Robust and Stochastic Receding Horizon Control

Shuhao Yan
Wolfson College
University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2021

To my family

Acknowledgements

Firstly and most importantly, my deepest gratitude goes to Professor Mark Cannon and Professor Paul Goulart. I am very fortunate to have them as the supervisors of my doctoral studies. I must thank them for their guidance, help, support and patience, without which this thesis would not have been possible. I also would like to thank Professor Kostas Margellos and Professor Lorenzo Fagiano for being my thesis examiners.

Further thanks go to previous and current members of the Control Group, especially Xiaonan Lu, Yang Zheng, Li Dai, Licio Romao, Xiaoqing Chen, Nikitas Rontsis, Michael Garstka, Bartolomeo Stellato, Goran Banjac and Sebastian East, for making it such a wonderful place to work and for fruitful discussions.

Finally, I very much appreciate the love and support of my family, especially my fiancée Yinru Chen. Thank you for being there for me. My life would not have been the same without you.

Abstract

Chance constraints, unlike robust constraints, allow constraint violation up to some predefined level and arise in numerous applications. They are often imposed in a pointwise-in-time fashion in control problems. This thesis considers a class of chance constraints imposed in an average-in-time fashion to focus more on aggregate behaviours and discounted to achieve trade-offs between short-term and long-term performance in the model predictive control (MPC) framework.

This thesis designs an MPC law for chance constrained stochastic systems with discrete-time linear dynamics and possibly unbounded additive disturbances. The chance constraint is defined as a discounted sum of violation probabilities over an infinite horizon. By penalising violation probabilities close to the initial time and assigning violation probabilities in the far future with vanishingly small weights, this form of constraints allows for an MPC law with guarantees of recursive feasibility by introducing an online constraint-tightening technique without an assumption of boundedness of the disturbance. We employ Chebyshev's inequality for constraint handling and formulate a computationally simple MPC optimisation problem. To mitigate the conservativeness of Chebyshev's inequality, a dynamic feedback gain is incorporated into the MPC law. This gain is selected online from a set of candidates generated by Pareto optimal solutions of a multiobjective optimisation problem. The closed loop system is guaranteed to satisfy the chance constraint and a quadratic stability condition. With dynamic feedback gain selection, the closed loop cost is reduced and a larger set of feasible initial conditions is obtained.

This thesis also considers an application of stochastic MPC in networked control systems, where constrained linear systems are subject to stochastic additive disturbances and noisy measurements transmitted over a lossy communication channel. An MPC controller is designed to minimise a discounted cost subject to a discounted expectation constraint. Sensor data is assumed to be lost with a known probability. Data losses are accounted for by expressing the predicted control policy as an affine function of future observations, resulting in a convex optimal control problem. Recursive feasibility of online optimisation problems and constraint satisfaction are ensured similarly via the constraint-tightening technique. We show that the discounted cost evaluated along trajectories of the closed loop system is bounded. Under certain conditions, the averaged undiscounted closed loop cost accumulated over an infinite horizon also remains bounded.

Contents

List of Figures

List of Tables

List of Notations and Abbreviations

Definitions and inequality

$A := B$ . . . . . A is defined by B

$A =: B$ . . . . . B is defined by A

$A \succ 0$ ($A \succeq 0$) . A is a symmetric positive definite (semidefinite) matrix

Sets and set operation

$\mathbb{R}$ . . . . the set of real numbers

$\mathbb{R}^n$ . . . . the set of n-dimensional real vectors

$\mathbb{R}^{m\times n}$ . . . . . . the set of m-by-n real matrices

$\mathcal{X} \oplus \mathcal{Y}$ . . . . . the Minkowski sum of sets $\mathcal{X}$ and $\mathcal{Y}$: $\mathcal{X} \oplus \mathcal{Y} := \{z|z = x+y, \forall\, x \in \mathcal{X}, y \in \mathcal{Y}\}$

Scalars and vectors

$|x|$ . . . . the absolute value of a real number x

$\lim_{x\to a^-} f(x)$ . . the one-sided limit of $f(x)$ as x approaches a from the left

$\|x\|$ . . . . the Euclidean norm of a vector x

$\|x\|_Q^2$ . . . . . . $x^\top Qx$

$\{x_i\}_{i=0}^{N-1}$ . . . . the sequence $\{x_0, \dots, x_{N-1}\}$

x_k . . . . the value of a variable x at time k

$x_{i|k}$ . . . . the i-step-ahead predicted value of x at time k

Matrices and matrix operations

$\mathbf{1}$ . . . . a matrix with all elements equal to 1 and a conformal dimension

$\mathbf{1}_{n\times m}$ . . . . . . the n-by-m matrix with all elements equal to 1

I . . . . an identity matrix with a suitable dimension

$I_{n\times n}$ . . . . . . the n-by-n identity matrix

$A^\dagger$ . . . . the Moore–Penrose inverse of a matrix A

$\mathrm{tr}(A)$ . . . . . . the trace of a square matrix A

$\mathrm{vec}(A)$ . . . . . the vectorised form of a matrix $A = [a_1 \ \cdots \ a_n]$: $[a_1^\top \ \cdots \ a_n^\top]^\top$

$A \otimes B$ . . . . . the Kronecker product of matrices A and B

$\mathrm{diag}\{A_1, \ldots, A_n\}$ the (block) diagonal matrix with main-diagonal elements (blocks) being (square matrices) $A_1, \ldots, A_n$

$\mathrm{mat}(x, [r, c])$ . . the operation to reshape a vector x of appropriate length into an r-by-c matrix

Probabilities and expectations

$\mathbb{P}\{\mathcal{E}\}$ . . . . . . the probability of an event $\mathcal{E}$

$\mathbb{P}\{\mathcal{E}|\mathcal{F}\}$ . . . . . the conditional probability of an event $\mathcal{E}$ given event $\mathcal{F}$

$\mathbb{P}_k\{\mathcal{E}\}$ . . . . . the conditional probability of an event $\mathcal{E}$ given information available at time k and $\mathbb{P}_0\{\mathcal{E}\}$ is equivalent to $\mathbb{P}\{\mathcal{E}\}$

$\mathbb{E}\{y\}$ . . . . . . the expectation of a random variable y

$\mathbb{E}\{y|x\}$ . . . . . the conditional expectation of y given x

$\mathbb{E}_k\{y\}$ . . . . . . the conditional expectation of y given information available at time k and $\mathbb{E}_0\{y\}$ is equivalent to $\mathbb{E}\{y\}$

Acronyms

DP . . . . Dynamic Programming

i.i.d. . . . . . . Independently and Identically Distributed

LHS . . . . . . Left Hand Side

LMI . . . . . . Linear Matrix Inequality

LQG . . . . . . Linear Quadratic Gaussian

LQR . . . . . . Linear Quadratic Regulation

MOP . . . . . Multiobjective Optimisation Problem

MPC . . . . . Model Predictive Control

MPI . . . . . . Maximal Positively Invariant

mRPI . . . . . Minimal Robust Positively Invariant

MSS . . . . . . Mean-square Stable

QCQP . . . . . Quadratically Constrained Quadratic Program

RHS . . . . . . Right Hand Side

SOCP . . . . . Second-order Cone Program

1
Introduction

Contents

1.1 Background

Optimal control problems have been well studied, and they can be seen as a derivative of classical calculus of variations [1]. Enormous contributions have been made since the 1950s, including Pontryagin's maximum principle [2] and dynamic programming (DP) [3]. The former gives the first necessary conditions of optimality for optimal control problems [4], and the latter provides techniques to divide a problem into sub-problems and solve them in a piecemeal fashion. Well-known optimal control problems include linear quadratic regulator (LQR) and linear quadratic Gaussian (LQG) problems. The LQR problem is concerned with minimising a quadratic cost function subject to a linear system dynamics, and the optimal control strategy is given by a linear state feedback control law. LQG problems consider linear systems with additive Gaussian disturbances, and, by the certainty equivalence principle [5], yield the same optimal control strategy as their deterministic counterparts, LQR problems.

Constrained optimal control problems are often solved in practical control applications, which are more challenging to deal with than unconstrained problems. Although dynamic programming can be used to compute optimal controllers in very general settings, it is computationally demanding and intractable for practical applications with large numbers of state variables. This is referred to as the "curse of dimensionality". Therefore, many efforts have been made to design suboptimal controllers for constrained optimal control problems, trading optimality for constraint satisfaction and stability. A successful and widely used example is model predictive control (MPC).

MPC is also referred to as receding horizon control, which repeatedly solves optimisation problems online at every sampling instant to compute a finite horizon control input sequence, in which only the first control action is executed. It is an advanced control method initially emerging in the petroleum industry, and now has a wide range of applications because of its capabilities of systematic constraint handling and controlling multivariable systems and its convenient implementation, yielding satisfactory performance without frequent specialist interventions [6]. MPC relies on accurate models to predict future behaviours of systems being controlled, while uncertainty is inevitably present, leading to discrepancies between predicted and actual responses.

Presence of uncertainty necessitates robust control strategies, since nominal MPC has limited inherent robustness with respect to uncertainty [7]. In the existing MPC literature, it is typically considered that uncertainty appears in system dynamics and includes additive process disturbances, which are possibly adverse exogenous inputs, and parametric uncertainty, which affects system and input matrices. Robust MPC ensures constraint satisfaction for all possible realisations of uncertainty, which is essential in safety-critical systems. It is rather involved to take into account uncertainty and the effects of its propagation and to ensure tractability of resulting optimisation problems. A large amount of work [e.g. 8–11] has been done on these problems. A main drawback of robust MPC that has been widely mentioned is its conservativeness, which is due to the effort made to account for rare eventualities.

To reduce conservativeness in controller design, it would be reasonable to consider chance constraints, which are also referred to as soft constraints. Comparing with hard constraints, chance constraints ensure that certain constraints are satisfied in expectation or with pre-specified probabilities. They are of practical relevance and are considered in many applications, such as building temperature control [12] and portfolio management [13]. In addition to imposing chance constraints, stochastic MPC also employs probabilistic descriptions for uncertainty. This enables explicit regulation of probability distributions of states and control inputs [14].

In the existing stochastic MPC literature, although probabilistic descriptions for uncertainty are used, it is often assumed that they have bounded support. Probability distributions with unbounded support are very common in physical models, either as a natural consequence of first-principle modelling or because of mathematical convenience. Some particular examples of the former are Weibull and Rayleigh distributions which are used to model wind speed, often with Dryden and von Karman turbulence models, in wind turbine and aerospace control problems. A well-known example of the latter is the normal distribution, which, in addition to being fundamental to many control problems due to the central limit theorem, has the property that the sum of two normal random variables is also normally distributed. Also, truncating an unbounded distribution to obtain a distribution with bounded support requires an important but essentially arbitrary decision on the level of confidence of the truncation. Nevertheless, it is understood that considering unbounded disturbances brings additional complications, notably recursive feasibility. Previous work [15–17] has resorted to a backup initialisation of the MPC optimisation problem when infeasibility occurs. This design can prevent feedback from the actual state measurement and could thus yield inadequate performance.

Discounting has a role to play in overcoming challenges associated with unbounded disturbances. The discount setting is common in many stochastic control problems [e.g. 18–22], as well as in reinforcement learning [23], financial engineering [e.g. 24–27] and ecosystem management [e.g. 28, 29]. Discount factors in optimal control problems allow performance in the near future to be prioritised over long-term behaviours. This shift of emphasis is vital for ensuring the recursive feasibility of chance constrained control problems involving possibly unbounded disturbances. In Dynamic Programming [e.g. 18, 19], discounting is commonly employed to ensure that infinite horizon problems with possibly unbounded cost per stage are well-defined. Also, given a discounted cost in MPC problems, stability guarantees can still be provided with an appropriate value of the discount factor [30, 31].

As a key feature in stochastic MPC, a large amount of work focuses on handling of chance constraints in the MPC framework. In general, chance constraints imply nonconvex feasible sets. Therefore, various approaches exist for providing tractable approximations of chance constraints, for example, exact evaluation using available information of probability distributions [32], deterministic approximation methods that are combined with confidence sets [33], stochastic tubes [34] and probabilistic reachable sets [16] respectively, conservative approximation using probabilistic inequalities [15, 35] and randomised methods [36, 37]. More importantly, a fundamental problem is how to incorporate a method providing nonconservative approximations of chance constraints into the MPC framework to ensure desirable closed loop properties, including recursive feasibility, constraint satisfaction and stability.

Chance constraints are often imposed in a pointwise-in-time fashion in control problems, where violation probabilities at each individual time step are equally taken into account. A different formulation is also widely considered, for example, in [38–41], where chance constraints are imposed in an average-in-time fashion. Unlike pointwise-in-time chance constraints, average-in-time constraints focus more on aggregate performance over a finite horizon, which leads to a less conservative controller design. In these cases, it also allows for adjusting control strategies online by observing past constraint violations to further reduce conservativeness [38, 42]. Average-in-time chance constraints also are more natural formulations of chance constraints in some practical applications, such as economics [43]. In this type of problems, discounting is naturally present and represents an interest rate. Discount factors enable the consideration of average-in-time chance constraints defined over longer and infinite horizons when disturbances are unbounded [44]. In addition, average-in-time chance constraints can be imposed specifically with discounting to suit design purposes. Examples include sustainable development [45] and public policy analysis [46], where the main argument is that near-term benefits are more important than those that would be gained in the future and current actions should not imperil possible future development [22].

This thesis aims to propose a holistic approach for chance constrained linear systems with possibly unbounded additive disturbances, which ensures desirable closed loop properties without breaking direct state feedback and reduces conservativeness in constraint handling.

1.2 Outline

This thesis is concerned with model predictive control for chance constrained discrete-time linear systems with possibly unbounded additive disturbances, which take the form

$$x_{k+1} = Ax_k + Bu_k + \omega_k.$$

The chance constraint considered is defined in an average-in-time fashion as an infinite discounted sum of violation probabilities/expectations. The designed MPC algorithms ensure recursive feasibility of online MPC optimisation problems and constraint satisfaction and stability for the closed loop system without bounds on disturbances. In the following, we summarise each chapter of this thesis.

Chapter 2: This chapter introduces basic definitions and results from the topics of model predictive control, chance constraints and Kalman filtering. It provides a brief introduction to material that is closely related to subsequent chapters.

Chapter 3: This chapter designs a model predictive control law to minimise a quadratic cost function subject to a chance constraint. The chance constraint is defined as a discounted sum of violation probabilities on an infinite horizon. By penalising violation probabilities close to the initial time and assigning violation probabilities in the far future with vanishingly small weights, this form of constraints enables the feasibility of the online optimisation to be guaranteed without an assumption of boundedness of the disturbance. A computationally convenient MPC optimisation problem is formulated using Chebyshev's inequality, and we introduce an online constraint-tightening technique to ensure recursive feasibility based on knowledge of a suboptimal solution. The closed loop system is guaranteed to satisfy the chance constraint and we show that it satisfies a quadratic stability condition. The results in this chapter have been published in [44].

Chapter 4: This chapter lays the foundation for further developments in Chapter 5. We consider a convex multiobjective optimisation problem subject to the dynamics of a discrete-time linear system. The multiobjective optimisation problem is cast as a single objective one via linear scalarisation and then solved numerically using dynamic programming. Their solutions yield a set of strictly stabilising feedback gains, denoted as $\mathcal{K}$, and sequences of solution matrices to algebraic Riccati equations. Properties of these solution matrices are exploited in Chapter 5 for the design of dynamic gain selection procedures that aim to reduce the conservativeness brought by the use of Chebyshev's inequality. The results in this chapter have been submitted for publication and appear in [47].

Chapter 5: This chapter revisits a similar control problem to that of Chapter 3 and designs a model predictive control law incorporating a dynamic feedback gain. The feedback gain is selected from the set $\mathcal{K}$ discussed in Chapter 4. We provide two methods for online gain selection based on minimising upper bounds of predicted costs, which both incur only a tiny amount of extra computation. The same constraint-tightening technique as in Chapter 3 is used to ensure recursive feasibility. The closed loop system is guaranteed to satisfy the chance constraint and a quadratic stability condition. With dynamic feedback gain selection, we mitigate the conservativeness of Chebyshev's inequality. The closed loop cost is reduced and a larger set of feasible initial conditions is obtained. Some of the results in this chapter appear in [47].

Chapter 6: This chapter considers an application of stochastic MPC that has been developed in previous chapters in communication networks. We propose a model predictive control law that minimises a discounted cost subject to a discounted expectation constraint, where discrete-time linear systems with stochastic additive disturbances and noisy measurements transmitted over a lossy communication channel are considered. It is assumed that sensor data is lost with a known probability, and data losses are accounted for by expressing the predicted control policy as an affine function of future observations, which results in a convex optimisation problem. A similar constraint-tightening technique to that of Chapter 3 is used to ensure recursive feasibility of online optimisation problems and satisfaction of the expectation constraint. The discounted cost evaluated along trajectories of the closed loop system is shown to be bounded by the initial optimal predicted cost. We also provide conditions under which the averaged undiscounted closed loop cost accumulated over an infinite horizon is finite. The results in this chapter have been summarised in [31], which has been submitted for publication, and preliminary results have been published in [48].

Chapter 7: In this chapter, we conclude the thesis by summarising its main contributions and then discussing some directions for future work.

Appendices: Appendix A and Appendix B collect some standard definitions and results from convex optimisation and probability theory, respectively.

To conclude, we provide a list of publications, which this thesis is based on, as follows:

- "*Stochastic model predictive control with discounted probabilistic constraints*", S.Yan, P.J.Goulart and M.Cannon, Proceedings of European Control Conference, Limassol, Cyprus, pp. 1003–1008, 2018. ([44])

- "*Stochastic MPC with dynamic feedback gain selection and discounted probabilistic constraints*", S.Yan, P.J.Goulart and M.Cannon, submitted to IEEE Transactions on Automatic Control, 2019. ([47])

- "*Output feedback stochastic MPC with packet losses*", S.Yan, M.Cannon and P.J.Goulart, IFAC-PapersOnLine, vol. 53, no. 2, pp. 7105-7110, 2020. ([48])
- "*Stochastic output feedback MPC with intermittent observations*", S.Yan, M.Cannon and P.J.Goulart, submitted to Automatica, 2020. ([31])

2
Background

Contents

This chapter introduces basic definitions and results from model predictive control, chance constraints and Kalman filtering. It serves as a brief introduction to material that is related to or directly used in subsequent chapters.

2.1 Model Predictive Control

Model predictive control (MPC), which is also called receding horizon control or moving horizon optimal control, is an optimisation-based control method. It is initially applied in the petrochemical industry, and the idea of the moving horizon approach can be traced back to the 1960s [49]. After some fundamental developments in this area, including model predictive heuristic control [50], dynamic matrix control [51] and generalised predictive control [52, 53], MPC has since received increasing research interest. Due to its capabilities of controlling multivariable systems and systematic handling of constraints, and its convenient operation providing high performance without specialist intervening for long periods of time [6], MPC has been proved a success for process control problems and now has a wider range of applications, for example, in autonomous driving [54, 55], power electronics control [56, 57], wind turbine control [58, 59], networked control systems [60, 61], building management [62, 63] and financial engineering [22, 64]. The basic idea of MPC is to solve online an optimisation problem, which is typically referred to as an MPC optimisation problem and is essentially an optimal control problem, repeatedly at each time step using current information of the plant. The MPC optimisation problem yields a sequence of control inputs, and only the first element of this sequence is applied to the plant. However, the remaining elements may also be applied in some variants of MPC algorithms [65, 66]. Next, we will introduce optimal control and dynamic programming, which are closely related to the developments of MPC, before presenting a more detailed introduction of MPC.

2.1.1 Optimal control and dynamic programming

Given the scope of this thesis, we restrict the following introduction to optimal control and dynamic programming to discrete-time cases.

Discrete-time constrained infinite horizon optimal control problems can be expressed in the form

$$\begin{aligned} \min \ & \sum_{i=0}^{\infty} \ell(x_i, u_i) \\ \text{s.t.} \ & x_{i+1} = f(x_i, u_i) \quad \forall\, i \geq 0,\ x_0 = x^0, \\ & x_i \in \mathcal{X},\ u_i \in \mathcal{U} \quad \forall\, i \geq 0, \end{aligned}$$

where $x_i \in \mathbb{R}^n$, $u_i \in \mathbb{R}^m$ and $\ell(x_i, u_i)$ is the stage cost. The initial condition is given by x^0, and $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{U} \subseteq \mathbb{R}^m$ are constraint sets. The finite horizon version of this problem can be expressed in the form

$$\begin{aligned} \min \ & \sum_{i=0}^{N-1} \ell(x_i, u_i) + g(x_N, u_N) \\ \text{s.t.} \ & x_{i+1} = f(x_i, u_i) \quad i = 0, \ldots, N-1,\ x_0 = x^0, \\ & x_i \in \mathcal{X},\ u_i \in \mathcal{U} \quad i = 0, \ldots, N, \end{aligned}$$

where $g(x_N, u_N)$ is the terminal cost.

We next give an important class of optimal control problem, linear quadratic regulation (LQR) problems. The unconstrained finite horizon linear quadratic optimal control problem is given by

$$\begin{aligned} V_N(x_0) := \min \ & \sum_{i=0}^{N-1} (\|x_i\|_Q^2 + \|u_i\|_R^2) + \|x_N\|_Q^2 \\ \text{s.t.} \ & x_{i+1} = Ax_i + Bu_i \quad i = 0, \dots, N-1, \end{aligned} \tag{2.1}$$

where $x_i \in \mathbb{R}^n$, $u_i \in \mathbb{R}^m$, and A, B, Q and R are known matrices with appropriate dimensions. The term $\|x_N\|_Q^2$ denotes the terminal cost. The infinite horizon version of this problem is given by

$$\begin{aligned} V_\infty(x_0) := \min \ & \sum_{i=0}^{\infty} (\|x_i\|_Q^2 + \|u_i\|_R^2) \\ \text{s.t.} \ & x_{i+1} = Ax_i + Bu_i \quad \forall\ i \geq 0. \end{aligned} \tag{2.2}$$

Problems (2.1) and (2.2) can be solved by dynamic programming as illustrated below.

Dynamic programming is based on Bellman's principle of optimality: "an optimal policy has the property that, whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" [67]. From this statement, it follows that

$$V_i(x_{N-i}) = \min_{\substack{u_{N-i} \\ x_{N-i+1} = Ax_{N-i} + Bu_{N-i}}} \left\{ \|x_{N-i}\|_Q^2 + \|u_{N-i}\|_R^2 + V_{i-1}(x_{N-i+1}) \right\} \ \forall\ i = 1, \dots, N, \tag{2.3}$$

hold for problem (2.1), where $V_i(x_{N-i})$ are optimal cost-to-go functions and $V_0(x_N) = \|x_N\|_Q^2$; and

$$\begin{aligned} V_\infty(x_0) &= \min_{\substack{u_0 \\ x_1 = Ax_0 + Bu_0}} \left\{ \|x_0\|_Q^2 + \|u_0\|_R^2 + \min_{\substack{\{u_i\}_{i=1}^{\infty} \\ x_{i+1} = Ax_i + Bu_i\ \forall i \geq 1}} \sum_{i=1}^{\infty} (\|x_i\|_Q^2 + \|u_i\|_R^2) \right\} \\ &= \min_{\substack{u_0 \\ x_1 = Ax_0 + Bu_0}} \left\{ \|x_0\|_Q^2 + \|u_0\|_R^2 + V_\infty(x_1) \right\} \end{aligned} \tag{2.4}$$

holds for problem (2.2). Equations (2.3) and (2.4) are called Bellman equations.

Therefore, by solving the recursion (2.3) backwards in time and defining $P_0 := Q$ as the weighting matrix associated with $V_0(x_N)$, we have

$$\begin{aligned} u^*_{N-i} :=& K_{N-i}x_{N-i}, \ \forall\, i = 1, \ldots, N, \\ V_i(x_{N-i}) =& \|x_{N-i}\|^2_{P_i}, \ \forall\, i = 1, \ldots, N, \end{aligned}$$

where $K_{N-i} := -(R + B^\top P_{i-1}B)^{-1}B^\top P_{i-1}A$ and P_i satisfies the discrete-time Riccati equation

$$P_i = Q + A^\top P_{i-1}A - A^\top P_{i-1}B(R + B^\top P_{i-1}B)^{-1}B^\top P_{i-1}A.$$

Similarly, since $V_\infty(x) = \lim_{N\to\infty} V_N(x)$ holds for all x for problem (2.2), it can be solved approximately by recursive equations (2.3) as $N \to \infty$. Then, under the conditions of controllability and observability [68], we have $P_\infty := \lim_{i\to\infty} P_i$ $\forall\ P_0 \succeq 0$ and it satisfies the algebraic Riccati equation

$$P_\infty = Q + A^\top P_\infty A - A^\top P_\infty B(R + B^\top P_\infty B)^{-1}B^\top P_\infty A$$

and the Lyapunov matrix equation

$$P_\infty = Q + K_0^\top RK_0 + (A + BK_0)^\top P_\infty(A + BK_0),$$

where $K_0 = -(R + B^\top P_\infty B)^{-1}B^\top P_\infty A$ and it is strictly stabilising. From this result, it is clear that, under the conditions of controllability and observability, employing an infinite horizon cost function imparts stability [7, 68].

DP algorithms are able to provide the optimal feedback policies in very general settings, including problems with uncertainty in system dynamics, such as linear quadratic Gaussian (LQG) problems. A discrete-time linear quadratic Gaussian control problem can be expressed in the form

$$\begin{aligned} \min\ & \mathbb{E}\Big\{\sum_{i=0}^{N-1}(\|x_i\|^2_Q + \|u_i\|^2_R) + \|x_N\|^2_Q\Big\} \\ \text{s.t.}\ \ & x_{i+1} = Ax_i + Bu_i + w_i \quad i = 0, \ldots, N-1, \end{aligned} \tag{2.5}$$

where w_i are i.i.d. Gaussian random variables.

It can be verified using the DP algorithm that certainty equivalence [69] holds. Specifically, the optimal feedback strategy remains unaffected when the additive disturbances are present, and the optimal policies of problems (2.1) and (2.5) are the same.

In contrast to the examples presented above, analytical solutions are often not achievable by DP approaches, and numerical solutions are used instead. As a powerful solution technique, DP still has limited applications in practice for problems involving constraints or nonlinear dynamics due to the curse of dimensionality [70]. During the execution of DP algorithms for such problems, their computational burden grows exponentially with state dimensions and also with the horizon length, and it becomes unaffordable in practical applications. Therefore, this motivates the use of pre-determined but suboptimal control parametrisation in a non-deterministic MPC framework to reduce online computational complexity, which will be introduced in Section 2.1.2.

2.1.2 Classical, robust and stochastic model predictive control

Unlike conventional feedback control policies that are usually computed offline, such as PID control, MPC has better capabilities of handling constraints and provides feedback based on the current information of systems at every sampling time. After several decades, deterministic MPC has been well developed, while there are more challenges concerning robust and stochastic MPC problems. In the following, we provide a brief introduction to classical, robust and stochastic MPC problems. Given the scope of this thesis, the consideration here is restricted to linear discrete-time systems subject to linear constraints. More comprehensive and extensive reviews of MPC literature can be found, for example, in [7, 14, 71–73].

Classical MPC

Considering discrete-time linear time invariant systems

$$x_{k+1} = Ax_k + Bu_k, \tag{2.6}$$

and a linear constraint on system states and control inputs

$$Fx_k + Gu_k \leq \mathbf{1}, \ \forall \ k \geq 0, \tag{2.7}$$

a deterministic MPC problem with a quadratic cost function can be expressed in the form

$$\begin{aligned} \min_{u_{0|k},\ldots,u_{N-1|k}} \quad & \sum_{i=0}^{N-1} (\|x_{i|k}\|_Q^2 + \|u_{i|k}\|_R^2) + \|x_{N|k}\|_P^2 \\ \text{s.t.} \quad & Fx_{i|k} + Gu_{i|k} \leq \mathbf{1}, \ i = 0, \ldots, N-1, \\ & x_{N|k} \in \mathcal{X}_T, \\ & x_{i+1|k} = Ax_{i|k} + Bu_{i|k}, \ i = 0, \ldots, N-1, \ x_{0|k} = x_k. \end{aligned} \tag{$\mathcal{P}^N$}$$

Here $x_{i|k}$ and $u_{i|k}$ are i-step-ahead predictions of x_{k+i} and u_{k+i} at time k, and Q, R are weighting matrices for the state and control input. Also, it is assumed that the pairs (A, B) and $(A, Q^{1/2})$ satisfy the conditions of controllability and observability or a slightly weakened condition (stabilisability and detectability), respectively. The notation $\mathcal{X}_T$ denotes a terminal constraint set, and N is a prediction horizon. The terminal matrix P is appropriately chosen. Its choice is closely related to closed loop stability, which is normally ensured in the MPC framework by employing the optimal value function of problem $(\mathcal{P}^N)$ as a Lyapunov function. The terminal cost $\|x_{N|k}\|_P^2$ serves as an upper bound on the cost-to-go function for stages from time N to infinity, and the finite horizon cost function then approximates the corresponding infinite horizon cost function, where the optimisation problem is infinite dimensional and is in general intractable to solve.

The predicted input sequence $\{u_{0|k}, \ldots, u_{N-1|k}, u_{N|k}, \ldots\}$ is parameterised, in many existing MPC work, by the dual mode prediction paradigm [74] that is given by

$$u_{i|k} = Kx_{i|k} + c_{i|k} \quad i = 0, \ldots, N-1, \tag{2.8a}$$

$$u_{i|k} = Kx_{i|k} \quad i = N, N+1, \ldots, \tag{2.8b}$$

where K is some strictly stabilising feedback gain and $\{c_{i|k}\}_{i=0}^{N-1}$, which is referred to as the perturbation sequence, are free variables. Typically, K is chosen as the unconstrained LQ optimal feedback gain. Therefore, under this predicted control law, the predicted state sequence is given by

$$x_{i+1|k} = \Phi x_{i|k} + Bc_{i|k}, \quad i = 0, \dots, N-1, \tag{2.9a}$$

$$x_{i+1|k} = \Phi x_{i|k} \quad i = N, N+1, \dots \tag{2.9b}$$

where $\Phi := A + BK$, and K, P must satisfy the Lyapunov matrix equation

$$P = Q + K^\top RK + (A + BK)^\top P(A + BK).$$

We denote $\mathbf{c}_k^* := \{c_{0|k}^*, \dots, c_{N-1|k}^*\}$ as the minimiser to problem $(\mathcal{P}^N)$, $\mathbf{x}_k^* := \{x_{0|k}^*, \dots, x_{N-1|k}^*, x_{N|k}^*, \dots\}$ and $\mathbf{u}_k^* := \{u_{0|k}^*, \dots, u_{N-1|k}^*, Kx_{N|k}^*, Kx_{N+1|k}^*, \dots\}$ as the corresponding predicted state and input sequences. Then, the closed loop system evolves as

$$x_{k+1} = \Phi x_k + Bc_{0|k}^*, \ \forall\ k \geq 0.$$

Terminal constraint sets are carefully designed as positively invariant sets such that the linear constraints imposed over the finite horizon in problem $(\mathcal{P}^N)$, together with the terminal constraint, is equivalent to linear constraints imposed over an infinite horizon.

Definition 2.1 (Positively invariant set)**.** *A set $\mathcal{X} \subseteq \mathbb{R}^n$ is invariant under the dynamics* (2.9b) *and with respect to the constraint* (2.7) *if*

$$\Phi x \in \mathcal{X} \text{ and } (F + GK)x \leq \mathbf{1}, \ \forall\ x \in \mathcal{X}.$$

In problem $(\mathcal{P}^N)$, $\mathcal{X}_T$ can be designed as the maximal positively invariant (MPI) set, which is the union of all sets that are also invariant under the dynamics (2.9b) and constraint (2.7), to increase the feasible set of initial conditions.

Theorem 2.2 ([75])**.** *The MPI set $\mathcal{X}^{MPI}$ under the dynamics* (2.9b) *and with respect to the constraint* (2.7) *can be obtained by finding the smallest integer $m > 0$ such that*

$$(F + GK)\Phi^{m+1}x \leq \mathbf{1}, \ \forall x \in \mathcal{X}^{MPI} := \{x \in \mathbb{R}^n | (F + GK)\Phi^i x \leq \mathbf{1}, \ i = 0, \dots, m\}.$$

A finite m necessarily exists if Φ is strictly stabilising and the pair $(\Phi, F + GK)$ is observable.

The procedure to find such an m is to solve a set of linear programs, which can be executed offline. In addition, in most MPC literature, terminal constraints are imposed also to ensure recursive feasibility of MPC optimisation problems, that is, there always exists a feasible solution to the MPC optimisation problem at each time step given initial feasibility. This feasible solution at time $k+1$ is $\{c^*_{1|k}, \ldots, c^*_{N-1|k}, 0\}$ or $\{u^*_{1|k}, \ldots, u^*_{N-1|k}, Kx^*_{N|k}\}$. These sequences are referred to as tail sequences and are the optimal solution at time k that is shifted forward by one time step.

The prediction horizon, as a tuning parameter, often affects considerably the performance of MPC algorithms. For nominal linear MPC problems, there exists a finite horizon N_∞ [76] such that if the prediction horizon, N, in problem $(\mathcal{P}^N)$ is greater or equal to N_∞, $(\mathcal{P}^N)$ is equivalent to the following infinite horizon optimal control problem $(\mathcal{P}^\infty)$

$$\begin{aligned} \min_{u_{0|k}, u_{1|k}, \cdots} \quad & \sum_{i=0}^{\infty} (\|x_{i|k}\|_Q^2 + \|u_{i|k}\|_R^2) \qquad (\mathcal{P}^\infty) \\ \text{s.t.} \quad & Fx_{i|k} + Gu_{i|k} \leq \mathbf{1}, \; i = 0, 1, \ldots, \\ & x_{i+1|k} = Ax_{i|k} + Bu_{i|k}, \; i = 0, 1, \ldots, \; x_{0|k} = x_k. \end{aligned}$$

However, there is generally no systematic way to determine a prediction horizon, especially when uncertainty is present. A choice is typically made to balance the performance against computational efficiency.

Robust MPC

It is rare that system models are perfect and match exactly the actual behaviours of plants. Uncertainty arises in many cases. For example, noise appears in the measurements of states or outputs, and system identification provides imperfect knowledge about model parameters. Under uncertainty, it is a fundamental problem in controller design to maintain desirable performance, namely robustness. Deterministic MPC possesses inherent robustness under certain circumstances. For instance, [77] shows an MPC algorithm designed for the nominal system ensures that its actual system with additive disturbances and only input constraints is input-to-state stable and closed loop states converge to a terminal set if the norm of the disturbance is bounded. Nevertheless, it is essential to design MPC algorithms that are intrinsically robust with respect to uncertainty, and they will have wider ranges of applications. Uncertainty in system models can be generally categorised into additive and multiplicative disturbances.

Robust MPC with additive disturbances An uncertain linear discrete-time system with additive disturbances can be described by

$$x_{k+1} = Ax_k + Bu_k + w_k,$$

where it is often assumed that w_k lies in a convex and compact set $\mathcal{W}$. Robust MPC algorithms have to satisfy constraints for all possible realisations of the additive disturbance. A typical approach to ensure this is to decompose the states and then tighten the constraints accordingly. By employing the predicted control law (2.8a)-(2.8b), which is referred to as an open loop control law since its feedback gain is fixed instead of a free variable, we have

$$x_{i+1|k} = \Phi x_{i|k} + Bc_{i|k} + w_{i|k}.$$

The states can be decomposed into a nominal part and an uncertain part as $x_{i+1|k} = z_{i+1|k} + e_{i+1|k}$, with

$$z_{i+1|k} = \Phi z_{i|k} + Bc_{i|k},\ i = 0, 1, \dots, \tag{2.10a}$$

$$e_{i+1|k} = \Phi e_{i|k} + w_{i|k},\ i = 0, 1, \dots, \tag{2.10b}$$

and $z_{0|k} = x_{0|k}$, $e_{0|k} = 0$. Considering the same constraint (2.7), we now write it as

$$(F + GK)z_{i|k} + Gc_{i|k} \leq \mathbf{1} - (F + GK)e_{i|k}.$$

To satisfy this constraint for all possible realisations of w_k in $\mathcal{W}$, a tightened constraint is imposed instead, which is given by

$$(F + GK)z_{i|k} + Gc_{i|k} \leq \mathbf{1} - \varepsilon_i,$$

where $\varepsilon_0 := 0$ and

$$\varepsilon_i := \max_{\{w_{0|k},\dots,w_{i-1|k}\}\in\mathcal{W}\times\cdots\times\mathcal{W}} (F + GK)e_{i|k}. \tag{2.11}$$

In this approach, predicted states can be seen as constrained in a sequence of tubes $\{\mathcal{X}_{0|k}, \mathcal{X}_{1|k}, \ldots\}$, characterised by its centre, the nominal state $z_{i|k}$, and its cross-section. These tubes are defined as $\mathcal{X}_{i|k} := \{z_{i|k}\} \oplus \mathcal{E}_{i|k}$, where it follows from (2.10b) that

$$\mathcal{E}_{i+1|k} = \Phi\mathcal{E}_{i|k} \oplus \mathcal{W}$$

and $\mathcal{E}_{0|k} = \{0\}$. For the case in which it is possible to evaluate $\mathcal{X}_{\infty|k} := \bigoplus_{j=0}^{\infty} \Phi^j \mathcal{W}$, which is referred to as the minimal robust positively invariant (mRPI) set, we can compute ε_∞ so that the constraints imposed over an infinite horizon are equivalent to a finite number of constraints. However, it is in general not tractable, and approximations for $\mathcal{X}_{\infty|k}$ and upper bounds for ε_∞ are constructed instead. Some methods for approximating the mRPI sets are discussed in [78, 79].

In contrast to the basic tube MPC method [80] introduced above, alternative tube-based approaches are proposed to bring more degrees of freedom into the tube parameterisation, such as rigid tubes [8], homothetic tubes [10], and tubes with more flexible cross-section representations [11]. Similar tube-based approaches have also been developed for multiplicative disturbances.

In robust MPC problems, due to the presence of uncertainty, the predicted system dynamics will be different from the actual dynamics. It is no longer the case that the open loop control law is equivalent to a feedback control policy [9]. Therefore, it is ideal to optimise over a general feedback policy to achieve optimal performance, which leads to the developments of min-max robust MPC [81] that can be solved by DP, but that would be computationally demanding. Alternatively, a pre-specified feedback control structure can be used, balancing computation against performance. One example is the affine state feedback policy

$$u_{i|k} = \sum_{j=0}^{i} K_{j,i|k} x_{j|k} + c_{i|k}, \quad i = 0, \ldots, N-1, \tag{2.12}$$

where $K_{j,i|k}$ and $c_{i|k}$ are free variables and $K_{j,i|k}$ form a block lower triangular matrix

$$\mathbf{K}_k := \begin{bmatrix} K_{0,0|k} & 0 & \cdots & 0 \\ K_{0,1|k} & K_{1,1|k} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ K_{0,N-1|k} & K_{1,N-1|k} & \cdots & K_{N-1,N-1|k} \end{bmatrix}.$$

This parameterisation results in nonlinear dependence of predicted states and inputs on the decision variables and a nonconvex optimisation problem. However, it can be reformed as linear functions of decision variables using a Youla parameter as shown in [82]. A similar result is given in [83], which proves the equivalence between (2.12) and the affine disturbance feedback policy

$$u_{i|k} = \sum_{j=0}^{i-1} L_{j,i|k} w_{j|k} + v_{i|k}, \quad i = 0, \ldots, N-1, \tag{2.13}$$

where $L_{j,i|k}$ and $v_{i|k}$ are free variables and $L_{j,i|k}$ form a strictly block lower triangular matrix

$$\mathbf{L}_k := \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ L_{0,1|k} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ L_{0,N-1|k} & \cdots & L_{N-2,N-1|k} & 0 \end{bmatrix}.$$

Robust MPC with multiplicative disturbances Uncertain linear discrete-time systems with multiplicative disturbances can be described by

$$x_{k+1} = A_k x_k + B_k u_k. \tag{2.14}$$

It is assumed that

$$(A_k, B_k) = \sum_{j=1}^{m} q_k^{(j)} (A^{(j)}, B^{(j)}),$$

where $(A^{(j)}, B^{(j)})$ are pairs of known matrices and $q_k^{(j)}$ are scalar random variables satisfying

$$\sum_{j=1}^{m} q_k^{(j)} = 1, \quad q_k^{(j)} \geq 0 \;\; \forall j \in \{1, \ldots, m\}.$$

Equivalently, (A_k, B_k), for all k, lies in the convex hull, $\mathcal{H}$, of vertices $(A^{(1)}, B^{(1)})$, …, $(A^{(m)}, B^{(m)})$. Equation (2.14) is referred to as a linear difference inclusion (LDI) model. This model enjoys the property that the convex hull of trajectories generated by each linear model associated with a vertex encloses the trajectory of the actual system [84], while it has limitations in capturing the behaviours of systems of which system matrices depend nonlinearly on the uncertain parameter.

An important early development in robust MPC dealing with multiplicative uncertainty is the use of LMI methods [85]. Employing a predicted linear state feedback law $u_{i|k} = K_k x_{i|k}$, where matrix K_k is a free variable, yields that

$$x_{i+1|k} = (A_{i|k} + B_{i|k} K_k) x_{i|k},\ (A_{i|k}, B_{i|k}) \in \mathcal{H}. \tag{2.15}$$

For this system, we can construct an ellipsoidal set $\mathcal{S}$ that is robust positively invariant, where $\mathcal{S} := \{x | \|x\|_P^2 \leq 1\}$. The set $\mathcal{S}$ is invariant for system (2.15) if

$$P - (A + BK_k)^\top P (A + BK_k) \succeq 0, \quad \forall (A, B) \in \mathcal{H}.$$

This condition is equivalent to

$$\begin{bmatrix} P & (A^{(j)} + B^{(j)} K_k)^\top \\ A^{(j)} + B^{(j)} K_k & P^{-1} \end{bmatrix} \succeq 0, \quad j = 1, \ldots, m$$

by Schur complements and the convexity of LMIs. Therefore, the linear constraint (2.7) can be ensured as shown in [72, Theorem 2.9] by the set $\mathcal{S}$ that is invariant for both dynamics (2.15) and constraint (2.7) if P satisfies additional conditions

$$\begin{bmatrix} Z & F + GK_k \\ (F + GK_k)^\top & P \end{bmatrix} \succeq 0, \quad e_i^\top Z e_i \leq 1,\ i = 1, \ldots, n_r,\ \exists Z = Z^\top,$$

where n_r is the number of rows in matrix F and e_i is the i-th column of a n_r-by-n_r identity matrix. This approach is computationally convenient since the quadratic bounds associated with the invariance conditions are translated into LMI conditions and the online MPC optimisation problem becomes a semidefinite program, despite the fact that quadratic bounds can be potentially conservative. More degrees of freedom can be introduced into the predicted controller parameterisation instead of optimising over only a single linear feedback gain. Some further work based on LMI methods aiming to reduce these restrictions can be found, for example, in [86–88].

Robust MPC algorithms ensure constraint satisfaction under all possible realisations of disturbances. It often requires a great effort to account for the worst case scenario that is a very rare eventuality. Deemphasising extreme disturbances would possibly enhance controller performance, and in some applications, it is acceptable if constraints are violated for a limited time period. Also, robust

MPC does not employ statistical information of disturbances, e.g. historical data, moments and probability distributions, if available. These drive the development of stochastic MPC, where the assumption in robust MPC can sometimes be removed that disturbances are bounded.

Stochastic MPC

There has been significant research interest in stochastic MPC during the past decade. Its key feature is the presence of soft constraints that are allowed to be violated up to a given limit, in contrast to hard constraints in the robust MPC framework. Employing probabilistic descriptions of uncertainty in the stochastic MPC framework reduces the conservativeness of controller design and allows for explicit regulation of probability distributions of states and control inputs [14]. These advantages also bring challenges. Chance constraints in many cases entail nonconvex feasible sets and it is generally difficult to evaluate them directly. These problems lead to the development of a distinct research topic, chance constrained optimisation [89], and various approaches to provide tractable approximations of chance constraints. The key challenge is how a method providing nonconservative approximations of chance constraints can be incorporated into the MPC framework to ensure desirable closed loop properties, for example, recursive feasibility, constraint satisfaction and stability.

