Approaches to the Design of Model Predictive Controller for Constrained Linear Systems With Bounded Disturbances

Wang Chen

Department of Mechanical Engineering

A thesis submitted to the National University of Singapore

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

2009

<u>Statement of Originality</u>

I hereby certify that the content of this thesis is the result of work done by me and has not been submitted for a higher degree to any other University or Institution.

. . . .

Date   Wang Chen

Acknowledgments

I would like to express my sincere appreciation to my supervisor, Assoc. Prof. Ong Chong-Jin, for his patient guidance, insightful comments, strong encouragements and personal concerns both academically and otherwise throughout the course of the research. I benefit a lot from his comments and critiques. I would also like to thank Assoc. Prof. Melvyn Sim, whom I feel lucky to known in NUS. A number of ideas in this thesis originate from the discussion with Melvyn.

I gratefully acknowledge the financial support provided by the National University of Singapore through Research Scholarship that makes it possible for me to study for academic purpose.

Thanks are also given to my friends and technicians in Mechatronics and Control Lab for their support and encouragement. They have provided me with helpful comments, great friendship and a warm community during the past few years in NUS.

Finally, my deepest thanks go to my parents and especially my wife Chang Hong for their encouragements, moral supports and loves. To support me, my wife gave up a lot and she is always by my side during the bitter times. I love you forever.

Abstract

This thesis is concerned with the Model Predictive Control (MPC) of linear discrete time-invariant systems with state and control constraints and subject to bounded disturbances.

This thesis proposes a new form of affine disturbance feedback control parametrization, and proves that this parametrization has the same expressive ability as the affine time-varying disturbance (state) feedback parametrization found in the recent literature. Consequently, the admissible sets of the finite horizon (FH) optimization problems under both parametrization are the same. Furthermore, by minimizing a norm-like cost function of the design variables, the MPC controller derived using the proposed parametrization steers the system state to the minimal disturbance invariant set asymptotically, and this minimal disturbance invariant set is associated with a feedback gain which is pre-chosen and fixed in the proposed control parametrization.

The second contribution of this thesis is a modification of the original proposed affine disturbance feedback parametrization. Specifically, the realized disturbances are not utilized in the parametrization. Hence, the resulting MPC controller is a purely state feedback law instead of a dynamic compensator in the previous case. It is proved that under the MPC controller derived using the new parametrization, the closed-loop system state converges to the same minimal disturbance invariant set with probability one if the distribution of the disturbance satisfies certain conditions. In the case where these conditions are not satisfied, the closed-loop system state can also converge to the same set if a less intuitive cost function is used in the FH optimization problem.

The third contribution of this thesis is the generalization of affine disturbance feedback parametrization to a piecewise affine function of disturbances. Hence, larger admissible set and better performance of the MPC controller could be expected under this parametrization. Unfortunately, the FH optimization problem under this parametrization is not directly computable. However, if the disturbance set is an absolute set, deterministic equivalence of the FH optimization problem can be determined and is solvable. Even if the disturbance set is not absolute, the FH optimization problem can still be solved by considering a larger disturbance set, and the resulting controller is not worse than the one under linear disturbance feedback law. In addition, minimal disturbance invariant set convergence stability is also achievable under this parametrization.

The fourth contribution of this thesis is a feedback gain design approach. Since asymptotic behavior of the closed-loop system under any of the proposed parametrization is determined by a fixed feedback gain chosen a priori in the parametrization, one method of designing this feedback gain is introduced to control the asymptotic behavior of the closed-loop system. The underlying idea of the method is that the support function of the minimal disturbance invariant set and its derivative with respect to the feedback gain can be evaluated as accurately as possible. Hence, an optimization problem with constraints imposed on the support function of the minimal disturbance invariant set can be solved. Therefore, a feedback gain can be designed by solving such an optimization problem so that the corresponding minimal disturbance invariant set has optimal supports along given directions.

Finally, MPC of systems with probabilistic constraints are considered. Properties of probabilistic constraint-admissible sets of such systems are studied and it turns out that such sets are generally non-convex, non-invariant and hard to determine. For the purpose of application, an inner invariant approximation is introduced. This is achieved by approximate probabilistic constraints by robust counter parts. It is shown that under certain conditions, the inner approximation can be finitely determined by a proposed algorithm. This inner approximation set is applied as a terminal set in the design of MPC controllers for probabilistically constrained systems. It is also proved that under the resulting controller, the closed-loop system is stable and all of the constraints, including both deterministic and probabilistic, are satisfied.

Table of Contents

List of Figures

List of Tables

Acronyms

FH	Finite Horizon
CTLD system	Constrained Time-invariant Linear Discrete-time system
LMI	Linear Matrix Inequality
LP	Linear Programming
MPC	Model Predictive Control
QP	Quadratic Programming
ISS	Input-to-State Stability
LQR	Linear Quadratic Regulator
SISO system	Single Input Single Output system

Nomenclature

A^T	transposed matrix (or vector)
$\rho(A)$	spectral radius of A
$A \succeq 0$	symmetric positive semi-definite matrix
$A \succ 0$	symmetric positive definite matrix
$\|\cdot\|_p$	p-norm of vector
$\|\cdot\|_F$	Frobenius norm
$\|\cdot\|$	Euclidean norm
I_n	n by n identity matrix
$\mathbb{R}$	set of real numbers
$\mathbb{R}^n$	n-dimensional real Euclidean space
$\mathbb{R}^{n\times m}$	set of $n \times m$ real matrix
$\mathrm{int}(\Omega)$	interior of Ω
$\lvert\Omega\rvert$	cardinality of Ω
$\mathbf{V}(\Omega)$	vertex set of Ω
$\mathrm{CH}(\Omega)$	convex hull of Ω
$B_\varsigma(\cdot)$	ς norm-ball

$\mathscr{I}$	index set
$\mathbb{Z}_k$	{0,1,2,...,k}
$\mathbb{Z}_k^+$	{1,2,...,k}
t	discrete-time index
F_∞	minimal disturbance invariant set
O_∞	maximal constraint admissible disturbance invariant set
X_f	terminal set
F	terminal cost
N	prediction horizon
$\mathbf{1}_r$	an r-vector with all elements being 1
$\delta_\Omega(\mu)$	support function of Ω, i.e. $\delta_\Omega(\mu) = \max_{\omega\in\Omega}\mu^T\omega$
$\Theta \oplus \Omega$	Minkowski sum of set Θ and Ω
$\Theta \ominus \Omega$	P-difference of set Θ and Ω
$\alpha\Omega$	scale of set Ω
μ_i	i^{th} element of vector μ
$A \otimes B$	Kronecker product of matrix A and B

Chapter 1

Introduction

This thesis is concerned with the control of systems under the Model Predictive Control (MPC) framework. It focuses on the design of MPC controller for a discrete time-invariant linear system with bounded additive disturbances while fulfilling state and control constraints. These constraints are either deterministic (hard) or probabilistic (soft) in nature. The rest of this chapter provides a review of the literature on this problem.

1.1 Background

Many control strategies developed around the 1960s do not explicitly take uncertainties into account. Typically, the robustness of the closed-loop system is described by notions such as gain margin and phase margin. Another common feature of those strategies is that constraints are also omitted in their design consideration. However disturbances and physical constraints, such as actuator saturation, maximal speed of a motor, minimal return of an investment, etc, are always important constraints in practice. Omitting these in the controller design may lead to a state or control action that violates them and result in unpredictable system behaviors or even physical damage to the systems.

Researchers began to focus on the control of constrained and disturbed systems after the 1960s. The control of such systems has been addressed intensively in the literature, and various methods have appeared, such as anti-windup control, reference governor, switching control and several others, see [1, 2, 3, 4, 5, 6, 7, 8]. Among them, a popular approach is Model Predictive Control, see [9, 10, 11, 12, 13, 14, 15, 16] and the references cited therein. This approach has been widely applied in industries [17], especially in the process industry since the 1980s. The basic idea of MPC is quite simple and can be found in several textbooks on optimal control theory [18, 19, 20]. In particular, Lee and Markus in [20] described the underlying idea of MPC as follows:

> "One technique for obtaining a feedback controller synthesis from knowledge of open-loop controllers is to measure the current control process state and then compute very rapidly for the open-loop control function. The first portion of this function is then used during a short time interval, after which a new measurement of the process state is made and a new open-loop control function is computed for this new measurement. The procedure is then repeated."

According to the above description, a model of the "control process" is available to predict the system behavior, and one practical and useful control process is that described by a linear time-invariant difference equation

$$x(t+1) = Ax(t) + Bu(t) \tag{1.1}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ are the state and control of the system at time t, respectively, (A, B) are appropriate matrices. The state and control are subject to a constraint

$$(x(t), u(t)) \in Y \subset \mathbb{R}^{n+m} \tag{1.2}$$

where Y represents the joint state and control constraint imposed on the system. The MPC approach designs a control law by looking ahead N steps at a time. Let the control in the N steps be

$$\mathbf{u}(t) := \{u(0|t), \cdots, u(N-1|t)\} \in \mathbb{R}^{Nm} \tag{1.3}$$

where $u(i|t) \in \mathbb{R}^m$ is the predicted control i steps from time t. Let $x(i|t)$ be the ith predicted state within the N steps and collect all the predicted states in,

$$\mathbf{x}(t) \;\; := \;\; \{x(0|t), \cdots, x(N|t)\} \in \mathbb{R}^{(N+1)n} \tag{1.4}$$

The MPC approach computes $\mathbf{u}(t)$ using a cost function of the form

$$J(\mathbf{x}(t),\mathbf{u}(t)) := \sum_{i=0}^{N-1} \ell(x(i|t),u(i|t)) + F(x(N|t)), \tag{1.5}$$

where $\ell(\cdot,\cdot)$ and $F(\cdot)$ are appropriate stage and terminal costs, respectively. The predicted control sequence can be determined by solving the following finite horizon (FH) optimization problem, referred to as $\mathscr{P}_N(x(t))$,

$$\begin{aligned}
\min_{\mathbf{u}(t)} \quad & J(\mathbf{x}(t),\mathbf{u}(t)) && \text{(1.6a)}\\
\text{s.t.} \quad & x(0|t) = x(t), && \text{(1.6b)}\\
& x(i+1|t) = Ax(i|t) + Bu(i|t),\ \forall i \in \mathbb{Z}_{N-1}, && \text{(1.6c)}\\
& (x(i|t),u(i|t)) \in Y,\ \forall i \in \mathbb{Z}_{N-1} && \text{(1.6d)}\\
& x(N|t) \in X_f && \text{(1.6e)}
\end{aligned}$$

where $\mathbb{Z}_k$ denotes the integer set $\{0,1,\ldots,k\}$ and X_f is an appropriate terminal constraint set. Based on the measurement of $x(t)$, $\mathscr{P}_N(x(t))$ yields an optimal control sequence

$$\mathbf{u}^*(t) := \{u^*(0|t),\cdots,u^*(N-1|t)\}. \tag{1.7}$$

The first control of $\mathbf{u}^*(t)$, $u^*(0|t)$, is then applied to system (1.1) as the control at time t.

Therefore, the MPC control law can be implicitly expressed as

$$\kappa(x(t)) := u^*(0|t). \tag{1.8}$$

At time instant $t+1$ when the measurement of $x(t+1)$ is available, $\mathscr{P}_N(x(t+1))$ is solved once again and the applied control is $u(t+1) = \kappa(x(t+1))$. By repeating this procedure at every time t, an MPC controller is implemented. One important measure of the performance of MPC that is mentioned frequently in this thesis is the admissible set. It is the set of system state within which controller (1.8) is defined and is given by

$$\mathscr{X}_N := \{x|\ \exists \mathbf{u} \text{ such that } \mathscr{P}_N(x) \text{ is feasible}\}. \tag{1.9}$$

Although MPC application dates back to the 1970s [17], its theoretical study only appeared in the late 1980s. One important requirement of MPC at that time is the stability of system (1.1) under the MPC control law (1.8). To ensure stability, the terminal constraint (1.6e) and the terminal cost $F(\cdot)$ in (1.5) play important roles. Specifically, the origin of the closed-loop system is asymptotically stable by applying either appropriate X_f set or $F(\cdot)$ or both based on the works of [21] by Bitmead *et al.*, [22] by Rawlings and Muske, [23] by Couchman *et al.*, [24] by Scokaert *et al.*, [25] by Sznaier and Damborg, [26] by De Nicolao *et al.* and others. The survey paper [6] by Mayne *et al.* summarizes the needed conditions for stability: X_f is a constraint-admissible invariant set under a local controller and the terminal cost function $F(\cdot)$ is a local Lyapunov function.

The MPC problem becomes more complicated when uncertainty in the form of additive disturbances are present. In this case, system (1.1) becomes

$$x(t+1) = Ax(t) + Bu(t) + w(t) \tag{1.10a}$$

$$w(t) \in W \tag{1.10b}$$

where $w(t) \in \mathbb{R}^n$ is the disturbance at time t and $w(t)$ is assumed to be bounded in the set $W \subset \mathbb{R}^n$. MPC of system (1.10) is the focus of this thesis. With disturbances in (1.10), the optimization problem $\mathscr{P}_N(x(t))$ defined by (1.6) has to be reformulated to take into account: (i) the effect of $w(t)$ and (ii) the interpretations of constraints (1.6d) and (1.6e) in the presence of $w(t)$.

For the control of system (1.10), one novel MPC approach that is closely related to the optimization (1.6) is proposed by Mayne *et al.* in [13]. In that work, it is assumed that a disturbance invariant set Z can be determined for the system (1.10) under a linear feedback law $u(t) = Kx(t)$ in the sense that $(A+BK)Z \oplus W \subseteq Z$, where $(A+BK)Z := \{z |\ z = (A+BK)\hat{z},\ \hat{z} \in Z\}$ and $\Omega_1 \oplus \Omega_2 := \{\omega = \omega_1 + \omega_2 |\ \omega_1 \in \Omega_1,\ \omega_2 \in \Omega_2\}$ is the Minkowski sum of sets Ω_1 and Ω_2. Using this set Z and feedback gain K, optimization problem (1.6) is reformulated as

$$\min_{(x(0|t),\mathbf{u}(t))} \quad J(\mathbf{x}(t),\mathbf{u}(t)) \tag{1.11a}$$

$$\text{s.t.} \quad x(t) \in x(0|t) \oplus Z, \tag{1.11b}$$

$$x(i+1|t) = Ax(i|t) + Bu(i|t),\ \forall i \in \mathbb{Z}_{N-1}, \tag{1.11c}$$

$$(x(i|t), u(i|t)) \in Y \ominus (Z \times KZ),\ \forall i \in \mathbb{Z}_{N-1} \tag{1.11d}$$

$$x(N|t) \in X_f \ominus Z \tag{1.11e}$$

where $\Omega_1 \ominus \Omega_2 := \{\omega |\ \omega + \omega_2 \in \Omega_1,\ \forall \omega_2 \in \Omega_2\}$ is the Pontryagin difference or P-difference between Ω_1 and Ω_2. Optimization (1.11) differs from (1.6) in that $x(0|t)$ is a design variable in (1.11) and $x(t)$, instead of being equal to $x(0|t)$ in (1.6), is only required to be in a neighborhood of $x(0|t)$ characterized by Z. Additionally, constraint sets in (1.11d) and (1.11e) are tightened so that the constraints are satisfied by the true states and controls. After solving (1.11), the MPC control law applied to system (1.10) is

$$\kappa(x(t)) = u^*(0|t) + K(x(t) - x^*(0|t)) \tag{1.12}$$

which is also different from (1.8). Mayne *et al.* [13] show that under mild assumptions of the cost function $J(\mathbf{x}(t),\mathbf{u}(t))$ and terminal set X_f the set Z is robustly exponentially stable for the closed-loop system under controller (1.12). A similar idea of introducing additional terms to the MPC controller can also be found in [8] by Langson *et al.* and

[27] by Mayne *et al.*

Other MPC approaches for the control of system (1.10) have also appeared and a popular one is the so called “min-max” approach, see for example [28] by Michalska and Mayne, [29] by Badgwell, [30] by Scokaert and Mayne, [31] by Bemporad *et al.*, [32] and [33]by Kerrigan and Maciejowski. The FH optimization problem minimizes the worst case that the disturbance can bring and takes the general form:

$$\min_{\mathbf{u}(t)} \quad \max_{\mathbf{w}(t)} J(\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)) \tag{1.13a}$$

$$\text{s.t.} \quad x(i+1|t) = Ax(i|t) + Bu(i|t) + w(i|t),\ \forall i \in \mathbb{Z}_{N-1} \tag{1.13b}$$

$$x(0|t) = x(t) \tag{1.13c}$$

$$(x(i|t), u(i|t)) \in Y,\ \forall i \in \mathbb{Z}_{N-1},\ \forall \mathbf{w}(t) \in W^N \tag{1.13d}$$

$$x(N|t) \in X_f,\ \forall \mathbf{w}(t) \in W^N \tag{1.13e}$$

where

$$\mathbf{w}(t) := \{w(0|t), \cdots, w(N-1|t)\} \in W^N \subset \mathbb{R}^{Nn} \tag{1.14}$$

and W^N is the N times cartesian product of W.

Although min-max MPC optimization problem (1.13) has precise interpretation, its computation is not easy: (i) the expression of the maximum of $J(\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t))$ with respect to $\mathbf{w}(t)$ is hard to determine in general; (ii) constraints (1.13d) and (1.13e) has an infinite number of constraints, one for each possible disturbance sequence, $\mathbf{w}(t) \in W^N$. Fortunately, if $J(\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t))$ is a convex function with respect to $\mathbf{w}(t)$ and W^N set is a polytope, the maximizer of $\max_{\mathbf{w}(t)} J(\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t))$ occurs at one of the vertices of W^N. Hence, the maximizer of $\max_{\mathbf{w}(t)} J(\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t))$ can be determined by searching over the vertices of W^N. The same strategy can also be applied to handle constraints (1.13d) and (1.13e). Namely, instead of considering all $\mathbf{w}(t) \in W^N$, we can consider $\mathbf{w}(t)$ generated from the vertices of W^N, avoiding the infinite number of constraints. However, even when vertices of W^N are considered, the number of constraints increases exponentially with the control horizon and the dimension of the system. Consequently, the computational burden of the resulting optimization problem can be extremely high when N or n is large. This usually limits min-max MPC to applications on small-scale problems.

Possible solution to the computation of MPC is to solve the FH optimization using off-line approaches or efficient on-line algorithms, such works includes [15] and [34] by Muñoz de la Peña *et al.*, [35] and [31] by Bemporad *et al.* and [36] by Goulart *et al.*

Besides the computational issue of min-max MPC problem (1.13), the resulting MPC controller derived using (1.13) can be very conservative since the optimal control sequence must stabilize the system against all possible disturbances while satisfying the state and control constraints. One direct result of conservatism of the controller is that the set of admissible initial state of problem (1.13) becomes small. A solution to reduce the conservatism is to parameterize the control by available information such as realized states or disturbances or both so that the influence of the disturbances on the system behavior can be compensated and reduced. The control parametrization and the associated stability of the corresponding closed-loop system have been research topics since the late 1990s and various results have appeared. Some of them are closely related to the work of this thesis. In the next section, a general review on control parametrization and closed-loop system stability is given.

1.2 Review of Control Parametrization in MPC

As discussed in the previous section, control parametrization plays an important role in MPC of systems with disturbances. It determines the degree of conservatism of the resulting MPC controller and the size of the admissible set $\mathscr{X}_N$. It is known that the most direct parametrization $\mathbf{u}(t) := \{u(0|t), \cdots, u(N-1|t)\}$ where $u(i|t)$ is a fixed value leads to a conservative MPC controller and a small admissible set for system (1.10). This is because a fixed value control sequence has limited flexibility to handle all possible sequences of $\mathbf{w}(t)$. Hence, $u(i|t)$ is ofen parameterized as functions of state x and/or disturbance w and when the function is more general, it is expected that $\mathscr{X}_N$ is larger. To reduce conservatism and enlarge $\mathscr{X}_N$, various control parametrization have been proposed in the literature and they are reviewed below.

Fixed Feedback Gain Parametrization

One of the most popular control parametrization, referred to as *fixed feedback gain parametrization*, is

$$u(i|t) = K_f x(i|t) + c(i|t),\ \forall i \in \mathbb{Z}_{N-1} \tag{1.15}$$

where K_f is a fixed feedback gain chosen a priori such that $A+BK_f$ is stable and $c(i|t) \in \mathbb{R}^m$, $i \in \mathbb{Z}_{N-1}$ are the new design variables, see [9] by Bemporad, [10] by Chisci *et al.*, [11] by Rossiter *et al.*, [12] by Lee and Kouvaritakis and [13] by Mayne *et al.* An advantage of this fixed feedback gain parametrization is the available characterization of the asymptotic behavior of the closed-loop system. This is well exemplified in the literature by the work of [10] by Chisci *et al.* In that work, the fulfillment of the state and control constraints is guaranteed by imposing tightened constraints on the nominal state and control variables. First, note that the predicted system state of (1.10) under control law (1.15) is

$$\begin{aligned} x(i|t) &= \underbrace{\Phi^i x(0|t) + \sum_{j=0}^{i-1} \Phi^{i-1-j} Bc(j|t)}_{\bar{x}(i|t)} + \underbrace{\sum_{j=0}^{i-1} \Phi^{i-1-j} w(j|t)}_{\in F_i} \\ &= \bar{x}(i|t) + \sum_{j=0}^{i-1} \Phi^{i-1-j} w(j|t) \end{aligned} \tag{1.16}$$

where $\Phi = A + BK_f$, $\bar{x}(i|t)$ is the nominal value of $x(i|t)$ or the state $x(i|t)$ with the absence of $w(i|t)$, the last term in (1.16) is due to the presence of disturbance and its value belongs to the set F_i defined by

$$F_0 := \varnothing, \;\; F_i := W \oplus \Phi W \oplus \cdots \oplus \Phi^{i-1} W. \tag{1.17}$$

Clearly, F_i characterizes the reachable set of state $x(i)$ of the system $x(i+1) = \Phi x(i) + w(i)$ with $x(0) = 0$. If Φ is asymptotically stable, F_∞ characterizes the asymptotic behavior of $x(i+1) = \Phi x(i) + w(i)$, see [37] by Kolmanovsky and Gillbert, [5] by Blanchini and the review in Section 2.2.1.

Let $\bar{u}(i|t) = K_f \bar{x}(i|t) + c(i|t)$, it follows that

$$x(i|t) \in \{\bar{x}(i|t)\} \oplus F_i \tag{1.18a}$$

$$u(i|t) \in \{\bar{u}(i|t)\} \oplus K_f F_i \tag{1.18b}$$

Hence,

$$(x(i|t), u(i|t)) \in Y, \; \forall \mathbf{w}(t) \in W^N \Leftrightarrow (\bar{x}(i|t), \bar{u}(i|t)) \in \bar{Y}_i := Y \ominus (F_i \times K_f F_i). \tag{1.19}$$

The terminal constraint can also be handled in the same way,

$$x(N|t) \in X_f, \; \forall \mathbf{w}(t) \in W^N \Leftrightarrow \bar{x}(N|t) \in \bar{X}_f := X_f \ominus F_N. \tag{1.20}$$

Collecting all the design variables within the control horizon in $\mathbf{c}(t) := \{c(0|t), \cdots, c(N-1|t)\}$ and using parametrization (1.15), the FH optimization problem, denoted by $\mathscr{P}_N^{FF}(x(t))$, is

$$\min_{\mathbf{c}(t)} \quad \sum_{i=0}^{N-1} \|c(i|t)\|_\Psi^2 \tag{1.21a}$$

$$\text{s.t.} \quad \bar{x}(0|t) = x(t) \tag{1.21b}$$

$$\bar{x}(i+1|t) = A\bar{x}(i|t) + B\bar{u}(i|t),\ \forall i \in \mathbb{Z}_{N-1} \tag{1.21c}$$

$$\bar{u}(i|t) = K_f\bar{x}(i|t) + c(i|t),\ \forall i \in \mathbb{Z}_{N-1} \tag{1.21d}$$

$$(\bar{x}(i|t), \bar{u}(i|t)) \in \bar{Y}_i,\ \forall i \in \mathbb{Z}_{N-1} \tag{1.21e}$$

$$\bar{x}(N|t) \in \bar{X}_f \tag{1.21f}$$

where $\Psi \succ 0$ and $\|c(i|t)\|_\Psi^2 := c^T(i|t)\Psi c(i|t)$. At time instant t, $\mathscr{P}_N^{FF}(x(t))$ is solved and the optimal solution $\mathbf{c}^*(t) = \{c^*(0|t), \cdots, c^*(N-1|t)\}$ is obtained. The very first term of the optimal solution is applied to the system, yielding the MPC controller,

$$\kappa^{FF}(x(t)) := Kx(t) + c^*(0|t) \tag{1.22}$$

Also let $\mathscr{X}_N^{FF}$ be the admissible set of this approach, i.e.,

$$\mathscr{X}_N^{FF} := \{x |\ \exists \mathbf{c} \text{ such that } \mathscr{P}_N^{FF}(x) \text{ is feasible}\}. \tag{1.23}$$

Feasibility of $\mathscr{P}_N^{FF}(x(t)),\ t \geq 0$ and stability of the closed-loop system is also investigated in [10] and they are summarized in the following property (Lemma 7 and Theorem 8 in [10]).

Property 1.2.1 *Provided that the initial state $x(0) \in \mathscr{X}_N^{FF}$, then under control law $u(t) = \kappa^{FF}(x(t))$ given in (1.22) problem $\mathscr{P}_N^{FF}(x(t))$ is feasible for all $t \geq 0$ and the system (1.10) satisfied the following properties: (i) $(x(t), u(t)) \in Y$ for all $t \geq 0$; (ii) $\lim_{t\to\infty} c^*(0|t) = 0$; (iii) $x(t) \to F_\infty(K_f)$ as t tends to infinity.*

In the above theorem, $F_\infty(K_f)$ set refers to the F_∞ set of the system $x(t+1) = (A + BK_f)x(t) + w(t)$, where $F_\infty := \lim_{i\to\infty} F_i$ with F_i defined in (1.17). Since the system asymptotic behavior described in (iii) of Property 1.2.1 is referred to many times in this thesis, it is defined by the following definition.

Definition 1.2.1 (F_∞ convergence) *A system is said to be $F_\infty(K)$ attractive if the system state converges to $F_\infty(K)$, the minimal disturbance invariant set of the system $x(t+1) = (A+BK)x(t) + w(t)$ asymptotically.*

Time-varying Affine State Feedback Parametrization

The advantages of parametrization (1.15) are the light computational burden of $\mathscr{P}_N^{FF}(x)$ and that the closed-loop system is F_∞ stable. However the pre-chosen feedback gain K_f restricts the expressive ability of the parametrization to some extent. To overcome this restriction, the following time-varying affine state feedback control parametrization has been proposed, see [9] by Bemporad and [38] by Smith,

$$u(i|t) = \sum_{j=0}^{i} L(i,j|t)x(j|t) + g(i|t), \;\; \forall i \in \mathbb{Z}_{N-1} \tag{1.24}$$

where $L(i,j|t),\ j \in \mathbb{Z}_i,\ g(i|t),\ i \in \mathbb{Z}_{N-1}$ are the design variables at time t. Unfortunately, the mapping between design variables and state and control variables is not linear. Therefore, the constraints on design variables are non-convex. This is verified by the following example.

Example 1.2.1 (Example 4 in [39]) *Consider the SISO system $x(t+1) = x(t) + u(t) + w(t)$ with constraint $|u(t)| \leq 3$, $|w(t)| \leq 1$ and initial state $x(t) = 0$. Follow parametrization (1.24) and let $g(i|t) \equiv 0$, $L(2,1|t) = 0$, it can be shown that*

$$u(1|t) = L(1,1|t)w(0|t) \tag{1.25a}$$

$$u(2|t) = L(2,2|t)(1 + L(1,1|t))w(0|t) + L(2,2|t)w(1|t) \tag{1.25b}$$

then the constraints on $u(1|t)$ and $u(2|t)$ is satisfied if and only if

$$|L(1,1|t)| \leq 3 \tag{1.26a}$$

$$|L(2,2|t)(1 + L(1,1|t))| + |L(2,2|t)| \leq 3 \tag{1.26b}$$

It can be verified $(L(1,1|t), L(2,2|t)) = (-3,1)$ and $(L(1,1|t), L(2,2|t)) = (-1,3)$ are feasible, while $(L(1,1|t), L(2,2|t)) = (-2,2)$ is not.

As a consequence of this non-linear mapping, the FH optimization problem under parametrization (1.24) is not computationally tractable. Several approximations [3, 40, 38] have been proposed to simplify the computation and this remains an open research issue.

Time-varying Affine Disturbance Feedback Parametrization

Instead of using time-varying state feedback parametrization (1.24), Löfberg [41] proposed a time-varying disturbance feedback parametrization,

$$u(i|t) = \sum_{j=1}^{i} M(i,j|t)w(i-j|t) + v(i|t), \ \ \forall i \in \mathbb{Z}_{N-1} \tag{1.27}$$

where $M(i,j|t) \in \mathbb{R}^{m\times n}, \ j \in \mathbb{Z}_i^+, \ v(i|t) \in \mathbb{R}^m, \ i \in \mathbb{Z}_{N-1}$ are design variables at time t. It is shown in [42] by Kerrigan and Maciejowksi that under this parametrization, $x(i|t)$ and $u(i|t)$ are affine functions of $M(i,j|t)$ and $v(i|t)$. Hence the FH optimization problem under this parametrization becomes convex and computationally tractable. The relationship between (1.24) and (1.27) was unclear until the work of Goulart *et al.* [39] in 2006. They show that parametrization (1.24) and (1.27) are equivalent in their expressive abilities, and this is summarized in the following property (Theorem 9 in [39]).

Property 1.2.2 *For any $L(i,j|t), \ j \in \mathbb{Z}_i, \ g(i|t), \ i \in \mathbb{Z}_{N-1}$ in (1.24), a set of $M(i,j|t), \ j \in \mathbb{Z}_i^+, \ v(i|t), \ i \in \mathbb{Z}_{N-1}$ in (1.27) can be found that yields the same control sequence for any disturbance sequence and vice-versa.*

Remark 1.2.1 *Property 1.2.2 implies that if an initial state is feasible for the FH op-*

timization problem under time-varying affine state feedback parametrization, it is also feasible under the time-varying affine disturbance feedback parametrization. Hence, both FH optimization problem share the same admissible set.

The cost function used in the FH optimization in [39] is the linear quadratic (LQ) cost function of nominal state and control variables,

$$J_{NLQ}(x(0|t),\{v(i|t)\}_{i=0}^{N-1}) := \sum_{i=0}^{N-1}(\|\bar{x}(i|t)\|_Q^2 + \|\bar{u}(i|t)\|_R^2) + \|\bar{x}(N|t)\|_P^2 \tag{1.28}$$

where $Q \succ 0,\ R \succ 0,\ P \succ 0$, $\bar{x}(0|t) = x(t)$, $\bar{x}(i+1|t) = A\bar{x}(i|t) + B\bar{u}(i|t)$ and $\bar{u}(i|t) = v(i|t)$. Collect all the design variables in $\mathbf{v}(t) := \{v(i|t)\ i \in \mathbb{Z}_{N-1}\}$ and $\mathbf{M}(t) := \{M(i,j|t)\ j \in \mathbb{Z}_i^+\ i \in \mathbb{Z}_{N-1}\}$. Then utilizing parametrization (1.27) and cost function (1.28), the FH optimization problem, referred to hereafter as $\mathscr{P}_N^{DF}(x(t))$, is

$$\begin{aligned}
\min_{\mathbf{v}(t),\mathbf{M}(t)} \quad & J_{NLQ}(x(0|t),\mathbf{v}(t)) && \text{(1.29a)}\\
\text{s.t.} \quad & x(0|t) = x(t) && \text{(1.29b)}\\
& x(i+1|t) = Ax(i|t) + Bu(i|t) + w(i|t),\ \ \forall i \in \mathbb{Z}_{N-1} && \text{(1.29c)}\\
& u(i|t) = \sum_{j=1}^{i} M(i,j|t)w(i-j|t) + v(i|t),\ \ \forall i \in \mathbb{Z}_{N-1} && \text{(1.29d)}\\
& (x(i|t),u(i|t)) \in Y,\ \forall i \in \mathbb{Z}_{N-1},\ \forall\ \mathbf{w}(t) \in W^N && \text{(1.29e)}\\
& x(N|t) \in X_f,\ \forall\ \mathbf{w}(t) \in W^N && \text{(1.29f)}
\end{aligned}$$

Optimization problem $\mathscr{P}_N^{DF}(x(t))$ can be solved using standard techniques of Robust Optimization, see [43] by Ben-Tal *et al.* and [44] by Ben-Tal and Nemirovski. A brief review of Robust Optimization techniques is given in Section 2.3 and details of solving $\mathscr{P}_N^{DF}(x(t))$ are postponed until Chapter 3, see also [39].

Solving optimization problem $\mathscr{P}_N^{DF}(x(t))$ yields the optimizer $(\mathbf{v}^*(t), \mathbf{M}^*(t))$ and the optimal control policy $\mathbf{u}^*(t)$. The very first control action of $\mathbf{u}^*(t)$ is applied to the system, yielding the MPC control law,

$$\kappa^{DF}(x(t)) := v^*(0|t). \tag{1.30}$$

The admissible set of $\mathscr{P}_N^{DF}(x(t))$ is defined in the same manner as $\mathscr{X}_N$ in (1.9),

$$\mathscr{X}_N^{DF} := \{x | \; \exists(\mathbf{M}, \mathbf{v}) \text{ such that } \mathscr{P}_N^{DF}(x) \text{ is feasible}\}. \tag{1.31}$$

It was also proved in [39] that under controller (1.30), the origin is input-to-state stable (ISS) for the closed-loop system under mild assumptions. Before introducing ISS, the following concepts are needed.

Definition 1.2.2 ($\mathscr{K}$-function) *A continuous function $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ is a $\mathscr{K}$-function if it is strictly increasing and $\gamma(0) = 0$; it is a $\mathscr{K}_\infty$-function if, in addition, $\gamma(s) \to \infty$ as $s \to \infty$.*

Definition 1.2.3 ($\mathscr{K}\mathscr{L}$-function) *A continuous function $\beta : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a $\mathscr{K}\mathscr{L}$-function if for all $k \geq 0$, the function $\beta(\cdot, k)$ is a $\mathscr{K}$-function and for each $s \geq 0$,*

$\beta(s,k) \to 0$ *as* $k \to \infty$.

The definition of input-to-state stability is given as follows [39, 45].

Definition 1.2.4 (Input-to-State Stability) *For system $x(t+1) = f(x(t), w(t))$, the origin is input-to-state stable with region of attraction $X \subseteq \mathbb{R}^n$, if there exist a $\mathscr{K}\mathscr{L}$-function $\beta(\cdot)$ and a $\mathscr{K}$-function $\gamma(\cdot)$ such that for all initial states $x(0) \in X$ and disturbance sequences $w(\cdot)$, the system state $x(t)$, for all $t \geq 0$, satisfies*

$$\|x(t)\| \leq \beta(\|x(0)\|, t) + \gamma(\sup\{\|w(\tau)\| | \ \tau \in \mathbb{Z}_{t-1}\}) \tag{1.32}$$

The feasibility and stability of system (1.10) under controller (1.30) is summarized in the following property (Proposition 13 and Theorem 23 in [39]).

Property 1.2.3 *Given $x(0) \in \mathscr{X}_N^{DF}$, $\mathscr{P}_N^{DF}(x(t))$ is feasible for all $t \geq 0$ for system (1.10) under MPC controller (1.30) and the closed-loop system has the following properties: (i) $(x(t), u(t)) \in Y$ for all $t \geq 0$; (ii) the origin is ISS for the closed-loop system.*

1.3 Motivations

Based on the review given in the previous section, a picture of the recent development of MPC for constrained linear systems with additive disturbances, together with its comparison with the Linear Quadratic Regulator (LQR) method, is shown in Figure 1.1.

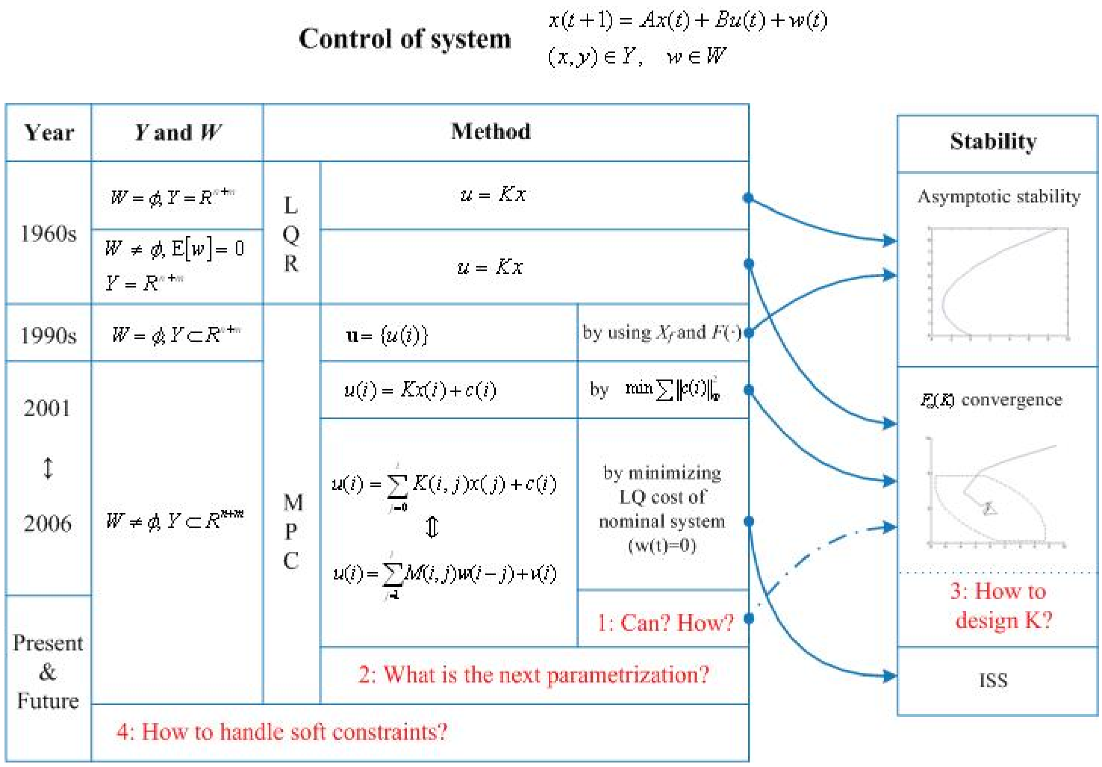

Figure 1.1: Recent development of MPC and comparison with LQR

LQR is one of the earliest optimal control methods for unconstrained linear systems, and the controller $u(t) = Kx(t)$ is obtained by minimizing the infinite horizon LQ cost $\sum_{t=0}^{\infty} \|x(t)\|_Q^2 + \|u(t)\|_R^2$ of the disturbance-free linear system $x(t+1) = Ax(t) + Bu(t)$. It is also known that under the optimal LQ feedback law, the closed-loop system is asymptotically stable [46]. When zero mean additive disturbance is present, controller $u(t) = Kx(t)$ is still optimal [47, 48], but the system state converges to the minimal disturbance invariant set $F_\infty(K)$ [37] in this case.

When no constraint is violated, it is desirable for the MPC controller to achieve the same closed-loop system behavior as the LQR controller. This is true for the MPC controller $\kappa^{FF}(x)$ in (1.22) under the fixed feedback gain parametrization (1.15) if the K_f in (1.15) is chosen to be the optimal LQ feedback gain, see Property 1.2.1. Time-varying state feedback parametrization (1.24) and time-varying disturbance feedback parametrization (1.27) generalize the control parametrization (1.15), and hence improve the MPC controller performance and admit a large admissible set than (1.15). However, a different stability result, ISS, is proved.

Several desirable properties of (1.10) under an MPC control law can be expected from the above discussions. These are listed as P1 to P3 below.

P1: F_∞ convergence for the closed-loop system under MPC feedback law.

P2: A control parametrization that has as general a representative ability as possible.

P3: Ways to influence the shape of F_∞ since it characterize the asymptotic behavior of the closed-loop system.

Properties P1 and P2 are discussed in Chapter 3, 4 and 5. Chapter 6 shows a design procedure for P3.

Besides P1-P3, it is observed that in almost all cases in the MPC literature, constraints are required to be satisfied at all times. This may be too restrictive for some applications. For some cases, it is acceptable that constraints hold at certain confidence levels [23, 49, 50, 51]. Such constraints are best represented by probabilistic constraints. Chapter 7 of this thesis shows a treatment of handling probabilistic constraint under the MPC framework.

1.4 Assumptions

This thesis focuses on the control of following constrained time-invariant linear discrete-time (CTLD) system,

$$x(t+1) = Ax(t) + Bu(t) + w(t) \tag{1.33a}$$

$$(x(t), u(t)) \in Y, \ \forall\, t \geq 0 \tag{1.33b}$$

$$w(t) \in W, \ \forall\, t \geq 0 \tag{1.33c}$$

where $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$ are the state and control of the system at time t, respectively, $w(t) \in \mathbb{R}^n$ is the disturbance on the system at time t and W is compact, and Y is the joint state and control constraint set imposed on the system. For the rest of this thesis, the CTLD system (1.33) is assumed to satisfy the following assumptions.

Assumption 1.4.1

(A1) *$x(t)$ is measurable at every time instant t, the system (1.33a) is stabilizable;*

(A2) *the set*

$$Y := \{(x,u) |\, Y_x x + Y_u u \leq \mathbf{1}_q\} \subset \mathbb{R}^{n+m} \tag{1.34}$$

for some $Y_x \in \mathbb{R}^{q\times n}$ and $Y_u \in \mathbb{R}^{q\times m}$, is compact and contains the origin;

(A3) *the disturbance $w(t)$ is bounded in the sense that*

$$w(t) \in W := \{w|\ Hw \leq \mathbf{1}_r\} \subset \mathbb{R}^n; \tag{1.35}$$

for some $H \in \mathbb{R}^{r\times n}$ and contains the origin;

(A4) *a constant feedback gain $K_f \in \mathbb{R}^{m\times n}$ is given such that $\Phi := A + BK_f$ has a spectral radius $\rho(\Phi) < 1$ and a constraint-admissible d-invariant set X_f in the following form*

$$X_f := \{x|\ Gx \leq \mathbf{1}_g\}, \tag{1.36}$$

for some $G \in \mathbb{R}^{g\times n}$, exists under controller $u(t) = K_f x(t)$ in the sense that $\Phi x + w \in X_f,\ (x, K_f x) \in Y$ for all $x \in X_f$ and for all $w \in W$.

In the above, $\mathbf{1}_k$ is a k-vector with all elements being 1. Assumption (A1) is standard, and using it, $w(t)$ can be obtained at time $t+1$ by $w(t) = x(t+1) - Ax(t) - Bu(t)$. Hence, a control parametrization based on past disturbances is possible. The characterizations of Y in (A2) is made out of the need for computational consideration. Other additional assumptions about $w(t)$ and W are also needed and will be introduced when they are used in the affected chapters. Under (A1)-(A3), results in [37, 52] show that for sufficiently small W, the maximal constraint-admissible disturbance invariant set O_∞ exists. This will be briefly reviewed in Section 2.2.2.

1.5 Organization of the Thesis

The following highlights the contents of each of the remaining chapters.

Chapter 2: This chapter reviews some basic concepts and methodologies needed in this thesis. These concepts include convex sets, operations on convex sets, constraint admissible robust disturbance invariant sets (both the maximal and the minimal) and robust optimization, etc.

Chapter 3: This chapter introduces a control parametrization based on time-varying disturbance feedback. The equivalence of the proposed parametrization to those that have appeared in the literature is established. Using this parametrization, the resulting FH optimization problem is a quadratic programming problem. Additionally, the control law approaches to an a prori chosen linear feedback law and the system state converges to the corresponding F_∞ set asymptotically.

Chapter 4: This chapter considers a simplification of the parametrization used in Chapter 3. This new parametrization does not rely on past realized disturbances. By minimizing a cost function similar to the one used in the Chapter 3, the closed-loop system state is shown to converge to the F_∞ set with probability one under certain assumptions. If a less intuitive cost function is optimized in the FH optimization problem, the convergence to F_∞ is shown to be deterministic.

Chapter 5: This chapter considers an even more general piecewise affine, control parametrization based on disturbance feedback. This parametrization includes affine disturbance feedback parametrization as a special case, and is expected to have the largest admissible set over the parametrizations used in Chapter 3 and 4. The rest of this chapter looks at the computational aspect of the resultant FH optimization problem. Particularly, if the disturbance set has certain properties, the FH optimization problem can be conveniently formulated as a convex optimization problem. The definition of this property and related properties are studied in detail.

Chapter 6: This chapter is concerned with the design of the F_∞ set using support function. For a given feedback gain, both the support function of the minimal disturbance invariant set and its derivative with respect to the feedback gain can be evaluated with arbitrary accuracy. Hence, an optimization problem with constraints imposed on the support functions may be solvable using gradient-based methods. Through an optimization problem, a feedback gain can be found so that the minimal disturbance invariant set satisfies certain constraints. Additional constraints are also introduced to guarantee the existence of the maximal constraints admissible disturbance invariant set.

