UNIVERSITY OF CALIFORNIA, SAN DIEGO

Healthcare Decision Making and Stochastic Model Predictive Control: Output-Feedback, Optimality, and Duality

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Engineering Sciences (Mechanical Engineering)

by

Martin Arno Sehr

Committee in charge:

Professor Robert Bitmead, Chair
Professor Jorge Cortés
Professor Bruce Driver
Professor John Fontanesi
Professor Maurício de Oliveira

2017

The dissertation of Martin Arno Sehr is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2017

DEDICATION

To my friends and family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Personal Acknowledgements

First and foremost, I have to thank my family for their unconditional love and support no matter what my decisions turn out to be. Thank you for unquestioningly accepting my path, even though it keeps requiring physical distance and other sacrifices. I especially want to thank my girlfriend, who has shared this journey with me from the very first day, with hopefully many more journeys to follow down the road. Life and work in San Diego have made for a special and unforgettable time largely because of the people I was very fortunate to meet along the way. The friendships I made during these last few years have really made this an incredible experience.

I would also like to specifically thank my PhD advisor, Professor Robert Bitmead, who has always given me all the flexibility I hoped for while helping and advising whenever needed. I truly appreciate Bob's efforts in guiding me towards growing professionally and as a person. I count myself very fortunate to have worked with an advisor who is as dedicated to his students as Bob has been throughout my time in San Diego. The exceptional degree of flexibility Bob granted me during my time at UC San Diego also allowed me to work with Professor Maurício de Oliveira, whom I wish to thank for helping me out countless times, not only in research matters, but also in making sure I never lacked for teaching positions when I needed them.

My time at UC San Diego was also influenced greatly by my remaining doctoral

committee members. Professor John Fontanesi has helped me make some sense of the intricate details and not least the politics in healthcare. Professor Bruce Driver has been an incredibly helpful and willing mentor during my endeavors into Mathematics. Last but not least, Professor Jorge Cortés has been an excellent instructor, enabling me to gain further insight into nonlinear systems and to take a very enjoyable excursion into cooperative control and multi-agent systems.

Chapter Acknowledgements

Chapter 1, in full, is a reprint of the material as it appears in: M.A. Sehr, R.R. Bitmead, J. Fontanesi, "Multi-Class Appointments in Individualized Healthcare: Analysis for Scheduling Rules", *Proc. 14th European Control Conference*, 2015, pp 1219–1224. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material as it appears in: M.A. Sehr, R.R. Bitmead, J. Fontanesi, "A Framework for Acuity-Based, Individualized Patient Scheduling", *Proc. 54th IEEE Conference on Decision and Control*, 2015, pp 1858–1863. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in: M.A. Sehr, R.R. Bitmead, "Sumptus Cohiberi: The Cost of Constraints in MPC with State Estimates", *Proc. American Control Conference*, 2016, pp 901–906. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, has been submitted for publication of the material as it may appear in: M.A. Sehr, R.R. Bitmead, "Stochastic Model Predictive Control: Output-

Feedback, Duality and Performance", *Automatica*. The dissertation author was the primary investigator and author of this paper.

Chapter 5, in full, is a reprint of the material as it will appear in: M.A. Sehr, R.R. Bitmead, "Particle Model Predictive Control: Tractable Stochastic Nonlinear Output-Feedback MPC", *Proc. IFAC World Congress*, 2017. The dissertation author was the primary investigator and author of this paper.

Chapter 6, in part, has been submitted for publication of the material as it may appear in: M.A. Sehr, R.R. Bitmead, "Performance of Model Predictive Control of POMDPs", *Proc. 56th IEEE Conference on Decision and Control*, 2017. The dissertation author was the primary investigator and author of this paper.

VITA

2010	Bachelor of Science in Mechanical and Process Engineering, TU Darmstadt, Germany.
2012	Master of Science in Mechanical and Process Engineering *with honors*, TU Darmstadt, Germany.
2014	Master of Arts in Applied Mathematics, University of California, San Diego.
2017	Doctor of Philosophy in Mechanical Engineering, University of California, San Diego.

PUBLICATIONS

M.A. Sehr, R.R. Bitmead, "Probing and Duality in Stochastic Model Predictive Control", *To appear in Handbook of Model Predictive Control*.

M.A. Sehr, R.R. Bitmead, "Stochastic Model Predictive Control: Output-Feedback, Duality and Performance", *Submitted to Automatica*.

M.A. Sehr, A.P. Pandey, M.C. de Oliveira, "Pre-Filtering in Continuous-Time Quadratic Gain-Scheduled and Robust State-Feedback Control". *Submitted to International Journal of Control*.

M.A. Sehr, K.D. Joshi, J.M. Fontanesi, R.J. Wong, R.R. Bitmead, R.G. Gish, "Markov Modeling in Hepatitis B Screening and Linkage to Care", *Theoretical Biology and Medical Modelling*, 2017, 14(1).

F.B. Becker, M.A. Sehr, S. Rinderknecht, "Vibration Isolation for Parameter-Varying Rotor Systems using Piezoelectric Actuators and Gain-Scheduled Control", *Journal of Intelligent Material Systems and Structures*, 2017.

M.A. Sehr, M.C. de Oliveira, "Pre- and Post-Filtering in Gain-Scheduled Output-Feedback H_∞ Control", *International Journal of Robust and Nonlinear Control*, 2016.

M.A. Sehr, R.R. Bitmead, "Performance of Model Predictive Control of POMDPs", *Submitted to 56th IEEE Conference on Decision and Control*, 2017.

M.A. Sehr, R.R. Bitmead, "Tractable Dual Optimal Stochastic Model Predictive Control: An Example in Healthcare", *Submitted to IEEE Conference on Control Technology and Applications*, 2017.

M.A. Sehr, R.R. Bitmead, "Particle Model Predictive Control: Tractable Stochastic Nonlinear Output-Feedback MPC", *To appear in Proc. IFAC World Congress*, 2017.

M.A. Sehr, F.B. Becker, M.C. de Oliveira, S. Rinderknecht, "A Catalog of LMI Conditions for Gain-Scheduled Output-Feedback H_∞-Control", *Proc. IEEE Multi-Conference on Systems and Control*, 2016, pp 1060–1065.

A.P. Pandey, M.A. Sehr, M.C. de Oliveira, "Pre-Filtering in Gain-Scheduled and Robust Control", *Proc. American Control Conference*, 2016, pp 3698–3703.

M.A. Sehr, R.R. Bitmead, "Sumptus Cohiberi: The Cost of Constraints in MPC with State Estimates", *Proc. American Control Conference*, 2016, pp 901–906.

M.A. Sehr, R.R. Bitmead, J. Fontanesi, "A Framework for Acuity-Based, Individualized Patient Scheduling", *Proc. 54th IEEE Conference on Decision and Control*, 2015, pp 1858–1863.

F.B. Becker, M.A. Sehr, S. Rinderknecht, "Gain-Scheduled H-Infinity-Control for Active Vibration Isolation of a Gyroscopic Rotor", *Proc. ASME Conference on Smart Materials, Adaptive Structures and Intelligent Systems*, 2015.

M.A. Sehr, R.R. Bitmead, J. Fontanesi, "Multi-Class Appointments in Individualized Healthcare: Analysis for Scheduling Rules", *Proc. 14th European Control Conference*, 2015, pp 1219–1224.

M.A. Sehr, A.P. Pandey, M.C. de Oliveira, "Robust Stabilization of Linear Continuous-Time Varying Systems", *Proc. American Control Conference*, 2015, pp 108–113.

A.P. Pandey, M.A. Sehr, M.C. de Oliveira, "Stability Criteria for Uncertain Time-Varying Systems", *Proc. 53rd IEEE Conference on Decision and Control*, 2014, pp 4795–4800.

B. Riemann, M.A. Sehr, R.S. Schittenhelm, S. Rinderknecht, "Real Gyroscopic Uncertainties in Robust Control of Flexible Rotors", *Proc. 52nd IEEE Conference on Decision and Control*, 2013, pp 3762–3769.

B. Riemann, M.A. Sehr, R.S. Schittenhelm, S. Rinderknecht, "Robust Control of Flexible High-Speed Rotors via Mixed Uncertainties", *Proc. 12th European Control Conference*, 2013, pp 2343–2350.

ABSTRACT OF THE DISSERTATION

Healthcare Decision Making and Stochastic Model Predictive Control: Output-Feedback, Optimality, and Duality

by

Martin Arno Sehr

Doctor of Philosophy in Engineering Sciences (Mechanical Engineering)

University of California, San Diego, 2017

Professor Robert Bitmead, Chair

Model Predictive Control has become a prevailing technique in practice by virtue of its natural inclusion of constraint enforcement in sub-optimal feedback design through repeated solution of finite-horizon, open-loop control problems. However, many approaches are lacking in proper accommodation of output feedback using imperfect measurements, as is normally required in practice. The conventional workaround for this disconnect between control theory and practice is the use of certainty equivalent control laws, which subsume best available state estimates in place of the system state in order

to salvage methods available for state-feedback Model Predictive Control.

This dissertation explores Stochastic Model Predictive Control in the general, nonlinear output-feedback setting. Starting the receding horizon development from Stochastic Optimal Control, we attain inherent accommodation of imperfect measurement data through propagation of the conditional state density, the information state. This setup further results in the control signals being of dual, probing nature: the control balances the typically antagonistic requirements of regulation and exploration. However, these conflicting tasks inherent to Stochastic Optimal Control also embody the associated computational intractability. While properties such as optimal probing and numerical performance bounds on the infinite time-horizon require solution of Stochastic Optimal Control problems, obtaining these solutions is typically not possible in practice due to the exorbitant computational demands.

We suggest two methods for tractable Stochastic Model Predictive Control. Firstly, we propose approximation of the information state update by a Particle Filter, which may be merged naturally with scenario optimization to generate control laws. While computationally tractable, this method does not maintain duality without additional measures. Alternatively, the nonlinear output-feedback problem can be approximated – or even cast – as a Partially Observable Markov Decision Process, a special class of systems for which Stochastic Optimal Control is numerically tractable for reasonable problem size, enabling dual optimal control with provable infinite-horizon properties.

Throughout this dissertation, we examine two classes of examples from healthcare: individualized appointment scheduling, a problem not requiring duality; medical

treatment decision making, where dual control decisions are often required to balance optimally when to order diagnostic tests and when to apply medical intervention.

Introduction

Background and Motivation

Model Predictive Control (MPC) is well applied and popular partially because of its capacity to handle constraints and its simple formulation as an open-loop, finite-horizon optimization problem evaluated on the receding horizon [1, 2]. There are a few areas in which MPC is wanting for more complete results, notably in the area of output feedback control and the associated requirement to manage the duality of the control signal in stochastic MPC (SMPC) problems.

While MPC, in its original formulation, is a full-state feedback law [3], there has been a number of approaches to output-feedback MPC. Most approaches hinge on replacement of the measured true state by a state estimate, which is computed via Kalman filtering (e.g. [4]), moving-horizon estimator (e.g. [5, 6]), tube-based minimax estimators (e.g. [7]), etc. These designs, often for linear systems, often separate the estimator design from the control design. The control problem may be altered to accommodate state estimation errors by methods such as: constraint tightening as in [4], chance/prob-

abilistic constraints as in [8, 9], and so forth. Likewise, for nonlinear problems, where the state estimation behavior is affected by control signal properties, the control may be modified to enhance the excitation properties of the estimator, as suggested in [10, 11]. Each of these aspects of accommodation is made in an isolation.

When output-feedback SMPC is developed as a logical extension of finite-horizon Stochastic Optimal Control, which demands computation of closed-loop policies, it inherits the computational intractability of this latter subject via the inclusion of the Bayesian filter, required to propagate the conditional state densities, and the stochastic dynamic programming equation. In the general nonlinear setting, Stochastic Optimal Control involves the propagation of the conditional probability density of the state given the input signal and output measurements. This density is known as the *information state* in controls and as the *belief state* in artificial intelligence and robotics. The choice of control signal affects the information state so that state observability becomes control-dependent. Thus, the feedback control law needs to include aspects of probing in addition to, or more accurately in competition with, its function in regulation. This is called *duality* of the control. In the linear case, this connection is not problematic since the control signal simply translates or recenters the conditional density without other effects. But for nonlinear systems, this complication renders all but the simplest optimal control problems computationally intractable. The usual recourse is to drop optimality and to use a more simply computed or approximated statistic from the conditional density, such as the conditional mean, and to move on from there.

This dissertation explores the origin and accommodations of duality in Stochas-

tic Optimal Control and, by implication, in SMPC. The development is guided by a number of practical examples in healthcare, which can be classified into two categories: problems not demanding duality in the control solution, and those that inherently require dual control laws. We start our discussion by exploring the particular issue of appointment scheduling in individualized healthcare, which constitutes Part I of this dissertation. Appointment scheduling requires solution of challenging optimization problems which do not lead to duality in the resulting open-loop decisions, the origin of which property we elucidate.

In contrast, Part II of the dissertation discusses problems involving duality explicitly and develops a set of theoretical results and practical algorithms. The second part of the dissertation starts with motivation for the development of SMPC through Stochastic Optimal Control in Chapter 3 and proofs of a number of key results for dual optimal SMPC laws in Chapter 4. Unfortunately, this variant of SMPC is computationally intractable, requiring further measures for practical implementation. To this end, we motivate two approaches. Firstly, we suggest a particle-based scenario approximation with loss of duality in Chapter 5. Secondly, we specifically discuss Partially Observable Markov Decision Processes (POMDPs) in Chapter 6. This class of systems makes up one of the few special cases in which Stochastic Optimal Control may be computationally tractable. Chapter 6 further explores a numerical example in medical treatment of a patient treated for a hypothetical disease, requiring the health professional to trade off diagnostics and medical intervention, an inherently dual decision problem. We next give a more detailled preview of the chapters making up this dissertation.

Chapter Contents and Contributions

In Chapter 1, we consider the problem of scheduling patients for visits at a cancer infusion room throughout a regular day. We suggest the use of high and low patient acuity indicators to account for punctuality and service uncertainties in the scheduling process. These supportive classifications can be used easily by schedulers to allow more efficient, individualized service. Based on patient acuity data and clinical observations, we propose two intuitive though somewhat conflicting scheduling guidelines on a qualitative basis and argue their benefits. We make use of analogies with standard queueing theory and strings of interconnected dynamic systems to introduce two separate surrogate problems allowing analysis of the effects resulting from our scheduling rules on the operation of the infusion room. As part of the theme of this thesis, acuity information is state information whose application leads to improved performance results.