Stochastic MPC algorithms that do not rely on disturbance bounds can consider unbounded uncertainty, while bringing additional challenges. For example, if additive disturbances have infinite support, it is difficult to provide guarantees of recursive feasibility, which typically depend on the use of tail sequences. Such guarantees are achieved in the existing stochastic MPC literature by using a backup initial condition for the online MPC optimisation problem. More specifically, when infeasibility occurs, $\bar{x}_{1|k}$, the nominal value of x_{k+1} predicted at time k, is chosen as the initial condition for the optimal control problem to be solved at time $k+1$ instead of the actual value of x_{k+1} [15–17]. This design can block feedback from the actual state measurement and therefore is likely to provide inadequate performance. Alternatively, we show in Chapter 3 that, with unbounded additive disturbances, recursive feasibility of online MPC optimisation problems is guaranteed by considering a discounted chance constraint, while updating the initial conditions of the online optimal control problems with actual states at each time step.

Stability is of paramount importance in controller design. In the stochastic MPC framework, different stability conditions are often used from those in robust MPC problems. When the cost function is subject to stochastic uncertainty, it is generally not expected that it will be monotonically non-increasing so as to allow a Lyapunov stability analysis. Also, it is not possible to construct a finite bound on the infinite sum of the norm of states using l_2 stability conditions [72], if disturbances are persistent (i.e. unless they decay sufficiently rapidly over time). Typically, a stability metric often used in the stochastic setting is the expected average cost over an infinite horizon [90]

$$\lim_{T\to\infty} \frac{1}{T} \mathbb{E}\Big\{\sum_{k=0}^{T-1} \ell(x_k, u_k)\Big\} \leq b,$$

where ℓ is the stage cost and a function of states and control inputs, the expectation is conditioned on initial conditions and b is some finite bound.

Chance constraints There are many different types of chance constraint. Some are defined in terms of probabilities, such as $\mathbb{P}\{g(x,\delta) \leq 0\} \geq 1-\epsilon$ where δ is a random variable and ϵ is a given violation threshold, or in terms of expectations, such as $\mathbb{E}\{g(x,\delta)\} \leq 0$. These constraints are often imposed in a pointwise-in-time fashion in control problems, such as $\mathbb{P}\{x_k \in \mathcal{X}\} \geq 1-\epsilon \ \forall\ k \geq 0$, where violations of the condition $x_k \in \mathcal{X}$ at every individual time step are equally taken into account. A different formulation is also widely considered, for example, in [38–41], where chance constraints are imposed in an average-in-time fashion as

$$\mathbb{E}\{\frac{1}{T}\sum_{k=0}^{T-1} \mathbb{I}_{\mathcal{X}}(x_k)\} \geq 1-\epsilon$$

where $\mathbb{I}_{\mathcal{X}}$ is an indicator function, or equivalently as

$$\frac{1}{T}\sum_{k=0}^{T-1} \mathbb{P}\{x_k \in \mathcal{X}\} \geq 1-\epsilon.$$

On the contrary, average-in-time chance constraints focus more on aggregate performance over a finite horizon, which lead to a less conservative controller design since satisfying pointwise-in-time chance constraints at each single time step is sufficient for the satisfaction of an average-in-time chance constraint. One approach to achieve less conservativeness is to adjust control strategies online by observing past constraint violations [38, 42]. Also, average-in-time chance constraints are more natural formulations of chance constraints in some practical applications, such as economics [43]. In these problems, discounting is naturally present and represents an interest rate. Discount factors enable the consideration of average-in-time chance constraints defined over longer and infinite horizons when disturbances are unbounded [44]. In addition, average-in-time chance constraints can be imposed specifically with discounting to suit design purposes. Examples include sustainable development [45] and public policy analysis [46], where the main argument is that near-term benefits are more important than those that would be gained in the future and current actions should not imperil possible future development [22]. Another two important classes of chance constraint are distributionally robust chance constraints [91] and its variant distributionally robust conditional value-at-risk constraints [92]. The former is defined as

$$\mathbb{P}\{x \in \mathcal{X}\} \geq 1 - \epsilon \quad \forall\, \mathbb{P} \in \mathcal{P},$$

where $\mathcal{P}$ is a set of probability distributions. This constraint can be seen as an intermediate version between robust and chance constraints, enforcing the satisfaction of a standard chance constraint for all probability distributions in the set $\mathcal{P}$. The advantage of this form is that it only requires limited knowledge about the true probability distribution of the disturbance. The latter is concerned with cases where $x \notin \mathcal{X}$ and it aims to bound expected constraint violations in the rest ϵ percentages [93].

Constraint handling Approximation methods for chance constraints can be broadly categorised into exact evaluation, deterministic approximation and sampling methods. When probability distributions are known, chance constraints can be evaluated exactly. [32] considers a controller synthesis problem concerning the stationary operation of discrete-time linear time invariant systems with Gaussian additive disturbances. Since the disturbance is Gaussian, the steady state x will have a unique stationary covariance $X \succ 0$ under stabilising control laws. Then the chance constraint $\mathbb{P}\{g^\top x \leq h\} \geq 1 - \epsilon_x$ on the stationary state x is equivalent to a linear constraint on X, given by $g^\top X g \leq \frac{h^2}{\alpha^2}$ where $\alpha := F_{nrm}^{-1}(1-\epsilon_x)$ and F_{nrm}^{-1} is the inverse cumulative distribution function of the standard normal distribution. However, the exact calculation of the distribution function of the sum of random variables requires the evaluation of a multivariate convolution integral, and it is generally impractical even if the probability distribution functions of each random variable are known [33].

Deterministic approximation methods replace chance constraints with deterministic constraints. In [33], with the assumption that elements of the disturbance are independent with zero mean and their finitely supported distribution functions are known, chance constraints on the state are turned into tightened deterministic constraints on nominal states, where tightening parameters are approximately calculated by discretising distributions of elements of the disturbance and then using univariate convolutions. In [34], tube-based methods are extended to the stochastic setting. By decomposing states into a nominal part and an uncertain part, ellipsoidal probabilistic tubes are constructed to contain the propagation of additive disturbances, and they are computed offline to reduce the online computational burden. Together with assumedly known probability distributions of the parameter defining the uncertainty set, these tubes are employed to determine the tightening parameters in the deterministic constraints on nominal states, which ensures the chance constraints on actual states. A similar approach is used in [16]. Under the assumption that the probability distribution of the possibly unbounded additive disturbance is unimodal, a constraint-tightening technique is utilised in combination with probabilistic reachable sets. Another typical approximation method is through the use of probabilistic inequalities, such as Chebyshev's inequality [35] and Cantelli's inequality [15]. They are capable of handling arbitrary probability distributions and result in tractable convex constraints, while only requiring the knowledge of the first two moments. Nevertheless, these inequalities are conservative except for specific probability distributions.

Sampling methods approximate chance constraints by using independent samples of the uncertainty and provide probabilistic guarantees of constraint satisfaction. A major contribution is made in [94], where confidence bounds for the scenario approach are proposed. The scenario approach replaces a chance constraint with a finite number of deterministic constraints using i.i.d. samples collected, and ensures that the optimal solution to the resulting problem is feasible to the original chance constrained problem with certain probability levels. Also, the scenario approach can handle a wide range of chance constraints, only requiring that the constraint function be a convex function of the decision variable for any realisations of the uncertainty. Thus, it has been widely used in the stochastic MPC literature, while they often lack guarantees of recursive feasibility. In the context of using sampling methods, a combination with robust bounds is necessary to achieve this guarantee [95, 96]. A variant of the scenario approach is proposed in [97]. It discards a number of the samples collected and observed to reduce conservativeness of optimal solutions to resulting scenario programs. To achieve high confidence levels that are sufficiently close to 1, the scenario approach often requires a large number of samples, which results in optimisation problems that are computationally demanding to solve and hence limits the applicability of the scenario approach. Therefore, many efforts [e.g. 98, 99] have been made to reduce the sample complexity. The method proposed in [41], which is based on the scenario approach, requires a considerably lower number of samples for establishing probabilistic guarantees of feasibility. Other sampling methods include sample average approximation [100] and sample approximation [101]. The latter is similar to the scenario approach and provides probabilistic guarantees of feasibility for the uncertain program, while it does not require the convexity assumption of constraint functions in decision variables.

2.2 Kalman Filtering with Intermittent Observations

2.2.1 Kalman filter

Consider a linear dynamical system

$$x_{k+1} = Ax_k + w_k,$$
$$y_k = Cx_k + v_k$$

where w_k is the additive disturbance and v_k is the measurement noise. The random sequence $\{w_k\}_{k=0}^{\infty}$ and $\{v_k\}_{k=0}^{\infty}$ are assumed to be i.i.d. with

$$\mathbb{E}\{w_k\} = 0, \qquad \mathbb{E}\{w_k w_k^\top\} =: \Sigma_w \succeq 0,$$
$$\mathbb{E}\{v_k\} = 0, \qquad \mathbb{E}\{v_k v_k^\top\} =: \Sigma_v \succ 0.$$

Assumption 2.1. *The pair (A, Σ_w) is controllable and the pair (A, C) is observable.*

For this system, the Kalman filter is given by

$$\hat{x}_{k+1} = A\hat{x}_k + K_k(y_k - C\hat{x}_k),$$
$$K_k := AP_kC^\top(CP_kC^\top + \Sigma_v)^{-1},$$

where $\hat{x}_k$ is the (prior) state estimate defined as $\hat{x}_k := \mathbb{E}\{x_k|y_0, \ldots, y_{k-1}\}$ and $\hat{x}_0 := \mathbb{E}\{x_0\}$ is assumed known. The matrix K_k denotes the Kalman filter gain. The matrix P_k denotes the (prior) estimation error covariance defined as $P_k := \mathbb{E}\{(x_k - \hat{x}_k)(x_k - \hat{x}_k)^\top|y_0, \ldots, y_{k-1}\}$, which satisfies the Riccati difference equation

$$P_{k+1} = AP_kA^\top - AP_kC^\top(CP_kC^\top + \Sigma_v)^{-1}CP_kA^\top + \Sigma_w, \tag{2.16}$$

and $P_0 := \mathbb{E}\{(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^\top\}$ is assumed known. By Assumption 2.1, P_k converges to a unique fixed point P satisfying the algebraic Riccati equation

$$P = APA^\top - APC^\top(CPC^\top + \Sigma_v)^{-1}CPA^\top + \Sigma_w \tag{2.17}$$

for any $P_0 \succeq 0$, and $A - KC$ is stabilising where $K := APC^\top(CPC^\top + \Sigma_v)^{-1}$. We next provide two lemmas about the monotonicity properties of the sequence $\{P_k\}_{k=0}^{\infty}$.

Lemma 2.3 ([102, Lemma 2]). *If there exists some j such that $P_j \succeq P_{j+1} \succeq 0$, then $P_{j+k-1} \succeq P_{j+k}$ $\forall k \geq 1$.*

Lemma 2.4 ([102, Lemma 1]). *If $P_0 \succeq P$, then $P_k \succeq P$ $\forall k \geq 1$.*

2.2.2 Packet loss and random Riccati equation

In sensor networks, communication delays and data losses often occur and substantial delays can also be seen as data losses from a control perspective [103]. The arrival of observed data can be modelled as a Bernoulli process $\{\gamma_k\}_{k=0}^{\infty}$ that is assumed to be i.i.d. with $\mathbb{P}\{\gamma_k = 0\} = 1 - \lambda$ and $\mathbb{P}\{\gamma_k = 1\} =: \lambda$, and the sensor measurement received is given by $z_k = \gamma_k y_k$. Therefore, the Kalman filter with intermittent observations is given by

$$\begin{aligned} \hat{x}_{k+1} =& A\hat{x}_k + \gamma_k K_k (y_k - C\hat{x}_k), \\ K_k :=& AP_k C^\top (CP_k C^\top + \Sigma_v)^{-1}, \end{aligned}$$

where P_k evolves according to

$$P_{k+1} = AP_kA^\top - \gamma_k AP_kC^\top(CP_kC^\top + \Sigma_v)^{-1}CP_kA^\top + \Sigma_w. \tag{2.18}$$

This equation is referred to as a random Riccati difference equation.

Theorem 2.5 (Existence of a critical value of arrival probabilities, [103, Theorem 2]). *If Assumption 2.1 holds and matrix A is unstable, there exists a critical value, $\lambda_c \in [0, 1)$, of observation arrival probabilities such that the following is true:*

- *If $0 \leq \lambda \leq \lambda_c$, $\lim_{k\to\infty} \mathbb{E}\{P_k\} = +\infty$ $\exists P_0 \succeq 0$;*
- *If $\lambda_c < \lambda \leq 1$, $\mathbb{E}\{P_k\}$ remains upper bounded for all k and for any initial condition $P_0 \succeq 0$ and this upper bound depends on P_0,*

where P_k is generated by (2.18)*.*

Theorem 2.6 (Bounds on the critical value of arrival probabilities, [103, Theorem 3])**.** *Let*

$$\underline{\lambda} := \arg\min\{\lambda | \exists Q, Q = (1-\lambda)AQA^\top + \Sigma_w\} = 1 - \frac{1}{\alpha^2},$$
$$\bar{\lambda} := \arg\min\{\lambda | \exists (K,X), X \succ (1-\lambda)AXA^\top + \lambda(A+KC)X(A+KC)^\top + \Sigma_w + \lambda K\Sigma_v K^\top\},$$

where $\alpha = \max_i |\sigma_i|$ *and* σ_i *are the eigenvalues of* A*. Then*

$$\underline{\lambda} \leq \lambda_c \leq \bar{\lambda}.$$

Theorem 2.7 (Mean-square stability, [104])**.** *For the system* $\xi_{i+1} = \Psi(\gamma)\xi_i$ *where the state matrix* Ψ *is determined by some random parameter* γ*, it is mean-square stable if there exists a matrix* $\Xi = \Xi^\top \succ 0$ *such that*

$$\Xi - \mathbb{E}\{\Psi(\gamma)\Xi\Psi^\top(\gamma)\} \succ 0.$$

3
Discounted Probabilistic Constraints

Contents

This chapter considers linear discrete-time systems with additive disturbances and designs a model predictive control (MPC) law to minimise a quadratic cost function subject to a chance constraint. The chance constraint is defined as a discounted sum of violation probabilities on an infinite horizon. By penalising violation probabilities close to the initial time and assigning violation probabilities in the far future with vanishingly small weights, this form of constraint enables the feasibility of the online optimisation to be guaranteed without an assumption of boundedness of the disturbance. A computationally convenient MPC optimisation problem is formulated using Chebyshev's inequality, and we introduce an online constraint-tightening technique to ensure recursive feasibility based on knowledge of a suboptimal solution. The closed loop system is guaranteed to satisfy the chance constraint, and we show that it satisfies a quadratic stability condition.

3.1 Introduction

Robust control design methods for systems with unknown disturbances must take into account the worst case disturbance bounds in order to guarantee the satisfaction of hard constraints on system states and control inputs [8, 85, 105]. However, for problems with stochastic disturbances and constraints that are allowed to be violated up to a specified probability, namely chance constraints, worst case control strategies can be unnecessarily conservative. This prompted the development of stochastic MPC, which addresses optimal control problems for systems with chance constraints by making use of information on the distribution of model uncertainty [33, 106]. Although capable of handling chance constraints, existing stochastic MPC algorithms that ensure constraint satisfaction in closed loop operation typically rely on knowledge of worst case disturbance bounds to obtain such guarantees [72]. For the algorithms proposed in [39, 40, 107] for example, which simultaneously ensure closed loop constraint satisfaction and recursive feasibility of the online MPC optimisation, the degree of conservativeness increases as the disturbance bounds become more conservative.

There are many different methods for approximating chance constraints, e.g. analytical approximation and sampling methods. The former provides tractable deterministic optimisation problems, while the latter generally results in randomised methods. In [15], Cantelli's inequality is used to turn the chance constraint on states into linear constraints. In [35] and [32], Chebyshev's inequality is used to reformulate chance constraints as a quadratic and a linear matrix inequality constraint, respectively. These two inequalities can handle a wide range of probability distributions and only require information on the first and second moments of additive disturbance distributions. However, the resulting constraints only provide tight bounds for specific probability distributions and otherwise are conservative. A scenario approach is used in [41] to impose time-average expectation constraints on system states. Although constraint satisfaction is demonstrated for the closed loop system, the recursive feasibility of online optimisation problems is assumed but not ensured. More generally, sample based methods are unable to ensure the recurrence of feasibility of receding horizon optimisation problems unless they are combined with robust bounds on model uncertainty. For example, a recursively feasible MPC strategy is proposed in [96], with chance constraints imposed using a scenario approach at the first prediction time step and replaced by robust constraints at later times, resulting in a conservative control law.

For problems involving stochastic uncertainty, the optimal expected value of the predicted cost is typically used to perform a Lyapunov analysis of closed loop stability. A performance metric widely used in this setting is the long-run expected average cost [90]. Although the vast majority of stability results are derived by imposing terminal constraints, there are a number of alternative approaches (e.g. [108–110]).

Discounting is present in many stochastic control settings (e.g. [18–21, 39]), as well as in reinforcement learning [23], financial engineering (e.g. [24–27]) and ecosystem management (e.g. [28, 29]). Discount factors in optimal control problems allow performance in the near future to be prioritised over long-term behaviours. This shift of emphasis is vital for ensuring recursive feasibility of chance constrained control problems involving possibly unbounded disturbances. In Dynamic Programming (e.g. [18, 19]), discounting is commonly employed to ensure that infinite horizon problems with possibly unbounded cost per stage are well-defined. In economics, discounting allows aggregation of current and potential future costs and revenues. For example, [28] shows that varying discount factors on future revenue can affect harvesting policies.

In this chapter, we ensure both closed loop satisfaction of chance constraints and recursive feasibility but do not rely on disturbance bounds, instead requiring knowledge of only the first and second moments of the disturbance input. This is achieved by formulating the chance constraint as the sum over an infinite horizon of discounted violation probabilities, and implementing the resulting constraints using Chebyshev's inequality. We describe an online constraint-tightening approach that guarantees the feasibility of the MPC optimisation, and, by considering the closed loop dynamics of the tightening parameters, we show that the closed loop system satisfies the discounted chance constraint as initially specified. The main features and contributions of this chapter are summarised as follows:

- We use a discount factor to ensure that the chance constraint is well-defined and to prioritise near-future system behaviours over steady state performance.
- A constraint-tightening technique is proposed to ensure recursive feasibility of online MPC optimisations and constraint satisfaction in closed loop without requiring disturbances to be bounded.
- The closed loop system satisfies a quadratic stability condition without the need for terminal constraints.

This chapter is organised as follows. The control problem is described and reformulated with a finite prediction horizon in Section 3.2. Section 3.3 proposes an online constraint-tightening method for guaranteeing recursive feasibility. Section 3.4 summarises the proposed MPC algorithm and derives bounds on closed loop performance. In Section 3.5, the closed loop behaviour of the tightening parameters is analysed and constraint satisfaction is proved. Section 3.6 gives a numerical example illustrating the results obtained and conclusions are drawn in Section 3.7.

3.2 Problem Description

Consider an uncertain linear system with model

$$x_{k+1} = Ax_k + Bu_k + \omega_k, \tag{3.1}$$

where $x_k \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$ are the system state and the control input respectively. The unknown disturbance input $\omega_k \in \Delta \subseteq \mathbb{R}^{n_x}$ is independently and identically distributed (i.i.d.) with known first and second moments:

$$\mathbb{E}\{\omega_k\} = 0, \quad \mathbb{E}\{\omega_k \omega_k^\top\} =: \Omega,$$

where Ω is upper bounded and time-invariant. Unlike the approaches of [33, 34], which assume the additive disturbance lies in a compact set, the disturbance ω_k is not assumed to be bounded and its distribution may have infinite support. It is assumed that the system state is measured directly and available to the controller at each sample instant.

The system (3.1) is subject to the constraint

$$\sum_{k=0}^{\infty} \gamma^k \mathbb{P}\{\|Cx_k\| \geq t\} \leq e, \tag{3.2}$$

for a given matrix $C \in \mathbb{R}^{n_c \times n_x}$, positive scalars t, e and discount factor $\gamma \in (0, 1)$. Note that the probability measure $\mathbb{P}$ denotes the product probability measure $\mathbb{P}^k$ that pertains to $(w_0, \ldots, w_{k-1}) \in \Delta^k$. For simplicity, it is written in the form of (3.2). This constraint gives a special feature to the control problem that the probability of future states violating the condition $\|Cx_k\| < t$ at time instants nearer to the initial time are weighted more heavily than those in the far future. For simplicity we refer to $\mathbb{P}\{\|Cx_k\| \geq t\}$ as a *violation probability.* Also, since $\mathbb{P}\{\|Cx_k\| \geq t\} = \mathbb{E}\{\mathbf{1}_{\mathbb{X}^c}(x_k)\}$, where $\mathbf{1}_{\mathbb{X}^c}$ is an indicator function and $\mathbb{X} := \{x \in \mathbb{R}^{n_x} | \, \|Cx\| < t\}$, (3.2) is equivalent to the constraint $\sum_{k=0}^{\infty} \gamma^k \mathbb{E}\{\mathbf{1}_{\mathbb{X}^c}(x_k)\} \leq e$. It is similar to the average-in-time chance constraint considered in [41], which treats violation probabilities at each time step with equal importance. The constraint (3.2) assigns different weights to each violation probability and can be interpreted as a constraint on weighted average-in-time violations.

The aim of this work is to design a controller that minimises the cost function

$$\mathbb{E}\Big\{\sum_{k=0}^{\infty} \|x_k - x^r\|_Q^2 + \|u_k - u^r\|_R^2\Big\} \tag{3.3}$$

and ensures a quadratic stability condition on the closed loop system while satisfying the constraint (3.2). The weighting matrices in (3.3) are assumed to satisfy $Q \succeq 0$

and $R \succ 0$, and we assume knowledge of reference targets x^r and u^r for the state and the control input that satisfy the steady state conditions

$$(I - A)\, x^r = Bu^r, \quad \|Cx^r\| < t. \tag{3.4}$$

Assumption 3.1. *(A, B) is controllable and $(A, Q^{\frac{1}{2}})$ is observable.*

3.2.1 Finite horizon formulation

The problem stated above employs an infinite horizon cost function and is subject to a constraint defined on an infinite horizon. If the infinite sequence of control inputs $\{u_k\}_{k=0}^{\infty}$ were considered to be decision variables, then clearly the optimisation problem would be infinite dimensional and thus in principle computationally intractable [72]. However, the use of an infinite horizon can impart desirable properties, notably stability [7, 111]. It is therefore beneficial to design an MPC law using a cost function and constraints that are defined over a finite horizon in such a way that they are equivalent to the infinite horizon cost and constraints of the original problem. The finite horizon cost function at time k for a prediction horizon of N steps is given by

$$\mathbb{E}\Big\{\sum_{i=0}^{N-1} \|x_{i|k} - x^r\|_Q^2 + \|u_{i|k} - u^r\|_R^2 + F(x_{N|k})\Big\} \tag{3.5}$$

where $\mathbb{E}\{F(x_{N|k})\}$ is the terminal cost and $F(x) \geq 0$ for all x. The constraint (3.2) is likewise truncated to a finite horizon:

$$\sum_{i=0}^{N-1} \gamma^i \mathbb{P}\{\|Cx_{i|k}\| \geq t\} + f(x_{N|k}) \leq \varepsilon_k. \tag{3.6}$$

Here $f(x_{N|k})$ is a terminal term chosen (as will be specified in (3.14) and Lemma 3.3) to approximate the infinite sum in (3.2) so that $\sum_{i=N}^{\infty} \gamma^i \mathbb{P}\{\|Cx_{i|k}\| \geq t\} \leq f(x_{N|k})$, and ε_k is a bound on the LHS of (3.6) that is achievable at time k. Although ε_k may increase or decrease over time since it is conditioned on the system state at time k, we show in Section 3.5 that (3.2) is satisfied if $\varepsilon_0 \leq e$ and ε_k is defined as described in Section 3.3.

Even with the cost and constraints defined as in (3.5)-(3.6) over a finite horizon, the probability distribution of states may be unknown at each time step and the finite horizon version of the problem is therefore still intractable in general. Even if the probability distribution of ω_k is known explicitly, computing (3.5) and (3.6) requires the solution of a set of multivariate convolution integrals, which in principle is difficult to manage [33].

3.2.2 Constraint handling and open loop optimisation

This section considers approximating the finite horizon constraint (3.6) using the two-sided Chebyshev inequality [112, Section V.7] and gives the explicit form of the MPC cost function. The cost and constraints are then combined to construct the MPC optimisation problem that is repeatedly solved online. We define the sequence of control inputs predicted at time k as

$$u_{i|k} = K(x_{i|k} - \bar{x}_{i|k}) + m_{i|k}, \qquad i = 0, \dots, N-1 \tag{3.7}$$

$$u_{N+i|k} = K(x_{N+i|k} - x^r) + u^r, \qquad i = 0, 1, \dots \tag{3.8}$$

where $m_{i|k}$ is the i-step-ahead prediction of the nominal control input given information at time k, that is, $\mathbb{E}_k\{u_{i|k}\} =: m_{i|k}$, and $\bar{x}_{i|k}$ is the i-step-ahead prediction of the nominal state given information at time k, that is, $\mathbb{E}_k\{x_{i|k}\} =: \bar{x}_{i|k}$.

Assumption 3.2. $\Phi := A + BK$ *is strictly stable.*

Given the predicted control law (3.7)-(3.8), the first two moments of the predicted state and control input sequences can be computed. Thus, the predicted nominal state trajectory is given by $\bar{x}_{0|k} = x_k$ and

$$\bar{x}_{i|k} = A^i \bar{x}_{0|k} + \sum_{j=0}^{i-1} A^{i-1-j} B m_{j|k}, \qquad i = 1, \dots, N, \tag{3.9}$$

$$\bar{x}_{N+i|k} = \Phi^i (\bar{x}_{N|k} - x^r) + x^r, \qquad i = 1, 2, \dots, \tag{3.10}$$

whereas the covariance matrix, $X_{i|k}$, of the i-step-ahead predicted state is given by $X_{0|k} = 0$ and

$$X_{i|k} = \sum_{j=0}^{i-1} \Phi^j \Omega (\Phi^j)^\top, \quad i = 1, 2, \dots. \tag{3.11}$$

Clearly $X_{i|k}$ is independent of k, and in the following development we simplify notation by letting $\hat{X}_i := X_{i|k}$.

In this chapter, we use Chebyshev's inequality to handle probabilistic constraints. The advantages of this approach are that it can cope with arbitrary or unknown disturbance probability distributions (the only information required being the first two moments of the predicted state trajectory), and furthermore it results in quadratic inequalities that are straightforward to implement. Approximating (3.6) by a direct application of the two-sided Chebyshev inequality [32], we obtain

$$\frac{\operatorname{tr}(C^\top C\hat{X}_i) + \|C\bar{x}_{i|k}\|^2}{t^2} \leq \beta_{i|k}, \qquad i = 0, \dots, N-1, \tag{3.12}$$

$$\sum_{i=0}^{N-1} \gamma^i \beta_{i|k} + f(\bar{x}_{N|k}) \leq \varepsilon_k, \tag{3.13}$$

where $\{\beta_{i|k}\}_{i=0}^{N-1}$ is a sequence of non-negative scalars. The terminal term $f(\bar{x}_{N|k})$ in (3.13) is chosen so that

$$f(\bar{x}_{N|k}) = \frac{\operatorname{tr}(C^\top C\widetilde{S})}{t^2} + \frac{\gamma^N}{t^2}\left(\|\bar{x}_{N|k} - x^r\|^2_{\widetilde{P}} + \frac{\|x^r\|^2_{C^\top C}}{1-\gamma}\right) + \frac{2\gamma^N (x^r)^\top C^\top C(I-\gamma\Phi)^{-1}(\bar{x}_{N|k} - x^r)}{t^2} \tag{3.14}$$

where $\widetilde{S} \succ 0$, $\widetilde{P} \succ 0$, and $I - \gamma\Phi$ is invertible since $\gamma\Phi$ is strictly stable. The design of $\widetilde{S}, \widetilde{P}$ is discussed in Section 3.5.

In terms of the predicted nominal state trajectory in (3.9)-(3.10), the predicted cost is defined as

$$J(\bar{x}_{0|k}, \{m_{i|k}\}_{i=0}^{N-1}, \varepsilon_k) := \|\bar{x}_{N|k} - x^r\|^2_P + \sum_{i=0}^{N-1}\left(\|\bar{x}_{i|k} - x^r\|^2_Q + \|m_{i|k} - u^r\|^2_R\right) \tag{3.15}$$

whenever a sequence $\{\beta_{i|k}\}_{i=0}^{N-1}$ exists satisfying (3.12)-(3.13) for the given $\bar{x}_{0|k}$, $\{m_{i|k}\}_{i=0}^{N-1}$ and ε_k. On the other hand, if $\bar{x}_{0|k}$, $\{m_{i|k}\}_{i=0}^{N-1}$ and ε_k are such that constraints (3.12)-(3.13) are infeasible, we set $J(\bar{x}_{0|k}, \{m_{i|k}\}_{i=0}^{N-1}, \varepsilon_k) := \infty$. Note that since the covariances of predicted states and control inputs are independent of $m_{i|k}$, minimising (3.5) is equivalent to minimising (3.15) in the sense that their optimal solutions are the same. In (3.15), $\|\bar{x}_{N|k} - x^r\|^2_P$ represents the terminal cost, and that $P \succ 0$. The choice of P is discussed in Section 3.4.

To summarise, the MPC optimisation solved at time k is

$$J^*(x_k, \varepsilon_k) := \min_{\{m_{i|k}\}_{i=0}^{N-1}} J(x_k, \{m_{i|k}\}_{i=0}^{N-1}, \varepsilon_k), \tag{3.16}$$

and its solution for any feasible x_k and ε_k is denoted

$$\left\{m^*_{i|k}(x_k, \varepsilon_k)\right\}_{i=0}^{N-1} := \arg \min_{\{m_{i|k}\}_{i=0}^{N-1}} J(x_k, \{m_{i|k}\}_{i=0}^{N-1}, \varepsilon_k). \tag{3.17}$$

For simplicity we write this solution as $\{m^*_{i|k}\}_{i=0}^{N-1}$, with the understanding that this sequence depends on x_k and ε_k. The corresponding nominal predicted state trajectory is given by

$$\bar{x}^*_{i|k} = A^i x_k + \sum_{j=0}^{i-1} A^{i-1-j} B m^*_{j|k}, \qquad i = 1, \dots, N, \tag{3.18}$$

$$\bar{x}^*_{N+i|k} = \Phi^i(\bar{x}^*_{N|k} - x^r) + x^r, \qquad i = 1, 2, \dots. \tag{3.19}$$

The MPC law at time k is defined by

$$u_k := m^*_{0|k}, \tag{3.20}$$

and the closed loop system dynamics are given by

$$x_{k+1} = Ax_k + Bm^*_{0|k}(x_k, \varepsilon_k) + \omega_k, \tag{3.21}$$

where ω_k is the disturbance realisation at time k.

In the remainder of this chapter, we discuss how to choose ε_k, K, P, $\widetilde{P}$ and $\widetilde{S}$ so as to guarantee a quadratic stability condition and satisfaction of the constraint (3.2) under the MPC law (3.20).

3.3 Recursive Feasibility

Recursively feasible MPC strategies have the property that the MPC optimisation problem is guaranteed to be feasible at every time step if it is initially feasible. This property can be ensured, for example, by imposing a terminal constraint that requires the predicted system state to lie in a particular set at the end of the prediction horizon [72]. For a deterministic MPC problem, if an optimal solution can be found at the current time, then the *tail sequence*, namely the optimal control sequence shifted forward by one time-step, will be a feasible, but possibly suboptimal, solution at the next time instant if the terminal constraint is defined in terms of a suitable invariant set for the predicted system state [15, 113]. For a robust MPC problem with bounded additive disturbances, recursive feasibility can likewise be guaranteed under either open or closed loop optimisation strategies by imposing a terminal constraint set that is robustly invariant. However, this approach is not generally applicable to systems with unbounded additive disturbances, and in general it is not possible to ensure recursive feasibility in this context while guaranteeing constraint satisfaction at every time instant.

In this section, we propose a method for guaranteeing recursive feasibility of the MPC optimisation that does not rely on terminal constraints. Instead recursive feasibility is ensured, despite the presence of unbounded disturbances, by allowing the constraint on the discounted sum of probabilities to be time-varying. For all time-steps $k > 0$, the approach uses the optimal sequence computed at time $k-1$ to determine a value of ε_k that is necessarily feasible at time k.

We use the notation $\mathscr{S}(\{m^*_{i|k}\}_{i=0}^{N-1})$ to denote a nominal control sequence derived from a time-shifted version of $\{m^*_{i|k}\}_{i=0}^{N-1}$, defined by

$$\mathscr{S}\Big(\{m^*_{i|k}\}_{i=0}^{N-1}\Big) := \{m^*_{i+1|k} + K\Phi^i\omega_k\}_{i=0}^{N-1}, \tag{3.22}$$

with $m^*_{N|k} := K(\bar{x}^*_{N|k} - x^r) + u^r$. Note that the disturbance realisation ω_k can be computed given the measured state x_{k+1} and hence the sequence $\mathscr{S}(\{m^*_{i|k}\}_{i=0}^{N-1})$ is available to the controller at time $k+1$.

Lemma 3.1. *The MPC optimisation* (3.16) *is recursively feasible if* ε_k *is defined at each time* $k = 1, 2, \ldots$ *as*

$$\varepsilon_k := \min\Big\{\varepsilon \mid J\Big(x_k, \mathscr{S}\Big(\{m^*_{i|k-1}\}_{i=0}^{N-1}\Big), \varepsilon\Big) < \infty\Big\}. \tag{3.23}$$

Proof. The definition of the MPC predicted cost implies that, for any given sequence $\{m_{i|k}\}_{i=0}^{N-1}$, there necessarily exists a value of ε such that $J(x_k, \{m_{i|k}\}_{i=0}^{N-1}, \varepsilon)$ is finite. Moreover $\mathscr{S}(\{m^*_{i|k-1}\}_{i=0}^{N-1})$ is (with probability 1) well-defined at time k if the MPC optimisation is feasible at time $k-1$. It follows that the minimum value of ε defining ε_k in (3.23) exists if the MPC optimisation is feasible at time $k-1$, and this establishes recursive feasibility. □

The sequence $\mathscr{S}(\{m^*_{i|k}\}_{i=0}^{N-1})$ can be regarded as the tail of the minimiser (3.17) with adjustments. With equations (3.9) and (3.10), the minimisation (3.23) can be solved to give an explicit expression for ε_k for all $k > 0$ as

$$\varepsilon_k = \sum_{i=0}^{N-1} \gamma^i \frac{\operatorname{tr}(C^\top C \hat{X}_i) + \|C(\bar{x}^*_{i+1|k-1} + \Phi^i \omega_{k-1})\|^2}{t^2} + f(\bar{x}^*_{N+1|k-1} + \Phi^N \omega_{k-1}). \quad (3.24)$$

Essentially, the optimisation problem to be solved at each time step is feasible because the parameter ε_k is updated via (3.24) using knowledge of the disturbance w_{k-1} obtained from the measurement of the current state x_k. In this respect, the approach is similar to constraint-tightening methods that have previously been applied in the context of stochastic MPC (e.g. [39, 40, 107]) in order to ensure recursive feasibility and constraint satisfaction in closed loop operation. However, each of these methods requires that the disturbances affecting the controlled system be bounded, and they become more conservative as the degree of conservativeness of the assumed disturbance bounds increases. The approach proposed here avoids this requirement.

The key to this method lies in the definition of the sequence $\mathscr{S}(\{m^*_{i|k}\}_{i=0}^{N-1})$. If this sequence were optimised simultaneously with ε_k, rather than defined by the suboptimal control sequence (3.22), then it would be possible to reduce the MPC cost (3.16). However, this would require more computational effort than is needed to evaluate (3.24) and lose the guarantee of satisfying (3.2) in closed loop.

3.4 Stochastic MPC Algorithm

For deterministic MPC, it can be shown by using the *tail sequence* that, with an appropriate terminal weighting matrix [7], optimal MPC predicted costs are monotonically non-increasing along closed loop trajectories. This property does not generally hold in the presence of unbounded disturbances, and in fact the optimal

MPC predicted cost defined by (3.16) is not necessarily monotonically non-increasing if ε_k is defined by (3.23). However, in this section, we show that the proposed approach ensures that the closed loop system satisfies a quadratic stability condition.

We first state the MPC algorithm based on the optimisation defined in (3.16).

Algorithm 3.1. *At each time step $k = 0, 1, \ldots$:*

(i) Measure x_k, and if $k > 0$, compute ε_k using (3.24)*.*

(ii) Solve the quadratically constrained quadratic programming (QCQP) problem:

$$\min_{\{m_{i|k}, \beta_{i|k}\}_{i=0}^{N-1}} \sum_{i=0}^{N-1} (\|\bar{x}_{i|k} - x^r\|_Q^2 + \|m_{i|k} - u^r\|_R^2) + \|\bar{x}_{N|k} - x^r\|_P^2$$

subject to (3.12), (3.13), *and* (3.9) *with $\bar{x}_{0|k} = x_k$.*

(iii) Apply the control law $u_k = m^*_{0|k}$.

Although the MPC optimisation in step (ii) involves a quadratic constraint as well as linear constraints, it can be solved efficiently, for example using a second-order conic program (SOCP) solver, since the objective and the quadratic constraint are both convex.

Theorem 3.2. *Given initial feasibility at $k = 0$, the minimisation in step (ii) of Algorithm 3.1 is feasible for $k = 1, 2, \ldots$ and the closed loop system satisfies the quadratic stability condition*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\{\|x_k - x^r\|_Q^2 + \|u_k - u^r\|_R^2\} \leq \mathrm{tr}(\Omega P) \tag{3.25}$$

provided K in (3.7)-(3.8) *and P in* (3.15) *are chosen so that*

$$P = \Phi^\top P \Phi + K^\top R K + Q. \tag{3.26}$$

Stability is the overriding requirement and in most recent MPC literature the optimal cost function is chosen as a Lyapunov function suitable for analysing closed loop stability [7]. Theorem 3.2 is proved via cost comparison, and, given the quadratic form of the cost function, this analysis results in the quadratic stability condition (3.25). Similar asymptotic bounds on the time average of quadratic expected stage costs are obtained in [33, 114]. However, in the current context, Theorem 3.2 demonstrates that an MPC algorithm can ensure closed loop stability *without* imposing terminal constraints based on an invariant set.

3.5 Sequence of Tightening Parameters and Constraint Satisfaction

This section considers the properties of the sequence $\{\varepsilon_k\}_{k=0}^{\infty}$ in closed loop operation under Algorithm 3.1. We first give expressions for the parameters $\widetilde{S}$ and $\widetilde{P}$ in the definition (3.14) of the terminal term $f(\bar{x}_{N|k})$. Then, using the explicit expression for ε_k in (3.24), we derive a recurrence equation relating the expected value of ε_{k+1} to x_k and ε_k. This allows an upper bound to be determined for the sum of discounted violation probabilities on the left hand side of (3.2). With this bound, we can show that the closed loop system under the control law of Algorithm 3.1 satisfies the chance constraint (3.2) if ε_k is initialised with $\varepsilon_0 = e$.

Lemma 3.3. *Let $\widetilde{S}$ and $\widetilde{P}$ be the solutions of*

$$\widetilde{P} = \gamma \Phi^\top \widetilde{P} \Phi + C^\top C, \tag{3.27}$$

$$\widetilde{S} = \gamma \Phi \widetilde{S} \Phi^\top + \frac{\gamma^{N+1}}{1-\gamma} \Omega + \gamma^N \hat{X}_N. \tag{3.28}$$

Then $f(\bar{x}_{N|k})$ defined in (3.14) satisfies

$$f(\bar{x}_{N|k}) = \sum_{i=N}^{\infty} \gamma^i \frac{\mathrm{tr}\Big(C^\top C \hat{X}_i\Big) + \|C\bar{x}_{i|k}\|^2}{t^2} \tag{3.29}$$

where $\bar{x}_{i|k}$ is given by (3.10) for all $i \geq N$.

The following result gives the relationship between ε_k and the expected value of ε_{k+1} for the closed loop system, which will be used to show Theorem 3.5.

Theorem 3.4. *If ε_k is defined by (3.24) at all times $k \geq 1$, then in closed loop operation under Algorithm 3.1 we have*

$$\gamma \mathbb{E}_k\{\varepsilon_{k+1}\} \leq \varepsilon_k - \frac{\|Cx_k\|^2}{t^2} \tag{3.30}$$

for all $k \geq 0$.

The main result of this section is given next.

Theorem 3.5. *The closed loop system under Algorithm 3.1 satisfies the chance constraint* (3.2) *if* $\varepsilon_0 = e$.

The presence of the factor $\gamma \in (0, 1)$ on the LHS of (3.30) implies that the expected value of ε_k can increase as well as decrease along closed loop system trajectories. In fact, for values of γ close to zero, a rapid initial growth in ε_k is to be expected, which is in agreement with the interpretation that the constraint (3.2) penalises violation probabilities more heavily at times closer to the initial time in this case. On the other hand, for values of γ close to 1, ε_k can be expected to decrease initially, implying a greater emphasis on the expected number of violations over some initial horizon.

3.6 Numerical Example

This section describes a numerical example illustrating the quadratic stability condition and constraint satisfaction of the closed loop system (3.21) under Algorithm 3.1. Consider a system with

$$A = \begin{bmatrix} 1 & 2 \\ 1.5 & 0.5 \end{bmatrix}, \quad B = \begin{bmatrix} 1.2 \\ 1.5 \end{bmatrix},$$

and Gaussian disturbance $\omega_k \sim \mathcal{N}(0, \Omega)$ with covariance matrix $\Omega = 0.2 I_{2\times 2}$. The constraint (3.2) is defined by $\gamma = 0.9$, $t = 1$, $e = 3.5$, $C = \begin{bmatrix} 0.6 & 0.52 \end{bmatrix}$, and the weighting matrices in the cost (3.3) are given by

$$Q = C^\top C = \begin{bmatrix} 0.3600 & 0.3120 \\ 0.3120 & 0.2704 \end{bmatrix}, \quad R = 1.$$

Input and state references are $u^r = -0.6$, $x^r = [0.72, 0.36]^\top$, and the prediction horizon is chosen as $N = 7$. The feedback gain is chosen as $K = [-0.92 \ -0.85]$ for the cost (3.3), and matrices P, $\widetilde{P}$ and $\widetilde{S}$ are chosen to satisfy (3.26), (3.27) and (3.28). The initial value for ε_k is $\varepsilon_0 = e = 3.5$.

Two sets of simulations (A and B) demonstrate the closed loop stability result in Theorem 3.2 and the constraint satisfaction result in Theorem 3.5, respectively.

Simulation A: To estimate empirically the average cost,

$$\bar{J} := \lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\Big\{ \|x_k - x^r\|_Q^2 + \|u_k - u^r\|_R^2 \Big\},$$

we consider the mean value of the stage cost over 100 simulations. Each simulation has a randomly selected initial condition ($x_0 \sim \mathcal{N}(0, I)$, with infeasible values discarded), and a length of $T = 500$ time steps. This gives the estimated average cost as $\bar{J} \approx 0.5036$, which is no greater than $\mathrm{tr}(\Omega P) = 0.5304$, and hence agrees with the bound (3.25). Moreover, the estimate of $\bar{J}$ decreases considerably more slowly as the simulation length T increases.