Chapter 7: In this chapter, the concept of a constraint admissible disturbance invariant set for a linear system with hard constraints and additive disturbances is generalized to the case where the CDTL system has probabilistic constraints and stochastic disturbances. As the most direct extension, a maximal constraint-admissible set for such a system can be defined. However, this set is non-invariant and non-convex in general and this limits its potential applications. An inner approximation of the maximal constraint-admissible set is proposed. This approximate set is convex and invariant under reasonable conditions. Properties and computation of this set is discussed. Its use as a terminal set in an MPC formulation is also introduced.

Chapter 8: This chapter summarizes the contributions of this thesis and outlines directions for future research.

Chapter 2

Review of Related Concepts and Properties

This chapter reviews some basic concepts and properties needed in the subsequent chapters. Definitions of convex sets and related operations are given in Section 2.1. Section 2.2 introduces definitions of two important sets of constrained linear systems with bounded additive disturbances together with their properties. Algorithms for determining or approximating these two sets are also provided. Section 2.3 briefly reviews results of Robust Optimization.

2.1 Convex Sets and Sets Operations

For a thorough review of the definitions and related issues, readers are referred to [53, 37].

2.1.1 Definitions of Convex Sets

Definition 2.1.1 (Convex Set) *A set Ω is convex if for any $\omega_1, \omega_2 \in \Omega$ and a scalar λ with $0 \leq \lambda \leq 1$, $\lambda \omega_1 + (1-\lambda)\omega_2 \in \Omega$.*

A convex set is usually defined by a convex function, the definition of which is given next.

Definition 2.1.2 (Convex Function) *A function $f(\cdot) : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is a convex function, if for any $x_1, x_2 \in X$ the following inequality holds for all $0 \leq \lambda \leq 1$,*

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

Most of convex sets dealt with in this thesis are polytopes which are defined by linear inequalities. The relevant definitions are stated below.

Definition 2.1.3 (Hyperplane) *A hyperplane in $\mathbb{R}^n$ is a set in the form*

$$\{\omega : \ \mu^T \omega = \theta\}$$

for some $\mu \in \mathbb{R}^n$, not all $\mu_i = 0$, $i \in \mathbb{Z}_n^+$ and some $\theta \in \mathbb{R}$.

Definition 2.1.4 (Half Space) *A closed half space in $\mathbb{R}^n$ is a set in the form*

$$\{\omega : \ \mu^T \omega \leq \theta\}$$

where $\mu \in \mathbb{R}^n$, not all $\mu_i = 0$, $i \in \mathbb{Z}_n^+$ and for some $\theta \in \mathbb{R}$.

Definition 2.1.5 (Polyhedron) *A convex set $\Omega \subset \mathbb{R}^n$ taking the form*

$$\Omega := \{\omega : \ A^T \omega \leq b\}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ is called a polyhedron. In this form, a polyhedron Ω is defined as the intersection of a finite number of closed half spaces. Each of the half spaces is define by a column of A and the corresponding element in b.

Definition 2.1.6 (Polytope) *A polytope is a bounded and closed polyhedron.*

2.1.2 Operations on Sets

Definition 2.1.7 (Convex Hull) *For a set of points $\Omega = \{\omega_1, \cdots, \omega_L\} \subset \mathbb{R}^n$, the convex hull of Ω is the smallest convex set that contains Ω and is defined by*

$$\mathrm{CH}(\Omega) := \{\omega = \sum_{i=1}^{L} \lambda_i \omega_i, \ \lambda_i \geq 0, \ \forall i \in \mathbb{Z}_L^+, \ \sum_{i=1}^{L} \lambda_i = 1\}$$

Definition 2.1.8 (Scaling of Set) *Given a set $\Omega \subset \mathbb{R}^n$ and $a \in \mathbb{R}$, a scaling of Ω by a is defined as*

$$a\Omega := \{z|\ z = a\omega,\ \omega \in \Omega\}.$$

Definition 2.1.9 (Linear Mapping of Set) *Given a set $\Omega \subset \mathbb{R}^n$ and $A \in \mathbb{R}^{m\times n}$, a linear mapping of Ω by A is*

$$A\Omega := \{z|\ z = A\omega,\ \omega \in \Omega\}.$$

Definition 2.1.10 (Minkowski Sum) *Given two sets $\Theta \subset \mathbb{R}^n$ and $\Omega \subset \mathbb{R}^n$, the Minkowski sum (vector sum) of Θ and Ω is defined as*

$$\Theta \oplus \Omega := \{z \in \mathbb{R}^n|\ z = z_1 + z_2,\ z_1 \in \Theta,\ z_2 \in \Omega\}.$$

It can be derived that

$$\begin{aligned}\Theta \oplus \Omega &= \{z \in \mathbb{R}^n|\ z - z_2 \in \Theta,\ z_2 \in \Omega\} \\ &= \text{Proj}_z\{(z, z_2)|\ z - z_2 \in \Theta,\ z_2 \in \Omega\}\end{aligned}$$

where Proj_z refers to the projection onto the z space from the (z, z_2) space.

Example 2.1.1 *Let*

$$\Theta = \left\{ z \;\middle|\; \begin{bmatrix} -0.5 & 0 \\ 0.5 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} z \le \mathbf{1}_4 \right\}, \quad \Omega = \left\{ z \;\middle|\; \begin{bmatrix} 5 & 5 \\ 5 & -5 \\ -5 & 5 \\ -5 & -5 \end{bmatrix} z \le \mathbf{1}_4 \right\}$$

Then $\Theta \oplus \Omega$ *is*

$$\Theta \oplus \Omega = \left\{ z \;\middle|\; \begin{bmatrix} -0.5 & 0 \\ 0.5 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} z \le \begin{bmatrix} 1.1 \\ 1.1 \\ 1.2 \\ 1.2 \end{bmatrix}, \begin{bmatrix} 5 & 5 \\ 5 & -5 \\ -5 & 5 \\ -5 & -5 \end{bmatrix} z \le \begin{bmatrix} 16 \\ 16 \\ 16 \\ 16 \end{bmatrix} \right\}$$

and the sets Θ, Ω *and* $\Theta \oplus \Omega$ *are plotted in Figure 2.1*

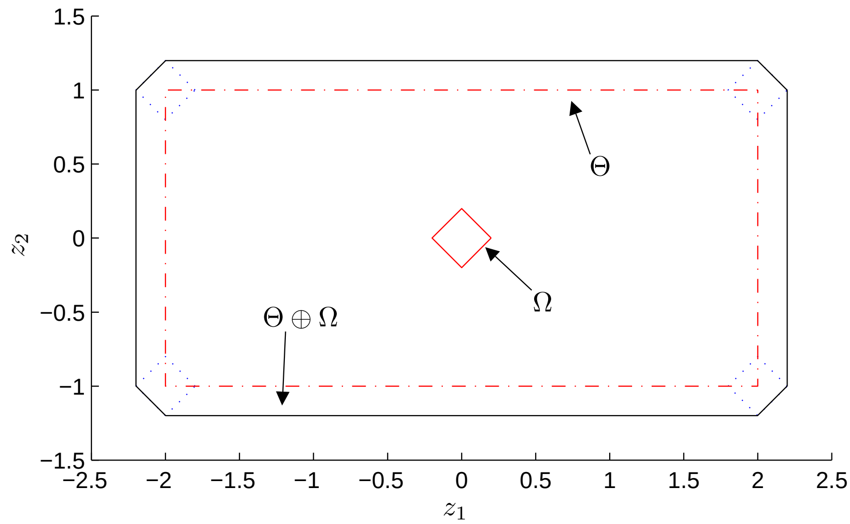

Figure 2.1: Example of Minkowski sum

Definition 2.1.11 (Pontryagin Difference) *Given two sets $\Theta \subset \mathbb{R}^n$ and $\Omega \subset \mathbb{R}^n$, the Pontryagin difference, also known as the Minkowski difference, of Θ and Ω is defined as*

$$\Theta \ominus \Omega := \{z \in \mathbb{R}^n |\, z + z_1 \in \Theta,\ \forall z_1 \in \Omega\}.$$

Definition 2.1.12 (Support Function) *Given a set $\Omega \subset \mathbb{R}^n$, the support function of Ω, evaluated at $y \in \mathbb{R}^n$ is defined as*

$$\delta_\Omega(y) := \sup_{\omega \in \Omega} y^T \omega.$$

Remark 2.1.1 *If Ω is a polytope, the support function of Ω can be computed as a linear programming (LP) problem. This is easy to see from $\delta_\Omega(y) = \max\{y^T w |\, A^T w \leq b\}$*

Property 2.1.1 *Some known properties of support function are: (i) $\delta_{A\Omega}(y) = \delta_\Omega(A^T y)$; (ii) $\delta_{\Omega_1 \oplus \Omega_2}(y) = \delta_{\Omega_1}(y) + \delta_{\Omega_2}(y)$.*

Proof: (i)$\delta_{A\Omega}(y) = \sup_{\omega \in \Omega} y^T (A\omega) = \sup_{\omega \in \Omega} (A^T y)^T \omega = \delta_\Omega(A^T y)$; (ii)$\delta_{\Omega_1 \oplus \Omega_2}(y) = \sup_{\omega_1 \in \Omega_1, \omega_2 \in \Omega_2} y^T (\omega_1 + \omega_2) = \sup_{\omega_1 \in \Omega_1} y^T \omega_1 + \sup_{\omega_2 \in \Omega_2} y^T \omega_2 = \delta_{\Omega_1}(y) + \delta_{\Omega_2}(y)$. ■

For computing Pontryagin difference of two polytopes, the support function operation is used, see the following properties.

Property 2.1.2 (Theorem 2.2 in [37]) *Suppose $\Theta \subset \mathbb{R}^n$ and $\Omega \subset \mathbb{R}^n$ are two polytopes,* *each contains the origin in its interior and Θ is given by*

$$\Theta = \{z|\ (\mu_\Theta^i)^T z \leq 1,\ \forall i \in \mathscr{I}_\Theta\}$$

where $\mu_\Theta^i \in \mathbb{R}^n$ and $\mathscr{I}_\Theta$ is an index set for Θ, then

$$\begin{aligned}
\Theta \ominus \Omega &= \{z|\ z+z_1 \in \Theta,\ \forall z_1 \in \Omega\} \\
&= \{z|\ (\mu_\Theta^i)^T (z+z_1) \leq 1,\ \forall z_1 \in \Omega,\ \forall i \in \mathscr{I}_\Theta\} \\
&= \{z|\ (\mu_\Theta^i)^T z \leq 1 - (\mu_\Theta^i)^T z_1,\ \forall z_1 \in \Omega,\ \forall i \in \mathscr{I}_\Theta\} \\
&= \{z|\ (\mu_\Theta^i)^T z \leq 1 - \max_{z_1 \in \Omega} (\mu_\Theta^i)^T z_1,\ \forall i \in \mathscr{I}_\Theta\} \\
&= \{z|\ (\mu_\Theta^i)^T z \leq 1 - \delta_\Omega(\mu_\Theta^i),\ \forall i \in \mathscr{I}_\Theta\}
\end{aligned}$$

Example 2.1.2 *Let Θ and Ω be the same as those in Example 2.1.1, using the result of Property 2.1.2 $\Theta \ominus \Omega$ is*

$$\Theta \ominus \Omega = \left\{ z \,\middle|\, \begin{bmatrix} -0.5 & 0 \\ 0.5 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} z \leq \begin{bmatrix} 0.9 \\ 0.9 \\ 0.8 \\ 0.8 \end{bmatrix} \right\}.$$

and the sets are plotted in Figure 2.2.

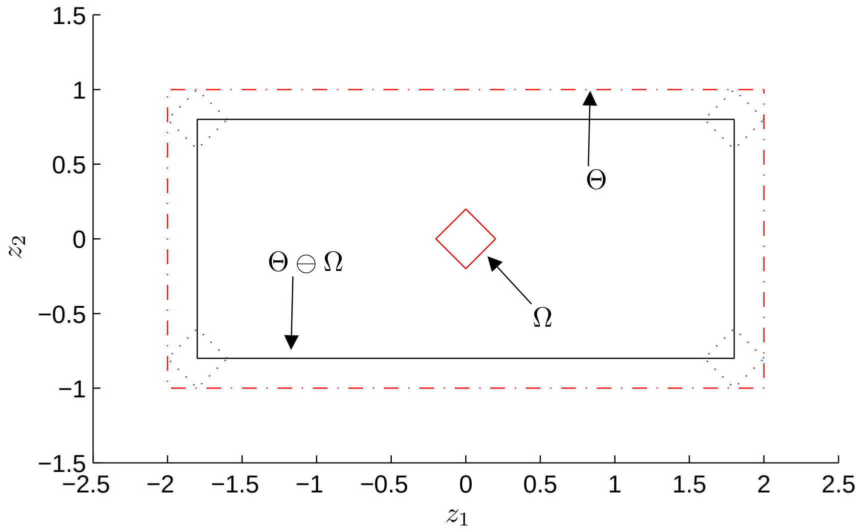

Figure 2.2: Example of Pontryagin difference

2.2 Robust Invariant Sets

The theory of set invariance plays a fundamental role in the control of constrained linear systems and has been a subject attracting much attention, see [5, 37] and references cited therein.

This thesis considers the constraint admissible disturbance invariant sets for linear systems taking the following form

$$x(t+1) = \Phi x(t) + w(t) \tag{2.1a}$$

$$x(t) \in X,\ w(t) \in W,\ \forall t \geq 0 \tag{2.1b}$$

where $x(t) \in \mathbb{R}^n$ is the system state, $w(t) \in \mathbb{R}^n$ is the additive disturbance, X is the state constraint set and W is a bounded disturbance set containing the origin as interior.

2.2.1 Minimal Disturbance Invariant Set

For system (2.1), the definition of a disturbance invariant set is the following,

Definition 2.2.1 (Disturbance Invariant Set) *A set $\mathscr{T}$ is a disturbance invariant set for system (2.1), also known as a d-invariant set, if*

$$\Phi x + w \in \mathscr{T},\ \forall\, w \in W \text{ and } \forall\, x \in \mathscr{T}.$$

One special d-invariant set is the minimal d-invariant set. Due to the existence of the disturbance, the state of (2.1a) does not converge to the origin but to a neighborhood of the origin. The set of all states of (2.1a) reachable at time t, starting from $x(0) = 0$, is

$$F_t := \{x |\ x = \sum_{i=0}^{t-1} \Phi^{(t-1-i)} w(i),\ w(i) \in W,\ i \in \mathbb{Z}_{t-1}\} \tag{2.2}$$

and using notation of Minkowski sum, it can also be written as

$$F_t = W \oplus \Phi W \oplus \cdots \oplus \Phi^{t-1} W \tag{2.3}$$

As t tends to infinity, F_t tends to F_∞ and its properties are summarized below, see Theorem 4.1 and Corollary 4.2 in [37] for more.

Property 2.2.1 *Assume matrix Φ in (2.1a) is asymptotically stable, then there exists a compact set, $F_\infty \in \mathbb{R}^n$, with the following properties:*

(i) *$0 \in F_t \subset F_\infty$;*

(ii) *$F_t \rightarrow F_\infty$ as t tends to infinity;*

(iii) *F_∞ is the smallest d-invariant set, i.e., if $\mathscr{T}$ is any closed d-invariant set, then $F_\infty \subseteq \mathscr{T}$.*

Although the characteristics of F_∞ is clear, unfortunately, there is no method for the exact determination of F_∞ for system (2.1) except for some special cases [37, 54]. In the literature, various methods of determining the outer bounds of F_∞ set have been proposed, see [5] by Blanchini, [55] by Hirata and Ohta, [56] by Raković *et al.* and [57] by Ong and Gilbert. In this thesis, the approach of [57] is used to compute the outer bounds of F_∞ set, and the details, based on the method in [57], are briefly reviewed below.

In the method, the outer approximation takes the form of $\sigma_k F_k$ for some scalar $\sigma_k > 1$ and some index k, with $F_\infty \subseteq \sigma_k F_k$. Since $F_k \rightarrow F_\infty$ as $k \rightarrow \infty$, $\sigma_k F_k$ is an excellent choice for an outer bound of F_∞ in the sense that the accuracy of the approximation increases with increasing k. Specifically, suppose F_k is expressed as

$$F_k = \{x \in \mathbb{R}^n | \ (\mu_{F_k}^j)^T x \leq 1, \ \forall j \in \mathscr{I}_{F_k}\} \tag{2.4}$$

where $\mu_{F_k}^j \in \mathbb{R}^n$ and $\mathscr{I}_{F_k}$ is the index set for the set F_k. The condition $F_\infty \subseteq \sigma_k F_k$ holds if and only if $\delta_{F_\infty}(\mu_{F_k}^j) \leq \sigma_k \delta_{F_k}(\mu_{F_k}^j) = \sigma_k, \forall j \in \mathscr{I}_{F_k}$. Let $\underline{\sigma}_k := \min_{j \in \mathscr{I}_{F_k}} \delta_{F_\infty}(\mu_{F_k}^j)$, then it is easy to see that $\underline{\sigma}_k F_k$ is an outer bound of F_∞. To compute $\underline{\sigma}_k$, note that

$$\delta_{F_\infty}(\mu_{F_k}^j) = \sum_{i=0}^{\infty} \delta_W((\Phi^i)^T \mu_{F_k}^j) \tag{2.5}$$

then it can be computed to a great degree of accuracy by its partial sum $\delta_{F_\infty}^L(\mu_{F_k}^j) = \sum_{i=0}^{L-1} \delta_W((\Phi^i)^T \mu_{F_k}^j)$. The error in the approximation of $\delta_{F_\infty}(\cdot)$ by $\delta_{F_\infty}^L(\cdot)$ can be bounded because there exist a $v > 0$ such that $0 < \delta_W((\Phi^i)^T \mu_{F_k}^j) < (\rho(\Phi))^i v$ for all $j \in \mathscr{I}_{F_k}$. Specifically,

$$v = (\max_{j \in \mathscr{I}_{F_k}} \|\mu_{F_k}^j\|_2) \cdot (\max_{w \in W} \|w\|_2) \cdot M \tag{2.6}$$

where $M(\rho(\Phi))^i \geq \|\Phi^i\|_2$ and $\|\Phi\|_2$ is the induced norm of Φ. Then the error is bounded by $v(1-\rho(\Phi))^{-1}(\rho(\Phi))^L$, and

$$\delta_{F_\infty}(\mu_{F_k}^j) < \underline{\sigma}_k^L := \min_{j \in \mathscr{I}_{F_k}} \delta_{F_\infty}^L(\mu_{F_k}^j) + v(1-\rho(\Phi))^{-1}(\rho(\Phi))^L. \tag{2.7}$$

Therefore, the set $\underline{\sigma}_k^L F_k$ is a tight outer bound of F_∞. The numerical results of computing the outer bounds of F_∞ set are illuminated through the following example.

Example 2.2.1 *Let the parameters in (2.1) be*

$$\Phi = \begin{bmatrix} 0.4 & -0.3 \\ -0.5 & -0.2 \end{bmatrix}, \quad w(t) \in W := \{w \in \mathbb{R}^2 |\ \|w\|_\infty \leq 1\}.$$

Two sets of approximation are determined. In the first set, L in (2.7) is chosen to be 7 *and k varies from 2 to 6. The corresponding $\underline{\sigma}_k^L$ are listed in Table 2.1 and $\underline{\sigma}_k^L F_k$ are plotted in Figure 2.3*

$L=7$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
$\underline{\sigma}_k^L$	1.5631	1.2855	1.1563	1.0950	1.0607

Table 2.1: Optimal scale with $L=7$ and $k=2,\ldots,6$

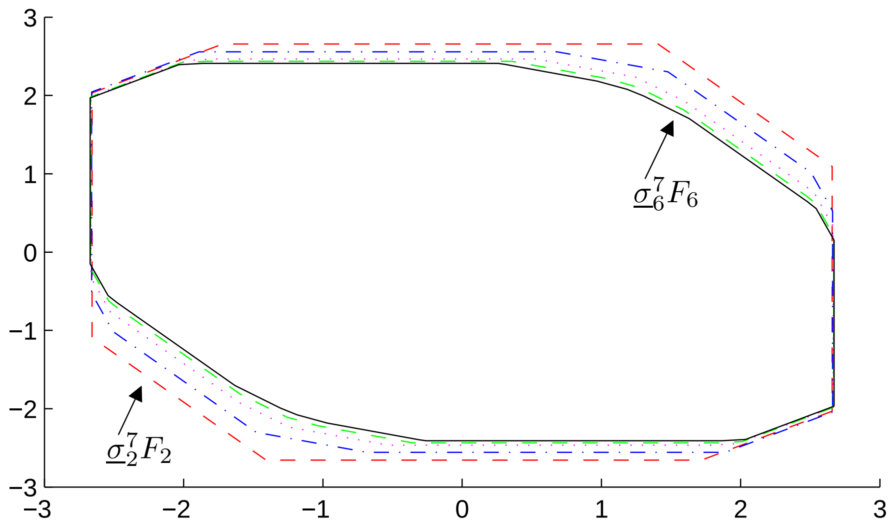

Figure 2.3: Approximation of F_∞ with $L=7$ and $k=2,\ldots,6$

In the second set, k is chosen to be 4 and L in (2.7) varies from 2 to 6. The corresponding $\underline{\sigma}_k^L$ are listed in Table 2.2 and $\underline{\sigma}_k^L F_k$ are plotted in Figure 2.4

$k=4$	$L=2$	$L=3$	$L=4$	$L=5$	$L=6$
$\underline{\sigma}_k^L$	1.4882	1.3355	1.2141	1.1836	1.1667

Table 2.2: Optimal scale with $k=4$ and $L=2,\ldots,6$

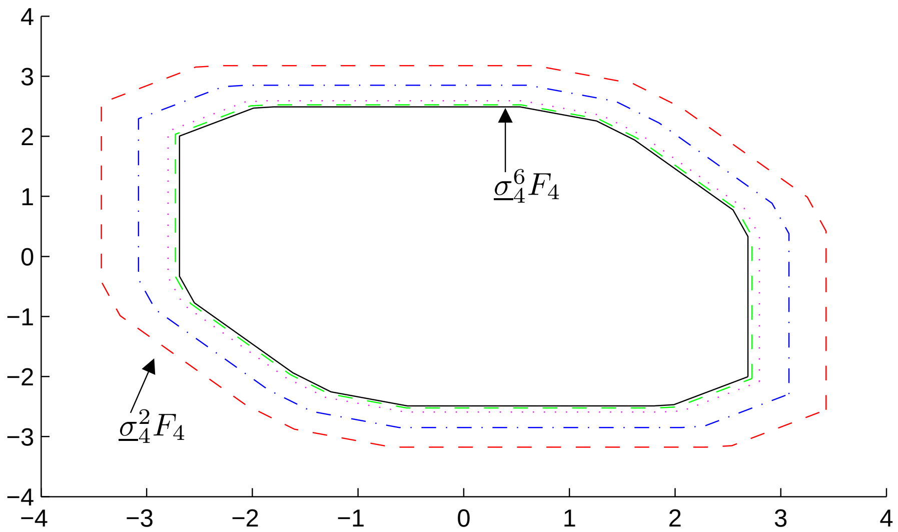

Figure 2.4: Approximation of F_∞ with $k = 4$ and $L = 2,\ldots,6$

2.2.2 Maximal Constraint Admissible Disturbance Invariant Set

This section briefly review another important set for system (2.1), the maximal constraint admissible d-invariant set. This set is usually chosen to be the terminal constraint set X_f in MPC approach and plays a key role in the proof of feasibility of the FH optimization problem and stability of the closed-loop system in the later chapters of this thsis.

Definition 2.2.2 (Constraint Admissible d-invariant Set) *A set $\Gamma \subset \mathbb{R}^n$ is a constraint admissible d-invariant set of system (2.1) if Γ is d-invariant and $\Gamma \subset X$.*

Definition 2.2.3 (Maximal Constraint Admissible d-invariant Set) *The maximal constraint admissible d-invariant set, denoted as O_∞, of system (2.1) is a constraint admissible d-invariant set that contains every closed, constraint admissible d-invariant set of system (2.1).*

Properties and computational issues of the O_∞ set of system (2.1) have already been intensively studied in the literature, see, e.g., [5] and [58] by Blanchini, [59] by Aubin, [60] by Bertsekas, [61] by Gilber and Tan. An algorithm, based on the results in [37] by Kolmanovsky and Gilbert, is given below.

Algorithm 2.2.1 (Determination of O_∞) *Given Φ, X and W in (2.1)*

Step 1: *Let $k = 0$, $O_0 = X$ and $X_0 = X$;*

Step 2: *Determine $X_{k+1} = X_k \ominus \Phi^k W$. If $X_{k+1} = \varnothing$, set $O_\infty = \varnothing$ and stop;*

Step 3: *Determine $O_{k+1} = O_k \cap \{x|\ \Phi^{k+1} x \in X_{k+1}\}$. If $O_{k+1} = \varnothing$, set $O_\infty = \varnothing$ and stop;*

Step 4: *If $O_{k+1} = O_k$, set $O_\infty = O_k$, $k^* = k$ and stop; otherwise let $k = k+1$ and got to Step 2.*

Property 2.2.2 (Theorem 6.3 in [37]) *Assume Φ is asymptotically stable and X_∞ contains origin in its interior, i.e. $0 \in \mathrm{int}(X_\infty)$, then O_∞ is finitely determined, i.e. Algorithm 2.2.1 stops in finite iterations.*

Example 2.2.2 *Consider the system*

$$x(t+1) = Ax(t) + Bu(t) + w(t)$$

$$u(t) = Kx(t)$$

where

$$A = \begin{bmatrix} 1.1 & 1 \\ 0 & 1.3 \end{bmatrix}, \; B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; K = [-0.4991 \;\; -0.9546].$$

and

$$x(t) \in \bar{X} = \{x| \; \|x\|_\infty \leq 5\}, \;\; u(t) \in U = \{u| \; \|u\|_\infty \leq 1\}, \;\; w(t) \in W = \{w| \; \|w\|_\infty \leq 0.2\}$$

This system can be converted to the form of (2.1) with

$$x(t) \in X := \{x| \; \|x\|_\infty \leq 5, \; \|Kx\|_\infty \leq 1\}$$

and $\Phi = A + BK$. *Using Algorithm 2.2.1, the* O_∞ *set, with* $k^* = 2$*, is plotted in Figure 2.5*

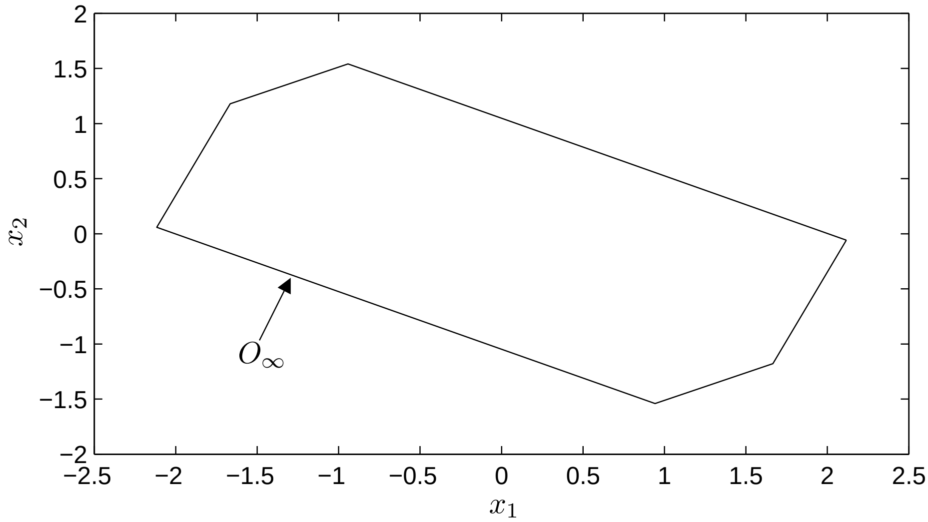

Figure 2.5: O_∞ set of the example system

2.3 Robust Optimization

This section reviews some results in Robust Optimization that are needed to solve the FH optimization problem of MPC in this thesis. For a thorough review of Robust Optimization, readers are referred to [62], [63] and [44] by Ben-Tal and Nemirovski.

2.3.1 Robust Linear Programming

Consider the following optimization problem,

$$\max_{x,y} \quad ax + c^T y \tag{2.8a}$$

$$\text{s.t.} \quad x + y^T w \le b,\ \forall\ w \in \{w : Lw \le l\} \tag{2.8b}$$

where $x \in \mathbb{R}$ and $y \in \mathbb{R}^n$ are decision variables, $w \in \mathbb{R}^n$ is uncertainty factor and (a, b, c, L, l) are appropriate parameters. Constraint (2.8b) can be equivalently written as

$$x + \max_{Lw \le l} y^T w \le b \tag{2.9}$$

Note that the maximum of $y^T w$ in (2.9) equals the minimum of its dual problem,

$$\max_{w} \{y^T w |\ Lw \le l\} = \min_{z} \{l^T z |\ L^T z = y,\ z \ge 0\}. \tag{2.10}$$

where z is the dual decision variable. Hence, constraint (2.9) can be equivalently expressed as

$$x + l^T z \leq b \tag{2.11a}$$

$$L^T z = y \tag{2.11b}$$

$$z \geq 0 \tag{2.11c}$$

Replacing (2.8b) with (2.11), the optimization problem (2.8) is equivalent to

$$\max_{x,y,z} \quad ax + c^T y \tag{2.12a}$$

$$\text{s.t.} \quad x + l^T z \leq b \tag{2.12b}$$

$$L^T z = y \tag{2.12c}$$

$$z \geq 0 \tag{2.12d}$$

and this LP is called the deterministic equivalent of (2.8).

2.4 Notations

Some notations that are to be used frequently are defined in this section.

Integer sets

$$\mathbb{Z}_i := \{0, 1, \cdots, i\} \tag{2.13}$$

$$\mathbb{Z}_i^+ := \{1, \cdots, i\} \tag{2.14}$$

Given two Matrices $A \in \mathbb{R}^{n \times m}$, $B \in \mathbb{R}^{p \times q}$ and a vector $v \in \mathbb{R}^n$. Let $A_i \in \mathbb{R}^n$ denote the ith column of A matrix and Let $a_{i,j}$ denote the (i, j) element of A. Let v_i, $i \in \mathbb{Z}_n^+$ denote the ith element of vector v.

Kronecker product

$$A \otimes B := \begin{bmatrix} a_{1,1}B & \cdots & a_{1,m}B \\ \vdots & \ddots & \vdots \\ a_{n,1}B & \cdots & a_{n,m}B \end{bmatrix} \in \mathbb{R}^{np \times mq} \tag{2.15}$$

Vector operation

$$\mathrm{vec}(A) := \begin{bmatrix} A_1^T & \cdots & A_m^T \end{bmatrix}^T \in \mathbb{R}^{nm} \tag{2.16}$$

Frobenius norm

$$\|A\|_F := \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} a_{i,j}^2} = \|\mathrm{vec}(A)\| \tag{2.17}$$

Absolute value of vector

$$|v| := [|v_1|\ |v_2|\ \cdots\ |v_n|]^T \in \mathbb{R}^n \tag{2.18}$$

Maximization of vectors

Given $v^1, v^2 \in \mathbb{R}^n$, $v = \max\{v^1,\ v^2\}$ means $v_i = \max\{v_i^1,\ v_i^2\}$, $\forall i \in \mathbb{Z}_n^+$.

Chapter 3

Stability of MPC Using Affine Disturbance Feedback Parametrization

In this chapter, an affine disturbance feedback control parametrization for an MPC formulation is proposed. Equivalence of the expressive abilities of the proposed parametrization and the time-varying disturbance feedback parametrization is shown. This leads to the conclusion that the closed-loop systems under both parametrizations share the same admissible set. Furthermore, by minimizing a norm-like cost function, the MPC controller derived under the proposed control law steers the system state to a minimal d-invariant set asymptotically.

This chapter is organized as follows. Section 3.1 proposes the new control parametrization and establishes the equivalence of its expressive ability to that of a time-varying affine disturbance feedback parametrization. The FH optimization and related definitions are also given in this section. The choice of objective function and its connection with the standard linear quadratic cost are given in Section 3.2. Computation of the resulting FH optimization problem is discussed in Section 3.3. Section 3.4 takes on the feasibility and stability issues of the closed-loop system. Numerical examples and summary are contents of the last two sections.

3.1 A New Affine Disturbance Feedback Parametrization

The CTLD system of (1.33), satisfying Assumption 1.4.1, is considered. For convenience, it is repeated here

$$x(t+1) = Ax(t) + Bu(t) + w(t) \tag{3.1a}$$

$$(x(t), u(t)) \in Y,\ w(t) \in W,\ \forall\, t \geq 0 \tag{3.1b}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $w(t) \in \mathbb{R}^n$ are the state, control and disturbance of the system at time t, respectively, W is a bounded disturbance set and Y is the joint state and control constraint imposed on the system.

The choice of $u(i|t)$ in this chapter, motivated by [10, 41, 39], takes the form

$$\begin{cases} u(i|t) = K_f x(i|t) + l(i|t) \\ l(i|t) = c(i|t) + \sum_{j=1}^{N-1} C(i,j|t) w(i-j|t) \end{cases} \quad \forall i \in \mathbb{Z}_{N-1} \tag{3.2}$$

where $l(i|t)$ is an affine function of the $N-1$ disturbances preceding time $t+i$, $C(i,j|t) \in \mathbb{R}^{m\times n}$ is the matrix of coefficients associated with the disturbance at time $t+i-j$ and $w(i-j|t)$ is the realized disturbance $w(t+i-j)$ if $i-j<0$ or an unknown future disturbance at time t if $i-j \geq 0$, see Figure 3.1. K_f is the specified feedback gain in (A4) of Assumption 1.4.1.

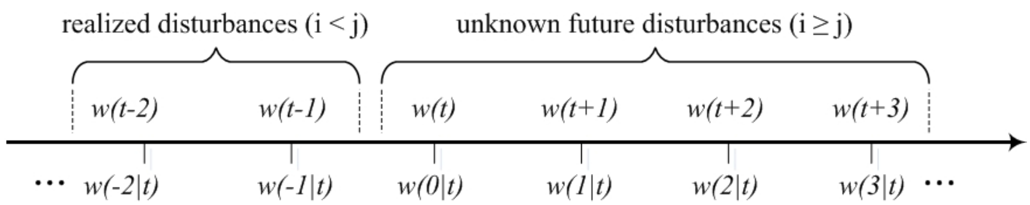

Figure 3.1: Disturbances in the parametrization

Parametrization (3.2) differs from time varying affine disturbance feedback parametrization (1.27), denoted hereafter by u^{DF}, in two ways:

(1) a fixed state feedback term $K_f x(i|t)$ is introduced;

(2) the index j runs from 1 to $N-1$ instead of i, hence realized disturbances are used in the parametrization.

The role of these two changes will become clear in Section 3.4. Collect the past $N-1$

realized disturbances mentioned above in

$$\mathbf{w}^r(t) := \begin{bmatrix} w(-(N-1)|t) \\ w(-(N-2)|t) \\ \vdots \\ w(-1|t) \end{bmatrix} = \begin{bmatrix} w(t-(N-1)) \\ w(t-(N-2)) \\ \vdots \\ w(t-1) \end{bmatrix} \in W^{N-1} \tag{3.3}$$

where W^{N-1} is the $N-1$ times cartesian product spaces of W. Let the predicted control sequence (1.3), predicted state sequence (1.4) and unrealized disturbances (1.14) have the following structure,

$$\mathbf{x}(t) := \begin{bmatrix} x(0|t) \\ x(1|t) \\ \vdots \\ x(N|t) \end{bmatrix}, \ \mathbf{u}(t) := \begin{bmatrix} u(0|t) \\ u(1|t) \\ \vdots \\ u(N-1|t) \end{bmatrix}, \text{ and } \mathbf{w}(t) := \begin{bmatrix} w(0|t) \\ w(1|t) \\ \vdots \\ w(N-1|t) \end{bmatrix}. \tag{3.4}$$

The rest of the variables within the control horizon in (3.2) can be collected in matrices $\mathbf{C}(t) \in \mathbb{R}^{Nm\times(2N-1)n}$ and $\mathbf{c}(t) \in \mathbb{R}^{Nm}$ in the following form

$$\mathbf{C}^-(t) := \begin{bmatrix} C(0,N-1|t) & C(0,N-2|t) & \cdots & C(0,2|t) & C(0,1|t) \\ 0 & C(1,N-1|t) & \cdots & C(1,3|t) & C(1,2|t) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & C(N-2,N-1|t) \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \tag{3.5}$$

$$\mathbf{C}^{+}(t) := \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ C(1,1|t) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C(N-2,N-2|t) & C(N-2,N-3|t) & \cdots & 0 & 0 \\ C(N-1,N-1|t) & C(N-1,N-2|t) & \cdots & C(N-1,1|t) & 0 \end{bmatrix}, \tag{3.6}$$

$$\mathbf{c}(t) := \begin{bmatrix} c(0|t) \\ c(1|t) \\ \vdots \\ c(N-1|t) \end{bmatrix}, \text{ and } \mathbf{C}(t) := [\mathbf{C}^{-}(t)\ \mathbf{C}^{+}(t)]. \tag{3.7}$$

Clearly, $\mathbf{C}^{-}(t)$ is the collection of the coefficient associated with the past disturbance $\mathbf{w}^{r}(t)$ and $\mathbf{C}^{+}(t)$ is the collection of the coefficient associated with the future disturbance $\mathbf{w}(t)$

Using these notations, the FH optimization using the control parametrization of (3.2), referred hereafter as $\mathscr{P}_N^{DFC}(x(t), \mathbf{w}^{r}(t))$, can be written as

$$\min_{\mathbf{c}(t),\mathbf{C}(t)} \quad J_{DFC}(\mathbf{c}(t), \mathbf{C}(t))$$

$$\text{s.t.} \quad \mathbf{x}(t) = \mathscr{A}x(t) + \mathscr{B}\mathbf{u}(t) + \mathscr{G}\mathbf{w}(t) \tag{3.8a}$$

$$\mathbf{u}(t) = \mathscr{K}\mathbf{x}(t) + \mathbf{c}(t) + \mathbf{C}^{-}(t)\mathbf{w}^{r}(t) + \mathbf{C}^{+}(t)\mathbf{w}(t) \tag{3.8b}$$

$$(x(i|t), u(i|t)) \in Y, \quad \forall\, \mathbf{w}(t) \in W^N,\ \forall i \in \mathbb{Z}_{N-1} \tag{3.8c}$$

$$x(N|t) \in X_f, \qquad \forall\, \mathbf{w}(t) \in W^N \tag{3.8d}$$

where X_f is the constraint-admissible d-invariant set as given in Assumption 1.4.1 (A4); $J_{DFC}(\mathbf{c}(t), \mathbf{C}(t))$ is an appropriate cost function details of which are discussed in Section 3.2 and

$$\mathscr{A} := \begin{bmatrix} I_n \\ A \\ A^2 \\ \vdots \\ A^N \end{bmatrix}, \ \mathscr{B} := \begin{bmatrix} 0 & 0 & \cdots & 0 \\ B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix}, \tag{3.9}$$

$$\mathscr{G} := \begin{bmatrix} 0 & 0 & \cdots & 0 \\ I_n & 0 & \cdots & 0 \\ A & I_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1} & A^{N-2} & \cdots & I_n \end{bmatrix}, \mathscr{K} := [I_N \otimes K_f \ \ \mathbf{0}], \tag{3.10}$$

$\otimes$ is Kronecker product of matrices defined in (2.15).

Let the feasible set of the FH optimization problem $\mathscr{P}_N^{DFC}(x, \mathbf{w}^r)$ be

$$\Pi_N^{DFC}(x, \mathbf{w}^r) := \{(\mathbf{c}, \mathbf{C}) | \ (\mathbf{c}, \mathbf{C}) \text{ is feasible for } \mathscr{P}_N^{DFC}(x, \mathbf{w}^r)\} \tag{3.11}$$

and the admissible set of $\mathscr{P}_N^{DFC}(x, \mathbf{w}^r)$ be defined as

$$\mathscr{X}_N^{DFC} := \{x | \ \Pi_N(x, \mathbf{w}^r) \neq \emptyset, \ \forall \ \mathbf{w}^r \in W^{N-1}\}. \tag{3.12}$$

The rest of the MPC formulation is standard: $\mathscr{P}_N^{DFC}(x(t), \mathbf{w}^r(t))$ is solved at each time t, yielding the optimal solution

$$(\mathbf{c}^*(t), \mathbf{C}^*(t)) := \arg\min \ \mathscr{P}_N^{DFC}(x(t), \mathbf{w}^r(t)) \tag{3.13}$$

and the very first term of the corresponding optimal control sequence is applied to system (3.1a). Hence, the MPC control law is

$$u(t) = \kappa^{DFC}(x(t), \mathbf{w}^r(t)) := K_f x(t) + c^*(0|t) + \sum_{j=1}^{N-1} C^*(0, j|t) w(t-j) \tag{3.14}$$

Although the information of $\mathbf{w}^r(t)$ is already captured in $x(t)$, it is added for stability consideration, see Theorem 3.4.2.

The remaining part of this section shows the equivalence of the expressive abilities of parametrization (3.2) and the time varying disturbance feedback parametrization u^{DF} in (1.27). Similar to (3.8b), the control sequence $\mathbf{u}(t)$ using parametrization (1.27), as proposed in [41, 39], can be written as

$$\mathbf{u}(t) = \mathbf{v}(t) + \mathbf{M}(t)\mathbf{w}(t) \tag{3.15}$$

where

$$\mathbf{v}(t) = \begin{bmatrix} v(0|t) \\ v(1|t) \\ \vdots \\ v(N-1|t) \end{bmatrix}, \mathbf{M}(t) = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ M(1,1|t) & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ M(N-1,N-1|t) & \cdots & M(N-1,1|t) & 0 \end{bmatrix}$$

are the variables of $\mathbf{u}(t)$. Under parametrization of (3.15), the admissible set of the FH optimization is $\mathscr{X}_N^{DF}$, see equation (1.31). Equivalence of $\mathscr{X}_N^{DF}$ and $\mathscr{X}_N^{DFC}$ is summarized in the following theorem.

Theorem 3.1.1 *Suppose $x(t) \in \mathbb{R}^n$ and a realization $\mathbf{w}^r(t)$ are given. Then, (i) for any $(\mathbf{c}(t), \mathbf{C}(t))$ and $\mathbf{w}(t)$ sequence that define $\mathbf{u}(t)$ in (3.8b) and $\mathbf{x}(t)$ in (3.8a), there exists a unique $(\mathbf{M}(t), \mathbf{v}(t))$ that yields the same $\mathbf{u}(t)$ and $\mathbf{x}(t)$ sequences. (ii) for any choice of $(\mathbf{M}(t), \mathbf{v}(t))$ that defines $\mathbf{u}(t)$ in (3.15) and $\mathbf{x}(t)$ in (3.8a), there exists at least one choice of $(\mathbf{c}(t), \mathbf{C}(t))$ of (3.8b) which yields the same $\mathbf{u}(t)$ and $\mathbf{x}(t)$ sequences for all $\mathbf{w}(t)$ sequence. Hence $\mathscr{X}_N^{DFC} = \mathscr{X}_N^{DF}$.*

Proof: See Appendix 3.A.1. ∎

Remark 3.1.1 *As $\mathscr{X}_N^{DFC} = \mathscr{X}_N^{DF}$, it may appear that the variable $\mathbf{C}^-$ is superfluous. Its inclusion is needed in ensuring stability of the closed-loop system and will become obvious in Section 3.4. See also Remark 3.4.1.*

A less obvious result regarding properties of $\mathscr{X}_N^{DFC}$ follows from Theorem 3.1.1 and is stated next.

Lemma 3.1.1 *For any pair of* $(\hat{\mathbf{w}}^r, \tilde{\mathbf{w}}^r) \in W^{N-1} \times W^{N-1}$, $\Pi_N^{DFC}(x, \tilde{\mathbf{w}}^r) \neq \emptyset$ *implies* $\Pi_N^{DFC}(x, \hat{\mathbf{w}}^r) \neq \emptyset$ *and vice versa.*

Proof: See Appendix 3.A.2. ■

An immediate consequence of Lemma 3.1.1 and Assumption 1.4.1 (A3) is that $\mathscr{X}_N^{DFC}$ of (3.12) can be equivalently stated as

$$\mathscr{X}_N^{DFC} := \{x|\ \Pi_N(x, \mathbf{0}) \neq \varnothing\}. \tag{3.16}$$

Remark 3.1.2 *Characterization (3.16) allows a simple verification of the condition* $x \in \mathscr{X}_N^{DFC}$. *Specifically,* $x \in \mathscr{X}_N^{DFC}$ *if and only if* $\mathscr{P}_N^{DFC}(x, \mathbf{0})$ *admits a feasible solution.*

3.2 Choice of Cost Function

To achieve the desired stability result, the cost function of $\mathscr{P}_N^{DFC}(x(t), \mathbf{w}^r(t))$ is chosen to be

$$J_{DFC}(\mathbf{c}(t), \mathbf{C}(t)) := \sum_{i=0}^{N-1} h(l(i|t)) := \sum_{i=0}^{N-1} \left[\|c(i|t)\|_\Psi^2 + \sum_{j=1}^{N-1} \|\text{vec}(C(i,j|t))\|_\Lambda^2 \right] \tag{3.17}$$

for some $\Psi \succ 0$ and $\Lambda \succ 0$ and $\text{vec}(C)$ is the vector operation of C defined in (2.16). This choice of cost function is motivated from consideration of the standard LQ cost and hence preserve the use of the 2 norm. It is possible to show, with additional notations, that the results of Theorem 3.4.2 remain true if the $1,\infty$ norm or any norm function of $l(i|t)$ is used for $h(\cdot)$ in (3.17). The choices of Ψ and Λ can be arbitrary so long as they are positive definite.

