsilon_\alpha/4$, (4.8) holds. Therefore, $2\epsilon_k \geq \epsilon_\alpha/4$. ■

4.5.2 Proof of Lemma 4.1

To begin with, we show that $x_k + \gamma(\epsilon_k)s_k^*$ is indeed strictly feasible. By using the mean value theorem and noticing that $\|s_k^*\| \leq \|s_k^*\|_1 \leq 1$, we have that for any $\gamma > 0$

$$\begin{aligned}
&f_i(x_k + \gamma s_k^*) \\
< &f_i(x_k) + \gamma \nabla f_i(x_k)^\top s_k^* + 2\gamma^2 M_{\max} \|s_k^*\|^2 \\
< &\gamma \nabla^{\nu_k^*(\epsilon_k)} f_i(x_k)^\top s_k^* + \gamma \|\Delta_i^{\nu_k^*(\epsilon_k)}(x)\| \cdot \|s_k^*\| + 2M_{\max}\gamma^2 \\
< &-2\epsilon_k \gamma + \epsilon_k \gamma + 2M_{\max}\gamma^2
\end{aligned}$$

$$
\begin{aligned}
&<2(M_{\max}+L_{\max})\gamma^2-\epsilon_k\gamma,\ \forall i\in\mathcal{A}(x_k,\epsilon_k), &(4.9)\\
&f_i(x_k+\gamma s_k^*)\\
&<f_i(x_k)+L_{\max}\gamma,\ \forall i\in\mathbb{Z}_1^m\setminus\mathcal{A}(x_k,\epsilon_k). &(4.10)
\end{aligned}
$$

Therefore, we have

$$f_i(x_k+\gamma(\epsilon_k)s_k^*)<-\epsilon_k^2/(8(M_{\max}+L_{\max}))<0$$

for any $i\in\mathcal{A}(x_k,\epsilon_k)$ and

$$f_i(x_k+\gamma(\epsilon_k)s_k^*)<-\epsilon_k/2$$

for any $i\in\mathbb{Z}_1^m\setminus\mathcal{A}(x_k,\epsilon_k)$. Hence, $x_k+\gamma(\epsilon_k)s_k^*$ is strictly feasible.

Similarly, we have that with $\gamma=\gamma(\epsilon_k)$ the objective function verifies

$$f_0(x_k+\gamma s_k^*)<f_0(x_k)+2(M_{\max}+L_{\max})\gamma^2-\epsilon_k\gamma.$$

Thus, $f_0(x_k+\gamma s_k^*)<f_0(x_k)-\epsilon_k^2/(8(M_{\max}+L_{\max}))$.

4.5.3 Proof of Theorem 4.2

We only consider the case where $\mathcal{A}(x_c,0)$ is not empty. The proof can be easily adapted for $\mathcal{A}(x_c,0)=\emptyset$.

We show the result through contradiction by assuming that there does not exist $\lambda_c\in\mathbb{R}^m_{\geq 0}$ such that (x_c,λ_c) is a KKT pair. Then, one and only one of the following arguments holds:

1) $\nabla f_0(x_c)$ is not a linear combination of $\nabla f_i(x_c)$, $i\in\mathcal{A}(x_c,0)$,

2) $\nabla f_0(x_c)=\sum_{i\in\mathcal{A}(x_c,0)}\lambda_i\nabla f_i(x_c)$ and there exists $i^*\in\mathcal{A}(x_c,0)$ such that $\lambda_{i^*}>0$.

We show in the following that no matter which argument holds, we can always find $s\in\mathbb{R}^d$ such that

$$\langle\nabla f_0(x_c),s\rangle<0,\ \ \langle\nabla f_i(x_c),s\rangle\leq 0,\ \forall i\in\mathcal{A}(x_c,0). \tag{4.11}$$

If 1) holds, we let $g_{\|}$ be the projection of $\nabla f_0(x_c)$ onto $\operatorname{span}\{\nabla f_i(x_c),i\in\mathcal{A}(x_c,0)\}$ and $g_\perp:=\nabla f_0(x_c)-g_{\|}$. Then $g_\perp\neq 0$, $\langle\nabla f_0(x_c),g_\perp\rangle>0$ and $\langle\nabla f_i(x_c),g_\perp\rangle=0$, $\forall i\in\mathcal{A}(x_c,0)$. Therefore, $s=-g_\perp$ satisfies (4.11).

If 2) holds, we assume without loss of generality that $\mathcal{A}(x_c,0)\neq\{i^*\}$. Then we let $h_{\|}$ be the projection of $\nabla f_{i^*}(x_c)$ onto $\operatorname{span}\{\nabla f_i(x_c),i\in\mathcal{A}(x_c,0)$ and $i\neq i^*\}$ and $h_\perp:=\nabla f_{i^*}(x_c)-h_{\|}$. Due to LICQ, $h_\perp\neq 0$. One can verify that $s=-h_\perp$ also satisfies (4.11).

