Doctoral Thesis

Sparse Optimal Control for Continuous-Time Dynamical Systems

by

Takuya Ikeda

Department of Applied Mathematics and Physics
Graduate School of Informatics
Kyoto University

Supervisor: Associate Professor, Kenji Kashima

Abstract

This thesis investigates an optimal control problem for continuous-time dynamical systems with the L^0 control cost. This cost functional penalizes the length of the support of control variables, and the optimization based on the criteria tends to make the control input identically zero on a set with positive measures. The main contributions of this thesis are threefold: Firstly, the sparse optimal control that steers the state from a point to a point on a finite horizon is investigated. An equivalence theorem among the L^p optimal control problems with $p \in [0, 1]$ is obtained. In addition, a numerical algorithm for the sparse optimization problem based on the alternating direction method of multipliers is provided. Secondly, the sparse optimization problem is applied to the control of multi-agent systems. A control node selection method for achieving high controllability by providing exogenous inputs equipped with small L^0 costs is proposed, and a distributed consensus algorithm based on the sparse optimal control via sampled-data state observation is obtained. Lastly, the sparse optimal control is characterized in a feedback framework by applying the dynamic programming approach. Since, in general, the value function is not differentiable in the domain because of the non-smoothness of the L^0 cost functional, the value function is characterized as a viscosity solution to the associated Hamilton-Jacobi-Bellman equation. Based on the result, the optimal feedback map is derived.

Table of Contents

Chapter 1

Introduction

Recent years have seen an explosion of scientific research in a field called *sparse modeling*. It is a powerful framework for data analysis and signal processing, and the underlying concept is that the signals or data of interest have a small number of essential features. The most famous one of the techniques in this field is probably *compressed sensing*, which is also known as *compressive sensing* or *compressive sampling* [1]. A Google Scholar search for papers including one of these three terms in their title returned about 15,500 hits at the time this thesis was written. This technique aims to reconstruct a finite-dimensional vector from linear measurements of dimension smaller than the size of the unknown vector. As a classical linear algebra indicates, such an undetermined linear equation has infinitely many solutions. Compressed sensing tackles with this issue by putting an additional assumption on the signals of interest. The underlying assumption is *sparsity*, and a signal is called *sparse* if most of its components are zero. Compressed sensing has brought successes into various scientific fields, such as medical imaging [2, 3, 4, 5, 6, 7, 8], machine learning [9, 10, 11, 12, 13, 14, 15, 16], communication engineering [17, 18, 19], and astronomy [20], to name a few. These applications have led to a breakthrough in sparse modeling.

More recently, sparse modeling has been applied to control of dynamical systems. For example, [21, 22, 23] report that the sparsity of signals is useful in communications in networked control systems, where a controller should communicate with a controlled plant over rate-limited networks and the data to be transmitted should satisfy the rate-limiting constraint. In [24], a control problem of linear systems when some of the sensors or actuators are corrupted by an attacker is considered, and a secure control loop that can improve the resilience of the system is proposed. In [25], some techniques are proposed to design feedback systems that sparsify the state in order to perfectly reconstruct it using compressive sensing algorithms. In [26, 27], a sparse feedback gain is designed to reduce the number of communication links or sensors. In addition, state estimation problems and system identification problems with connections to sparse modeling are studied in [28, 29, 30] and in [31, 32, 33], respectively.

Most of the aforementioned literature consider the minimization problem of *the number of nonzero components* of finite-dimensional vectors of interest, and their results heavily depend on the discussion in compressed sensing. Therefore, literature in control theory related to sparse modeling almost consider finite-dimensional optimization problems that have sparse solutions. In other words, there seem to be very little related control literature that consider the sparsity of continuous-time signals in a functional space, as far

as the author knows. Then, this thesis defines the sparsity of a continuous-time signal as *the length of the time duration where the signal takes zero value* and focuses on the minimization problem of such a length. This problem is then mathematically defined as an infinite-dimensional optimization problem, compared to the existing literature. In particular, considering the fact that a convex relaxation method is widely used in order to circumvent the computation burden in sparse modeling [34, 35, 36], this thesis shows the equivalence between the main problem and the associated convex problem. This is the main theoretical contribution of this thesis. The more precise overview of this thesis is as follows.

Chapter 2 investigates an optimal control problem in which control variables are penalized by the length of the support. The optimization tends to make the control input identically zero on a set with positive measures, and the optimal control is then switched off completely on parts of the time domain. This is why the problem is referred to as *sparse optimal control problem* or *maximum hands-off control problem*. Alternatively, such a problem is also called L^0 *optimal control problem*, since the penalty cost is mathematically defined by the L^0 *norm* of the decision variables. This control strategy, which is also known as gliding or coasting, is actually used in hybrid/electric vehicles [37, 38, 39], railway vehicles [40, 41, 42, 43], free-flying robots [44], stop-start systems [45], and sleep mode operation in wireless communication systems [46, 47] for saving fuel or electricity consumptions and reducing CO_2 emissions or vibrations. Also, as will be shown in this chapter, the L^0 optimal control is theoretically equivalent to the L^1 optimal control, which is a classical control minimizing fuel consumption in aerial vehicles and space rockets and is known as *minimum fuel control* [48]. In addition, this kind of control can extract the time duration that has a large impact on the system, which is useful in time-varying nodes selection problem in multi-agent systems, as discussed in Chapter 3. Chapter 2 considers continuous-time linear systems and the control object is to steer the state from an initial state to a target state over a finite horizon by using the L^0 optimal control, i.e. the control input having the minimum length of the support. For the characterization of the L^0 optimal control, the relationship among the L^p optimal solutions with $p \in [0, 1]$ is discussed. The main theorem in this chapter is derived based on an insight from existing literature for finite-dimensional systems. Therefore, the obtained theorem naturally extends the existing results to general linear systems including infinite-dimensional systems, which are not addressed in the literature.

Chapter 3 and Chapter 4 discuss the application of the sparse optimal control to multi-agent systems (MAS). In Chapter 2, the sparse optimal control satisfying a given state transition from a point to another point is theoretically analyzed. On the other hand, it is also important to investigate a time duration over which control inputs can realize efficient state transitions towards all directions. This enables us to find *controller activation schedule* that does not depend on the target state. In addition to input sparsity, one might be concerned with the energy required to steer the system. For this issue, Chapter 3 investigates a novel optimal control problem that aims to maximize a quantitative metric of the controllability when control inputs are constrained in terms of the L^0 norm. For the metric of the controllability, the trace of the controllability Gramian is adopted from the analytical perspective. This quantity is closely related to the average energy required to steer the system in all directions in the state space. The control inputs are thus penalized by not only the L^0 cost but also the L^2 cost, which is related to the

sparse quadratic regulator [49]. Here, it should be emphasized that this optimal control problem is compatible with the node selection problem in MAS. In the context, the optimal solution is interpreted as an answer to the question of when and where exogenous control inputs should be provided for high controllability. In other words, it characterizes essential time-varying control nodes for the system of interest. To the best of the author's knowledge, time-varying node selection problems for continuous-time systems have not yet been proposed at the time this thesis was written.

Chapter 4 investigates a consensus problem in MAS. This problem focuses on the mechanism of how agents can reach an agreement at a common value through local interactions. In MAS, it is often the case that agents have limits on their power sources. For example, small unmanned ground vehicles can only use cheap energy sources on many occasions, due to the size and weight restrictions. Moreover, the local interactions in MAS are usually performed on a wireless network. This imposes agents on communications with their neighbors by transmitting *sampled data* of the states (and other values as necessary) of them. The control system thus becomes a sampled-data control system [50], and only sampled values of information can be used to determine the control. Therefore, these considerations are needed to be reflected in the problem formulation. For this issue, Chapter 4 proposes a novel design method of distributed control input for consensus that is based on maximum hands-off control and sampled-data state observation. This formulation is motivated by three reasons: Firstly, the hands-off control has an ecological aspect as discussed above. MAS consisting of agents whose energy sources are constrained may benefit from the *sparse property* of the maximum hands-off control. A similar strategy can be found in event-triggered (or self-triggered) control, which aims at increasing "zero control (or observation) time" to reduce computational/communication resources [51, 52, 53]. Secondary, the maximum hands-off control takes only three values of zero and the upper bound and the lower bound of the magnitude constraint, as shown in Chapter 2. Such a discreteness is called bang-off-bang property. This property is preferable for cheap actuators since the control values are already quantized. Thirdly, the maximum hands-off control is easily obtained from the results in Chapter 2. In particular, analytical optimal solutions for simple systems such as first- and second-order systems can be described in a closed form. This less computation burden is advantageous for MAS, since heavy computation burden could break the consensus.

Chapter 5 investigates an optimal control problem for non-linear systems with the L^0 control cost. Compared to the chapters above, this chapter characterizes the sparse optimal control in a feedback framework. The approach is based on the dynamic programming approach and the optimal control is analyzed via the value function. Due to the non-smoothness of the L^0 cost functional, in general, the value function is not differentiable in the domain. Then, the value function is characterized as a viscosity solution to the associated Hamilton-Jacobi-Bellman (HJB) equation. This result derives a sufficient and necessary condition for the L^0 optimality, which immediately gives the optimal feedback map. In addition, the relationship with L^1 optimal control problem is considered and an equivalence theorem is shown. Note that, while most of the literature consider the L^p optimal feedback controls with $p \in (0, 1]$ to generate sparse feedback controls, the approach in this thesis directly addresses the exact L^0 optimal feedback control.

This thesis ends with Chapter 6 offering concluding remarks. The contributions and their limitations are outlined, and possible future developments are then provided.

Chapter 2

On Sparse Optimal Control for General Linear Systems

2.1 Introduction

This chapter considers a sparse optimal control problem with the L^0 control cost. Such optimization problems involving the L^0 norm have the combinatorial structure, which causes heavy computational burden. Then, the associated convex relaxation problem has been solved instead of the sparse optimization problem in existing literature. In particular, the ℓ^1 norm is adopted as the cost function in compressed sensing, and the sparsity of the ℓ^1 optimal solution is often investigated [34, 35, 36]. Similarly, the conditions under which the L^0 optimal control is exactly solved by the associated L^1 optimal control problem are investigated in [54] and an equivalence theorem is derived. In addition, considering how the L^p penalization promotes the sparsity as p goes to 0, the L^0 optimal control problem has been analyzed in view of the optimal control problem penalized by the L^p cost with $p \in (0, 1)$ in [55, 56, 57]. On the other hands, the L^0 optimal controls is directly analyzed in [58, 59]. The optimization problem is also discussed in the context of partial differential equations [60, 61, 62].

The previous works [54, 55, 58] have studied the L^0 optimal control that steers the state from a point to a point. Here, the constraint for linear systems that requires a state transition from an initial state to a target state is described by a Volterra integral equation related to the convolution of a control variable and the impulse response of the system. In other words, the system dynamics affects the optimization problem not through its realization, but through the impulse response. Nevertheless, the proofs in the previous results, which are based on the Pontryagin's minimum principle, heavily depended on the finite dimensionality of the system realization. Then, this observation naturally raises the question of how the L^0 optimal solution is characterized if the kernel of the Volterra integral is arbitrarily taken from a functional space. In view of this, this paper thus investigates an L^0 optimization problem with a constraint in a form of Volterra integral whose kernel belongs to the space L^∞. Since the formulation is based on the kernel function, it addresses a wide range of minimization problems with the L^0 cost. Indeed, the formulation includes the analysis for not only finite-dimensional linear systems but also infinite-dimensional linear systems whose transfer functions are irrational, e.g., time-delay systems, spatially distributed systems. It should be emphasized that the proofs in

[54, 55, 58] are no longer applicable for such systems. Furthermore, this framework can also address output constraints. These extensions not only lead to better understanding of the sparse optimal control, but also improve its practical usefulness. These advantages over existing results are main contribution of this study.

In the light of the fact that the L^0 optimal solution has been analyzed in view of the L^p optimal solution with $p \in (0, 1]$, we investigate the relationship among the L^p optimizations with $p \in [0, 1]$. For the analysis, we first characterize the relationship between the L^0 optimization and the L^1 optimization. More precisely, we give the condition that guarantees the L^0 optimization is exactly solved by the L^1 optimization. The result shows that the L^0 cost is not necessarily replaced by the L^1 cost. In order to bridge the gap, we also give the inclusion theorem claiming that any L^0 optimal solution is an L^1 optimal solution. We next consider the relationship between the L^0 optimization and the L^p optimization with $p \in (0, 1)$. The equivalence for any $p \in (0, 1)$ is proved. The results also show the existence theorem on the L^0 optimal solution.

The remainder of this chapter is organized as follows. In Section 2.2, we give mathematical preliminaries for our subsequent discussion. Section 2.3 defines the main problem and briefly reviews the existing results for finite-dimensional linear systems. An assumption for the equivalence between the L^0 optimization and the L^1 optimization is also introduced based on the review. Section 2.4 is the main section of this chapter. We first characterize the L^1 optimal solutions in view of the existence and the structure. We then give a theorem on the equivalence. We also mention the optimization involving output constraints and show that the sufficient condition for the equivalence is closely related to the *output controllability*. In addition, we describe the inclusion between the L^0 optimization and the L^1 optimization, which is followed by the equivalence among the L^p optimizations with $p \in [0, 1)$. Section 2.5 proposes a numerical optimization algorithm to solve the L^0 optimization problem. The proposed algorithm is based on the alternating direction method of multipliers. Section 2.6 presents an example to illustrate the equivalence between the L^0 optimization and the L^1 optimization for a spatially distributed system. In Section 2.7, we offer concluding remarks.

2.2 Mathematical Preliminaries

This section reviews basic definitions and notations that will be used throughout this thesis. As the chapter progresses, the notations will be gradually introduced.

The set of all real numbers and the set of all positive integers are denoted by $\mathbb{R}$ and $\mathbb{N}$, respectively. $\mathbb{R}^n$, $n \in \mathbb{N}$, denotes the n-dimensional Euclidean space. The Euclidean norm in $\mathbb{R}^n$ is denoted by $\|\cdot\|$, i.e., $\|x\| \triangleq \sqrt{x^\top x}$ for $x \in \mathbb{R}^n$, where $\top$ denotes the transpose. The ℓ^p norms in $\mathbb{R}^n$ with $p \in [0, 1]$ or $p = \infty$ are respectively defined by

$$\|x\|_{\ell^0} \triangleq \sum_{i=1}^{n} |x_i|^0,$$

$$\|x\|_{\ell^p} \triangleq \left\{ \sum_{i=1}^{n} |x_i|^p \right\}^{1/p}, \quad p \in (0, 1]$$

$$\|x\|_{\ell^\infty} \triangleq \max_{1 \le i \le n} |x_i|$$

for $x = [x_1, x_2, \ldots, x_n]^\top \in \mathbb{R}^n$, where $0^0 = 0$.

Let $n \in \mathbb{N}$. For a vector $x \in \mathbb{R}^n$ and a scalar $\varepsilon > 0$, the *ε-neighborhood* of x is defined by

$$\mathbb{B}(x, \varepsilon) \triangleq \{y \in \mathbb{R}^n : \|y - x\| < \varepsilon\}.$$

Let $\mathcal{X}$ be a subset of $\mathbb{R}^n$. A point $x \in \mathcal{X}$ is called an *interior point* of $\mathcal{X}$ if there exists $\varepsilon > 0$ such that $\mathbb{B}(x, \varepsilon) \subset \mathcal{X}$. The *interior* of $\mathcal{X}$ is the set of all interior points of $\mathcal{X}$, and we denote the interior of $\mathcal{X}$ by $\mathrm{int}\mathcal{X}$. A point $x \in \mathbb{R}^n$ is called an *adherent point* of $\mathcal{X}$ if $\mathbb{B}(x, \varepsilon) \cap \mathcal{X} \neq \emptyset$ for every $\varepsilon > 0$, and the *closure* of $\mathcal{X}$ is the set of all adherent points of $\mathcal{X}$, which is denoted by $\overline{\mathcal{X}}$. The *boundary* of $\mathcal{X}$ is the set of all points in the closure of $\mathcal{X}$, not belonging to the interior of $\mathcal{X}$, and we denote the boundary of $\mathcal{X}$ by $\partial\mathcal{X}$, i.e., $\partial\mathcal{X} = \overline{\mathcal{X}} - \mathrm{int}\mathcal{X}$, where $\mathcal{X}_1 - \mathcal{X}_2$ is the set of all points which belong to the set $\mathcal{X}_1$ but not to the set $\mathcal{X}_2$. A vector $\rho \in \mathbb{R}^n$ is said to be *normal* to a convex set $\mathcal{C} \subset \mathbb{R}^n$ at a point $a \in \mathcal{C}$, if $\rho^\top x \leq \rho^\top a$ for every $x \in \mathcal{C}$.

Let $T > 0$ and m, $r \in \mathbb{N}$. For $p \in [0, 1]$ or $p = \infty$, $L^p([0, T], \mathbb{R}^m)$ denotes the set of all continuous-time signals $u(t) = [u_1(t), u_2(t), \ldots, u_m(t)]^\top \in \mathbb{R}^m$ over a time interval $[0, T]$ such that $\|u\|_p < \infty$, where $\|\cdot\|_p$, referred to as L^p *norm*, is defined by

$$\|u\|_0 \triangleq \sum_{j=1}^{m} \mu_L(\{t \in [0, T] : u_j(t) \neq 0\}), \tag{2.2.1}$$

$$\|u\|_p \triangleq \left\{ \sum_{j=1}^{m} \int_0^T |u_j(t)|^p dt \right\}^{1/p}, \tag{2.2.2}$$

$$\|u\|_\infty \triangleq \max_{j=1,2,\ldots,m} \operatorname*{ess\,sup}_{0 \leq t \leq T} |u_j(t)|, \tag{2.2.3}$$

with the Lebesgue measure μ_L on $\mathbb{R}$. We denote by $L^\infty([0, T], \mathbb{R}^{r \times m})$ the set of all functions $\phi(t) \in \mathbb{R}^{r \times m}$ that satisfy $\operatorname*{ess\,sup}_{0 \leq t \leq T} |\phi_{ij}(t)| < \infty$ for all i and j, where $\phi_{ij}(t)$ is the (i, j)-component of $\phi(t)$. If no confusion may arise, we simply denote the spaces by L^p. We denote the sign function by sgn, i.e.,

$$\mathrm{sgn}(w) \triangleq \begin{cases} |w|/w, & \text{if } w \neq 0, \\ 0, & \text{if } w = 0. \end{cases}$$

2.3 Problem Formulation

This chapter characterizes the solutions to an L^0 optimization problem with constraints on a Volterra integral equation and the L^∞ norm. In this section, we first formulate the main problem in this chapter. After a brief review of the result for finite-dimensional systems, we introduce an assumption under which we will show the equivalence between the L^0 optimization and the L^1 optimization (we call the equivalence *L^0- and L^1-equivalence* for the simplicity). Note that given two optimization problems are said to be *equivalent* if each of the optimizations gives the optimal solution to the other.

The main problem in the chapter is given as follows. We call the problem *L^p optimization problem* and the solution is called *L^p optimal solution.*

Problem 2.3.1
Given a number $p \in [0,1]$, a vector $a \in \mathbb{R}^r$, a time duration $T > 0$, and a function $\phi \in L^\infty([0,T], \mathbb{R}^{r\times m})$, find a function $u(t) \in \mathbb{R}^m$ on $[0,T]$ that solves

$$\begin{aligned} \underset{u}{\text{minimize}} \quad & \|u\|_p \\ \text{subject to} \quad & \int_0^T \phi(t)u(t)dt = a, \quad \|u\|_\infty \leq 1. \end{aligned}$$

Remark 2.3.2
Throughout the chapter, we confine ourselves to the class $L^\infty([0,T], \mathbb{R}^{r\times m})$. The assumption on ϕ is used to construct the L^0 optimal solution, as illustrated in Theorem 2.4.6. More precisely, it is needed for the existence of $\alpha \geq 0$ such that $f(\alpha) = \theta$ for any $\theta \in [0, mT]$ in Lemma 2.4.3.

Remark 2.3.3
The impulse response of a given linear system corresponds to $\phi(T-t)$. This formulation includes the L^0 optimal control problem for finite-dimensional continuous-time linear systems, which is investigated in existing literature such as [54], as a special case. The details are illustrated below.

Remark 2.3.4
Although Problem 2.3.1 constrains the L^∞ norm by 1, it does not lose generality. That is, if we have $\|u\|_\infty \leq U_{max}$ with $U_{max} > 0$, then it is enough to use a/U_{max} instead of a in the equality constraint.

This chapter investigates the condition that guarantees L^0 optimization is exactly solved by the associated L^p optimization problem with $p \in (0,1]$. For the analysis, we here briefly review an existing result for finite-dimensional linear systems, based on the discussion in [54]. Let us consider an optimal control problem that seeks a control that minimizes the L^1 norm among all controls steering the system represented by

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

to a given target state $x_1 \in \mathbb{R}^n$ from a given initial state $x_0 \in \mathbb{R}^n$ at a given time $T > 0$ and satisfying $\|u\|_\infty \leq 1$, where $A \in \mathbb{R}^{n\times n}$ and $B = [b_1, b_2, \ldots, b_m] \in \mathbb{R}^{n\times m}$ are constant matrices. Note that this problem corresponds to a special case of Problem 2.3.1 when $p = 1$, in which $\phi(t) = e^{A(T-t)}B$ and $a = x_1 - e^{AT}x_0$. In [54], the authors applied Pontryagin's minimum principle for the characterization of the L^1 optimal control. According to the principle, if there exists an L^1 optimal control $u^*(t) = [u_1^*(t), u_2^*(t), \ldots, u_m^*(t)]^\top$, then there exists a scalar η equal to 0 or 1 and a vector $p_0 \in \mathbb{R}^n$ such that *costate* $p(t) \triangleq e^{A^\top(T-t)}p_0$ satisfies

1. $(\eta, p(t)) \neq 0 \quad \forall t \in [0,T]$,
2. $u_j^*(t) = \underset{u_j\in[-1,1]}{\arg\max}(p(t)^\top b_j u_j - \eta|u_j|) \quad$ a.e.

This implies that if

$$\mu_L(\{t \in [0,T] : p_0^\top e^{A(T-t)}b_j = \alpha\}) = 0 \tag{2.3.1}$$

for any $p_0 \neq 0$ and $\alpha \in \mathbb{R}$, then the L^1 optimal control takes only the values of 0 and ± 1 almost everywhere. This discreteness is called *bang-off-bang property* and this indicates that the L^1 optimization enhances the sparsity. Based on this, the work [54] has established an L^0- and L^1-equivalence theorem under an assumption on the discreteness. Furthermore, the work [55] has shown that the equivalence may fail if the discreteness assumption fails. Thus, the condition (2.3.1) is closely related to the L^0- and L^1-equivalence.

Motivated by this result, we next introduce the following assumption on $\phi = [\phi_1, \phi_2, \ldots, \phi_m]$ for Problem 2.3.1.

Assumption 2.3.5
For any nonzero $\rho \in \mathbb{R}^r$, any $\alpha \in \mathbb{R}$, and any $j \in \{1, 2, \ldots, m\}$,

$$\mu_L(\{t \in [0, T] : \rho^\top \phi_j(t) = \alpha\}) = 0.$$

We will show that the L^0- and L^1-equivalence for general linear systems is guaranteed under this assumption in Theorem 2.4.6, which is a natural extension of existing results in [54]. The equivalence for $p \in (0, 1)$ follows from the relationship between the L^0 optimization and L^1 optimization, as proved in Theorem 2.4.11.

For ease of notation, we here define two sets associated with Problem 2.3.1. One is the set of all vectors a for which there exists u satisfying the constraints in Problem 2.3.1, and the other is the set of all such u:

Definition 2.3.6
Given $T > 0$ and $\phi \in L^\infty([0, T], \mathbb{R}^{r \times m})$,

$$\mathcal{R}(\phi, T) \triangleq \left\{ \int_0^T \phi(t)u(t)dt : \|u\|_\infty \leq 1 \right\},$$

and for $\alpha > 0$

$$\mathcal{R}_\alpha(\phi, T) \triangleq \left\{ \int_0^T \phi(t)u(t)dt : \|u\|_\infty \leq 1, \|u\|_1 \leq \alpha \right\}.$$

Definition 2.3.7
Given $a \in \mathbb{R}^r$, $T > 0$, and $\phi \in L^\infty([0, T], \mathbb{R}^{r \times m})$,

$$\mathcal{U}(a, \phi, T) \triangleq \left\{ u : \int_0^T \phi(t)u(t)dt = a, \|u\|_\infty \leq 1 \right\}.$$

By definition, $\mathcal{U}(a, \phi, T)$ is non-empty if and only if $a \in \mathcal{R}(\phi, T)$. The set $\mathcal{R}_\alpha(\phi, T)$ is a subset of $\mathcal{R}(\phi, T)$ for any α, and we have $\mathcal{R}(\phi, T) = \mathcal{R}_\alpha(\phi, T)$ for $\alpha \geq mT$.

2.4 Analysis

This section establishes the relationship among the L^p optimizations with $p \in [0, 1]$. For the purpose, we first characterize the L^1 optimal solution in view of the existence, uniqueness, discreteness, and the structure.

2.4.1 Characterization of the L^1 Optimal Solution

We first prepare three lemmas. The first one states the existence of the L^1 optimal solution, and the others are used to elucidate the form of the L^1 optimal solution in Theorem 2.4.4.

Lemma 2.4.1
Given $T > 0$ and $\phi \in L^\infty([0,T], \mathbb{R}^{r\times m})$, there exists at least one L^1 optimal solution for each $a \in \mathcal{R}(\phi, T)$.

Proof. Fix any $a \in \mathcal{R}(\phi, T)$. Since the set $\mathcal{U}(a, \phi, T)$ is non-empty, we can define

$$\theta \triangleq \inf_{u\in\mathcal{U}(a,\phi,T)} \|u\|_1.$$

Then there exists a sequence $\{u^{(l)}\}_{l\in\mathbb{N}} \subset \mathcal{U}(a, \phi, T)$ such that

$$\lim_{l\to\infty} \|u^{(l)}\|_1 = \theta, \quad \|u^{(l)}\|_\infty \le 1,$$

and

$$\int_0^T \phi(t)u^{(l)}(t)dt = a. \tag{2.4.1}$$

Since the set $\{u \in L^\infty : \|u\|_\infty \le 1\}$ is sequentially compact in the weak* topology of L^∞ [63], there exist a measurable function $u^{(\infty)}$ with $\|u^{(\infty)}\|_\infty \le 1$ and a subsequence $\{u^{(l')}\}$ such that each component $\{u_j^{(l')}\}$ converges to $u_j^{(\infty)}$, the j-th component of $u^{(\infty)}$, in the weak* topology of L^∞, that is, we have

$$\lim_{l'\to\infty} \int_0^T \big(u_j^{(l')}(t) - u_j^{(\infty)}(t)\big) f(t)dt = 0 \tag{2.4.2}$$

for any $f \in L^1$ and $j = 1, 2, \ldots, m$. By (2.4.1) and (2.4.2),

$$\int_0^T \phi(t)u^{(\infty)}(t)dt = a$$

and hence $u^{(\infty)} \in \mathcal{U}(a, \phi, T)$. It follows from (2.4.2) that

$$\lim_{l'\to\infty} \int_0^T u_j^{(l')}(t)\,\mathrm{sgn}(u_j^{(\infty)}(t))dt = \int_0^T u_j^{(\infty)}(t)\,\mathrm{sgn}(u_j^{(\infty)}(t))dt = \int_0^T |u_j^{(\infty)}(t)|dt$$

for all $j = 1, 2, \ldots, m$. We also have

$$\sum_{j=1}^m \int_0^T u_j^{(l')}(t)\,\mathrm{sgn}(u_j^{(\infty)}(t))dt \le \sum_{j=1}^m \int_0^T |u_j^{(l')}(t)|dt = \|u^{(l')}\|_1$$

for all $j = 1, 2, \ldots, m$. Hence

$$\begin{aligned}
\|u^{(\infty)}\|_1 &= \sum_{j=1}^m \int_0^T |u_j^{(\infty)}(t)|dt \\
&= \lim_{l'\to\infty} \sum_{j=1}^m \int_0^T u_j^{(l')}(t)\,\mathrm{sgn}(u_j^{(\infty)}(t))dt \\
&\le \lim_{l'\to\infty} \|u^{(l')}\|_1.
\end{aligned}$$

Here, the subsequence $\{\|u^{(l')}\|_1\}$ converges to θ, since the sequence $\{\|u^{(l)}\|_1\}$ converges to θ as $l \to \infty$. Therefore, we have

$$\|u^{(\infty)}\|_1 \leq \lim_{l' \to \infty} \|u^{(l')}\|_1 = \theta. \tag{2.4.3}$$

On the other hand, since $u^{(\infty)} \in \mathcal{U}(a, \phi, T)$, we have $\|u^{(\infty)}\|_1 \geq \theta$. This with (2.4.3) gives $\|u^{(\infty)}\|_1 = \theta$. Thus, $u^{(\infty)}$ is an optimal solution for a. □

Lemma 2.4.2
Given $T > 0$ and $\phi \in L^\infty([0, T], \mathbb{R}^{r \times m})$, any $a \in \mathcal{R}(\phi, T)$ satisfies $a \in \partial\mathcal{R}_\theta(\phi, T)$, where θ is the L^1 optimal value for a.