In Chapter 2, we consider scheduling finite numbers of patients for visits to a medical clinic throughout a workday. Use of individual patient acuity level indicators in the scheduling process is explored to accomodate fluctuations in service requirements. These acuity classifications can be used by schedulers to enable more efficient, individualized service. After analyzing sojourn time behavior for patient populations with arbitrary deterministic arrivals and Erlang service time distributions in a queueing setup, we motivate and develop a Markov model suitable to account for general time-varying service time distributions in the scheduling process. The significance of this work is that it provides a mathematical basis for the analysis of acuity-based scheduling rules

currently applied and proven effective in healthcare. This provides a foundation from which to develop and validate new scheduling rules in individualized medicine. The linkage to the proposed Markov model permits the conversion of particular scheduling questions into Markov Decision Problems.

There are several aspects of Model Predictive Control (MPC) which are often ignored: use of state estimates, stochastic disturbances, robustness outside of full state availability. In Chapter 3, we raise awareness of one of these issues. The appeal of MPC in applications rests primarily with its capacity to accommodate constraints, which in turn equips the designer with both an objective function and a higher-priority set of constraints, which meshes well with the engineering control formulation. Yet, MPC in industrial applications is principally a disturbance rejection controller targeted at the regulation of plant set points in the face of stochastic environmental disturbances. Perversely for such an implementation, MPC is also posed as a full-state feedback problem, where this state should include the disturbance process state, necessitating the use of approximate state estimates. The chapter considers the interplay between state estimation errors and constraints in MPC and exposes the feedthrough of these errors to the MPC input signals resulting from the solution of finite-horizon constrained optimization problems. We show how the MPC solution injects measurement noise directly into the control signal entering the plant and demonstrate the increased sensitivity to this noise when the plant is operating on active constraints. This reveals a downside of the use of constrained control with state estimation that is generally flouted in MPC.

In Chapter 4, a new formulation of Stochastic Model Predictive Output Feed-

back Control is presented and analyzed as a translation of Stochastic Optimal Output Feedback Control into a receding horizon setting. This requires lifting the design into a framework involving propagation of the conditional state density, the *information state*, via the Bayesian Filter and solution of the Stochastic Dynamic Programming Equation for an optimal feedback policy, both stages of which are computationally challenging in the general, nonlinear setup. The upside is that the clearance of three bottleneck aspects of Model Predictive Control is connate to the optimality: output feedback is incorporated naturally; dual regulation and probing of the control signal is inherent; closed-loop performance relative to infinite-horizon optimal control is guaranteed. While the methods are numerically formidable, our aim is to develop an approach to Stochastic Model Predictive Control with guarantees and, from there, to seek a less onerous approximation.

In Chapter 5, we combine conditional state density construction with an extension of the Scenario Approach for stochastic Model Predictive Control to nonlinear systems to yield a novel particle-based formulation of stochastic nonlinear output-feedback Model Predictive Control. Conditional densities given noisy measurement data are propagated via the Particle Filter as an approximate implementation of the Bayesian Filter. This enables a particle-based representation of the conditional state density, or information state, which naturally merges with scenario generation from the current system state. This approach attempts to address the computational tractability questions of general nonlinear stochastic optimal control. The Particle Filter and the Scenario Approach are shown to be fully compatible and – based on the time- and measurement-update

stages of the Particle Filter – incorporated into the optimization over future control sequences. A numerical example is presented and examined for the dependence of solution and computational burden on the sampling configurations of the densities, scenario generation and the optimization horizon.

In Chapter 6, we revisit closed-loop performance guarantees for Model Predictive Control in the deterministic and stochastic cases, which extend to novel performance results applicable to receding horizon control of Partially Observable Markov Decision Processes. While performance guarantees similar to those achievable in deterministic Model Predictive Control can be obtained even in the stochastic case, the presumed stochastic optimal control law is intractable to obtain in practice. However, this intractability relaxes for a particular instance of stochastic systems, namely Partially Observable Markov Decision Processes, provided reasonable problem dimensions are taken. This motivates extending available performance guarantees to this particular class of systems, which may also be used to approximate general nonlinear dynamics via gridding of state, observation, and control spaces. We demonstrate applicability of the novel closed-loop performance results on a particular example in healthcare decision making, which relies explicitly on the duality of the control decisions associated with Stochastic Optimal Control in weighing appropriate appointment times, diagnostic tests, and medical intervention for treatment of a disease modeled by a Markov Chain.

In summary, the contributions of this thesis are twofold: we first discuss open-loop decision problems in healthcare and then transition to closed-loop stochastic output-feedback problems, which provide motivation for the novel work on Stochastic Model

Predictive Control. The appointment scheduling problems discussed in Part I of the thesis do not require dual solutions and probing because the full state is accessible whenever decisions have to be made. This property is in contrast with the examples discussed in Part II of the thesis and in particular Chapter 6, where instantaneous medical treatment decisions have to be made, requiring careful balancing of diagnostic tests and medical intervention, an inherently dual problem. Open-loop solutions such as those observed in appointment scheduling do not enable the necessary use of diagnostic tests, which serve to enhance our state knowledge at a given cost without altering the patient state. We use these *dual* decision problems as motivation for developing the theoretical results in Chapters 4 and 6 as well as the practical algorithm presented in Chapter 5.

Part I

Modeling for Analysis in Healthcare

Chapter 1

Multi-class Appointments in Individualized Healthcare: Analysis for Scheduling Rules

1.1 Introduction

We consider the problem of scheduling a certain number of patients for visits to a cancer infusion room throughout a usual business day. The infusion room has limited numbers of chairs and nurses to service patients and every patient has to receive an individual, prescribed number of infusion bags, all of which require the same amount of time. Some patients have to go through phlebotomy and additional health checks to ensure their medical condition permits the infusion procedure. Currently, patients are scheduled for infusion room visits on a first-come, first-served basis for a time deter-

mined by their prescription. Individual patient characteristics apart from the number of prescribed infusion bags are not taken into account by the scheduler when organizing the workload throughout the day. This scheduling policy results in significant fluctuations when comparing waiting times, patient throughput or satisfaction and other performance measures over the course of multiple days. On some days, all chairs are occupied for most of the time and the nurses spend little to no time idle. However, on other days, approximately one third of the chairs are unoccupied while all nurses are at their constrained maximal workload and the waiting room is overflowing.

A cause for these discrepancies may be the unavailability of individual patient acuity information in the scheduling procedure and the unawareness of the appropriate scheduling action were it to be available. While most physicians have a good idea of which of their patients may present a higher uncertainty in the sense of fluctuations in punctuality and unexpected service requirements, the scheduler does not currently have access to this information. We suggest the classification of the patient population into acuity classes, which can be performed reliably by the corresponding doctor. High acuity in the infusion room setup refers to high severity of the prescribed medication and usually equates with advanced stages of disease. High-acuity patients frequently have; a level of frailty and impairment which causes them to fail pre-infusion lab tests, which themselves introduce an uncontrolled delay; and, a propensity to require nurse intervention during the infusion. Individual patient acuity is a professional assessment which is instinctive for treating physicians but largely undocumented for the schedulers. From the perspective of the analysis in this chapter, high acuity is equated with high

variability and vice versa. Successful strategies for the inclusion of individual patient acuity into scheduling are proposed and studied, while the prime focus is on unearthing the appropriate mathematical tools capable of yielding these strategies.

Using patient acuity information in the scheduling process allows the scheduler to ensure a steadier operation of the infusion room. In the following, we consider the special case of two distinct acuity classes, namely high- and low-acuity. The difference between these two classes is that high-acuity patients are less likely to be on time and have less predictable service requirements than low-acuity patients. That is, both service and arrival variances are higher for high-acuity patients. To use patient acuity data, we propose the following qualitative scheduling rules:

Rule 1: Avoid scheduling many high-acuity patients at once.

Rule 2: Schedule high-acuity patients later in the day.

Rule 1 aims to avoid periods of large accumulated uncertainty, while Rule 2 avoids disturbances from being propagated through the workday[1]. Issues closely related to Rule 2 are also discussed in the literature concerning surgical procedure scheduling, where it was observed that performing high-variance procedures after their low-variance counterparts yields better operating room performance (e.g., [12, 13]). As we are going to illustrate in this chapter, Rules 1 and 2 for appointment scheduling permit more efficient service, directly leading to reduced and more predictable waiting times. In combination

[1]Interestingly, physicians often prefer having their harder cases scheduled early in the day, contrary to the observation that their performance tends to be better at later times. The preference for treating harder cases early in the day comes from the impression of being more alert at that time. Given that the operation of the infusion room does not actively involve the treating doctors, we do not discuss this issue any further.

with improved exchange of patient data, the two rules were implemented temporarily at UCSD Moores Cancer Center, leading to improvements of approximately 26% in the total number of patient visits per day, while the number of no-shows decreased from approximately 13% to below 1% and the number of late arrivals (15 minutes or more) dropped from 12% to less than 1%. While improvements in late arrivals and no-shows were attributed mainly to an increased exchange of information regarding other, potentially conflicting, appointments patients had on the same day, the increased numbers of patients seen per day were ascribed to the two scheduling rules above. After resuming with the conventional scheduling procedure (i.e., first-come, first-served appointments), the numbers of patients seen per day returned to the previous levels.

Perhaps surprisingly, mathematical analysis suggesting Rules 1 and 2 for appointment scheduling is not straightforward and even though the setup appears to be similar to standard queueing problems, there are some key differences requiring a different type of analysis, which is the aim of this chapter. Core reasons why the scheduling problems discussed here are different from standard queueing problems are the explicit finite-horizon and transient characteristics of the problem, while queueing theory deals primarily with infinite-horizon, steady-state results (e.g., [14, 15]). For instance, Rule 2 relies explicitly on an analysis capturing transition from busy operation of the infusion room to the end of the workday, which cannot be captured using an infinite-horizon stationary analysis. However, we can use an analysis strongly related to queueing problems but allowing for non-stationary behavior to estimate the effects of Rule 1 on the scheduling process, which is discussed in Section 1.2. Another body of analysis displaying par-

allels to the problem at hand is the study of interconnected dynamic systems (e.g., [16]), where disturbances are propagated through leader-follower strings of coupled dynamic systems. Although these problems may appear farther from the multi-class appointment scheduling problem discussed here, they show some useful similarities with respect to finite-horizon patient ordering as required to synthesize Rule 2, which will be examined in Section 1.3. For practical implementation, the scheduling rules introduced here are conflicting to some degree, requiring the scheduler to perform a tradeoff decision when adding appointments to a given day. Such implementation issues will be the focus of future studies and only hinted at in this chapter.

1.2 Rule 1: Balancing Workload

The aim of this section is the synthesis of Rule 1 using ideas from calculus of variations, stochastic differential equations and queueing theory. To approach this goal, we consider the surrogate problem displayed in Figure 1.1, where two streams of customers arrive continuously at a single deterministic server with incremental service rate

$$S = m + \varepsilon$$

for some number $\varepsilon > 0$ and $\varepsilon \ll m$. The two streams of customers are mixed and arrive with random incremental amounts of workload to be serviced in the same buffer, which operates on a first-come first-served basis. The customer streams are distinguished by having the same mean incremental workload but differing variances; the aim is to model

scheduling of patients with high and low acuity levels. The low-variance stream of customers has a workload generation rate of X_t for $t \geq 0$, where $(X_t)_{t\geq 0}$ is a stationary nonnegative left-continuous stochastic process with mean value $\mathbb{E}[X_t] = m$ and variance $\mathrm{Var}(X_t) = \sigma_l^2$ for all $t \geq 0$. Analogously, customers from the high-variance stream have workload generation rate Y_t, where $(Y_t)_{t\geq 0}$ is another stationary nonnegative left-continuous stochastic process, independent from $(X_t)_{t\geq 0}$, with mean value $\mathbb{E}[Y_t] = m$ and variance $\mathrm{Var}(Y_t) = \sigma_h^2 \geq \sigma_l^2$. The long-run average fraction of customers from the highly variable stream is μ and that from the less variable stream is $1 - \mu$, respectively. Workload and arrival processes for this setup are described by the stochastic differential equation

$$\frac{dW_t}{dt} = \max\{A_t - S, -W_t\}, \tag{1.1}$$

$$A_t = (1 - f_t)X_t + f_t Y_t, \tag{1.2}$$

for some $W_0 \geq 0$ and a deterministic mixture process f_t with $t \in [0, T]$ satisfying

$$f_t \in [0, 1]\ \forall t \in [0, T], \qquad \frac{1}{T}\int_0^T f_t\, dt = \mu. \tag{1.3}$$

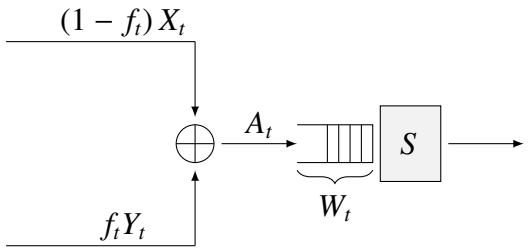

Figure 1.1: Workload process with mixed arrival streams.

