Modeling, Robustness and Stability for Sparse Optimal Control of Dynamical Systems

Submitted to
Graduate School of Information Science and Technology
Osaka University

January 2024

Zhicheng ZHANG

Abstract of thesis entitled

Modeling, Robustness and Stability for Sparse Optimal Control of Dynamical Systems

Submitted by

Zhicheng ZHANG

for the degree of Doctor of Philosophy in Informatics

Osaka University

Sparse optimization with uncertainty has been widely applied in operations research and industrial engineering. For instance, in control systems, sparsity effectively minimizes the control efforts by deactivating control actuators through holding a long-period of zero-valued control inputs – a paradigm known as sparse optimal control. In practical applications, dynamical systems often encounter diverse uncertainties, such as noise, disturbances, parameter mismatches, or missing data, which significantly affect their stability and reliability. These uncertainties pose challenges in modeling systems, where solutions become highly sensitive to uncertain variables. Therefore, pursuing a robust, stable, and optimal solution for uncertain models is a critical issue for optimization theory and its control applications.

This dissertation aims to address these challenges. Firstly, by implementing the dynamic linear compensator for system modeling, an explicit sparse feedback controller can be inferred from its open-loop optimal solution to closed-loop realization, which also provides the initialization robustness guarantees for control systems. Secondly, a chance-constrained sparse optimization problem is proposed by modeling the stochastic dynamics, where the uncertain parameters are assumed to be random variables. By means of convex relaxation and data-driven sampling technique, the sparse random convex program and risk-aware sparse optimal (predictive) control problems are presented. This framework not only delivers a randomized sparse solution but also ensures robustness with a high level of confidence in probabilistic guarantees. Thirdly, a data-driven framework for a discrete linear time invariant system is employed in conjunction with sparse feedback control synthesis. Instead of a priori knowledge of true system model, the black-box control systems can be purely exploiting experimental input/state/output data samples. Fourthly, a methodology for linear quadratic sparse optimal control is devised to tackle a continuous-time master-slave tracking issue, employing a framework grounded in the non-smooth maximum principle. Finally, numerical benchmarks illustrates the effectiveness of the proposed theoretical results and its control technologies.

Modeling, Robustness and Stability for Sparse Optimal Control of Dynamical Systems

by

Zhicheng ZHANG

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy
in Informatics

Osaka University

January 2024

To my lovely family for their remarkable endurance
to grandmother Xingdi, parents Heng & Huizhen
and to the memory of grandfather Caibao
forever no matter what
Z.Z.

Acknowledgements

At this juncture, as I compose my PhD dissertation, I reflect on the journey that led me here. Undertaking a Ph.D. stands as one of the most rewarding decisions I have made. Despite its myriad challenges, this unique voyage has been an invaluable source of learning and fulfillment. Thus, I extend my heartfelt gratitude to all those who supported me throughout this pivotal chapter of my life. Without your invaluable assistance, I wouldn't have reached this significant milestone.

Foremost, I wish to express my deepest appreciation to my supervisor, Prof. Yasumasa Fujisaki, who served not only as my mentor but also as my daily guide. Yasumasa, your decision to oversee my Ph.D. research is a privilege I deeply cherish. Your patience, unwavering enthusiasm, incisive critique, and vast reservoir of knowledge have consistently steered me towards the culmination of this dissertation. Your meticulous feedback and our stimulating discussions have been enlightening. Moreover, I am grateful for the freedom you accorded me to explore my research interests.

I am grateful to extend my heartfelt appreciation to the rest members of my Ph.D. dissertation committee: Prof. Hiroshi Morita, Prof. Hideyuki Suzuki, and Associate Prof. Takayuki Wada. Their extremely kind supports throughout my general exam, Ph.D. proposal defense, dissertation defense, and subsequent review has been immensely valuable. I am sincerely thankful for the abundance of valuable feedback and constructive comments they have provided.

I would like to express my deepest gratitude to my co-supervisor Associate Prof. Takayuki Wada for his elaborated guidance, considerable encouragement, and invaluable discussion. His insightful comments and constructive suggestions improved my research study. I owe a huge debt of gratitude to Assistant Prof. Tomoyuki Iori for his continuous support and encouragement. My sincere thanks also go to Prof. Masaaki Nagahara at Hiroshima University.

I would like to acknowledge all past and present members of the Operations Research Laboratory for their invaluable contributions, insightful discussions, and enduring friendship. Finally, I am grateful to my family for their profound understanding and unwavering encouragement. Their steadfast support has been a constant source of motivation and strength, for which I am truly grateful.

Osaka, January 9, 2024
Zhicheng ZHANG

List of Publications

JOURNALS:

[1] Z. Zhang and Y. Fujisaki, " Sparse feedback controller: From open-loop solution to closed-loop realization," *SICE Journal of Control, Measurement, and System Integration*, 2023, vol. 16, No. 1, pp. 286–296.
Doi: 10.1080/18824889.2023.2237234

CONFERENCES:

[1] Z. Zhang and M. Nagahara, "Linear quadratic tracking control with sparsity-promoting regularization," *2021 American Control Conference*, IEEE, New Orleans, May. 2021, pp. 3812–3817.

[2] Z. Zhang and Y. Fujisaki, "Sparse robust control design via scenario program," *Proceeding of the 53rd ISCIE International Symposium on Stochastic Systems Theory and Its Applications*, Kusatsu, Oct. 2021, pp. 61-64.

[3] Z. Zhang and Y. Fujisaki, "Sparse feedback control realization using a dynamic linear compensator," *SICE International Symposium on Control Systems*, Kusatsu, Mar. 2023, 3M1-4.

[4] Z. Zhang and Y. Fujisaki, "Risk-aware sparse predictive control," *Preprints of the 22nd IFAC World Congress*, Yokohama, Jul., 2023, pp. 1477-1480.

[5] Z. Zhang and Y. Fujisaki, "Risk assessment for sparse optimization with relaxation," *The 55th ISCIE International Symposium on Stochastic Systems Theory and Its Applications*, Tokyo, Nov., 2023, pp. 61-62.

[6] Z. Zhang and Y. Fujisaki, "Data-driven sparse feedback with Schur-α stability," *SICE International Symposium on Control Systems*, Hiroshima, Mar., 2024 (accepted and to be presented).

Contents

List of Figures

List of Symbols

$\mathbb{R}^n$	space of real n-dimensional vectors
$\mathbb{R}^{n\times m}$	space of real n-by-m real matrices
e_i	a standard basis with appropriate dimension
$\mathbf{1}_N$	a vector for all one values
$\otimes$	kronecker product
I_n	n-by-n identity matrix
$0_{n\times m}$	n-by-m zero matrix
$X^\top$	transpose of matrix X
X^{-1}	inverse of (nonsingular) matrix X
$X^\dagger$	Moore-Penrose pseudo-inverse of matrix X
$\mathrm{rank}(X)$	rank of matrix X
$\det(X)$	determinant of matrix X
$\mathrm{Im}(X)$	image of matrix X
abs(X)	absolute value of all entries for matrix X
$\mathrm{vec}(X)$	column vectorization of matrix X
$\lambda_i(X)$	i-th eigenvalue of matrix A
$X \succ 0$	positive definite symmetric matrix
$X \succeq 0$	positive semidefinite symmetric matrix
$X \prec 0$	negative definite symmetric matrix
$X \preceq 0$	negative semidefinite symmetric matrix
$\mathbb{1}_S(\cdot)$	indicator function of the set S
$\mathrm{card}(S)$	cardinality of the set S
$\mathrm{supp}(x)$	support of a vector x
$\mu_m(\mathcal{S})$	Lebesgue measure of the set S
$\\|x\\|_0$	ℓ_0 pseudo-norm of a vector x
	(resp., $\mathcal{L}_0$ norm of a continuous-time signal)
$\|x\|$	absolute value of a vector x

Chapter 1

Introduction

1.1 Background

A philosophical principle of parsimony, is commonly referred to as *Ockham's razor*. This principle advocating the preference for the simplest explanation of the problem's structure should take over more complicated ones, which means that the solution is *sparse*. Put differently, "*simplest is the best*." Within the realm of machine learning and statistics modeling [19], it manifests in the practice of variable or feature selection, principle component analysis [151], commonly applied in two situations. Firstly, it aims to render the model or prediction more interpretable or computationally efficient. For instance, even in cases where the underlying problem is not sparse, one looks for the best optimal sparse approximation, like regression problem by Lasso regularization [127]. Secondly, the sparsity is embedded when prior knowledge requires *sparse modeling*. In connection to the regressions of machine learning, the sparse regularization can efficiently enhance more zero coefficients to prevent the overfitting learning [112, 139]. In bio-informatics, sparsity-inducing plays an important role in finding meaningful fragments in gens [124]. Also, when sparsity is in charge of investment management that selects a relatively small number of invested asserts in portfolio optimization [29]. Another important field that can benefit from sparsity is signal processing and its extension *compressed sensing* [41, 55]. The fundamental concept behind compressive sensing is the recognition that most real-world signals, such as image, audio, or video data can be well approximated by sparse vectors within some suitable basis. Leveraging this sparse representation can dramatically diminish the signal acquisition cost. Oftentimes, the accurate signal can be reconstructed in a computationally feasible way, by means of *sparse optimization* methods.

As we all know, energy-saving is critical for system performance guarantees in optimal control problems, which is closely related to minimum energy control [7, Chap. 8] by performing ℓ_2 norm (resp., $\mathcal{L}_2$ norm in continuous-time) of control energy transition to prevent engine overheating and to reduce transmission cost. Another alternative strategy is minimum fuel control [7, Chap. 6], that is, the total consumption of fuel is minimized with the ℓ_1 norm (resp., $\mathcal{L}_1$ norm in continuous-time) of the control signals.

A famous real-world application showcasing fuel minimization is in spacecraft guidance, such as rocket flight and landing control from the earth to the moon. This practice aims to reduce the total expenditure of fuel as soon as possible, yielding cost savings.

Sparse optimal control is a new paradigm for control effort minimization [102, 105], and thus it can be viewed as a "green control" that achieves energy or fuel conservation as well as (CO or CO_2) emission reduction (cf., in power systems, smart grids, transportation systems, electric or hybrid vehicles, etc), which is vital to pave the road to Net-Zero objectives, utilizing automation technology as a part of the solution. In general, the sparse optimal control objective is to find a sparest control sequences that drives the system state from an initial state to a prescribed terminal state. Here, the sparsity is characterized as the ℓ_0 norm on discrete-time control vector (or $\mathcal{L}_0$ norm over continuous-time control signals) by means of minimizing the $\ell_0/\mathcal{L}_0$ norm objective. In essence, sparsity is characterized by the ℓ_0 norm, which quantifies the number of nonzero components (or, equivalently, the Lebesgue measure in $\mathbb{R}$ with the $\mathcal{L}_0$ norm). This metric serves as a guiding principle in sparse optimization techniques. Sparse optimization problems have been a subject of study since the emergence of compressed sensing [55, 130] in signal processing. Over the last decade, the concept of "sparsity" has significantly evolved and gained prominence across various cutting-edge research domains, including machine learning, statistics, data science, and control engineering [19, 105, 127, 151]. Sparse optimization has the following remarkable benefits. First, adopting a precise $\ell_0/\mathcal{L}_0$ (or relaxed $\ell_1/\mathcal{L}_1$) norm regularizes the optimization problem and thereby mitigates the optimizer's dimensions characteristic for modern programming. Second, sparse convex models are often tractable even though the exact sparse model with the non-convex nature (which is generically non-continuous) are hard. Miscellaneous methods are employed to address sparse optimization, such as convex optimization, gradient projection method, difference-of-convex functions (DC) approach [64], mixed-integer programming (MIP) [26], submodular optimization [53], and majorization-minimization (MM) [94], among others.

This thesis focuses on sparse optimization and its applications in controlling dynamical systems, encompassing aspects such as modeling, robustness, and stability.

1.2 Related literature

In control community, sparsity-promoting strategy has thrived in various directions to distributed control [59, 73, 97], tracking control [149], and predictive control [3, 52, 104, 77]. Indeed, there are two distinct scenarios for sparse control: "control structured sparsity" versus "control input sparsity", each serving different purposes. When the sparsity is imposed on the control structure (i.e., generating *structured sparsity*) whose controller depends on feedback policy [8, 85, 113, 114], then the sparse control is recast into a distributed control, which attempts to *reduce the number of communication links or network flows* [73] in cyber–physical systems, and thereby achieving resource-aware allocation and consensus or synchronization. Another appealing alternative is to penalize sparsity on control signals (i.e., implementing *input sparsity*) that *maximizes the time duration over which the control value is exactly zero*, the resulting control paradigm is referred to as "maximum hands-off control", or "$\mathcal{L}_0/\ell_0$ optimal control" [102, 104, 105]. Spuriously, for a continuous-time system, the work [105] has revealed that $\mathcal{L}_1$ optimal control with "*bang-off-bang*" property (i.e., taking $\{\pm 1, 0\}$) is equivalent to $\mathcal{L}_0$ optimal control, while such a wonderful result falters in the discrete-time case, as ensuring equivalence between ℓ_0 and ℓ_1 optimal control problems require the "*restricted isometry property*" (RIP) [39, 104].

The primary focus of this dissertation revolves around generating a sparse optimal control signal for a class of discrete-time systems. Specifically, the control objective involves exploring the minimization problem related to *the number of nonzero components* within finite-dimensional control vectors, inspired by concepts in compressed sensing [40, 55, 130]. As a matter of fact, the optimal control design of (sparse) control input is of great significance for dynamical systems and directly affects the dynamic process of the systems. Many practical systems of interest are dependent on a feedback mechanism to achieve closed-loop stability. However, closed-loop realization is more challenging to *constrained optimal control problem* because *determining the feedback gain (matrix)* is a non-trivial task [23]. Indeed, in closed-loop sparse control design, almost all existing results have been focused on discussing "structured sparsity" [59, 73, 97] by optimizing linear quadratic state feedback cost [8, 85, 113] rather than pursuing our expected sparse control inputs (i.e., ℓ_0 optimal control) [42, 104]. Furthermore, most successful stories on sparse control adopting convex relaxed ℓ_1 cost of discrete (resp., $\mathcal{L}_1$ cost of continuous)-time systems have been extensively treated in *open-loop solutions* [10, 104, 105, 116]. It is noteworthy that, in the literature mentioned above, on one hand, sparse feedback control exhibits closed-loop stability and structured sparsity but lacks input sparsity. On the other hand, open-loop sparse optimal control enjoys input sparsity but lacks closed-loop stability.

Although "real-time control" bridges the gap between the open-loop and closed-loop solutions, schemes such as self-triggered sparse control [105] and sparse predictive control [104] can, and often do, emerge feedback solutions. Obviously, these iterative feedback algorithms perform *online optimization*, leading to the computationally burden, especially when the decision variable is high dimension. In [6, 134], a *system level synthesis* (SLS) scheme shifts the controller synthesis task from the design of a controller to the design of the entire closed-loop system, and highlight the benefits of this approach in terms of scalability and transparency. Neither exploring sparsity on the structure of feedback gain matrix or exploiting real-time control, we intend to immediately promotes sparsity on the control inputs with a closed-loop response. Inspired by seminal works [21, 23], a "*relatively optimal control*" technique paves the way towards the open-loop solution to the closed-loop solution by means of linear implementation. Therefore, the challenge lies in designing an optimal sparse feedback control that encompasses both input sparsity and closed-loop stability, highlighting its significance is one of the interests of this dissertation.

Similar optimal tracking control problem has been arisen in a variety of industrial applications including automatic river navigation [132], super-tankers [48], UAVs (unmanned aerial vehicle) [30], tokamaks [107], and motion systems [95]. On the one hand, optimal tracking can be viewed as *reference and command governors* [63], which are add-on control strategies for reference supervision and constraint enforcement in closed-loop feedback control systems. Here, the reference governor plays the role of pre-filter that, based on the current value of the desired reference command and of the (measured or estimated) state, generates a modified reference command whenever the reference command without modifications may lead to constraints violations [61]. Note that the idea of pre-stabilizing a (closed-loop) system and then using gain scheduling or cancellation and convergence of the desired reference governor or trajectory has subsequently become popular also in the servo control [83].

On the other hand, optimal tracking aims at minimizing the tracking error between the plant state (or output) and a desired reference trajectory with state and control constraints, which has been appeared in consensus or synchronization control of master-slave tracking system, multi-agent system, and large scale network systems [31]. In particular, the linear quadratic (LQ) tracker has been extensively researched, where the plant is linear and the cost function is a quadratic function [84, Chap 4]. This LQ formulation adds the $\mathcal{L}_2$ norm of the control input to the cost function as a regularization term. The LQ tracking control problem has been investigated for infinite-horizon control with a constant state target [135], receding horizon control [108], time-invariant control consisting of a static feedback and a static pre-filter [17], and discrete-time stochastic control against lossy channels [86]. In this regard, this dissertation focuses on the above mentioned two categories of tracking problems, addressing the design of sparse optimal tracking control techniques, called as minimum attention (tracking) control [28, 103, 146] and LQ sparse optimal tracking control [149], respectively.

Moreover, the presence of uncertainty in system modeling has remained a pivotal concern in control theory. For example, the uncertainty can be characterized as "additive" disturbances or "model parametric" instances of the plant [150]. The earliest attempts to address uncertainty is worst-case setting, and the uncertain instances may be extremely malicious, and thereby the idea is to hedge against the worst-case scenario, even if it may be unlikely to occur. Previous studies [76, 149] have presented sparse optimal control methods emphasizing robustness, including extensions to sparse predictive control [3, 52, 104]. However, the imposition of worst-case constraints demands satisfaction under *all* uncertainties, potentially leading to excessively conservative and pessimistic solutions. Lots of momentums for sparse optimal control problems are subject to *deterministic constraints*, such as hard state and input constraints, as far as the author knows. In this regard, this dissertation relaxes these "hard constraints" as the "soft constraints" by introducing a *probabilistic* or *stochastic* modeling. Consequently, the derived solution guarantees *probabilistic robustness* when the constraint is satisfied for the "most" of the realizations of the uncertainty. The problem is then mathematically defined as *chance-constrained optimization* problems [45] with sparse decision-making, referring to as *chance-constrained sparse programs* and giving rise to *risk-aware sparse optimal control* [77, 144], compared to the existing literature [71, 76, 105, 149]. While the solutions are often computational intractable even if the problem setups are well-defined. Fortunately, considering the fact that a convex relaxation method is commonly utilized in order to mitigate the computational burden in sparse modeling [55, 102], as well as a data-driven random sampling technique (a.k.a., scenario approach [37], sample average approximation (SAA) [91], randomized algorithms [126]) is popularly accepted in handling stochastic modeling [14, 122]. Therefore, the risk-aware sparse (predictive) optimal control problems can be addressed in (online) sparse random convex programs, and this dissertation provides a probabilistic robustness guarantees for a class of sparse optimal control problems.

In real-world control systems, the modeling of underlying linear time invariant (LTI) system is always high dimensionality and complexity, leading to priori knowledge of the "actual" system model is often *unknown* or partial available. Instead of controlling an identified system model, a data-driven paradigm has attracted much attention on "*model-free*" systems by purely relying on input-state/output data [58, 96, 137]. The essential procedure of data-driven method is how to encourage data "active", namely, learning system behavior and controller from informative data. Roughly speaking, all trajectories of system behavior can be recast through a finite samples of its *offline* input-output data stored in Hankel matrices. Under assumption of *persistently exciting* (PE) input sequence [96], a celebrated *Willems' fundamental lemma* was established [131, 137] using behavioral system theory [136]. From this, the space of all LTI system trajectories can be *directly* produced from a single experiment or simulation of input/state/output data trajectory samples, which captures the whole behavior of the controlled system, resulting in a "data-based" system representation [15, 16, 96, 131]. Moreover, additional methods such as data-driven minimum energy control [9], data-driven relative optimal control [22], data-drievn system level synthesis [140] and data-driven model predictive control [16, 50] have undergone extensive development. In [32, 113], the model-based results induce a desired sparse control input by taking a *row-sparsity* (i.e., an $\ell_{1,\infty}$ matrix norm [115, 128]) on the static state feedback gain, which implies a structured sparsity on channels. Building upon the valuable insights from prior works, this thesis aims to develop a data-driven sparse feedback controller by employing the input/state/output sampled data trajectories, all without priori knowledge of model information [142].

This dissertation focuses on the *modeling, robustness, and stability* of sparse optimal control for dynamical systems, the following concerns and comparisons are proposed:

- Open-loop versus Closed-loop models
  - Stability of sparse feedback systems (Chapter 2)
- Deterministic versus Stochastic models

    - Probabilistic robustness guarantees (Chapters 3 & 4)
- Model-based versus Model-free frameworks
    - Data-driven sparse feedback control synthesis (Chapter 5)
- Discrete-time versus Continuous-time models
    - LQ Sparse Optimal Control (Chapter 6)

Each proposed concern and comparison has its respective chapter, which involves the methodology and a comprehensive analysis of its properties, including system modeling, robustness evaluation and stability analysis. Numerical simulations are also provided for demonstrating the effectiveness of the proposed theoretical results.

1.3 Contribution overview

The main contribution of this dissertation is to demonstrate that the optimal sparse feedback controller can directly be inferred efficiently from its open-loop optimal solution to closed-loop realization. It further shows that an equivalence for the open-loop sparse control and the closed-loop sparse control under a specified basis. Furthermore, we present a sparse random convex program for dealing with a chance-constrained sparse optimization problem that attains a feasible solution with a high probability or confidence level, which also ensured that such a randomized sparse solution is computationally tractable. We also make the trade-off between the sparse cost performance and the violated constraints for the sparse random convex optimization with relaxation. As a by-product, our results are applied to robust control applications that includes risk-aware sparse optimal control and predictive control for the discrete-time uncertain dynamical systems. We also follow the trend of the hot research topics and exploit the state-of-the-art data-driven control technique to construct a direct data-driven sparse feedback control for a model-free linear time invariant system, where the sparse feedback control is purely data-enable through input/state/output data samples. The synthesized data-driven controller enjoys the input sparsity by implementing a row-sparsity structured feedback gain matrix. Without loss of generality, we finally develop a linear quadratic (LQ) sparse optimal control for a continuous-time system, and establish the theoretical guarantees for the proposed control based on *non-smooth maximum principle* [49]. In addition, the robustness of the LQ sprase optimal control in the presence of perturbations in the initial states and the state-space matrices are explored. The more concrete overview of this dissertation is described as follows.

Chapter 2 concerns on optimal sparse feedback control synthesis from its open-loop solution to closed-loop realization, where the feedback matrices of sparse optimal control can be derived by means of a dynamic linear compensator [145, 147]. In other words, a constrained sparse optimization is firstly performed and then the feedback realization is achieved, thereby the sparse feedback controller can be derived from its open-loop optimal solutions. One of its benefits is its reliance solely on *offline optimization*, which is more efficient than establishing real-time control with iterative processes to create the closed-loop response, such as predictive control [52, 104], self-triggered control [104], and dynamic programming [83]. Thus,the proposed sparse feedback controller can often offer a significant computational saving and control effort minimization. Also, such a sparse feedback control is analytical and explicit, which also provides the stability, optimality, and sparsity for the closed-loop augmented system, and displays that the designed sparse feedback control is essentially a deadbeat control. The relationship among the sparse feedback control and the open-loop sparse optimal is discussed. Also, as will be shown in this chapter, the synthesized sparse feedback control is equivalent to open-loop sparse optimal control under a specified basis. In addition, the similar principle is useful to design a minimum attention feedback control by using a dynamic tracking compenstor, which is useful to tracking control problem.

Chapter 3 and Chapter 4 discuss the sparse decision making with stochastic uncertainty and the applications of the risk-aware sparse optimal (predictive) control problems to the uncertain control systems. Chapter 3 presents a novel chance-constrained sparse optimization problem that aims to perform sparse decision-making when the constraints are in terms of chance or probabilistic model [148]. For the metric of the probabilistic problem setup, a small fraction of constraint violations is allowed governed by a given risk level. This uncertainty quantification is closely related to robust optimization [14], chance-constrained program [45], and stochastic programming [122]. Here, it should be emphasized that although the proposed chance-constrained sparse optimization problem is well-defined, the solution is computationally infeasible since the “nonconvex” structure and multiple-integral calculation. By virtue of convex relaxation and data-driven sampling techniques [37, 126], the intractable optimization problem can be converted into the finite sparse random convex program that the obtained sparse randomized solution is feasible to original chance-constrained sparse optimization with a priori and high confidence probabilistic robustness guarantees. In addition, a trade-off between the sparse cost performance and the risk level is studied, giving rise to a relaxed sparse optimization problem [143]. The objective function is thus penalized by not only the sparse cost but also the violated regrets, and a posteriori probabilistic guarantee is ensured, and the related work is similar to the sparse support vector machine [60]. Here, it should be emphasized that the results of sparse random convex program can be successfully applied to the robust control applications, referred to as risk-aware sparse optimal control.

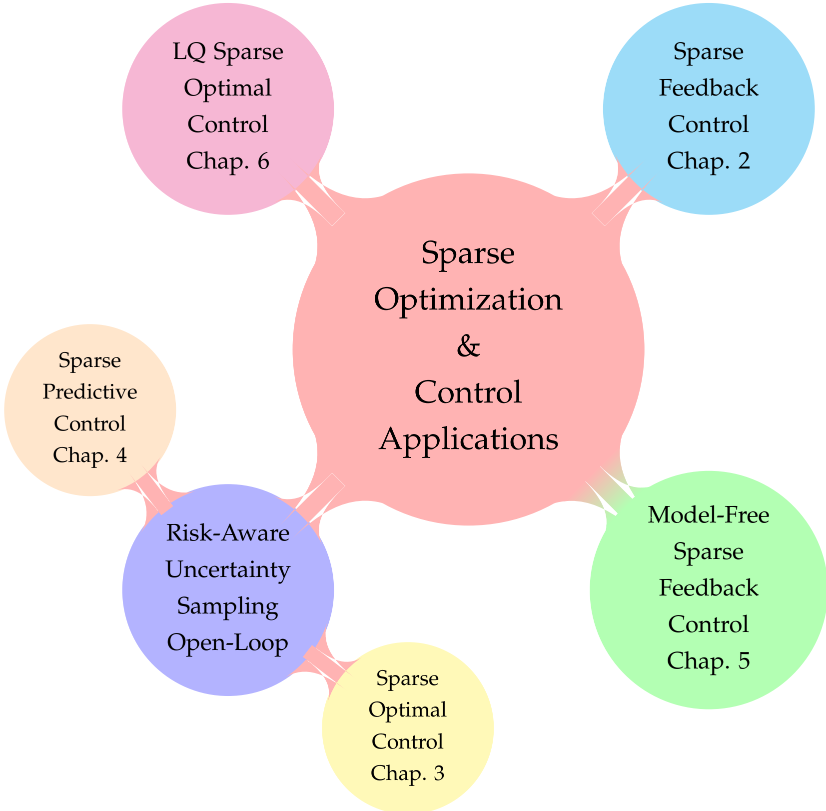
Chapter 4 investigates a risk-aware sparse predictive control problem for a discrete-time uncertain system subject to external stochastic noise and model parametric uncertainty [144], which is related to stochastic model predictive control [33, 90, 98] and sparse predictive control problems [77, 104]. This problem focuses on the control mechanism of how system state can drive from an initial state near to a prescribed terminal set through a minimum control effort of the predictive control sequences. In model predictive control, the principle of a receding horizon control strategy is analogous to playing chess game, that is, only the first control action of all predicted control inputs is applied to the dynamical systems, which requires the control system to repeatedly solve an open-loop optimal control problem over a prediction horizon. Thus, the risk-aware sparse predictive control becomes a real-time control problem by performing an *online sparse random convex optimization* problem. As shown in Chapter 3, the result can also provides the probabilistic robustness guarantees for the spare predictive control.

Chapter 5 explores the data-driven sparse feedback control for a *model-free* discrete-time linear time-invariant (LTI) system, where the true dynamics of the system is *unknown* and the control inputs of the feedback response is *sparse*. More concretely, the control objective tends to *learn sparse feedback controllers directly from the input/state/output data trajectories*, in contrast to the common approaches where a priori knowledge of the actual system model is given or identified to synthesize a "model-based" sparse (feedback) controller, discussed in Chapter 2 – 4, and the literature therein. Assuming that the input signals for a single or multiple experiments [131] either hold *persistently exciting* (PE) condition [54, 96, 137] or go beyond one using *data informativity* [129], then LTI system can be successfully reconstructed by the input/state/output data samples, resulting in a data assisted LTI systems representation. Similar to Chapter 2, here Chapter 5 also intends to pursue a sparse feedback controller with input sparsity and closed-loop stability. However, there are some distinctions here: Firstly, the considered dynamical system operates under a black-box setup, meaning it lacks a known model but is accessible through input/state/output data based on a PE condition [54]. Secondly, the input sparsity is achieved by employing a row-sparsity structured feedback gain matrix [113, 114], accomplished through penalizing an $\ell_{1,\infty}$ norm on itself [128]. Therefore, the derived controller naturally is data-driven sparse feedback controller.

Chapter 6 proposes a linear quadratic (LQ) sparse optimal control, also named as *LQ hands-off control*, for the continuous-time master-slave systems that combines the LQ cost with a cardinality penalty to promote the sparse input in quadratic performance as well as reach trajectory tracking [149]. The articles [71, 105] have obtained the necessary conditions of sparse control problems, and they were devoted to produce the sparse solutions via convex surrogate, that is, $\mathcal{L}_1$ optimal solutions. Observe that the exact $\mathcal{L}_0$ measure leads to non-convex and discontinuous optimal problem, this dissertation tackles this obstacle directly without any approximation in theoretical analysis, as mentioned in [46, 79]. From this, it is possible to derive the precise $\mathcal{L}_0$ optimality conditions for LQ hands-off control based on the *non-smooth maximum principle* [49]. As a result, the proposed LQ hands-off control may not be continuous due to the sparsity pattern, while the standard LQ control is continuous and smooth in general. In order to compute numerical solutions efficiently, the time-discretization method and $\mathcal{L}_1$ relaxation is employed, and the proposed LQ sparse optimal problem can be reduced to a finite-dimensional convex problem. In addition, the robustness of the LQ hands-off control in the presence of perturbations in the initial states and the state-space matrices are explored. Simulation results are shown to illustrate the effectiveness of the proposed control method, and the robustness of the LQ hands-off control.

This thesis ends with Chapter 7 offering concluding remarks. The contributions and their limitations are outlined, and possible future developments are then provided.

More specifically, the interconnection between each chapter of this dissertation is depicted in the following diagram.



1.4 Mathematical preliminaries

Throughout this dissertation, we introduce the following mathematical preliminaries for sparse signal in discrete-time and continuous-time, respectively.

For a vector $x \in \mathbb{R}^n$, $\mathrm{supp}(x)$ denotes the support set of x, that is, the set of non-zero elements of $x = [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$:

$$\mathrm{supp}(x) \triangleq \{i \in \{1, \ldots, n\} : x_i \neq 0\}.$$

Definition 1.1 (ℓ_0 norm) *An ℓ_0 "quasi-norm"[1] of the vector $x \in \mathbb{R}^n$ is defined by its support*

$$\|x\|_0 = \#\{\mathrm{supp}(x)\} = \#\{i : x_i \neq 0, i = 1, \cdots, n\},$$

[1]Noticed that here the exact ℓ_0 "pseudo-norm" is not properly (or mathematically) a norm, since it lacks the absolute homogeneity, namely, $\|u\|_0 = \|\alpha u\|_0 \neq |\alpha| \|u\|_0$ for $|\alpha| \neq 1$; Similarity, the class of ℓ_p norm with $0 < p < 1$ is also a pseudo norm, a graphical illustration for sparsity is shown in Figure 1.3. For the sake of simplicity, we adopt quotation mark on quasi-norm.

where the symbol # returns the number of elements if the argument set.

Definition 1.2 (Sparsity) *A sparse vector is a vector having a relatively small number of nonzero elements.*

Example 1.1 (Sparse Solution) Find sparsest solution that minimizes

$$\min_{x} \quad \|x\|_0 \quad \text{s.t.} \quad x \in \mathcal{X} \subseteq \mathbb{R}^n. \tag{1.1}$$

For instance, a visual representation of a sparse vector $x \in \mathbb{R}^{33}$ is given as follows

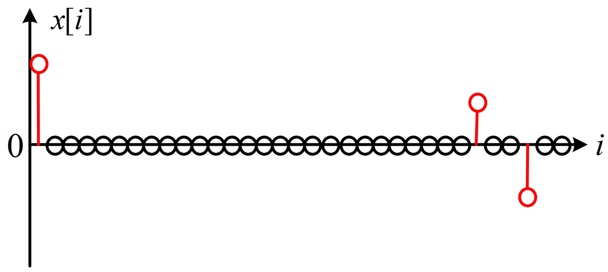


Figure 1.1: Sketch of sparse vector $\|x\|_0 = 3$ with $x \in \mathbb{R}^{33}$, where only three entries are active (i.e., non-zero) elements marked as "red-o" and the others are all zero elements marked as "black-o", respectively.

Definition 1.3 (ℓ_p norm) *The ℓ_p norm with $p \geq 1$ of a vector $x \in \mathbb{R}^n$ is defined by*

$$\|x\|_p \triangleq \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}, \tag{1.2}$$

and the ℓ_∞ norm is defined by

$$\|x\|_\infty \triangleq \max_{i=1,\ldots,n} |x_i|. \tag{1.3}$$

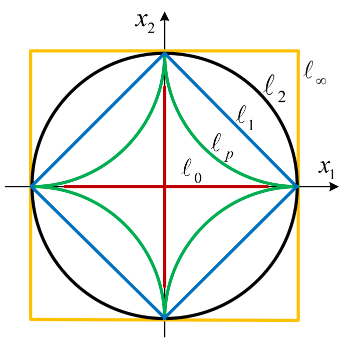


Figure 1.2: ℓ_p norms for a vector $x \in \mathbb{R}^2$ under $p = \{0, 1, 2, \infty\}$, where different norms are highlighted as different colors: ℓ_0 "norm" (red), ℓ_p norm with $0 < p < 1$ (green), ℓ_1 norm (blue), ℓ_2 norm (black), and ℓ_∞ norm (yellow), respectively.

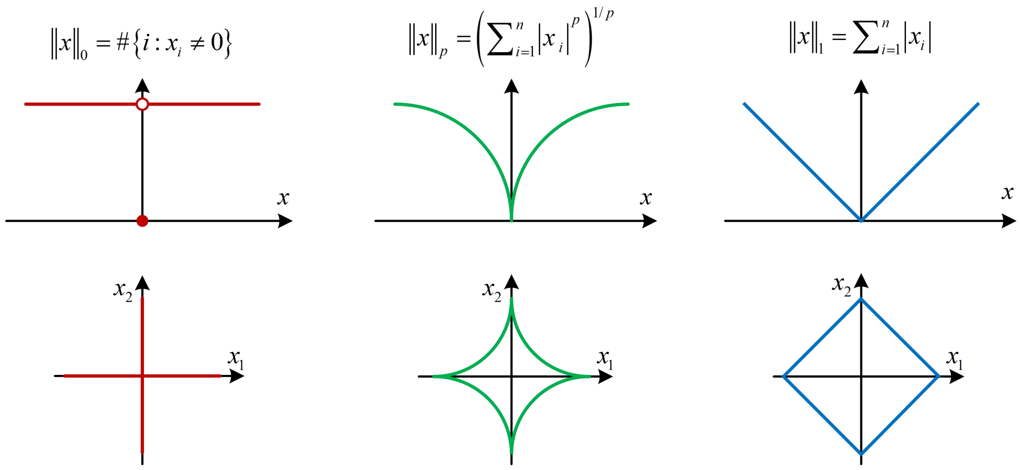


Figure 1.3: Exactness versus Approximation versus Relaxation for sparsity of a vector $x \in \mathbb{R}^n$: exact ℓ_0 "norm" (left), approximated ℓ_p "norm" with $0 < p < 1$ (middle), relaxed ℓ_1 norm (right), respectively.

Definition 1.4 (Support) *Let $f : [0, t_f] \to \mathbb{R}$ be a measurable function with $t_f > 0$, the support of a function f over the interval $[0, t_f]$ is defined by*

$$\text{supp}(f) \triangleq \{t \in [0, t_f] : f(t) \neq 0\}. \tag{1.4}$$

Utilizing the support of function f, we can define the $\mathcal{L}_0$ norm of f as follows.

$$\|f\|_0 = \mu_m(\text{supp}(f)), \tag{1.5}$$

where $\mu_m(\mathcal{S})$ is the Lebesgue measure of a subset $\mathcal{S} \subset [0, t_f]$.