Proof. Fix any $a \in \mathcal{R}(\phi, T)$ and let the optimal value be given by θ. Then $a \in \mathcal{R}_\theta(\phi, T)$ holds obviously. Let us assume that $a \in \mathrm{int}\mathcal{R}_\theta(\phi, T)$ and show a contradiction. Since $a \in \mathrm{int}\mathcal{R}_\theta(\phi, T)$, there exists a positive number $\varepsilon > 0$ such that

$$\mathbb{B}(a, \varepsilon) \subset \mathcal{R}_\theta(\phi, T).$$

Hence, there exists a vector $b \in \mathcal{R}_\theta(\phi, T)$ and a positive number $\lambda \in (0, 1)$ such that $a = \lambda b$. Since $b \in \mathcal{R}_\theta(\phi, T)$, there exists a function $u_b \in \mathcal{U}(b, \phi, T)$ such that $\|u_b\|_1 \leq \theta$. Then the function $u_a \triangleq \lambda u_b$ satisfies that

$$\int_0^T \phi(t) u_a(t) dt = \lambda \int_0^T \phi(t) u_b(t) dt = \lambda b = a$$

and

$$\|u_a\|_\infty = \lambda \|u_b\|_\infty \leq \lambda < 1.$$

Hence $u_a \in \mathcal{U}(a, \phi, T)$. In adittion,

$$\|u_a\|_1 = \lambda \|u_b\|_1 \leq \lambda\theta < \theta.$$

Thus, we obtain a contradiction since θ is the L^1 optimal value. Therefore, $a \in \partial\mathcal{R}_\theta(\phi, T)$. □

Lemma 2.4.3
Suppose $T > 0$ and $\phi \in L^\infty([0, T], \mathbb{R}^{r \times m})$ satisfy Assumption 2.3.5. Fix any $\theta \in [0, mT]$ and any $\rho \neq 0$. Take $\alpha \geq 0$ such that

$$f(\alpha) \triangleq \sum_{j=1}^m \mu_L(\{t \in [0, T] : |\rho^\top \phi_j(t)| \geq \alpha\}) = \theta, \tag{2.4.4}$$

and define a function $\bar{u} = [\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m]^\top$ by

$$\bar{u}_j(t) \triangleq \begin{cases} \mathrm{sgn}(\rho^\top \phi_j(t)), & \text{if } |\rho^\top \phi_j(t)| \geq \alpha, \\ 0, & \text{else} \end{cases} \tag{2.4.5}$$

for all j. Then, the function $\bar{u}$ is a unique (in the a.e. sense) maximizer of

$$\int_0^T \rho^\top \phi(t) u(t) dt \tag{2.4.6}$$

among all u with $\|u\|_\infty \leq 1$ and $\|u\|_1 \leq \theta$.

Proof. First, Assumption 2.3.5 is used to prove that $\bar{u}$ is uniquely (in the a.e. sense) determined by θ. The continuity of $f(\alpha)$ follows from this assumption. Note that $f(0) = mT$ and $f(\alpha) = 0$ for α sufficiently larger than $\max_{i,j} \operatorname*{ess\,sup}_{0\le t\le T} |\phi_{ij}(t)|$. Hence there exists a set $\mathcal{A}$ such that $f(\alpha) = \theta$ for all $\alpha \in \mathcal{A}$. Further, for any $\alpha_1, \alpha_2 \in \mathcal{A}$ such that $\alpha_1 < \alpha_2$ and $j = 1, 2, \ldots, m$, we have

$$\mu_L\big(\{t \in [0,T] : \alpha_1 < |\rho^\top \phi_j(t)| \le \alpha_2\}\big) = 0.$$

This shows the desired uniqueness.

Let us fix $\alpha \in \mathcal{A}$ arbitrarily. Then, the function $\bar{u}$ defined by (2.4.5) satisfies $\|\bar{u}\|_1 = \theta$, since

$$\|\bar{u}\|_1 = \sum_{j=1}^{m} \int_{\{t\in[0,T]:|\rho^\top\phi_j(t)|\ge\alpha\})} 1dt = f(\alpha).$$

In addition, we also have $\|\bar{u}\|_\infty \le 1$.

We first show that $\bar{u}$ maximizes (2.4.6). Put

$$\begin{aligned} E_j &\triangleq \{t \in [0,T] : |\rho^\top \phi_j(t)| \ge \alpha\}, \\ F_j &\triangleq \{t \in [0,T] : |\rho^\top \phi_j(t)| < \alpha\} \end{aligned}$$

for all j. Then, $E_j \cup F_j = [0,T]$, $E_j \cap F_j = \emptyset$, and

$$\sum_{j=1}^{m} \mu_L(E_j) = \theta, \tag{2.4.7}$$

obviously. It follows from the definition of $\bar{u}$ that

$$\int_0^T \rho^\top \phi_j(t)\bar{u}_j(t)dt = \int_{E_j} |\rho^\top \phi_j(t)|dt.$$

Then, for any u with $\|u\|_\infty \le 1$ and $\|u\|_1 \le \theta$, we have

$$\begin{aligned} \int_0^T \rho^\top \phi_j(t)u_j(t)dt &\le \int_{E_j} |\rho^\top \phi_j(t)u_j(t)|dt + \int_{F_j} |\rho^\top \phi_j(t)u_j(t)|dt \\ &\le \int_{E_j} |\rho^\top \phi_j(t)u_j(t)|dt + \alpha \int_{F_j} |u_j(t)|dt \\ &= \int_{E_j} \big(|\rho^\top \phi_j(t)| - \alpha\big)|u_j(t)|dt + \alpha \int_0^T |u_j(t)|dt \\ &\le \int_{E_j} \big(|\rho^\top \phi_j(t)| - \alpha\big)dt + \alpha \int_0^T |u_j(t)|dt \\ &= \int_{E_j} |\rho^\top \phi_j(t)|dt - \alpha\mu_L(E_j) + \alpha \int_0^T |u_j(t)|dt \\ &= \int_0^T \rho^\top \phi_j(t)\bar{u}_j(t)dt - \alpha\mu_L(E_j) + \alpha \int_0^T |u_j(t)|dt \end{aligned} \tag{2.4.8}$$

for all j. Hence

$$\begin{aligned}\int_0^T \rho^\top \phi(t)u(t)dt &= \sum_{j=1}^m \int_0^T \rho^\top \phi_j(t)u_j(t)dt \\ &\le \int_0^T \rho^\top \phi(t)\bar{u}(t)dt - \alpha \sum_{j=1}^m \mu_L(E_j) + \alpha\|u\|_1 \\ &\le \int_0^T \rho^\top \phi(t)\bar{u}(t)dt\end{aligned}$$

from (2.4.7).

We finally show the uniqueness of such a maximizer. Suppose v satisfies $\|v\|_\infty \le 1$, $\|v\|_1 \le \theta$, and

$$\int_0^T \rho^\top \phi(t)v(t)dt = \int_0^T \rho^\top \phi(t)\bar{u}(t)dt.$$

This implies that equality holds at each evaluation in (2.4.8) when we substitute $u_j = v_j$. In particular, we have

$$\begin{aligned}&\int_{F_j} |\rho^\top \phi_j(t)v_j(t)|dt = \alpha \int_{F_j} |v_j(t)|dt, \\ &\int_{E_j} \big(|\rho^\top \phi_j(t)| - \alpha\big)|v_j(t)|dt = \int_{E_j} \big(|\rho^\top \phi_j(t)| - \alpha\big)dt, \\ &\int_{E_j} \rho^\top \phi_j(t)v_j(t)dt = \int_{E_j} |\rho^\top \phi_j(t)v_j(t)|dt\end{aligned}$$

for all j. The first and second equalities show

$$v_j(t) = 0 \text{ a.e. on } F_j, \quad |v_j(t)| = 1 \text{ a.e. on } E_j.$$

Then the third one implies $\mathrm{sgn}(v_j(t)) = \mathrm{sgn}(\rho^\top \phi_j(t))$ a.e. on E_j. This completes the proof. □

Based on the lemmas, we show the structure of the L^1 optimal solution under Assumption 2.3.5. To put it briefly, the j-th component of the L^1 optimal solution takes the maximum magnitude when $|\rho^\top \phi_j(t)|$ is larger than a constant and otherwise the component takes 0, where ρ is a normal vector to a set at a. The following theorem characterizes the L^1 optimal solution in terms of the impulse response without assuming its finite-dimensional realization, unlike existing results, e.g. [48].

Theorem 2.4.4
Suppose $T > 0$ and $\phi \in L^\infty([0,T], \mathbb{R}^{r\times m})$ satisfy Assumption 2.3.5. Then, for each $a \in \mathcal{R}(\phi, T)$, there exists a unique L^1 optimal solution, which is given by the form of (2.4.5) for a parameter $\rho \ne 0$ and $\alpha \ge 0$.

Proof. Fix any $a \in \mathcal{R}(\phi, T)$ and let the L^1 optimal value for a be θ; see Lemma 2.4.1. Take any L^1 optimal solution $u^* \in \mathcal{U}(a, \phi, T)$, which, of course, satisfies $\|u^*\|_\infty \le 1$ and $\|u^*\|_1 = \theta$. It follows from Lemma 2.4.2 that $a \in \partial\mathcal{R}_\theta(\phi, T)$, and hence there are nonzero

normal vectors to $\mathcal{R}_\theta(\phi, T)$ at a by [64, COROLLARY 11.6.1] from the convexity of the set $\mathcal{R}_\theta(\phi, T)$. Take arbitrarily such a vector $\rho \neq 0$. By the definition of ρ, we have

$$\rho^\top z \leq \rho^\top a$$

for all $z \in \mathcal{R}_\theta(\phi, T)$, and hence

$$\int_0^T \rho^\top \phi(t) u(t) dt \leq \int_0^T \rho^\top \phi(t) u^*(t) dt \tag{2.4.9}$$

for all u with $\|u\|_\infty \leq 1$ and $\|u\|_1 \leq \theta$. It follows from Lemma 2.4.3 that $u^* = \bar{u}$, where $\bar{u}$ is defined by (2.4.5) accordingly to ρ. Hence any L^1 optimal solution is represented by the form of (2.4.5). Here, note that the maximizer of (2.4.6) is unique regardless of the selection of ρ. Indeed, for any nonzero normal vector ρ_1 and ρ_2 to $\mathcal{R}_\theta(\phi, T)$ at a, the equation (2.4.9) holds, and then $\bar{v}$ and $\bar{w}$ defined by (2.4.5) accordingly to ρ_1 and ρ_2 satisfy $\bar{v} = u^* = \bar{w}$. Thus the uniqueness of the optimal solution follows. □

By combining [65] and [66], we obtain the following corollary, which considers the case that Assumption 2.3.5 does not hold.

Corollary 2.4.5
There exists at least one L^1 optimal solution that takes only the values of 0 and ± 1, even if Assumption 2.3.5 does not hold.

Proof. This follows from [65, Lemma 1] and the proof of [66, Theorem]. Note that although the work [66] considers an L^1 optimal control problem for a linear system (i.e., ϕ is given by a form of $\phi(t) = \Phi^{-1}(t, 0)B(t)$, where $\Phi(t, 0)$ is the state transition matrix), the proof of [66] is also applicable to our problem. □

2.4.2 Characterization of the L^0 Optimal Solution

From now on, we characterize the L^0 optimal solutions in view of the relationship with the L^p-optimization for $p \in (0, 1]$. Theorem 2.4.6 (resp. Theorem 2.4.10) refers to the relationship with the L^1-optimization with (resp. without) Assumption 2.3.5. Theorem 2.4.11 considers the L^0- and L^p-equivalence for $p \in (0, 1)$.

Theorem 2.4.6
Fix $T > 0$, $\phi \in L^\infty([0, T], \mathbb{R}^{r \times m})$, and $a \in \mathcal{R}(\phi, T)$. Let the L^1 optimal value be given by θ. If Assumption 2.3.5 holds, then the L^0 optimal solution is unique and given by the L^1 optimal solution $\bar{u}$ defined by (2.4.5), where $\rho \neq 0$ is a normal vector to $\mathcal{R}_\theta(\phi, T)$ at a and α is taken so that (2.4.4) holds.

Proof. We first note that there uniquely exists an L^1 optimal solution $\bar{u}$ from Lemma 2.4.1, and it is given by the form of (2.4.5) from Theorem 2.4.4.

Let us show that $\bar{u}$ is also L^0 optimal. The following discussion is based on [54]. Since $\bar{u}(t)$ takes only the values of 0 and ± 1 on $[0, T]$, we have

$$\|\bar{u}\|_1 = \|\bar{u}\|_0. \tag{2.4.10}$$

On the other hand, $\|u\|_\infty \le 1$ readily implies

$$\|u\|_1 \le \|u\|_0. \tag{2.4.11}$$

From (2.4.10), (2.4.11) and the L^1 optimality of $\bar{u}$, we have

$$\|\bar{u}\|_0 = \|\bar{u}\|_1 \le \|u\|_1 \le \|u\|_0$$

for any $u \in \mathcal{U}(a, \phi, T)$. This implies that $\bar{u}$ is L^0 optimal. In addition, this guarantees the existence of L^0 optimal solutions.

We next take any L^0 optimal solution $\hat{u}$ and show $\hat{u} = \bar{u}$. From (2.4.10), (2.4.11), and the optimality of $\hat{u}$ and $\bar{u}$, we have

$$\|\bar{u}\|_1 \le \|\hat{u}\|_1 \le \|\hat{u}\|_0 = \|\bar{u}\|_0 = \|\bar{u}\|_1,$$

which yields $\|\hat{u}\|_1 = \|\bar{u}\|_1$. This implies that $\hat{u}$ is L^1 optimal. Thanks to the uniqueness of the L^1 optimal solution, the L^0 optimal solution is unique and given by $\bar{u}$. □

Corollary 2.4.7
Consider the following L^0 optimal control problem involving output constraints:

$$\begin{aligned} \underset{u}{\text{minimize}} \quad & \|u\|_0 \\ \text{subject to} \quad & \dot{x}(t) = Ax(t) + Bu(t), \\ & y(t) = Cx(t), \\ & x(0) = x_0, \quad y(T) = y_f, \quad \|u\|_\infty \le 1, \end{aligned}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the input, $y(t) \in \mathbb{R}^r$ is the output, $T > 0$ is the given final time of control, x_0 is the given initial state, y_f is the given target state, and $A \in \mathbb{R}^{n\times n}$, $B = [b_1, b_2, \ldots, b_m] \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{r\times n}$ are given constant matrices. Assume there exists at least one control that satisfies the constraints. If the matrix

$$\mathcal{M}_j \triangleq CA[b_j, Ab_j, \ldots, A^{n-1}b_j]$$

has row-full rank for all $j = 1, 2, \ldots, m$, then the optimization is equivalent to the L^1 optimal control problem.

Proof. Note that this optimization problem is a special case of Problem 2.3.1 with $p = 0$, $\phi(t) = Ce^{A(T-t)}B$, and $a = y_f - Ce^{AT}x_0$. Hence there exists a control satisfying the constraints if and only if

$$y_f - Ce^{AT}x_0 \in \left\{ \int_0^T Ce^{A(T-t)}Bu(t)dt : \|u\|_\infty \le 1 \right\}.$$

We show that Assumption 2.3.5 holds if $\mathcal{M}_j$ has row-full rank for all j, by which the proof is completed from Theorem 2.4.6. Let us suppose Assumption 2.3.5 fails. Then we have $\mu_L(\{t \in [0, T] : \rho^\top \phi_j(t) = \alpha\}) > 0$ for some $\rho \ne 0$, $\alpha \ge 0$, and $j \in \{1, 2, \ldots, m\}$. This implies

$$g(t) \triangleq [1, \rho^\top C]\exp\left(\begin{bmatrix} 0 & 0 \\ 0 & A(T-t) \end{bmatrix} \right)[-\alpha, b_j]^\top = 0$$

on a set with positive measures. Hence $g \equiv 0$ by [67]. From $\dfrac{d^k g}{dt^k}(T) = 0$ for any $k = 1, 2, \ldots$, we have $\rho^\top CA^k b_j = 0$ for any $k = 1, 2, \ldots$. It follows that $\mathcal{M}_j^\top \rho = 0$. Since ρ is nonzero, $\text{rank}(\mathcal{M}_j) < r$. This completes the proof. □

Remark 2.4.8
The equivalence theorem in [54, 55] considers the case of $C = I_{n\times n} \in \mathbb{R}^{n\times n}$, where $I_{n\times n}$ is the identity matrix, and claims that if the matrix $A[b_j, Ab_j, \ldots, A^{n-1}b_j]$ is nonsingular for all j, then the L^0- and L^1-equivalence holds. On the other hand, this corollary extends the analysis to systems having any output matrix C in $\mathbb{R}^{r\times n}$. The condition that $\mathcal{M}_j$ has row-full rank is known as the sufficient and necessary condition for the system

$$\begin{aligned}\dot{x}(t) &= Ax(t) + b_j u(t),\\ y(t) &= CAx(t)\end{aligned}$$

to be *output controllable* [68]. The notion focuses on the possibility of the system to steer the output to any value from any initial state in a finite time interval.

Remark 2.4.9
For example, when $C = I_{n\times n}$, the L^0- and L^1-equivalence for the double integrator system may fail because of the singularity of the A-matrix. Indeed, there exists an L^1 optimal control that is not L^0 optimal [55]. On the other hand, when $C = [c_1, c_2]$ with $c_1 \neq 0$, the condition in Corollary 2.4.7 holds, and hence the L^0- and L^1-equivalence is guaranteed. This indicates that the gap between the L^0 optimization and the L^1 optimization for $C = I_{n\times n}$ arises since the transfer function includes the first integrator.

As seen in the proof of Theorem 2.4.6, any L^1 optimal solution equipped with the bang-off-bang property is also L^0 optimal. On the other hand, any L^0 optimal solution is also L^1 optimal and takes only the values of 0 and ± 1 almost everywhere. Hence, Corollary 2.4.5 immediately shows the next theorem, which characterizes the relationship in the absence of Assumption 2.3.5.

Theorem 2.4.10
Fix $T > 0$, $\phi \in L^\infty([0,T], \mathbb{R}^{r\times m})$, and $a \in \mathcal{R}(\phi, T)$. Then there exists at least one L^0 optimal solution. In addition, the set of all L^0 optimal solutions is a subset of the set of all L^1 optimal solutions, even if Assumption 2.3.5 does not hold. More precisely, the set of all L^0 optimal solutions is given by the set of all L^1 optimal solutions that take only the values of 0 and ± 1 almost everywhere.

The remainder of this section examines the relationship between the L^0 optimization and the L^p optimization with $p \in (0,1)$. We here recall that Assumption 2.3.5 on ϕ is needed for the L^0- and L^1-equivalence. Indeed, the equivalence may fail in the absence of the assumption, as illustrated in Remark 2.4.9. On the other hand, the next theorem guarantees the L^0 optimization is equivalent to the L^p optimization for any $p \in (0,1)$ in general.

Theorem 2.4.11
Fix $T > 0$, $\phi \in L^\infty([0,T], \mathbb{R}^{r\times m})$, and $a \in \mathcal{R}(\phi, T)$. The L^0 optimization is equivalent to the L^p optimization for any $p \in (0,1)$.

Proof. It follows from Theorem 2.4.10 that there exists at least one L^0 optimal solution. Let us denote the set of all L^0 optimal solutions by $\mathcal{U}_0^*(a, \phi, T)$. Take any $\hat{u} \in \mathcal{U}_0^*(a, \phi, T)$. We first show that $\hat{u}$ is also L^p optimal for any $p \in (0,1)$. Fix any $p \in (0,1)$. Since $\hat{u}$ is also

L^1 optimal and takes only the values of 0 and ± 1 almost everywhere from Theorem 2.4.10, we have

$$\|\hat{u}\|_p^p = \|\hat{u}\|_1 \leq \|u\|_1 \leq \|u\|_p^p \tag{2.4.12}$$

for any $u \in \mathcal{U}(a, \phi, T)$. Note that the last inequality follows immediately from $\|u\|_\infty \leq 1$. Hence $\hat{u}$ is L^p optimal, and there exist L^p optimal solutions.

Let us denote the set of all L^p optimal solutions by $\mathcal{U}_p^*(a, \phi, T)$. Take any $\tilde{u} \in \mathcal{U}_p^*(a, \phi, T)$. We next show $\tilde{u}$ is also L^0 optimal. We now have

$$\|\tilde{u}\|_1 \leq \|\tilde{u}\|_p^p, \tag{2.4.13}$$

$$\|\hat{u}\|_1 \leq \|\tilde{u}\|_1, \tag{2.4.14}$$

$$\|\hat{u}\|_p^p = \|\hat{u}\|_0 = \|\hat{u}\|_1, \tag{2.4.15}$$

$$\|\tilde{u}\|_p^p \leq \|\hat{u}\|_p^p. \tag{2.4.16}$$

Note that the first inequality (2.4.13) follows from $\|\tilde{u}\|_\infty \leq 1$, the second inequality (2.4.14) follows from the L^1 optimality of $\hat{u}$, the third equalities (2.4.15) follows from $\hat{u}(t) \in \{0, \pm 1\}$ a.e. $t \in [0, T]$, and the forth inequality (2.4.16) follows from the L^p optimality of $\tilde{u}$. Then equality holds at all inequalities in (2.4.13), (2.4.14), and (2.4.16). In particular, we have

$$\|\tilde{u}\|_p^p = \|\tilde{u}\|_1, \tag{2.4.17}$$

$$\|\tilde{u}\|_1 = \|\hat{u}\|_0. \tag{2.4.18}$$

It follows from (2.4.17) that

$$\int_0^T \big(|\tilde{u}_j(t)|^p - |\tilde{u}_j(t)|\big) dt = 0$$

for all $j = 1, 2, \ldots, m$. Since $|\tilde{u}_j(t)|^p - |\tilde{u}_j(t)| \geq 0$ on $[0, T]$, we have $|\tilde{u}_j(t)|^p = |\tilde{u}_j(t)|$ a.e. on $[0, T]$. This implies that $\tilde{u}_j(t) \in \{0, \pm 1\}$ a.e. on $[0, T]$. Hence $\|\tilde{u}\|_0 = \|\tilde{u}\|_1$. Taking into account (2.4.18), we have $\|\tilde{u}\|_0 = \|\hat{u}\|_0$. This means $\tilde{u} \in \mathcal{U}_0^*$. This completes the proof. □

2.5 Numerical Algorithm

This section proposes an algorithm to numerically solve the sparse optimization problem. For this purpose, we adopt a time discretization approach, which is standard for numerical optimization [69, Sec. 2.3].

In this approach, we first divide the interval $[0, T]$ into N subintervals, $[0, T] = [0, h) \cup \cdots \cup [(N-1)h, Nh]$, and assume (or approximate) that the control $u(t)$ are constant over each subinterval. where h is the discretization step chosen such that $T = Nh$. Put the control vector

$$z \triangleq \big[u(0)^\top, u(h)^\top, \ldots, u((N-1)h)^\top\big]^\top \in \mathbb{R}^{mN}.$$

Then, the L^0 optimization problem is reformulated by

$$\begin{aligned} &\underset{z \in \mathbb{R}^{mN}}{\text{minimize}} && \|z\|_{\ell^0} \\ &\text{subject to} && \Phi z = a, \quad \|z\|_{\ell^\infty} \leq 1, \end{aligned} \tag{2.5.1}$$

where $\|\cdot\|_{\ell^0}$ and $\|\cdot\|_{\ell^1}$ are the ℓ^0 and ℓ^∞ norms in $\mathbb{R}^{mN}$, respectively, and

$$\Phi \triangleq \left[\int_0^h \phi(t)dt, \quad \int_h^{2h} \phi(t)dt, \ldots, \quad \int_{(N-1)h}^{Nh} \phi(t)dt\right] \in \mathbb{R}^{r\times mN}.$$

While it would be desirable to solve this optimization, the complexity of this problem grows exponentially with the number of the dimension N [70]. We then avoid to directly solve this problem and treat the L^1 optimization or L^p optimization with $p \in (0,1)$. This relaxation is motivated from Theorem 2.4.6 and Theorem 2.4.11.

2.5.1 via the L^1 Optimization Problem

When Assumption 2.3.5 is satisfied, the L^0 optimal solution is equal to the L^1 optimal solution by Theorem 2.4.6. Hence, it is enough to solve the L^1 optimization problem. Here, the L^1 cost function is approximated by $\|z\|_{\ell^1}$, where $\|\cdot\|_{\ell^1}$ is the ℓ^1 norm in $\mathbb{R}^{mN}$. Then, the L^1 optimization problem is reformulated by

$$\begin{aligned} &\underset{z\in\mathbb{R}^{mN}}{\text{minimize}} && \|z\|_{\ell^1} \\ &\text{subject to} && \Phi z = a, \qquad \|z\|_{\ell^\infty} \le 1. \end{aligned} \tag{2.5.2}$$

The optimization problem (2.5.2) is reducible to linear programming [71] and can be solved by standard numerical software packages, such as `cvx` with MATLAB [72, 73], based on the interior point method. However, for large scale problems, the computational burden of such an algorithm becomes heavy. We then give a more efficient algorithm based on the *alternating direction method of multipliers* (ADMM) [74, 75, 76].

Alternating direction method of multipliers

We here briefly review the ADMM algorithm. The ADMM solves the following type of convex optimization.

$$\begin{aligned} &\underset{z\in\mathbb{R}^{N_1}, y\in\mathbb{R}^{N_2}}{\text{minimize}} && f(z) + g(y) \\ &\text{subject to} && y = \Psi z \end{aligned} \tag{2.5.3}$$

where $f : \mathbb{R}^{N_1} \mapsto \mathbb{R}\cup\{\infty\}$ and $g : \mathbb{R}^{N_2} \mapsto \mathbb{R}\cup\{\infty\}$ are proper lower semi-continuous convex functions, and $\Psi \in \mathbb{R}^{N_2\times N_1}$. The algorithm of ADMM is given, for $y[0], w[0] \in \mathbb{R}^{N_2}$ and $\gamma > 0$, by

$$\begin{cases} z[j+1] \leftarrow \underset{z\in\mathbb{R}^{N_1}}{\arg\min}\left\{f(z) + \frac{1}{2\gamma}\left\|y[j] - \Psi z - w[j]\right\|^2\right\} \\ y[j+1] \leftarrow \mathrm{prox}_{\gamma g}\big(\Psi z[j+1] + w[j]\big) \\ w[j+1] \leftarrow w[j] + \Psi z[j+1] - y[j+1] \end{cases} \tag{2.5.4}$$

for $j = 0, 1, 2, \ldots$, where $\mathrm{prox}_{\gamma g}$ denotes the *proximity operator* of γg defined by

$$\mathrm{prox}_{\gamma g}(z) \triangleq \underset{y\in\mathbb{R}^{N_2}}{\arg\min}\, \gamma g(y) + \frac{1}{2}\|z - y\|^2.$$

We recall a convergence analysis of ADMM by Eckstein-Bertsekas [75].

Theorem 2.5.1 (Convergence of ADMM [75])
Consider the optimization problem (2.5.3). Assume that $\Psi^{\mathrm{T}}\Psi$ is invertible and that a saddle point of its unaugmented Lagrangian $\mathcal{L}_0(z, y, w) \triangleq f(z) + g(y) - (\Psi z - y)^{\mathrm{T}} w$ exists. Then the sequence $\{(z[j], y[j])\}_{j\in\mathbb{N}}$ generated by Algorithm (2.5.4) converges to a solution of (2.5.3).

Reformulation into ADMM-applicable form

In what follows, we reformulate our optimization problem described in (2.5.2) into the standard form in (2.5.3) to apply ADMM.