Interpretation of this workload process with mixed arrivals is the potentially non-stationary scheduling of high- and low-acuity patients in the operation of the infusion room via f_t, where the highly variable stream Y_t with mean value m and variance σ_h^2 models high-acuity and the less variable stream X_t with mean value m and variance σ_l^2 models low-acuity patients, respectively. In particular, this auxiliary model allows us to synthesize the stationary Rule 1. Keeping the problem of unsteady operation of the infusion room in mind, we are interested in finding a mixture function f_t that minimizes the variance of the workload buffer length (i.e., total amount of workload in the system) over the time interval $t \in [0, T]$. As is the case in standard queueing theory,

$$\mathbb{E}[S - A_t] = \varepsilon > 0$$

for all $t \geq 0$ implies that the workload process is stable over time for any mixing process f_t satisfying the constraints (1.3) for $t \in [0, T]$. However, since $\varepsilon \ll m$, the server is busy most of the time, which is the case when operating the infusion room near its maximum capacity and suggests that the variance of dW_t/dt and thus that of W_t is dominated by the variance of the arrival process A_t. From (1.2), we have

$$\begin{aligned}
\mathrm{Var}(A_t) &= \mathbb{E}\left[((1 - f_t)X_t + f_t Y_t - m)^2\right] \\
&= \mathbb{E}\left[((1 - f_t)X_t + f_t Y_t)^2\right] - m^2 \\
&= (1 - f_t)^2 \left(\sigma_l^2 + m^2\right) + f_t^2 \left(\sigma_h^2 + m^2\right) \\
&\quad + 2 f_t (1 - f_t) m^2 - m^2 \\
&= \left(\sigma_l^2 + \sigma_h^2\right) f_t^2 - 2 f_t \sigma_l^2 + \sigma_l^2
\end{aligned}$$

for all $t \geq 0$. Having deterministic service at the buffer, we can indirectly minimize the overall workload variance by choosing a mixture process $(f_t)_{t\geq 0}$ satisfying (1.3) and minimizing the average variance of the arrival process over time. On the fixed time horizon $t \in [0, T]$, this yields the finite-horizon functional optimization problem

$$\begin{aligned} &\min_{(f_t)_{t=0}^T} \int_0^T \left(\left(\sigma_l^2 + \sigma_h^2\right) f_t^2 - 2\sigma_l^2 f_t + \sigma_l^2\right) dt \\ &\text{s.t.} \quad 0 \leq f_t \leq 1 \ \forall t \in [0, T], \\ &\qquad \frac{1}{T} \int_0^T f_t \, dt = \mu. \end{aligned} \tag{1.4}$$

The global solution f_t^* of this functional optimization problem minimizes the cumulative and thus average arrival variance over the time interval $t \in [0, T]$. The solution to this problem is summarized by the following result.

Theorem 1.1. *Given $\mu \in [0, 1]$, the unique global minimizer of problem* (1.4) *for all parameters $\sigma_h^2 \geq \sigma_l^2 \geq 0$ and $T \geq 0$ is $f_t^* = \mu$ for all $t \in [0, T]$. The resulting average variance of the arrival process A_t over the interval $t \in [0, T]$ is then given by*

$$\frac{1}{T} \int_0^T \mathrm{Var}(A_t) \, dt = (\mu - 1)^2 \, \sigma_l^2 + \mu^2 \sigma_h^2.$$

Proof. Define

$$J(f_t) := \frac{1}{T} \int_0^T \mathrm{Var}(A_t) \, dt$$

for all mixing processes f_t satisfying the constraints (1.3) on the interval $t \in [0, T]$. Moreover, write

$$f_t := \mu + \eta_t = f_t^* + \eta_t \tag{1.5}$$

for some function η_t and $t \in [0, T]$. Notice that the construction (1.5) encompasses all mixing processes satisfying the constraints (1.3) on the compact set $0 \leq t \leq T$ if and only if the function η_t is confined to

$$\eta_t \in [-\mu, 1-\mu] \; \forall t \in [0, T], \qquad \int_0^T \eta_t \, dt = 0.$$

Evaluating $J(f_t)$ for any feasible function η_t, we then have

$$\begin{aligned} J(f_t) &= \left((\mu-1)^2 \sigma_l^2 + \mu^2 \sigma_h^2\right) T + 2\left((\mu-1)\sigma_l^2 + \mu\sigma_h^2\right) \int_0^T \eta_t \, dt + \left(\sigma_l^2 + \sigma_h^2\right) \int_0^T \eta_t^2 \, dt \\ &= \left((\mu-1)^2 \sigma_l^2 + \mu^2 \sigma_h^2\right) T + \left(\sigma_l^2 + \sigma_h^2\right) \int_0^T \eta_t^2 \, dt. \end{aligned}$$

Clearly, $\eta_t^2 \geq 0$ for all $t \geq 0$, implying that $J(f_t)$ has a global minimum

$$J(f_t^* = \mu) = \left((\mu-1)^2 \sigma_l^2 + \mu^2 \sigma_h^2\right) T$$

with $\eta_t = 0$ for all $t \in [0, T]$, which concludes the proof. □

As anticipated, the analysis of the surrogate problem introduced in this section leads to Rule 1 for patient appointment scheduling in the infusion room setup, which serves as the running example in this chapter. Given a large population with a certain fraction of high-acuity patients who can be mixed in continuously, the optimal scheduling strategy when concerned with average arrival variance is to distribute the high-acuity patients evenly throughout the day. Rule 1 says there is no benefit in average arrival variance to lumping together high-acuity patients in a non-stationary arrival scenario with continuous mixing. The function f_t is the open-loop control which admits the possibility, but not the optimality, of leaving the high-acuity patients unserved until late

in the day. The formulation operates under a heavy-traffic assumption as common in scheduling theory, so that the workload rarely returns to zero. This is consistent with the efficient operation of the infusion room. Under stationary arrivals of high- and low-acuity patients and continuous, potentially non-stationary mixing via f_t, the resulting optimal open-loop control with respect to average workload arrival variance is precisely the stationary Rule 1.

While the analysis of the auxiliary problem introduced in this section justifies Rule 1 for patient appointment scheduling, it does not serve as a tool to synthesize Rule 2. The central reason for being unable to distinguish effects of Rule 2 using the mixing process scenario is the limitation of the analysis to the workload arrival process through our heavy-traffic assumption. This assumption is based on the service rate being only slightly higher than the mean arrival rate, which is constant for any open-loop control f_t. Using this assumption, we were able to conclude that the workload rarely returns to zero, such that we diverted our analysis from the variance of the workload process W_t to the variance of the arrival process A_t. While this indeed reasonable assumption allows for simplified analysis of the problem, it blends out downstream-effects of large accumulated workload early in the day on the total workload variance late in the day through the system dynamics (1.1). Hence, a different type of modeling and analysis is required in order to examine the effects of Rule 2, which is discussed in the following section.

When concerned with scheduling small numbers of patients, further need for a non-stationary scheduling guideline such as Rule 2 is related to implementation is-

sues arising with respect to the range of the open-loop control f_t in (1.4), allowing for arbitrary fractions of low- and high-acuity patients to be scheduled at any point in time. Contrary to the setup used in this section, implementing a scheduling rule requires scheduling either a high- or a low-acuity patient for service with a single nurse at any given time, not fractions of both. The assumption of scheduling fractions of patients becomes more reasonable when the server in Figure 1.1 models nurses helping multiple patients simultaneously. However, even in this case only a finite range for the function f_t would be admissible for a given finite number of nurses. That is, the solution in Theorem 1.1 may not produce a feasible scheduling policy for operation of the infusion room, consequently requiring us to come up with an additional, non-stationary scheduling guideline such as Rule 2. Finally, notice that the above mixing process scenario extends easily to the case of discrete-time arrivals and mixing, provided the structural properties of the problem do not change drastically (ie. stable process under heavy traffic etc.). The resulting discrete-time version of Theorem 1.1 may be more intuitive in relation to the queueing relationships of the examined patient scheduling problem. However, we chose to present the more general continuous-time result and leave the details of this extension to the reader. Another aspect that should be kept in mind is that we considered the continuous, cumulative arrival of workload at the server, while in reality there may be a high number of patients associated with little cumulative workload and vice versa, which becomes a factor when taking limited numbers of nurses and chairs in waiting and infusion rooms into account. However, this issue is not thematized further in this chapter.

1.3 Rule 2: Disturbance Propagation

As we have seen in the previous section, we need to take the dynamics of our workload buffer into account when attempting to analyze Rule 2 for patient scheduling in the infusion room setup. In this section, we first show how to extend the mixing process introduced above for a particular example. To solve for workload variances at different time instances, we have to take the distributions of the two patient arrival processes into account. To get a more general and easily applicable setup, we make use of parallels with strings of vehicles with different dynamics in leader-follower formation. This setup results in Rule 2 for individual patient appointment scheduling and gives rise to approaches for combining the two scheduling rules.

1.3.1 Discrete-Time Mixing Process

Consider a particular instance of the discrete-time version of (1.4) with confined range of f_k corresponding to eight patients being scheduled for service by two nurses sharing the workload buffer (see Figure 1.1) over four time instances. If one of the eight patients has high acuity, we have $\mu = 1/8$, leading to the discrete optimization problem

$$\begin{aligned} \min_{f_1,\dots,f_4} \ & \sum_{k=1}^{4} \left(\left(\sigma_l^2 + \sigma_h^2 \right) f_k^2 - 2\sigma_l^2 f_k + \sigma_l^2 \right) \\ \text{s.t.} \ & f_k \in \{0, 1/2, 1\} \ k \in \{1, \dots, 4\}, \\ & \sum_{k=1}^{4} f_k = 1/2. \end{aligned} \tag{1.6}$$

By the range constraints on f_k, this problem does not admit the solution suggested by Rule 1, requiring us instead to find a non-stationary policy f_k, where $k \in \{1, \dots, 4\}$. In

total, there exist four feasible, non-stationary policies for problem (1.6) with

$$\begin{aligned} f^{(1)} &= \{1/2, 0, 0, 0\}, \qquad f^{(2)} = \{0, 1/2, 0, 0\}, \\ f^{(3)} &= \{0, 0, 1/2, 0\}, \qquad f^{(4)} = \{0, 0, 0, 1/2\}. \end{aligned} \tag{1.7}$$

One can easily see that these four policies are permutations of each other and indistinguishable through the cost

$$J(f_k^{(i)}) = \sum_{k=1}^{4} \left(\left(\sigma_l^2 + \sigma_h^2\right) \left(f_k^{(i)}\right)^2 - 2\sigma_l^2 f_k^{(i)} + \sigma_l^2 \right). \tag{1.8}$$

However, if we analyze the variance of the discrete-time workload process

$$W_{k+1} = \max\{W_k + A_{k+1} - S, 0\} \quad \text{with} \quad W_0 \geq 0$$

instead of the variance of the discrete-time arrival process

$$A_{k+1} = \left(1 - f_{k+1}^{(i)}\right) X_{k+1} + f_{k+1}^{(i)} Y_{k+1}$$

corresponding to (1.2), the policies $f^{(i)}$ are indeed distinguishable in that the variances of initial arrivals have influence on the variances of the overall workload at later times. However, solving for $\mathrm{Var}(W_k)$ requires knowledge of the distributions of the discrete-time arrival processes X_k and Y_k, not only their respective means and variances. Consider for instance problem (1.6) with initial condition $W_0 = m/2$ and the independent and finitely distributed arrival processes

$$\begin{aligned} P(X_k = 0) &= 1/8, \qquad P(Y_k = 0) &= 3/8, \\ P(X_k = m) &= 3/4, \qquad P(Y_k = m) &= 1/4, \\ P(X_k = 2m) &= 1/8, \qquad P(Y_k = 2m) &= 3/8, \end{aligned}$$

where $k \in \{0, 1, 2, 3\}$. For the four feasible scheduling policies (1.7), we can use $\varepsilon \ll m$ to obtain the approximate workload variances summarized in Table 1.1. These variances indicate clearly how policy $f^{(4)}$ yields the lowest workload variance at any given time. Hence, scheduling policy $f^{(4)}$ is the preferable strategy in this case, even though the four policies are indistinguishable through the cost (1.8).

Table 1.1: Workload variances for policies (1.7) and $W_0 = m/2$.

Schedule	$\mathrm{Var}(W_1)$	$\mathrm{Var}(W_2)$	$\mathrm{Var}(W_3)$	$\mathrm{Var}(W_4)$
$\{1/2, 0, 0, 0\}$	$0.214\, m^2$	$0.341\, m^2$	$0.459\, m^2$	$0.570\, m^2$
$\{0, 1/2, 0, 0\}$	$0.152\, m^2$	$0.328\, m^2$	$0.450\, m^2$	$0.564\, m^2$
$\{0, 0, 1/2, 0\}$	$0.152\, m^2$	$0.289\, m^2$	$0.435\, m^2$	$0.553\, m^2$
$\{0, 0, 0, 1/2\}$	$0.152\, m^2$	$0.289\, m^2$	$0.415\, m^2$	$0.539\, m^2$

The average variances

$$\mathrm{Var}(W^{(i)}) := \frac{1}{4} \sum_{k=1}^{4} \mathrm{Var}(W_k(f^{(i)}))$$

for the four scheduling policies are

$$\mathrm{Var}(W^{(1)}) \approx 0.396\, m^2, \qquad \mathrm{Var}(W^{(2)}) \approx 0.374\, m^2,$$

$$\mathrm{Var}(W^{(3)}) \approx 0.357\, m^2, \qquad \mathrm{Var}(W^{(4)}) \approx 0.349\, m^2.$$

Hence, we observe the trend that for the given setup, the average workload variance decreases as the single high-acuity patient is scheduled later in the day, which is precisely the motivation of Rule 2 for patient appointment scheduling. However, to arrive at this conclusion, we had to take the explicit workload arrival distributions corresponding to the two patient classes into account. We seek to find a more general, qualitative answer

to this issue that does not involve distributions of random processes that would have to be estimated for practical implementation. Thus, we proceed by introducing a second surrogate problem.

1.3.2 Vehicle Traffic

To examine how variances in the operation of the infusion room are propagated downstream in time, we make use of parallels with interconnected dynamic systems theory, which finds application in a number of problems including highway traffic [17], which in turn has similar characteristics to the appointment scheduling issue. In the appointment scheduling case, we try to find desirable time arrangements of different patients each subject to individual disturbances and following one another, whereas in the vehicle traffic case one may venture to find preferable relative orderings of individual vehicles each subject to their own perturbations.