Then we notice that since the set $\{s:(4.11)\text{ holds}\}$ is non-empty, there exist $\epsilon>0$ and

s_ϵ with $\|s_\epsilon\|_1 = 1$ such that

$$\langle \nabla f_i(x_c), s_\epsilon \rangle \le -4\epsilon, \ \forall i \in \mathcal{A}(x_c, 0) \cup \{0\}. \tag{4.12}$$

To see this result, we assume $\bar{s} \in \mathbb{R}^d$ satisfies $(\nabla f_0(x_c))^\top \bar{s} < 0$ and $(\nabla f_i(x_c))^\top \bar{s} \le 0$ for any $i \in \mathcal{A}(x_c, 0)$. We let $\mathcal{A}(x_c, 0) = \{i_1, \ldots, i_l\}$. There exists $y \in \mathbb{R}^d$ such that

$$Jy = \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix}, \text{ where } J = \begin{bmatrix} \nabla f_{i_1}(x_c)^\top \\ \vdots \\ \nabla f_{i_l}(x_c)^\top \end{bmatrix}, \tag{4.13}$$

because J is full row rank due to LICQ. Therefore, there exists $\sigma > 0$ such that

$$\delta_i := -(\nabla f_i(x_c))^\top (\bar{s} + \sigma y) > 0, \ \forall i \in \mathcal{A}(x_c, 0) \cup \{0\}.$$

Then $s_\epsilon = s_\epsilon^* := (\bar{s} + \sigma y)/\|\bar{s} + \sigma y\|_1$ and $\epsilon = \epsilon^* := \frac{1}{4} \min_i \delta_i / \|\bar{s} + \sigma y\|_1$ satisfy (4.12).

Due to the continuity of $\nabla f_i(x)$ for $i \in \mathbb{Z}_0^m$, there exists $r > 0$ such that any $x \in \mathcal{B}_r(x_c)$ verifies that

$$\langle \nabla f_i(x), s_\epsilon^* \rangle \le -3\epsilon^*, \ \forall i \in \mathcal{A}(x_c, 0) \cup \{0\}. \tag{4.14}$$

Since x_c is an accmulation point and $\{\epsilon_k\}_{k \ge 1}$ converges to 0, there exist infinitely many k such that

$$\begin{aligned} &k > K_{\text{switch}}, \ x_k \ne x_{k+1}, \\ &\mathcal{A}(x_c, \epsilon_k) \subset \mathcal{A}(x_c, 0), \ x_k \in \mathcal{B}_r(x_c). \end{aligned} \tag{4.15}$$

For any of these ks, considering (4.14) and for any i

$$\langle \Delta_i^{\nu_k^*(\epsilon^*)}(x), s_\epsilon^* \rangle \le |\Delta_i^{\nu_k^*(\epsilon^*)}(x)| \cdot \|s_\epsilon^*\| \le \epsilon^*,$$

we have

$$\langle \nabla^{\nu_k^*(\epsilon^*)} f_i(x), s_\epsilon^* \rangle \le -2\epsilon^*, \ \forall i \in \mathcal{A}(x_c, 0) \cup \{0\}. \tag{4.16}$$

From Algorithm 3, we see that, for any k satisfying (4.15), x_{k+1} is derived through Line 12 and s_k^* through Line 7. Therefore, we can use Proposition 4.3 and (4.16) to conclude that $\epsilon_k \ge \epsilon^*/8$ for infinitely many k. However, this conclusion contradicts with Point 3 of Theorem 4.1.

4.6 Final Remarks

In this chapter, we proposed a safe zeroth-order method SZO-LP, which iteratively solves linear programs to obtain descent directions and determines the step lengths. We showed that, under mild conditions, the iterates of SZO-LP have an accumulation point that is also the primal of a KKT pair. Through an experiment where we use SZO-LP to solve an OPF problem on the IEEE 30-bus system and compare with three other methods, we see that SZO-LP is both computation-efficient and sample-efficient. However, currently we do not have sufficient results to rigorously explain these advantages. In Algorithm 3, the termination condition "$\epsilon_k > \epsilon_{\min}$" is heuristic since it is not clear how to select $\epsilon_{\min}$ for deriving a η-KKT pair. This should be the first step to do for conducting a complete complexity analysis.

Part III
Conclusions and Further Directions

5 Conclusions and Further Directions

In this thesis, we delved into the realm of data-driven control and optimization methods, with a central focus on the challenges posed by uncertainty attenuation and quantification, while sidestepping the use of predefined parameterized models.

In Part I, we designed controllers for data-driven robust control in the presence of measurement noise. By rigorously quantifying the predictive errors stemming from data-driven system representations, we were able to ensure robust constraint satisfaction and establish upper bounds for suboptimality gaps. Our methodologies, showcased in Chapters 1 and 2, differ in their approaches to error quantification. Chapter 1 leveraged bootstrapping to upperbound the prediction error. This method is computationally expensive and lacks finite-sample guarantees. On the other hand, Chapter 2 implemented active experiment design and harnessed perturbation analysis for error quantification, which relies solely on noise range. While the error upper bounds in Chapter 2 hold almost surely, they can exhibit conservatism, as illustrated in Section 2.6.2, since the worst-case errors might arise from highly improbable noise realizations.

Given that the efficacy of robust control schemes hinges on how the uncertainties are characterized, a pivotal avenue lies in devising a data-driven trajectory prediction method that boasts rigorous and non-conservative error bounds. From the perspective of the author, three directions might be worth exploring:

(a) **Experiment design for increasing data informativity.** For example, from Section 2.3.2, if the collected historical data leads to a high value of $\sigma_{\min}(\overline{H})$, then the influence of measurement will become weaker. Traditional experiment design methods are based on good enough prior knowledge of the system model [145]. In the data-driven framework, one can use collected data to form rough system representations to assist formulating and solving experiment design problems.