Let $\Omega_1 \triangleq \{z \in \mathbb{R}^{mN} : \|z\|_{\ell^\infty} \leq 1\}$ be the unit-ball of the infinity norm, and $\Omega_2 \triangleq \{a\}$ be the singleton consisting of the vector a. Define the indicator function of a nonempty closed convex set by

$$\iota_\Omega(z) \triangleq \begin{cases} 0, & \text{if } z \in \Omega, \\ \infty, & \text{otherwise.} \end{cases}$$

Then, we can rewrite the optimization problem (2.5.2) as

$$\underset{z\in\mathbb{R}^{mN}}{\text{minimize}} \quad \|z\|_{\ell^1} + \iota_{\Omega_1}(z) + \iota_{\Omega_2}(\Phi z). \tag{2.5.5}$$

Introducing new variables y_1, y_2 and y_3 such that $y_i = z$ $(i = 1, 2)$, and $y_3 = \Phi z$, we can translate (2.5.5) into

$$\begin{aligned} \underset{z\in\mathbb{R}^{N_1}, y\in\mathbb{R}^{N_2}}{\text{minimize}} \quad & \|y_1\|_{\ell^1} + \iota_{\Omega_1}(y_2) + \iota_{\Omega_2}(y_3) \\ \text{subject to} \quad & y = \Psi z \end{aligned} \tag{2.5.6}$$

where $N_1 \triangleq mN$, $N_2 \triangleq 2mN + r$, $y \triangleq [y_1^\top y_2^\top y_3^\top]^\top \in \mathbb{R}^{N_2}$, and

$$\Psi \triangleq \begin{bmatrix} I & I & \Phi^\top \end{bmatrix}^\top \in \mathbb{R}^{N_2 \times N_1}.$$

Finally, by setting

$$\begin{aligned} f(z) &\triangleq 0, \\ g(y) &\triangleq \|y_1\|_{\ell^1} + \iota_{\Omega_1}(y_2) + \iota_{\Omega_2}(y_3) \end{aligned}$$

the optimization problem (2.5.6) is reduced to the standard form of (2.5.3).

Computation

Since $f = 0$, the first step of (2.5.4) becomes strictly convex quadratic minimization, which boils down to solving linear equations, that is,

$$\begin{aligned} z[j+1] &= \underset{z\in\mathbb{R}^{mN}}{\arg\min} \tfrac{1}{2\gamma}\|y[j] - \Psi z - w[j]\|^2 \\ &= (\Psi^\top\Psi)^{-1}\Psi^\top(y[j] - w[j]) \\ &= \left(2I + \Phi^\top\Phi\right)^{-1} v[j] \end{aligned}$$

where

$$v[j] \triangleq \sum_{i=1}^{2}(y_i[j] - w_i[j]) + \Phi^\top(y_3[j] - w_3[j]).$$

On the other hand, the second step of (2.5.4) can be separated with respect to each y_i. For y_1, we have to compute the proximity operator of the ℓ_1 norm, which is reduced to a simple soft-thresholding operation [77]: for $l = 1, 2, \dots, mN$,

$$\begin{aligned}\left[\mathrm{prox}_{\gamma\|\cdot\|_{\ell^1}}(z)\right]_{(l)} &= \mathrm{prox}_{\gamma|\cdot|}(z_{(l)}) \\ &= \mathrm{sgn}(z_{(l)}) \max\{|z_{(l)}| - \gamma, 0\}\end{aligned}$$

where $(\cdot)_{(l)}$ denotes the l-th entry of a vector.

For y_2 and y_3, the computation of the proximity operators of the indicator functions are required. Since the proximity operator of the indicator function of a nonempty closed convex set Ω equals to the metric projection P_Ω onto Ω, the updates of y_2 and y_3 are reduced to calculating P_{Ω_1} and P_{Ω_2}, respectively. We can compute P_{Ω_1} as follows:

$$P_{\Omega_1}(z) \triangleq \begin{cases} z, & \text{if } \|z\|_{\ell^\infty} \le 1, \\ \tilde{z}, & \text{otherwise} \end{cases}$$

where

$$\begin{aligned}&\tilde{z} \triangleq [r_1 \dots r_N]^\top, \\ &r_i \triangleq \mathrm{sgn}(z_{(i)}) \min\{|z_{(i)}|, 1\}, \quad i = 1, 2, \dots, mN.\end{aligned}$$

Meanwhile, $P_{\Omega_2} = P_{\{a\}}$ is simply given by $P_{\Omega_2}(z) \triangleq a$.

2.5.2 via the L^p Optimization Problem, $p \in (0, 1)$

If Assumption 2.3.5 is not satisfied, then the L^1 optimal solution is not necessarily equal to the L^0 optimal solution, and hence the algorithm described above may not yield the sparse solution. Alternatively, we here solve the L^p optimization problem. Now, the L^p cost function is approximated by $\|z\|_{\ell^p}^p$, and hence the L^p optimization problem is reformulated by

$$\begin{aligned}\underset{z\in\mathbb{R}^{mN}}{\text{minimize}} \quad & \|z\|_{\ell^p}^p \\ \text{subject to} \quad & \Phi z = a, \quad \|z\|_{\ell^\infty} \le 1.\end{aligned} \tag{2.5.7}$$

Although this optimization is still NP-hard, to solve this type of concave optimization an effective algorithm called *Successive Linearization Algorithm* (SLA) is proposed in [78].

The algorithm of the SLA is given as follows: Choose a positive integer $q \in \{1, 2, \dots\}$ and a sufficiently small number $\delta > 0$. Solve the ℓ^1 minimization problem (2.5.2) in the following equivalent form:

$$\begin{aligned}\underset{z,w\in\mathbb{R}^{mN}}{\text{minimize}} \quad & 1_N^\top w \\ \text{subject to} \quad & \begin{cases} \Phi z = a, \quad \|z\|_{\ell^\infty} \le 1, \\ -w \le z \le w, \end{cases}\end{aligned} \tag{2.5.8}$$

and get the solution (z^0, w^0). Having (z^i, w^i) determine (z^{i+1}, w^{i+1}) by solving the following linear program:

$$\begin{aligned} \underset{z,w\in\mathbb{R}^{mN}}{\text{minimize}} \quad & (w^i)^{\top\left(\frac{1}{q}-1\right)} w \\ \text{subject to} \quad & \begin{cases} \Phi z = a, \quad \|z\|_{\ell^\infty} \le 1, \\ -w \le z \le w, \\ \delta 1_N \le w \le 1_N \end{cases} \end{aligned} \tag{2.5.9}$$

Stop when

$$(w^i)^{\top\left(\frac{1}{q}-1\right)} (w^i - w^{i+1}) = 0.$$

2.6 Example

We here illustrate the L^0- and L^1-equivalence, taking an infinite-dimensional system with an output constraint. Let us consider the electromagnetic molding machine [79], which is modeled by

$$\begin{cases} \dfrac{\partial \mathcal{X}(t,\xi)}{\partial t} = 2\gamma \dfrac{\partial}{\partial \xi}\left(\xi \dfrac{\partial \mathcal{X}(t,\xi)}{\partial \xi}\right), & \xi \in (0,1) \\ \mathcal{X}(t,\xi) = \dfrac{\delta}{\gamma} u(t), & \xi = 1 \\ \xi \dfrac{\partial \mathcal{X}(t,\xi)}{\partial \xi} = 0, & \xi = 0 \\ y(t) = \int_0^1 \mathcal{X}(t,\xi) d\xi \end{cases} \tag{2.6.1}$$

where $u(t)$ is the input current in the coil and $y(t)$ is the output suction force generated by magnetic flux; see the chapter for what the other parameters mean. We define an L^0 optimization problem as follows:

$$\begin{aligned} \underset{u}{\text{minimize}} \quad & \|u\|_0 \\ \text{subject to} \quad & (2.6.1), \\ & \mathcal{X}(0,\xi) = 0, \quad y(T) = y_f, \quad \|u\|_\infty \le 1, \end{aligned}$$

where $T > 0$ is a given final time of control and $y_f \in \mathbb{R}$ is a given target output. Note that the transfer function of the system from u to y is given by

$$G(s) = \frac{\delta}{\gamma} \frac{J_1\left(2\sqrt{-\dfrac{s}{2\gamma}}\right)}{\sqrt{-\dfrac{s}{2\gamma}} J_0\left(2\sqrt{-\dfrac{s}{2\gamma}}\right)},$$

where $J_k(z)$ is the Bessel function of the first kind of order k, i.e.,

$$J_k(z) \triangleq \sum_{q=0}^{\infty} \frac{(-1)^q}{q!(q+k)!} \left(\frac{z}{2}\right)^{2q+k}, \quad k \in \{1,2\}.$$

The system thus has an irrational transfer function, which has not been addressed in [54, 55, 58].

For the check of the L^0- and L^1-equivalence, we need to examine the shape of the impulse response of the system (2.6.1). Fig. 2.1 shows the impulse response when $\gamma = 10^{-5}$ and $\delta = 10^{-2}$. Obviously, there are no parts on which the response is constant. Hence this example satisfies Assumption 2.3.5 and the L^0 optimization is equivalent to the L^1 optimization. For example, let $T = 10$ and $y_f = 3$. Fig. 2.2 shows the optimal control input via the L^1 optimization and the L^2 optimal control input (Problem 2.3.1 with $p = 2$) for comparison. Fig. 2.3 shows the output $y(t)$ corresponding to the controls. Both controls achieve the given state transition $y(T) = 3$ while satisfying the magnitude constraint $\|u\|_\infty \leq 1$. We can also verify that the control obtained by the L^1 optimization exactly takes zero over a long time interval, while the L^2 optimization does not promote the sparsity. Thus, we can confirm the validity of the L^1 optimization.

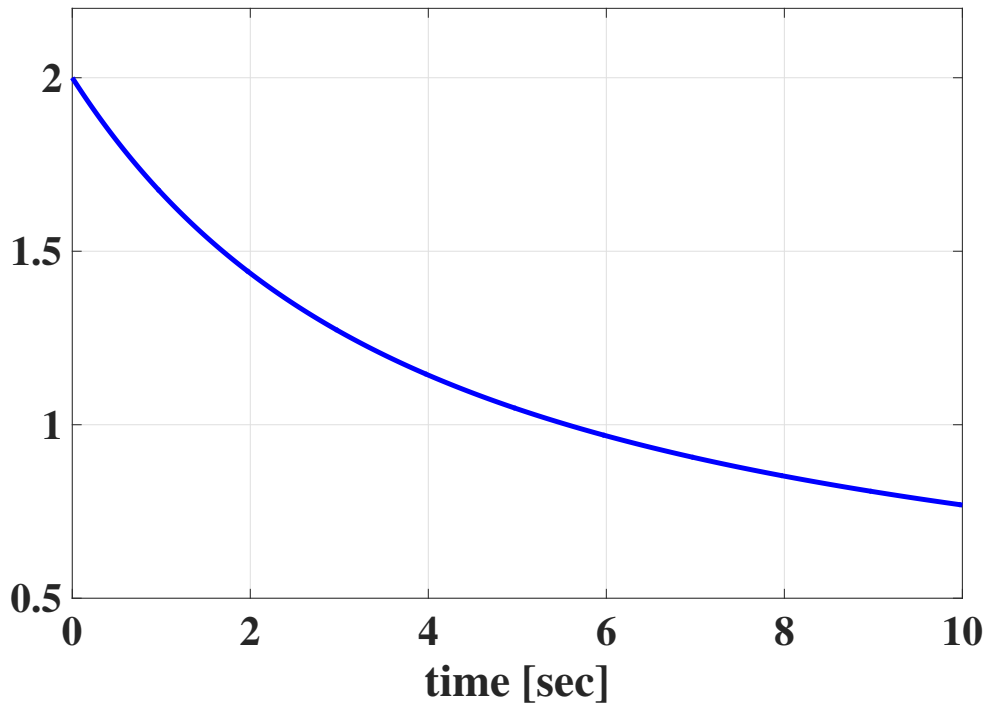

Figure 2.1: Impulse response of the electromagnetic molding machine

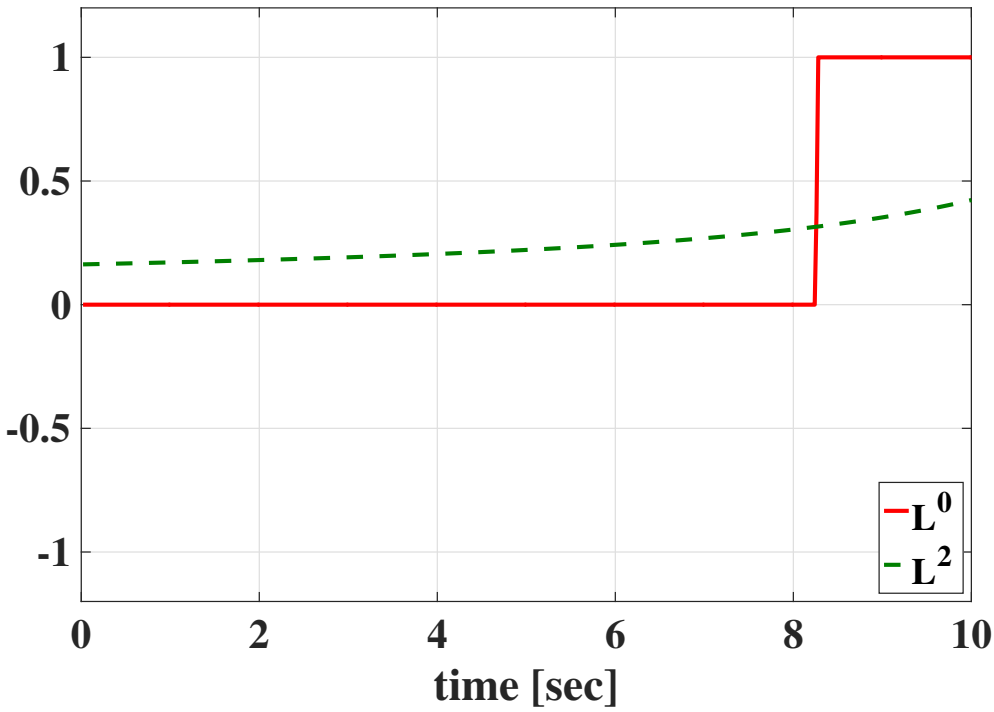

Figure 2.2: L^0 optimal control (solid line) and L^2 optimal control (dashed line).

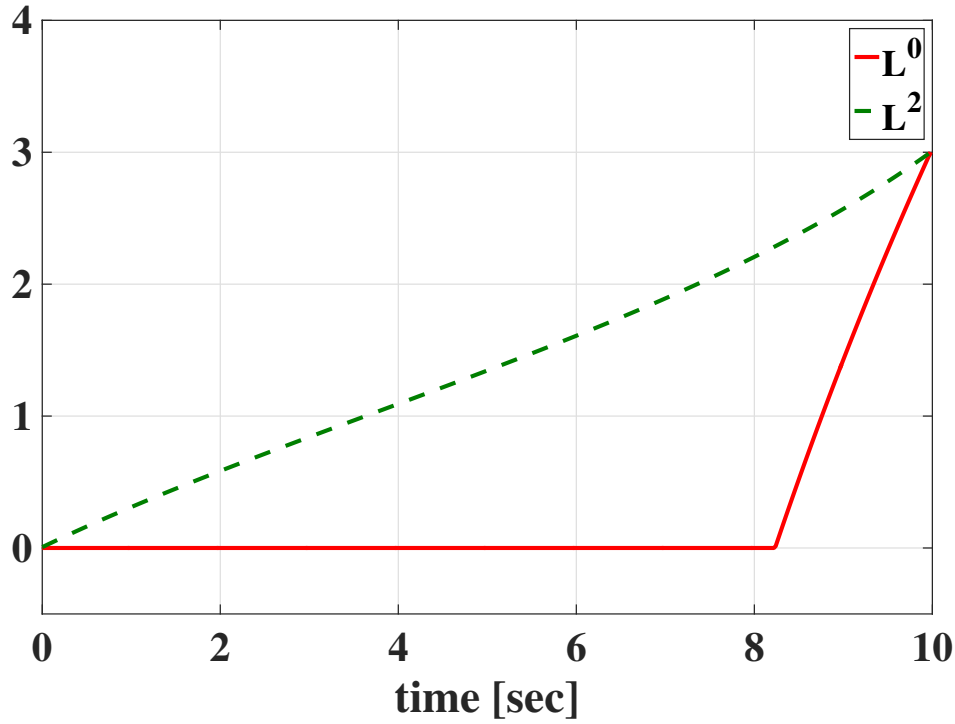

Figure 2.3: Output corresponding to L^0 optimal control (solid line) and L^2 optimal control (dashed line).

2.7 Conclusions

We have analyzed the relationship among the L^p optimizations with $p \in [0, 1]$ when the constraint is imposed in terms of a Volterra integral equation. The framework can address infinite-dimensional linear systems and output constraints. Motivated by an insight drawn from Pontryagin's minimum principle for finite-dimensional linear systems, we have derived a sufficient condition for the L^0- and L^1-equivalence and also clarified the construction of the L^0 optimal solutions. In particular, when the state is partially constrained, it has been proved that the output controllability guarantees the equivalence. In addition, we have proved that there exists at least one L^1 optimal solution that is also L^0 optimal, in general. We have finally given the L^0- and L^p-equivalence for any $p \in (0, 1)$.

Chapter 3

Sparsity-Constrained Controllability Maximization with Application to Time-Varying Control Node Selection

3.1 Introduction

As introduced in Chapter 1, this chapter investigates a control problem that maximizes a controllability metric so that the system can be steered towards all directions in the state space with small energy. This framework can also address nodes selection problem. The remainder of this chapter is organized as follows. Section 3.2 gives mathematical preliminaries. Section 3.3 formulates our optimal control problem after a review of the controllability metric. Section 3.4 gives a theoretical analysis. For the characterization of the optimal solution, we introduce a convex optimization problem. After that, we reformulate the convex optimization problem so that Pontryagin's minimum principle is applicable. Then, the optimal solution is characterized by using the costate vector. Based on this, we establish a condition for the main problem to be exactly solved via the convex optimization problem. Thanks to the result, the existence of optimal solutions to the main problem is also shown. Section 3.5 illustrates the application of the main problem to nodes selection problem, and Section 3.6 offers concluding remarks.

3.2 Mathematical Preliminaries

This section introduces notations that will be used throughout the chapter.

Let $m \in \mathbb{N}$ and $\Omega \subset \mathbb{R}$. For a vector $a = [a_1, a_2, \ldots, a_m]^\top \in \mathbb{R}^m$, $\mathrm{diag}(a)$ denotes the diagonal matrix whose (i,i)-component is given by a_i, and $a \in \Omega^m$ means $a_i \in \Omega$ for all i. Let N_1, $N_2 \in \mathbb{N}$. For a matrix $M \in \mathbb{R}^{N_1 \times N_2}$, $\mathrm{Tr}M$ denotes the trace of M.

3.3 Problem Formulation

3.3.1 Controllability Metrics

This chapter investigates an optimal control problem that maximizes a metric of controllability with a sparse control. Let us first define a linear model as follows:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad 0 \leq t \leq T, \tag{3.3.1}$$

where $x(t) \in \mathbb{R}^n$ is the state; $u(t) \in \mathbb{R}^m$ is the control input; $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are constant matrices; and $T > 0$ is the final time of control. This chapter is interested in a sparse control constrained in terms of the L^0 norm by $\|u\|_0 \leq \alpha$ with a given positive number $\alpha > 0$. This results in the following expression of the system (3.3.1):

$$\begin{aligned} &\dot{x}(t) = Ax(t) + BV(t)u(t), \quad 0 \leq t \leq T, \\ &V(t) \triangleq \mathrm{diag}(v(t)), \end{aligned} \tag{3.3.2}$$

where $v(t) \triangleq [v_1(t), v_2(t), \ldots, v_m(t)]^\top \in \mathbb{R}^m$ is a time-varying vector satisfying $\|v\|_0 \leq \alpha$ and $v(t) \in \{0,1\}^m$ for all $t \in [0,T]$. Note that the function $v(t)$ represents the activation schedule of the control input. More precisely, the j-th variable of the control input is able to affect the system at time t if and only if $v_j(t) = 1$.

In what follows, we introduce a metric of the controllability for the system (3.3.2). The classical notion of controllability denotes whether the system can be driven to any desired state from any initial state by using an appropriate control input. More precisely, the system (3.3.2) is said to be controllable on $[0,T]$ if for any $x_0, x_f \in \mathbb{R}^n$ there exists a control input $u(t)$ such that the state of the system $x(t)$ is driven from $x(0) = x_0$ to $x(T) = x_f$ by using the control input $u(t)$. It is well known that the system is controllable if and only if the following matrix called *controllability Gramian* is non-singular:

$$W_c \triangleq \int_0^T e^{At}\mathcal{B}(t)\mathcal{B}(t)^\top e^{A^\top t}dt,$$

where $\mathcal{B}(t) \triangleq BV(t) \in \mathbb{R}^{n\times m}$. Note that this classical controllability is a binary measure that determines whether the system is controllable or not, and it does not evaluate how easy the system is to control. Even if the system is theoretically controllable, control inputs might require high energy cost, which fails to realize desired state transitions in practice. Then, controllability measures that quantify the required energy cost of steering the system have been analyzed. We here recall the minimum-energy control problem:

$$\begin{aligned} \underset{u}{\text{minimize}} \quad & \int_0^T \|u(t)\|^2 dt \\ \text{subject to} \quad & \dot{x}(t) = Ax(t) + \mathcal{B}(t)u(t), \\ & x(0) = x_0, \quad x(T) = 0. \end{aligned}$$

The minimum control energy is then given by $x_0^\top W_c^{-1} x_0$ [80]. Based on this, recent works have been considered to make W_c as "large" as possible, and a number of controllability metrics have been proposed, including $\mathrm{Tr}W_c$, $\lambda_{\min}(W_c)$, $\det(W_c)$, and $\mathrm{rank}(W_c)$; see e.g. [81, 82]. While each measure has its own advantage, this chapter adopts the trace

of the controllability Gramian $\mathrm{Tr}W_c$ from the analytical perspective. This quantity is closely related to the average controllability on the ball $\{x_0 : \|x_0\| = 1\}$ by the following equation [83]:

$$\frac{\int_{\|x_0\|=1} x_0^\top W_c^{-1} x_0 dx_0}{\int_{\|x_0\|=1} dx_0} = \frac{1}{n}\mathrm{Tr}W_c^{-1}.$$

3.3.2 Main Problem

In this chapter, we consider one problem: how should we provide control inputs to the system in order to realize the best controllability when control inputs are constrained in terms of the L^0 norm? The problem is formulated as the following optimization:

Problem 3.3.1
Given $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $T > 0$, and $\alpha > 0$, find a time-varying matrix $V(t) \triangleq \mathrm{diag}(v(t))$, $v(t) \triangleq [v_1(t), v_2(t), \ldots, v_m(t)]^\top$, that solves

$$\begin{aligned} \underset{v_1,v_2,\ldots,v_m}{\text{maximize}} \quad & J(V) \triangleq \mathrm{Tr}\int_0^T e^{At}BV(t)V(t)^\top B^\top e^{A^\top t}dt \\ \text{subject to} \quad & v(t) \in \{0,1\}^m \quad \forall t \in [0,T], \quad \|v\|_0 \le \alpha. \end{aligned}$$

We are now interested in the average controllability as mentioned in subsection 3.3.1. Hence the cost function in the optimization problem is defined by the trace of the controllability Gramian according to the system (3.3.2).

Let us now reformulate Problem 3.3.1 for the subsequent theoretical analysis.

Lemma 3.3.2
Problem 3.3.1 is equivalent to the following problem:

$$\begin{aligned} \underset{v_1,v_2,\ldots,v_m}{\text{maximize}} \quad & J_1(V) \\ \text{subject to} \quad & v(t) \in \{0,1\}^m \quad \forall t \in [0,T], \quad \|v\|_0 \le \alpha. \end{aligned} \tag{3.3.3}$$

where

$$\begin{aligned} J_1(V) &\triangleq \int_0^T \big[f_1(t), f_2(t), \ldots, f_m(t)\big] v(t)dt, \\ f_j(t) &\triangleq b_j^\top e^{A^\top t} e^{At} b_j, \quad j = 1, 2, \ldots, m, \\ v(t) &\triangleq [v_1(t), v_2(t), \ldots, v_m(t)]^\top, \end{aligned}$$

and b_j is the j-th column of B.

Proof. It is enough to show $J(V) = J_1(V)$ for any v that satisfies the constraints. It follows from a property of the trace operator that

$$\begin{aligned} J(V) &= \mathrm{Tr}\int_0^T e^{At}BV(t)V(t)^\top B^\top e^{A^\top t}\, dt \\ &= \mathrm{Tr}\int_0^T B^\top e^{A^\top t} e^{At} BV(t)^2\, dt \end{aligned}$$

Here, we have $V(t)^2 = V(t)$, since a constraint in Problem 3.3.1 imposes the variables v_i's to take only the values of 0 and 1. We thus obtain the equivalent form (3.3.3) of Problem 3.3.1. □

In this chapter, we show that optimal solutions to Problem 3.3.1 can be obtained by solving a convex optimal control problem.

3.4 Analysis

We first introduce a convex relaxed problem of Problem 3.3.1.

Problem 3.4.1
Given $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $T > 0$, and $\alpha > 0$, find a time-varying matrix $V(t) \triangleq \mathrm{diag}(v(t))$, $v(t) \triangleq [v_1(t), v_2(t), \ldots, v_m(t)]^\top$, that solves

$$\begin{aligned} &\underset{v_1,v_2,\ldots,v_m}{\text{maximize}} \quad J_1(V) \\ &\text{subject to} \quad v(t) \in [0,1]^m \quad \forall t \in [0,T], \quad \|v\|_1 \le \alpha. \end{aligned}$$

We first show the discreteness of solutions to Problem 3.4.1, which guarantees that the optimal solutions to Problem 3.4.1 satisfy the constraints in Problem 3.3.1, which will be illustrated in the proof of Theorem 3.4.5. For this purpose, we recall Pontryagin's maximum principle [84, Theorem 22.2].

Proposition 3.4.2
Consider the following optimal control problem:

$$\begin{aligned} &\underset{w}{\text{minimize}} \quad \int_0^T \ell(t, w(t))dt \\ &\text{subject to} \quad \dot{z}(t) = Fz(t) + Gw(t),\ t \in [0,T] \text{ a.e.} \\ &\qquad\qquad\quad w(t) \in [0,1]^m,\ t \in [0,T] \text{ a.e.} \\ &\qquad\qquad\quad z(0) = z_0,\ z(T) \in E, \end{aligned} \tag{OC}$$

where ℓ is continuous, $z(t) \in \mathbb{R}$, $w(t) \in \mathbb{R}^m$, $F \in \mathbb{R}$, $G \in \mathbb{R}^{1\times m}$, $T > 0$, $z_0 \in \mathbb{R}$, and $E \subset \mathbb{R}$. Note that (ℓ, F, G, T, z_0, E) is given. Define the Hamiltonian function $H^\eta : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^m$ associated to Problem (OC) by

$$H^\eta(t, z, p, w) \triangleq p(Fz + Gw) - \eta\ell(t, w),$$

where η is either 0 or 1. Let the process (z_*, w_*) be a local minimizer for Problem (OC). Then, there exists an arc $p : [0,T] \to \mathbb{R}$ and a scalar η equal to 0 or 1 satisfying the nontriviality condition:

$$(\eta, p(t)) \neq 0 \quad \forall t \in [0,T], \tag{3.4.1}$$

the adjoint equation for almost every t:

$$-\dot{p}(t) = \frac{\partial H^\eta}{\partial z}(t, z_*(t), p(t), w_*(t)), \tag{3.4.2}$$

as well as the maximum condition for almost every t:

$$H^\eta(t, z_*(t), p(t), w_*(t)) = \sup_{w\in[0,1]^m} H^\eta(t, z_*(t), p(t), w). \tag{3.4.3}$$

From now on, we show the discreteness of optimal solutions to Problem 3.4.1.

Theorem 3.4.3
Assume that $f_j(t)$ defined in Lemma 3.3.2 is not constant on $[0,T]$ for all $j \in \{1,2,\dots,m\}$. Then any solution to Problem 3.4.1 takes only 0 and 1 almost everywhere.

Proof. We first reformulate Problem 3.4.1 into a form to which Pontryagin's maximum principle is applicable. For any v such that $v(t) \in [0,1]^m$ on $[0,T]$, we have

$$\|v\|_1 = \sum_{j=1}^{m} \int_0^T |v_j(t)|dt = \sum_{j=1}^{m} \int_0^T v_j(t)dt = \int_0^T [1,1,\dots,1]v(t)dt,$$

where $v(t) = [v_1(t), v_2(t), \dots, v_m(t)]^\top$. Note that the value is equal to the final state $y(T)$ of the system

$$\dot{y}(t) = [1,1,\dots,1]v(t)$$

with $y(0) = 0$. Hence the set of all variables satisfying the constraints in Problem 3.4.1 is equivalent to the set

$$\begin{aligned}\{v :\ & v(t) \in [0,1]^m \quad \forall t \in [0,T], \\ & \dot{y}(t) = [1,1,\dots,1]v(t), \quad y(0) = 0, \quad y(T) \le \alpha\},\end{aligned}$$

and Problem 3.4.1 is equivalently expressed by the following problem:

$$\begin{aligned}\underset{v_1,v_2,\dots,v_m}{\text{maximize}} \quad & J_1(V) \\ \text{subject to} \quad & \dot{y}(t) = [1,1,\dots,1]v(t), \\ & y(0) = 0, \quad y(T) \le \alpha, \\ & v(t) \in [0,1]^m \quad \forall t \in [0,T].\end{aligned} \tag{3.4.4}$$

We thus obtain an optimal control problem to which Pontryagin's maximum principle is applicable.