In the following, we attempt to extract Rule 2 from a highway traffic platooning example, which is related to the example discussed in [17]. This problem serves as a parallel replacement problem to analyze effects of having larger disturbances at the front or end of a string of interconnected dynamic systems. However, we are interested in leader-follower formations with different individual dynamics, whereas [17] examines the case of formations with identical dynamics. Consider now a leader-follower platoon of vehicles on a freeway modeled as

$$m_i \ddot{x}_i(t) = u_i(t) + w_i(t), \qquad i = 1, \ldots, N,$$

where m_i, $x_i(t)$, $u_i(t)$ and $w_i(t)$ are the i^{th} vehicle's mass, position, control input and perturbation, respectively. That is, the vehicle dynamics differ in their masses and thus susceptibilities to the disturbances $w_i(t)$, where the lighter vehicle is more sensitive and thus associated with higher state variance. We interpret these vehicles as more or less variable patients scheduled for sequential service with a nurse. A state-space representation of this system with state

$$z_1(t) = \begin{bmatrix} x_1(t) \\ \\ \dot{x}_1(t) \end{bmatrix}, \qquad z_i(t) = \begin{bmatrix} x_i(t) - x_{i-1}(t) \\ \\ \dot{x}_i(t) - \dot{x}_{i-1}(t) \end{bmatrix}$$

for $i = 2, \ldots, N$ is

$$\dot{z}(t) = \begin{bmatrix} 0 & I \\ \\ 0 & 0 \end{bmatrix} z(t) + \begin{bmatrix} 0 \\ \\ B \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ \\ B \end{bmatrix} w(t),$$

where

$$b_{ij} = \begin{cases} 1/m_i, & \text{for } i = j \\ -1/m_{i-1}, & \text{for } i = j + 1 \\ 0, & \text{else} \end{cases} .$$

We assume that each vehicle's control input $u_i(t)$ depends only on relative distance and velocity information with respect to the vehicle ahead or a potential tracking reference. We thus introduce a decentralized feedback control law of the form

$$u(t) = \begin{bmatrix} K & K \end{bmatrix} z(t), \qquad K = -\operatorname{diag}(m_1, \ldots, m_N).$$

We then have

$$\dot{z}(t) = \begin{bmatrix} 0 & I \\ \\ BK & BK \end{bmatrix} z(t) + \begin{bmatrix} 0 \\ \\ B \end{bmatrix} w(t)$$

with the particular structure

$$BK = \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{bmatrix}.$$

The individual vehicle positions and velocities can be recovered from $z(t)$ using the transformation

$$\begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & T \end{bmatrix} z(t),$$

where T is a lower triangular matrix of ones. Notice that we have $T^{-1} = -BK$, such that

$$\begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ BK & BK \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ -K^{-1} \end{bmatrix} w(t)$$

and the steady-state covariance matrix $P \succ 0$ for stationary white noise $w(t)$ with covariance $W \succ 0$ can be obtained by solving the Lyapunov matrix equation

$$\begin{bmatrix} 0 & I \\ BK & BK \end{bmatrix} P + P \begin{bmatrix} 0 & K^T B^T \\ I & BK \end{bmatrix} = \begin{bmatrix} 0 \\ K^{-1} \end{bmatrix} W \begin{bmatrix} 0 & K^{-1} \end{bmatrix},$$

where

$$P = \begin{bmatrix} X & Z^T \\ Z & V \end{bmatrix}.$$

Figure 1.2 shows the velocity variances diag(V) for a platoon of four vehicles with $W = I$, ordered by increasing (blue curves) and decreasing (green curves) vehicle

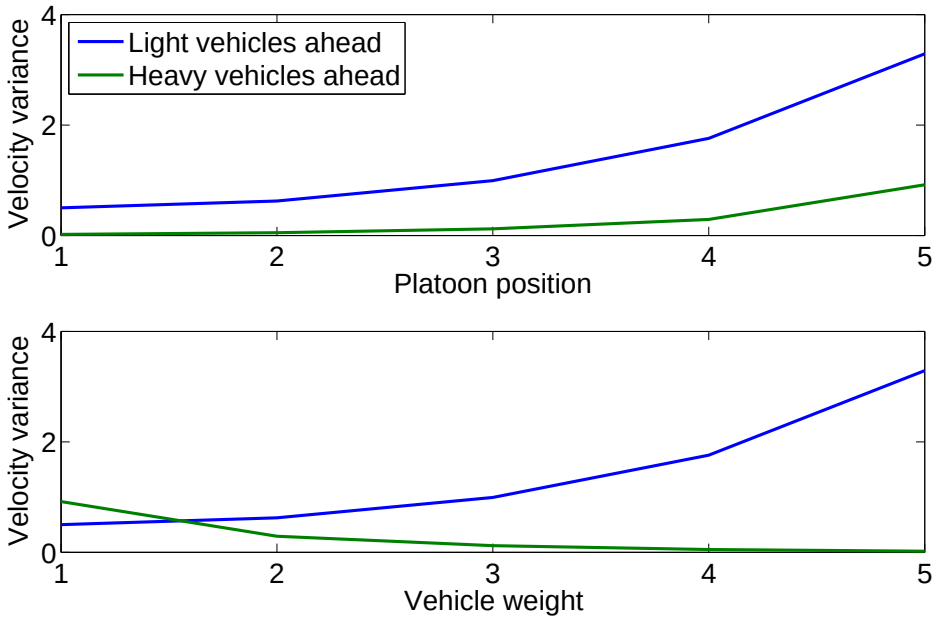

Figure 1.2: Vehicle velocity variances diag(V) arranged by platoon position (upper plot) and by vehicle weight (lower plot) with the blue curves depicting the case where the vehicles are arranged by increasing weight and the green curves by decreasing weight.

weights, respectively. As one can see clearly in the bottom picture, the average variance decreases drastically when heavy vehicles are placed ahead of lighter vehicles. This effect is caused by the variances being propagated downstream, where the base variances of heavier vehicles are lower due to their smaller susceptibility to the input noise $w(t)$. In this setup, only the lightest vehicle is slightly better off when placed ahead of the other vehicles, while the remaining variances increase significantly. This hints at the trend that less variable patients should be placed ahead of more variable patients when scheduling appointments to the infusion room, as is suggested in Rule 2.

The trend observed above is illustrated further in Figure 1.3, which compares the

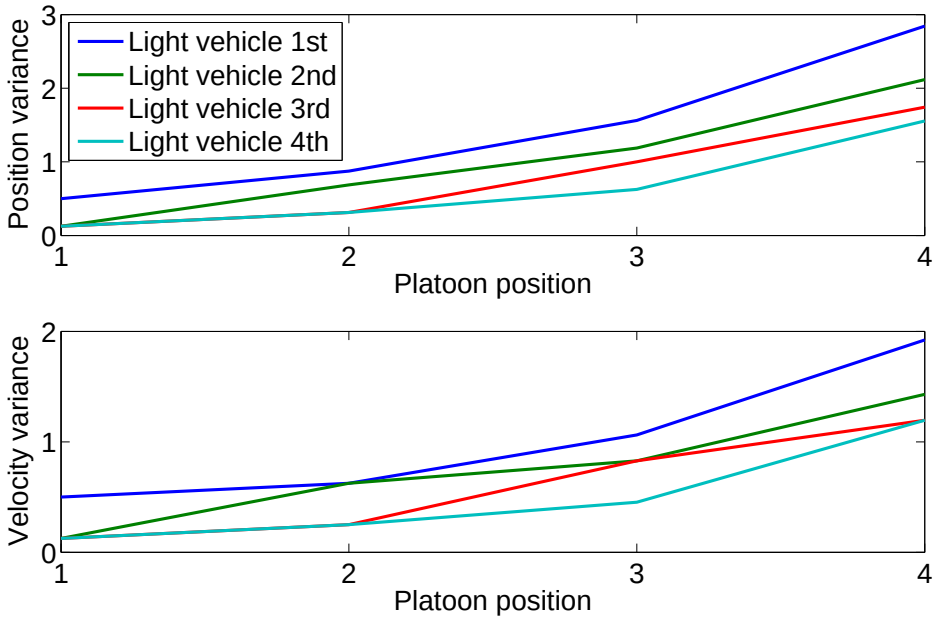

Figure 1.3: Vehicle position variances diag(X) and velocity variances diag(V) for platoons with a single light vehicle of mass 1 and three heavier vehicles with mass 2.

four possible relative orderings of four vehicles, one of which is lighter and thus more susceptible to its individual disturbance $w_i(t)$, where as in the previous example, we have $W = I$. Clearly, this setup corresponds to the initial example considered in this section, where four possible scheduling policies were compared with respect to the resulting workload variances. The four policies in Figure 1.3 are precisely those corresponding to (1.7) and Table 1.1. The figure confirms that the relative ordering with the highly variable patient at the end is preferable from an average variance standpoint, which is the motivation for Rule 2 in the scheduling problem.

While we were able to conclude Rule 2 with this second surrogate problem, we are unable to capture the benefits of Rule 1 in this setup. This issue is connected to string

instability effects in the vehicle platoon (see [16]), which dominate the behavior for larger strings of vehicles. A solution for practical implementation could be scheduling in multiple strings separated by a period of low workload arrival, for example during lunch break.

1.4 Conclusions

We introduced two surrogate problems that illustrate the benefits of acuity-based scheduling rules for visits to an infusion room. The first rule suggests distributing high-acuity patients evenly throughout given periods of time and can be analyzed using methods from calculus of variations applied to a stochastic mixing process under heavy traffic. The second rule requires scheduling high-acuity patients late in the day and is analyzed using parallels to strings of interconnected systems and a vehicle platoon in particular.

Clearly, all observations made regarding our two scheduling rules and in particular the corresponding surrogate problems extend beyond the infusion room, which served as the running example in this chapter. However, a number of extensions are of interest in particular for practical implementation of our scheduling rules. For instance, a tradeoff has to be made between violations of the two rules in the actual scheduling process. One could think of summarizing violations in each of the guidelines in penalty functions and summarizing those in a weighted cost functional evolving as the schedule is made based on patient requests. Moreover, not all patients may be compliant with

ideal schedule positioning for a multitude of reasons such as other appointments on the same day, such that their restrictions may have to be taken into account additionally. One could also think of extensions requiring further transient rules comparable to Rule 2, including variable numbers of nurses throughout the day, for example to include lunch times into the scheduling procedure. Another factor is that in a real workday, the fraction of patients from the high-acuity group may not be precisely μ but lower or higher, depending on daily fluctuations and the quality of the estimate μ for the actual patient population. Another direction for future work is the extension to results from queueing theory such as in [18] to the appointment scheduling problems occurring in healthcare.

Acknowledgements

Chapter 1, in full, is a reprint of the material as it appears in: M.A. Sehr, R.R. Bitmead, J. Fontanesi, "Multi-Class Appointments in Individualized Healthcare: Analysis for Scheduling Rules", *Proc. 14th European Control Conference*, 2015, pp 1219–1224. The dissertation author was the primary investigator and author of this paper.

Chapter 2

A Framework for Acuity-Based, Individualized Patient Scheduling

2.1 Introduction

Scheduling policies for patient visits at health centers are well known to have significant effects on various performance measures such as patient waiting times and idle times of medical personnel. The surveys [19] and [20] give a comprehensive overview of literature on medical appointment scheduling. In [19], references are grouped into analytic, simulation-based and case studies, where perhaps surprisingly, the majority of analytic approaches listed focus on scheduling patients drawn from homogeneous populations for visits to a clinic. That is, most analytic studies are aimed at determining desirable arrival times and inherent appointment lengths for consecutive service of patients with independent and identically distributed service times. Patient arrivals are

usually assumed to occur deterministically at scheduled times, in particular for elective patients, who do not require immediate medical intervention and can be scheduled for service well in advance. Analysis with respect to scheduling patients from heterogeneous populations is closely related to scheduling different surgical procedures in sequence. The survey paper [21] and the discussion in [22] give an overview of recent developments and open problems in surgical procedure scheduling.

In most common approaches to patient scheduling, individual patient characteristics apart from durations of certain procedures are not taken into account by the scheduler when organizing workload throughout a particular day. This scheduling approach often results in significant fluctuations when comparing waiting times, patient satisfaction and other performance measures over the course of multiple days. Causes for these discrepancies may be unavailability of information regarding individual patient characteristics in the scheduling procedure and unawareness of the appropriate scheduling action were it to be available. In many cases, knowledge of patient characteristics allowing further individualization can be made available rather easily given that doctors generally have a good idea of which patients are more uncertain in terms of fluctuations in punctuality and service requirements. We suggest the classification of patient populations into *acuity* classes, which can be performed reliably by the corresponding physician and then provided to the scheduler. High acuity could refer to high severity of prescribed treatment and usually equates with advanced stages of disease. From the perspective of the analysis in this chapter, high acuity is equated with high variability and vice versa. This is related to surgical procedure scheduling, where classification of

procedures into ones with low and high variances is discussed for instance in [13].

Strategies for the inclusion of patient acuity data into scheduling were proposed and studied in [23] for the special case of two acuity classes. The main results of this study are the following two qualitative ad hoc scheduling rules.

Rule 1: Avoid scheduling many high-acuity patients at once.

Rule 2: Schedule high-acuity patients later in the day.

Rule 1 aims to avoid periods of large accumulated uncertainty, which might cause overflowing waiting rooms and inherently long patient waiting times. Rule 2 hinders disturbances from being propagated through the workday and is closely related to the discussion in [13] and observations reported in [12]. The advantages of each rule were demonstrated in [23] using closely related surrogate problems. While these two scheduling rules require tradeoff decisions, they yield easily applicable heuristics for appointment scheduling.

This chapter aims to formulate a mathematical basis to address scheduling issues when dealing with multiple classes of individual patients which can be attributed with distinct service demand characteristics. While the analysis in Section 2.2 benefits the general understanding of scheduling systems with deterministic arrivals and independent Erlang and exponential service times, we use Section 2.3 to pose a Markov chain model that can be used in appointment systems with more general service time distributions. We anticipate models of the form motivated below to yield scheduling rules such as those reported in [23] when equipped with reasonable cost functions, which will be

part of a future study.