(b) **Filter design for noise attenuation in recent measurements.** Several works in the literature focus on this aspect, including [146, Chapter 6] and [51]. However, it is still open how to derive error bounds for the prediction based on the filtered recent data.

(c) **Less conservative upper bounds for prediction errors that hold for a high probability.** Identifying and excluding rare noise realizations responsible for worst-case prediction errors could potentially shrink the uncertainty set underpinning robust control tasks. If it is too hard to locate these worst-case scenarios, one can consider using empirical methods (e.g., bootstrapping) and constructing a metric assessing how trustworthy the empirical results are.

In Part II, we designed zeroth-order methods for safely optimizing problems with unknwon objective and constriaint functoins. Our methods stand apart from traditional numerical optimization due to the guaranteed feasibility throughout iterations. We achieved this feature by constructing local feasible sets based on function smoothness. The main challenge lies in how to enhance computation and sampling efficiency. In Section 4, we propose SZO-LP such that the linear programming subproblems can be efficiently solved and the iterates tend to stay away from the feasible region boundary. Although the performance is improved compared with SZO-QQ in Section 3 when applied to solve the OPF problem, there are several open problems worth further exploration:

(a) **Non-asymptotic convergence properties.** In Section 4, we only provide asymptotic results. The key ingredients lacking are how the termination condition in Algorithm 3 is related to of the solution being an approximate KKT pair and how many iterations needed before termination. The non-asymptotic results in Section 3 rely on the regulation term. Moreover, the convergence rates of SZO-QQ and SZO-LP are still lacking. Convergence rate is a standard property in first-order methods [147, Section 2.2]. To prove this property, one may have to assume local strong convexity for the objective function.

(b) **Harnessing accelerated gradient descent** [148] **and quasi-Newton** [149, Chapter 6] **methods.** Both methods give faster convergence rate than standard gradient-based methods while the second is advantageous especially when the objective function has a Hessian matrix with a large condition number. In the literature, there are several works on these two methods for constrained optimization [150, 151]. It is still open regarding how they can be used for black-box optimization while ensuring sample feasibility.

(c) **Sample complexity analysis in the presence of measurement noise.** It is beneficial to learn function gradients based on the collected data. However, regarding how many samples are needed to achieve a certain level of accuracy, many open questions might arise, e.g., what is a good sampling strategy for attenuation of noise, which data subset is the most beneficial for function proxy construction and what reasonable assumptions are needed to ensure error bounds.

(d) **Incorporating local Lipschitz/smoothness constants or other prior knowledge.** In SZO-QQ and SZO-LP, the regularity properties on the objective/constraint functions we use are the global Lipschitz/smoothness constants (see Assumption 3.1). Since both SZO-QQ and SZO-LP are based on local models, certainly, using local information/prior knowledge will reduce the conservativeness and thus accelerate the convergence.

(e) **Online implementation.** In reality, most safety-critical systems are subject to time-varying disturbances. Therefore, online/real-time implementation is essential to practical applications of optimization algorithms [152, 153]. For example, in real-life OPF problems, the loads vary following daily and seasonal patterns. In this case, it is challenging to maintain sample feasibility since the learned optimum at one time instant might turn infeasible in the next. Therefore, constraint tightening is needed against any unknown disturbance. However, doing so can compromise the optimality. Moreover, the transient of the iterates between different time instants might exhibit too much overshoot.

Besides the abovementioned theoretical problems, there might be issues related to the practical applications of our zeroth-order algorithms. For example, given a specific problem, the feature of the involved system might raise new requirements for the queries of the unknown objective/constraint functions. In OPF problems, the varying power generation must satisfy certain ramping constraints [154] and conform to the grid harmonics standards [155, Chapter 3.1]. Therefore, the sampling strategies in SZO-QQ and SZO-LP might not be feasible for real-life power system applications.

Index

Bibliography

[1] L. Ljung, “System identification,” *Wiley encyclopedia of electrical and electronics engineering*, pp. 1–19, 1999.

[2] K. Zhou and J. C. Doyle, *Essentials of robust control*, vol. 104. Prentice hall Upper Saddle River, NJ, 1998.

[3] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, “On the sample complexity of the Linear Quadratic Regulator,” *Foundations of Computational Mathematics*, pp. 1–47, 2019.

[4] P. Rao, M. Crow, and Z. Yang, “Statcom control for power system voltage control applications,” *IEEE Transactions on power delivery*, vol. 15, no. 4, pp. 1311–1317, 2000.

[5] T. Yuan and R. Zhao, “Lqr-mpc-based trajectory-tracking controller of autonomous vehicle subject to coupling effects and driving state uncertainties,” *Sensors*, vol. 22, no. 15, p. 5556, 2022.

[6] S. Dean, S. Tu, N. Matni, and B. Recht, “Safely learning to control the constrained Linear Quadratic Regulator,” in *2019 American Control Conference (ACC)*, pp. 5582–5588, IEEE, 2019.

[7] S. Oymak and N. Ozay, “Non-asymptotic identification of LTI systems from a single trajectory,” in *2019 American Control Conference (ACC)*, pp. 5655–5661, IEEE, 2019.

[8] T. de Jong, V. Breschi, M. Schoukens, and S. Formentin, “Data-driven model-reference control with closed-loop stability: the output-feedback case,” *IEEE Control Systems Letters*, 2023.

[9] Y. Lian, J. Shi, M. Koch, and C. N. Jones, “Adaptive robust data-driven building control via bilevel reformulation: An experimental result,” *IEEE Transactions on Control Systems Technology*, 2023.