Besides, for a continuous-time control signal $f(t)$, $t \in [0, t_f]$, the sparse control signal governed by a $\mathcal{L}_0$ norm in (1.5) can be rewritten as follows

$$\|f\|_0 = \int_0^{t_f} \phi_0(f(t))dt \tag{1.6}$$

where $\phi_0(\cdot)$ is the kernel function, described by

$$\phi_0(f) \triangleq \begin{cases} 1, & \text{if } f \neq 0, \\ 0, & \text{if } f = 0. \end{cases} \tag{1.7}$$

Chapter 2

Sparse Feedback Control by Dynamic Linear Compensator

This chapter explores the discrete time sparse feedback control for a linear time invariant (LTI) system, where the proposed optimal feedback controller enjoys input sparsity by using a dynamic linear compensator, namely, the components of feedback control signal having the smallest possible nonzero values. The resulting augmented dynamics ensures closed-loop stability, which infers sparse feedback controller from open-loop solution to closed-loop realization. In particular, the implemented sparse optimal feedback (closed-loop) control solution is equivalent to the original (open-loop) sparse optimal control solution under a specified basis. Moreover, an extension result with respect to a feedforward tracking control problem based on tracking dynamic compensator will also be discussed in this chapter. Finally, numerical examples demonstrate the effectiveness of proposed control approaches.

2.1 Review of open-loop sparse optimal control

Consider a discrete-time LTI system described by

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t), \quad x(0) = x_0, \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{2.1}$$

where $u(t) \in \mathbb{R}^m$ is the control input with $m \leq n$, $x(t) \in \mathbb{R}^n$ is the state with an initial value x_0, $y(t) \in \mathbb{R}^p$ is the output, and A, B, C, and D are real constant matrices of appropriate sizes. Throughout this chapter, we assume that the pair (A, B) is *reachable*.

In this chapter, we focus on sparse optimal control problem, and the control objective is to seek a control sequence $\{u(0), u(1), \cdots, u(N-1)\}$ such that it drives the resultant state $x(t)$ from an initial state $x(0) = x_0$ to the origin in a finite N steps (i.e., $x(N) = 0$) with minimum or sparse control effort. Such control paradigm also known as "maximum hands-off control" that *maximizes the time duration over which the control value is exactly zero* [10, 104, 105].

Based on concept of sparsity (see Definition 1.1 in Chapter 1), we now formally describe the following sparse optimal control problem.

Problem 2.1 (Open-Loop Sparse Optimal Control) *Given a discrete-time LTI dynamics* (2.1)*, an open-loop sparse optimal control aims to minimize*

$$\begin{aligned} \min_{u} \quad & c(u) = \|u\|_0 \\ \text{s.t.} \quad & x(t+1) = Ax(t) + Bu(t), \\ & x(0) = x_0, \quad x(N) = 0, \\ & x(t) \in \mathbb{X}, \quad u(t) \in \mathbb{U}, \quad \forall t = 0, 1, \cdots, N-1 \end{aligned} \tag{2.2}$$

where the cost function is related to precise sparse control input, i.e., ℓ_0 "quasi-norm" minimization of control vector (see Definition 1.1) that counts the number of nonzero elements of the control vector $u \in \mathbb{R}^{mN}$. Here the first constraint represents the system behavior, and the second constraints stand for a fixed initial state and terminal state, respectively, and the last constraints are the magnitude of path (or state-input) constraints of the system.

Remark 2.1 (Exact Sparsity) As indicated in [55], an exact sparsity is achieved by penalizing an ℓ_0 "quasi-norm" on decision variables. However, computing the ℓ_0 norm precisely is challenging due to its "non-convex" and "non-smooth" nature, often resulting in an NP-hard problem [106]. Therefore, most literature on non-convex sparse optimization have been devoted to *inexact* polynomial-time or greedy algorithms, which are not guaranteed to find optimal solutions: such as orthogonal matching pursuit (OMP) [109], iterative hard thresholding (IHT) [24], and hard thresholding pursuit (HTP) [57]. Such ℓ_0 minimization is related to subset selection [69], for example, given a vector $u \in \mathbb{R}^{n_u}$, the cardinality $\|u\|_0 \leq s$ exhibits the "s-sparse" condition for $s \ll n_u$, leading to the worst case is, $\binom{n_u}{s}$ possible combinations, especially when the dimension of decision variable n_u is high, like 100 million and even 100 billion.

Remark 2.2 (Sparse Optimization via MIP) Seeking the optimal solution to Problem 2.1 equivalents to solving a constrained sparse optimization problem, the "exact" sparse solution for minimizing $\|u\|_0$ quasi-norm can be converted to the mixed-integer programming (MIP) using modern solvers (e.g., Gubori and CPLEX) and employing sophisticated search strategies such as the branch-and-bound (BnB) exploration [26].

To reduce the computational complexity from exponential to polynomial time, one suggests replacing the ℓ_0 norm with the ℓ_1 norm, $\|x\|_1 = \sum_{i=1}^{n} |x_i|$, as ℓ_1 norm is a tight convex relaxation of ℓ_0 norm, which still generates the relative sparse solution. For more details with respect to sparse knowledge, we refer to Appendix A.

In this chapter, we shift the principle from compressed sensing to sparse optimal control problem, which is to seek an "open-loop" ℓ_1 optimal control action u^* for a discrete time LTI control system (2.27). Going back to Problem 2.2, we replace the exact sparse cost $\|u\|_0$ by a convex relaxation ℓ_1 norm (i.e., $\|u\|_1$), the following sparse optimization is central and can be described by [10, 104]

$$u^* = \arg\min_{u \in \mathcal{U}} \|u\|_1 = \arg\min_{u \in \mathcal{U}} \sum_{t=0}^{N-1} \|u(t)\|_1, \tag{2.3}$$

where $\|u\|_1$ indicates the ℓ_1 norm of input vector u

$$u \triangleq \left[u^\top(0)\ u^\top(1)\ \cdots\ u^\top(N-1)\right]^\top \in \mathbb{R}^{mN},$$

that *sums the absolute values of its elements*.

For program (2.3), a feasible control set is defined as

$$\mathcal{U} \triangleq \left\{u \in \mathbb{R}^{mN} : \Phi_N u = -A^N x_0\right\}, \quad \Phi_N = \left[A^{N-1}B|\cdots|AB|B\right] \in \mathbb{R}^{n\times mN}, \tag{2.4}$$

where Φ_N is an N-step reachability matrix satisfying full row rank, i.e., $\operatorname{rank}(\Phi_N) = n$. Occasionally, the state and input constraints are necessarily taken into account, for instance, a mixed input-state constraint can be concluded as an output constraint $y(t) \in \mathcal{Y}$, where $\mathcal{Y}$ is a convex and closed set. Besides, the horizon N should be sufficiently long so that the admissible set of u is *non-empty*, which implies the existences of sparse control solution u^*.

We sometimes call ℓ_1 optimal control problem (2.3) as *minimum fuel control problem*. In particular, the restricted isometry property (RIP) [39] reveals an equivalence between ℓ_0 optimal control (2.2) and ℓ_1 optimal control (2.3) for discrete-time LTI systems [104]. Obviously, the convex "open-loop" ℓ_1 optimal solution u^* can be calculated by some optimization methods, such as, linear program [1, 116] or convex optimization, etc.

However, for discrete-time systems, sparse optimal control would be deprived of some wonderful results, such as it lacks bang-off-bang property (in continuous-time systems) [104, 105], and it is also hard to offer the explicit open-loop solution. To intuitively recognize non-feedback control, let us first reduce the Problem 2.1 to a well-known ℓ_2 optimal control (a.k.a., minimum energy control) problem [7], see a preview Example 2.1, which provides an explicit, open-loop, analytic solution.

Example 2.1 (Minimum Energy Control) Given a finite control horizon $N \in \mathbb{N}_{\geq 0}$ and a desired finial state $x(N) = x_f$ (does not need to be zero), the minimum energy control asks for finding the control sequence $\{u(t)\}_{t=0}^{N-1}$ such that it steers the system state $x(t)$ in (2.1) from x_0 to x_f in N steps, which becomes to solve the following problem

$$\begin{aligned} \min_{u} \quad & c(u) = \sum_{t=0}^{N-1} \|u(t)\|_2 \\ \text{s.t.} \quad & x(t+1) = Ax(t) + Bu(t), \quad \forall t = 0, 1, \ldots, N-1, \\ & x(0) = x_0, \quad x(N) = x_f. \end{aligned} \tag{2.5}$$

As established in a classic result [74], the energy minimization problem in Example 2.1 is feasible if and only if $(x_f - A^N x_0) \in \mathrm{Im}(\mathcal{W}_N)$, where

$$\mathcal{W}_N = \sum_{t=0}^{N-1} A^t B B^\top (A^\top)^t$$

is the N steps *controllability Gramian* and $\mathrm{Im}(\mathcal{W}_N)$ means the image of the matrix $\mathcal{W}_N$. Furthermore, the solution to Problem 2.5 is

$$u^*(t) = B^\top (A^\top)^{N-t-1} \mathcal{W}_N^\dagger (x_f - A^N x_0), \tag{2.6}$$

where $\mathcal{W}_N^\dagger$ the is the *generalized inverse* or Moore-Penrose pseudo-inverse of $\mathcal{W}_N$ [12].

From the vector form perspective, the minimum-energy control problem is to find the solution to the following equation

$$x_f = A^N x_0 + \underbrace{\begin{bmatrix} B & AB & \dots & A^{N-1}B \end{bmatrix}}_{\mathcal{C}_N} \underbrace{\begin{bmatrix} u(N-1) \\ u(N-2) \\ \vdots \\ u(0) \end{bmatrix}}_{u_N},$$

where x_f is reachable, $\mathcal{C}_N$ denotes the N-steps controllability matrix and $u_N \in \mathbb{R}^{mN}$ is the control vector over finite-horizon $[0, N-1]$. Thus, the minimum-energy control input to arrive at x_f is derived as

$$u_N^* = \mathcal{C}_N^\dagger (x_f - A^N x_0). \tag{2.7}$$

Example 2.1 is a quadratic convex program, and its extension is linear quadratic regulator (LQR) control, i.e., taking $c(x,u) = x^\top Q x + u^\top R u$ with $Q \succ 0$ and $R \succ 0$. In general, no matter for ℓ_1 optimal control in Problem 2.1 or more common ℓ_2 optimal control in Example 2.1, the resulting literature are almost related to open-loop solution [10] rather than the feedback solution, since determining the feedback solution for optimal control under some constraints is a not a simple task.

2.2 Closed-loop solution via dynamic programming

In discrete-time system, a standard smooth feedback maps $\kappa : \mathbb{R}^n \to \mathbb{R}^m$ is defined as $u = \kappa(x)$, and can be applied to the plant (2.1), then the closed-loop system becomes

$$x(t+1) = f_{cl}(x(t), \kappa(x(t))), \tag{2.8}$$

with a *smooth* $f_{cl}(x, u) = Ax + Bu$ such that

$$\mathrm{J}(x) : \min_{u(0),\cdots,u(N-1)} \sum_{t=0}^{N-1} c(x(t), u(t)) = \sum_{t=0}^{N-1} c(x(t), \kappa(x(t))),$$

for any initial state x_0 for which $\mathrm{J}(x_0)$ is well-posed, here $\mathrm{J}(x)$ is called *value function* or optimal cost function that satisfies Bellman's dynamic programming (DP) equation

$$\mathrm{J}(x) = \min_{u} \left[\mathrm{J}(f_{cl}(x, u)) + c(x, u) \right], \tag{2.9}$$

which implies the optimal feedback control

$$\kappa^*(x) = \arg\min_{u} \left[\mathrm{J}(f_{cl}(x, u)) + c(x, u) \right]. \tag{2.10}$$

Hence, we have

$$\mathrm{J}^*(x) = \mathrm{J}(f_{cl}(x, \kappa^*(x))) + c(x, \kappa^*(x)),$$

and the closed-loop system response (2.8) has the origin (i.e. $f_{cl}(0, 0) = 0$) as a stable equilibrium with Lyapunov function $\mathrm{J}(x)$. This gives a *first-order necessary condition* for a minimum of the optimal feedback control

$$0 = \frac{\partial \mathrm{J}}{\partial x}(f_{cl}(x, \kappa^*(x))) \frac{\partial f_{cl}}{\partial u}(x, \kappa^*(x)) + \frac{\partial c}{\partial u}(x, \kappa^*(x)). \tag{2.11}$$

Unfortunately, the DP method [83, Chapter 5] fails to dervie the *explicit* and *closed-form* solution from condition (2.11), and thus one way is to compute numerical solution via patch domain construction [2].

To tackle this difficulty, a seminal work by [23], they build on a relatively optimal control from open-loop solution to closed-loop solution.

2.3 Sparse feedback control problem

Before proceeding with the "feedback realization", a compensator $\mathcal{K}$ which we demand to design is a *dynamic* and *linear* state feedback controller, depending on the evolution

$$\mathcal{K} : \begin{aligned} z(t+1) &= Fz(t) + Gx(t), \quad z(0) = 0, \\ u(t) &= Hz(t) + Kx(t), \end{aligned} \tag{2.12}$$

where $z(t)$ is the compensator's state and F, G, H, and K are real matrices of compatible sizes. Note that we set the initial state $z(0)$ of the compensator as *zero* (i.e., $z(0) = 0$).

The advantages of the dynamic compensator $\mathcal{K}$ introduced in (2.12) are threefold. First, it brings the *linear* and *dynamic* fashions to the sparse feedback control realization, which allows for computationally tractable compensator gain matrices. Second, it promotes *input/temporal* sparsity [10, 104, 105], rather than structured or spatial sparsity for the controller [8, 59, 73, 113, 85, 97]. Lastly, it ensures *internal stability* for the synthesized closed-loop behavior.

Remark 2.3 (Initialization Robustness) Notice that the requirement $z(0) = 0$ is not overly restrictive in our context. In fact, in the following section, we further impose $z(N) = 0$, along with $x(N) = 0$, as part of the sparse control implementation. This means that we consider the closed-loop system through "deadbeat control". In this case, once the stable closed-loop reaches the zero state within a finite time, the condition $z(0) = 0$ is automatically fulfilled whenever we have another $x(0) \neq 0$ due to a new disturbance.

In what follows, we focus on optimal sparse feedback control synthesis *from open-loop solution to closed-loop realization*. A natural question is whether the obtained optimal input-state trajectories alone implies "optimal feedback" about sparse control (2.12).

Our answer is yes, but with regards to a more general initial condition (i.e., *a set of initial guesses*, see (2.13)) rather than a single initial condition stated in Problem 2.1, which is based on the technique called *relatively optimal control* [23]. Thereby, determining the explicit matrices of F, G, H, K for dynamic compensator (2.12) is of primary interest in designing sparse feedback controller.

Firstly, we introduce *a general initial scenario* for discrete LTI system (2.1), as follows

$$x_0 \in \mathcal{X}_0 \doteq \{\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n\}, \tag{2.13}$$

which is used to *generate all n possible input-state trajectories*, here e_i can be viewed as the canonical basis vector of $\mathbb{R}^n$, for example, $\mathrm{e}_1 = [1\ 0\ \cdots\ 0]^\top \in \mathbb{R}^n$. Therefore, the constrained sparse feedback control problem is formally stated follows.

Problem 2.2 (Sparse Feedback Control) *Find F, G, H, and K of controller* (2.12) *such that*

(i) the designed dynamic compensator (2.12) *stabilizes the discrete LTI plant* (2.1) *and*

(ii) for any $x_0 \in \mathcal{X}_0$ with zero initialization $z(0) = 0$, the controller (2.12) *generates an input sequence $\{u(t)\}_{t=0}^{N-1}$, which minimizes $\sum_{t=0}^{N-1} \|u(t)\|_1$ subject to the terminal constraints $x(N) = 0$ and $z(N) = 0$ for a positive integer N, as well as the mixed state and input (or output) constraints*

$$y(t) \in \mathcal{Y}, \qquad \mathcal{Y} = \{y \in \mathbb{R}^p : -s \leq y(t) \leq s\}, \tag{2.14}$$

where $s \in \mathbb{R}^p$ is a given positive vector.

Remark 2.4 The input sparsity can be easily performed by minimizing the convex ℓ_1 norm instead of the non-convex ℓ_0 norm, as seen in the previous section. Moreover, although we select only a subset of the initial states of the plant, i.e., $x_0 \in \mathcal{X}_0$, which is sufficient for our purposes. In fact, suppose that the given constrained sparse control problem is solved. Since the resultant closed-loop system composed of (2.1) and (2.12) is *linear*, it means that for any $x_0 \in \mathbb{R}^n$ with $z(0) = 0$, the controller (2.12) generates a linear combination of the input sequences corresponding to $x_0 = \mathrm{e}_1$, $x_0 = \mathrm{e}_2$, ..., $x_0 = \mathrm{e}_n$, thereby achieving sparsity and satisfying $x(N) = 0$ and $z(N) = 0$.

2.4 Closed-loop stability for augmented system

As mentioned in the celebrated works [21, 23], the linear implementation can be built from the relatively optimal control strategy. For a compact closed-loop system representation, we accordingly introduce a stable matrix P, which is an N-Jordan block associated with 0 eigenvalue, defined by

$$P = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \triangleq \begin{bmatrix} 0 & 0 \\ I_{N-1} & 0 \end{bmatrix} \in \mathbb{R}^{N \times N}. \tag{2.15}$$

Obviously, P is a nilpotent matrix with property $P^N = 0$

This oracle suggests us to investigate an augmented closed-loop system composed of the discrete dynamics (2.1) and the dynamic compensator (2.12) of the form

$$\begin{aligned} \psi(t+1) &= (\mathcal{A} + \mathcal{B}\mathcal{K})\psi(t), \quad \psi(0) = \psi_0, \\ y(t) &= (\mathcal{C} + \mathcal{D}\mathcal{K})\psi(t), \end{aligned} \tag{2.16}$$

where the corresponding state and the gain matrices are given by

$$\begin{aligned} \psi(t) &= \begin{bmatrix} x(t) \\ z(t) \end{bmatrix}, & \psi_0 &= \begin{bmatrix} x_0 \\ 0 \end{bmatrix} & & \\ \mathcal{A} &= \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}, & \mathcal{B} &= \begin{bmatrix} B & 0 \\ 0 & I \end{bmatrix}, & \mathcal{K} &= \begin{bmatrix} K & H \\ G & F \end{bmatrix}, \\ \mathcal{C} &= \begin{bmatrix} C & 0 \end{bmatrix}, & \mathcal{D} &= \begin{bmatrix} D & 0 \end{bmatrix}. & & \end{aligned}$$

Based on the previous problem setup, the sparse optimal control with a general initial condition (i.e., $x_0 \in \mathcal{X}_0$), we thus need to solve n sparse optimal control problems simultaneously, which can be recast as the constrained sparse matrix optimization.

Problem 2.3 (Constrained Sparse Matrix Optimization) *Find the matrices $X \in \mathbb{R}^{n \times nN}$ and $U \in \mathbb{R}^{m \times nN}$ such that the obtained U is sparse, which amounts to solve a constrained convex matrix optimization*

$$\begin{aligned} \min_{X,U} \quad & \|U\|_1 = \sum_i \sum_j |U_{ij}| \\ \text{s.t.} \quad & AX + BU = X(P \otimes I_n), \end{aligned}$$

$$I_n = X(\mathrm{e}_1 \otimes I_n),$$
$$\mathrm{abs}(CX + DU) \leq s(\mathbf{1}_n \otimes \mathbf{1}_N)^\top,$$

where $\mathrm{e}_1 = [1\ 0\ \cdots\ 0]^\top \in \mathbb{R}^N$, *the symbol* $\otimes$ *is the kronecker product, and* $\mathrm{abs}(\cdot)$ *returns the absolute value of each element in a matrix.*

Problem 2.3 reduces to a convex optimization, and hence the solution is computationally tractable by means of the off-the-shelf packages, such as `CVX` [67] or `YALMIP` [88]. Once the open-loop optimal solution (X, U) of Problem 2.3 is attained, we then proceed the second step that tackles the following sparse feedback realization problem.

Problem 2.4 (Feedback Realization) *Based on the solution* (X, U) *of Problem 2.3, the feedback realization needs to solve a linear matrix equation*

$$\begin{bmatrix} K & H \\ G & F \end{bmatrix} \begin{bmatrix} X \\ Z \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} \tag{2.17}$$

with respect to (K, H, G, F) *and determine the compensator's gain matrices, where*

$$Z = \begin{bmatrix} 0_{n(N-1)\times n} & I_{n(N-1)} \end{bmatrix}, \qquad V = Z(P \otimes I_n). \tag{2.18}$$

The approach to sparse feedback control design makes use of the above discussed sparse optimization (Problem 2.3) and feedback realization (Problem 2.4), where the dynamic compensator ensures the internally stability of the closed-loop augmented system (2.16). The result is summarized in the following theorem and corollary.

Theorem 2.1 (Sparse Feedback Control Realization) *Suppose that Problem 2.3 has the minimizer* (X, U)*. Then the equation (2.17) has the unique solution* (K, H, G, F) *and the resulting compensator (2.12) with* $z(0) = 0$ *generates the input sequence* $u(t) = U(\mathrm{e}_{t+1} \otimes x_0)$*,* $t = 0, 1, \ldots, N-1$*, for* $x_0 \in \mathcal{X}_0$*, which drives the plant state* $x(t)$ *from the initial state* $x(0) = x_0$ *to the terminal state* $x(N) = 0$ *under the output constraint (2.14). Furthermore, the closed-loop system (2.16) is internally stable.*

Proof 2.1 We first describe the matrices X and U[1]

$$X = \begin{bmatrix} X_0 & X_1 & \cdots & X_{N-1} \end{bmatrix}, \qquad U = \begin{bmatrix} U_0 & U_1 & \cdots & U_{N-1} \end{bmatrix},$$

where $X_t \in \mathbb{R}^{n\times n}$, $U_t \in \mathbb{R}^{m\times n}$, and $t = 0, 1, \ldots, N-1$. Notice that checking the second constraint of Problem 2.3 gives rise to the result $X_0 = I_n$. With this fact, it admits that the matrix Ψ is *non-singular*. In other words, the matrix determinant claims that

$$\det(\Psi) \neq 0, \qquad \Psi = \begin{bmatrix} X \\ Z \end{bmatrix} = \left[\begin{array}{c|c} I_n & X_1 \ \cdots \ X_{N-1} \\ \hline 0_{n(N-1)\times n} & I_{n(N-1)} \end{array}\right]. \tag{2.19}$$

[1] For $x_0 \in \mathcal{X}_0$, the state trajectory is $X = [x^1(0)\ x^2(0)\ \ldots\ x^n(0)\ x^1(1)\ x^1(2)\ \ldots\ x^n(N-1)]$ and the input trajectory is $U = [u^1(0)\ u^2(0)\ \ldots\ u^n(0)\ u^1(1)\ u^1(2)\ \ldots\ u^n(N-1)]$.

Since the matrix Ψ is invertible, it follows that the equation (2.17) has the unique solution (K, H, G, F) associated with the dynamic compensator (2.12). Meanwhile, the first constraint of Problem 2.3 with (2.17) and (2.18) asserts that

$$\begin{aligned}
(\mathcal{A} + \mathcal{B}\mathcal{K})\Psi &= \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X \\ Z \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} \\
&= \begin{bmatrix} AX + BU \\ V \end{bmatrix} \\
&= \begin{bmatrix} X \\ Z \end{bmatrix} (P \otimes I_n) \\
&= \Psi(P \otimes I_n).
\end{aligned}$$

This implies that the closed-loop augmented matrix $\mathcal{A} + \mathcal{B}\mathcal{K} = \Psi(P \otimes I_n)\Psi^{-1}$, so that it is similar to a nilpotent matrix $(P \otimes I_n)$. Hence, the closed-loop system (2.16) is internally stable and the zero terminal state $x(N) = 0$ is achieved for any initial state $\psi(0)$ of the system.

Moreover, we see that the sequences

$$\begin{aligned}
x(t) &= X(\mathrm{e}_{t+1} \otimes x_0) = X_t x_0, \\
u(t) &= U(\mathrm{e}_{t+1} \otimes x_0) = U_t x_0, \\
z(t) &= Z(\mathrm{e}_{t+1} \otimes x_0) = Z_t x_0
\end{aligned}$$

are indeed generated by the system (2.1) with $x(0) = x_0$ and the controller (2.12) with $z(0) = 0$. As a matter of fact, apparently $x(0) = X(\mathrm{e}_1 \otimes x_0) = X_0 x_0 = x_0$ and $z(0) = Z(\mathrm{e}_1 \otimes x_0) = 0$. Furthermore, based on the fact that $(P \otimes I_n)(\mathrm{e}_{t+1} \otimes x_0) = \mathrm{e}_{t+2} \otimes x_0$, we have the system behavior

$$\begin{aligned}
Ax(t) + Bu(t) &= (AX + BU)(\mathrm{e}_{t+1} \otimes x_0) \\
&= X(P \otimes I_n)(\mathrm{e}_{t+1} \otimes x_0) \\
&= X(\mathrm{e}_{t+2} \otimes x_0) \\
&= x(t+1), \qquad (2.20)
\end{aligned}$$

and the dynamic control evolution

$$\begin{aligned}
Fz(t) + Gx(t) &= (FZ + GX)(\mathrm{e}_{t+1} \otimes x_0) \\
&= Z(P \otimes I_n)(\mathrm{e}_{t+1} \otimes x_0) \\
&= Z(\mathrm{e}_{t+2} \otimes x_0) \\
&= z(t+1), \qquad (2.21)
\end{aligned}$$

which further gives rise to the control action

$$\begin{aligned} Hz(t) + Kx(t) &= (HZ + KX)(\mathrm{e}_{t+1} \otimes x_0) \\ &= U(\mathrm{e}_{t+1} \otimes x_0) \\ &= u(t). \end{aligned} \tag{2.22}$$

We next consider the third constraint (i.e., output constraint) of Problem 2.3, whose validity can be inspected by assessing the two-side inequalities

$$-\mathrm{abs}(CX + DU) \leq CX + DU \leq \mathrm{abs}(CX + DU)$$

hold true. Therefore, for a given positive vector $s \in \mathbb{R}^p$ and $x_0 \in \mathcal{X}_0$, we have

$$\begin{aligned} (CX + DU)(\mathrm{e}_{t+1} \otimes x_0) &\leq \mathrm{abs}(CX + DU)(\mathrm{e}_{t+1} \otimes x_0) \\ &\leq s(\mathbf{1}_n \otimes \mathbf{1}_N)^\top (\mathrm{e}_{t+1} \otimes x_0) = s. \end{aligned} \tag{2.23}$$

Notice that $Cx(t) + Du(t) = (CX + DU)(\mathrm{e}_{t+1} \otimes x_0) = y(t)$, then the output constraint of the form $\{-s \leq y(t) \leq s\}$ in (2.14) is verified.

According to the above arguments, we claim that the sequences (2.20), (2.21), (2.22), and (2.23) indeed satisfy the input-state trajectories of LTI dynamics (2.1) under output constraint (2.14), which proves the theorem for realizing sparse feedback control. ■

Remark 2.5 (Offline vs. Online) Realizing sparse feedback control in Theorem 2.1 is based on "offline computation", and hence the computational complexity is low. Compared with sparse predictive control [3, 72, 104], a real-time feedback control is employed to ensure closed-loop dynamics and "online optimization" is repeatedly performed as a feedback controller to calculate sparse iterative solutions. Beyond all doubt, predictive feedback control naturally leads to computational burden when the sizes of controlled system is high (e.g., the curse of dimensionality), even for using a fast alternating direction method of multipliers (ADMM) algorithm [52, 104].

A direct corollary can be derived from proposed Theorem 2.1, which establishes the "equivalence" connection between the open-loop sparse optimal control solution and the closed-loop sparse optimal control solution.

Corollary 2.1 (Equivalence) *Suppose that Problem 2.3 has the minimizer (X, U). Let $u^*_{\mathcal{K}}$ be optimal sparse feedback control (i.e., closed-loop ℓ_1 optimal control) solution using a dynamic linear compensator $\mathcal{K}$ (2.12), and u^* be open-loop ℓ_1 optimal control solution u^* of program (2.3) with output constraint (2.14), respectively. Then, for $x_0 \in \mathcal{X}_0$, it holds that*

$$u^* = u^*_{\mathcal{K}} = Hz + Kx^*. \tag{2.24}$$

In addition, we claim that if the synthesized sparse feedback control is available, then it is essentially a deadbeat control.

Corollary 2.2 (Deadbeat Control) *Suppose that Problems 2.3 and 2.4 have solved, then the implemented sparse feedback controller* $u^*_{\mathcal{K}} = Hz + Kx^*$ *(i.e., closed-loop* ℓ_1 *optimal control) of discrete LTI plant* (2.1) *is essentially an N-step "deadbeat controller".*

Remark 2.6 Regarding the deadbeat control, since the designed compensator $\mathcal{K}$ brings the system state $x(t)$ to the origin exactly in N steps (satisfying $x(N) = 0$), which places all of the eigenvalues of the augmented closed-loop system matrix $\mathcal{A} + \mathcal{B}\mathcal{K}$ at the origin in the complex plane.

2.5 Extension: Tracking problem

In this section, we extend the result of sparse feedback control to tracking control problem [83, Chapter 8], also known as *reference governor* [61, 63]. In other words, reference and command governors are add-on control schemes which enforce state and control constraints on pre-stabilized systems by modifying, whenever necessary, the reference input to the closed-loop systems.

We start by giving a step-type reference signal $r(t) \in \mathbb{R}^p$ as

$$r(t) = \begin{cases} r_-, & t < 0 \\ r_+, & t \geq 0, \end{cases} \tag{2.25}$$

where $r_- \in \mathbb{R}^p$ and $r_+ \in \mathbb{R}^p$ are constant vectors as reference governors. The purpose of tracking problem is to design a dynamic tracking compensator such that the performance output tracks a reference input with zero steady-state error by using additional feedforward gains. For this reason, we define the tracking error by $e(t) = y(t) - r(t)$, where $y(t)$ is a performance output signal stated in LTI plant (2.1). Meanwhile, we make the following assumption before giving an effective tracking controller.

Assumption 2.1 *For a tracking problem, assume that the performance output signal* $y(t) \in \mathbb{R}^p$ *and the control signal* $u(t) \in \mathbb{R}^m$ *in LTI dynamics* (2.1) *be of the same size (i.e.,* $p = m$*) and take the matrix* $D = 0_m$.

It is known that the performance output $y(t) \in \mathbb{R}^m$ can track any reference signal $r(t) \in \mathbb{R}^m$ of (2.25) in the steady-state sense if and only if

$$\operatorname{rank} \begin{bmatrix} I - A & B \\ C & 0 \end{bmatrix} = n + m. \tag{2.26}$$

As already reported in Section 2.4, an analogous dynamic tracking compensator $\mathcal{K}_r$ can be applied to the discrete LTI plant (2.1) by adding a prescribed reference input $r(t) \in \mathbb{R}^m$ to the control actuator (2.12).

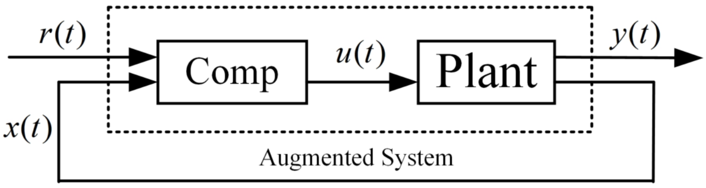


Figure 2.1: Feedforward tracking control system: $r(t)$ is reference signal; $y(t)$ is the performance output which must track a specified reference input $r(t)$; "Comp" represents a dynamic tracking compensator (2.27) applied to a discrete LTI plant (2.1).

To this end, a dynamic tracking compensator $\mathcal{K}_r$ for the plant can be designed as

$$\mathcal{K}_r : \quad \begin{aligned} z_r(t+1) &= Fz_r(t) + Gx(t) + Lr(t), \qquad z_r(0) = 0, \\ u(t) &= Hz_r(t) + Kx(t) + Mr(t), \end{aligned} \tag{2.27}$$

where L and M represent the feedforward gain matrices with suitable sizes, and $r(t) \in \mathbb{R}^m$ is a specific reference input (2.25). Notice that here the initial value $z_r(0)$ of tracking compensator $\mathcal{K}_r$ is also set to zero (i.e., $z_r(0) = 0$). Figure 2.1 shows the closed-loop augmented system composed of the plant (2.1) and the tracking compensator (2.27).

For a preferable reference input tracking, we employ *the difference or variation of the control inputs* $\sum_{t=0}^{N-1} \|u(t+1) - u(t)\|_1$ as the performance index, referred to as minimum attention control (MaC) [28, 80, 103]. We slightly relax the constraints in the previous sections by removing the state and input constraints (2.14).

As a result, we formulate the following *tracking MaC feedback realization problem* that we intend to to solve here:

Problem 2.5 (Tracking MaC) *Find F, G, H, and K of tracking controller* (2.27) *such that*

(i) the dynamic tracking compensator (2.27) *stabilizes the discrete LTI plant* (2.1) *and*

(ii) for any $x_0 \in \mathcal{X}_0$ with zero initialization $z_r(0) = 0$ and $r(t) \equiv 0$, the controller (2.27) *generates an input sequence $\{u(t)\}_{t=0}^{N-1}$, which minimizes $\sum_{t=0}^{N-1} \|u(t+1) - u(t)\|_1$ subject to $x(N) = 0$ and $z(N) = 0$ for a positive N.*

Then, determine L and M of compenstor (2.27) *such that the steady state gain of the closed-loop system from $r(t)$ to $y(t)$ is the identity and that from $r(t)$ to $z_r(t)$ is zero.*

Remark 2.7 When we have a solution to the above problem, we see that $y(t)$ tracks $r(t)$ without steady state error owing to the selected steady state gain. We also observe that $u(t)$ achieves MaC for any $x_0 \in \mathbb{R}^n$ due to linearity of the system. Moreover, since $z(N) = 0$ is achieved in the steady state for any r_+, the condition $z(0) = 0$ is automatically satisfied whenever we have another r_+ as a new reference signal.

The closed-loop behavior can be described in an augmented description

$$\begin{aligned} \psi_r(t+1) &= (\mathcal{A} + \mathcal{B}\mathcal{K})\psi_r(t) + \mathcal{M}_r r(t), \quad \psi_r(0) = \psi_{r0}, \\ y(t) &= \mathcal{C}\psi_r(t), \end{aligned} \tag{2.28}$$

where

$$\psi_r(t) = \begin{bmatrix} x(t) \\ z_r(t) \end{bmatrix}, \qquad \mathcal{M}_r = \begin{bmatrix} BM \\ L \end{bmatrix},$$

and the other matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{K}$, and $\mathcal{C}$ have been defined for (2.16).

According to the problem setup above, we consider the following MaC problem.

Problem 2.6 (Minimum Attention Control) *Find the matrices $X \in \mathbb{R}^{n \times nN}$, $U \in \mathbb{R}^{m \times nN}$ such that the obtained U achieves a minimum attention, which is equivalent to solve a matrix optimization problem*

$$\begin{aligned} \min_{X,U} \quad & \|U(P \otimes I_n) - U\|_1 \\ \text{s.t.} \quad & AX + BU = X(P \otimes I_n) \\ & I_n = X(\mathrm{e}_1 \otimes I_n). \end{aligned}$$

where the control cost penalizes the ℓ_1 norm for the control variations.

Looking for the solution (X, U) of Problem 2.6 is always accessible because, the above problem is a convex program. Then, with (K, H, G, F) of (2.17) and (2.18), the resultant closed-loop system (2.28) always assures internal stability, that is, the condition $\mathcal{A} + \mathcal{B}\mathcal{K} = \Psi(P \otimes I_n)\Psi^{-1}$ holds, as discussed in Section 2.4.

We next deal with tracking problem. Due to the fact that the closed-loop augmented system (2.28) is internally stable with matrices (K, H, G, F), it admits a unique steady-state ψ_∞ for a desired reference input $r_+ = \lim_{t\to\infty} r(t)$.

More precisely, we have the following matrix equation

$$\begin{bmatrix} \psi_\infty \\ y_\infty \end{bmatrix} = \begin{bmatrix} \mathcal{A} + \mathcal{B}\mathcal{K} & \mathcal{M}_r \\ \mathcal{C} & 0 \end{bmatrix} \begin{bmatrix} \psi_\infty \\ r_+ \end{bmatrix}, \quad \text{for} \quad \psi_\infty = \begin{bmatrix} x_\infty \\ z_\infty \end{bmatrix}. \tag{2.29}$$

If $y_\infty = Cx_\infty = r_+$ for any reference r_+, the output $y(t)$ tracks reference $r(t)$ with no steady-state tracking error. If $z_\infty = 0$ for any r_+, we can have $z(0) = 0$ whenever the reference signal changes again after a steady-state is achieved.

In what follows, we are going to achieve tracking error elimination with $z_\infty = 0$ by assigning feedforward tracking gains M and L, bringing about the following lemma.

Lemma 2.2 (Steady-State Tracking) *Supposed that Assumption 2.1 and rank condition (2.26) of steady-state tracking hold. If Problem 2.6 has the miniminer (X, U), then the priori gain matrices (K, H, G, F) with respect to compensator $\mathcal{K}_r$ (2.27) can be uniquely determined through feedback realization (2.17) and (2.18). More precisely, if the following conditions*

(i) $\det\,(I - (A + BK)) \neq 0$,

(ii) $\det\,(I - F) \neq 0$,

hold true, then the feedforward tracking gain matrices (M, L) in compensator $\mathcal{K}_r$ (2.27) can be easily derived with the forms

$$M = \big(C\big(I - (A+BK)\big)^{-1}B\big)^{-1}, \tag{2.30}$$
$$L = -G\big(I - (A+BK)\big)^{-1}BM. \tag{2.31}$$

Based on the obtained matrices (K, H, G, F, M, L), for all initial state $x_0 \in \mathbb{R}^n$ and any reference $r_+ \in \mathbb{R}^m$, there exist the unique steady-state values (x_∞, z_∞) such that $y_\infty = r_+$ and $z_\infty = 0$ are achieved in the steady-state.