It is of interest to note that (3.17) can be related to the standard infinite horizon LQ cost if stronger assumptions are made. While not needed for stability consideration, these assumptions are:

Assumption 3.2.1 *Disturbance $w(t)$, cost weight Q, R, P and terminal feedback gain K_f are assumed to satisfy*

(A5) *$w(t)$ is a random vector, uncorrelated from instant to instant, has zero mean and covariance matrix Σ_w, i.e.*

$$\mathrm{E}\left[w(t)\right] = 0, \;\; \mathrm{E}\left[w^T(t)w(t)\right] = \Sigma_w; \tag{3.18}$$

(A6) *Suppose $Q \succeq 0$, $R \succ 0$ are given and $(Q^{\frac{1}{2}}, A)$ is detectable. Let $P = A^T PA - A^T PB(R+B^T PB)^{-1}B^T PA + Q$, the solution of the algebraic Riccati equation, $K_f = -(R+B^T PB)^{-1}B^T PA$, $\Psi = R+B^T PB$ and $\Lambda = \Sigma_w \otimes \Psi$.*

Consider the expected LQ cost of

$$J_{ELQ}(\mathbf{x}(t),\mathbf{u}(t)) := \mathrm{E}_{(\mathbf{w}^r(t),\mathbf{w}(t))}[\sum_{i=0}^{N-1}(\|x(i|t)\|_Q^2+\|u(i|t)\|_R^2)+\|x(N|t)\|_P^2]$$

where $\|x(i|t)\|_Q^2+\|u(i|t)\|_R^2$ and $\|x(N|t)\|_P^2$ are the stage and terminal costs respectively. Theorem 11.2 of [47] shows that under Assumption 3.2.1 (A6)

$$\begin{aligned}
&\sum_{i=0}^{N-1}(x^T(i|t)Qx(i|t)+u^T(i|t)Ru(i|t))+x^T(N|t)Px(N|t) \\
= \;& x^T(0|t)Px(0|t)+\sum_{i=0}^{N-1}\|u(i|t)-K_f x(i|t)\|_{(R+B^TPB)}^2+\sum_{i=0}^{N-1}w^T(i|t)Pw(i|t) \\
&+\sum_{i=0}^{N-1}\left[(Ax(i|t)+Bu(i|t))^T Pw(i|t)+w^T(i|t)P(Ax(i|t)+Bu(i|t))\right]. \qquad (3.19)
\end{aligned}$$

Following control parametrization (3.2), the terms $x(i|t)$ and $u(i|t)$ on the right hand side of the preceding equation are linear functions of past disturbances $w(j|t), j<i$. Taking the expected value over $(\mathbf{w}^r(t),\mathbf{w}(t))$, the last term on the right hand side of (3.19) vanishes following Assumption 3.2.1 (A5). In addition, the first and third terms of the right hand side are constants. This yields

$$J_{ELQ}(\mathbf{x}(t),\mathbf{u}(t)) := x^T(0|t)Px(0|t)+N\cdot\mathrm{trace}(\Sigma_w P)+\mathrm{E}_{(\mathbf{w}^r(t),\mathbf{w}(t))}(\sum_{i=0}^{N-1}l^T(i|t)\Theta l(i|t)) \qquad (3.20)$$

where $\Theta = R + B^T PB$. The last term of (3.20) is exactly $J_{DFC}(\mathbf{c}(t), \mathbf{C}(t))$ in (3.17) if

$$\Psi = \Theta, \;\; \Lambda = \Sigma_w \otimes \Theta. \tag{3.21}$$

This can be seen from

$$\begin{aligned}
&\sum_{i=0}^{N-1} \mathrm{E}_{(\mathbf{w}^r(t), \mathbf{w}(t))} \left[(c(i|t) + \sum_{j=1}^{N-1} C(i,j|t) w(i-j|t))^T \Theta (c(i|t) + \sum_{j=1}^{N-1} C(i,j|t) w(i-j|t)) \right] \\
=& \sum_{i=0}^{N-1} \left[c^T(i|t) \Theta c(i|t) + \sum_{j=1}^{N-1} \mathrm{trace} \left[\Sigma_w (C^T(i,j|t) \Theta C(i,j|t)) \right] \right] \\
=& \sum_{i=0}^{N-1} \left[c^T(i|t) \Theta c(i|t) + \sum_{j=1}^{N-1} \mathrm{vec}(C(i,j|t))^T (\Sigma_w \otimes \Theta) \mathrm{vec}(C(i,j|t)) \right]
\end{aligned} \tag{3.22}$$

The last line results from $\mathrm{E}\left[w^T X w\right] = \mathrm{trace}(X\Sigma_w) = \mathrm{trace}(\Sigma_w X) = \mathrm{vec}(X^T)^T \mathrm{vec}(\Sigma_w)$ and $\mathrm{vec}(AXB) = (B^T \otimes A)\mathrm{vec}(X)$. Θ is positive definite since R and P are positive definite. $\Sigma_w \otimes \Theta$ is positive (semi)definite since Kronecker product of two positive (semi)definite matrices is also positive (semi)definite according to Theorem 4.2.12 of [64].

As the first two terms of (3.20) are independent of $(\mathbf{c}(t), \mathbf{C}(t))$, the minimization of $J_{DFC}(\mathbf{c}(t), \mathbf{C}(t))$ in (3.17) is equivalent to the minimization of expected LQ cost (3.20).

The above equations of (3.20)-(3.22) follows a similar development in [65, 66, 67] by Goulart and Kerrigan, but is adapted for the proposed parametrization (3.2).

3.3 Computation of the FH Optimization Problem

This section discusses the computation of the the FH optimization problem $\mathscr{P}_N^{DFC}(x(t),\mathbf{w}^r(t))$. Inequalities (3.8c) and (3.8d) can be restated, using characterization of Y and X_f given in Assumption 1.4.1, as

$$\begin{bmatrix} \bar{Y}_x & 0 & \bar{Y}_u \\ 0 & G & 0 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}(t) \\ x(N|t) \\ \mathbf{u}(t) \end{bmatrix} \leq 1_{Nq+g}, \ \ \forall \, \mathbf{w}(t) \in W^N \tag{3.23}$$

where $\bar{Y}_x := I_N \otimes Y_x$, $\bar{Y}_u := I_N \otimes Y_u$, $\tilde{\mathbf{x}}(t) := [x^T(0|t) \ \cdots \ x^T(N-1|t)]^T$ and the dependence of $\mathbf{x}(t), \mathbf{u}(t)$ on $\mathbf{w}(t)$ are shown explicitly. Using expressions of $\mathbf{x}(t)$ and $\mathbf{u}(t)$ from (3.29), (3.23) can be written as

$$\bar{\mathscr{A}} x(t) + \bar{\mathscr{B}} \mathbf{c}(t) + \bar{\mathscr{F}} \mathrm{vec}(\mathbf{C}^-(t)) + \max_{\mathbf{w}(t) \in W^N} \left[\bar{\mathscr{B}} \mathbf{C}^+(t) + \bar{\mathscr{G}} \right] \mathbf{w}(t) \leq 1_s \tag{3.24}$$

where $s := Nq + g$, the max operator is taken element-wise and

$$Y := \begin{bmatrix} \bar{Y}_x & \mathbf{0} & \bar{Y}_u \\ \mathbf{0} & G & \mathbf{0} \end{bmatrix}, \ \ \bar{\mathscr{A}} := Y \begin{bmatrix} \mathscr{A}_x \\ \mathscr{A}_u \end{bmatrix}, \ \ \bar{\mathscr{B}} := Y \begin{bmatrix} \mathscr{B}_x \\ \mathscr{B}_u \end{bmatrix} \tag{3.25}$$

$$\bar{\mathscr{G}} := Y \begin{bmatrix} \mathscr{G}_x \\ \mathscr{G}_u \end{bmatrix}, \ \ \bar{\mathscr{F}} := Y((\mathbf{w}^r(t))^T \otimes \begin{bmatrix} \mathscr{B}_x \\ \mathscr{B}_u \end{bmatrix}) \tag{3.26}$$

Using the expression of (1.35), $W^N = \{\mathbf{w} | \bar{H}\mathbf{w} \leq 1_\ell, \bar{H} := I_N \otimes H, \ell := Nr\}$. Let μ_i be the i^{th} row of $(\bar{\mathscr{B}}\mathbf{C}^+(t) + \bar{\mathscr{G}})$, then each row i of the maximization in (3.24) can be related to a linear programming problem in $\mathbf{w}(t)$,

$$\max_{\mathbf{w}(t)}\{\mu_i \mathbf{w}(t) |\ \bar{H}\mathbf{w}(t) \leq 1_\ell\} \tag{3.27}$$

and following the standard procedure in robust optimization reviewed in Section 2.3.1 (see [62, 63] for more details),

$$\max_{\mathbf{w}(t)}\{\mu_i \mathbf{w}(t) | \bar{H}\mathbf{w}(t) \leq 1_\ell\} = \min_{z_i}\{1_\ell^T z_i | \bar{H}^T z_i = \mu_i^T, z_i \geq 0\}$$

where $z_i \in \mathbb{R}^\ell$ is the dual variable of the primal LP. Let $Z := [z_1 \cdots z_s] \in R^{\ell \times s}$. By standard duality results given in Section 2.3.1, constraint (3.23) can be written as a set of linear inequalities in $\mathbf{c}(t), \mathbf{C}(t)$ and Z as follows,

$$\begin{cases} \bar{\mathscr{A}} x(t) + \bar{\mathscr{B}} \mathbf{c}(t) + \bar{\mathscr{F}} vec(\mathbf{C}^-(t)) + Z^T \cdot 1_\ell \leq 1_s, \\ Z^T \bar{H} = \bar{\mathscr{B}} \mathbf{C}^+(t) + \bar{\mathscr{G}}, \\ z_i \geq 0,\ i \in \{1, \cdots, s\} \end{cases} \tag{3.28}$$

It is known that $\mathscr{P}_N^{DFC}(x(t), \mathbf{w}^r(t))$ is feasible if and only if (3.28) is feasible. With these results, the computation of the solution to $\mathscr{P}_N^{DFC}(x(t), \mathbf{w}^r(t))$ corresponds to solving a convex quadratic programming problem with convex quadratic cost function (3.17) and linear constraints (3.28) in $\mathbf{c}(t), \mathbf{C}(t)$ and Z.

3.4 Feasibility and Stability of the Closed-Loop System

The feasibility of $\mathscr{P}_N^{DFC}(x(t), \mathbf{w}^r(t))$ and stability of the closed-loop system under the feedback law $\kappa^{DFC}(\cdot,\cdot)$ of (3.14) are addressed in this section.

Theorem 3.4.1 *If $\mathscr{P}_N^{DFC}(x(t), \mathbf{w}^r(t))$ admits an optimal solution, so does the FH optimization problem $\mathscr{P}_N^{DFC}(x(t+1), \mathbf{w}^r(t+1))$ under the feedback law (3.14).*

Proof: See Appendix 3.A.3. ■

Theorem 3.4.2 *Suppose $x(0) \in \mathscr{X}_N^{DFC}$ and Assumption 1.4.1 is satisfied. System (3.1a) under MPC control law (3.14) has the following properties: (i) $(x(t), u(t)) \in Y$ for all $t \geq 0$, (ii) $\lim_{t\to\infty} l(t) = 0$ element-wise, (iii) $x(t) \to F_\infty(K_f)$ as $t \to \infty$, (iv) If (A5) of Assumption 3.2.1 is satisfied, $\lim_{t\to\infty} \mathrm{E}\left[x(t)x(t)^T\right] = \Sigma_\infty$, where $\Sigma_\infty = \Sigma_\infty^T$, $\Phi\Sigma_\infty\Phi^T = \Sigma_\infty - \Sigma_w$ and $\Phi = A + BK_f$.*

Proof: See Appendix 3.A.4. ■

Remark 3.4.1 *The presence of $\mathbf{C}^-(t)$ in (3.2) is important in Theorems 3.4.1 and 3.4.2. Specifically, $(\hat{\mathbf{c}}(t+1), \hat{\mathbf{C}}(t+1))$ of (3.31) may not be feasible if $\mathbf{w}^r$ is not included in the parametrization of (3.2), see also Remark 3.A.1. Other feasible $(\mathbf{c}, \mathbf{C})$ exist and they are discussed in Chapter 4.*

It is interesting to note, from Theorems 3.1.1 and 3.4.2, that the asymptotic behavior of the closed-loop system is determined only by the terminal feedback gain K_f. Hence K_f

offers freedom to control the system asymptotic behavior. One method of designing K_f to control the system asymptotic behavior is introduced in Chapter 6.

3.5 Numerical Examples

In this section, the performance of the proposed MPC control law is illustrated on an example system with $n = 2, m = 1$. The system parameters and constraints are:

$$A = \begin{bmatrix} 1.1 & 1 \\ 0 & 1.3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad W = \{w| \; \|w\|_\infty \le 0.2\}$$

$$Y = \{(x,u)| \; |u| \le 1, \; \|x\|_\infty \le 10\}, \; K_f = [-0.4991 \; -0.9546]$$

where K_f is the LQR feedback gain with

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad R = 1, \quad P = \begin{bmatrix} 2.6093 & 0.21 \\ 0.21 & 2.1837 \end{bmatrix}$$

Terminal set X_f is the corresponding maximal constraint-admissible disturbance invariant set of (3.1a) under $u(t) = K_f x(t)$.

In the first simulation, $w(t)$ is uniformly distributed over the W set. The proposed approach is simulated with Ψ, Λ chosen according to (3.21), $x(0) = [-3.87 \; 2.18]^T$ and $N = 5$. The simulation results are shown in Figure 3.2 and 3.3. In Figure 3.2, X_f and an outer bound of F_∞, $\hat{F}_\infty$, are plotted. The outer bound $\hat{F}_\infty$ is used because the exact F_∞ is not computable. The procedure for computing $\hat{F}_\infty$ follows that given in Section 2.2.1. It is clear from these figures that $x(t)$ stays within $\hat{F}_\infty$ and $l(t) \to 0$ as $t \to \infty$ and that the constraints are satisfied at all time.

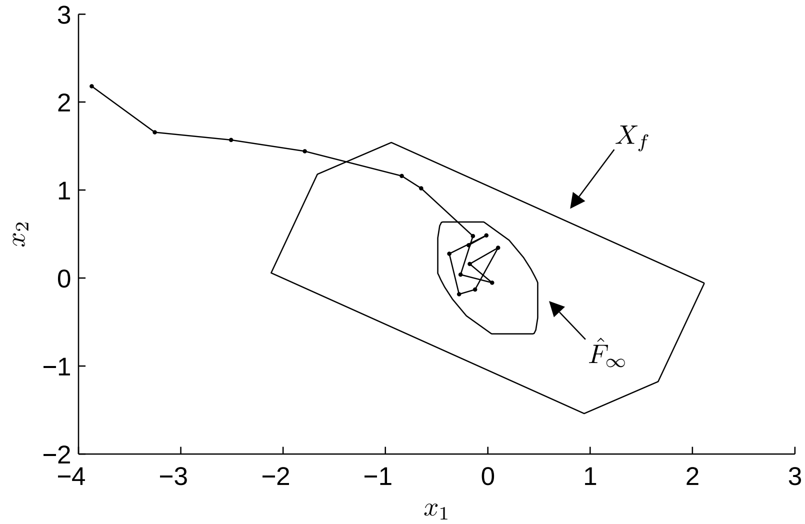

Figure 3.2: State trajectory of the first simulation

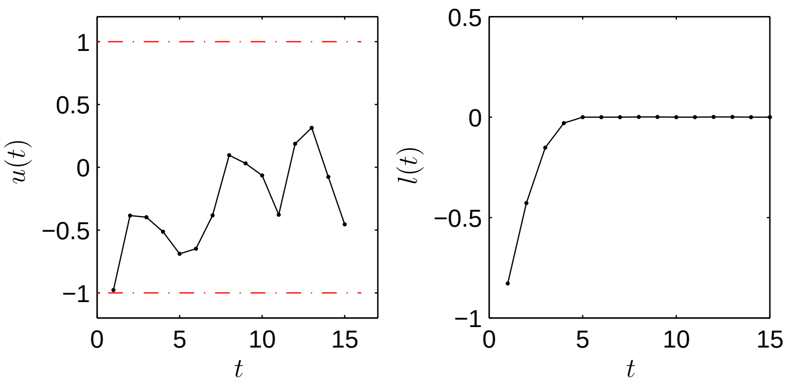

Figure 3.3: Control trajectory of the first simulation

The second simulation attempts to show the difference between the stability result of this work and ISS proved in [39]. In order to highlight the difference, $w(t)$ is assumed to be

$$w(t) = \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix} - (0.9)^t \times 0.4 \times \begin{bmatrix} \theta^1(t) \\ \theta^2(t) \end{bmatrix}$$

where $\theta^1(t)$ and $\theta^2(t)$ are uniformly distributed random variables between 0 and 1. And two sets of parameters in the cost functions of two approaches are used. The two sets used for (3.17) are $S_1 = \{\Psi, \Lambda\}$ (the one used in the first simulation) and $S_2 = \{\Psi = 1,\ \Lambda = I_2\}$. The two sets used for the cost function in [39] are $G_1 = \{Q, R, P\}$ (the one used in the first simulation) and $G_2 = \{0.01 \times Q, R, P\}$ which satisfy the assumption (Assumption 2) in [39]. Performances of the various cases are simulated with $x(0) = [-1.5\ 1.5]^T$, $N = 5$ and the same disturbance realization. The results are shown in Figure 3.4 and 3.5. It can be observed that the asymptotic behavior of the system using the proposed approach depends on the choice of K_f only, but that of the other approach is also affected by the choice of the parameters of objective function.

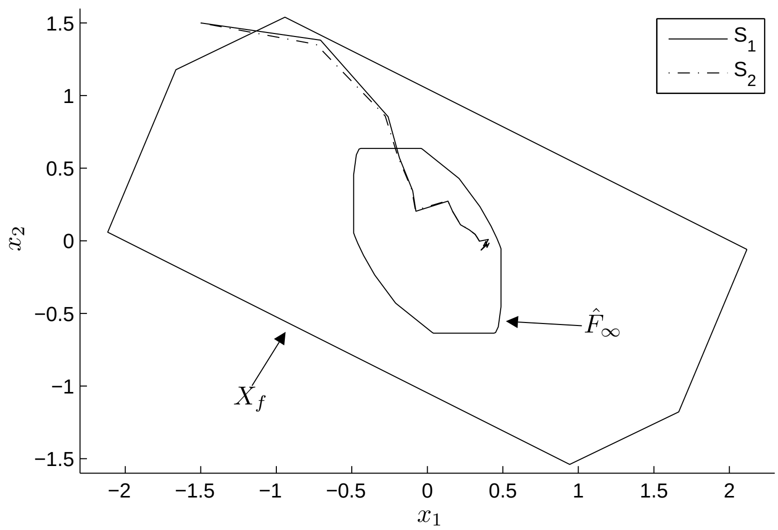

Figure 3.4: State trajectories of the proposed approach

The next simulation compares the sizes of $\mathscr{X}_N^{DFC}$ for the proposed approach and that

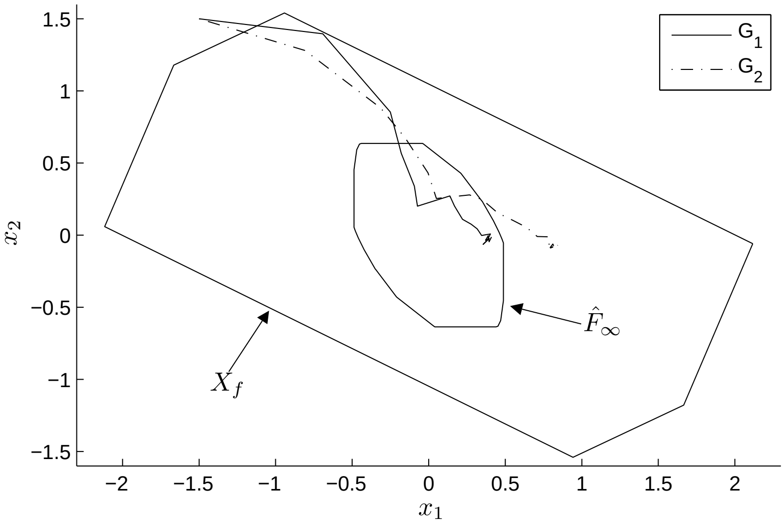

Figure 3.5: State trajectories of the other approach

of [10]. The intention is to show the differences between the parametrization of (3.2) and the parametrization of $u(t) = K_f x + c(t)$ where $c(t)$ is a direct variable of the FH problem, exemplified by the work of [10] and others. For comparison purpose, let $\mathrm{Pre}(X_f) = \{x|\exists u, (x,u) \in Y, s.t.\ Ax + Bu + w \in X_f, \forall w \in W\}$, the set of states that can be brought into X_f in one step and $\mathrm{Pre}^r = \mathrm{Pre} \cdots \mathrm{Pre}(\mathrm{Pre}(X_f))$, the r-times repeated application of $\mathrm{Pre}(\cdot)$. In general, the computation of $\mathscr{X}_N^{DFC}$ is expensive. An estimate of it can be obtained by checking over a grid of points in the x space according to Remark 3.1.2. Figure 3.6 shows 2 sets: $\mathscr{X}_5^{RPC}$, the admissible set using approach proposed in [10] and $\mathrm{Pre}^5(X_f)$. A "$\cdot$" point stands for a feasible initial state of $\mathscr{P}_N^{DFC}(x, \mathbf{0})$ and a "$\times$" stands for an infeasible initial state. As shown from the figure, $\mathscr{X}_5^{DFC}$ is almost indistinguishable from $\mathrm{Pre}^5(X_f)$ but is appreciably larger than $\mathscr{X}_5^{RPC}$.

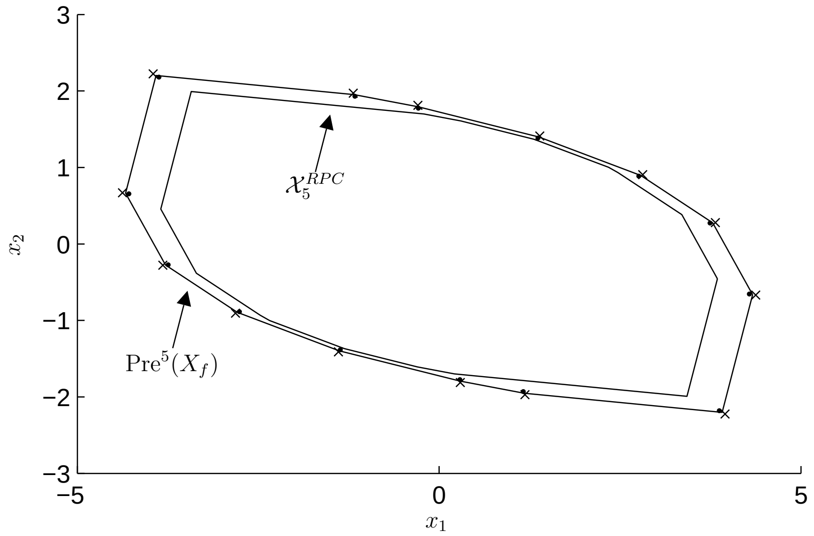

Figure 3.6: Comparison of admissible sets

3.6 Summary

A new affine disturbance feedback control parametrization is proposed in this chapter. This new parametrization is shown to have the same expressive ability as the time-varying affine disturbance feedback parametrization proposed in [41, 39]. Therefore the admissible set under this parametrization is equivalent to that under time-varying feedback parametrization. The advantage of this parametrization is that F_∞ stability is achievable if a norm-like cost function of the design variables is minimized. In addition, the size of the consequent F_∞ is adjustable by regulating the terminal feedback gain K_f. Finally, the performance of the controller derived under the new parametrization is illustrated by numerical examples and from the results it seems that the admissible set under the new parametrization is larger than the one under fixed feedback gain parametrization.

3.A Appendix

3.A.1 Proof of Theorem 3.1.1

For notational simplicity, the index $|t$ is dropped from all variables and $x(t)$ is denoted, without loss of generality, as $x(0)$. Equations (3.8a) and (3.8b) can be rearranged as

$$\begin{aligned}
\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} &= \begin{bmatrix} I & -\mathscr{B} \\ -\mathscr{K} & I \end{bmatrix}^{-1} \begin{bmatrix} \mathscr{A}x(0) + \mathscr{G}\mathbf{w} \\ \mathbf{c} + \mathbf{C}^{-}\mathbf{w}^{r} + \mathbf{C}^{+}\mathbf{w} \end{bmatrix} \\
&= \begin{bmatrix} \varphi & \varphi\mathscr{B} \\ \mathscr{K}\varphi & \mathscr{K}\varphi\mathscr{B} + I \end{bmatrix} \begin{bmatrix} \mathscr{A}x(0) + \mathscr{G}\mathbf{w} \\ \mathbf{c} + \mathbf{C}^{-}\mathbf{w}^{r} + \mathbf{C}^{+}\mathbf{w} \end{bmatrix} \\
&= \begin{bmatrix} \mathscr{A}_x x(0) + \mathscr{B}_x\mathbf{c} + \mathscr{B}_x\mathbf{C}^{-}\mathbf{w}^{r} + (\mathscr{B}_x\mathbf{C}^{+} + \mathscr{G}_x)\mathbf{w} \\ \mathscr{A}_u x(0) + \mathscr{B}_u\mathbf{c} + \mathscr{B}_u\mathbf{C}^{-}\mathbf{w}^{r} + (\mathscr{B}_u\mathbf{C}^{+} + \mathscr{G}_u)\mathbf{w} \end{bmatrix}.
\end{aligned} \tag{3.29}$$

where

$$\varphi := (I - \mathscr{B}\mathscr{K})^{-1} = \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ BK_f & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Phi^{N-2}BK_f & \Phi^{N-3}BK_f & \cdots & I & 0 \\ \Phi^{N-1}BK_f & \Phi^{N-2}BK_f & \cdots & BK_f & I \end{bmatrix} \text{ and}$$

$\mathscr{A}_x := \varphi\mathscr{A}$, $\mathscr{B}_x := \varphi\mathscr{B}$, $\mathscr{G}_x := \varphi\mathscr{G}$, $\mathscr{A}_u := \mathscr{K}\varphi\mathscr{A}$, $\mathscr{B}_u := I + \mathscr{K}\varphi\mathscr{B}$, $\mathscr{G}_u := \mathscr{K}\varphi\mathscr{G}$.

Comparing the expressions of $\mathbf{u}$ of (3.29) with (3.15), it follows that to get the same $\mathbf{u}$ following equations are needed.

$$\mathscr{B}_u(\mathbf{c}+\mathbf{C}^-\mathbf{w}^r) = \mathbf{v}-\mathscr{A}_u x(0) \tag{3.30a}$$

$$\mathscr{B}_u\mathbf{C}^+ = \mathbf{M}-\mathscr{G}_u. \tag{3.30b}$$

In the above, $\mathbf{C}^+$, $\mathscr{B}_u$, $\mathbf{M}$ and $\mathscr{G}_u$ are block lower triangular matrices. Hence, $(\mathbf{M},\mathbf{v})$ can be expressed in terms of $(\mathbf{c},\mathbf{C})$. This proves (i). To show that multiple $(\mathbf{c},\mathbf{C})$ exist for one choice of $(\mathbf{M},\mathbf{v})$, note that $\mathscr{B}_u$ is invertible following $\mathscr{B}_u^{-1} = I-\mathscr{K}(\varphi^{-1}+\mathscr{B}\mathscr{K})^{-1}\mathscr{B} = I-\mathscr{K}\mathscr{B}$. Then $\mathbf{C}^+ = (\mathscr{B}_u)^{-1}(\mathbf{M}-\mathscr{G}_u)$ and $\mathbf{c} = (\mathscr{B}_u)^{-1}(\mathbf{v}-\mathscr{A}_u x(0)) - \mathbf{C}^-\mathbf{w}^r$ for any choice of $\mathbf{C}^-$. Then the equivalence of $\mathscr{X}_N^{DFC}$ and $\mathscr{X}_N^{DF}$ follows directly from (i) and (ii).

3.A.2 Proof of Lemma 3.1.1

$(\Rightarrow)\Pi_N^{DFC}(x,\tilde{\mathbf{w}}^r) \neq \emptyset$ implies that there exists $(\tilde{\mathbf{c}},\tilde{\mathbf{C}}^-,\tilde{\mathbf{C}}^+)$ such that $\mathscr{P}_N^{DFC}(x,\tilde{\mathbf{w}}^r)$ is feasible. This also means that there exists $(\tilde{\mathbf{v}},\tilde{\mathbf{M}})$ such that (3.30) hold for $(\tilde{\mathbf{c}},\tilde{\mathbf{C}}^-,\tilde{\mathbf{C}}^+)$. Let $\mathbf{C}^- = 0$, $\mathbf{c} = \mathscr{B}_u^{-1}(\tilde{\mathbf{v}}-\mathscr{A}x)$ and $\mathbf{C}^+ = \mathscr{B}_u^{-1}(\tilde{\mathbf{M}}-\mathscr{G}_u)$ and they are feasible to $\mathscr{P}_N^{DFC}(x,\hat{\mathbf{w}}^r)$ following Theorem 3.1.1. $(\Leftarrow)$ Obvious by the symmetry of $(\hat{\mathbf{w}}^r,\tilde{\mathbf{w}}^r)$.

3.A.3 Proof of Theorem 3.4.1

The proof follows standard arguments in [10]. Let

$$(\mathbf{c}^*(t), \mathbf{C}^*(t)) = \arg\min \mathscr{P}_N^{DFC}(x(t), \mathbf{w}^r(t)).$$

Choose the feasible control at time $t+1$ $(\hat{\mathbf{c}}(t+1), \hat{\mathbf{C}}(t+1))$ as

$$\hat{c}(i|t+1) = c^*(i+1|t),\ \forall i \in \mathbb{Z}_{N-2}, \quad \hat{c}(N-1|t+1) = 0 \tag{3.31a}$$

$$\hat{C}(i,j|t+1) = \begin{cases} C^*(i+1,j|t) & \forall (i,j) \in \mathbb{Z}_{N-2} \times \mathbb{Z}_{N-1}^+ \\ 0 & i = N-1, \forall j \in \mathbb{Z}_{N-1}^+ \end{cases} \tag{3.31b}$$

and it is feasible to $\mathscr{P}_N^{DFC}(x(t+1), \mathbf{w}^r(t+1))$ due to the disturbance invariance of X_f under control law $u(t) = K_f x(t)$. Let $T_{(c,C)} = \{(\mathbf{c}, \mathbf{C}) | \exists \mathbf{Z} \text{ s.t. } (\mathbf{c}, \mathbf{C}, \mathbf{Z}) \in T\}$ where T is the polyhedron defined by (3.28). Since T is a polyhedron, so is $T_{(c,C)}$. As $J_{DFC}(\mathbf{c}, \mathbf{C})$ is a norm function, the set $\{(\mathbf{c}, \mathbf{C}) \in T_{(c,C)} | J_{DFC}(\mathbf{c}, \mathbf{C}) \leq J_{DFC}(\hat{\mathbf{c}}(t+1), \hat{\mathbf{C}}(t+1))\}$ is compact. Hence, the optimum of $\mathscr{P}_N^{DFC}(x(t+1), \mathbf{w}^r(t+1))$ exists, following Weierstrass' theorem.

Remark 3.A.1 *It is to be noted that the choice of (3.31) as the feasible control is only possible under the proposed parametrization (3.2). Specifically, j runs from 0 to $N-1$. If j runs from 0 to i as in the case of Chapter 4, a different feasible control at time $t+1$ is needed, see the proof of Theorem 4.4.1.*

3.A.4 Proof of Theorem 3.4.2

(i) The stated result follows directly from Lemma 3.1.1, Remark 3.1.2 and Theorem 3.4.1.

(ii) Let $J^*(t) := J_{DFC}(\mathbf{c}^*(t), \mathbf{C}^*(t))$ and $\hat{J}(t+1) := J_{DFC}(\hat{\mathbf{c}}(t+1), \hat{\mathbf{C}}(t+1))$ where $(\hat{\mathbf{c}}(t+1), \hat{\mathbf{C}}(t+1))$ are given by (3.31). Then it follows that

$$J^*(t) - J^*(t+1) \geq J^*(t) - \hat{J}(t+1) = h(l(t)) = h(l^*(0|t)) \geq 0, \; \forall\, t \geq 0 \tag{3.32}$$

where $h(\cdot)$ is as defined in (3.17). Hence, $\{J^*(t)\}$ is a monotonic non-increasing sequence and is bounded from below by zero. This means that

$$J^*_\infty := \lim_{t\to\infty} J^*(t) \geq 0 \tag{3.33}$$

exists. Repeating (3.32) for t from 0 to ∞ and summing them up, it follows that

$$\infty > J^*(0) - J^*_\infty \geq \sum_{t=0}^{\infty} h(l(t)) \geq 0 \tag{3.34}$$

This implies that $\lim_{t\to\infty} h(l(t)) = 0$. Since Ψ and Λ are positive definite, this implies that

$$\lim_{t\to\infty} c^*(0|t) = 0 \text{ and } \lim_{t\to\infty} C^*(0, j|t) = 0 \; \forall\, j \in \mathbb{Z}^+_{N-1} \tag{3.35}$$

and the stated result follows.

(iii) The system state under (3.14) can be written as

$$x(t) = \Phi^t x(0) + \sum_{i=0}^{t-1} \Phi^{t-1-i} Bl(i) + \sum_{i=0}^{t-1} \Phi^{t-1-i} w(i). \tag{3.36}$$

The first term on the right approaches zero as $t \to \infty$ since $\rho(\Phi) < 1$ and the second term approaches zero following property (ii). The last term corresponds to a point in the set $F_t := W \oplus \cdots \oplus \Phi^{t-1} W$, which approaches F_∞ as $t \to \infty$. Hence the stated result follows.

(iv) Let $x_\infty := \Sigma_{i=0}^{\infty} \Phi^i w(i)$. Then $\mathrm{E}(x_\infty) = 0$ and

$$\mathrm{E}(x_\infty x_\infty^T) = \Sigma_\infty = \Sigma_w + \Phi \Sigma_w \Phi^T + \cdots \tag{3.37}$$

following the assumptions in (A5) By pre- and post-multiplications of Φ and Φ^T of (3.37) respectively, the stated result follows.

Chapter 4

Probabilistic Convergence under Affine Disturbance Feedback

In this chapter, an affine disturbance feedback control parametrization is proposed. This parametrization differs from that proposed in Chapter 3 in that it does not use past realized disturbances. However, it has the same expressive ability and yields the same admissible set. The use of this parametrization and an appropriate cost function yields a different closed-loop convergence property: the state of the closed-loop system converges to a minimal d-invariant set with probability one. Deterministic convergence to the same minimal d-invariant set is also possible if a less intuitive cost function is used.

4.1 Introduction and Assumption

This chapter continues to consider the CTLD system,

$$x(t+1) = Ax(t) + Bu(t) + w(t) \tag{4.1a}$$

$$(x(t), u(t)) \in Y,\ w(t) \in W,\ \forall t \geq 0 \tag{4.1b}$$

which satisfies Assumption 1.4.1. Additionally, $w(t)$ is assumed to satisfy the following assumption for the discussion in this chapter.

Assumption 4.1.1

(A3a) *the disturbances $w(t)$ $t \geq 0$ are independent and identically distributed (i.i.d.) with zero mean.*

One other technical condition is also needed but its discussion is postponed until Section 4.2.

The MPC controller $\kappa^{DFC}(x, \mathbf{w}^r)$ in (3.14) of Chapter 3 requires the knowledge of realized disturbances $\mathbf{w}^r$ for its computation and is different from a traditional MPC controller. This chapter relaxes the utilization of this realized disturbances $\mathbf{w}^r$ while preserving similar convergence result.

The rest of this chapter is organized as follows. Section 4.2 states the proposed control parametrization, the FH optimization problem and the cost function used. Computation of the FH optimization problem is briefly discussed in Section 4.3. The result of probabilistic convergence of the closed-loop system state is given in Section 4.4. Section 4.5 discusses a formulation that strengthens the convergence result under a weaker set of assumptions. This, however, requires the use of a somewhat less intuitive cost function. Numerical examples are the contents of Section 4.6 and they are followed by the summary of this chapter.

4.2 Control Parametrization and MPC Formulation

The proposed control parametrization within the FH optimization problem takes the form

$$u(i|t) = K_f x(i|t) + d(i|t) + \sum_{j=1}^{i} D(i,j|t) w(i-j|t) \qquad \forall i \in \mathbb{Z}_{N-1} \tag{4.2}$$

where $d(i|t) \in \mathbb{R}^m, D(i,j|t) \in \mathbb{R}^{m \times n}, j \in \mathbb{Z}_i^+, i \in \mathbb{Z}_{N-1}$ are design variables and K_f is the feedback gain given in (A4) of Assumption 1.4.1. Since $j \in \mathbb{Z}_i^+$, $w(i-j|t)$ are all predicted disturbances and no elements of $\mathbf{w}^r(t)$ are used in (4.2). In this regard, (4.2) is similar to u^{DF} in (1.27) in that only predicted disturbances are used in the parametrization. In addition, parametrization (4.2) is equivalent to u^{DF} and (3.2), denoted hear after as u^{DFC}, in terms of the family of functions that can be represented.

Remark 4.2.1 *To see the equivalence of (4.2) to u^{DF} in (1.27) and u^{DFC} in (3.2), set $C(i,j|t) = 0$ for all $j > i$ in (3.2) and it follows that $u(i|t)$ in (4.2) is a special case of $u(i|t)$ in (3.2). To show the converse, let*

$$\begin{cases} d(i|t) = c(i|t) + \sum_{j=i+1}^{N-1} C(i,j|t)w(i-j|t), & \forall i \in \mathbb{Z}_{N-1} \\ D(i,j|t) = C(i,j|t), & \forall j \leq i,\ \forall i \in \mathbb{Z}_{N-1} \end{cases} \tag{4.3}$$

for any $c(i|t)$, $C(i,j|t)$ that defines $u(i|t)$ in (3.2). This establishes the equivalence of (3.2) and (4.2). The expressive equivalence of (3.2) and (1.27), has already been established in Theorem 3.1.1 and [68, 69]. With the above result, the representative abilities of (1.27), (3.2) and (4.2) are all equivalent.

Let $\mathbf{x}(t)$, $\mathbf{u}(t)$ and $\mathbf{w}(t)$ be defined in the same way as in (3.4) and collect all the design variables in (4.2) within the control horizon N in

$$\mathbf{D}(t) := \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ D(1,1|t) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ D(N-2,N-2|t) & D(N-2,N-3|t) & \cdots & 0 & 0 \\ D(N-1,N-1|t) & D(N-1,N-2|t) & \cdots & D(N-1,1|t) & 0 \end{bmatrix},$$

$$\mathbf{d}(t) := \begin{bmatrix} d(0|t) \\ d(1|t) \\ \vdots \\ d(N-1|t) \end{bmatrix}.$$

Then the control sequence defined by (4.2) is

$$\mathbf{u}(t) = \mathscr{K}\mathbf{x}(t) + \mathbf{d}(t) + \mathbf{D}(t)\mathbf{w}(t) \tag{4.4}$$

where $\mathscr{K} := [I_N \otimes K_f \;\; \mathbf{0}]$.

Remark 4.2.2 *Remark 4.2.1 implies that by letting* $\mathbf{c} = \mathbf{d}$, $\mathbf{C}^- = \mathbf{0}$ *and* $\mathbf{C}^+ = \mathbf{D}$, *(3.8b) becomes (4.4). Then due to (3.30),* $(\mathbf{d}, \mathbf{D})$ *and* $(\mathbf{v}, \mathbf{M})$ *must satisfy*

$$\mathscr{B}_u\mathbf{d} = \mathbf{v} - \mathscr{A}_u x(0) \tag{4.5a}$$

$$\mathscr{B}_u\mathbf{D} = \mathbf{M} - \mathscr{G}_u. \tag{4.5b}$$

to obtain the same $\mathbf{u}(t)$ *of (4.4) and (3.15). Additionally, the mapping between* $(\mathbf{d}, \mathbf{D})$ *and* $(\mathbf{v}, \mathbf{M})$ *is one-to-one since* $\mathscr{B}_u$ *is a lower triangular matrix with all diagonal elements being* 1 *and hence invertible as shown in Section 3.A.1.*

Using the notations defined in (3.9), (3.10) and (4.4), the FH optimization problem under parametrization (4.2), referred hereafter as $\mathscr{P}_N^{DFD}(x(t))$, is

$$\min_{\mathbf{d}(t),\mathbf{D}(t)} \quad J_{DFD}(\mathbf{d}(t),\mathbf{D}(t)) \tag{4.6a}$$

$$\text{s.t.} \quad \mathbf{x}(t) = \mathscr{A}x(t) + \mathscr{B}\mathbf{u}(t) + \mathscr{G}\mathbf{w}(t) \tag{4.6b}$$

$$\mathbf{u}(t) = \mathscr{K}\mathbf{x}(t) + \mathbf{d}(t) + \mathbf{D}(t)\mathbf{w}(t) \tag{4.6c}$$

$$(x(i|t), u(i|t)) \in Y, \ \ \forall\, \mathbf{w}(t) \in W^N, \ \ \forall i \in \mathbb{Z}_{N-1} \tag{4.6d}$$

$$x(N|t) \in X_f, \qquad \forall\, \mathbf{w}(t) \in W^N \tag{4.6e}$$

where the terminal set X_f is given in (A4) of Assumption 1.4.1. The cost function $J_{DFD}(\mathbf{d}(\mathbf{t}),\mathbf{D}(\mathbf{t}))$ takes the form

$$J_{DFD}(\mathbf{d}(t),\mathbf{D}(t)) := \sum_{i=0}^{N-1}\left[\|d(i|t)\|_{\Psi}^{2} + \sum_{j=1}^{i}\|\mathrm{vec}(D(i,j|t))\|_{\Lambda}^{2}\right] \tag{4.7}$$

for any choice of Ψ and Λ that satisfy

$$\Psi = \Psi^T \succ 0, \ \ \Lambda \succeq \Sigma_w \otimes \Psi \tag{4.8}$$

where Σ_w is the covariance matrix of disturbance w and $\mathrm{vec}(\cdot)$ is stacking operator defined in (2.16). Clearly, $J_{DFD}(\mathbf{d},\mathbf{D})$ is a measure of the deviation of $u(i|t)$ from the linear control law $K_f x(i|t)$ and the motivation for it as the objective function is clear: penalizing the use of non-zero $(\mathbf{d},\mathbf{D})$ forces the asymptotic behavior of the closed-loop system to approach that of $x(t+1) = (A+BK_f)x(t)$. The technical condition (4.8) is to ensure convergence of the closed-loop states and its exact role will become clear in the proof of Theorem 4.4.2. Several comments on $J_{DFD}(\mathbf{d},\mathbf{D})$ are in order.

Remark 4.2.3 *Following the same procedure of Section 3.2, a connection between* $J_{DFD}(\mathbf{d}(t),\mathbf{D}(t))$ *and the standard LQ cost can also be established. Specifically, under Assumption 3.2.1*

$$\begin{aligned}
&\mathrm{E}_{\mathbf{w}(t)}\left[\sum_{i=0}^{N-1}(\|x(i|t)\|_Q^2+\|u(i|t)\|_R^2)+\|x(N|t)\|_P^2\right]\\
&=x(t)^T Px(t)+N\mathrm{trace}(\Sigma_w P)+J_{DFD}(\mathbf{d}(t),\mathbf{D}(t)).
\end{aligned} \tag{4.9}$$

Since the first two terms on the right hand side of (4.9) are independent of $(\mathbf{d}(t),\mathbf{D}(t))$, *minimizing* $J_{DFD}(\mathbf{d}(t),\mathbf{D}(t))$ *is equivalent to minimizing the expected infinite horizon LQ cost over the disturbance input.*

Remark 4.2.4 *From (4.8) and (4.9), it may appear that* Σ_w *is needed for the determination of* Λ. *However, this is not true. The choice of* Λ *can be made to satisfy (4.8) even when* Σ_w *is not known accurately. One simple choice is to let* $\Lambda = \alpha^2 I_n \otimes \Psi$ *where* $\alpha := \max_{w\in W}\|w\|$. *Then it follows that* $\Lambda \succeq \Sigma_w \otimes \Psi$ *because* $\alpha^2 I_n \succeq ww^T$ *for all* $w \in W$ *and* $\alpha^2 I_n \otimes \Psi \succeq \mathrm{E}[ww^T]\otimes\Psi$. *Consequently, (A3) of Assumption 1.4.1 provides conditions that guarantee the computability of* $\max_{w\in W}\|w\|$.