[10] L. Furieri, Y. Zheng, A. Papachristodoulou, and M. Kamgarpour, “An Input-Output Parametrization of stabilizing controllers: amidst Youla and System Level Synthesis,” *IEEE Control Systems Letters*, vol. 3, no. 4, pp. 1014–1019, 2019.

[11] M. Yin, A. Iannelli, and R. S. Smith, "Maximum likelihood estimation in data-driven modeling and control," *IEEE Transactions on Automatic Control*, 2021.

[12] L. Furieri, B. Guo, A. Martin, and G. Ferrari-Trecate, "A behavioral input-output parametrization of control policies with suboptimality guarantees," in *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 2539–2544, 2021.

[13] L. Furieri, B. Guo, A. Martin, and G. Ferrari-Trecate, "Near-optimal design of safe output-feedback controllers from noisy data," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 2699–2714, 2022.

[14] L. Xu, M. S. Turan, B. Guo, and G. Ferrari-Trecate, "A data-driven convex programming approach to worst-case robust tracking controller design," *arXiv preprint arXiv:2102.11918*, 2021.

[15] Y. Lian, J. Shi, M. Koch, and C. N. Jones, "Adaptive robust data-driven building control via bilevel reformulation: An experimental result," *IEEE Transactions on Control Systems Technology*, 2023.

[16] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1702–1717, 2021.

[17] A. Alanwar, Y. Stürz, and K. H. Johansson, "Robust data-driven predictive control using reachability analysis," *European Journal of Control*, vol. 68, p. 100666, 2022.

[18] V. Breschi, A. Chiuso, and S. Formentin, "Data-driven predictive control in a stochastic setting: A unified framework," *Automatica*, vol. 152, p. 110961, 2023.

[19] L. Huang, J. Zhen, J. Lygeros, and F. Dörfler, "Robust data-enabled predictive control: Tractable formulations and performance guarantees," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 3163–3170, 2023.

[20] B. Guo, Y. Jiang, C. N. Jones, and G. Ferrari-Trecate, "Data-driven robust control using prediction error bounds based on perturbation analysis," *arXiv preprint arXiv:2308.14178*, 2023.

[21] R. M. Lewis and V. Torczon, "Pattern search methods for linearly constrained minimization," *SIAM Journal on Optimization*, vol. 10, no. 3, pp. 917–941, 2000.

[22] S. Lucidi, M. Sciandrone, and P. Tseng, "Objective-derivative-free methods for constrained optimization," *Mathematical Programming*, vol. 92, pp. 37–59, 2002.

[23] Y. Shu, Z. Dai, W. Sng, A. Verma, P. Jaillet, and B. K. H. Low, "Zeroth-order optimization with trajectory-informed derivative estimation," in *The Eleventh International Conference on Learning Representations*, 2022.

[24] I. Usmanova, Y. As, M. Kamgarpour, and A. Krause, “Log barriers for safe black-box optimization with application to safe reinforcement learning,” *arXiv preprint arXiv:2207.10415*, 2022.

[25] Y. Sui, A. Gotovos, J. Burdick, and A. Krause, “Safe exploration for optimization with gaussian processes,” in *International Conference on Machine Learning*, pp. 997–1005, PMLR, 2015.

[26] B. Guo, Y. Jiang, M. Kamgarpour, and G. Ferrari-Trecate, “Safe zeroth-order convex optimization using quadratic local approximations,” in *2023 European Control Conference (ECC)*, pp. 1–8, IEEE, 2023.

[27] B. Guo, Y. Jiang, G. Ferrari-Trecate, and M. Kamgarpour, “Safe zeroth-order optimization using quadratic local approximations,” *arXiv preprint arXiv:2303.16659*, 2023.

[28] B. Guo, Y. Wang, Y. Jiang, M. Kamgarpour, and G. Ferrari-Trecate, “Safe zeroth-order optimization using linear programs,” in *in 62th IEEE Conference on Decision and Control (to appear)*, IEEE, 2023.

[29] B. Guo, O. Karaca, S. Azhdari, M. Kamgarpour, and G. Ferrari-Trecate, “Actuator placement for structural controllability beyond strong connectivity and towards robustness,” in *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 5294–5299, IEEE, 2021.

[30] B. Guo, O. Karaca, T. Summers, and M. Kamgarpour, “Actuator placement under structural controllability using forward and reverse greedy algorithms,” *IEEE Transactions on Automatic Control*, vol. 66, no. 12, pp. 5845–5860, 2020.

[31] B. Guo, O. Karaca, T. Summers, and M. Kamgarpour, “Actuator placement for optimizing network performance under controllability constraints,” in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 7140–7147, IEEE, 2019.

[32] L. Xu, B. Guo, and G. Ferrari-Trecate, “Finite-sample-based spectral radius estimation and stabilizability test for networked control systems,” in *2022 European Control Conference (ECC)*, pp. 2087–2092, IEEE, 2022.

[33] L. Xu, M. S. Turan, B. Guo, and G. Ferrari-Trecate, “Non-conservative design of robust tracking controllers based on input-output data,” in *Learning for Dynamics and Control*, pp. 138–149, PMLR, 2021.

[34] L. Xu, B. Guo, C. Galimberti, M. Farina, R. Carli, and G. F. Trecate, “Suboptimal distributed lqr design for physically coupled systems,” *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 11032–11037, 2020.