Simulation B: To test numerically whether the chance constraint (3.2) is satisfied, we estimate the discounted sum of violation probabilities on the LHS of (3.2),

$$V := \sum_{k=0}^{\infty} \gamma^k \mathbb{P}\{\|Cx_k\| \geq t\},$$

by counting the number of violations at $k \in \{0, \ldots, T-1\}$, for 10^3 simulations with $x_0 = [-1.1130, 1.1156]^\top$ and $T = 100$. We run three tests, each of which chooses a different value for e. Simulation results are summarised in Table 3.1, which confirms the satisfaction of constraint (3.2) and shows that the relative level of violation rates with respect to the value of e remains almost unchanged. For this example, we have $\gamma^{100} \approx 10^{-5}$, so increasing T beyond 100 time steps has negligible effect on the estimate of V. It is noticed that there are discrepancies between constraint violation rates observed and the limit imposed. There are three factors that may account for this: (1) Chebyshev's inequality is used to handle the chance constraint, bringing substantial conservativeness; (2) constraint-tightening is used to ensure recursive feasibility; (3) e is an upper bound on the discounted sum of violation probabilities and there may exist solutions yielding small rates of violation. In addition, if the unconstrained LQ-optimal feedback law $u_k = K_{LQ}(x_k - x^r) + u^r$ were employed, the value of the bound $\sum_{k=0}^{\infty} \gamma^k \mathbb{E}_k\{\|Cx_k\|^2\}/t^2$ in (3.35) would be 4.6998, which exceeds e. Hence this control law may not satisfy (3.2) and is worse than the MPC law (3.20) in terms of this bound.

V	e
0.8328	3.5
0.8088	3.3
0.7357	3.1

Table 3.1: Violation rates for Simulation B

3.7 Conclusion

A stochastic MPC algorithm that imposes constraints on the sum of discounted future constraint violation probabilities can ensure recursive feasibility of the online optimisation and closed loop constraint satisfaction without the boundedness assumption on additive disturbances. Key features are the design of a constraint-tightening procedure and closed loop analysis of the tightening parameters. The MPC algorithm requires knowledge of the first and second moments of the disturbance, and is implemented as a convex QCQP problem.

3.8 Proofs of Chapter 3

3.8.1 Proof of Theorem 3.2

From Lemma 3.1, the sequence $\mathscr{S}\Big(\{m^*_{i|k}\}_{i=0}^{N-1}\Big)$ provides a feasible solution at time $k+1$. Hence, by optimality, we necessarily have

$$J^*(x_{k+1}, \varepsilon_{k+1}) \leq J(x_{k+1}, \mathscr{S}(\{m^*_{i|k}\}_{i=0}^{N-1}), \varepsilon_{k+1}),$$

and since this inequality holds for every realisation of ω_k, by taking expectations conditioned on the state x_k, we obtain

$$\mathbb{E}_k\{J^*(x_{k+1}, \varepsilon_{k+1})\} \leq \mathbb{E}_k\{J(x_{k+1}, \mathscr{S}(\{m^*_{i|k}\}_{i=0}^{N-1}), \varepsilon_{k+1})\}. \tag{3.31}$$

Evaluating $\bar{x}_{i|k+1}$ by setting $\bar{x}_{0|k+1} = x_{k+1}$ and $m_{i|k+1} = m^*_{i+1|k} + K\Phi^i\omega_k$ in (3.9)-(3.10) gives the feasible sequence

$$\bar{x}_{i|k+1} = \bar{x}^*_{i+1|k} + \Phi^i\omega_k, \qquad i = 0, \dots, N,$$

and, from (3.26) and (3.31), it follows that

$$\mathbb{E}_k\{J^*(x_{k+1}, \varepsilon_{k+1})\} \leq J^*(x_k, \varepsilon_k) - \|x_k - x^r\|_Q^2 - \|u_k - u^r\|_R^2 + \mathrm{tr}(\Omega P). \tag{3.32}$$

Summing both sides of this inequality over $k \geq 0$ after taking expectations given information available at time $k = 0$, and making use of the property that $\mathbb{E}_0\{\mathbb{E}_k\{J^*(x_{k+1}, \varepsilon_{k+1})\}\} = \mathbb{E}_0\{J^*(x_{k+1}, \varepsilon_{k+1})\}$, we obtain (3.25).

3.8.2 Proof of Lemma 3.3

Writing $\|C\bar{x}_{i|k}\|^2 = \|C(\bar{x}_{i|k} - x^r) + Cx^r\|^2$ and using (3.10), we obtain

$$\|C\bar{x}_{i|k}\|^2 = \|C\Phi^{i-N}(\bar{x}_{N|k} - x^r)\|^2 + 2(x^r)^\top C^\top C\Phi^{i-N}(\bar{x}_{N|k} - x^r) + \|Cx^r\|^2$$

for all $i \geq N$, and since $\widetilde{P}$ satisfies (3.27), we have

$$\sum_{i=N}^{\infty} \frac{\gamma^i}{t^2}\|C\bar{x}_{i|k}\|^2 = \frac{\gamma^N}{t^2}\|\bar{x}_{N|k} - x^r\|^2_{\widetilde{P}} + \frac{\gamma^N}{1-\gamma}\frac{\|x^r\|^2_{C^\top C}}{t^2} + \frac{2\gamma^N(x^r)^\top C^\top C(I-\gamma\Phi)^{-1}(\bar{x}_{N|k} - x^r)}{t^2}. \quad (3.33)$$

Furthermore, if $\widetilde{S} = \sum_{i=N}^{\infty} \gamma^i \hat{X}_i$, then $\widetilde{S}$ is the solution of the Lyapunov equation (3.28) since (3.11) implies

$$\gamma\Phi\widetilde{S}\Phi^\top = \sum_{i=N}^{\infty} \gamma^{i+1}\Phi\hat{X}_i\Phi^\top = \sum_{i=N}^{\infty} \gamma^{i+1}(\hat{X}_{i+1} - \Omega) = \widetilde{S} - \gamma^N\hat{X}_N - \frac{\gamma^{N+1}}{1-\gamma}\Omega,$$

and it follows that

$$\sum_{i=N}^{\infty} \frac{\gamma^i}{t^2}\operatorname{tr}(C^\top C\hat{X}_i) = \frac{\operatorname{tr}(C^\top C\widetilde{S})}{t^2}. \quad (3.34)$$

Combining (3.33) and (3.34), it is clear that (3.29) is equivalent to (3.14) if $\widetilde{P}$ and $\widetilde{S}$ are defined by (3.27) and (3.28).

3.8.3 Proof of Theorem 3.4

Evaluating ε_{k+1} using (3.24) and (3.29) gives

$$\varepsilon_{k+1} = \sum_{i=0}^{\infty} \gamma^i \frac{\operatorname{tr}(C^\top C\hat{X}_i) + \|C(\bar{x}^*_{i+1|k} + \Phi^i\omega_k)\|^2}{t^2},$$

where $\bar{x}^*_{i|k}$ is given by (3.18)-(3.19) and ω_k is the realisation of the disturbance at time k. Taking expectations conditioned on information available at time k, we have

$$\gamma\mathbb{E}_k\{\varepsilon_{k+1}\} = \sum_{i=0}^{\infty} \gamma^{i+1}\frac{\operatorname{tr}(C^\top C\hat{X}_i) + \|C\bar{x}^*_{i+1|k}\|^2}{t^2} + \sum_{i=0}^{\infty} \gamma^{i+1}\frac{\operatorname{tr}(C^\top C\Phi^i\Omega(\Phi^i)^\top)}{t^2},$$

but feasibility of the sequence $\{m^*_{i|k}\}_{i=0}^{N-1}$ at time k implies $\sum_{i=0}^{\infty} \frac{\gamma^i}{t^2}\Big(\mathrm{tr}(C^\top C\hat{X}_i) + \|C\bar{x}^*_{i|k}\|^2\Big) \leq \varepsilon_k$. Thus,

$$\gamma\mathbb{E}_k\{\varepsilon_{k+1}\} \leq \varepsilon_k - \frac{\|Cx_k\|^2}{t^2} + \sum_{i=0}^{\infty}\frac{\gamma^i}{t^2}\,\mathrm{tr}\Big(C^\top C\Big(\gamma\Phi^i\Omega\Phi^{i\top} + (\gamma-1)\hat{X}_i\Big)\Big).$$

To complete the proof, we note that the sum on the RHS of this inequality is zero since

$$\gamma^{i+1}\Phi^i\Omega\Phi^{i\top} + (\gamma^{i+1} - \gamma^i)\hat{X}_i = \gamma^{i+1}\hat{X}_{i+1} - \gamma^i\hat{X}_i,$$

and because $\hat{X}_0 = 0$ and $\lim_{i\to\infty}\gamma^i\hat{X}_i = 0$.

3.8.4 Proof of Theorem 3.5

Theorem 3.4 implies that the closed loop evolution of ε_k satisfies

$$\gamma^{i+1}\mathbb{E}_k\{\varepsilon_{k+i+1}\} \leq \gamma^i\mathbb{E}_k\{\varepsilon_{k+i}\} - \frac{\gamma^i}{t^2}\mathbb{E}_k\{\|Cx_{k+i}\|^2\}$$

for all non-negative integers k, i. Summing both sides of this equation over $i \in \{0, 1, \ldots\}$ gives

$$\varepsilon_k \geq \sum_{i=0}^{\infty}\gamma^i\frac{\mathbb{E}_k\{\|Cx_{k+i}\|^2\}}{t^2} + \lim_{i\to\infty}\gamma^i\mathbb{E}_k\{\varepsilon_{k+i}\}. \tag{3.35}$$

But $\gamma^i\mathbb{E}_k\{\varepsilon_{k+i}\}$ is necessarily non-negative for all $k, i \geq 0$, so by Chebyshev's inequality this implies

$$\sum_{i=0}^{\infty}\gamma^i\mathbb{P}\{\|Cx_{k+i}\| \geq t\} \leq \varepsilon_k \tag{3.36}$$

for all $k \geq 0$. An obvious consequence of the bound (3.36) is that the closed loop system will satisfy the chance constraint (3.2) if ε_0 is chosen to be equal to e.

4
Multiobjective Optimisation and Dynamic Programming

Contents

This chapter lays the foundation for further developments in Chapter 5. We consider a convex multiobjective optimisation problem (MOP) subject to the dynamics of a discrete-time linear system. The MOP is cast as a single objective problem via linear scalarisation and then solved numerically using dynamic programming (DP). Their solutions yield a set of strictly stabilising feedback gains, denoted as $\mathcal{K}$, and sequences of solution matrices to algebraic Riccati equations. Properties of these solution matrices are exploited in Chapter 5 for the design of dynamic gain selection procedures, based on which a suitable feedback gain is selected from the set $\mathcal{K}$ at each time step.

4.1 Introduction

Multiobjective optimisation minimises or maximises more than one objective function simultaneously. It has a wide range of applications, such as finance, inventory control, manufacturing and game theory. Although optimisation problems in engineering applications are commonly cast as scalar objective problems, which have been well studied, they often need to consider several possibly conflicting requirements at the same time. Unlike scalar objective optimisation problems where the optimal solution can be determined by comparing values of the objective function, there generally is no unique optimiser and many optimal points exist instead. These optimal points are referred to as Pareto optimal solutions [115]. A set of Pareto optimal solutions is called a Pareto front, and finding the complete Pareto front is fundamental in solving MOPs.

Various methods exist for solving MOPs. For example, the weighted sum method [116] assigns nonnegative weights to every objective function of the MOP and defines a single objective as the weighted sum. This formulation only provides a sufficient condition for Pareto optimality [117] and solutions of this single objective problem lie in the Pareto optimal set. For convex MOPs, this method also provides a necessary condition [118] and then a complete Pareto front can be captured. The ϵ-constraint method [119] also turns an MOP into a single objective problem, where one objective function $f_k(x)$ is minimised and other objective functions are reformulated as constraints $f_i(x) \leq \epsilon_i \; i = 1, \ldots, k-1, k+1, \ldots, p$. Parameters ϵ_i are appropriately selected based on prior information about the Pareto front, which may not be available. Changing their values within certain ranges yields different points in the Pareto optimal set [120].

In this chapter, we solve a multiobjective linear quadratic optimal control problem by using the weighted sum method. Given a pair of nonnegative weights, the resulting single objective problem is solved via dynamic programming and a strictly stabilising feedback gain is obtained. To explore trade-offs between objective functions, we solve the corresponding single objective problems numerically with many different pairs of weighting parameters within a certain range. The set, $\mathcal{K}$, of strictly stabilising feedback gains and the sets of solution matrices to algebraic Riccati equations are then obtained. We also show that these solution matrices are monotonic and Lipschitz continuous functions of the weighting parameter.

This chapter is organised as follows. The MOP is described in Section 4.2 and reformulated as a single objective problem, which is then solved via dynamic programming, in Section 4.3. In Section 4.4, we analyse the properties of the solution matrices to algebraic Riccati equations. A numerical example is provided in Section 4.5 and this chapter is concluded in Section 4.6.

4.2 Problem Description

Consider a discrete-time linear system

$$x_{i+1} = Ax_i + Bu_i + \omega_i, \tag{4.1}$$

where $x_i \in \mathbb{R}^{n_x}$, $u_i \in \mathbb{R}^{n_u}$ are the system state and control input. The additive disturbance $\omega_i \in \mathbb{R}^{n_x}$ is assumed to be independently and identically distributed with zero mean and $\mathbb{E}\{\omega_k \omega_k^\top\} =: \Omega \succ 0$. In this chapter, we solve the following problem

$$\begin{aligned} \min_{u_0,u_1,\ldots} \ & \mathbf{f} := [f_1, f_2]^\top \\ \text{s.t. } & (4.1), \end{aligned} \tag{4.2}$$

where

$$f_1 := \sum_{i=0}^{\infty} \gamma^i \mathbb{E}\{\|Cx_i\|^2\}, \quad f_2 := \sum_{i=0}^{\infty} \mathbb{E}\{\|x_i\|_Q^2 + \|u_i\|_R^2\}.$$

Note that $\gamma \in (0,1)$ is a discount factor and it is assumed $Q \succeq 0$ and $R \succ 0$. The objective function f_1 corresponds to a reformulated form of the constraint function in (3.2) of Chapter 3 via Chebyshev's inequality and f_2 corresponds to a similar version of the cost function in (3.3) of Chapter 3. By solving this MOP, we aim to compute linear state feedback gains that explore trade-offs between these two conflicting objectives.

Assumption 4.1. *(A, B) is controllable and $(A, Q^{\frac{1}{2}})$ is observable.*

4.3 Dynamic Programming

In this section, we propose to solve problem (4.2) using the weighted sum method and formulate a single objective optimisation problem $(\mathcal{P}_1)$. Its optimal solution is a linear state feedback law in the form $u_i = Kx_i \ \forall i \geq 0$. To explore possible trade-offs between the two competing objectives, problem $(\mathcal{P}_1)$ is solved with different pairs of weighting parameters and we obtain a set of different stabilising feedback gains, denoted as $\mathcal{K}$.

The single objective optimisation problem is written for a given $x_0 \in \mathbb{R}^{n_x}$ in the form

$$\begin{aligned} \min_{u_0,u_1,\dots} \ & (1-\mu)\sum_{i=0}^{\infty}\gamma^i \mathbb{E}\Big\{\|Cx_i\|^2\Big\} + \mu \sum_{i=0}^{\infty}\mathbb{E}\Big\{\|x_i\|_Q^2 + \|u_i\|_R^2\Big\} \qquad (\mathcal{P}_1) \\ \text{s.t.} \ & (4.1) \end{aligned}$$

where $\mu \in (0,1]$ is a weighting parameter. We solve problem $(\mathcal{P}_1)$ by dynamic programming [3] via its equivalent deterministic counterpart in the form of

$$\begin{aligned} \min_{\bar{u}_0,\bar{u}_1,\dots} \ & (1-\mu)\sum_{i=0}^{\infty}\gamma^i \|C\bar{x}_i\|^2 + \mu \sum_{i=0}^{\infty}\Big(\|\bar{x}_i\|_Q^2 + \|\bar{u}_i\|_R^2\Big) \qquad (\mathcal{P}_2) \\ \text{s.t.} \ & \bar{x}_{i+1} = A\bar{x}_i + B\bar{u}_i, \ i = 0,1,\dots \qquad (4.3) \end{aligned}$$

where $\bar{x}_i$ and $\bar{u}_i$ are nominal values of x_i and u_i respectively, with $\bar{x}_0 = x_0$. From the quadratic form of the objective, the optimal solution to $(\mathcal{P}_2)$ is a linear feedback control law, and its corresponding feedback gain has desirable properties as will be shown in Theorem 4.2. In the first instance, we show the equivalence between $(\mathcal{P}_1)$ and $(\mathcal{P}_2)$ in Theorem 4.1, which is in agreement with the certainty equivalence principle [5, Section 8.4] that relies on the purely additive nature of the disturbance.

Theorem 4.1. *Problems $(\mathcal{P}_1)$ and $(\mathcal{P}_2)$ are equivalent in the sense that their optimal solutions are linear feedback control laws with the same feedback gains.*

Proof. We begin by deriving the solution of $(\mathcal{P}_1)$. Let

$$\bar{J}_0(x_0) := \begin{bmatrix} x_0 \\ 1 \end{bmatrix}^\top \begin{bmatrix} \bar{S}_0 & 0 \\ 0 & r_0 \end{bmatrix} \begin{bmatrix} x_0 \\ 1 \end{bmatrix}, \ \hat{J}_0(x_0) := \begin{bmatrix} x_0 \\ 1 \end{bmatrix}^\top \begin{bmatrix} \hat{S}_0 & 0 \\ 0 & v_0 \end{bmatrix} \begin{bmatrix} x_0 \\ 1 \end{bmatrix},$$

$$
\begin{aligned}
&(1-\mu)\bar{J}_0(x_0) + \mu \hat{J}_0(x_0) := \\
&\lim_{T\to\infty} \min_{u_0,\ldots,u_T} \mathbb{E}\Big\{(1-\mu)\sum_{i=0}^{T} \gamma^i \|Cx_i\|^2 + \mu \sum_{i=0}^{T} (\|x_i\|_Q^2 + \|u_i\|_R^2)\Big\} \\
&\qquad \text{s.t. } (4.1)
\end{aligned}
$$

where $\bar{J}_0(x_0)$ and $\hat{J}_0(x_0)$ denote the optimal values of $\lim_{T\to\infty}\sum_{i=0}^{T}\gamma^i\mathbb{E}\{\|Cx_i\|^2\}$ and $\lim_{T\to\infty}\sum_{i=0}^{T}\mathbb{E}\{\|x_i\|_Q^2+\|u_i\|_R^2\}$ respectively. Each of $\bar{J}_0(x_0)$ and $\hat{J}_0(x_0)$ consists of a term associated with the initial condition x_0 (weighted by $\bar{S}_0 \succeq 0$ and $\hat{S}_0 \succeq 0$ respectively) and a term associated with the covariance of x_i, $i = 1, \ldots, T$ (denoted by r_0 and v_0 respectively). Due to their quadratic form and the assumption of zero-mean disturbance ω_i, the optimal control input u_i is a linear function of x_i for each i. Denoting the optimal feedback gain as H_i and using the Bellman principle of optimality [3], we obtain two separate sequences of Bellman equations coupled via feedback gains H_i as

$$
\begin{aligned}
\bar{J}_i(x_i) &= \|Cx_i\|^2 + \gamma\mathbb{E}_i\{\bar{J}_{i+1}(x_{i+1})\} = \|Cx_i\|^2 + \gamma\mathbb{E}_i\Big\{\begin{bmatrix} x_{i+1} \\ 1\end{bmatrix}^\top \begin{bmatrix} \bar{S}_{i+1} & 0 \\ 0 & r_{i+1}\end{bmatrix}\begin{bmatrix} x_{i+1} \\ 1\end{bmatrix}\Big\} \\
&= \begin{bmatrix} x_i \\ 1\end{bmatrix}^\top \begin{bmatrix} C^\top C + \gamma(A+BH_i)^\top \bar{S}_{i+1}(A+BH_i) & 0 \\ 0 & \gamma\Big(r_{i+1} + \mathrm{tr}(\bar{S}_{i+1}\Omega)\Big)\end{bmatrix}\begin{bmatrix} x_i \\ 1\end{bmatrix}
\end{aligned}
$$

for $i = 0, \ldots, T-1, \forall x_i \in \mathbb{R}^{n_x}$,

$$
\begin{aligned}
\hat{J}_i(x_i) &= \|x_i\|_Q^2 + \|H_i x_i\|_R^2 + \mathbb{E}_i\{\hat{J}_{i+1}(x_{i+1})\} \\
&= \|x_i\|^2_{Q+H_i^\top R H_i} + \mathbb{E}_i\Big\{\begin{bmatrix} x_{i+1} \\ 1\end{bmatrix}^\top \begin{bmatrix} \hat{S}_{i+1} & 0 \\ 0 & v_{i+1}\end{bmatrix}\begin{bmatrix} x_{i+1} \\ 1\end{bmatrix}\Big\} \\
&= \begin{bmatrix} x_i \\ 1\end{bmatrix}^\top \begin{bmatrix} Q + H_i^\top R H_i + (A+BH_i)^\top \hat{S}_{i+1}(A+BH_i) & 0 \\ 0 & v_{i+1} + \mathrm{tr}\Big(\hat{S}_{i+1}\Omega\Big)\end{bmatrix}\begin{bmatrix} x_i \\ 1\end{bmatrix}
\end{aligned}
$$

for $i = 0, \ldots, T-1, \forall x_i \in \mathbb{R}^{n_x}$.

These two sequences of Bellman equations yield the following DP iteration for $i = T-1, \ldots, 0$:

$$\bar{S}_i = C^\top C + \gamma(A+BH_i)^\top \bar{S}_{i+1}(A+BH_i), \tag{4.4}$$

$$\hat{S}_i = Q + H_i^\top R H_i + (A+BH_i)^\top \hat{S}_{i+1}(A+BH_i), \tag{4.5}$$

$$r_i = \gamma\left(r_{i+1} + \mathrm{tr}(\bar{S}_{i+1}\Omega)\right), \tag{4.6}$$

$$v_i = \mathrm{tr}\Big(\hat{S}_{i+1}\Omega\Big) + v_{i+1}, \tag{4.7}$$

where we choose $\bar{S}_T = \hat{S}_T = 0$ and $r_T = v_T = 0$. The corresponding feedback gains H_i can be computed as

$$H_i = -\Big(\mu R + B^\top \Delta_{i+1} B\Big)^{-1} B^\top \Delta_{i+1} A$$

for $i = T-1, \ldots, 0$ with $\Delta_{i+1} = (1-\mu)\gamma \bar{S}_{i+1} + \mu \hat{S}_{i+1}$. To obtain H_i, $\bar{S}_i$ and $\hat{S}_i$ it is only necessary to perform the iterations in (4.4)-(4.5), since these are independent of r_i and v_i. Considering the limit as $T \to \infty$ and using the similar reasoning, it can be shown that $x_0^\top \Big((1-\mu)\bar{S}_0 + \mu \hat{S}_0\Big) x_0$ is the optimal objective of problem $(\mathcal{P}_2)$ and the corresponding optimal control law is given for all $i \geq 0$ by $\bar{u}_i = H_i \bar{x}_i$. $\square$

Theorem 4.1 demonstrates that we only need to solve problem $(\mathcal{P}_2)$ to obtain the optimal feedback gains for $(\mathcal{P}_1)$. To suit our purposes, we reverse the time indexing in (4.4) and (4.5) and define $\bar{P}_i := \bar{S}_{T-i}$, $\hat{P}_i := \hat{S}_{T-i}$ for $i = 0, \ldots, T$ and $L_i := H_{T-i}$ for $i = 1, \ldots, T$. The resulting DP iteration is

$$L_{i+1} = -\Big(\mu R + B^\top \Sigma_i B\Big)^{-1} B^\top \Sigma_i A, \tag{4.8}$$

$$\bar{P}_{i+1} = C^\top C + \gamma (A + BL_{i+1})^\top \bar{P}_i (A + BL_{i+1}), \tag{4.9}$$

$$\hat{P}_{i+1} = Q + L_{i+1}^\top R L_{i+1} + (A + BL_{i+1})^\top \hat{P}_i (A + BL_{i+1}), \tag{4.10}$$

with $\Sigma_i := \gamma(1-\mu)\bar{P}_i + \mu \hat{P}_i$, for $i = 0, \ldots, T-1$. These provide the solution of $(\mathcal{P}_2)$ in the limit as $T \to \infty$.

Theorem 4.2. *Consider equations* (4.8)-(4.10). *Under Assumption 4.1: (a) there exist matrices $\bar{P}_\infty \succeq 0$, $\hat{P}_\infty \succ 0$ such that for arbitrary positive semidefinite matrices $\bar{P}_0$ and $\hat{P}_0$ we have*

$$\lim_{i \to \infty} \bar{P}_i = \bar{P}_\infty \text{ and } \lim_{i \to \infty} \hat{P}_i = \hat{P}_\infty,$$

(b) for all $\mu \in (0, 1]$ the feedback gain L_∞ is strictly stabilising, where

$$L_\infty := -\Big(\mu R + B^\top \Sigma_\infty B\Big)^{-1} B^\top \Sigma_\infty A \tag{4.11}$$

with $\Sigma_\infty := \gamma(1-\mu)\bar{P}_\infty + \mu \hat{P}_\infty$, and (c) the matrices $\bar{P}_\infty$ and $\hat{P}_\infty$ satisfy the following Lyapunov matrix equations:

$$\bar{P}_\infty = C^\top C + \gamma \left(A + BL_\infty\right)^\top \bar{P}_\infty \left(A + BL_\infty\right), \tag{4.12}$$

$$\hat{P}_\infty = Q + L_\infty^\top R L_\infty + \left(A + BL_\infty\right)^\top \hat{P}_\infty \left(A + BL_\infty\right). \tag{4.13}$$

Proof. The proof uses standard DP arguments, which can be found, for example, in [19], and therefore is omitted here. □

The solutions L_∞, $\bar{P}_\infty$, $\hat{P}_\infty$ are functions of μ and are denoted $L_\infty(\mu)$, $\bar{P}_\infty(\mu)$, $\hat{P}_\infty(\mu)$ in the remainder of this chapter. With $\mu = 1$, $L_\infty(1) = K_{LQ}$, where K_{LQ} is the LQ-optimal feedback gain that minimises the second part of the objective in $(\mathcal{P}_2)$, whereas with $\mu = 0$ $L_\infty(0) = -\left(B^\top \bar{P}_\infty(0) B\right)^\dagger B^\top \bar{P}_\infty(0) A$, which is optimal with respect to the first part of the objective. However, the gain $L_\infty(0)$ may not be stabilising and unique due to discounting and the absence of penalties on control inputs if $\mu = 0$, and we therefore restrict the weighting parameter μ to the interval $(0, 1]$.

Remark 4.3. *The fixed point* $\left(L_\infty(\mu), \bar{P}_\infty(\mu), \hat{P}_\infty(\mu)\right)$ *of the iteration* (4.8)-(4.10) *coincides with the minimising argument of*

$$\begin{aligned}
\min_{G \in \mathbb{R}^{n_u \times n_x}, Z_1 \succeq 0, Z_2 \succeq 0} \quad & \operatorname{tr}((1-\mu) Z_1 + \mu Z_2) \\
s.t. \quad & Z_1 = C^\top C + \gamma (A + BG)^\top Z_1 (A + BG) \\
& Z_2 = Q + G^\top R G + (A + BG)^\top Z_2 (A + BG).
\end{aligned}$$

However, the iteration is generally preferred over solving this equivalent problem directly since it is nonconvex in variables G, Z_1, Z_2, with no obvious convexifying transformation.

Next, we give procedures to generate feedback gains based on a sequence of positive weighting parameters for gain selection methods in Section 5.4 of Chapter 5.

Algorithm 4.1 (Offline generation of the set $\mathcal{K}$)**.**

(1) Choose appropriately a sequence $\{\mu_i\}_{i=1}^m$, with $\mu_i \in (0, 1]$ in ascending order, and $\mu_m = 1$;

(2) For each $i = 1, \dots, m$, solve problem $(\mathcal{P}_2)$ with $\mu = \mu_i$ by executing iterations (4.8)-(4.10)*;*

(3) Obtain the set of strictly stabilising feedback gains $\mathcal{K} := \{L_\infty(\mu_i)\}_{i=1}^m$ and sets $\{\bar{P}_\infty(\mu_i)\}_{i=1}^m$, $\{\hat{P}_\infty(\mu_i)\}_{i=1}^m$.

In Algorithm 4.1, there should be a sufficiently large number of elements in the sequence $\{\mu_i\}_{i=1}^m$ so that L_∞ can be adequately approximated on the intervals between consecutive points in this sequence. More importantly, μ_1 should be appropriately chosen close to 0 while ensuring that $L_\infty(\mu_1) \neq L_\infty(0)$ and $\bar{P}_\infty(\mu_1) \neq \bar{P}_\infty(0)$. Furthermore, step (2) can be warm-started by using $\bar{P}_\infty(\mu_i)$ and $\hat{P}_\infty(\mu_i)$ to initialise the iteration with weighting parameter μ_{i+1}, thereby reducing considerably the time required to solve $(\mathcal{P}_2)$ for each value of μ.

Remark 4.4. *The setting in Algorithm 4.1 that* $L_\infty(\mu_1) \neq L_\infty(0)$ *and* $\bar{P}_\infty(\mu_1) \neq \bar{P}_\infty(0)$ *ensures* $L_\infty(\mu) \neq L_\infty(0)\ \forall \mu \geq \mu_1$*. This relation obviously holds if* $L_\infty(0)$ *is not stabilising. For other cases, we prove this relation by contradiction. Supposing there exists a scalar* $\mu_2 \geq \mu_1$ *such that* $L_\infty(\mu_2) = L_\infty(0)$*, then* $\bar{P}_\infty(\mu_2) = \bar{P}_\infty(0)$*. By the monotonicity property of* $\bar{P}_\infty$ *that will be given in Lemma 4.5,* $\bar{P}_\infty(\mu_2) \succeq \bar{P}_\infty(\mu_1)$ *and therefore*

$$\bar{P}_\infty(0) \succeq \bar{P}_\infty(\mu_1). \tag{4.14}$$

In addition, by the optimality condition, it is clear that

$$\bar{P}_\infty(0) \preceq \bar{P}_\infty(\mu) \quad \forall \mu \in [0,1]. \tag{4.15}$$

Combining equations (4.14) *and* (4.15) *implies that* $\bar{P}_\infty(0) = \bar{P}_\infty(\mu_1)$*. This contradicts the requirement on* μ_1 *and therefore proves the statement* $L_\infty(\mu) \neq L_\infty(0)$ $\forall \mu \geq \mu_1$*.*

4.4 Properties of Steady-State Solutions

In this section, we derive properties of $\bar{P}_\infty(\mu)$ and $\hat{P}_\infty(\mu)$, namely that $\bar{P}_\infty(\mu)$ and $\hat{P}_\infty(\mu)$ are monotonic in μ and that $(1-\mu)\bar{P}_\infty(\mu) + \mu\hat{P}_\infty(\mu)$ is concave for $\mu \in (0,1)$, implying the Lipschitz continuity of $\bar{P}_\infty(\cdot)$ and $\hat{P}_\infty(\cdot)$ on $(0,1)$. These properties will be exploited in the design of gain selection methods in Section 5.4 and facilitate the proofs of convergence results, i.e. Lemma 5.3 and Theorem 5.5, in Chapter 5.

Lemma 4.5. *For all* μ_1, μ_2 *such that* $0 < \mu_1 \le \mu_2 \le 1$, $\bar{P}_\infty(\cdot)$, $\hat{P}_\infty(\cdot)$ *satisfy*

$$\bar{P}_\infty(\mu_1) \preceq \bar{P}_\infty(\mu_2),$$
$$\hat{P}_\infty(\mu_1) \succeq \hat{P}_\infty(\mu_2).$$

Lemma 4.6. *Let* $S(\mu) := (1-\mu)\bar{P}_\infty(\mu) + \mu\hat{P}_\infty(\mu)$, *then* $S(\cdot)$ *is concave on* $(0, 1)$.

The main result of this section is given next, which is established based on Lemmas 4.5 and 4.6.

Lemma 4.7. $\bar{P}_\infty(\cdot)$, $\hat{P}_\infty(\cdot)$ *are Lipschitz continuous on* $(0, 1)$.

4.5 Numerical Example

This section provides an example to demonstrate the numerical computation of the set $\mathcal{K}$. Considering the model parameters

$$A = \begin{bmatrix} 0.8207 & 0.04 \\ 0.0799 & 0.7808 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0454 & 0.0011 \\ 0.0022 & 0.0443 \end{bmatrix}, \quad x_0 = \begin{bmatrix} -1 \\ 3 \end{bmatrix},$$

and problem parameters $\gamma = 0.9$, $C = \left[\begin{smallmatrix} 0.3 & 0.15 \\ 0.1 & -0.1 \end{smallmatrix}\right]$, $Q = R = I$, and appropriately choosing a sequence, $\{\mu_i\}_{i=1}^m$, of weighting parameters with $m = 290320$ and $\mu_m = 1$, we solve coupled algebraic Riccati equations (4.12)-(4.13) approximately by executing DP iterations (4.8)-(4.10) with every weighting parameter in the sequence $\{\mu_i\}_{i=1}^m$. Setting a tolerance level $\alpha = 10^{-12}$, we say that $L_i(\mu)$, $\bar{P}_i(\mu)$ and $\hat{P}_i(\mu)$ have converged to $L_\infty(\mu)$, $\bar{P}_\infty(\mu)$ and $\hat{P}_\infty(\mu)$ respectively for a given value of μ if for some i

$$\max\left\{\frac{\|L_{i+1}(\mu) - L_i(\mu)\|_2}{\|L_i(\mu)\|_2}, \frac{\|\bar{P}_{i+1}(\mu) - \bar{P}_i(\mu)\|_2}{\|\bar{P}_i(\mu)\|_2}, \frac{\|\hat{P}_{i+1}(\mu) - \hat{P}_i(\mu)\|_2}{\|\hat{P}_i(\mu)\|_2}\right\} < \alpha.$$

Since their convergence is ensured under Assumption 4.1, it is not necessary to set a mamximum number of iterations. DP iterations are initialised with $\bar{P}_0(\mu) = C^\top C$, $\hat{P}_0(\mu) = Q$ when $\mu \leftarrow \mu_1$ and with $\bar{P}_0(\mu) = \bar{P}_\infty(\mu_i)$, $\hat{P}_0(\mu) = \hat{P}_\infty(\mu_i)$ when $\mu \leftarrow \mu_{i+1}$ $\forall i \ge 1$. This warm-starting considerably reduces computation time. As shown in Figure 4.3, when the DP recursion is executed with warm-starting the number of iterations required for convergence is reduced by at least 50% on average, comparing with that of the execution without warm-starting. Although the fixed points

$\bar{P}_\infty(\mu)$ and $\hat{P}_\infty(\mu)$ are functions of μ, implying that the value of μ has an indirect impact on the number of iterations required for convergence, we can see that the number of iterations required for convergence is more directly related to the initialisations, $\bar{P}_0(\mu)$ and $\hat{P}_0(\mu)$, of DP recursion. Note also that the number of iterations required for convergence sharply increases at some values of μ when the DP recursion is executed with warm-starting. This is probably because the sequence $\{\mu_i\}_{i=1}^m$ is chosen in such a way that its elements are grouped in different intervals, in which points are evenly distributed with small gaps, and there are relatively longer distances between adjacent intervals.

With sets $\{\bar{P}_\infty(\mu_i)\}_{i=1}^m$ and $\{\hat{P}_\infty(\mu_i)\}_{i=1}^m$, we plot a Pareto front in Figure 4.1, and plot values of $x_0^\top \bar{P}_\infty(\mu) x_0$ against μ and values of $x_0^\top \hat{P}_\infty(\mu) x_0$ against μ in Figures 4.2(a)-(b) respectively. It can be seen that the curves of objective functions $\sum_{i=0}^\infty \gamma^i \|C\bar{x}_i\|^2$ and $\sum_{i=0}^\infty \left(\|\bar{x}_i\|_Q^2 + \|\bar{u}_i\|_R^2\right)$ are in line with the monotonic properties of $\bar{P}_\infty(\cdot)$ and $\hat{P}_\infty(\cdot)$.

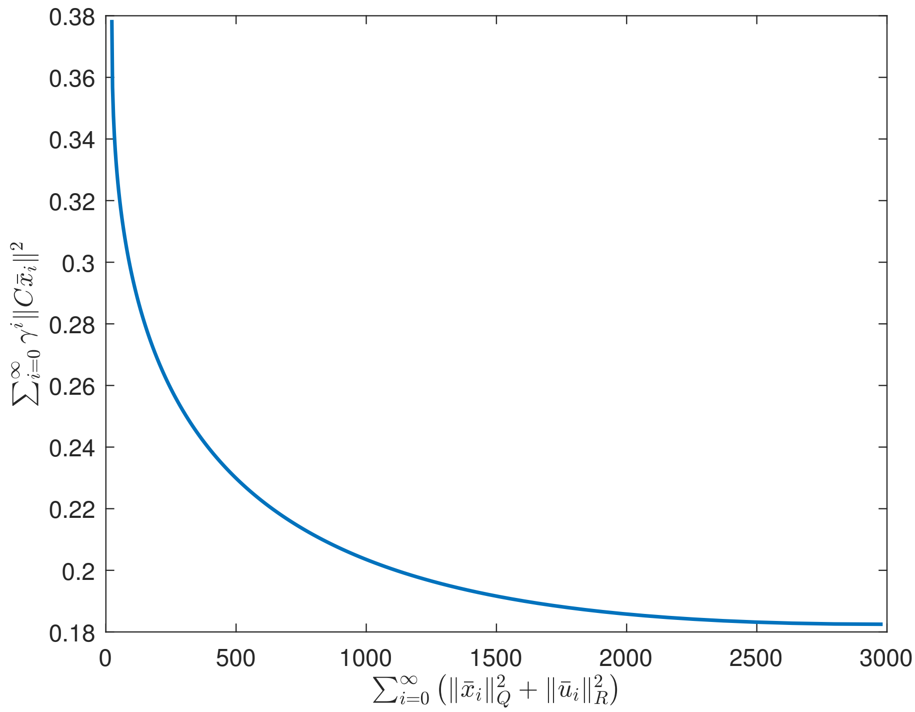


Figure 4.1: Pareto front

4.6 Conclusion

We use the weighted sum method to solve a convex MOP and this method is able to capture the complete Pareto front in this case. The resulting single objective problems are then solved via dynamic programming iterations that are guaranteed to converge, and their solutions yield strictly stabilising feedback gains. These gains will be used for online dynamic gain selection in Chapter 5 and they also coincide with the minimiser of a nonconvex problem, therefore providing an alternative approach for solving these problems exactly. Future work will search for conditions under which the feedback gain $L_\infty(0)$ is stabilising.

4.7 Proofs of Chapter 4

4.7.1 Proof of Lemma 4.5

A symmetric matrix Σ necessarily satisfies $\Sigma \preceq 0$ if $\mathrm{tr}(Z\Sigma) \leq 0$ for all $Z = Z^\top \succeq 0$. Therefore, the optimality of $\mathrm{tr}\Big[Z\Big((1-\mu)\bar{P}_\infty(\mu) + \mu\hat{P}_\infty(\mu)\Big)\Big]$ implies

$$
\begin{aligned}
(1-\mu_1)\bar{P}_\infty(\mu_1) + \mu_1\hat{P}_\infty(\mu_1) &\preceq (1-\mu_1)\bar{P}_\infty(\mu_2) + \mu_1\hat{P}_\infty(\mu_2), \\
(1-\mu_2)\bar{P}_\infty(\mu_2) + \mu_2\hat{P}_\infty(\mu_2) &\preceq (1-\mu_2)\bar{P}_\infty(\mu_1) + \mu_2\hat{P}_\infty(\mu_1)
\end{aligned}
$$

and, defining $\Delta\bar{P}$, $\Delta\hat{P}$ by $\Delta\bar{P} = \bar{P}_\infty(\mu_2) - \bar{P}_\infty(\mu_1)$, $\Delta\hat{P} = \hat{P}_\infty(\mu_2) - \hat{P}_\infty(\mu_1)$, it follows that

$$
\begin{aligned}
-(1-\mu_1)\Delta\bar{P} &\preceq \mu_1\Delta\hat{P}, \\
(1-\mu_2)\Delta\bar{P} &\preceq -\mu_2\Delta\hat{P}.
\end{aligned} \tag{4.16}
$$

Combining these inequalities yields $\Delta\bar{P} \succeq 0$ and $\Delta\hat{P} \preceq 0$.

4.7.2 Proof of Lemma 4.6

Let $0 < \mu_1 \leq \mu_2 \leq \mu_3 \leq 1$ with $\mu_3 - \mu_2 = \mu_2 - \mu_1$, and let $\Delta^2\bar{P} = \bar{P}_\infty(\mu_3) - 2\bar{P}_\infty(\mu_2) + \bar{P}_\infty(\mu_1)$, $\Delta^2\hat{P} = \hat{P}_\infty(\mu_3) - 2\hat{P}_\infty(\mu_2) + \hat{P}_\infty(\mu_1)$ and $\Delta^2 S = S(\mu_3) - 2S(\mu_2) + S(\mu_1)$. Then by definition we have $\Delta^2 S = (1-\mu_3)\Delta^2\bar{P} + \mu_3\Delta^2\hat{P} - (\mu_3-\mu_1)(\Delta\bar{P} - \Delta\hat{P})$, where

$\Delta\bar{P} = \bar{P}_\infty(\mu_2) - \bar{P}_\infty(\mu_1)$, $\Delta\hat{P} = \hat{P}_\infty(\mu_2) - \hat{P}_\infty(\mu_1)$. But the optimality property of $\mathrm{tr}\Big(ZS(\mu)\Big)$ implies $(1-\mu_3)\bar{P}_\infty(\mu_3) + \mu_3\hat{P}_\infty(\mu_3) \preceq (1-\mu_3)\hat{P}_\infty(\mu_2) + \mu_3\hat{P}_\infty(\mu_2)$, so

$$(1-\mu_3)\Big(\bar{P}_\infty(\mu_3) - \bar{P}_\infty(\mu_2)\Big) \preceq -\mu_3\Big(\hat{P}_\infty(\mu_3) - \hat{P}_\infty(\mu_2)\Big).$$

Combining this inequality with (4.16), we obtain

$$(1-\mu_3)\Delta^2\bar{P} + \mu_3\Delta^2\hat{P} - (\mu_3-\mu_1)(\Delta\bar{P} - \Delta\hat{P}) \preceq 0,$$

which implies $\Delta^2 S \preceq 0$ and hence $S(\mu)$ is midpoint concave for all $\mu \in (0,1]$. Furthermore, $S(\mu)$ is positive definite (bounded below) for all $\mu \in (0,1]$. Therefore, $S(\cdot)$ is continuous and hence concave on $(0,1)$ in the sense that $\mathrm{tr}(ZS(\cdot))$ is concave on $(0,1)$ for all $Z = Z^\top \succeq 0$ [121, Section 72].