Define the Hamiltonian function $H^\eta : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^m$ associated to the problem (3.4.4) by

$$H^\eta(t,y,p,v) \triangleq p[1,1,\dots,1]v - \eta\ell(t,v),$$

where η is either 0 or 1 and

$$\begin{aligned}\ell(t,v) &\triangleq -\big[f_1(t), f_2(t), \dots, f_m(t)\big]v, \\ f_j(t) &\triangleq b_j^\top e^{A^\top t} e^{At} b_j, \quad j = 1,2,\dots,m.\end{aligned}$$

Let the process (y_*, v_*) be a local maximizer for the problem (3.4.4). Then there exists an arc $p : [0,T] \to \mathbb{R}$ and a scalar η equal to 0 or 1 satisfying the conditions (3.4.1), (3.4.2), and (3.4.3). Put the set on which (3.4.1), (3.4.2), and (3.4.3) hold by $I \subset [0,T]$. Note that $\mu_L(I) = T$. Since we now have $\frac{\partial H^\eta}{\partial y} = 0$, it follows from (3.4.2) that there exists $p_0 \in \mathbb{R}$ and $p(t) = p_0$ for $t \in I$. Then, it follows from (3.4.3) that

$$\big(p_0 + \eta f_j(t)\big)v_j^*(t) = \sup_{v_j \in [0,1]} \big(p_0 + \eta f_j(t)\big)v_j$$

for each $j \in \{1, 2, \ldots, m\}$ and $t \in I$. Note that the supremum is attained by a point in $[0, 1]$, since the right hand side is a linear function of v_j on a closed interval. Hence

$$v_j^*(t) = \arg\max_{v_j \in [0,1]} \big(p_0 + \eta f_j(t)\big) v_j.$$

The characterization is divided into the following two cases.

1. If $\eta = 0$, then $p_0 \neq 0$ from (3.4.1). Hence

$$v_j^*(t) = \arg\max_{v_j \in [0,1]} p_0 v_j = \begin{cases} 1, & \text{if } p_0 > 0, \\ 0, & \text{if } p_0 < 0. \end{cases}$$

2. If $\eta = 1$, then

$$v_j^*(t) = \arg\max_{v_j \in [0,1]} \big(p_0 + f_j(t)\big) v_j.$$

   Hence

$$v_j^*(t) = \begin{cases} 1, & \text{if } p_0 + f_j(t) > 0, \\ 0, & \text{if } p_0 + f_j(t) < 0, \end{cases}$$

   and $v_j^*(t)$ is not determined if $p_0 + f_j(t) = 0$. Here, put $I_j \triangleq \{t \in [0, T] \cap I : p_0 + f_j(t) = 0\}$.

In what follows, we show that for each j, if $f_j(t)$ is not constant, then v_j^* takes only 0 and 1 almost everywhere. For the purpose, we show that if $\eta = 1$ and $\mu_L(I_{j_0}) > 0$ for some j_0, then $f_{j_0}(t)$ is constant. Put

$$\phi_{j_0}(t) \triangleq p_0 + f_{j_0}(t). \tag{3.4.5}$$

Then, $\phi_{j_0}(t) = 0$ for $t \in I_{j_0}$. Since $\phi_{j_0}(t)$ is an analytic function, it follows from $\mu_L(I_{j_0}) > 0$ that $\phi_{j_0} \equiv 0$, as described in [85]. This implies that $f_{j_0}(t)$ is constant, and then completes the proof. □

Remark 3.4.4
As illustrated in the proof, the j-th component $v_j^*(t)$ of the optimal solution may switch the values only at time t such that $f_j(t) + p_0 = 0$. Since the function $f_j(t) + p_0$ is analytic, the set of all zeros of the function does not contain accumulation points, which ensures the nonexistence of Zeno phenomena.

The following theorem is the main result, which shows the equivalence between Problem 3.3.1 and Problem 3.4.1.

Theorem 3.4.5
Assume that $f_j(t)$ defined in Lemma 3.3.2 is not constant for all j. Denote the set of all solutions to Problem 3.3.1 and Problem 3.4.1 by $\mathcal{V}_1^*$ and $\mathcal{V}_2^*$, respectively. If $\mathcal{V}_2^* \neq \emptyset$, then $\mathcal{V}_1^* = \mathcal{V}_2^*$.

Proof. Let $\hat{v} \triangleq [\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_m]^\top$ be an optimal solution to Problem 3.4.1. Then all $\hat{v}_j$'s take only the values of 0 and 1 almost everywhere by Theorem 3.4.3. Note that the null set $\cup_{j=1}^m \{t \in [0,T] : \hat{v}_j(t) \notin \{0,1\}\}$ does not affect the cost, and hence we can adjust the variables so that $\hat{v}_j(t) \in \{0,1\}$ on $[0,T]$ for all j, without loss of the optimality. We have

$$\sum_{j=1}^m \|\hat{v}_j\|_1 = \sum_{j=1}^m \int_{\{t\in[0,T]:\hat{v}_j(t)\neq 0\}} |\hat{v}_j(t)|dt = \sum_{j=1}^m \|\hat{v}_j\|_0, \tag{3.4.6}$$

where we used the discreteness of $\hat{v}_j$. Since $\hat{v}$ satisfies the constraints in Problem 3.4.1, we have $\sum_{j=1}^m \|\hat{v}_j\|_0 \leq \alpha$ from (3.4.6). Thus, $\hat{v}$ also satisfies the constraints in Problem 3.3.1.

Here,

$$\max_{\substack{v(t)\in\{0,1\}^m,\forall t\\ \|v\|_0\leq\alpha}} J_1(V) \leq \max_{\substack{v(t)\in[0,1]^m,\forall t\\ \|v\|_0\leq\alpha}} J_1(V). \tag{3.4.7}$$

Since we have $\|g\|_1 \leq \|g\|_0$ for any measurable function $g(t) \in \mathbb{R}$ with $g(t) \in [0,1]$ on $[0,T]$, we have

$$\{v : v(t) \in [0,1]^m \ \forall t, \ \|v\|_0 \leq \alpha\} \subset \{v : v(t) \in [0,1]^m \ \forall t, \ \|v\|_1 \leq \alpha\}.$$

Hence

$$\max_{\substack{v(t)\in[0,1]^m,\forall t\\ \|v\|_0\leq\alpha}} J_1(V) \leq \max_{\substack{v(t)\in[0,1]^m,\forall t\\ \|v\|_1\leq\alpha}} J_1(V). \tag{3.4.8}$$

It follows from inequalities (3.4.7) and (3.4.8) that

$$\max_{\substack{v(t)\in\{0,1\}^m,\forall t\\ \|v\|_0\leq\alpha}} J_1(V) \leq \max_{\substack{v(t)\in[0,1]^m,\forall t\\ \|v\|_1\leq\alpha}} J_1(V). \tag{3.4.9}$$

Here, $\hat{v}$ is an optimal solution to Problem 3.4.1, and hence

$$\max_{\substack{v(t)\in[0,1]^m,\forall t\\ \|v\|_1\leq\alpha}} J_1(V) = J_1(V_2^*) \tag{3.4.10}$$

where $V_2^* \triangleq \mathrm{diag}(\hat{v})$. On the other hand, since $\hat{v}$ satisfies the constraints in Problem 3.3.1, we have

$$J_1(V_2^*) \leq \max_{\substack{v(t)\in\{0,1\}^m,\forall t\\ \|v\|_0\leq\alpha}} J_1(V). \tag{3.4.11}$$

Hence, we have

$$J_1(V_2^*) = \max_{\substack{v(t)\in\{0,1\}^m,\forall t\\ \|v\|_0\leq\alpha}} J_1(V) \tag{3.4.12}$$

by (3.4.9), (3.4.10), and (3.4.11). This means that $\hat{v}$ is a solution to Problem 3.3.1. Hence we have $\mathcal{V}_2^* \subset \mathcal{V}_1^*$ and $\mathcal{V}_1^* \neq \emptyset$.

Let us take any optimal solution $\tilde{v} \triangleq [\tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_m]^\top \in \mathcal{V}_1^*$. Then we have $\tilde{v}_j(t) \in \{0,1\}$ on $[0,T]$ for all j and

$$\sum_{j=1}^m \|\tilde{v}_j\|_1 = \sum_{j=1}^m \|\tilde{v}_j\|_0 \leq \alpha.$$

Hence $\tilde{v}$ also satisfies the constraints in Problem 3.4.1. In addition, it follows from (3.4.12) that $J_1(V_1^*) = J_1(V_2^*)$, where $V_1^* \triangleq \mathrm{diag}(\tilde{v})$. Therefore, $\tilde{v} \in \mathcal{V}_2^*$ and $\mathcal{V}_1^* \subset \mathcal{V}_2^*$. This gives $\mathcal{V}_1^* = \mathcal{V}_2^*$. □

Remark 3.4.6
When a function f_j is constant for some j, optimal solutions to Problem 3.4.1 do not necessarily take only values of 0 and 1. Indeed, once the j-th component of an optimal solution to Problem 2 takes the value 1 on an interval, there exists an optimal solution whose j-th component takes besides the values of 0 and 1. This implies that the set $\mathcal{V}_2^*$ is not necessarily included in the set $\mathcal{V}_1^*$ when a function f_j is constant.

In Theorem 3.4.5, we assume the existence of optimal solutions to Problem 3.4.1 in order to discuss the equivalence $\mathcal{V}_1^* = \mathcal{V}_2^*$. We next show the existence of solutions to Problem 3.3.1 and Problem 3.4.1. For the purpose, we first show that there exist solutions to Problem 3.4.1. Note that this guarantees the existence of solutions to Problem 3.3.1 if $f_j(t)$ is not constant for all j by Theorem 3.4.5. We then finally show that there exists at least one solution to Problem 3.3.1, even if $f_j(t)$ is constant for some j.

Theorem 3.4.7
For any $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $T > 0$, and $\alpha > 0$, there exists at least one optimal solution to Problem 3.3.1 and Problem 3.4.1, respectively.

Proof. We first show the existence of solutions to Problem 3.4.1, which is similar to the proof of Lemma 2.4.1. For this, we consider the equivalent problem (3.4.4). Let us denote the set of all variables satisfying the constraints in the problem (3.4.4) by $\mathcal{V}$, i.e,

$$\mathcal{V} \triangleq \left\{ v \in L^1 : \int_0^T [1, 1, \ldots, 1] v(t) dt \leq \alpha,\ v(t) \in [0,1]^m\ \forall t \in [0,T] \right\}.$$

Since $\mathcal{V} \neq \emptyset$, we can define

$$\theta \triangleq \sup\{J_1(V) : v \in \mathcal{V}\}.$$

Then there exists a sequence $\{v^{(l)}\}_{l\in\mathbb{N}} \subset \mathcal{V}$ such that $\lim_{l\to\infty} J_1(v^{(l)}) = \theta$. Here, the set $\{w \in L^\infty : \|w\|_\infty \leq 1\}$ is sequentially compact in the weak* topology of L^∞ [63]. Then it follows that there exist a measurable function $v^{(\infty)}$ with $v^{(\infty)}(t) \in [0,1]^m$ almost everywhere and a subsequence $\{v^{(l')}\}$ such that each component $v_j^{(l')}$ converges to $v_j^{(\infty)}$, the j-th component of $v^{(\infty)}$, in the weak* topology of L^∞, that is, we have

$$\lim_{l'\to\infty} \int_0^T \big(v_j^{(l')}(t) - v_j^{(\infty)}(t)\big) g(t) dt = 0$$

for any $g \in L^1$ and $j = 1, 2, \ldots, m$. Then

$$\int_0^T [1, 1, \ldots, 1] v^{(\infty)}(t) dt = \lim_{l'\to\infty} \int_0^T [1, 1, \ldots, 1] v^{(l')}(t) dt \leq \alpha$$

and hence $v^{(\infty)} \in \mathcal{V}$. In addition, we have

$$\begin{aligned} J_1(V^{(\infty)}) &= \int_0^T \big[f_1(t), f_2(t), \ldots, f_m(t)\big] v^{(\infty)}(t)dt \\ &= \lim_{l' \to \infty} \int_0^T \big[f_1(t), f_2(t), \ldots, f_m(t)\big] v^{(l')}(t)dt \\ &= \lim_{l' \to \infty} J_1(V^{(l')}) = \theta, \end{aligned}$$

where $V^{(\infty)} \triangleq \mathrm{diag}(v^{(\infty)})$ and $V^{(l')} \triangleq \mathrm{diag}(v^{(l')})$. This shows that $v^{(\infty)}$ is an optimal solution to Problem 3.4.1.

We next show that there exists at least one optimal solution to Problem 3.4.1 that takes only the values of 0 and 1 *even if $f_j(t)$ is constant for some j*. Here, let us denote the set of all solutions to Problem 3.4.1 by $\mathcal{V}_2^*$, which is not empty from the discussion above, and take any optimal solution $v^* \in \mathcal{V}_2^*$. Now, it is enough to consider the case that $\eta = 1$ and $K \triangleq \{j \in \{1, 2, \ldots, m\} : \mu_L(I_j) > 0\} \neq \emptyset$, where η and I_j are defined in the proof of Theorem 3.4.3, since for the other cases v^* obviously take only the values of 0 and 1. Note that the j-th component v_j^* of v^* satisfies $v_j^*(t) \in \{0, 1\}$ on $[0, T]$ for $j \notin K$. For any $j \in K$, it follows from the proof of Theorem 3.4.5 that $\phi_j(t) \equiv 0$, where ϕ_j is defined by (3.4.5). Hence $f_j(t) = -p_0$ on $[0, T]$ for all $j \in K$. Here we take a measurable function v such that

$$\begin{aligned} &\int_0^T v_j(t)dt = \int_0^T v_j^*(t)dt, \\ &v_j(t) \in \{0, 1\} \quad \forall t \in [0, T] \end{aligned}$$

for $j \in K$ and $v_j = v_j^*$ for $j \notin K$. Then the function $v \triangleq [v_1, v_2, \ldots, v_m]^\top$ satisfies $J_1(V) = J_1(V^*)$, $\|v\|_1 = \|v^*\|_1 \leq \alpha$, and $v(t) \in \{0, 1\}^m$ on $[0, T]$, where $V = \mathrm{diag}(v)$ and $V^* = \mathrm{diag}(v^*)$. Hence v is a solution to Problem 3.4.1 that takes only 0 and 1.

Finally, the set of all solutions to Problem 3.4.1 that take only 0 and 1 is equal to $\mathcal{V}_1^*$, which is the set of all solutions to Problem 3.3.1. This can be verified by the proof of Theorem 3.4.5. Thus, we obtain $\mathcal{V}_1^* \neq \emptyset$. □

3.5 Application to Control Node Selection

In this section, we illustrate the application of Problem 3.3.1 to control node selection. We first briefly review the node selection problem. The purpose of the problem is to identify the set of nodes with exogenous control inputs that can effectively guide the system's entire dynamics. The selected nodes are called *control nodes*, and the selection techniques of control nodes have been extensively studied in the context of complex networks. In recent works, control nodes are chosen in order to optimize a metric of controllability [86, 87, 81, 82, 88, 89]. For example, [86] considers the minimum set of control nodes that guarantees the classical controllability proposed in [90]; [87] considers the structural controllability; and [81, 82] introduce quantities that evaluate how much the system is easy to control, such as the trace, the minimum eigenvalue, and the rank of the controllability Gramian. While these works investigate the selection problem in which the set of control nodes is

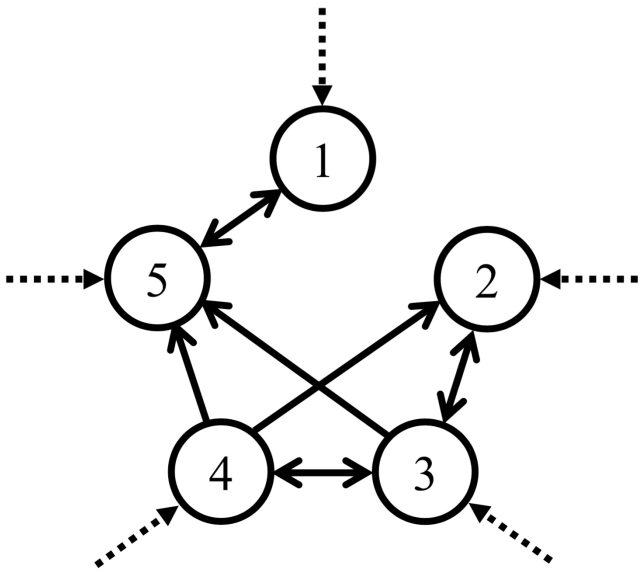

Figure 3.1: Graph associated with A; Dashed lines show nodes that can be provided exogenous inputs.

fixed over the time, more recent works [91, 92] alternatively select the time-varying set of control nodes for discrete-time systems.

In the context of the selection problem, our model (3.3.2) is interpreted as follows: $x(t) \triangleq [x_1(t), x_2(t), \ldots, x_n(t)]^\top$ is the state vector of the network consisting of n nodes, where $x_i(t)$ is the state of the i-th node at time t; $u(t)$ is the exogenous control input that influences the network dynamics; A is the dynamics matrix that represents the information flow among nodes. In particular, the node selection problem is interested in the design of the B-matrix, since it locates control nodes. In our case, $BV(t)$ plays the role. More precisely, if $v_j(t) = 1$, then the input $u_j(t)$ is provided to the system through the vector b_j at time t, and if $v_j(t) = 0$, then the input $u_j(t)$ is not provided to the system. Thus Problem 3.3.1 can answer to a question of when and where exogenous control inputs should be provided. In other words, Problem 3.3.1 is considered as a node selection problem that extracts time-varying control nodes. To the best of our knowledge, the time-varying control node selection for continuous-time systems has not yet been proposed.

We next give an example of node selection problem based on the proposed method. We consider the network consisting of 5 nodes. Let A in (3.3.2) be the adjacency matrix defined by

$$A \triangleq \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

and take B as $I_{5\times 5}$, which is the identity matrix of dimension 5. Fig. 3.1 shows the graph according to the adjacency matrix A.

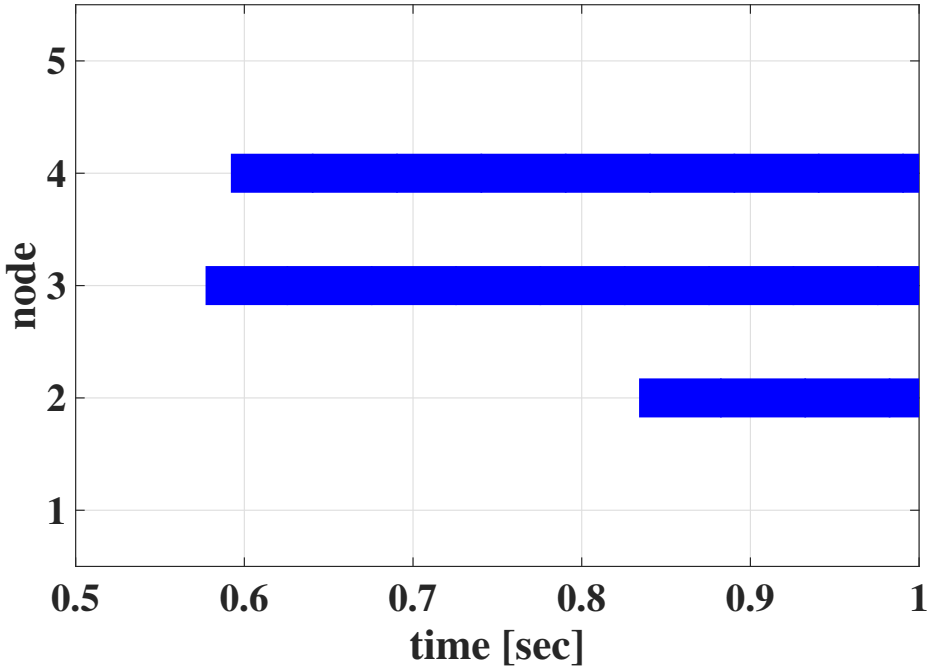

Figure 3.2: Profile of control nodes

For this network, we simulated the proposed method with $T = 1$ and $\alpha = 1$. In this example, each node can be a control node since $B = I_{5\times 5}$, but the constraint $\|v\|_0 \leq \alpha = 1$ imposes us to provide control inputs at most 1 sec. Note that this example satisfies the assumption of Theorem 3.4.5, and hence the optimal solution can be obtained by solving Problem 3.4.1, for which we can use, for example, CVX in MATLAB [72] after applying a time discretization.

Fig. 3.2 shows the profile of control nodes selected on the interval $[0, T]$. Note that we plot the profile only on $[0.5, T]$, since no node is activated on $[0, 0.5]$ in this example. We can see that the set of control nodes is not constant on $[0, T]$. Indeed, the sets of control nodes on $[0.576, 0.591)$, $[0.591, 0.833)$, and $[0.833, 1]$ are $\{3\}$, $\{3, 4\}$, and $\{2, 3, 4\}$, respectively. Note that our framework may select several nodes at the same time since $\|v\|_0 \triangleq \sum_{j=1}^{m} \mu_L(\{t \in [0, T] : v_j(t) \neq 0\})$.

3.6 Conclusions

In this chapter, we have analyzed an optimal control problem that maximizes a controllability metric when a sparse control is applied. This analysis enables us to find an activation schedule of control inputs that can steer the system while saving energy. We have analytically shown the existence of the optimal solutions and proved that the solutions can be obtained via a convex optimization if each function (i.e. f_j) in front of decision variables is not constant. We have also illustrated the application of the optimization problem addressed in the chapter to the control node selection problem. Our optimization can select time-varying control nodes for continuous-time systems.

Chapter 4

Maximum Hands-off Distributed Control for Consensus of Multi-Agent Systems with Sampled-Data State Observation

4.1 Introduction

A multi-agent system (MAS) consists of distributed autonomous agents that exchange information with their neighbors on a network and update their own controls in order to perform a shared task. The MAS framework has been analyzed in a wide variety of applications, such as formation controls [93], sensor networks [94], and flocking of animals [95], to name a few. Among the numerous research topics in MAS, the *consensus* problem particularly focuses on the mechanism of how agents can reach an agreement at a common value through local interactions [96, 97, 98].

This chapter proposes a consensus algorithm based on maximum hands-off control, or sparse optimal control, and sampled-state observation. The remainder of this chapter is organized as follows. Section 4.2 gives mathematical preliminaries. Section 4.3 reviews the maximum hands-off control introduced in Chapter 2 for the readability. Sections 4.4 and 4.5 are the main sections of this chapter. We first formulate the consensus problem in Subsection 4.4.1. In the subsequent subsections in Section 4.4, we consider the first-order consensus problem. We give a consensus scheme based on the idea of the maximum hands-off control with sampled-data state observation in Subsection 4.4.2 and discuss the feasibility and characterization of the proposed consensus scheme in Subsection 4.4.3. Section 4.5 gives a consensus scheme for the second-order consensus problem. Section 4.6 presents two examples to illustrate the proposed maximum hands-off distributed control for the consensus. Although this chapter analyzes noiseless systems, we also consider a numerical example in the presence of noise. The simulation results show robustness against noise by the proposed distributed control, while we note that it is also implied that a feasibility problem may occur under noise. Section 4.7 offers concluding remarks.

4.2 Mathematical Preliminaries

This section reviews basic definitions, facts, and notation that will be used throughout the chapter.

We denote by $\mathbf{1}_N$ the vector in $\mathbb{R}^N$ whose components are equal to 1. For two vectors z, $w \in \mathbb{R}^N$, the notation $z \leq w$ corresponds to the component-wise inequality, that is, $z \leq w$ if $z_k \leq w_k$ for $k = 1, 2, \ldots, N$.

A *digraph* $\mathcal{G}$ consists of a pair $(\mathcal{V}, \mathcal{E})$ with a finite nonempty set of *nodes* $\mathcal{V} = \{1, 2, \ldots, N\}$ and a finite set of *edges* $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. The edge set $\mathcal{E}$ represents the structure of information flow of the digraph, i.e., if information flows from $j \in \mathcal{V}$ to $i \in \mathcal{V}$, then $(j, i) \in \mathcal{E}$, otherwise $(j, i) \notin \mathcal{E}$. The node $j \in \mathcal{V}$ is said to be a *neighbor* of a node $i \in \mathcal{V}$ if $(j, i) \in \mathcal{E}$, and the set of all neighbors of node i is denoted by $\mathcal{N}_i$, that is,

$$\mathcal{N}_i \triangleq \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}.$$

We denote by $|\mathcal{N}_i|$ the number of elements of the set $\mathcal{N}_i$. The *maximum degree* of a digraph $\mathcal{G}$ is defined as $\max_{i \in \mathcal{V}} |\mathcal{N}_i|$ and we denote it by Δ. A sequence of nodes $(i_0, i_1, \ldots, i_\ell)$ of a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is said to be a *directed path* if $\ell \geq 1$ and $(i_k, i_{k+1}) \in \mathcal{E}$ for all $k = 0, 1, 2, \ldots, \ell - 1$. A *directed tree* is a connected digraph in which every node has exactly one neighbor, except for one node called the *root*, which has no neighbor and from which every other node is reachable through a directed path. A digraph $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ is said to be a *subgraph* of a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ if $\mathcal{V}' \subset \mathcal{V}$ and $\mathcal{E}' \subset \mathcal{E}$. A subgraph $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ of a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is said to be a *spanning tree* of $\mathcal{G}$ if $\mathcal{V}' = \mathcal{V}$ and $\mathcal{G}'$ is a directed tree. The digraph $\mathcal{G}$ is said to have a spanning tree if a spanning tree is a subgraph of $\mathcal{G}$. The *adjacency matrix* $A = (a_{ij}) \in \mathbb{R}^{N \times N}$ of a digraph $\mathcal{G}$ is defined as

$$a_{ij} = \begin{cases} 1, & \text{if } (j, i) \in \mathcal{E} \text{ and } i \neq j, \\ 0, & \text{otherwise.} \end{cases}$$

A digraph is said to be *balanced* if

$$\sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} a_{ji}$$

for all $i \in \mathcal{V}$, where $A = (a_{ij})$ is the adjacency matrix of the digraph. The *graph Laplacian* $L = (l_{ij}) \in \mathbb{R}^{N \times N}$ of a digraph is defined as

$$l_{ij} = \begin{cases} -a_{ij}, & \text{if } i \neq j, \\ |\mathcal{N}_i|, & \text{if } i = j. \end{cases}$$

A matrix $P \triangleq I - \epsilon L$ is called a *Perron matrix* with a parameter $\epsilon > 0$.

4.3 Maximum Hands-off Control

Here, we briefly review the maximum hands-off control defined in Chapter 2. Although we focus on first- and second-order agents in this chapter, we here discuss a more general linear time-invariant system modeled by

$$\dot{x}(t) = Fx(t) + Gu(t), \quad 0 \leq t \leq T_f, \tag{4.3.1}$$

where $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{n\times 1}$, and $T_f > 0$ is a fixed final time of control. Note that the first- and second-order systems discussed in the main part of this chapter are modeled with

$$F = 0, \quad G = 1 \quad \text{and} \quad F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

respectively.

For the system (4.3.1), we call a control $u \in L^1$ *feasible* if it steers the state $x(t)$ from a given initial state $x(0) = x_0 \in \mathbb{R}^n$ to a target state x^f at time T_f (i.e., $x(T_f) = x^f$) and satisfies the magnitude constraint $\|u\|_\infty \leq 1$. We denote by $\mathcal{U}(x_0, x^f, T_f)$ the set of all feasible controls for an initial state $x_0 \in \mathbb{R}^n$, target state $x^f \in \mathbb{R}^n$, and the final time of control $T_f > 0$.

As introduced in Chapter 2, the maximum hands-off control is a control $u \in \mathcal{U}(x_0, x^f, T_f)$ that has the maximum time duration on which $u(t)$ takes the value 0. Although the problem to seek such a control is formulated in Problem 2.3.1, we here again describe it in a form appropriate to the context. This optimization problem will be solved for each agent in our proposed consensus scheme as described in Sections 4.4 and 4.5.

Problem 4.3.1
For a given initial state $x_0 \in \mathbb{R}^n$, target state $x^f \in \mathbb{R}^n$, and the final time of control $T_f > 0$, find a feasible control $u \in \mathcal{U}(x_0, x^f, T_f)$ that minimizes $J(u) \triangleq \|u\|_0$.

For the readability, the following theorem summarizes some results in Chapter 2. This will be used to show a condition under which the feasibility of the proposed control scheme is guaranteed; see Theorem 4.4.4 and Theorem 4.5.1.

Theorem 4.3.2
For a given final time of control $T_f > 0$, define the set

$$\mathcal{R} \triangleq \left\{ \int_0^{T_f} e^{-Ft} G u(t) dt : \|u\|_\infty \leq 1 \right\}. \tag{4.3.2}$$

Then, there exist maximum hands-off controls that steer the state from an initial state $x_0 \in \mathbb{R}^n$ to a target state $x^f \in \mathbb{R}^n$ if and only if the inclusion

$$x_0 - e^{-FT_f} x^f \in \mathcal{R}$$

holds. Moreover, the maximum hands-off controls are given by the L^1 optimal controls that take only values ± 1 and 0.