Further discussion on the choice of Ψ and Λ and their influence on closed-loop system trajectories are discussed in Section 4.4. Several associated sets, needed to facilitate the discussions in the sequel, are introduced. Let the feasible set of optimization problem $\mathscr{P}_N^{DFD}(x)$ be

$$T_N^{DFD} := \{(x, \mathbf{d}, \mathbf{D}) |\ (\mathbf{d}, \mathbf{D}) \text{ is feasible to } \mathscr{P}_N^{DFD}(x)\} \tag{4.10}$$

and the set of admissible initial states, or equivalently, the domain of attraction of the MPC controller is

$$\mathscr{X}_N^{DFD} := \{x |\ \exists\ (\mathbf{d}, \mathbf{D}) \text{ such that } (x, \mathbf{d}, \mathbf{D}) \in T_N^{DFD}\}. \tag{4.11}$$

Remark 4.2.5 *One direct result following Theorem 3.1.1 and Remark 4.2.1 is that* $\mathscr{X}_N^{DFD} = \mathscr{X}_N^{DFC} = \mathscr{X}_N^{DF}$.

As usual, $\mathscr{P}_N^{DFD}(x(t))$ is solved at each time t to obtain the optimizer $(\mathbf{d}^*(t), \mathbf{D}^*(t))$ and the first control, $u^*(0|t)$, is applied to (4.1) at time t resulting in the MPC control law,

$$u(t) = \kappa^{DFD}(x(t)) := K_f x(t) + d^*(0|t). \tag{4.12}$$

4.3 Computation of the FH Optimization

Similar to constraint (3.24), constraints (4.6b)-(4.6e) can be written as

$$\bar{\mathscr{A}}x(0)+\bar{\mathscr{B}}\mathbf{d}(t)+\max_{\mathbf{w}(t)\in\mathbf{W}^N}\left[\bar{\mathscr{B}}\mathbf{D}(t)+\bar{\mathscr{G}}\right]\mathbf{w}(t)\leq\mathbf{1}_s \tag{4.13}$$

where $s=Nq+g$, $\mathbf{W}^N:=\{\mathbf{w}|\bar{H}\mathbf{w}\leq 1_\ell\}$, $\bar{\mathscr{A}}$, $\bar{\mathscr{B}}$ and $\bar{\mathscr{G}}$ are given in (3.25) and (3.26). Using the procedure given in Section 3.3, $\mathscr{P}_N^{DFD}(x(t))$ can be equivalently stated as

$$\min_{\mathbf{d(t)},\mathbf{D(t)},\mathbf{Z}} \quad J_{DFD}(\mathbf{d(t)},\mathbf{D}(t)) \tag{4.14a}$$

$$\text{s.t.} \quad \bar{\mathscr{A}}x(t)+\bar{\mathscr{B}}\mathbf{d}(t)+\mathbf{Z}^T 1_\ell\leq 1_s \tag{4.14b}$$

$$\mathbf{Z}^T\bar{H}=\bar{\mathscr{B}}\mathbf{D}(t)+\bar{\mathscr{G}} \tag{4.14c}$$

$$\mathbf{z}_i\geq 0,\ i=1,\ldots,s \tag{4.14d}$$

where $\mathbf{Z}=[\mathbf{z}_1\cdots\mathbf{z}_s]\in\mathbb{R}^{\ell\times s}$. In the conversion from (4.13) to (4.14b)-(4.14d), strong duality of linear programming is used. This duality results can be extended to W sets that are non-polyhedral. See, for example, treatments of such sets in [44] by Ben-Tal and Nemirovski and [70] by Nemirovski. If W is a second-order cone [71, 72] representable bounded set with non-empty interior such that $\mathbf{W}^N=\{\mathbf{w}|\ \|L_i\mathbf{w}-l_i\|\leq\lambda_i^T\mathbf{w}-\theta_i,\ i\in\mathbb{Z}_k^+\}$ for some matrices L_i, l_i, λ_i and θ_i, $i\in\mathbb{Z}_k^+$, then it follows from duality that

$$\max\{e^T\mathbf{w}|\ \mathbf{w}\in\mathbf{W}^N\}=\min_{(\mu_i,\eta_i)}\{\sum_{i=1}^k(\mu_i^T l_i-\eta_i\theta_i)|\ \sum_{i=1}^k(L_i^T\mu_i-\eta_i\lambda_i)=e,\ \|\mu_i\|\leq\eta_i,\ \forall i\in\mathbb{Z}_k^+\}.$$

$$(4.15)$$

In this case, $\mathscr{P}_N^{DFD}(x(t))$ is a second-order cone programming problem.

Similarly, if W is a bounded semi-definite cone representable set with non-empty interior such that $\mathbf{W}^N = \{\Omega \in R^{Nn} | \sum_{i=1}^{Nn} \Omega_i C_i - F \succcurlyeq 0,\ \forall i \in \mathbb{Z}_{Nn}^+\}$ where C_i and F are symmetrical matrices of appropriate dimension, then

$$\max\{e^T\mathbf{w} |\ \mathbf{w} \in \mathbf{W}^N\} = \min_{\Upsilon}\{\mathrm{Trace}(F\Upsilon) |\ \mathrm{Trace}(C_i\Upsilon) = e_i,\ \forall i \in \mathbb{Z}_{Nn}^+,\ \Upsilon \preccurlyeq 0\}. \quad (4.16)$$

In this case, $\mathscr{P}_N^{DFD}(x(t))$ is a semi-definite programming problem.

Remark 4.3.1 *While the duality result is available for W being a second-order or semi-definite cone representable set, the availability of X_f satisfying (A4) of Assumption 1.4.1 deserves some clarifications. When W is non-polyhedral, computation of a constraint-admissible disturbance invariant set X_f may not be easy. A simple approach is to construct a polytope W_p such that $W_p \supset W$ and $W_p \approx W$. In that case, a X_f can be constructed using W_p following existing computational methods [37]. Using this X_f in (4.6e), $\mathscr{P}_N^{DFD}(x)$ becomes either a second-order cone or a semi-definite cone programming problem. It is worthy to note that the use of such an X_f in (4.6e) and with $\mathbf{w}(t) \in W^N$ in both (4.6d) and (4.6e) is less conservative than replacing W by W_p throughout (4.6d)-(4.6e). An example using such an approach is illustrated in Section 4.6.*

4.4 Feasibility and Probabilistic Convergence

The feasibility of $\mathscr{P}_N^{DFD}(x(t))$ at different time instants and stability of the closed-loop system under the feedback law (4.12) are addressed in this section.

Theorem 4.4.1 *Suppose Assumption 1.4.1 is satisfied, the FH optimization problem* $\mathscr{P}_N^{DFD}(x)$ *has the following properties (i)* T_N^{DFD} *is convex and compact. (ii) If* $x \in \mathscr{X}_N^{DFD}$*, the optimal solution of* $\mathscr{P}_N^{DFD}(x)$ *exists. (iii) If* $\mathscr{P}_N^{DFD}(x(t))$ *admits an optimal solution, so does* $\mathscr{P}_N^{DFD}(x(t+1))$ *under the feedback law (4.12) for all possible* $w(t) \in W$.

Proof: See Appendix 4.A.1. ■

Remark 4.4.1 *The set* $\mathscr{X}_N^{DFD}$ *can also be proved to be convex and compact. Interested readers can refer to Section 3.4 of [73].*

The main result of probabilistic convergence of the closed-loop system is stated in the next theorem. Such a convergence is achieved under the Assumption 1.4.1, Assumption 4.1.1 and condition (4.8).

Theorem 4.4.2 *Suppose* $x(0) \in \mathscr{X}_N^{DFD}$ *and Assumption 1.4.1 and Assumption 4.1.1 are satisfied. System (4.1a) under MPC control law (4.12) obtained from the solution of* $P_N^{DFD}(x)$ *using cost function (4.7) satisfying (4.8) has the following properties: (i)* $(x(t), u(t)) \in Y$ *for all* $t \geq 0$*, (ii)* $x(t) \to F_\infty(K_f)$ *with probability one as* $t \to \infty$.

Proof: See Appendix 4.A.2. ■

One associated issue in the formulation of $\mathscr{P}_N^{DFD}(x)$ is the choices of Ψ and Λ in $J_{DFD}(\mathbf{d},\mathbf{D})$. How should Ψ and Λ be chosen and how do these choices affect the closed-loop system trajectories? As $x(t) \to F_\infty$ from result (ii) of Theorem 4.4.2, it implies that $x(t)$ enters X_f with probability one and stays within thereafter since $F_\infty \subset X_f$. When this happens, the optimal $(\mathbf{d},\mathbf{D})$ are zero in $\mathscr{P}_N^{DFD}(x)$ and the MPC control law becomes $u(t) = K_f x(t)$ for all t thereafter. The closed-loop system behavior then corresponds to that of the system $x(t+1) = (A+BK_f)x(t)+w(t)$. Clearly, the choices of Λ and Ψ does not affect the asymptotic behavior of the system but only the transient when $x(t) \notin X_f$.

Suppose $\Lambda = \Sigma_w \otimes \Psi$. Then admissible changes in Ψ will not result in changes in the system behavior since $\Sigma_w \otimes \Psi$ is linear in Ψ. On the other hand, if Ψ is fixed, Λ can be chosen to be increasingly "larger" than $\Sigma_w \otimes \Psi$. In loose terms, a "larger" choice of Λ penalizes the use of $\mathbf{D}$ versus the use of $\mathbf{d}$ in $J(\mathbf{d},\mathbf{D})$. Such a preference would mitigate the effect of the disturbance feedback component in the control parametrization, resulting in a parametrization that is closer in spirit to $u(t) = K_f x(t) + c(t)$ of [10]. When this happens, the transient response for the system may become slower even though the domain of attraction $\mathscr{X}_N^{DFD}$ remains unaffected. This observation together with the associated details used in the numerical examples are discussed in Section 4.6.

4.5 Deterministic Convergence

While the assumption of W being a compact set is reasonable, the assumption of $w(t)$ being zero mean and i.i.d. is harder to verify in practice. This section is concerned with the relaxation of Assumption 4.1.1 while achieving a stronger convergence result than that of Theorem 4.4.2. To this end, the following new cost function is needed.

$$V_{DFD}(\mathbf{d}(t),\mathbf{D}(t)) := \sum_{i=0}^{N-1}\left[\|d(i|t)\|_{\Psi}^2 + \sum_{j=1}^{i}(\gamma_1\|\text{vec}(D(i,j|t))\|^2 + \gamma_2\|\text{vec}(D(i,j|t))\|)\right] \tag{4.17}$$

for some constants γ_1 and γ_2 satisfying

$$\gamma_1 \geq \alpha^2\|\Psi\|,\ \gamma_2 \geq 2\alpha\beta\|\Psi\| \tag{4.18}$$

where $\alpha := \max_{w\in W}\|w\|$ and $\beta := \max_{(x(t),\mathbf{d}(t),\mathbf{D}(t))\in T_N^{DFD}, i\in\mathbb{Z}_{N-1}}\|d(i|t)\|$. The existence of α and β are guaranteed by compactness of the W and T sets, provided for in (A3) of Assumption 1.4.1 and part (i) of Theorem 4.4.1 respectively. Furthermore, the computation of β can be simplified to $\beta = \max_{(x(t),\mathbf{d}(t),\mathbf{D}(t))\in T_N^{DFD}}\|d(0|t)\|$, see Appendix 4.A.3.

Let $J_{DFD}(\mathbf{d}(t),\mathbf{D}(t))$ in $\mathscr{P}_N^{DFD}(x(t))$ be replaced by $V_{DFD}(\mathbf{d}(t),\mathbf{D}(t))$, and denote the resulting FH optimization problem by $\mathscr{P}_N^{DFDV}(x(t))$. Since only the cost function is replaced, the admissible set of $\mathscr{P}_N^{DFDV}(x(t))$, denoted by $\mathscr{X}_N^{DFDV}$, is equivalent to $\mathscr{X}_N^{DFD}$. Also denote the corresponding MPC control law by $\kappa^{DFDV}(\cdot)$. The convergence result under controller $u(t) = \kappa^{DFDV}(x(t))$ is summarized in the following theorem.

Theorem 4.5.1 *Suppose $x(0) \in \mathscr{X}_N^{DFDV}$, Assumption 1.4.1 is satisfied, then system (4.1a) under the MPC controller $u(t) = \kappa^{DFDV}(x(t))$ satisfies (i) $(x(t), u(t)) \in Y$ for all $t \geq 0$, (ii) $x(t) \to F_\infty(K_f)$ as $t \to \infty$.*

Proof: See Appendix 4.A.4. ■

Remark 4.5.1 *Several choices of the cost function of (4.17) are possible. For example, the results of Theorem remain true if $\|vec(D(i,j|t))\|$ is replaced by $\|D(i,j|t)\|$. This may appear more appealing as less conservative bounds on γ_1 and γ_2 can be found to ensure the non-negativity of $p(w(t))$ in (4.32). However, its use will result in a semi-definite programming problem for $\mathscr{P}_N^{DFDV}(x(t))$, which is less desirable computationally. The use of $\|vec(D(i,j|t))\|$ or $\|D(i,j|t)\|_F$ results in a second-order cone programming for $\mathscr{P}_N^{DFDV}(x(t))$ and is computationally more amiable.*

4.6 Numerical Examples

Four examples are presented to validates the results of this chapter. The system parameters and constraints of the system are:

$$A = \begin{bmatrix} 1.1 & 1 \\ 0 & 1.3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad K_f = [-0.7434 \ -1.0922],$$

$$Y = \{(x,u)|\ |u| \le 1,\ \|x\|_\infty \le 8\},$$

$$W = \{H\tilde{w}|H = \begin{bmatrix} 1 & -0.2 \\ 0 & 1 \end{bmatrix} \text{ and } \|\tilde{w}\|_\infty \le 0.2\}$$

where $\tilde{w} \in \mathbb{R}^2$ is a random vector uniformly distributed over $[-0.2, 0.2] \times [-0.2, 0.2]$ with covariance matrix $\Sigma_{\tilde{w}} = 0.0133 I_2$. Terminal set X_f is the corresponding maximal constraint-admissible disturbance invariant set of (4.1a) under $u(t) = K_f x(t)$, specifically

$$X_f = \{x|\ Gx \le 1_4\},\ G = \begin{bmatrix} -0.7434 & -1.0922 \\ 0.7434 & 1.0922 \\ 0.8252 & -0.2391 \\ -0.8252 & 0.2391 \end{bmatrix}.$$

The weight matrices in the cost function (4.7) are chosen to be

$$\Psi = 1,\ \ \Lambda = \Lambda_{op} := \Sigma_w \otimes \Psi = \begin{bmatrix} 0.0139 & -0.0027 \\ -0.0027 & 0.0133 \end{bmatrix}$$

The algorithm using cost function (4.20) is simulated with $N = 8$ and $x(0) = [-4\ \ 2]^T$ over 15 realizations of disturbance sequences and the results are shown in Figures 4.1 to 4.4 in solid lines. It is clear from Figures 4.1 and 4.2 that both the state and control constraints are satisfied by all trajectories, in accordance to property (i) of Theorem 4.4.2. In addition, Figure 4.1 shows the convergence of $x(t)$ into $\hat{F}_\infty(K_f)$, an outer bound of the minimal invariant set associated with the choice of K_f. This convergence is further verified by solid lines in Figures 4.3 and 4.4 where the plots of $dis(x(t), \hat{F}_\infty) :=$

$\min_{x \in \hat{F}_\infty} \|x - x(t)\|$, the minimum distance to $\hat{F}_\infty$, and $d(t) := d^*(0|t)$ against increasing t are shown.

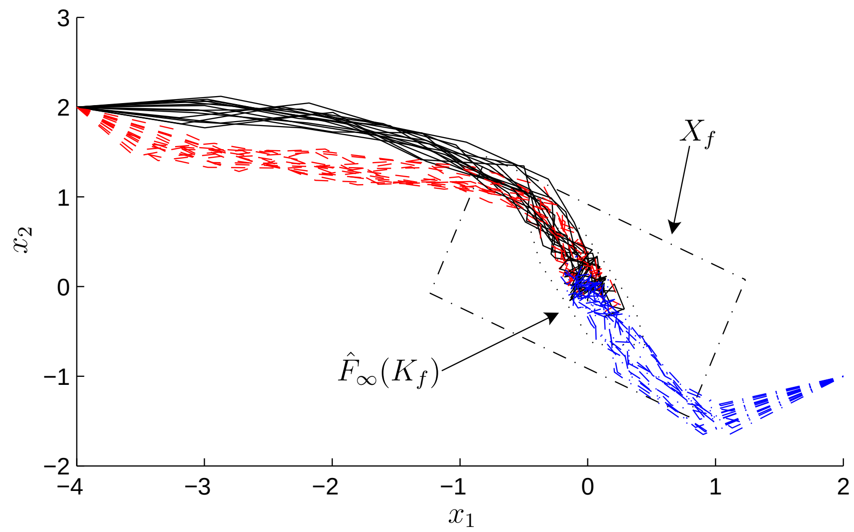

Figure 4.1: State trajectories of the first three simulations

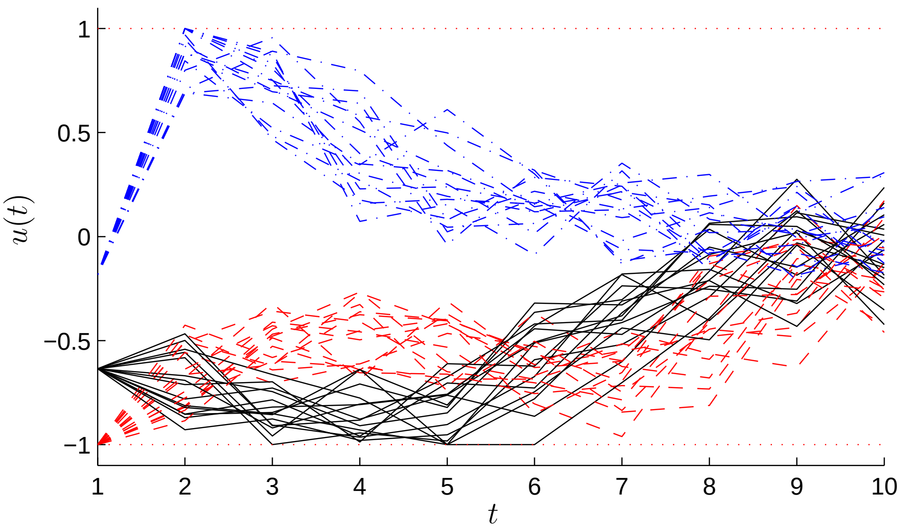

Figure 4.2: Control trajectories of the first three simulations

The case where W is non-polyhedral is shown in Simulation II, in connection to Section 4.3. The system considered is that of Simulation I but has a different disturbance characteristic: w is uniformly distributed over $\bar{W} := \{\bar{w} | \ \|S_1\bar{w}\| \leq 1, \ \|S_2\bar{w}\| \leq 1\}$ where

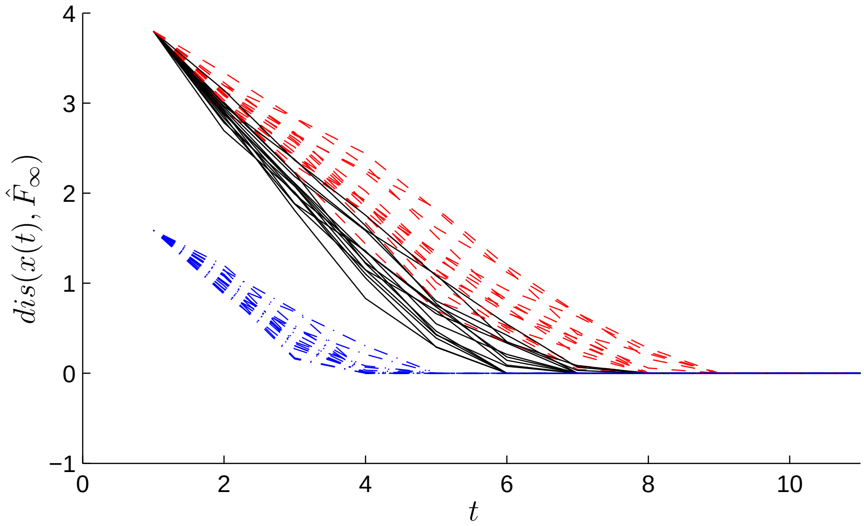

Figure 4.3: Distance between states and $F_\infty(K_f)$ of the first three simulations

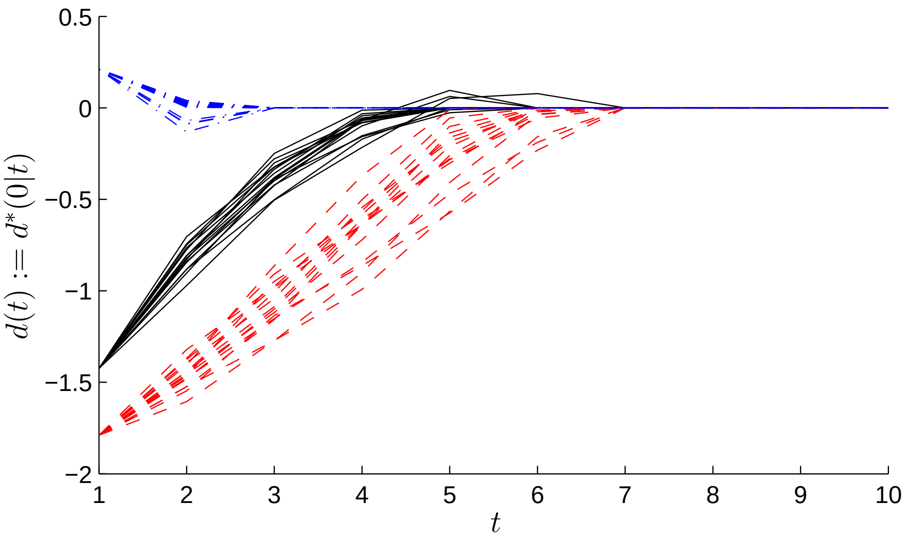

Figure 4.4: Values of $d(t)$ of the first three simulations

$S_1 = [5\ \ 1;\ 0\ \ 2.5]$ and $S_1 = [2.5\ \ 0.5;\ 0\ \ 5]$. Note that a tight bounding polytope, W_p, such that $W_p \supset \bar{W}$, is the W set of Simulation I (see Figure 4.5) and (A4) of Assumption 1.4.1 is satisfied using X_f of the first simulation. Also, Ψ and Λ of the first simulation are used and it is easy to verify that condition (4.8) remains true because $\Sigma_w \succ \Sigma_{\bar{w}}$. In this case, $\bar{W}$ is a second-order cone representable set and the conversion of (4.6d) and (4.6e) for all $w(i|t) \in \bar{W}$ follows the procedure in Section 4.3, results in $\mathscr{P}_N^{DFD}(x)$ being a second-order cone programming problem. The simulation results with $N = 8$ and $x(0) = [2\ \ -1]^T$ for 15 different realizations of $\{w(t)\}$ are plotted in Figure 4.1 to 4.4 using dash-dot lines.

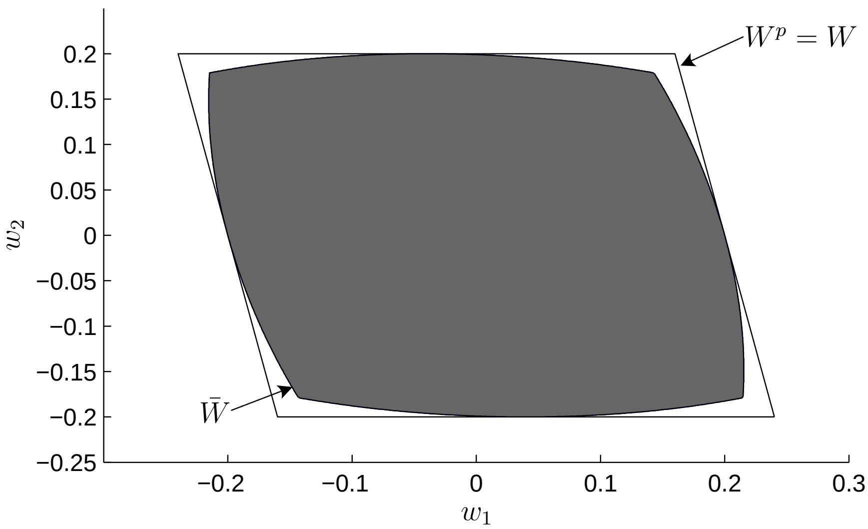

Figure 4.5: W_p set and $\bar{W}$ set.

Simulation III attempts to understand the influence of the weight matrices, Λ and Ψ of (4.7), on the performance of the closed-loop system. As stated in Section 4.4, choices of these matrices affect only the transient behavior when $x(t) \notin X_f$ and not the asymptotic behavior of the closed-loop system. To quantify the transient, the average number of time step, $t_f(x(0))$, taken to enter X_f from a given $x(0)$ is reported. Here, the average is taken over different realizations of the disturbances. Without loss of generality, values of Λ is varied following the discussions in Section 4.4. Table 4.7 shows the $t_f(x(0))$ and the associated standard deviation over 20 disturbance realizations for several choices of $x(0)$ and Λ. For each $x(0)$, the same 20 disturbance realizations are used for the different Λ in computing $t_f(x(0))$ and the standard deviations. From the table, $t_f(x(0))$ generally increases when Λ increases. For comparison purpose, the corresponding trajectories of the system under same settings as the first simulation except for $\Lambda = 10^4\Lambda_{op}$ are plotted in Figure 4.1 to 4.4 using dash lines. From Figure 4.3 and 4.4, the slower convergence of the state and control trajectories are clearly evident.

The last simulation, Simulation IV, considers the case discussed in Section 4.5. The system parameters are the same as those in the first simulation except that the distribution of $\tilde{w}$ is assumed to be unknown. Hence, α is 0.3124 and β of (4.18) are determined to be 2.7307 (when $N = 8$) and 3.5425 (when $N = 10$). Correspondingly, the weight matrices of (4.17) are chosen to be $\Psi = 1$, $\gamma_1 = \gamma_1^{op} := \alpha^2\|\Psi\| = 0.0976$, $\gamma_2 = \gamma_2^{op} := 2\alpha\beta\|\Psi\| = 1.7059$ $(2.213$ when $N = 10)$. The values of γ_1 and γ_2 are increased separately and jointly to assess their influence on the system behavior. Similar to the third simulation, the response of the system with several choices of the weight matrices are simulated over 20 series of disturbance realizations.

The general effect of increasing values of γ_1 and γ_2 appears to have similar trend on the system as the increase in Λ. The time taken to reach X_f from any given $x(0)$ increases, although of a lesser percentage than that by Λ, with increasing values of γ_1 and γ_2 with

γ_2 having a heavier influence.

4.7 Summary

A control parametrization that does not make use of past realized disturbances is proposed in this chapter. This parametrization preserves the expressive ability of the parametrization in Chapter 3. Using this parametrization and a proposed cost function under the MPC framework, the closed-loop system state converges to the minimal robust invariant set F_∞ with probability one. Deterministic convergence to F_∞ is also possible using a less intuitive cost function.

Initial Condition		Λ			(γ_1, γ_2)			
$x(0)$	N	Λ_{op}	$10^2\Lambda_{op}$	$10^4\Lambda_{op}$	$(\gamma_1^{op}, \gamma_2^{op})$	$(10\gamma_1^{op}, \gamma_2^{op})$	$(\gamma_1^{op}, 10\gamma_2^{op})$	$(10\gamma_1^{op}, 10\gamma_2^{op})$
$[-4\ 2]^T$	8	4.2 (0.4104)	4.9 (0.4472)	5.65 (0.4894)	5.2 (0.4104)	5.4 (0.5026)	5.65 (0.4894)	5.65 (0.4894)
$[-2.5\ -1.2]^T$	8	4.9 (0.5525)	5.15 (0.4894)	5.15 (0.4894)	5.15 (0.4894)	5.15 (0.4894)	5.15 (0.4894)	5.15 (0.4894)
$[-4\ -1]^T$	10	6.3 (0.7327)	6.75 (0.6387)	7.45 (0.8256)	7.2 (0.6959)	7.25 (0.7164)	7.45 (0.8256)	7.45 (0.8256)
$[-6\ 2]^T$	10	6.45 (0.6863)	7.15 (0.6708)	8.55 (0.6863)	7.9 (0.4472)	8.05 (0.6863)	8.6 (0.5982)	8.6 (0.5982)

Table 4.1: Average time step, $t_f(x(0))$, and its standard deviation

4.A Appendix

4.A.1 Proof of Theorem 4.4.1

(i) It is easy to see that for each $i \in Z_{N-1}$, $(x(i|t), u(i|t))$ are affine functions of $(x(t), \mathbf{d}(t), \mathbf{D}(t))$ from (4.6b)-(4.6c). This, together with the fact that Y and X_f are convex and compact sets, means that the feasible set T_N^{DFD}, as defined by constraints (4.6b)-(4.6e) is convex and compact. (ii) Since $x \in \mathscr{X}_N^{DFD}$, $\mathscr{P}_N^{DFD}(x)$ is feasible. From (i), this means that $\Pi_N^{DFD}(x) := \{(\mathbf{d}, \mathbf{D}) | (x, \mathbf{d}, \mathbf{D}) \in T_N^{DFD}\}$ is compact. This, together with the fact that $J_{DFD}(\mathbf{d}, \mathbf{D})$ is continuous with respect to $(\mathbf{d}, \mathbf{D})$ means that the optimal solution exists by Weierstrass' Theorem [74]. (iii) The proof follows standard arguments but the details are given for their relevance to Theorem 4.5.1. Let $(\mathbf{d}^*(t), \mathbf{D}^*(t))$ denote the optimal solution of $\mathscr{P}_N^{DFD}(x(t))$. At time $t+1$ when $w(t)$ is realized, choose $(\hat{\mathbf{d}}(t+1), \hat{\mathbf{D}}(t+1))$ by letting

$$\hat{d}(i|t+1) = \begin{cases} d^*(i+1|t) + D^*(i+1, i+1|t)w(t) & \forall i \in \mathbb{Z}_{N-2} \\ 0 & i = N-1 \end{cases} \tag{4.19a}$$

$$\hat{D}(i, j|t+1) = \begin{cases} (D^*(i+1, j|t) & \forall j \in \mathbb{Z}_i^+,\ \forall i \in \mathbb{Z}_{N-2}^+ \\ 0 & \forall j \in \mathbb{Z}_{N-1}^+,\ i = N-1 \end{cases} \tag{4.19b}$$

and it is feasible to $\mathscr{P}_N^{DFD}(x(t+1))$ for all possible $w(t) \in W$ due to the disturbance invariance of X_f for system (4.1a) under control law $u(t) = K_f x(t)$. It is clear that

$\Pi_N^{DFD}(x)$ is compact for all $x \in \mathscr{X}_N^{DFD}$. Since W is bounded and J_{DFD} is a norm function, $\max_{w(t)} J_{DFD}(\hat{\mathbf{d}}(t+1), \hat{\mathbf{D}}(t+1)) < \infty$ and the set $\{(\mathbf{d}, \mathbf{D}) \in \Pi_N^{DFD}(x(t+1)) | J(\mathbf{d}, \mathbf{D}) \leq \max_{w(t)} J_{DFD}(\hat{\mathbf{d}}(t+1), \hat{\mathbf{D}}(t+1))\}$ is compact. Hence, the optimum of $\mathscr{P}_N^{DFD}(x(t+1))$ exists, following the Weierstrass' theorem.

4.A.2 Proof of Theorem 4.4.2

(i) The stated result follows directly from Theorem 4.4.1. (ii) Let $J^*(t) := J_{DFD}(\mathbf{d}^*(t), \mathbf{D}^*(t))$ and $\hat{J}(t+1) := J_{DFD}(\hat{\mathbf{d}}(t+1), \hat{\mathbf{D}}(t+1))$ where $(\hat{\mathbf{d}}(t+1), \hat{\mathbf{D}}(t+1))$ are given by (4.19). Then it follows that

$$\begin{aligned}
& J^*(t) - \hat{J}(t+1) \\
= & \sum_{i=0}^{N-1} (\|d^*(i|t)\|_\Psi^2 - \|\hat{d}(i|t+1)\|_\Psi^2) + \sum_{i=1}^{N-1} \|\mathrm{vec}(D^*(i,i|t))\|_\Lambda^2 \\
= & \|d^*(0|t)\|_\Psi^2 + \sum_{i=1}^{N-1} \left(\|d^*(i|t)\|_\Psi^2 - \|\hat{d}(i-1|t+1)\|_\Psi^2\right) + \sum_{i=1}^{N-1} \|\mathrm{vec}(D^*(i,i|t))\|_\Lambda^2 \\
= & \|d^*(0|t)\|_\Psi^2 + \sum_{i=1}^{N-1} \left(\|d^*(i|t)\|_\Psi^2 - \|d^*(i|t) + D^*(i,i|t)w(t)\|_\Psi^2\right) + \sum_{i=1}^{N-1} \|\mathrm{vec}(D^*(i,i|t))\|_\Lambda^2 \\
= & \|d^*(0|t)\|_\Psi^2 + g(w(t))
\end{aligned} \tag{4.20}$$

where

$$g(w(t)) = \sum_{i=1}^{N-1} (\|\mathrm{vec}(D^*(i,i|t))\|_\Lambda^2 - 2d^*(i|t)^T \Psi D^*(i,i|t)w(t) - \|(D^*(i,i|t)w(t)\|_\Psi^2). \tag{4.21}$$

Taking the expectation of (4.20) over $w(t)$, it follows that

$$\begin{aligned} J^*(t) - \|d^*(0|t)\|_\Psi^2 &= \mathrm{E}_{w(t)}\left[\hat{J}(t+1)\right] + \mathrm{E}_{w(t)}[g(w(t))] \\ &\geq \mathrm{E}_{w(t)}\left[\hat{J}(t+1)\right] \qquad (4.22) \\ &\geq \mathrm{E}_{w(t)}\left[J^*(t+1)\right] = \mathrm{E}_t\left[J^*(t+1)\right]. \qquad (4.23) \end{aligned}$$

where E_t in (4.23) is the expectation taken over $w(i),\ i \geq t$. Inequality (4.22) follows from the fact that $E_{w(t)}[g(w(t))] \geq 0$. This is true because by taking the expectation of (4.21), one gets

$$\begin{aligned} & \mathrm{E}_{w(t)}[g(w(t))] \\ = & \sum_{i=1}^{N-1}(\|\mathrm{vec}(D^*(i,i|t))\|_\Lambda^2 - \|\mathrm{vec}(D^*(i,i|t))\|_{\Sigma_w\otimes\Psi}^2 - 2(d^*(i|t))^T\Psi D^*(i,i|t)E[w(t)]) \qquad (4.24) \end{aligned}$$

where the last term is zero due to Assumption 4.1.1 and the rest is non-negative due to (4.8).

Inequality (4.23) follows from the fact that $\hat{J}(t+1) \geq J^*(t+1)$ for every $w(t) \in W$ which implies that $\mathrm{E}_{w(t)}[\hat{J}(t+1)] \geq \mathrm{E}_{w(t)}[J^*(t+1)]$. The last equality of (4.23) follows from the fact that $J^*(t+1)$ depends on $w(t)$ only and not on any $w(i), i > t$.

Repeating the inequality of (4.23) for increasing t, one gets,

$$J^*(t+1) - \|d^*(0|t+1)\|_\Psi^2 \geq \mathrm{E}_{w(t+1)}\left[J^*(t+2)\right]$$

Since the two terms on the left hand side depends on $w(t)$ and the term on the right depends on $w(t)$ and $w(t+1)$, the above can be equivalently written as

$$J^*(t+1)(w(t)) - \|d^*(0|t+1)(w(t))\|_\Psi^2 \geq \mathrm{E}_{w(t+1)}\left[J^*(t+2)(w(t),w(t+1))\right]. \quad (4.25)$$

The above inequality holds true for all possible $w(t)$, hence

$$\begin{aligned} & \mathrm{E}_{w(t)}[J^*(t+1)(w(t))] - \mathrm{E}_{w(t)}[\|d^*(0|t+1)(w(t))\|_\Psi^2] \\ \geq\ & \mathrm{E}_{w(t)}[\mathrm{E}_{w(t+1)}\left[J^*(t+2)(w(t),w(t+1))\right]] = \mathrm{E}_t[J^*(t+2)(w(t),w(t+1))] \end{aligned} \quad (4.26)$$

or equivalently

$$\mathrm{E}_t[J^*(t+1)] - \mathrm{E}_t[\|d^*(0|t+1)\|_\Psi^2] \geq \mathrm{E}_t[J^*(t+2)] \quad (4.27)$$

The equality in (4.26) follows from the i.i.d. in Assumption 4.1.1, particularly,

$$\begin{aligned} & \mathrm{E}_{w(t)}[\mathrm{E}_{w(t+1)}\left[J^*(t+2)(w(t),w(t+1))\right]] \\ =\ & \mathrm{E}_{w(t)}\left[\int J^*(t+2)(w(t),w(t+1))f_{w(t+1)}(w(t+1))dw(t+1)\right] \\ =\ & \int\int J^*(t+2)(w(t),w(t+1))f_{w(t+1)}(w(t+1))dw(t+1)f_{w(t)}(w(t))dw(t) \\ =\ & \int\int J^*(t+2)(w(t),w(t+1))f_{w(t),w(t+1)}(w(t),w(t+1))dw(t+1)dw(t) \\ =\ & \mathrm{E}_{w(t),w(t+1)}[J^*(t+2)(w(t),w(t+1))] = \mathrm{E}_t[J^*(t+2)(w(t),w(t+1))] \end{aligned}$$

where $f_{w(t)}(\cdot)$, $f_{w(t+1)}(\cdot)$ and $f_{w(t),w(t+1)}(\cdot,\cdot)$ are density functions of $w(t)$, $w(t+1)$ and their joint density function respectively and $f_{w(t),w(t+1)}(\cdot,\cdot) = f_{w(t)}(\cdot)f_{w(t+1)}(\cdot)$ from i.i.d. Summing (4.23) and (4.27) leads to

$$J^*(t) \geq \|d^*(0|t)\|_\Psi^2 + \mathrm{E}_t[\|d^*(0|t+1)\|_\Psi^2] + \mathrm{E}_t[J^*(t+2)] \tag{4.28}$$

Repeating the above procedure infinite times leads to

$$\infty > J^*(t) \geq \sum_{i=t}^{\infty} \mathrm{E}_t \left[\|d^*(0|i)\|_\Psi^2\right] \tag{4.29}$$

By applying Markov bound (given non-negative random variable η and any $\varepsilon \geq 0$, $\mathrm{E}[\eta] \geq \varepsilon \Pr\{\eta \geq \varepsilon\}$), we have

$$\infty > \varepsilon \sum_{i=t}^{\infty} \Pr(\|d^*(0|i)\|_\Psi^2 \geq \varepsilon) \tag{4.30}$$

for any arbitrary small $\varepsilon > 0$. From the First Borel-Cantelli Lemma [75], this implies that $\lim_{i\to\infty} \Pr(\|d^*(0|i)\|_\Psi^2 \geq \varepsilon) = 0$. Hence $d(0|i) \to 0$ with probability one as t increases. Consequently, the MPC control law (4.12) converges to $K_f x(t)$ with probability one. When this happens, the closed-loop system converges to $x(t+1) = \Phi x(t) + w(t)$ and, hence, $x(t)$ converges to $F_\infty(K_f)$ with probability one.

4.A.3 Computation of β

For notational simplicity, notations (t) and $|t$ are omitted here.

To see $\beta := \max_{(x,\mathbf{d},\mathbf{D})\in T_N^{DFD}, i\in\mathbb{Z}_{N-1}} \|d(i)\| = \max_{(x,\mathbf{d},\mathbf{D})\in T_N^{DFD}} \|d(0)\|$, it is needed to show that for any $(x,\mathbf{d},\mathbf{D}) \in T_N^{DFD}$ and any integer $i \in \mathbb{Z}_{N-1}^+$, there exists $(\tilde{x},\tilde{\mathbf{d}},\tilde{\mathbf{D}}) \in T_N^{DFD}$ such that $\tilde{d}(0) = d(i)$. Suppose $(x,\mathbf{d},\mathbf{D})$ defines the state sequence $\{x(0),\dots,x(N)\}$ and control sequence and $\{u(0),\dots,u(N-1)\}$. Also let $\{\bar{x}(0),\dots,\bar{x}(N)\}$ denotes the nominal value of $\{x(0),\dots,x(N)\}$. Define $(\tilde{x},\tilde{\mathbf{d}},\tilde{\mathbf{D}})$ to be

$$\tilde{x} = \bar{x}(i) = \Phi^i x + \sum_{j=0}^{i-1} \Phi^{i-1-j} B d(j), \quad \tilde{d}(j) = \begin{cases} d(j+i) & \forall j \in \mathbb{Z}_{N-1-i} \\ 0 & N-i \leq j \leq N-1 \end{cases},$$

and

$$\tilde{D}(j,k) = \begin{cases} D(j+i,k) & \forall j \in \mathbb{Z}_{N-1-i}^+ \\ 0 & N-i \leq j \leq N-1 \end{cases} \quad k \in \mathbb{Z}_j^+.$$

It can be verified that $(\tilde{x},\tilde{\mathbf{d}},\tilde{\mathbf{D}})$ defines the state sequence

$$\{x(i),\cdots,x(N),\Phi x(N)+w(N),\cdots,\Phi^i x(N) + \sum_{j=0}^{i-1} \Phi^{i-1-j} w(N+j)\}$$

and control sequence

$$\{u(i),\cdots,u(N-1),K_f x(N),\cdots,K_f(\Phi^i x(N) + \sum_{j=0}^{i-1} \Phi^{i-1-j} w(N+j))\}.$$

According to (4.6e), $x(N) \in X_f$. According to (A4) of Assumption 1.4.1 under controller $u(i) = K_f x(i)$, all the constraints are satisfied and $x(i) \in X_f$ for $i \geq N$ since $x(N) \in X_f$. Therefore, $(\tilde{x},\tilde{\mathbf{d}},\tilde{\mathbf{D}})$ satisfies (4.6b)-(4.6e) and hence $(\tilde{x},\tilde{\mathbf{d}},\tilde{\mathbf{D}}) \in T_N^{DFD}$.

Note that any upper bound of β can be used to guarantee the results of Theorem 4.5.1. One such upper bound is $\tilde{\beta} := \|\sigma\|$ where $\sigma(j) := \max_{(x,\mathbf{d},\mathbf{D})\in T} |d_j(0)|$ and $d_j(0)$ is the j^{th} element of $d(0)$. Hence, the value of $\tilde{\beta}$ can be computed efficiently by solving n LPs.

4.A.4 Proof of Theorem 4.5.1

(i) The replacement of cost function $J_{DFD}(\mathbf{d}(t),\mathbf{D}(t))$ by $V_{DFD}(\mathbf{d}(t),\mathbf{D}(t))$ does not affect the feasibility of problem $\mathscr{P}_N^{DFD}(x(t))$. This means that part (i) of Theorem 4.4.2 remains valid. (ii) Let $V^*(t)$ and $\hat{V}(t+1)$ be defined in the same manner as $J^*(t)$ and $\hat{J}(t+1)$ in the statement of proofs of Theorem 4.4.2. Following the same reasoning until (4.20), it can be shown that

$$V(t)^* - \hat{V}(t+1) = \|d^*(0|t)\|_\Psi^2 + p(w(t)) \tag{4.31}$$

where

$$\begin{aligned} p(w(t)) \quad = \quad & \sum_{i=1}^{N-1} (\gamma_1 \|\mathrm{vec}(D^*(i,i|t))\|^2 + \gamma_2 \|\mathrm{vec}(D^*(i,i|t))\| \\ & -2(d^*(i|t))^T \Psi D^*(i,i|t) w(t) - \|D^*(i,i|t) w(t)\|_\Psi^2). \end{aligned} \tag{4.32}$$

Hence

$$
\begin{aligned}
& p(w(t)) \\
\geq & \sum_{i=1}^{N-1} (\gamma_1 \|\mathrm{vec}(D^*(i,i|t))\|^2 + \gamma_2 \|\mathrm{vec}(D^*(i,i|t))\| - 2\|d^*(i|t)\|\|\Psi\|\|w(t)\|\|(D^*(i,i|t))\| \\
& \qquad -\|\Psi\|\|w(t)\|^2\|(D^*(i,i|t))\|^2) \\
\geq & \sum_{i=1}^{N-1} (\gamma_1 \|\mathrm{vec}(D^*(i,i|t))\|^2 + \gamma_2 \|\mathrm{vec}(D^*(i,i|t))\| - 2\alpha\beta\|\Psi\|\|(D^*(i,i|t))\| \\
& \qquad -\alpha^2\|\Psi\|\|(D^*(i,i|t))\|^2) \\
\geq & \sum_{i=1}^{N-1} (\gamma_1 \|\mathrm{vec}(D^*(i,i|t))\|^2 + \gamma_2 \|\mathrm{vec}(D^*(i,i|t))\| - 2\alpha\beta\|\Psi\|\|(D^*(i,i|t))\|_F \\
& \qquad -\alpha^2\|\Psi\|\|(D^*(i,i|t))\|_F^2) \qquad (4.33) \\
= & \sum_{i=1}^{N-1} (\gamma_1 \|\mathrm{vec}(D^*(i,i|t))\|^2 + \gamma_2 \|\mathrm{vec}(D^*(i,i|t))\| - 2\alpha\beta\|\Psi\|\|\mathrm{vec}(D^*(i,i|t))\| \\
& \qquad -\alpha^2\|\Psi\|\|\mathrm{vec}(D^*(i,i|t))\|^2) \qquad (4.34) \\
= & \sum_{i=1}^{N-1} ((\gamma_1 - \alpha^2\|\Psi\|)\|\mathrm{vec}(D^*(i,i|t))\|^2 + (\gamma_2 - 2\alpha\beta\|\Psi\|)\|\mathrm{vec}(D^*(i,i|t))\|) \\
\geq & \ 0 \qquad (4.35)
\end{aligned}
$$

where $\|D^*(i,i|t)\|_F$ is the Frobenius norm of $D^*(i,i|t)$ and the facts $\|D^*(i,i|t)\| \leq \|D^*(i,i|t)\|_F$ and $\|D^*(i,i|t)\|_F = \|\mathrm{vec}(D^*(i,i|t))\|$ are used in (4.33) and (4.34). Hence, $p(w(t)) \geq 0$ under (4.18). As a consequence, equation (4.31) implies

$$V^*(t) - \|d^*(0|t)\|_\Psi^2 \geq V^*(t+1) \geq 0 \qquad (4.36)$$

Hence, $\{V^*(t)\}$ is a monotonic non-increasing sequence and is bounded from below by zero. This means that $V_\infty := \lim_{t\to\infty} V^*(t) \geq 0$ exists. Repeating (4.36) for t from 0 to ∞ and summing them up, it follows that

$$\infty > V^*(0) - V_\infty \geq \sum_{t=0}^{\infty} \|d^*(0|t)\|_\Psi^2 \tag{4.37}$$

Since Ψ is positive definite, this implies that $\lim_{t\to\infty} d^*(0|t) = 0$ and $\lim_{t\to\infty} u(t) = K_f x(t)$. Therefore, the stated result follows.