4.7.3 Proof of Lemma 4.7

Let $\Sigma(\mu) := \gamma(1-\mu)\bar{P}_\infty(\mu) + \mu\hat{P}_\infty(\mu)$, and $\Phi(\mu) := A + BL_\infty(\mu)$ and consider the effect of an infinitesimal change $\delta\mu$ in the value of $\mu \in (0,1)$, where $\delta\mu \ll 1-\gamma$ and $\delta\mu \ll 1-\mu$. Let $\Delta\Phi$, $\Delta\Sigma$, ΔL, $\Delta\bar{P}$, $\Delta\hat{P}$ denote the corresponding changes in the values of Φ, Σ, L_∞, $\bar{P}_\infty$ and $\hat{P}_\infty$ respectively, so that $\Delta\Phi = \Phi(\mu+\delta\mu) - \Phi(\mu)$, $\Delta\Sigma = \Sigma(\mu+\delta\mu) - \Sigma(\mu)$, $\Delta L = L_\infty(\mu+\delta\mu) - L_\infty(\mu)$, $\Delta\bar{P} = \bar{P}_\infty(\mu+\delta\mu) - \bar{P}_\infty(\mu)$, and $\Delta\hat{P} = \hat{P}_\infty(\mu+\delta\mu) - \hat{P}_\infty(\mu)$. To simplify notation we omit the argument μ for the remainder of the proof (so that $\Phi = \Phi(\mu)$, $\Sigma = \Sigma(\mu)$, $L_\infty = L_\infty(\mu)$, etc.). Lemma 4.6 implies that $(1-\mu)\bar{P}_\infty + \mu\hat{P}_\infty$ is Lipschitz continuous on $(0,1)$, so $(1-\mu)\Delta\bar{P} + \mu\Delta\hat{P} - \delta\mu(\bar{P}_\infty(\mu+\delta\mu) - \hat{P}_\infty(\mu+\delta\mu)) = O(\delta\mu)$ and

$$\begin{aligned} O(\delta\mu) = (1-\mu)\Delta\bar{P} + \mu\Delta\hat{P} &= (\Phi+\Delta\Phi)^\top(\Sigma+\Delta\Sigma)(\Phi+\Delta\Phi) \\ &\quad - \Phi^\top\Sigma\Phi + \mu(L_\infty+\Delta L)^\top R(L_\infty+\Delta L) - \mu L_\infty^\top R L_\infty. \end{aligned} \tag{4.17}$$

But $L_\infty + \Delta L = -\Big[(\mu+\delta\mu)R + B^\top(\Sigma+\Delta\Sigma)B\Big]^{-1}B^\top(\Sigma+\Delta\Sigma)A$ implies

$$L_\infty + \Delta L = L_\infty - \Gamma^{-1}(B^\top\Delta\Sigma\Phi + RL_\infty\delta\mu) = L_\infty - \Gamma^{-1}B^\top\Delta\Sigma\Phi + O(\delta\mu), \tag{4.18}$$

where $\Gamma = (\mu+\delta\mu)R + B^\top(\Sigma+\Delta\Sigma)B$, and

$$\Phi + \Delta\Phi = \Phi - B\Gamma^{-1}B^\top\Delta\Sigma\Phi + O(\delta\mu). \tag{4.19}$$

Combining (4.18) and (4.19) with (4.17) and using $L_\infty = -(\mu R + B^\top \Sigma B)^{-1} B^\top \Sigma A$, we obtain

$$\begin{aligned} O(\delta\mu) &= \Phi^\top \Delta\Sigma\Phi + \Phi^\top \Delta\Sigma B \Gamma^{-1}(B^\top \Sigma B + \mu R - 2\Gamma)\Gamma^{-1} B^\top \Delta\Sigma\Phi \\ &\quad + \Phi^\top \Delta\Sigma B \Gamma^{-1} B^\top \Delta\Sigma B \Gamma^{-1} B^\top \Delta\Sigma\Phi \\ &= \Phi^\top \Delta\Sigma\Phi - \Phi^\top \Delta\Sigma B \Gamma^{-1} B^\top \Delta\Sigma\Phi - \Phi^\top \Delta\Sigma B \Gamma^{-1} R \Gamma^{-1} B^\top \Delta\Sigma\Phi\delta\mu, \end{aligned}$$

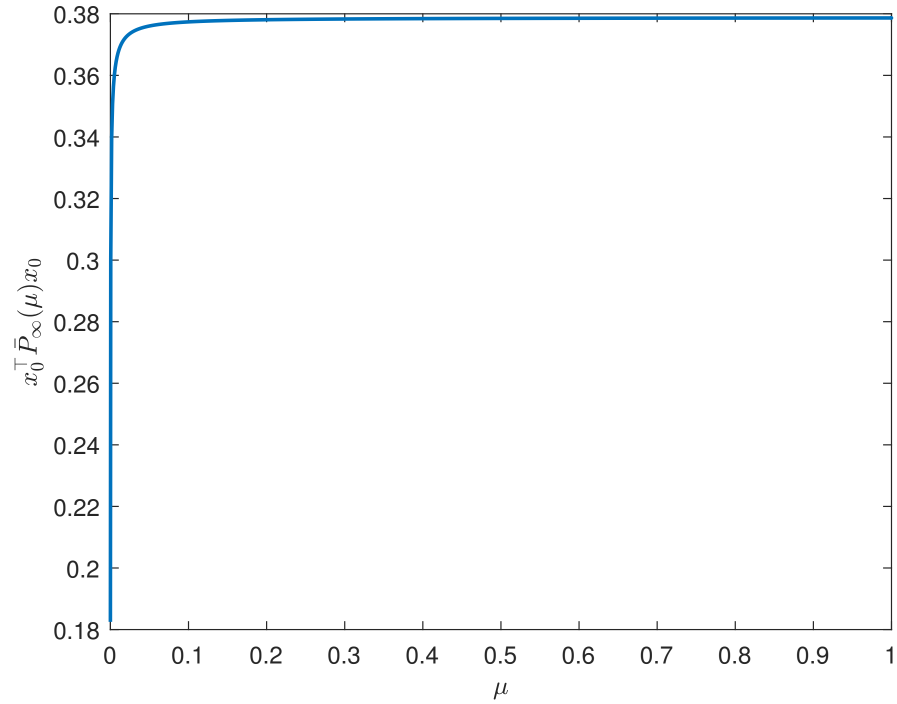
and hence

$$\Phi^\top \Delta\Sigma\Phi - \Phi^\top \Delta\Sigma B \Gamma^{-1} B^\top \Delta\Sigma\Phi = O(\delta\mu).$$

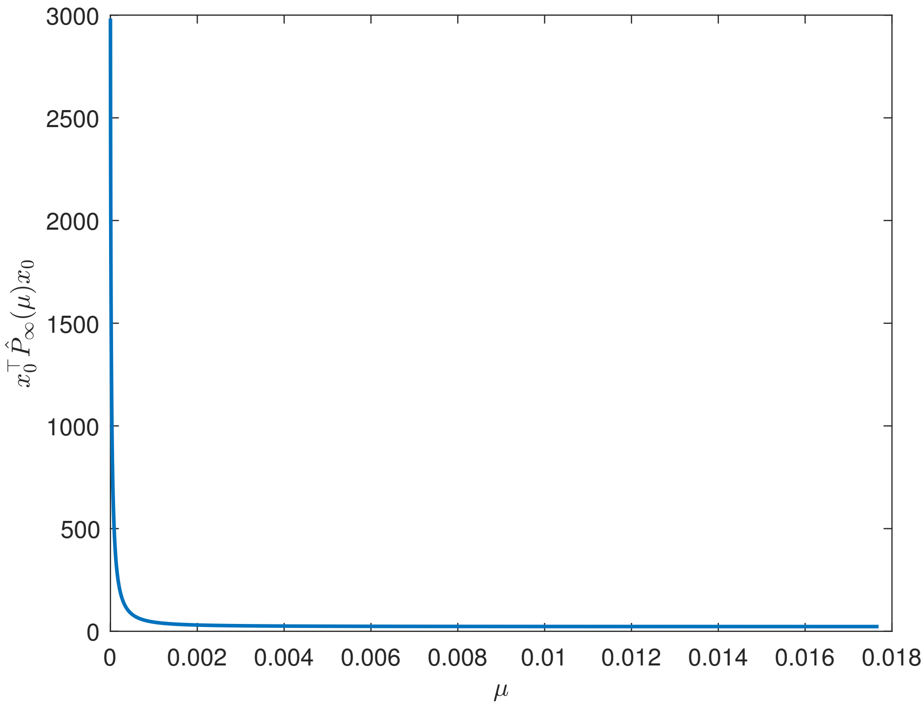
But $\Sigma = (1-\mu)\bar{P}_\infty + \mu\hat{P}_\infty - (1-\gamma)(1-\mu)\bar{P}_\infty$ implies $\Delta\Sigma = O(\delta\mu) - (1-\gamma)(1-\mu)\Delta\bar{P}$, so

$$(1-\gamma)(1-\mu)\Phi^\top \Big[\Delta\bar{P} + (1-\gamma)(1-\mu)\Delta\bar{P} B \Gamma^{-1} B^\top \Delta\bar{P}\Big]\Phi = O(\delta\mu).$$

This implies $\Phi^\top \Delta\bar{P}\Phi = O(\delta\mu)$ since $\delta\mu \ll 1-\gamma$, $\delta\mu \ll 1-\mu$ and $\Gamma \succeq 0$, and since $\Delta\bar{P} \succeq 0$ by Lemma 4.5. Furthermore, $\Delta\bar{P}$ is symmetric so we must have $\Delta\bar{P}\Phi = O(\delta\mu)$ and hence $\Delta\Sigma\Phi = O(\delta\mu)$. From (4.18)-(4.19) it follows that $\Delta L = O(\delta\mu)$ and $\Delta\Phi = O(\delta\mu)$, and hence the solutions, $\bar{P}_\infty(\cdot)$, $\hat{P}_\infty(\cdot)$ of the Lyapunov equations (4.12)-(4.13) are Lipschitz continuous on $(0, 1)$.



(a) Objective function $\sum_{i=0}^{\infty} \gamma^i \|C\bar{x}_i\|^2$



(b) Objective function $\sum_{i=0}^{\infty} \left(\|\bar{x}_i\|_Q^2 + \|\bar{u}_i\|_R^2\right)$

Figure 4.2: Pareto optimal values of objective functions in problem $(\mathcal{P}_2)$

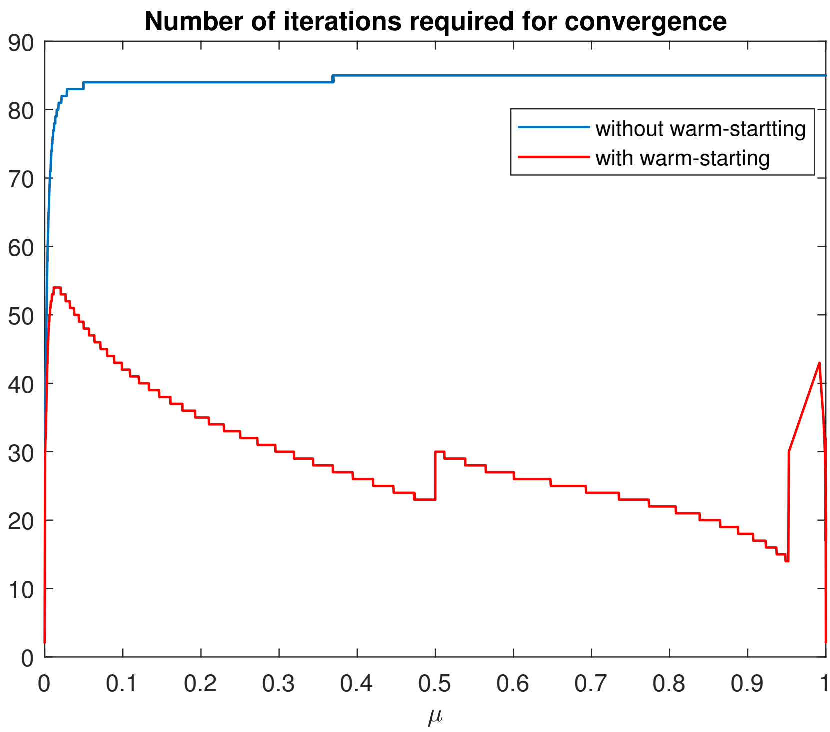


Figure 4.3: Number of iterations required for convergence

5
Dynamic Feedback Gain Selection

Contents

This chapter revisits a similar control problem to that of Chapter 3 and designs a model predictive control (MPC) law incorporating a dynamic feedback gain. The feedback gain is selected from the set $\mathcal{K}$ discussed in Chapter 4. We provide two methods for online gain selection based on minimising upper bounds of predicted costs, which incur only a tiny amount of extra computation. A computationally convenient MPC optimisation problem is formulated using Chebyshev's inequality, and we introduce an online constraint-tightening technique to ensure recursive feasibility. The closed loop system is guaranteed to satisfy the chance constraint and a quadratic stability condition. With dynamic feedback gain selection, the conservativeness of Chebyshev's inequality is mitigated and closed loop cost is reduced. Also, a larger set of feasible initial conditions can be obtained. A numerical example is given to show these properties.

5.1 Introduction

Stochastic MPC employs probabilistic descriptions of uncertainty and allows soft constraints, which are of practical relevance. In some applications, such as building temperature control [12] and portfolio management [13], it is tolerable if there are constraints being violated for a limited time period.

In order to provide guarantees of recursive feasibility and constraint satisfaction in closed loop operation with reduced conservativeness, online constraint-tightening techniques are proposed in [39, 40, 107]. These methods rely on knowledge of worst case disturbance bounds, and their degree of conservativeness increases as the disturbance bounds become more conservative. An adaptive approach is developed in [42], which aims to avoid this problem by introducing a scaling factor for tightening parameters. This factor is adapted online, making use of observations of past constraint violations. The authors show that the time average constraint violation rate converges in probability to a specified limit. For the case that disturbance bounds are known, hard constraints on control inputs can be incorporated. However, without bounded disturbance distributions or an assumption that the open loop system is stable (as in [122], for example), it is not possible to guarantee the satisfaction of hard constraints.

In contrast to existing work on stochastic MPC, this chapter considers linear discrete-time systems subject to possibly unbounded additive disturbances. We propose an MPC strategy incorporating a dynamic feedback gain to minimise a quadratic cost while satisfying a chance constraint. The constraint combines long-term and short-term considerations by imposing a bound on discounted violation probabilities accumulated over an infinite horizon. The chance constraint is reformulated via Chebyshev's inequality using knowledge of only the first and second moments of the disturbance input to obtain a convex optimisation problem.

This chapter reduces the conservativeness of Chebyshev's inequality by using a dynamic feedback gain in the definition of the MPC law. Alternatively, the affine disturbance feedback structure [83] can be used to define predicted control sequences, whereas it leads to higher computational complexity of online MPC optimisation problems. The first step of this proposed approach is to pre-compute the set, $\mathcal{K}$, of strictly stabilising feedback gains offline as introduced in Chapter 4. Then, for the online gain selection, two methods are designed to minimise upper bounds on the optimal predicted cost while retaining guarantees of recursive feasibility and computationally simple MPC optimisation. We show that the gain selection procedure can be configured so that the feedback gain converges almost surely to the unconstrained LQ-optimal, K_{LQ}, and the set of admissible initial conditions can be enlarged by choosing an appropriate initial gain. The MPC algorithm significantly improves closed loop performance in terms of the long-run expected average cost and reduces the conservativeness of closed loop constraint handling compared to MPC laws based on fixed feedback gains.

This chapter is organised as follows. The control problem is described and the controller structure is formulated in Section 5.2. Section 5.3 proposes an online constraint-tightening method for guaranteeing recursive feasibility. Online gain selection methods are proposed in Section 5.4 to choose feedback gains defining predicted control sequences. Section 5.5 summarises the proposed MPC algorithm and derives a bound on closed loop performance. In Section 5.6, the closed loop behaviour of the tightening parameters is analysed and constraint satisfaction is proved. Section 5.7 gives a numerical example illustrating the results obtained and this chapter is concluded in Section 5.8.

5.2 Problem Description

Consider an uncertain linear system

$$x_{k+1} = Ax_k + Bu_k + \omega_k, \tag{5.1}$$

where $x_k \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$ are the system state and control input respectively. The unknown disturbance input $\omega_k \in \mathbb{R}^{n_x}$ is independently and identically distributed (i.i.d.) with known first and second moments

$$\mathbb{E}\{\omega_k\} = 0, \quad \mathbb{E}\{\omega_k \omega_k^\top\} =: \Omega \succ 0. \tag{5.2}$$

The disturbance distribution may have infinite support, and we assume that measurements of the system state are available to the controller at each sample instant.

The system (5.1) is subject to the constraint

$$\sum_{k=0}^{\infty} \gamma^k \mathbb{P}\left\{\|Cx_k + Du_k\| \geq 1\right\} \leq e$$

for given $C \in \mathbb{R}^{n_c \times n_x}$, $D \in \mathbb{R}^{n_c \times n_u}$, a positive scalar e and a discount factor $\gamma \in (0,1)$. To simplify presentation, we set $D = 0$ for the remainder of this chapter, noting that all of the results given in Sections 5.3-5.6 apply to the case of non-zero D, and in the sequel the constraint

$$\sum_{k=0}^{\infty} \gamma^k \mathbb{P}\left\{\|Cx_k\| \geq 1\right\} \leq e \tag{5.3}$$

is considered. We refer to $\mathbb{P}\{\|Cx_k\| \geq 1\}$ as a *violation probability.*

In this chapter, we design a controller to solve the following problem

$$\begin{aligned} \min \ & \mathbb{E}\Big\{\sum_{k=0}^{\infty} \|x_k\|_Q^2 + \|u_k\|_R^2\Big\} \\ \text{s.t.} \ & (5.3). \end{aligned} \tag{5.4}$$

The weighting matrices in the cost function of problem (5.4) are assumed to satisfy $Q \succeq 0$ and $R \succ 0$.

Assumption 5.1. *(A, B) is controllable and $(A, Q^{\frac{1}{2}})$ is observable.*

As in Chapter 3, the discounting in this problem introduces a special feature that the probabilities of violating the condition $\|Cx_k\| < 1$ at time instants k nearer to the initial time are weighted more heavily than those in the far future. Costs defined over infinite horizons (with no discounting) over-emphasise steady state performance at the expense of performance during transients, and finite horizon costs are biased towards only the short-term performance. Therefore the discount factor allows for a trade-off between short-term and long-term behaviours and essentially determines how much priority the algorithm gives to constraint violations in the immediate future relative to those in the distant future.

We propose to solve problem (5.4) using a receding horizon approach wherein our control law is parameterised at each stage with a finite number of decision variables, and the constraint (5.3) is approximated conservatively using the two-sided Chebyshev inequality [112, Section V.7]. The resulting optimisation problem to be solved at each time step is then both finite dimensional and computationally tractable.

5.2.1 Predicted nominal control input and state sequences

Before deriving the finite horizon expressions of the cost and the constraint, this section defines predicted nominal control input and state sequences.

The sequence of nominal control inputs predicted at time k is given by

$$\bar{u}_{i|k} = K_k \bar{x}_{i|k} + c_{i|k}, \quad i = 0, \dots, N-1, \tag{5.5}$$

$$\bar{u}_{N+i|k} = K_k \bar{x}_{N+i|k}, \quad i = 0, 1, \dots, \tag{5.6}$$

where $\bar{x}_{i|k}$ is the i-step-ahead prediction of the nominal state given information at time k, that is, $\mathbb{E}_k\{x_{i|k}\} = \bar{x}_{i|k}$. The matrix-valued term K_k is a stabilising feedback gain that is selected online at time k from amongst the precomputed set $\mathcal{K}$ obtained by Algorithm 4.1 in Chapter 4. The selection procedure will be detailed in Section 5.4. After selecting a gain K_k, the *perturbation sequence* $\{c_{0|k}, c_{1|k}, \dots, c_{N-1|k}\}$ then constitutes the decision variables in the MPC optimisation problem to be solved at time k.

Given the predicted nominal control law (5.5)-(5.6), the predicted nominal state trajectory is given by $\bar{x}_{0|k} = x_k$ and

$$\bar{x}_{i|k} = \Phi_k^i \bar{x}_{0|k} + \sum_{j=0}^{i-1} \Phi_k^{i-1-j} B c_{j|k}, \qquad i = 1, \ldots, N, \tag{5.7}$$

$$\bar{x}_{N+i|k} = \Phi_k^i \bar{x}_{N|k}, \qquad i = 1, 2, \ldots, \tag{5.8}$$

where $\Phi_k := A + BK_k$. The covariance matrix, $X_{i|k}$, of the i-step-ahead predicted state is given by $X_{0|k} = 0$ and

$$X_{i|k} = \sum_{j=0}^{i-1} \Phi_k^j \Omega \left(\Phi_k^j\right)^\top, \quad i = 1, 2, \ldots. \tag{5.9}$$

We rewrite (5.7) in a compact form as

$$\bar{\mathbf{x}}_k = M_x \left(K_k\right) \bar{x}_{0|k} + M_c \left(K_k\right) \mathbf{c}_k, \tag{5.10}$$

$$\begin{bmatrix} \bar{x}_{1|k} \\ \bar{x}_{2|k} \\ \vdots \\ \bar{x}_{N|k} \end{bmatrix} = \begin{bmatrix} \Phi_k \\ \Phi_k^2 \\ \vdots \\ \Phi_k^N \end{bmatrix} \bar{x}_{0|k} + \begin{bmatrix} B & & \\ \Phi_k B & & \\ \vdots & \ddots & \\ \Phi_k^{N-1} B & \cdots & B \end{bmatrix} \begin{bmatrix} c_{0|k} \\ c_{1|k} \\ \vdots \\ c_{N-1|k} \end{bmatrix}, \tag{5.11}$$

and $M_x \left(K_k\right) \in \mathbb{R}^{Nn_x \times n_x}$, $M_c \left(K_k\right) \in \mathbb{R}^{Nn_x \times Nn_u}$. For simplicity we write these two matrices as M_x and M_c, with the understanding that they depend on K_k.

5.2.2 Online MPC optimisation

Based on predicted sequences defined in Section 5.2.1 and employing Chebyshev's inequality, we give finite horizon expressions of the cost and the constraint and formulate an MPC optimisation problem to be solved repeatedly online.

Minimising the predicted cost $\mathbb{E}_k\{\sum_{i=0}^{\infty} \|x_{i|k}\|_Q^2 + \|u_{i|k}\|_R^2\}$ at time k over the optimisation variable $\mathbf{c}_k$ is equivalent to minimising the cost defined in terms of the predicted nominal input sequence (5.5)-(5.6) and state trajectory (5.7)-(5.8) by

$$J(\bar{x}_{0|k}, K_k, \mathbf{c}_k) := \sum_{i=0}^{N-1} \left(\|\bar{x}_{i|k}\|_Q^2 + \|K_k \bar{x}_{i|k} + c_{i|k}\|_R^2\right) + \|\bar{x}_{N|k}\|_{P_k}^2. \tag{5.12}$$

Here $\|\bar{x}_{N|k}\|_{P_k}^2$ is the terminal cost and $P_k \succ 0$ is chosen as the solution of

$$P_k = Q + K_k^\top R K_k + \Phi_k^\top P_k \Phi_k. \tag{5.13}$$

Using (5.10)-(5.11), we rewrite (5.12) in a compact form as

$$J(\bar{x}_{0|k}, K_k, \mathbf{c}_k) = \begin{bmatrix} \bar{x}_{0|k} \\ \mathbf{c}_k \end{bmatrix}^\top W_2(K_k) \begin{bmatrix} \bar{x}_{0|k} \\ \mathbf{c}_k \end{bmatrix}, \tag{5.14}$$

where $W_2(K_k)$ is a function of K_k and its expression is omitted here for simplicity.

Approximating the LHS of (5.3) at time k by a direct application of the two-sided Chebyshev inequality, we obtain

$$\sum_{i=0}^{\infty} \gamma^i \Big(\|C\bar{x}_{i|k}\|^2 + \operatorname{tr}(C^\top C X_{i|k}) \Big) \leq \varepsilon_k. \tag{5.15}$$

Here ε_k replaces e in (5.3) as a threshold on the resulting constraint function, and it is a design parameter to be chosen at time k in some way such that we can ensure recursive feasibility of online MPC optimisations under possibly unbounded disturbances. By the following lemma, we show that the LHS of (5.15) is equivalent to a finite horizon expression.

Lemma 5.1. *Let $\widetilde{P}_k$ be the solution of*

$$\widetilde{P}_k = \gamma \Phi_k^\top \widetilde{P}_k \Phi_k + C^\top C. \tag{5.16}$$

Then

$$\sum_{i=N}^{\infty} \gamma^i \left\| C\bar{x}_{i|k} \right\|^2 = \gamma^N \left\| \bar{x}_{N|k} \right\|^2_{\widetilde{P}_k},$$
$$\sum_{i=0}^{\infty} \gamma^i \operatorname{tr}\Big(C^\top C X_{i|k} \Big) = \frac{\gamma}{1-\gamma} \operatorname{tr}\Big(\Omega \widetilde{P}_k \Big),$$

where $\bar{x}_{i|k}$ is given by (5.8) *for all $i \geq N$, $X_{0|k} = 0$ and $X_{i|k}$ is given by* (5.9) *for all $i \geq 1$.*

By Lemma 5.1, (5.15) is equivalent to

$$\sum_{i=0}^{N-1} \gamma^i \|C\bar{x}_{i|k}\|^2 + \gamma^N \|\bar{x}_{N|k}\|^2_{\widetilde{P}_k} + \frac{\gamma}{1-\gamma} \operatorname{tr}\Big(\Omega \widetilde{P}_k \Big) \leq \varepsilon_k, \tag{5.17}$$

where $\gamma^N \|\bar{x}_{N|k}\|^2_{\widetilde{P}_k}$ is the terminal term of the infinite discounted sum associated with predicted nominal states and $\frac{\gamma}{1-\gamma} \operatorname{tr}\Big(\Omega \widetilde{P}_k\Big)$ is the infinite discounted sum associated with covariance, which remains finite due to the discount factor. Using (5.10)-(5.11), we rewrite (5.17) in a more compact form as

$$\begin{bmatrix}\bar{x}_{0|k}\\ \mathbf{c}_k\end{bmatrix}^\top W_1\left(K_k\right)\begin{bmatrix}\bar{x}_{0|k}\\ \mathbf{c}_k\end{bmatrix} + \frac{\gamma}{1-\gamma}\operatorname{tr}\left(\Omega\widetilde{P}_k\right) \le \varepsilon_k, \tag{5.18}$$

where $W_1(K_k)$ is a function of K_k defined as

$$W_1\left(K_k\right) := \begin{bmatrix}C^\top C + M_x^\top H M_x & M_x^\top H M_c\\ M_c^\top H M_x & M_c^\top H M_c\end{bmatrix}, \tag{5.19}$$

and $H := \operatorname{diag}\{\gamma C^\top C, \ldots, \gamma^{N-1}C^\top C, \gamma^N \widetilde{P}_k\} \succeq 0$.

To summarise, the MPC optimisation solved at time k is

$$J^*(x_k, K_k) := \min_{\mathbf{c}_k}\{J(\bar{x}_{0|k}, K_k, \mathbf{c}_k) \mid (5.18) \text{ with } \bar{x}_{0|k} = x_k\}, \tag{5.20}$$

and its solution for any feasible x_k, K_k and ε_k is denoted $\mathbf{c}_k^*\left(x_k, K_k, \varepsilon_k\right)$. For simplicity we write this solution as $\mathbf{c}_k^*$, with the understanding that this vector depends on x_k, K_k and ε_k. The corresponding predicted nominal state trajectory is given by

$$\bar{x}_{i|k}^* = \Phi_k^i x_k + \sum_{j=0}^{i-1}\Phi_k^{i-1-j}Bc_{j|k}^*, \qquad i = 1, \ldots, N, \tag{5.21}$$

$$\bar{x}_{N+i|k}^* = \Phi_k^i \bar{x}_{N|k}^*, \qquad i = 1, 2, \ldots. \tag{5.22}$$

The MPC law at time k is defined by

$$u_k := K_k x_k + c_{0|k}^*, \tag{5.23}$$

and the closed loop system dynamics are given by

$$x_{k+1} = \Phi_k x_k + Bc_{0|k}^*(x_k, K_k, \varepsilon_k) + \omega_k, \tag{5.24}$$

where ω_k is the disturbance realisation at time k.

In the remainder of this chapter, we discuss how to choose ε_k so as to guarantee recursive feasibility in Section 5.3; how to select K_k from the set $\mathcal{K}$ (that is described in Chapter 4) to minimise predicted costs while retaining the recursive feasibility guarantee of online MPC optimisations in Section 5.4; and how the choices of P_k and $\widetilde{P}_k$ given in (5.13) and (5.16) allow for a guarantee of quadratic stability and satisfy constraint (5.3) respectively under the MPC law (5.23) in Section 5.5 and Section 5.6.

5.3 Recursive Feasibility

In this section, we use the same method for guaranteeing recursive feasibility of the MPC optimisation as proposed in Chapter 3, which does *not* rely on terminal constraints. For every time step $k > 0$, this approach uses the optimal sequence computed at time $k-1$ to determine a value of ε_k that is necessarily feasible at time k.

We use the notation $\widetilde{\mathbf{c}}_{k+1}$ to denote a time-shifted version of $\mathbf{c}_k^*$, defined by

$$\widetilde{\mathbf{c}}_{k+1} := E\mathbf{c}_k^*, \tag{5.25}$$

where E is the matrix such that $E\mathbf{c} = [c_1^\top, \ldots, c_{N-1}^\top, 0]^\top$ if $\mathbf{c} = [c_0^\top, \ldots, c_{N-1}^\top]^\top$.

Lemma 5.2. *The MPC optimisation* (5.20) *is recursively feasible if* ε_k *is defined at each time* $k = 1, 2, \ldots$ *as*

$$\varepsilon_k := \begin{bmatrix} x_k \\ \widetilde{\mathbf{c}}_k \end{bmatrix}^\top W_1\left(K_{k-1}\right) \begin{bmatrix} x_k \\ \widetilde{\mathbf{c}}_k \end{bmatrix} + \frac{\gamma}{1-\gamma} \operatorname{tr}\Big(\Omega \widetilde{P}_{k-1}\Big). \tag{5.26}$$

Equation (5.26) provides an explicit expression for ε_k for all $k > 0$ in terms of (5.21), (5.22) and disturbance realisations as

$$\begin{aligned} \varepsilon_k = \sum_{i=0}^{N-1} \gamma^i \Big\| C\Big(\bar{x}^*_{i+1|k-1} + \Phi^i_{k-1}\omega_{k-1}\Big)\Big\|^2 \\ + \gamma^N \Big\| \bar{x}^*_{N+1|k-1} + \Phi^N_{k-1}\omega_{k-1}\Big\|^2_{\widetilde{P}_{k-1}} + \frac{\gamma}{1-\gamma} \operatorname{tr}(\Omega \widetilde{P}_{k-1}). \end{aligned} \tag{5.27}$$

Here

$$\bar{x}^*_{i+1|k-1} + \Phi^i_{k-1}\omega_{k-1} := \bar{x}_{i|k}, \quad i = 0, \ldots, N \tag{5.28}$$

defines a feasible nominal state sequence predicted at time k, obtained by setting $\bar{x}_{0|k} = x_k$ and $\bar{u}_{i|k} = K_{k-1}\bar{x}_{i|k} + \widetilde{c}_{i|k}$. Note that in constructing this feasible sequence, we still use K_{k-1} and therefore $\widetilde{P}_{k-1}$ as the corresponding terminal matrix, and that the feedback gain is updated after the update of ε_k.

5.4 Dynamic Feedback Gain Selection

In this section, we provide two methods for dynamically selecting feedback gains from the set $\mathcal{K}$ obtained by Algorithm 4.1 in Chapter 4. These methods exploit monotonicity and continuity results proved in Section 4.4 of Chapter 4, and they are designed such that the recursive feasibility guarantee of problem (5.20) and a computationally simple online optimisation are retained. Both methods determine feedback gains that minimise upper bounds on the optimal predicted cost in (5.20). Method 1 requires less online computation and is equivalent to a binary search, while guaranteeing almost sure asymptotic convergence of K_k to K_{LQ}. Method 2 is more intuitive and in many cases gives better closed loop performance over short time horizons, but requires slightly more online computation since it requires the online evaluation of a function of $\mu \in \{\mu_i\}_{i=1}^m$.

5.4.1 Gain selection method 1

This section describes a method for selecting the gain K_k online and determines the properties of the sequence $\{\bar{\mu}_k\}_{k=0}^{\infty}$ generated by the following procedure.

Method 1. *At each time step $k = 1, 2, \dots$*

(1) Compute

$$\mathbf{c}^o(x_k) := \arg\min_{\mathbf{c}} \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix}^\top W_1\left(K_{k-1}\right) \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix} \tag{5.29}$$

where $W_1(K_{k-1})$ is defined in (5.19)*;*

(2) Compute

$$\begin{aligned} \bar{\mu}_k := \max \Big\{ \arg \max_{\mu \in \{\mu_i\}_{i=1}^m} \ & \frac{\gamma}{1-\gamma} \operatorname{tr}(\Omega \bar{P}_\infty(\mu)) \\ \textit{s.t.} \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}^\top W_1\big(K_{k-1}\big) \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix} & + \frac{\gamma}{1-\gamma} \operatorname{tr}(\Omega \bar{P}_\infty(\mu)) \le \varepsilon_k \Big\} \end{aligned} \tag{5.30}$$

where ε_k, $\bar{P}_\infty(\mu)$ are defined in (5.26), (4.12) *respectively;*

(3) Set $K_k = L_\infty(\bar{\mu}_k)$, $\widetilde{P}_k = \bar{P}_\infty(\bar{\mu}_k)$ and $P_k = \hat{P}_\infty(\bar{\mu}_k)$, where $L_\infty(\mu)$, $\bar{P}_\infty(\mu)$ and $\hat{P}_\infty(\mu)$, $\forall \mu \in \{\mu_i\}_{i=1}^m$, are obtained by Algorithm 4.1 in Chapter 4.

Step (1) determines the *perturbation sequence* $\mathbf{c}^o(x_k)$ that minimises the LHS of constraint (5.18) with $K_k = K_{k-1}$ and $\widetilde{P}_k = \widetilde{P}_{k-1}$. Step (2) then chooses $\bar{\mu}_k$ as the largest element of the sequence $\{\mu_i\}_{i=1}^m$ such that $\mathbf{c}^o(x_k)$ is feasible for the constraint (5.18). To show this, note that by combining (5.25), (5.26) and step (3) we obtain

$$\varepsilon_k = \begin{bmatrix} x_k \\ E\mathbf{c}^*_{k-1} \end{bmatrix}^\top W_1\Big(L_\infty(\bar{\mu}_{k-1})\Big) \begin{bmatrix} x_k \\ E\mathbf{c}^*_{k-1} \end{bmatrix} + \frac{\gamma}{1-\gamma}\operatorname{tr}(\Omega \bar{P}_\infty(\bar{\mu}_{k-1})).$$

From Lemma 4.5, (5.30) is therefore equivalent to

$$\bar{\mu}_k := \max_{\mu \in \{\mu_i\}_{i=1}^m} \mu \tag{5.31a}$$

$$\text{s.t.} \quad \frac{\gamma}{1-\gamma}\operatorname{tr}\Big(\Omega\left(\bar{P}_\infty(\mu) - \bar{P}_\infty(\bar{\mu}_{k-1})\right)\Big) \le \begin{bmatrix} x_k \\ E\mathbf{c}^*_{k-1} \end{bmatrix}^\top W_1\Big(L_\infty(\bar{\mu}_{k-1})\Big) \begin{bmatrix} x_k \\ E\mathbf{c}^*_{k-1} \end{bmatrix} - \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}^\top W_1\Big(L_\infty(\bar{\mu}_{k-1})\Big) \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}, \tag{5.31b}$$

where the RHS of (5.31b) is nonnegative due to the definition of $\mathbf{c}^o(x_k)$ in step (1). The aim of Method 1 is therefore to use the slack introduced into the constraint through the choice of $\mathbf{c}^o(x_k)$ in step (1) in order to maximise $\bar{\mu}_k$ in step (2) subject to

$$\begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}^\top W_1\Big(K_{k-1}\Big) \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix} + \frac{\gamma}{1-\gamma}\operatorname{tr}\Big(\Omega \widetilde{P}_k\Big) \le \varepsilon_k.$$

This implies that a *perturbation sequence* $\mathbf{c}'$ exists satisfying

$$\begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}^\top W_1\Big(K_{k-1}\Big) \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix} = \begin{bmatrix} x_k \\ \mathbf{c}' \end{bmatrix}^\top W_1\Big(K_k\Big) \begin{bmatrix} x_k \\ \mathbf{c}' \end{bmatrix}, \tag{5.32}$$

so that problem (5.20) remains feasible at time k. Therefore Method 1 retains the recursive feasibility guarantee. From (5.32), it also follows that $J(x_k, K_k, \mathbf{c}')$ is an upper bound on the optimal value of the cost in (5.20) at time k, and from (5.31a) and Lemma 4.5 it follows that $\bar{\mu}_k$ defined in step (2) minimises the trace of the terminal matrix $P_k = \hat{P}_\infty(\bar{\mu}_k)$ in this cost.

By Lemma 4.5, step (2) is simply a binary search to determine the largest $\mu \in \{\mu_i\}_{i=1}^m$ satisfying the constraint in (5.30). Since the values of $\operatorname{tr}(\Omega \bar{P}_\infty(\mu_i))$ can be calculated offline, Method 1 can be implemented very efficiently. Note that the initial value $\bar{\mu}_0$ is not determined by Method 1, and we therefore use other means of choosing

$\bar{\mu}_0 \in \{\mu_i\}_{i=1}^m$ and $K_0 = L_\infty(\bar{\mu}_0)$ to make the MPC optimisation (5.20) initially feasible (assuming such a $\bar{\mu}_0$ exists); this is discussed in Section 5.7, Simulation D.

In the remainder of this section, we analyse the properties of the sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ generated by Method 1. Since $\bar{\mu}_{k-1}$ is a feasible solution to problem (5.30) at time k, which implies $\bar{\mu}_{k-1} \le \bar{\mu}_k$, and since $\bar{\mu}_k$ is upper bounded by 1 for all k, the sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ generated by Method 1 is monotonically non-decreasing and convergent. To derive a stronger convergence result, we make a simplifying assumption.

Assumption 5.2. *The maximisation* (5.30) *can be solved with the optimisation variable μ varying continuously in the interval* $[\bar{\mu}_0, 1]$ *rather than being constrained to the discrete set $\{\mu_i\}_{i=1}^m$.*

The purpose of Assumption 5.2 is to ensure that $\bar{\mu}_k > \bar{\mu}_{k-1}$ whenever the RHS of (5.31b) is positive, and hence that $\bar{\mu}_k$ does not converge to a value less than 1 as a result of the constraint $\mu \in \{\mu_i\}_{i=1}^m$ in step (2).

Lemma 5.3. *Provided that Method 1 is used for time step $k = 1, 2, \dots$ and that the Assumption 5.2 holds, the equality $\bar{\mu}_{k-1} = \bar{\mu}_k (\neq 1)$ only holds for a finite number of consecutive time steps if* $\operatorname{rank}(CB) = n_u$.

We next derive a stochastic convergence result requiring less restrictive assumptions than Lemma 5.3. Considering the RHS of (5.31b), we can see that, by definition, $W_1(K_{k-1}) = W_1(L_\infty(\bar{\mu}_{k-1}))$ is at least positive semidefinite. Hence the minimisation in step (1) is well-defined and $\mathbf{c}^o(x_k)$ can be defined uniquely as

$$\mathbf{c}^o(x_k) = -\left(W_{cc}(\bar{\mu}_{k-1})\right)^\dagger W_{cx}(\bar{\mu}_{k-1}) x_k.$$

Here $W_{cc}(\mu)$, $W_{cx}(\mu)$ are blocks of $W_1\Big(L_\infty(\mu)\Big)$ in the partition (5.19) such that $W_{cc} = M_c^\top H M_c$, $W_{cx} = M_c^\top H M_x$, and $W_{cc}^\dagger$ is the pseudoinverse of W_{cc}. Since $x_k = \bar{x}^*_{1|k-1} + \omega_{k-1}$, it follows that the RHS of (5.31b) is equal to

$$\left\| z_{k-1} + \Big(W_{cc}(\bar{\mu}_{k-1})\Big)^\dagger W_{cx}(\bar{\mu}_{k-1}) \omega_{k-1} \right\|^2_{W_{cc}(\bar{\mu}_{k-1})}, \tag{5.33}$$

where $z_{k-1} := E\mathbf{c}^*_{k-1} + \Big(W_{cc}(\bar{\mu}_{k-1})\Big)^\dagger W_{cx}(\bar{\mu}_{k-1}) \bar{x}^*_{1|k-1}$.

Lemma 5.4. *There exist $\delta > 0$ and $p_\delta > 0$ such that*

$$\inf_{\substack{z\in\mathbb{R}^{Nn_u}\\ \mu\in[\bar{\mu}_0,1]}} \mathbb{P}\Big\{\Big\|z + \Big(W_{cc}(\mu)\Big)^\dagger W_{cx}(\mu)\omega_k\Big\|^2_{W_{cc}(\mu)} \geq \delta\Big\} \geq p_\delta. \tag{5.34}$$

This lemma implies that the RHS of (5.31b) takes positive values with a non-zero probability, which thus results in an increase of $\bar{\mu}_k$ from $\bar{\mu}_{k-1}$ under Assumption 5.2.

Based on Lemma 5.4, we next conclude that K_k converges almost surely to K_{LQ}.

Theorem 5.5. *Under Assumption 5.2, $\bar{\mu}_k \to 1$ as $k \to \infty$ with probability 1.*

5.4.2 Gain selection method 2

We first give a monotonicity result that can be used to prove the order-preserving property of Riccati Difference Equations (RDE) [123], which is exploited in Method 2.

Lemma 5.6. *$V(x_0, \check{P}_1) \leq V(x_0, \check{P}_2)$ $\forall x_0 \in \mathbb{R}^{n_x}$ if $\check{P}_1 \preceq \check{P}_2$, where*

$$V(x_0, \check{P}) := \min_{u_0,\dots,u_{N-1}} \Big\{\widetilde{V}(x_0, \mathbf{u}, \check{P}) \text{ s.t. } x_{i+1} = Ax_i + Bu_i,\ i = 0, \dots, N-1\Big\}, \tag{5.35}$$

and $\widetilde{V}(x_0, \mathbf{u}, \check{P}) := \sum_{i=0}^{N-1}\Big(\|x_i\|^2_Q + \|u_i\|^2_R\Big) + \|x_N\|^2_{\check{P}}$, $\mathbf{u} := [u_0^\top, \dots, u_{N-1}^\top]^\top$ for $Q, R, \check{P} \succeq 0$.

Proof. Let $\mathbf{u}^*(\check{P}) := \arg\min_{\mathbf{u}}\Big\{\widetilde{V}(x_0, \mathbf{u}, \check{P}) \text{ s.t. } x_{i+1} = Ax_i + Bu_i,\ i = 0, \dots, N-1\Big\}$, then by the ordering between $\check{P}_1$ and $\check{P}_2$ and optimality, we have $V(x_0, \check{P}_2) = \widetilde{V}\Big(x_0, \mathbf{u}^*(\check{P}_2), \check{P}_2\Big) \geq \widetilde{V}\Big(x_0, \mathbf{u}^*(\check{P}_2), \check{P}_1\Big) \geq \widetilde{V}\Big(x_0, \mathbf{u}^*(\check{P}_1), \check{P}_1\Big) = V(x_0, \check{P}_1)$. □

Remark 5.7. *Note that Lemma 5.6 holds if Q and R are time-varying. Let $\check{P}_1', \check{P}_2'$ be the matrices satisfying $V(x_0, \check{P}_1) = x_0^\top \check{P}_1' x_0$ and $V(x_0, \check{P}_2) = x_0^\top \check{P}_2' x_0$ for all $x_0 \in \mathbb{R}^{n_x}$, then Lemma 5.6 implies $\check{P}_1' \preceq \check{P}_2'$. Therefore the monotonicity property of RDE can be shown by solving (5.35) using DP recursion with $N = 1$ sequentially backwards in time. This property is also proved in [102, 124] with the assumption that B has full rank, which is not required here.*

We give Method 2 as follows.