Proof 2.2 Since the gain matrices (K, H, G, F) can previously be calculated by (2.17) and (2.18), the augmented closed-loop system (2.28) is internally stable, as reported in Theorem 2.1. We here reformulate the augmented system (2.29) as

$$\begin{bmatrix} x_\infty \\ z_\infty \end{bmatrix} = \begin{bmatrix} A+BK & BH \\ G & F \end{bmatrix} \begin{bmatrix} x_\infty \\ z_\infty \end{bmatrix} + \begin{bmatrix} BM \\ L \end{bmatrix} r_+. \tag{2.32}$$

Suppose that the zero steady-state $z_\infty = 0$, we then have

$$x_\infty = (A+BK)x_\infty + BMr_+.$$

This implies that, if the matrix $I - (A+BK)$ is invertible,

$$x_\infty = (I - (A+BK))^{-1}BMr_+.$$

Therefore, we see the result $y_\infty = Cx_\infty = r_+$ if M is selected as (2.30).

On the other hand, the steady-state of the tracking compensator z_∞ (2.32) satisfies

$$z_\infty = Gx_\infty + Fz_\infty + Lr_+.$$

Consequently, we derive

$$\begin{aligned} z_\infty &= (I-F)^{-1}(Gx_\infty + Lr_+) \\ &= (I-F)^{-1}\big(G(I-(A+BK))^{-1}BM + L\big)r_+ \end{aligned}$$

when the matrix $(I - F)$ is invertible. Thus, we see that $z_\infty = 0$ if the feedforward gain matrix L meets (2.31). Therefore, the feedforward gains (2.30) and (2.31) make error cancellation for achieving tracking MaC. ■

Remark 2.8 (Illustrative conditions) Checking determinant conditions (i) and (ii) in Lemma 2.2 can be equivalently formulated with optimal solution X to Problem 2.6. In fact, we have

$$I - (A+BK) = I_n - X_1,$$

$$I - F = I_{n(N-1)} - \left[\begin{array}{ccc|c} -X_1 & \cdots & -X_{N-2} & -X_{N-1} \\ \hline & I_{n(N-2)} & & 0_{n(N-2)\times n} \end{array}\right],$$

with $\mathcal{A} + \mathcal{B}\mathcal{K} = \Psi(P \otimes I_n)\Psi^{-1}$ and (2.19).

Theorem 2.3 *Under Assumption 2.1 and Lemma 2.2, the output $y(t)$ tracks reference $r(t)$ with no steady-state error via dynamic tracking compensator (2.27), where the tracking control input realizes minimum attention.*

Proof 2.3 The proof can be easily derived from Lemma 2.2. ■

2.6 Numerical benchmarks

In this section, we perform several numerical benchmarks to illustrate the effectiveness of the synthesized sparse feedback controller, which successfully gives a closed-loop solution with sparsity and optimality. Without loss of generality, we verify the different system models, such as single input single output (SISO) and multiple inputs and multiple outputs (MIMO) plants, respectively, based on the dynamic state feedback control. Also, a dynamic tracking compensator is applied to the tracking problem.

2.6.1 Single input

At first, we consider a single-input control system modeled as a linearized cart-pole system, and the parameters benchmark are similar to [23, Section VI], in which the mass of the cart is 0.29 [kg], mass of the pole is 0.1 [kg], length of the pole is 1 [m], gravity acceleration is 9.81 $[\mathrm{m/s^2}]$, and friction is neglected.

We then execute time-discretization of the continuous system using zero-order hold (ZoH) with sampling $\Delta t = 0.3$ [s], then the system matrices of form (2.1) are as

$$A = \begin{bmatrix} 1 & 0.3 & 0.1377 & 0.0143 \\ 0 & 1 & 0.8256 & 0.1377 \\ 0 & 0 & 0.4628 & 0.2441 \\ 0 & 0 & -3.2198 & 0.4628 \end{bmatrix}, \quad B = \begin{bmatrix} 0.1514 \\ 0.9850 \\ -0.1404 \\ -0.8416 \end{bmatrix}.$$

Meanwhile, the output matrices with respect to state-input constraints (2.14) are set to

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

then we have $y^\top(t) = [x_3(t)\ u(t)]$, which means that the enforced constrained state is only third component of the state $x_3(t)$ and the imposed constrained input is $u(t)$. By selecting the suitable variable s, it gives state-input constraints $|x_3(t)| \leq 1, |u(t)| \leq 1$.

Next, we simulate the discrete-time controlled system, the target is to drive the cart state from a non-zero initial state $x_0^\top = [0.9453,\ 0.7465,\ 0.7506,\ 0.4026]$ to the zero terminal state $x(N) = 0$ in a finite $N = 40$ steps (i.e., taking Tfinal $= (5 * N * \Delta t)/4 = 15$ [s] in a state-space representation). In order to realize the sparse feedback controller, we need to solve Problems 2.3 and 2.4 to seek the closed-loop ℓ_1 optimal solution. By computing, the total CPU time in PC is 0.38 [s] in MATLAB using `CVX` [67], and the found optimal value $\|U^*\|_1 = 6.4204$.

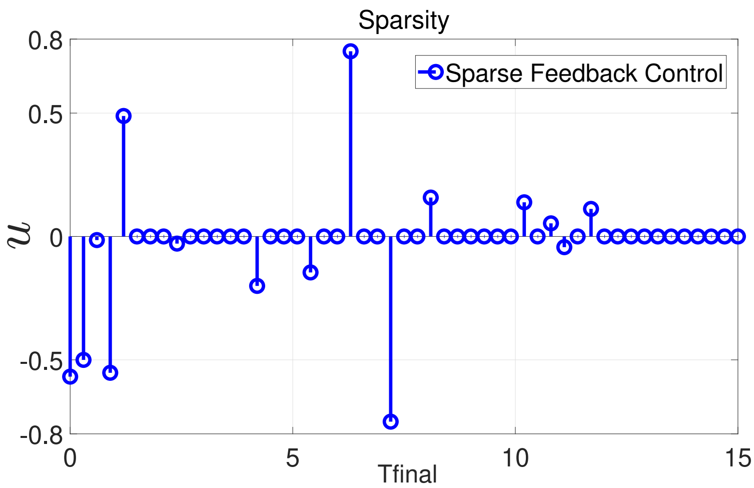


Figure 2.2: Single input case: optimal sparse feedback control inputs $u^*(t)$.

Figure 2.2 illustrates the optimal sparse feedback control inputs, and it reflects the input sparsity on sparse feedback control, in which the control sequence is with less active components, and the optimal control meets constraint $|u(t)| \leq 1$. In addition, Figure 2.3 plots the optimal state trajectories, where the pole angle x_3 fluctuates between the bounds -1 and 1, satisfying the prescribed state constraint $|x_3(t)| \leq 1$. Meanwhile, the trajectories of four different states start from an initial state x_0 and eventually converge to zero state as time tends to a fixed steps under the dynamic compensator (2.12), this implies that the closed-loop stabilization is achieved.

2.6.2 Multiple inputs numerical benchmark

As a second numerical simulation, we show the result that the synthesized dynamic controller (2.12) is useful for sparse feedback control of multi-input control system. We here consider a discretized version of third-order system with two control inputs. Using a ZoH sampling time of $\Delta t = 0.1$ $[s]$, the discrete system matrices are given by

$$A = \begin{bmatrix} 1.1133 & 0.0177 & -0.1478 \\ 0.0177 & 1.4517 & 0.2514 \\ 0.0418 & 0.2758 & 0.9208 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0031 & 0.5218 \\ 0.0121 & 0.1486 \\ 0.0957 & 0.1202 \end{bmatrix},$$

and output matrices is chosen as $C = 0$ and $D = I$, which means that only the restriction on the control inputs $|u_i(t)| \leq 10$ by taking $s_i = 10$, $\forall i = 1, 2$. We now randomly generate initial data $x_0 \in [-1, 1]^3$ and each input channel of time length is $N = 4$, then the related final time in state-space is as Tfinal $= 0.5$ $[s]$.

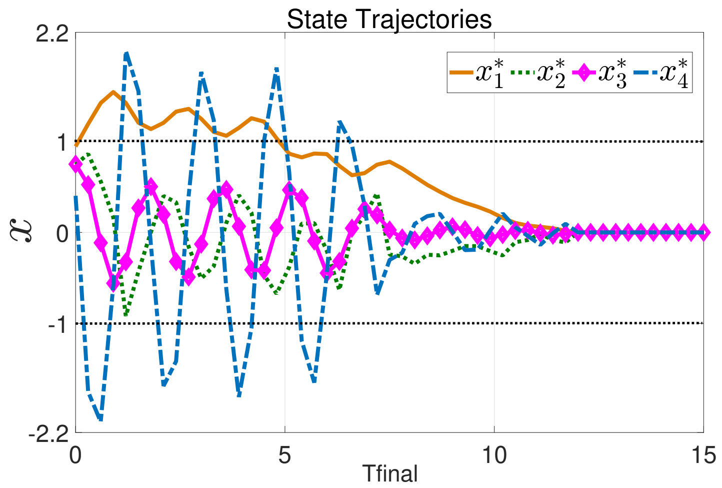


Figure 2.3: Single input case: optimal state trajectories $x^*(t)$, where the black dot line represents the constraint on the pole angle $|x_3| \leq 1$.

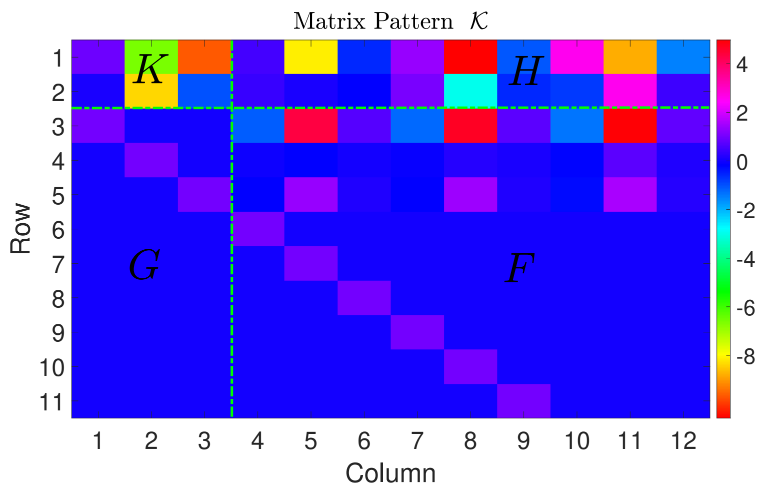


Figure 2.4: The pattern of dynamic linear Compensator $\mathcal{K}$.

After the state-input matrices X and U were calculated, see in (A.1), we obtained the optimal value $\|U^*\|_1 = 24.1544$. Due to the fact that the augmented matrix Ψ is invertible (2.19), then the controller $\mathcal{K}$ with real matrices (K, H, G, F) is computationally efficient. Figure 2.4 reveals the pattern of the compensator $\mathcal{K}$, in which the color-bar reports the level of real values of the correlation elements in matrix $\mathcal{K}$. Theorem 2.1 implies that we require the knowledge of the matrices K and H, see (A.2) to synthesize the sparse feedback control, as follows

$$u_{\mathcal{K}}^* = \begin{bmatrix} 1.8798 & 0.0000 & -0.0000 & 2.8111 & 0.0000 & 0.0000 \\ -2.7970 & 0.0000 & -0.0000 & 1.7122 & -0.0000 & 0.0000 \end{bmatrix}.$$

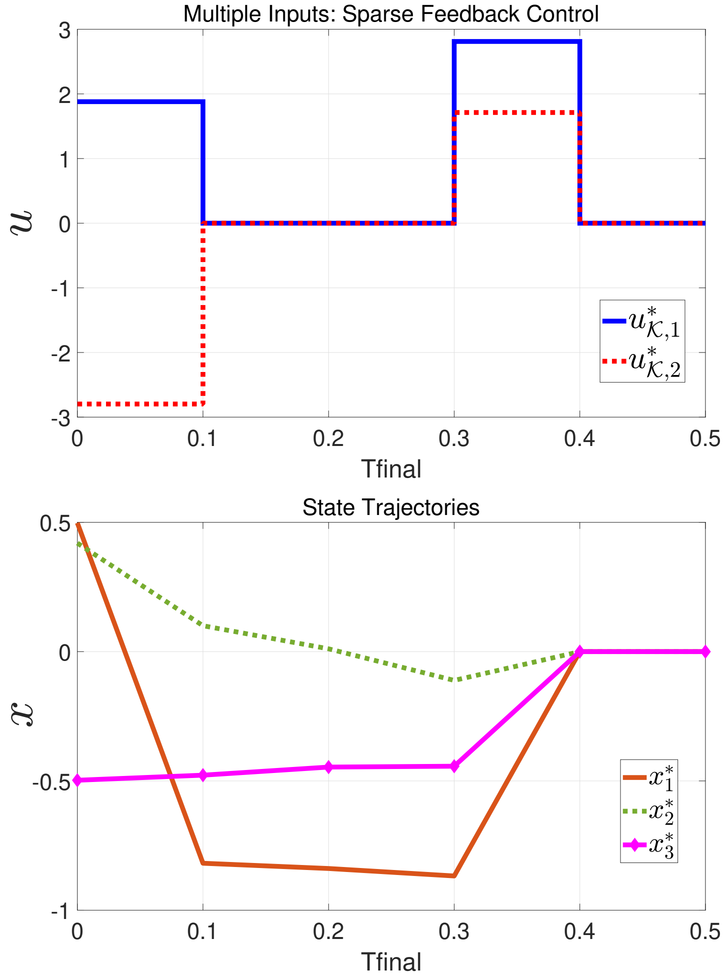


Figure 2.5: Multiple control inputs case: optimal sparse feedback control inputs $u^*_{\mathcal{K}}(t)$ (top) and the corresponding optimal state trajectories $x^*(t)$ (bottom).

As shown in Figure 2.5, the optimal feedback control signals contain two components, where both control inputs are along the input constraints $|u_{\mathcal{K},i}(t)| \leq 10, \forall i = 1, 2$. Clearly, the inferred feedback control sequences are sparse as desired. From this figure, it appears that the optimal state trajectories converge to zeros with minimum/sparse control effort.

2.6.3 Tracking problem

Finally, we show a numerical example to illustrate the effectiveness of our extended dynamic tracking compensator (2.27) for tracking problem, in Section 2.5. By taking a continuous second-order harmonic oscillator as

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} -2 \\ 1 \end{bmatrix} u(t), \qquad x(0) = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix},$$
$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t),$$

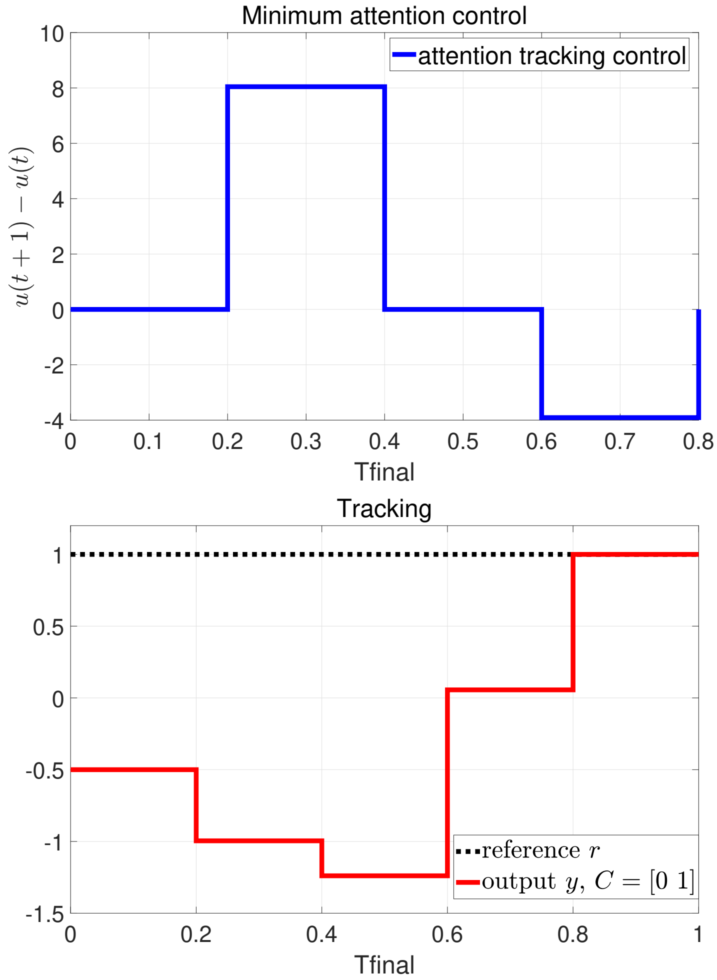


Figure 2.6: Tracking Minimum Attention Control (i.e., the difference of control input) (top) and the tracking trajectories of performance y with respect to a step reference r (bottom).

we then discretize the plant under a ZoH sampling and set the time period $\Delta t = 0.2\ [s]$ for time horizon $N = 5$. For the sake of simplicity, a step reference signal (2.25) is modeled as $r_+ = 1$ for all $t \geq 0$. In this case, the default value of steady-state of reference is $r_+ = r_0 = 1$.

Solving Problem 2.6, the optimal solution (X, U) is determined as follows

$$X = \begin{bmatrix} 1.0000 & 0 & -0.8758 & -2.1177 & -2.5880 & -4.0772 & -1.2121 & -1.9096 \\ 0 & 1.0000 & 1.0001 & 1.9972 & 1.6427 & 2.5880 & 0.6924 & 1.0908 \end{bmatrix},$$
$$U = \begin{bmatrix} 4.9553 & 6.0613 & 4.9553 & 6.0613 & -2.8673 & -4.5173 & -2.8673 & -4.5173 \end{bmatrix}.$$

We then make use of the feedback realization technique to derive the feedback matrices (K, H, G, F), where the relevant numerical results are given as

$$
\begin{aligned}
K &= \begin{bmatrix} 4.9553 & 6.0613 \end{bmatrix}, \\
H &= \begin{bmatrix} 3.2331 & 4.4499 & 0.0000 & 0.0000 & -1.0576 & -1.6662 \end{bmatrix} \\
G &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}^\top, \\
F &= \begin{bmatrix} 0.8758 & 2.1177 & 2.5880 & 4.0772 & 1.2121 & 1.9096 \\ -1.0001 & -1.9972 & -1.6427 & -2.5880 & -0.6924 & -1.0908 \\ 1.0000 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.0000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0000 & 0 & 0 \end{bmatrix}.
\end{aligned}
\tag{2.33}
$$

The problem at hand is to achieve tracking, we found that the determinants $\det(I - (A + BK)) = \det(I - X_1) = 0.2475 \neq 0$ and $\det(I - F) = 2.4902 \neq 0$, fitting two conditions in Lemma 2.2, and the related feedforward gains can be selected as

$$M = -3.0837, \quad L = \begin{bmatrix} 0.5 & -1 & 0 & 0 & 0 & 0 \end{bmatrix}^\top.$$

Based on the above analysis, the implemented dynamic tracking compensator

$$u^*_{\mathcal{K}_r} = \begin{bmatrix} -3.6367 & -3.6367 & 4.4087 & 4.4087 & 0.5000 & 0.5000 \end{bmatrix}$$

can be easily applied to the discrete LTI plant so as to achieve tracking. Figure 2.6 depicts the evolution of minimum attention tracking control (top) and the corresponding tracking trajectories. It can be see that the performance output signal $y(t)$ gradually tracks a step reference signal $r(t) = 1$ under the specified time steps.

2.7 Summary

In this chapter, we have proposed a sparse feedback controller from open-loop solution to closed-loop realization. By means of implementing a dynamic linear compensator, the stability, optimality, and sparsity of the closed-loop ℓ_1 optimal control are ensured. Besides, we extended the result to a tracking problem to achieve the minimum attention control. Finally, the simulations illustrated the effectiveness of the proposed sparse feedback control.

$$X = \begin{bmatrix} 1.0000 & 0 & 0 & 1.1753 & -4.2597 & -0.7724 & 1.2898 & -4.5326 & -0.8472 & 1.4109 & -4.8391 & -0.9265 \\ 0 & 1.0000 & 0 & 0.0443 & 0.1526 & -0.0283 & 0.1183 & -0.2060 & -0.0772 & 0.2406 & -0.7377 & -0.1578 \\ 0 & 0 & 1.0000 & 0.1322 & -1.4008 & -0.0892 & 0.1830 & -1.4257 & -0.1222 & 0.2551 & -1.5589 & -0.1692 \end{bmatrix},$$
$$U = \begin{bmatrix} 0.8012 & -7.2752 & -9.1168 & 0.0000 & -0.0000 & -0.0000 & -0.0000 & 0.0000 & 0.0000 & -0.0626 & 6.8408 & 0.0578 \\ 0.1140 & -8.1530 & -1.1419 & -0.0000 & 0.0000 & 0.0000 & -0.0000 & 0.0000 & 0.0000 & -2.9457 & 9.8661 & 1.9337 \end{bmatrix}. \quad \text{(A.1)}$$

$$K = \begin{bmatrix} 0.8012 & -7.2752 & -9.1168 \\ 0.1140 & -8.1530 & -1.1419 \end{bmatrix},$$
$$H = \begin{bmatrix} 0.5857 & -8.2481 & -0.4003 & 1.4963 & -10.8652 & -0.9970 & 2.8829 & -8.8614 & -1.8905 \\ 0.3782 & 0.1298 & -0.2447 & 1.0268 & -2.7911 & -0.6724 & -0.8534 & 2.6232 & 0.5596 \end{bmatrix}. \quad \text{(A.2)}$$

Chapter 3

Probabilistic Robustness Guarantees for Sparse Optimal Control

In this chapter, we are interested in sparse optimization with uncertainty that optimizes the sparse decision-making subject to uncertain constraints. In order to alleviate the conservative solution for worst-case robust counterpart, a chance constrained representation for uncertainty is modeled that a small violation of the robust feasibility is allowed governing by a user-defined risk level. We thus propose a well-defined chance constrained sparse optimization problem framework, which can not only measure the sparse cost performance but also evaluate the risk of constraints violation. To make problem computationally tractable, a data-driven sampling method called scenario approach is exploited to generate a randomized solution by solving a sparse random convex program. A high confidence probability for the approximated solution is attained based on a finite sample guarantee. Furthermore, we make a trade-off bridge between the sparse cost performance and the risk level by relaxing the constraint violations. Finally, the theoretical results are applied to the sparse robust control design by performing numerical simulations.

3.1 Sparse decision-making with uncertainty

Let us start from a general sparse optimization problem subject to uncertainty, that is, sparse decision making with uncertainty, which handles the (worst-case) robust sparse programs of the following form:

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & \|x\|_0 \\ \text{s.t.} \quad & x \in \mathcal{X}_\delta, \quad \forall \delta \in \Delta, \end{aligned} \tag{3.1}$$

where the exact sparse cost function is penalized by ℓ_0 "norm" that counts the number of nonzero elements of the decision variable $x \in \mathcal{X} \subseteq \mathbb{R}^n$, and to every uncertain parameter $\delta \in \Delta \subseteq \mathbb{R}^{n_\delta}$ there is associated a robust counterpart set $\mathcal{X}_\delta$. Here $\mathcal{X}$ and $\mathcal{X}_\delta$ are convex and closed sets. Oftentimes, the set Δ contains infinite cardinality. For any robust feasibility $\mathcal{X}_\delta$, $\delta \in \Delta$ is regarded as an additional constraint to the nominal or deterministic sparse program $\min_{x \in \mathcal{X}} \ \|x\|_0$. Therefore, if the solution of program (3.1) is accessible, then it is safeguarded against all the possible uncertainties from the nominal sparse decision-making.

In practical applications, enforcing uncertain constraints $\mathcal{X}_\delta$ for *all* possible $\delta \in \Delta$ may be overly conservative and overkill, and violating a small fraction of uncertain constraints over a small subset Δ_ϵ^c of Δ is allowed and accepted in general, referred to chance-constrained program [45] or stochastic programming [122]. In other words, we leverage a probabilistic or stochastic model to relax the worst-case uncertainty based on the following assumptions for the uncertainty, where Δ is endowed with a probability $\mathbb{P}$, and this $\mathbb{P}$ can be given various distributions, depending on the users, see Figure 3.2.

Assumption 3.1 (Probability Space) *Assume that $\delta \in \Delta \subseteq \mathbb{R}^{n_\delta}$ be a random variable defined on a probability space $(\Delta, \mathcal{F}, \mathbb{P})$, where Δ is a metric space with respect to Borel σ-algebra $\mathcal{F}$ and the probability distribution $\mathbb{P}$, which measures the chance of a constraint set $\mathcal{X}_\delta$ to occur.*

We thus weaken worst-case robust sparse program (3.1) based on Assumption 3.1 that a subset Δ_ϵ of Δ admits a small constraints violation, giving rise to the following *chance-constrained sparse optimization* problem:

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & \|x\|_0 \\ \text{s.t.} \quad & \mathbb{P}\left\{\delta \in \Delta : x \in \mathcal{X}_\delta\right\} \geq 1 - \epsilon, \end{aligned} \tag{3.2}$$

where δ is random vector whose probability distribution is supported on set $\delta \in \Delta$, and $\epsilon \in (0, 1)$ is a given risk (or accuracy) parameter, also known as violation of probability or significance level. Obviously, such a chance-constrained problem setup implies that the reliability $x \in \mathcal{X}_\delta$ occurs with probability at least $1 - \epsilon$.

We introduce the following definition of risk.

Definition 3.1 (Violation of probability) Given a decision variable x, the probability of violation (i.e., risk) is defined as

$$V(x) \triangleq \mathbb{P}\big\{\delta \in \Delta : x \notin \mathcal{X}_\delta\big\}. \tag{3.3}$$

Note that $V(x)$ is the probability with which the constraint is not satisfied by the decision variable x. We say that x is an *ϵ-level probabilistic robustness* if the violation probability holds that $V(x) \leq \epsilon$.

Chance-constrained sparse optimization is now well-defined, while the program (3.2) is a *non-convex* optimization, leading to NP-hard problem [92]. Due to the fact that both the cost function and chance constraints keep non-convex structure to measure the sparsity and probability, respectively.

In light of the above challenges, existing methods for chance-constrained programs can be solved by convex approximation of non-convex chance constraint utilizing randomized algorithms via sampling method [126]. Thus, the produced solutions are feasible, or at least with a high-confidence guarantee on feasibility, to the original problem.

Assumption 3.2 (Data-driven Sampling) *Suppose that the experimenter can observe a finite samples $\delta^{(1)}, \ldots, \delta^{(N)}$ of random variable δ, which is randomly independent and identically distributed (i.i.d.) generated from constraint sets $\{\mathcal{X}_{\delta^{(i)}}\}_{i=1}^{N}$, according to probability $\mathbb{P}^N$.*

The principle of sampling technique claims that the uncertain constraint set $\mathcal{X}_\delta$ can be approximated by a random problem with finitely many constraints $\{\mathcal{X}_{\delta^{(i)}}\}_{i=1}^{N}$. More precisely, the random variable δ is replaced by a finite scenarios $\{\delta^{(1)}, \ldots, \delta^{(N)}\}$, to let the uncertain data speak. Such a paradigm is known as *data-driven robust optimization*, encompassing methodologies like Monte-Carlo simulations, sample average approximation [91], and scenario approach [37], etc.

Hence, the chance-constrained sparse optimization (3.2) can be further approximated as the following *sceanrio-based sparse program*:

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & \|x\|_0 \\ \text{s.t.} \quad & x \in \bigcap_{i=1}^{N} \mathcal{X}_{\delta^{(i)}}, \end{aligned} \tag{3.4}$$

where each element $\mathcal{X}_{\delta^{(i)}}$ of the finite sample is so-called *scenario*, and thus program (3.4) is named as *scenario-based sparse program* or *data-driven sparse optimization*.

As a matter of fact, checking robust feasibility is closely related to uncertainty quantification, we refer to see Section 3.2 and reference therein for more details.

3.2 Uncertainty quantification

In general, there exist lots of momentums on decision-making under uncertainty [14, 110, 126, 122]. We now itemize some examples for uncertainty quantification of robustness analysis. To concretizately the uncertainty evaluation mentioned in Section 3.1, that the robust feasibility problem can be further characterized as

$$\text{find } x \quad \text{s.t.} \quad h(x,\delta) \le 0, \qquad \forall \delta \in \Delta, \tag{3.5}$$

where uncertain constraint function $h(x,\delta) : \mathcal{X} \times \Delta \to \mathbb{R}$ is a scalar-valued[1], and it is *convex* in the first argument x and *bounded* in the second argument for $\delta \in \Delta$. Without loss of generality, throughout this chapter, we denote the measurable reliability set by

$$x \in \mathcal{X}_\delta \triangleq \{\delta : h(x,\delta) \le 0\},$$

[1]If the constraint mapping $h(x,\delta) : \mathcal{X} \times \Delta \to \mathbb{R}^m$, then there exist a number of constraints $h_j(x,\delta) \le 0$, $j = 1, \cdots, m$, which can be equivalently replaced by one constraint $h(x,\delta) \triangleq \max_{1<j<m} h_j(x,\delta) \le 0$.

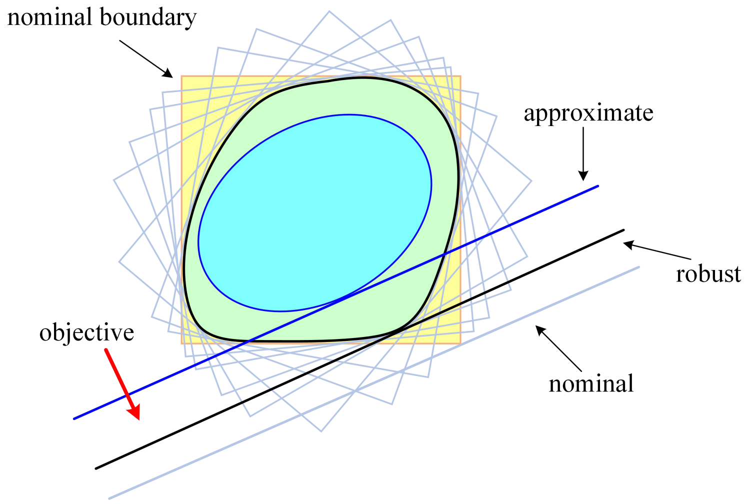


Figure 3.1: Robust feasibility regions: A nominal boundary (the field of yellow square) gives the nominal or deterministic solution. When encountering with the uncertainty, the robust counterpart (the area of light green) involves the infinite constraints (a bunch of light blue squares), leading to the conservative robust solution. Alternatively, an inner approximation (the cyan ellipsoid) returns the robust approximate solution.

and its complementary (i.e., violation, see Definition 3.1) is given as

$$x \notin \mathcal{X}_\delta \quad \Leftrightarrow \quad x \in \mathcal{X}_\delta^c \triangleq \{\delta : h(x,\delta) > 0\}\,.$$

Example 3.1 (Worst Case) A standard uncertainty description in (3.5) is *worst-case*, for which the uncertain constraint satisfaction is enforced for *all* possible realizations $\delta \in \Delta$

$$h(x,\delta) \leq 0, \quad \forall \delta \in \Delta. \tag{3.6}$$

Here the term (3.6) is so-called *robust counterpart* [14]. Indeed, the constraint (3.6) is *semi-infinite program* since it contains infinite number of constraints, namely, the cardinality $\text{card}(\Delta)$ is uncountable, and hence it is hard to tackle expect for some special structures for uncertainty set, like interval or box constraints [89], see Figure 3.1.

Alternatively, for a given nonempty, compact set Δ satisfies a regularity condition[2], then the robust inequality (3.6) can be written as a *finite* and *deterministic* form

$$\max_{\delta \in \Delta} h(x,\delta) \leq 0. \tag{3.7}$$

On the other hand, it has been proved that if the uncertain constraint $h(x,\cdot)$ is concave for all $x \in \mathbb{R}^n$, then the worst-case constraint (3.7) is equivalent to [13]

$$\max_{\delta \in \Delta} h(u,\delta) = \sup_{\delta \in \mathbb{R}^{n_\delta}} \left\{h(x,\delta) - \sigma(\nu|\Delta)\right\} = \inf_{\nu \in \mathbb{R}^{n_\delta}} \left\{\sigma^\star(\nu|\Delta) - h_\star(x,\nu)\right\}$$

[2]The regularity condition means that $\text{ri}(\Delta) \cup \text{ri}(\text{dom}(h(x,\cdot)) \neq \emptyset$, $\forall x \in \mathbb{R}^n$, where $\text{ri}(\Delta)$ is the relative interior of set Δ.

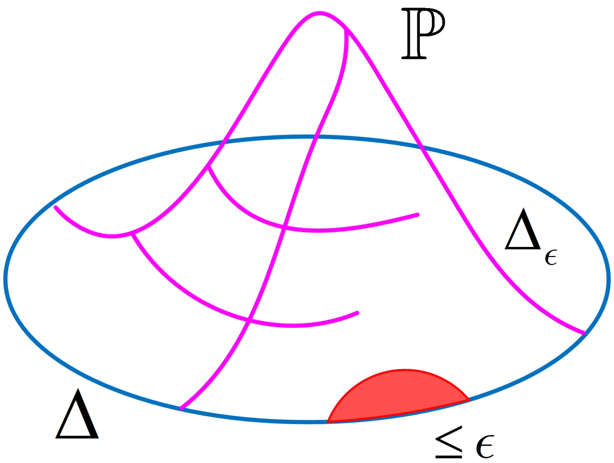


Figure 3.2: Probabilistic or chance constrained framework for the uncertainty Δ over the probability $\mathbb{P}$, where the red region represents a small violated subset of Δ with a prescribed risk level $\epsilon \in (0, 1)$.

where the first equality holds from the property of the characteristic function

$$\sigma(v|\Delta) \triangleq \begin{cases} 0, & \text{if } \delta \in \Delta, \\ +\infty, & \text{otherwise,} \end{cases} \tag{3.8}$$

and the last passage holds according to Fenchel's duality theorem [118]. Here $h_\star(x, v)$ is the *partial concave conjugate* of $h(x, v)$ and $\sigma^\star(v|\Delta)$ is the *support function* of the set Δ (i.e., the conjugate of characteristic function $\sigma(\delta|\Delta)$ in (3.8)), defined respectively, as

$$h_\star(x, v) = \inf_{v \in \mathbb{R}^{n_\delta}} \left\{ v^\top \delta - h(x, \delta) \right\}, \quad \sigma^\star(v|\Delta) = \sup_{v \in \mathbb{R}^{n_\delta}} v^\top \delta. \tag{3.9}$$

However, computing the support function $\sigma^\star$ requires the knowledge of structure setup for uncertainty set Δ. For instance, a common uncertainty set is selected as

$$\Delta = \left\{ \delta \in \Delta : \delta = \delta_0 + \sum_{i=1}^{n_q} G_i q_i, \ q \in \mathcal{Q} \subseteq \mathbb{R}^{n_q}, \ G \in \mathbb{R}^{n_\delta \times n_q} \right\}, \tag{3.10}$$

where $\delta_0 \in \mathbb{R}^{n_\delta}$ is "nominal" vector satisfying regularity, that is, $\delta_0 \in \text{ri}(\text{dom } h(x, \cdot))$, $\forall x$, and the matrix $G \in \mathbb{R}^{n_\delta \times n_q}$ is a given column-wise $G = \begin{bmatrix} G_1 & \dots & G_{n_q} \end{bmatrix}$, q is the vector of primitive uncertainties.

To get rid of the worst-case and stringent constraint, it is attractive to treat the robust counterpart (3.6) as *chance constraint* rather than as deterministic constraint. Therefore we convert the deterministic uncertainty into a probabilistic or chance constrained (cf., Example 3.2) or stochastic modeling (cf., Example 3.3). Notice that in this chapter, we do *not* require any structure assumption for uncertainty set Δ, like the type (3.10).

Example 3.2 (Chance Constraint) A probabilistic relaxation for uncertainty (3.6) can be modeled as the *random variables* δ governed by the probability $\mathbb{P}$ (i.e., Assumption 3.1). Formally, one should ascertain the following *safety* (resp., *violation*) event

$$\mathbb{P}\{\delta : h(x,\delta) \leq 0\} \geq 1-\epsilon \quad \Leftrightarrow \quad \mathbb{P}\{\delta : h(x,\delta) > 0\} \leq \epsilon, \tag{3.11}$$

where ϵ assures that the violation event $h(x,\delta) > 0$ occurring does not exceed a *risk* level $\epsilon \in (0,1)$. Figure 3.2 illustrates the probabilistic model for uncertainty set Δ. As shown in [92], chance-constrained program (3.11) is an NP hard problem. Besides, checking the feasibility of (3.11) is general difficult either analytically or numerically due to three major reasons:

(i) Firstly, the feasible region together with the probabilistic model is usually *non-convex* even under the convexity assumption for uncertain function $h(x,\cdot)$ other than some special structure cases, e.g., closed form.