Proof. See Theorem 2.4.10. □

4.4 Consensus by Maximum Hands-off Control for First-Order Systems

In this chapter, we adopt the idea of the maximum hands-off control to the consensus control problem in multi-agent systems. We begin with the problem formulation.

4.4.1 Problem Formulation

Let us consider a network model $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ on which dynamical agents are defined. In this chapter, we consider the following two cases with regard to the behavior of agents:

(i) Each agent $i \in \mathcal{V} \triangleq \{1, 2, \ldots, N\}$ has dynamics

$$\dot{x}_i(t) = u_i(t), \tag{4.4.1}$$

(ii) Each agent $i \in \mathcal{V} \triangleq \{1, 2, \ldots, N\}$ has dynamics

$$\ddot{x}_i(t) = u_i(t), \tag{4.4.2}$$

where $x_i(t) \in \mathbb{R}$ and $u_i(t) \in \mathbb{R}$ respectively denote the state and the control of agent i. We first consider the case (i) in this section, and we then discuss the extension to the case (ii) in Section 4.5.

We use a *distributed* control for $u_i(t)$, which depends only on the i's state $x_i(t)$ and the states $x_j(t)$ of the neighboring agents $j \in \mathcal{N}_i$. Moreover, we consider a *sampled-data* control system where the observation of the states $x_i(t)$ can be executed only at the sampling time kT, $k = 0, 1, 2, \ldots$, where $T > 0$ denotes the sampling period. Each agent determines the control $u_i(t)$, $t \in [kT, (k+1)T)$, based on $x_i(kT)$ and $\{x_j(kT) : j \in \mathcal{N}_i\}$. We also assume that the magnitude of the control is restricted to be bounded by 1, that is,

$$\|u_i\|_\infty \leq 1, \quad \forall i \in \mathcal{V}. \tag{4.4.3}$$

In addition, it is important to take into account of the control effort reduction in real systems. This chapter adopts the maximum hands-off control idea to the sampled-data consensus. Let us denote the control u_i on the sampling interval $[kT, (k+1)T)$ by $u_i[k]$, that is

$$u_i[k](t) \triangleq u_i(t + kT) \tag{4.4.4}$$

on $[0, T)$. Then we choose the control $u_i[k]$ at every discrete time k, that minimizes its L^0 norm $\|u_i[k]\|_0$ satisfying the above control objectives and constraints.

In summary, we formulate the *sampled-data maximum hands-off consensus control* problem as follows:

Problem 4.4.1
Given a sampling period $T > 0$, find a control scheme in which the control $\{u_i(t) : t \in [0, \infty)\}$ for each agent $i \in \mathcal{V}$ satisfies the following:

1. It achieves
$$\lim_{t \to \infty} \|x_i(t) - x_j(t)\| = 0$$
for any i, $j \in \mathcal{V}$.

2. It satisfies the magnitude constraint (4.4.3).

3. It is determined by sampled states $x_i(kT)$ and $\{x_j(kT) : j \in \mathcal{N}_i\}$, $k = 0, 1, 2, \ldots$.

4. It minimizes $\|u_i[k]\|_0$ at each $k = 0, 1, 2, \ldots$.

Remark 4.4.2
Since the first-order system is not steered at all by zero control, we can obtain a sparse distributed control equipped with arbitrary sparsity for the system once a consensus control scheme is found. However, this approach obviously requires a lot of time to reach consensus. Moreover, second-order agents move around according to their velocity while zero control is applied. Hence, control input should be properly switched off, otherwise agents may diverge from each other, which may require high L^0 cost of control to achieve consensus as a result. From these reasons, in this chapter we try to find a control scheme in which the state information is regularly exchanged among agents and sparse control is applied according to the state information. Note that since each agent does not know the consensus value in our setting, we consider a sparse distributed control that is L^0 optimal on each interval $[kT, (k+1)T]$ but is not necessarily equipped with the minimum L^0 cost among all controls that steer the agent from the initial state to the consensus value on the time axis $[0, \infty)$. The agents' target states on each interval $[kT, (k+1)T]$ are designed in the following subsection.

4.4.2 Control Protocol

Here we show a solution to Problem 4.4.1. First, we review the discrete-time consensus [99].

Lemma 4.4.3
Consider a directed network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of discrete-time dynamical agents represented by

$$z_i[k+1] = z_i[k] + w_i[k], \quad i \in \mathcal{V}, \quad k = 0, 1, 2, \ldots, \tag{4.4.5}$$

with distributed control

$$w_i[k] = -\epsilon \sum_{j \in \mathcal{N}_i} \big(z_i[k] - z_j[k]\big). \tag{4.4.6}$$

If $0 < \epsilon < 1/\Delta$ (Δ is the maximum degree of $\mathcal{G}$) and $\mathcal{G}$ has a spanning tree, then the multi-agent system asymptotically reaches consensus for all initial states with the consensus value

$$\alpha = \sum_{i \in \mathcal{V}} v_i z_i[0], \tag{4.4.7}$$

where v_i is the i-th element of the normalized left eigenvector associated with the eigenvalue 0 of the Laplacian matrix of $\mathcal{G}$.

From (4.4.5) and (4.4.6), we have

$$z_i[k+1] = z_i[k] - \epsilon \sum_{j \in \mathcal{N}_i} \big(z_i[k] - z_j[k]\big).$$

Based on this, we design a continuous-time control $u_i[k]$ in (4.4.4) that achieves the control objectives in Problem 4.4.1. Let $z_i[k]$ be the state $x_i(kT)$ of the i-th agent at sampling time kT, $k = 0, 1, 2, \ldots$. If we choose a control $u_i[k]$ that drives the state $x_i(t)$ from $z_i[k] = x_i(kT)$ to

$$\begin{aligned} x_i^f[k] &\triangleq z_i[k] - \epsilon \sum_{j \in \mathcal{N}_i} \big(z_i[k] - z_j[k]\big) \\ &= x_i(kT) - \epsilon \sum_{j \in \mathcal{N}_i} \big(x_i(kT) - x_j(kT)\big), \end{aligned} \tag{4.4.8}$$

then the multi-agent control system reaches a consensus at least on the sampling instants $t = 0, kT, 2kT, \dots$ from Lemma 4.4.3. Note that the control $u_i[k]$ depends only on the i-th state $x_i(kT)$ and the neighbor's states $x_j(kT)$, $j \in \mathcal{N}_i$. Moreover, we seek a control that has the minimum L^0 norm $\|u_i[k]\|_0$ (i.e. the maximum hands-off control) among such controls with the magnitude constraint (4.4.3). We describe the control algorithm in Algorithm 1.

Algorithm 1 Maximum hands-off distributed control for agent $i \in \mathcal{V}$

Given a sampling period T and an initial state $x_i(0)$
for $k = 0, 1, 2, \dots$ **do**
  Observe $x_i(kT)$ and $x_j(kT)$, $j \in \mathcal{N}_i$.
  Compute a maximum hands-off control

$$u_i[k] = \underset{u \in \mathcal{U}(x_i(kT), x_i^f[k], T)}{\arg\min} \|u\|_0$$

  where $x_i^f[k]$ is given in (4.4.8).
  Apply $u_i(t) = u_i[k](t - kT),\ t \in [kT, (k+1)T]$ to the agent i.
end for

4.4.3 Analysis

Here we give analysis of the proposed control. Our approach is as follows: Briefly speaking, the proposed control tracks the non-restricted control in (4.4.6) while minimizing the L^0 cost of control inputs. Then, we first need to analyze the existence of the maximum hands-off control on each sampling interval $[kT, (k+1)T]$. For this, a necessary and sufficient condition for the existence is given in Theorem 4.4.4. Note that under this condition, all agents converge to a common value at least on sampling instants by Lemma 4.4.3. We then investigate the consensus in the continuous-time domain in Theorem 4.4.7.

The following Theorem shows the existence of a maximum hands-off distributed control $u_i[k]$ defined in Algorithm 1.

Theorem 4.4.4
Define

$$x(t) \triangleq \begin{bmatrix} x_1(t) & x_2(t) & \dots & x_N(t) \end{bmatrix}^\top,$$

and let L be the Laplacian matrix of $\mathcal{G}$. Take $0 < \epsilon < 1/\Delta$. Then there exist a maximum hands-off distributed control $u_i[k]$ in Algorithm 1 at each step for each agent if and only if the initial state $x(0)$ satisfies

$$-T\mathbf{1}_N \le \epsilon L x(0) \le T\mathbf{1}_N. \tag{4.4.9}$$

Proof. It follows from Theorem 4.3.2 that a sufficient and necessary condition for agent $i \in \mathcal{V}$ to have a maximum hands-off distributed control $u_i[k]$ at step $k \in \{0, 1, 2, \dots\}$ is given by

$$x_i(kT) - x_i^f[k] \in \mathcal{R}, \tag{4.4.10}$$

where $\mathcal{R} = [-T, T]$, which is obtained by (4.3.2) with $F = 0$, $G = 1$, and $T_f = T$. We define

$$x^f[k] \triangleq \begin{bmatrix} x_1^f[k] & x_2^f[k] & \dots & x_N^f[k] \end{bmatrix}^\top.$$

Then, the condition that (4.4.10) holds for all k is equivalent to

$$-T\mathbf{1}_N \leq \epsilon Lx(kT) \leq T\mathbf{1}_N \tag{4.4.11}$$

for all k, because we have

$$x(kT) - x^f[k] = x(kT) - (x(kT) - \epsilon Lx(kT)) = \epsilon Lx(kT).$$

Hence we show that the condition (4.4.9) is equivalent for (4.4.11) to hold for all k. For this, it is enough to show that (4.4.9) implies that (4.4.11) holds for all k. In other words, we show that if (4.4.11) holds for $k = 0$, then (4.4.11) holds for all subsequent $k = 1, 2, \dots$. In order to prove this, we use mathematical induction.

Fix any $k_0 \in \{0, 1, 2, \dots\}$ and we assume that (4.4.11) holds for $k = k_0$. We show that (4.4.11) holds for $k = k_0 + 1$. The assumption guarantees the existence of $u_i[k_0]$ for all i. Since we consider the system without disturbances, $u_i[k_0]$ steers the state $x_i(t)$ from $x_i(k_0T)$ to $x_i^f[k_0]$ on the time interval $[k_0T, (k_0+1)T]$, and hence we have $x((k_0+1)T) = x^f[k_0] = Px(k_0T)$ by (4.4.8), where P is the Perron matrix according to the graph $\mathcal{G}$. Then

$$\epsilon Lx((k_0+1)T) = \epsilon LPx(k_0T) = \epsilon PLx(k_0T),$$

where we used the commutativity between P and L. Here, the matrix P is row stochastic, i.e., all of its row-sums are 1, and all of the components are non-negative from $0 < \epsilon < 1/\Delta$, and hence each component of $\epsilon PLx(k_0T)$ is a convex combination of the components of $\epsilon Lx(k_0T)$. Since each component of $\epsilon Lx(k_0T)$ is in the interval $[-T, T]$ from the assumption (4.4.11), each component of $\epsilon PLx(k_0T)$ is in the interval $[-T, T]$ from the convexity of $[-T, T]$. It follows that

$$-T\mathbf{1}_N \leq \epsilon Lx((k_0+1)T) \leq T\mathbf{1}_N.$$

Thus, (4.4.11) holds for $k = k_0 + 1$.

Hence if (4.4.11) holds for the initial step $k = 0$, then (4.4.11) holds for all subsequent steps. This gives the result. □

The following lemma gives characterization of the maximum hands-off distributed control.

Lemma 4.4.5

Assume that (4.4.9) holds. Then, the set of all maximum hands-off controls that steer the state $x_i(t)$ of agent $i \in \mathcal{V}$ from $x_i(kT)$ to the state $x_i^f[k]$ defined in (4.4.8) is given by

$$\begin{aligned} \mathcal{U}_i^*[k] \triangleq \Big\{ -v \operatorname{sgn}(\theta_i[k]) : \ & v \in L^1[kT, (k+1)T], \\ & v(t) \in \{0, 1\} \text{ for all } t \in [kT, (k+1)T], \\ & \int_{kT}^{(k+1)T} v(t)dt = |\theta_i[k]| \Big\}, \end{aligned} \tag{4.4.12}$$

where $\theta_i[k] \triangleq x_i(kT) - x_i^f[k]$.

Proof. Note that from Theorem 4.4.4 for any agent $i \in \mathcal{V}$ there exists a maximum hands-off control that steers the state $x_i(t)$ from $x_i(kT)$ to $x_i^f[k]$ for $k = 0, 1, 2, \ldots$. Since any maximum hands-off control is an L^1 optimal control that takes only values ± 1 and 0 from Theorem 4.3.2, we obtain the result by the discussion in [48, Sec. 8-2]. □

Remark 4.4.6
Lemma 4.4.5 shows that the maximum hands-off distributed control is piecewise constant. We can easily show that the number of switches is finite. In fact, the number of switches is at most 1 on each sampling interval $[kT, (k+1)T)$. On the other hand, to avoid discontinuity and undesired high-frequency oscillation as chattering, CLOT (Combined L-One and Two) optimization can be used to obtain continuous control signals as proposed in [100]. It is also worth to note that we have obtained a ternary controller belonging to $\{-1, 0, 1\}$ as a result of minimizing the L^0 norm on sampling intervals, even though the controller has a convex constraint $\|u_i\|_\infty \leq 1$. On the other hand, there exist works considering a coordination scheme by using a ternary controller in order to address quantization issues. For example, the authors of [101] propose a robust consensus algorithm with a ternary controller in a self-triggered manner. Note that the work leads to a zero control, but the maximization of time duration of zero control is not addressed, which is different from our work.

The maximum hands-off control thus takes its values on $\{-1, 0, 1\}$. This is a favorable property when the controller is implemented on a cheap computer on the controlled agent, since the control values are already quantized.

The lemma is thanks to the simple plant model in (4.4.1). If we consider more general plant, such as given in (4.3.1), one should rely on numerical computation to obtain the maximum hands-off control. After time discretization, the approximated optimization problem is described as an ℓ^0 optimization problem, which is known to be NP-hard [102]. That is, to numerically solve the maximum hands-off control, it may take a lot of time to compute. If the time constant of the agents is very short (e.g. an unmanned aerial vehicle), then the delay due to the computation will easily break the consensus and it may cause instability. To avoid this, one can use ℓ^1 relaxation for the ℓ^0 optimization and adopt a fast algorithm like the alternating direction method of multipliers (ADMM) [76]. The details of the algorithm are illustrated in Section 2.5.

Finally, we give a theorem to guarantee a consensus by the maximum hands-off distributed control in Algorithm 1 in the continuous-time domain.

Theorem 4.4.7
Assume that the digraph $\mathcal{G}$ has a spanning tree. Assume also that the gain ϵ is chosen to satisfy $0 < \epsilon < 1/\Delta$ and the initial state $x(0)$ satisfies (4.4.9). Under the maximum hands-off distributed control in Algorithm 1, all agents converge to the value α defined in (4.4.7) in the continuous-time domain.

Proof. By Lemma 4.4.3 and Theorem 4.4.4, the maximum hands-off distributed control achieves consensus on sampling instants. Fix arbitrarily $i \in \mathcal{V}$. Then, for any $\eta > 0$, there exists $k_i \in \mathbb{N}$ such that $k \geq k_i$ implies $|x_i(kT) - \alpha| < \eta$, where α is given in (4.4.7). From Lemma 4.4.5, the maximum hands-off control $u_i[k]$ also limits the state $x_i(t)$ on any time interval $[kT, (k+1)T]$ between $x_i(kT)$ and $x_i^f[k]$, that is,

$$x_i(t) \in \big[\min\{x_i(kT), x_i^f[k]\}, \max\{x_i(kT), x_i^f[k]\}\big]$$

for all $t \in [kT, (k+1)T]$. Hence for any $k \geq k_i$ we have $|x_i(t) - \alpha| < \eta$ for all $t \in [kT, (k+1)T]$.

Then, define

$$T^\star \triangleq \max_{i \in \mathcal{V}} k_i T.$$

It follows that we have $|x_i(t) - \alpha| < \eta$ for any $i \in \mathcal{V}$ and any $t \geq T^\star$. This means $\lim_{t\to\infty} |x_i(t) - x_j(t)| = 0$ for any $i, j \in \mathcal{V}$, and hence the maximum hands-off distributed control achieves consensus in the continuous-time domain. □

Remark 4.4.8
If the sampling time T is large, then from Theorem 4.4.4, we can take a large step size $\epsilon \in (0, 1/\Delta)$ that satisfies (4.4.9), and the convergence can be fast. Also, if T is large, the sparsity of the control can be easily promoted. On the other hand, if T is large, then the time duration becomes long on which the control is open loop, and hence the system may lose the robustness against unpredicted disturbances during this duration. Hence, for the robustness T should be small.

4.5 Extension to Second-Order Consensus Problem

In this section, we extend the consensus algorithm based on the maximum hands-off control to the multi-agent system consisting of agents with second-order dynamics, which is given by (4.4.2). Let us first define variables $x_{i1}(t)$ and $x_{i2}(t)$ by

$$x_{i1}(t) \triangleq x_i(t), \quad x_{i2}(t) \triangleq \dot{x}_{i1}(t) \tag{4.5.1}$$

for all $i \in \mathcal{V}$. Then the dynamics of agents (4.4.2) are represented by

$$\begin{bmatrix} \dot{x}_{i1}(t) \\ \dot{x}_{i2}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{i1}(t) \\ x_{i2}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_i(t). \tag{4.5.2}$$

Regarding $x_i(t)$ as the position of agent i, each agent is controlled by manipulating acceleration in the system.

The control objective is consensus in the sense that they reach a common position and stop (i.e. $x_i = \text{const.}$ and $\dot{x}_i = 0$) asymptotically with a distributed control satisfying the magnitude constraint (4.4.3) and reducing the control effort $\|u_i[k]\|_0$, where $u_i[k]$ is defined by (4.4.4). From the discussion in Section 4.4, the agents reach a consensus at least on the sampling instants, by steering the state $\begin{bmatrix} x_{i1}(t) & x_{i2}(t) \end{bmatrix}^\top$ from $\begin{bmatrix} x_{i1}(kT) & x_{i2}(kT) \end{bmatrix}^\top$ to $\begin{bmatrix} x_{i1}^f[k] & x_{i2}^f[k] \end{bmatrix}^\top$ with

$$x_{i1}^f[k] \triangleq x_{i1}(kT) - \epsilon \sum_{j \in \mathcal{N}_i} \big(x_{i1}(kT) - x_{j1}(kT) \big), \tag{4.5.3}$$

$$x_{i2}^f[k] \triangleq x_{i2}(kT) - \epsilon \sum_{j \in \mathcal{N}_i} \big(x_{i2}(kT) - x_{j2}(kT) \big), \tag{4.5.4}$$

on each time interval $[kT, (k+1)T]$, where $\epsilon > 0$. Then, we adopt the following algorithm (Algorithm 2).

In what follows, we first give a theorem regarding the feasibility of Algorithm 2 and then prove that it guarantees the consensus in the continuous-time domain.

Algorithm 2 Maximum hands-off distributed control for agent $i \in \mathcal{V}$ in the case of second-order problem

Given a sampling period T and an initial state $x_i[0] \triangleq \begin{bmatrix} x_i(0) & \dot{x}_i(0) \end{bmatrix}^\top$
for $k = 0, 1, 2, \ldots$ **do**
Observe $x_i[k] \triangleq \begin{bmatrix} x_i(kT) & \dot{x}_i(kT) \end{bmatrix}^\top$ and $x_j[k] \triangleq \begin{bmatrix} x_j(kT) & \dot{x}_j(kT) \end{bmatrix}^\top$, $j \in \mathcal{N}_i$.
Compute a maximum hands-off control

$$u_i[k] = \underset{u \in \mathcal{U}(x_i[k], x_i^f[k], T)}{\arg\min} \|u\|_0$$

where $x_i^f[k] \triangleq \begin{bmatrix} x_{i1}^f[k] & x_{i2}^f[k] \end{bmatrix}^\top$ is given in (4.5.3) and (4.5.4).
Apply $u_i(t) = u_i[k](t - kT),\ t \in [kT, (k+1)T]$ to the agent i.
end for

Theorem 4.5.1
Define

$$\begin{aligned} x(t) &\triangleq \begin{bmatrix} y_1^\top(t) & y_2^\top(t) \end{bmatrix}^\top, \\ y_1(t) &\triangleq \begin{bmatrix} x_{11}(t) & x_{21}(t) & \ldots & x_{N1}(t) \end{bmatrix}^\top, \\ y_2(t) &\triangleq \begin{bmatrix} x_{12}(t) & x_{22}(t) & \ldots & x_{N2}(t) \end{bmatrix}^\top, \end{aligned}$$

and

$$\mathcal{R}^2 \triangleq \left\{ \begin{bmatrix} r_1 & r_2 & \ldots & r_{2N} \end{bmatrix}^\top \in \mathbb{R}^{2N} : \begin{bmatrix} r_i & r_{i+N} \end{bmatrix}^\top \in \mathcal{R} \text{ for all } i = 1, 2, \ldots, N \right\},$$

where $\mathcal{R}$ is the reachable set defined in (4.3.2) with

$$T_f = T, \quad F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Let P be a Perron matrix according to the graph $\mathcal{G}$ with a parameter $\epsilon \in (0, 1/\Delta)$. Then the set $\mathcal{U}(x_i[k], x_i^f[k], T)$ in Algorithm 2 is not empty at each step k for all agents if and only if the initial state $x(0)$ satisfies

$$\begin{bmatrix} I - P & -TP \\ 0 & I - P \end{bmatrix} x(0) \in \mathcal{R}^2. \tag{4.5.5}$$

Proof. Fix any k. Algorithm 2 steers the state from $\begin{bmatrix} x_{i1}(kT) & x_{i2}(kT)) \end{bmatrix}^\top$ to $\begin{bmatrix} x_{i1}^f[k] & x_{i2}^f[k] \end{bmatrix}^\top$ on the time interval $[kT, (k+1)T]$ at step k for each agent. This is equivalent to the transition from $\begin{bmatrix} x_{i1}(kT) & x_{i2}(kT)) \end{bmatrix}^\top - \exp\left(\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} T \right) \begin{bmatrix} x_{i1}^f[k] & x_{i2}^f[k] \end{bmatrix}^\top$ to the origin on the time interval by the linearity of the system (4.5.2). Hence feasible controls at step k exist for all agents if and only if

$$\begin{bmatrix} x_{i1}(kT) \\ x_{i2}(kT) \end{bmatrix} - \begin{bmatrix} 1 & -T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{i1}^f[k] \\ x_{i2}^f[k] \end{bmatrix} \in \mathcal{R} \tag{4.5.6}$$

for all i. Here

$$\begin{bmatrix} x_{i1}^f[k] \\ x_{i2}^f[k] \end{bmatrix} = \begin{bmatrix} P_i & 0 \\ 0 & P_i \end{bmatrix} x(kT)$$

from the definitions (4.5.3) and (4.5.4), where P_i is the i-th row of the Perron matrix P. Hence, it follows from this with (4.5.6) that the set $\mathcal{U}(x_i[k], x_i^f[k], T)$ in Algorithm 2 is not empty at each step k for all agents if and only if

$$\begin{bmatrix} I-P & TP \\ 0 & I-P \end{bmatrix} x(kT) \in \mathcal{R}^2 \tag{4.5.7}$$

for all k.

Then it is enough to show that if the inclusion (4.5.7) holds for step $k = 0$, then (4.5.7) holds for all subsequent steps $k = 1, 2, \ldots$. The proof follows by mathematical induction. We assume that (4.5.7) holds for $k = 0$ and show (4.5.7) holds for $k = 1$. Since we consider the system in the absence of noise, we have $x_{i1}(T) = x_{i1}^f[0]$, and $x_{i2}(T) = x_{i2}^f[0]$ for all i. Hence

$$x(T) = \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} x(0).$$

Then

$$\begin{aligned} \begin{bmatrix} I-P & TP \\ 0 & I-P \end{bmatrix} x(T) &= \begin{bmatrix} I-P & TP \\ 0 & I-P \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} x(0) \\ &= \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} I-P & TP \\ 0 & I-P \end{bmatrix} x(0). \end{aligned}$$

Here we define

$$\begin{bmatrix} a_i \\ b_i \end{bmatrix} \triangleq \begin{bmatrix} e_i^\top \\ e_{i+N}^\top \end{bmatrix} \begin{bmatrix} I-P & TP \\ 0 & I-P \end{bmatrix} x(0)$$

for all i, where e_i is the i-th canonical basis in $\mathbb{R}^{2N}$. From the assumption we now have $\begin{bmatrix} a_i & b_i \end{bmatrix}^\top \in \mathcal{R}$ for all i. Then we have

$$\begin{bmatrix} e_i^\top \\ e_{i+N}^\top \end{bmatrix} \begin{bmatrix} I-P & TP \\ 0 & I-P \end{bmatrix} x(T) = \sum_{j=1}^{N} p_{ij} \begin{bmatrix} a_i \\ b_i \end{bmatrix}, \tag{4.5.8}$$

where p_{ij} is the j-th column of P_i. Here, the matrix P is row stochastic and all of the components are non-negative from $0 < \epsilon < 1/\Delta$, and hence the left hand side of (4.5.8) exists in $\mathcal{R}$ for each i by the convexity of the reachable set $\mathcal{R}$. This means that (4.5.7) holds for $k = 1$.

The consideration can also be used to show that the assumption for an arbitrarily fixed step k_0 implies the inclusion (4.5.7) for step $k_0 + 1$. In summary, the inclusion (4.5.7) for $k = 0$ is equivalent to the inclusion for all k. Thus the proof is completed. □

Remark 4.5.2

Theorem 4.5.1 shows the existence of the maximum hands-off distributed control $u_i[k]$ at each step for all agents from Theorem 4.3.2. In other words, this guarantees the feasibility of Algorithm 2. We also note that the reachable set $\mathcal{R}$ for the double integrator system (4.5.2) is analytically calculated as in [48]. Hence we can easily check the feasibility by

(4.5.5). In addition, a closed form of maximum hands-off control $u_i[k]$ can be obtained as discussed in [103]. According to the work, there exists a maximum hands-off control of which the number of switches is at most 2 on each interval $[kT, (k+1)T)$.

Finally, we prove a theorem to mention the consensus in the continuous-time domain. For the purpose, we introduce the following lemma.

Lemma 4.5.3
Consider the double integrator system (4.5.2). For a given final time of control $T > 0$, define a function

$$V(\xi) \triangleq \min_{u \in \mathcal{U}(\xi, 0, T)} \|u\|_0$$

on the reachable set $\mathcal{R}$. Then the function $V(\xi)$ is continuous on the reachable set $\mathcal{R}$ and $V(0) = 0$.

Proof. For the double-integral system (4.5.2), the maximum hands-off control is given in a closed form and the function $V(\xi)$ can be analytically calculated as in [103, Example 2], in which we can confirm the result. □

The following theorem is the main result of this section.

Theorem 4.5.4
Assume that the digraph $\mathcal{G}$ has a spanning tree, the gain ϵ is chosen to satisfy $0 < \epsilon < 1/\Delta$, the initial state $x(0)$ satisfies (4.5.5), and $x_{i2}(0) = 0$ for each agent i. Under the maximum hands-off distributed control in Algorithm 2, all agents satisfy

$$\lim_{t \to \infty} x_{i1}(t) = \alpha, \quad \lim_{t \to \infty} x_{i2}(t) = 0,$$

where α is defined in (4.4.7).

Proof. Define

$$\begin{aligned} \mathrm{Box}(a) &\triangleq \left\{ \begin{bmatrix} \xi_1 & \xi_2 \end{bmatrix}^\top \in \mathbb{R}^2 : |\xi_1| \le (T+2)a, |\xi_2| \le 2a \right\}, \\ M(a) &\triangleq \max_{\xi \in \mathrm{Box}(a) \cap \mathcal{R}} V(\xi) \end{aligned}$$

for $a > 0$. Note that $M(a)$ is well-defined because the set $\mathcal{R}$ is closed [104, Lemma 12.1] and $V(\xi)$ is continuous on the set $\mathcal{R}$ from Lemma 4.5.3. Take any $\eta > 0$. Since we have $\lim_{a \to 0+}(M(a) + a) = 0$ by Lemma 4.5.3, there exists $\eta_0 > 0$ such that

$$0 < M(a) + a < \eta, \quad \forall a \in (0, \eta_0). \tag{4.5.9}$$

Here, it follows from Theorem 4.5.1 and Lemma 4.4.3 that we have $\lim_{k \to \infty} x_{i1}(kT) = \alpha$ and $\lim_{k \to \infty} x_{i2}(kT) = 0$ for each agent i. Hence, for an arbitrarily fixed $\eta_1 \in (0, \min\{\eta_0, \eta\})$, there exist $p_i \in \mathbb{N}$ and $v_i \in \mathbb{N}$ for each i such that $k \ge p_i$ implies $|x_{i1}(kT) - \alpha| < \eta_1$ and $k \ge v_i$ implies $|x_{i2}(kT)| < \eta_1$. Define

$$K^* \triangleq \max\{p_1, \ldots, p_N, v_1, \ldots, v_N\}.$$

Then $k \ge K^*$ implies

$$|x_{i1}(kT) - \alpha| < \eta_1, \quad |x_{i2}(kT)| < \eta_1 \tag{4.5.10}$$

for all i.