2.2 Sojourn Time Analysis for Erlang Service Time Distributions

Suppose we have a number of patients arriving for consecutive service at a clinic throughout a day. We assume patient arrivals to occur deterministically with scheduled interarrival times $T_k \geq 0$, where $T_k = 0$ for some k indicates multiple patients arriving simultaneously. Patient service times X_k are assumed independent and Erlang distributed [24] with constant rate parameter $\mu \in \mathbb{R}_+$ and varying shape parameters $d_k \in \mathbb{N}$ used to model individual patient acuities. Probability density and cumulative distribution functions of Erlang random variables with rate μ and shape d for $x \geq 0$ are

$$\gamma_{\mu,d}(x) = \frac{\mu^d x^{d-1} e^{-\mu x}}{(d-1)!}, \qquad \Gamma_{\mu,d}(x) = 1 - \sum_{n=0}^{d-1} \frac{(\mu x)^n}{n!} e^{-\mu x}. \tag{2.1}$$

The Erlang distribution is a special case of the Gamma distribution with $d \in \mathbb{N}$ instead of $d \in \mathbb{R}_+$. Moreover, there are two appealing links between the Erlang and exponential distributions. Firstly, the Erlang distribution subsumes the exponential distribution in the special case $d = 1$. Secondly, an Erlang random variable with shape d and rate μ is equal in distribution to the sum of d independent and exponentially distributed random variables with rate μ.

Using different rate and shape parameters, one can construct Erlang random variables with arbitrary positive means and variances, making them a powerful distribution

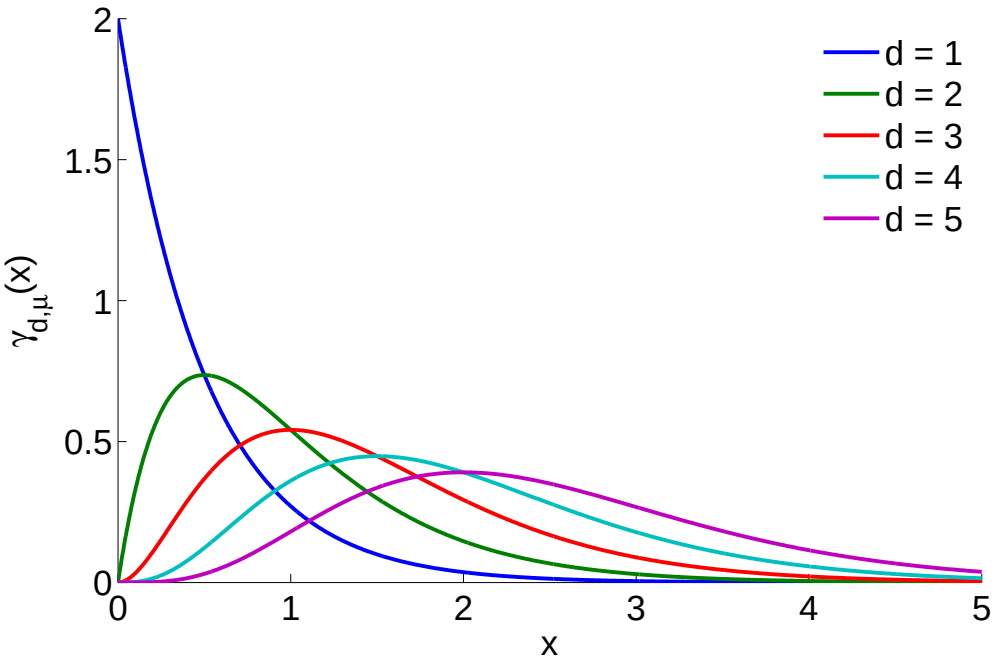

Figure 2.1: Probability density functions of Erlang random variables with varying shape parameters and rate parameter $\mu = 2$.

in modeling service requirements of patients with different acuity levels. This is further underlined by Figure 2.1, which shows probability density functions of Erlang random variables for a set of shape parameters and rate $\mu = 2$. Shapes $d > 1$ (i.e., those deviating from the exponential distribution) allow for service time distributions that seem reasonable in describing patient time demands modulo the unbounded tails.

In the following, let the *sojourn time* S_k denote the random total duration patient number k has to spend at the clinic. That is, S_k is the combined *waiting time* W_k and *service time* X_k encountered by the k^{th} patient arriving for service. Hinting at performance of particular scheduling rules, we are particularly interested in the cumulative

distribution functions $F_k(x) \triangleq P(S_k \leq x)$, mean values $\mathbb{E}[S_k]$ and variances $\mathrm{Var}(S_k)$ for all patients scheduled throughout the day.

2.2.1 Stationary Behavior for Exponential Service Times

First consider stationary behavior of the underlying appointment system in the special case $d = 1$ (i.e., exponentially distributed service times) with constant rate $\mu > 0$ and interarrival times T. This particular case is interesting in that it provides a link to classic queueing and scheduling literature, where such D/M/1 queues are frequently used. Moreover, results about long-run stationary behavior can be used to approximate sections of schedules filled with small numbers of patients. The remainder of this subsection is based on the following set of assumptions.

A.I: Patient arrivals occur deterministically with constant interarrival time $T > 0$.

A.II: Patient service times X_k are independent and exponentially distributed with rate parameter $\mu > 0$.

A.III: T and μ satisfy the stability condition $\mu T > 1$.

Notice that in particular Assumption A.A.III: is often undesirable in scheduling as it goes in hand with an increased fraction of idle time for medical personnel. Given Assumptions A.I:-A.III:, the sojourn times have a stationary distribution

$$F_\infty(x) \triangleq P(S_\infty \leq x), \qquad \text{where} \qquad S_\infty \triangleq \lim_{k\to\infty} S_k\,.$$

The following result characterizes the stationary distribution of the closely related waiting times W_k.

Theorem 2.1 (Jansson, 1966 [25])**.** *Suppose Assumptions A.A.I:-A.III: hold and let* $W_\infty \triangleq \lim_{k\to\infty} W_k$. *The stationary distribution of the waiting times* W_∞ *is then given by*

$$P(W_\infty = 0) = 1 - \delta_0,$$

$$P(0 < W_\infty \leq x) = \delta_0(1 - e^{-\mu x(1-\delta_0)})$$

for $x \geq 0$*, where* δ_0 *is the root of the equation*

$$\delta = e^{-\mu T(1-\delta)} \tag{2.2}$$

which has the smallest absolute value.

Stationary behavior of the sojourn times S_k under the stable D/M/1 queueing setup with $\mu T > 1$ follows directly from the above result.

Corollary 2.1. *Under Assumptions A.A.I:-A.III:, the stationary cumulative distribution function of the sojourn times* S_∞ *is*

$$F_\infty(x) = 1 - e^{-\mu x(1-\delta_0)}$$

for $x \geq 0$*, where* δ_0 *is the root of* (2.2) *with smallest absolute value. Moreover, we have*

$$\mathbb{E}[S_\infty] = \frac{1}{\mu(1-\delta_0)}, \quad \mathrm{Var}(S_\infty) = \frac{1}{\mu^2(1-\delta_0)^2}.$$

Proof. The sojourn time of patient k is $S_k = W_k + X_k$, where W_k and X_k are the waiting time and the exponentially distributed service time for patient k, respectively. By Theorem 2.2, we get the stationary distribution for the sojourn time using the convolution

$$\begin{aligned} P(S_\infty \leq x) &= \frac{d}{dx}\left(P(W_\infty \leq x) * P(X_k \leq x)\right) \\ &= \frac{d}{dx}\int_0^x (1-\delta_0 e^{-\mu z(1-\delta_0)})(1-e^{-\mu(x-z)})\,dz \\ &= 1 - e^{-\mu x(1-\delta_0)}. \end{aligned}$$

Mean value $\mathbb{E}[S_\infty]$ and variance $\mathrm{Var}(S_\infty)$ follow directly. □

2.2.2 Transient Behavior for Erlang Service Times

So far, we have discussed stationary behavior in the special case of exponential service distributions under the stability condition $\mu T > 1$. However, in a scheduling setup with limited numbers of patients over a finite time horizon, we are usually more interested in the unstable case $\mu T \leq 1$, corresponding to lower idling times of medical personnel at the cost of patient waiting times potentially building up faster throughout the day. In this case, stationary distributions do not exist. Moreover, given the explicit transient nature of the scheduling problem, we are particularly interested in distribution functions $F_k(x)$ for rather small integers $k \in \mathbb{N}$. For the remainder of this discussion about transient behavior of the scheduled queueing system with Erlang service demands, we make the following assumptions.

B.I: Patient arrivals occur deterministically with interarrival times $T_k \geq 0$, where T_k is the time between consecutive appointments for patients k and $k+1$.

B.II: Patient service times X_k are independent and Erlang distributed with rate $\mu > 0$ and shapes $d_k \in \mathbb{N}$.

Notice that Assumptions A.A.I:-A.II: in the previous section are a special case of Assumptions B.B.I:-B.II: with $d_k = 1$ and $T_k = T$ for $k \in \mathbb{N}$. However, Assumptions B.B.I:-B.II: allow for a substantially larger set of possible scheduling scenarios to be analyzed. For instance, $T_k = 0$ for some $k \in \mathbb{N}$ could be imposed to account for multiple patients arriving at once.

In the following, we replace the analysis of Erlang distributed patient service times with shapes d_k by an equivalent non-stationary problem in which each patient is decomposed into a sum of d_k individual *sub-patients* arriving simultaneously and each of whom has service time which is exponentially distributed with rate μ. The sequence $\{\tilde{T}_j\}$ is the sequence of interarrival times of these collections of sub-patients.

$$\tilde{T}_j \triangleq \sum_{i=1}^{\infty} \mathbb{1}\{j = \tau_i\} T_i, \quad \text{where} \quad \tau_i \triangleq \sum_{k=1}^{i} d_k \tag{2.3}$$

and $\mathbb{1}\{A\}$ denotes the indicator function of event A. The interarrival time between sub-patients n and j is then

$$\tilde{T}_n^j \triangleq \sum_{i=n}^{j} \tilde{T}_i.$$

This problem with exponentially distributed sub-patients is equivalent to the original problem with Erlang distributed patients and allows application of the memoryless property of the exponential distribution. The transient behavior of the scheduled queueing model given Assumptions B.B.I:-B.II: is then described by the following result for the sojourn times S_k.

Theorem 2.2. *Suppose Assumptions B.B.I:-B.II: hold. The transient cumulative distribution functions $F_k(x)$ of the sojourn times S_k are then given by*

$$F_k(x) = \tilde{F}_{\tau_k}(x), \tag{2.4}$$

where

$$\tilde{F}_j(x) \triangleq 1 - \left(1 + \sum_{n=1}^{j-1} \left(1 - \tilde{F}_{j-n}\left(\tilde{T}_{j-n}^{j-1}\right)\right) \frac{(\mu x)^n}{n!}\right) e^{-\mu x} \tag{2.5}$$

for $x \geq 0$ and $j \in \mathbb{N}$.

Proof. Following the above discussion, we can equivalently analyze a problem with interarrival times (2.3) and exponentially distributed service times with rate μ. In the following, let $\tilde{S}_j$ and $\tilde{F}_j(x)$ denote the sojourn times and their cumulative distribution functions in this surrogate problem. Reapplying the correspondence between exponential and Erlang distributions (2.1) in combination with the memoryless property of the exponential distribution yields

$$\begin{aligned}\tilde{F}_j(x) = {} & P\left(\tilde{S}_{j-1} \leq \tilde{T}_{j-1}^{j-1}\right) \Gamma_{\mu,1}(x) \\ & + \sum_{n=2}^{j-1} P\left(\tilde{S}_n > \tilde{T}_n^{j-1}, \tilde{S}_{n-1} \leq \tilde{T}_{n-1}^{j-1}\right) \Gamma_{\mu,j+1-n}(x) + P\left(\tilde{S}_1 > \tilde{T}_1^{j-1}\right) \Gamma_{\mu,j}(x). \end{aligned} \tag{2.6}$$

That is, $\tilde{F}_j(x)$ is a convex combination of cumulative distribution functions for Erlang random variables with increasing shape parameters but equal rate parameter μ. The coefficients in this convex combination for $\tilde{F}_j(x)$ depend on $\{\tilde{F}_1(x), \ldots, \tilde{F}_{j-1}(x)\}$. Now notice that by (2.1) the cumulative distribution functions $\Gamma_{\mu,j}(x)$ of Erlang distributed random variables satisfy the recursion

$$\Gamma_{\mu,1}(x) = 1 - e^{-\mu x}, \qquad \Gamma_{\mu,j+1}(x) = \Gamma_{\mu,j}(x) - \frac{(\mu x)^j}{j!} e^{-\mu x}$$

for $j \in \mathbb{N}$. Hence, the cumulative distribution function of the sojourn time for patient j in the surrogate problem with exponential service demands can be written

$$\tilde{F}_j(x) = 1 - e^{-\mu x} - \sum_{n=1}^{j-1} c_n \frac{(\mu x)^n}{n!} e^{-\mu x},$$

where the coefficients c_n follow from (2.6) as

$$\begin{aligned}
c_1 &= 1 - P\left(\tilde{S}_{j-1} \leq \tilde{T}_{j-1}^{j-1}\right) \\
&= 1 - \tilde{F}_{j-1}\left(\tilde{T}_{j-1}^{j-1}\right), \\
c_2 &= c_1 - P\left(\tilde{S}_{j-1} > \tilde{T}_{j-1}^{j-1}, \tilde{S}_{j-2} \leq \tilde{T}_{j-2}^{j-1}\right) \\
&= 1 - P\left(\tilde{S}_{j-2} \leq \tilde{T}_{j-2}^{j-1}\right) \\
&= 1 - \tilde{F}_{j-2}\left(\tilde{T}_{j-2}^{j-1}\right), \\
c_3 &= 1 - \tilde{F}_{j-3}\left(\tilde{T}_{j-3}^{j-1}\right), \\
&\vdots \\
c_{j-1} &= 1 - \tilde{F}_1\left(\tilde{T}_1^{j-1}\right).
\end{aligned}$$

□

As demonstrated above, Theorem 2.2 allows for two distinct interpretations of the scheduled Erlang queueing process. Firstly, we can interpret the process as patients with Erlang service distributions of equal rate μ but potentially varying shapes d_k arriving at deterministic times for consecutive service. This interpretation allows for analysis of a multi-class scheduled queue. Secondly, using the correspondence between Erlang distributed random variables and sums of i.i.d. exponentials with equal rate μ, we can interpret the process as numbers d_k of sub-patients with independent exponential service

demands arriving simultaneously in each time period. This duality allows for a variety of scheduling problems to be analyzed using Theorem 2.2. This is further demonstrated in the following example.