[35] 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.

[36] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, "Global convergence of policy gradient methods for the linear quadratic regulator," in *International Conference on Machine Learning*, pp. 1467–1476, PMLR, 2018.

[37] Y. Zheng, L. Furieri, M. Kamgarpour, and N. Li, "Sample complexity of linear quadratic gaussian (lqg) control for output feedback systems," in *Learning for Dynamics and Control*, pp. 559–570, PMLR, 2021.

[38] M. Simchowitz, K. Singh, and E. Hazan, "Improper learning for non-stochastic control," in *Conference on Learning Theory*, pp. 3320–3436, PMLR, 2020.

[39] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, "Logarithmic regret bound in partially observable linear dynamical systems," *arXiv preprint arXiv:2003.11227*, 2020.

[40] K. Zhang, B. Hu, and T. Basar, "Policy optimization for $\mathcal{H}_2$ linear control with $\mathcal{H}_\infty$ robustness guarantee: Implicit regularization and global convergence," in *Learning for Dynamics and Control*, pp. 179–190, PMLR, 2020.

[41] A. Tsiamis, N. Matni, and G. Pappas, "Sample complexity of kalman filtering for unknown systems," in *Learning for Dynamics and Control*, pp. 435–444, PMLR, 2020.

[42] S. Fattahi, N. Matni, and S. Sojoudi, "Efficient learning of distributed linear-quadratic control policies," *SIAM Journal on Control and Optimization*, vol. 58, no. 5, pp. 2927–2951, 2020.

[43] L. Furieri, Y. Zheng, and M. Kamgarpour, "Learning the globally optimal distributed LQ regulator," in *Learning for Dynamics and Control*, pp. 287–297, PMLR, 2020.

[44] G. Baggio, D. S. Bassett, and F. Pasqualetti, "Data-driven control of complex networks," *Nature communications*, vol. 12, no. 1, pp. 1–13, 2021.

[45] J. C. Willems and J. W. Polderman, *Introduction to mathematical systems theory: a behavioral approach*, vol. 26. Springer Science & Business Media, 1997.

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

[47] J. Coulson, J. Lygeros, and F. Dörfler, "Distributionally robust chance constrained data-enabled predictive control," *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3289–3304, 2021.

[48] F. Dörfler, J. Coulson, and I. Markovsky, "Bridging direct & indirect data-driven control formulations via regularizations and relaxations," *arXiv preprint arXiv:2101.01273*, 2021.

[49] C. De Persis and P. Tesi, “Formulas for data-driven control: Stabilization, optimality, and robustness,” *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 909–924, 2020.

[50] A. Russo and A. Proutiere, “Poisoning attacks against data-driven control methods,” *arXiv preprint arXiv:2103.06199*, 2021.

[51] D. Alpago, F. Dörfler, and J. Lygeros, “An extended Kalman filter for data-enabled predictive control,” *IEEE Control Systems Letters*, vol. 4, no. 4, pp. 994–999, 2020.

[52] H. J. Vanwaarde, M. K. Camlibel, and M. Mesbahi, “From noisy data to feedback controllers: non-conservative design via a matrix S-lemma,” *IEEE Transactions on Automatic Control, to appear*, 2020.

[53] A. Sassella, V. Breschi, and S. Formentin, “Data-driven design of explicit predictive controllers,” in *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 2821–2826, IEEE, 2022.

[54] A. Xue and N. Matni, “Data-driven system level synthesis,” in *Learning for Dynamics and Control*, pp. 189–200, PMLR, 2021.

[55] C. De Persis and P. Tesi, “Low-complexity learning of linear quadratic regulators from noisy data,” *Automatica*, vol. 128, p. 109548, 2021.

[56] F. Dörfler, P. Tesi, and C. De Persis, “On the role of regularization in direct data-driven lqr control,” in *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 1091–1098, IEEE, 2022.

[57] F. Zhao, F. Dörfler, and K. You, “Data-enabled policy optimization for the linear quadratic regulator,” *arXiv preprint arXiv:2303.17958*, 2023.

[58] V. G. Lopez, M. Alsalti, and M. A. Müller, “Efficient off-policy q-learning for data-based discrete-time lqr problems,” *IEEE Transactions on Automatic Control*, 2023.

[59] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: Theory—Implementation—Applications.* Springer Science & Business Media, 2012.

[60] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski, “Optimization over state feedback policies for robust control with constraints,” *Automatica*, vol. 42, no. 4, pp. 523–533, 2006.

[61] J. Sieber, S. Bennani, and M. N. Zeilinger, “A system level approach to tube-based model predictive control,” *IEEE Control Systems Letters*, 2021.

[62] P. J. Goulart and E. C. Kerrigan, “Output feedback receding horizon control of constrained systems,” *International Journal of Control*, vol. 80, no. 1, pp. 8–20, 2007.

[63] Y. Zheng, L. Furieri, A. Papachristodoulou, N. Li, and M. Kamgarpour, "On the equivalence of Youla, System-level and Input-output parameterizations," *IEEE Transactions on Automatic Control*, pp. 1–8, 2020.

[64] A. Bemporad, "Reducing conservativeness in predictive control of constrained systems with disturbances," in *Proceedings of the 37th IEEE Conference on Decision and Control (Cat. No. 98CH36171)*, vol. 2, pp. 1384–1389, IEEE, 1998.