Method 2. *At each time step $k = 1, 2, \dots$*

(1) For all $\mu \in (0,1]$, *define*

$$\underline{\mathbf{c}}(x_k,\mu) := \arg\min_{\mathbf{c}} \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix}^\top W_1\left(L_\infty(\mu)\right) \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix}, \tag{5.36}$$

where W_1 *is defined in* (5.19)*;*

(2) Compute

$$\begin{aligned}\tilde{\mu} := \max\Big\{ \arg\max_{\mu\in\{\mu_i\}_{i=1}^m} & \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k,\mu) \end{bmatrix}^\top W_1(L_\infty(\mu)) \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k,\mu) \end{bmatrix} + \star \\ s.t. & \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k,\mu) \end{bmatrix}^\top W_1(L_\infty(\mu)) \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k,\mu) \end{bmatrix} + \frac{\gamma}{1-\gamma}\operatorname{tr}(\Omega \bar{P}_\infty(\mu)) \le \varepsilon_k \Big\},\end{aligned} \tag{5.37}$$

where $\star$ *here denotes* $\frac{\gamma}{1-\gamma}\operatorname{tr}(\Omega\bar{P}_\infty(\mu))$*;*

(3) Compute

$$\begin{aligned}\bar{\mu}_k := \max\Big\{ \arg\min_{\mu\in\{\mu_i\}_{i=1}^m} & \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k,\mu) \end{bmatrix}^\top W_2(L_\infty(\mu)) \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k,\mu) \end{bmatrix} \\ & s.t.\ \tilde{\mu} \ge \mu \ge \bar{\mu}_{k-1} \Big\},\end{aligned} \tag{5.38}$$

where W_2 *is defined in* (5.14)*;*

(4) Set $K_k = L_\infty(\bar{\mu}_k)$, $\widetilde{P}_k = \bar{P}_\infty(\bar{\mu}_k)$ *and* $P_k = \hat{P}_\infty(\bar{\mu}_k)$, *where* $L_\infty(\mu)$, $\bar{P}_\infty(\mu)$ *and* $\hat{P}_\infty(\mu)$, $\forall \mu \in \{\mu_i\}_{i=1}^m$, *are obtained by Algorithm 4.1 in Chapter 4.*

Comparing step (1) of Methods 1 and 2, here $L_\infty(\mu)$ replaces K_{k-1} and $\underline{\mathbf{c}}(x_k,\mu)$ is a function of both x_k and μ. Recalling the equivalence between (5.17) and (5.18), we have that

$$\begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix}^\top W_1\left(L_\infty(\mu)\right) \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix} = \sum_{i=0}^{N-1} \gamma^i \|C\bar{x}_{i|k}\|^2 + \gamma^N \|\bar{x}_{N|k}\|^2_{\bar{P}_\infty(\mu)},$$

where $\bar{x}_{i|k}$ evolves according to $\bar{\mathbf{x}}_k = M_x\left((L_\infty(\mu)\right) x_k + M_c\left((L_\infty(\mu)\right)\mathbf{c}$. Therefore, by Lemmas 4.5 and 5.6, the objective in the maximisation (5.37) is monotonic in μ and step (2) can be performed as a binary search. The constraint in (5.37) makes the constraint (5.18) as tight as possible with $K_k = L_\infty(\tilde{\mu})$, and the constraints of (5.38) imply that (5.20) is recursively feasible with Method 2. Note that the initial value $\bar{\mu}_0$ is not determined by Method 2. Therefore we choose $\bar{\mu}_0 \in \{\mu_i\}_{i=1}^m$ and $K_0 = L_\infty(\bar{\mu}_0)$ to make the MPC optimisation (5.20) initially feasible (if such $\bar{\mu}_0$ exists), as discussed in Section 5.7, Simulation D.

Comparing with Method 1, both terms appearing on the LHS of the constraint in problem (5.37) are monotonic in μ, whereas only second term on the LHS of the constraint in problem (5.30) increases with μ and the first term is fixed. It follows that $\bar{\mu}_k$ obtained by Method 1 is necessarily greater than or equal to $\tilde{\mu}$ defined in (5.37), and hence it is also necessarily greater than or equal to $\bar{\mu}_k$ obtained by Method 2. The sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ generated by Method 1 is therefore likely to converge more quickly to 1. On the other hand, since its objective function has more degrees of freedom to vary with μ, it is likely that problem (5.38) results in a smaller upper bound on the optimal predicted cost (5.20) when a relatively large prediction horizon N is chosen and hence Method 2 provides better closed loop performance over the time period required for the sequence $\{K_k\}_{k=0}^\infty$ determined by Method 1 to converge to K_{LQ}. Over a longer time interval however, Method 1 is likely to perform better due to its earlier convergence to the LQ-optimal feedback gain K_{LQ}. These observations are supported by the numerical example in Section 5.7.

Based on this discussion, we provide the following guideline for choosing the online gain selection method. If the discount factor γ in (5.3) is close to 1, indicating that performance over a long horizon is a priority, then Method 1 should be chosen. Alternatively if γ is close to 0 and performance in the immediate future is important, then Method 2 is preferable. Finally, we note that Method 2 requires more computation than Method 1 since (5.38) (the objective of which is not necessarily monotonic in μ) requires an additional search over $\{\mu_i\}_{i=1}^m$.

Remark 5.8. *Since $\bar{\mu}_{k-1}$ is a feasible solution to the optimisation problem* (5.38) *at time k and $\bar{\mu}_k$ is upper bounded by* 1 *for all k, the sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ generated by Method 2 is monotonically non-decreasing and convergent.*

Several factors influence the implementation of online gain selection methods. The choice of the sequence $\{\mu_i\}_{i=1}^m$ can affect convergence of $\{\bar{\mu}_k\}_{k=0}^\infty$ and hence closed loop performance, since large gaps between successive elements of the sequence $\{\mu_i\}_{i=1}^{m}$ reduce the likelihood of convergence $\bar{\mu}_k \to 1$. Therefore it is desirable to choose m to be as large as possible, subject to offline computation and online storage constraints. To ensure a large feasible set and to steer the closed loop system away from feedback gains K_k that give worse closed loop performance, it is desirable to choose μ_1 close to 0 with $\mu_{i+1} - \mu_i$ increasing for larger values of i. Alternatively, if the initial conditions of the MPC problem are known offline when the set $\mathcal{K}$ is generated, then it is obviously advantageous to set μ_1 equal to a value $\bar{\mu}_0$ that makes the MPC optimisation initially feasible.

5.5 MPC Algorithm with Gain Selection and Stability Condition Revisited

In this section, we summarise an MPC algorithm and prove that the closed loop system satisfies a quadratic stability condition.

Algorithm 5.1. *(SMPC Algorithm) At each time-step $k = 0, 1, \dots$:*

(i) Measure x_k;

(ii) If $k > 0$, compute ε_k using (5.26) *and determine K_k using Method 1 or 2;*

(iii) Solve the quadratically constrained quadratic programming (QCQP) problem (5.20)*;*

(iv) Apply the control law (5.23).

We choose $\varepsilon_0 = e$ in step (ii) as will be explained in Section 5.6. In step (iii), the MPC optimisation can be solved efficiently since there is only one quadratic constraint in (5.20), for example using a second-order conic program (SOCP) solver or using the algorithm proposed in [125], which is based on the Newton-Raphson method.

Theorem 5.9. *Given initial feasibility at time $k = 0$, by using Algorithm 5.1, the closed loop system satisfies the quadratic stability condition*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\{\|x_k\|_Q^2 + \|u_k\|_R^2\} \leq \lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathrm{tr}\Big(\Omega \mathbb{E}\{P_k\}\Big) \leq \mathrm{tr}\Big(\Omega P_0\Big) \tag{5.39}$$

provided P_k satisfies (5.13)*.*

5.6 Constraint Satisfaction

This section proves constraint satisfaction in closed loop by analysing the behaviour of the sequence $\{\varepsilon_k\}_{k=0}^{\infty}$.

Using a similar approach to the analysis performed in Section 3.5 of Chapter 3, we first provide a result describing the relationship between ε_k and the expected value of ε_{k+1} for the closed loop system.

Lemma 5.10. *If ε_k is defined by* (5.26) *at all times $k \geq 1$ and $\widetilde{P}_k$ satisfies* (5.16) *for all $k \geq 0$, then in closed loop operation under Algorithm 5.1 we have*

$$\gamma \mathbb{E}_k\{\varepsilon_{k+1}\} \leq \varepsilon_k - \|Cx_k\|^2, \ \forall k \geq 0. \tag{5.40}$$

The main result of this section is given next, which is developed based on Lemma 5.10.

Theorem 5.11. *The closed loop system under Algorithm 5.1 satisfies the constraint* (5.3) *if $\varepsilon_0 = e$.*

Proof. It is similar to the proof of Theorem 3.5 in Chapter 3 and therefore is omitted here. □

5.7 Numerical Example

This section describes a numerical example to illustrate the following points: (i) the closed loop system (5.24) satisfies the quadratic stability condition (5.39) and the constraint (5.3) when Algorithm 5.1 is used without dynamic gain selection or with either gain selection Method 1 or Method 2; (ii) the degree of conservativeness of Chebyshev's inequality is mitigated by using online gain selection procedures in the sense that the long-run expected average costs are improved and the observed constraint violation rates are closer to the imposed limit; (iii) gain selection Method 2 provides better closed loop performance over short time intervals than Method 1; (iv) K_k converges to K_{LQ} with high probability for Method 1; (v) the set of feasible initial conditions is enlarged with an appropriate initial feedback gain. We also discuss the computation times required by the online MPC optimisation problem (5.20) and the gain selection methods.

We consider the discrete-time linearised model derived from the continuous-time model of a coupled-tank system [126] with a sampling interval of 0.05 sec. It has model parameters

$$A = \begin{bmatrix} 0.8207 & 0.04 \\ 0.0799 & 0.7808 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0454 & 0.0011 \\ 0.0022 & 0.0443 \end{bmatrix},$$

and disturbance ω_k having a Laplace distribution with zero mean and covariance $\Omega = I$. The constraint (5.3) is defined by $\gamma = 0.9$, $e = 1.5$ and $C = \left[\begin{smallmatrix} 0.3 & 0.15 \\ 0.1 & -0.1 \end{smallmatrix}\right]$, and the weighting matrices in the cost of problem (5.4) are given by $Q = R = I$. We choose a prediction horizon $N = 10$, and the initial value for ε_k is $\varepsilon_0 = e = 1.5$. We appropriately choose a sequence $\{\mu_i\}_{i=1}^m$ with $m = 290320$, where $0 < \mu_i \leq 1$ for all i and $\mu_m = 1$. Using this sequence, we generate the set of feedback gains $\mathcal{K} = \{L_\infty(\mu_i)\}_{i=1}^m$ offline and store the sets of matrices $\{\bar{P}_\infty(\mu_i)\}_{i=1}^m$, $\{\hat{P}_\infty(\mu_i)\}_{i=1}^m$. We choose $\bar{\mu}_0 = 10^{-15}$ so that

$$K_0 = \begin{bmatrix} -18.0749 & -0.4626 \\ -0.9251 & -17.6123 \end{bmatrix},$$

and $\Phi_0 = A + BK_0$ is strictly stable. We run four sets of simulations, all using the initial condition $x_0 = [-1, 3]^\top$ and the same sequences of disturbances. Note that $K_k = K_{LQ}$ is infeasible for the online MPC optimisation (5.20) at time $k = 0$.

Simulation A (demonstrating (i) and (ii)): To estimate empirically the average cost, denoted J_{average}, and the discounted sum of *violation probabilities*, denoted $P_{\text{violation}}$, we run 10^4 simulations, each of which has a length of 10^4 time steps, using Algorithm 5.1 with fixed $K_k = K_0$ $\forall k \geq 0$, with gain selection Method 1, and with gain selection Method 2, respectively. We compute the mean value of stage costs over these 10^4 simulations and count the number of violations in each simulation up to 150 time steps. Simulation results are summarised in Table 5.1.

	$K_k = K_0$	Method 1	Method 2	
J_{average}	639.9	8.4	96.2	$\text{tr}(\Omega P_0) = 640.0$
$P_{\text{violation}}$	0.448	0.914	0.769	$e = 1.5$

Table 5.1: Average costs and violation rates for Simulation A

Table 5.1 confirms that the three empirical cost estimates agree with the bound (5.39) and the three estimates for the discounted sum of *violation probabilities* are all smaller than $e = 1.5$, implying that constraint (5.3) is satisfied. Note that the cost estimates decrease more slowly as the simulation horizon length continues to increase, and that $\gamma^{150} \approx 1.37 \times 10^{-7}$ so the number of violations occurring at time steps $k > 150$ has negligible effect on the estimate of $P_{\text{violation}}$. Also, cost estimates obtained using the dynamic gain selection methods are considerably smaller than that obtained using a fixed feedback gain. Estimates for $P_{\text{violation}}$ are much closer to the maximum allowed level $e = 1.5$ when gain selection methods are used. More specifically, the estimate obtained using gain selection Method 1 is more than double the estimate obtained with a fixed feedback gain. Hence these results show that the conservativeness of Chebyshev's inequality is mitigated by dynamic gain selection. Moreover, the cost estimate obtained using Method 1 is smaller than that obtained using Method 2 since Method 1 achieves convergence of $\bar{\mu}_k$ to 1 (and hence convergence of K_k to K_{LQ}) earlier. This observation supports the statement in Section 5.4 that Method 1 is likely to provide better closed loop performance over longer time periods.

Simulation B (demonstrating (iii)): We run 10^4 simulations, each of which has a length of 40 time steps, using Algorithm 5.1 with fixed $K_k = K_0 \; \forall k \geq 0$, with gain selection Method 1, and with gain selection Method 2, respectively. We compute the mean value of stage costs over these 10^4 simulations, as shown in Table 5.2.

It is clear from Table 5.2 that Method 2 gives a smaller average cost estimate, which is in agreement with the statement in Section 5.4 that Method 2 is preferable

	$K_k = K_0$	Method 1	Method 2
J_{average}	807.2	575.9	500.8

Table 5.2: Average costs for Simulation B

if short-term performance is prioritised. This is because $\bar{\mu}_k$ generated by Method 1 generally does not converge to 1 within 40 time steps in this set of simulations, so the average costs are dominated by transient behaviours. We show an example of sequences $\{\bar{\mu}_k\}_{k=0}^{\infty}$ generated by Methods 1 and 2 in Figure 5.2, where Figure 5.2(b) plots the two sequences up to time $k = 150$ and Figure 5.2(a) zooms in on the initial 40 time steps. These two figures show that $\bar{\mu}_k$ obtained from Method 1 is greater than or equal to that obtained from Method 2 at all times, which is in agreement with the analysis in Section 5.4.

Simulation C (demonstrating (iv)): We run 10^4 simulations, each of which has a length of 200 time steps. When Method 1 is used with Algorithm 5.1, $\bar{\mu}_{200} = 1$ is obtained for every simulation, implying that K_k converges to K_{LQ} in every simulation. On the other hand, when Method 2 is used, the mean value of $\bar{\mu}_{200}$ over these 10^4 simulations is 0.8578.

Simulation D (demonstrating (v)): Minimising the LHS of (5.18) over $\mathbf{c}_k$ yields the largest set of feasible initial conditions

$$\mathcal{X}_0(\mu) := \Big\{x | x^\top \Big(W_{xx}(\mu) - W_{xc}(\mu)W_{cc}^\dagger(\mu)W_{cx}(\mu)\Big)x + \frac{\gamma}{1-\gamma}\,\mathrm{tr}(\Omega \bar{P}_\infty(\mu)) \leq \varepsilon_0\Big\},$$

where $W_{xx}(\mu)$, $W_{xc}(\mu)$, $W_{cx}(\mu)$, $W_{cc}(\mu)$ are the blocks of $W_1\Big(L_\infty(\mu)\Big)$ in (5.19). The feasible set $\mathcal{X}_0$ is plotted for μ taking values of $\mu_1 = 10^{-15}$, $\mu_2 = 10^{-4}$, $\mu_3 = 2.5\times 10^{-4}$ in Figure 5.1. Clearly the feasible set is enlarged as μ is reduced.

Computation times: Simulations are run in MATLAB R2019a on a computer with 2.20GHz Intel Core i7-8750H CPU and 16GB RAM, and the online MPC optimisation (5.20) is solved using the root-finding algorithm proposed in [125]. Each online MPC optimisation is solved within 1 millisecond (ms), with an average time of 0.35 ms. Gain selection Method 1 requires less than 0.003 ms at each iteration of Algorithm 5.1, which represents a tiny fraction of the time needed to solve an MPC optimisation. Gain selection Method 2 requires less than 0.05 ms at most

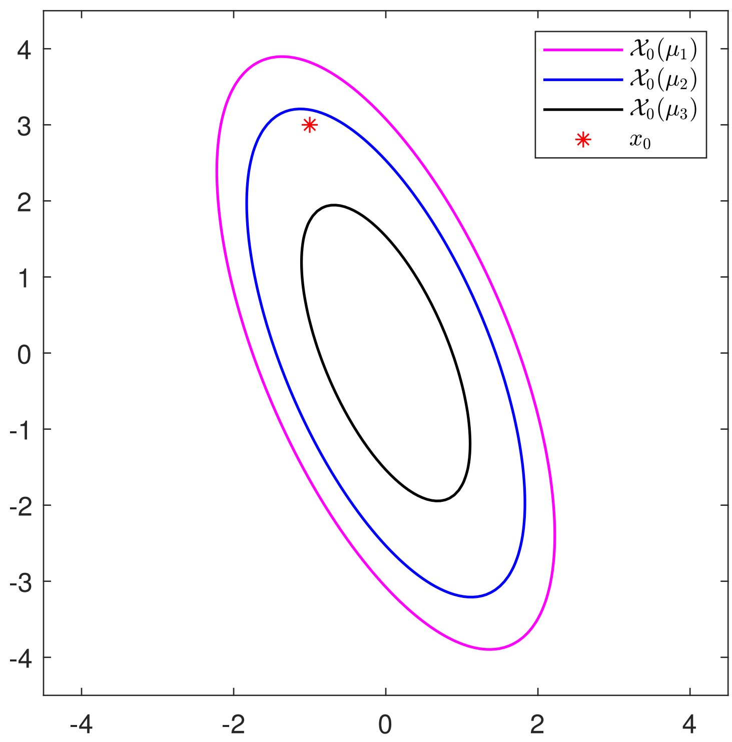


Figure 5.1: Largest feasible sets of initial conditions

time steps, which remains a small fraction of the time needed to solve an MPC optimisation. The maximum time observed for implementation of gain selection Method 2 is 70 ms. After $\bar{\mu}_k$ has converged to 1, no gain selection method is executed and no additional online computation other than the MPC optimisation is required.

5.8 Conclusion

We propose an improved stochastic MPC algorithm with dynamic gain selection. Gain selection procedures are executed online and lead to extra computation time that is a tiny fraction of the time for solving an MPC optimisation problem. These selection procedures can be configured so that the feedback gain in the MPC law converges to the unconstrained LQ-optimal and the set of feasible initial conditions is enlarged. By using dynamic gain selection, the conservativeness in controller design is considerably reduced and the closed loop cost is improved. Also, recursive feasibility of MPC optimisation problems, constraint satisfaction and closed loop stability are retained using a similar approach to that of Chapter 3. The results in Sections 5.3, 5.5 and 5.6 of this chapter can be extended to problems with multiple chance constraints. However, solving the corresponding multiobjective optimisation problems to generate stabilising feedback gains for dynamic gain selection becomes more challenging in this case.

5.9 Proofs of Chapter 5

5.9.1 Proof of Lemma 5.1

Equation (5.16) implies that $\widetilde{P}_k = \sum_{i=0}^{\infty} \gamma^i (\Phi_k^\top)^i C^\top C \Phi_k^i$. Combining this with (5.8) shows that

$$\sum_{i=N}^{\infty} \gamma^i \|C\bar{x}_{i|k}\|^2 = \gamma^N \bar{x}_{N|k}^\top \Big(\sum_{j=0}^{\infty} (\gamma^{\frac{1}{2}} \Phi_k^\top)^j C^\top C (\gamma^{\frac{1}{2}} \Phi_k)^j \Big) \bar{x}_{N|k} = \gamma^N \|\bar{x}_{N|k}\|^2_{\widetilde{P}_k}.$$

Furthermore, let $\widetilde{S}_k = \sum_{i=0}^{\infty} \gamma^i X_{i|k}$, then (5.9) implies

$$\gamma \Phi_k \widetilde{S}_k \Phi_k^\top = \sum_{i=0}^{\infty} \gamma^{i+1} \Phi_k X_{i|k} \Phi_k^\top = \sum_{i=0}^{\infty} \gamma^{i+1} \Big(X_{i+1|k} - \Omega \Big) = \widetilde{S}_k - X_{0|k} - \frac{\gamma}{1-\gamma} \Omega,$$

and $\widetilde{S}_k$ satisfies the Lyapunov equation $\widetilde{S}_k = \gamma \Phi_k \widetilde{S}_k \Phi_k^\top + \frac{\gamma}{1-\gamma} \Omega$. Therefore,

$$\begin{aligned} \operatorname{tr}\Big(\sum_{i=0}^{\infty} \gamma^i C^\top C X_{i|k} \Big) = \operatorname{tr}\Big(C^\top C \widetilde{S}_k \Big) &= \operatorname{tr}\Big(C^\top C \sum_{i=0}^{\infty} \gamma^i \Phi_k^i \Big(\frac{\gamma}{1-\gamma} \Omega \Big) (\Phi_k^\top)^i \Big) \\ &= \frac{\gamma}{1-\gamma} \operatorname{tr}\Big(\Omega \widetilde{P}_k \Big). \end{aligned}$$

5.9.2 Proof of Lemma 5.3

We begin the proof with showing that the bottom right block $M_c^\top H M_c$ of $W_1(K_k)$ defined in (5.19) is positive definite. Writing H as $H = h^\top h$, where

$$h := \operatorname{diag}\{\gamma^{\frac{1}{2}} C, \ldots, \gamma^{\frac{N-1}{2}} C, \gamma^{\frac{N}{2}} \widetilde{P}_k^{\frac{1}{2}}\},$$

we have $\operatorname{rank}(M_c^\top H M_c) = \operatorname{rank}(M_c^\top h^\top h M_c) = \operatorname{rank}(h M_c)$, and

$$h M_c = \begin{bmatrix} \gamma^{\frac{1}{2}} CB & & & \\ \gamma C \Phi_k B & \gamma CB & & \\ \vdots & \vdots & \ddots & \\ \gamma^{\frac{N}{2}} \widetilde{P}_k^{\frac{1}{2}} \Phi_k^{N-1} B & \cdots & \cdots & \gamma^{\frac{N}{2}} \widetilde{P}_k^{\frac{1}{2}} B \end{bmatrix}.$$

The matrix $hM_c \in \mathbb{R}^{Nn_x \times Nn_u}$ is a block triangular matrix and it follows that

$$\operatorname{rank}(M_c^\top HM_c) = \operatorname{rank}(hM_c) \geq (N-1)\operatorname{rank}(CB) + \operatorname{rank}\Big(\widetilde{P}_k^{\frac{1}{2}} B\Big).$$

From (5.16), it follows that $\widetilde{P}_k \succeq C^\top C \ \forall k = 0, \ldots$, which, together with that $\operatorname{rank}(CB) = n_u$, implies $\operatorname{rank}\Big(\widetilde{P}_k^{\frac{1}{2}} B\Big) = n_u$. Then $M_c^\top HM_c$ has full rank and is positive definite. Therefore, minimisation (5.29) has a unique minimiser. We next prove this lemma by contradiction. By Lemma 4.7 and Assumption 5.2, the constraint in (5.30) is satisfied with equality if $\bar{\mu}_{k-1} \neq 1$ and $\bar{\mu}_k$ will not reach 1. Therefore, if $\bar{\mu}_{k-1} = \bar{\mu}_k \neq 1$, it implies that

$$\begin{bmatrix} x_k \\ E\mathbf{c}_{k-1}^* \end{bmatrix}^\top W_1\Big(L_\infty(\bar{\mu}_{k-1})\Big) \begin{bmatrix} x_k \\ E\mathbf{c}_{k-1}^* \end{bmatrix} = \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}^\top W_1\Big(L_\infty(\bar{\mu}_{k-1})\Big) \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}$$

and that $E\mathbf{c}_{k-1}^* = \mathbf{c}^o(x_k)$. Recalling the definition of ε_k in (5.26), the feasible sequence $E\mathbf{c}_{k-1}^*$ then becomes the unique minimiser and the only *perturbation sequence* that can satisfy the constraint (5.18) in the online MPC optimisation (5.20) at time k. This gives that $E\mathbf{c}_{k-1}^* = \mathbf{c}_k^*$. If $\bar{\mu}_{k-1} = \bar{\mu}_k \neq 1$ holds for N consecutive time steps, $E^N\mathbf{c}_{k-1}^* = \mathbf{c}_{k+N-1}^* = 0$. This implies that $L_\infty(\bar{\mu}_{k+N-1})$ is optimal with respect to the cost $\sum_{i=0}^\infty \gamma^i \|C\bar{x}_i\|^2$ and that $L_\infty(\bar{\mu}_{k+N-1}) = L_\infty(0)$, which contradicts with the statement proved in Remark 4.4 of Chapter 4. Therefore, the equality $\bar{\mu}_{k-1} = \bar{\mu}_k$ only holds for a finite number of consecutive time steps if $\bar{\mu}_k \neq 1$.

5.9.3 Proof of Lemma 5.4

We prove this lemma by contradiction. Suppose that a pair of $\delta > 0$ and $p_\delta > 0$ does not exist such that (5.34) holds. In this case, for some $z \in \mathbb{R}^{Nn_u}$ and $\mu \in [\bar{\mu}_0, 1]$ we must have

$$\mathbb{P}\Big\{\Big\|z + \Big(W_{cc}(\mu)\Big)^\dagger W_{cx}(\mu)\omega_k\Big\|_{W_{cc}(\mu)}^2 \geq \delta\Big\} < p_\delta \tag{5.41}$$

for all $\delta > 0$ and all $p_\delta > 0$, which implies that $W_{cx}(\mu)\omega_k + W_{cc}(\mu)z = 0$ with probability 1. Here $W_{cx}(\mu)$ must be non-zero since otherwise $\mathbf{c}^o(x_k) = 0$, implying that by its definition in (5.29) the feedback gain used is optimal with respect to $\sum_{i=0}^\infty \gamma^i \|C\bar{x}_i\|_Q^2$ and then $L_\infty(\mu) = L_\infty(0)$ would be obtained, which contradicts the fact that $\mu \geq \bar{\mu}_0 \geq \mu_1$ and Remark 4.4 in Chapter 4. Furthermore, z and μ are by assumption independent of the realisation ω_k. Therefore (5.41) contradicts the assumption in (5.2) that $\mathbb{E}\{\omega_k \omega_k^\top\}$ is positive definite implying ω_k is not a constant vector or in any subspaces of $\mathbb{R}^{n_x}$, and hence (5.34) must hold for some $\delta > 0$ and $p_\delta > 0$.

5.9.4 Proof of Theorem 5.5

We first show convergence of $\bar{\mu}_k$ to 1 in probability (i.e. $\lim_{k\to\infty} \mathbb{P}\{\bar{\mu}_k = 1\} = 1$) by splitting an infinite horizon into intervals of N_f time steps and providing an upper bound on $\mathbb{P}\{\bar{\mu}_k < 1\}$, which is parameterised by p_δ, N_f and k, and which converges to 0 as $k \to \infty$. Let N_f be defined in terms of a $\delta > 0$ satisfying (5.34) for some $p_\delta > 0$ by

$$N_f = \left\lceil \frac{\gamma}{1-\gamma} \frac{\operatorname{tr}\left(\Omega \bar{P}_\infty(1) - \Omega \bar{P}_\infty(\bar{\mu}_0)\right)}{\delta} \right\rceil$$

where it is assumed that $\bar{\mu}_0 < 1$, and let $\mathcal{E}_k$ denote the event that

$$\left\| z_{k-1} + \left(W_{cc}(\bar{\mu}_{k-1})\right)^\dagger W_{cx}(\bar{\mu}_{k-1})\omega_{k-1} \right\|^2_{W_{cc}(\bar{\mu}_{k-1})} \geq \delta.$$

Under Assumption 5.2, Lemma 4.7 implies that before $\bar{\mu}_k$ reaches 1, (5.31b) is satisfied with equality and every possible increment in $\bar{\mu}_k$ is attained if (5.33) is positive. It follows that $\bar{\mu}_{N_f} = 1$ if $\mathcal{E}_k$ occurs for $k = 1, \ldots, N_f$. Furthermore,

$$\mathbb{P}\{\mathcal{E}_1 \cap \cdots \cap \mathcal{E}_{N_f}\} = \mathbb{P}\{\mathcal{E}_1\}\mathbb{P}\{\mathcal{E}_2 | \mathcal{E}_1\} \cdots \mathbb{P}\{\mathcal{E}_{N_f} | \mathcal{E}_1 \cap \mathcal{E}_2 \cap \cdots \cap \mathcal{E}_{N_f - 1}\},$$

where $\mathbb{P}\{\mathcal{E}_1\} \geq p_\delta$ and

$$\mathbb{P}\{\mathcal{E}_k | \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_{k-1}\} \geq \inf_{z \in \mathbb{R}^{Nn_u}, \mu \in [\bar{\mu}_0, 1]} \mathbb{P}\left\{ \left\| z + \left(W_{cc}(\mu)\right)^\dagger W_{cx}(\mu)\omega_{k-1} \right\|^2_{W_{cc}(\mu)} \geq \delta \right\} \geq p_\delta$$

for $k = 2, \ldots, N_f$. Also, these conditional probabilities are well defined by Lemma 5.4. Therefore,

$$\mathbb{P}\{\bar{\mu}_{N_f} < 1\} \leq 1 - p_\delta^{N_f},$$

and more specifically we rewrite it as

$$\mathbb{P}\{\bar{\mu}_{N_f} < 1 \,|\, \bar{\mu}_0 < 1\} \leq 1 - p_\delta^{N_f}.$$

Similarly, given $\bar{\mu}_{(i-1)N_f} < 1$ for $i = 2, \ldots,$ we have

$$\mathbb{P}\{\bar{\mu}_{iN_f} = 1\} \geq \Big(\inf_{z\in\mathbb{R}^{Nn_u},\ \mu\in[\bar{\mu}_0,1]} \mathbb{P}\{\star \geq \delta_i\}\Big)^{N_f} \geq \Big(\inf_{z\in\mathbb{R}^{Nn_u},\ \mu\in[\bar{\mu}_0,1]} \mathbb{P}\{\star \geq \delta\}\Big)^{N_f} \geq p_\delta^{N_f},$$

where $\star$ here denotes $\Big\|z + \Big(W_{cc}(\mu)\Big)^{\dagger} W_{cx}(\mu)\omega_{iN_f-1}\Big\|^2_{W_{cc}(\mu)}$, and δ_i is the counterpart of δ for $\bar{\mu}_{iN_f}$ that is sufficient for $\bar{\mu}_k$ to reach 1 within N_f time steps from time $k = (i-1)N_f$ and is smaller than δ since the sequence $\{\bar{\mu}_k\}_{k=0}^{\infty}$ is monotonically non-decreasing. Then

$$\mathbb{P}\{\bar{\mu}_{iN_f} < 1 \,|\, \bar{\mu}_{(i-1)N_f} < 1\} \leq 1 - p_\delta^{N_f}, \quad \forall\, i = 2, \ldots.$$

For given $k > N_f$, we choose integers $j_1, j_2, \ldots, j_{\lfloor k/N_f \rfloor}$ so that $j_1 = N_f$, $j_{\lfloor k/N_f \rfloor} \leq k$ and $j_{i+1} - j_i = N_f$ for all i. Then $\bar{\mu}_k < 1$ only if $\bar{\mu}_{j_i} < 1$ for all $i = 1, \ldots, \lfloor k/N_f \rfloor$ and hence

$$\begin{aligned}\mathbb{P}\{\bar{\mu}_k < 1\} &\leq \mathbb{P}\{\bar{\mu}_{j_1} < 1\}\mathbb{P}\{\bar{\mu}_{j_2} < 1 \,|\, \bar{\mu}_{j_1} < 1\} \cdots \mathbb{P}\{\bar{\mu}_{j_{\lfloor k/N_f \rfloor}} < 1 \,|\, \bar{\mu}_{j_{\lfloor k/N_f \rfloor -1}} < 1\} \\ &\leq \Big(1 - p_\delta^{N_f}\Big)^{\lfloor k/N_f \rfloor}. \end{aligned} \tag{5.42}$$

It follows that $\mathbb{P}\{\bar{\mu}_k = 1\} \geq 1 - (1 - p_\delta^{N_f})^{\lfloor k/N_f \rfloor}$, which implies $\lim_{k\to\infty} \mathbb{P}\{\bar{\mu}_k = 1\} = 1$. To complete the proof, we use the Borel-Cantelli lemma to show the almost sure convergence of $\bar{\mu}_k$ to 1, that is, $\mathbb{P}\{\lim_{k\to\infty} \bar{\mu}_k = 1\} = 1$. Let $\mathcal{F}_k$ denote the event that $\bar{\mu}_k < 1$, then (5.42) ensures that

$$\sum_{k=1}^{\infty} \mathbb{P}\{\mathcal{F}_k\} \leq \sum_{k=1}^{\infty} (1 - p_\delta^{N_f})^{\lfloor k/N_f \rfloor} = N_f p_\delta^{-N_f} - 1 < \infty.$$

Therefore the Borel-Cantelli lemma implies that $\mathbb{P}\{\bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} \mathcal{F}_j\} = 0$. But $\bar{\mu}_k < 1$ only if $\bar{\mu}_{k-1} < 1$, so $\mathcal{F}_k \subseteq \mathcal{F}_{k-1}$ and $\bigcup_{j=k}^{\infty} \mathcal{F}_j = \mathcal{F}_k$. Therefore, $\mathbb{P}\{\lim_{k\to\infty} \mathcal{F}_k\} = 0$ or equivalently $\mathbb{P}\{\lim_{k\to\infty} \bar{\mu}_k = 1\} = 1$.

5.9.5 Proof of Theorem 5.9

Given initial feasibility at time $k = 0$, step (ii) of Algorithm 5.1 ensures that problem (5.20) is always feasible for time $k = 1, 2, \ldots$. From Lemma 5.2, the vector $\tilde{\mathbf{c}}_{k+1}$ provides a feasible but possibly suboptimal solution at time $k + 1$. Hence by optimality we necessarily have

$$J^*(x_{k+1}, K_{k+1}) \leq J(x_{k+1}, K_k, \tilde{\mathbf{c}}_{k+1}),$$

and since this inequality holds for every realisation of ω_k, by taking expectations conditioned on the state x_k we obtain

$$\mathbb{E}_k\{J^*(x_{k+1}, K_{k+1})\} \leq \mathbb{E}_k\{J(x_{k+1}, K_k, \tilde{\mathbf{c}}_{k+1})\}. \tag{5.43}$$

From (5.28) we have the feasible sequence

$$\bar{x}_{i|k+1} = \bar{x}^*_{i+1|k} + \Phi^i_k \omega_k, \qquad i = 0, \ldots, N,$$

and from (5.13) and (5.43) it follows that

$$\mathbb{E}_k\{J^*(x_{k+1}, K_{k+1})\} \leq J^*(x_k, K_k) - \|x_k\|^2_Q - \|u_k\|^2_R + \operatorname{tr}(\Omega P_k). \tag{5.44}$$

Summing both sides of this inequality over $k \geq 0$ after taking expectations given information available at time $k = 0$, and making use of the property that $\mathbb{E}_0\{\mathbb{E}_k\{J^*(x_{k+1}, K_{k+1})\}\} = \mathbb{E}_0\{J^*(x_{k+1}, K_{k+1})\}$, we obtain the first inequality of (5.39). Note that K_k and P_k depend on x_k, which implies that they are random variables, but they are uncorrelated with ω_k. Moreover, the second inequality in (5.39) follows from the choice of P_k given in step (3) of Method 1 or in step (4) of Method 2 and the monotonicity result in Lemma 4.5.

5.9.6 Proof of Lemma 5.10

Since (5.26) is equivalent to (5.27), expanding the terms in (5.27) yields

$$\begin{aligned}
\varepsilon_{k+1} = &\sum_{i=0}^{N-1} \gamma^i \left\| C\bar{x}^*_{i+1|k} \right\|^2 + \gamma^N \left\| \bar{x}^*_{N+1|k} \right\|^2_{\widetilde{P}_k} + \frac{\gamma}{1-\gamma} \operatorname{tr}\!\left(\Omega \widetilde{P}_k\right) \\
&+ \sum_{i=0}^{N-1} \gamma^i \left\| C\Phi^i_k \omega_k \right\|^2 + \gamma^N \left\| \Phi^N_k \omega_k \right\|^2_{\widetilde{P}_k} \\
&+ \sum_{i=0}^{N-1} 2\gamma^i \left(\Phi^i_k \omega_k\right)^\top C^\top C \bar{x}^*_{i+1|k} + 2\gamma^N \left(\Phi^N_k \omega_k\right)^\top \widetilde{P}_k \, \bar{x}^*_{N+1|k},
\end{aligned} \tag{5.45}$$

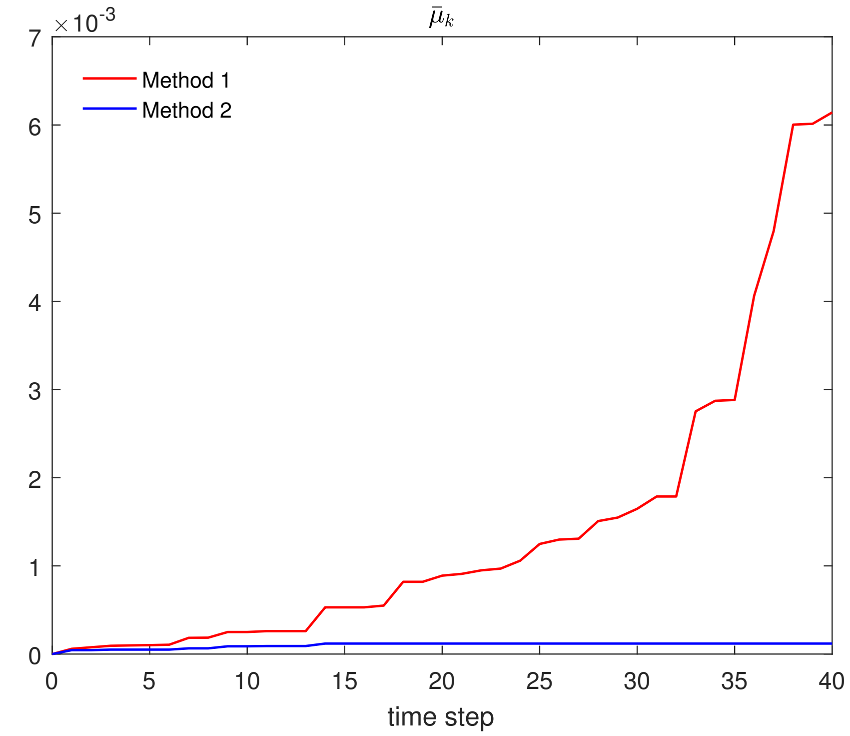
where $\bar{x}^*_{i|k}$ is given by (5.21)-(5.22) and ω_k is the realisation of the disturbance at time k. From (5.16) and $\bar{x}^*_{N+1|k} = \Phi_k \bar{x}^*_{N|k}$, the sum of the first three terms on the RHS of (5.45) is

$$\gamma^{-1}\Big(\sum_{i=0}^{N-1} \gamma^i \|C\bar{x}^*_{i|k}\|^2 + \gamma^N \|\bar{x}^*_{N|k}\|^2_{\widetilde{P}_k} + \frac{\gamma}{1-\gamma} \operatorname{tr}(\Omega \widetilde{P}_k) - \|Cx_k\|^2\Big) - \operatorname{tr}(\Omega \widetilde{P}_k),$$

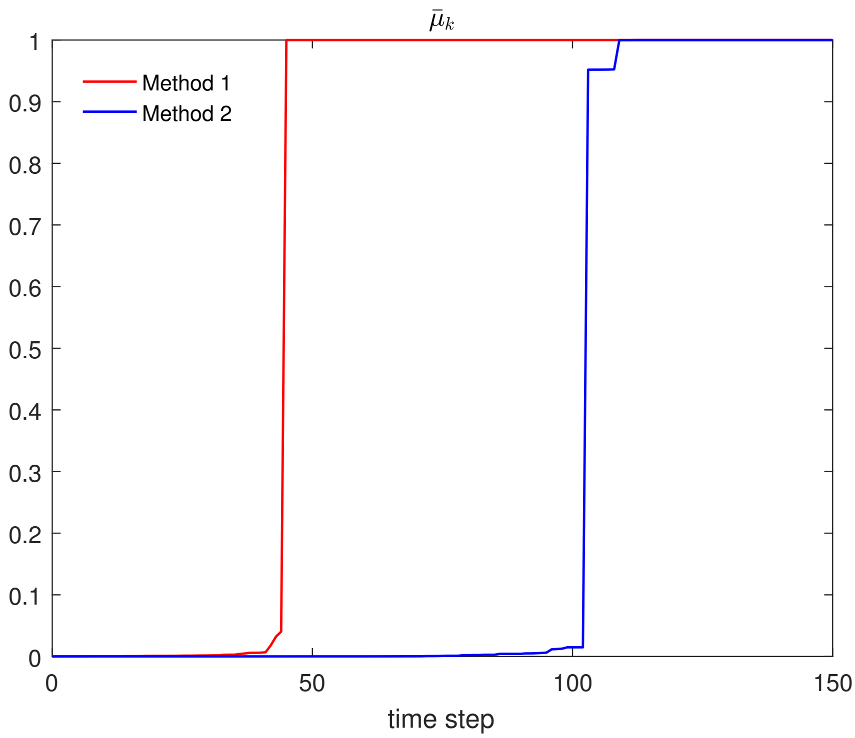
and from (5.16) the sum of the next two terms is $\omega_k^\top \widetilde{P}_k \omega_k$. Noting that K_k, $\widetilde{P}_k$ and x_k are independent of ω_k, taking the expectation of ε_{k+1} conditioned on information available at time k therefore gives

$$\gamma \mathbb{E}_k\{\varepsilon_{k+1}\} = \sum_{i=0}^{N-1} \gamma^i \|C\bar{x}^*_{i|k}\|^2 + \gamma^N \|\bar{x}^*_{N|k}\|^2_{\widetilde{P}_k} + \frac{\gamma}{1-\gamma}\,\mathrm{tr}(\Omega \widetilde{P}_k) - \|Cx_k\|^2.$$

This equation, together with feasibility of the sequence $\{\bar{x}^*_{i|k}\}_{i=0}^N$ at time k, proves (5.40).



(a) Short simulation horizon



(b) Long simulation horizon

Figure 5.2: Evolution of $\bar{\mu}_k$ in Simulation B

6

Output Feedback Control with Intermittent Observations

Contents

This chapter considers an application of stochastic model predictive control

(MPC) developed in previous chapters in communication networks. We propose an MPC law that minimises a discounted cost subject to a discounted expectation constraint, where discrete-time linear systems with stochastic additive disturbances and noisy measurements transmitted over a lossy communication channel are considered. It is assumed that sensor data is lost with a known probability, and data losses are accounted for by expressing the predicted control policy as an affine function of future observations, which results in a convex optimisation problem. A constraint-tightening technique ensures recursive feasibility of the online optimisation problem and satisfaction of the expectation constraint without imposing bounds on the distributions of the noise and disturbance inputs. The discounted cost evaluated along trajectories of the closed loop system is shown to be bounded by the initial optimal predicted cost. We also provide conditions under which the averaged undiscounted closed loop cost accumulated over an infinite horizon is bounded. Numerical simulations are described to illustrate these results.

6.1 Introduction

Applications of wireless sensor networks are becoming omnipresent, for example, in healthcare monitoring, weather forecasting, autonomous driving, and building management systems. These networks can be time-varying and subject to congestion. Networked control systems that rely on sensor signals transmitted over communication channels must tolerate communication delays and data losses. These issues pose additional challenges for estimator and controller design, especially when constraints need to be met.

A large body of work exists on estimation problems involving intermittent observations. They are largely concerned with the Random Riccati Equation [127], and the main issue addressed is the boundedness of the state estimation error covariance. In [103], packet losses are modelled by an independent and identically distributed (i.i.d.) Bernoulli process and fundamental results are derived, such as the existence and bounds of the critical value for the arrival probability of observation updates and convergence properties of the algebraic Riccati equation for Kalman filters with intermittent observations. This is the basis of further work in [128] showing that the estimation error covariance sequence $\Sigma_0, \Sigma_1, \ldots$ converges in distribution to a unique invariant probability distribution with finite mean $\bar{\Sigma}$ when the observation arrival probability exceeds the critical value. In [129], the boundedness of the estimation error covariance is studied from a probabilistic perspective and bounds on the probability of $\Sigma_k \preceq \bar{\Sigma}$ (henceforth denoted $\mathbb{P}\{\Sigma_k \preceq \bar{\Sigma}\}$) are derived. A similar metric is used in [130]. It considers the tail distribution of the trace of the estimation error covariance and the decay rate of $\mathbb{P}\{\mathrm{tr}(\Sigma_k) > \mathrm{tr}(\bar{\Sigma})\}$ (where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix) with packet losses modelled by a Markov process. For the case of Markovian packet losses, it is possible to account for the temporal correlation of communication channel conditions, but the stability analysis is more involved without an i.i.d. assumption [131, 132]. In contrast to the aforementioned work focusing on limiting behaviours of the estimation error covariance, [133] provides a different perspective by analysing transient properties. This work considers an intermittent quantised communication link and characterises the cumulative probability distribution function of the escape time at which system states or state estimates first exit a given compact set.