(ii) Secondly, the chance constraint involves the multiple-integral computing.

(iii) Lastly, the underlying knowledge of probability distribution $\mathbb{P}$ is typically *unknown*, or partially known.

Example 3.3 (Stochastic Constraint) Another perspective of probabilistic description for uncertainty (3.6) is stochastic modeling, which leads to an average or empirical uncertainty quantification, i.e.,

$$\mathbb{E}_{\mathbb{P}}\left[h(x,\delta)\right] = \int_{\Delta} h(x,\delta)\mathbb{P}(d\delta), \tag{3.12}$$

where $\mathbb{E}_{\mathbb{P}}[\cdot]$ denotes the expectation operator with respect to the uncertainty Δ over distribution $\mathbb{P}$. As a matter a fact, we define an indicator function associated the set

$$\mathbb{P}\{\delta : h(x,\delta) > 0\} = \mathbb{E}\left[\mathbb{1}_{\{h(x,\delta)>0\}}\right].$$

Recently, significant attention has been directed towards a distributionally robust chance constraint framework for quantifying uncertainty. This concept is introduced through the following illustrative examples.

Example 3.4 (Distributionally Robust Chance Constraint) A more general case is the extension of the chance constraints (3.11), called *ambiguous* or *distributionally robust chance constraints* (DRCC) [99] to evaluate the *worst-case of the violation event* under a low certificate level ϵ such that

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\{h(x,\delta) > 0\} < \epsilon \quad \Leftrightarrow \quad \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\{h(x,\delta) \leq 0\} \geq 1-\epsilon, \tag{3.13}$$

where the probability $\mathbb{P}$ is taken over an ambiguity set $\mathcal{P}$, i.e., *a family of probability distributions* that can be defined by moment ambiguity set (moment and variance) [47, 152], divergence ambiguity set (entropy) [56], and Wassersetin ambiguity set [20, 100], etc.

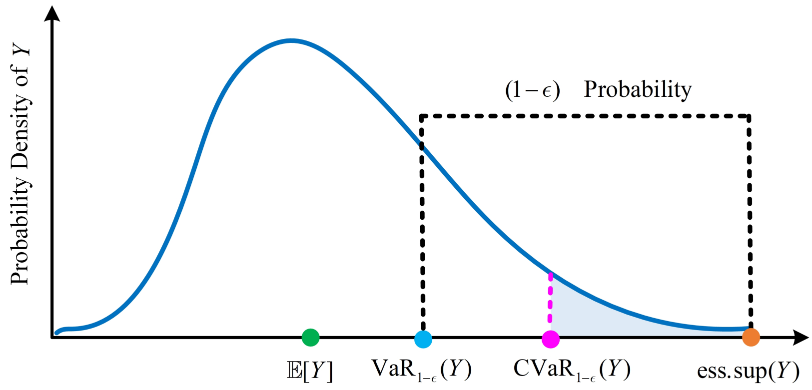


Figure 3.3: Conditional Value at risk [121].

Example 3.5 (Value at Risk) So far we have shown the chance constraint (3.11) in Example 3.2. An analogous statement is that it is able to develop as a *value at risk* (VaR) case [122, Chapter 4], yielding "percentile" optimization

$$\begin{aligned} \mathrm{VaR}_{1-\epsilon}(h(x,\delta)) &= \sup_{\gamma\in\mathbb{R}} \{\gamma : \mathbb{P}\{h(x,\delta) \leq \gamma\} \geq 1-\epsilon\} \\ &= \inf_{\gamma\in\mathbb{R}} \{\gamma : \mathbb{P}\{h(x,\delta) > \gamma\} \leq \epsilon\}, \end{aligned} \tag{3.14}$$

where VaR is positively homogenous in u, but is non-convex. Put differently, VaR minimizes one performance level $\gamma \in \mathbb{R}$ (percentile) that evaluates a "soft" violation event of term (3.11) defined by $\{h(x,\delta) > \gamma\}$ (i.e., make a relaxation from "0" to "γ") with a probability not larger than the risk level $\epsilon \in (0,1)$.

However, the true probability distribution $\mathbb{P}$ outlined above is usually inaccessible, which enlightens us about ambiguity set, leading to the worst case distributionally VaR [138] is

$$\begin{aligned} &\sup_{\mathbb{P}\in\mathcal{P}} \mathrm{VaR}_{\epsilon}(h(x,\delta)) \\ &= \inf_{\gamma\in\mathbb{R}} \left\{\gamma : \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\{h(x,\delta) > \gamma\} \leq \epsilon\right\}. \end{aligned} \tag{3.15}$$

Suppose that the uncertainty set Δ is non-empty, convex and compact. For every $\nu \in \mathbb{R}^{n_\delta}$, if the support function (3.9) and VaR function satisfy

$$\sigma^\star(\nu|\Delta) \geq \mathrm{VaR}_{1-\epsilon}(\nu^\top\delta), \quad \forall \nu \in \mathbb{R}^{n_\delta}, \tag{3.16}$$

then, the set Δ implies a probabilistic guarantee at a risk level $\epsilon \in (0,1)$ for the "true" probability distribution $\mathbb{P}$.

Example 3.6 (Conditional VaR) Different from the chance constraint description (3.11) and VaR (3.14) applying probabilistic form, another method is to use *stochastic programming* modeling [122] (i.e., take expected value of random variables, see Example 3.3), resulting in the *conditional value at risk* (CVaR) is as follows [119]

$$\begin{aligned} \mathrm{CVaR}_{1-\epsilon}(h(x,\delta)) &= \inf_{\tau\in\mathbb{R}} \left\{ \tau + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}\left[\left(h(x,\delta)-\tau\right)^{+} \right] \right\} \\ &= \inf_{\tau\in\mathbb{R}} \left\{ \tau\varepsilon + \mathbb{E}\left[(h(x,\delta)-\tau)^{+} \right] \right\}, \end{aligned} \tag{3.17}$$

where $(a)^{+} = \max\{a, 0\}$.

It further gives the worst case distributionally CVaR [152] as follows

$$\begin{aligned} &\sup_{\mathbb{P}\in\mathcal{P}} \mathrm{CVaR}_{1-\epsilon}(h(x,\delta)) \\ &= \sup_{\mathbb{P}\in\mathcal{P}} \inf_{\tau\in\mathbb{R}} \left\{ \tau + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}\left[(h(x,\delta)-\tau)^{+} \right] \right\} \\ &= \inf_{\tau\in\mathbb{R}} \left\{ \tau + \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}\left[\frac{(h(x,\delta)-\tau)^{+}}{\epsilon} \right] \right\} \end{aligned} \tag{3.18}$$

the last equality holds according to the min-max theory of semi-infinite program [123, Proposition 2.3].

Remark 3.1 As a consequence of uncertainty analysis, we remark that for any (continuous) random variable Y, the following revelation can be derived from [119], i.e.,

$$\begin{aligned} \mathrm{ess.\,sup}(Y) :&= \mathrm{CVaR}_{1}(Y) \\ &\geq \mathrm{CVaR}_{1-\epsilon}(Y) \qquad \big(= \mathbb{E}_{\mathbb{P}}[Y|\mathrm{VaR}_{1-\epsilon}(Y) \geq Y] \big) \\ &= \mathrm{VaR}_{1-\epsilon}(Y) + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}\left[(Y - \mathrm{VaR}_{1-\epsilon}(Y))^{+} \right] \\ &\geq \mathrm{VaR}_{1-\epsilon}(Y) \qquad \big(\Leftrightarrow \mathbb{P}(Y \leq 0) \geq 1-\epsilon \big) \end{aligned}$$

and hence it is easy to show some *inner* and *outer* approximations for chance constraint form (3.11), and these relationships (cf., Figure 3.3) can be exploited to diverse distributionally robust programs, such as, moment and metric ambiguity sets [99, 138, 152].

3.3 Sparse random convex program

In this section, we reduce the "non-convex" chance-constrained sparse optimization model (3.2) to a *sparse random convex program*, which is essentially based on the "convex relaxation" and "data-driven sampling" methods, respectively.

More precisely, for program (3.2), we replace an exact non-convex ℓ_0 norm for decision variable x by a convex ℓ_1 norm, optimizing the following program

$$\begin{aligned} \min_{x\in\mathcal{X}} \quad & \|x\|_1 \\ \text{s.t.} \quad & \mathbb{P}\left\{ \delta \in \Delta : x \in \mathcal{X}_{\delta} \right\} \geq 1-\epsilon. \end{aligned} \tag{3.19}$$

Let x^*_ϵ denote the optimal solution to program (3.19) if it is available.

On the other hand, the uncertain parameter δ is modeled as the random variable satisfying Assumption 3.2, namely, i.i.d. draws N "scenarios" $\delta^{(1)}, \ldots, \delta^{(N)}$ from the probability $\mathbb{P}$. We propose the following *sparse random convex program*, formulated as

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & \|x\|_1 \\ \text{s.t.} \quad & h(x, \delta^{(i)}) \leq 0, \quad i = 1, \ldots, N. \end{aligned} \tag{3.20}$$

Obviously, program (3.20) is an ℓ_1 norm convex relaxation for *scenario-based sparse optimization problem* (3.4). The optimal solution (resp., optimal value) to program (3.20), denoted by x_N^* (resp., J_N^*), is a *data-driven, randomized, sparse solution* based on the N random observations or scenarios.

Assumption 3.3 (Existence and Uniqueness) For every N and for every multisample $(\delta^{(1)}, \cdots, \delta^{(N)})$, the optimal solution x_N^* of program (3.20) exists and is unique.

Note that we here assume that the optimal solution x_N^* to program (3.20) is unique. Otherwise, if program (3.20) more than one optimal solution exists, we can ensure the uniqueness of solution x_N^* after the implementation of the tie-breaking rule [35, 120].

Assumption 3.4 (Non-accumulation) For every $x \in \mathcal{X}$, $\mathbb{P}\{\delta : h(x, \delta) = 0\} = 0$.

Furthermore, for $k = 1, \cdots, N$, program (3.20) induces a new program by discarding the k-th constraint, which gives a "removal" sparse random convex program as

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & \|x\|_1 \\ \text{s.t.} \quad & h(x, \delta^{(i)}) \leq 0, \quad i = \{1, \ldots, N\} \setminus k, \end{aligned} \tag{3.21}$$

where the optimal solution $x_{N,k}^*$ is obtained from program (3.21) associated with an optimal value J_k^*.

Definition 3.2 (Support Constraint) Let J_N^* and J_k^* be the optimal (objective) value of sparse random convex program (3.20) and (3.21), respectively. The k-th constraint in (3.20) is said to be a *support constraint* if $J_k^* < J_N^*$.

It is clearly that when a support constraint k is removed from the sample set $\mathcal{N} = \{1, \cdots, N\}$, then the optimal solution $x_{N,k}^*$ would be improved as well since $J_k^* < J_N^*$. In other words, a constraint $h(x, \delta^{(k)})$, $k = 1, \ldots, N$ of sparse random convex program (3.21) is a support constraint if its removes improves the solution of program (3.20). For simplicity, denote the support constraints set of program (3.20) as $\mathcal{S}$, where $\mathcal{S} \subset \mathcal{N}$. As mentioned in Levin [82]:

"*the number of support scenarios for convex program is at most n, the number of decision variables x.*"

Therefore, no matter what the structure of problem is or how large the sample complexity N is, the cardinality of support constraints set for (3.20) is less than or equal to n, the size of decision variable x [35], that is to say, with probability 1, we have

$$|\mathcal{S}| \leq \dim(x) = n.$$

In particular, we call program (3.20) is a *fully-supported*, if $s_N^* = |\mathcal{S}| = n$ [38]. The program (3.20) is said to be *non-degenerate*, if $J_{\mathcal{N}}^* = J_{\mathcal{S}}^*$, or equivalently, $\|x_{\mathcal{N}}^*\|_1 = \|x_{\mathcal{S}}^*\|_1$.

Lemma 3.1 (A Priori Probabilistic Guarantee [38]) *Under Assumptions 3.1-3.3, given a risk level $\epsilon \in (0,1)$ and the optimal solution x_N^* to program (3.20), it holds that*

$$\mathbb{P}^N\{V(x_N^*) > \epsilon\} \leq \sum_{i=1}^{n-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}. \tag{3.22}$$

For a given confidence parameter $\beta \in (0,1)$, we choose the sample complexity N such that

$$\sum_{i=1}^{n-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta. \tag{3.23}$$

Then, the optimal solution x_N^* is feasible to program (3.19) with a high probability, that is,

$$\mathbb{P}^N\{V(x_N^*) \leq \epsilon\} \geq 1 - \beta, \tag{3.24}$$

where the probability $\mathbb{P}^N$ is taken with respect to N observed random samples $\{\delta^{(i)}\}_{i=1}^N$.

The program (3.20) typically provides *a priori* probabilistic robustness for the randomized sparse solution x_N^*, and the corresponding feasibility of solution x_N^* is guaranteed *before* obtaining x_N^*. From the relation between the chance-constrained sparse convex program (3.19) and sparse random convex program (3.20), it is clear that the approximated solution u_N^* of program (3.20) is also a feasible solution to program (3.19), with a high confidence $1 - \beta$.

Notice that the left-hand-side term in (3.23) tends to zero as the N increases. The implies that, for a large number scenarios N, the likelihood of obtaining an unfriendly sparse solution x_N^* with a violation probability $V(x_N^*)$ exceeding a risk threshold ϵ is relatively low.

From the estimation of randomized sparse solution x_N^* perspective, we require an explicit expression of (3.23) with respect to N. In fact, Lemma 3.1 can provide a finite-ample guarantee for the obtained solution x_N^*. More specifically, for a given risk level $\epsilon \in (0,1)$, confidence level $\beta \in (0,1)$, and the size of decision variable $\dim(x) = n$, then the sample complexity N satisfies the following lower bound [38]

$$N \geq \frac{2}{\epsilon}\left(\ln\frac{1}{\beta} + n - 1\right), \tag{3.25}$$

then the sparse random solution x_N^* is ϵ-level probabilistically robust with probability at least $1 - \beta$.

Remark 3.2 (Sample complexity) Naturally, employing various probability inequalities allows us to derive distinct (precise) sample complexities. For a comprehensive understanding of sample complexity, further insights are available in references [4, 36, 37, 43, 62, 126].

In particular, Calafiore [34] examined the Helly's dimension case of scenario program that for a smallest integer ζ

$$\operatorname{ess\,sup}_{\delta\in\Delta^N} |\mathcal{S}(\delta)| \le \zeta < \bar{\zeta} \tag{3.26}$$

holds for any finite $N \ge 1$, where $\mathcal{S}$ is the support constraint scenarios. Therefore, the similar theoretical guarantees on sample complexity can be derived easily by replacing n with Helly's dimension h in (3.25). However, Helly's dimension is usually hard to compute, while identifying the upper bounds $\bar{\zeta}$ on Helly's dimension ζ is much easier to calculate. Moreover, when uncertain function $h(x,\delta)$ enjoys structured dependence on the uncertainty [141], a lower bound for Helly's dimension ζ can be determined, which that further quantifies the bound for the number of decision variable $n = \dim(x)$, and it is useful to reduce the sample complexity N.

Based on these, we conclude a procedure for a priori feasibility guarantee on general scenario/random convex optimization (three steps):

1. Exploring the problem structure of program and obtain the number of support scenarios $|\mathcal{S}| = s_N^*$;
2. Select the sample complexity N using the parameter pair (ϵ, β, n)
3. Obtain optimal solution x_N^* and optimal value J_N^*.

3.4 Sparse random convex program with relaxation

An intriguing and profound question arises: How can one effectively balance the *sparse cost performance* and the associated *violation of constraints* in program (3.20) ? Motivated by the scenario program having a linear objective function [60, Section 4], it suggests us exploring a flexible paradigm that minimizes the cost as well as the additional "soft" constraints for program (3.20).

Adapted from program (3.20), this section focuses on risk assessment for *sparse random convex optimization with relaxation*, giving rise to a *relaxed* scenario-based sparse convex optimization problem, described by

$$\begin{aligned}
\min_{\substack{x\in\mathcal{X}\\ \xi^i,\ i=1,\dots,N}} \quad & \|x\|_1 + \rho \sum_{i=1}^{N} \xi^i \\
\text{s.t.} \quad & h(x,\delta^{(i)}) \le \xi^i, \quad i = 1,\dots,N, \\
& \xi^i \ge 0,
\end{aligned} \tag{3.27}$$

where $\xi^i \geq 0$ are the "relaxed" slack variables that indicate the maximum constraint violation over all possible uncertain parameters $\delta^{(i)}$, and the allowed violations ξ^i's are penalized to the original cost function.

This moment program (3.27) has $n + N$ decision variables, including $x \in \mathbb{R}^n$ and $\xi = [\xi^1, \ldots, \xi^N]^\top \in \mathbb{R}^N$. The interpretation of program 3.27 is that some constraints $f(x, \delta^{(i)})) \leq 0$ can be violated for the purpose purpose of improving the cost value, but constraints violation has itself a cost as expressed by the auxiliary optimization variables ξ^i: if $\xi^i \geq 0$, then constraint $f(x, \delta^{(i)}) \leq 0$ is relaxed to $f(x, \delta^{(i)}) \leq \xi^i$ and this generates the regret ξ^i, which adds to the original cost$\|x\|_1$.

Meanwhile, the penalty weight ρ is used to achieve the trade-offs between minimizing the sparse cost and the regrets for constraint violations. For instance, $\rho \to 0$ implies the non-regrets; $\rho = 1/N$ indicates the empirical regrets; and $\rho \to \infty$ denotes the infinite regrets, then program (3.27) recovers to original program (3.20). In general, ρ must larger than the dual norm of the original Lagrange multiplier for the constraint $\max_i \; h(x, \delta^{(i)})$.

We now introduce the Lagrange multipliers $\mu_i, \lambda_i \geq 0$ for $i = 1, \cdots, N$, then the Lagrange (primal) objective function $\mathcal{L}$ for program (3.27) can be formulated as follows

$$\mathcal{L}(x, \xi, \mu, \lambda) = \|x\|_1 + \rho \sum_{i=1}^{N} \xi^i + \sum_{i=1}^{N} \mu_i \left(h(x, \delta^i) - \xi^i \right) - \sum_{i=1}^{N} \lambda_i \xi^i, \tag{3.28}$$

which we attempt to minimize with respect to decision variables x and $\xi^i \geq 0$. Taking the respective derivatives to zero, we have

$$\frac{\partial \mathcal{L}}{\partial x} = 0 \quad \Longleftrightarrow \quad \partial \|x\|_1 = \sum_{i=1}^{N} \mu_i \frac{\partial}{\partial u} h(x, \delta^{(i)})), \tag{3.29}$$

$$\frac{\partial \mathcal{L}}{\partial \xi^i} = 0 \quad \Longleftrightarrow \quad \mu_i = \rho - \lambda_i, \quad \forall i = 1, \cdots, N, \tag{3.30}$$

where μ_i, λ_i, $\xi^i \geq 0$ are all positive constraints. Notice that the ℓ_1 norm is *non-smooth* and *non-differentiable* function, while it is possible to give its *subdifferential* $\partial \|x\|_1$ as

$$\partial \|x\|_1 = \frac{\partial}{\partial x_i} \sum_{i=1}^{n} |x_i| = \{ v \mid v_i = \text{sign}(x_i), \; \forall i \in \{1, \cdots, n\}, \; \|v\|_\infty \leq 1 \}, \tag{3.31}$$

here $\text{sign}(v)$ is a sign function defined as

$$\text{sign}(x) = \begin{cases} -1, & \text{if } x < 0, \\ 0, & \text{if } x = 0, \\ 1, & \text{if } x > 0. \end{cases}$$

On the other hand, by substituting (3.29) and (3.30) into (3.28), we obtain the Lagrange Wolfe (dual) objective function $\mathfrak{L}$ as follows

$$\mathfrak{L}(\mu) = \sum_{i=1}^{N} \mu_i - \sum_{i=1}^{N} \sum_{i'=1}^{N} \mu_i \mu_{i'} \frac{\partial}{\partial u} h^\top(x, \delta^i) \frac{\partial}{\partial x} h(x, \delta^i) \tag{3.32}$$

which gives a lower bound on the objective function (3.27) for any feasible solution.

Before proceeding the relaxed sparse convex program (3.27), let us consider a similar one called *sample average approximation* problem [91]

$$\begin{aligned} \min_{u} \quad & \|x\|_1 \\ \text{s.t.} \quad & h(x, \delta^{(i)}) \leq \mathrm{M}\xi^{(i)}, \quad i = 1, \cdots, N, \\ & \sum_{i=1}^{N} \xi^i \leq \mathrm{K}, \quad \xi^i \in \{0, 1\}^N \end{aligned} \tag{3.33}$$

Here M is a positive constant greater than the maximum possible $h(x, \delta)$. The binary variables ξ^i allow the violation up to K constraints, and thus program (3.33) becomes a mixed integer program problem, which requires to solve a big-M problem [122].

Using Assumptions 3.3 and 3.4, the program (3.27) is computationally tractable, and a general theory for scenario (linear) program with relaxation (but not involves sparse cost) has been proven in [60, Theroem 4], stated as follows

Lemma 3.2 (Violation [60]) *Under Assumptions 3.3 and 3.4, consider sparse random convex program with relaxation problem (3.27), given a confidence $\beta \in (0, 1)$, the violation of probability $V(x_N^*)$ is evaluated as follows*

$$\mathbb{P}^N\{\underline{\epsilon}(s_N^*) \leq V(x_N^*) \leq \overline{\epsilon}(s_N^*)\} \geq 1 - \beta \tag{3.34}$$

where $\underline{\epsilon}(\cdot) = \max\{0, 1 - \overline{t}(k)\}$, $\overline{\epsilon}(\cdot) = \max\{0, 1 - \underline{t}(k)\}$, s_N^* is the number of samples q^i for which $h(x, \delta^{(i)}) \geq 0$ at $x = x_N^*$, $V(x) = \mathbb{P}\{\delta : f(x, \delta) > 0\}$, and for $k = 0, 1, \cdots, N-1$ the pair $\{\underline{t}(k), \overline{t}(k)\}$ is the solution of the polynomial equation in t variable

$$\mathfrak{B}_N(t; k) = \frac{\beta}{2N} \sum_{j=k}^{N-1} \mathfrak{B}_j(t; k) + \frac{\beta}{6N} \sum_{j=N+1}^{4N} \mathfrak{B}_j(t; k) \tag{3.35}$$

where $\mathfrak{B}_j(t; k) = \binom{j}{k} t^{j-k}$ represents a binomial expansion. For $k = N$, the upper bound is set to $\overline{\epsilon}(k) = 1$ and the lower bound is derived by

$$1 = \frac{\beta}{6N} \sum_{j=N+1}^{4N} \mathfrak{B}_j(t; N). \tag{3.36}$$

Lemma 3.2 establishes an interval $[\underline{\epsilon}(s_N^*), \overline{\epsilon}(s_N^*)]$ for violation $V(x_N^*)$ that lies with a high confidence level $1 - \beta$. The value of complexity s_N^* can be calculated according to the observed data samples $\delta^{(1)}, \ldots, \delta^{(N)}$, which is equivalent to the number of violated constraints plus at most n scenarios (i.e., the number of decision variables).

Fig. 3.4 depicts the curves $\underline{\epsilon}(k)$ and $\overline{\epsilon}(k)$ for risk $V(x_N^*)$ with $N = 1000$, and $\beta = 10^{-5}, 10^{-7}, 10^{-9}$, respectively.

Remark 3.3 (Risk Estimation) A common practice to estimate the risk of the sparse solution $x_{N_\mathrm{T}}^*$ based on the observations $\delta^{(i)}$, where $i \in \mathcal{N}_\mathrm{T} = \{1, \cdots, N_\mathrm{T}\}$ is the training samples, and $i \in \mathcal{N}_\mathrm{V} = \{N_\mathrm{T}+1, \cdots, N_\mathrm{T}+N_\mathrm{V}\}$ stands for the validation or out-of-sample samples, then the violation $V(x_{N_\mathrm{T}}^*)$ is estimated form the ratio

$$\frac{\#\{i,\ h(x_{N_\mathrm{T}}^*, \delta^{(i)}) > 0,\ i = N_\mathrm{T}+1, \cdots, N_\mathrm{T}+N_\mathrm{V}\}}{N_\mathrm{V}}.$$

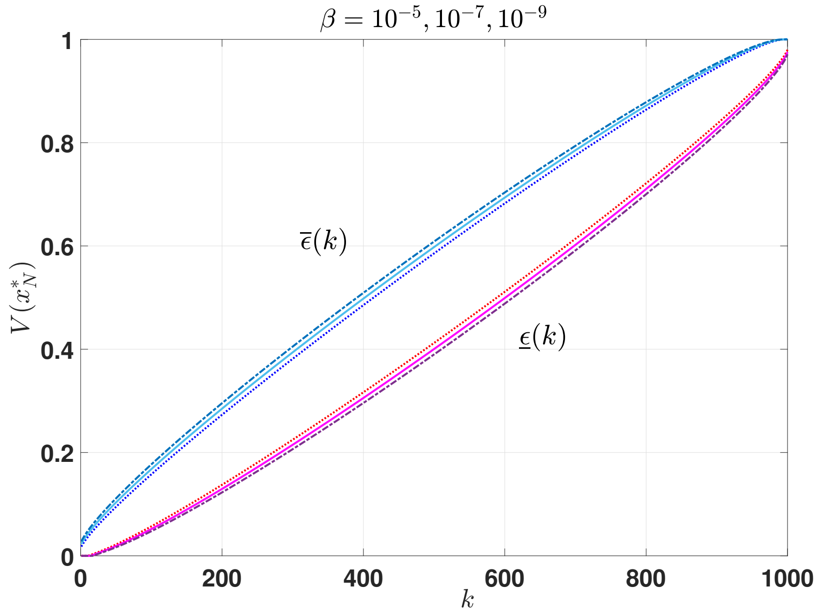


Figure 3.4: Profile of curves $\underline{\epsilon}(k)$ and $\overline{\epsilon}(k)$ for the violation of probability $V(x_N^*)$ with $N = 1000$, and $\beta = 10^{-5}, 10^{-7}, 10^{-9}$, respectively, where $k = 0, 1, \cdots, N$.

3.5 Sparse robust control applications

Motivated by these seminal works for robust control design [105, 76, 35], in this section, we consider the sparse robust control design for dealing with the uncertain discrete-time system (3.37). We here delve into the sparse robust control problems concerning discrete-time uncertain dynamical systems. All the formulations discussed herein are built upon the principles outlined in program (3.20) and program (3.27), respectively.

Give a discrete-time uncertain dynamical system, the state process is modeled as

$$x(t+1) = A(\delta)x(t) + B(\delta)u(t), \quad x(0) = x_0, \tag{3.37}$$

for time series $t = 0, 1, \cdots, T-1$, where $x(t) \in \mathbb{R}^n$ is the state variable and $u(t) \in \mathbb{R}^m$ is the control input. Meanwhile, the coefficients of matrices $A(\delta) \in \mathbb{R}^{n \times n}$ and $B(\delta) \in \mathbb{R}^{n \times 1}$ are the functions of uncertain parameters $\delta \in \Delta \subseteq \mathbb{R}^{n_\delta}$ (also known as, model parametric uncertainties), in which *arbitrary* dependence is allowed.

For an uncertain system (3.37), one can considers designing robust control to provide probabilistic robustness for uncertain parameters $\delta \in \Delta$. Due to the impact of uncertainty propagation, it is not possible to drives the system state (3.37) from a given initial state x_0 to the origin exactly. Therefore, the control goal is to find a sparse sequence input $\{u(t)\}_{t=0}^{T-1}$ that brings the terminal state $x_T(u, \delta)$ near the target state $\bar{x}$ (e.g., taking the origin $\bar{x} = 0$) over a finite time horizon T with a small metric $\gamma > 0$.

Based on the sequential iterations for plant (3.37), the terminal state x_T at final time T can eventually be expressed by

$$\begin{aligned} x_T(u,\delta) &= A(\delta)^T x_0 + \sum_{t=0}^{T-1} A(\delta)^{T-1-t} B(\delta) u(t) \\ &= A(\delta)^T x_0 + \underbrace{\begin{bmatrix} A(\delta)^{T-1}B(\delta) & \cdots & A(\delta)B(\delta) & B(\delta) \end{bmatrix}}_{\Phi_T(\delta)} \underbrace{\begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(T-2) \\ u(T-1) \end{bmatrix}}_{u} \end{aligned} \tag{3.38}$$

where $\Phi_M(\delta) \in \mathbb{R}^{n \times mT}$ stands for the *reachability matrix* and $u \in \mathbb{R}^{mT}$ represent the desired control input vector. Without loss of generality, assume that the pair $(A(\delta), B(\delta))$ is robustly reachable, that is, $\text{rank}(\Phi_T(\delta)) = n$ for all $\delta \in \Delta$. This reminds us to prescribe a suitable scalar-valued uncertain function $h(u, \delta) : \mathcal{U} \times \Delta \to \mathbb{R}$, defined by

$$\begin{aligned} h(u,\delta) &\triangleq \|x_T(u,\delta) - \bar{x}\|_2 - \gamma \\ &= \|A(\delta)^T x_0 + \Phi_T(\delta) u - \bar{x}\|_2 - \gamma \le 0, \quad \forall \delta \in \Delta, \end{aligned} \tag{3.39}$$

which is used to evaluate the "price of robustness" for the uncertainty.

3.5.1 Risk-aware sparse optimal control

With the help of convex and probabilistic relaxation techniques, we can characterize the risk-aware sparse optimal control problem as a sparse random convex program (3.20). More specifically, we employ an ℓ_1 norm convex relaxation for on control inputs whose naturally induces the sparse control signal. Adapted from [35, 104], the centerpiece of this control problem is called *risk-aware sparse optimal control* (a.k.a., ℓ_1 optimal robust control), formulated as follows.

Problem 3.1 (Risk-aware sparse optimal control) *Consider a reachable uncertain dynamical system (3.37). For given parameters $x_0, T, \rho, \gamma, \bar{x}$, one aims to seek a sequence $\{u(t)\}_{t=0}^{T-1}$*

with input sparsity such that it drives the state near to the target $\bar{x} \in \mathbb{R}^n$ *as well as hedges against the uncertainty* $\{\delta^{(i)}\}_{i=1}^N$. *The risk-aware sparse optimal control problem amounts to solving the following sparse random convex program with control variable* $u \in \mathbb{R}^{mT}$

$$\begin{aligned} \min_{u\in\mathbb{R}^{mL}} \quad & \|u\|_1 \\ \text{s.t.} \quad & h(u,\delta^{(i)}) \leq \xi^i, \quad i = 1, \cdots, N. \end{aligned} \tag{3.40}$$

where the scalar-valued uncertain constraint $h(u,\delta)$ *is measured by a distance between the terminal state* $x_L(u,\delta)$ *and the target* $\bar{z}$ *with a metric* $\gamma \geq 0$, *defined in* (3.39).

Then we formally state the following result, which is a restatement of the results of [38] in our Lemma 3.1.

Theorem 3.3 *Assume that the pair* $(A(\delta), B(\delta))$ *is robustly reachable and* $mT > n$. *Given a robust level* $\epsilon \in (0,1)$ *and a confidence parameter* $\beta \in (0,1)$, *choose the number of scenarios* $N > mT$ *such that*

$$\sum_{i=0}^{mT-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta. \tag{3.41}$$

Let u_N^* *be the feasible and unique solution to the risk-aware sparse optimal control of Problem 3.1. Then it holds that*

$$\mathbb{P}^N\{V(u_N^*) > \epsilon\} \leq \beta, \tag{3.42}$$

for any given probability $\mathbb{P}$ *over uncertainty* Δ. *In other words, with probability (N-fold product probability measure) at least* $1-\beta$, *the solution* u_N^* *is the* ϵ*-probability robust design.*

3.5.2 Risk assessment sparse optimal control with relaxation

In this subsection, we shift the principle from relaxed sparse random convex program (3.27) to sparse optimal control problem, where the control objective attempts to make the trade-offs between the sparse performance and violation constraints. It is well known that optimal control problem with constraints relaxation plays a vital role in control design, since the decision-makers have to simultaneously take the control performance and feasibility into account in practical applications. We now consider the following risk assessment sparse optimal control problem.

Problem 3.2 (Risk Assessment Sparse Optimal Control) *Given a discrete-time reachable uncertain LTI system* (3.37), *and the parameters* $x_0, \bar{x}, T, \rho, \gamma$, *the risk-aware sparse optimal control is to achieve the trade-off between the sparse control inputs and the risk assessment, which amounts to solving the following relaxed sparse random convex program, minimizing*

$$\begin{aligned} \min_{u\in\mathbb{R}^{mL},\xi^i\geq 0} \quad & \|u\|_1 + \rho \sum_{i=1}^N \xi^i \\ \text{s.t.} \quad & h(u,\delta^{(i)}) \leq \xi^i, \quad i = 1, \cdots, N. \end{aligned} \tag{3.43}$$

Then we formally state the following result, which is a restatement of the results of Lemma 3.2 in our context.

Theorem 3.4 (Trade-off sparse control and risk) *With $\underline{\epsilon}(\cdot)$ and $\overline{\epsilon}(\cdot)$ in Lemma 3.2, given a confidence $\beta \in (0,1)$, the risk assessment for violation $V(u_N^*)$ in risk assessment sparse optimal control of Problem 3.2 is as follows*

$$\mathbb{P}^N\{\underline{\epsilon}(s_N^*) \le V(u_N^*) \le \overline{\epsilon}(s_N^*)\} \ge 1-\beta. \tag{3.44}$$

where s_N^* is defined as the number of the violated constraints $h(u_N^*, \delta^{(i)}) \ge 0$.

Theorem 3.4 offers the lower and upper bounds for risk-aware sparse optimal control on its violation $V(u_N^*)$, which holds with a high confidence level $1-\beta$. Besides, s_N^* indicates the violated constraints $h(u_N^*, \delta^{(i)}) > 0$, adding those active constraints $h(u_N^*, \delta^{(i)}) = 0$.

3.6 Numerical Simulations

In this section, we conduct numerical simulations for the sparse robust control applications by means of program 3.1 and program 3.2, respectively.

3.6.1 Sparse robust control design

Let us first extend the sparse random convex program (3.20) to risk aware sparse optimal control of Problem 3.1. Consider an uncertain discrete-time plant

$$x(t+1) = A(\delta)x(t) + B(\delta)u(t), \quad x(0) = \xi \neq 0,$$

where $x(t) \in \mathbb{R}^4$ and $u(t) \in \mathbb{R}$ are the state variable and control input, respectively. The coefficients $A(\delta) \in \mathbb{R}^{4\times 4}$ and $B(\delta) \in \mathbb{R}^4$ are uncertain matrice with uncertain parameters $\delta \in \mathbb{Q} \subseteq \mathbb{R}^{20}$ of the form

$$A(\delta) = A_0 + A_\delta, \quad B(\delta) = B_0 + B_\delta,$$

in which A_0 and B_0 are nominal matrices as follows

$$A_0 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad b_0 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

They are perturbed by a uniform distribution, that is, the elements of the perturbation matrix A_δ are i.i.d. sampled from a uniform distribution over the interval $[-0.05, 0.05]$ and the entries of the perturbation matrix B_δ are i.i.d. generated from a uniform distribution over (0, 0.2).

The initial state of the plant is $x_0^\top = [0.1\ 0.1\ 0.1\ 0.1]$, and the terminal time $T = 20$. We then set the desired level of probabilistic robustness to $1 - \epsilon = 0.95$ (i.e., the tolerance level $\epsilon = 0.05$), and choose the confidence level $\beta = 10^{-5}$. Then using the simplified condition in (3.25), we need *a priori* scenarios $N \geq 888$. Without loss of generality, we assign the target as origin (i.e., $\bar{x} = 0$). We set a small gap $\gamma = 0.02$, and then choose $N = 900$ scenarios of the terminal state constraints to Problem 3.1

The goal is to find a sparse control input u such that steer the terminal state x_T (with respect to the various model parametric uncertainties $A(\delta^{(i)})$ and $B(\delta^{(i)})$) near the origin with a specified gap $\gamma = 0.02$ in (3.39). The sparse random convex program (3.20) can be solved by means of the off-the-shelf packages like `CVX` or `YALMIP` in `MATLAB`. We here adopted `YALMIP` [88] to perform the calculation. Then the optimal objective value (ℓ_1 norm of control input) was $\|u_{900}^*\|_1 = 0.0884$.