[65] Y. Zheng, L. Furieri, M. Kamgarpour, and N. Li, "System-level, input–output and new parameterizations of stabilizing controllers, and their numerical computation," *Automatica*, vol. 140, p. 110211, 2022.

[66] L. Furieri and M. Kamgarpour, "Unified approach to convex robust distributed control given arbitrary information structures," *IEEE Transactions on Automatic Control*, 2019.

[67] H. Mania, S. Tu, and B. Recht, "Certainty equivalence is efficient for linear quadratic control," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[68] Y. Zheng and N. Li, "Non-asymptotic identification of linear dynamical systems using multiple trajectories," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1693–1698, 2020.

[69] M. Yin, A. Iannelli, and R. S. Smith, "Data-driven prediction with stochastic data: Confidence regions and minimum mean-squared error estimates," *arXiv preprint arXiv:2111.04789*, 2021.

[70] M. Tanaskovic, L. Fagiano, R. Smith, and M. Morari, "Adaptive receding horizon control for constrained mimo systems," *Automatica*, vol. 50, no. 12, pp. 3019–3029, 2014.

[71] E. Terzi, L. Fagiano, M. Farina, and R. Scattolini, "Learning-based predictive control for linear systems: A unitary approach," *Automatica*, vol. 108, p. 108473, 2019.

[72] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization.," *Journal of machine learning research*, vol. 13, no. 2, 2012.

[73] S. Chen, H. Wang, M. Morari, V. M. Preciado, and N. Matni, "Robust closed-loop model predictive control via system level synthesis," in *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 2152–2159, IEEE, 2020.

[74] A. Agrawal and S. Boyd, "Disciplined quasiconvex programming," *Optimization Letters*, pp. 1–15, 2020.

[75] J. Kiefer, "Sequential minimax search for a maximum," *Proceedings of the American mathematical society*, vol. 4, no. 3, pp. 502–506, 1953.

[76] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

[77] N. Matni, Y.-S. Wang, and J. Anderson, "Scalable system level synthesis for virtually localizable systems," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 3473–3480, IEEE, 2017.

[78] MOSEK Aps, "The MOSEK optimization toolbox for MATLAB manual. Version 8.1.," 2017.

[79] J. Löfberg, "YALMIP : A Toolbox for Modeling and Optimization in MATLAB," in *In Proc. of the CACSD Conf.*, (Taipei, Taiwan), 2004.

[80] A. Iannelli, M. Yin, and R. S. Smith, "Experiment design for impulse response identification with signal matrix models," *IFAC-PapersOnLine*, vol. 54, no. 7, pp. 625–630, 2021.

[81] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.

[82] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.

[83] P. Tokekar and V. Isler, "Sensor placement and selection for bearing sensors with bounded uncertainty," in *2013 IEEE international conference on robotics and automation*, pp. 2515–2520, IEEE, 2013.

[84] J. Coulson, H. van Waarde, and F. Dörfler, "Robust fundamental lemma for data-driven control," *International Symposium on Mathematical Theory of Networks and Systems*, 2022.

[85] J. Coulson, H. J. Van Waarde, J. Lygeros, and F. Dörfler, "A quantitative notion of persistency of excitation and the robust fundamental lemma," *IEEE Control Systems Letters*, vol. 7, pp. 1243–1248, 2022.

[86] J. Coulson, J. Lygeros, and F. Dörfler, "Distributionally robust chance constrained data-enabled predictive control," *arXiv preprint arXiv:2006.01702*, 2020.

[87] L. Huang, J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control for grid-connected power converters," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 8130–8135, IEEE, 2019.

[88] R.-C. Li, "Relative perturbation theory: I. eigenvalue and singular value variations," *SIAM Journal on Matrix Analysis and Applications*, vol. 19, no. 4, pp. 956–982, 1998.

[89] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Foundations of Computational Mathematics*, vol. 20, no. 4, pp. 633–679, 2020.

[90] S. Lu, I. Tsaknakis, and M. Hong, "Block alternating optimization for non-convex min-max problems: algorithms and applications in signal processing and communications," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4754–4758, IEEE, 2019.

[91] J. Guthrie, M. Kobilarov, and E. Mallada, "Closed-form minkowski sum approximations for efficient optimization-based collision avoidance," in *2022 American Control Conference (ACC)*, pp. 3857–3864, IEEE, 2022.

[92] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.

[93] F. Oldewurtel, C. N. Jones, and M. Morari, "A tractable approximation of chance constrained stochastic mpc based on affine disturbance feedback," in *2008 47th IEEE conference on decision and control*, pp. 4731–4736, IEEE, 2008.

[94] G. W. Stewart, "On the perturbation of pseudo-inverses, projections and linear least squares problems," *SIAM review*, vol. 19, no. 4, pp. 634–662, 1977.

[95] J. Nocedal and S. J. Wright, *Numerical optimization.* Spinger, 2006.

[96] S. Ghadimi and G. Lan, "Stochastic first-and zeroth-order methods for nonconvex stochastic programming," *SIAM Journal on Optimization*, vol. 23, no. 4, pp. 2341–2368, 2013.

[97] A. R. Conn, N. I. Gould, and P. L. Toint, *Trust region methods.* SIAM, 2000.

[98] I. Usmanova, A. Krause, and M. Kamgarpour, "Safe convex learning under uncertain constraints," in *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 2106–2114, PMLR, 2019.