Example 2.1. Consider scheduling four patients for successive service at a clinic. Of the four patients, one is classified as a high-acuity patient and all others as low-acuity patients, respectively. In addition to the four possible orders in which the patients can be scheduled, we have to choose a sequence of interarrival times. In this example, we use mean service times of the patients to determine desirable interarrival times. Using Erlang rate parameter $\mu = 2$ for all patients, shape parameter $d_h = 3$ for the high-acuity patient and shape parameter $d_l = 2$ for the low-acuity patients, mean service times are

$$\mathbb{E}[X_l] = \frac{d_l}{\mu} = 1 \qquad \text{and} \qquad \mathbb{E}[X_h] = \frac{d_h}{\mu} = 3/2.$$

Now suppose we choose interarrival times to match these mean service times multiplied by a constant scaling factor $\alpha \geq 0$. That is, patient $k+1$ is scheduled to arrive

$$T_k = \alpha\, \mathbb{E}[X_k]$$

later than patient k, where $\mathbb{E}[X_k]$ is the expected service duration for patient k. Using this scheduling approach leaves us with choice of α and the order in which patients are scheduled to arrive. Mean sojourn times corresponding to $\alpha = 0.75$ (left) and $\alpha = 1.0$ (right) are displayed for the patient orders in Figure 2.2. Based on these results, we can make a number of observations about the scheduled queues. For instance, we see that for either choice of α, Rule 2 from [23] transpires. That is, all patients except for the high-acuity patient are better off provided the high-acuity patient is scheduled later in the day.

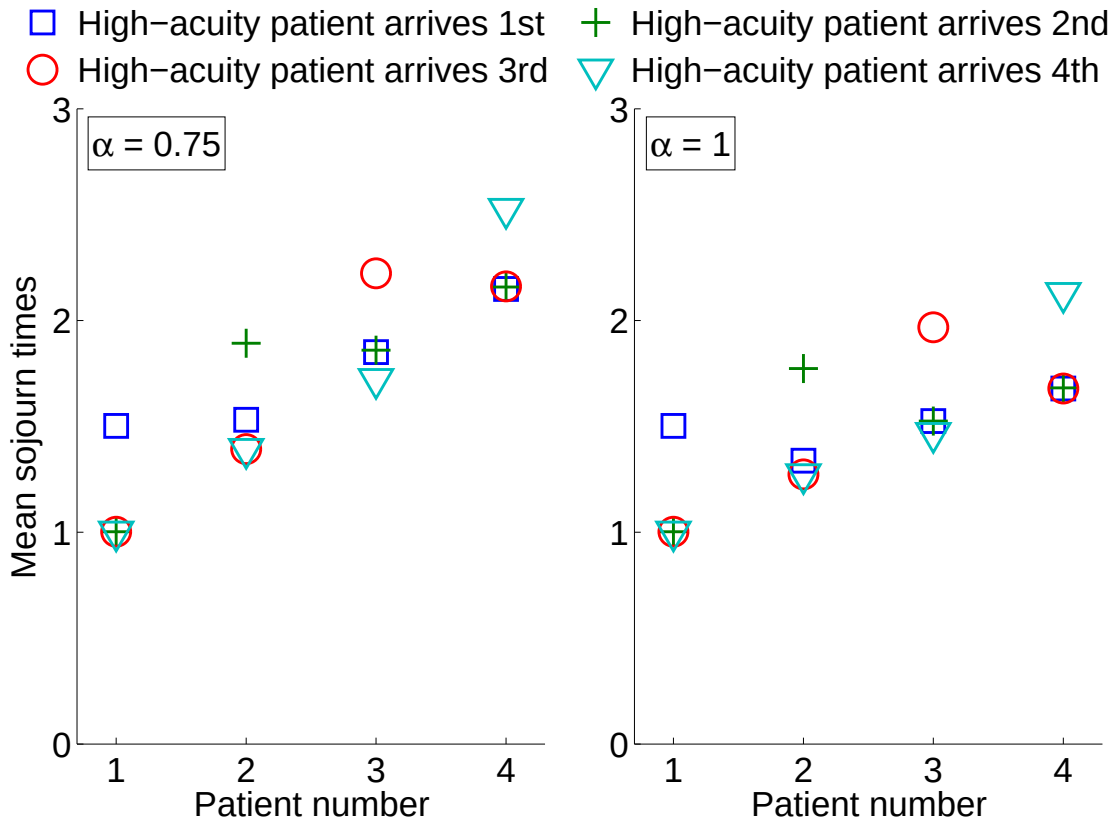

Figure 2.2: Mean sojourn times for different patient orders and interarrival time scaling factors α.

This property is most apparent for small scaling factors α and disappears gradually as α grows. Moreover, we can see how mean sojourn times build up throughout the schedule, with a large increase where the high-acuity patient is placed. For larger α, the sojourn times recover slightly after the high-acuity patient.

The following result is an immediate consequence of Theorem 2.2, and (2.5) in particular, when dealing with a setup as in Corollary 2.1. However, the stability condition $\mu T > 1$ is not required at this point.

Corollary 2.2. *Under Assumptions A.A.I:-A.II:, the cumulative distribution functions* $F_k(x)$ *of the sojourn times* S_k *are*

$$F_k(x) = 1 - \left(1 + \sum_{n=1}^{k-1}(1 - F_{k-n}(nT))\frac{(\mu x)^n}{n!}\right)e^{-\mu x}$$

for $x \geq 0$ *and integers* $k \in \mathbb{N}$*. The corresponding transient means and variances of the sojourn times are*

$$\mathbb{E}[S_k] = \frac{1}{\mu}\left(k - \sum_{n=1}^{k-1} F_{k-n}(nT)\right),$$
$$\mathrm{Var}(S_k) = \frac{1}{\mu^2}\left(k + 2\sum_{n=1}^{k-1}(k-n-1)F_{k-n}(nT) - \left(\sum_{n=1}^{k-1} F_{k-n}(nT)\right)^2\right).$$

Example 2.2. To illustrate Corollaries 2.1 and 2.2 in the special case of independent exponential service time distributions, consider the following example with fixed service rate $\mu = 4/3$, corresponding to an average service duration of $3/4$ and stable behavior of the queue given interarrival times $T > 3/4$. Figure 2.3 shows cumulative distribution functions of the sojourn times for patients $k = 1, 5, 10$ and the corresponding stationary distribution in the stable case with $T = 1$. As we can see, the distributions $F_k(x)$ strongly resemble the stationary distribution.

A central point unaddressed so far is the behavior of the scheduled queueing system when varying the rate parameter μ of the underlying Erlang service time distributions. The following result summarizes a scaling property of the cumulative distribution functions (2.4) for different rate parameters.

Theorem 2.3. *Suppose Assumptions B.B.I:-B.II: hold for a rate parameter* μ_1 *and a*

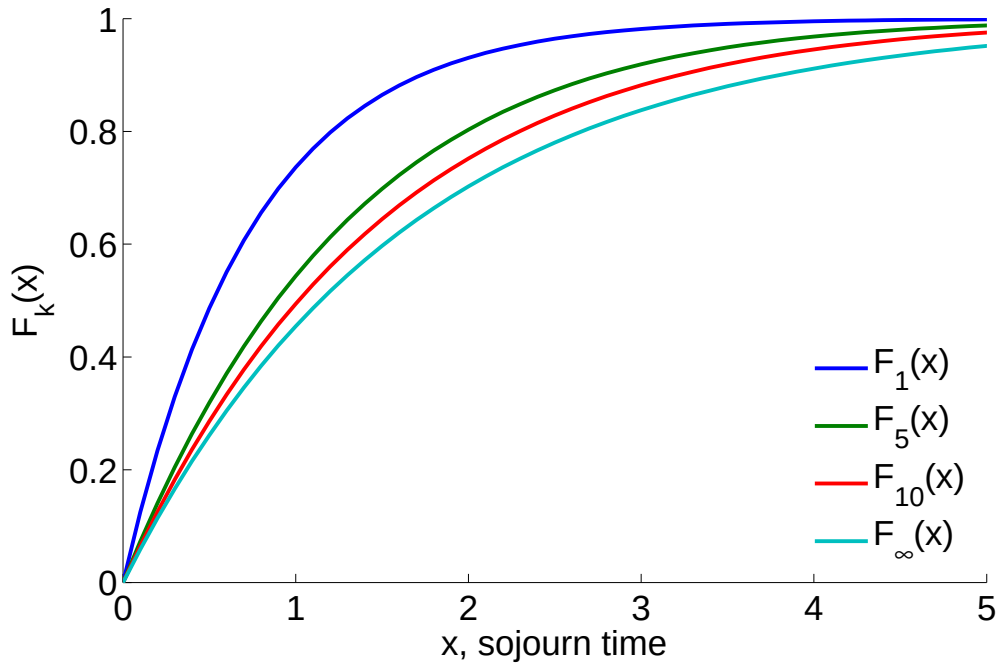

Figure 2.3: Transient and steady-state distributions of sojourn times S_k for stable exponential configuration $\mu = 4/3$ and $T = 1$.

sequence of interarrival times $T_{k,1}$. Moreover, let $\beta > 0$ and

$$\mu_2 = \mu_1/\beta, \qquad T_{k,2} = \beta\, T_{k,1}, \qquad d_k = d_{k,2} = d_{k,1}$$

for $k \in \mathbb{N}$. If $F_{k,1}(x)$ denotes the cumulative distribution function of the sojourn time for patient k in the scheduled queue with interarrival times $T_{k,1}$ and rate parameter μ_1 and $F_{k,2}(x)$ its counterpart with interarrival times $T_{k,2}$ and rate parameter μ_2, then

$$F_{k,1}(x) = F_{k,2}(\beta x) \tag{2.7}$$

for $x \geq 0$ and $k \in \mathbb{N}$.

Proof. Recall that in the proof of Theorem 2.2, we defined a surrogate problem with exponential service times. If suffices to verify (2.7) for this equivalent problem which

has cumulative distribution functions $\tilde{F}_j$ as in (2.5). First notice that

$$\tilde{F}_{1,2}(\beta x) = \Gamma_{\mu_2,1}(\beta x) = 1 - e^{-\mu_2 \beta x} = \Gamma_{\mu_1,1}(x) = \tilde{F}_{1,1}(x),$$

verifying (2.7) in the auxiliary problem for $j = 1$. Now suppose (2.7) holds for all positive integers $j \leq m - 1$ in the auxiliary problem. We then have

$$\tilde{F}_{m,2}(\beta x) = 1 - \left(1 + \sum_{n=1}^{m-1} \left(1 - \tilde{F}_{m-n,2}\left(\tilde{T}_{n,2}^{m-1}\right)\right) \frac{(\mu_1 x)^n}{n!}\right) e^{-\mu_1 x},$$

where

$$\tilde{F}_{m-n,2}\left(\tilde{T}_{n,2}^{m-1}\right) = \tilde{F}_{m-n,2}\left(\beta\, \tilde{T}_{n,1}^{m-1}\right) = \tilde{F}_{m-n,1}\left(\tilde{T}_{n,1}^{m-1}\right)$$

for $n \in \{1, 2, \ldots, m-1\}$, which completes the proof. □

The next result is a consequence of Theorem 2.3 for means and variances of the sojourn times $S_{k,1}$ and $S_{k,2}$.

Corollary 2.3. *Suppose the assumptions in Theorem 2.3 hold. Then*

$$\mathbb{E}[S_{k,2}] = \beta\, \mathbb{E}[S_{k,1}], \qquad \mathrm{Var}(S_{k,2}) = \beta^2\, \mathrm{Var}(S_{k,1})$$

for $k \in \mathbb{N}$.

In this section, we have reported a number of results for transient behavior of sojourn times in scheduled queues with Erlang service distributions, which can be used to model either patients with varying service demand characteristics or patient groups of varying size but equal service characteristics arriving at a clinic. The results in this section allow for analysis of a wide range of scheduling problems. We can also use the

results as a first approximation to the optimization problems of interest when searching for optimal scheduling policies, where especially Theorem 2.3 gives some insight into choosing relative appointment lengths if patients with distinct characteristics are to be scheduled. However, the iterative form of the results reported in this section is not very suitable for computing optimal policies with multiple acuity classes, where both ideal arrival orders and arrival times need to be determined. An alternative model more suitable to finding optimal scheduling policies for general service time distributions is suggested below.

2.3 Markov Model for Appointment Scheduling

As discussed in the previous section, the iterative form of the result reported in Theorem 2.2 is more suitable for analysis than determination of desirable scheduling policies. To address this issue, we can approximate the behavior of our queueing system with scheduled arrivals from different patient classes using a finite number D of waiting time intervals of length Δ each. This time-discretization allows for approximate analysis using a simple Markov chain model to propagate the probabilities of patient k's waiting time being within one of a number of fixed time intervals. For the following discussion of this Markov chain model, we restrict ourselves to scheduling a total number $N+1 \in \mathbb{N}$ of patients for the workday of interest and impose the following assumptions.

C.I: There are D waiting time intervals of length Δ each, where D is a sufficiently large integer multiple of N.

C.II: Patient arrivals occur deterministically with interarrival times T_k, where T_k is an integer multiple of Δ and measures the time difference between appointments for patient k and patient $k+1$.

C.III: Patient service times X_k are independent random variables satisfying $P(X_k \geq T_k + \Delta D/N) = 0$. That is, the scheduled service period cannot be exceeded by $\Delta D/N$ or more.

C.IV: Regardless of when within a given time period patient k finishes service, patient $k+1$ is not served until the beginning of the following time period.