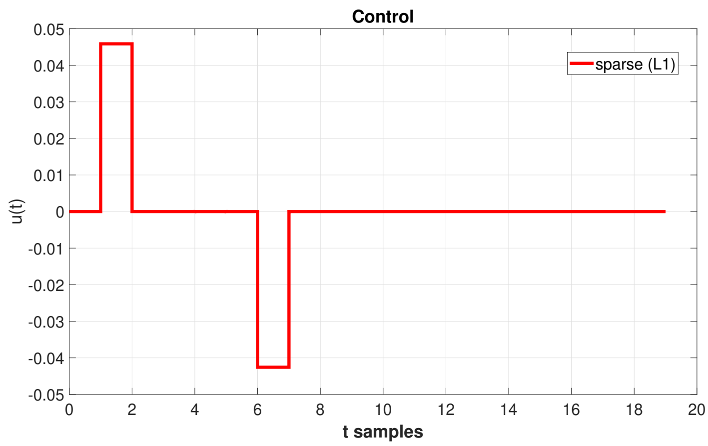


Figure 3.5: Sparse (ℓ_1 norm) optimal control

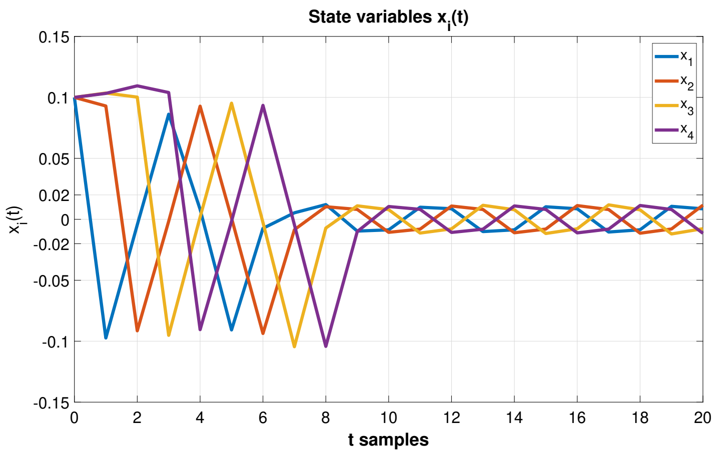


Figure 3.6: State trajectories $x_i(t)$, $i = 1, 2, 3, 4$

.

Fig. 3.5 displays the sparse optimal control input. From this figure, we can see that the designed ℓ_1 norm control is quite sparse. The four state variables generated by the

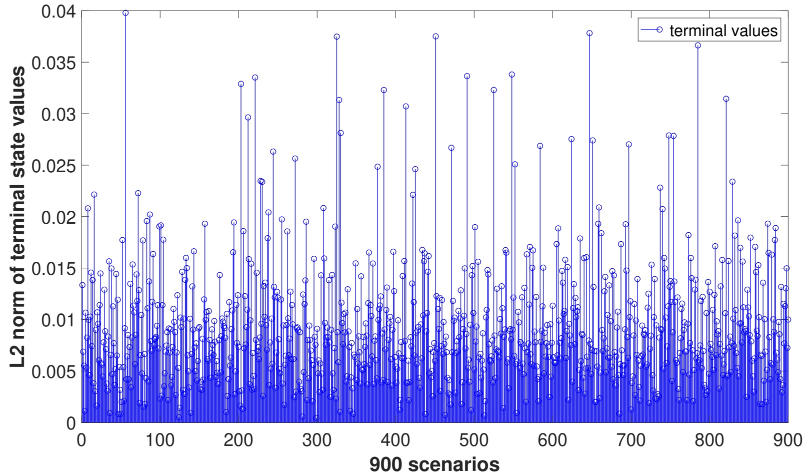


Figure 3.7: ℓ_2 norm of terminal state values for 900 scenarios testing.

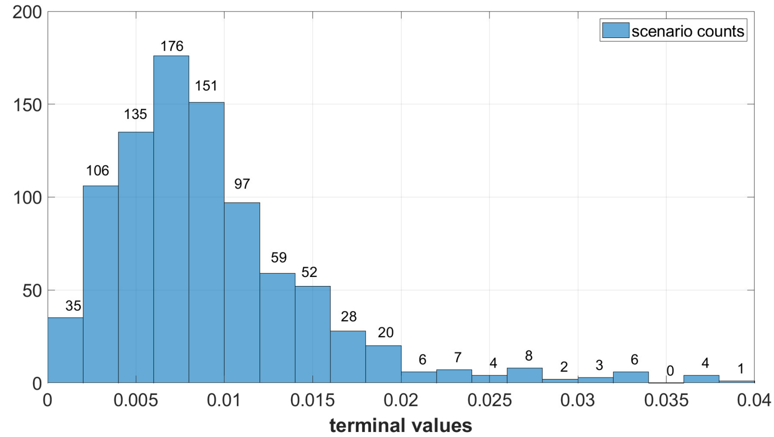


Figure 3.8: Terminal state value testing for 900 scenarios: validation test.

sparse optimal control is illustrated in Fig. 3.6. It is clear that the terminal state of four states are all close to the origin (the desired target), and they are fluctuated under a specified gap value $\gamma = 0.02$.

For the sparse random convex program (3.20), Theorem 1 asserts that with the confidence $1 - 10^{-5}$, the sparse scenario solution u^*_{900} is ϵ-probabilistic robustness. It means that the violation case $\|x_T(u, \delta^{(i)})\|_2 > \gamma$ happens with probability at most $\epsilon = 0.05$, that is, $\mathbb{P}\big\{\|x_T(u, \delta^{(i)})\|_2 > 0.02\big\} \leq 5\%$. In other words, the ℓ_2 norm of terminal state values x_T under 900 scenarios in Fig. 3.7 fall into the origin neighborhood with a radius of $\gamma = 0.02$ with probability is at least 95%.

Fig. 3.8 shows the validation test of ℓ_2 norm of terminal state values by $N = 900$ simulations, the violated value counts are 41 (i.e., less than the prescribed risk level $5\% \times 900 = 45$), which implies that the designed sparse control activates with a high probabilistic robustness.

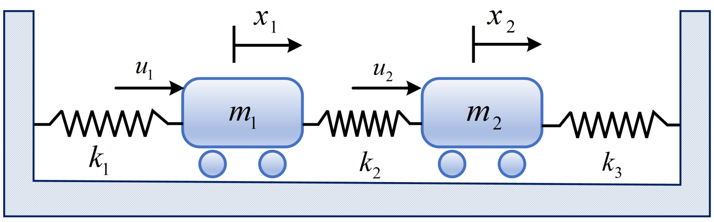


Figure 3.9: A fourth order mass-spring system with two control inputs

3.6.2 Sparse cost-risk trade-off

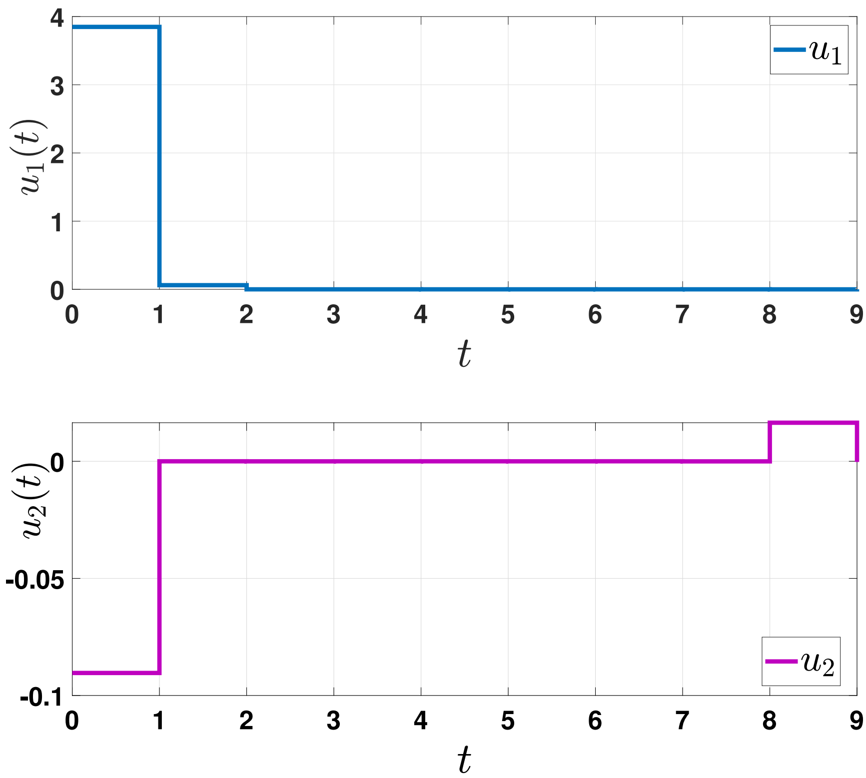


Figure 3.10: Risk-aware sparse optimal control inputs.

We now present a numerical example to assess the trade-off between sparse cost and risk for the proposed risk-aware sparse optimal control of Problem 3.43. Given an uncertain mass-spring system, which can be modeled by a linear system (3.37) using ZoH sampling Ts = 0.05, with

$$A(\delta) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-(k_1+k_2)}{m_1} & 0 & \frac{k_2}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_2}{m_1} & 0 & \frac{-(k_2+k_3)}{m_2} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ \frac{1}{m_1} & 0 \\ 0 & 0 \\ 0 & \frac{1}{m_2} \end{bmatrix}$$

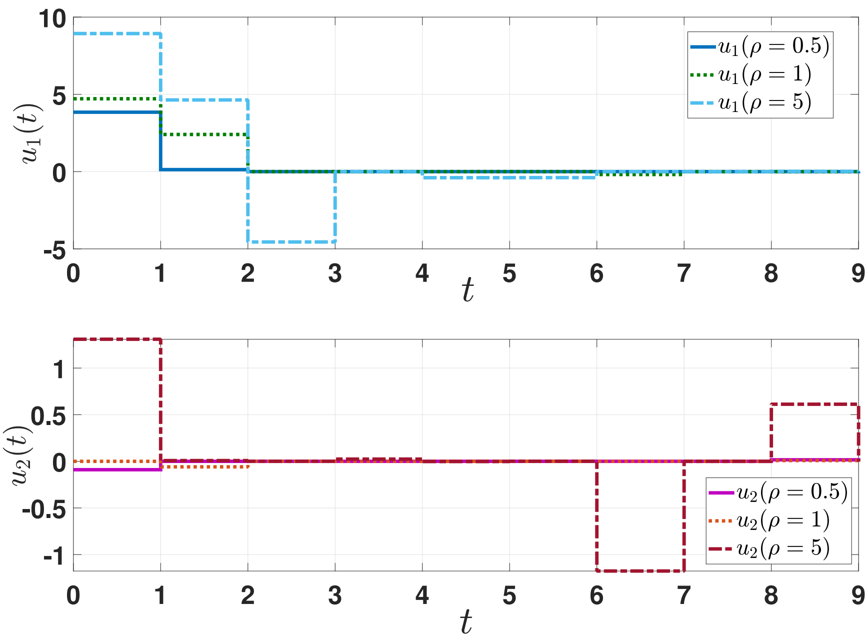


Figure 3.11: Risk-aware sparse optimal control inputs with different penalty weights $\rho = 0.5, 1, 5$, respectively.

where $m_1 = 1$, $m_2 = 2$ are the masses, and $k_1, k_2, k_3 \in \mathbb{R}$ are the flexibility of three springs, respectively. They can be collected by a vector $\delta = [k_1\ k_2\ k_3]^\top \in \mathcal{Q}$ to represent the uncertain parameters. Besides, the elements of the uncertain vector δ are i.i.d. randomly sampled from a uniform distribution over the interval $[0.1, 1]$. For this system, we compute the risk-aware sparse optimal control of Problem 3.27 over time horizon $T = 10$ with random samples $N = 1000$, and select the initial condition as $x_0 = [1\ 0\ 1\ 0]^\top$.

The objective is to seek a risk-aware control input u such that steers the terminal state $x_T(u, q)$ near to the target with a specified gap γ. We thus assign the target as the origin $\bar{x} = 0$ and set a small gap metric $\gamma = 0.5$ in (3.39). Note that this example is a convex optimization under randomized platform, and hence we applied the off-the-shelf package CVX in MATLAB to Problem 3.27.

Fig. 3.10 displays the risk-aware sparse optimal control with two inputs. It is clearly that the obtained optimal control signals (resp., above and bottom) are truly sparse, and the computed optimal value was $\|u_N^*\|_1 + \frac{1}{5}\sum_i \xi_*^i = 247.083$. Fig. 3.11 depicts the risk-aware sparse optimal control under different penalty weights $\rho = 0.5, 1, 5$, respectively. From this figure, we can see that both of two control inputs (the above and bottom) are sparse when choose the appropriate value ρ, which is used to make the trade-off between the sparse performance and the violated constraints.

Fig. 3.12 describes the sparse controls with additional scenarios validation $i = 1, \cdots, 2000$, it shows that the control signals enjoy sparsity and posterioi guarantees.

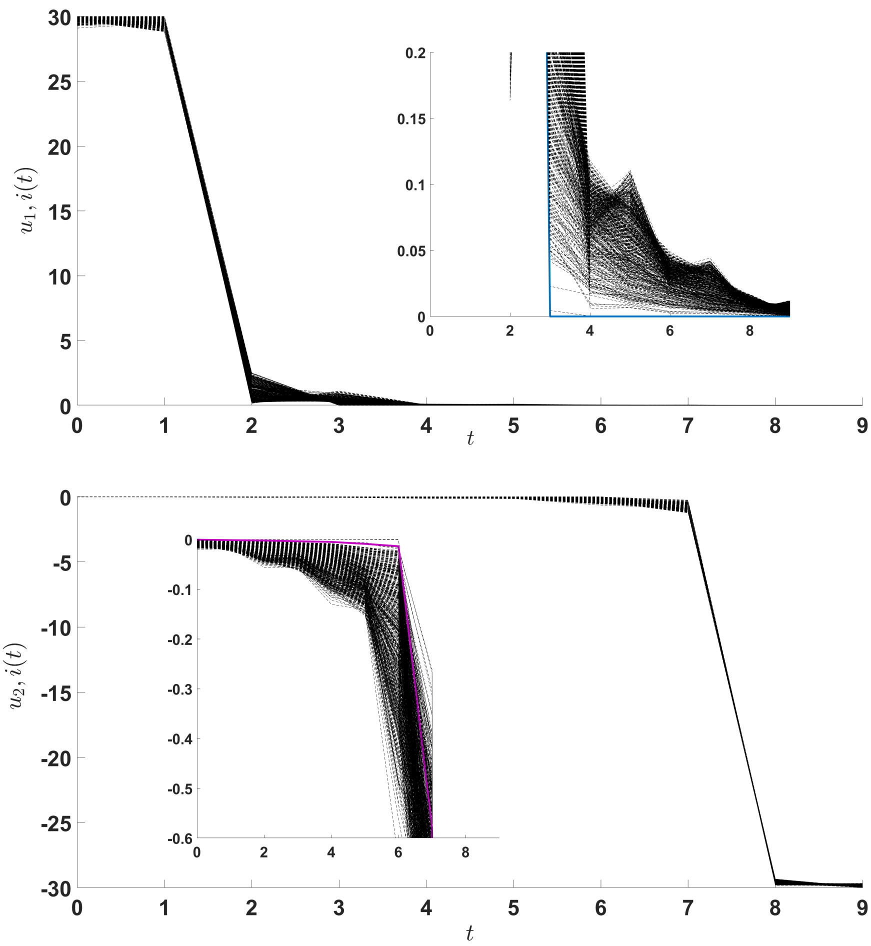


Figure 3.12: Risk-aware sparse optimal control inputs with additional scenarios $N_V = 2000$.

3.7 Summary

In this chapter, we have proposed a sparse robust control design for uncertain linear discrete-time systems by mean of sparse random convex program. This program provides a finite-sample guarantee to make the designed sparse optimal control with a high probabilistic robustness. With the help of relaxation techniques, we have converted the nonconvex ℓ_0 cost to convex ℓ_1 objective as well as have transformed the uncertain variable $\delta \in \delta$ into a probabilistic counterpart such that the sparse random convex program would be computational tractable. Furthermore, a relaxation for the constraints of sparse random convex program was investigated that makes a trade-off between the sparse cost and the violated constraints. Finally, we have performed the numerical examples which demonstrate the effectiveness for the designed risk-aware sparse optimal control.

Chapter 4

Model-Based Risk-Aware Sparse Predictive Control

In this chapter, we present a risk-aware sparse predictive control for linear discrete-time systems subject to model parametric uncertainty and additive disturbances. We consider the case of probabilistic constraints on system state and hard constraints on control actions, which is equivalent to solving a chance-constrained sparse optimization problem. We then adopt a data-driven sampling approach to seek a "randomized" solution by solving a online sparse random convex program that approximates the risk-aware sparse solution with a high confidence probability. Furthermore, we provide an explicit finite sample complexity to ensure the probabilistic robustness. Finally, the numerical example illustrates the effectiveness of the proposed control strategy.

4.1 Problem formulation

4.1.1 System description

Let us consider an uncertain linear discrete-time system

$$x_{t+1} = A(\delta)x_t + B(\delta)u_t + Ew_t, \quad x_0 \neq 0, \tag{4.1}$$

where $x_t \in \mathcal{X} \subseteq \mathbb{R}^n$ is the system state, $u_t \in \mathcal{U} \subseteq \mathbb{R}^m$ is the control input, $w_t \in \mathcal{W} \subseteq \mathbb{R}^{n_w}$ is the additive disturbance with $E \in \mathbb{R}^{n \times n_w}$, and $\delta \in \Delta \subseteq \mathbb{R}^{n_\delta}$ stands for time invariant modeling uncertain parameters of $A(\delta) \in \mathbb{R}^{n \times n}$ and $B(\delta) \in \mathbb{R}^{n \times m}$.

Without loss of generality, throughout this chapter, we make the following mild assumptions for the studied uncertain dynamics (4.1).

Assumption 4.1 (Stabilizability) *Assume that the pair $(A(\delta), B(\delta))$ is stabilizable for any uncertain realizations $\delta \in \Delta$.*

Assumption 4.2 (Stochastic Uncertainty) *Assume that the uncertainty sets Δ and $\mathcal{W}$ are bounded, and the uncertain realizations δ and w_t are independent and identically distributed* *(i.i.d.), randomly extracting from the probability distributions $\mathbb{P}_\delta$ supported on Δ and $\mathbb{P}_w$ supported on $\mathcal{W}$, respectively.*

In what follows, we present a discrete-time prediction system for the uncertain plant (4.1). Given the state x_t observed at sampling time $t \in \mathbb{N}$, let $j \in \mathbb{N}$ be *prediction horizon*, then the predicted, future states can be modelled as

$$x_{j+1|t} = A(\delta)x_{j|t} + B(\delta)u_{j|t} + Ew_{j|t}, \quad j = 0, 1, \ldots, N-1. \tag{4.2}$$

Here, we mark the subscript $\bullet_{j|t}$ as predictive instants for $j = 0, 1, \ldots, N-1$. For example, $x_{j|t}$ denotes the j-th step forward prediction of the state at sampling time t that is initialized at $x_{0|t} = x_t$; and similarly, $u_{j|t}$ and $w_{j|t}$ are the corresponding j-th predictive control input and external disturbance at sampling time t, respectively.

In this chapter, the uncertain prediction system (4.2) is subject to the *soft state* constraint (a.k.a., *probabilistic* or *chance* constraint) and the *hard input* constraint, that is,

$$\mathbb{P}\big\{x_{j|t} \in \mathcal{X},\ j = 0, 1, \ldots, N-1\big\} \geq 1 - \epsilon, \quad \text{and} \quad u_{j|t} \in \mathcal{U}, \tag{4.3}$$

or equivalently, $\mathbb{P}\big\{x_{j|t} \notin \mathcal{X},\ j = 0, 1, \ldots, N-1\big\} \leq \epsilon$, where $\epsilon \in (0, 1)$ is a user-defined risk level or violation of probability, and a joint probability $\mathbb{P} = \mathbb{P}_\delta \times \mathbb{P}_w$ satisfies Assumption 4.2.

To make the problem setup more clearly, we redefine the (predictive) state and input constraints for the plant (4.1) or (4.2) as the convex (polytopic) sets, described by

$$\begin{aligned} x \in \mathcal{X} &\triangleq \{x : Cx \leq c\}, \\ u \in \mathcal{U} &\triangleq \{u : Du \leq d\}, \end{aligned} \tag{4.4}$$

where C and D are state and input constraint matrices governed by the regarding vectors c, d with appropriate sizes. Therefore, the soft state and hard input constraints (4.3) with explicit linear constraints (4.4) are replaced with the *joint state chance constraints*[1] and the deterministic/hard control input constraint of the polytopic set forms

$$\mathbb{P}\big\{Cx_{j+1|t} \leq c,\ j = 0, 1, \ldots, N-1\big\} \geq 1 - \epsilon, \tag{4.5a}$$

$$Du_{j|t} \leq d. \tag{4.5b}$$

For each step $j = 0, 1, \ldots, N-1$ of the dynamics (4.2), we assemble the trajectories of the state $x_{j|t}$, control input $u_{j|t}$, and disturbance $w_{j|t}$, respectively, as follows

$$\bar{x} \triangleq \begin{bmatrix} x_{1|t}^\top & x_{2|t}^\top & \cdots & x_{N|t}^\top \end{bmatrix}^\top \in \mathbb{R}^{nN}, \tag{4.6a}$$

[1] Using Boole's inequality to bound the probability of violation of joint chance constraint $\mathbb{P}\{x_{j|t} \notin \mathcal{X}\} = \mathbb{P}\{x \in \bigcup_{j=1}^N \{Cx_j > c\}\} \leq \sum_{j=1}^N \mathbb{P}\{Cx_j > c\}$ as N individual chance constraints of forms $\mathbb{P}\{Cx_j > c\} \leq \varepsilon_j$, $j = 1, \ldots, N$, where $\varepsilon_j \in [0, \epsilon]$ denotes the risk for the j-th individual chance constraint satisfying $\sum_{j=1}^N \varepsilon_j \leq \epsilon$.

$$\bar{u} \triangleq \begin{bmatrix} u_{0|t}^\top & u_{1|t}^\top & \cdots & u_{N-1|t}^\top \end{bmatrix}^\top \in \mathbb{R}^{mN}, \tag{4.6b}$$

$$\bar{w} \triangleq \begin{bmatrix} w_{0|t}^\top & w_{1|t}^\top & \cdots & w_{N-1|t}^\top \end{bmatrix}^\top \in \mathcal{W}^N \subseteq \mathbb{R}^{n_w N}. \tag{4.6c}$$

To alleviate the burden and abuse of notations, in the following context, we collect the random variables $\delta, \bar{w}$ of plant (4.2) as a pair $\theta \triangleq (\delta, \bar{w})$, where θ is supported on the set $\Theta \triangleq \Delta \times \mathcal{W}^N$ (i.e., $\theta \in \Theta$) underlying a probability distribution $\mathbb{P} \triangleq \mathbb{P}_\delta \times \mathbb{P}_w^N$.

Due to the presence of the uncertainties $\theta \in \Theta$, regulation of the system state to the origin can not be precisely achieved in general. Conversely, we intend to seek a "risk-aware" predictive control sequences $\{u_{0|t}, \ldots, u_{N-1|t}\}$ of length N predictions for each sampling time t such that it can hedge against the stochastic perturbations θ and can drive the uncertain prediction state (4.2) from the initial state $x_{0|t}$ towards a prescribed terminal set $\mathcal{X}_f \triangleq \{x_f : C_f x_f \le c_f\}$ satisfying $x_f \in \mathcal{X}_f$ with a high probability. In other words, we replace (4.5a) with

$$\mathbb{P}\{C_f x_{N|t} \le c_f,\, C x_{j+1|t} \le c,\, j = 0, 1, \ldots, N-1\} \ge 1 - \epsilon, \tag{4.7}$$

by adding an additional terminal state constraint. We now need to inspect whether the condition (4.7) holds or not, where C_f is a state constraint matrix governed by the regarding a vector c_f with an appropriate size. The essence of the term (4.7) claims that the state constraints allow a small risk level to violate itself, i.e.,

$$\mathbb{P}\{C_f x_{N|t} > c_f,\, C x_{j+1|t} > c,\, j = 0, 1, \ldots, N-1\} \le \epsilon, \tag{4.8}$$

and it is still feasible for *most* of the uncertain instances.

4.1.2 Sparse predictive control

Similar to [52, 77], the control objective of this chapter aims to implement a sparse predictive control, we then introduce a sparsity-promoting predictive control cost as

$$J(\bar{u}) = \sum_{j=0}^{N-1} \|u_{j|t}\|_1 = \|\bar{u}\|_1. \tag{4.9}$$

Note that this control cost is a convex ℓ_1 optimal (predictive) control that maximizes the time interval over which the predictive signals is exactly zero under a certain condition [104]. Besides, here the predictive control cost (4.9) is a direct real-time control iterations rather than the stochastic tubed-based or parametric affine control schemes [98], since determining a feedback gain of pure sparse optimal (predictive) control problem is a nontrivial task.

As indicated in [93, 117], a *receding horizon strategy* is used to execute model predictive control. More precisely, no matter how many steps the controller can predict, only the *first control action* $u_{0|t}$ *of all predicted inputs* (4.6b) is applied to the prediction system

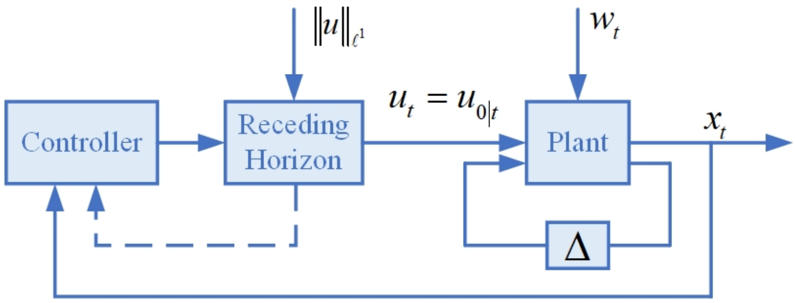


Figure 4.1: Sparse predictive control based on the receding horizon policy.

(4.2) at each sampling instant t. We thus formulate the predictive control policy as

$$u_t = u_{0|t} = \underbrace{\begin{bmatrix} I_m & 0_{m\times(N-1)m} \end{bmatrix}}_{F} \underbrace{\begin{bmatrix} u_{0|t} \\ u_{1|t} \\ \vdots \\ u_{N-1|t} \end{bmatrix}}_{\bar{u}} \tag{4.10}$$

where $F \in \mathbb{R}^{m\times Nm}$ is called a moving horizon scheme matrix that consists of the identify matrix $I_m \in \mathbb{R}^{m\times m}$ and zero matrix $0_{m\times(N-1)m}$.

Remark 4.1 (Sparse Predictive Control) When considering the sparsity-promoting on predictive control (a.k.a., sparse predictive control, see Figure 4.1), the efficient of the receding horizon policy implies the first control action should be sparse at each sampling time $t \in \mathbb{N}$, namely, conducting $\|u_t\|_1 = \|u_{0|t}\|_1$.

4.2 Risk-aware sparse predictive control

This section studies the main results of identifying the risk-aware sparse optimal solution $\bar{u}^*$ that minimizes an ℓ_1 optimal control cost over a finite prediction horizon. This problem is equivalent to dealing with a chance-constrained sparse optimization problem. Meanwhile, we follow a data-driven sampling approach to seek a "randomized" solution $\bar{u}_K^*$ that approximates the original chance-constrained or risk-aware sparse solution with a high probability

we now present a novel sparse predictive control that adapts from the model predictive control with *risk-aware* sense and *sparse* manner, which can be formulated as the following problem.

Problem 4.1 (Risk-aware sparse predictive control) *A risk-aware sparse predictive control for prediction system* (4.2) *subject to soft state constraints* (4.5a) *and hard input constraint* (4.5b)*, is amounts to solving an online chance-constrained sparse optimization problem*

$$\min_{\bar{x},\bar{u}} \quad \|\bar{u}\|_1 \tag{4.11a}$$

$$\text{s.t.} \quad x_{j+1|t} = A(\delta)x_{j|t} + B(\delta)u_{j|t} + Ew_{j|t},$$
$$x_{0|t} = x_t, \quad j = 0, 1, \ldots, N-1, \tag{4.11b}$$
$$\mathbb{P}\{C_f x_{N|t} \leq c_f,\ Cx_{j+1|t} \leq c,\ j = 0, 1, \ldots, N-1\} \geq 1 - \epsilon, \tag{4.11c}$$
$$Du_{j|t} \leq d, \quad j = 0, 1, \ldots, N-1, \tag{4.11d}$$

Although the proposed Problem 4.1 is well-defined, such chance-constrained solution $\bar{u}_\epsilon^*$ is extremely difficult to calculate exactly. In particular, the feasible set described by chance constraints (4.11c) is general difficult to evaluate either analytically or numerically, since it involves the multivariate integrals computing and the relevant set is *non-convex* even under the convexity assumption for hard constraint set.

4.3 Data-driven sampling approach

In what follows, we introduce a "data-driven" sampling approach called *scenario optimization* [37] to deal with the risk-aware sparse predictive control problem, i.e., chance-constrained sparse program (4.11), which is essentially a *randomized algorithms* [125] that generates a finite number of random instances of an uncertain variable θ according to the probability $\mathbb{P}$, see Assumption 3.2 in Chapter 3.

First, we randomly collect a finite number of "samples" or "realizations" or "scenarios" $\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(K)}\}$ from $\theta \doteq (\delta, \bar{w})$ in an i.i.d. fashion according to the probability distribution $\mathbb{P}$. We now substitute the scenario counterpart into the chance constrained sparse optimization (4.11). In this way, we explore the risk-aware sparse predictive control via data-driven scenario approach as

$$\min_{\bar{x}, \bar{u}} \quad \|\bar{u}\|_1 \tag{4.12a}$$
$$\text{s.t.} \quad x^{(i)}_{j+1|t} = A(\delta^{(i)})x^{(i)}_{j|t} + B(\delta^{(i)})u_{j|t} + Ew^{(i)}_{j|t},$$
$$x^{(i)}_{0|t} = x_t, \quad j = 0, 1, \ldots, N-1, \quad i = 1, 2, \ldots, K, \tag{4.12b}$$
$$C_f x^{(i)}_{N|t} \leq c_f, \quad Cx^{(i)}_{j|t} \leq c, \quad j = 0, 1, \ldots, N-1, \quad i = 1, 2, \ldots, K, \tag{4.12c}$$
$$Du_{j|t} \leq d, \quad j = 0, 1, \ldots, N-1, \tag{4.12d}$$

where the superscript i denotes the random sampling with K scenarios. It is clear that the data-driven scenario approach gives rise to a " sparse random convex program" that can be solved by means of the off-the-shelf packages, like `CVX` or `YALMIP` in MATLAB. This data-driven sampling method generates a sparse scenario solution $\bar{u}_K^*$ to problem (4.12), and $\bar{u}_K^*$ is a "randomized" solution since it depends on the collected random samples $\theta^{(1)}$, $\theta^{(2)}$, . . ., $\theta^{(K)}$. Without loss of generality, we here assume that the random convex program (4.12) admits a unique solution. Besides, the feasible set of (4.12) is an "inner approximation" of the feasible set of the original chance-constrained sparse optimization (4.11).

We next focus on the uncertain quantification.

Lemma 4.1 *The uncertain part of the robust feasibility set defined by (4.12b) and (4.12c) is represented as*

$$h(\bar{u}, \theta) \leq 0, \tag{4.13}$$

with a function $h(\bar{u}, \theta) : \mathbb{R}^{mN} \times \Theta \to \mathbb{R}$ which is convex in $\bar{u}$ for any fixed $\theta \in \Theta$.

Proof 4.1 For a given prediction horizon N, substituting the prediction dynamics (4.2) recursively, we have the evolution of system in a compact form

$$\bar{x} = \bar{A}_\delta x_{0|t} + \bar{B}_\delta \bar{u} + \bar{E}_\delta \bar{w},$$

where parametric uncertain matrices $\bar{A}_\delta$, $\bar{B}_\delta$ and $\bar{E}_\delta$ are defined as follows, respectively,

$$\bar{A}_\delta = \begin{bmatrix} A(\delta) \\ A^2(\delta) \\ \vdots \\ A^N(\delta) \end{bmatrix}, \quad \bar{B}_\delta = \begin{bmatrix} B(\delta) & 0 & \cdots & 0 \\ A(\delta)B(\delta) & B(\delta) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}(\delta)B(\delta) & A^{N-2}(\delta)B(\delta) & \cdots & B(\delta) \end{bmatrix},$$

$$\bar{E}_\delta = \begin{bmatrix} E & 0 & \cdots & 0 \\ A(\delta)E & E & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}(\delta)E & A^{N-2}(\delta)E & \cdots & E \end{bmatrix}.$$

In addition, gathering the constraint matrices C, C_f and vectors c, c_f stated in (4.12c)

$$\bar{C} = \begin{bmatrix} C & 0 & \cdots & 0 \\ 0 & C & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C \\ 0 & 0 & \cdots & C_f \end{bmatrix}, \quad \bar{c} = \begin{bmatrix} c \\ c \\ \vdots \\ c \\ c_f \end{bmatrix},$$

we then have

$$h(\bar{u}, \theta) \triangleq \max\left(\bar{C}(\bar{A}_\delta x_{0|t} + \bar{B}_\delta \bar{u} + \bar{E}_\delta \bar{w}) - \bar{c}\right),$$

where $\max(v)$ for a vector v means that the greatest element of v, which completes the proof of Lemma 4.1. ■

4.4 A finite-sample guarantee

In general, the obtained data-driven solution $\bar{u}_K^*$ does not always successfully approximate the risk-aware solution $\bar{u}^*$ thanks to the randomness of the data sampling. In order to hold the probabilistic robustness guarantee with a risk ϵ and a confidence level β, sufficient data-driven samples must be collected. In what follows, we state the sample complexity to ensure the data-driven sparse solution $\bar{u}_K^*$ is a good approximation for risk-aware sparse solution $\bar{u}^*$.

Theorem 4.2 (Finite-sample guarantee) *Suppose that an uncertain constraint function $h(\bar{u}, \theta)$ defined in (4.13) is given. Assume that the uncertain variable $\theta \doteq (\delta, \bar{\omega})$ has the probability distribution $\mathbb{P}$ supported on Θ. Let $\Theta^K \doteq \{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(K)}\}$ be a multi-sample of θ, where K is selected so that it satisfies*

$$\sum_{i=0}^{mN-1} \binom{K}{i} \epsilon^i (1-\epsilon)^{K-i} \leq \beta \tag{4.14}$$

for given specified risk $\epsilon \in (0,1)$ and confidence $\beta \in (0,1)$. Suppose that there exists a unique optimal solution for sparse random convex program (4.12), i.e., solving the program

$$\begin{aligned} \min_{\bar{x}, \bar{u}} \quad & \|\bar{u}\|_1 \\ \text{s.t.} \quad & h(\bar{u}, \theta^{(i)}) \leq 0, \quad i = 1, 2, \dots, K \\ & Du_{j|t} \leq d, \quad j = 0, 1, \dots, N-1. \end{aligned} \tag{4.15}$$

Then, with confidence $1 - \beta$, a probabilistic robustness guarantee

$$\mathbb{P}^K\{V(\bar{u}_K^*) \leq \epsilon\} \geq 1 - \beta$$

holds true for the optimal input $\bar{u}_K^*$ obtained by (4.15).

Corollary 4.1 (Sample Complexity) *The satisfaction of term (4.14) implies that a priori and explicit sample complexity $K \geq \mathcal{K}(mN, \epsilon, \beta)$ holds true, where*

$$\mathcal{K}(mN, \epsilon, \beta) \triangleq \frac{mN - 1 + \ln(1/\beta) + \sqrt{2(mN-1)\ln(1/\beta)}}{\epsilon}. \tag{4.16}$$

Proof 4.2 The proof is directly derived from the result [4, Therorem 1]. ■

Then a direct result is summarized as follows.

Proposition 4.3 (Risk-aware sparse predictive control realization) *Let $\bar{u}_K^*$ be a feasible optimal control for Problem (4.12) satisfying sample complexity $K \geq \mathcal{K}(mN, \epsilon, \beta)$. If $u_{j|t}^*$ is applied to the uncertain system (4.2) with a finite prediction horizon N, then, with probability $1 - \beta$, the risk-aware sparse predictive control (4.11) is achieved for $j = 0, 1, \dots, N-1$.*

4.5 Numerical examples

In this section, we illustrate the risk-aware sparse predictive control on a benchmark two-mass-spring system [78] shown in Figure 4.2. The system matrices of discrete dynamics (4.1) by Euler's approximation with ZoH sampling time $t_s = 0.1$ s are of the

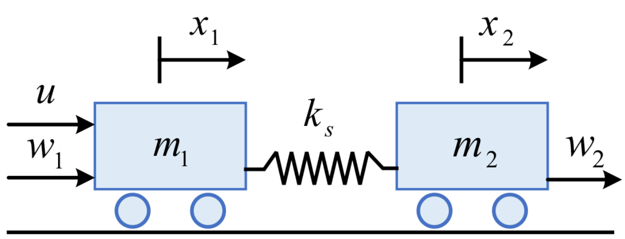


Figure 4.2: Two-mass-spring systems

forms

$$A(\delta) = \begin{bmatrix} 1 & 0 & t_s & 0 \\ 0 & 1 & 0 & t_s \\ -\frac{k_s t_s}{m_1} & \frac{k_s t_s}{m_1} & 1 & 0 \\ \frac{k_s t_s}{m_2} & -\frac{k_s t_s}{m_2} & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \frac{t_s}{m_1} \\ 0 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{t_s}{m_1} & 0 \\ 0 & \frac{t_s}{m_2} \end{bmatrix},$$

$$x_t = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad w_t = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

The model state x_t has four dimensions, in which the first two components $\{x_1, x_2\}$ stand for positions of two masses, the rest components $\{x_3, x_4\}$ represent velocities.