Problems involving constrained estimation and control under intermittent and noisy observations, on the other hand, have received less attention. In [134], the problem of minimising an average estimation error covariance over a finite horizon is studied subject to an energy constraint on remote sensors. [135] considers a similar problem subject to an average transmission power constraint, while the probability of packet losses at each time step is conditional on transmission power. In [61], the problem of controlling stable linear systems with Gaussian disturbances and measurement noises is considered subject to hard input constraints, where both sensor and control channels are unreliable. This work assumes that remote sensors equipped with computing capabilities send state estimates instead of raw sensor measurements, which could be costly and unreasonable in certain situations.

Assuming sensor measurements will be lost with a known probability according to a Bernoulli process, we propose an output-feedback MPC algorithm to minimise a discounted cost subject to a discounted expectation constraint. The system model has additive disturbances and measurement noise with probability distributions that are not assumed to be known exactly and may not have bounded support. The discount setting is common to many control problems [e.g. 19–21, 45, 136], and an appropriate discount factor can provide stability guarantees [30]. The main contributions of this chapter are as follows:

- The discount factors ensure that both the cost function and constraint are well defined without bounds on the disturbance and measurement noise.
- Instead of choosing the control policy as pre-stabilising linear state feedback laws with perturbations [e.g. 34], we parameterise predicted control inputs as affine functions of future output measurements and show that the problem of optimising the associated feedback gains is convex. This allows the distributions of future states to be controlled explicitly even when output measurements are lost.
- We employ a constraint-tightening technique [44] to guarantee recursive feasibility of online MPC optimisation problems in this stochastic setting, and we derive bounds on the discounted cost and constraint for the closed loop system.
- We also show that an appropriate discount factor leads to a stochastic stability guarantee in closed loop under reasonable assumptions.
- Different discount factors in the cost and constraint, respectively, are considered. We demonstrate that a cost discount factor sufficiently close to 1 ensures stochastic stability and provide additional insights into the role of cost and constraint discount factors.

This chapter is organised as follows. The control problem is described in Section 6.2, and the controller parameterisation and implementation introduced in Section 6.3. We compute predicted state and control sequences via their first and second moments in Section 6.4. In Section 6.5, terminal conditions and explicit expressions for the cost and constraints are derived. Section 6.6 proves a closed loop discounted cost bound and analyses constraint satisfaction. The averaged undiscounted closed loop cost accumulated over an infinite horizon is shown to be finite and an explicit bound is given in Section 6.7. Sections 6.8 and 6.9 provide a numerical example and conclusions.

6.2 Problem Description

6.2.1 System model and feedback information

We assume a system with linear discrete-time dynamics

$$x_{k+1} = Ax_k + Bu_k + Dw_k, \tag{6.1a}$$

$$y_k = Cx_k + v_k, \tag{6.1b}$$

$$z_k = \gamma_k y_k, \tag{6.1c}$$

where $x_k \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$, $y_k \in \mathbb{R}^{n_y}$, $z_k \in \mathbb{R}^{n_y}$ are the state, control input, sensor measurement, and the measurement information received by the controller respectively. The disturbance, measurement noise and packet loss sequences, $\{w_k\}_{k=0}^{\infty}$, $\{v_k\}_{k=0}^{\infty}$ and $\{\gamma_k\}_{k=0}^{\infty}$, are assumed to be i.i.d. with

$$\begin{aligned} &\mathbb{E}\{w_k\} = 0, && \mathbb{E}\{w_k w_k^\top\} =: \Sigma_w \succeq 0, \\ &\mathbb{E}\{v_k\} = 0, && \mathbb{E}\{v_k v_k^\top\} =: \Sigma_v \succ 0, \\ &\mathbb{P}\{\gamma_k = 0\} = 1 - \lambda, && \mathbb{P}\{\gamma_k = 1\} =: \lambda. \end{aligned}$$

The variable $\gamma_k \in \{0, 1\}$ indicates whether sensor data at the kth sampling instant is received by the controller. The information assumed available to the controller at the kth time step consists of $\{u_i\}_{i=0}^{k-1}$, $\{(z_i, \gamma_i)\}_{i=0}^{k}$, the initial mean $\mathbb{E}\{x_0\} =: \hat{x}_0$, and model state covariance $\mathbb{E}\{(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^\top\} =: \Sigma_0$.

We define the information sets

$$\mathcal{I}_k := \{\mathcal{I}_{k-1}, (z_k, \gamma_k)\}, \quad \mathcal{U}_k := \{\mathcal{U}_{k-1}, u_k\},$$

for all $k \geq 0$, where $\mathcal{I}_{-1} := \{\hat{x}_0, \Sigma_0\}$, $\mathcal{U}_{-1} := \{\,\}$. Finally, we define conditional expectation operators as

$$\mathbb{E}_k\{\cdot\} := \mathbb{E}\{\cdot \,|\, \mathcal{U}_{k-1}, \mathcal{I}_{k-1}\}, \quad \mathbb{E}\{\cdot\} := \mathbb{E}_0\{\cdot\}.$$

Assumption 6.1. *The pair (A, B) is stabilisable, and (A, C) is detectable.*

6.2.2 Optimal control problem

We consider the problem of minimising the discounted sum of expected future values of $\|x_k\|_Q^2 + \|u_k\|_R^2$, where $Q \succeq 0$ and $R \succ 0$. The minimisation is performed subject to a constraint on the discounted sum of second moments of an auxiliary output, defined for a given matrix H by $\xi_k = Hx_k$, so the optimal control problem is

$$\begin{aligned} \min\ & \sum_{k=0}^{\infty} \beta_1^k \mathbb{E}\{\|x_k\|_Q^2 + \|u_k\|_R^2\} \\ \text{s.t.}\ & \sum_{k=0}^{\infty} \beta_2^k \mathbb{E}\{\|Hx_k\|^2\} \leq \epsilon. \end{aligned} \tag{6.2}$$

Here $\beta_1 \in (0, 1)$ and $\beta_2 \in (0, 1)$ are discount factors and ϵ is a given bound on the infinite discounted sum of second moments. For simplicity (and with the understanding that the general case of $\beta_1 \neq \beta_2$ can be treated similarly), we set $\beta_1 = \beta_2 = \beta$ for the remainder of this chapter with the exception of Sections 6.7.2 and 6.8.

We will employ a receding horizon control law based on a finite horizon control policy in the form

$$u_{i|k} = \kappa_i(\theta_k, \mathcal{U}_{k+i-1}, \mathcal{I}_{k+i}) \tag{6.3}$$

where $u_{i|k}$ for $i = 0, 1, \ldots$ is the prediction of u_{k+i} at time k, and θ_k is a vector of controller parameters at time k. The dependence of $\kappa_i(\cdot)$ on the sets $\mathcal{U}_{k+i-1}$ and $\mathcal{I}_{k+i}$ ensures causality and the dependence on θ_k is chosen so that the optimal parameter vector, denoted θ_k^*, will be the solution of a convex problem.

Assumption 6.2. *(i) The probability, λ, of successfully receiving sensor measurements is known and is time invariant. (ii) When θ_k^* is computed, (z_{k+i}, γ_{k+i}) are unknown for all $i \geq 0$.*

Thus θ_k^* is computed online prior to the kth sampling instant using knowledge of $\mathcal{U}_{k-1}$ and $\mathcal{I}_{k-1}$. However (z_k, γ_k) is known when the control law is implemented, via

$$u_k = \kappa_0(\theta_k^*, \mathcal{U}_{k-1}, \mathcal{I}_k).$$

In order to address (6.2), we define an MPC optimisation problem to be solved at time k as follows

$$\begin{aligned} \theta_k^* = \arg\min_{\theta_k} \ & \sum_{i=0}^{\infty} \beta^i \mathbb{E}_k\{\|x_{i|k}\|_Q^2 + \|u_{i|k}\|_R^2\} \\ \text{s.t.} \ & \sum_{i=0}^{\infty} \beta^i \mathbb{E}_k\{\|Hx_{i|k}\|^2\} \leq \mu_k. \end{aligned} \tag{6.4}$$

Here $\mu_0 = \epsilon$, and μ_k for $k > 0$ is chosen as described in Section 6.6 to ensure that (6.4) is recursively feasible and that the closed loop system satisfies the constraint in (6.2).

6.3 Controller Design

6.3.1 Parameterisation

Consider the output feedback control law defined by an observer and an affine feedback law:

$$\hat{x}_k = A\tilde{x}_{k-1} + Bu_{k-1}, \tag{6.5a}$$

$$\tilde{x}_k = \hat{x}_k + \gamma_k M(y_k - C\hat{x}_k), \tag{6.5b}$$

$$u_k = K\tilde{x}_k + c_k. \tag{6.5c}$$

with $\hat{x}_0 = \mathbb{E}\{x_0\}$, where $\hat{x}_k$ and $\tilde{x}_k$ are the prior estimate and the posterior estimate of x_k, respectively. A simplistic parameterisation of the predicted control law $\kappa_i(\cdot)$ in (6.3) could be obtained if the observer gain M and feedback gain K were fixed and the optimisation variables in problem (6.4) were defined as $\theta_k = \{c_{0|k}, \ldots, c_{N-1|k}\}$ for some fixed N, with the predicted control sequence defined as $u_{i|k} = K\tilde{x}_{i|k} + c_{i|k}$. Although this would require a number of optimisation variables that grows only linearly with N, the parameters $\{c_{0|k}, \ldots, c_{N-1|k}\}$ constitute an open loop control sequence that does not vary with the future measurement noise and disturbance realisations. This is likely to provide poor performance and small sets of feasible initial conditions when the probability of packet loss is non-zero.

By using a parameterisation that allows the dependence of the predicted control sequence on future realisations of model uncertainty to be optimised, the predicted probability distributions of states and control inputs can be controlled explicitly. This provides flexibility to balance conflicting requirements for performance and constraint satisfaction. However, similarly to the case of predicted control laws in which state feedback gains are decision variables [83, 137], the cost and constraints of problem (6.4) are nonconvex if time-varying gains M, K are considered as optimisation variables. On the other hand, if predicted control inputs are parameterised in terms of affine functions of the future output measurements received by the controller [138], then the dependence of the first and second moments of predicted states and inputs on controller parameters is convex. Moreover, by incorporating affine terms in the future innovation sequence, a predicted control law with arbitrary linear dependence of $\kappa_i(\cdot)$ on the received sensor measurements can be obtained. This approach allows the future control sequence to be optimised at every sampling instant, including those at which information from sensors is lost.

We therefore express the i-step-ahead predicted control input $u_{i|k}$, for all $i = 0, 1, \ldots,$ as

$$u_{i|k} = K\hat{x}_{i|k} + c_{i|k} + d_{i|k}, \tag{6.6a}$$

$$d_{i|k} = \gamma_{0|k}L_{i,0|k}(y_{0|k} - C\hat{x}_{0|k}) + \gamma_{1|k}L_{i,1|k}(y_{1|k} - C\hat{x}_{1|k}) + \cdots + \gamma_{i|k}L_{i,i|k}(y_{i|k} - C\hat{x}_{i|k}), \tag{6.6b}$$

$$\hat{x}_{i+1|k} = A\hat{x}_{i|k} + Bu_{i|k} + \gamma_{i|k}AM(y_{i|k} - C\hat{x}_{i|k}) \tag{6.6c}$$

where $c_{i|k} = 0$ and $L_{i,j|k} = 0$ for all $i \geq N$. Here $\gamma_{i|k}$ and $y_{i|k}$ are random variables, denoting the i-step-ahead predicted packet loss and sensor measurement at time k, respectively. Then, for all $i = 0, 1, \ldots$ the predicted state estimate satisfies

$$\hat{x}_{i+1|k} = \Phi\hat{x}_{i|k} + B(c_{i|k} + d_{i|k}) + \gamma_{i|k}AM(y_{i|k} - C\hat{x}_{i|k}) \tag{6.7}$$

with $\Phi := A + BK$. Since $x_{i+1|k} = Ax_{i|k} + Bu_{i|k} + Dw_{i|k}$, the predicted estimation error evolves according to

$$x_{i+1|k} - \hat{x}_{i+1|k} = \Psi_{i|k}(x_{i|k} - \hat{x}_{i|k}) - \gamma_{i|k} AMv_{i|k} + Dw_{i|k} \tag{6.8}$$

where $\Psi_{i|k} := A(I - \gamma_{i|k}MC)$. These relationships allow the first and second moments of $x_{i|k}$ to be determined in terms of the decision variable θ_k, which consists of the parameters $\{c_{0|k}, \ldots, c_{N-1|k}\}$ and feedback gains $L_{0,0|k}$, $\{L_{1,0|k}, L_{1,1|k}\}$, $\ldots, \{L_{N-1,0|k}, \ldots, L_{N-1,N-1|k}\}$.

The gains K and M in the predicted control law (6.6a-c) are fixed and chosen offline, satisfying the following assumption.

Assumption 6.3. *Gains K and M are chosen so that $\xi_{i+1} = (A + BK)\xi_i$ is asymptotically stable and $\xi_{i+1} = A(I - \gamma_i MC)\xi_i$ is mean-square stable (MSS) [104].*

Remark 6.1. *Matrices K and M exist satisfying Assumption 6.3 if Assumption 6.1 holds and if the probability, λ, of successfully receiving a sensor measurement is greater than the critical value [e.g. 103]. Suitable choices for K, M are the optimal gains for (6.4) in the absence of constraints, or the certainty equivalent LQ feedback gain for a problem with state and control weighting matrices Q and R and the steady state Kalman filter gain [103]. Note that time-varying gains K_k, M_k can be used within the framework of this chapter, provided the dependence on γ_k is known in advance.*

6.3.2 Implementation

The control law is implemented as follows:

(i) Given $\mathcal{U}_{k-1}$ and $\mathcal{I}_{k-1}$, solve problem (6.4) for θ_k^*.

(ii) Given γ_k and $z_k = \gamma_k y_k$:

(a) apply the control input

$$u_k = K\hat{x}_k + c^*_{0|k} + \gamma_k L^*_{0,0|k}(y_k - C\hat{x}_k),$$

(b) update the state estimate

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + \gamma_k AM(y_k - C\hat{x}_k).$$

Note that this receding horizon control law includes (6.5) as a special case, since u_k and $\hat{x}_{k+1}$ in step (ii) are equal to their counterparts in (6.5) if $(c^*_{0|k}, L^*_{0,0|k}) = (c_k, KM)$.

6.4 Predicted State and Control Sequences

To simplify notation, we express the predicted control law in terms of vectorised sequences, with $\mathbf{x}_k$ denoting the vectorised true state sequence $\{x_{i|k}\}_{i=0}^{N-1}$, $\hat{\mathbf{x}}_k$ the estimated state sequence $\{\hat{x}_{i|k}\}_{i=0}^{N-1}$, $\mathbf{u}_k$ the predicted control sequence $\{u_{i|k}\}_{i=0}^{N-1}$, $\mathbf{c}_k$ the predicted control perturbations $\{c_{i|k}\}_{i=0}^{N-1}$, $\mathbf{w}_k$ the disturbance sequence $\{w_{i|k}\}_{i=0}^{N-1}$, $\mathbf{v}_k$ the sensor noise sequence $\{v_{i|k}\}_{i=0}^{N-1}$, and $\boldsymbol{\zeta}_k$ the future innovation sequence $\{\gamma_{i|k}(y_{i|k} - C\hat{x}_{i|k})\}_{i=0}^{N-1}$ at time k. For a given sequence of matrices $\{\Psi_{i|k}\}_{i=0}^{N-1}$ and matrix B let

$$\mathbf{S}_\Psi := \begin{bmatrix} I \\ \Psi_{0|k} \\ \vdots \\ \prod_{i=N-2}^{0} \Psi_{i|k} \end{bmatrix}, \ \mathbf{T}_{(\Psi,B)} := \begin{bmatrix} 0 & \cdots & 0 \\ B & & \\ \vdots & \ddots & \\ \prod_{i=N-2}^{1} \Psi_{i|k}B & \cdots \ B & 0 \end{bmatrix},$$

$$S_\Psi^N := \prod_{i=N-1}^{0} \Psi_{i|k}\,, \qquad T_{(\Psi,B)}^N := \begin{bmatrix} \prod_{i=N-1}^{1} \Psi_{i|k}B & \cdots & B \end{bmatrix},$$

where $\prod_{i=m}^{n} \Psi_{i|k} := \Psi_{m|k} \cdots \Psi_{n|k}$ for $m \geq n$, and define

$$\mathbf{L}_k := \begin{bmatrix} L_{0,0|k} & & & \\ L_{1,0|k} & L_{1,1|k} & & \\ \vdots & \vdots & \ddots & \\ L_{N-1,0|k} & L_{N-1,1|k} & \cdots & L_{N-1,N-1|k} \end{bmatrix},$$

$$\boldsymbol{\Gamma}_k := \mathrm{diag}\{\gamma_{0|k}, \ldots, \gamma_{N-1|k}\} \otimes I_{n_y \times n_y},$$

$\mathbf{K} := I_{N\times N} \otimes K$, $\mathbf{M} := I_{N\times N} \otimes M$ and $\mathbf{C} := I_{N\times N} \otimes C$. Then from (6.8) we have

$$\mathbf{x}_k - \hat{\mathbf{x}}_k = \mathbf{S}_\Psi(x_k - \hat{x}_k) - \mathbf{T}_{(\Psi,A)}\mathbf{M}\boldsymbol{\Gamma}_k\mathbf{v}_k + \mathbf{T}_{(\Psi,D)}\mathbf{w}_k \tag{6.9}$$

while (6.7) and (6.6b) give

$$\hat{\mathbf{x}}_k = \mathbf{S}_\Phi \hat{x}_k + \mathbf{T}_{(\Phi,B)}(\mathbf{c}_k + \mathbf{L}_k \boldsymbol{\zeta}_k) + \mathbf{T}_{(\Phi,A)}\mathbf{M}\boldsymbol{\zeta}_k.$$

Here $\boldsymbol{\zeta}_k = \boldsymbol{\Gamma}_k \mathbf{C}(\mathbf{x}_k - \hat{\mathbf{x}}_k) + \boldsymbol{\Gamma}_k \mathbf{v}_k$ and $\mathbf{S}_\Phi$, $\mathbf{T}_{(\Phi,B)}$ are defined (analogously to $\mathbf{S}_\Psi$, $\mathbf{T}_{(\Psi,B)}$) in terms of Φ and B. Hence

$$\hat{\mathbf{x}}_k = \mathbf{S}_\Phi \hat{x}_k + \mathbf{T}_{(\Phi,B)}\mathbf{c}_k + (\mathbf{T}_{(\Phi,B)}\mathbf{L}_k + \mathbf{T}_{(\Phi,A)}\mathbf{M})\boldsymbol{\zeta}_k, \tag{6.10a}$$

$$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_k + \mathbf{c}_k + \mathbf{L}_k \boldsymbol{\zeta}_k. \tag{6.10b}$$

Clearly the predicted estimation error, state and control sequences in (6.9) and (6.10a,b) depend linearly on the decision variables $\theta_k := (\mathbf{c}_k, \mathbf{L}_k)$.

6.4.1 First and second moments of predicted sequences

In order to express the cost and constraints of problem (6.4) in terms of the parameterisation introduced in Section 6.3.1, we derive in this section expressions for the means and variances of predicted state and control sequences.

We first consider the state x_k of the plant in (6.1a) and the state estimate update $\hat{x}_k$ in step (ii) of the controller implementation in Section 6.3.2. By assumption, $\mathbb{E}\{x_0\} = \hat{x}_0$ and $\mathbb{E}\{w_k\} = 0$, $\mathbb{E}\{v_k\} = 0$ for all $k \geq 0$, and hence the update of state estimates $\hat{x}_k$ in step (ii)(b) ensures that

$$\mathbb{E}_k\{x_k\} = \hat{x}_k \tag{6.11}$$

for all $k \geq 1$. Furthermore, from (6.1a) we have

$$x_k - \hat{x}_k = \Psi_{k-1}(x_{k-1} - \hat{x}_{k-1}) - \gamma_{k-1} A M v_{k-1} + D w_{k-1}$$

for all $k \geq 1$. Let Σ_k denote the second moment of the state estimate error at time k:

$$\Sigma_k := \mathbb{E}_k\Big\{(x_k - \hat{x}_k)(x_k - \hat{x}_k)^\top\Big\}.$$

Then Σ_k evolves according to

$$\Sigma_k = \Psi_{k-1}\Sigma_{k-1}\Psi_{k-1}^\top + \gamma_{k-1} A M \Sigma_v M^\top A^\top + D \Sigma_w D^\top \tag{6.12}$$

for all $k \geq 1$, with initial condition Σ_0, and, by Assumption 6.3, $\mathbb{E}\{\Sigma_k\}$ remains upper bounded $\forall k \geq 0$.

We can now derive the first and second moments of the predicted state sequence $\mathbf{x}_k$ and control sequence $\mathbf{u}_k$.

Proposition 6.2. *Let $\boldsymbol{\pi}_k$, $\boldsymbol{\Pi}_k$, and $\boldsymbol{\Omega}_k$ be defined*

$$\boldsymbol{\pi}_k := \mathbf{S}_\Phi \hat{x}_k + \mathbf{T}_{(\Phi,B)}\mathbf{c}_k, \quad \boldsymbol{\Pi}_k := \mathbf{T}_{(\Phi,B)}\mathbf{L}_k + \mathbf{T}_{(\Phi,A)}\mathbf{M},$$
$$\boldsymbol{\Omega}_k := \mathbb{E}_k\left\{\begin{bmatrix}\mathbf{x}_k - \hat{\mathbf{x}}_k \\ \boldsymbol{\zeta}_k\end{bmatrix}\begin{bmatrix}\mathbf{x}_k - \hat{\mathbf{x}}_k \\ \boldsymbol{\zeta}_k\end{bmatrix}^\top\right\}.$$

Then

$$\mathbb{E}_k\{\mathbf{x}_k\} = \mathbb{E}_k\{\hat{\mathbf{x}}_k\} = \boldsymbol{\pi}_k, \tag{6.13a}$$
$$\mathbb{E}_k\{\mathbf{u}_k\} = \mathbf{K}\mathbb{E}_k\{\hat{\mathbf{x}}_k\} + \mathbf{c}_k = \mathbf{K}\boldsymbol{\pi}_k + \mathbf{c}_k, \tag{6.13b}$$

and

$$\mathbb{E}_k\{\mathbf{x}_k\mathbf{x}_k^\top\} = \boldsymbol{\pi}_k\boldsymbol{\pi}_k^\top + \begin{bmatrix}I & \boldsymbol{\Pi}_k\end{bmatrix}\boldsymbol{\Omega}_k\begin{bmatrix}I \\ \boldsymbol{\Pi}_k^\top\end{bmatrix}, \tag{6.14a}$$
$$\mathbb{E}_k\{\mathbf{u}_k\mathbf{u}_k^\top\} = (\mathbf{K}\boldsymbol{\pi}_k + \mathbf{c}_k)(\mathbf{K}\boldsymbol{\pi}_k + \mathbf{c}_k)^\top + \begin{bmatrix}0 & \mathbf{L}_k + \mathbf{K}\boldsymbol{\Pi}_k\end{bmatrix}\boldsymbol{\Omega}_k\begin{bmatrix}0 \\ (\mathbf{L}_k + \mathbf{K}\boldsymbol{\Pi}_k)^\top\end{bmatrix}. \tag{6.14b}$$

Proof. From (6.9) and (6.11), we have $\mathbb{E}_k\{\mathbf{x}_k - \hat{\mathbf{x}}_k\} = 0$. Therefore $\boldsymbol{\zeta}_k = \boldsymbol{\Gamma}_k\mathbf{C}(\mathbf{x}_k - \hat{\mathbf{x}}_k) + \boldsymbol{\Gamma}_k\mathbf{v}_k$ implies $\mathbb{E}_k\{\boldsymbol{\zeta}_k\} = 0$ and (6.13a,b) follow from the expectations of (6.10a,b). To determine the second moments of $\mathbf{x}_k$ and $\mathbf{u}_k$, let

$$\mathbf{X}_k := \mathbb{E}_k\left\{\begin{bmatrix}\mathbf{x}_k - \hat{\mathbf{x}}_k \\ \hat{\mathbf{x}}_k\end{bmatrix}\begin{bmatrix}\mathbf{x}_k - \hat{\mathbf{x}}_k \\ \hat{\mathbf{x}}_k\end{bmatrix}^\top\right\}.$$

Then from (6.9) and (6.10a) we have

$$\mathbf{X}_k = \begin{bmatrix}0 & 0 \\ 0 & \boldsymbol{\pi}_k\boldsymbol{\pi}_k^\top\end{bmatrix} + \begin{bmatrix}I & 0 \\ 0 & \boldsymbol{\Pi}_k\end{bmatrix}\boldsymbol{\Omega}_k\begin{bmatrix}I & 0 \\ 0 & \boldsymbol{\Pi}_k\end{bmatrix}^\top, \tag{6.15}$$

and (6.14a,b) follow from $\mathbb{E}_k\{\mathbf{x}_k\mathbf{x}_k^\top\} = \begin{bmatrix}I & I\end{bmatrix}\mathbf{X}_k\begin{bmatrix}I & I\end{bmatrix}^\top$ and (6.10a,b), respectively. □

Since $\boldsymbol{\pi}_k$ and $\boldsymbol{\Pi}_k$ are linear in $(\mathbf{c}_k, \mathbf{L}_k)$ and $\boldsymbol{\Omega}_k$ is independent of $(\mathbf{c}_k, \mathbf{L}_k)$ (as will be shown by (6.16)), it is clear from (6.13a,b) and (6.14a,b) that the first moments of the predicted state and input sequences are linear in $\theta_k = (\mathbf{c}_k, \mathbf{L}_k)$ while their second moments are quadratic functions of θ_k.

To determine $\boldsymbol{\Omega}_k$, note that $\mathbf{x}_k - \hat{\mathbf{x}}_k$ and $\boldsymbol{\zeta}_k$ can be written as

$$\mathbf{x}_k - \hat{\mathbf{x}}_k = F(\boldsymbol{\Gamma}_k) q_k, \quad \boldsymbol{\zeta}_k = G(\boldsymbol{\Gamma}_k) q_k, \quad q_k := \begin{bmatrix} x_k - \hat{x}_k \\ \mathbf{v}_k \\ \mathbf{w}_k \end{bmatrix},$$

with $F(\boldsymbol{\Gamma}_k) := \begin{bmatrix} \mathbf{S}_\Psi & -\mathbf{T}_{(\Psi,A)} \mathbf{M} \boldsymbol{\Gamma}_k & \mathbf{T}_{(\Psi,D)} \end{bmatrix}$ and $G(\boldsymbol{\Gamma}_k) := \boldsymbol{\Gamma}_k \mathbf{C} F(\boldsymbol{\Gamma}_k) + \begin{bmatrix} 0 & \boldsymbol{\Gamma}_k & 0 \end{bmatrix}$. Hence, by the law of total expectation,

$$\boldsymbol{\Omega}_k = \sum_j \begin{bmatrix} F(\boldsymbol{\Gamma}^{(j)}) \\ G(\boldsymbol{\Gamma}^{(j)}) \end{bmatrix} \mathbb{E}_k\{q_k q_k^\top\} \begin{bmatrix} F(\boldsymbol{\Gamma}^{(j)}) \\ G(\boldsymbol{\Gamma}^{(j)}) \end{bmatrix}^\top \mathbb{P}\{\boldsymbol{\Gamma}_k = \boldsymbol{\Gamma}^{(j)}\}, \tag{6.16}$$

where $\mathbb{E}_k\{q_k q_k^\top\}$ is the block-diagonal matrix:

$$\mathbb{E}_k\{q_k q_k^\top\} = \mathrm{diag}\{\Sigma_k, \bar{\Sigma}_v, \bar{\Sigma}_w\}, \quad \bar{\Sigma}_v := I_{N\times N} \otimes \Sigma_v, \quad \bar{\Sigma}_w := I_{N\times N} \otimes \Sigma_w,$$

and where $\boldsymbol{\Gamma}^{(j)}$ for $j = 1, \ldots, 2^N$ enumerates the 2^N matrices with binary-valued diagonal elements defined by

$$\begin{aligned} &\boldsymbol{\Gamma}^{(1)} = 0, \quad \boldsymbol{\Gamma}^{(2)} = \mathrm{diag}\{0, \ldots, 0, 1\} \otimes I_{n_y \times n_y} \quad \cdots \\ &\cdots \quad \boldsymbol{\Gamma}^{(2^N - 1)} = \mathrm{diag}\{1, \ldots, 1, 0\} \otimes I_{n_y \times n_y}, \quad \boldsymbol{\Gamma}^{(2^N)} = I. \end{aligned}$$

Remark 6.3. *The matrix $\boldsymbol{\Omega}_k$ in (6.16) can be computed conveniently via*

$$\mathrm{vec}(\boldsymbol{\Omega}_k) = \Big(\sum_j \begin{bmatrix} F(\boldsymbol{\Gamma}^{(j)}) \\ G(\boldsymbol{\Gamma}^{(j)}) \end{bmatrix} \otimes \begin{bmatrix} F(\boldsymbol{\Gamma}^{(j)}) \\ G(\boldsymbol{\Gamma}^{(j)}) \end{bmatrix} \mathbb{P}\{\boldsymbol{\Gamma}_k = \boldsymbol{\Gamma}^{(j)}\} \Big) \mathrm{vec}\Big(\begin{bmatrix} \Sigma_k & & \\ & \bar{\Sigma}_v & \\ & & \bar{\Sigma}_w \end{bmatrix} \Big)$$

where the first term on the RHS can be determined offline given the probability distribution of γ_k. This allows $\boldsymbol{\Omega}_k$ to be computed online using the current value of Σ_k with a single matrix-vector multiplication.

Using the same arguments as the proof of Proposition 6.2, it can be verified that

$$X_{N|k} := \mathbb{E}_k \Big\{ \begin{bmatrix} x_{N|k} - \hat{x}_{N|k} \\ \hat{x}_{N|k} \end{bmatrix} \begin{bmatrix} x_{N|k} - \hat{x}_{N|k} \\ \hat{x}_{N|k} \end{bmatrix}^\top \Big\}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & \pi_{N|k}\pi_{N|k}^\top \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & \Pi_{N|k} \end{bmatrix} \Omega_{N|k} \begin{bmatrix} I & 0 \\ 0 & \Pi_{N|k} \end{bmatrix}^\top$$

where

$$\Pi_{N|k} := T_{(\Phi,B)}^N \mathbf{L}_k + T_{(\Phi,A)}^N \mathbf{M},\ \pi_{N|k} := S_\Phi^N \hat{x}_k + T_{(\Phi,B)}^N \mathbf{c}_k,$$
$$\Omega_{N|k} = \sum_j \begin{bmatrix} F_N(\mathbf{\Gamma}^{(j)}) \\ G(\mathbf{\Gamma}^{(j)}) \end{bmatrix} \left[\begin{smallmatrix} \Sigma_k & & \\ & \bar{\Sigma}_v & \\ & & \bar{\Sigma}_w \end{smallmatrix}\right] \begin{bmatrix} F_N(\mathbf{\Gamma}^{(j)}) \\ G(\mathbf{\Gamma}^{(j)}) \end{bmatrix}^\top \mathbb{P}\{\mathbf{\Gamma}_k = \mathbf{\Gamma}^{(j)}\}$$

with $F_N(\mathbf{\Gamma}_k) := \begin{bmatrix} S_\Psi^N & -T_{(\Psi,A)}^N \mathbf{M}\mathbf{\Gamma}_k & T_{(\Psi,D)}^N \end{bmatrix}$ and S_Φ^N, $T_{(\Phi,A)}^N$, $T_{(\Phi,B)}^N$ being defined (analogously to S_Ψ^N and $T_{(\Psi,B)}^N$) in terms of Φ, A and B.

6.5 Cost and Constraints

We next show that the cost and constraints of (6.4) can be expressed as convex functions of $\theta_k = (\mathbf{c}_k, \mathbf{L}_k)$. First note that the objective in (6.4) can be written as

$$\sum_{i=0}^{\infty} \beta^i \mathbb{E}_k\Big\{\|x_{i|k}\|_Q^2 + \|u_{i|k}\|_R^2\Big\} = \mathrm{tr}(\mathbf{Q}_\beta \mathbf{X}_k) + \mathrm{tr}(\mathbf{R}_\beta \mathbf{U}_k) + f_N(\theta_k, \hat{x}_k, \Sigma_k) \tag{6.17}$$

where $\mathbf{X}_k$ is given by (6.15), and

$$\mathbf{Q}_\beta := \mathbf{1}_{2\times 2} \otimes \mathrm{diag}\{Q, \beta Q, \ldots, \beta^{N-1} Q\}, \quad \mathbf{R}_\beta := \mathrm{diag}\{R, \beta R, \ldots, \beta^{N-1} R\},$$
$$\mathbf{U}_k := \mathbb{E}_k\{\mathbf{u}_k \mathbf{u}_k^\top\}, \quad f_N(\theta_k, \hat{x}_k, \Sigma_k) := \sum_{i=N}^{\infty} \beta^i \mathbb{E}_k\Big\{\|x_{i|k}\|_Q^2 + \|u_{i|k}\|_R^2\Big\}.$$

Since $\mathbf{Q}_\beta \succeq 0$ and $\mathbf{R}_\beta \succ 0$, the term $\mathrm{tr}(\mathbf{Q}_\beta \mathbf{X}_k) + \mathrm{tr}(\mathbf{R}_\beta \mathbf{U}_k)$ in (6.17) can be expressed as a convex quadratic function of $\theta_k = (\mathbf{c}_k, \mathbf{L}_k)$ using (6.14b) and (6.15). To determine the terminal term, $f_N(\theta, \hat{x}_k, \Sigma_k)$, let

$$P_k := \sum_{i=N}^{\infty} \beta^i X_{i|k}, \tag{6.18}$$

where

$$X_{i|k} := \mathbb{E}_k\left\{ \begin{bmatrix} x_{i|k} - \hat{x}_{i|k} \\ \hat{x}_{i|k} \end{bmatrix} \begin{bmatrix} x_{i|k} - \hat{x}_{i|k} \\ \hat{x}_{i|k} \end{bmatrix}^\top \right\}.$$

Then, for $i \geq N$, we have

$$X_{i+1|k} = \mathbb{E}\Big\{\tilde{\Psi}(\gamma) X_{i|k} \tilde{\Psi}^\top(\gamma)\Big\} + \mathbb{E}\Big\{\tilde{D}(\gamma) \left[\begin{smallmatrix} \Sigma_v & \\ & \Sigma_w \end{smallmatrix}\right] \tilde{D}^\top(\gamma)\Big\} \tag{6.19}$$

where

$$\tilde{\Psi}(\gamma) := \begin{bmatrix} A(I-\gamma MC) & 0 \\ \gamma AMC & \Phi \end{bmatrix}, \quad \tilde{D}(\gamma) := \begin{bmatrix} -\gamma AM & D \\ \gamma AM & 0 \end{bmatrix},$$

and γ is a random variable identically distributed as γ_k. Hence

$$\begin{aligned} \mathbb{E}\Big\{\tilde{\Psi}(\gamma)P_k\tilde{\Psi}^\top(\gamma)\Big\} &= \sum_{i=N}^{\infty}\beta^i\Big(X_{i+1|k} - \mathbb{E}\{\tilde{D}(\gamma)\big[\begin{smallmatrix}\Sigma_v & \\ & \Sigma_w\end{smallmatrix}\big]\tilde{D}^\top(\gamma)\}\Big) \\ &= \beta^{-1}(P_k - \beta^N X_{N|k}) - \frac{\beta^N}{1-\beta}\mathbb{E}\{\tilde{D}(\gamma)\big[\begin{smallmatrix}\Sigma_v & \\ & \Sigma_w\end{smallmatrix}\big]\tilde{D}^\top(\gamma)\}, \end{aligned}$$

and the terminal term $f_N(\theta_k, \hat{x}_k, \Sigma_k)$ in (6.17) is equal to $\mathrm{tr}(Z_1 P_k)$ where $Z_1 := \left[\begin{smallmatrix} Q & Q \\ Q & Q+K^\top RK\end{smallmatrix}\right]$ and P_k is the solution to the stochastic Lyapunov equation

$$P_k = \beta\mathbb{E}\Big\{\tilde{\Psi}(\gamma)P_k\tilde{\Psi}^\top(\gamma)\Big\} + \beta^N X_{N|k} + \frac{\beta^{N+1}}{1-\beta}\mathbb{E}\Big\{\tilde{D}(\gamma)\Big[\begin{smallmatrix}\Sigma_v & \\ & \Sigma_w\end{smallmatrix}\Big]\tilde{D}^\top(\gamma)\Big\}. \tag{6.20}$$

Lemma 6.4. *By Assumption 6.3, the linear system $\xi_{i+1} = \tilde{\Psi}(\gamma)\xi_i$ is mean-square stable.*

Let $\bar{X} \succeq 0$ denote the steady state solution to (6.19). Then, by Lemma 6.4, it is ensured that $X_{i|k}$ converges to $\bar{X}$ as $i \to \infty$ for any $X_{N|k} \succeq 0$ and $k \geq 0$. Also, Lemma 6.4 implies $\beta^{\frac{1}{2}}\tilde{\Psi}(\gamma)$ is MSS since $\beta \in (0,1)$ and thus P_k in (6.18) is well defined and finite.

Re-writing the constraints of problem (6.4) using the matrix, $X_{i|k}$, of second moments yields the condition

$$\sum_{i=0}^{\infty}\beta^i\,\mathrm{tr}\Big((\mathbf{1}_{2\times 2}\otimes H^\top H)X_{i|k}\Big) \leq \mu_k,$$

which is equivalent to the constraint

$$\mathrm{tr}(\mathbf{H}_\beta\mathbf{X}_k) + \mathrm{tr}(Z_2 P_k) \leq \mu_k, \tag{6.21}$$

where $\mathbf{H}_\beta := \mathbf{1}_{2\times 2}\otimes\mathrm{diag}\{H^\top H, \beta H^\top H, \ldots, \beta^{N-1}H^\top H\}$ and $Z_2 := \mathbf{1}_{2\times 2}\otimes H^\top H$.

The expressions for the cost and constraints in (6.17), (6.20), (6.21) allow the optimisation (6.4) defining θ_k^* to be formulated as

$$\begin{aligned} (\theta_k^*, P_k^*) = \arg\min_{\theta_k, P_k}\ & \mathrm{tr}(\mathbf{Q}_\beta\mathbf{X}_k) + \mathrm{tr}(\mathbf{R}_\beta\mathbf{U}_k) + \mathrm{tr}(Z_1 P_k) \\ \text{s.t. } & (6.21), \\ & P_k \succeq \beta\mathbb{E}\Big\{\tilde{\Psi}(\gamma)P_k\tilde{\Psi}^\top(\gamma)\Big\} + \beta^N X_{N|k} \\ & \quad + \frac{\beta^{N+1}}{1-\beta}\mathbb{E}\Big\{\tilde{D}(\gamma)\Big[\begin{smallmatrix}\Sigma_v & \\ & \Sigma_w\end{smallmatrix}\Big]\tilde{D}^\top(\gamma)\Big\}. \end{aligned} \tag{6.22}$$

Remark 6.5. *The relaxation of (6.20) as a linear matrix inequality (LMI) in problem* (6.22) *does not introduce any conservativeness as it can be shown that there always exists a solution that satisfies this LMI with equality. Suppose there does not exist such a solution to problem* (6.22)*. Let* $(\theta_k^*, P_k^{(0)})$ *be a minimiser and*

$$P_k^{(1)} = \beta\mathbb{E}\Big\{\tilde{\Psi}(\gamma)P_k^{(0)}\tilde{\Psi}^\top(\gamma)\Big\} + \beta^N X_{N|k} + \frac{\beta^{N+1}}{1-\beta}\mathbb{E}\Big\{\tilde{D}(\gamma)\begin{bmatrix}\Sigma_v & \\ & \Sigma_w\end{bmatrix}\tilde{D}^\top(\gamma)\Big\}.$$

Then $P_k^{(0)} \succ P_k^{(1)}$ *and* $(\theta_k^*, P_k^{(1)})$ *also satisfies the LMI with strict inequality because of the supposition. Continuing this procedure, we will have that* $\forall i \geq 1$ $P_k^{(i)} \succ P_k^{(i+1)}$ *and that* $(\theta_k^*, P_k^{(i+1)})$ *is a minimiser and is at least as good as* $(\theta_k^*, P_k^{(i)})$ *in terms of values of the cost function. From the mean-square stability of* $\beta^{\frac{1}{2}}\tilde{\Psi}(\gamma)$*, it follows that* $P_k^* := \lim_{i\to\infty} P_k^{(i)}$ *exists and satisfies this LMI with equality. This contradicts the supposition and proves our argument. Therefore, we let* (θ_k^*, P_k^*) *be the minimiser that satisfies this LMI constraint with equality and* $\theta_k^* = (\mathbf{c}_k^*, \mathbf{L}_k^*)$ *where* $\mathbf{c}_k^*$ *and* $\mathbf{L}_k^*$ *are the corresponding collections of perturbation vectors and feedback gains. Note that if* $\begin{bmatrix}Q^{\frac{1}{2}} & 0 \\ Q^{\frac{1}{2}} & K^\top R^{\frac{1}{2}}\end{bmatrix}$ *has full column rank,* (θ_k^*, P_k^*) *is unique.*

6.6 Closed Loop Properties

This section considers the performance of the closed loop system (6.1) with the control law of Section 6.3.2. We use the solution $\theta_k^* = \{\mathbf{c}_k^*, \mathbf{L}_k^*\}$ of (6.4) at time k to construct a feasible, but possibly suboptimal, solution for (6.4) at time $k+1$ (i.e. given $\mathcal{U}_k$, $\mathcal{I}_k$), which we denote as $\theta_{k+1}^\circ := \{\mathbf{c}_{k+1}^\circ, \mathbf{L}_{k+1}^\circ\}$, where

$$\mathbf{c}_{k+1}^\circ := \begin{bmatrix} c_{1|k}^* \\ \vdots \\ c_{N-1|k}^* \\ 0 \end{bmatrix} + \begin{bmatrix} L_{1,0|k}^* \\ \vdots \\ L_{N-1,0|k}^* \\ 0 \end{bmatrix} \gamma_k(y_k - C\hat{x}_k), \tag{6.23a}$$

$$\mathbf{L}_{k+1}^\circ := \begin{bmatrix} L_{1,1|k}^* & & & \\ \vdots & \ddots & & \\ L_{N-1,1|k}^* & \cdots & L_{N-1,N-1|k}^* & \\ 0 & \cdots & 0 & 0 \end{bmatrix}. \tag{6.23b}$$

Following [44], we define the constraint threshold μ_k in (6.4) for all $k > 0$ in terms of θ_k°. This ensures recursive feasibility of the MPC optimisation without requiring bounds on the noise v_k and disturbance w_k. Thus

$$\mu_k := \begin{cases} \epsilon, & k = 0 \\ \mathrm{tr}(\mathbf{H}_\beta \mathbf{X}_k^\circ) + \mathrm{tr}(Z_2 P_k^\circ), & k > 0 \end{cases} \tag{6.24}$$

where

$$\begin{aligned}
P_k^\circ &:= \beta \mathbb{E}\Big\{\tilde{\Psi}(\gamma) P_k^\circ \tilde{\Psi}^\top(\gamma)\Big\} + \beta^N X_{N|k}^\circ + \frac{\beta^{N+1}}{1-\beta}\mathbb{E}\Big\{\tilde{D}(\gamma)\begin{bmatrix}\Sigma_v & \\ & \Sigma_w\end{bmatrix}\tilde{D}^\top(\gamma)\Big\}, \\
\mathbf{X}_k^\circ &:= \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{\pi}_k^\circ \boldsymbol{\pi}_k^{\circ\top} \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & \mathbf{\Pi}_k^\circ \end{bmatrix} \mathbf{\Omega}_k \begin{bmatrix} I & 0 \\ 0 & \mathbf{\Pi}_k^\circ \end{bmatrix}^\top, \\
X_{N|k}^\circ &:= \begin{bmatrix} 0 & 0 \\ 0 & \pi_{N|k}^\circ \pi_{N|k}^{\circ\top} \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & \Pi_{N|k}^\circ \end{bmatrix} \Omega_{N|k} \begin{bmatrix} I & 0 \\ 0 & \Pi_{N|k}^\circ \end{bmatrix}^\top
\end{aligned}$$

with

$$\begin{aligned}
\boldsymbol{\pi}_k^\circ &:= \mathbf{S}_\Phi \hat{x}_k + \mathbf{T}_{(\Phi,B)} \mathbf{c}_k^\circ, & \mathbf{\Pi}_k^\circ &:= \mathbf{T}_{(\Phi,B)} \mathbf{L}_k^\circ + \mathbf{T}_{(\Phi,A)} \mathbf{M}, \\
\pi_{N|k}^\circ &:= S_\Phi^N \hat{x}_k + T_{(\Phi,B)}^N \mathbf{c}_k^\circ, & \Pi_{N|k}^\circ &:= T_{(\Phi,B)}^N \mathbf{L}_k^\circ + T_{(\Phi,A)}^N \mathbf{M}.
\end{aligned}$$

The main results of this section are given next, demonstrating constraint satisfaction and satisfaction of a stability condition in closed loop.