Chapter 5

Segregated Disturbance Feedback Parametrization

This chapter proposes a new control parametrization for MPC of constrained linear discrete time systems with bounded additive disturbances. This parametrization takes the form of a special piecewise affine disturbance feedback and is a generalization of the affine disturbance feedback discussed in Chapters 3 and 4. Thus, the domain of attraction of the resulting closed-loop system using the proposed parametrization is expected to be larger than those using affine disturbance feedback. Properties and the numerical computation of MPC under the proposed parametrization are discussed. Under mild assumptions of the disturbance set, the associated FH optimization can be computed efficiently. Stability of the closed-loop system with the proposed parametrization is also ensured.

5.1 Introduction

Like Chapters 3 and 4, this chapter is concerned with the CTLD system

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

$$(x(t), u(t)) \in Y,\ w(t) \in W,\ \forall\, t \geq 0 \tag{5.1b}$$

where the variables have the usual meaning and the system satisfies Assumption 1.4.1. Additionally, $w(t)$ is assumed to satisfy the following assumption

Assumption 5.1.1

(A3b) *the disturbances $w(t)$ $t \geq 0$ are independent and identically distributed (i.i.d.).*

The rest of this chapter is organized as follows. Details of the new control parametrization and the MPC framework together with the cost function are given in Section 5.2. A convex reformulation and computational issues are introduced in Section 5.3. Section 5.4 discuss the feasibility of the FH optimization problem and stability of the closed-loop system. Numerical examples and summary are the contents of the last two sections.

5.2 Control Parametrization and MPC Framework

This chapter generalizes the results and the parametrization given in Chapter 4 (a similar generalization of the parametrization u^{DFC} in (3.2) of Chapter 3 is available in [76]).

5.2.1 Control Parametrization

The proposed control parametrization is a special piecewise affine function of disturbance w. To be precise, let $w \in \mathbb{R}^n$ be segregated into its positive and negative parts via

$$w^p(w) := \max\{w, 0\}, \;\; w^m(w) := \max\{-w, 0\} \tag{5.2}$$

where the max operation is taken component-wise, i.e., the ith element of $w^p(w)$, denoted by $w_i^p(w) = \max\{w_i, 0\}$. With this definition, it follows that for any $w \in \mathbb{R}^n$, $w^p(w), w^m(w) \in \mathbb{R}^n$, $w^p(w) \geq 0$, $w^m(w) \geq 0$ and $w = w^p(w) - w^m(w)$. The following values are also needed for the new parametrization.

$$\bar{w}^p := \mathrm{E}_w[w^p(w)], \;\; \bar{w}^m := \mathrm{E}_w[w^m(w)], \tag{5.3}$$

and define

$$\hat{w}^p(w) := w^p(w) - \bar{w}^p, \;\; \hat{w}^m(w) := w^m(w) - \bar{w}^m. \tag{5.4}$$

Clearly, $\hat{w}^p(w)$ and $\hat{w}^m(w)$ have zero means. Hereafter, the dependence of w in $w^p(w)$, $w^m(w)$, $\hat{w}^p(w)$, $\hat{w}^m(w)$ is omitted for notational convenience, except when warranted by context. Using $(\hat{w}^p, \hat{w}^m)$ and similar to parametrization u^{DFD} in (4.2), the proposed $u(i|t)$ takes the form

$$\begin{cases} u(i|t) = K_f x(i|t) + l(i|t), & \forall i \in \mathbb{Z}_{N-1} \\ l(i|t) = d(i|t) + \sum_{j=1}^{i} D^p(i,j|t)\hat{w}^p(i-j|t) + \sum_{j=1}^{i} D^m(i,j|t)\hat{w}^m(i-j|t) \end{cases} \tag{5.5}$$

where $d(i|t) \in \mathbb{R}^m$, $D^p(i,j|t)$, $D^m(i,j|t) \in \mathbb{R}^{m\times n}$ are the optimization variables, K_f is the specified state feedback gain in (A4) of Assumption 1.4.1 and the disturbances $\hat{w}^p(i-j|t)$ and $\hat{w}^m(i-j|t)$ are obtained from $w(i-j|t)$ using (5.4).

Like $w^p(w)$, $w^m(w)$, $u(i|t)$ in (5.5) is a special piecewise function of w and, because of the particular choice of the pieces, is termed *Segregated Disturbance Feedback*. Clearly, it is a generalization of linear disturbance feedback parametrization u^{DFD} in (4.2) (choose $D^p(i,j|t)$ and $D^m(i,j|t)$ in (5.5) to be $D^p(i,j|t) = -D^m(i,j|t) = D(i,j|t)$ and the same $d(i|t)$ to recover (4.2)) or those in (3.2) and (1.27). This is shown in the following example.

Example 5.2.1 *Consider the system*

$$x(t+1) = \begin{bmatrix} 0 & 0.5 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} w(t)$$

where $x(0) = [0;0]$, $w(t)$ *is independent, identical and uniformly-distributed over* $W =$

$\{w|\ |w| \leq 1\}$ and the constraint set $Y = \{(x,u)|\ x_1 \geq -1.1,\ x_2 \geq -1.1,\ u \geq -0.2\}$. Suppose it is required that the sequence $x(\cdot)$ has zero mean. Then, it can be verified that no linear disturbance feedback law in the form of u^{DFD}, u^{DFC} or u^{DF} (respectively (4.2), (3.2) or (1.27)) can simultaneously satisfy the constraints and the zero mean requirement. Since u^{DFD}, u^{DFC} and u^{DF} have the same expressive ability, consider u^{DF} as a representative case. Zero mean of $x(1)$ with $x(0) = [0;0]$ implies $u(0) = 0$. Also, then $u(1) = v + Mw(0)$. The zero mean requirement of $x(2)$ with $w(1)$ being zero-mean means that $\mathrm{E}[u(1)] = 0$ and hence $v = 0$, leading to $u(1) = Mw(0)$. The choice of M is further limited to $|M| \leq 0.2$ since $u \geq -0.2$ is a constraint and $|w| \leq 1$. The first component of $x(2)$ is $x_1(2) = w(1) + (0.5 + M)w(0)$ and no M exists that can satisfy $x_1(2) \geq -1.1$ and $|M| \leq 0.2$ simultaneously. However, the segregated disturbance feedback law, $u(0) = 0$, $u(i) = 0.5\hat{w}^m(i-1) - 0.1\hat{w}^p(i-1)$ is feasible to all constraints and requirement.

Let $\mathbf{x}(t)$, $\mathbf{u}(t)$ and $\mathbf{w}(t)$ be defined in the same way as in (3.4). Using (5.4), $\mathbf{w}(t)$ is separated into $\hat{\mathbf{w}}^p(t)$, $\hat{\mathbf{w}}^m(t)$. The rest of the variables in (5.5) are collected in

$$\mathbf{D}^p(t) := \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ D^p(1,1|t) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ D^p(N-2,N-2|t) & D^p(N-2,N-3|t) & \cdots & 0 & 0 \\ D^p(N-1,N-1|t) & D^p(N-1,N-2|t) & \cdots & D^p(N-1,1|t) & 0 \end{bmatrix},$$

$$\mathbf{D}^m(t) := \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ D^m(1,1|t) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ D^m(N-2,N-2|t) & D^m(N-2,N-3|t) & \cdots & 0 & 0 \\ D^m(N-1,N-1|t) & D^m(N-1,N-2|t) & \cdots & D^m(N-1,1|t) & 0 \end{bmatrix},$$

$$\mathbf{d}(t) := \begin{bmatrix} d(0|t) \\ d(1|t) \\ \vdots \\ d(N-1|t) \end{bmatrix}.$$

Using these notations, the control parametrization of (5.5) within the control horizon becomes

$$\mathbf{u}(t) = \mathscr{K}\mathbf{x}(t) + \mathbf{d}(t) + \mathbf{D}^m(t)\hat{\mathbf{w}}^m(t) + \mathbf{D}^p(t)\hat{\mathbf{w}}^p(t) \tag{5.6}$$

where $\mathscr{K} := \begin{bmatrix} I_N \otimes K_f & \mathbf{0} \end{bmatrix}$.

5.2.2 MPC Formulation

Using the above-mentioned notations, the FH optimization based on the control parametrization of (5.5) can be summarized as the following problem $\mathscr{P}_N^{SDF}(x(t))$:

$$
\begin{aligned}
&\min_{\mathbf{d}(t),\mathbf{D}(t)} \; J_{SDF}(\mathbf{d}(t),\ \mathbf{D}(t)) && \text{(5.7a)}\\
\text{s.t.}\quad & \mathbf{x}(t) = \mathscr{A}x(t) + \mathscr{B}\mathbf{u}(t) + \mathscr{G}(\mathbf{w}^p(t) - \mathbf{w}^m(t)) && \text{(5.7b)}\\
& \mathbf{u}(t) = \mathscr{K}\mathbf{x}(t) + \mathbf{d}(t) + \mathbf{D}^m(t)\hat{\mathbf{w}}^m(t) + \mathbf{D}^p(t)\hat{\mathbf{w}}^p(t) && \text{(5.7c)}\\
& (x(i|t), u(i|t)) \in Y, \qquad \forall\, \mathbf{w}(t) \in W^N,\ \forall i \in \mathbb{Z}_{N-1} && \text{(5.7d)}\\
& x(N|t) \in X_f, \qquad \forall\, \mathbf{w}(t) \in W^N && \text{(5.7e)}
\end{aligned}
$$

where $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{G}$ are the same as in (3.9) and (3.10), $J_{SDF}(\mathbf{d}(t),\ \mathbf{D}(t))$ is an appropriate cost function details of which are discussed in the next subsection. Let the feasible set of the FH optimization problem be

$$\Theta_N^{SDF}(x) = \{(\mathbf{d},\mathbf{D}) \mid (\mathbf{d},\mathbf{D}) \text{ is feasible for } \mathscr{P}_N^{SDF}(x)\} \tag{5.8}$$

The set of admissible initial states to the FH problem is then

$$\mathscr{X}_N^{SDF} := \{x|\ \Theta_N^{SDF}(x) \neq \varnothing\}. \tag{5.9}$$

Remark 5.2.1 *Since by letting $D^p(i,j|t) = -D^m(i,j|t) = D(i,j|t)$ parametrization (5.5) becomes (4.2), the expressive ability of the former is stronger than that of the latter. Hence, $\mathscr{X}_N^{DFD} \subseteq \mathscr{X}_N^{SDF}$ and this, together with Remark 4.2.5, implies $\mathscr{X}_N^{DF} = \mathscr{X}_N^{DFC} = \mathscr{X}_N^{DFD} \subseteq \mathscr{X}_N^{SDF}$.*

The rest is as usually: the FH optimization problem is solved at each time t and the very first term of $(\mathbf{d}^*(t), \mathbf{D}^*(t)) = \arg\min\ \mathscr{P}_N^{SDF}(x(t))$ is applied to system (5.1a) yielding the MPC control law

$$u(t) = \kappa^{SDF}(x(t)) := K_f x(t) + d^*(0|t) \tag{5.10}$$

5.2.3 Cost Function

The cost function used in this chapter is similar to that used in Chapter 4 and hence its discussion here is brief. Specifically, the cost function is

$$J_{SDF}(\mathbf{d}(t), \mathbf{D}(t)) := \sum_{i=0}^{N-1} \left[\|d(i|t)\|_\Psi^2 + \sum_{j=1}^{i} \|\mathrm{vec}([D^p(i,j|t)\ \ D^m(i,j|t)])\|_\Lambda^2 \right] \tag{5.11}$$

for any choice of

$$\Psi \succ 0,\ \ \Lambda \succeq \Sigma_v \otimes \Psi \tag{5.12}$$

where Σ_v is the covariance matrix of $[\hat{w}^p(w);\ \hat{w}^m(w)] := [(\hat{w}^p(w))^T\ \ (\hat{w}^m(w))^T]^T$.

If the density function of w is known, Σ_v can be found. However, the knowledge of Σ_v is not needed to satisfy (5.12). For example, if Σ_v is not known, let $\Lambda = 2\alpha^2 I_{2n} \otimes \Psi$ where $\alpha := \max_{w \in W} \|w\|$ and I_{2n} is the identity matrix of dimension $2n$. Then, it can be shown, in Appendix 5.A.1, that (5.12) is satisfied.

5.3 Convex Reformulation and Computation

This section focuses on the computation of the FH optimization problem $\mathscr{P}_N^{SDF}(x)$. Since (5.7d) and (5.7e) have to hold for all $w(i|t) \in W$ and $(x(i|t), u(i|t))$ are piecewise affine functions of $w(j|t)$ $j < i$, $\mathscr{P}_N^{SDF}(x)$ is not directly solvable using standard techniques in Robust Optimization and a reformulation is needed. To this end, define

$$v := [w^p;\ w^m] = [(w^p)^T \ \ (w^m)^T]^T \tag{5.13}$$

which belongs to the set

$$V_W := \{v = [v^1;\ v^2] |\ v^1 - v^2 \in W,\ v \geq 0,\ (v^1)^T v^2 = 0\} \subset \mathbb{R}^{2n}. \tag{5.14}$$

The condition of $(v^1)^T v^2 = 0$ comes from definition (5.2) and it implies, under (A3) of Assumption 1.4.1, that V_W is a non-convex set even when W is convex. Clearly, there is a one-to-one mapping between V_W and W: for any $w \in W$, $v^1 = w^p$ and $v^2 = w^m$ while for any $[v^1; v^2] \in V_W$, $w = v^1 - v^2$. Define $\mathbf{v}^1(t)$ and $\mathbf{v}^2(t)$ in the same structure as $\mathbf{w}(t)$ and let

$$\bar{\mathbf{v}}^p := \mathbf{1}_N \otimes \bar{w}^p,\ \bar{\mathbf{v}}^m := \mathbf{1}_N \otimes \bar{w}^m,$$

then $\mathscr{P}_N^{SDF}(x(t))$ of (5.7) can be equivalently formulated as

$$\min_{\mathbf{d}(t),\mathbf{D}(t)} \; J_{SDF}(\mathbf{d}(t),\, \mathbf{D}(t)) \tag{5.15a}$$

$$\text{s.t.} \quad \mathbf{x}(t) = \mathscr{A}x(t) + \mathscr{B}\mathbf{u}(t) + \mathscr{G}(\mathbf{v}^1(t) - \mathbf{v}^2(t)) \tag{5.15b}$$

$$\mathbf{u}(t) = \mathscr{K}\mathbf{x}(t) + \mathbf{d}(t) + \mathbf{D}^p(t)(\mathbf{v}^1(t) - \bar{\mathbf{v}}^p) + \mathbf{D}^m(t)(\mathbf{v}^2(t) - \bar{\mathbf{v}}^m) \tag{5.15c}$$

$$(x(i|t), u(i|t)) \in Y, \qquad \forall\, [\mathbf{v}^1(t);\, \mathbf{v}^2(t)] \in (V_W)^N,\ \forall i \in \mathbb{Z}_{N-1} \tag{5.15d}$$

$$x(N|t) \in X_f, \qquad \forall\, [\mathbf{v}^1(t);\, \mathbf{v}^2(t)] \in (V_W)^N \tag{5.15e}$$

Since V_W is generally non-convex, $\mathscr{P}_N^{SDF}(x)$ is still not computationally tractable by standard techniques because of (5.15d) and (5.15e). The next subsection shows an additional assumption on W and an associated definition that helps towards this end.

5.3.1 Absolute Set

Definition 5.3.1 (Absolute set) *A set Ω is an absolute set if it is compact, convex, contains the origin in its interior and $\omega \in \Omega$ if and only if $|\omega| \in \Omega$ where $|\cdot|$ is taken element-wise.*

Examples of absolute sets include those generated by the L_p norms and their intersections, e.g. $\{\omega : \|\omega\|_p \le a\}$, $\{\omega : \|\omega\|_\infty \le a, \|\omega\|_2 \le b, \|\omega\|_1 \le c\}$. The use of absolute set as disturbance model is also quite common [77, 62, 63, 78] and it is stated here as an assumption.

Assumption 5.3.1

(A3c) *W set is an absolute set.*

Theorem 5.3.1 *Suppose W satisfy Assumption 5.3.1. The set*

$$V_W^C := \{v = [v^1; v^2] | \; v^1 + v^2 \in W, \; v \geq 0\} \tag{5.16}$$

is the convex hull of V_W.

Proof: See Appendix 5.A.2. ∎

Figure 5.1 shows the sets W, V_W and V_W^C for a simple one dimensional $W = \{w| \; |w| \leq 0.2\}$. That V_W^C is the convex hull of V_W as stated in Theorem 5.3.1 can be easily verified.

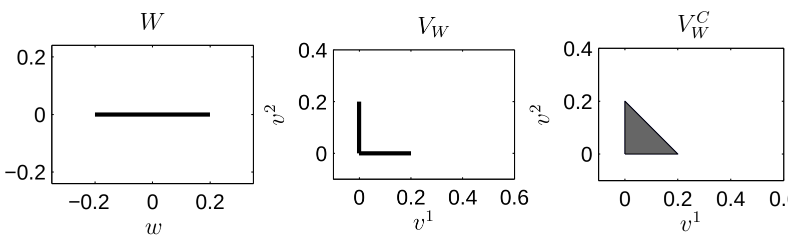

Figure 5.1: Disturbance set and segregated disturbance set.

Remark 5.3.1 *W being an absolute set is not as restrictive as it may appear. Many non-symmetrical disturbances or disturbances generated from a set with dimension different from $\mathbb{R}^n$ can be represented as $\{w | w = E\bar{w} + e, \bar{w} \in \bar{W} \subset \mathbb{R}^\ell\}$ where $\bar{W}$ is an absolute set* *and E and e are some appropriate matrices. For such disturbance models, the exposition hereafter remains valid but with w replaced by $E\bar{w}+e$.*

If W is a polytope set, so is V_W^C. It is well known that the solution of $\mathscr{P}_N^{SDF}(x)$ is unaffected by the replacement of V_W by V_W^C in (5.15d) and (5.15e) (Exercise 3.35 in [79]). Since V_W^C is a polytope, its use in (5.15d) and (5.15e) in place of V_W leads to a computable $\mathscr{P}_N^{SDF}(x)$ using standard techniques in robust optimization, see Section 3.3 and [39, 69, 68, 63, 44] for details.

However, it is possible for $\mathscr{P}_N^{SDF}(x)$ to handle a more general class of absolute sets other than polytopes. The next subsection introduces a definition of absolute norm, its dual and an additional result needed for this purpose.

5.3.2 Absolute Norm

Definition 5.3.2 (Absolute norm) *A function $\eta : \mathbb{R}^n \to \mathbb{R}$ is an absolute norm function if $\eta(\cdot)$ satisfies the three standard properties of a norm and the additional property of $\eta(v) = \eta(|v|)$.*

Clearly, all polynomial norms or L_p norms are absolute. However, a polynomial norm induced by an invertible matrix, is not necessarily absolute. Absolute norm function can also be defined from other absolute norm functions. For example,

$$\zeta(w) := \max_{l=1,\ldots,L} \{a_l \eta_l(w)\}, \tag{5.17}$$

in which $\eta_l(\cdot)$ are absolute norms with $a_l > 0$ for all $l \in \mathbb{Z}_L^+$, is absolute. Hence, $\{w: \|w\|_\infty \leq 1,\ \|w\|_2 \leq r\}$ can be expressed as $\{w: \eta(w) \leq 1\}$ with $\eta(w) = \max\{\|w\|_2/r,\ \|w\|_\infty\}$.

Given an absolute norm $\eta(\cdot)$, its dual norm, $\eta^*: \mathbb{R}^n \to \mathbb{R}$, is defined as

$$\eta^*(y) := \max_{\eta(w)\leq 1} y^T w. \tag{5.18}$$

Some useful and relevant properties of the dual norm are collected below.

Lemma 5.3.1 *Suppose $\eta(\cdot)$ and $\eta^*(\cdot)$ are an absolute norm and its dual. Then (i) $\eta^*(\cdot)$ is also an absolute norm function (ii) $\eta^{**}(\cdot) = \eta(\cdot)$. (iii) The dual norm of the L_p norm $\|\cdot\|_p$, $p \geq 1$, is the L_q norm $\|\cdot\|_q$ with $q = 1 + 1/(p-1)$. (iv) The dual norm of the composite norm (5.17) is $\zeta^*(y) = \min\left\{\sum_{l=1}^L \frac{1}{a_l}\eta^*(y^l)\ :\ \sum_{l=1}^L y^l = y, y^l \in \mathbb{R}^n\ \forall l \in \mathbb{Z}_L^+\right\}$.*

Proof: See Appendix 5.A.3. ∎

The following example demonstrates (iv) of Lemma 5.3.1. Let $W = \{w \in \mathbb{R}^2 |\ \|w\|_\infty \leq 0.2,\ \|w\|_2 \leq 0.224\}$ which is shown in Figure 5.2. The corresponding absolute norm is

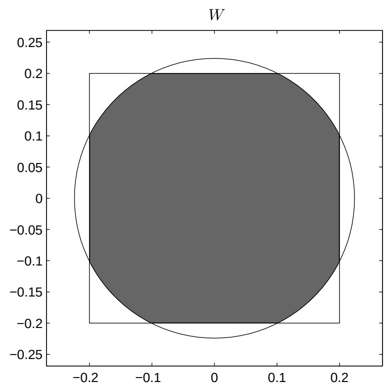

Figure 5.2: W defined by composite norm

$\eta(w) = \max\{5\|w\|_\infty, 4.64\|w\|_2\}$ and its dual norm is $\eta^*(z) = \min\{0.2\|y^1\|_1 + 0.224\|y^2\|_2 : y^1 + y^2 = z\}$.

Theorem 5.3.2 *A set W is an absolute set if and only if there exists an absolute norm function, $\eta_w(\cdot)$, such that $W = \{w : \eta_w(w) \leq 1\}$.*

Proof: See Appendix 5.A.4. ■

5.3.3 Deterministic Equivalence

One additional result, needed to show the deterministic equivalence of $\mathscr{P}_N^{SDF}(x)$, is given in the following theorem.

Theorem 5.3.3 *Let $W = \{w : \eta_w(w) \leq 1\} \subset \mathbb{R}^n$ be an absolute set for some absolute norm function $\eta_w(\cdot)$, $\eta_w^*(\cdot)$ be the corresponding dual norm and V_W^C be as defined by (5.16). The two sets*

$$\mathscr{C}_1 := \{(x,y,z) \in \mathbb{R}^{2n+1} \mid x^T v^1 + y^T v^2 \leq z, \ \forall\, [v^1;\, v^2] \in V_W^C\} \tag{5.19}$$

$$\mathscr{C}_2 := \{(x,y,z) \in \mathbb{R}^{2n+1} \mid \eta_w^*(\tau) \leq z, \ \tau \geq x, \ \tau \geq y \text{ for some } \tau\} \tag{5.20}$$

are equivalent.

Proof: See Appendix 5.A.5. ■

Remark 5.3.2 *Following the definition of the dual norm of (5.18), the constraint $\eta_w^*(\tau) \leq z$ arising in the expression for $\mathscr{C}_2$ in Theorem 5.3.3 is equivalent to*

$$w^T \tau \leq z \qquad \forall w \in W$$

whose tractability and explicit formulation can be found in [62]. In particular, if the set W is conic quadratic representable, which includes sets prescribed by intersections of L_p norms, p being a rational number, the resulting robust counterpart is also conic quadratic representable. The representative power of conic quadratic constraints can be found in [44]. Software involving conic quadratic representable constraints includes SDPT3 and MOSEK.

Under Assumption 5.3.1 and the result of Theorem 5.3.2, there exists an absolute norm function η_w such that $W = \{w|\ \eta_w(w) \leq 1\}$. Define $\eta_{w^N} : \mathbb{R}^{Nn} \to \mathbb{R}$ as $\eta_{w^N}(\mathbf{p}) := \max_{i \in \mathbb{Z}_{N-1}}\{\eta_w(p(i))\}$ where $\mathbf{p} := [p^T(1)\ p^T(2)\ \cdots\ p^T(N)]^T$ and $p(i) \in \mathbb{R}^n$. Then, it follows that W^N can be represented by $W^N = \{\mathbf{w}|\eta_{w^N}(\mathbf{w}) \leq 1\}$. The corresponding dual norm of $\eta_{w^N}(\cdot)$, $\eta^*_{w^N} : \mathbb{R}^{Nn} \to \mathbb{R}$, is given by

$$\eta^*_{w^N}(\mathbf{q}) = \max\{\mathbf{q}^T\mathbf{p}|\eta_{w^N}(\mathbf{p}) \leq 1\} = \sum_{i=1}^{N} \max\{(q(i))^T p(i)|\eta_w(p(i)) \leq 1\} = \sum_{i=1}^{N} \eta^*_w(q(i)) \tag{5.21}$$

where $\mathbf{q} := [q^T(1)\ \cdots\ q^T(N)]^T$ and $q(i) \in \mathbb{R}^n$.

Using these quantities and the characterizations of Y and X_f in Assumption 1.4.1, constraints (5.15b)-(5.15e) can be restated as

$$\bar{\mathscr{A}}x(t)+\bar{\mathscr{B}}\mathbf{d}(t)-\bar{\mathscr{B}}(\mathbf{D}^p(t)\bar{\mathbf{v}}^p+\mathbf{D}^m(t)\bar{\mathbf{v}}^m)+(\bar{\mathscr{B}}\mathbf{D}^p(t)+\bar{\mathscr{G}})\mathbf{v}^1(t)$$
$$+(\bar{\mathscr{B}}\mathbf{D}^p(t)-\bar{\mathscr{G}})\mathbf{v}^2(t)\leq 1_s,\ \forall\ [\mathbf{v}^1(t);\mathbf{v}^2(t)]\in(V_W^C)^N \quad (5.22)$$

where $s=qN+g$, $\bar{\mathscr{A}}$, $\bar{\mathscr{B}}$ and $\bar{\mathscr{G}}$ are the same as in (3.25) and (3.26).

To apply the result of Theorem 5.3.3, let

$$\tau(t):=1_s-\bar{\mathscr{A}}x(t)-\bar{\mathscr{B}}\mathbf{d}(t)+\bar{\mathscr{B}}(\mathbf{D}^p(t)\bar{\mathbf{v}}^p+\mathbf{D}^m(t)\bar{\mathbf{v}}^m)\in\mathbb{R}^s \quad (5.23)$$

Then (5.22) is equivalent to

$$\max_{[\mathbf{v}^1;\mathbf{v}^2]\in(V_W^C)^N}\left[(\bar{\mathscr{B}}\mathbf{D}^p(t)+\bar{\mathscr{G}})\mathbf{v}^1+(\bar{\mathscr{B}}\mathbf{D}^m(t)-\bar{\mathscr{G}})\mathbf{v}^2\right]\leq\tau(t). \quad (5.24)$$

There are s inequalities in (5.24), and each of them is in the same form as the inequality in (5.19), with $\bar{\mathscr{B}}\mathbf{D}^{p+}(t)+\bar{\mathscr{G}}$, $\bar{\mathscr{B}}\mathbf{D}^{m+}(t)-\bar{\mathscr{G}}$ and $\tau(t)$ analogous respectively to x^T, y^T and z. Then by introducing $T_i(t)\in\mathbb{R}^{nN}$, $i\in\mathbb{Z}_s^+$ (analogous to τ in (5.20)) for each row of (5.24) and applying result of Theorem 5.3.3, (5.24) is equivalent to

$$\begin{cases} \bar{\mathscr{A}}x(t)+\bar{\mathscr{B}}\mathbf{d}(t)-\bar{\mathscr{B}}(\mathbf{D}^p(t)\bar{\mathbf{v}}^p+\mathbf{D}^m(t)\bar{\mathbf{v}}^m)+\mu(t)\leq \mathbf{1}_s \\ \mathbf{T}^T(t)\geq \bar{\mathscr{B}}\mathbf{D}^p(t)+\bar{\mathscr{G}} \\ \mathbf{T}^T(t)\geq \bar{\mathscr{B}}\mathbf{D}^m(t)-\bar{\mathscr{G}} \\ \mu(t)=\left[\eta^*_{w^N}(T_1(t))\ \cdots\ \eta^*_{w^N}(T_s(t))\right]^T \\ \mathbf{T}(t)=[T_1(t)\ \cdots\ T_s(t)] \end{cases} \tag{5.25}$$

where $\eta^*_{w^N}(\cdot)$ is that given in (5.21).

Remark 5.3.3 *The adaptation of Theorem 5.3.3 to a disturbance set defined by the intersection of L_p norm sets is also quite easy. For example, if $W=\{w|\ \|w\|_\infty\leq 1,\ \|w\|_2\leq r\}$, then $\eta_w(w)=\max\{\frac{1}{r}\|w\|_2,\|w\|_\infty\}$, $\eta^*_w(w)=\min\{r\|w^1\|_2+\|w^2\|_1, w^1+w^2=w\}$ and the deterministic equivalence of $\mathscr{C}_1$ in Theorem 5.3.3 is $\mathscr{C}_2=\{(x,y,z)|\ \exists\tau,\tau^1,\tau^2\in\mathbb{R}^n, \|\tau^2\|_1+r\|\tau^1\|_2\leq z, \tau^1+\tau^2=\tau,\ \tau\geq x,\ \tau\geq y\}$. In such a case, (5.21) and (5.25) remains correct using the new expression of η^*_w.*

Remark 5.3.4 *For some class of disturbances where W is convex but not absolute and cannot be represented using Remark 5.3.1, the set V_W in (5.15) may be relaxed and be approximated by*

$$V_W^R=\{v=[v^1;v^2]|\ v^1-v^2\in W, v\geq 0\} \tag{5.26}$$

in which the constraint $(v^1)^T v^2\geq 0$ in V_W is removed. Therefore, $V_W\subseteq V_W^R$ and V_W^R

is convex (since W is convex). Suppose $\mathscr{P}_N^{SDFR}$ *and* $\mathscr{X}_N^{SDFR}$ *are the corresponding FH problem and the admissible initial set when* V_W *is replaced by* V_W^R *in (5.15d) and (5.15e). Then,* $\mathscr{P}_N^{SDFR}$ *is computationally more amiable as* V_W^R *is convex. Also, since* V_W *is defined by having an additional condition to* V_W^R*,* $V_W \subseteq V_W^R$*. Consequently,* $\mathscr{X}_N^{SDFR} \subseteq \mathscr{X}_N^{SDF}$ *as the control law obtained is more conservative.*

Remark 5.3.5 *While more conservative than* $\mathscr{P}_N^{SDF}$*,* $\mathscr{P}_N^{SDFR}$ *is less conservative than* $\mathscr{P}_N^{DFD}$*, the FH problem when parametrization (4.2) is used. Again, this is true because* $\mathscr{P}_N^{DFD}$ *is a special case of* $\mathscr{P}_N^{SDFR}$*. Hence, if a feasible solution exists for* $\mathscr{P}_N^{DFD}$ *for all* $w \in W$*, a feasible solution exists for* $\mathscr{P}_N^{SDFR}$ *for all* $v \in V_W^R$*. This, together with Remark 5.3.4, means that* $\mathscr{X}_N^{DFD} \subseteq \mathscr{X}_N^{SDFR} \subseteq \mathscr{X}_N^{SDF}$*.*

5.4 Feasibility and Stability

This section deals with the feasibility and stability of system (5.1a) under the control law (5.10). Since the results follow the same arguments as in Section 4.4, both the theorem and its proof are brief.

Theorem 5.4.1 *Suppose* $x(0) \in \mathscr{X}_N^{SDF}$ *and Assumptions 1.4.1, 5.1.1 and 5.3.1 are satisfied. The closed-loop system using the MPC control law (5.10) has the following properties:*

(i) $\mathscr{X}_N^{SDF}$ *and* $\Theta_N^{SDF}(x)$ *are convex sets;*

(ii) $\mathscr{P}_N^{SDF}(x(t))$ *is feasible for all* $t \geq 0$*;*

(iii) $(x(t), u(t)) \in Y$ *for all* $t \geq 0$;

(iv) $x(t) \to F_\infty$ as $t \to \infty$ *with probability one.*

Proof: See Appendix 5.A.6. ■

5.5 Numerical Examples

The proposed approach is illustrated using an example. The system parameters and constraints are given by:

$$A = \begin{bmatrix} 1.1 & 1 \\ 0 & 1.3 \end{bmatrix}, \ B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ W = \{w| \ \|w\|_\infty \leq 0.2\},$$

$$Y = \{(x,u)| \ |u| \leq 1, \|x\|_\infty \leq 6\},$$

and $w(t)$ is uniformly distributed over W. To implement the MPC controller, $K_f = [-0.4991 \ -0.9546]$ is chosen as the optimal feedback gain for the unconstrained LQR problem when $Q = I_2$ and $R = 1$. Terminal set X_f is chosen to be the maximal constraint-admissible disturbance invariant set of (5.1a) under $u = K_f x$. The proposed approach is simulated from $x(0) = [-5 \ 2]'$ for the case where $N = 8$ and its result is shown in Figure 5.3 and Figure 5.4.

In Figure 5.3, the outer bound $\hat{F}_\infty$ is used because the exact F_∞ is not computable. The

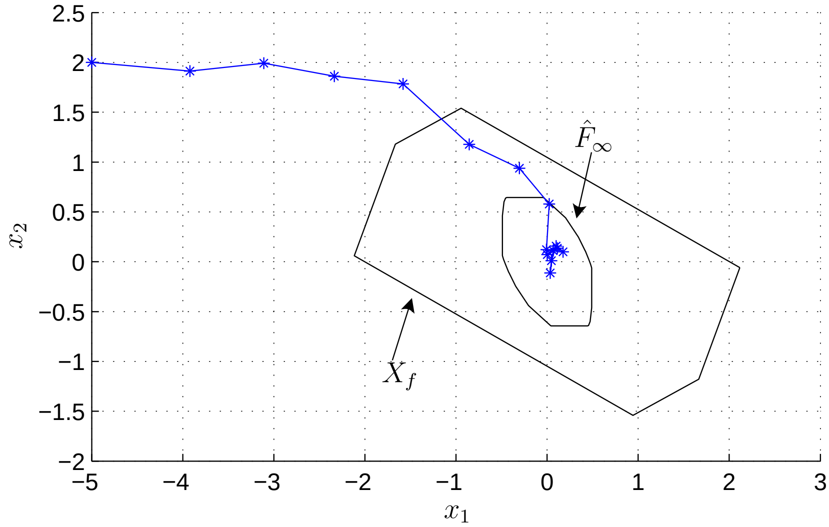

Figure 5.3: State trajectory of example one

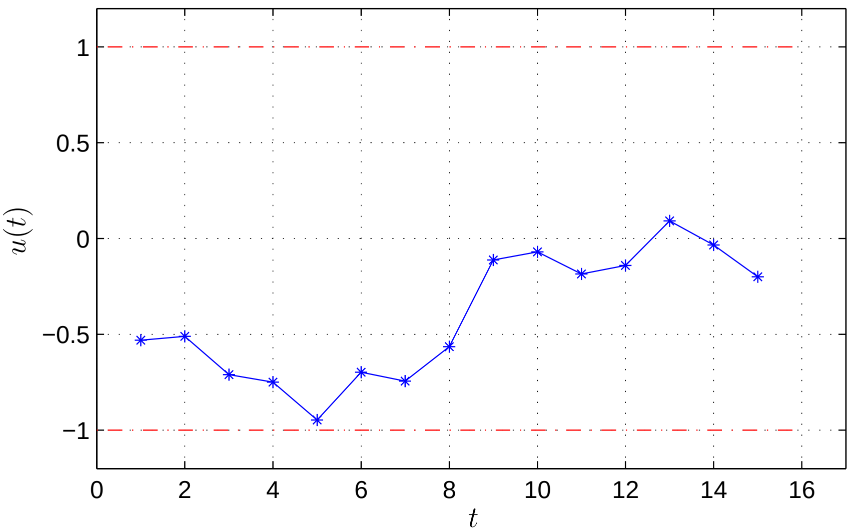

Figure 5.4: Control trajectory of example one

procedure for computing $\hat{F}_\infty$ follows that given in Section 2.2.1, also see [80]. It can be observed that the state converges to F_∞ and all the constraints are satisfied all the time.

The next simulation compares the optimal costs of the FH optimization problems using parameterizations (4.2) and (5.5) for the case where $N = 8$. For a fair comparison, the cost functions of the two parameterizations should be consistent. For this purpose, the weight matrices of $J_{SDF}(\mathbf{d},\mathbf{D})$ is chosen according to (5.12) while the cost function associated with parametrization (4.2) is chosen according to chapter 4. Our simulation with (4.2) uses the cost function $\sum_{i=0}^{N-1}\left[\|d(i|t)\|_\Lambda^2+\sum_{j=1}^{i}\|\mathrm{vec}(D(i,j|t))\|_\Upsilon^2\right]$ where $\Lambda = R + B^T PB$ and $\Upsilon = \Sigma_w \otimes \Lambda$. Under such choices, both cost functions are equivalent to the expected value of the same LQ cost. The optimal costs of both problems over the admissible region are compared and the result is shown in Figure 5.5 and Figure 5.6 where J_N^L is the optimal cost under parametrization (4.2) and J_N^S is that under (5.5). Clearly, parametrization (5.5) always yields a better cost than (4.2).

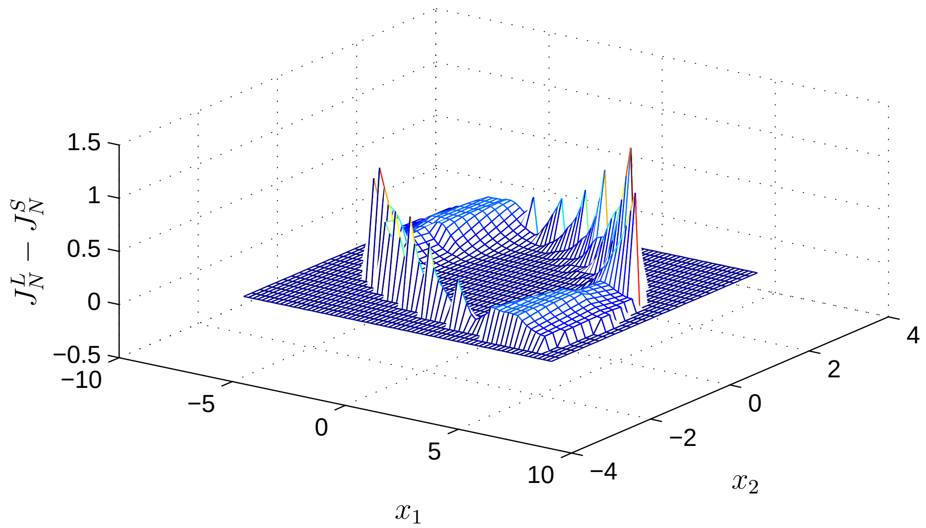

Figure 5.5: Difference between the two optimal costs

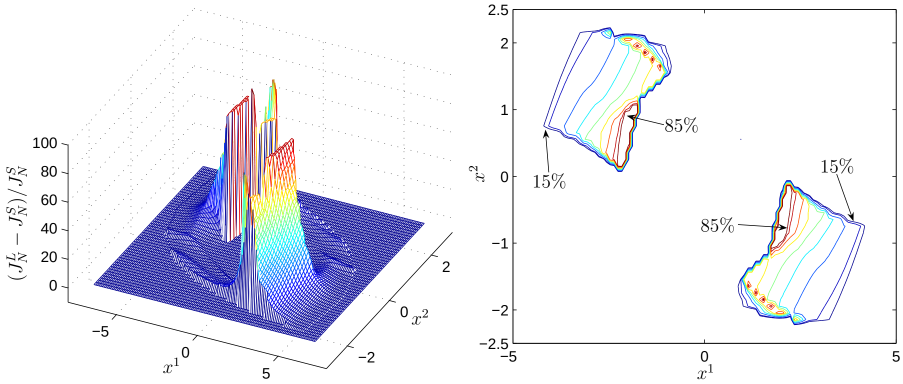

Figure 5.6: Plots of percentage of $\frac{J_N^L - J_N^S}{J_N^S}$ over the admissible set

5.6 Summary

A piecewise affine disturbance feedback parametrization is proposed under the MPC formulation of constrained linear systems with disturbances. This parametrization includes the standard affine disturbance feedback as a special case, and hence, has a stronger expressive ability. When the disturbance set is absolute, the FH optimization problem can be converted into an equivalent convex problem solvable with existing solvers. Even when the disturbance set is not absolute, the new parametrization still results in a MPC controller that is less conservative than the one derived from using affine disturbance feedback. The stability of the closed-loop system under the proposed parametrization is shown and the asymptotic behavior of the system is characterized by $F_\infty(K_f)$ where K_f is a user-defined constant feedback gain.

5.A Appendix

5.A.1 Choice of Λ

$$
\begin{aligned}
\Sigma_v &= \mathrm{E}_w \left[\begin{bmatrix} \hat{w}^p(w) - \bar{w}^p \\ \hat{w}^m(w) - \bar{w}^m \end{bmatrix} \begin{bmatrix} \hat{w}^p(w) - \bar{w}^p \\ \hat{w}^m(w) - \bar{w}^m \end{bmatrix}^T \right] \\
&= \mathrm{E}_w \left[\begin{bmatrix} \hat{w}^p(w) \\ \hat{w}^m(w) \end{bmatrix} \begin{bmatrix} \hat{w}^p(w) \\ \hat{w}^m(w) \end{bmatrix}^T \right] - \begin{bmatrix} \bar{w}^p \\ \bar{w}^m \end{bmatrix} \begin{bmatrix} \bar{w}^p \\ \bar{w}^m \end{bmatrix}^T \\
&\prec \mathrm{E}_w \left[\begin{bmatrix} \hat{w}^p(w) \\ \hat{w}^m(w) \end{bmatrix} \begin{bmatrix} \hat{w}^p(w) \\ \hat{w}^m(w) \end{bmatrix}^T \right] \preceq \max_{w \in W} \left\| \begin{matrix} \hat{w}^p(w) \\ \hat{w}^m(w) \end{matrix} \right\|^2 I_{2n} \\
&\preceq (\max_{w \in W} \|\hat{w}^p(w)\|^2 + \max_{w \in W} \|\hat{w}^m(w)\|^2) I_{2n} \\
&\preceq 2\alpha^2 I_{2n}
\end{aligned}
$$

5.A.2 Proof of Theorem 5.3.1

Proof: ($\Rightarrow$)Consider $[v^1; v^2] \in V_W$. It follows that $v^1 \geq 0, v^2 \geq 0$ and $(v^1)^T v^2 = 0$. Therefore, $v^1 + v^2 = |v^1 - v^2|$. Since W is absolute and $v^1 - v^2 \in W$, we have $v^1 + v^2 = |v^1 - v^2| \in W$ which implies that $[v^1; v^2] \in V_W^C$. Since the set V_W^C is convex, we have $\mathrm{CH}(V_W) \subseteq V_W^C$.