Fix arbitrarily $i \in \mathcal{V}$. It follows from Theorem 4.5.1 that there exists the maximum hands-off distributed control $u_i[k]$ for any k, and $u_i[k]$ steers the state from

$$\xi_k \triangleq \begin{bmatrix} x_{i1}(kT) \\ x_{i2}(kT) \end{bmatrix} - \begin{bmatrix} 1 & -T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{i1}^f[k] \\ x_{i2}^f[k] \end{bmatrix}$$

to the origin in time T for the system (4.5.2). Hence $\xi_k \in \mathcal{R}$ for any step k. In particular, if $k \geq K^*$, then we have

$$|\xi_{k,1}| \leq |x_{i1}(kT) - \alpha| + |\alpha - x_{i1}^f[k]| + T|x_{i2}^f[k]| < (T+2)\eta_1$$

and

$$|\xi_{k,2}| \leq |x_{i2}(kT)| + |x_{i2}^f[k]| < 2\eta_1$$

from (4.5.10), where $\xi_k \triangleq [\xi_{k,1}, \xi_{k,2}]^\top$. It follows that $\xi_k \in \mathrm{Box}(\eta_1)$, and hence $\xi_k \in \mathrm{Box}(\eta_1) \cap \mathcal{R}$. Hence for $k \geq K^*$, we have

$$\|u_i[k]\|_0 \leq M(\eta_1) \tag{4.5.11}$$

by the definition of M.

Here if $k \geq K^*$, then we have

$$\begin{aligned} |x_{i2}(t) - x_{i2}(kT)| &\leq \int_{kT}^{t} |u_i(\tau)| d\tau \\ &\leq \int_{kT}^{(k+1)T} |u_i(\tau)| d\tau \\ &= \int_{\{\tau : \tau \in [kT,(k+1)T], u_i(\tau) \neq 0\}} |u_i(\tau)| d\tau \\ &\leq \|u_i[k]\|_0 \\ &\leq M(\eta_1) \end{aligned} \tag{4.5.12}$$

for all $t \in [kT, (k+1)T]$ from (4.5.2) and (4.5.11). We also have

$$\begin{aligned} |x_{i1}(t) - x_{i1}(kT)| &\leq \int_{kT}^{t} |x_{i2}(\tau)| d\tau \\ &\leq \int_{kT}^{(k+1)T} |x_{i2}(\tau)| d\tau \\ &\leq \int_{kT}^{(k+1)T} |x_{i2}(\tau) - x_{i2}(kT)| d\tau + \int_{kT}^{(k+1)T} |x_{i2}(kT)| d\tau \\ &\leq T(M(\eta_1) + \eta_1) \end{aligned} \tag{4.5.13}$$

for all $t \in [kT, (k+1)T]$ from (4.5.12) and (4.5.10). Since $0 < \eta_1 < \eta_0$,

$$0 < M(\eta_1) + \eta_1 < \eta \tag{4.5.14}$$

by (4.5.9). Hence we have

$$|x_{i1}(t) - x_{i1}(kT)| \leq T(M(\eta_1) + \eta_1) < T\eta \tag{4.5.15}$$

and

$$|x_{i2}(t) - x_{i2}(kT)| \leq M(\eta_1) < M(\eta_1) + \eta_1 < \eta \tag{4.5.16}$$

for all $t \in [kT, (k+1)T]$ with $k \geq K^*$ by (4.5.13), (4.5.14), and (4.5.12).

Therefore, we have

$$|x_{i1}(t) - \alpha| \leq |x_{i1}(t) - x_{i1}(kT)| + |x_{i1}(kT) - \alpha| < T\eta + \eta_1 < (T+1)\eta$$

and

$$|x_{i2}(t)| \leq |x_{i2}(t) - x_{i2}(kT)| + |x_{i2}(kT)| < \eta + \eta_1 < 2\eta$$

for all $t \geq kT$ with $k \geq K^*$ from (4.5.10), (4.5.15), and (4.5.16). This implies that for all $t \geq K^*T$ and any i, $j \in \mathcal{V}$,

$$|x_{i1}(t) - x_{j1}(t)| \leq |x_{i1}(t) - \alpha| + |x_{j1}(t) - \alpha| < 2(T+1)\eta$$

and

$$|x_{i2}(t) - x_{j2}(t)| \leq |x_{i2}(t)| + |x_{j2}(t)| < 4\eta,$$

and hence we have

$$\lim_{t\to\infty} |x_{i1}(t) - x_{j1}(t)| = 0, \quad \lim_{t\to\infty} |x_{i2}(t) - x_{j2}(t)| = 0.$$

Moreover,

$$\lim_{t\to\infty} x_{i1}(t) = \alpha, \quad \lim_{t\to\infty} x_{i2}(t) = 0$$

for all i. Thus we obtain the result. □

Remark 4.5.5
While we showed the consensus is achieved for first-order systems under the assumption that guarantees the feasibility of Algorithm 1 and the consensus on sampling instants in Theorem 4.4.7, Theorem 4.5.4 further assumes a condition on the initial velocity $x_{i2}(0)$. This additional assumption is imposed in order to guarantee all agents asymptotically stop and converge to a common position.

4.6 Example

In this section, we show simulation results of the maximum hands-off distributed control for the first-order systems and the second-order systems.

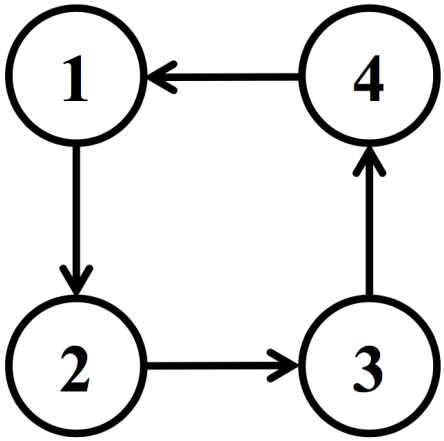

Figure 4.1: Digraph $\mathcal{G}$ associated with A

4.6.1 Example 1

We first consider the first-order systems given by

$$\dot{x}_i(t) = u_i(t),$$

where $i \in \mathcal{V} \triangleq \{1, 2, 3, 4\}$ and $x_i(t) \in \mathbb{R}$. We assume the initial states are $x_1(0) = 1$, $x_2(0) = 2$, $x_3(0) = 0$, $x_4(0) = -1$, and the adjacency matrix A is given by

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The digraph $\mathcal{G}$ associated with this adjacency matrix is shown in Fig. 4.1. For this network, we compute the maximum hands-off distributed control with the sampling period $T = 1$ [sec] and the gain $\epsilon = 0.5$.

Fig. 4.2 shows the obtained maximum hands-off distributed control $u_1(t)$ for the first agent. We can see that the control is sufficiently sparse (i.e. $\|u_1\|_0$ is sufficiently small compared to the whole time length 15 [sec]). In fact, the obtained controls $u_i(t)$, $i = 1, 2, 3, 4$ take 0 respectively for 83.44%, 83.41%, 83.44%, and 83.41% of 15 [sec]. Fig. 4.3 shows the state behavior of all agents corresponding to $u_i(t)$. We see all agents achieve consensus.

As mentioned in subsection 4.4.3, the proposed control tracks the control in (4.4.6). Then, we also compute the control in (4.4.6). Fig. 4.4 shows the control $w_1(t)$ for the first agent defined by

$$w_1(t) = -\frac{\epsilon}{T} \sum_{j \in \mathcal{N}_1} \big(x_1(kT) - x_j(kT)\big) \tag{4.6.1}$$

on $[kT, (k+1)T)$, $k = 0, 1, 2, \ldots$. The corresponding states $x_i(t)$ by $w_i(t)$, $i = 1, 2, 3, 4$, are shown in Fig. 4.5. We first note that the states in Fig. 4.3 and 4.5 have the same

value on sampling instants kT, which is obvious from our proposed algorithm. Note also that our proposed control takes only three discrete values of $\{0, \pm 1\}$, compared to the control (4.6.1).

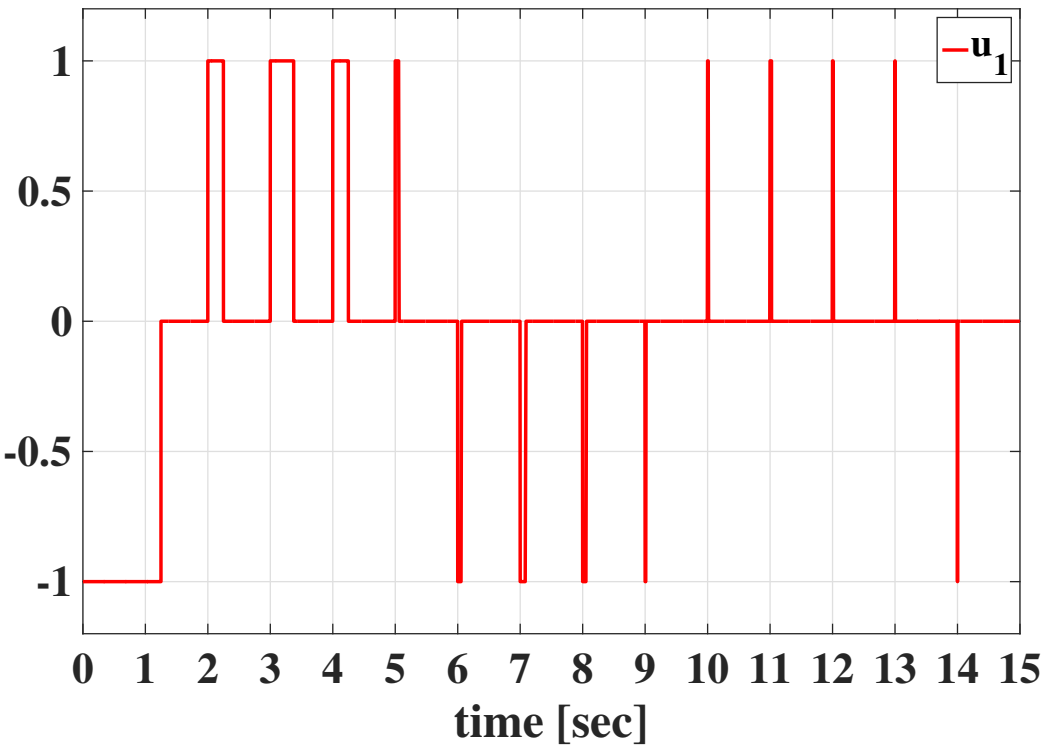

Figure 4.2: Control input $u_1(t)$

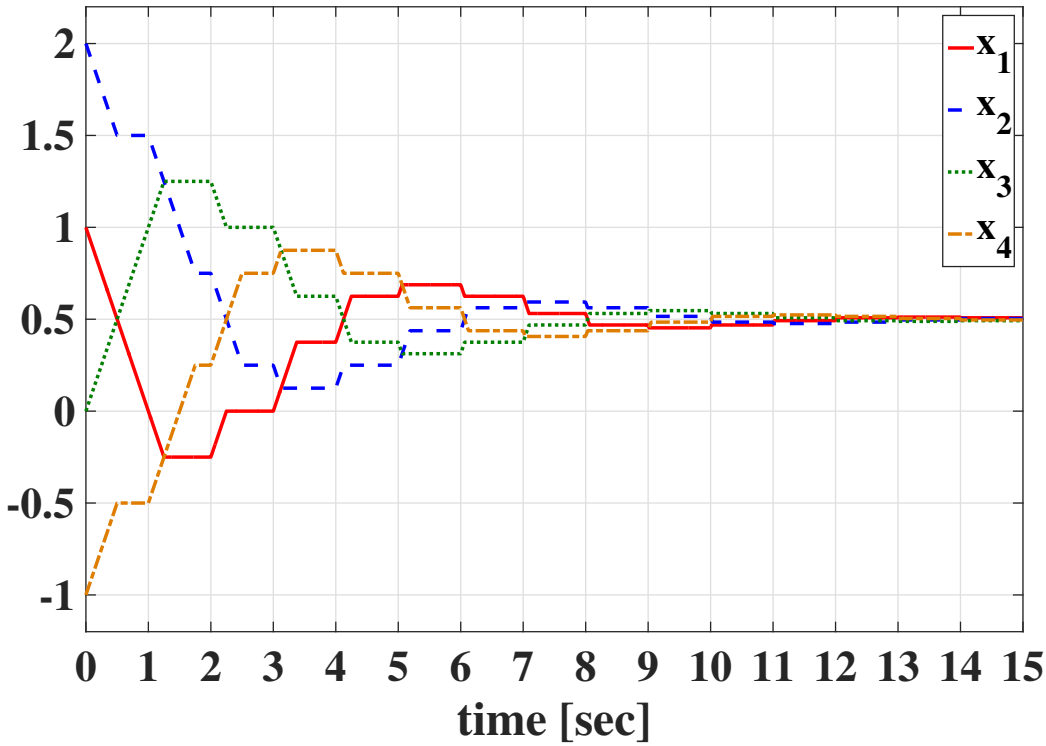

Figure 4.3: State trajectories $x_i(t)$ by $u_i(t)$, $i = 1, 2, 3, 4$

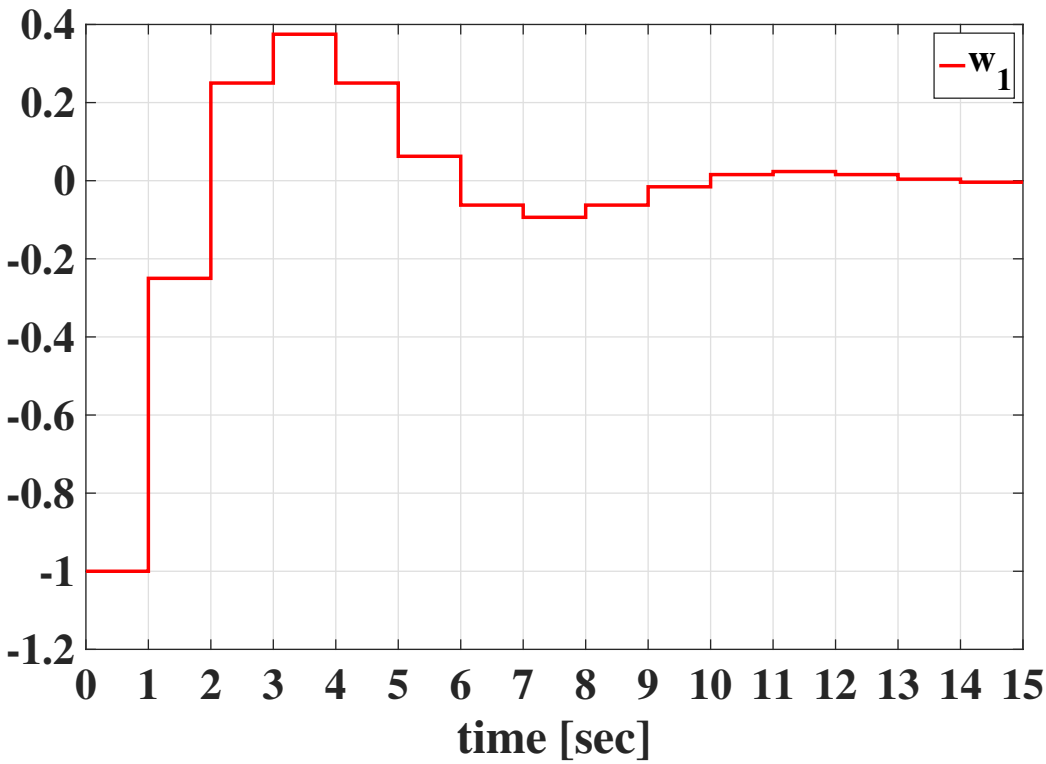

Figure 4.4: Control input $w_1(t)$ defined by (4.6.1)

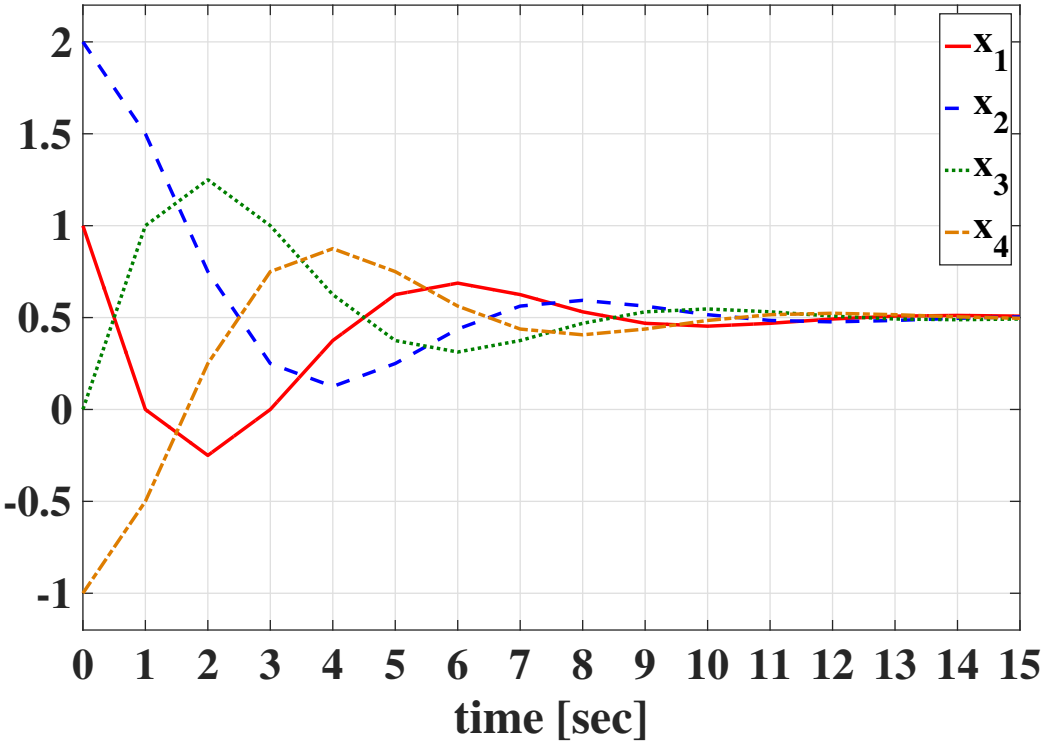

Figure 4.5: State trajectories $x_i(t)$ by $w_i(t)$, $i = 1, 2, 3, 4$

4.6.2 Example 2

We next consider the maximum hands-off distributed control of multi vehicle formations [99, Section 2.7]. We consider 4 vehicles connected through a wireless network as shown in Fig. 4.6. The dynamics of the i-th vehicle is given from Newton's law by

$$\ddot{p}_i(t) = u_i(t) + d_i(t), \tag{4.6.2}$$

where p_i is the i-th vehicle's position, u_i is the acceleration input, and $d_i(t)$ is the disturbance drawn from a normal distribution having a mean of 0 and standard deviation of σ. Define a vector $x_i = [p_i, \dot{p}_i]^\top = [x_{i1}, x_{i2}]^\top \in \mathbb{R}^2$. The initial state is given by $x_{11}(0) = 0$, $x_{21}(0) = 2$, $x_{31}(0) = 3$, $x_{41}(0) = -1$, and $x_{i2}(0) = 0$ for all i. The adjacency matrix A is defined by

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

from Fig. 4.6. According to Algorithm 2, the maximum hands-off distributed control $u_i(t)$ was computed on each sampling interval so that it steers the state from $[x_{i1}(kT), x_{i2}(kT)]^\top$ to $[x^f_{i1}[k], x^f_{i2}[k]]^\top$ for non-disturbed model and has the minimum L^0 cost among all controls satisfying the state transition. In this setting, we simulated the maximum hands-off distributed control with $T = 3$, $\epsilon = 1/6$, and $\sigma = 0.1$.

Fig. 4.7 shows the resulting distributed control $u_1(t)$ for the first agent. We can see that the control $u_1(t)$ is sufficiently sparse. In fact, the controls $u_i(t)$ for $i = 1, 2, 3, 4$ take 0 respectively for 99.65%, 98.57%, 97.74%, and 98.18% of the whole time length 100 [sec]. Each agent is subject to a force only on a small duration of the interval. Fig. 4.8 shows the state trajectories corresponding to the controls.

We can see that our hands-of control keeps each state variable $x_{i1}(t)$ and $x_{i2}(t)$ around 0. Note that, in this example, if the noise is absent (i.e., $d_i(t) = 0$), then each state $x_i(t)$ converges to $[0, 0]^\top$ by Theorem 4.5.4. Although this example includes noise, agents are certainly kept around the consensus value. We also note that the proposed scheme solves an optimization problem (i.e., maximum hands-off control problem) in which the time-horizon is finite and the control input is bounded in the L^∞ norm. Hence a drawback of our scheme is that the feasibility may not be guaranteed for large noise. We then tested the feasibility of the proposed scheme for the system (4.6.2) with noise having $\sigma = 1$. We simulated the proposed method over the time interval $[0, 100]$, 500 times. As a result, each control u_i on the interval was obtained 352 out of 500 times. Theoretical analysis on how to address noise would be included in future work.

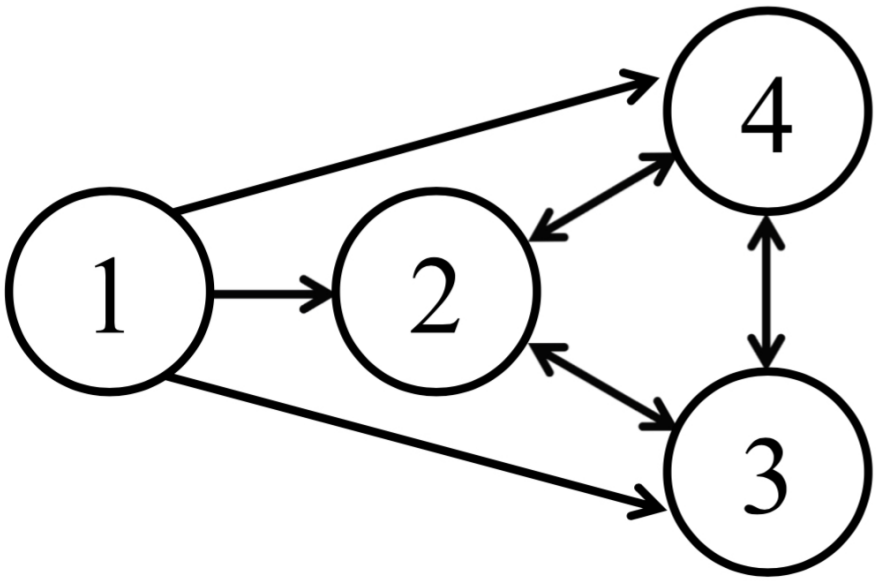

Figure 4.6: Vehicle network

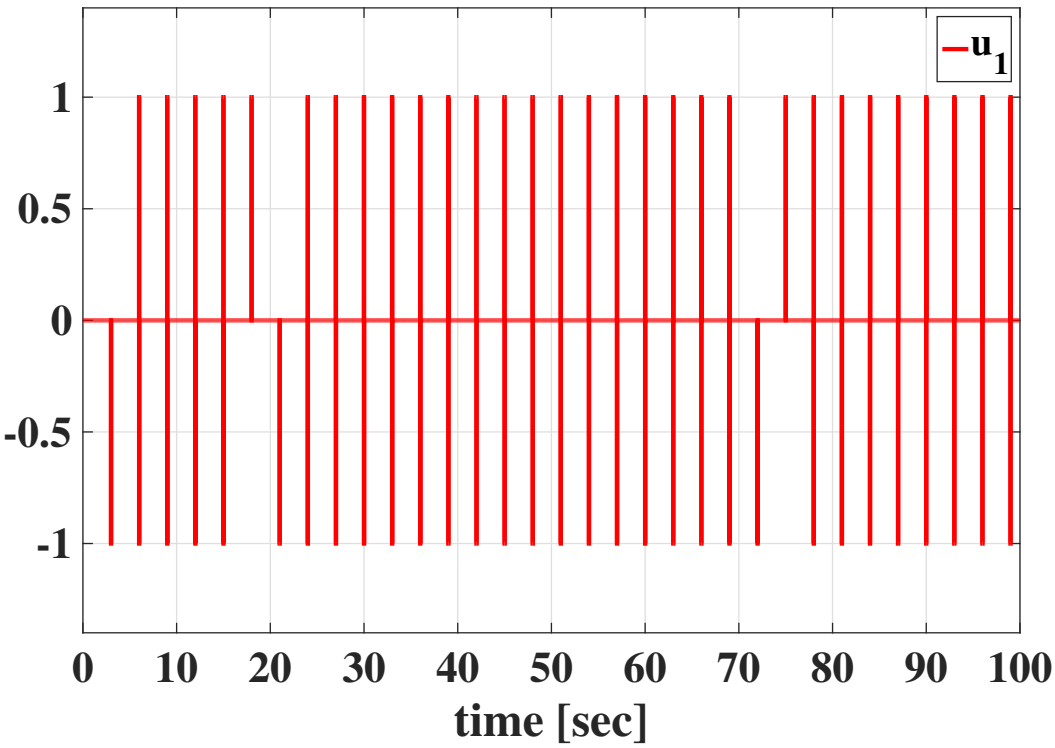

Figure 4.7: Control input $u_1(t)$

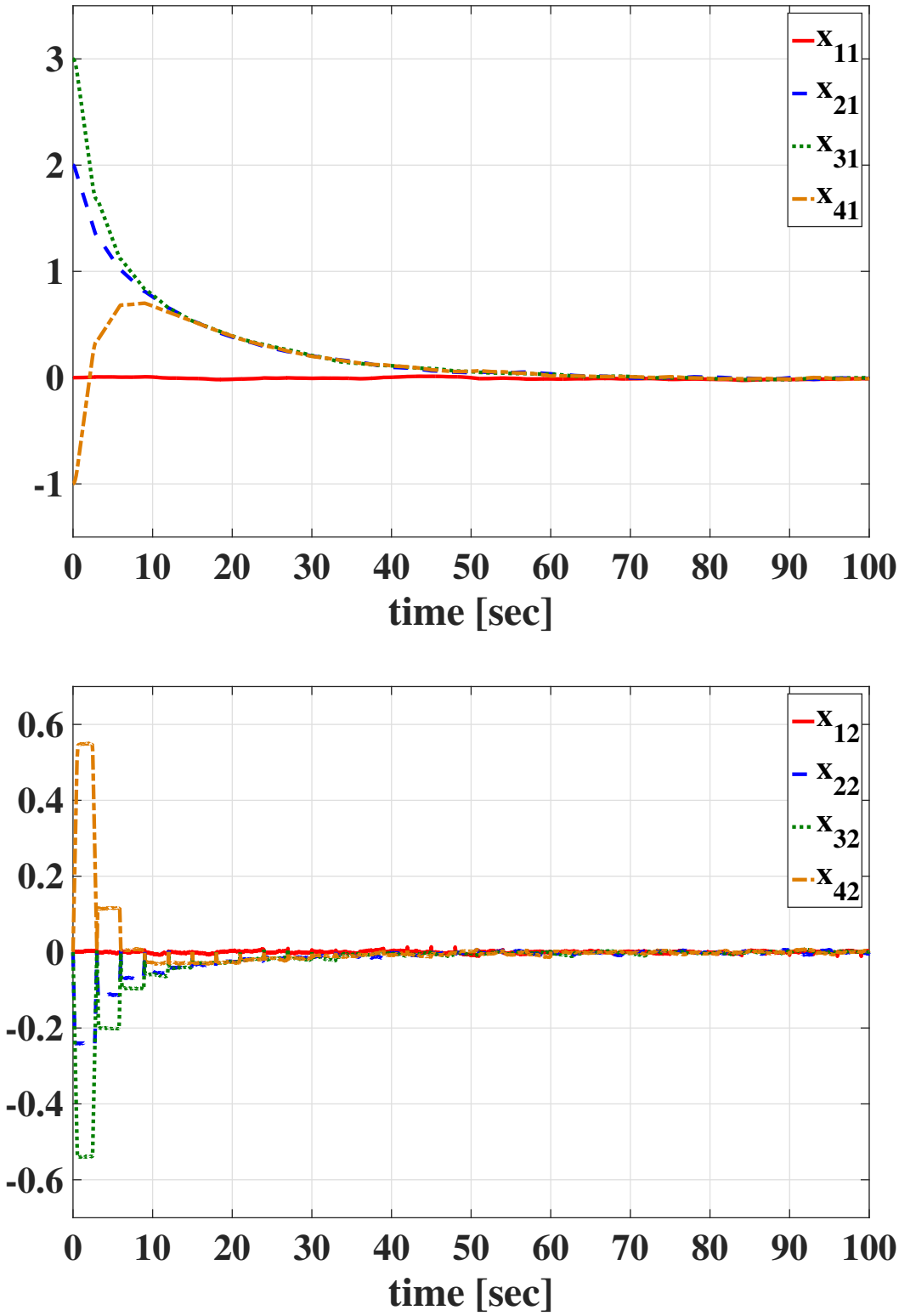

Figure 4.8: State trajectories corresponding to u_i; $x_{i1}(t)$ (top) and $x_{i2}(t)$ (bottom)

4.7 Conclusions

In this chapter, we have proposed a distributed control algorithm based on maximum hands-off control that leads to a sparse control satisfying a given magnitude constraint. We have given a feasibility condition for the distributed control, and a stability result to reach a consensus by the proposed control. Simulation results have been shown to illustrate the effectiveness of the proposed method. Although we have considered the first- or second-order integrator, extension to higher-order systems can be obtained by adopting efficient numerical optimization methods.