Notice that the last assumption above increases the waiting time of patient $k+1$ by some random value $X < \Delta$, resulting in a potentially conservative approximation of the waiting times for larger time increments Δ and higher patient numbers N. However, the assumption is reasonable provided time increments and schedule lengths are sufficiently short. Based on Assumptions C.C.I:-C.IV:, we can cast a Markov chain model with state

$$W_k \triangleq \begin{bmatrix} W_{0,k} \\ W_{1,k} \\ W_{2,k} \\ \vdots \\ W_{D,k} \end{bmatrix} = \begin{bmatrix} P(k \text{ starts on time}) \\ P(k \text{ starts } \Delta \text{ late}) \\ P(k \text{ starts } 2\Delta \text{ late}) \\ \vdots \\ P(k \text{ starts } D\Delta \text{ late}) \end{bmatrix},$$

initial conditions $W_{0,0} = 1$ and $W_{i,0} = 0$ for $i \in \{1, 2, \ldots, D\}$ and $W_{k+1}^T = W_k^T P_k$ for a potentially time-varying state transition matrix with elements

$$(P_{i,j})_k = P(W_{j,k+1} = 1 \mid W_{i,k} = 1).$$

Notice at this point that if we allow for a countable state space, Assumption C.C.III: can be dropped. However, having an upper bound on patient service times does in fact coincide with actual scheduling applications. Notice also that given the total number of patients $N + 1$ and the correspondence of the time index k in this Markov model with the successive patient index, it only makes sense to evolve the state W_k through N time steps. That is, we initialize the model at $W_{\cdot,0}$ and terminate the state evolution at $W_{\cdot,N}$. The elements of the transition matrices P_k follow as

$$(P_{i,j})_k = P\left(X_k \in [T_k + (j - i - 1)\Delta, T_k + (j - i)\Delta)\right).$$

While the last D/N rows of the state transition matrix do not sum up to one, we only evolve the state over N iterations from the given initial conditions, such that these elements of the state transition matrix never come into play. The following example illustrates how the Markov model based on Assumptions C.C.I:-C.IV: can be used to approximate the behavior observed in Example 2.2 above.

Example 2.3. Consider the following setup in the spirit of Example 2.2. To approximate the queueing process from the previous section, we choose interarrival times $T_k = T$ and let the service times X_k be independent and identically distributed truncated exponential random variables with support on $(0, b\,]$, such that the cumulative distribution function for each service time is

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{1-e^{-\mu b}}\left(1 - e^{-\mu x}\right), & x \in (0, b] \\ 1, & x > b \end{cases} .$$

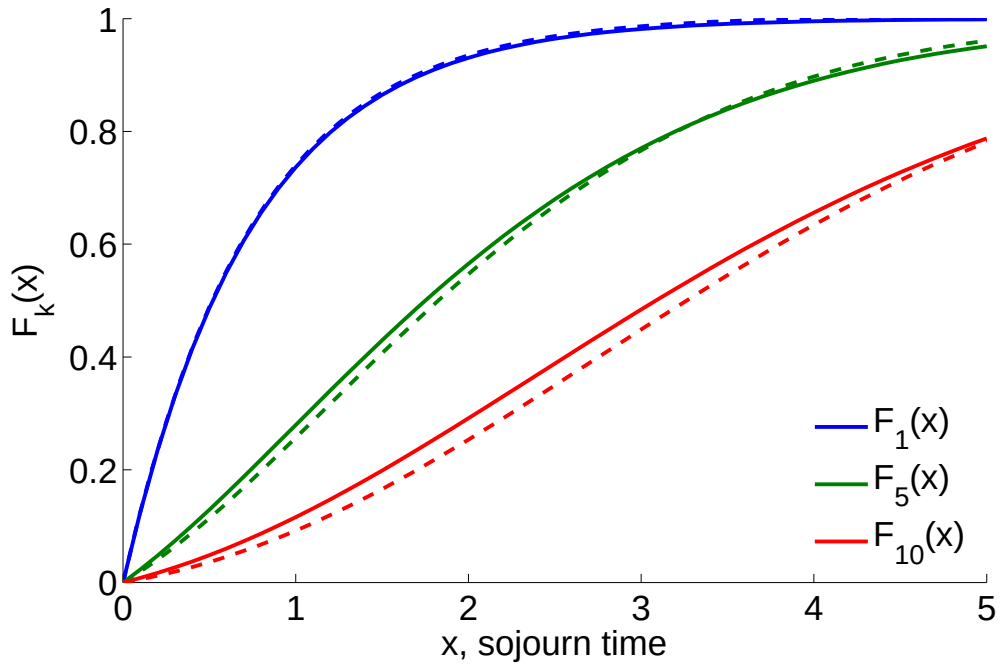

Figure 2.4: Transient distributions of sojourn times S_k for unstable exponential configuration $\mu = 4/3$ and $T = 1/2$. Solid lines for exact distributions via Corollary 2.2; dashed lines for approximation using Markov model with $\Delta = T/6 = 1/12$ and truncated exponential service times with $b = 4$.

Choosing the parameters $\mu = 4/3$ and $T = 1/2$, we can approximate the behavior of our scheduled queueing system with $N = 9$ and $D = N(b - T)/\Delta$ by evolving the state N times and taking convolutions with the distribution of the truncated exponential service times X_k to obtain distributions of the sojourn times per patient. For $b = 4$ and $\Delta = T/6 = 1/12$, we obtain the cumulative distribution functions displayed in Figure 2.4, which closely resemble those in the exact case computed using Corollary 2.2.

2.4 Conclusions

We have reported a number of results regarding analysis of scheduled queueing systems, where arrivals occur deterministically and service times are random. In particular, we have focused on setups allowing for patients with a number of different service distributions within the same queue. The significance of this work is that it provides a mathematical basis for the analysis of acuity-based scheduling rules applied in medicine. This provides a foundation from which to develop new scheduling rules in individualized medicine.

Acknowledgements

Chapter 2, in full, is a reprint of the material as it appears in: M.A. Sehr, R.R. Bitmead, J. Fontanesi, "A Framework for Acuity-Based, Individualized Patient Scheduling", *Proc. 54th IEEE Conference on Decision and Control*, 2015, pp 1858–1863. The dissertation author was the primary investigator and author of this paper.

Part II

Stochastic Model Predictive Control

Chapter 3

Sumptus Cohiberi: The Cost of Constraints in MPC with State Estimates

3.1 Introduction

Model Predictive Control (MPC) invites the practitioner to insert constraints, a natural technique for engineering design. In this chapter we offer a cautionary analysis of the potential performance cost of active constraints when state estimates are used in control implementation; the active constraints can amplify the extant state estimate error in the control signal and inject this into the closed-loop system affecting achieved performance. MPC is posed as full-state feedback control even though its application invariably uses erroneous state estimates. The performance of MPC controlled systems

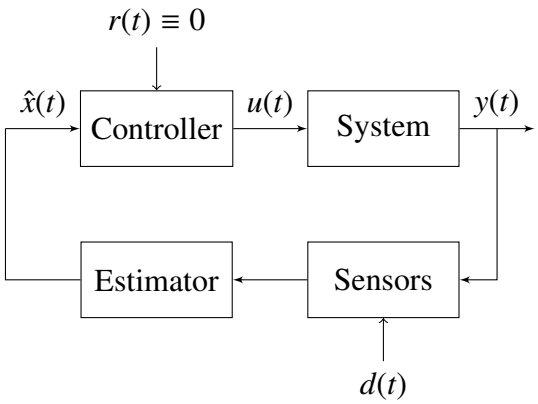

Figure 3.1: Closed-loop system architecture for state-estimate feedback control based on perturbed output measurements.

has been addressed by only a few, including [26, 27, 28], and similarly the effect of state estimation errors has been largely eschewed or circumvented [2]. We analyze how the number of constraints influences sensitivity of the control inputs to estimation errors for general closed-loop architectures such as displayed in Figure 3.1.

In [4], it is explored how stochastic output-feedback MPC can be tackled using Gaussian assumptions on the state-estimates. The stochastic problem is converted to a deterministic counterpart with tightened constraints to ensure all constraints can be satisfied with sufficiently large probability. The procedure prescribed leads to controllers depending explicitly on state-estimate covariance data. Robust output feedback MPC is explored in [29], where a simple Luenberger state observer is paired with tube-based MPC. Using tightened constraints for an invariant bound on the estimation error, the authors establish robust exponential closed-loop stability of an invariant set. The results are extended in [7] to drop steady-state assumptions on the underlying state observer. More recently, [6] presents an approach to nonlinear output feedback MPC based on Moving

Horizon Estimation (MHE, [30]), where state-estimates are based on a fixed number of past measurements subject to state-estimate constraints. Linear output-feedback MPC via MHE is discussed for instance in [5]. In [31], robust output feedback MPC for systems with unstructured model uncertainty is considered. Based on a linear state estimator and quadratic cost, the authors propose Linear Matrix Inequality (LMI) conditions involving the cost parameters to guarantee robust closed-loop stability.

To analyze how the interplay of estimation errors and constraints can impact closed-loop MPC behavior, we make use of several results from the explicit MPC literature. In case system dynamics, constraints and value function are time-invariant, explicit solutions are appealing for problems of sufficiently small size. That is, given that the optimal control inputs for this setup depend only on the current state, there is a static mapping from feasible states to feasible controls summarizing all solutions of the MPC problem. This parametrization of the MPC optimization problem by feasible states leads to multiparametric programs. In [32, 33], the explicit MPC problem was famously solved for linear time-invariant systems with quadratic cost and polytopic constraints. The optimal control inputs are piecewise affine over polytopic regions, where each polytope corresponds to a set of active constraints at the optimal solution. This allows reduction of online quadratic programming to evaluation of a piecewise affine function. Parallel results for 1- and ∞-norms in the cost function have been presented in [34], where on-line solution of linear programs reduces to evaluation of a static piecewise affine function as in [33]. An overview of explicit MPC and multiparametric programming is given in the surveys [35, 36]. In practice, computing explicit MPC solutions for

most problems still is an exceedingly demanding task computationally, making such solutions viable only for low-dimensional problems. Applicability can be extended based on approximate methods.

3.2 Linear Quadratic Model Predictive Control

3.2.1 Motivation and Problem Setup

We consider model predictive control based on *noisy* state-estimate feedback using an implicit state estimator which is not further specified for the remainder of this chapter – it could, for instance, be a Kalman filter. We examine interactions between state estimation errors and active constraints, leading to perturbed control inputs generated by the MPC solutions. We consider time-invariant systems mostly to simplify notation in the technical discussion. However, the observations generally hold for time-varying systems, constraints and cost functions. At time instant t, a linear system is modeled by linear difference equation

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

where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$. Within the MPC optimization problem, future states are propagated as

$$\begin{aligned} \hat{x}_{k+1}(t) &= A\hat{x}_k(t) + Bu_k(t), \\ \hat{x}_0(t) &= x(t) - \tilde{x}(t), \end{aligned} \tag{3.2}$$

where the signal $\tilde{x}(t)$ is the state estimation error and $u_k(t)$ denotes the prospective MPC control action k time instants along the control horizon. As indicated above, the current state estimate $\hat{x}_0(t)$ is generated by an arbitrary state estimator, e.g. a Kalman Filter. Below, we omit dependencies on time t when possible to simplify notation. Further denoting

$$\mathbf{u} = \begin{bmatrix} u_0 & u_1 & \cdots & u_{N-1} \end{bmatrix}^T,$$

we aim to describe perturbations of the optimal receding horizon solution arising from the MPC optimization problem

$$\begin{aligned} \min_{\mathbf{u}} & \left\{ \hat{x}_N^T P \hat{x}_N + \sum_{k=0}^{N-1} \hat{x}_k^T Q \hat{x}_k + u_k^T R u_k \right\}, \\ \text{s.t.} \quad & C_u u_k \leq d_u, && k = 0, \ldots, N-1; \\ & C_x \hat{x}_k \leq d_x, && k = 1, \ldots, N-1; \\ & C_f \hat{x}_k \leq d_f, && k = N; \\ & \hat{x}_{k+1} = A\hat{x}_k + Bu_k, && k = 0, \ldots, N-1; \\ & \hat{x}_0 = x(t) - \tilde{x}(t). \end{aligned} \tag{3.3}$$

The first element of this optimal solution is the control input at time instant t, i.e.

$$u(t) = u_0^*(t).$$

The MPC optimization problem (3.3) at time t can equivalently be posed as a multi-parametric quadratic program (mpQP) dependent on $\hat{x}_0$. The question analyzed in this chapter is how the control inputs of such state estimate based MPC schemes differ from

the control inputs generated by problem (3.3) were the real state to be known precisely, i.e. $\hat{x}_0 = x(t)$.

3.2.2 Perturbed mpQP Formulation

In [33], it was shown that the optimal control input $u^*(t)$ arising from the MPC problem (3.3) is piecewise affine over a finite polyhedral partition, $\{\mathrm{R}_i\}$, of the set of states for which feasible solutions exist. Each polytope in this partition is characterized by a particular set of active constraints. Transitioning a boundary between adjacent polytopes results in additional constraints becoming active or some of the previously active constraints becoming inactive. These observations allow us to reduce the multiparametric quadratic program arising from (3.3) to an equality-constrained multiparametric program with identical cost but a reduced set of constraints corresponding to the constraints active in the polytopic region enclosing $\hat{x}_0$. The equivalent equality-constrained mpQP with state estimate confined to the i^{th} polytope, i.e. $\hat{x}(t) \in \mathrm{R}_i$, can be written in the form

$$\begin{aligned} \mathrm{QP}_i(\hat{x}(t)): \ \min_{\mathbf{u}} \left\{ \frac{1}{2}\mathbf{u}^T\mathbf{H}\mathbf{u} + \mathbf{f}^T\mathbf{u} + \mathbf{c} \right\}, \\ \text{s.t.} \quad \mathbf{C}_i\mathbf{u} - \mathbf{d}_i = 0. \end{aligned} \tag{3.4}$$

The parameters defining the cost in this problem are

$$\begin{aligned} \mathbf{H} &= 2\left(\mathcal{B}^T \operatorname{diag}\left(I_{N-1} \otimes Q, P\right)\mathcal{B} + I_N \otimes R\right), \\ \mathbf{f} &= 2\mathcal{B}^T \operatorname{diag}\left(I_{N-1} \otimes Q, P\right)\mathcal{A}\hat{x}_0, \\ \mathbf{c} &= \hat{x}_0^T\left(Q + \mathcal{A}^T \operatorname{diag}\left(I_{N-1} \otimes Q, P\right)\mathcal{A}\right)\hat{x}_0, \end{aligned} \tag{3.5}$$

where I_N is the N-dimensional identity matrix, $\text{diag}(X, Y)$ denotes the block-diagonal composition of matrices X and Y, $(X \otimes Y)$ is the Kronecker product of X and Y, and

$$\mathcal{A} = \begin{bmatrix} A \\ \vdots \\ A^N \end{bmatrix}, \qquad \mathcal{B} = \begin{bmatrix} B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix},$$

such that

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 & \hat{x}_2 & \cdots & \hat{x}_N \end{bmatrix}^T = \mathcal{A}\hat{x}_0 + \mathcal{B}\mathbf{u}.$$