($\Leftarrow$) To show $V_W^C \subseteq \mathrm{CH}(V_W)$, consider $[u^1;u^2] \in V_W^C$ and let $S^0 = \{[u^1;u^2]\}$. For all $i \in \mathbb{Z}_n^+$, let

$$S^i = \bigcup_{[v^1;v^2]\in S^{i-1}} \{[v^1 - e^i v_i^1; v^2 + e^i v_i^1], [v^1 + e^i v_i^2; v^2 - e^i v_i^2]\}$$

where e^i denotes a unit vector in $\mathbb{R}^n$, with one at the i^{th} element and zeros otherwise. Observe that for all $[v^1;v^2] \in S^i$, $[v^1;v^2] \in \mathrm{CH}(S^{i+1})$. Indeed, if $v_i^1 + v_i^2 > 0$, let $\lambda = v_i^1/(v_i^1 + v_i^2)$ and it follows that

$$[v^1;v^2] = \lambda[v^1 - e^i v_i^1; v^2 + e^i v_i^1] + (1-\lambda)[v^1 + e^i v_i^2; v^2 - e^i v_i^2].$$

Otherwise, if $v_i^1 + v_i^2 = 0$, we have $[v^1;v^2] \in S^{i+1}$. Therefore, by induction, we have $[u^1;u^2] \in \mathrm{CH}(S^n)$. We can also induce that each $[v^1;v^2] \in S^n$ satisfies $v^1, v^2 \geq 0$, $v^1 + v^2 = u^1 + u^2$ and $v_i^1 v_i^2 = 0$, $i \in \mathbb{Z}_n^+$. Hence, $|v^1 - v^2| = v^1 + v^2 = u^1 + u^2 \in W$. Since W is an absolute set, we have $v^1 - v^2 \in W$ and $[v^1;v^2] \in V_W$. Therefore, $[u^1;u^2] \in \mathrm{CH}(V_W)$. ∎

5.A.3 Proof of Lemma 5.3.1

(i) The proof can be found in [81].

(ii)-(iii)The first two results are well known, see [82].

(iv): From (5.17),

$$
\begin{aligned}
\zeta^*(y) &= \max\{y^T x | a_i \eta_i(x) \le 1, i \in \mathbb{Z}_L^+\} \\
&= \max(y^T x | x \in \bar{C}_1 \cap \cdots \bar{C}_L\}) \\
&= \min\{\delta(y^1|\bar{C}_1) + \cdots + \delta(y^L|\bar{C}_L) | \sum_{i=1}^{L} y^i = y\} && (5.27) \\
&= \min\{\sum_{i=1}^{L} \frac{1}{a_i} \delta(y^i|C_i) | y^1 + \cdots + y^L = y\} && (5.28) \\
&= \min\{\sum_{i=1}^{L} \frac{1}{a_i} \eta^*(y^i) | y^1 + \cdots + y^L = y\} && (5.29)
\end{aligned}
$$

where $\bar{C}_i = \frac{1}{a_i} C_i$ and $C_i = \{x | \eta_i(x) \le 1\}$ $\forall i \in \mathbb{Z}_L^+$. Equation (5.27) follows a property of support function (Corollary 16.4.1 of [53]). Specifically, suppose $C_1, C_2, \cdots C_m$ are non-empty convex sets in $\mathbb{R}^n$ such that $C_1 \cap C_2 \cdots \cap C_m \neq \emptyset$, then $\delta(y | C_1 \cap C_2 \cdots \cap C_m) = \min\{\delta(y^1|C_1) + \cdots + \delta(y^m|C_m) | y^1 + \cdots + y^m = x\}$. Equation (5.28) holds because $\delta(x|\alpha C) = \alpha \delta(x|C)$ for any $\alpha > 0$ while (5.29) follows from the definition of C_i.

5.A.4 Proof of Theorem 5.3.2

One direction is trivial. Conversely, it suffices to show that for any absolute set, V, an absolute norm function $\eta(\cdot)$ exists such that

$$\max\{y^T v \,:\, \eta(v) \le 1\} = \max\{y^T v \,:\, v \in V\}$$

for all $y \in \mathbb{R}^n$. Consider the support function of V, $\delta(y|V)$, which by inspection is an absolute norm. Note that any convex, compact and symmetric set with zero in the interior has support function that satisfies the properties of a norm. Hence, from (5.18) and property (i) of *Lemma* 5.3.1, the corresponding dual norm function $\delta^*(v) = \max\{v^T y \,:\, \delta(y|V) \leq 1\}$ is also an absolute norm function. Let $\eta(\cdot) = \delta^*(\cdot)$. Hence, by strong duality of norms, we have for all $y \in \mathbb{R}^n$,

$$\max\{y^T v : \eta(v) \leq 1\} = \eta^*(y) = \delta^{**}(y) = \delta(y|V) = \max\{y^T v : v \in V\}. \tag{5.30}$$

5.A.5 Proof of Theorem 5.3.3

Proof: ($\Rightarrow$) Let (x,y,z) be an element of $\mathscr{C}_1$. It follows from (5.16) that

$$\begin{aligned}
z \geq \quad & \max\{x^T v^1 + y^T v^2 | v^1 \geq 0, v^2 \geq 0, \eta_w(v^1 + v^2) \leq 1\} \\
= \quad & \max\{x^T v^1 + y^T v^2 |\; v^1 \geq 0, v^2 \geq 0,\; u = v^1 + v^2,\; \eta_w(u) \leq 1\} \\
= \quad & \max\{\bar{\tau}^T u |\; u \geq 0,\; \eta_w(u) \leq 1,\; \bar{\tau}_i = \max\{x_i, y_i\},\; \forall i \in \mathbb{Z}_n^+ \} && \text{(5.31a)} \\
= \quad & \max\{\bar{\tau}^T |u| \mid \eta_w(u) \leq 1,\; \bar{\tau}_i = \max\{x_i, y_i\},\; \forall i \in \mathbb{Z}_n^+\} && \text{(5.31b)} \\
= \quad & \max\{\tau^T |u| \mid \eta_w(u) \leq 1,\; \tau_i = \max\{0, \bar{\tau}_i\},\; \forall i \in \mathbb{Z}_n^+\} && \text{(5.31c)} \\
= \quad & \max\{\tau^T u |\; \eta_w(u) \leq 1,\; \tau_i = \max\{0, \bar{\tau}_i\},\; \forall i \in \mathbb{Z}_n^+\} && \text{(5.31d)} \\
\Rightarrow \quad & (x,y,z) \in \mathscr{C}_2
\end{aligned}$$

where the subscript i denotes the i^{th} element of the corresponding vector. The first two relations come from the definitions of V_W^C, W and the re-organization of the constraints. Equation (5.31a) comes from the fact that the optimal value can be achieved by considering v^1 and v^2 where $v_i^1 v_i^2 = 0$ for all i. This is true because the optimal u^* is such that $u_i^* = v_i^{1*}$ if $x_i > y_i$ and $u_i^* = v_i^{2*}$ if $x_i \leq y_i$ for all i. Equation (5.31b) follows because W is an absolute set. Equation (5.31c) comes from the fact that if $\tau_i < 0$, the optimal u_i^* must be 0. Hence, the maximum value can be obtained by letting $\tau_i = \max\{0, \bar{\tau}_i\}$. Since $\tau \geq 0$, the absolute sign on u can be relaxed based on property (1) of Lemma 5.3.1. The last implication follows since the existence of τ, $\tau \geq x$ and $\tau \geq y$ is established.

($\Leftarrow$) Let (x, y, z) be an element of $\mathscr{C}_2$ with a suitable $\tau \in \mathbb{R}^n$. Then, from the definition of $\eta_w^*(\cdot)$,

$$
\begin{aligned}
z &\geq \max\{\tau^T (v^1 + v^2) |\ (v^1 + v^2) \in W, \tau \geq x, \tau \geq y\} \\
&\geq \max\{\tau^T (v^1 + v^2) |\ (v^1 + v^2) \in W, v^1 \geq 0, v^2 \geq 0,\ \tau \geq x, \tau \geq y\} \\
&\geq \max\{x^T v^1 + y^T v^2 |\ (v^1 + v^2) \in W, v^1 \geq 0, v^2 \geq 0\} \\
&= \max\{x^T v^1 + y^T v^2 |\ [v^1; v^2] \in V_W^C\} \\
&\Rightarrow (x, y, z) \in \mathscr{C}_1.
\end{aligned}
$$

Again, the first inequality holds from definition. The second inequality follows from the imposition of two additional constraints $v^1 \geq 0, v^2 \geq 0$. The third inequality follows from the fact that $\tau^T v^1 \geq x^T v^1$ and $\tau^T v^2 \geq y^T v^2$ for all $v^1, v^2 \geq 0$ since $\tau \geq x$ and $\tau \geq y$.

The last equality is from the definition of V_W^C which implies the inclusion. ∎

5.A.6 Proof of Theorem 5.4.1

(i) From (A2) of Assumption 1.4.1 and property (i) of Lemma 5.3.1, $\eta^*_{w^N}(\cdot)$ is a convex function. This means that the set of feasible $(x, \mathbf{d}, \mathbf{D}, \mathbf{T})$ of (5.25) is a convex set. The sets $\Theta_N^{SDF}(x)$ and $\mathscr{X}_N^{SDF}$ are projections of this convex set onto the $(\mathbf{d}, \mathbf{D})$ and x space respectively and are hence convex sets.

The proof of (ii)-(v) follows essentially the arguments in section 4.4.

(ii)If $(\mathbf{d}^*(t), \mathbf{D}^*(t))$ is the optimal control at time t, choose the feasible control at time $t+1$ by

$$\hat{d}(i|t+1) = \begin{cases} d^*(i+1|t) + (D^p(i+1, i+1|t))^* \hat{w}^p(t) \\ \qquad\qquad + (D^m(i+1, i+1|t))^* \hat{w}^m(t) & \forall i \in \mathbb{Z}_{N-2} \\ 0 & i = N-1 \end{cases} \tag{5.32a}$$

$$\hat{D}^k(i, j|t+1) = \begin{cases} (D^k(i+1, j|t))^* & \forall j \in \mathbb{Z}_i^+,\ \forall i \in \mathbb{Z}_{N-2}^+ \\ 0 & \forall j \in \mathbb{Z}_{N-1}^+,\ i = N-1 \end{cases} \quad \forall k \in \{p, m\} \tag{5.32b}$$

The feasibility of $(\hat{\mathbf{d}}(t+1), \hat{\mathbf{C}}(t+1))$ for $\mathscr{P}_N(x(t+1))$ follows from constraint (5.7e) and (A4) of Assumption 1.4.1.

(iii) The result follows directly from (ii).

(iv) Note that $(\hat{w}^p, \hat{w}^m)$ have zero mean. Hence, following the same argument in the proof of Theorem 4.4.2, it can be also shown that under condition (5.12) the expected value of the optimum of $\mathscr{P}_N^{SDF}(x(t))$ decreases with respect to time. Hence, the difference between the costs at successive times, which is $\|d^*(0|t)\|_\Psi$, converges to zero with probability one. As a result, the feedback law in (5.10) converges to $K_f x(t)$, leading to the stated result.

Chapter 6

Design of Feedback Gain

The stability results of the MPC method in chapters 3, 4 and 5 show that the closed-loop system state converges to the minimal d-invariant set $F_\infty(K_f)$ where K_f is some fixed state feedback gain. This chapter introduces a procedure for designing the terminal feedback gain K_f that characterizes the $F_\infty(K_f)$ set.

6.1 Introduction and Problem Statement

The theory of set invariance plays a fundamental role in the control of constrained dynamical systems, see [5] by Blanchini and [59] by Aubin, and a large number of works are based on set invariance, see [6] by Mayne *et al.*, [37] by Kolmanovsky and Gilbert, [61] by Gilber and Tan, [54] by Mayne and Schroeder, [83] by Caravani and De Santis, [84] by Raković *et al.*, [85] by Blanchini, [86] by Dórea and Hennet, [87] by De Santis *et al.*, [56] by Raković *et al.* and the references therein. However, most of these works are on the determination of an invariant set for a given system. Little attention has been paid to the research along the reverse direction, i.e. the design of system parameters that affect the invariant set. This is important since the invariant sets of one particular system may differ dramatically from each other under different controllers and this is shown by the following example.

Example 6.1.1 *Consider the system* $x(t+1) = Ax(t) + Bu(t) + w(t)$ *with*

$$A = \begin{bmatrix} 1.2 & 2 \\ 0 & 1.5 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}, \; w(t) \in W = \{w| \ \|w\|_\infty \leq 0.2\}$$

and three linear feedback laws:

$$K_{LQR} = [-0.6232 \quad -1.9678], \; K_{PP1} = [-0.1714 \quad -1.6571], \; K_{PP2} = [-3.3 \quad -0.4]$$

where K_{LQR} *is obtained using the LQR method with* $Q = I_2$, $R = 1$, K_{PP1} *and* K_{PP2} *are designed using pole placement method with poles being* $[0.8 \quad 0.9]$ *and* $[0.1 \quad -0.9]$, *respectively. The outer approximations of the minimal d-invariant sets of the system under these three linear feedback controllers are plotted in Figure 6.1. It can be seen that the minimal d-invariant sets are dramatically different.*

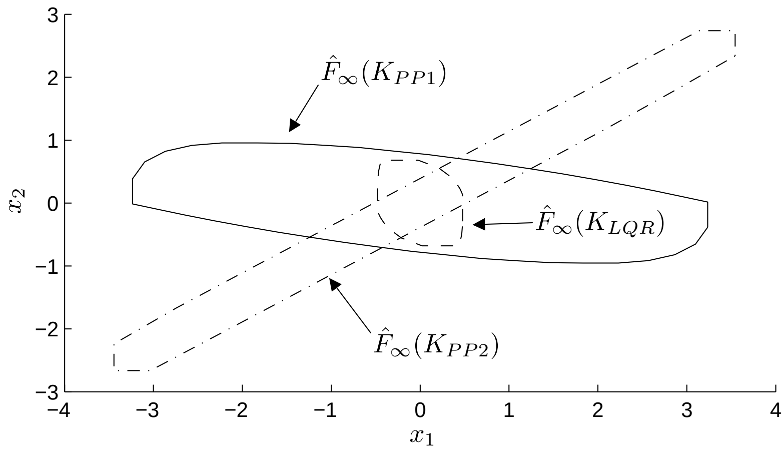

Figure 6.1: F_∞ sets under different controllers

This chapter emphasizes the design of feedback laws and consideres the following system,

$$x(t+1) = Ax(t) + Bu(t) + w(t) \tag{6.1a}$$

$$u(t) = Kx(t) \tag{6.1b}$$

$$w(t) \in W \tag{6.1c}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $K \in \mathbb{R}^{m\times n}$, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $w(t) \in \mathbb{R}^n$ is bounded in a convex compact set W that contains the origin in its interior. Denote the minimal d-invariant set of system (6.1) by $F_\infty(K)$. Note that any convex compact set can be approximated arbitrarily accurately by an appropriate polytope and any polytope can be expressed by the support function evaluated in some directions. The rest of this chapter shows the evaluation, with arbitrary accuracy, of the support function of $F_\infty(K)$,

$$\delta_{F_\infty(K)}(\eta) := \max\{\eta^T x |\ x \in F_\infty(K)\},$$

its derivative with respect to feedback gain K,

$$\frac{\partial \delta_{F_\infty(K)}(\eta)}{\partial K}$$

for any given $\eta \in \mathbb{R}^n$ and the design of K, using $\delta_{F_\infty(K)}(\eta)$ and $\partial \delta_{F_\infty(K)}(\eta)/\partial K$.

6.2 Support Function of $F_\infty(K)$ and Its Derivative

Without loss of generality, consider system (6.1) with $m = 1$, $B = b \in \mathbb{R}^{n\times 1}$ and $K = k \in \mathbb{R}^{1\times n}$ for the convenience of presentation. It can be shown that the results in the rest of this chapter can be generalized to the multiple input case with minor changes. System (6.1) can be equivalently written as

$$x(t+1) = \Phi x(t) + w(t) \tag{6.2}$$

where $\Phi = A + bk$. As discussed in Section 2.2.1, the reachable set of $x(t)$ when $x(0) = 0$ is $F_t = W \oplus \Phi W \oplus \cdots \oplus \Phi^{t-1} W$. When time tends to infinity, the reachable set is the minimal d-invariant set, $F_\infty(k)$.

6.2.1 Evaluation of Support Function

Using the properties that (i) $\delta_{A\Omega}(y) = \delta_\Omega(A^T y)$, (ii) $\delta_{\Omega_1 \oplus \Omega_2}(y) = \delta_{\Omega_1}(y) + \delta_{\Omega_2}(y)$ from Property 2.1.1, and the expression of $F_\infty = W \oplus \Phi W \oplus \cdots$, given $\eta \in \mathbb{R}^n$, the support function of F_∞ can be expressed as

$$\begin{aligned}
\delta_{F_\infty}(\eta) &= \delta_W(\eta) + \delta_{\Phi W}(\eta) + \delta_{\Phi^2 W}(\eta) + \cdots && \text{(6.3a)}\\
&= \delta_W(\eta) + \delta_W(\Phi^T \eta) + \delta_W((\Phi^2)^T \eta) + \cdots && \text{(6.3b)}\\
&= \sum_{\ell=0}^{\infty} \delta_W((\Phi^\ell)^T \eta) && \text{(6.3c)}\\
&= \sum_{\ell=0}^{L} \delta_W((\Phi^\ell)^T \eta) + \sum_{\ell=L+1}^{\infty} \delta_W((\Phi^\ell)^T \eta) && \text{(6.3d)}
\end{aligned}$$

Hence, $\delta_{F_\infty}(\eta)$ can be approximated by the first term in (6.3d).

$$\delta_{F_L}(\eta) := \sum_{\ell=0}^{L} \delta_W((\Phi^\ell)^T \eta) \tag{6.4}$$

with a large enough integer L because the second term in (6.3d) is bounded by zero from below and bounded from above since

$$\delta_W((\Phi^\ell)^T \eta) \le \|\eta\| \cdot \max_{w \in W} \|w\| \cdot (\|\Phi^\ell\|) \le v_s \cdot (\rho(\Phi))^\ell \tag{6.5}$$

where $v_s = \|\eta\| \cdot \max_{w\in W} \|w\| \cdot M$, $\rho(\Phi)$ is the spectral radius of Φ and M is a constant such that $M(\rho(\Phi))^i \geq \|\Phi^i\|,\ \forall i \geq 0$. Hence, the last term in (6.3d) satisfies

$$0 \leq \sum_{\ell=L+1}^{\infty} \delta_W((\Phi^\ell)^T\eta) \leq \sum_{\ell=L+1}^{\infty} v_s \cdot (\rho(\Phi))^\ell = \frac{v_s \cdot (\rho(\Phi))^{L+1}}{1-\rho(\Phi)}. \tag{6.6}$$

Let the error of approximation be

$$\varepsilon_L := \frac{v_s \cdot (\rho(\Phi))^{L+1}}{1-\rho(\Phi)}, \tag{6.7}$$

then

$$\delta_{F_\infty}(\eta) \in [\delta_{F_L}(\eta),\ \ \delta_{F_L}(\eta) + \varepsilon_L] \tag{6.8}$$

For any $\varepsilon > 0$, L can be chosen to be $\lceil \ln_{\rho(\Phi)}(\varepsilon(1-\rho(\Phi))/v) \rceil - 1$ so that $\varepsilon_L \leq \varepsilon$.

Example 6.2.1 *Consider a system in the form of (6.1a)-(6.1b) with parameters*

$$A = \begin{bmatrix} 1.2 & 2 \\ 0 & 1.5 \end{bmatrix},\ \ b = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix},\ k = [-0.6232\ \ -1.9678]$$

$$w(t) \in W = \{w|\ \|w\|_\infty \leq 0.2\}.$$

Hence

$$\Phi = A + bk = \begin{bmatrix} 0.5768 & 0.0322 \\ -0.3116 & 0.5161 \end{bmatrix}$$

η *is chosen to be* $[1\ \ 1]^T$ *and* v_s *is* 1.6 *with* $M = 4$. *The approximations of* $\delta_{F_\infty}(\eta)$ *and* *the corresponding errors with different value of L are listed in the following table and plotted in Figure 6.2. From Figure 6.2, we can see that the approximation of* $\delta_{F_\infty}(\eta)$

$\eta = [1\ \ 1]^T$	$L=3$	$L=4$	$L=5$	$L=6$	$L=7$	$L=8$	$L=9$	$L=10$
$\delta_{F_L}(\eta)$	0.6748	0.7107	0.7344	0.7492	0.7582	0.7634	0.7664	0.7682
ε_L	0.3402	0.1887	0.1047	0.0581	0.0322	0.0179	0.0099	0.0055

Table 6.1: Approximation of $\delta_{F_\infty}(\eta)$ with $L = 3, \ldots, 10$

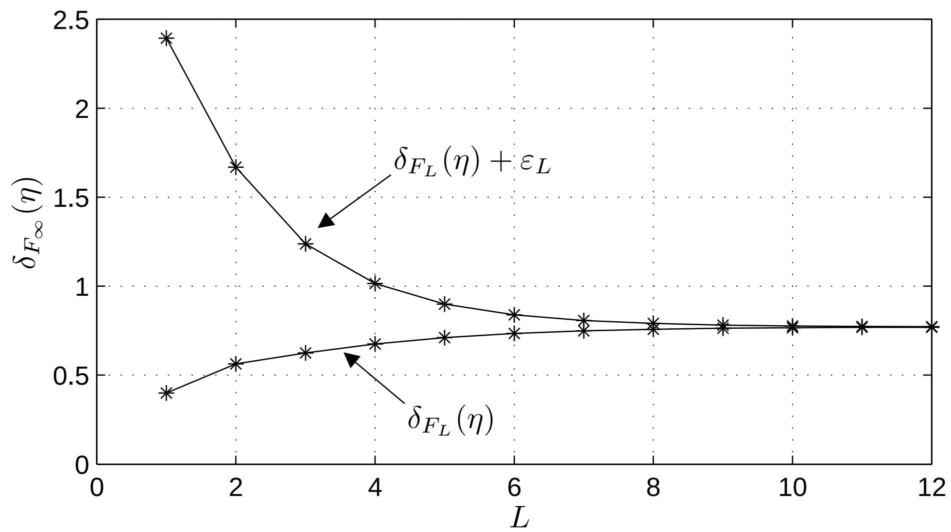

Figure 6.2: Approximation of $\delta_{F_\infty}(\eta)$ with different L

converges to 0.7705 *as L increases.*

The next section shows the evaluation of $\partial \delta_{F_\infty(k)}(\eta)/\partial k$.

6.2.2 Evaluation of the Derivative of the Support Function

For any integer ℓ, let $\bar{\eta}_\ell := (\Phi^\ell)^T \eta$ and

$$\delta_W((\Phi^\ell)^T \eta) = \delta_W(\bar{\eta}_\ell) = \bar{\eta}_\ell^T \cdot w_\ell^* \tag{6.9}$$

where w_ℓ^* is the maximizer of $\delta_W(\bar{\eta}_\ell)$. Therefore

$$\frac{\partial \delta_W(\bar{\eta}_\ell)}{\partial k_j} = (w_\ell^*)^T \frac{\partial \bar{\eta}_\ell}{\partial k_j} + \bar{\eta}_\ell^T \frac{\partial w_\ell^*}{\partial k_j} \tag{6.10}$$

The term $\partial w_\ell^*/\partial k_j$ in (6.10) depends on the characterization of W. Since $w_\ell^* = \arg\max_{w\in W} \bar{\eta}_\ell^T w$, w_ℓ^* is a function of $\bar{\eta}_\ell$. For the special case when W is a polytope in $\mathbb{R}^n$, $\partial w_\ell^*/\partial \bar{\eta}_\ell$ is piecewise constant with respect to $\bar{\eta}_\ell$ where the various pieces are zero in value. In particular, $\partial w_\ell^*/\partial \bar{\eta}_\ell$ has a unique value when the set $S(\bar{\eta}_\ell) := \{w|\ \bar{\eta}_\ell^T w = \delta_W(\bar{\eta}_\ell)\}$ is a singleton. When this is the case $\partial w_\ell^*/\partial \bar{\eta}_\ell = 0$, see Figure 6.3a. For the value of $\bar{\eta}_\ell$

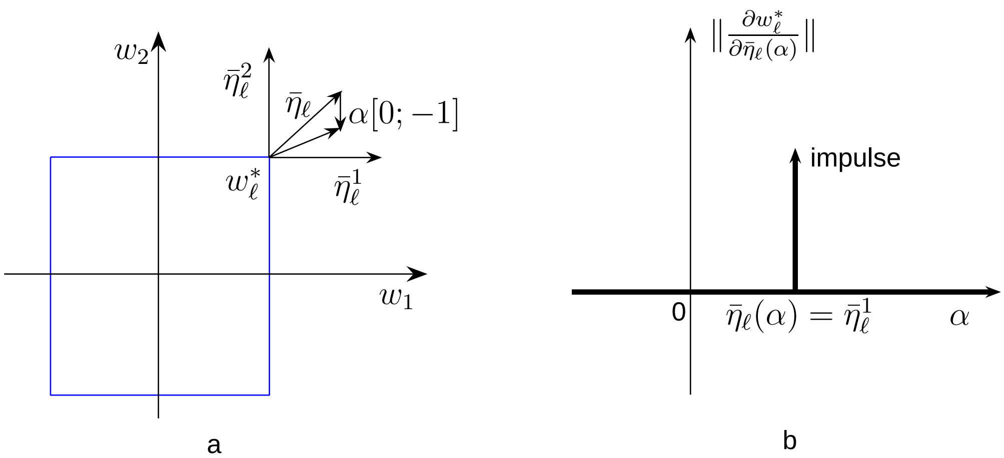

Figure 6.3: Derivative of support function

shown, w_ℓ^* is the unique minimizer of $\delta_W(\bar{\eta}_\ell)$. For small perturbation around $\bar{\eta}_\ell$, so long as $\bar{\eta}_\ell$ lies in the cone of $\bar{\eta}_\ell^1$ and $\bar{\eta}_\ell^2$, w_ℓ^* remains the same and $\partial w_\ell^*/\partial \bar{\eta}_\ell = 0$. Figure 6.3b shows value of $\|\partial w_\ell^*/\partial \bar{\eta}_\ell(\alpha)\|$ where $\bar{\eta}_\ell(\alpha) = \bar{\eta}_\ell + \alpha \begin{bmatrix} 0 \\ -1 \end{bmatrix}$. As shown, $\|\partial w_\ell^*/\partial \bar{\eta}_\ell(\alpha)\|$ is zero almost everywhere except when $\bar{\eta}_\ell(\alpha) = \bar{\eta}_\ell^1$. Hence, we assume that $\partial w_\ell^*/\partial \bar{\eta}_\ell = 0$ throughout. Since $\bar{\eta}_\ell$ is a function of k_j, $\partial w_\ell^*/\partial k_j = 0$.

Remark 6.2.1 *It is possible that multiple maximizers exist in (6.9). In this case, there are two possible scenarios: either the maximizer is about to change or remains as maximizer. In the former scenarios the derivative of $\delta_W(\bar{\eta}_\ell)$ does not exist actually and we can use any maximizer in (6.10) to compute an assumed derivative. The latter scenario is the case which we have discussed before and additionally $(\hat{w}_\ell^* - \check{w}_\ell^*)^T \partial \bar{\eta}_\ell / \partial k_j = 0$ for any two different maximizers $\hat{w}_\ell^*$ and $\check{w}_\ell^*$.*

Using (6.3c),

$$\frac{\partial \delta_{F_\infty}(\eta)}{\partial k_j} = \frac{\partial}{\partial k_j} \sum_{\ell=0}^{\infty} \delta_W((\Phi^\ell)^T \eta) = \sum_{\ell=0}^{\infty} \frac{\partial \delta_W(\bar{\eta}_\ell)}{\partial k_j} = \sum_{\ell=0}^{\infty} (w_\ell^*)^T \cdot \frac{\partial \bar{\eta}_\ell}{\partial k_j} \tag{6.11}$$

Let $\hat{B}_j$ is a $n \times n$ square matrix of all zeros except the j^{th} column being b, then

$$\begin{aligned} \frac{\partial \bar{\eta}_\ell}{\partial k_j} &= \frac{\partial((\Phi^\ell)^T \eta)}{\partial k_j} = \frac{\partial((\Phi^\ell)^T)}{\partial k_j} \cdot \eta && \text{(6.12a)} \\ &= (\hat{B}_j \Phi^{\ell-1} + \Phi \hat{B}_j \Phi^{\ell-2} + \Phi^2 \hat{B}_j \Phi^{\ell-3} + \cdots + \Phi^{\ell-1} \hat{B}_j)^T \cdot \eta && \text{(6.12b)} \\ &= (\sum_{i=0}^{\ell-1} \Phi^i \hat{B}_j \Phi^{\ell-1-i})^T \cdot \eta && \text{(6.12c)} \end{aligned}$$

where (6.12b) follows from $\Phi^\ell = (A+bk)^\ell$. Using (6.12c) in (6.11) yields

$$\begin{aligned}
&\frac{\partial \delta_{F_\infty}(\eta)}{\partial k_j} \\
= \;& \sum_{\ell=0}^{\infty} (w_\ell^*)^T \cdot \frac{\partial \bar{\eta}_\ell}{\partial k_j} && (6.13a)\\
= \;& \sum_{\ell=0}^{\infty} \eta^T \cdot \left[\sum_{i=0}^{\ell-1} \Phi^i \hat{B}_j \Phi^{\ell-1-i}\right] \cdot w_\ell^* && (6.13b)\\
= \;& \sum_{\ell=0}^{L} \eta^T \cdot \left[\sum_{i=0}^{\ell-1} \Phi^i \hat{B}_j \Phi^{\ell-1-i}\right] \cdot w_\ell^* + \sum_{\ell=L+1}^{\infty} \eta^T \cdot \left[\sum_{i=0}^{\ell-1} \Phi^i \hat{B}_j \Phi^{\ell-1-i}\right] \cdot w_\ell^* && (6.13c)
\end{aligned}$$

Given k and a direction η, the value of $\partial \delta_{F_\infty}(\eta)/\partial k_j$ can be determined as accurate as possible from (6.13c) by choosing a large enough L. This is true since the second term in (6.13c) can be bounded and decreases to zero as L increases. To show this, consider

$$\begin{aligned}
&\left|\eta^T \cdot \left[\sum_{i=0}^{\ell-1} \Phi^i \hat{B}_j \Phi^{\ell-1-i}\right] \cdot w_\ell^*\right| && (6.14a)\\
\leq \;& \|\eta\| \cdot \max_{w\in W} \|w\| \cdot \sum_{i=0}^{\ell-1} (\|\Phi^i\| \cdot \|\hat{B}_j\| \cdot \|\Phi^{\ell-1-i}\|) && (6.14b)\\
\leq \;& \|\eta\| \cdot \max_{w\in W} \|w\| \cdot \ell \cdot \|b\| \cdot M^2 \cdot (\rho(\Phi))^{\ell-1} && (6.14c)\\
= \;& v_d \cdot (\rho(\Phi))^{\ell-1} \cdot \ell && (6.14d)
\end{aligned}$$

where $v_d = \|\eta\| \cdot \max_{w\in W} \|w\| \cdot \|b\| \cdot M^2$ and $M(\rho(\Phi))^i \geq \|\Phi^i\|,\ \forall i \geq 0$, then the second term in (6.13c) satisfies

$$\left| \sum_{\ell=L+1}^{\infty} \eta^T \cdot \left[\sum_{i=0}^{\ell-1} \Phi^i \hat{B}_j \Phi^{\ell-1-i} \right] \cdot w_\ell^* \right| \quad \leq \quad v_d \cdot \sum_{\ell=L+1}^{\infty} (\rho(\Phi))^{\ell-1} \cdot \ell \tag{6.15}$$

Next we show that the right hand side of inequality (6.15) is bounded. To this end, let

$$\Delta = \sum_{\ell=L+1}^{\infty} (\rho(\Phi))^{\ell-1} \cdot \ell = \sum_{\ell=L+1}^{\infty} \rho^{\ell-1} \cdot \ell \tag{6.16}$$

where ρ denotes $\rho(\Phi)$ for notational simplicity, then Δ satisfies

$$\Delta \quad = \quad \rho^L(L+1) + \rho^{L+1}(L+2) + \rho^{L+2}(L+3) + \cdots \tag{6.17}$$

$$\rho\Delta \quad = \quad \rho^{L+1}(L+1) + \rho^{L+2}(L+2) + \rho^{L+3}(L+3) + \cdots \tag{6.18}$$

When (6.17) minus (6.18), we get

$$(1-\rho)\Delta = \rho^L(L+1) + (\rho^{L+1} + \rho^{L+2} + \rho^{L+3} + \cdots = \rho^L(L+1) + \frac{\rho^{L+1}}{1-\rho}. \tag{6.19}$$

Hence,

$$\Delta = \frac{\rho^L(L+1)}{1-\rho} + \frac{\rho^{L+1}}{(1-\rho)^2} = \frac{\rho^L}{1-\rho}(L + \frac{1}{1-\rho}) \tag{6.20}$$

and (6.15) becomes

$$\left|\sum_{\ell=L+1}^{\infty}\eta^T\cdot\left[\sum_{i=0}^{\ell-1}\Phi^i\hat{B}_j\Phi^{\ell-1-i}\right]\cdot w_\ell^*\right|\leq \xi_L := v_d\frac{\rho^L}{1-\rho}(L+\frac{1}{1-\rho}) \tag{6.21}$$

Therefore,

$$\frac{\partial\delta_{F_\infty}(\eta)}{\partial k_j}\in\left[\frac{\partial\delta_{F_L}(\eta)}{\partial k_j}-\xi_L,\ \frac{\partial\delta_{F_L}(\eta)}{\partial k_j}+\xi_L\right]. \tag{6.22}$$

Since Φ must be stable, $\rho(\Phi)<1$ and from (6.21)

$$\lim_{L\to\infty}\xi_L=0 \tag{6.23}$$

The above method of evaluating the derivative is illustrated in the following example.

Example 6.2.2 *Consider a system in the form of (6.1a)-(6.1b) with parameters*

$$A=\begin{bmatrix}1.2 & 2\\ 0 & 1.5\end{bmatrix},\ b=\begin{bmatrix}1\\ 0.5\end{bmatrix},\ k=[-0.6232\ \ -1.9678]$$

$$w(t)\in W=\{w|\ \|w\|_\infty\leq 0.2\},\ \ \Phi=A+bk=\begin{bmatrix}0.5768 & 0.0322\\ -0.3116 & 0.5161\end{bmatrix}.$$

η is chosen to be $[1\ \ 1]^T$ and $v_d=7.1554$ with $M=4$. The approximations of $\partial\delta_{F_\infty}(\eta)/\partial k_j$, $j=1,2$ and the corresponding errors with different L are listed in the following table and plotted in Figure 6.4. From Figure 6.4, we can see that the approximation of $d\delta_{F_\infty}(\eta)/dk$ converges to $[0.1776\ \ -1.3281]$ as L increases.

$\eta = [1\ \ 1]^T$	$L=11$	$L=12$	$L=13$	$L=14$	$L=15$	$L=16$	$L=17$
$\partial\delta_{F_L}(\eta)/\partial k_1$	0.1933	0.1876	0.1837	0.1812	0.1796	0.1787	0.1782
$\partial\delta_{F_L}(\eta)/\partial k_2$	-1.3361	-1.3358	-1.3344	-1.3327	-1.3313	-1.3301	-1.3293
ξ_L	0.3258	0.1943	0.1154	0.0682	0.0402	0.0236	0.0138

Table 6.2: Approximation of $\partial\delta_{F_\infty}(\eta)/\partial k_j$ with different L

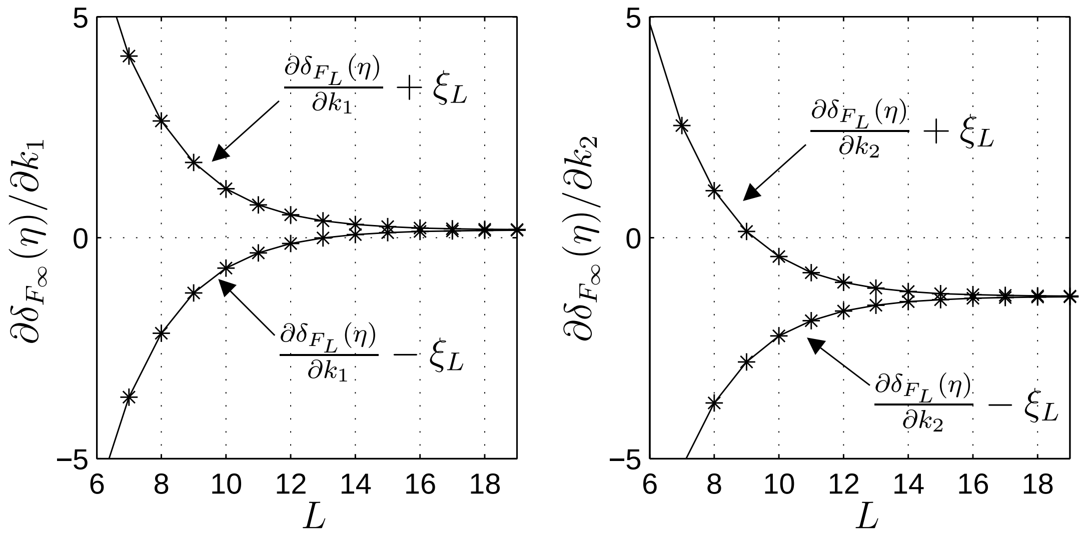

Figure 6.4: Approximation of $\partial\delta_{F_\infty}(\eta)/\partial k_j$ with different L

6.3 Design of Feedback Gain

Using the results introduced in Section 6.2, it is possible to design the feedback gain k by solving optimization problems with constraints imposed on the support function of $F_\infty(k)$ since both the support and its derivative can be evaluated.

Consider the simplest case where the only concern is the "shape" of the resulting $F_\infty(k)$. Specifically, we want $F_\infty(k)$ to have minimal supports along some given directions $\mu_s^i,\ i \in \mathscr{I}_s$. One possible way to achieve this is by minimizing α subject to the condition that $\delta_{F_\infty(k)}(\mu_s^i) \leq \alpha$ in the form of

$$\min_{k,\alpha} \quad \alpha \tag{6.24a}$$

$$\text{s.t.} \quad \delta_{F_\infty(k)}(\mu_s^i) \leq \alpha, \quad \forall i \in \mathscr{I}_s \tag{6.24b}$$

The numerical solution of (6.24) can be obtained using a standard non-linear optimization solver. For this to happen, values of $\delta_{F_\infty(k)}(\eta)$ and $\partial \delta_{F_\infty(k)}(\eta)/\partial k$ at different value of k are needed and they are given by (6.8) and (6.22).

Another potential application of the techniques developed in this chapter is to ensure the existence of the maximal disturbance invariant set O_∞ defined in Section 2.2.2 in the design of the feedback gain k. This is important in MPC since the O_∞ set is usually used in MPC as the terminal set.

Suppose system (6.1) is subject to the following joint state and control constraints

$$Y := \{(x,u) | \; (\mu_x^i)^T x + (\mu_u^i)^T u \leq 1, \;\; \forall i \in \mathscr{I}_y\}. \tag{6.25}$$

One necessary condition of the existence of O_∞ set is that

$$(x, kx) \in Y, \;\; \forall x \in F_\infty(k) \tag{6.26}$$

or equivalently using $u = kx$ and (6.25),

$$\delta_{F_\infty(k)}(\mu_x^i + k^T \mu_u^i) \leq 1, \ \ \forall i \in \mathscr{I}_y. \tag{6.27}$$

Hence, in designing k, it is necessary to take this condition into the formulation of optimization (6.24). More exactly,

$$\min_{k,\alpha} \quad \alpha \tag{6.28a}$$

$$\text{s.t.} \quad \delta_{F_\infty(k)}(\mu_s^i) \leq \alpha, \quad \forall i \in \mathscr{I}_s \tag{6.28b}$$

$$\delta_{F_\infty(k)}(\mu_x^i + k^T \mu_u^i) \leq 1, \ \ \forall i \in \mathscr{I}_y \tag{6.28c}$$

In the above, constraint (6.28b) can be handled by applying the results in Section 6.2 directly, but handling (6.28c) needs slight modification since η is parameterized by k. To show the modification, let $g_i^T := \mu_x^i + k^T \mu_u^i$, $\forall i \in \mathscr{I}_y$. Similar to (6.3d),

$$\delta_{F_\infty}(g_i^T) = \sum_{\ell=0}^{L} \delta_W((g_i \Phi^\ell)^T) + \sum_{\ell=L+1}^{\infty} \delta_W((g_i \Phi^\ell)^T) \tag{6.29}$$

and similar to (6.6), the second term on the right hand side of (6.29) satisfies

$$0 \leq \sum_{\ell=L+1}^{\infty} \delta_W((g_i \Phi^\ell)^T) \leq \frac{\bar{v}_s \cdot (\rho(\Phi))^{L+1}}{1-\rho(\Phi)} \tag{6.30}$$

where $\bar{v}_s := \|g_i\| \cdot \max_{w \in W} \|w\| \cdot M$ and $M(\rho(\Phi))^i \geq \|\Phi^i\|$, $\ \forall i \geq 0$. For the support function to be bounded, we should have $\rho(\Phi) < 1$. As a result, the second term on the right hand side of (6.29) tends to zero if L is large enough, so $\delta_{F_\infty}(g_i^T)$ can be evaluated as accurate as possible. Now consider $\partial\delta_W((g_i\Phi^\ell)^T)/\partial k_j$. Let

$$\delta_W((g_i\Phi^\ell)^T) = (g_i\Phi^\ell)\cdot w_\ell^*, \tag{6.31}$$

then using the same argument below (6.10), we have $\partial w_\ell^*/\partial k_j = 0$ and

$$\frac{\partial\delta_W((g_i\Phi^\ell)^T)}{\partial k_j} = \frac{\partial g_i}{\partial k_j}\Phi^\ell w_\ell^* + g_i\frac{\partial\Phi^\ell}{\partial k_j}w_\ell^* \tag{6.32}$$

Note that

$$\frac{\partial\delta_{F_\infty}(g_i^T)}{\partial k_j} = \sum_{\ell=0}^{\infty}\frac{\partial\delta_W((g_i\Phi^\ell)^T)}{\partial k_j}, \tag{6.33}$$

following a similar procedure from (6.11) to (6.21), it can be obtained that

$$\frac{\partial\delta_{F_\infty}(g_i^T)}{\partial k_j} \in \left[\frac{\partial\delta_{F_L}(g_i^T)}{\partial k_j} - \zeta_L,\ \ \frac{\partial\delta_{F_L}(g_i^T)}{\partial k_j} + \zeta_L\right] \tag{6.34}$$

where $\zeta_L := \hat{v}_d\frac{\rho^{L+1}}{1-\rho} + \check{v}_d\frac{\rho^L}{1-\rho}(L+\frac{1}{1-\rho})$, $\hat{v}_d := |\mu_u^i|\cdot\max_{w\in W}\|w\|\cdot M$ and $\check{v}_d := \max_{w\in W}\|w\|\cdot\|g_i\|\cdot\|b\|\cdot M^2$.

Remark 6.3.1 *For multiple input cases where $u(t)\in\mathbb{R}^m$ $m\geq 2$, the results introduced in Section 6.2 and 6.3 can be easily applied by considering $x(t+1) = Ax(t) + Bu(t) + w(t),\ u(t) = Kx(t)$, where $B = [b_1\cdots b_m]$ and $K = \left[k_1^T\cdots k_m^T\right]^T$.*

6.4 Numerical examples

In this section, the design approach described in Section 6.3 is demonstrated by examples. Consider the system (6.1a) with the following parameters,

$$A = \begin{bmatrix} 1.1 & 1 \\ 0 & 1.3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad W = \{w| \ \|w\|_\infty \leq 0.12\}$$

and the μ_s^i are

$$\mu_s^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mu_s^2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \mu_s^3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mu_s^4 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

The objective is to find the smallest bounding square box that contains F_∞ set. Solving optimization problem (6.24) yields

$$k_s = [-1.5111 \quad -0.8889], \ \alpha_s = 0.3511$$

The corresponding $F_\infty(k_s)$ set is plotted in Figure 6.5. For the purpose of comparison, $F_\infty(k_{LQR})$ is also plotted, where $k_{LQR} = [-0.4991 \quad -0.9546]$ is obtained using the LQR method with $Q = I_2$ and $R = 1$.