The proposed method needs only state observations, data transmissions, and input updates on sampling instants, which can reduce the burden in terms of the communication and controller actuation. Although we have assumed a periodic scheme in this chapter, introducing some conditions to input updates may need less network and computational resources, as seen in event-triggered control [52, 53]. Future work includes an extension to the consensus based on the event-triggered fashion with hands-off control. Also, future work includes analysis of consensus when there is noise in the state observation, and some agents are lost from the network because of their limits of fuel or electricity.

Chapter 5

Sparse Optimal Feedback Control for Continuous-Time Systems

5.1 Introduction

This chapter investigates an L^0 optimal control problem for non-linear systems. This optimal control involves the discontinuous and non-convex cost functional. Then, in order to deal with the difficulty of analysis, some relaxed problems with the L^p cost functional have been often investigated, as discussed so far. In [54], the L^1 cost functional is analyzed with an aim to show the relationship between the L^0 optimality and the L^1 optimality, and an equivalence theorem is derived. In [105], the result is extended to general linear systems including infinite-dimensional systems. The L^1 control cost is also considered in [44, 48, 106]. In [105], the sparsity properties of optimal controls for the L^p cost with $p \in (0, 1)$ is discussed. While the literature aforementioned have investigated open loop optimal controls, more recently, optimal controls in the feedback framework have been also considered in [57, 107]. The work [57] studies an infinite horizon optimal control problem with the L^p cost functional, where $p \in (0, 1]$, and derives the existence and the discreteness results for the time-discretized problem. In [107], a finite horizon optimal control problem with the L^1 cost functional is discussed for stochastic control-affine dynamical systems, and a sampling-based algorithm utilizing forward and backward stochastic differential equations is proposed.

In the previous studies, the exact L^0 optimal feedback control was not investigated. While there are studies that directly address the L^0 cost in the context of the open loop design, e.g. [58, 108], the counterparts in the context of the feedback control seem to be very few. Then, the purpose of this chapter is to directly deal with the underlying non-smooth and non-convex L^0 optimal control problem without the aid of any L^p relaxations. For this purpose, we adopt dynamic programming approach and investigate the sparse optimal control via the value function. Due to the non-smoothness of the L^0 cost functional, our value function is not differentiable, and hence it does not satisfy the associated HJB equation in the classical sense. Then, we first characterize the value function as a viscosity solution of the HJB equation. Based on the result, we show a sufficient and necessary condition for the L^0 optimality, which immediately gives an optimal feedback map. In addition, motivated by the discussion given in previous studies and previous chapters, we also consider the relationship between the L^0 optimality and the L^1 optimal-

ity and show an equivalence theorem by utilizing the uniqueness theorem of the viscosity solution. Moreover, the discreteness property of the sparse control is often reported in some frameworks. Then, we further show this property of our sparse optimal control for control-affine systems with a box constraint.

The remainder of this chapter is organized as follows: Section 5.2 gives mathematical preliminaries. Section 5.3 formulates our optimal control problem. Section 5.4 is devoted to the theoretical analysis. We first characterize the value function as a viscosity solution to the associated HJB equation, and next show a sufficient and necessary condition for the L^0 optimality. We also mention the relationship with the L^1 optimization problem and some basic properties of the sparse optimal control for control-affine systems with box constraints. In Section 5.5 we offer concluding remarks.

5.2 Mathematical Preliminaries

This section reviews notation that will be used throughout the chapter.

Let $N \in \mathbb{N}$. For a vector $a = [a_1, a_2, \ldots, a_N]^\top \in \mathbb{R}^N$, the open ball with center at a and radius $r > 0$ by $\mathbb{B}(a, r)$, i.e. $\mathbb{B}(a, r) \triangleq \{x \in \mathbb{R}^N : \|x - a\| < r\}$, and the closed ball with center at a and radius $r > 0$ by $\bar{\mathbb{B}}(a, r)$, i.e., $\bar{\mathbb{B}}(a, r) \triangleq \{x \in \mathbb{R}^N : \|x - a\| \leq r\}$. We denote the inner product of $a \in \mathbb{R}^N$ and $b \in \mathbb{R}^N$ by $a \cdot b$. For scalars $a \in \mathbb{R}$ and $b \in \mathbb{R}$, $\min\{a, b\}$ (resp. $\max\{a, b\}$) returns the smaller one (resp. larger one) of a and b.

Let $T > 0$ and $m \in \mathbb{N}$. The L^0 norm defined by (2.2.1) is also expressed as

$$\|u\|_0 = \int_0^T \psi_0(u(t))dt,$$

where $\psi_0 : \mathbb{R}^m \to \mathbb{R}$ is a function that returns the number of nonzero components, i.e., for $a = [a_1, a_2, \ldots, a_m]^\top \in \mathbb{R}^m$,

$$\psi_0(a) \triangleq \sum_{j=1}^{m} |a_j|^0$$

with $0^0 = 0$.

For a given set $\Omega \subset \mathbb{R}^N$, $C(\Omega)$ denotes the set of all continuous functions on Ω, and $C^1(\Omega)$ denotes the set of all differentiable functions with continuous partial derivatives in Ω.

For a locally Lipschitz continuous function $v : \Omega \to \mathbb{R}$ with an open set $\Omega \subset \mathbb{R}^N$, the *lower Dini derivative* and the *upper Dini derivative* of v at $x \in \Omega$ in the direction $q \in \mathbb{R}^N$ are defined by

$$\partial^- v(x; q) \triangleq \liminf_{h \to 0+} \frac{v(x + hq) - v(x)}{h},$$
$$\partial^+ v(x; q) \triangleq \limsup_{h \to 0+} \frac{v(x + hq) - v(x)}{h}.$$

If these two values coincide, then we denote the derivative by $\partial v(x; q)$. (Note that the symbol ∂ was used for the boundary in Chapter 2. However, the notation is also often used for the derivative and no confusion may arise since their arguments are quite different.) We also denote the gradient of v at x by $Dv(x)$.

Given a partial differential equation

$$F(x, v(x), Dv(x)) = 0, \quad x \in \Omega, \tag{5.2.1}$$

where Ω is an open set of $\mathbb{R}^N$ and $Dv(x)$ is the gradient of v at x, a function $v \in C(\Omega)$ is said to be a *viscosity subsolution* of (5.2.1) if, for any $\phi \in C^1(\Omega)$,

$$F(x_0, v(x_0), D\phi(x_0)) \leq 0$$

at any local maximum point $x_0 \in \Omega$ of $v - \phi$. Similarly, a function $v \in C(\Omega)$ is said to be a *viscosity supersolution* of (5.2.1) if, for any $\phi \in C^1(\Omega)$,

$$F(x_0, v(x_0), D\phi(x_0)) \geq 0$$

at any local minimum point $x_0 \in \Omega$ of $v - \phi$. Finally, v is said to be a *viscosity solution* of (5.2.1), if it is simultaneously a viscosity subsolution and supersolution.

5.3 Problem Formulation

We consider the following control system:

$$\begin{aligned} \dot{y}(t) &= f(y(t), u(t)), \quad t > 0, \\ y(0) &= x, \end{aligned} \tag{5.3.1}$$

where $y(t) \in \mathbb{R}^n$ is the state variable, $x \in \mathbb{R}^n$ is the initial state, and $u(t) \in \mathbb{R}^m$ is the control variable. For given initial state $x \in \mathbb{R}^n$ and final time of control $t \in (0, \infty)$, we consider the cost functional

$$J(x, t, u) \triangleq \int_0^t \psi_0(u(s))ds + g(y(t)),$$

where y and u satisfy the equation (5.3.1), g is the terminal cost, and the function $\psi_0 : \mathbb{R}^m \to \mathbb{R}$ returns the number of nonzero components, which is defined in the Section 5.2. Note that the first term expresses the L^0 cost of the control input, and hence the minimization of J enhances the sparsity. We assume the range of the control u is constrained in a compact set $\mathbb{U} \subset \mathbb{R}^m$ that contains $0 \in \mathbb{R}^m$, i.e., $u(t) \in \mathbb{U}$ for all t, and we denote the set of all such functions by $\mathcal{U}$, i.e.,

$$\mathcal{U} \triangleq \{u \in L^\infty : u(t) \in \mathbb{U} \text{ for all } t\}.$$

In other words, the main problem is formulated as follows:

Problem 5.3.1
Given $x \in \mathbb{R}^n$ and $t \geq 0$, find a control input u on $[0, t]$ that solves

$$\begin{aligned} \underset{u}{\text{minimize}} \quad & J(x, t, u) \\ \text{subject to} \quad & \dot{y}(s) = f(y(s), u(s)), \\ & y(0) = x, \\ & u \in \mathcal{U}. \end{aligned}$$

In this chapter, we assume the following conditions for the dynamic $f(y,u)$ and the terminal cost $g(y)$:

(A_1) $f : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}^n$ is continuous;

(A_2) f is Lipschitz continuous in the state variable, uniformly in the control variable, i.e., there exists a constant L such that

$$\|f(y,u) - f(z,u)\| \leq L\|y - z\| \tag{5.3.2}$$

for all $y, z \in \mathbb{R}^n$ and $u \in \mathbb{U}$;

(A_3) $g : \mathbb{R}^n \to \mathbb{R}$ is continuous.

Assumptions (A_1) and (A_2) guarantee the existence and the uniqueness of a solution to the differential equation (5.3.1). For given $x \in \mathbb{R}^n$ and $u \in \mathcal{U}$, we denote the state at time τ by $y(x;\tau;u)$, or briefly by $y(\tau)$ if no confusion may arise. As will be seen in Theorem 5.4.2, Assumption (A_3) is used to get the continuity of the *value function*, which is defined by

$$V(x,t) \triangleq \inf_{u \in \mathcal{U}} J(x,t,u), \quad x \in \mathbb{R}^n, \quad t \in [0,\infty).$$

5.4 Analysis

In this section, we derive the associated HJB equation and a sufficient and necessary condition for the L^0 optimality, which gives an optimal feedback map for Problem 5.3.1. Our analysis is based on the dynamic programming approach, and hence we first investigate the value function V.

5.4.1 Characterization of the Value Function

We here characterize the value function as a viscosity solution to HJB equation. For the purpose, we first show the continuity of the value function, which is a fundamental property since we investigate the solutions to the HJB equation in the class $C(\mathbb{R}^n \times [0,T])$ for some $T > 0$. Here, the basic estimates of the solution y to (5.3.1) are the following [109]:

Lemma 5.4.1
Fix any $x, z \in \mathbb{R}^n$. If (A_1) and (A_2) are satisfied, then

$$\|y(x;t;u)\| \leq (\|x\| + \sqrt{2Kt})e^{Kt} \quad \forall u \in \mathcal{U}, \forall t > 0, \tag{5.4.1}$$

$$\|y(x;t;u) - y(z;t;u)\| \leq e^{Lt}\|x - z\| \quad \forall u \in \mathcal{U}, \forall t > 0, \tag{5.4.2}$$

$$\|y(x;t;u) - x\| \leq F_x t \quad \forall u \in \mathcal{U}, \forall t \in [0, 1/F_x], \tag{5.4.3}$$

where $K \triangleq L + \sup\{f(0,u) : u \in \mathbb{U}\}$, L is a constant that satisfies (5.3.2), and $F_x \triangleq \sup\{\|f(w,u)\| : \|x - w\| \leq 1, u \in \mathbb{U}\}$.

Theorem 5.4.2
Fix $T > 0$. Under assumptions (A_1), (A_2), and (A_3), the value function V is continuous on $\mathbb{R}^n \times [0,T]$. If in addition the terminal cost g is Lipschitz continuous, then V is locally Lipschitz continuous.

Proof. Fix any $(x,t) \in \mathbb{R}^n \times [0,T]$ and $\varepsilon > 0$. Then, the set $\tilde{D}_{x,t} \triangleq \{y(x;t;u) : u \in \mathcal{U}\} \subset \mathbb{R}^n$ is bounded from (5.4.1) in Lemma 5.4.1. Hence, we can take a compact set $D_{x,t} \subset \mathbb{R}^n$ that contains $\tilde{D}_{x,t}$ in its interior. Since g is continuous on $\mathbb{R}^n$, g is uniformly continuous on $D_{x,t}$. It follows that there exists $\delta > 0$ depending on x, t, ε such that, for any $c, d \in D_{x,t}$ with $\|c - d\| < \delta$, $|g(c) - g(d)| < \varepsilon$ holds. In addition, from (5.4.2), there exists $r > 0$ depending on x, t, ε such that, for any $(z,\tau) \in \mathbb{B}((x,t),r)$, $\|y(x;t;u) - y(z;\tau;u)\| < \delta$ and $y(z;\tau;u) \in D_{x,t}$ hold for all $u \in \mathcal{U}$. This can be observed from

$$\begin{aligned}\|y(x;t;u) - y(z;\tau;u)\| &\le \|y(x;t;u) - y(x;\tau;u)\| + \|y(x;\tau;u) - y(z;\tau;u)\| \\ &\le \left\| \int_\tau^t f(y(x;s;u),u(s))ds \right\| + e^{L\tau}\|x - z\| \\ &\le C_x|t-\tau| + e^{LT}\|x - z\|\end{aligned}$$

for all $u \in \mathcal{U}$, where

$$C_x \triangleq \max\{\|f(y,u)\| : (y,u) \in \bar{\mathbb{B}}(0,\bar{x}) \times \mathbb{U}\} \tag{5.4.4}$$

with $\bar{x} \triangleq (\|x\| + \sqrt{2KT})e^{KT}$ and $C_x < \infty$ from the assumption (A_1). Thus, for $(z,\tau) \in \mathbb{B}((x,t),r)$,

$$|g(y(x;t;u)) - g(y(z;\tau;u))| < \varepsilon$$

for all $u \in \mathcal{U}$.

Here, fix any $(z,\tau) \in \mathbb{B}((x,t),\min\{r,\varepsilon\})$. Then, by definition of V, there exists a control $\bar{u} \in \mathcal{U}$ such that

$$V(z,\tau) + \varepsilon \ge J(z,\tau,\bar{u}). \tag{5.4.5}$$

Then,

$$\begin{aligned}V(x,t) - V(z,\tau) &\le J(x,t,\bar{u}) - J(z,\tau,\bar{u}) + \varepsilon \\ &= \int_{\min\{t,\tau\}}^{\max\{t,\tau\}} \psi_0(\bar{u}(s))ds + g(y(x;t;\bar{u})) - g(y(z;\tau;\bar{u})) + \varepsilon \\ &< m|t-\tau| + \varepsilon + \varepsilon < (m+2)\varepsilon,\end{aligned}$$

where we used the boundedness of ψ_0 in the third estimation, i.e., $|\psi_0(a)| \le m$ for all $a \in \mathbb{R}^m$. Similarly, we can also show $V(z,\tau) - V(x,t) < (m+2)\varepsilon$ by taking a control $\tilde{u} \in \mathcal{U}$ such that $V(x,t) + \varepsilon \ge J(x,t,\tilde{u})$. This shows the continuity of V.

If in addition g is Lipschitz continuous, then the result follows from the local boundedness of f in the state variable. More precisely, take any bounded neighborhood $D_{x,t}$ that contains (x,t). Define $C_x^* \triangleq \sup\{C_w : (w,s) \in D_{x,t} \text{ for some } s\}$, where C_x is defined in (5.4.4) and $C_x^* < \infty$. Take any $(w,s) \in D_{x,t}$ and $(z,\tau) \in D_{x,t}$. Fix any $\varepsilon > 0$ and take $\bar{u}$ that satisfies (5.4.5) for (z,τ). Let G be the Lipschitz constant of g. Then, from the estimation above, we have

$$\begin{aligned}V(w,s) - V(z,\tau) &\le m|s-\tau| + |g(y(w;s;\bar{u})) - g(y(z;\tau;\bar{u}))| + \varepsilon \\ &\le m|s-\tau| + G\|y(w;s;\bar{u}) - y(z;\tau;\bar{u})\| + \varepsilon \\ &\le (m + GC_x^*)|s-\tau| + Ge^{LT}\|w - z\| + \varepsilon.\end{aligned}$$

The arbitrarity of ε and the similar discussion above complete the proof. □

For the value function, we have another fundamental lemma.

Lemma 5.4.3
Assume (A_1), (A_2), and (A_3). Fix any $t \in (0, \infty)$. Then

$$V(x,t) = \inf_{u \in \mathcal{U}} \left\{ \int_0^\tau \psi_0(u(s))ds + V(y(x;\tau;u), t-\tau) \right\}$$

for all $x \in \mathbb{R}^n$ and $\tau \in (0, t]$.

Proof. The result is obvious for $\tau = t$, since we have $V(x, 0) = g(x)$ for all $x \in \mathbb{R}^n$. Then, we take any $\tau \in (0, t)$. Fix any $u \in \mathcal{U}$, and denote $\bar{u}(s) \triangleq u(s+\tau)$, then we have

$$\begin{aligned} J(x,t,u) &= \int_0^\tau \psi_0(u(s))ds + \int_0^{t-\tau} \psi_0(\bar{u}(s))ds + g(y(t)) \\ &= \int_0^\tau \psi_0(u(s))ds + J(y(x;\tau;u), t-\tau, \bar{u}) \\ &\geq \int_0^\tau \psi_0(u(s))ds + V(y(x;\tau;u), t-\tau). \end{aligned}$$

Taking the infimum over $u \in \mathcal{U}$, we get the inequality

$$V(x,t) \geq \inf_{u \in \mathcal{U}} \left\{ \int_0^\tau \psi_0(u(s))ds + V(y(x;\tau;u), t-\tau) \right\}.$$

We next show the inverse inequality. Fix any $u \in \mathcal{U}$ and $\varepsilon > 0$. Denote $z \triangleq y(x;\tau;u)$. Then, there exists $\tilde{u} \in \mathcal{U}$ such that

$$V(z, t-\tau) + \varepsilon \geq J(z, t-\tau, \tilde{u}).$$

Put $\hat{u}(s) \triangleq u(s)$ for $0 \leq s \leq \tau$ and $\hat{u}(s) \triangleq \tilde{u}(s-\tau)$ for $\tau < s \leq t$. Then,

$$\begin{aligned} V(x,t) &\leq J(x,t,\hat{u}) \\ &= \int_0^\tau \psi_0(u(s))ds + \int_0^{t-\tau} \psi_0(\tilde{u}(s))ds + g(y(x;t;\hat{u})). \end{aligned}$$

Here, $y(x;t;\hat{u}) = y(z;t-\tau;\tilde{u})$, and hence

$$\begin{aligned} V(x,t) &\leq \int_0^\tau \psi_0(u(s))ds + J(z, t-\tau, \tilde{u}) \\ &\leq \int_0^\tau \psi_0(u(s))ds + V(z, t-\tau) + \varepsilon. \end{aligned}$$

The arbitrarity of ε and $u \in \mathcal{U}$ show the desired inequality. □

The next theorem is one of the main results of this chapter.

Theorem 5.4.4
Assume (A_1), (A_2), and (A_3). Fix $T > 0$. Then, the value function V is a viscosity solution of the following Hamilton-Jacobi-Bellman equation with an initial condition:

$$\begin{cases} v_t(x,t) + H(x, D_x v(x,t)) = 0 \quad \text{in } \mathbb{R}^n \times (0,T), & (5.4.6) \\ v(x,0) = g(x) \quad \text{in } \mathbb{R}^n, & (5.4.7) \end{cases}$$

where $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is defined by

$$H(x,p) \triangleq \sup_{u \in \mathbb{U}} \{-f(x,u) \cdot p - \psi_0(u)\},$$

v_t denotes the partial derivative with respect to the last variable and $D_x v$ denotes the gradient with respect to the first n variables.

Proof. We first show that V is a viscosity subsolution of the equation (5.4.6). Fix any $\phi \in C^1(\mathbb{R}^n \times (0,T))$, and let (x,t) be a local maximum point of $V - \phi$. Then, there exists $r > 0$ such that, for all $z \in \mathbb{R}^n$ and $s \in (0,T)$ with $\|x - z\| < r$ and $|t - s| < r$,

$$\phi(x,t) - \phi(z,s) \leq V(x,t) - V(z,s). \tag{5.4.8}$$

Take any $u \in \mathbb{U}$ and define $\bar{u}(\tau) = u$ for all τ. Then, by continuity of y, for sufficiently small τ, we have $\|y(x;\tau;\bar{u}) - x\| < r$ and $\tau < r$. It follows from (5.4.8) that

$$\phi(x,t) - \phi(y(x;\tau;\bar{u}), t-\tau) \leq V(x,t) - V(y(x;\tau;\bar{u}), t-\tau).$$

Here, we have

$$\begin{aligned} V(x,t) &\leq \int_0^\tau \psi_0(\bar{u}(s))ds + V(y(x;\tau;\bar{u}), t-\tau) \\ &= \tau\psi_0(u) + V(y(x;\tau;\bar{u}), t-\tau) \end{aligned}$$

from Lemma 5.4.3. Hence,

$$\phi(x,t) - \phi(y(x;\tau;\bar{u}), t-\tau) \leq \tau\psi_0(u).$$

Devide τ and let $\tau \to 0$, then

$$-D_x\phi(x,t) \cdot f(x,u) + \phi_t(x,t) \leq \psi_0(u).$$

This inequality holds for all $u \in \mathbb{U}$. This means

$$\phi_t(x,t) + \sup_{u \in \mathbb{U}} \{-D_x\phi(x,t) \cdot f(x,u) - \psi_0(u)\} \leq 0.$$

We next show that V is the viscosity supersolution of (5.4.6). Fix any $\phi \in C^1(\mathbb{R}^n \times (0,T))$, and let (x,t) be a local minimum point of $V - \phi$. Then, there exists $r > 0$ such that, for all $z \in \mathbb{R}^n$ and $s \in (0,T)$ with $\|x - z\| < r$ and $|t - s| < r$,

$$V(x,t) - V(z,s) \leq \phi(x,t) - \phi(z,s). \tag{5.4.9}$$

Here, fix any $\varepsilon > 0$ and $\tau \in (0, t]$. Then, from Lemma 5.4.3, there exists $\tilde{u} \in \mathcal{U}$, which depends on ε and τ, such that

$$V(x,t) + \tau\varepsilon \geq \int_0^\tau \psi_0(\tilde{u}(s))ds + V(y(x;\tau;\tilde{u}), t-\tau).$$

Denote simply $y(x;s;\tilde{u})$ by $\tilde{y}(s)$ for all s. For sufficiently small $\tau > 0$, we have $\|\tilde{y}(\tau) - x\| < r$ and $\tau < r$ by (5.4.3). It follows from (5.4.9) that

$$\begin{aligned} \phi(x,t) - \phi(\tilde{y}(\tau), t-\tau) &\geq V(x,t) - V(\tilde{y}(\tau), t-\tau) \\ &\geq \int_0^\tau \psi_0(\tilde{u}(s))ds - \tau\varepsilon. \end{aligned} \tag{5.4.10}$$

Here,

$$\begin{aligned} \phi(x,t) - \phi(\tilde{y}(\tau), t-\tau) &= -\int_0^\tau \frac{d}{ds}\phi(\tilde{y}(s), t-s)ds \\ &= -\int_0^\tau D_x\phi(\tilde{y}(s), t-s)\cdot f(\tilde{y}(s), \tilde{u}(s))ds \\ &\quad + \int_0^\tau \phi_t(\tilde{y}(s), t-s)ds. \end{aligned} \tag{5.4.11}$$

Note that

$$\int_0^\tau D_x\phi(\tilde{y}(s), t-s)\cdot f(\tilde{y}(s), \tilde{u}(s))ds = \int_0^\tau D_x\phi(x,t)\cdot f(x, \tilde{u}(s))ds + o(\tau),$$

where $o(\tau)$ denotes a function $h(\tau)$ such that $\lim_{\tau\to 0+} h(\tau)/\tau = 0$, and we used $\phi \in C^1$, (5.4.1), (A_1), and (5.4.3) with a calculation

$$\begin{aligned} &\int_0^\tau \big\{D_x\phi(\tilde{y}(s), t-s)\cdot f(\tilde{y}(s), \tilde{u}(s)) - D_x\phi(x,t)\cdot f(x, \tilde{u}(s))\big\}ds \\ &\qquad = \int_0^\tau \big\{D_x\phi(\tilde{y}(s), t-s) - D_x\phi(x,t)\big\}\cdot f(\tilde{y}(s), \tilde{u}(s))ds \\ &\qquad\quad + \int_0^\tau D_x\phi(x,t)\cdot\big\{f(\tilde{y}(s), \tilde{u}(s)) - f(x, \tilde{u}(s))\big\}ds. \end{aligned}$$

Similarly,

$$\int_0^\tau \phi_t(\tilde{y}(s), t-s)ds = \phi_t(x,t) + o(\tau).$$

It follows from (5.4.10) and (5.4.11) that

$$\begin{aligned} -\tau\varepsilon &\leq \int_0^\tau \big\{-D_x\phi(x,t)\cdot f(x, \tilde{u}(s)) - \psi_0(\tilde{u}(s))\big\}ds + \phi_t(x,t) + o(\tau) \\ &\leq \tau \sup_{u\in\mathbb{U}}\big\{-D_x\phi(x,t)\cdot f(x,u) - \psi_0(u)\big\} + \phi_t(x,t) + o(\tau). \end{aligned}$$

Devide $\tau > 0$ and let $\tau \to 0$, then

$$-\varepsilon \leq \sup_{u\in\mathbb{U}}\{-D_x\phi(x,t)\cdot f(x,u) - \psi_0(u)\} + \phi_t(x,t).$$

The arbitrarity of ε shows that V is the viscosity supersolution of (5.4.6). Thus, V is a viscosity solution of (5.4.6). Moreover, V satisfies $V(x,0) = g(x)$ in $\mathbb{R}^n$ by definition, and hence V is a viscosity solution of (5.4.6) with an initial condition (5.4.7). $\square$

We next mention the connection to the L^1 optimization problem, which is based on the uniqueness theorem of the viscosity solution.

Theorem 5.4.5
Assume (A_1), (A_2), (A_3), and fix $T > 0$. If the dynamics f is affine in u, i.e.,

$$f(y,u) = f_0(y) + \sum_{j=1}^{m} f_j(y)u_j \tag{5.4.12}$$

for some $f_j : \mathbb{R}^n \to \mathbb{R}^n$, $j = 0, 1, 2, \ldots, m$, and $\mathbb{U} = \{u \in \mathbb{R}^m : |u_j| \leq 1, \forall j\}$, then the value function V coincides with the value function of the corresponding L^1 *optimal control problem*. More precisely, the L^1 optimal control problem is defined by the optimization problem in which the cost functional of Problem 5.3.1 is replaced with

$$J_1(x,t,u) \triangleq \sum_{j=1}^{m} \int_0^t |u_j(s)|ds + g(y(t)),$$

and the value function V_1 of this problem is defined by

$$V_1(x,t) \triangleq \inf_{u \in \mathcal{U}} J_1(x,t,u).$$

Then $V = V_1$ holds on $\mathbb{R}^n \times [0,T)$.