Theorem 6.6. *If problem* (6.22) *is feasible at* $k = 0$*, then* (6.22) *remains feasible for all* $k > 0$ *and the state of (6.1) under the control law of Section 6.3.2 satisfies*

$$\sum_{k=0}^{\infty} \beta^k \mathbb{E}\{\|H x_k\|^2\} \le \epsilon. \tag{6.25}$$

Corollary 6.7. *Let* $J_k := J(\theta_k^*, \hat{x}_k, \Sigma_k)$ *denote the optimal cost of problem* (6.22)*. Then, under the control law of Section 6.3.2, the trajectories of (6.1) satisfy*

$$\sum_{k=0}^{\infty} \beta^k \mathbb{E}\Big\{\|x_k\|_Q^2 + \|u_k\|_R^2\Big\} \le J_0. \tag{6.26}$$

6.7 Strengthened Stability Conditions

In Corollary 6.7, we provide an upper bound on the closed loop cost that is discounted, while this may not prevent closed loop states from gradually diverging to infinity if the closed loop cost is heavily discounted. Therefore, in the following, we provide stability conditions in terms of bounds on the closed loop cost that is undiscounted.

6.7.1 Closed loop cost bound via a numerical check

In this section, we derive a bound on the averaged undiscounted closed loop cost accumulated over an infinite horizon by running a numerical check on β. Specifically, if some condition holds for values of β in $(0, 1)$, we can provide a bound that is parametric in β on this closed loop cost. To this end, we first establish an upper bound on J_k.

An upper bound on J_k can be achieved by finding a feasible solution to problem (6.22), and such a solution can be easily obtained if we eliminate variable P_k and rewrite problem (6.22) as follows. Using standard matrix vectorisation identities, we rewrite (6.20) as

$$\operatorname{vec}(P_k) = W_1 \operatorname{vec}(\beta^N X_{N|k} + W_2), \tag{6.27}$$

where

$$\begin{aligned} W_1 :=& \Big[I - \beta(1-\lambda)\tilde{\Psi}(0) \otimes \tilde{\Psi}(0) - \beta\lambda\tilde{\Psi}(1) \otimes \tilde{\Psi}(1)\Big]^{-1}, \\ W_2 :=& \frac{\beta^{N+1}}{1-\beta}\mathbb{E}\Big\{\tilde{D}(\gamma)\begin{bmatrix} \Sigma_v & \\ & \Sigma_w \end{bmatrix}\tilde{D}^\top(\gamma)\Big\}. \end{aligned}$$

Note that in the definition of W_1, the matrix inverse is guaranteed to exist since $\beta^{\frac{1}{2}}\tilde{\Psi}(\gamma)$ is MSS. From (6.27) we have

$$\begin{aligned} \operatorname{tr}(Z_1 P_k) = \operatorname{vec}(Z_1^\top)^\top \operatorname{vec}(P_k) &= \operatorname{vec}(Z_1)^\top W_1 \operatorname{vec}(\beta^N X_{N|k} + W_2) \\ &= \operatorname{tr}(\beta^N \tilde{Z}_1 X_{N|k}) + \operatorname{tr}(\tilde{Z}_1 W_2), \end{aligned} \tag{6.28}$$

where $\tilde{Z}_1^\top := \operatorname{mat}\Big((\operatorname{vec}(Z_1)^\top W_1)^\top, [2n_x, 2n_x]\Big)$. Similarly,

$$\operatorname{tr}(Z_2 P_k) = \operatorname{tr}(\beta^N \tilde{Z}_2 X_{N|k}) + \operatorname{tr}(\tilde{Z}_2 W_2),$$

where $\tilde{Z}_2^\top := \operatorname{mat}\Big((\operatorname{vec}(Z_2)^\top W_1)^\top, [2n_x, 2n_x]\Big)$. Therefore, an equivalent form of problem (6.22) is given by

$$\begin{aligned} \theta_k^* = \arg\min_{\theta_k} & \operatorname{tr}(\mathbf{Q}_\beta \mathbf{X}_k) + \operatorname{tr}(\mathbf{R}_\beta \mathbf{U}_k) + \beta^N \operatorname{tr}(\tilde{Z}_1 X_{N|k}) + \operatorname{tr}(\tilde{Z}_1 W_2) \\ \text{s.t. } & \operatorname{tr}(\mathbf{H}_\beta \mathbf{X}_k) + \beta^N \operatorname{tr}(\tilde{Z}_2 X_{N|k}) + \operatorname{tr}(\tilde{Z}_2 W_2) \le \mu_k, \end{aligned} \tag{6.29}$$

which has a single convex quadratic constraint.

Minimising the constraint function in (6.29) over $\theta_k = \{\mathbf{c}_k, \mathbf{L}_k\}$ necessarily yields a feasible, but possibly suboptimal, solution $\theta_k^f := \{\mathbf{c}_k^f, \mathbf{L}_k^f\}$. This solution has an explicit form, in which $\mathbf{c}_k^f$ is given by $\mathbf{c}_k^f := K_c \hat{x}_k$ with

$$K_c := -\left[\mathbf{T}_{(\Phi,B)}^\top H_b \mathbf{T}_{(\Phi,B)} + \beta^N (T_{(\Phi,B)}^N)^\top \tilde{Z}_{2,22}^\top T_{(\Phi,B)}^N\right]^\dagger \cdot \left[\mathbf{T}_{(\Phi,B)}^\top H_b \mathbf{S}_\Phi + \frac{\beta^N}{2}(T_{(\Phi,B)}^N)^\top (\tilde{Z}_{2,22}^\top + \tilde{Z}_{2,22}) S_\Phi^N\right]$$

and $H_b := \operatorname{diag}\{H^\top H, \beta H^\top H, \dots, \beta^{N-1} H^\top H\}$. The matrix $\tilde{Z}_{2,22}$ is the bottom-right block of $\tilde{Z}_2$ partitioned as $\tilde{Z}_2 = \begin{bmatrix} \tilde{Z}_{2,11} & \tilde{Z}_{2,12} \\ \tilde{Z}_{2,21} & \tilde{Z}_{2,22} \end{bmatrix}$. The expression for $\mathbf{L}_k^f$ is omitted here for simplicity but we note that it is a function of Σ_k. Therefore, evaluating the cost function at $\theta_k^f = \{\mathbf{c}_k^f, \mathbf{L}_k^f\}$ gives that

$$J(\theta_k^f, \hat{x}_k, \Sigma_k) = \hat{x}_k^\top P_{\hat{x}} \hat{x}_k + \operatorname{tr}(Q\Sigma_k) + s(\beta, \Sigma_k) + \operatorname{tr}(\tilde{Z}_1 W_2),$$

where $s(\beta, \Sigma_k)$ is a nonnegative scalar function of β and Σ_k, taking finite values when Σ_k is finite, and $\operatorname{tr}(\tilde{Z}_1 W_2)$ is a constant depending on β. The matrix $P_{\hat{x}}$ is given by

$$\begin{aligned} P_{\hat{x}} := {} & (\mathbf{S}_\Phi + \mathbf{T}_{(\Phi,B)} K_c)^\top Q_b (\mathbf{S}_\Phi + \mathbf{T}_{(\Phi,B)} K_c) \\ & + (\mathbf{K}(\mathbf{S}_\Phi + \mathbf{T}_{(\Phi,B)} K_c) + K_c)^\top \mathbf{R}_\beta (\mathbf{K}(\mathbf{S}_\Phi + \mathbf{T}_{(\Phi,B)} K_c) + K_c) \\ & + \frac{\beta^N}{2} (S_\Phi^N + T_{(\Phi,B)}^N K_c)^\top (\tilde{Z}_{1,22}^\top + \tilde{Z}_{1,22})(S_\Phi^N + T_{(\Phi,B)}^N K_c) \end{aligned}$$

where $Q_b := \operatorname{diag}\{Q, \beta Q, \dots, \beta^{N-1} Q\}$ and $\tilde{Z}_{1,22}$ is the bottom-right block of $\tilde{Z}_1 = \begin{bmatrix} \tilde{Z}_{1,11} & \tilde{Z}_{1,12} \\ \tilde{Z}_{1,21} & \tilde{Z}_{1,22} \end{bmatrix}$. Moreover, if there exists a scalar σ such that

$$P_{\hat{x}} \preceq \sigma Q, \tag{6.30}$$

then we have

$$\begin{aligned} J_k \le J(\theta_k^f, \hat{x}_k, \Sigma_k) & \le \sigma \operatorname{tr}(Q \hat{x}_k \hat{x}_k^\top) + \operatorname{tr}(Q\Sigma_k) + s(\beta, \Sigma_k) + \operatorname{tr}(\tilde{Z}_1 W_2) \\ & \le \sigma \mathbb{E}_k\{\|x_k\|_Q^2\} + s(\beta, \Sigma_k) + \operatorname{tr}(\tilde{Z}_1 W_2), \end{aligned} \tag{6.31}$$

where the first inequality holds by optimality. From the definitions of $P_{\hat{x}}$, $\mathbf{S}_\Phi$ and Q_b, we have $P_{\hat{x}} \succeq Q$. This implies that, if a scalar σ exists such that (6.30) is satisfied, then it must be greater than or equal to 1. Thus the third inequality in (6.31) follows. A possible choice of σ is the largest eigenvalue of $P_{\hat{x}}Q^{-1}$ if $Q \succ 0$.

We are now able to identify a parametric bound on the averaged undiscounted closed loop cost accumulated over an infinite horizon in Theorem 6.8.

Theorem 6.8. *Provided that optimisation* (6.22) *is feasible at time* $k = 0$ *and a scalar* σ *exists such that* (6.30) *is satisfied, if the discount factor* $\beta \in (0, 1)$ *in optimisation* (6.22) *satisfies*

$$\sigma < \frac{1}{1-\beta} \tag{6.32}$$

and the random sequence $\{\Sigma_k\}_{k=0}^{\infty}$ *is upper bounded by some matrix* $\bar{\Sigma}$*, then*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\Big\{\|x_k\|_Q^2 + \|u_k\|_R^2\Big\} \leq \frac{1}{\frac{1}{1-\beta} - \sigma}\Big(\sup_{0\preceq\Sigma\preceq\bar{\Sigma}} s(\beta, \Sigma) + \operatorname{tr}\Big(\tilde{Z}_1 W_2\Big)\Big). \tag{6.33}$$

Remark 6.9. *Theorem 6.8 provides a sufficient condition for* (6.33)*, which requires* $\Sigma_k \preceq \bar{\Sigma}$ $\forall k \geq 0$*. This requirement is reasonable in real world applications, while mean-square stability is not enough to guarantee it. This can be ensured, for example, if a common quadratic Lyapunov equation [139] exists for the subsystems of* $\xi_{i+1} = A(I - \gamma_i MC)\xi_i$*, that is, if there exists a matrix* $P = P^\top \succ 0$ *such that*

$$\begin{aligned} &P - A^\top P A \succ 0, \\ &P - (A - AMC)^\top P (A - AMC) \succ 0 \end{aligned} \tag{6.34}$$

hold simultaneously. Condition (6.34) *implies that* (6.12) *is input-to-state stable and since* w_k *and* v_k *have finite second moments, the sequence* $\{\Sigma_k\}_{k=0}^{\infty}$ *is necessarily upper bounded.*

6.7.2 Closed loop cost bound via asymptotic limits

In this section, we consider different discount factors β_1 and β_2 in the objective and the constraint in problem (6.2). The corresponding online optimisation problem defining θ_k^* then becomes

$$\min_{\theta_k, P_{\beta_1,k}, P_{\beta_2,k}} \operatorname{tr}(\mathbf{Q}_{\beta_1}\mathbf{X}_k) + \operatorname{tr}(\mathbf{R}_{\beta_1}\mathbf{U}_k) + \operatorname{tr}(Z_1 P_{\beta_1,k})$$

$$
\begin{aligned}
\text{s.t. } & \operatorname{tr}(\mathbf{H}_{\beta_2}\mathbf{X}_k) + \operatorname{tr}(Z_2 P_{\beta_2,k}) \le \mu_k, \\
& P_{\beta_1,k} \succeq \beta_1 \mathbb{E}\Big\{\tilde{\Psi}(\gamma) P_{\beta_1,k} \tilde{\Psi}^\top(\gamma)\Big\} + \beta_1^N X_{N|k} + \frac{\beta_1^{N+1}}{1-\beta_1}\mathbb{E}\Big\{\tilde{D}(\gamma)\Big[\begin{smallmatrix}\Sigma_v & \\ & \Sigma_w\end{smallmatrix}\Big]\tilde{D}^\top(\gamma)\Big\}, \\
& P_{\beta_2,k} \succeq \beta_2 \mathbb{E}\Big\{\tilde{\Psi}(\gamma) P_{\beta_2,k} \tilde{\Psi}^\top(\gamma)\Big\} + \beta_2^N X_{N|k} + \frac{\beta_2^{N+1}}{1-\beta_2}\mathbb{E}\Big\{\tilde{D}(\gamma)\Big[\begin{smallmatrix}\Sigma_v & \\ & \Sigma_w\end{smallmatrix}\Big]\tilde{D}^\top(\gamma)\Big\},
\end{aligned} \tag{6.35}
$$

where $\mathbf{Q}_{\beta_1}$, $\mathbf{R}_{\beta_1}$, $\mathbf{H}_{\beta_2}$ are constant matrices similarly constructed to $\mathbf{Q}_{\beta}$, $\mathbf{R}_{\beta}$, $\mathbf{H}_{\beta}$. To analyse the averaged undiscounted closed loop cost, we keep the discount factor β_2 fixed, while taking the left-hand limit of the discount factor in the cost at $\beta_1 = 1^-$ and then solving the optimisation (6.35). Given that the limit at $\beta_1 = 1^-$ implies $\beta_1 < 1$, hence (6.35) remains solvable.

We next give a bound on the closed loop cost.

Theorem 6.10. *Provided the optimisation problem* (6.35) *is feasible at time* $k = 0$, *if the discount factor* $\beta_1 \in (0, 1)$ *in* (6.35) *is arbitrarily close to* 1, *then the trajectories of* (6.1) *under the control law of Section 6.3.2 satisfy*

$$
\lim_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} \mathbb{E}\Big\{\|x_k\|_Q^2 + \|u_k\|_R^2\Big\} \le \operatorname{tr}(Z_1 \bar{X}). \tag{6.36}
$$

The analysis in Theorem 6.10 provides some insights into the role that the discount factors play. With discount factors $\beta_1, \beta_2 \in (0, 1)$, we not only ensure the objective and constraint in problem (6.2) are well-defined under possibly unbounded disturbances and measurement noise, but we also obtain a trade-off between transient and steady state behaviours. When β_1 takes values far from 1, greater emphasis is put on the cost accumulated over a short horizon near the initial time, and the cost corresponding to later times is more heavily discounted. On the other hand, as β_1 approaches 1, the closed loop cost is dominated by steady state behaviours.

6.8 Numerical Example

In this section, we run three sets of simulations: (A) demonstrates that the closed loop system satisfies the bounds in (6.25) and (6.26), and compares performance with the unconstrained optimal LQG controller; (B) shows that the averaged undiscounted closed loop cost accumulated over an infinite horizon is finite under the conditions of Theorem 6.8; and (C) demonstrates that the closed loop system satisfies (6.36).

Simulation A : We consider a system obtained by discretising a linearised continuous-time model of a double inverted pendulum with a sample time of 0.01 s as in [140]. The system matrices are

$$A = \begin{bmatrix} 1.0005 & 0.01 & -0.0005 & 0 \\ 0.098 & 1.0005 & -0.0981 & -0.0005 \\ -0.0005 & 0 & 1.0015 & 0.01 \\ -0.0981 & -0.0005 & 0.2942 & 1.0015 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.0001 & -0.0001 \\ 0.01 & -0.02 \\ -0.0001 & 0.0003 \\ -0.02 & 0.05 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad D = I,$$

and $\omega_k \sim \mathcal{N}(0, \Sigma_w)$, $v_k \sim \mathcal{N}(0, \Sigma_v)$, $\lambda = 0.6$. Here $\Sigma_w = \operatorname{diag}\{0.5, 0.2, 0.9, 0.3\}$ and $\Sigma_v = 1.1I$. Initial conditions are given by

$$x_0 = \begin{bmatrix} -0.8 \\ 0.4 \\ 0.55 \\ -0.5 \end{bmatrix}, \quad \hat{x}_0 = \begin{bmatrix} 0.1 \\ 0.05 \\ 0.1 \\ 0.05 \end{bmatrix}, \quad \Sigma_0 = \begin{bmatrix} 0.5 & -0.5 & -0.5 & 0.5 \\ -0.5 & 0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & 0.5 & -0.5 \\ 0.5 & -0.5 & -0.5 & 0.5 \end{bmatrix}.$$

The constraint of (6.2) is defined by $\beta = 0.8$, $\epsilon = 2$ and

$$H = \begin{bmatrix} 0 & 0.1 & 0 & -0.1 \\ 0.1 & 0 & -0.1 & 0 \end{bmatrix}.$$

The weighting matrices in the cost function of (6.2) are given by $R = 0.01I$, $Q = \operatorname{diag}\{10, 0.1, 10, 0.1\}$. We choose a prediction horizon $N = 5$, K as the unconstrained LQ-optimal, K_{LQ}, with respect to (A, B, Q, R), and $M = \hat{\Sigma}C^\top(C\hat{\Sigma}C^\top + \Sigma_v)^{-1}$, where $\hat{\Sigma}$ is the solution of the algebraic Riccati equation

$$\hat{\Sigma} = A\hat{\Sigma}A^\top + \Sigma_w - \lambda A\hat{\Sigma}C^\top(C\hat{\Sigma}C^\top + \Sigma_v)^{-1}C\hat{\Sigma}A^\top.$$

For this system, cost and constraint, the solution of (6.22) yields $J_0 = 9.0757 \times 10^5$.

To verify (6.25) and (6.26) experimentally, we consider the average cost and constraint values over 10^3 simulations, each of which has a run time of 150 time steps. For comparison, we run the same number of simulations with the same $\{\omega_k\}$, $\{v_k\}$, $\{\gamma_k\}$ sequences using the unconstrained optimal LQG controller, where $u_k = K_{LQ}\hat{x}_k$

and the estimator gain is time-varying and given by $M = \Sigma_k C^\top (C\Sigma_k C^\top + \Sigma_v)^{-1}$. Here Σ_k evolves as

$$\Sigma_{k+1} = A\Sigma_k A^\top + \Sigma_w - \gamma_k A\Sigma_k C^\top (C\Sigma_k C^\top + \Sigma_v)^{-1} C\Sigma_k A^\top.$$

The results summarised in Table 6.1 agree with the bounds in (6.25) and (6.26) and show that, although the LQG controller gives a smaller closed loop cost (as expected), it violates the constraint. Note that $\beta^{150} = 2.9074 \times 10^{-15}$, so a further increase in the simulation run time has negligible effect on the cost and constraint estimates.

	MPC controller	LQG controller
empirical cost value	$2.0545 \times 10^4 < J_0$	893.5569
empirical constraint value	$1.5175 < \epsilon$	$3.9272 > \epsilon$

Table 6.1: Average discounted cost and constraint values for Simulation A

Simulation B: For this simulation, the system model, cost and constraint parameters are the same as for Simulation A, except that the state matrix is redefined as $A \leftarrow A - 1.175I$ and the unconstrained LQ-optimal feedback gain K_{LQ}, the steady state Kalman filter gain M are modified accordingly. The modified matrices A and M satisfy (6.34) and there exists a scalar $\sigma = 4.2759$ such that (6.30) holds and $\sigma < \frac{1}{1-\beta} = 5$. To estimate empirically the LHS of (6.33), we run 10^2 simulations, each of which has a run time of 10^4 time steps, and thus obtain an estimate of the average value as 14.5375. For this example therefore, the LHS of (6.33) is finite as implied by Theorem 6.8. Further increases in the simulation run time cause negligible changes in this estimate.

Simulation C: In this simulation, we consider different discount factors β_1 and $\beta_2 (= \beta = 0.8)$ in problem (6.2) and we set $\epsilon = 3.8$. The weighting matrix for control inputs is given by $R = 0.001I$ and the unconstrained LQ-optimal feedback gain, K_{LQ}, changes accordingly. All other model and problem parameters are the same as for Simulation A. We solve the steady state equation of (6.19) and obtain the RHS of (6.36) as 518.3913. To estimate empirically the LHS of (6.36) as β_1 approaches 1 while keeping β_2 fixed, we run three tests with β_1 equal to 0.98, 0.99

and 0.999, respectively. Each test consists of 100 simulations, each of which has a run time of 10^4 time steps. Empirical values of the LHS of (6.36) corresponding to different values of β_1 are summarised in Table 6.2, and are in agreement with the bound (6.36). Note also that a further increase of the simulation run time only results in small variations in these estimates.

	$\beta_1 = 0.98$	$\beta_1 = 0.99$	$\beta_1 = 0.999$	$\mathrm{tr}(Z_1 \bar{X})$
empirical cost value	483.3557	472.8744	468.7458	518.3913

Table 6.2: Average undiscounted cost values for Simulation C

6.9 Conclusion

This chapter describes an output feedback MPC algorithm for linear discrete-time systems with additive disturbances and noisy sensor measurements transmitted over a packet-dropping communication channel. By designing a control policy with an affine dependence on future observations, we provide a convex formulation of a stochastic quadratic regulation problem subject to a discounted expectation constraint. Our controller parameterisation guarantees recursive feasibility of the MPC optimisation problem and ensures constraint satisfaction and a discounted cost bound in closed loop operation. We provide a sufficient condition on the discount factor to ensure that the averaged undiscounted closed loop cost is finite. Also, different discount factors in the cost and the expectation constraint, respectively, are considered. By allowing the discount factor in the cost to approach 1 arbitrarily closely while keeping the discount factor in the expectation constraint fixed, we obtain an upper bound on the averaged undiscounted closed loop cost. Future work will investigate the impact of uncertainties in the probability, λ, of successfully receiving a sensor measurement on closed loop properties.

6.10 Proofs of Chapter 6

6.10.1 Proof of Lemma 6.4

In Assumption 6.3, it is assumed that $\Psi(\gamma)$ is MSS and Φ is asymptotically stable, which are equivalent to

$$\exists\ \Xi_1 = \Xi_1^\top \succ 0, \Xi_1 - \mathbb{E}\{\Psi(\gamma)\Xi_1\Psi^\top(\gamma)\} \succ 0,$$
$$\exists\ \Xi_2 = \Xi_2^\top \succ 0, \Xi_2 - \Phi\Xi_2\Phi^\top \succ 0,$$

respectively. It can be shown that $\begin{bmatrix}\Xi_1 & \\ & \Xi_2\end{bmatrix}$ satisfies

$$\begin{bmatrix}\Xi_1 & \\ & \Xi_2\end{bmatrix} - \mathbb{E}\{\tilde{\Psi}(\gamma)\begin{bmatrix}\Xi_1 & \\ & \Xi_2\end{bmatrix}\tilde{\Psi}^\top(\gamma)\} \succ 0,$$

and this implies Lemma 6.4.

6.10.2 Proof of Theorem 6.6

The definition (6.24) of μ_k trivially ensures feasibility for all $k > 0$. The definitions (6.23a,b) ensure that, at time k (given $\mathcal{U}_{k-1}$, $\mathcal{I}_{k-1}$), the distributions of the state and control sequences $\{x_{i|k+1}\}_{i=0}^\infty$ and $\{u_{i|k+1}\}_{i=0}^\infty$ are identical to the distributions of $\{x_{i+1|k}\}_{i=0}^\infty$ and $\{u_{i+1|k}\}_{i=0}^\infty$. Therefore

$$\sum_{i=0}^\infty \beta^i \mathbb{E}_k\{\|Hx_{i|k}\|^2\} = \mathrm{tr}(\mathbf{H}_\beta \mathbf{X}_k) + \mathrm{tr}(Z_2 P_k)$$

implies

$$\beta\mathbb{E}_k\{\mu_{k+1}\} \leq \mu_k - \mathbb{E}_k\{\|Hx_{0|k}\|^2\} = \mu_k - \mathbb{E}_k\{\|Hx_k\|^2\}.$$

Hence the trajectories of the closed loop system satisfy

$$\sum_{i=0}^\infty \beta^i \mathbb{E}_k\{\|Hx_{k+i}\|^2\} \leq \mu_k - \lim_{i\to\infty} \beta^i \mathbb{E}_k\{\mu_{k+i}\} \leq \mu_k$$

for all $k \geq 0$.

6.10.3 Proof of Corollary 6.7

Applying the same argument used in the proof of Theorem 6.6 to the definition of the objective in (6.22) yields

$$\beta\mathbb{E}_k\{J(\theta_{k+1}^\circ, \hat{x}_{k+1}, \Sigma_{k+1})\} = J_k - \mathbb{E}_k\{\|x_k\|_Q^2 + \|u_k\|_R^2\}. \tag{6.37}$$

Also, since $J_k \leq J(\theta_k^\circ, \hat{x}_k, \Sigma_k)$ $\forall k$ by optimality, the bound in (6.26) follows.

6.10.4 Proof of Theorem 6.8

From (6.37), it follows that

$$\beta\mathbb{E}_k\{J_{k+1}\} \leq J_k - \mathbb{E}_k\{\|x_k\|_Q^2 + \|u_k\|_R^2\}.$$

Multiplying both sides of this inequality by β^{-1} and then subtracting J_k from both sides, we obtain

$$\mathbb{E}_k\{J_{k+1}\} - J_k \leq (\beta^{-1} - 1)J_k - \beta^{-1}\mathbb{E}_k\{\|x_k\|_Q^2 + \|u_k\|_R^2\}. \tag{6.38}$$

From (6.31) and (6.38), it can be concluded that

$$\begin{aligned}\mathbb{E}_k\{J_{k+1}\} - J_k \leq \Big((\beta^{-1} - 1)\sigma - \beta^{-1}\Big)\mathbb{E}_k\{\|x_k\|_Q^2\} - \beta^{-1}\mathbb{E}_k\{\|u_k\|_R^2\} \\ + (\beta^{-1} - 1)\Big(s(\beta, \Sigma_k) + \operatorname{tr}(\tilde{Z}_1 W_2)\Big).\end{aligned}$$

Since $(\beta^{-1} - 1)\sigma > 0$, we then have

$$\begin{aligned}(\beta^{-1} - (\beta^{-1} - 1)\sigma)\mathbb{E}_k\{\|x_k\|_Q^2 + \|u_k\|_R^2\} \leq J_k - \mathbb{E}_k\{J_{k+1}\} \\ + (\beta^{-1} - 1)\Big(s(\beta, \Sigma_k) + \operatorname{tr}(\tilde{Z}_1 W_2)\Big).\end{aligned} \tag{6.39}$$

Furthermore, (6.32) implies that $\beta^{-1} - (\beta^{-1} - 1)\sigma > 0$. Summing both sides of (6.39) over $k = 0, 1, \ldots, T-1$, dividing by T and considering the limit as $T \to \infty$, we get

$$\lim_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} \mathbb{E}\Big\{\|x_k\|_Q^2 + \|u_k\|_R^2\Big\} \leq \frac{1}{\frac{1}{1-\beta} - \sigma}\Big(\operatorname{tr}(\tilde{Z}_1 W_2) + \lim_{T\to\infty}\frac{1}{T}\sum_{k=0}^{T-1}\mathbb{E}\{s(\beta, \Sigma_k)\}\Big).$$

Also, from that $\Sigma_k \preceq \bar{\Sigma}$ for all $k \geq 0$, it follows

$$\lim_{T\to\infty}\frac{1}{T}\sum_{k=0}^{T-1}\mathbb{E}\{s(\beta, \Sigma_k)\} \leq \sup_{0\preceq\Sigma\preceq\bar{\Sigma}} s(\beta, \Sigma).$$

By the definition of function $s(\beta, \Sigma)$, this supremum is finite, which completes the proof.

6.10.5 Proof of Theorem 6.10

We denote as $(\hat{\theta}_k^*, \hat{P}_k^*, \tilde{P}_k^*)$ the minimiser to (6.35) with β_1 being arbitrarily close to 1 from the left, where $\hat{P}_k^*$ and $\tilde{P}_k^*$ are the terminal matrices in the predicted cost and the discounted sum constraint, respectively, and $\hat{\mathbf{X}}_k^*$, $\hat{\mathbf{U}}_k^*$, $\hat{X}_{i|k}^*$ as the quantity of $\mathbf{X}_k$, $\mathbf{U}_k$, $X_{i|k}$ that correspond to the minimiser $\hat{\theta}_k^*$, respectively. Let $\hat{P}_k^{(T)} := \sum_{i=N}^{N+T-1} \hat{X}_{i|k}^*$. Then we have

$$\hat{P}_k^{(T)} = \sum_{i=N}^{N+T-1} (\hat{X}_{i|k}^* - \bar{X}) + T\bar{X}, \tag{6.40}$$

where $\bar{X}$, the steady state solution to (6.19), exists and is unique by Lemma 6.4. Dividing both sides of (6.40) by T and taking the limit at $T = \infty$, we have

$$\begin{aligned} \lim_{T\to\infty} \frac{1}{T} \hat{P}_k^{(T)} &= \lim_{T\to\infty} \frac{1}{T} \sum_{i=N}^{N+T-1} (\hat{X}_{i|k}^* - \bar{X}) + \bar{X} = \bar{X} && (6.41) \\ &= \lim_{T\to\infty} \frac{1}{T} \sum_{i=N}^{N+T-1} \lim_{\beta_1\to 1^-} \beta_1^i \hat{X}_{i|k}^* = \lim_{T\to\infty} \frac{1}{T} \hat{P}_k^*. && (6.42) \end{aligned}$$

The second equality in (6.41) follows from the facts that $\hat{X}_{i|k}^*$ converges to $\bar{X}$ $\forall k \geq 0$ as $i \to \infty$ and hence $\sum_{i=N}^{\infty}(\hat{X}_{i|k}^* - \bar{X})$ is finite. We next show that the second equality in (6.42) holds. Whenever the minimiser $(\hat{\theta}_k^*, \hat{P}_k^*)$ satisfies its corresponding LMI constraint in problem (6.35) with equality, $\hat{P}_k^*$ satisfies (6.18). Also, since we solve problem (6.35) after taking the limit at $\beta_1 = 1^-$, each term of that infinite sum on the RHS of (6.18) is evaluated at $\beta_1 = 1^-$.

Taking the limit at $\beta_1 = 1^-$, we obtain a slightly different version of (6.37) as

$$\mathbb{E}_k\{J(\hat{\theta}_{k+1}^*, \hat{x}_{k+1}, \Sigma_{k+1})\} \leq J(\hat{\theta}_k^*, \hat{x}_k, \Sigma_k) - \mathbb{E}_k\{\|x_k\|_Q^2 + \|u_k\|_R^2\},$$

and summing both sides of this inequality over $k = 0, \ldots, T-1$ yields that

$$\sum_{k=0}^{T-1} \mathbb{E}\Big\{\|x_k\|_Q^2 + \|u_k\|_R^2\Big\} \leq J(\hat{\theta}_0^*, \hat{x}_0, \Sigma_0) - \mathbb{E}\{J(\hat{\theta}_T^*, \hat{x}_T, \Sigma_T)\}. \tag{6.43}$$

Dividing both sides of (6.43) by T and taking the limit at $T = \infty$, we obtain

$$\begin{aligned} \lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\Big\{\|x_k\|_Q^2 + \|u_k\|_R^2\Big\} &\leq \lim_{T\to\infty} \frac{1}{T}\Big(J(\hat{\theta}_0^*, \hat{x}_0, \Sigma_0) - \mathbb{E}\{J(\hat{\theta}_T^*, \hat{x}_T, \Sigma_T)\}\Big) \\ &\leq \lim_{T\to\infty} \frac{1}{T}\Big(\mathrm{tr}\Big(\mathbf{Q}_{\beta_1}\hat{\mathbf{X}}_0^*\Big) + \mathrm{tr}\Big(\mathbf{R}_{\beta_1}\hat{\mathbf{U}}_0^*\Big) + \mathrm{tr}\Big(Z_1\hat{P}_0^*\Big)\Big) \end{aligned}$$

$$= \lim_{T\to\infty} \frac{1}{T} \operatorname{tr}(Z_1 \hat{P}_0^*) = \operatorname{tr}(Z_1 \lim_{T\to\infty} \frac{1}{T} \hat{P}_0^*).$$

This, together with (6.41) and (6.42), implies (6.36).

7
Conclusion

Contents

A large amount of work in the existing stochastic MPC literature considers discrete-time linear systems with bounded additive disturbances. This thesis takes advantage of a discounted chance constraint to propose a holistic approach for constrained discrete-time linear systems with unbounded additive disturbances. This thesis ensures recursive feasibility, constraint satisfaction, stability conditions (without breaking direct state feedback) and mitigation of conservativeness in constraint handling. Throughout this thesis, there are a number of assumptions made that are necessary for main results to hold, including i.i.d. additive disturbances (and measurement noise) with known and finite first and second moments, controllability and observability (stabilisability and detectability), asymptotic stability of state feedback gains and mean-square stability of estimator gains, and the quadratic form of MPC cost functions. Moreover, although this thesis considers specific types of constraint functions, results can be extended in principle to other types of constraints. We now conclude this thesis by summarising its main contributions and then discussing some directions for future work.

7.1 Contributions of This Thesis

Chapter 3 designs a stochastic MPC algorithm that imposes constraints on the sum of discounted future constraint violation probabilities. It ensures recursive feasibility of the online optimisation problem and constraint satisfaction in closed loop without bounds on the additive disturbance. Key features are the design of a constraint-tightening procedure and closed loop analysis of tightening parameters. This MPC algorithm requires knowledge of the first and second moments of the disturbance, and is implemented as a convex QCQP problem.

Chapter 4 uses the weighted sum method to solve a convex MOP, and this method is able to capture the complete Pareto front in this case. The resulting single objective problems are then solved via dynamic programming iterations that are guaranteed to converge, and their solutions yield strictly stabilising state feedback gains. These gains are used for online dynamic gain selection in Chapter 5, and they coincide with the minimiser of a nonconvex problem, therefore providing an alternative approach for solving this class of problems exactly. Also, it is possible to consider more than two individual objective functions (with different set-points) in the MOP. Provided that these individual objective functions are quadratic, main statements of Theorems 4.1 and 4.2 continue to hold. However, assigning values to the weighting parameters associated with each individual objective requires an important and non-trivial decision, and is closely related to the implementation of online gain selection methods in Chapter 5.

Chapter 5 proposes an improved stochastic MPC algorithm incorporating dynamic gain selection. Gain selection procedures are executed online and lead to extra computation time that is a tiny fraction of the time for solving an MPC optimisation problem. These selection procedures can be configured so that the feedback gain in the MPC law converges to the unconstrained LQ-optimal and the set of feasible initial conditions is enlarged. By using dynamic gain selection, the conservativeness in controller design is considerably reduced and the closed loop cost is improved. Also, recursive feasibility of MPC optimisation problems, constraint satisfaction and closed loop stability are retained using a similar approach to that of Chapter 3.

Chapter 6 describes an output feedback MPC algorithm for linear discrete-time systems with additive disturbances and noisy sensor measurements transmitted over a packet-dropping communication channel. By designing a control policy with an affine dependence on future observations, we provide a convex formulation of a stochastic quadratic regulation problem subject to a discounted expectation constraint. Our controller parameterisation ensures recursive feasibility of the MPC optimisation problem and ensures constraint satisfaction and a discounted cost bound in closed loop operation. We provide a sufficient condition on the discount factor to ensure that the averaged undiscounted closed loop cost is finite. Also, different discount factors in the cost and the expectation constraint, respectively, are considered. By allowing the discount factor in the cost to approach 1 arbitrarily closely while keeping the discount factor in the expectation constraint fixed, we obtain an upper bound on the averaged undiscounted closed loop cost.

7.2 Future Research Directions

Sampling methods to impose chance constraints As discussed in Sections 2.1.2 and 3.1, the scenario approach can handle a wide range of chance constraints, while it is challenging to guarantee recursive feasibility of online optimisation problems in the MPC framework. It is not possible to use a similar constraint-tightening technique to that of Chapter 3 to redefine constraint thresholds, because this would break the i.i.d. assumption on collected samples of uncertainty. However, using a combination of robust optimisation and scenario approach as proposed in [37] can circumvent this issue. With this mixed approach, the scenario approach is used offline, where randomness lies, to obtain an uncertainty set with a certain confidence level, and we obtain a robust MPC problem online. The difficulty is how a sequence of individual probabilistic guarantees at each time step can be connected to yield desirable results given that states are updated at each time step, affected by actual disturbances that can lie outside of the uncertainty set.

Robustness analysis of observation arrival probabilities In exiting literature on Kalman filtering with intermittent observations, many model the data arrivals as Bernoulli processes and the arrival probability is assumed known and denoted as λ. Their main focuses are on the analysis of behaviours of estimation error covariance matrices, while it is worth investigating the impact of a small unknown variation in λ on estimation/control performance. More specifically, if a nominal value λ is used in the estimator and controller design in Chapter 6 and the actual arrival probability is $\lambda + \delta\lambda$ that is unknown, the future work is to study whether a similar version of the closed loop cost bound (6.33) still exists and the resulting changes in this bound from uncertainty $\delta\lambda$.

Discounted cheap discrete-time LQR In Chapter 4, we study a multiobjective optimisation problem and solve its single objective version with different weighting parameters that are restricted to $(0, 1]$. This restriction is because $L_\infty(0)$ is not necessarily stabilising due to the presence of a discount factor. Before analysing its stability, we consider the weighting matrix $C^\top C$. If $C^\top C \succ 0$, then the discounted cheap discrete-time LQR problem becomes a regular one since the discount factor can be translated onto the pair (A, B) as shown in [20]. If $C^\top C$ is only positive semidefinite, then $L_\infty(0)$ is not unique and it is the case worth exploring. We will work on a LQR problem with prioritised objectives, which is similar to the one considered in [141]. With the remaining degree of freedom in $L_\infty(0)$, we minimise the second objective function in problem $(\mathcal{P}_2)$ and necessarily obtain a unique feedback gain, denoted as $L^*_\infty(0)$, that gives the same value for the first objective function in problem $(\mathcal{P}_2)$ as any other $L_\infty(0)$. It is interesting to know under what conditions $L^*_\infty(0)$ is strictly stabilising.

Learning-based MPC Machine learning has been an active research area for the last decade, which uses past data to learn about/act upon the world. The rapid development of machine learning also has a marked impact on the control community and leads to more research interest in data-driven control. In general, learning methods applied to control problems can be categorised into two types:

model-based methods and model-free methods. The former fits a dynamical model to data and the latter could be preferable when learning the system dynamics is too costly. However, [142] shows that, for a class of LQR problems, the model-free methods require considerably more data than model-based methods to achieve the same estimation performance and no method can be more sample efficient than model-based methods.

Reinforcement learning algorithms not only require a massive amount of data to achieve desirable performance, but also are sensitive to initial conditions [143]. With different realisations of randomness, the same algorithm could yield dramatically different performance. This implies that machine learning itself is not a robust technology. Therefore, a combination of learning methods and robust control would be a more promising research direction, which can provide safety guarantees while exploiting available data. In this framework, model-based methods are preferable, since in many cases they perform better than model-free ones in theory and in practice [142]. Also, obtaining a dynamical model enables many readily available and powerful tools from control theory. A recent work on ball catching with a cup has shown the great potential of learning-based MPC [144], comparing with the deep reinforcement learning algorithm [145]. Future work is to develop learning-based MPC algorithms that deal with uncertainty using learning-based methods and provide robustness and safety guarantees.

Appendices

A
Convex Optimisation

Contents

This appendix collects some standard definitions and basic results from convex optimisation, which can be found, for example, in [121, 146–148].

A.1 Convex Sets

Definition A.1 (Convex set)**.** *A set $\mathcal{C} \subseteq \mathbb{R}^n$ is convex if*

$$\mu x_1 + (1-\mu)x_2 \in \mathcal{C}, \quad \forall x_1, x_2 \in \mathcal{C},\ \forall \mu \in [0,1].$$

Definition A.2 (Convex hull)**.** *For a set $\mathcal{C} \subseteq \mathbb{R}^n$, its convex hull is defined as the smallest convex set encompassing $\mathcal{C}$.*

Example A.3 (Polytopic set)**.** *A set $\mathcal{C} \subseteq \mathbb{R}^n$ is a polytope if it is defined by a set of linear inequalities, i.e.*

$$\mathcal{C} := \{x \in \mathbb{R}^n | Ax \leq b\}$$

for some matrix $A \in \mathbb{R}^{m\times n}$ *and vector* $b \in \mathbb{R}^m$.

Example A.4 (Ellipsoidal set)**.** *A set* $\mathcal{C} \subseteq \mathbb{R}^n$ *is an ellipsoid if it is defined by a set of quadratic inequalities, i.e.*

$$\mathcal{C} := \{x \in \mathbb{R}^n | x^\top A x + 2b^\top x + c \leq 0\}$$

for some symmetric positive definite matrix $A \in \mathbb{R}^{n\times n}$, *vector* $b \in \mathbb{R}^n$ *and* $c \in \mathbb{R}$.

Both polytopic and ellipsoidal sets are convex, and they have alternative parameterisations. For example, a polytopic set can be described by the convex hull of its vertices, and an ellipsoidal set can be described by its centre and directions and lengths of its semi-axes, i.e.

$$\mathcal{C} := \{x \in \mathbb{R}^n | (x - x_c)^\top P^{-1} (x - x_c) \leq 1\}$$

for some symmetric positive definite matrix $P \in \mathbb{R}^{n\times n}$ and vector $x_c \in \mathbb{R}^n$. The matrix P defines the orientation and shape of the ellipsoid, and the vector x_c is the centre of the ellipsoid.

A.2 Convex and Midpoint-convex Functions and Lipschitz Continuity

Definition A.5 (Convex function)**.** *Given a convex set* $\mathcal{C} \subseteq \mathbb{R}^n$, *a function* $f : \mathcal{C} \mapsto \mathbb{R}$ *is convex if*

$$f(\mu x_1 + (1-\mu)x_2) \leq \mu f(x_1) + (1-\mu) f(x_2), \quad \forall x_1, x_2 \in \mathcal{C}, \forall \mu \in [0,1]. \quad \text{(A.1)}$$

If this condition is satisfied with a strict inequality, the function f *is strictly convex.*

Concave functions are similarly defined. If f is a (strictly) convex function, then $-f$ is a (strictly) concave function. Also, note that affine functions are both convex and concave. Equation (A.1) is sometimes called Jensen's inequality, which can be developed to more general cases, and systematic studies of convex functions can be traced back to Jensen [149, 150].