[99] N. Echebest, M. L. Schuverdt, and R. P. Vignau, "An inexact restoration derivative-free filter method for nonlinear programming," *Computational and Applied Mathematics*, vol. 36, no. 1, pp. 693–718, 2017.

[100] I. Bajaj, S. S. Iyer, and M. F. Hasan, "A trust region-based two phase algorithm for constrained black-box and grey-box optimization with infeasible initial point," *Computers & Chemical Engineering*, vol. 116, pp. 306–321, 2018.

[101] M. Turchetta, F. Berkenkamp, and A. Krause, "Safe exploration for interactive machine learning," *Advances in Neural Information Processing Systems*, vol. 32, p. 2887–2897, 2019.

[102] L. Sabug Jr, F. Ruiz, and L. Fagiano, "Smgo-δ: Balancing caution and reward in global optimization with black-box constraints," *Information Sciences*, vol. 605, pp. 15–42, 2022.

[103] A. P. Vinod, A. Israel, and U. Topcu, "Constrained, global optimization of unknown functions with lipschitz continuous gradients," *SIAM Journal on Optimization*, vol. 32, no. 2, pp. 1239–1264, 2022.

[104] R. M. Lewis and V. Torczon, "A globally convergent augmented lagrangian pattern search algorithm for optimization with general constraints and simple bounds," *SIAM Journal on Optimization*, vol. 12, no. 4, pp. 1075–1089, 2002.

[105] C. Audet and J. E. Dennis Jr, "A progressive barrier for derivative-free nonlinear programming," *SIAM Journal on optimization*, vol. 20, no. 1, pp. 445–472, 2009.

[106] D. B. Arnold, M. Negrete-Pincetic, M. D. Sankur, D. M. Auslander, and D. S. Callaway, "Model-free optimal control of var resources in distribution systems: An extremum seeking approach," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3583–3593, 2015.

[107] J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin, "A general safety framework for learning-based control in uncertain robotic systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 2737–2752, 2018.

[108] J. Achiam, D. Held, A. Tamar, and P. Abbeel, "Constrained policy optimization," in *International conference on machine learning*, pp. 22–31, PMLR, 2017.

[109] Z. Qin, D. Sun, and C. Fan, "Sablas: Learning safe control for black-box dynamical systems," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1928–1935, 2022.

[110] W. Xu, C. N. Jones, B. Svetozarevic, C. R. Laughman, and A. Chakrabarty, "Vabo: Violation-aware bayesian optimization for closed-loop control performance optimization with unmodeled constraints," in *2022 American Control Conference (ACC)*, pp. 5288–5293, IEEE, 2022.

[111] F. Berkenkamp, A. P. Schoellig, and A. Krause, "Safe controller optimization for quadrotors with gaussian processes," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 491–496, IEEE, 2016.

[112] C. König, M. Turchetta, J. Lygeros, A. Rupenyan, and A. Krause, "Safe and efficient model-free adaptive control via bayesian optimization," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 9782–9788, IEEE, 2021.

[113] H. Abdi, S. D. Beigvand, and M. La Scala, "A review of optimal power flow studies applied to smart grids and microgrids," *Renewable and Sustainable Energy Reviews*, vol. 71, pp. 742–766, 2017.

[114] P. Jain, P. Kar, *et al.*, "Non-convex optimization for machine learning," *Foundations and Trends® in Machine Learning*, vol. 10, no. 3-4, pp. 142–363, 2017.

[115] J. B. Rawlings, D. Q. Mayne, and M. Diehl, *Model predictive control: theory, computation, and design*, vol. 2. Nob Hill Publishing Madison, WI, 2017.

[116] P. S. Kundur and O. P. Malik, *Power system stability and control.* McGraw-Hill Education, 2022.

[117] J. Dutta, K. Deb, R. Tulshyan, and R. Arora, "Approximate kkt points and a proximity measure for termination," *Journal of Global Optimization*, vol. 56, no. 4, pp. 1463–1499, 2013.

[118] G. N. Grapiglia and Y. Nesterov, "Tensor methods for finding approximate stationary points of convex functions," *Optimization Methods and Software*, vol. 37, no. 2, pp. 605–638, 2022.

[119] J.-P. Dussault, M. Haddou, A. Kadrani, and T. Migot, "On approximate stationary points of the regularized mathematical program with complementarity constraints," *Journal of Optimization Theory and Applications*, vol. 186, no. 2, pp. 504–522, 2020.

[120] A. Wibisono, M. J. Wainwright, M. Jordan, and J. C. Duchi, "Finite sample convergence rates of zero-order stochastic optimization methods," *Advances in Neural Information Processing Systems*, vol. 25, 2012.

[121] A. R. Conn, K. Scheinberg, and L. N. Vicente, *Introduction to derivative-free optimization.* SIAM, 2009.

[122] A. Rupenyan, M. Khosravi, and J. Lygeros, "Performance-based trajectory optimization for path following control using bayesian optimization," in *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 2116–2121, IEEE, 2021.

[123] A. S. Berahas, L. Cao, K. Choromanski, and K. Scheinberg, "A theoretical and empirical comparison of gradient approximations in derivative-free optimization," *Foundations of Computational Mathematics*, vol. 22, no. 2, pp. 507–560, 2022.

[124] A. Y. Aravkin, R. Baraldi, and D. Orban, "A proximal quasi-newton trust-region method for nonsmooth regularized optimization," *SIAM Journal on Optimization*, vol. 32, no. 2, pp. 900–929, 2022.