Next, consider the case where state and control constraints are presented. μ_s^i are chosen to be the same as in the previous example and let the constraint set Y be

$$Y = \{(x,u)| \ \mu_x x \leq 1, \ \mu_u^1 u \leq 1, \ \mu_u^2 u \leq 1\}$$

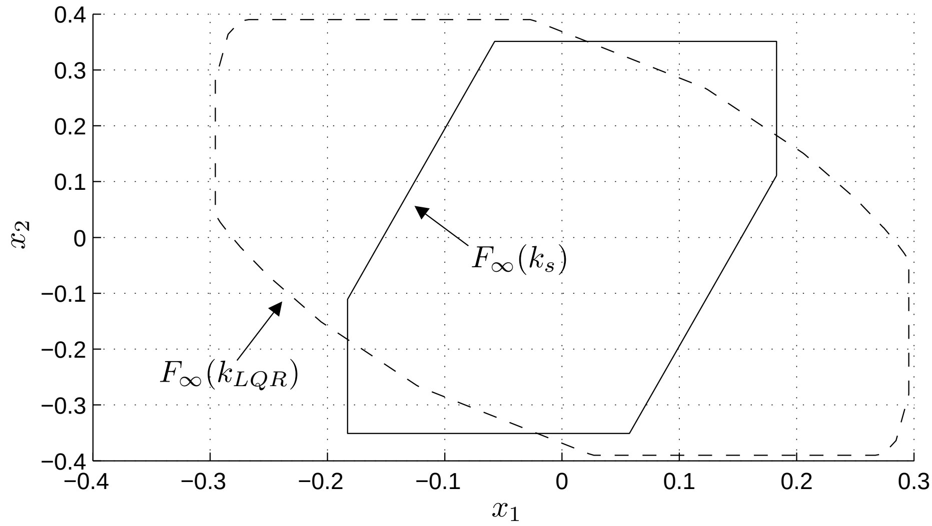

Figure 6.5: Comparison of F_∞ sets

where $\mu_x = [2.4\ \ 2.4]$, $\mu_u^1 = 1$, $\mu_u^2 = -1$. Solving optimization problem (6.28) yields

$$k_{sxu} = [-0.6133\ \ -1.1150],\ \alpha_{sxu} = 0.3838.$$

The corresponding $F_\infty(k_{sxu})$ is plotted in Figure 6.6. For the purpose of comparison, $F_\infty(k_s)$, which is the optimal F_∞ set without state and control constraints, is also plotted. It can be observed that k_s becomes infeasible since it violates the state constraint, while $F_\infty(k_{sxu})$ satisfies both state and control constraints. Not surprisingly, the optimum also becomes larger. Hence, under the controller $u(t) = k_{sxu}x(t)$, O_∞ set exists and it is plotted in Figure 6.7.

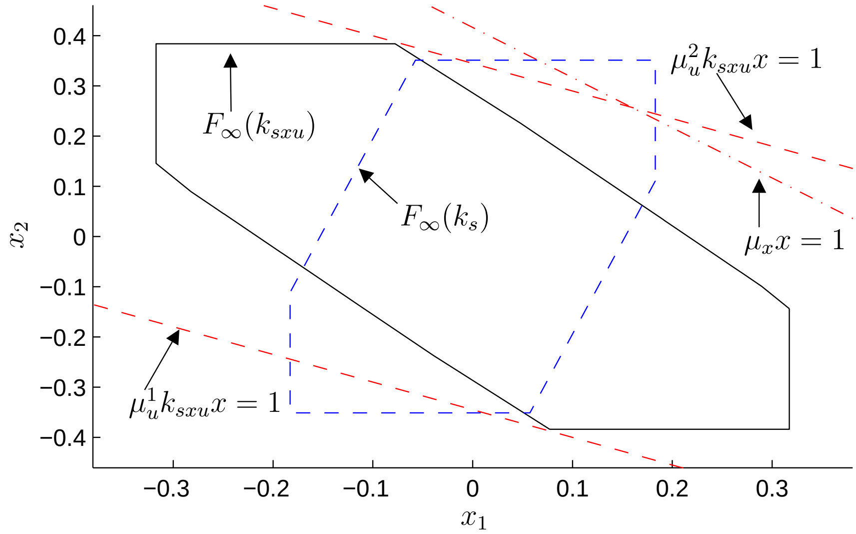

Figure 6.6: Optimal F_∞ sets with state and control constraints

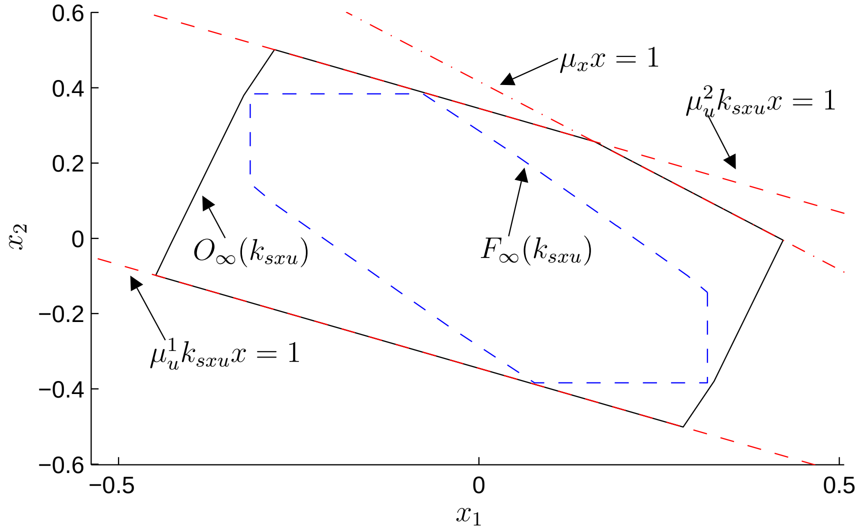

Figure 6.7: Optimal $F_\infty(k_{sxu})$ and $O_\infty(k_{sxu})$

6.5 Summary

Methods of approximating the support function of $F_\infty(k)$ and its derivative with respect to the feedback gain k are introduced in this chapter. It has been shown that both the support function and its derivative can be evaluated as accurate as possible for the case where W is a polytope. With these values, an optimization problem with constraints imposed on the support function becomes numerically solvable. Two optimization problems are set up for the design of feedback gains. One aims to find a feedback gain that yields F_∞ having a certain optimal shape; while the other guarantees the satisfaction of state and control constraints in addition. These design methods are illustrated by two numerical examples.

Chapter 7

Probabilistically Constraint-Admissible Set for Linear Systems with Disturbances and Its Application

The maximal constraint-admissible set of linear systems with hard constraints has been widely studied and applied in the design of controller for such systems. This chapter generalizes the concept of maximal constraint-admissible set to the case where probabilistic constraints, also known as chance constraints or soft constraints, are present. Properties of probabilistically constraint-admissible sets are studied in this chapter and it is shown that the maximal probabilistically constraint-admissible set is not invariant in general. An invariant inner approximation of the set is then proposed. This approximate set is used as the terminal set in the design of an MPC controller for a linear system with additive disturbances and probabilistic constraints. Feasibility and stability of the resulting closed-loop system are also discussed. The effectiveness of the proposed approach is illustrated via numerical examples.

7.1 Introduction

Constraint-admissible invariant (CAI) sets play an important role in the study of constrained systems [61, 5, 59]. These sets have been used in many approaches for the control of constrained systems. For example, CAI sets are used as terminal sets in MPC, they are also used to characterize the domain of attractions of nonlinear control laws [88, 89]. Other uses of CAI sets include [3, 90, 91, 92]. Many results have been obtained for the case of linear discrete time systems with polyhedral constraints (see [93] and [94] by Bitsoris and [95] by Vassilaki *et al*). A notable contribution is the characterization of the maximal invariant CAI sets by Gilbert and Tan [61]. Several nonlinear feedback controllers have been designed based on this characterization [88]. More recently, CAI sets for the case where a disturbance is present has also been studied [5] by Blanchini, [37] by Kolmanovsky and Gilbert, [96] by Kerrigan and Maciejowski and [97] by Raković *et al.*

Almost all studies on CAI sets have been for systems subject to hard constraints. Typically, these constraints are imposed on both the state and control of the system and they have to be satisfied at all time instances. In some applications, constraints need not be satisfied for all times but are probabilistic in nature. Problems of such nature are typically studied under the broader domain of stochastic programming [98]. Indeed, some problems are better modeled by chance, or probabilistic, constraints or a mixture of probabilistic and hard constraints. Examples of such systems include the water level control in a distillation column, [49, 51], risk management on sustainable development [50, 23], temperature control in buildings [99] and others [100].

This chapter generalizes the concept of constraint-admissible sets to the case where probabilistic constraints are present. Definition and properties of such a constraint-admissible set for a linear system with probabilistic constraints are discussed. An inner approximation of the maximal constraint-admissible set is proposed which is invariant. Its use as a terminal set under the MPC framework where some constraints are probabilistic in nature are included. The feasibility of the MPC finite horizon optimization problem and the closed-loop stability under this setting are also discussed.

The rest of this chapter is organized as follows. Properties of stochastic linear systems are studied in Section 7.2. Definition and properties of probabilistically CAI sets are discussed in Section 7.3. Section 7.4 proposes a method of determining an inner invariant approximation of the probabilistically CAI sets. The computation of this approximate set is given in Section 7.5. The application of this set in MPC is discussed in Section 7.6. Numerical examples and summary are the contents of the last two sections.

7.2 Probabilistic Constraint and Stochastic System

Consider the linear discrete-time system

$$x(t+1) = \Phi x(t) + w(t), \quad x(0) = x_0 \tag{7.1}$$

where $x(t) \in \mathbb{R}^n$ is the system state and $w(t) \in W \subset \mathbb{R}^n$ is a disturbance input. It is assumed that (7.1) satisfy the following assumption:

Assumption 7.2.1

(A1) *System (7.1) is stable, or equivalently,* Φ *has a spectral radius* $\rho(\Phi) < 1$*;*

(A2) $w(t) \in W$, $t \geq 0$ *are independent and identically distributed (i.i.d.) random vectors with a continuous probability density function* $f_w : W \to \mathbb{R}^+$ *and* $\int_W f_w(\omega) d\omega = 1$*. Additionally,* W *is compact and contains the origin in its interior.*

The $x(t)$ of system (7.1) is an affine function of disturbances $w(0), \cdots, w(t-1)$ given by

$$x(t) = \Phi^t x_0 + \sum_{j=0}^{t-1} \Phi^{t-j-1} w(j). \tag{7.2}$$

Since W is compact and $\rho(\Phi) < 1$, $x(t) \in F_t \oplus \{\Phi^t x_0\}$ where

$$F_t = W \oplus \Phi W \oplus \cdots \oplus \Phi^{t-1} W. \tag{7.3}$$

Also, $x(t)$ is a random variable following (7.2). Suppose the density function of $x(t)$ is $f_t(x;x_0)$. Then $f_t(x;x_0)$ can be obtained using $f_w(\cdot)$ under (A2) of Assumption 7.2.1.

Theorem 7.2.1 *Suppose system (7.1) satisfies Assumption 7.2.1. The density function $f_t(x;x_0)$ of state $x(t)$ is defined for all t, converges in the sense that $\lim_{t\to\infty}(f_{t+1}(x;x_0) - f_t(x;x_0)) = 0$ for any x and x_0.*

Proof: See Appendix 7.A.1. ∎

The following example shows the validity of Theorem 7.2.1.

Example 7.2.1 *Consider system (7.1) with $n = 1, \Phi = 0.5$ and assume that w is uniformly distributed on the set $W = \{w \mid |w| \leq 1\}$. The density function $f_t(x;0)$, for $t = 2,\ldots,7$ is shown in Figure 7.1, which shows the convergence of the density functions as t increases.*

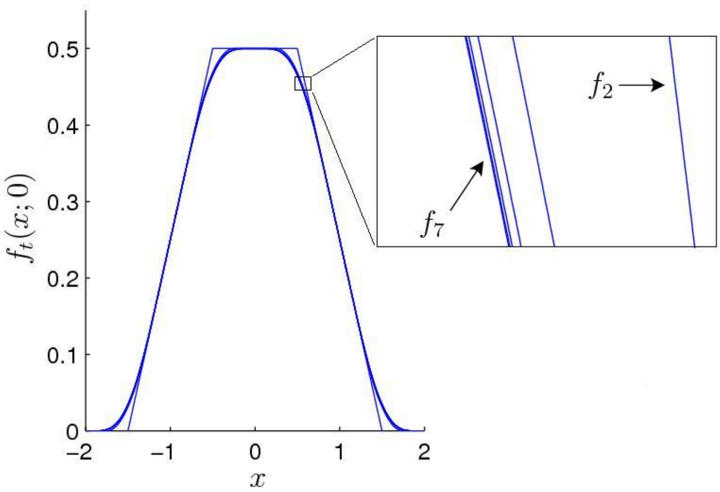

Figure 7.1: Probability density function of $x(2)$ to $x(7)$

Suppose a constraint set, $X_s \subset \mathbb{R}^n$, is given. The probability that $x(t)$ ($t > 0$) of (7.1) lies in X_s can be stated as $\Pr(x(t) \in X_s | x(0) = x_0)$ and evaluated from the fact that

$$\Pr(x(t) \in X_s | x_0) = \int_{X_s} f_t(x; x_0) dx, \tag{7.4}$$

if $f_t(x; x_0)$ is available. Hence, a constraint on $x(t)$ not to lie outside of X_s with probability more than ε, $0 < \varepsilon < 1$, can be imposed as

$$\Pr(x(t) \in X_s | x_0) \geq 1 - \varepsilon. \tag{7.5}$$

Consider all x_0 that satisfy (7.5) and collect them as

$$P_t^{\varepsilon}(X_s, \Phi, f_w) := \{x_0 | \Pr(x(t) \in X_s | x_0) \geq 1 - \varepsilon\} \tag{7.6a}$$

$$= \{x_0 | \int_{X_s} f_t(x; x_0) dx \geq 1 - \varepsilon\}. \tag{7.6b}$$

Clearly, $P_t^{\varepsilon}(X_s, \Phi, f_w)$ is a set of all points from which the probabilistic constraint is satisfied t steps from the current time. Indeed, probabilistic constraints like (7.5) are only meaningful for states in the future. They are not so when applied on past states as these are no longer stochastic in nature. In (7.6b), the dependence of parameters X_s, Φ and f_w are shown. These system parameters are assumed fixed and references to them will generally be omitted for notational simplicity in the sequel, unless warranted by context.

In general, the numerical characterization of P_t^ε is not easy as it involves a multidimensional integration and several convolution operations (7.6b). Procedures that avoid these expensive operations are discussed in the next two sections.

7.3 Maximal Probabilistically Constraint-Admissible Set and Its Properties

With (7.5), it is possible to characterize the set that satisfy the probabilistic constraint from $t = 1$ till $t = k$ as

$$O_k^\varepsilon(X_s, \Phi, f_w) := \bigcap_{t=1}^{k} P_t^\varepsilon = \{x|\Pr(x(t) \in X_s|x(0) = x) \geq 1-\varepsilon,\ t = 1, \cdots, k\} \quad (7.7)$$

and the maximal probabilistically constraint-admissible (PCA) set, $O_\infty^\varepsilon(X_s, \Phi, f_w)$, as

$$O_\infty^\varepsilon(X_s, \Phi, f_w) := \lim_{k\to\infty} O_k^\varepsilon(X_s, \Phi, f_w). \quad (7.8)$$

Since probabilistic constraints are imposed for future states, t starts from 1 instead of 0 in (7.7) and (7.8). As defined above, O_∞^ε can be seen as the generalization of the standard maximal disturbance invariant set [37] for system (7.1) with hard constraint $x(t) \in X_s$. Recall that if $x(t) \in X_s$ has to be satisfied at all times, the maximal disturbance invariant set of (7.1) is

$$O_\infty := \{x(0) |\, x(t) \in X_s,\ t = 0, 1, \cdots\} \tag{7.9}$$

Hence, besides handling probabilistic constraint, O_∞^ε differs from O_∞ in that the constraint at $t = 0$ is excluded from consideration. More exactly, O_∞ is equivalent to $X_s \cap O_\infty^\varepsilon$ with $\varepsilon = 0$ and $O_\infty \neq O_\infty^0$.

While being the most direct generalization of O_∞, O_∞^ε does not share many of its nice properties. These limitations are best illustrated using examples. Consider again the simple example of system (7.1) with $n = 1$, $\Phi = 0.5$, $f_w(\cdot)$ being a constant over $W = \{w|\ |w| \le 1\}$, $X_s = \{x|\ |x| \le 1\}$ and $\varepsilon = 0.5$. From (7.2) and Example 7.2.1, $x(t)$ has a symmetric probability density function with respect to its mean, $\Phi^t x(0)$, see Figure 7.2. Also, taking expectation over $w(i)$ $i \in \mathbb{Z}_{t-1}$ of $x(t)$ in (7.2), $\mathrm{E}[x(t)] = \Phi^t x(0)$ since

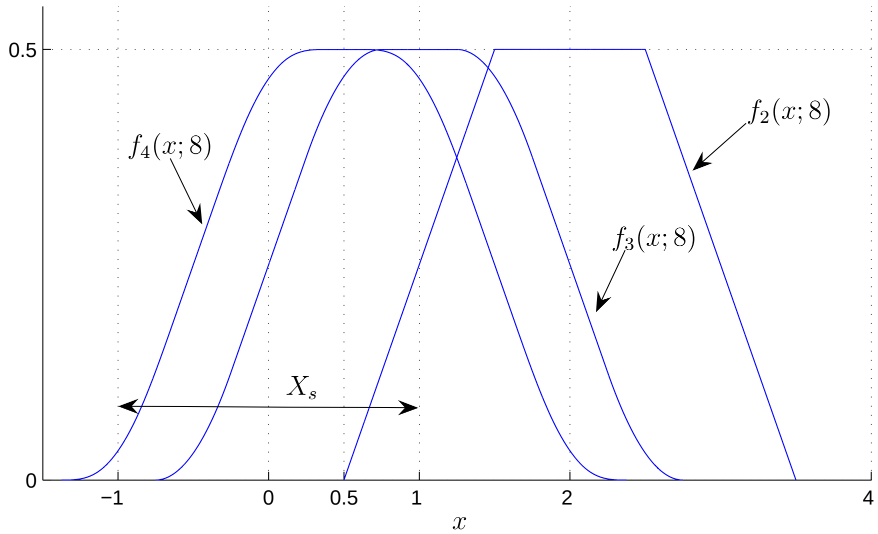

Figure 7.2: Density function of $x(2)$, $x(3)$ and $x(4)$ with $x(0) = 8$ including the location of the $\Phi^t x(0)$ in the figure

$\mathrm{E}[w(t)] = 0$. Consider a probabilistic constraint of the form $\Pr(x(t) \in X_s | x(0)) \geq 0.5$. This constraint can be restated for the example system as a deterministic constraint, $\Phi^t x(0) \in X_s$, following the symmetry of $f_t(x, x_0)$, $\mathrm{E}[x(t)] = \Phi^t x(0)$ and the range of X_s. Using (7.6a) and the above observations, $P_t^\varepsilon = \{x | \ \Phi^t x \in X_s\}$ for this example. Similarly, $O_\infty^\varepsilon = \bigcap_{t=1}^\infty P_t^\varepsilon = \{x | \ \Phi^t x \in X_s, \ t = 1, \cdots, \infty\} = \{x | \Phi^t |x| \leq 1, \ t \geq 1\} = \{x : \ |x| \leq 2\}$.

Example 7.3.1 (Non-invariance) *In general, O_∞^ε is not invariant, that is $x(t) \in O_\infty^\varepsilon \not\Rightarrow x(t+1) \in O_\infty^\varepsilon$. Consider the system with $\Phi = -0.5$, $w(t)$ uniformly distributed over $W = \{w | \ |w| \leq 1\}$, $X_s = \{x | \ -3 \leq x \leq 6\}$ and $\varepsilon = 0.5$. Using the analysis in the preceding paragraph, $P_t^\varepsilon = \{x | \Phi^t x \in X_s\}$, so $P_1^\varepsilon = \{x | -12 \leq x \leq 6\}$, $P_2^\varepsilon = \{x | -12 \leq x \leq 24\}$, $P_3^\varepsilon = \{x | -48 \leq x \leq 24\}$, $\cdots$. Hence, $O_\infty^\varepsilon = \bigcap_{t=1}^\infty P_t^\varepsilon = \{x | \ -12 \leq x \leq 6\}$. Let $x(0) = -12 \in O_\infty^\varepsilon$ and suppose $w(0) = 1$. Then, $x(1) = -0.5 \times (-12) + 1 = 7 \notin O_\infty^\varepsilon$ and it shows the non-invariance of O_∞^ε.*

Remark 7.3.1 *The non-invariance of O_∞^ε deserves further comments. The main reason of this non-invariance follows from the fact that* $\Pr(x(2) \in X_s | x(0)) \geq 1 - \varepsilon$ *does not imply* $\Pr(x(2) \in X_s | x(1) = \Phi x(0) + w(0)) \geq 1 - \varepsilon$*, for all $w(0) \in W$, since the value of* $\Pr(x(2) \in X_s | x(1) = \Phi x(0) + w(0))$ *depends on the realization of $w(0)$. Using the definition of (7.6a), this also means that $x(0) \in P_2^\varepsilon$ does not imply $x(1) \in P_1^\varepsilon$. Hence, O_∞^ε is not invariant.*

In general, O_∞^ε can be an empty set. But it is non-empty if X_s contains any robust invariant set of (7.1). For example, if $F_\infty \subseteq X_s$, then O_∞^ε is non-empty because $F_\infty := \lim_{t \to \infty} F_t$ where F_t is given by (7.3) must be part of O_∞^ε. Also, the above example has a convex O_∞^ε set. As shown in the following example, this is not the general case.

Example 7.3.2 (Non-convexity) *We show the non-convexity of O_∞^ε by showing the non-convexity of P_i^ε. Consider the system with $\Phi = 0.5$, $X_s = \{x|\ |x| \leq 0.5\}$, $\varepsilon = 0.5$, $W = \{w|\ |w| \leq 1\}$ and $f_w(\cdot)$ as shown in Figure 7.3. Let $x^a = 2$ and $x^b = -2$. It can be verified from Figure 7.3 that $x^a \in P_1^\varepsilon$, $x^b \in P_1^\varepsilon$ since $\Pr(x(1) \in X_s|x(0) = x^a) = \Pr(-1.5 \leq w \leq -0.5) = 0.5 \geq 1-\varepsilon$ and $\Pr(x(1) \in X_s|x(0) = x^b) = \Pr(0.5 \leq w \leq 1.5) = 0.5 \geq 1-\varepsilon$. However, $\Pr(x(1) \in X_s|x(0) = 0.5x^a + 0.5x^b) = \Pr(-0.5 \leq w \leq 0.5) = 0 < 1-\varepsilon$ which implies that $0.5x^a + 0.5x^b \notin P_1^\varepsilon$.*

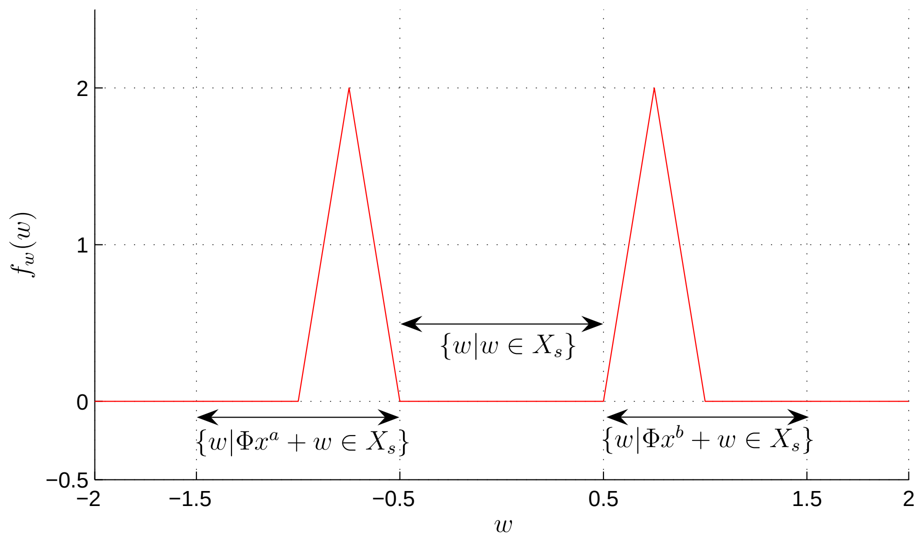

Figure 7.3: Probability density function of w

7.4 An Inner Approximation of O_∞^ε

As shown in the previous section, O_∞^ε is not invariant, convex or easily computed in general. This lack of nice properties prevents it from being useful in applications. This section reviews the general treatment of probabilistic constraint and, exploiting the inherent freedom, proposes an inner approximation of O_∞^ε, $\hat{O}_\infty^\varepsilon$, which is convex and invariant with respect to (7.1).

Consider the probabilistic constraint

$$\Pr(v \in \Omega) \geq 1-\varepsilon \tag{7.10}$$

where v is a vector of random variables and Ω is some appropriate set. Define

$$S_v(\varepsilon) := \{\Omega| \Pr(v \in \Omega) \geq 1-\varepsilon\}. \tag{7.11}$$

Clearly, $S_v(\varepsilon)$ is the collection of sets that have a probability measure greater than $1-\varepsilon$ and $\Pr(v \in \bar{\Omega}) \geq 1-\varepsilon$ for any $\bar{\Omega} \in S_v(\varepsilon)$. Obvious properties of $S_v(\varepsilon)$ following its definition are (i) $S_v(\varepsilon_1) \subseteq S_v(\varepsilon_2)$ if $\varepsilon_1 \geq \varepsilon_2$; (ii) Suppose $\Omega_1 \in S_v(\varepsilon)$ and $\Omega_1 \subseteq \Omega_2$, then $\Omega_2 \in S_v(\varepsilon)$. In general, $S_v(\varepsilon)$ can have many or infinite elements and this freedom can be exploited in the approximation of O_∞^ε.

Consider (7.6a) and restate it in terms of $w(j), j = 0, \cdots, t-1$ using (7.2). Specifically, (7.6a) becomes

$$\begin{aligned} 1-\varepsilon \leq \Pr(x(t) \in X_s|\, x(0) = x_0) &= \Pr(\Phi^t x_0 + \sum_{j=0}^{t-1} \Phi^{t-j-1} w(j) \in X_s|x_0) && (7.12a)\\ &= \Pr(\hat{\Phi}\mathbf{w}_t \in X_s \ominus \{\Phi^t x_0\}|x_0) && (7.12b)\\ &= \Pr(\mathbf{w}_t \in \Omega_t(x_0)) && (7.12c) \end{aligned}$$

where $\mathbf{w}_t := [(w(0))^T \cdots (w(t-1))^T]^T \in W^t$, $\hat{\Phi} = [\Phi^{t-1} \; \Phi^t \; \cdots \; \Phi^0]$ and $\Omega_t(x_0) := \{\mathbf{w}_t | \hat{\Phi}\mathbf{w}_t \in X_s \ominus \{\Phi^t x_0\}\}$. Following (7.11) and its ensuing discussion, $\Omega_t(x_0)$ is only one choice in $S_{\mathbf{w}_t}(\varepsilon)$. Other choices exist. In particular, let

$$V_{t,\varepsilon} := \underbrace{W \times \cdots \times W}_{t-1} \times W_\varepsilon \tag{7.13}$$

where W_ε is any set such that $W_\varepsilon \subset W$ with the property that $\int_{W_\varepsilon} f_w(w)dw \geq 1-\varepsilon$. It is easy to see that $V_{t,\varepsilon} \in S_{\mathbf{w}_t}(\varepsilon)$ under the i.i.d. assumption in (A2) of Assumption 7.2.1. Using $V_{t,\varepsilon}$, an inner approximation of P_t^ε is chosen as

$$\begin{aligned} \hat{P}_t^\varepsilon &:= \{x | \, V_{t,\varepsilon} \subseteq \Omega_t(x)\} \\ &= \{x | \, x(0) = x, \; x(t) \in X_s, \; \forall w(t-1) \in W_\varepsilon, \; \forall w(j) \in W, \; j \in \mathbb{Z}_{t-2}\} \end{aligned} \tag{7.14}$$

Accordingly, O_∞^ε can be approximated by

$$\begin{aligned} \hat{O}_\infty^\varepsilon := \cap_{t=1}^\infty \hat{P}_t^\varepsilon = \{x | \, & x(0) = x, \; x(t) \in X_s, \\ & \forall w(t-1) \in W_\varepsilon, \; \forall w(j) \in W \;\; j \in \mathbb{Z}_{t-2}, \; \forall t \geq 1\} \end{aligned} \tag{7.15}$$

As $\hat{P}_t^\varepsilon$ is an inner approximation of P_t^ε, $\hat{O}_\infty^\varepsilon$ is an inner approximation of O_∞^ε. Properties of $\hat{O}_\infty^\varepsilon$ are summarized in the following theorem.

Theorem 7.4.1 *If non-empty, $\hat{O}_\infty^\varepsilon$ of (7.15) has the following properties: (i) it is convex* *if X_s is convex; (ii) it is robustly invariant set with respect to system (7.1) in the sense that $x(t) \in \hat{O}^{\varepsilon}_{\infty}$ implies $x(t+1) \in \hat{O}^{\varepsilon}_{\infty}$.*

Proof: See Appendix 7.A.2. ■

7.5 Numerical Computation of $\hat{O}^{\varepsilon}_{\infty}$

Using (7.2) and (7.14), it is easy to see that $\hat{P}^{\varepsilon}_1 = \{x | \Phi x + w \in X_s, \forall w \in W_\varepsilon\} = \{x | \; \Phi x \in X_s \ominus W_\varepsilon\}$, $\hat{P}^{\varepsilon}_2 = \{x | \; \Phi^2 x \in X_s \ominus W_\varepsilon \ominus \Phi W\}$ and, in general, $\hat{P}^{\varepsilon}_i = \{x | \; \Phi^i x \in X_s \ominus W_\varepsilon \ominus \cdots \ominus \Phi^{i-1} W\}$. These expressions form the basis for the numerical computation of $\hat{O}^{\varepsilon}_{\infty}$. For this purpose, assume

Assumption 7.5.1

(A3) *X_s and W are polytopes and contains the origin in their respective interior, and one choice of $W_\varepsilon \in S_w(\varepsilon)$ with W_ε being a polytope containing the origin is found.*

Assumption (A3) is more stringent than required for reason of simplicity in presentation. For example, it is possible to assume that X_s is a polyhedron but some other assumptions are needed [101] to ensure the finite termination of the algorithm for the computation of $\hat{O}^{\varepsilon}_{\infty}$. This algorithm is given below.

Algorithm 7.5.1 (Determination of $\hat{O}^{\varepsilon}_{\infty}$)

(1) *Let* $Y_1 = X_s \ominus W_\varepsilon$*; if* $Y_1 = \emptyset$*, then* $\hat{O}_\infty^\varepsilon = \emptyset$ *and stop; otherwise let* $\hat{O}_1^\varepsilon = \{x|\ \Phi x \in Y_1\}$ *and* $i = 1$.

(2) *Compute* $Y_{i+1} = Y_i \ominus \Phi^i W$*. If* $Y_{i+1} = \emptyset$*, then* $\hat{O}_\infty^\varepsilon = \emptyset$ *and stop; otherwise continue.*

(3) *Compute* $\hat{O}_{i+1}^\varepsilon = \hat{O}_i^\varepsilon \cap \{x|\ \Phi^{i+1} x \in Y_{i+1}\}$*. If* $\hat{O}_{i+1}^\varepsilon = \hat{O}_i^\varepsilon$*, then* $\hat{O}_\infty^\varepsilon = \hat{O}_i^\varepsilon$ *and stop; otherwise let* $i = i+1$ *and go to step (2).*

$\hat{O}_\infty^\varepsilon$ is said to be finitely determined if there exist a finite i such that $\hat{O}_{i+1}^\varepsilon = \hat{O}_i^\varepsilon$. The following theorem gives a sufficient condition that guarantees finite determination of $\hat{O}_\infty^\varepsilon$.

Theorem 7.5.1 *Suppose Assumption 7.2.1 and 7.5.1 are satisfied and* Φ *has full rank. If* $Y_\infty = \lim_{t\to\infty} Y_t$ *is non-empty, then* $\hat{O}_\infty^\varepsilon$ *is finitely determined.*

Proof: See Appendix 7.A.3. ■

The following example verifies Theorems 7.4.1 and 7.5.1.

Example 7.5.1 *Consider the example given in Example 7.3.1 and choose* $W_\varepsilon = \{w|\ |w| \le 0.5\}$*. Following the procedure of Algorithm 7.5.1,* $\hat{O}_1^\varepsilon = \{x|\ -11 \le x \le 5\}$ *and* $\hat{O}_2^\varepsilon = \hat{O}_3^\varepsilon = \{x|\ -8 \le x \le 5\}$*. Therefore,* $\hat{O}_\infty^\varepsilon = \{x|\ -8 \le x \le 5\}$*. That this is an invariant set with respect to the system can also be easily verified.*

Remark 7.5.1 *From Algorithm 7.5.1, it follows that*

$$Y_\infty = X_s \ominus W_\varepsilon \ominus \Phi W \ominus \Phi^2 W \ominus \Phi^3 W \cdots = X_s \ominus W_\varepsilon \ominus \Phi F_\infty = X_s \ominus (W_\varepsilon \oplus \Phi F_\infty) \supset X_s \ominus F_\infty \quad (7.16)$$

where F_∞ *is* $\lim_{t\to\infty} F_t$ *and* F_t *is that given by (7.3). The last superset inclusion follows from the fact that* $W_\varepsilon \subset W$, $W_\varepsilon \oplus \Phi F_\infty \subset W \oplus \Phi F_\infty = F_\infty$. *Using this observation, another sufficient condition for finite determination is that* $F_\infty \subset \mathrm{int}(X_s)$. *This follows because if* $F_\infty \subset \mathrm{int}(X_s)$, $0 \in X_s \ominus F_\infty$ *which from (7.16) implies* $0 \in \mathrm{int}(Y_\infty)$. *This observation can be useful at times since it does not require the characterization of* Y_∞ *and accurate bounds on* F_∞ *can be computed [80].*

Remark 7.5.2 *If hard constraints given by* $x(t) \in X_h$, $\forall\, t \geq 0$ *is present in system (7.1) in addition to the probabilistic constraints, Algorithm 7.5.1 can be modified slightly to handle such a case. Replace step (1) of Algorithm 7.5.1 with*

(1) Let $\hat{O}^{\varepsilon}_0 = X_h$ *and* $Y_1 = (X_h \ominus W) \cap (X_s \ominus W_\varepsilon)$. *If* $Y_1 = \emptyset$, *then* $\hat{O}^{\varepsilon}_\infty = \emptyset$ *and stop; otherwise let* $\hat{O}^{\varepsilon}_1 = \hat{O}^{\varepsilon}_0 \cap \{x \mid \Phi x \in Y_1\}$ *and set* $i = 1$.

The rest of the algorithm remains unchanged. The resultant algorithm determines a $\hat{O}^{\varepsilon}_\infty$ *set incorporating both hard and probabilistic constraints.*

7.6 The MPC Formulation with Probabilistic Constraint

An obvious application of the $\hat{O}_\infty^\varepsilon$ set is its use in the Model Predictive Control framework where probabilistic and hard constraints are present. Consider

$$x(t+1) = Ax(t) + Bu(t) + w(t) \tag{7.17}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $w_t \in \mathbb{R}^n$ are the state, control and disturbance of the system at time t. Suppose the system is subject to control and state constraints in the form of

$$(x(t), u(t)) \in X_h,\ t \geq 0 \tag{7.18}$$

$$\Pr\{x(t) \in X_s\} \geq 1 - \varepsilon,\ t \geq 1 \tag{7.19}$$

where X_h and X_s are appropriate sets for the hard and soft constraints respectively. This model includes the typical situation of the hard constraints for control and soft for state. The objective is to find a state feedback MPC control law that will steer the state to the neighborhood of origin while satisfying all the constraints.

In addition, system (7.17) is assumed to satisfy the following assumptions:

Assumption 7.6.1

(B1) *the system (A,B) is stabilizable;*

(B2) *$w(t)$, $t \geq 0$ are i.i.d. with zero mean, W is a polytope and a $W_\varepsilon \in S_w(\varepsilon)$ satisfying (A3) of Assumption 7.5.1 is known;*

(B3) *Constraint sets X_h and X_s are compact polytopes and contain the origin in their respective interiors.*

These assumptions are quite standard for MPC and are needed for computational requirement of $\hat{O}_\infty^\varepsilon$. Clearly, (B1) implies the existence of Φ that satisfies (A1). Similarly, (B2) implies (A2) and (A3). Suppose the finite horizon (FH) problem has a horizon length N. Let $x(i|t)$, $u(i|t)$, $w(i|t)$ be the i^{th} predicted state, control and disturbance respectively within the horizon at time t. The parametrization of $u(i|t)$ is chosen to be the one proposed in Chapter 4 and it takes the form

$$u(i|t) = K_f x(i|t) + d(i|t) + \sum_{j=1}^{i} D(i,j|t)w(i-j|t),\ i \in \mathbb{Z}_{N-1} \tag{7.20}$$

where $d(i|t) \in \mathbb{R}^m$, $D(i,j|t) \in \mathbb{R}^{m\times n}$ are design variables and $K_f \in \mathbb{R}^{m\times n}$ is chosen so that $\Phi := A + BK_f$ satisfies (A1) of Assumption 7.2.1.

For notation simplicity, collect all the design variables within the control horizon in the following form,

$$\mathbf{d}(t) = \{d(i|t)\}_{i=0}^{N-1},\ \mathbf{D}(t) = \{\{D(i,j|t)\}_{j=1}^{i}\}_{i=1}^{N-1}. \tag{7.21}$$

The cost function that is to be optimized in the FH optimization problem is also the same as that in Chapter 4 and is

$$J_{DFD}(\mathbf{d}(t),\mathbf{D}(t)) = \sum_{i=0}^{N-1}\left[\|d(i|t)\|_{\Psi}^2 + \sum_{j=1}^{i}\|\mathrm{vec}(D(i,j|t))\|_{\Lambda}^2\right] \tag{7.22}$$

where Ψ and Λ are arbitrary matrices as long as they satisfy

$$\Psi \succ 0,\ \Lambda \succeq \Sigma_w \otimes \Psi \tag{7.23}$$

where Σ_w is the covariance matrix of $w(t)$. For more information about choosing Ψ and Λ, see Chapter 4.

With control parametrization (7.20) and cost function (7.22), the FH optimization problem, referred to hereafter as $\mathscr{P}_N^{PCA}(x(t))$, is

$$\min_{(\mathbf{d}(t),\mathbf{D}(t))} J_{DFD}(\mathbf{d}(t),\mathbf{D}(t))$$

$$\text{s.t. } x(0|t) = x(t),\ x(i+1|t) = Ax(i|t) + Bu(i|t) + w(i|t) \tag{7.24a}$$

$$u(i|t) = K_f x(i|t) + d(i|t) + \sum_{j=1}^{i} D(i,j|t)w(i-j|t) \tag{7.24b}$$

$$(x(i|t),u(i|t)) \in X_h,\ \ \forall\, w(i|t) \in W,\ \forall i \in \mathbb{Z}_{N-1} \tag{7.24c}$$

$$x(i|t) \in X_s,\ \forall\, w(i-1|t) \in W_\varepsilon,\ \forall w(j|t) \in W, j \in \mathbb{Z}_{i-2},\ \forall i \in \mathbb{Z}_N^+ \tag{7.24d}$$

$$x(N|t) \in \hat{O}_\infty^\varepsilon,\ \forall\, w(j|t) \in W, j \in \mathbb{Z}_{N-1} \tag{7.24e}$$

Constraints (7.24a)-(7.24c) are standard in the FH optimization where all constraints are hard. Constraint (7.24d) guarantees $\Pr(x(i|t) \in X_s|x_0) \geq 1-\varepsilon$ for all $i \in \mathbb{Z}_N^+$ following the discussion in Section 7.4. The last constraint (7.24e) ensures that both the soft and hard constraints are satisfied at all times beyond the horizon. This is true because $x(N|t) \in \hat{O}^{\varepsilon}_{\infty}$ means that $\Pr(x(N+i|t) \in X_s | x(N|t)) \geq 1-\varepsilon$ for all $i \geq 1$ following the characterization of $\hat{O}^{\varepsilon}_{\infty}$ in (7.15) and $x(N+i|t) \in X_h$ of Remark 7.5.2.

The computation of $\mathscr{P}_N^{PCA}(x(t))$ has already been discussed in Section 3.3 and Section 4.3 and hence is omitted here. The rest of the MPC formulation is standard: $\mathscr{P}_N^{PCA}(x(t))$ is solved at each time t and the very first term of

$$(\mathbf{d}^*(t), \mathbf{C}^*(t)) = \arg\min \ \mathscr{P}_N^{PCA}(x(t))$$

is applied to system (7.17). Hence, the MPC control law is

$$u(t) = \kappa^{PCA}(x(t)) := K_f x(t) + d(t) := K_f x(t) + d^*(0|t) \tag{7.25}$$

Issues of existence of feasible solution of $\mathscr{P}_N^{PCA}(x(t))$ at various t and the stability result of overall system under control law (7.25) are summarized in the following Theorem.

Theorem 7.6.1 *Suppose $\mathscr{P}_N^{PCA}(x(0))$ is feasible and Assumption 7.6.1 is satisfied. $\mathscr{P}_N^{PCA}(x(t))$ and system (7.17) under MPC control law (7.25) has the following properties: (i) $\mathscr{P}_N^{PCA}(x(t))$ admits an optimal solution for all t, (ii) $(x(t), u(t)) \in X_h$ for all $t \geq 0$ and $\Pr\{x(t+i) \in X_s | x(t)\} \geq 1-\varepsilon$ for all $i \geq 1$ and $t \geq 0$, (iii) $x(t)$ tends to F_{∞} set with probability one as t tends to infinity.*

Proof: See Appendix 7.A.4. ■

7.7 Numerical Examples

An example is used for the numerical simulation in this section. Its system parameters are:

$$A = \begin{bmatrix} 1.5 & 0.6 \\ 0 & 1.2 \end{bmatrix}, \ B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ K_f = \begin{bmatrix} -1.0912 & -0.6113 \end{bmatrix}, W = \{w | \ \|w\|_\infty \leq 0.1\}.$$

Here, K_f is the optimal LQ feedback gain with $Q = I_2$ and $R = 1$ and $w(t)$ is uniformly distributed on W.

The first simulation aims to understand the implications of hard and soft constraints on constraint admissible sets. In this regard, (7.1) is obtained with $\Phi = A + BK_f$ and O_∞ and $\hat{O}^\varepsilon_\infty$ are computed using Algorithm 7.5.1 and Remark 7.5.2. To be consistent with typical settings of a physical system, the control constraints are modeled as hard while the states are soft. Correspondingly, these sets are

$$U = \{u | \ |u| \leq 1\}, \ X_s = \{x | \ |x_1| \leq 1.95, -1.95 \leq x_2 \leq 1.05\}, \ \varepsilon = 0.3$$

and $W_\varepsilon \subset W$ is chosen to be

$$W_\varepsilon = \{w | \ |w_1| \leq 0.1, \ -0.1 \leq w_2 \leq 0.04\}.$$

Figure 7.4 shows three constraint-admissible sets: O_∞, $\hat{O}^0_\infty$ and $\hat{O}^\varepsilon_\infty$. Here, O_∞ is computed by treating X_s as hard constraint to be satisfied at all times. Correspondingly, constraints $-1.95 \leq x_2 \leq 1.05$ of X_s appear as binding constraints of O_∞. Also, O_∞ and $\hat{O}^0_\infty$ are not the same as the latter considers soft constraints imposed from $t \geq 1$ onwards. Figure 7.4 also include an outer bound of F_∞, the set to which the closed-loop system state converges to under the MPC control law as shown in (iii) of Theorem 7.6.1.

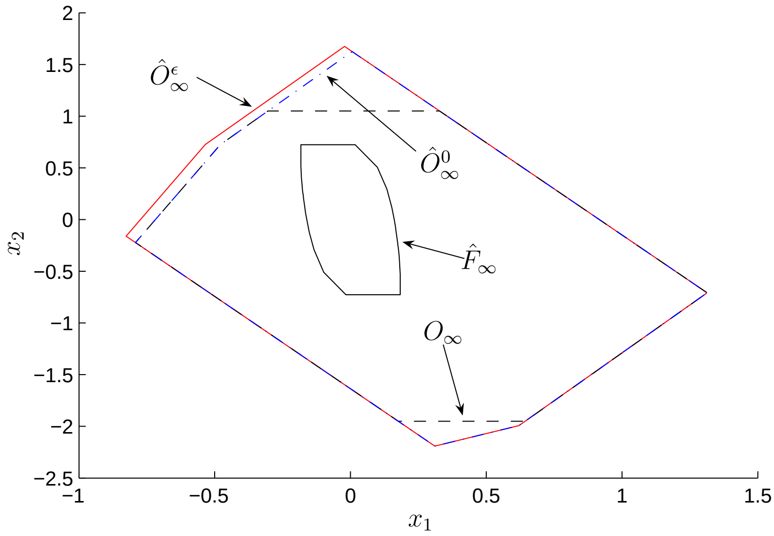

Figure 7.4: $\hat{O}^\varepsilon_\infty$, $\hat{O}^0_\infty$ and $\hat{O}_\infty$ set of the example system

Figure 7.5 shows the admissible set X^ε_N for the example with $N = 2$. As a comparison, it also shows X_N, the admissible set for the FH optimization problem where W_ε and $\hat{O}^\varepsilon_\infty$ of (7.24d) and (7.24e) are replaced by W and O_∞ respectively. Hence, X_N is the admissible set where X_s is treated as hard constraint.