Definition A.6 (Midpoint-convex function)**.** *Given a convex set $\mathcal{C} \subseteq \mathbb{R}^n$, a function $f : \mathcal{C} \mapsto \mathbb{R}$ is midpoint-convex if*

$$f(\frac{x_1 + x_2}{2}) \leq \frac{1}{2}(f(x_1) + f(x_2)), \quad \forall x_1, x_2 \in \mathcal{C}.$$

Lemma A.7 ([121, Section 72 Theorem A])**.** *A midpoint-convex (midpoint-concave) function $f : [x_1, x_2] \mapsto \mathbb{R}$ is continuous and hence convex (concave) if it is bounded from above (below) on (x_1, x_2).*

Definition A.8 (Lipschitz continuity)**.** *A function $f : \mathbb{R} \mapsto \mathbb{R}$ is Lipschitz continuous if*

$$\|f(x_1) - f(x_2)\| \leq L\|x_1 - x_2\|, \quad \forall x_1, x_2 \in \mathbb{R}, \exists L > 0,$$

where L is a Lipschitz constant.

A.3 Convex Optimisation Problems

Quadratically constrained quadratic programs (QCQPs) are optimisation problems that can be expressed in the form

$$\begin{aligned} \min_x \ & x^\top A_0 x + b_0^\top x + c_0 \\ \text{s.t.} \ & x^\top A_i x + b_i^\top x + c_i \leq 0, \ i = 1, \dots, m, \\ & Fx = g, \end{aligned} \tag{A.2}$$

where $A_i \in \mathbb{R}^{n \times n}$ are symmetric positive semidefinite matrices for $i = 0, \dots, m$, $b_i \in \mathbb{R}^n$ for $i = 0, \dots, m$, $c_i \in \mathbb{R}$ for $i = 0, \dots, m$, $F \in \mathbb{R}^{p \times n}$ and $g \in \mathbb{R}^p$. If $A_i = 0$ for $i = 0, \dots, m$, then this problem reduces to a linear program. If $A_i = 0$ for $i = 1, \dots, m$, this problem is a quadratic program.

Second-order cone programs (SOCPs) are optimisation problems that can be expressed in the form

$$\begin{aligned} \min_x \ & f^\top x \\ \text{s.t.} \ & \|A_i x + b_i\|_2 \leq c_i^\top x + d_i, \quad i = 1, \dots, m, \\ & Fx = g, \end{aligned}$$

where $f \in \mathbb{R}^n$, $F \in \mathbb{R}^{p\times n}$, $g \in \mathbb{R}^p$, and, for $i = 1, \ldots, m$, $A_i \in \mathbb{R}^{n_i \times n}$, $b_i \in \mathbb{R}^{n_i}$, $c_i \in \mathbb{R}^n$ and $d_i \in \mathbb{R}$. SOCPs are more general than QCQPs and can be efficiently solved, for example, by commercial solvers Gurobi [151] and MOSEK [152]. QCQPs can also be expressed in the form of SOCPs. This involves introducing a quadratic constraint to bound the quadratic objective function by a linear function which then acts as the new objective function. Then quadratic constraints can be equivalently written in the form of second-order conic constraints and (A.2) is equivalent to

$$\left\| \begin{matrix} (1 + b_i^\top x + c_i)/2 \\ A_i^{\frac{1}{2}} x \end{matrix} \right\|_2 \leq (1 - b_i^\top x - c_i)/2.$$

A.4 Multiobjective Optimisation

A multiobjective optimisation problem (MOP) can be expressed in the form

$$\begin{aligned} \min_x \ & \mathbf{f}(x) := [f_1(x), \ldots, f_p(x)]^\top \\ \text{s.t.} \ & g_i(x) \leq 0, \quad i = 1, \ldots, n, \\ & h_i(x) = 0, \quad i = 1, \ldots, l, \end{aligned} \tag{A.3}$$

where $\mathbf{f}(x)$ is a vector of objective functions.

Definition A.9 (Pareto Optimality)**.** *A feasible solution, x^*, of problem* (A.3) *is Pareto optimal if and only if there is no other feasible point, y, satisfying $\mathbf{f}(y) \leq \mathbf{f}(x^*)$ (in an elementwise fashion) with strict inequalities for at least one objective function.*

A common approach for solving MOPs is the weighted sum method, resulting in a single objection problem. It is given by

$$\begin{aligned} \min_x \ & \sum_{i=1}^p w_i f_i(x) \\ \text{s.t.} \ & g_i(x) \leq 0, \quad i = 1, \ldots, n, \\ & h_i(x) = 0, \quad i = 1, \ldots, l, \end{aligned}$$

where $\{w_i\}_{i=1}^p$ is a set of weights chosen by decision makers such that $\sum_{i=1}^p w_i = 1$ and $w_i > 0 \ \forall i$.

A.5 Linear Matrix Inequality and Schur Complement

Definition A.10 (Linear matrix inequality). *A linear matrix inequality (LMI) can be expressed in the form*

$$A(x) := A_0 + \sum_{i=1}^{n} x_i A_i \succeq 0,$$

where $x \in \mathbb{R}^n$ is the variable and $A_i \in \mathbb{R}^{m \times m},\ i = 0, \dots, n$ are symmetric matrices.

Nonlinear inequalities can be reformulated as LMI conditions using Schur complements as shown in Lemmas (A.11) and (A.12).

Lemma A.11 ([148, Section 2.1]). *The LMI*

$$\begin{bmatrix} A(x) & B(x) \\ B^\top(x) & C(x) \end{bmatrix} \succ 0,$$

where $A(x) = A^\top(x)$, $C(x) = C^\top(x)$ and $A(x)$, $B(x)$, and $C(x)$ are affine functions of x, is equivalent to the following conditions

$$C(x) \succ 0, \quad A(x) - B(x)C(x)^{-1}B(x)^\top \succ 0.$$

Lemma A.12 ([148, Section 2.7]). *For $R \succ 0$, let*

$$P - A^\top PA + Q - A^\top PB(R + B^\top PB)^{-1}B^\top PA \succeq 0,$$

where $P = P^\top$ is an optimisation variable and A, B, $Q = Q^\top$, $R = R^\top$ are given matrices with conformal dimensions, then this condition is equivalent to the LMI

$$\begin{bmatrix} P - A^\top PA + Q & A^\top PB \\ B^\top PA & R + B^\top PB \end{bmatrix} \succeq 0.$$

B

Probability Theory and Random Variables

Contents

This appendix collects some standard definitions and results in probability theory, which can be found, for example, in [112, 153, 154].

B.1 Probability Measure and Probability Space

Definition B.1 (σ-field)**.** *A σ-field is defined as a set, $\mathcal{F}$, of subsets of sample space Ω, enjoying the following properties:*

(i) empty set $\emptyset \in \mathcal{F}$;

(ii) if sets $A_1, A_2, \ldots \in \mathcal{F}$, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$;

(iii) if set $A \in \mathcal{F}$, its complement $A^c \in \mathcal{F}$.

Definition B.2 (Probability measure). *Given a pair* $(\Omega, \mathcal{F})$*, a function* $\mathbb{P} : \mathcal{F} \mapsto [0,1]$ *is called a probability measure defined on the pair* $(\Omega, \mathcal{F})$*, if it satisfies the following conditions:*

(i) $\mathbb{P}\{\emptyset\} = 0$ *and* $\mathbb{P}\{\Omega\} = 1$*, where* $\emptyset$ *denotes the empty set and* Ω *is the sample space;*

(ii) Given a set, $\{A_1, A_2, \ldots\}$*, of members of* $\mathcal{F}$ *that are mutually disjoint for any pairs,* $\mathbb{P}\{\bigcup_{i=1}^{\infty} A_i\} = \sum_{i=1}^{\infty} \mathbb{P}\{A_i\}$.

Definition B.3 (Probability space). *The probability space is defined as the triplet* $(\Omega, \mathcal{F}, \mathbb{P})$ *that consists of a sample space* Ω*, a* σ*-field* $\mathcal{F}$ *and a probability measure* $\mathbb{P}$.

Definition B.4 (Conditional probability). *The probability that event* E_1 *happens given that event* E_2 *has already occurred is called a conditional probability given by*

$$\mathbb{P}\{E_1|E_2\} = \frac{\mathbb{P}\{E_1 \cap E_2\}}{\mathbb{P}\{E_2\}}.$$

Note that this conditional probability is well defined only if $\mathbb{P}\{E_2\} > 0$.

Lemma B.5 ([153, Section 1.4]). *Given a collection of events* A_1, A_2, $\ldots$, A_n *and that* $\mathbb{P}\{\bigcap_{i=1}^{n-1} A_i\} > 0$,

$$\mathbb{P}\{\bigcap_{i=1}^{n} A_i\} = \mathbb{P}\{A_1\}\mathbb{P}\{A_2|A_1\}\mathbb{P}\{A_3|A_1 \cap A_2\} \cdots \mathbb{P}\{A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}\}.$$

B.2 Random Variables and Probability Distributions

Definition B.6 (Random variable). *A random variable* X *on sample space* Ω *is a function defined on* Ω *that assumes real values, i.e.* $X : \Omega \mapsto R$.

Definition B.7 (Expectation). *For a continuous random variable* X*, its expectation is defined as*

$$\mathbb{E}\{X\} = \int x f(x) dx$$

where $f(x)$ *is the probability density function of* X*; and for a discrete random variable, it is defined as*

$$\mathbb{E}\{X\} = \sum x\mathbb{P}\{X = x\}.$$

Also, given random variables X, Y on the same probability space, the conditional expectation of X given Y is written as $\mathbb{E}\{X|Y\}$.

Lemma B.8 ([153, Section 3.7 Theorem 4]). *Given random variables X, Y on the same probability space,*

$$\mathbb{E}\{\mathbb{E}\{X|Y\}\} = \mathbb{E}\{X\}.$$

Definition B.9 (Cumulative distribution function). *For a continuous random variable X, its cumulative distribution function is given by*

$$F(x) = \mathbb{P}\{X \leq x\} = \int_{-\infty}^{x} f(t)dt,$$

where $f(t)$ is the probability density function of X.

Definition B.10 (Gaussian distribution). *A Gaussian distribution, which is also called a normal distribution, is a continuous probability distribution characterised by its mean and variance. Its probability density function is given by*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{1}{2}\Big(\frac{x-\mu}{\sigma}\Big)^2\Big),$$

where μ is the mean and σ^2 is the variance.

For a random variable X that is normally distributed with mean μ and variance σ^2, it can be written as $X \sim \mathcal{N}(\mu, \sigma^2)$.

Definition B.11 (Bernoulli distribution). *A Bernoulli distribution is a probability distribution for a discrete random variable that only assumes values 0 (with probability $1-\lambda$) and 1 (with probability λ).*

A sequence of independently and identically distributed (i.i.d.) Bernoulli random variables is called a Bernoulli process.

Definition B.12 (Binomial distribution). *A Binomial distribution is a discrete probability distribution characterised by a number, n, of independent Bernoulli trials and a success rate λ. Its probability mass function $\mathbb{P}\{X = k\}$ denotes the probability of achieving k successes in n experiments and is given by*

$$\mathbb{P}\{X = k\} = \binom{n}{k} \lambda^k (1-\lambda)^{n-k}.$$

Lemma B.13 (Markov's inequality, [153, Section 7.2 Lemma 7]). *For any random variable X with finite mean, the following holds:*

$$\mathbb{P}\{|X| \geq a\} \leq \frac{\mathbb{E}\{|X|\}}{a} \quad \forall\ a > 0.$$

Lemma B.14 (Chebyshev's inequality, [153, Section 7.3 Example 3]). *For any random variable X with finite mean and variance, the following holds:*

$$\mathbb{P}\{|X| \geq a\} \leq \frac{\mathbb{E}\{X^2\}}{a^2} \quad \forall\ a > 0.$$

B.3 Convergence of Sequences of Random Variables

For a random variable X and a sequence of random variables $\{X_1, X_2, \ldots\}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, several notions of convergence of random variables are introduced as follows.

Definition B.15 (Convergence in distribution). *If*

$$\lim_{n\to\infty} F_n(x) = F(x)$$

for all points $x \in \mathbb{R}$ at which $F(x)$ is continuous, where $F(x)$ and $F_n(x)$ are the cumulative distribution functions of random variables X and X_n, then the sequence $\{X_1, X_2, \ldots\}$ converges in distribution to the random variable X.

Definition B.16 (Convergence in probability). *If*

$$\lim_{n\to\infty} \mathbb{P}\{|X_n - X| > \varepsilon\} = 0 \quad \forall\ \varepsilon > 0,$$

then the sequence $\{X_1, X_2, \ldots\}$ converges in probability to the random variable X.

Definition B.17 (Convergence in mean). *Given a real number $r \geq 1$, if $\mathbb{E}\{|X_n|^r\}$ for $n = 1, 2, \ldots$ and $\mathbb{E}\{|X|^r\}$ exist, and*

$$\lim_{n\to\infty} \mathbb{E}\{|X_n - X|^r\} = 0,$$

then the sequence $\{X_1, X_2, \ldots\}$ converges in the r-th mean to the random variable X.

Definition B.18 (Almost sure convergence). *If*

$$\mathbb{P}\{\lim_{n\to\infty} X_n = X\} = 1,$$

or equivalently, by expressing random variables as real-valued functions,

$$\mathbb{P}\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = 1,$$

then the sequence $\{X_1, X_2, \ldots\}$ *converges almost surely/with probability* 1 *to the random variable* X.

Lemma B.19 (Borel–Cantelli lemmas, [153, Section 7.3 Theorem 10]). *Given a sequence of events* $\{E_1, E_2, \ldots\}$,

(1) $\mathbb{P}\{\limsup_{n\to\infty} E_n\} = \mathbb{P}\{\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m\} = 0$, *if* $\sum_{n=1}^{\infty} \mathbb{P}\{E_n\} < \infty$;

(2) $\mathbb{P}\{\limsup_{n\to\infty} E_n\} = 1$, *if* $\sum_{n=1}^{\infty} \mathbb{P}\{E_n\} = \infty$ *and* E_1, E_2, $\ldots$ *are independent events.*

Lemma B.20 (Strong law of large numbers, [153, Section 7.4 Theorem 3]). *Given a sequence,* $\{X_n\}_{n=1}^{\infty}$, *of independent and identically distributed random variables with* $\mathbb{E}\{X_n\} = \bar{X}$ *and a finite second moment,*

$$\mathbb{P}\Big\{\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} X_n = \bar{X}\Big\} = 1.$$

References

[1] I.M. Gelfand and S.V. Fomin. *Calculus of Variations*. Ed. by R. A. Silverman. Dover Books on Mathematics. Dover Publications, 2000.

[2] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko. *Mathematical Theory of Optimal Processes {in Russian}*. 1961.

[3] R.E. Bellman. *Dynamic Programming*. Princeton, NJ: Princeton University Press, 1957.

[4] R. Vinter. *Optimal Control*. Modern Birkhäuser Classics. Birkhäuser Basel, 2010.

[5] Karl J. Åström. *Introduction to Stochastic Control Theory*. Dover Books on Electrical Engineering. Dover Publications, 2006.

[6] C.E. García, D.M. Prett, and M. Morari. “Model predictive control: Theory and practice—A survey”. In: *Automatica* 25.3 (1989), pp. 335–348.

[7] D.Q. Mayne, J.B. Rawlings, C.V. Rao, and P.O.M. Scokaert. “Constrained model predictive control: Stability and optimality”. In: *Automatica* 36.6 (2000), pp. 789–814.

[8] D.Q. Mayne, M.M. Seron, and S.V. Raković. “Robust model predictive control of constrained linear systems with bounded disturbances”. In: *Automatica* 41.2 (2005), pp. 219–224.

[9] D.Q. Mayne, E.C. Kerrigan, and P. Falugi. “Robust model predictive control: advantages and disadvantages of tube-based methods”. In: *IFAC Proceedings Volumes* 44.1 (2011). 18th IFAC World Congress, pp. 191–196.

[10] S.V. Raković, B. Kouvaritakis, R. Findeisen, and M. Cannon. “Homothetic tube model predictive control”. In: *Automatica* 48.8 (2012), pp. 1631–1638.

[11] X. Lu, M. Cannon, and D. Koksal-Rivet. “Robust adaptive model predictive control: performance and parameter estimation”. In: *International Journal of Robust and Nonlinear Control (Early Access)* (2020), pp. 1–22. URL: https://ora.ox.ac.uk/objects/pubs:1071417.

[12] Y. Ma, S. Vichik, and F. Borrelli. “Fast stochastic MPC with optimal risk allocation applied to building control systems”. In: *2012 IEEE 51st IEEE Conference on Decision and Control*. IEEE. 2012, pp. 7559–7564.

[13] A. Bemporad, L. Bellucci, and T. Gabbriellini. “Dynamic option hedging via stochastic model predictive control based on scenario simulation”. In: *Quantitative Finance* 14.10 (2014), pp. 1739–1751.

[14] A. Mesbah. “Stochastic Model Predictive Control: An Overview and Perspectives for Future Research”. In: *IEEE Control Systems Magazine* 36.6 (Dec. 2016), pp. 30–44.

[15] M. Farina, L. Giulioni, L. Magni, and R. Scattolini. “A probabilistic approach to Model Predictive Control”. In: *IEEE Conference on Decision and Control.* 2013, pp. 7734–7739.

[16] L. Hewing and M.N. Zeilinger. “Stochastic model predictive control for linear systems using probabilistic reachable sets”. In: *2018 IEEE Conference on Decision and Control.* IEEE. 2018, pp. 5182–5188.

[17] L. Hewing, K.P. Wabersich, and M.N. Zeilinger. “Recursively feasible stochastic model predictive control using indirect feedback”. In: *Automatica* 119 (2020), p. 109095.

[18] D. Blackwell. “Discounted dynamic programming”. In: *The Annals of Mathematical Statistics* 36.1 (1965), pp. 226–235.

[19] D.P. Bertsekas. *Dynamic Programming and Optimal Control.* 3rd. Belmont, Massachusetts: Athena Scientific, 2005.

[20] B.P.G. Van Parys, P.J. Goulart, and M. Morari. “Infinite Horizon Performance Bounds for Uncertain Constrained Systems”. In: *IEEE Transactions on Automatic Control* 58.11 (Nov. 2013), pp. 2803–2817.

[21] M. Kamgarpour and T. Summers. “On infinite dimensional linear programming approach to stochastic control”. In: *Proceedings of the IFAC World Congress.* 2017.

[22] B. Kouvaritakis, M. Cannon, and P. Couchman. “MPC as a tool for sustainable development integrated policy assessment”. In: *IEEE Transactions on Automatic Control* 51.1 (2006), pp. 145–149.

[23] D.P. Bertsekas. *Reinforcement learning and optimal control.* Belmont, Massachusetts: Athena Scientific, 2019.

[24] S. Barratt and S.P. Boyd. “Multi-Period Liability Clearing via Convex Optimal Control”. In: *Available at SSRN 3604618* (2020).

[25] P. Nystrup, S.P. Boyd, E. Lindström, and H. Madsen. “Multi-period portfolio selection with drawdown control”. In: *Annals of Operations Research* 282.1-2 (2019), pp. 245–271.

[26] A. Frankel. “Discounted quotas”. In: *Journal of Economic Theory* 166 (2016), pp. 396–444.

[27] S.P. Boyd, M.T. Mueller, B. O’Donoghue, and Y. Wang. *Performance bounds and suboptimal policies for multi-period investment.* Now Publishers, 2014.

[28] C.W. Clark. “Profit maximization and the extinction of animal species”. In: *Journal of Political Economy* 81.4 (1973), pp. 950–961.

[29] D. Ludwig, W.A. Brock, and S.R. Carpenter. “Uncertainty in discount models and environmental accounting”. In: *Ecology and Society* 10.2 (2005).

[30] R. Postoyan, L. Busoniu, D. Nesic, and J. Daafouz. “Stability analysis of discrete-time infinite-horizon optimal control with discounted cost”. In: *IEEE Transactions on Automatic Control* 62.6 (2017), pp. 2736–2749.

[31] S. Yan, M. Cannon, and P.J. Goulart. “Stochastic output feedback MPC with intermittent observations”. In: (2021). Under review. URL: https://arxiv.org/abs/arxiv:2009.11071.

[32] G. Schildbach, P.J. Goulart, and M. Morari. "Linear Controller Design for Chance Constrained Systems". In: *Automatica* 51 (2015), pp. 278–284.

[33] B. Kouvaritakis, M. Cannon, S.V. Raković, and Q. Cheng. "Explicit use of probabilistic distributions in linear predictive control". In: *Automatica* 46.10 (2010), pp. 1719–1724.

[34] M. Cannon, B. Kouvaritakis, S.V. Raković, and Q. Cheng. "Stochastic Tubes in Model Predictive Control With Probabilistic Constraints". In: *IEEE Transactions on Automatic Control* 56.1 (2011), pp. 194–200.

[35] T. Hashimoto. "Probabilistic constrained model predictive control for linear discrete-time systems with additive stochastic disturbances". In: *IEEE Conference on Decision and Control*. 2013, pp. 6434–6439.

[36] G.C. Calafiore and L. Fagiano. "Robust model predictive control via scenario optimization". In: *IEEE Transactions on Automatic Control* 58.1 (2012), pp. 219–224.

[37] K. Margellos, P.J. Goulart, and J. Lygeros. "On the road between robust optimization and the scenario approach for chance constrained optimization problems". In: *IEEE Transactions on Automatic Control* 59.8 (2014), pp. 2258–2263.

[38] M. Korda, R. Gondhalekar, F. Oldewurtel, and C. N. Jones. "Stochastic model predictive control: Controlling the average number of constraint violations". In: *2012 IEEE 51st IEEE Conference on Decision and Control*. 2012, pp. 4529–4536.

[39] M. Korda, R. Gondhalekar, F. Oldewurtel, and C.N. Jones. "Stochastic MPC framework for controlling the average constraint violation". In: *IEEE Transactions on Automatic Control* 59 (2014), pp. 1706–1721.

[40] J. Fleming and M. Cannon. "Time-average constraints in stochastic Model Predictive Control". In: *American Control Conference*. 2017, pp. 5648–5653.

[41] G. Schildbach, L. Fagiano, C. Frei, and M. Morari. "The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations". In: *Automatica* 50.12 (2014), pp. 3009–3018.

[42] D. Muñoz-Carpintero, G. Hu, and C.J. Spanos. "Stochastic Model Predictive Control with adaptive constraint tightening for non-conservative chance constraints satisfaction". In: *Automatica* 96 (2018), pp. 32–39.

[43] I. Palacios-Huerta and A. Pérez-Kakabadse. *Consumption and portfolio rules with stochastic hyperbolic discounting*. IKERLANAK 72. Universidad del País Vasco - Departamento de Fundamentos del Análisis Económico I, 2013.

[44] S. Yan, P.J. Goulart, and M. Cannon. "Stochastic Model Predictive Control with Discounted Probabilistic Constraints". In: *2018 European Control Conference*. 2018, pp. 1003–1008.

[45] B. Kouvaritakis, M. Cannon, and G. Huang. "MPC as a Tool for Sustainable Development Integrated Policy Assessment". In: *IFAC Proceedings Volumes* 36.18 (2003), pp. 513–518.

[46] P. Dasgupta. "Discounting climate change". In: *Journal of risk and uncertainty* 37.2 (2008), pp. 141–169.

[47] S. Yan, P.J. Goulart, and M. Cannon. "Stochastic MPC with dynamic feedback gain selection and discounted probabilistic constraints". In: (2021). Under review. URL: https://arxiv.org/abs/arxiv:2007.07134.

[48] S. Yan, M. Cannon, and P.J. Goulart. "Output feedback stochastic MPC with packet losses". In: vol. 53. 2. 21th IFAC World Congress. 2020, pp. 7105–7110.

[49] A.I. Propoi. "Application of linear programming methods for the synthesis of automatic sampled-data systems". In: *Avtomat. i Telemekh* 24.7 (1963), pp. 912–920.

[50] J. Rault, A. Richalet, J.L. Testud, and J. Papon. "Model predictive heuristic control: application to industrial processes". In: *Automatica* 14.5 (1978), pp. 413–428.

[51] C.R. Cutler and B.L. Ramaker. "Dynamic matrix control- A computer control algorithm". In: *joint automatic control conference*. 17. 1980, p. 72.

[52] D.W. Clarke, C. Mohtadi, and P.S. Tuffs. "Generalized predictive control—Part I. The basic algorithm". In: *Automatica* 23.2 (1987), pp. 137–148.

[53] D.W. Clarke, C. Mohtadi, and P.S. Tuffs. "Generalized predictive control—part II extensions and interpretations". In: *Automatica* 23.2 (1987), pp. 149–160.

[54] M. Babu, Y. Oza, A.K. Singh, K.M. Krishna, and S. Medasani. "Model Predictive Control for Autonomous Driving Based on Time Scaled Collision Cone". In: *2018 European Control Conference*. 2018, pp. 641–648.

[55] A. Gray, Y. Gao, T. Lin, J.K. Hedrick, and F. Borrelli. "Stochastic predictive control for semi-autonomous vehicles with an uncertain driver model". In: *16th International IEEE Conference on Intelligent Transportation Systems*. IEEE. 2013, pp. 2329–2334.

[56] S. Vazquez, J.I. Leon, L.G. Franquelo, J. Rodriguez, H.A. Young, A. Marquez, and P. Zanchetta. "Model predictive control: A review of its applications in power electronics". In: *IEEE industrial electronics magazine* 8.1 (2014), pp. 16–31.

[57] B. Stellato, T. Geyer, and P.J. Goulart. "High-Speed Finite Control Set Model Predictive Control for Power Electronics". In: *IEEE Transactions on Power Electronics* 32.5 (2017), pp. 4007–4020.

[58] M.A. Evans, M. Cannon, and B. Kouvaritakis. "Robust MPC Tower Damping for Variable Speed Wind Turbines". In: *IEEE Transactions on Control Systems Technology* 23.1 (2015), pp. 290–296.

[59] M. Mirzaei, N.K. Poulsen, and H.H. Niemann. "Robust model predictive control of a wind turbine". In: *2012 American Control Conference*. IEEE. 2012, pp. 4393–4398.

[60] D.E. Quevedo and D. Nesić. "Input-to-state stability of packetized predictive control over unreliable networks affected by packet-dropouts". In: *IEEE Transactions on Automatic Control* 56.2 (2010), pp. 370–375.

[61] P.K. Mishra, D. Chatterjee, and D.E. Quevedo. "Stochastic predictive control under intermittent observations and unreliable actions". In: *Automatica* 118 (2020), p. 109012.

[62] F. Oldewurtel, C.N. Jones, A. Parisio, and M. Morari. “Stochastic model predictive control for building climate control”. In: *IEEE Transactions on Control Systems Technology* 22.3 (2013), pp. 1198–1205.

[63] J. Drgona et al. “All you need to know about model predictive control for buildings”. In: *Annual Reviews in Control* (2020).

[64] F. Herzog, G. Dondi, and H.P. Geering. “Stochastic model predictive control and portfolio optimization”. In: *International Journal of Theoretical and Applied Finance* 10.2 (2007), pp. 203–233.

[65] L. Dai, M. Cannon, F. Yang, and S. Yan. “Fast Self-Triggered MPC for Constrained Linear Systems with Additive Disturbances”. In: *IEEE Transactions on Automatic Control* (2020). Early Access.

[66] F.D. Brunner, M. Heemels, and F. Allgöwer. “Robust self-triggered MPC for constrained linear systems: A tube-based approach”. In: *Automatica* 72 (2016), pp. 73–83.

[67] R.E. Bellman. “Dynamic Programming”. In: *Science* 153.3731 (1966), pp. 34–37.

[68] R.E. Kalman. “Contributions to the theory of optimal control”. In: *Boletin Sociedad Matematica Mexicana* 5 (1960), pp. 102–119.

[69] H. Van de Water and J. Willems. “The certainty equivalence property in stochastic control theory”. In: *IEEE Transactions on Automatic Control* 26.5 (1981), pp. 1080–1087.

[70] R.E. Bellman. *Adaptive control processes: a guided tour*. Princeton, NJ: Princeton University Press, 1961.

[71] A. Bemporad and M. Morari. “Robust model predictive control: A survey”. In: *Robustness in identification and control*. Springer, 1999, pp. 207–226.

[72] B. Kouvaritakis and M. Cannon. *Model Predictive Control: Classical, Robust and Stochastic*. Springer, 2015.

[73] M. Diehl J.B. Rawlings D.Q. Mayne. *Model Predictive Control: Theory, Computation, and Design*. 2nd edition. Nob Hill Publishing, 2017.

[74] H. Michalska and D.Q. Mayne. “Robust receding horizon control of constrained nonlinear systems”. In: *IEEE transactions on automatic control* 38.11 (1993), pp. 1623–1633.

[75] E.G. Gilbert and K.T. Tan. “Linear systems with state and control constraints: The theory and application of maximal output admissible sets”. In: *IEEE Transactions on Automatic control* 36.9 (1991), pp. 1008–1020.

[76] D. Chmielewski and V. Manousiouthakis. “On constrained infinite-time linear quadratic optimal control”. In: *Systems & Control Letters* 29.3 (1996), pp. 121–129.

[77] S. Yu, M. Reble, H. Chen, and F. Allgöwer. “Inherent robustness properties of quasi-infinite horizon nonlinear model predictive control”. In: *Automatica* 50.9 (2014), pp. 2269–2280.

[78] S.V. Rakovic, E.C. Kerrigan, K.I. Kouramas, and D.Q. Mayne. “Invariant approximations of the minimal robust positively invariant set”. In: *IEEE Transactions on automatic control* 50.3 (2005), pp. 406–410.

[79] I. Kolmanovsky and E.G. Gilbert. “Theory and computation of disturbance invariant sets for discrete-time linear systems”. In: *Mathematical problems in engineering* 4.4 (1998), pp. 317–367.

[80] W. Langson, I. Chryssochoos, S.V. Raković, and D.Q. Mayne. “Robust model predictive control using tubes”. In: *Automatica* 40.1 (2004), pp. 125–133.

[81] J. Löfberg. “Minimax Approaches to Robust Model Predictive Control”. PhD thesis. Linköping University, 2003.

[82] D.H. Van Hessem and O.H. Bosgra. “A conic reformulation of model predictive control including bounded and stochastic disturbances under state and input constraints”. In: *Proceedings of the 41st IEEE Conference on Decision and Control.* Vol. 4. IEEE. 2002, pp. 4643–4648.

[83] P.J. Goulart, E.C. Kerrigan, and J.M. Maciejowski. “Optimization Over State Feedback Policies for Robust Control with Constraints”. In: *Automatica* 42.4 (2006), pp. 523–533.

[84] R.W. Liu. “Convergent systems”. In: *IEEE Transactions on Automatic Control* 13.4 (1968), pp. 384–391.

[85] M.V. Kothare, V. Balakrishnan, and M. Morari. “Robust constrained model predictive control using linear matrix inequalities”. In: *Automatica* 32.10 (1996), pp. 1361–1379.

[86] J. Schuurmans and J.A. Rossiter. “Robust predictive control using tight sets of predicted states”. In: *IEE proceedings-Control theory and applications* 147.1 (2000), pp. 13–18.

[87] M. Cannon and B. Kouvaritakis. “Optimizing prediction dynamics for robust MPC”. In: *IEEE Transactions on Automatic Control* 50.11 (2005), pp. 1892–1897.

[88] Q. Cheng, M. Cannon, and B. Kouvaritakis. “The design of dynamics in the prediction structure of robust MPC”. In: *International Journal of Control* 86.11 (2013), pp. 2096–2103.

[89] A. Nemirovski and A. Shapiro. “Convex approximations of chance constrained programs”. In: *SIAM Journal on Optimization* 17.4 (2007), pp. 969–996.

[90] D. Chatterjee and J. Lygeros. “On Stability and Performance of Stochastic Predictive Control Techniques”. In: *IEEE Transactions on Automatic Control* 60.2 (2015), pp. 509–514.

[91] S. Zymler, D. Kuhn, and B. Rustem. “Distributionally robust joint chance constraints with second-order moment information”. In: *Mathematical Programming* 137.1 (2013), pp. 167–198.

[92] S. Zhu and M. Fukushima. “Worst-case conditional value-at-risk with application to robust portfolio management”. In: *Operations research* 57.5 (2009), pp. 1155–1168.

[93] B.P.G Van Parys, D. Kuhn, P.J. Goulart, and M. Morari. “Distributionally robust control of constrained stochastic systems”. In: *IEEE Transactions on Automatic Control* 61.2 (2015), pp. 430–442.

[94] G.C. Calafiore and M.C. Campi. “The scenario approach to robust control design”. In: *IEEE Transactions on automatic control* 51.5 (2006), pp. 742–753.

[95] M. Lorenzen, F. Allgöwer, F. Dabbene, and R. Tempo. “Scenario-based stochastic MPC with guaranteed recursive feasibility”. In: *2015 54th IEEE conference on decision and control*. IEEE. 2015, pp. 4958–4963.

[96] J. Fleming and M. Cannon. “Stochastic MPC for Additive and Multiplicative Uncertainty Using Sample Approximations”. In: *IEEE Transactions on Automatic Control* 64.9 (2019), pp. 3883–3888.

[97] M.C. Campi and S. Garatti. “A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality”. In: *Journal of optimization theory and applications* 148.2 (2011), pp. 257–280.

[98] G. Schildbach, L. Fagiano, and M. Morari. “Randomized Solutions to Convex Programs with Multiple Chance Constraints”. In: *SIAM Journal on Optimization* 23.4 (2013), pp. 2479–2501.

[99] X. Zhang, S. Grammatico, G. Schildbach, P.J. Goulart, and J. Lygeros. “On the sample size of random convex programs with structured dependence on the uncertainty”. In: *Automatica* 60 (2015), pp. 182–188.

[100] A.J. Kleywegt, A. Shapiro, and T. Homem-de-Mello. “The sample average approximation method for stochastic discrete optimization”. In: *SIAM Journal on Optimization* 12.2 (2002), pp. 479–502.

[101] J. Luedtke and S. Ahmed. “A sample approximation approach for optimization with probabilistic constraints”. In: *SIAM Journal on Optimization* 19.2 (2008), pp. 674–699.

[102] R.R. Bitmead, M.R. Gevers, I.R. Petersen, and R.J. Kaye. “Monotonicity and stabilizability- properties of solutions of the Riccati difference equation: Propositions, lemmas, theorems, fallacious conjectures and counterexamples”. In: *Systems and Control Letters* 5 (1985), pp. 309–315.

[103] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan, and S.S. Sastry. “Kalman Filtering With Intermittent Observations”. In: *IEEE Transactions on Automatic Control* 49.9 (2004), pp. 1453–1464.

[104] H.J. Kushner. *Introduction to stochastic control*. Holt, Rinehart and Winston, 1971.

[105] F. Blanchini. “Constrained control for uncertain linear systems”. In: *J. Optim. Theory Appl.* 71 (1991), pp. 465–484.

[106] A.T. Schwarm and M. Nikolaou. “Chance-constrained model predictive control”. In: *AIChE Journal* 45.8 (1999), pp. 1743–1752.

[107] M. Lorenzen, F. Dabbene, R. Tempo, and F. Allgöwer. “Constraint-Tightening and Stability in Stochastic Model Predictive Control”. In: *IEEE Transactions on Automatic Control* 62 (2017), pp. 3165–3177.

[108] M. Lorenzen, M.A. Müller, and F. Allgöwer. “Stochastic model predictive control without terminal constraints”. In: *International Journal of Robust and Nonlinear Control* 29.15 (2019), pp. 4987–5001.

[109] D. Limón, T. Alamo, F. Salas, and E.F. Camacho. “On the stability of constrained MPC without terminal constraint”. In: *IEEE Transactions on Automatic Control* 51.5 (2006), pp. 832–836.

[110] Camile and J.M. Maciejowski. "Robust finite horizon MPC without terminal constraints". In: *Proceedings of the 39th IEEE Conference on Decision and Control*. Vol. 1. 2000, pp. 166–171.

[111] P.O.M. Scokaert. "Infinite horizon generalized predictive control". In: *Int. J. Control* 66.1 (1997), pp. 161–175.

[112] W. Feller. *An Introduction to Probability Theory and Its Applications*. 2nd. Vol. 2. John Wiley & Sons, 1971.

[113] M. Lorenzen, F. Allgöwer, F. Dabbene, and R. Tempo. "An improved constraint-tightening approach for Stochastic MPC". In: *American Control Conference*. 2015, pp. 944–949.

[114] M. Cannon, B. Kouvaritakis, and X. Wu. "Probabilistic Constrained MPC for Multiplicative and Additive Stochastic Uncertainty". In: *IEEE Transactions on Automatic Control* 54 (2009), pp. 1626–1632.

[115] A. Chinchuluun and P.M. Pardalos. "A survey of recent developments in multiobjective optimization". In: *Annals of Operations Research* 154.1 (2007), pp. 29–50.

[116] R.T. Marler and J.S. Arora. "The weighted sum method for multi-objective optimization: new insights". In: *Structural and multidisciplinary optimization* 41.6 (2010), pp. 853–862.

[117] L. Zadeh. "Optimality and non-scalar-valued performance criteria". In: *IEEE transactions on Automatic Control* 8.1 (1963), pp. 59–60.

[118] K. Miettinen. *Nonlinear multiobjective optimization*. Vol. 12. Springer Science & Business Media, 2012.

[119] Z. Yang, X. Cai, and Z. Fan. "Epsilon Constrained Method for Constrained Multiobjective Optimization Problems: Some Preliminary Results". In: *Proceedings of the Companion Publication of the 2014 Annual Conference on Genetic and Evolutionary Computation*. 2014, pp. 1181–1186.

[120] M.T.M. Emmerich and A.H. Deutz. "A tutorial on multiobjective optimization: fundamentals and evolutionary methods". In: *Natural computing* 17.3 (2018), pp. 585–609.

[121] A.W. Roberts and D.E. Varberg. *Convex functions*. Pure and applied mathematics ; 57. New York: Academic Press, 1973.

[122] P. Hokayem, D. Chatterjee, and J. Lygeros. "On stochastic receding horizon control with bounded control inputs". In: *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*. IEEE. 2009, pp. 6359–6364.

[123] S. Bittanti, A.J. Laub, and J.C. Willems. *The Riccati Equation*. Springer Science & Business Media, 2012.

[124] T. Nishimura. "On the a priori Information in Sequential Estimation Problems". In: *IEEE Transactions on Automatic Control* 11 (1966), pp. 197–204.

[125] B. Kouvaritakis, M. Cannon, and J.A. Rossiter. "Who needs QP for linear MPC anyway?" In: *Automatica* 38 (2002), pp. 879–884.

[126] S.T. Lian, K. Marzuki, and Y. Rubiyah. "Tuning of a neuro-fuzzy controller by genetic algorithms with an application to a coupled-tank liquid-level control system". In: *Engineering Applications of Artificial Intelligence* 11.4 (1998), pp. 517–529.

[127] Y. Wang and L. Guo. "On stability of random Riccati equations". In: *Science in China Series E: Technological Sciences* 42.2 (1999), pp. 136–148.

[128] S. Kar, B. Sinopoli, and J.M. Moura. "Kalman filtering with intermittent observations: Weak convergence to a stationary distribution". In: *IEEE Transactions on Automatic Control* 57.2 (2012), pp. 405–420.

[129] L. Shi, M. Epstein, and R.M. Murray. "Kalman filtering over a packet-dropping network: A probabilistic perspective". In: *IEEE Transactions on Automatic Control* 55.3 (2010), pp. 594–604.

[130] Y. Mo and B. Sinopoli. "Kalman filtering with intermittent observations: Tail distribution and critical value". In: *IEEE Transactions on Automatic Control* 57.3 (2012), pp. 677–689.

[131] K. You, M. Fu, and L. Xie. "Mean square stability for Kalman filtering with Markovian packet losses". In: *Automatica* 47.12 (2011), pp. 2647–2657.

[132] M. Huang and S. Dey. "Stability of Kalman filtering with Markovian packet losses". In: *Automatica* 43.4 (2007), pp. 598–607.

[133] C.C. Huang and R.R. Bitmead. "Escape time formulation of state estimation and stabilization with quantized intermittent communication". In: *Automatica* 61 (2015), pp. 201–210.

[134] Z. Ren, P. Cheng, J. Chen, L. Shi, and Y. Sun. "Optimal periodic sensor schedule for steady-state estimation under average transmission energy constraint". In: *IEEE Transactions on Automatic Control* 58.12 (2013), pp. 3265–3271.

[135] A.S. Leong and S. Dey. "Power allocation for error covariance minimization in Kalman filtering over packet dropping links". In: *51st IEEE Conference on Decision and Control.* Maui, HI, USA, 2012, pp. 3335–3340.

[136] E.A. Feinberg and A. Shwartz. "Constrained dynamic programming with two discount factors: applications and an algorithm". In: *IEEE Transactions on Automatic Control* 44.3 (1999), pp. 628–631.

[137] J. Löfberg. "Approximations of closed-loop minimax MPC". In: *42nd IEEE Conference on Decision and Control.* Maui, HI, USA, 2003, pp. 1438–1442.

[138] A. Ben-Tal, S.P. Boyd, and A. Nemirovski. "Extending scope of robust optimization: Comprehensive robust counterparts of uncertain problems". In: *Mathematical Programming* 107.1-2 (2006), pp. 63–89.

[139] H. Lin and P.J. Antsaklis. "Stability and Stabilizability of Switched Linear Systems: A Survey of Recent Results". In: *IEEE Transactions on Automatic Control* 54.2 (2009), pp. 308–322.

[140] E.J. Davison. *Benchmark Problems for Control System Design: Report of the IFAC Theory Committee.* IFAC, 1990.

[141] E.C. Kerrigan and J.M. Maciejowski. "Designing model predictive controllers with prioritised constraints and objectives". In: *Proceedings. IEEE International Symposium on Computer Aided Control System Design.* IEEE. 2002, pp. 33–38.

[142] S. Tu and B. Recht. "The Gap Between Model-Based and Model-Free Methods on the Linear Quadratic Regulator: An Asymptotic Viewpoint". In: vol. 99. Proceedings of Machine Learning Research. Phoenix, USA: PMLR, June 2019, pp. 3036–3083.

[143] P. Henderson, R. Islam, P. Bachman, J. Pineau, D. Precup, and D. Meger. "Deep Reinforcement Learning That Matters". In: *Conference on Artificial Intelligence (AAAI)*. 2018.

[144] M. Bujarbaruah, T. Zheng, A. Shetty, M. Sehr, and F. Borrelli. "Learning to Play Cup-and-Ball with Noisy Camera Observations". In: *2020 IEEE 16th International Conference on Automation Science and Engineering*. 2020, pp. 372–377.

[145] D. Schwab, J.T. Springenberg, M.F. Martins, T. Lampe, M. Neunert, A. Abdolmaleki, T. Hertweck, R. Hafner, F. Nori, and M.A. Riedmiller. "Simultaneously Learning Vision and Feature-based Control Policies for Real-world Ball-in-a-Cup". In: *CoRR* abs/1902.04706 (2019). URL: `http://arxiv.org/abs/1902.04706`.

[146] S.P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge: Cambridge University Press, 2004.

[147] D.P. Bertsekas. *Convex optimization theory*. Belmont, Massachusetts, 2009.

[148] S.P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics, 1994.

[149] J.L.W.V. Jensen. "Om konvekse Funktioner og Uligheder imellem Middelværdier". In: *Nyt tidsskrift for matematik* 16 (1905), pp. 49–68.

[150] J.L.W.V. Jensen. "Sur les fonctions convexes et les inégalités entre les valeurs moyennes". In: *Acta Mathematica* 30.1 (1906), pp. 175–193.

[151] Gurobi Optimization LLC. *Gurobi Optimizer Reference Manual*. 2020. URL: `http://www.gurobi.com`.

[152] MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 9.0*. 2019. URL: `http://docs.mosek.com/9.0/toolbox/index.html`.

[153] G. Grimmett and D. Stirzaker. *Probability and random processes*. 3rd ed. Oxford: Oxford University Press, 2001.

[154] M. Mitzenmacher and E. Upfal. *Probability and computing : randomized algorithms and probabilistic analysis*. Cambridge, 2005.