[125] M. Fukushima, Z.-Q. Luo, and P. Tseng, "A sequential quadratically constrained quadratic programming method for differentiable convex minimization," *SIAM Journal on Optimization*, vol. 13, no. 4, pp. 1098–1119, 2003.

[126] F. Messerer and M. Diehl, "Determining the exact local convergence rate of sequential convex programming," in *2020 European Control Conference (ECC)*, pp. 1280–1285, IEEE, 2020.

[127] F. Messerer, K. Baumgärtner, and M. Diehl, "Survey of sequential convex programming and generalized gauss-newton methods," *ESAIM. Proceedings and Surveys*, vol. 71, pp. 64–88, 2021.

[128] G. Frison, J. Frey, F. Messerer, A. Zanelli, and M. Diehl, "Introducing the quadratically-constrained quadratic programming framework in hpipm," in *2022 European Control Conference (ECC)*, pp. 447–453, IEEE, 2022.

[129] G. Wachsmuth, "On LICQ and the uniqueness of lagrange multipliers," *Operations Research Letters*, vol. 41, no. 1, pp. 78–80, 2013.

[130] I. Usmanova, M. Kamgarpour, A. Krause, and K. Levy, "Fast projection onto convex smooth constraints," in *International Conference on Machine Learning*, pp. 10476–10486, PMLR, 2021.

[131] M. Zhu and E. Frazzoli, "Distributed robust adaptive equilibrium computation for generalized convex games," *Automatica*, vol. 63, pp. 82–91, 2016.

[132] M. Diehl, H. G. Bock, and J. P. Schlöder, "A real-time iteration scheme for nonlinear optimization in optimal feedback control," *SIAM Journal on control and optimization*, vol. 43, no. 5, pp. 1714–1736, 2005.

[133] A. F. Izmailov and M. V. Solodov, "Newton-type methods for optimization problems without constraint qualifications," *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 210–228, 2004.

[134] D. D. Fan, A.-a. Agha-mohammadi, and E. A. Theodorou, "Deep learning tubes for tube mpc," in *Proceedings of Robotics: Science and Systems XVI*, 2020.

[135] L. Hewing, J. Kabzan, and M. N. Zeilinger, "Cautious model predictive control using gaussian process regression," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 6, pp. 2736–2743, 2019.

[136] R. Christie, "Power systems test case archive," *U of Washington*, 2017.

[137] J. Das, *Load flow optimization and optimal power flow.* Crc Press, 2017.

[138] Z. Chu, S. Lakshminarayana, B. Chaudhuri, and F. Teng, "Mitigating load-altering attacks against power grids using cyber-resilient economic dispatch," *IEEE Transactions on Smart Grid*, 2022. Early access.

[139] D. Lee, K. Turitsyn, D. K. Molzahn, and L. A. Roald, "Robust AC optimal power flow with robust convex restriction," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 4953–4966, 2021.

[140] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2010.

[141] I. Gurobi Optimization, "Gurobi optimizer reference manual," 2016.

[142] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in matlab," in *In Proceedings of the CACSD Conference*, (Taipei, Taiwan), 2004.

[143] J. F. Bonnans and A. Shapiro, *Perturbation analysis of optimization problems.* Springer Science & Business Media, 2013.

[144] F. Berkenkamp, A. Krause, and A. P. Schoellig, "Bayesian optimization with safety constraints: safe and automatic parameter tuning in robotics," *Machine Learning*, pp. 1–35, 2021.

[145] L. Ljung, "System identification," *Wiley Encyclopedia of Electrical and Electronics Engineering*, 2001.

[146] I. Markovsky, *Low-Rank Approximation: Algorithms, Implementation, Applications.* Springer, 2nd edition ed., 2019.

[147] Y. Nesterov *et al.*, *Lectures on convex optimization*, vol. 137. Springer, 2018.

[148] Y. E. Nesterov, "A method of solving a convex programming problem with convergence rate o\bigl(kˆ2\bigr)," in *Doklady Akademii Nauk*, vol. 269, pp. 543–547, Russian Academy of Sciences, 1983.

[149] J. Nocedal and S. Wright, *Numerical optimization.* Springer Science & Business Media, 2006.

[150] L. Chen, K.-U. Bletzinger, N. R. Gauger, and Y. Ye, "A gradient descent akin method for constrained optimization: algorithms and applications," *arXiv preprint arXiv:2302.11898*, 2023.

[151] P. T. Boggs, J. W. Tolle, and P. Wang, "On the local convergence of quasi-newton methods for constrained optimization," *SIAM journal on control and optimization*, vol. 20, no. 2, pp. 161–171, 1982.

[152] D. Krishnamoorthy, B. Foss, and S. Skogestad, "Steady-state real-time optimization using transient measurements," *Computers & Chemical Engineering*, vol. 115, pp. 34–45, 2018.

[153] V. Häberle, A. Hauswirth, L. Ortmann, S. Bolognani, and F. Dörfler, "Non-convex feedback optimization with input and output constraints," *IEEE Control Systems Letters*, vol. 5, no. 1, pp. 343–348, 2020.

[154] X. Guan, A. Svoboda, and C.-a. Li, "Scheduling hydro power systems with restricted operating zones and discharge ramping constraints," *IEEE transactions on Power Systems*, vol. 14, no. 1, pp. 126–131, 1999.

[155] T. Geyer, *Model predictive control of high power converters and industrial drives.* John Wiley & Sons, 2016.