Assume that the additive disturbance w_t is perturbed by a Gaussian distribution $w \sim \mathcal{N}(0, \Sigma_w)$ with zero mean and covariance matrix $\Sigma_w = \text{diag}(0.02^2, 0.02^2)$. The masses of two spring systems are set as $m_1 = m_2 = 1$ kg, and the "elastic constant" k_s is related to parametric uncertainty that follows a uniform probability distribution $k_s \sim \text{Unif}([0.5, 2.0])$. We take the initial state as

$$x_0 = \begin{bmatrix} 0.15 & 0.15 & -0.15 & -0.1 \end{bmatrix}^\top,$$

and the prediction horizon is $N = 6$, and the sampling time $t = 50$. For the data-driven sampling setup, we choose the risk $\epsilon = 0.05$, confidence level $\beta = 10^{-6}$ and by means of sample complexity (4.16), we take scenarios $K = 695$. Next, we use these parameters to prescribe the constraints.

In two-mass-spring system, the standard setting for the mass positions (i.e., the first two states x_1 and x_2) are often free. Hence, we only give the sampled-based velocities constraints $|x_3| \leq 0.15$, $|x_4| \leq 0.15$, and hard control inputs $|u_t| \leq 1$. Meanwhile, we conduct a simulation for quadratic model predictive control (MPC) based on ℓ_2 optimal control (i.e., $\min_{\bar{u}} \|\bar{u}\|_2$) so as to make a comparison with sparse predictive control. Figure 4.3 displays the control inputs with ℓ_2 optimal and ℓ_1 optimal controllers, respectively. Clearly, the sparse predictive control promotes more zero inputs than quadratic MPC, therefore it enjoys sparsity.

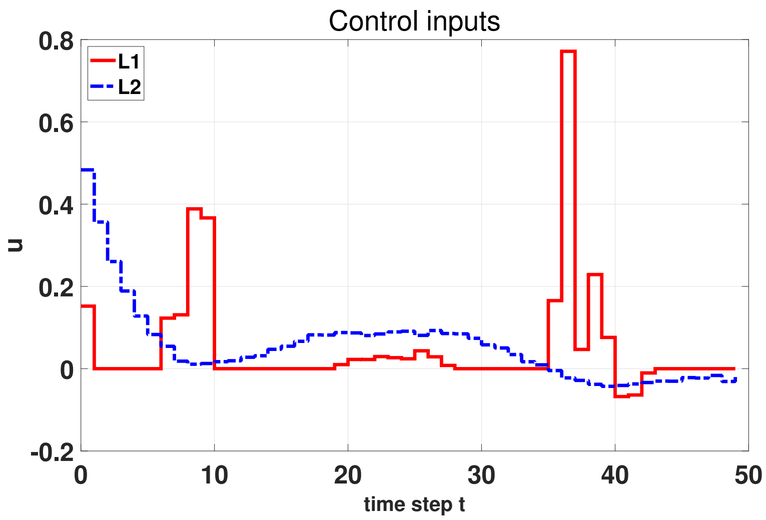


Figure 4.3: Profiles of sparse (ℓ_1) predictive control (solid) and quadratic (ℓ_2) model predictive control (dashed).

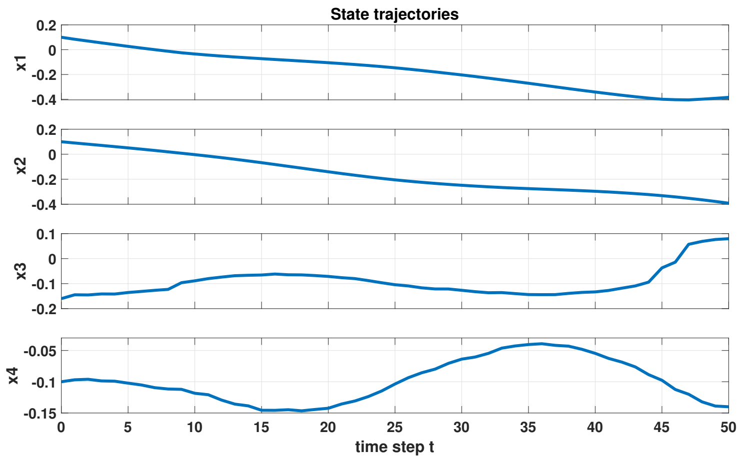


Figure 4.4: State trajectories with prediction horizon $N = 6$ and time step $t = 50$ s.

Fig. 4.4 illustrates the four different state trajectories under risk-aware sparse predictive control. Due to the nonzero uncertainties, the state trajectories will tend to the neighborhood of the origin. In particular, we can see that the velocity x_3 tends to the assigned upper bound 0.15 as well as the velocity x_4 gets close to the pre-specified lower bound -0.15. This implies that the sparse predictive control offers a good robustness.

4.6 Summary

In this chapter, we presented a risk-aware sparse predictive control, where the state constraints are formed as probabilistic model. To solve such chance-constrained sparse optimization problem, we utilized a data-driven scenario approach to seek a randomized solution by solving a sparse random convex program such that it approximates the original risk-aware sparse solution with a high probability. Meanwhile, we provided a finite sample guarantee for the obtained data-driven scenario sparse solution. Finally, the numerical example showed that the usefulness of proposed control approach.

Chapter 5

Data-Driven Sparse Optimal Control Synthesis

A data-driven framework for a discrete linear time invariant (LTI) system is employed in conjunction with sparse feedback control synthesis. Instead of a priori knowledge of actual system model, this chapter concerns on the black-box control systems by purely exploiting experimental input/state/output data samples under Willem's lemma. An $\ell_{1,\infty}$ matrix norm on feedback gain is penalized to promote a row-sparse structure that maximizes the number of zero-valued rows within the feedback matrix itself, resulting in the generation of sparse control signals.

5.1 Recap of model-based sparse feedback control

Consider a discrete linear time invariant (LTI) system

$$x(t+1) = Ax(t) + Bu(t), \quad x(0) = x_0, \tag{5.1a}$$

$$y(t) = Cx(t), \quad t \in \mathbb{N}_{\geq 0}, \tag{5.1b}$$

where $x(t) \in \mathbb{R}^n$ is the system state, $u(t) \in \mathbb{R}^m$ is the control action, and $y(t) \in \mathbb{R}^p$ stands for the output with $p \geq 2$. Besides, $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are system matrices, and the output matrix $C \in \mathbb{R}^{p\times n}$ has full row rank. Assume that the pairs (A, B) is *reachable* and (A, C) is *observable*.

5.1.1 Sparse state feedback control design

Let us introduce a static state feedback control $K \in \mathbb{R}^{m\times n}$, expressed by

$$u(t) = Kx(t), \tag{5.2}$$

for which the control signal u is generated to ensure that a discrete LTI dynamics (5.1) is stable (i.e., closed-loop stability) and minimizes the control effort (i.e., input sparsity) [32, 104, 113].

Definition 5.1 ($\ell_{1,\infty}$ norm for a Matrix [115, 128]) *The $\ell_{1,\infty}$ matrix norm for a matrix $K \in \mathbb{R}^{m\times n}$ (aka. a ℓ_1 norm of matrix rows) is defined as*

$$\|K\|_{1,\infty} = \sum_{i=1}^{m} \max_{1\leq j\leq n} |K_{i,j}|$$

which penalizes the sum of maximum absolute values of each row. Obviously, $\ell_{1,\infty}$ norm promotes row sparsity, i.e., it encourages entire rows of the matrix to have zero-valued elements. It is a convex relaxation for $\|K\|_{r_0} = \|K\|_{0,\infty}$.

Problem 5.1 (Model-Based Sparse Feedback Control) *For a discrete LTI plant (5.1a), finding a static state feedback gain K meets (5.2), then the sparse feedback control u with input sparsity simultaneously satisfies the following two conditions*

(c-i) the discrete time closed-loop (or feedback) system

$$x(t+1) = A_{cl}x(t), \quad A_{cl} \doteq (A + BK) \in \mathbb{R}^{n\times n}, \tag{5.3}$$

is Schur stable *if for all eigenvalues have $|\lambda_i(A_{cl})| < 1$;*

(c-ii) and feedback matrix $K \in \mathbb{R}^{m\times n}$ enjoys "row-sparsity" that penalizes the $\ell_{1,\infty}$ norm on the matrix, defined as $\|K\|_{1,\infty}$ by promoting the number of zero-valued rows.

Remark 5.1 (Motivation) Distinct from prior methods that directly apply sparsity to the control signal, here an motivation for integrating sparsity into control design is illuminated by considering the interplay of *Control, Communication, and Computation* from a unified standpoint [113]. Within this framework, reducing the number of states required for plant control corresponds to a decrease in the number of sensors utilized. Likewise, the number of controls relates to the quantity of actuators, and minimizing the number of outputs aligns with reducing the information transmitted through control channels, we refer to see Example 5.1.

Remark 5.2 (Schur α-Stability) A modified Shur α-stability (i.e., an expected level of stability α, $0 < \alpha < 1$) in (c-i) for the plant (5.3) can be evaluated, which *measures the minimal distance between the magnitude of the eigenvalues $\lambda_i(A_{cl})$ and the unit disk*. Let $\alpha = 1 - \bar{\rho}$, here $\bar{\rho} \doteq \max_{1\leq i\leq n} |\lambda_i(A_{cl})| < 1$ denotes the spectral radius of the closed-loop matrix A_{cl}.

The subsequent linear matrix inequality (LMI) lemma is useful to solve the condition (c-i) or Remark 5.2 in Problem 5.1.

Lemma 5.1 (An LMI for Schur α-Stability) *A discrete-time closed-loop matrix A_{cl} (5.3) consisting of the LTI plant (5.1a) and state feedback (5.2) is* Schur α-stabilizable *if and only if*

$$\exists P \succ 0, \quad s.t. \quad A_{cl}PA_{cl}^\top - (1-\alpha)^2 P \preceq 0. \tag{5.4}$$

Proof 5.1 Suppose that there exists a positive-definite matrix $P = P^\top \succ 0$ such that the LMI $A_{cl}^\top PA_{cl} - \bar{\rho}^2 P \prec 0$ holds. Construct a Lyapunov function $V(x) = x^\top Px \succ 0$, satisfying $x \neq 0$ and $V(0) = 0$. By difference, for all $t \in \mathbb{N}_{\geq 0}$, it gives

$$\begin{aligned}
& V(x(t+1)) - V(x(t)) \\
&= x^\top(t+1)Px(t+1) - x^\top(t)Px(t) \\
&= x^\top(t)A_{cl}^\top PA_{cl}x(t) - x^\top(t)Px(t) \\
&= x^\top(t)\big(A_{cl}^\top PA_{cl} - P\big)x(t) \prec 0
\end{aligned} \tag{5.5}$$

for all non-zero $x(t)$.

We now slightly revise the matrix form A_{cl} as $A_{cl}/(1-\alpha)$ by taking a degree of stability $\alpha = 1 - \bar{\rho}$ into account, which does not change the stability from a theoretical point of view. Obviously, the inequality $V(x(t+1)) \prec V(x(t))$ for all $t \in \mathbb{N}_{\geq 0}$ means that the conditions (5.4) hold true, and hence for any $x_0 \in \mathbb{R}^n$ the system (5.3) is Schur or asymptotic stable.

On the other hand, the α-stability implies that discrete time closed-loop system with matrix A_{cl} is Schur α-stabilizable. ∎

With relation to Lemma 5.1, introducing new variables $Q = P^{-1}$, $V = KQ$ and the recovered the expression $A_{cl} = A + BK$, the semi-definite program (SDP) of (5.4) is announced by

$$\begin{bmatrix} Q & Q^\top A^\top + V^\top B^\top \\ AQ + BV & \bar{\rho}^2 Q \end{bmatrix} \succ 0, \quad Q \succ 0. \tag{5.6}$$

This result derives from Schur complement. The existence of solutions for LMI hinges on the controllability of pair (A, B), which admits the controller $K = VQ^{-1}$.

In what follows, we deal with (c-ii) in Problem 5.1 that infers sparse control signals from its state feedback policy $u = Kx$. A desired scenario we would like a solution matrix K with a few non-zero rows (i.e., a few active features). When solution $K = VQ^{-1}$ is eager to structural sparsity in rows, then the matrix V naturally exhibits the identical row sparsity, that is, *row-sparsity invariance*, see Example 5.1. We thus replace the $\|K\|_{1,\infty}$ regularizer by $\|V\|_{1,\infty}$ penalty [113, 32]. Therefore, the next result which comes directly from Lemma 5.1 or LMI (5.6), making Problem 5.1 is computationally tractable in terms of SDP with the lens of `CVX` or `SeDuMi`.

Example 5.1 (Row-sparse matrix) *Identifying a row-sparsity feedback matrix $K \in \mathbb{R}^{m \times n}$ is equivalent to find a matrix $V = KQ$, which shares a common row-sparse pattern, that is,*

$$\underbrace{\begin{bmatrix} * \\ 0 \\ \vdots \\ * \\ 0 \end{bmatrix}}_{u} = \underbrace{\begin{bmatrix} * & * & \cdots & * \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{K} \times \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \cdot \\ x_n \end{bmatrix}}_{x}, \quad K = \underbrace{\begin{bmatrix} * & * & \cdots & * \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{V} \times \underbrace{\begin{bmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \\ * & * & \cdots & * \end{bmatrix}}_{Q^{-1}}.$$

Proposition 5.2 (Sparse State Feedback Control [32]) *To address Problem 5.1 is equivalent to finding the solutions V, Q to perform constrained row sparse matrix optimization subject to LMI constraints as follows*

$$\min_{V,Q} \quad \|V\|_{1,\infty} \quad s.t. \quad LMIs\ (5.6), \tag{5.7}$$

where the new matrix variable $V = KQ$ induces zero-valued rows, giving rise to a row sparse feedback matrix $K = VQ^{-1}$, which stabilizes the interconnected systems (5.1a) and (5.2) with a desired level of α-stability for the closed-loop behavior (5.3).

First, we assure the sufficient stabilization (5.6), then performing $\ell_{1,\infty}$ matrix norm on V (or K) yields row sparse solutions whose most rows of the matrix should be zeros (resp., zero or "rest" components in feedback control u). Conversely, the nonzero rows of the matrix should contain many nonzero entries (resp., nonzero or "active" inputs over control signal).

Before closing this subsection, we review some definition and remarks withe respect to structured closed-loop implementation.

Definition 5.2 (Sparse Pattern) *Finding a feedback matrix $K \in \mathbb{R}^{m\times n}$ that assigns the closed-loop eigenvalues (5.3) at some desired locations given by set $\Lambda = \{\lambda_1, \cdots, \lambda_n\}$, and satisfies the row-sparsity. Let $\mathbf{K} \in \{0,1\}^{m\times n}$ denote a binary mask matrix that specifies the row-sparsity structure of the feedback matrix K. For row sparse matrix, If $\mathbf{K}_{i,:} = 0$ (resp., $\mathbf{K}_{i,:} = 1$), then computing i-th row is as follow*

$$\mathbf{K}_{i,:} = \begin{cases} \mathbf{0}, & \text{if } \mathbf{K}_{ij} = 0, \\ \star, & \text{if } \mathbf{K}_{ij} = 1, \end{cases}$$

where $\star$ denotes a real number. Let $\mathbf{K}^c = 1_{m\times n} - \mathbf{K}$ denote the complementary *sparse structure matrix.*

Definition 5.3 (Fixed Modes [75, 133]) *The fixed modes of the pair (A, B) regarding (row) sparsity structure $\mathcal{K}$ are those eigenvalues of matrix A that can not be changed by LTI static (or dynamic) feedback gain K, and are typically denoted as*

$$\Gamma(A, B, \mathbf{K}) = \bigcap_{K:K\circ\mathbf{K}=0} \Gamma(A + BK). \tag{5.8}$$

Remark 5.3 (Eigenvalues) Let $\Lambda = \{\lambda_1, \cdots, \lambda_n\}$ be the set of desired closed-loop *eigenvalues* and $\mathrm{V} = \{v_1, \cdots, v_n\}$ be the set of desired closed-loop *eigenvectors* for (5.3), respectively, with $\lambda_i \in \mathbb{R}$ and $v_i \in \mathcal{V}_i \subseteq \mathbb{R}^n$ for all $i \in \{1, \cdots, n\}$. Thus,

$$(A + BK)v_i = \lambda_i v_i, \quad \forall v_i \in \mathcal{V}_i,\ \lambda_i \in \Lambda, \tag{5.9}$$

where $\mathcal{V}_i$ indicates n-dimensional *allowable eigenvector subspace*. Also, for some K it derives the spectrum $\rho(A + BK) = \Lambda$, and $\mathcal{V}_i$ is independent of K [87].

Remark 5.4 (Pole Placement) Checking $|\lambda_i(A_{cl})| < 1$ is analogous to pole placement [75, 44] that is the closed-loop action coincides with a desired eigenvalues from Λ

$$(A + BK)X = X\Xi \tag{5.10}$$

where $\Xi = \text{diag}\left([\lambda_1, \cdots, \lambda_n]^\top\right)$ is the diagonal matrix of the desired eigenvalues.

In fact the state of a system can be restricted within a subspace $\mathcal{W}$ via a state feedback control K *if and only if* $\mathcal{W}$ is a controlled invariant space.

Remark 5.5 (Controlled Invariant Subspace) The subspace $\mathcal{W} \subseteq \mathbb{R}^n$ is an $(A, \text{Im}(B))$-*controlled invariant*, that is, $A\mathcal{W} \subseteq \mathcal{W} + \text{Im}(B)$, then there exists a matrix K (called a friend of $\mathcal{W}$) such that

$$(A + BK)\mathcal{W} \subseteq \mathcal{W} \tag{5.11}$$

$$\exists X, U \quad AW = WX + BU \tag{5.12}$$

where W is a *basis* of subspace $\mathcal{W}$, this is related to geometric control approach [11]

5.1.2 Sparse output feedback control synthesis

Considering the system state (5.1a) may *not* be measured in real-world scenario. We extend the sparse *state feedback* result in Section 5.1.1 to *output feedback* case. Given a discrete LTI system (5.1), there exists a static output feedback $F \in \mathbb{R}^{m \times p}$ such that

$$u(t) = Fy(t), \tag{5.13}$$

and the resultant discrete-time feedback response

$$x(t + 1) = \hat{A}_{cl}x(t), \quad \hat{A}_{cl} \doteq A + BFC \in \mathbb{R}^{n \times n}, \tag{5.14}$$

with a closed-loop matrix $\hat{A}_{cl}$ is stable.

As similar to Problem 5.1, the goal is to *find a sparse static output feedback controller* (5.13) *and ensure the dynamics* (5.14) *stable*. Namely, determining sparse control signal for some output feedback gain $u = Fy$ can be converted into tackling a row-sparse matrix optimization $\|F\|_{1,\infty}$ under LMI stability constraints, the details will be presented in Proposition 5.3.

In view of the stability criteria (5.4) or (5.6), we take the output feedback control (5.13) by using $u = FCx$, then there exist $\hat{P} \succ 0$ and F such that output stabilization condition [27]

$$(A + BFC)^\top \hat{P}(A + BFC) - \hat{\rho}^2 \hat{P} \prec 0,$$

holds, in which $\hat{\rho} = \max_i |\lambda_i(\hat{A}_{cl})|$ denotes the spectral radius of the matrix $\hat{A}_{cl}$ in (5.14).

Alternatively, we recommend a new variable $\hat{Q} = \hat{P}^{-1}$, then the LMI reformulation is as follows

$$\begin{bmatrix} \hat{Q} & \hat{Q}A^\top + \hat{Q}C^\top F^\top B \\ A\hat{Q} + BFC\hat{Q} & \hat{\rho}^2\hat{Q} \end{bmatrix} \succ 0, \qquad \hat{Q} \succ 0. \tag{5.15}$$

However, computing the above LMI (5.15) is difficult because it involves *nonconvex* and *bilinear matrix inequality* (BMI). A replaced convex one for (5.15) was proposed in [51], that is, given system model (A, B, C), where C has *full row rank*. If feasible matrix solutions $\hat{Q}, \hat{V}, \hat{R}$ can be solved by LMIs

$$\begin{bmatrix} -\hat{Q} & \hat{Q}A^\top + C^\top \hat{V}^\top B^\top \\ A\hat{Q} + B\hat{V}C & -\hat{\rho}^2\hat{Q} \end{bmatrix} \prec 0, \qquad \hat{R}C = C\hat{Q}, \tag{5.16}$$

then a static output feedback control $F = \hat{V}\hat{R}^{-1}$ stabilizes the discrete LTI systems (5.1). We then assert the following result.

Proposition 5.3 (Sparse Output Feedback Control) *For the discrete LTI systems (5.1), identifying a stable and sparse static output feedback control $u = Fy$ equals to solving the solutions $\hat{V}, \hat{Q}, \hat{R}$ of the following row-sparse matrix optimization*

$$\min_{\hat{V},\hat{Q},\hat{R}} \quad \|\hat{V}\|_{1,\infty} \quad s.t. \quad LMIs\ (5.16). \tag{5.17}$$

Specially, if take $C = I \in \mathbb{R}^{n\times n}$, then we have $FC = K$, and the output feedback control reduces to state feedback control.

5.2 Data-driven meets sparse feedback control

In what follows, we discard the "model-based" frameworks, namely, assume that the knowledge of "true" system matrices (A, B, C) of LTI dynamics (5.1) is partially or even completely *unknown* (aka., "*model-free*", "*black-box*"). A state-of-the-art method is direct *data-driven technique* that the controlled LTI behavioral systems (5.1) can be *directly reconstructed through a finite samples of experimental input-state/output data trajectories*, stemmed from Willems' fundamental lemma [54, 96, 137].

For the sake of simplicity, throughout this chapter, let $w : \mathbb{Z} \to \mathbb{R}^p$ be a signal, and use $w_{[i,j]}$ as shorthand for the signal of w to the interval $[i, j]$, where $i, j \in \mathbb{Z}$, namely, $w_{[i,j]} = \begin{bmatrix} w^\top(i) \ w^\top(i+1) \ \cdots \ w^\top(j) \end{bmatrix}^\top$. Then, we further define a *Hankel matrix* of depth L associated with $w_{[i,j]}$ as

$$W_{i,L,j} = \begin{bmatrix} w(i) & w(i+1) & \cdots & w(j-L+1) \\ w(i+1) & w(i+2) & \cdots & w(j-L+2) \\ \vdots & \vdots & & \vdots \\ w(i+L-1) & w(i+L) & \cdots & w(i+L+j-2) \end{bmatrix}, \tag{5.18}$$

where the block rows of $W_{i,L,j}$ correspond to the windows of length L and satisfy $L \leq j - i + 1$, associated with signal $w_{[i,j]}$. When $L = 1$, the matrix

$$W_{i,j} = \begin{bmatrix} w(i) & w(i+1) & \cdots & w(i+j-1) \end{bmatrix}$$

has only one block row.

Definition 5.4 (Persistently Exciting) *A sequence $w_{[0,T-1]}$ is* persistently exciting *(PE) of order L if its Hankel matrix $W_{0,L,T-L+1}$ has full row rank.*

5.2.1 Data-driven behavioral system of LTI dynamics

We now harvest a finite input/state/output measurements over time T from actual system (5.1), recorded by the following matrices

$$\begin{aligned} U_{0,T} &= \begin{bmatrix} u_d(0) & u_d(1) & \cdots & u_d(T-1) \end{bmatrix} \in \mathbb{R}^{m \times T}, \\ X_{0,T} &= \begin{bmatrix} x_d(0) & x_d(1) & \cdots & x_d(T-1) \end{bmatrix} \in \mathbb{R}^{n \times T}, \\ Y_{0,T} &= \begin{bmatrix} y_d(0) & y_d(1) & \cdots & y_d(T-1) \end{bmatrix} \in \mathbb{R}^{p \times T}, \\ \hat{X}_{0,T} &= \begin{bmatrix} \chi_d(0) & \chi_d(1) & \cdots & \chi_d(T-1) \end{bmatrix} \in \mathbb{R}^{\hat{n} \times T}, \end{aligned} \tag{5.19}$$

where the subscript d in u_d, x_d, y_d, χ_d stands for the data samples collected from the system performing a single experiment in offline. Besides, $X_{1,T} = [x_d(1) \ \cdots \ x_d(T)]$ denotes forward step of matrix $X_{0,T}$.

Offline experiments (5.19) can be assembled in Hankel matrices (5.18), and the next rank condition (5.20) claims that if PE input data is sufficiently rich, adopting $T \geq (p+1)L - 1$, then the windows of the signal could span the whole system behavioral trajectories. We then show a celebrated Willems' fundamental lemma [137].

Lemma 5.4 (Fundamental Lemma [54]) *Given the LTI systems* (5.1)*. If the input signal $u_{d,[0,T-1]}$ is persistently exciting of order $n + L$, then*

(i) A full row rank condition satisfies

$$\text{rank}\left(\begin{bmatrix} U_{0,L,T-L+1} \\ X_{0,T-L+1} \end{bmatrix} \right) = Lm + n. \tag{5.20}$$

(ii) Any input-output signal $\{\bar{u}_{[0,L-1]}, \bar{y}_{[0,L-1]}\}$ is L-length feasible trajectories of LTI systems (5.1) if and only if *there exists a $g \in \mathbb{R}^{T-L+1}$ such that a linear combination*

$$\begin{bmatrix} \bar{u}_{[0,L-1]} \\ \bar{y}_{[0,L-1]} \end{bmatrix} = \begin{bmatrix} U_{0,L,T-L+1} \\ Y_{0,L,T-L+1} \end{bmatrix} g. \tag{5.21}$$

Clearly, Lemma 5.4 establishes a *data-based representation* for the model-free LTI systems (5.1). More precisely, if input signals satisfy PE [96], then the space of any input-output trajectories of the systems can be spanned by a finite time-shift collection of data sampls, stored in the Hankel matrices through some trials. Note that the condition related to (5.21) is also effective for any "input-state" trajectories of the system behaviors (5.1a).

A PE input experiment $u_{d,[0,T-1]}$ of order $n + L$ with $L = 1$ should be sufficient long, i.e., $T \geq (m+1)n + m$. From (5.20),

$$\text{rank}\left(\begin{bmatrix} U_{0,1,T} \\ X_{0,T} \end{bmatrix}\right) = n + m, \tag{5.22}$$

where $U_{0,1,T} = U_{0,T}$, and thus a compact state evolution of a discrete LTI plant (5.1a) with T-length window is rephrased as

$$X_{1,T} = AX_{0,T} + BU_{0,1,T} = \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} U_{0,1,T} \\ X_{0,T} \end{bmatrix}. \tag{5.23}$$

Let us now turn to the central problem of this chapter that how to develop the sparse feedback control (see Section 5.1) for the discrete LTI systems from model-based to data-based methodology, i.e., learning the dynamics and controller from data. We aim to provide the *data-driven sparse feedback controllers* for "unknown" LTI plant (5.1) by purely leveraging input-state/output data samples built in Section 5.2.1.

5.2.2 Data-driven sparse state feedback controller

Different from the model-based controller in Section 5.1.1, we here propose a data-driven sparse state feedback control for black-box LTI dynamics (5.1a). With the help of Lemma 5.4 and data-enabled behavior (5.23), it is sufficient to establish the closed-loop data-driven representation of (5.3), stated as follows:

Lemma 5.5 (Closed-Loop Data-Driven Representation [54]) *Let Lemma 5.4 holds true. Then, the discrete LTI system (5.1a) for some state feedback $u = Kx$ (5.2) in closed-loop response (5.3) is equivalent to the following data-based closed-loop behavior*

$$x(t+1) = A_{cl}x(t) = X_{1,T}G_K x(t), \tag{5.24}$$

in which $A_{cl} \doteq X_{1,T}G_K$. Based on Rouché-Capelli theorem, there exists a matrix $G_K \in \mathbb{R}^{T\times n}$, such that it satisfies

$$\begin{bmatrix} K \\ I_n \end{bmatrix} = \begin{bmatrix} U_{0,1,T} \\ X_{0,T} \end{bmatrix} G_K, \tag{5.25}$$

then a data-driven state feedback control can be formulated as

$$u(t) = U_{0,1,T}G_K x(t). \tag{5.26}$$

Combing the Proposition 5.2 and Lemma 5.5, we give rise to the following result of data-driven sparse state feedback controller.

Theorem 5.6 (Data-Driven Sparse State Feedback Control) *Let Lemma 5.4 with condition (5.22) hold. Then, the data-driven sparse state feedback for Problem 5.1 can be obtained via program (5.7) and Lemma 5.5, and the data-driven description $K = U_{0,1,T}V(X_{0,T}V)^{-1}$ ensures that the closed-loop process (5.24) is stable, where V is a row sparse matrix that optimizes*

$$\begin{aligned} \min_{V,Q} \quad & \|U_{0,1,T}V\|_{1,\infty} \\ \text{s.t.} \quad & \begin{bmatrix} X_{0,T}V & X_{1,T}V \\ V^\top X_{1,T}^\top & \bar{\rho}^2 X_{0,T}V \end{bmatrix} \succeq 0, \\ & \bar{\rho} = \max_i |\lambda_i(X_{1,T}V(X_{0,T}V)^{-1})| < 1. \end{aligned} \tag{5.27}$$

Proof 5.2 By Lemma 5.5, the data-based closed-loop description (5.24) with a linear static state feedback control $u = Kx$ is Schur α-stable (or Lyapunov stable) if and only if there exists a matrix variable $Q \succ 0$ with $Q = P^{-1}$ such that

$$X_{1,T}G_K Q G_K^\top X_{1,T}^\top - \bar{\rho}^2 Q \preceq 0, \tag{5.28}$$

in which the matrix G_K meets the equality (5.25). Define a new matrix variable $V \doteq G_K Q$, the stability implies that the matrices V and $Q \succ 0$ exist and admit the LMIs

$$X_{1,T}VQ^{-1}V^\top X_{1,T}^\top - \bar{\rho}^2 Q \preceq 0, \tag{5.29a}$$
$$X_{0,T}V = Q, \tag{5.29b}$$
$$U_{0,1,T}V = KQ, \tag{5.29c}$$

where the last two inequalities (5.29b), (5.29c) are derived from expression (5.25). Especially, the stability relies on the existence of a matrix V under constraint (5.29b), which further yields

$$X_{1,T}V(X_{0,T}V)^{-1}V^\top X_{1,T}^\top - \bar{\rho}^2 X_{0,T}V \preceq 0, \tag{5.30a}$$
$$X_{0,T}V \succ 0, \tag{5.30b}$$
$$U_{0,1,T}V = KX_{0,T}V. \tag{5.30c}$$

The term (5.30a) is based on the facts (5.29a) and (5.29b), thus data-driven feedback gain is taken as $K = U_{0,1,T}V(X_{0,T}V)^{-1}$.

From (5.25), a data-assistant testing for spectral radius of A_{cl} is offered by

$$\rho = \max_i |\lambda_i(X_{1,T}V(X_{0,T}V)^{-1})| < 1.$$

More concretely, we have $A + BK = X_{1,T}V(X_{0,T}V)^{-1}$ and $G_K = V(X_{0,T}V)^{-1}$. Meanwhile, the control cost requires the feedback matrix to impose row-sparsity, i.e., $\|K\|_{1,\infty} =$

$\|U_{0,1,T}G_K\|_{1,\infty} = \|U_{0,1,T}VQ^{-1}\|_{1,\infty}$ ■

5.2.3 Data-driven sparse output feedback controller

Motivated by the synthesis of data-driven output feedback control [54, 70], the MIMO state-space model (5.1) and (5.13) can be rewritten in an input-output signal with an ARX form [111]

$$\begin{aligned} y(t) + \alpha_n y(t-1) + \cdots + \alpha_2 y(t-n+1) + \alpha_1 y(t-n) \\ = \beta_n u(t-1) + \cdots \beta_2(t-n+1) + \beta_1 u(t-n), \end{aligned} \tag{5.31}$$

where $\alpha_i \in \mathbb{R}^{aa}$

$$\begin{aligned} \chi(t) = \big[&u^\top(t-n), u^\top(t-n+1), \cdots, u^\top(t-1), \\ &y^\top(t-n), y^\top(t-n+1), \cdots, y^\top(t-1)\big]^\top \end{aligned} \tag{5.32}$$

where the state $\chi(t) \in \mathbb{R}^{\hat{n}}$ with $\hat{n} = (m+p)n$ consists of the input and output vectors

Lemma 5.7 *Suppose that the PE condition with $L = 1$ holds. A data-driven representation of the discrete time closed-loop system* (5.14) *consisting of the LTI dynamics* (5.1) *and output feedback control $u = Fy$* (5.13) *is given by*

$$x(t+1) = \hat{X}_{n+1,T+1}G_F x(t) \tag{5.33a}$$

$$y(t) = Y_{n,T}G_F x(t), \tag{5.33b}$$

with a matrix $G_F \in \mathbb{R}^{(T-n+1)\times n}$ meeting

$$\begin{bmatrix} FC \\ I \end{bmatrix} = \begin{bmatrix} U_{n,T} \\ \hat{X}_{n,T} \end{bmatrix} G_F. \tag{5.34}$$

The proof sketch is similar to that of [54, Th. 2]. In fact, we have $I_n = \hat{X}_{n,T}G_F$ and $Y_{n,T} = C\hat{X}_{n,T}$. Hence, we get $C = Y_{n,T}G_F$.

5.3 Numerical benchmark

In this section, we conduct the numerical benchmarks for synthesizing data-driven sparse feedback control.

We discretize a continuous-time helicopter model (HE) with eighth-order and four inputs called "HE4" from the *COMPL$_e$ib* benchmark example [81, Chapter 2.1.2].

$$\dot{x}_{\mathrm{HE4}}(t) = Ax(t) + Bu(t)$$

which is the *eighth-order* linear model of a two-engine multipurpose helicopter. For this model, the sampling interval takes $\delta = 0.2$ and employs $A_d = e^{A\delta}$, $B_d = \int_0^\delta e^{A\tau}d\tau B$.

Given a degree of stability $\alpha = 0.05$ (i.e., $\bar{\rho} = 0.95$), we solve program (7) to obtain the following numerical values of the maxima for the rows of the solution matrix V :

$$0.0286 \times \begin{bmatrix} 1.9 \times 10^{-5} & 1.0000 & 0.6538 & 8.0 \times 10^{-6} \end{bmatrix}.$$

The first and fourth rows are seen to be smaller in absolute value than the other two rows. We make them to zero and solve the stabilization problem (5.27), i.e., check the feasibility of the LMI (5.27) subject to the sparse structure of the matrix variable V. The sparse controller thus obtained is determined by the row-sparse matrix K_d and the degree of stability of the related closed-loop system is equal to $\alpha = 0.0565$; that is, it is no less than the prespecified value.

Note that a solution of sparsity-free stabilization (5.27) is given by a feasible controller K

$$K_\mathrm{d} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -58.9943 & 3.5782 & -0.1078 & -6.3773 & -0.3368 & 6.1727 & 0.2342 & 0.2592 \\ 0.5493 & 27.9995 & 1.4598 & -0.7429 & 1.0118 & -0.0482 & 2.2801 & 0.0545 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$K = \begin{bmatrix} -0.3015 & 0.1713 & 0.0191 & -0.0616 & -0.0335 & -0.0263 & -0.0281 & 1.0043 \\ -58.8741 & 1.8028 & -0.1973 & -6.3837 & -0.2185 & 6.1438 & 0.0126 & 0.3935 \\ 0.7322 & 22.6563 & 1.1560 & -0.6804 & 1.9832 & -0.0416 & 1.7090 & 0.1365 \\ 3.3135 & -56.1465 & -1.9597 & 0.3678 & 9.6554 & -0.1794 & -11.0744 & 0.7465 \end{bmatrix},$$

whose nonzero elements are of the same order of magnitude as those of the controller K_d.

5.4 Summary

This chapter introduces data-driven sparse feedback controllers tailored for model-free LTI systems. In this approach, the static state feedback controller was derived from input-state data samples, while the state-feedback matrix was penalized based on structured row-sparsity. Moreover, the method ensured Schur-α stability for the synthesized data-driven sparse feedback controller. Finally, the numerical benchmarks illustrated the effectiveness of the proposed control policy.

Chapter 6

Linear Quadratic Sparse Optimal Control

In this chapter, we propose a novel linear quadratic (LQ) tracking control problem for the continuous-time systems with sparsity regularization using $\mathcal{L}_0$ norm, we refer to as LQ hands-off control, also known as LQ sparse optimal control. Sparsity regularization leads to sparse control, which has a significant length of time over which the control is exactly zero. Since the $\mathcal{L}_0$ cost is non-convex and discontinuous, we introduce $\mathcal{L}_1$ relaxation to make the optimization numerically tractable. The numerical solution is obtained with the aid of time-discretization to derive a discrete-time optimal control problem with ℓ_1 regularization. We also give an upper bound of the terminal state under perturbations in the initial states and the state-space matrices. Numerical examples illustrate the effectiveness of the proposed control.