The next simulation shows the states of the system under MPC control law for the case where $x(0) = x^a = [-0.4 \ 1.05]^T$, $x(0) = x^b = [1.34 \ 0.3]^T$, $N = 2$ and

$$\Psi = 7.8842, \ \Lambda = \begin{bmatrix} 0.0263 & 0 \\ 0 & 0.0263 \end{bmatrix},$$

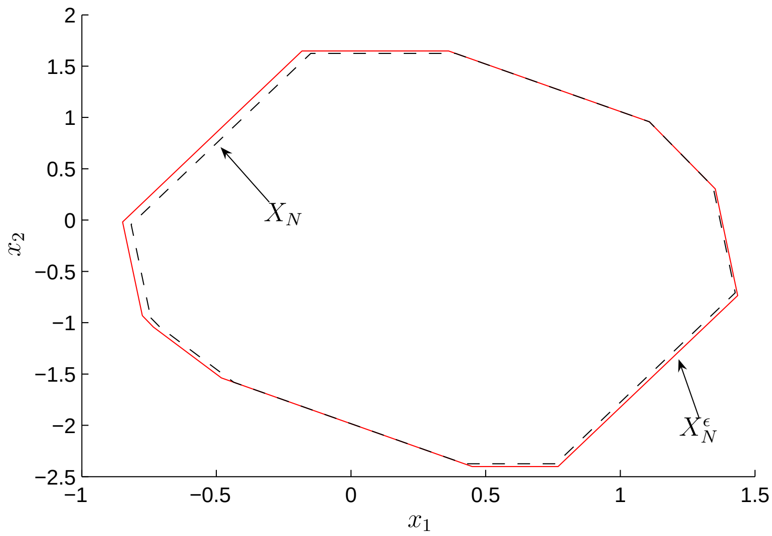

Figure 7.5: Comparison of X_N^ε and X_N sets

satisfying (7.23). The system is simulated over 200 disturbance sequence realizations and eight of the simulation results are shown in Figure 7.6 to avoid clutter. As shown, the hard constraints (those associated with $|u(t)| < 1$) are satisfied and $x(t)$ converges to the $\hat{F}_\infty$ set.

As shown in Figure 7.6, $x \in X_s$ does not hold for all time. However, the percentage of time that $x(t) \notin X_s$ for $t = 1, \cdots, 4$ over the 200 runs are shown in Table 7.1. Clearly, the constraint violations never exceed ε, verifying result (ii) of Theorem 7.6.1. The amount of violation also decreases with increasing t, a behavior that is expected following (iii) of Theorem 7.6.1 as F_∞ lies in the interior of X_s.

Table 7.1: Statics Results

State		$x(1)$	$x(2)$	$x(3)$	$x(4)$
Percentage of $x(t) \notin X_s$	$x(0) = x^a$	13.5%	1.5%	0%	0%
	$x(0) = x^b$	0%	0%	0%	0%

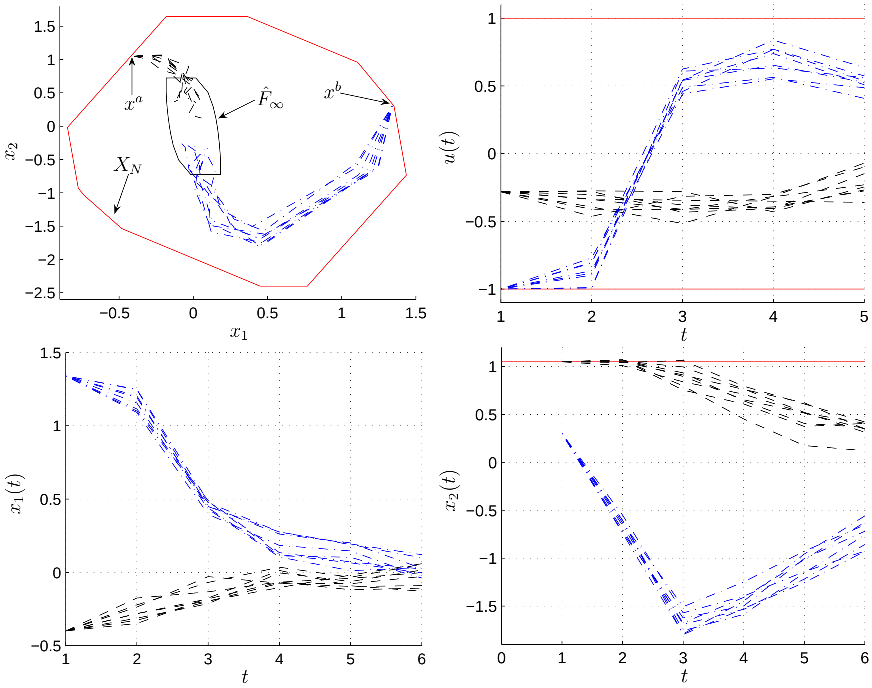

Figure 7.6: State and control trajectories

7.8 Summary

This chapter proposes an approach for characterizing a constraint admissible invariant set for a linear discrete system subjected to hard and probabilistic constraints. When the constraint and disturbance sets are described by linear inequalities, so is the proposed characterization. Properties and computations of this set are discussed. Using this as the terminal set, a state feedback control law is designed under the Model Predictive Control framework for a system that has both soft and hard constraints. The availability of the constraint admissible invariant set and the treatment of probabilistic constraints allow for greater design freedom in dealing with constraints of different importance.

7.A Appendix

7.A.1 Proof of Theorem 7.2.1

Proof: To derive the density function of $x(t)$, consider the case where $x(0) = 0$ first. Then the predicted state of (7.1) is

$$x(t) = w(t-1) + \Phi w(t-2) + \cdots + \Phi^{t-1} w(0). \tag{7.26}$$

and denote the density function of $x(t)$ in (7.26) by $f_t^0(\cdot)$ where the superscript 0 means $x(0) = 0$. Clearly when $t = 1$, the density function of $x(1)$ is the same as the density of $w(0)$ since $x(1) = w(0)$, therefore, $f_1^0(\cdot) = f_w(\cdot)$. Define

$$y_t := w(t) + \Phi w(t-1) + \cdots + \Phi^{t-1} w(1),$$

then due to the fact that $w(i)$, $i \geq 0$ are i.i.d. y_t has the same density function as $x(t)$ in (7.26). Note that random vector $x(t+1)$ can be equivalent expressed as

$$x(t+1) = y_t + \Phi^t w(0), \tag{7.27}$$

Therefore, the density function of $x(t+1)$ can be determined using those of y_t and $w(0)$. By applying the result in Section 8.16 of [102], the density function of $x(t+1)$ in (7.27) is

$$f_{t+1}^0(x) = \int f_{(y_t,w(0))}(x-\Phi^t w, w)dw = \int f_t^0(x-\Phi^t w) f_w(w)dw \tag{7.28}$$

where $f_{(y_t,w(0))}(\cdot)$ is the joint density function of y_t and $w(0)$, and the second equation is due to the independence of y_t and $w(0)$. Equation (7.28), together with $f_1^0 = f_w$, yields the expression of the density function of $x(t)$ in (7.26) as

$$\begin{cases} f_1^0(x) = f_w(x) & t=1 \\ f_t^0(x) = \int f_{t-1}^0(x-\Phi^{t-1}w) f_w(w)dw & t \geq 2 \end{cases} \tag{7.29}$$

Next, we show that f_t^0 is continuous for all t. To this end, consider the difference between $f_2^0(x)$ and $f_2^0(x+\xi)$ as $\xi \to 0$, we have

$$\begin{aligned} & \lim_{\xi\to 0}(f_2^0(x+\xi) - f_2^0(x)) \\ = \ & \lim_{\xi\to 0}\int (f_w(x+\xi-\Phi w) - f_w(x-\Phi w)) f_w(w)dw && (7.30a) \\ = \ & \int \lim_{\xi\to 0}(f_w(x+\xi-\Phi w) - f_w(x-\Phi w)) f_w(w)dw && (7.30b) \\ = \ & 0 && (7.30c) \end{aligned}$$

where the last equation is due to the continuity of $f_w(\cdot)$ made in (A2) of Assumption 7.2.1. Using a similar argument as (7.30a)-(7.30c), it can be shown that $f_t^0(\cdot)$, $t \geq 2$ are all continuous.

Now consider the case where $x(0) = x_0 \neq 0$. According to (7.2) the predicted state in this case is the sum of $x(t)$ in (7.26) and $\Phi^t x_0$, then its density function can be expressed as

$$f_t(x;x_0) = f_t^0(x - \Phi^t x_0). \tag{7.31}$$

Then the continuity of $f_t(x;x_0)$ follows directly from the continuity of $f_t^0(\cdot)$ and function $z(x,x_0) = x - \Phi^t x_0$. To show the convergence of $f_t(x;x_0)$, $t \geq 1$, consider

$$\begin{aligned} & \lim_{t\to\infty}(f_{t+1}(x;x_0) - f_t(x;x_0)) \\ = & \lim_{t\to\infty}\int (f_t^0(x - \Phi^{t+1}x_0 - \Phi^t w) - f_t^0(x - \Phi^t x_0))f_w(w)dw = 0 \end{aligned} \tag{7.32}$$

The second equation holds true due to continuity of $f_t^0(\cdot)$ and (A1) of Assumption 7.2.1.

■

7.A.2 Proof of Theorem 7.4.1

Proof: (i) For each $\mathbf{w}_t \in V_{t,\varepsilon}$, the set of x_0 such that $x(t;\mathbf{w}_t,x_0) \in X_s$ is a convex set following (7.2) and convexity of X_s. Since $\hat{P}_t^\varepsilon = \bigcap_{\mathbf{w}_t \in V_{t,\varepsilon}} \{x_0 |\; x(t;\mathbf{w}_t,x_0) \in X_s\}$ and intersection of convex set is convex, $\hat{P}_t^\varepsilon$ is convex. Similarly, $\hat{O}_\infty^\varepsilon$, as the intersection of $\hat{P}_t^\varepsilon$ for all $t \geq 1$, is also a convex set.

(ii) Suppose $x_0 \in \hat{O}_\infty^\varepsilon$. The following shows that $x_1 \in \hat{O}_\infty^\varepsilon$ where $x_1 = x(1;w(0),x_0) =$

$\Phi x_0 + w(0)$ for any $w(0) \in W$. Given $x_0 \in \hat{O}^{\varepsilon}_{\infty}$, it follows from (7.14) that $x(t;\mathbf{w}_t,x_0) \in X_s$, $\forall \mathbf{w}_t \in V_{t,\varepsilon}$ and $\forall t \geq 1$. In particular, for any specific choice of $t = \bar{t}$ with $\bar{t} \geq 2$, this means that

$$x(\bar{t};\mathbf{w}_{\bar{t}},x_0) \in X_s, \text{ for any } \mathbf{w}_{\bar{t}} = [w(0),\cdots,w(\bar{t}-1)] \in V_{t,\varepsilon}$$

Let $w_0 = w(0)$ be a particular realization of $w(0)$, then there exists a $x(\bar{t};\mathbf{w}_{\bar{t}},x_0) = x(\bar{t}-1;\mathbf{w}_{\bar{t}-1},x_1)$ and $x(\bar{t}-1;\mathbf{w}_{\bar{t}-1}, x(\bar{t}-1;\mathbf{w}_{\bar{t}-1},x_1) \in X_s$. Let $\tilde{t} = \bar{t}-1$, then

$$x(\bar{t}-1;\mathbf{w}_{\bar{t}-1},x_1) \in X_s \forall \bar{t} \geq 2 \Rightarrow x(\tilde{t};\mathbf{w}_{\tilde{t}},x_1) \in X_s \forall \tilde{t} \geq 1$$

Hence, $x_1 \in \hat{O}^{\varepsilon}_{\infty}$. ■

7.A.3 Proof of Theorem 7.5.1

Proof: From the assumption that Y_∞ is non-empty, it follows from $Y_{i+1} = Y_i \ominus \Phi^i W$ and Assumption 7.5.1 that Y_i, $i \geq 1$ are non-empty and compact. In addition, $0 \in \text{int}(Y_i)$ since 0 is inside X_s, W and W_ε. From step (1) of Algorithm 7.5.1, $\hat{O}^{\varepsilon}_1 = \{x|\Phi x \in Y_1\}$ is compact since Y_1 is compact and Φ has full rank. From $\hat{O}^{\varepsilon}_{i+1} = \hat{O}^{\varepsilon}_i \cap \{x|\ \Phi^{i+1}x \in Y_{i+1}\}$ of step (3) of Algorithm 7.5.1,

$$\hat{O}^{\varepsilon}_{i+1} \subseteq \hat{O}^{\varepsilon}_i \ \ \forall i. \tag{7.33}$$

Since $0 \in \text{int}(Y_\infty)$ and $\rho(\Phi) < 1$ from (A1) of Assumption 7.2.1, there exist a finite integer $k \geq 1$, such that $\Phi^{k+1}\hat{O}_1^\varepsilon \subset Y_\infty$. This fact, together with $\hat{O}_k^\varepsilon \subseteq \hat{O}_1^\varepsilon$, imply that

$$\Phi^{k+1}\hat{O}_k^\varepsilon \subseteq \Phi^{k+1}\hat{O}_1^\varepsilon \subset Y_\infty \subset Y_{k+1}. \tag{7.34}$$

Since $\hat{O}_{i+1}^\varepsilon = \hat{O}_i^\varepsilon \cap \{x | \ \Phi^{i+1}x \in Y_{i+1}\}$, equation (7.34) implies that any $x \in \hat{O}_k^\varepsilon$ is also in $\hat{O}_{k+1}^\varepsilon$ or $\hat{O}_k^\varepsilon \subseteq \hat{O}_{k+1}^\varepsilon$. This and (7.33) implies that $\hat{O}_k^\varepsilon = \hat{O}_{k+1}^\varepsilon$. ■

7.A.4 Proof of Theorem 7.6.1

Proof: The proof follows those in Chapter 4, and hence is brief.

(i) Let $(\mathbf{d}^*(t), \mathbf{D}^*(t))$ denote the optimal solution of $\mathscr{P}_N^{PCA}(x(t))$. At time $t+1$ when $w(t)$ is realized, choose $(\hat{\mathbf{d}}(t+1), \hat{\mathbf{D}}(t+1))$ by letting

$$\hat{d}(i|t+1) = \begin{cases} d^*(i+1|t) + D^*(i+1, i+1|t)w(t), \ \forall i \in \mathbb{Z}_{N-2} \\ 0 \qquad\qquad i = N-1 \end{cases} \tag{7.35}$$

$$\hat{D}(i,j|t+1) = \begin{cases} D^*(i+1, j|t) \quad \forall j \in \mathbb{Z}_i^+, \ \forall i \in \mathbb{Z}_{N-2}^+ \\ 0 \qquad\qquad \forall j \in \mathbb{Z}_{N-1}^+, \ i = N-1 \end{cases} \tag{7.36}$$

Due to the result of Theorem 7.4.1 and the definition of $\hat{O}_\infty^\varepsilon$, constraints (7.24d) and (7.24e) are all satisfied by $(\hat{\mathbf{d}}(t+1), \hat{\mathbf{D}}(t+1))$. Hence, it is feasible to $\mathscr{P}_N^{PCA}(x(t+1))$ for all possible $w(t) \in W$. Let $\Pi_N^{PCA}(x)$ denote the feasible set of $\mathscr{P}_N^{PCA}(x)$. It is clear that $\Pi_N^{PCA}(x)$ is compact for all admissible x. Since W is bounded and J_{DFD} is a norm function, $\max_{w(t)} J_{DFD}(\hat{\mathbf{d}}(t+1),\hat{\mathbf{D}}(t+1)) < \infty$ and the set $\{(\mathbf{d},\mathbf{D}) \in \Pi_N^{PCA}(x(t+1))|J_{DFD}(\mathbf{d},\mathbf{D}) \leq \max_{w(t)} J(\hat{\mathbf{d}}(t+1),\hat{\mathbf{D}}(t+1))\}$ is compact. Hence, the optimum of $\mathscr{P}_N^{PCA}(x(t+1))$ exists, following the Weierstrass' theorem.

(ii) Following from (i), the hard constraints are satisfied all the time. $x(N|t) \in \hat{O}_\infty^\varepsilon$ implies that $\Pr(x(N+i|t) \in X_s|x(N|t)) \geq 1-\varepsilon$ for all $i \geq 1$. This together with constraint (7.24e) implies $\Pr(x(N+i|t) \in X_s|x(t)) \geq 1-\varepsilon$ for all $i \geq 1$. This and constraint (7.24d) implies $\Pr(x(i|t) \in X_s|x(t)) \geq 1-\varepsilon$ for all $i \geq 1$. Then the stated result follows this and (i).

(iii) Let $J_t^* := J_{DFD}(\mathbf{d}^*(t),\mathbf{D}^*(t))$ and $\hat{J}_{t+1}(w(t)) := J_{DFD}(\hat{\mathbf{d}}(t+1),\hat{\mathbf{D}}(t+1))$ where $(\hat{\mathbf{d}}(t+1),\hat{\mathbf{D}}(t+1))$ are given by (7.35)-(7.36). Then it can be shown that

$$J_t^* - \hat{J}_{t+1}(w(t)) = \|d^*(0|t)\|_\Psi^2 + g(w(t)) \tag{7.37}$$

where

$$g(w(t)) := \sum_{i=1}^{N-1} (\|\mathrm{vec}(D^*(i,i|t))\|_\Lambda^2 - 2(d^*(i|t))^T \Psi D^*(i,i|t)w(t) - \|D^*(i,i|t)w(t)\|_\Psi^2). \tag{7.38}$$

Due to (B2) of Assumption 7.6.1 and (7.23), $\mathrm{E}_{w(t)}[g(w(t))] \geq 0$. Hence, taking the expectation of (7.37) over $w(t)$, it follows that

$$J_t^* - \|d^*(0|t)\|_\Psi^2 \geq \mathrm{E}_t\left[J_{t+1}^*\right], \tag{7.39}$$

where E_t is the expectation taken over $w(i),\ i \geq t$. Repeating the inequality of (7.39) for increasing t, one gets,

$$\mathrm{E}_t[J^*_{t+1}] - \mathrm{E}_t[\|d^*(0|t+1)\|^2_\Psi] \geq \mathrm{E}_t[J^*_{t+2}]. \tag{7.40}$$

Summing (7.39) and (7.40) leads to

$$J^*_t \geq \|d^*(0|t)\|^2_\Psi + \mathrm{E}_t[\|d^*(0|t+1)\|^2_\Psi] + \mathrm{E}_t[J^*_{t+2}]$$

Repeating the above procedure infinite times leads to

$$\infty > J^*_t \geq \sum_{i=t}^{\infty} \mathrm{E}_t\left[\|d^*(0|i)\|^2_\Psi\right]$$

By applying Markov bound (given non-negative random variable R and any $\varepsilon \geq 0$, $\mathrm{E}[R] \geq \varepsilon \Pr\{R \geq \varepsilon\}$), we have

$$\infty > \varepsilon \sum_{i=t}^{\infty} \Pr(\|d^*(0|i)\|^2_\Psi \geq \varepsilon) \tag{7.41}$$

for any arbitrary small $\varepsilon > 0$. From the First Borel-Cantelli Lemma [75], this implies that $\lim_{i\to\infty} \Pr(\|d^*(0|i)\|^2_\Psi \geq \varepsilon) = 0$. Hence $d^*(0|i)$ approaches zero with probability one as t increases. Consequently, the MPC control law (7.25) converges to $K_f x(t)$ with probability one. When this happens, the closed-loop system converges to $x(t+1) = \Phi x(t) + w(t)$ and, hence, $x(t)$ converges to $F_\infty(K_f)$ with probability one. ■

Chapter 8

Conclusions

The main focus of this dissertation is on the design of control laws under the Model Predictive Control framework for constrained linear discrete-time systems with bounded disturbances. Emphasis is placed on the admissible set, stability results under several control parameterizations, computational considerations and the handling of probabilistic constraints. Specifically, this thesis contributes towards the four issues stated in Section 1.3.

8.1 Contributions of This Dissertation

P1: F_∞ convergence under MPC control law.

This thesis contributes towards achieving closed-loop F_∞ convergence by proposing

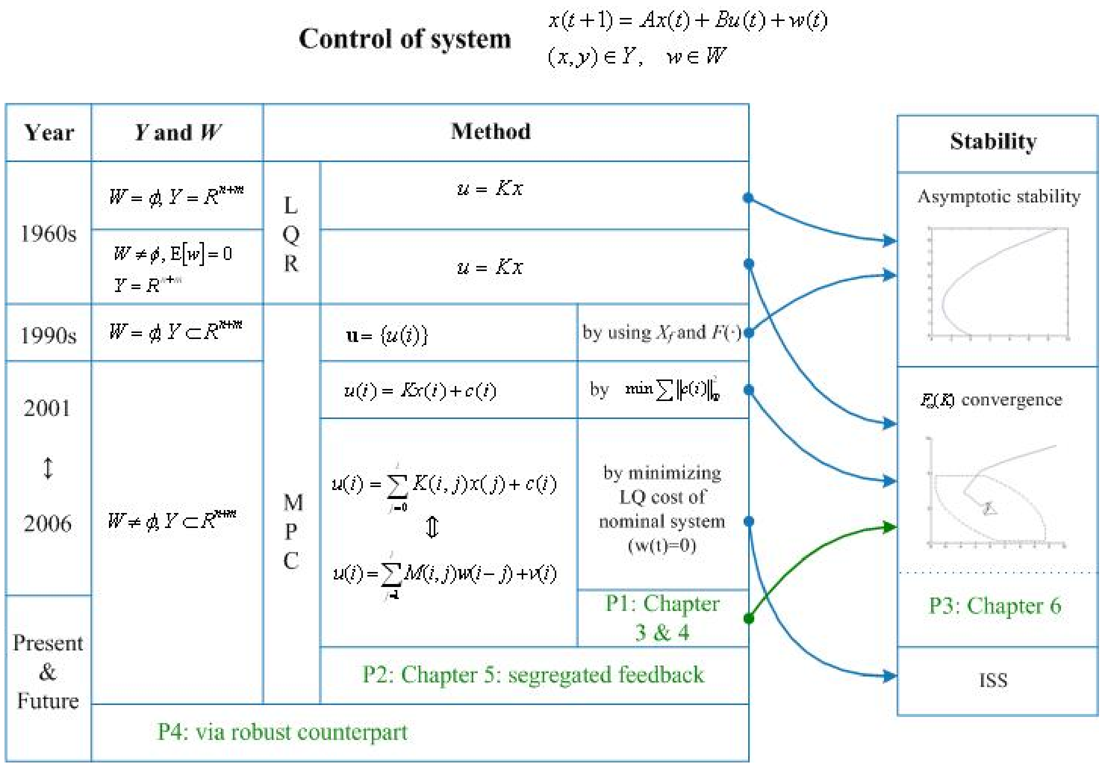

Figure 8.1: Contributions towards the issues in Section 1.3

three control parametrizations:

$$u^{DFC}(i|t) = K_f x(i|t) + c(i|t) + \sum_{j=1}^{N-1} C(i,j|t)w(i-j|t),$$

$$u^{DFD}(i|t) = K_f x(i|t) + d(i|t) + \sum_{j=1}^{i} D(i,j|t)w(i-j|t),$$

$$u^{SDF}(i|t) = K_f x(i|t) + d(i|t) + \sum_{j=1}^{i} D^p(i,j|t)\hat{w}^p(i-j|t) + \sum_{j=1}^{i} D^m(i,j|t)\hat{w}^m(i-j|t).$$

While the details can be different, these three parametrizations all achieve F_∞ convergence in some way. The use of u^{DFC} and the corresponding cost function results in $x(t)$ converging to F_∞ set as t tends to infinity, while the use of u^{DFD} and u^{SDF} results in $x(t)$ converging to F_∞ with probability one as t tends to infinity. Also u^{DFC} differs form u^{DFD} and u^{SDF} in that past realized disturbances are not needed for u^{DFD} and u^{SDF} in the feedback law. Parametrization u^{SDF} is a generalization of the other two and hence is expected to have the largest admissible set.

P2: A general control parametrization.

u^{SDF} also contributes towards the goal of having a more general parametrization under the MPC framework in the sense that it has the most expressive ability. However, its use in the FH optimization problem is not directly computable in general. Fortunately, this is not an issue when the disturbance set is an absolute set.

P3: The ability to control the F_∞ set.

The third contribution of this thesis is the approach towards designing a feedback gain K such that the minimal d-invariant set, $F_\infty(K)$ is well bounded in some given directions. The approach is to choose K such that the support functions of $F_\infty(K)$, $\delta_{F_\infty(K)}(\eta)$, for some given directions η are under some specific bound. When a set of η_i are used, the design problem is cast as an optimization problem to be solved by standard numerical techniques. This thesis provides expressions for the evaluation of $\delta_{F_\infty(K)}(\eta)$ and $\partial\delta_{F_\infty(K)}(\eta)/\partial K$ for those numerical techniques.

4: Probabilistic constraints in MPC.

The last contribution of this thesis is the treatment of probabilistic constraints in general and the formulation of the FH optimization problem of MPC of systems having probabilistic constraints. This leads to the concept of constraint-admissible set for system with probabilistic constraints.

8.2 Directions of Future Work

Several directions are available for future research and applications based on the work in this dissertation.

8.2.1 Output Feedback Parametrization

In this thesis, the system state is assumed to be measurable. This situation is not always possible as only output variables are available in some applications. Hence, extending the work of this thesis to the case where the control is parameterized by output feedback is one future direction. A more important issue is to guarantee the feasibility and stability of the system under an output feedback parametrization, especially when measurement errors are presented.

8.2.2 Computation of Admissible Set

The admissible set under a time-varying disturbance feedback control parametrization and segregated disturbance feedback is larger than that under a fixed feedback gain parametrization. The current approach of computing the admissible set relies on the projection algorithm [103] by Keerthi and Gilbert. However, this algorithm is not efficient especially when the dimension of the system is large. For this reason, it is significant if an efficient algorithm can be established. As far as the author can see, this may need knowledge in set operations and may be developed based on the work in [104] by

Kerrigan.

8.2.3 Distributed MPC

For application of MPC to large scale systems, decentralized approaches appear appealing. Decentralized control approaches date back to the seventies of the last century and are found in a broad spectrum of applications ranging from robotics, flight control to paper making industries [105]. For example, in a robotic football game, each robot has to plan its own actions and cooperate with others to score a goal; in the control of unmanned aerial vehicles (UAVs), each vehicle plans its own trajectory so that together they achieve certain formation and avoid collision onto each other. These applications fall under the general class of optimal control problems for a set of decoupled dynamical sub-systems where cost functions and constraints of the sub-systems are coupled. One possibility of extending the work presented in this thesis is to the area of distributed MPC.

Bibliography

[1] D. Bertsekas and I. Rhodes. Sufficiently informative functions and the minimax feedback control of uncertain dynamic systems. *IEEE Transactions on Automatic Control*, 18(2):117–124, 1973.

[2] M. A. Dahleh and I. J. Diaz-Bobillo. *Control of uncertain systems: a linear programming approach*. Prentice Hall, 1995.

[3] M. V. Kothare, V. Balakrishnan, and M. Morari. Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32(10):1361–1379, 1996.

[4] A. Bemporad and M. Morari. Fulfilling hard constraints in uncertain linear systems by reference managing. *Automatica*, 34(4):451–462, 1998.

[5] F. Blanchini. Set invariance in control. *Automatica*, 35(11):1747–1767, 1999.

[6] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.

[7] J. M. Maciejowski. *Predictive control with constraints*. UK: Prentice Hall, 2002.

[8] W. Langson, I. Chryssochoos, S. V. Raković, and D. Q. Mayne. Robust model predictive control using tubes. *Automatica*, 40(1):125–133, 2004.

[9] A. Bemporad. Reducing conservativeness in predictive control of constrained systems with disturbances. In *Proceedings of 37th Conference on Decision and Control*, pages 1384–1389, Tampa, Florida, 1998.

[10] L. Chisci, J. A. Rossiter, and G. Zappa. Systems with persistent disturbances: predictive control with restricted constraints. *Automatica*, 37(7):1019–1028, 2001.

[11] J. A. Rossiter, B. Kouvaritakis, and M. J. Rice. A numerically robust state-space approach to stable predictive control strategies. *Automatica*, 34(1):65–73, 1998.

[12] Y. I. Lee and B. Kouvaritakis. Constrained receding horizon predictive control for systems with disturbances. *International Journal of Control*, 72(11):1027–1032, 1999.

[13] D. Q. Mayne, M. M. Seron, and S. V. Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219–224, 2005.

[14] T. Alamo, D. Muñoz de la Peña, D. Limon, and E. F. Camacho. Constrained min-max predictive control: modifications of the objective function leading to polynomial complexity. *IEEE Transactions on Automatic Control*, 50(5):710–714, 2005.

[15] D. Muñoz de la Peña, T. Alamo, A. Bemporad, and E. F. Camacho. A decompo-

sition algorigthm for feedback min-max model predictive control. *IEEE Transactions on Automatic Control*, 51(10):1688–1692, 2006.

[16] R. Cagienard, P. Grieder, E. C. Kerrigan, and M. Morari. Move blocking strategies in receding horizon control. *Journal of Process Control*, 17(6):563–570, 2007.

[17] S. J. Qina and T. A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764, 2003.

[18] M. Athans and Falb P. L. *Optimal Control*. McGraw-Hill, New York, 1966.

[19] A. E. Bryson and Y. C. Ho. *Applied Optimal Control*. Hemisphere Publishing Co., Washington, 1975.

[20] E. B. Lee and L. Markus. *Foundations of Optimal Control Theory*. Wiley, New York, 1967.

[21] R. R. Bitmead, M. Gevers, and V. Wertz. *Adaptive optimal control: The thinking man's GPC*. Prentice-Hall, Englewood Cliffs, NJ, 1990.

[22] J. B. Rawlings and K. R. Muske. The stability of constrained receding horizon control. *IEEE Transactions on Automatic Control*, 38(10):1512–1516, 1993.

[23] P. Couchman, B. Kouvaritakis, and M. Cannon. Dual receding horizon control of constrained discrete time systems. *European Journal of Control*, 2(4):278–285, 1996.

[24] P. O. M. Scokaert, D. Q. Mayne, and J. B. Rawlings. Suboptimal model predictive control (feasibility implies stability). *IEEE Transactions on Automatic Control*, 44(3):648–654, 1999.

[25] D. Sznaier and M. J. Damborg. Suboptimal control of linear systems with state and control inequality constraints. In *Proceedings of the 26th IEEE Conference on Decision and Control*, pages 761–762, Los Angeles, USA, 1987.

[26] G. De Nicolao, L. Magni, and R. Scattolini. Stabilizing nonlinear receding horizon control via a nonquadratic terminal state penalty. In *Proceedings of the IMACS multiconference CESA*, pages 185–187, Lille, France, 1996.

[27] D. Q. Mayne, S. V. Raković, R. Findeisen, and F. Allgöwer. Robust output feedback model predictive control of constrained linear systems. *Automatica*, 42(7):1217–1222, 2006.

[28] H. Michalska and D. Q. Mayne. Robust receding horizon control of constrained nonlinear systems. *IEEE Transactions on Automatic Control*, 38(11):1623–1633, 1993.

[29] T. A. Badgwell. Robust model predictive control of stable linear systems. *International Journal of Control*, 68(4):797–818, 1997.

[30] P. O. M. Scokaert and D. Q. Mayne. Min-max feedback model predictive control for constrained linear systems. *IEEE Transactions on Automatic Control*, 43(8):1136–1142, 1998.

[31] A. Bemporad, F. Borrelli, and M. Morari. Min-max control of constrained uncertain discrete-time linear systems. *IEEE Transactions on Automatic Control*, 48(9):1600–1606, 2003.

[32] E. C. Kerrigan and J. M. Maciejowski. Robustly stable feedback min-max model predictive control. In *Proceedings of the 2003 American Control Conference*, pages 3490–3495, Denver, Colorado, USA, 2003.

[33] E. C. Kerrigan and J. M. Maciejowski. Feedback min-max model predictive control using a single linear program: Robust stability and the explicit solution. *International Journal of Robust and Nonlinear Control*, 14(4):395–413, 2004.

[34] D. Muñoz de la Peña, D. R. Ramírez, E. F. Camacho, and T. Alamo. Explicit solution of mincmaxmpcwith additive uncertainties and quadratic criterion. *Automatica*, 55(4):266C274, 2006.

[35] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 38(1):3–20, 2002.

[36] P. J. Goulart, E. C. Kerrigan, and D. Ralph. Efficient robust optimization for robust control with constraints. *Mathematical Programming*, 114(1):115–147, 2008.

[37] I. V. Kolmanovsky and E. G. Gilbert. Theory and computation of disturbance invariant sets for discrete-time linear systems. *Mathematical Problems in Engineering*, 4(4):317–367, 1998.

[38] R. S. Smith. Robust model predictive control of constrained linear systems. In

Proceedings of the 2004 American Control Conference, pages 245–250, Boston, MA, USA, 2004.

[39] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski. Optimization over state feedback policies for robust control with constraints. *Automatica*, 42(4):523–533, 2006.

[40] D. Jia, B. H. Krogh, and O. Stursberg. LMI approach to robust model predictive control. *Journal of Optimization Theory and Applications*, 127(2):347–365, 2005.

[41] J. Löfberg. Approximations of closed-loop minimax MPC. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, pages 1438–1442, Maui, Hawaii, USA, 2003.

[42] E. C. Kerrigan and J. M. Maciejowski. Properites of a new parameterization for the control of constrained systems with disturbances. In *Proceedings of the 2004 American Control Conference*, pages 4669–4674, Boston, Massachusetts, USA, 2004.

[43] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Ajustable robust solutions of uncertaint linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

[44] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, Engineering Applications*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

[45] Z. Jiang and Y. Wang. Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 37(6):857–869, 2001.

[46] V. Kucera. *Discrete Linear Control*. New York: Wiley, 1979.

[47] K. J. Åström and B. Wittenmark. *Computer-Controlled Systems: Theory and Design*. Prentice Hall, 3 edition, 1997.

[48] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[49] L. Pu, M. Wendt, and G. Wozny. A probabilistically constrainted model predictive controller. *Automatica*, 38(7):1171–1176, 2002.

[50] B. Kouvaritakis, M. Cannon, and V. Tsachouridis. Recent developments in stochastic MPC and sustainable development. *Annual Revies in Control*, 28(1):23–35, 2004.

[51] Alexander T. Schwarm and Michael Nikolaou. Chance-constrained model predictive control. *AIChE Journal*, 45(8):1743–1752, 1999.

[52] I. V. Kolmanovsky and E. G. Gilbert. Maximal output admissible sets for discrete-time systems with disturbance inputs. In *Proceedings of the 1995 American Control Conference*, pages 1995–1999, Seattle, 1995.

[53] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[54] D. Q. Mayne and W. R. Schroeder. Robust time-optimal control of constrained linear systems. *Automatica*, 33(12):2103–2118, 1997.

[55] K. Hirata and Y. Ohta. ε-feasible approximation of the state reachable set for discrete-time systems. In *Proceedings of the 42nd Conference on Decision and Control*, pages 5520–5525, 2003.

[56] S. V. Raković, E. C. Kerrigan, K. I. Kouramas, and D. Q. Mayne. Invariant approximations of the minimal robust positively invariant set. *IEEE Transactions on Automatic Control*, 50(3):406–410, 2005.

[57] C.-J. Ong and E. G. Gilbert. The minimal disturbance invariant set: Outer approximations via its partial sums. *Automatica*, 42(9):1563–1568, 2006.

[58] F. Blanchini. Feedback control for linear time-invariant systems with state and control bounds in the presence of disturbances. *IEEE Transactions on Automatic Control*, 35(11):1231–1234, 1990.

[59] J. P. Aubin. *Viability Theory*. Wiley-Interscience, 1991.

[60] D. P. Bertsekas. Infinite-time reachability of state-space regions by using feedback control. *IEEE Transactions on Automatic Control*, 17(5):604– 613, 1972.

[61] E. G. Gilbert and K. T. Tan. Linear systems with state and control constraints: the theory and applications of maximal output admissible sets. *IEEE Transactions on Automatic Control*, 36(9):1008–1020, 1991.

[62] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[63] A. Ben-Tal and A. Nemirovski. Robust solutions to uncertain linear programs. *OR Letters*, 25, 1999.

[64] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[65] P. J. Goulart and E. C. Kerrigan. On the stability of a class of robust receding horizon control laws for constrained systems. Technical report, Department of Engineering, University of Cambridge, August 2005. CUED/F-INFENG/TR.532.

[66] P. J. Goulart and E. C. Kerrigan. Robust receding horizon control with an expected value cost. In *Proc. UKACC International Conference (Control 2006)*, Glasgow, Scotland, August 2006.

[67] P. J. Goulart and E. C. Kerrigan. Input-to-state stability of robust receding horizon control with an expected value cost. *Automatica*, 44(4):1171–1174, 2008.

[68] C. Wang, C.-J. Ong, and M. Sim. Model predictive control using affine disturbance feedback for constrained linear system. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 1275–1280, New Orleans, Louisiana, USA, 2007.

[69] C. Wang, C.-J. Ong, and M. Sim. Constrained linear system with disturbance: Convergence under disturbance feedback. *Automatica*, 44(10):2583–2587, 2008.

[70] Arkadi Nemirovski. Advances in convex optimization: conic programming. In *Proceedings of the International Congress of Mathematicians*, pages 413–444, Madrid, 2006.

[71] M. S. Lobo, L. Vandenberghe, S. Boyd, and H Lebret. Applications of second-

order cone programming. *Linear Algebra and its Applications*, 284:193–228, 1998.

[72] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Methimetical Programming*, 95(1):3–51, 2003.

[73] P. J. Goulart. *Affine Feedback Policies for Robust Control with Constraints*. PhD thesis, Department of Engineering, University of Cambridge, UK, January 2007.

[74] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 2nd edition, 2003.

[75] David Williams. *Probability with Martingales*. Cambridge University Press, 1991.

[76] C. Wang, C.-J. Ong, and M. Sim. Model predictive control using segregated disturbance feedback. In *Proceedings of the 2008 American Control Conference*, pages 3566–3571, Seattle, Washington, USA, 2008.

[77] A. L. Soyster. Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming. *Operations Research*, 21(5):1154–1157, 1973.

[78] D. Bertsimas and M. Sim. The Price of Robustness. *Operations Research*, 52(1):35–53, 2004.

[79] S. P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[80] C.-J. Ong and E. G. Gilbert. The minimal disturbance invariant set: Outer approximations via its partial sums. *Automatica*, 42(9):1563–1568, 2006.

[81] D. Bertsimas and M. Sim. Tractable approximations to robust conic optimization problems. *Mathematical Programming*, 107(1):5–36, 2006.

[82] P. D. Lax. *Linear Algebra*. John Wiley, 1997.

[83] P. Caravani and E. De Santis. A polytopic game. *Automatica*, 36(7):973–981, 2000.

[84] S. V. Raković, E. De Santis, and P. Caravani. Invariant equilibria of polytopic games via optimized robust control invariance. In *Proceedings of the 44th Conference on Decision and Control*, pages 7686–7691, Seville, Spain, 2005.

[85] F. Blanchini. Ultimate boundness control for uncertain discrete-time systems via set-induced lyapunov functions. *IEEE Transactions on Automatic Control*, 39(2):428–433, 1994.

[86] C. E. T. Dórea and J. C. Hennet. (A,B)-invariant polyhedral sets of linear discrete-time systems. *Journal of Optimization Theory and Applications*, 103(3):521–542, 1999.

[87] E. De Santis, M. D. Di Benedetto, and L. Berardi. Computation of maximal safe sets for switching systems. *IEEE Transactions on Automatic Control*, 49(2):184–195, 2004.

[88] E. G. Gilbert, I. V. Kolmanovsky, and K. T. Tan. Nonlinear control of discrete-time systems with state and control constraints: a reference governor with global

convergence properties. In *Proceedings of the 33rd Conference on Decision and Control*, pages 144–149, 1994.

[89] C.-J. Ong and E. G. Gilbert. Constrained linear systems with disturbances: enlargement of their maximal invariant sets by nonlinear feedback. In *Proceedings of the 2006 American Control Conference*, pages 5245–5250, Minneapolis, Minnesota, USA, 2006.

[90] A. Bemporad, F. D. Torris, and M. Morari. Performance analysis of piecewise linear systems and model predictive control systems. In *Proceedings of 39th Conference on Decision and Control*, pages 4957–4962, Sydney, Australia, 2000.

[91] K. Hirata and Y. Ohta. Exact determinations of the maximal output admissible set for a class of nonlinear systems. In *Proceedings of the 44th Conference on Decision and Control*, pages 8276–8281, Seville, Spain, 2005.

[92] B. Pluymers, J. A. Rossiter, J. A. K. Suykens, and B. De Moor. Interpolation based MPC for LPV systems using polyhedral invariant sets. In *Proceedings of the 2005 American Control Conference*, pages 810–815, Portland, OR, USA, 2005.

[93] G. Bitsoris. Positive invariant polyhedral sets of discrete-time linear systems. *International Journal of Control*, 47:1713–1726, 1988.

[94] G. Bitsoris. On positive invariance of polyhedral sets for discrete-time systems. *System and Control Letters*, 11:243–248, 1988.

[95] M. Vassilaki, J. C. Hennet, and G. Bitsoris. Feedback control of linear discrete-time systems under state and control contraints. *International Journal on Control*, 47:1727–1735, 1988.

[96] E. C. Kerrigan and J. M. Maciejowski. Invariant sets for constrained nonlinear discrete-time systems with application to feasibility in model predictive control. In *Proceedings of 39th Conference on Decision and Control*, pages 4951–4956, Sydney, Australia, 2000.

[97] S. V. Raković, E. C. Kerrigan, D. Q. Mayne, and K. I. Kouramas. Optimized robust control invariance for linear discrete-time systems: Theoretical foundations. *Automatica*, 43(5):831–841, 2007.

[98] John R. Birge and Francois Louveaux. *Introduction to Stochastic Programming*. Springer Verlag, New York, 1997.

[99] F. Oldewurtel, C. N. Jones, and M. Morari. A tractable approximation of chance constrained stochastic mpc based on affine disturbance feedback. In *Proceedings of the 47th IEEE Conference on Decision and Control*, pages 4731–4736, Cancun, Mexico, 2008.

[100] James A. Primbs. A soft constraint approach to stochastic receding horizon control. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 4797 – 4802, New Orleans, Louisiana, USA, 2007.

[101] E. G. Gilbert and C.-J. Ong. Linear systems with hard constraints and variable set points: Their robustly invariant sets. Technical report, De-

partment of Aerospace Engineering, University of Michigan and Department of Mechanical Engineering, National University of Singapore, available at http://guppy.mpe.nus.edu.sg/ mpeongcj/ongcj.html, Jan 2008. C08-002.

[102] J. Hoffmann-Jørgensen. *Probability with a View Toward Statistics, Volumes I and II*. Chapman-Hall, New York., 1994.

[103] S. S. Keerthi and E. G. Gilbert. Computation of minimum-time feedback control laws for discrete-time systems with state-control constraints. *IEEE Transactions on Automatic Control*, AC-32(5):432–435, 1987.

[104] E. C. Kerrigan. *Robust Constraint Satisfaction: Invariant Sets and Predictive Control*. PhD thesis, Department of Engineering, University of Cambridge, UK, November 2000.

[105] F. Borrelli, T. Keviczky, G. J. Balas, G. Stewart, K. Fregene, and D. Godbole. Hybrid decentralized control of large scale systems. *Lecture Notes in Computer Science*, 3414:168–183, 2005.

Author's Publications

Journal Paper

1. C. Wang, C.-J. Ong, and M. Sim. Constrained linear system with disturbance: Convergence under disturbance feedback. *Automatica*, pages 2583-2587, 44(10), 2008.

2. C. Wang, C.-J. Ong, and M. Sim. Convergence properties of constrained linear system under MPC control law using affine disturbance feedback. *Automatica*,45(7), 1715-1720, 2009.

3. C. Wang, C.-J. Ong and M. Sim, Model predictive control using segregated disturbance feedback. To appear in *IEEE Transactions on Automatic Control*, April 2010.

Conference Papers

1. C. Wang, C. J. Ong, and M. Sim. Model predictive control using affine disturbance feedback for constrained linear system. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 1275-1280, New Orleans, Louisiana, USA, 2007.

2. C. Wang, C. J. Ong, and M. Sim. Model predictive control using segregated disturbance feedback. In *Proceedings of the 2008 American Control Conference*, pages 3566-3571, Seattle, Washington, USA, 2008.

3. C. Wang, C. J. Ong, and M. Sim. Constrained linear system under disturbance feedback: convergence with probability one. In *Proceedings of the 47th IEEE Conference on Decision and Control*, pages 2820-2825, Cancun, Mexico, 2008.

4. C. Wang, C. J. Ong, and M. Sim. Support function of minimal disturbance invariant set and its derivative: application in designing feedback gain. In *Proceedings of 10th Singapore-MIT-Alliance Anniversary Symposium*, Singapore, 2009.

5. C. Wang, C.-J. Ong, and M. Sim, Linear systems with chance constraints: constraint-admissible set and applications in predictive control. Accepted in *Proceedings of the 48th IEEE Conference on Decision and Control*, Shanghai, China, 2009.

Technical Report

1. C. Wang, C. J. Ong, and M. Sim. Convergence Properties of Constrained Linear System under MPC Control Law using Affine Disturbance Feedback. National University of Singapore, available at http://guppy.mpe.nus.edu.sg/mpeongcj/ongcj .html, C09-001, Jan 2009.

2. C. Wang and C.-J. Ong. Linear Systems with Soft Constraints and Stochastic Disturbances. National University of Singapore, available at http://guppy.mpe.nus.edu .sg/mpeongcj/ongcj.html, C09-003, August 2009.