Proof. In this case, for any $x, p \in \mathbb{R}^n$,

$$\begin{aligned} H(x,p) &= \sup_{u \in \mathbb{U}} \left\{ -\sum_{j=1}^{m} f_j(x)u_j \cdot p - \sum_{j=1}^{m} |u_j|^0 \right\} - f_0(x) \cdot p \\ &= \sum_{j=1}^{m} \sup_{u_j \in \mathbb{U}_j} \left\{ -\big(f_j(x) \cdot p\big)u_j - |u_j|^0 \right\} - f_0(x) \cdot p, \end{aligned}$$

where $\mathbb{U}_j$ is the set of all j-th components of $\mathbb{U}$, i.e., $\mathbb{U}_j \triangleq \{a \in \mathbb{R} : |a| \leq 1\}$. Here, it follows from an elementary calculation that

$$\sup_{u_j \in \mathbb{U}_j} \left\{ -a^j_{x,p}u_j - |u_j|^0 \right\} = \sup_{u_j \in \mathbb{U}_j} \left\{ -a^j_{x,p}u_j - |u_j| \right\}$$

for all $x, p \in \mathbb{R}^n$ and $j = 1, 2, \ldots, m$, where $a^j_{x,p} \triangleq f_j(x) \cdot p$. Indeed, the supremum of both sides is given by

$$\begin{cases} a^j_{x,p} - 1, & \text{if} \quad a^j_{x,p} > 1, \\ 0, & \text{if} \quad a^j_{x,p} = 1, \\ 0, & \text{if} \quad -1 < a^j_{x,p} < 1, \\ 0, & \text{if} \quad a^j_{x,p} = -1, \\ -a^j_{x,p} - 1, & \text{if} \quad a^j_{x,p} < -1. \end{cases}$$

Hence, the equation (5.4.6) is equivalent to

$$v_t(x,t) + H_1(x, D_x v(x,t)) = 0, \tag{5.4.13}$$

where

$$H_1(x,p) \triangleq \sup_{u\in\mathbb{U}}\{-f(x,u)\cdot p - \psi_1(u)\}, \quad \forall x, \forall p \in \mathbb{R}^n,$$
$$\psi_1(a) \triangleq \sum_{j=1}^{m} |a_j|, \quad \forall a \in \mathbb{R}^m.$$

Note that the equation (5.4.13) is the HJB equation for the L^1 optimal control problem, and it is known that the value function V_1 is the unique viscosity solution of the equation with initial condition (5.4.7) in the class $C(\mathbb{R}^n \times (0,T))$ [109]. This means $V = V_1$. □

5.4.2 Optimality of a Control

We here derive a sufficient and necessary condition for the L^0 optimality, which is the second main result.

Theorem 5.4.6
Assume (A_1), (A_2), and g is Lipschitz continuous. Fix any initial state $x \in \mathbb{R}^n$ and final time of control $T > 0$. Then, a control u is an optimal solution to Problem 5.3.1 if and only if

$$\partial V(y(t), T-t; f(y(t),u(t)), -1) + \psi_0(u(t)) = 0 \tag{5.4.14}$$

almost everywhere on $(0,T)$.

Proof. We first show the sufficiency. We assume (5.4.14) and take any u that satisfies (5.4.14). Put

$$h(t) \triangleq \int_0^t \psi_0(u(s))ds + V(y(t), T-t) \tag{5.4.15}$$

for $t \in (0,T)$. Since ψ_0 is integrable and the value function V is locally Lipschitz continuous by Theorem 5.4.2, $h(t)$ is differentiable almost everywhere, and

$$\begin{aligned}\frac{d}{dt}h(t) &= \psi_0(u(t)) + \frac{d}{dt}V(y(t), T-t)\\ &= \psi_0(u(t)) + \partial V(y(t), T-t; f(y(t),u(t)), -1).\end{aligned} \tag{5.4.16}$$

It follows from (5.4.14) that $\frac{d}{dt}h(t) = 0$ almost everywhere on $(0,T)$. Then, the Lipschitz continuity of h shows $h(0) = h(T) = J(x,T,u)$. On the other hand, $h(0) = V(x,T)$ by definition of h. This means $J(x,T,u) = V(x,T)$, and hence u is optimal.

We next show the necessity. We assume u is optimal. Then, from Lemma 5.4.3, $h(t)$ defined by (5.4.15) is constant. Hence, $\frac{d}{dt}h(t) = 0$. This with (5.4.16) completes the proof. □

Remark 5.4.7
Theorem 5.4.6 mentions the relationship with the optimal control value and the state value at the current time, and hence the optimal control is immediately characterized in terms of the feedback control.

Remark 5.4.8
In addition, Theorem 5.4.6 reveals the discreteness of the sparse optimal control for control-affine systems with a box constraint, i.e., f is given by (5.4.12) and $\mathbb{U} = \{u \in \mathbb{R}^m : U_j^- \leq u_j \leq U_j^+, \forall j\}$ for some $U_j^- < 0$ and $U_j^+ > 0$. Here we take any (z,t) such that the value function V is differentiable at $(z, T-t)$. Note that such points exist almost everywhere, since V is locally Lipschitz continuous from Theorem 5.4.2. Denote the optimal input value for the current point (z,t) by $u(z,t)$. Then, the differentiability of V and (5.4.14) imply that

$$\begin{aligned}
-\psi_0(u(z,t)) &= \partial V(z, T-t; f(z, u(z,t)), -1) \\
&= DV(z, T-t) \cdot (f(z, u(z,t)), -1) \\
&= D_x V(z, T-t) \cdot f(z, u(z,t)) - V_t(z, T-t).
\end{aligned}$$

Here, V is a viscosity solution of (5.4.6), and hence V satisfies the equation at any point where V is differentiable. In other words,

$$V_t(z, T-t) - D_x V(z, T-t) \cdot f(z,u) - \psi_0(u) \leq 0$$

for all $u \in \mathbb{U}$. Hence,

$$u(z,t) \in \arg\min_{u \in \mathbb{U}} \{D_x V(z, T-t) \cdot f(z,u) + \psi_0(u)\}.$$

Here, if the system is control-affine, such as (5.4.12), then the j-th component $u_j(z,t)$ of $u(z,t)$ is given by

$$u_j(z,t) \in \arg\min_{u_j \in \mathbb{U}_j} \{D_x V(z, T-t) \cdot f_j(z) u_j + |u_j|^0\}, \tag{5.4.17}$$

where $\mathbb{U}_j$ is the set of all j-th components of $\mathbb{U}$, i.e., $\mathbb{U}_j = \{a \in \mathbb{R} : U_j^- \leq a \leq U_j^+\}$, and define $b_{z,t}^j \triangleq D_x V(z, T-t) \cdot f_j(z)$. Then, it follows from (5.4.17) that

$$u_j(z,t) \in \begin{cases} \{U_j^-\}, & \text{if} \quad b_{z,t}^j > -1/U_j^-, \\ \{U_j^-, 0\}, & \text{if} \quad b_{z,t}^j = -1/U_j^-, \\ \{0\}, & \text{if} \quad -1/U_j^+ < b_{z,t}^j < -1/U_j^-, \\ \{0, U_j^+\}, & \text{if} \quad b_{z,t}^j = -1/U_j^+, \\ \{U_j^+\}, & \text{if} \quad b_{z,t}^j < -1/U_j^+. \end{cases}$$

Thus, the optimal control takes only three values of $\{U_j^-, 0, U_j^+\}$.

5.5 Conclusions

We have investigated a finite horizon optimal control problem with the L^0 control cost functional. We have first characterized the value function as a viscosity solution to the associated HJB equation, and shown an equivalence theorem between the L^0 optimality and the L^1 optimality via the uniqueness theorem of the viscosity solution. In addition, we have derived a sufficient and necessary condition for the L^0 optimality that connects the current state and the current optimal control value. We have finally revealed the discreteness property of the sparse optimal control for control-affine systems.

Chapter 6

Conclusions

This thesis investigated sparse optimal control problems for dynamical continuous-time systems. The optimization problems find the control with the minimum length of support for given control objects. The problems are defined on infinite-dimensional spaces, compared to the standard optimization problem in the field of sparse modeling.

In Chapter 2, the sparse optimal control that achieves a state transition for linear systems from an initial state to a target state over a finite horizon was considered. This chapter showed a relationship among the L^p optimal control problems with $p \in [0, 1]$. Based on the relationship, two numerical algorithms to solve the sparse optimal control were also proposed. Chapter 3 considered a maximization problem of the controllability when the control inputs are constrained in terms of the L^0 norm. The trace of the controllability Gramian was adopted as the controllability metric. In order to characterize the optimal solution, the main problem was approximated to an optimization problem for which Pontryagin's minimum principle is applicable, and the validity of the approximation was then discussed. The framework answers when and where control inputs should be provided for high controllability. Hence, the result is also effective for time-varying node selection problem. Chapter 4 proposed a consensus protocol that promotes the sparsity of distributed control inputs. The proposed protocol was derived from a discrete-time consensus algorithm. The first- and second-order systems were assumed for agents. From the simplicity of dynamics, the distributed control was described in a closed form. Chapter 5 directly considered a sparse optimal control in a feedback framework, compared to the previous chapters. The analysis was based on the dynamical programing approach, for which the value function was characterized as a viscosity solution of the associated Hamilton-Jacobi-Bellman equation. The relationship between the sparse optimal control and the corresponding L^1 optimal control was also given in terms of the value function.

This thesis revealed the discreteness property of the sparse optimal controls. This indicates that the optimal control switches the value according to the location of the state. The analysis of the boundary in the state space where sparse optimal controls switch is left for future work. Another direction of future work includes the extension to stochastic control systems. Throughout this thesis, deterministic control systems were assumed and influence of disturbances was not covered. For example, the feasibility of the proposed control protocol in Chapter 4 could be broken by disturbances, which was illustrated through a numerical example. Also, related estimation problems when the system is corrupted by impulsive noise are in scope.

Bibliography

[1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] M. Lustig, D. L. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[3] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, 2008.

[4] J. P. Haldar, D. Hernando, and Z.-P. Liang, "Compressed-sensing MRI with random encoding," *IEEE Transactions on Medical Imaging*, vol. 30, no. 4, pp. 893–903, 2011.

[5] M. Murphy, M. Alley, J. Demmel, K. Keutzer, S. Vasanawala, and M. Lustig, "Fast ℓ_1-SPIRiT compressed sensing parallel imaging MRI: Scalable parallel implementation and clinically feasible runtime," *IEEE Transactions on Medical Imaging*, vol. 31, no. 6, pp. 1250–1262, 2012.

[6] S. S. Vasanawala, M. T. Alley, B. A. Hargreaves, R. A. Barth, J. M. Pauly, and M. Lustig, "Improved pediatric MR imaging with compressed sensing," *Radiology*, vol. 256, no. 2, pp. 607–616, 2010.

[7] D. J. Holland, M. J. Bostock, L. F. Gladden, and D. Nietlispach, "Fast multidimensional NMR spectroscopy using compressed sensing," *Angewandte Chemie*, vol. 123, no. 29, pp. 6678–6681, 2011.

[8] Y. Shrot and L. Frydman, "Compressed sensing and the reconstruction of ultrafast 2D NMR data: Principles and biomolecular applications," *Journal of Magnetic Resonance*, vol. 209, no. 2, pp. 352–358, 2011.

[9] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, p. 607, 1996.

[10] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[11] M. Ishikawa, "Structural learning with forgetting," *Neural Networks*, vol. 9, no. 3, pp. 509–521, 1996.

[12] C. De Mol, E. De Vito, and L. Rosasco, "Elastic-net regularization in learning theory," *Journal of Complexity*, vol. 25, no. 2, pp. 201–230, 2009.

[13] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

[14] M. Girolami, "A variational method for learning sparse and overcomplete representations," *Neural Computation*, vol. 13, no. 11, pp. 2517–2532, 2001.

[15] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[16] J. Zhu, S. Rosset, R. Tibshirani, and T. J. Hastie, "1-norm support vector machines," in *Advances in Neural Information Processing Systems 16*, 2004, pp. 49–56.

[17] K. Hayashi, M. Nagahara, and T. Tanaka, "A user's guide to compressed sensing for communications systems," *IEICE Transactions on Communications*, vol. 96, no. 3, pp. 685–712, 2013.

[18] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[19] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[20] M. Honma, K. Akiyama, M. Uemura, and S. Ikeda, "Super-resolution imaging with radio interferometry using sparse modeling," *Publications of the Astronomical Society of Japan*, vol. 66, no. 5, 2014.

[21] M. Nagahara and D. E. Quevedo, "Sparse representations for packetized predictive networked control," in *Proceedings of the IFAC 18th World Congress*, 2011, pp. 84–89.

[22] M. Nagahara, D. E. Quevedo, and J. Østergaard, "Sparse packetized predictive control for networked control over erasure channels," *IEEE Transactions on Automatic Control*, vol. 59, no. 7, pp. 1899–1905, 2014.

[23] H. Kong, G. C. Goodwin, and M. M. Seron, "A cost-effective sparse communication strategy for networked linear control systems: an SVD-based approach," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 14, pp. 2223–2240, 2015.

[24] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[25] S. Bhattacharya and T. Başar, "Sparsity based feedback design: A new paradigm in opportunistic sensing," in *Proceedings of the American Control Conference*, 2011, pp. 3704–3709.

[26] F. Lin, M. Fardad, and M. R. Jovanović, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2426–2431, 2013.

[27] B. Polyak, M. Khlebnikov, and P. Shcherbakov, "An LMI approach to structured sparse feedback design in linear control systems," in *Proceedings of the European Control Conference*, 2013, pp. 833–838.

[28] M. B. Wakin, B. M. Sanandaji, and T. L. Vincent, "On the observability of linear systems from random, compressive measurements," in *Proceedings of the 49th IEEE Conference on Decision and Control*, 2010, pp. 4447–4454.

[29] B. M. Sanandaji, M. B. Wakin, and T. L. Vincent, "Observability with random observations," *IEEE Transactions on Automatic Control*, vol. 59, no. 11, pp. 3002–3007, 2014.

[30] A. Charles, M. Asif, J. Romberg, and C. Rozell, "Sparsity penalties in dynamical system estimation," in *Proceedings of the 45th Annual Conference on Information Sciences and Systems*, 2011, pp. 1–6.

[31] T. Chen, M. S. Andersen, L. Ljung, A. Chiuso, and G. Pillonetto, "System identification via sparse multiple kernel-based regularization using sequential convex optimization techniques," *IEEE Transactions on Automatic Control*, vol. 59, no. 11, pp. 2933–2945, 2014.

[32] R. Toth, B. Sanandaji, K. Poolla, and T. Vincent, "Compressive system identification in the linear time-invariant framework," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 783–790.

[33] B. M. Sanandaji, T. L. Vincent, M. B. Wakin, R. Toth, and K. Poolla, "Compressive system identification of LTI and LTV ARX models," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 791–798.

[34] E. J. Candès and Y. Plan, "Near-ideal model selection by ℓ_1 minimization," *The Annals of Statistics*, vol. 37, no. 5A, pp. 2145–2177, 2009.

[35] M. Elad, *Sparse and Redundant Representations.* Springer, 2010.

[36] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing.* Birkhäuser Basel, 2013, vol. 1, no. 3.

[37] C. Chan, "The state of the art of electric, hybrid, and fuel cell vehicles," *Proceedings of the IEEE*, vol. 95, no. 4, pp. 704–718, 2007.

[38] P. Shakouri, A. Ordys, P. Darnell, and P. Kavanagh, "Fuel efficiency by coasting in the vehicle," *International Journal of Vehicular Technology*, vol. 2013, 2013.

[39] M. Nalbach, A. Korner, and S. Kahnt, "Active engine-off coasting using 48V: Economic reduction of CO_2 emissions," in *Proceedings of the 17th International Congress ELIV*, 2015, pp. 41–51.

[40] E. Khmelnitsky, "On an optimal control problem of train operation," *IEEE Transactions on Automatic Control*, vol. 45, no. 7, pp. 1257–1266, 2000.

[41] C. Chang and S. Sim, "Optimising train movements through coast control using genetic algorithms," *IEE Proceedings-Electric Power Applications*, vol. 144, no. 1, pp. 65–73, 1997.

[42] M. Ogasa, "Energy saving and environmental measures in railway technologies," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 3, no. 1, pp. 15–20, 2008.

[43] R. Liu and I. M. Golovitcher, "Energy-efficient operation of rail vehicles," *Transportation Research Part A: Policy and Practice*, vol. 37, no. 10, pp. 917–932, 2003.

[44] G. Vossen and H. Maurer, "On L^1-minimization in optimal control and applications to robotics," *Optimal Control Applications and Methods*, vol. 27, no. 6, pp. 301–321, 2006.

[45] B. Dunham, "Automatic on/off switching gives 10-percent gas saving," *Popular Science*, vol. 205, no. 4, p. 170, 1974.

[46] D. Jeong and W. Jeon, "Performance of adaptive sleep period control for wireless communications systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 3012–3016, 2006.

[47] L. Kong, G. Wong, and D. Tsang, "Performance study and system optimization on sleep mode operation in IEEE 802.16e," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4518–4528, 2009.

[48] M. Athans and P. L. Falb, *Optimal Control: An Introduction to the Theory and Its Applications*. Dover Publications, 1966.

[49] M. R. Jovanović and F. Lin, "Sparse quadratic regulator," in *Proceedings of the European Control Conference*, 2013, pp. 1047–1052.

[50] T. Chen and B. A. Francis, *Optimal sampled-data control systems*. Springer Science & Business Media, 2012.

[51] W. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in *Proceedings of the 51st IEEE Conference on Decision and Control*, 2012, pp. 3270–3285.

[52] D. V. Dimarogonas, E. Frazzoli, and K. H. Johansson, "Distributed event-triggered control for multi-agent systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1291–1297, 2012.

[53] G. Guo, L. Ding, and Q.-L. Han, "A distributed event-triggered transmission strategy for sampled-data consensus of multi-agent systems," *Automatica*, vol. 50, no. 5, pp. 1489–1496, 2014.

[54] M. Nagahara, D. E. Quevedo, and D. Nešić, "Maximum hands-off control: a paradigm of control effort minimization," *IEEE Transactions on Automatic Control*, vol. 61, no. 3, pp. 735–747, 2016.

[55] T. Ikeda and M. Nagahara, "Maximum hands-off control without normality assumption," in *Proceedings of the American Control Conference*, 2016, pp. 2050–2055.

[56] K. Ito and K. Kunisch, "Optimal control with $L^p(\Omega)$, $p \in [0, 1)$, control cost," *SIAM Journal on Control and Optimization*, vol. 52, no. 2, pp. 1251–1275, 2014.

[57] D. Kalise, K. Kunisch, and Z. Rao, "Infinite horizon sparse optimal control," *Journal of Optimization Theory and Applications*, vol. 172, no. 2, pp. 481–517, 2017.

[58] D. Chatterjee, M. Nagahara, D. E. Quevedo, and K. M. Rao, "Characterization of maximum hands-off control," *Systems & Control Letters*, vol. 94, pp. 31–36, 2016.

[59] Z. Rao, "A dynamic programming approach for L^0 optimal control design," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 2886–2891, 2017.

[60] R. Herzog, G. Stadler, and G. Wachsmuth, "Directional sparsity in optimal control of partial differential equations," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 943–963, 2012.

[61] K. Kunisch, K. Pieper, and B. Vexler, "Measure valued directional sparsity for parabolic optimal control problems," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 3078–3108, 2014.

[62] G. Wachsmuth and D. Wachsmuth, "Convergence and regularization results for optimal control problems with sparsity functional," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 17, no. 3, pp. 858–886, 2011.

[63] E. Hille and R. S. Phillips, *Functional Analysis and Semi-groups.* American Mathematical Society, 1996, vol. 31.

[64] R. T. Rockafellar, *Convex Analysis.* Princeton University Press, 1970.

[65] J. LaSalle, "Time optimal control systems," *Proceedings of the National Academy of Sciences*, vol. 45, no. 4, pp. 573–577, 1959.

[66] W. C. Grimmell, "The existence of piecewise continuous fuel optimal controls," *SIAM Journal on Control*, vol. 5, no. 4, pp. 515–519, 1967.

[67] O. Hájek, "On the number of roots of exp-trig polynomials," *Computing*, vol. 18, no. 2, pp. 177–183, 1977.

[68] C. Chen, *Introduction to linear system theory.* Holt, Rinehart and Winston, 1970.

[69] R. F. Stengel, *Optimal Control and Estimation.* Dover Publications, 1994.

[70] M. Elad, J.-L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Applied and Computational Harmonic Analysis*, vol. 19, no. 3, pp. 340–358, 2005.

[71] M. Nagahara, "Discrete signal reconstruction by sum of absolute values," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1575–1579, 2015.

[72] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[73] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*. Springer, 2008, pp. 95–110.

[74] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[75] J. Eckstein and D. Bertsekas, "On the Douglas-Rachford splitting method and proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, 1992.

[76] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[77] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds. New York: Springer-Verlag, 2011, pp. 185–212.

[78] G. Fung and O. Mangasarian, "Equivalence of minimal ℓ_0-and ℓ_p-norm solutions of linear equalities, inequalities and linear programs for sufficiently small p," *Journal of optimization theory and applications*, vol. 151, no. 1, pp. 1–10, 2011.

[79] T. Ishizaki, K. Kashima, J.-i. Imura, A. Katoh, H. Morita, and K. Aihara, "Distributed parameter modeling and finite-frequency loop-shaping of electromagnetic molding machine," *Control Engineering Practice*, vol. 21, no. 12, pp. 1735–1743, 2013.

[80] T. Kailath, *Linear systems*. Prentice-Hall Englewood Cliffs, NJ, 1980, vol. 156.

[81] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 40–52, 2014.

[82] T. H. Summers, F. L. Cortesi, and J. Lygeros, "On submodularity and controllability in complex dynamical networks," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 91–101, 2016.

[83] P. Müller and H. Weber, "Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems," *Automatica*, vol. 8, no. 3, pp. 237–246, 1972.

[84] F. Clarke, *Functional analysis, calculus of variations and optimal control.* Springer Science & Business Media, 2013, vol. 264.

[85] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions.* Springer Science & Business Media, 2002.

[86] A. Olshevsky, "Minimal controllability problems," *IEEE Transactions on Control of Network Systems*, vol. 1, pp. 249–258, 2014.

[87] S. Assadi, S. Khanna, Y. Li, and V. M. Preciado, "Complexity of the minimum input selection problem for structural controllability," in *IFAC-PapersOnLine*, 2015, pp. 70–75.

[88] A. Clark, L. Bushnell, and R. Poovendran, "On leader selection for performance and controllability in multi-agent systems," in *Proceedings of the 51st IEEE Conference on Decision and Control*, 2012, pp. 86–93.

[89] W. Yang, X. Wang, and H. Shi, "Optimal control nodes selection for consensus in multi-agent systems," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 11 697–11 702, 2014.

[90] R. E. Kalman, "Mathematical description of linear dynamical systems," *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, vol. 1, no. 2, pp. 152–192, 1963.

[91] Y. Zhao, F. Pasqualetti, and J. Cortés, "Scheduling of control nodes for improved network controllability," in *Proceedings of the 55th IEEE Conference on Decision and Control*, 2016, pp. 1859–1864.

[92] E. Nozari, F. Pasqualetti, and J. Cortés, "Time-invariant versus time-varying actuator scheduling in complex networks," in *Proceedings of the American Control Conference*, 2017, pp. 4995–5000.

[93] Z. Lin, B. Francis, and M. Maggiore, "Necessary and sufficient graphical conditions for formation control of unicycles," *IEEE Transactions on Automatic Control*, vol. 50, no. 1, pp. 121–127, 2005.

[94] R. Olfati-Saber, "Distributed kalman filtering for sensor networks," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 5492–5498.

[95] G. F. Young, L. Scardovi, A. Cavagna, and I. Giardina, "Starling flock networks manage uncertainty in consensus at low cost," *PLoS Computational Biology*, vol. 9, no. 1, pp. 1–7, 2013.

[96] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[97] W. Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," in *Proceedings of the American Control Conference*, 2005, pp. 1859–1864.

[98] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[99] F. Lewis, H. Zhang, K. Hengster-Movric, and A. Das, *Cooperative control of multi-agent systems: optimal and adaptive design approaches.* London: Springer-Verlag, 2014.

[100] N. Challapalli, M. Nagahara, and M. Vidyasagar, "Continuous hands-off control by CLOT norm minimization," in *Proceedings of the 20th IFAC World Congress*, 2017, pp. 15 019–15 024.

[101] C. D. Persis and P. Frasca, "Robust self-triggered coordination with ternary controllers," *IEEE Transactions on Automatic Control*, vol. 58, no. 12, pp. 3024–3038, 2013.

[102] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 227–234, 1995.

[103] T. Ikeda and M. Nagahara, "Value functions in sparse optimal and L^1 optimal controls," in *Proceedings of the SICE International Symposium on Control Systems*, 2015.

[104] H. Hermes and J. P. Lasalle, *Function Analysis and Time Optimal Control.* Academic Press, 1969.

[105] T. Ikeda and K. Kashima, "On sparse optimal control for general linear systems," *IEEE Transactions on Automatic Control*, 2018.

[106] W. Alt and C. Schneider, "Linear-quadratic control problems with L^1-control cost," *Optimal Control Applications and Methods*, vol. 36, no. 4, pp. 512–534, 2015.

[107] I. Exarchos, E. A. Theodorou, and P. Tsiotras, "Stochastic L^1-optimal control via forward and backward sampling," *Systems & Control Letters*, vol. 118, pp. 101–108, 2018.

[108] S. Sukumar and D. Chatterjee, "A jammer's perspective of reachability and LQ optimal control," *Automatica*, vol. 70, pp. 295–302, 2016.

[109] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Springer Science & Business Media, 2008.

List of Figures

Acknowledgements

I would like to express my heartfelt appreciation to Professor Yoshito Ohta for his enthusiastic and widespread advice. His educational philosophy impressed me and developed my attitude toward science. I also appreciate him for providing me this precious study opportunity in his laboratory. I would like to express my deepest gratitude to my supervisor Associate Professor Kenji Kashima for his elaborated guidance, considerable encouragement, and invaluable discussion. His insightful comments and constructive suggestions improved my study and led me to the completion of this thesis. I owe a huge debt of gratitude to Assistant Professor Kentaro Ohki for his continuous support and encouragement. My sincere thanks also go to Professor Masaaki Nagahara at The University of Kitakyushu.

I would like to record my gratitude to Japan Society for the Promotion of Science, Marubun Research Promotion Foundation, and The NEC C&C Foundation for the funding supports.

I would like to acknowledge past and present members of the control systems theory group laboratory for their helpful and interesting discussion and daily friendship. Finally, I am grateful to my family for their deep understanding and encouragement.

Publications

Refereed Journal Articles

1. T. Ikeda, M. Nagahara, and K. Kashima, Maximum hands-off distributed control for consensus of multi-agent systems with sampled-data state observation, *IEEE Transactions on Control of Network Systems*. (to appear)

2. T. Ikeda and K. Kashima, On sparse optimal control for general linear systems, *IEEE Transactions on Automatic Control*. (to appear)

3. T. Ikeda and M. Nagahara, Time-optimal hands-off control for linear time-invariant systems, *Automatica*, vol. 99, pp. 54–58, 2019.

4. T. Ikeda and K. Kashima, Sparsity-constrained controllability maximization with application to time-varying control node selection, *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 321–326, 2018.

5. T. Ikeda and M. Nagahara, Discrete-valued model predictive control using sum-of-absolute-values optimization, *Asian Journal of Control*, vol. 20, no. 1, pp. 196–206, 2018.

6. T. Ikeda, M. Nagahara, and S. Ono, Discrete-valued control of linear time-invariant systems by sum-of-absolute-values optimization, *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 2750–2763, 2017.

7. T. Ikeda and M. Nagahara, Value function in maximum hands-off control for linear systems, *Automatica*, vol. 64, pp. 190–195, 2016.

Presentations at Refereed International Conferences

1. T. Ikeda and K. Kashima, Sparsity-constrained controllability maximization with application to time-varying control node selection, *57th IEEE Conference on Decision and Control*, Miami, Dec. 17, 2018.

2. T. Ikeda and K. Kashima, Path planning for linear systems based on sparsity-in-time, *SICE Annual Conference*, Nara, Sep. 12, 2018.

3. T. Ikeda and K. Kashima, Optimal trajectories and state estimation based on sparsity-in-time, *20th IFAC World Congress*, Toulouse, Jul. 14, 2017.

4. T. Ikeda, M. Nagahara, and K. Kashima, Consensus by maximum hands-off distributed control with sampled-data state observation, *55th IEEE Conference on Decision and Control*, Las Vegas, Dec. 12, 2016.

5. T. Ikeda, M. Nagahara, and D. E. Quevedo, Quantized self-triggered control by sum-of-absolute-values optimization, *22nd International Symposium on Mathematical Theory of Networks and Systems*, Minneapolis, Jul. 14, 2016.

6. T. Ikeda and M. Nagahara, Maximum hands-off control without normality assumption, *American Control Conference*, Boston, Jul. 6, 2016.

7. T. Ikeda and M. Nagahara, Computation of maximum hands-off control, *SICE International Symposium on Control Systems*, Nagoya, Mar. 9, 2016.

8. T. Ikeda and M. Nagahara, Discrete model predictive control by sum-of-absolute-values optimization, *The 47th ISCIE International Symposium on Stochastic Systems Theory and Its Applications*, Hawaii, Dec. 5, 2015.

9. T. Ikeda and M. Nagahara, Continuity of the value function in sparse optimal control, *The 10th Asian Control Conference*, Kota Kinabalu, Jun. 2, 2015.

10. T. Ikeda and M. Nagahara, Value functions in sparse optimal and L^1-optimal controls, *SICE International Symposium on Control Systems*, Tokyo, Mar. 5, 2015.

Copyright Notice