6.1 Problem formulation

Let us consider the tracking problem, with a continuous-time master system

$$\dot{z}_{\mathrm{m}}(t) = Az_{\mathrm{m}}(t), \quad t \geq 0, \quad z_{\mathrm{m}}(0) = \xi_{\mathrm{m}} \in \mathbb{R}^n \tag{6.1}$$

and the corresponding slave system

$$\dot{z}_{\mathrm{s}}(t) = Az_{\mathrm{s}}(t) + Bu(t), \quad t \geq 0, \quad z_{\mathrm{s}}(0) = \xi_{\mathrm{s}} \in \mathbb{R}^n \tag{6.2}$$

where $z_{\mathrm{s}}(t), z_{\mathrm{m}}(t) \in \mathbb{R}^n$ are the state variables and $u(t) \in \mathbb{R}$ is a single control input. We assume the two systems share the same matrix A. In the case where there is a gap, we will discuss the robustness of the control system in Section 6.4.

The problem of the tracking control is to seek a control $u(t)$ that achieves tracking for given initial states $z_{\mathrm{s}}(0) = \xi_{\mathrm{s}}$ and $z_{\mathrm{m}}(0) = \xi_{\mathrm{m}}$. To consider this, we define the tracking state error by

$$x(t) \triangleq z_{\mathrm{s}}(t) - z_{\mathrm{m}}(t). \tag{6.3}$$

Then, from (6.1) and (6.2), we have the following linear model:

$$\dot{x}(t) = Ax(t) + Bu(t). \tag{6.4}$$

We here consider a finite horizon control problem. Namely, we find a finite-horizon control $\{u(t) : 0 \le t \le t_f\}$ with $t_f > 0$ that achieves

$$x(t_f) = z_{\mathrm{s}}(t_f) - z_{\mathrm{m}}(t_f) = 0, \tag{6.5}$$

from a given initial state

$$x(0) = \xi = \xi_{\mathrm{s}} - \xi_{\mathrm{m}}. \tag{6.6}$$

Also, it is practical to assume that $u(t)$ is bounded, that is,

$$|u(t)| \le 1, \quad \forall t \in [0, t_f]. \tag{6.7}$$

A control $u(t)$, $t \in [0, t_f]$ is said to be *feasible* if the control satisfies (6.4), (6.5), (6.6), and (6.7). We assume there exists at least one admissible control since the feasible region is *non-empty* in general.

The LQ (Linear Quadratic) tracking control is the control that minimizes the following LQ cost function among feasible controls:

$$J_{LQ}(u) = \frac{1}{2}\int_0^{t_f} \{x(t)^\top Qx(t) + Ru(t)^2\}dt, \tag{6.8}$$

with positive semi-definite matrix $Q \succeq 0$ and positive number $R > 0$.

To promote the sparsity of the control input, we penalize the $\mathcal{L}_0$ norm on the control signal, we refer to see Definition 1.4 in Chapter 1.

To promote the sparsity of the control input, we recommend to penalize the $\mathcal{L}_0$ norm to the cost function as a regularization term:

$$J_0(u) = J_{LQ}(u) + \lambda\|u\|_0, \tag{6.9}$$

where $\|u\|_0$ is the $\mathcal{L}_0$ norm and $\lambda > 0$ is the penalty weight [105]. We refer to this proposed control policy as *LQ hands-off control*.

In summary, our control problem is formulated as follows:

Problem 6.1 (LQ Hands-off Control) *Assume that $t_f > 0$ and $\xi \in \mathbb{R}^n$ are given and there exists at least one feasible control. Find an optimal control that solves the following optimal control problem:*

$$\begin{aligned} \min_{x,u} \quad & \frac{1}{2}\int_0^{t_f} \{x(t)^\top Qx(t) + Ru(t)^2\}dt + \lambda\|u\|_0 \\ \text{s.t.} \quad & \dot{x}(t) = Ax(t) + Bu(t), \quad t \in [0, t_f] \end{aligned}$$

$$\begin{aligned} & x(0) = \xi, \quad x(t_f) = 0 \\ & |u(t)| \leq 1, \quad \forall t \in [0, t_f]. \end{aligned} \tag{6.10}$$

For the sake of simplicity, we here consider a single-input plant, which also provides an paradigm for extending to a multiple-input plant. In fact, there are some works [79, 105] have provided the approaches for multiplexing signals in sparse control. Meanwhile, sparse control for time-varying system is also an open and challenging issue for sparsity architecture, and it is possible to tackle this topic, see [5].

6.2 Characterization of LQ hands-off control

Here, we analyze the LQ hands-off control, the optimal solution of Problem 6.1. We first show necessary conditions for the optimal control. For this, we first define the Hamiltonian $H^\eta(x, p, u)$ by

$$H^\eta(x, p, u) \triangleq p^\top(Ax + Bu) - \eta\Lambda(x, u), \tag{6.11}$$

where $p \in \mathbb{R}^n$, $\eta \in \{0, 1\}$, and $\Lambda(x, u)$ is the running cost for (6.9) defined by

$$\Lambda(x, u) \triangleq \frac{1}{2}(x^\top Qx + Ru^2) + \lambda\phi_0(u). \tag{6.12}$$

Since $\Lambda(x, u)$ is not continuous in u, we adopt the non-smooth maximum principle [49]. The following gives necessary conditions for Problem 6.1.

Lemma 6.1 (Necessary condition) *Let (x^*, u^*) be a local minimizer for Problem 6.1. Then there exist an arc $\{p(t) \in \mathbb{R}^n : t \in [0, t_f]\}$ and a scalar $\eta \in \{0, 1\}$ satisfying the following properties:*

1. *the* **non-triviality condition:**

$$(\eta, p(t)) \neq 0, \quad \forall t \in [0, t_f]. \tag{6.13}$$

2. *the* **adjoint equation:**

$$\begin{aligned} \dot{p}(t) &= -\partial_x H^\eta(x^*(t), p(t), u^*(t)) \\ &= -A^\top p(t) + \eta Q x^*(t), \quad \forall t \in [0, t_f], \end{aligned} \tag{6.14}$$

3. *the* **maximum condition:**

$$\begin{aligned} u^*(t) &\in \arg\max_{u \in [-1,1]} H^\eta(x^*(t), p(t), u) \\ &= \arg\max_{u \in [-1,1]} f_0(t, u), \quad \forall t \in [0, t_f], \\ f_0(t, u) &\triangleq p(t)^\top Bu - \eta(\frac{1}{2}Ru^2 - \lambda\phi_0(u)). \end{aligned} \tag{6.15}$$

4. *the* **constancy of the Hamiltonian:**

$$H^{\eta}(x^*(t), p(t), u^*(t)) = h, \quad \forall t \in [0, t_f], \tag{6.16}$$

for some constant $h \in \mathbb{R}$.

Proof 6.1 This is a direct result from [49, Theorem 22.26]. ■

Remark 6.1 We notice that for fixed p, u, the function $x \mapsto H^{\eta}(x, p, u)$ is smooth, then we derive the adjoint state differential equation (6.14). However, if the controlled system (6.4) is *non-smooth*, then one would contain a differential inclusion, referred to as *adjoint inclusion*. This is reflected by the generalized gradient ∂_C that appears in the adjoint inclusion rather than the usual (sub)gradient ∂_x in adjoint equation (6.14).

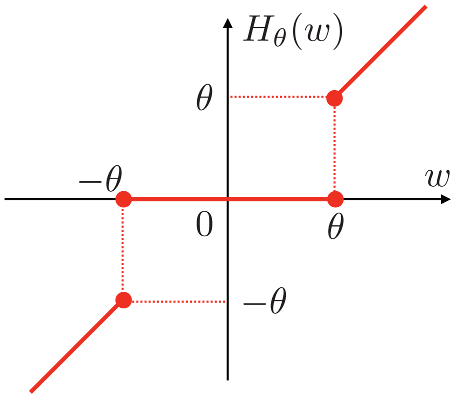


Figure 6.1: Hard-thresholding function $\mathbf{H}_\theta(w)$

Figure 6.1 shows the graph of the hard-thresholding operator. From the necessary conditions, we obtain an important property of LQ hands-off control.

Theorem 6.2 (LQ hands-off control) *The optimal control $u^*(t)$ (if it exists) satisfies*

1. *If $\eta = 1$, then*

$$u^*(t) = \mathrm{sat}\big(\mathbf{H}_\theta(R^{-1}B^\top p(t))\big), \tag{6.17}$$

where $\theta = \sqrt{2\lambda/R}$, $\mathrm{sat}(\cdot)$ is the saturation function defined by

$$\mathrm{sat}(v) \triangleq \begin{cases} -1, & \text{if } v < -1 \\ v, & \text{if } -1 \le v \le 1 \\ 1, & \text{if } v > 1. \end{cases} \tag{6.18}$$

and $\mathbf{H}_\theta(\cdot)$ is the hard-thresholding function *defined by*

$$\mathbf{H}_\theta(w) \triangleq \begin{cases} 0, & \text{if } -\theta < w < \theta \\ w, & \text{if } w < -\theta \text{ or } \theta < w, \end{cases} \tag{6.19}$$

and $\mathbf{H}_\theta(w) \in \{0, w\}$, if $w = \pm\theta$.

Unlike the (weighted) $\mathcal{L}_1$ optimal solution can be obtained for the *dead-zone* or *proximal* operator in [105], here the precise $\mathcal{L}_0$ (or cardinality) solution can be determined analytically for the hard-thresholding operator in [25]. From this theorem, the LQ hands-off control $u^*(t)$ may not be continuous in t while the mixed $\mathcal{L}_1$-$\mathcal{L}_2$ (LASSO) control [105] and the CLOT (Combined L-One and Two) control [101] are continuous.

Remark 6.2 We claim that the *abnormal case* (i.e., $\eta = 0$) does not arise in Problem 6.1. More specifically, the maximum condition (6.15) asserts that, the point $u^*(t)$ maximizes the function $f_0(t, u)$. If $\eta = 0$, then the solution form of adjoint function is $p(t) = e^{-A^\top t} p_0$ for some $p_0 \in \mathbb{R}^n$, which also obtain $B^\top p(t) = B^\top e^{-A^\top t} p_0 \equiv 0$. For then we would have $p_0 = 0$, whence $p \equiv 0$, violating the non-triviality condition (6.13). Hence, we only consider the normal case, that is $\eta = 1$.

In fact, when $\eta = 0$, it induces a piecewise constant control signal that takes ± 1 (i.e., bang-bang control). This control is not sparse at all, that is, $\|u^*\|_0 = t_f$. Typically, the abnormal case happens when the horizon length t_f is equal to the minimum time t_f^* to drive the state from the initial state ξ to the origin by a control that satisfies (6.7).

6.3 Numerical computation

In this section, we show a numerically tractable computation method for Problem 6.1. For this, we apply convex relaxation and time-discretization to convert the optimal control problem into a finite-dimensional convex optimization problem.

First, we adopt the convex relaxation using the $\mathcal{L}_1$ norm to avoid computational difficulties due to the non-convexity and discontinuity of the $\mathcal{L}_0$ norm in the cost function (6.9). Define $\mathcal{L}_1$ norm for a continuous-time control signal $u(t)$ over $[0, t_f]$ by

$$\|u\|_1 \triangleq \int_0^{t_f} |u(t)| dt.$$

Then, the relaxed $\mathcal{L}_1$ cost function with weight λ is given by

$$J_1(u) = \frac{1}{2} \int_0^{t_f} \left(x(t)^\top Q x(t) + R u(t)^2 \right) dt + \lambda \int_0^{t_f} |u(t)| dt \tag{6.20}$$

as a convex relaxation of (6.9). With the help of $J_1(u)$ proposed in (6.20), the relaxed LQ hands-off control problem reduces to the following optimal control problem.

Problem 6.2 (Relaxed LQ Hands-off Control) *Assume that $t_f > 0$ and $\xi \in \mathbb{R}^n$ are given and there exists at least one feasible control. Find a sparse optimal control that solves the following optimal control problem:*

$$\begin{aligned} \textit{minimize} \quad & \frac{1}{2} \int_0^{t_f} \left(x(t)^\top Q x(t) + R u(t)^2 \right) dt + \lambda \|u\|_1 \\ \textit{subject to} \quad & \dot{x}(t) = A x(t) + B u(t), \quad t \in [0, t_f] \\ & x(0) = \xi, \quad x(t_f) = 0 \end{aligned} \tag{6.21}$$

$$|u(t)| \leq 1, \quad \forall t \in [0, t_f].$$

Next, we apply time-discretization for Problem 6.2 to reduce the problem into a finite-dimensional convex optimization problem.

Let $m \in \mathbb{N}$ be the number of short intervals with which we divide the time interval $[0, t_f]$ into the following subintervals

$$[0, t_f] = [0, h) \cup [h, 2h) \cup \cdots \cup [(m-1)h, mh], \tag{6.22}$$

where $h = t_f/m$ is the discretization step. Then, we assume that the control is constant over each subinterval (i.e. the zero-order hold). Namely, we assume

$$u(t) = u_d^k, \quad t \in [kh, (k+1)h), \quad k = 0, 1, \ldots, m-1, \tag{6.23}$$

where $u_d^k \in [-1, 1]$ since the control constraints (6.7). This zero-order hold assumption is realistic since we often use a digital device to produce control signals. Then the feasible control is parametrized by a finite-dimensional vector $u_d \triangleq [u_d^0, u_d^1, \cdots, u_d^{m-1}]^\top$.

From the time discretization, the continuous-time system (6.4) is discretized as

$$x_d^{k+1} = A_d x_d^k + B_d u_d^k, \quad k = 0, 1, \ldots, m-1, \tag{6.24}$$

where $x_d^k = x(kh)$ for $k = 0, 1, \ldots, m-1$, and

$$A_d \triangleq e^{Ah}, \quad B_d \triangleq \int_0^h e^{At} B dt. \tag{6.25}$$

The terminal and initial conditions (6.5) and (6.6) are now described as

$$x_d^m = 0, \quad x_d^0 = \xi. \tag{6.26}$$

Also, the terminal state $x(t_f)$ is given by

$$x(t_f) = x_d^m = A_d^m \xi + \Upsilon_m u_d, \tag{6.27}$$

where

$$\Upsilon_m \triangleq \begin{bmatrix} A_d^{m-1} B_d, & A_d^{m-2} B_d, & \ldots, & B_d \end{bmatrix}.$$

Accordingly, the discretized feasible control set is defined as follows discretized as

$$\mathcal{U}_m(\xi) \triangleq \{ u_d \in \mathbb{R}^m : A_d^m \xi + \Upsilon_m u_d = 0, \ \|u_d\|_{\ell^\infty} \leq 1 \}. \tag{6.28}$$

Next, we discretize the quadratic term of the cost function $J_1(u)$ in (6.20). For this,

$$\begin{aligned}\phi_1 &\triangleq \frac{1}{2}\int_0^{t_f} (x^\top(t)Qx(t) + Ru(t)^2)dt \\ &= \frac{1}{2}\sum_{k=0}^{m-1}\int_{kh}^{(k+1)h} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^\top \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt \\ &= \frac{1}{2}\sum_{k=0}^{m-1}\int_0^h \begin{bmatrix} x(kh+\tau) \\ u(kh+\tau) \end{bmatrix}^\top \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} x(kh+\tau) \\ u(kh+\tau) \end{bmatrix} d\tau. \end{aligned} \tag{6.29}$$

From (6.23), we obtain

$$u(kh+\tau) = u_d^k, \quad \tau \in [0, h), \tag{6.30}$$

and also from the solution of (6.4), we have

$$x(kh+\tau) = e^{A\tau}x_d^k + \left(\int_0^\tau e^{A\sigma}Bd\sigma\right)u_d^k. \tag{6.31}$$

By substituting (6.30) and (6.31) into (6.29), we have

$$\phi_1 = \frac{1}{2}\sum_{k=0}^{m-1}\begin{bmatrix} x_d^k \\ u_d^k \end{bmatrix}^\top \begin{bmatrix} Q_d & S_d \\ S_d^\top & R_d \end{bmatrix} \begin{bmatrix} x_d^k \\ u_d^k \end{bmatrix} \tag{6.32}$$

where

$$\begin{bmatrix} Q_d & S_d \\ S_d^\top & R_d \end{bmatrix} \triangleq \int_0^h \Gamma^\top(\tau) \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \Gamma(\tau)d\tau, \tag{6.33}$$

and

$$\Gamma(\tau) \triangleq \begin{bmatrix} e^{A\tau} & \int_0^\tau e^{A\sigma}Bd\sigma \\ 0 & I \end{bmatrix}.$$

For the $\mathcal{L}_1$ term in (6.20), we have

$$\lambda\|u\|_1 = \lambda\sum_{k=0}^{m-1}\int_{kh}^{(k+1)h}|u(t)|\,dt = \lambda h\|u_d\|_{\ell_1}, \tag{6.34}$$

where $\|u_d\|_{\ell^1}$ is the ℓ_1 norm of vector u_d defined by

$$\|u_d\|_{\ell^1} \triangleq \sum_{k=0}^{m-1}|u_d^k|. \tag{6.35}$$

Finally, since the control is assumed to be piecewise constant, the magnitude constraint (6.7) on $u(t)$ is equivalently described as

$$|u_d^k| \le 1, \quad k = 0, 1, \ldots, m-1. \tag{6.36}$$

In summary, from (6.24), (6.26), (6.32), (6.34), and (6.36), we obtain the following finite-dimensional optimization problem:

$$\begin{aligned} \min_{x,u} \quad & \frac{1}{2}\sum_{k=0}^{m-1} \begin{bmatrix} x_d^k \\ u_d^k \end{bmatrix}^\top \begin{bmatrix} Q_d & S_d \\ S_d^\top & R_d \end{bmatrix} \begin{bmatrix} x_d^k \\ u_d^k \end{bmatrix} + \lambda \sum_{k=0}^{m-1} |u_d^k| \\ \text{s.t.} \quad & x_d^{k+1} = A_d x_d^k + B_d u_d^k, \quad k = 0, 1, \dots, m-1 \\ & x_d^0 = \xi, \quad x_d^m = 0 \\ & |u_d^k| \le 1, \quad k = 0, 1, \dots, m-1 \end{aligned} \tag{6.37}$$

The optimization problem in (6.37) is a convex optimization problem for finite-dimensional vector $u_d \in \mathbb{R}^m$, which is efficiently solved by numerical optimization toolboxes such as `CVX` or `YALMIP` with MATLAB [65, 66].

Remark 6.3 By the computational method proposed in this section, we obtain a finite-horizon control over time interval $[0, t_f]$. This is a feed-forward control using the initial state observation (6.6). In practice, it is preferable to implement the tracking control as a feedback control. To do this, one can adapt the receding horizon (or model predictive) control scheme [108, 104] or the self-triggered strategy [68, 105].

6.4 Robustness

We here consider the robustness of the control system when the initial state observation is perturbed, or there exists a gap between the A matrices of the master system (6.1) and the slave system (6.2).

6.4.1 Uncertainties in the initial states

First, we consider the case that there is uncertainty in the initial state in the master system (6.1) or the slave system (6.2). We assume that the initial states are perturbed as

$$z_{\mathrm{m}}(0) = z_{\mathrm{m}} + \delta_{\mathrm{m}}, \quad z_{\mathrm{s}}(0) = z_{\mathrm{s}} + \delta_{\mathrm{s}}, \tag{6.38}$$

where δ_{m} and δ_{s} are uncertain vectors in $\mathbb{R}^n$. Define

$$\delta \triangleq \delta_{\mathrm{s}} - \delta_{\mathrm{m}}. \tag{6.39}$$

Then, the initial state $x(0)$ in (6.4) is described as

$$x(0) = \xi + \delta. \tag{6.40}$$

Then we have the following lemma.

Lemma 6.3 *Let $u^*(t)$ be the hands-off control that solves Problem 6.2 with initial state $x(0) = \xi$. Let $x(t;\delta)$ denote the state variable of (6.4) with the optimal control u^* from the perturbed* *initial state in* (6.40). *Then we have*

$$\|x(t_f;\delta)\|_{\ell_2} \leq \|e^{At_f}\|\|\delta\|_{\ell_2}, \tag{6.41}$$

where $\|e^{At_f}\|$ *is the maximum singular value of* e^{At_f}, *and* $\|\delta\|_{\ell_2}$ *is the* ℓ_2 *norm of vector* δ *defined by* $\|\delta\|_{\ell_2} \triangleq \sqrt{\delta^\top \delta}$.

This lemma guarantees that the terminal tracking error $x(t_f;\delta)$ is bounded if the perturbation δ is bounded.

6.4.2 Gap between the A matrices

Now we consider the case when there is a gap between the A matrices in (6.1) and (6.2). We denote the gap matrix by Δ, that is, we consider the following state-space equation instead of (6.4):

$$\dot{x}(t;\Delta) = (A+\Delta)x(t;\Delta) + Bu(t). \tag{6.42}$$

We have the following lemma:

Lemma 6.4 *Let* $u^*(t)$ *be the hands-off control that solves Problem* 6.2 *for the ideal plant* (6.4). *Then we have*

$$\begin{aligned} \|x(t_f;\Delta)\|_{\ell_2} &\leq \alpha(\Delta)(\|\xi\|_{\ell_2} + \|B\|\|u^*\|_1) \\ &\leq \alpha(\Delta)(\|\xi\|_{\ell_2} + \|B\|\|u^*\|_0) \end{aligned} \tag{6.43}$$

where

$$\alpha(\Delta) \triangleq e^{\min\{\|A\|,\|A+\Delta\|\}t_f}(e^{\|\Delta\|t_f} - 1). \tag{6.44}$$

Proof 6.2 Since the optimal control u^* is a feasible control for the nominal system (6.4), it satisfies

$$x(t_f) = e^{At_f}\xi + \int_0^{t_f} e^{A(t_f-\tau)}Bu^*(\tau)d\tau = 0. \tag{6.45}$$

Then, from (6.42) and (6.45), we have

$$x(t_f;\Delta) = e^{(A+\Delta)t_f}\xi + \int_0^{t_f} e^{(A+\Delta)(t_f-\tau)}Bu^*(\tau)d\tau \tag{6.46}$$

and hence we obtain the terminal state gap

$$\begin{aligned} x(t_f;\Delta) &= x(t_f;\Delta) - x(t_f) \\ &= D(t_f)\xi + \int_0^{t_f} D(t_f-\tau)Bu^*(\tau)d\tau, \end{aligned} \tag{6.47}$$

where

$$D(t) \triangleq e^{(A+\Delta)t} - e^{At}. \tag{6.48}$$

From [18, Fact11.16.8], we have

$$\|D(t)\| \leq e^{\min\{\|A\|,\|A+\Delta\|\}t}(e^{\|\Delta\|t} - 1). \tag{6.49}$$

Then, from (6.47) and (6.49), we have

$$\begin{aligned}\|x(t_f;\Delta)\|_{\ell_2} \leq \|D(t_f)\|\|\xi\|_{\ell^2} \\ + \int_0^{t_f} \|D(t_f-\tau)\|\|B\||u^*(\tau)|d\tau \\ \leq \alpha(\Delta)(\|\xi\|_{\ell_2} + \|B\|\|u^*\|_1).\end{aligned} \tag{6.50}$$

Also, since u^* is a feasible control, it satisfies $|u^*(t)| \leq 1$ for all $t \in [0, t_f]$. It follows that

$$\|u^*\|_1 \leq \int_{\text{supp}(u^*)} 1 d\tau = \|u^*\|_0. \tag{6.51}$$

■

From Lemma 6.4, we can say that a smaller $\|u\|_1$ (or $\|u\|_0$) leads to a smaller upper bound of $\|x(t_f;\Delta)\|_{\ell_2}$. This is another merit of hands-off control in view of robustness against the gap between master and slave systems.

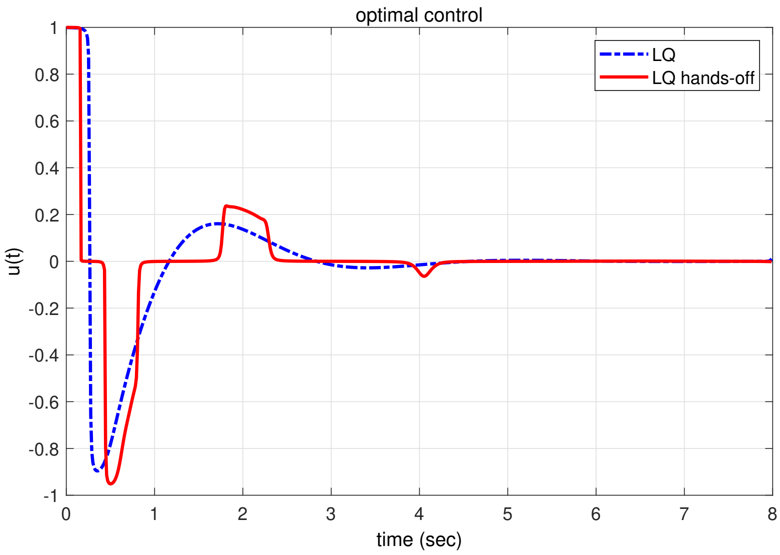


Figure 6.2: LQ control (dashed) and LQ hands-off control (solid).

6.5 Numerical examples

In this section, we give two numerical examples of the LQ hands-off control.

6.5.1 Nominal LQ hands-off control

We first consider the LQ hands-off control problem without perturbations, that is, nominal or deterministic case. The continuous dynamics (6.4) can be modeled as a fourth-order inverted pendulum-cart system with signal input

$$\begin{cases} \ddot{\epsilon} = \frac{(M_c+m_p)g}{M_c l_p}\epsilon - \frac{1}{M_c l_p}u, \\ \ddot{q} = -\frac{m_p g}{M_c}\epsilon + \frac{1}{M_c}u, \end{cases} \tag{6.52}$$

where ϵ is the pendulum angle referenced to the vertical axis, q is the cart position and u is the control force on the cart, i.e., $x = \begin{bmatrix}\theta & \dot{\theta} & q & \dot{q}\end{bmatrix}$. The typical parameters of system are selected as: cart's mass ($M_c = 2.4$ kg), pendulum's mass ($m_p = 0.23$ kg), pendulum's length ($l_p = 0.36$ m), cart track's length ($L_c = \pm 0.5$ m) and the gravity force ($g = 9.81$ m/s^2). The weights in the cost function (6.9) are chosen as: $Q = 5I$, $R = 1$, and the initial states are $x(0) = [-\pi/120, \pi/12, 0, 0]^\top$. We take the time horizon $t_f = 8$ s and the control constraint $|u(t)| \leq 1$. By performing time discretization based on ZoH sampling with step size $h = t_f/665$, we have a linearized discrete-time model of (6.24)

$$A_d = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 29.8615 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -0.9401 & 0 & 0 & 0 \end{bmatrix}, \quad B_d = \begin{bmatrix} 0 \\ -1.1574 \\ 0 \\ 0.4167 \end{bmatrix}.$$

We then convert the LQ hands-off control to the relaxed $\mathcal{L}^1$ optimal control (Problem 6.2), and solve the convex optimization (6.37) with `CVX` toolbox on MATLAB©[65, 66].

Fig. 6.2 shows the control signals of LQ control (i.e., $\lambda = 0$), and LQ hands-off control with weight $\lambda = 0.3$, respectively. We can see that LQ hands-off control is discontinuous at some time instants (that is, may not be continuous), but much sparser than the LQ control. Fig. 6.3 reflects that the tracking error trajectories $x_i(t)$, $i = 1, 2, 3, 4$ eventually convergence to zero under the proposed LQ hands-off control.

Fig. 6.4 shows the control inputs of the LQ hands-off control with different regularization weights $\lambda = 0.02$, 0.2, 2, and 20. We can see that the sparsity of the LQ control with $\mathcal{L}_1$ cost relying on the regularization weight λ. As the larger the weight λ of the regularization parameter become, the more the sparsity of LQ control is promoted, resulting in more zero inputs and, in turn, LQ hands-off (or sparse) control.

6.5.2 Perturbed plant

We next analyze the robustness of LQ hands-off control under uncertain factors. Let us consider a situation that there exists a gap between state matrices A in master-slaver system (6.1) and (6.2). For convenience, we focus on the related error system (6.42), and the dynamics is govern by a second-order linearized inverted pendulum system with

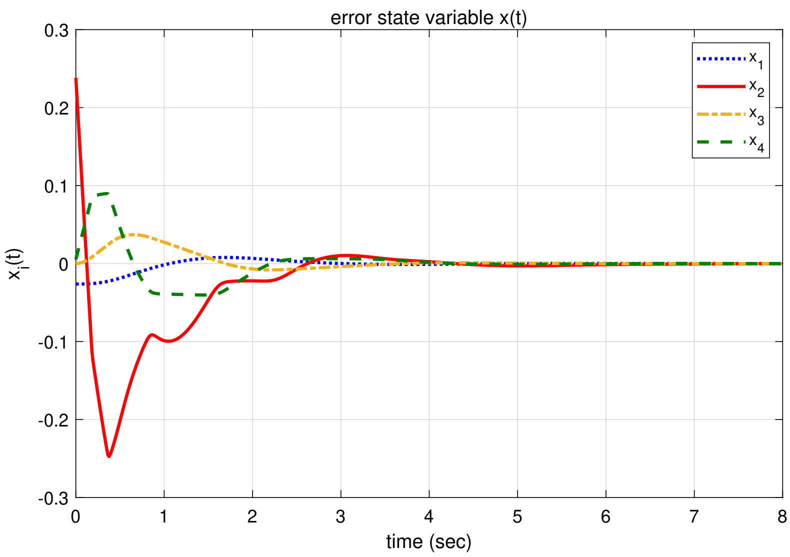


Figure 6.3: Tracking error states $x_i(t)$ by LQ hands-off control with $\lambda = 0.05$.

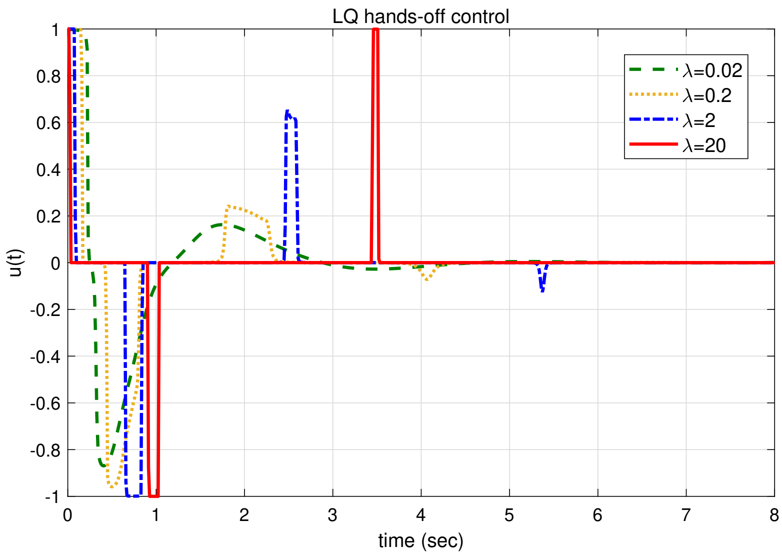


Figure 6.4: LQ hands-off control with different regularization weights.

a small disturbance

$$A = \begin{bmatrix} 0 & 1 \\ \frac{m_p l_p g}{L_p} & \frac{-b}{L_p} \end{bmatrix}, \ \Delta = \begin{bmatrix} 0 & 0 \\ 0 & 0.1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

The parameters of dynamics are chosen as: $m_p = 3$ kg, $l_p = 1.5$ m, $g = 9.81$ m/s^2, $L_p = m_p l_p^2/2$ kgm^2 and $b = 0.06$. Also, the weights of cost function are selected as: $\lambda = 1$, $Q = 3I$, $R = 1$ and the initial conditions $x(0) = [0.02, 0.1]^\top$. We take the terminal time $t_f = 1.2$ s and use the discretization step $h = t_f/500$ to execute the time-discretization

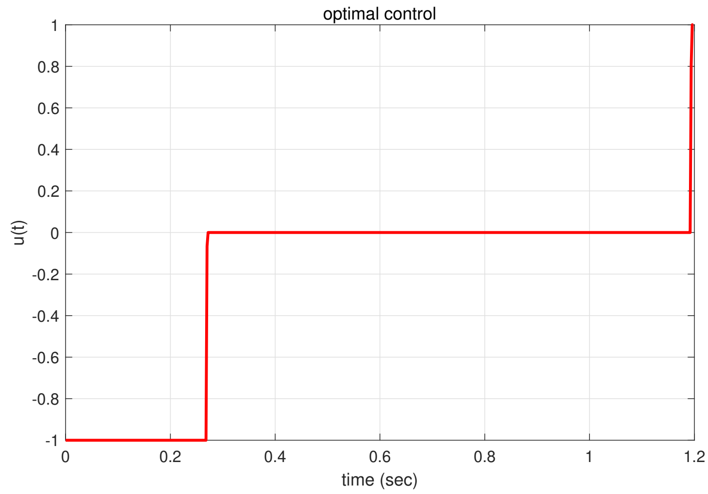


Figure 6.5: LQ hands-off control for the second-order plant

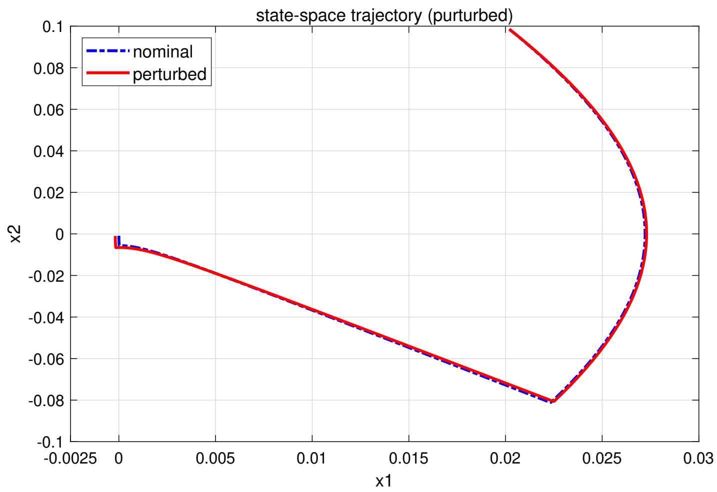


Figure 6.6: Phase portraits for the inverted pendulum with a gap matrix Δ.

procedure.

Fig. 6.5 shows the obtained LQ hands-off control signal, and it is sufficiently sparse thanks to the sparsity-promoting regularization. Then, Fig. 6.6 displays the phase portraits of the perturbed inverted pendulum system for LQ hands-off control with a bounded gap matrix Δ (i.e., $\|\Delta\| = 0.1$).

As shown in Lemma 6.4, the terminal state $x(t;\Delta)$ is bounded if Δ is bounded. The trajectory shown in Fig. 6.6 well illustrates this property. The gap $\|x(t_f) - x(t_f;\Delta)\|_{\ell_2} = \|x(t_f;\Delta)\|_{\ell_2}$ in this simulation is 0.0010, while the theoretical upper bound in (6.43) is 1.0273 in this case. This shows that the estimate (6.43) is very conservative.

6.6 Summary

In this chapter, we have investigated finite-horizon LQ tracking control with sparsity-promoting regularization on control using the $\mathcal{L}_0$ norm. We have shown necessary conditions for $\mathcal{L}_0$ optimality. For the computation of optimal control, we have proposed convex surrogate by using the $\mathcal{L}_1$ norm, instead of the non-convex and discontinuous $\mathcal{L}_0$ norm, which leads to finite-dimensional convex optimization after time discretization. We have also analyzed the robustness of the LQ hands-off control in the presence of perturbations in the initial states and the state-space matrices. Finally, simulation results have been shown to reveal the effectiveness of the proposed method. Future work includes the feedback control formulation of the LQ hands-off control.

Chapter 7

Conclusion and Future Research Directions

7.1 Summary

This dissertation investigated modeling, robustness and stability for sparse optimal control problems for dynamical systems. The constrained optimization problems find the control with minimum control efforts for given control objects, generating sparse optimal control inputs. The problems are defined on uncertain situations characterized as diverse robust counterparts, chance-constraints, and data-driven scenario realizations, compared to the deterministic (convex) optimization problem in the field of sparse modeling.

Previous research on sparse optimal control problems has extensively explored various methodologies, including open-loop dynamical system models, deterministic models, model-based frameworks, and LQ regulator approaches. Nevertheless, these investigations into sparse optimal control problems lack critical aspects such as the stability of closed-loop system models, the probabilistic robustness of stochastic models, the implementation of data-driven sparse control for model-free or black-box system models, and the Pontryagin's maximum principle of LQ sparse optimal control for continuous-time system models. As a result, this dissertation sets the stage for further exploration of modeling, robustness, and stability within the realm of sparse optimal control for dynamical systems. The detailed are summarized as follows:

Regarding the stability of the closed-loop system model, the optimal sparse feedback controller was derived from its original open-loop optimal solution by implementing a dynamic linear compensator. The implemented sparse feedback control can ensure initialization robustness, closed-loop stability, input sparsity and optimality guarantees, and the feedback solution was described in an explicit closed form. Moreover, this approach revealed the equivalent relationship between the sparse feedback control, open-loop optimal control and deadbeat control among the specific basis. Hence, the result was also useful to minimum attention feedback control in tracking problem.

From the perspective of probabilistic robustness in stochastic models, the risk-aware sparse optimal control was presented that achieves a state transition for a discrete-time uncertain/stochastic linear systems from an initial state near to a target state over a finite horizon. Taking probabilistic or chance-constrained form into account, the "hard" state constraints were relaxed as "soft" state constraints. To numerically compute the sparse solution, the main chance-constrained sparse optimization problem was approximated to the finite sparse random convex program for which convex relaxation data-driven scenario approach is applicable, and the feasibility of the approximation was then evaluated. The framework answered the question of how many finite samples (i.e., sample complexity) were needed to provide a prirori and high-confidence probabilistic robustness guarantees for the proposed risk-aware sparse control. Besides, the trade-off between the sparse cost performance and violated constraints of risk level was also discussed, which further offered a posterioi probabilistic robustness guarantee.

Additionally, a risk-aware sparse predictive control policy was developed to facilitate the generation of sparse predictive control inputs. This approach was achieved through the resolution of an online sparse random convex program, enabling the implementation of sparse predictive control. Notably, this sparse predictive control methodology accounted for external stochastic noise and model parametric uncertainties within the context of discrete-time dynamical systems.

Form the viewpoint of model-free problem setup, a data-driven sparse feedback control for model-free systems was proposed, compared to the previous chapters. The analysis, rooted in Willems' fundamental lemma, assumes that the length of input data samples is sufficiently long to satisfy the persistently exciting condition. Subsequently, this enables the derivation of a direct data-driven sparse feedback approach, and the synthesized data-driven sparse feedback control satisfies Schur-α stability. Sparse control input was attained by imposing a penalty on the structured row sparsity of the state-feedback matrix. The state feedback can be directly represented as experimental data samples of input-state trajectories, eliminating the necessity for prior knowledge of the system model.

In the realm of continuous-time modeling, an exploration of linear quadratic sparse optimal control has been undertaken for a master-slave tracking system. This differs from the previous chapters, which primarily focused on sparse optimal control for the discrete-time models. The non-smooth minimum principle was proposed to guarantee the theoretical results, implying that LQ sparse optimal control might not maintain continuity. Moreover, this well-defined problem leads to infinite-dimensional control problems, which can be computationally addressed through time-discretization techniques. Furthermore, LQ sparse optimal control demonstrates robustness in the face of worst-case uncertainties within the gaps of initial values and state-space matrices.

7.2 Future works

The proposed sparse optimal control methodologies can be further expanded to encompass large-scale network systems, including cyber-physical systems, multi-agent systems involving robots or vehicles, and power systems. This expansion involves designing a distributed sparse optimal control that not only adheres to structured sparsity but also accounts for input sparsity. This approach can be seen as a "spatial-temporal" control design, offering utility in diminishing communication flows and reducing control cost consumption. As society progresses into the era of Society 6.0, novel targets like achieving Net-zero emissions, developing smart cities, and embracing the Internet of Things (IoT) are emerging. These real-world applications present opportunities to model distributed control within networked systems. Here, each city or thing can be perceived as an agent or subsystem. Exploring distributed sparse optimal control methodologies within these complex network systems stands as an intriguing avenue for future research.

Although the sparse optimal control problems for model-based/free dynamical system with/without state-feedback response have been investigated in this dissertation, the extension to sparse output feedback case using relatively optimal control technique/data-driven method can be another significant works. Obviously, directly replicating the current results may be challenging due to the influence of generic initial conditions and the initialization of the dynamic output compensator, which are associated with the model matching problem.

Last but not least, the presented risk-aware sparse predictive control requires further exploration in theoretical aspects, particularly concerning recursive feasibility and (input-to-state) stability. This evaluation is crucial in assessing the effectiveness of sparse optimal control under probabilistic or chance-constraints. Although numerical benchmarks for risk-aware sparse predictive control have been demonstrated in this dissertation, the theoretical analysis is lacking due to the utilization of a data-driven scenario approach, which does not always ensure feasibility. Therefore, investigating and establishing theoretical frameworks for recursive feasibility and stability becomes essential for a comprehensive understanding of the proposed control method.

Appendix A

Sparsity

A.1 Null space recovery

Given a reachability matrix $\Phi_N \in \mathbb{R}^{n\times mN}$ with $mN \gg n$, a desired sparse control input $u \in \mathbb{R}^{mN}$ and an initial condition $x_0 \in \mathbb{R}^n$, which gives rise to a reference information vector $r \triangleq -A^N x_0 \in \mathbb{R}^n$ such that

$$r = -A^N x_0 = \Phi_N u.$$

We now recast the program 2.3 as a decoder (de-reachablility) $\Delta(\Phi_N u)$ that maps from $\mathbb{R}^n \to \mathbb{R}^{mN}$, with

$$\Delta(\Phi_N u) = \underset{\{u\in\mathbb{R}^{mN}:\Phi_N u=-A^N x_0\}}{\arg\min} \|u\|_1, \tag{A.1}$$

The typical paradigm for seeking a sparse control signal u is to Suppose that the desired control signal u is sparse, then the corresponding decoder requires to perfectly promote such a sparse inputs. More precisely, for a reachable matrix $\Phi_N \in \mathbb{R}^{n\times mN}$ and $0 < k < mN$, we can quantify the ℓ_1 error of a decoder $\Delta(\Phi_N u)$ by calculating the smallest constant $\mathscr{C} \geq 0$ such that for all $u \in \mathbb{R}^{mN}$

$$\|u - \Delta(\Phi_N u)\|_1 \leq \mathscr{C}\sigma_k(u), \tag{A.2}$$

where the ℓ_1 error of the best k-term approximation of signal u can be computed as the ℓ_1 norm of the $mN - k$ smallest coefficients of $u \in \mathbb{R}^{mN}$, defined by

$$\sigma_k(u) \triangleq \min_{\{v\in\mathbb{R}^{mN}:\|v\|_0=k\}} \|z - v\|_1$$

Definition A.1 (Null Space Property) *A (reachability) matrix $\Phi_N \in \mathbb{R}^{n\times mN}$ satisfies null space property (in ℓ_1 norm) of order k with constant $\mathscr{C}_k$ if and only if for all $v \in \mathbb{R}^{mN}$*

$$\|v\|_1 \leq \mathscr{C}_k \|v_{\mathcal{S}^c}\|_1 \tag{A.3}$$

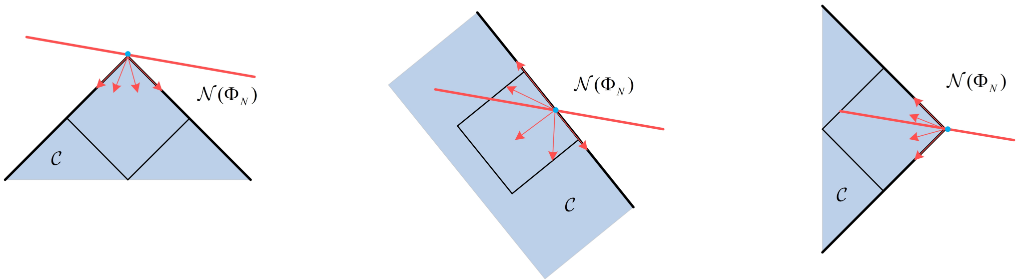


Figure A.1: Null space of Φ_N: left, middle, right

holds, where a null space $\mathcal{N}(\Phi_N) \triangleq \{v \in \mathbb{R}^{mN} : \Phi_N v = 0\}$ *and the index subset* $\mathcal{S} \subseteq \mathcal{T} = \{1, 2, \cdots, mN\}$ *of cardinality* $\text{card}(\mathcal{S}) \leq k$*, where* $\mathcal{S}^c$ *stands for the complement of* $\mathcal{S}$ *in* $\mathcal{T}$*.*

Suppose that we observe $r = \Phi_N$ without any reference information on u, which spans a null space $\mathcal{N}(\Phi_N)$ or kernel $\text{Ker}(\Phi_N)$,

From the perspective of geometry, we consider a tangent cone $\mathcal{T}_f(v)$ over f, that is, the set of descent directions of f at v, also known as *cone of descent*, defined as follows

$$\mathcal{T}_f(v) = \mathcal{C}(v) = \{e : f(v + te) \leq f(v), \text{ for some } t > 0\}.$$

For example taking $f = \|v\|_1$, we then have the result that we have

$$\mathcal{C}(v) \cap \mathcal{N}(\Phi_N) = \{0\}. \tag{A.4}$$

Definition A.2 (Helly's Theorem) *Let* $\Omega := \{\Omega_1, \cdots, \Omega_N\}$ *be a finite collection of convex subsets of* $\mathbb{R}^d$ *with* $N \geq d + 1$ *(i.e.,* $\text{card}(\Omega) \geq N + 1$*). If the intersection of these sets is nonempty, then the whole collection* Ω *haves a non-empty intersection. More formally*

$$\bigcap_{i=1}^{N} \Omega_i \neq \emptyset$$

A.2 Convex relaxation

Assume that $\|u\|_\infty \leq 1$ holds, then the biconjugate function of ℓ_0 norm gives the result

$$\|u\|_0^{\star\star} = \|u\|_1.$$

Indeed, its conjugate function is defined by

$$\|u\|_0^{\star}(v) := \sup_u \left\{ \langle u, v \rangle - \|u\|_0 \right\} = \max\left(\textstyle\sum_i |v_i|, 0 \right), \tag{A.5}$$

and it is always a convex form even the $\|x\|_0$ is non-convex. Besides, it derives the biconjugate case

$$\|u\|_0^{\star\star}(r) := \sup_v \{\langle r, v\rangle - \|u\|_0^{\star}(v)\} = \|r\|_1, \tag{A.6}$$

which concludes that ℓ_1 norm is a *convex relaxation* of ℓ_0 norm in some sense due to the fact that $\|u\|_0^{\star\star} \leq \|u\|_0$.

Bibliography

[1] N. N. Abdelmalek. "Solutions of minimum time problem and minimum fuel problem for discrete linear admissible control systems". In: *I. J. Sys. Sci.* 9.8 (1978), pp. 849–855.

[2] C. O. Aguilar and A. J. Krener. "Numerical solutions to the Bellman equation of optimal control". In: *J. Optim Theory Appl.* 160.2 (2014), pp. 527–552.

[3] R. P. Aguilera, G. Urrutia, R. A. Delgado, D. Dolz, and J. C. Agüero. "Quadratic model predictive control including input cardinality constraints". In: *IEEE Trans. Autom. Contr.* 62.6 (2017), pp. 3068–3075.

[4] T. Alamo, R. Tempo, and A. Luque. "On the sample complexity of randomized approaches to the analysis and design under uncertainty". In: *Proceedings of the 2010 American Control Conference*. IEEE. 2010, pp. 4671–4676.

[5] W. Alt and C. Schneider. "Linear-quadratic control problems with L1-control cost". In: *Optimal control applications and methods* 36.4 (2015), pp. 512–534.

[6] J. Anderson, J. C. Doyle, S. H. Low, and N. Matni. "System level synthesis". In: *Annual Reviews in Control* 47 (2019), pp. 364–393.

[7] M. Athans and P. L. Falb. *Optimal Control*. Dover Publications, 1966.

[8] M. Babazadeh and A. Nobakhti. "Sparsity promotion in state feedback controller design". In: *IEEE Trans. Autom. Contr.* 62.8 (2016), pp. 4066–4072.

[9] G. Baggio, D. S. Bassett, and F. Pasqualetti. "Data-driven control of complex networks". In: *Nature Communi.* 12.1 (2021), pp. 1–13.

[10] E. Bakolas. "On the computation of sparse solutions to the controllability problem for discrete-time linear systems". In: *J. Optim. Theory Appl.* 183.1 (2019), pp. 292–316.

[11] G. Basile and G. Marro. *Controlled and conditioned invariants in linear system theory*. Prenticle-Hall, 1992.

[12] A. Ben-Israel and T. N. Greville. *Generalized inverses: theory and applications*. Vol. 15. Springer Science & Business Media, 2003.

[13] A. Ben-Tal, D. Den Hertog, and J.-P. Vial. "Deriving robust counterparts of nonlinear uncertain inequalities". In: *Math. Program.* 149.1 (2015), pp. 265–299.

[14] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Vol. 28. Princeton Press, Princeton, 2009.

[15] J. Berberich and F. Allgöwer. "A trajectory-based framework for data-driven system analysis and control". In: *2020 Europ. Contr. Conf.* 2020, pp. 1365–1370.

[16] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer. "Data-driven model predictive control with stability and robustness guarantees". In: *IEEE Trans. Automa. Contr.* 66.4 (2020), pp. 1702–1717.

[17] S. Bernhard. "Time-Invariant Control in LQ Optimal Tracking: An Alternative to Output Regulation". In: *IFAC-Papers OnLine* 50.1 (July 2017), pp. 4912–4919.

[18] D. S. Bernstein. *Matrix Mathematics*. 2nd. Princeton University Press, 2009.

[19] D. Bertsimas, A. King, and R. Mazumder. "Best subset selection via a modern optimization lens". In: (2016).

[20] J. Blanchet and K. Murthy. "Quantifying distributional model risk via optimal transport". In: *Math. Oper. Res.* 44.2 (2019), pp. 565–600.

[21] F. Blanchini, P. Colaneri, Y. Fujisaki, S. Miani, and F. A. Pellegrino. "A Youla–Kučera parameterization approach to output feedback relatively optimal control". In: *Sys. & Control Lett.* 81 (2015), pp. 14–23.

[22] F. Blanchini, F. Dabbene, G. Fenu, F. A. Pellegrino, and E. Salvato. "Model-free feedback control synthesis from expert demonstration". In: *IEEE Contr. Sys. Lett.* (2023).

[23] F. Blanchini and F. A. Pellegrino. "Relatively optimal control and its linear implementation". In: *IEEE Trans. Autom. Contr.* 48.12 (2003), pp. 2151–2162.

[24] T. Blumensath and M. E. Davies. "Iterative hard thresholding for compressed sensing". In: *Applied and computational harmonic analysis* 27.3 (2009), pp. 265–274.

[25] T. Blumensath and M. E. Davies. "Iterative hard thresholding for compressed sensing". In: *Applied and Computational Harmonic Analysis* 27.3 (2009), pp. 265–274.

[26] S. Bourguignon, J. Ninin, H. Carfantan, and M. Mongeau. "Exact sparse approximation problems via mixed-integer programming: Formulations and computational performance". In: *IEEE Trans. Signal Proces.* 64.6 (2015), pp. 1405–1419.

[27] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear matrix inequalities in system and control theory*. SIAM, 1994.

[28] W. Brockett. "Minimum attention control". In: *Proceedings of the 36th IEEE Conference on Decision and Control*. Vol. 3. IEEE. 1997, pp. 2628–2632.

[29] J. Brodie, I. Daubechies, C. De Mol, D. Giannone, and I. Loris. "Sparse and stable Markowitz portfolios". In: *Proceedings of the National Academy of Sciences* 106.30 (2009), pp. 12267–12272.

[30] A. Budiyono and S. S. Wibowo. "Optimal Tracking Controller Design for a Small Scale Helicopter". In: *Journal of Bionic Engineering* 4.4 (2007), pp. 271–280.

[31] F. Bullo. *Lectures on network systems*. Vol. 1. 3. Kindle Direct Publishing Seattle, DC, USA, 2020.

[32] A. Bykov and P. S. Shcherbakov. "Sparse feedback design in discrete-time linear systems". In: *Autom. Remote Contr.* 79.7 (2018), pp. 1175–1190.

[33] G. C. Calafiore and L. Fagiano. "Stochastic model predictive control of LPV systems via scenario optimization". In: *Automatica* 49.6 (2013), pp. 1861–1866.

[34] G. C. Calafiore. "Random convex programs". In: *SIAM Journal on Optimization* 20.6 (2010), pp. 3427–3464.

[35] G. C. Calafiore and M. C. Campi. "The scenario approach to robust control design". In: *IEEE Trans. Automa. Contr.* 51.5 (2006), pp. 742–753.

[36] M. C. Campi and A. Caré. "Random convex programs with L_1-regularization: sparsity and generalization". In: *SIAM J. Contr. and Optim* 51.5 (2013), pp. 3532–3557.

[37] M. C. Campi and S. Garatti. *Introduction to the scenario approach*. SIAM, 2018.

[38] M. C. Campi and S. Garatti. "The exact feasibility of randomized solutions of uncertain convex programs". In: *SIAM J. Optimi.* 19.3 (2008), pp. 1211–1230.

[39] E. J. Candes. "The restricted isometry property and its implications for compressed sensing". In: *Comptes rendus. Mathematique* 346.9-10 (2008), pp. 589–592.

[40] E. J. Candes, M. B. Wakin, and S. P. Boyd. "Enhancing sparsity by reweighted ℓ_1 minimization". In: *Journal of Fourier analysis and applications* 14.5 (2008), pp. 877–905.

[41] E. J. Candès, J. Romberg, and T. Tao. "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information". In: *IEEE Transactions on information theory* 52.2 (2006), pp. 489–509.

[42] M. Caponigro, B. Piccoli, F. Rossi, and E. Trélat. "Sparse Jurdjevic–Quinn stabilization of dissipative systems". In: *Automatica* 86 (2017), pp. 110–120.

[43] A. Carè, S. Garatti, and M. C. Campi. "FAST—Fast algorithm for the scenario technique". In: *Operations Research* 62.3 (2014), pp. 662–671.

[44] F. Celi, G. Baggio, and F. Pasqualetti. "Data-driven Eigenstructure Assignment for Sparse Feedback Design". In: *arXiv:2303.18209* (2023).

[45] A. Charnes and W. W. Cooper. "Chance-constrained programming". In: *Manage. Sci.* 6.1 (1959), pp. 73–79.

[46] D. Chatterjee, M. Nagahara, D. E. Quevedo, and K. S. M. Rao. "Characterization of maximum hands-off control". In: *Systems & Control Letters* 94 (Aug. 2016), pp. 31–36.

[47] Z. Chen, S. Peng, and A. Lisser. "A sparse chance constrained portfolio selection model with multiple constraints". In: *J. Glob. Optim.* 77.4 (2020), pp. 825–852.

[48] T. Çimen and S. P. Banks. "Nonlinear optimal tracking control with application to super-tankers for autopilot design". In: *Automatica* 40.11 (2004), pp. 1845–1863.

[49] F. Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*. Springer, 2013.

[50] J. Coulson, J. Lygeros, and F. Dörfler. "Distributionally robust chance constrained data-enabled predictive control". In: *IEEE Trans. Automa. Contr.* 67.7 (2021), pp. 3289–3304.

[51] C. A. Crusius and A. Trofino. "Sufficient LMI conditions for output feedback control problems". In: *IEEE Trans. Autom. Contr.* 44.5 (1999), pp. 1053–1057.

[52] M. S. Darup, G. Book, D. E. Quevedo, and M. Nagahara. "Fast hands-off control using ADMM real-time iterations". In: *IEEE Trans. Autom. Contr.* 67.10 (2021), pp. 5416–5423.

[53] A. Das and D. Kempe. “Approximate submodularity and its applications: Subset selection, sparse approximation and dictionary selection”. In: *The Journal of Machine Learning Research* 19.1 (2018), pp. 74–107.

[54] C. De Persis and P. Tesi. “Formulas for data-driven control: Stabilization, optimality, and robustness”. In: *IEEE Trans. Autom. Contr.* 65.3 (2019), pp. 909–924.

[55] D. L. Donoho. “Compressed sensing”. In: *IEEE Trans. Inform. Theory* 52.4 (2006), pp. 1289–1306.

[56] E. Erdoğan and G. Iyengar. “Ambiguous chance constrained problems and robust optimization”. In: *Math. Program.* 107.1 (2006), pp. 37–61.

[57] S. Foucart. “Hard thresholding pursuit: an algorithm for compressive sensing”. In: *SIAM Journal on numerical analysis* 49.6 (2011), pp. 2543–2563.

[58] Y. Fujisaki, Y. Duan, and M. Ikeda. “System representation and optimal control in input-output data space”. In: *IFAC Proc. Vol.* 37.11 (2004), pp. 185–190.

[59] L. Furieri, Y. Zheng, A. Papachristodoulou, and M. Kamgarpour. “Sparsity invariance for convex design of distributed controllers”. In: *IEEE Trans. Contr. Netw. Syst.* 7.4 (2020), pp. 1836–1847.

[60] S. Garatti and M. C. Campi. “Risk and complexity in scenario optimization”. In: *Mathematical Programming* 191.1 (2022), pp. 243–279.

[61] E. Garone, S. Di Cairano, and I. Kolmanovsky. “Reference and command governors for systems with constraints: A survey on theory and applications”. In: *Automatica* 75 (2017), pp. 306–328.

[62] X. Geng and L. Xie. “Data-driven decision making in power systems with probabilistic guarantees: Theory and applications of chance-constrained optimization”. In: *Annual reviews in control* 47 (2019), pp. 341–363.

[63] E. G. Gilbert, I. Kolmanovsky, and K. T. Tan. “Discrete-time reference governors and the nonlinear control of systems with state and control constraints”. In: *International Journal of Robust and Nonlinear Control* 5.5 (1995), pp. 487–504.

[64] J.-y. Gotoh, A. Takeda, and K. Tono. “DC formulations and algorithms for sparse optimization problems”. In: *Mathematical Programming* 169 (2018), pp. 141–176.

[65] M. Grant and S. Boyd. “Graph implementations for nonsmooth convex programs”. In: *Recent Advances in Learning and Control*. Ed. by V. Blondel, S. Boyd, and H. Kimura. Vol. 371. Lecture Notes in Control and Information Sciences. Splinger, 2008, pp. 95–110.

[66] M. Grant and S. Boyd. *CVX: Matlab Software for Disciplined Convex Programming, version 2.1*. http://cvxr.com/cvx. Mar. 2014.

[67] M. Grant, S. Boyd, and Y. Ye. *CVX: Matlab software for disciplined convex programming*. 2008.

[68] W. P. M. H. Heemels, K. H. Johansson, and P. Tabuada. “An introduction to event-triggered and self-triggered control”. In: *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. 2012, pp. 3270–3285.

[69] R. R. Hocking and R. Leslie. “Selection of the best subset in regression analysis”. In: *Technometrics* 9.4 (1967), pp. 531–540.

[70] K. Hu and T. Liu. "Data-Driven Output Feedback Control for Unknown Switched Linear Systems". In: *IEEE Contr. Sys. Lett.* (2023).

[71] T. Ikeda and M. Nagahara. "Value function in maximum hands-off control for linear systems". In: *Automatica* 64 (Feb. 2016), pp. 190–195.

[72] T. Iwata, Y. Oishi, and M. Nagahara. "Realization of sparse control using the model predictive control scheme". In: *Transactions of the Society of Instrument and Control Engineers* 56.3 (2020), pp. 74–80.

[73] M. R. Jovanović and N. K. Dhingra. "Controller architectures: Tradeoffs between performance and structure". In: *Eur. J. Contr.* 30 (2016), pp. 76–91.

[74] T. Kailath. *Linear systems*. Vol. 156. Prentice-Hall Englewood Cliffs, NJ, 1980.

[75] V. Katewa and F. Pasqualetti. "Minimum-gain pole placement with sparse static feedback". In: *IEEE Trans. Automa. Contr.* 66.8 (2020), pp. 3445–3459.

[76] M. Kishida, M. Barforooshan, and M. Nagahara. "Hands-off control for discrete-time linear systems subject to polytopic uncertainties". In: *IFAC-PapersOnLine* 51.23 (2018), pp. 355–360.

[77] M. Kishida and M. Nagahara. "Risk-aware maximum hands-off control using worst-case conditional value-at-risk". In: *IEEE Trans. Automa. Contr.* (2023).

[78] M. V. Kothare, V. Balakrishnan, and M. Morari. "Robust constrained model predictive control using linear matrix inequalities". In: *Automatica* 32.10 (1996), pp. 1361–1379.

[79] Y. Kumar, S. Srikant, and D. Chatterjee. "Optimal multiplexing of sparse controllers for linear systems". In: *Automatica* 106 (2019), pp. 134–142.

[80] P. Lee and F. Park. "On the Existence and Computation of Minimum Attention Optimal Control Laws". In: *IEEE Trans. Automa. Contr.* 67.5 (2021), pp. 2576–2581.

[81] F. Leibfritz and W. Lipinski. *Description of the benchmark examples in COMPleib 1.0*. Tech. rep. 2003. URL: http://www.complib.de/.

[82] V. L. Levin. "Application of E. Helly's theorem to convex programming, problems of best approximation and related questions". In: *Math. USSR Sbornik* 8.2 (1969), p. 235.

[83] F. L. Lewis, D. Vrabie, and V. L. Syrmos. *Optimal control*. John Wiley & Sons, 2012.

[84] F. L. Lewis, D. L. Vrabie, and V. L. Syrmos. *Optimal Control*. 3rd. Chapter 4. Wiley, 2012.

[85] F. Lin, M. Fardad, and M. R. Jovanović. "Design of optimal sparse feedback gains via the alternating direction method of multipliers". In: *IEEE Trans. Autom. Contr.* 58.9 (2013), pp. 2426–2431.

[86] R. Y. Ling, Y. Feng, F. Claveau, and P. Chevrel. "Stochastic LQ control under asymptotic tracking for discrete systems over multiple lossy channels". In: *IET Control Theory & Application* 13.18 (Dec. 2019), pp. 3107–3116.

[87] G. P. Liu and R. Patton. *Eigenstructure assignment for control system design*. John Wiley & Sons, Inc., 1998.

[88] J. Lofberg. "YALMIP: A toolbox for modeling and optimization in MATLAB". In: *2004 IEEE international conference on robotics and automation (IEEE Cat. No. 04CH37508)*. IEEE. 2004, pp. 284–289.

[89] M. López and G. Still. "Semi-infinite programming". In: *European J. Oper. Res.* 180.2 (2007), pp. 491–518.

[90] M. Lorenzen, F. Dabbene, R. Tempo, and F. Allgöwer. "Stochastic MPC with offline uncertainty sampling". In: *Automatica* 81 (2017), pp. 176–183.

[91] J. Luedtke and S. Ahmed. "A sample approximation approach for optimization with probabilistic constraints". In: *SIAM J. Optim.* 19.2 (2008), pp. 674–699.

[92] J. Luedtke, S. Ahmed, and G. L. Nemhauser. "An integer programming approach for linear programs with probabilistic constraints". In: *Math. Program.* 122.2 (2010), pp. 247–272.

[93] J. M. Maciejowski. *Predictive control: with constraints*. Pearson education, 2002.

[94] J. Mairal. "Incremental majorization-minimization optimization with application to large-scale machine learning". In: *SIAM Journal on Optimization* 25.2 (2015), pp. 829–855.

[95] A. Mannava, S. N. Balakrishnan, L. Tang, and R. G. Landers. "Optimal Tracking Control of Motion Systems". In: *IEEE Trans. Control Systems Technology* 20.6 (Nov. 2012), pp. 1548–1558.

[96] I. Markovsky and P. Rapisarda. "Data-driven simulation and control". In: *Int. J. Contr.* 81.12 (2008), pp. 1946–1959.

[97] N. Matni and V. Chandrasekaran. "Regularization for design". In: *IEEE Trans. Automa. Contr.* 61.12 (2016), pp. 3991–4006.

[98] A. Mesbah. "Stochastic model predictive control: An overview and perspectives for future research". In: *IEEE Control Systems Magazine* 36.6 (2016), pp. 30–44.

[99] P. Mohajerin Esfahani and D. Kuhn. "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations". In: *Math. Program.* 171.1 (2018), pp. 115–166.

[100] P. Mohajerin Esfahani, T. Sutter, and J. Lygeros. "Performance bounds for the scenario approach and an extension to a class of non-convex programs". In: *IEEE Trans. Autom. Contr.* 60.1 (2015), pp. 46–58.

[101] M. Nagahara, D. Chatterjee, N. Challapalli, and M. Vidyasagar. "CLOT norm minimization for continuous hands-off control". In: *Automatica* 113.108679 (Mar. 2020), pp. 1–7.

[102] M. Nagahara. *Sparsity methods for systems and control*. Now Publishers, 2020.

[103] M. Nagahara and D. Nešić. "An approach to minimum attention control by sparse derivative". In: *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE. 2020, pp. 5005–5010.

[104] M. Nagahara, J. Østergaard, and D. E. Quevedo. "Discrete-time hands-off control by sparse optimization". In: *EURASIP J. Adv. Signal Proce.* 2016.1 (2016), pp. 1–8.

[105] M. Nagahara, D. E. Quevedo, and D. Nešić. "Maximum hands-off control: a paradigm of control effort minimization". In: *IEEE Trans. Automa. Contr.* 61.3 (2015), pp. 735–747.

[106] B. K. Natarajan. "Sparse approximate solutions to linear systems". In: *SIAM J. Computing* 24.2 (1995), pp. 227–234.

[107] Y. Ou, C. Xu, E. Schuster, T. C. Luce, J. R. Ferron, M. L. Walker, and D. A. Humphreys. "Optimal Tracking Control of Current Profile in Tokamaks". In: *IEEE Trans. Control Systems Technology* 19.2 (Mar. 2011), pp. 432–441.

[108] J. H. Park, S. Han, and W. H. Kwon. "LQ tracking controls with fixed terminal states and their application to receding horizon controls". In: *Systems & Control Letters* 57 (Sept. 2008), pp. 772–777.

[109] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition". In: *Proceedings of 27th Asilomar conference on signals, systems and computers*. IEEE. 1993, pp. 40–44.

[110] I. R. Petersen and R. Tempo. "Robust control of uncertain systems: Classical results and recent developments". In: *Automatica* 50.5 (2014), pp. 1315–1335.

[111] M. Phan and R. Longman. "Relationship between state-space and input-output models via observer Markov parameters". In: *WIT Trans. Built Environ* 22 (1970).

[112] M. Pilanci, M. J. Wainwright, and L. El Ghaoui. "Sparse learning via Boolean relaxations". In: *Math. Program.* 151.1 (2015), pp. 63–87.

[113] B. Polyak, M. Khlebnikov, and P. Shcherbakov. "An LMI approach to structured sparse feedback design in linear control systems". In: *2013 European control conference (ECC)*. IEEE. 2013, pp. 833–838.

[114] B. T. Polyak, M. V. Khlebnikov, and P. S. Shcherbakov. "Sparse feedback in linear control systems". In: *Autom. Remote Contr.* 75.12 (2014), pp. 2099–2111.

[115] A. Quattoni, X. Carreras, M. Collins, and T. Darrell. "An efficient projection for $\ell_{1,\infty}$ regularization". In: *Proc. of 26th Annu. Int. Conf. Mach. Learn.* 2009, pp. 857–864.

[116] C. V. Rao. "Sparsity of Linear Discrete-Time Optimal Control Problems With l_1 Objectives". In: *IEEE Trans. Autom. Contr.* 63.2 (2017), pp. 513–517.

[117] J. B. Rawlings, D. Q. Mayne, and M. Diehl. *Model predictive control: theory, computation, and design*. Vol. 2. Nob Hill Publishing Madison, WI, 2017.

[118] R. T. Rockafellar. *Convex analysis*. Vol. 18. Princeton Press, Princeton, 1970.

[119] R. T. Rockafellar and S. Uryasev. "Optimization of conditional value-at-risk". In: *J. Risk* 2 (2000), pp. 21–42.

[120] S. M. Ryan, J. C. Bean, and R. L. Smith. "A tie-breaking rule for discrete infinite horizon optimization". In: *Operations Research* 40.1-supplement-1 (1992), S117–S126.

[121] S. Sarykalin, G. Serraino, and S. Uryasev. "Value-at-risk vs. conditional value-at-risk in risk management and optimization". In: *State-of-the-art decision-making tools in the information-intensive age*. Informs, 2008, pp. 270–294.

[122] A. Shapiro, D. Dentcheva, and A. Ruszczynski. *Lectures on stochastic programming: Modeling and theory, 3rd Ed.* SIAM, 2021.

[123] A. Shapiro and A. Kleywegt. “Minimax analysis of stochastic problems”. In: *Optim. Meth. Soft.* 17.3 (2002), pp. 523–542.

[124] S. K. Shevade and S. S. Keerthi. “A simple and efficient algorithm for gene selection using sparse logistic regression”. In: *Bioinformatics* 19.17 (2003), pp. 2246–2253.

[125] R. Tempo, G. Calafiore, and F. Dabbene. *Randomized algorithms for analysis and control of uncertain systems: with applications*. Springer, 2012.

[126] R. Tempo, G. Calafiore, and F. Dabbene. *Randomized algorithms for analysis and control of uncertain systems: with applications 2nd Ed.* Springer, 2013.

[127] R. Tibshirani. “Regression shrinkage and selection via the lasso”. In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 58.1 (1996), pp. 267–288.

[128] J. A. Tropp. “Algorithms for simultaneous sparse approximation. Part II: Convex relaxation”. In: *Signal Process.* 86.3 (2006), pp. 589–602.

[129] H. J. Van Waarde, J. Eising, H. L. Trentelman, and M. K. Camlibel. “Data informativity: A new perspective on data-driven analysis and control”. In: *IEEE Trans. Automa. Contr.* 65.11 (2020), pp. 4753–4768.

[130] M. Vidyasagar. *An Introduction to Compressed Sensing*. SIAM, 2020.

[131] H. J. van Waarde, C. De Persis, M. K. Camlibel, and P. Tesi. “Willems’ fundamental lemma for state-space systems and its extension to multiple datasets”. In: *IEEE Contr. Syst. Lett.* 4.3 (2020), pp. 602–607.

[132] A. Wahl and E. D. Gilles. “Model predictive versus linear quadratic control for the tracking problem of automatic river navigation”. In: *1999 European Control Conference* (1999), pp. 1137–1142.

[133] S.-H. Wang and E. Davison. “On the stabilization of decentralized control systems”. In: *IEEE Trans. Automa. Contr.* 18.5 (1973), pp. 473–478.

[134] Y.-S. Wang, N. Matni, and J. C. Doyle. “A system-level approach to controller synthesis”. In: *IEEE Trans. Autom. Contr.* 64.10 (2019), pp. 4079–4093.

[135] J. L. Willems and I. M. Y. Mareels. “A rigorous solution of the infinite time interval LQ problem with constant state tracking”. In: *Systems & Control Letters* 52 (July 2004), pp. 31–36.

[136] J. C. Willems and J. W. Polderman. *Introduction to mathematical systems theory: a behavioral approach*. Vol. 26. Springer, 1997.

[137] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor. “A note on persistency of excitation”. In: *Syst. & Contr. Lett.* 54.4 (2005), pp. 325–329.

[138] W. Xie. “On distributionally robust chance constrained programs with Wasserstein distance”. In: *Math. Program.* 186.1 (2021), pp. 115–155.

[139] H. Xu, C. Caramanis, and S. Mannor. “Robust Regression and Lasso”. In: *IEEE Trans. Inf. Theory* 7.56 (2010), pp. 3561–3574.

[140] A. Xue and N. Matni. “Data-driven system level synthesis”. In: *Learning for Dynamics and Control*. PMLR. 2021, pp. 189–200.

[141] X. Zhang, S. Grammatico, G. Schildbach, P. Goulart, and J. Lygeros. "On the sample size of random convex programs with structured dependence on the uncertainty". In: *Automatica* 60 (2015), pp. 182–188.

[142] Z. Zhang and Y. Fujisaki. "Data-driven sparse feedback with Schur α-stability". In: *SICE International Symposium on Control Systems*. 2024.

[143] Z. Zhang and Y. Fujisaki. "Risk assessment for sparse optimization withe relaxation". In: *55th International Symposium on Stochastic Systems Theory and Its Applications (SSS'23)*. ISCIE. 2023, pp.4A3.

[144] Z. Zhang and Y. Fujisaki. "Risk-aware sparse predictive control". In: *Preprints of the 22nd IFAC World Congress*. 2023, pp. 1477–1480.

[145] Z. Zhang and Y. Fujisaki. "Sparse feedback control realization using a dynamic linear compensator". In: *SICE International Symposium on Control Systems*. 2023, pp. 3M1–4.

[146] Z. Zhang and Y. Fujisaki. "Sparse Feedback Controller: From Open-loop Solution to Closed-loop Realization". In: *arXiv:2303.15175* (2023).

[147] Z. Zhang and Y. Fujisaki. "Sparse feedback controller: from open-loop solution to closed-loop realization". In: *SICE Journal of Control, Measurement, and System Integration* 16.1 (2023), pp. 286–296.

[148] Z. Zhang and Y. Fujisaki. "Sparse Robust Control Design via Scenario Program". In: *Proceedings of the ISCIE International Symposium on Stochastic Systems Theory and its Applications (SSS'21)* 2022 (2022), pp. 61–64.

[149] Z. Zhang and M. Nagahara. "Linear Quadratic Tracking Control with Sparsity-Promoting Regularization". In: *2021 American Control Conference (ACC)*. IEEE. 2021, pp. 3812–3817.

[150] K. Zhou and J. C. Doyle. *Essentials of robust control*. Vol. 104. Prentice hall Upper Saddle River, NJ, 1998.

[151] H. Zou, T. Hastie, and R. Tibshirani. "Sparse principal component analysis". In: *Journal of computational and graphical statistics* 15.2 (2006), pp. 265–286.

[152] S. Zymler, D. Kuhn, and B. Rustem. "Distributionally robust joint chance constraints with second-order moment information". In: *Math. Program.* 137.1 (2013), pp. 167–198.