Notice that only the linear term $\mathbf{f}$ and constant term $\mathbf{c}$ of the mpQP cost depend on the state estimate $\hat{x}_0$. While perturbations in the quadratically dependent constant term $\mathbf{c}$ do not alter the optimal solution $\mathbf{u}^*$ of (3.4), disturbances in the linearly dependent linear term $\mathbf{f}$ cause translational shifts of the level sets associated with the mpQP cost. This causes explicit dependency of $\mathbf{u}^*$ and thus $u(t)$ on the estimation error $\tilde{x}(t)$. The constant matrix describing the quadratic terms in (3.4) satisfies $\mathbf{H} \succ 0$ under standard assumptions on the MPC problem (3.3), implying convexity of the mpQP (3.4) for all $\hat{x}_0 \in \mathrm{R}_i$. In fact, we know from [33] that the problem is convex for all $\hat{x}_0$ in the feasible region. The equality constraints in problem (3.4) are parametrized by the quantities

$$\begin{aligned} \mathbf{C}_i &= C_{x,i}\mathcal{B} + C_{u,i}, \\ \mathbf{d}_i &= d_i - C_{x,i}\mathcal{A}\hat{x}_0, \end{aligned} \tag{3.6}$$

where the matrices $C_{x,i}, C_{u,i}$ and vectors d_i depend on which sets of constraints in (3.3) are active at the solution for $\hat{x}_0$ throughout the control horizon. That is, the matrices

$\mathbf{C}_i$ remain constant and the vectors $\mathbf{d}_i$ depend linearly on the state estimate $\hat{x}_0$ for all $\hat{x}_0 \in \mathrm{R}_i$. We assume without loss of generality that $\mathbf{C}_i$ has linearly independent rows, as otherwise the pair $\begin{bmatrix} \mathbf{C}_i & \mathbf{d}_i \end{bmatrix}$ could be replaced by another pair $\begin{bmatrix} \hat{\mathbf{C}}_i & \hat{\mathbf{d}}_i \end{bmatrix}$ with fewer rows without altering the set of feasible control sequences $\mathbf{u}$.

3.2.3 Effects of Estimation Errors

Only the parameters $\mathbf{C}_i, \mathbf{d}_i$ and $\mathbf{f}$ in (3.4) depend on the estimation error $\tilde{x}(t)$ through the state estimate $\hat{x}_0$. The two fundamental means by which the estimation error perturbs the solution of (3.4), and thus the closed-loop control signals, are as following.

1. State and state estimate may lie within the same polytopic region R_i. If $\tilde{x}$ is within a – potentially small – polytopic neighborhood enclosing the origin, the estimation error affects the vectors $\mathbf{d}_i$ and $\mathbf{f}$ linearly.[1]

2. State and state estimate may lie within different polytopic regions. If $\tilde{x}$ is such that $\hat{x}_0 \in \mathrm{R}_j$ and $x(t) \in \mathrm{R}_i$ with $i \neq j$, the mpQP (3.4) is additionally perturbed by a change in the equality constraints with $\mathbf{C}_i \to \mathbf{C}_j$ and $\mathbf{d}_i \to \mathbf{d}_j$.

We explore these two mechanisms in the following. Suppose first that the state and its estimate belong to the same polytopic region, R_i. Within this polytope, the estimation

[1] Joint perturbation in the linear cost term $\mathbf{f}$ and the constant term $\mathbf{d}$ of the equality constraints corresponds to a right-hand side perturbation of the underlying KKT system

$$\begin{bmatrix} \mathbf{H} & \mathbf{C}^T \\ \mathbf{C} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} -\mathbf{f} \\ \mathbf{d} \end{bmatrix}.$$

Sensitivity of the cost objective to such perturbations is analyzed e.g. in [37].

error perturbs the vectors $\mathbf{d}_i$ and $\mathbf{f}$ linearly. For $\hat{x}_0 \in \mathrm{R}_i$, all $\mathbf{u}$ satisfying the equality-constraints in (3.4) are given by

$$\mathbf{u} = \mathbf{C}_i^T \left(\mathbf{C}_i \mathbf{C}_i^T\right)^{-1} \mathbf{d}_i + \mathbf{Z}_i \mathbf{w} = \mathbf{C}_i^+ \mathbf{d}_i + \mathbf{Z}_i \mathbf{w}, \tag{3.7}$$

where $\mathbf{w} \in \mathbb{R}^{N \cdot m - p_i}$, $\mathbf{Z}_i \in \mathbb{R}^{N \cdot m \times (N \cdot m - p_i)}$ can be any basis for the null space $\mathcal{N}\{\mathbf{C}_i\}$ of $\mathbf{C}_i$ and the integer-valued p_i denotes the number of active constraints associated with the region R_i, i.e. the number of equality-constraints in (3.4). We can now rewrite the mpQP for $\hat{x}_0 \in \mathrm{R}_i$ as an equivalent lower-dimensional, unconstrained quadratic program in $\mathbf{w}$ with cost

$$\frac{1}{2}\mathbf{w}^T \mathbf{Z}_i^T \mathbf{H} \mathbf{Z}_i \mathbf{w} + \left(\mathbf{d}_i^T (\mathbf{C}_i^+)^T \mathbf{H} + \mathbf{f}^T\right) \mathbf{Z}_i \mathbf{w} + \frac{1}{2} \mathbf{d}_i^T (\mathbf{C}_i^+)^T \mathbf{H} \mathbf{C}_i^+ \mathbf{d}_i + \mathbf{f}^T \mathbf{C}_i^+ \mathbf{d}_i + \mathbf{c},$$

where the quantity $\mathbf{Z}_i^T \mathbf{H} \mathbf{Z}_i$ is the reduced Hessian. Since $\mathbf{H} \succ 0$ and $\mathbf{Z}_i$ has full rank, we have $\mathbf{Z}_i^T \mathbf{H} \mathbf{Z}_i \succ 0$, which yields the optimal solution, $\mathbf{w}^*$, of the subspace QP as

$$\mathbf{w}^* = -\left(\mathbf{Z}_i^T \mathbf{H} \mathbf{Z}_i\right)^{-1} \mathbf{Z}_i^T \left(\mathbf{f} + \mathbf{H} \mathbf{C}_i^+ \mathbf{d}_i\right),$$

and thus the optimal predicted control trajectory

$$\mathbf{u}^* = \mathbf{C}_i^+ \mathbf{d}_i - \mathbf{Z}_i \left(\mathbf{Z}_i^T \mathbf{H} \mathbf{Z}_i\right)^{-1} \mathbf{Z}_i^T \left(\mathbf{f} + \mathbf{H} \mathbf{C}_i^+ \mathbf{d}_i\right) = \alpha_i + \beta_i \hat{x}_0,$$

where

$$\begin{aligned} \alpha_i &= \left(I_m - \mathbf{Z}_i \left(\mathbf{Z}_i^T \mathbf{H} \mathbf{Z}_i\right)^{-1} \mathbf{Z}_i^T \mathbf{H}\right) \mathbf{C}_i^+ d_i, \\ \beta_i &= -\left(\mathbf{C}_i^+ C_{x,i} + \mathbf{Z}_i \left(\mathbf{Z}_i^T \mathbf{H} \mathbf{Z}_i\right)^{-1} \mathbf{Z}^T \left(2\mathcal{B}^T \operatorname{diag}\left(I_{N-1} \otimes Q, P\right) - \mathbf{H} \mathbf{C}_i^+ C_{x,i}\right)\right) \mathcal{A}. \end{aligned} \tag{3.8}$$

Hence, provided $\tilde{x}(t)$ is sufficiently small such that $\hat{x}_0$ remains within the region associated with $x(t)$, the anticipated optimal control sequence $\mathbf{u}^*$ and thus the injected control

inputs $u(t)$ are perturbed linearly in $\tilde{x}$. Notice that when more constraints are imposed in the MPC problem (3.3) or the control horizon N is increased, the polytopic regions partitioning the feasible subset of the state-space shrink. That is, both control horizon and primal inequality constraints influence sizing of the polytopic regions for which the solution of (3.4) is computed. The following corollary to the results from [33] summarizes the effects of $\tilde{x}$ on the closed-loop control inputs as outlined in this section.

Corollary 3.1. *If the estimation errors $\tilde{x}(t)$ are such that $x(t)$ and $\hat{x}_0$ lie within the same region R_i, then the computed optimal control sequence $\mathbf{u}$ is perturbed by the linear function*

$$\tilde{\mathbf{u}}(\tilde{x}(t)) = \mathbf{u}(x(t)) - \mathbf{u}(\hat{x}_0) = \beta_i \tilde{x}(t) \tag{3.9}$$

with β_i as in (3.8).

Example 3.1. To illustrate the effects of small estimation errors such that $\{x(t), \hat{x}_0\} \subset \mathrm{R}_i$, consider the following example of a multiparametric quadratic program as in (3.4). The problem could correspond to a single-input single-output MPC problem with horizon $N = 2$ and a single active constraint on the polytopic region R_i. Suppose

$$\mathbf{H} = \begin{bmatrix} 2 & 1 \\ & \\ 1 & 1 \end{bmatrix} \succ 0, \qquad\qquad \mathbf{C}_i = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

are constant matrices. Given that we are not interested in the optimal cost but the optimal control sequence $\mathbf{u}$, we choose $\mathbf{c} = 0$ without effect on the remainder of this example. Perturbations enter linearly through the vectors $\mathbf{f}$ and $\mathbf{d}_i$. Changes of $\mathbf{f}$ amount to translations of the level sets associated with the cost index, whereas perturbations in $\mathbf{d}_i$ equal

orthogonal translations of constraint hyperplanes. To capture these effects, we compare the two linear cost terms

$$\mathbf{f} = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix}, \qquad \hat{\mathbf{f}} = \mathbf{f} + \tilde{\mathbf{f}} = -\begin{bmatrix} 10 \\ \\ 6 \end{bmatrix} \tag{3.10}$$

as well as the two constraint right-hand sides

$$\mathbf{d}_i = 2, \qquad \hat{\mathbf{d}}_i = \mathbf{d}_i + \tilde{\mathbf{d}} = -3. \tag{3.11}$$

The quantities $\mathbf{f}$ and $\mathbf{d}_i$ are associated with the original, unperturbed quadratic program. That is, they are associated with a equality-constrained mpQP problem at time t were the state $x(t)$ to be known perfectly. Conversely, the quantities $\hat{\mathbf{f}}$ and $\hat{\mathbf{d}}_i$ are associated with a perturbed version of this quadratic program under the assumption that $\mathbf{C}_i$ remains constant.

By linearity, we can interpret the joint perturbation geometrically as a sequence of the effects with perturbing $\mathbf{f}$ and $\mathbf{d}_i$. This sequence is illustrated for the perturbations in (3.10) and (3.11) in Figure 3.2. The change $\mathbf{f} \to \hat{\mathbf{f}}$ by itself results in a translation of the level sets with shape associated only with the same $\mathbf{H}$. The resulting perturbation of the anticipated control sequence $\mathbf{u}$ is a projection of the variation $-\mathbf{H}^{-1}\tilde{\mathbf{f}}$ associated with the unconstrained problem onto the single active constraint

$$u_1 = 1 - 0.5\, u_0.$$

Similarly, if only the right-hand side of the active constraint is perturbed via $\mathbf{d}_i \to \hat{\mathbf{d}}_i$, we experience a change of the optimal control sequence along a line connecting the

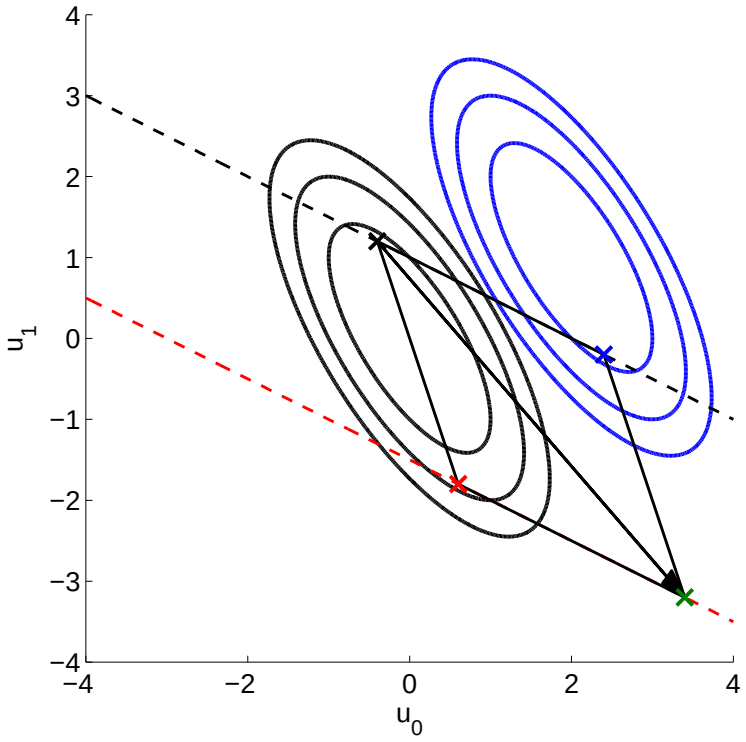

Figure 3.2: Shifts of level sets, constraint and optimal solution for linearly perturbed QP in Example 3.1. Unperturbed level sets and constraint in black, perturbed level sets in blue, perturbed constraint in red. Optimal solutions are marked with x's, where black corresponds to the unperturbed problem, blue to an intermediate problem with only the level sets perturbed, red to an intermediate problem with only the constraint perturbed and green to the full perturbed quadratic program.

unconstrained optimizer and the unperturbed constrained optimizer. As in (3.7), this corresponds to picking a point on the constraint and moving it with respect to a reduced-order unconstrained quadratic program on the constraint manifold.

The outcome of the joint perturbation $(\mathbf{f}, \mathbf{d}_i) \to (\hat{\mathbf{f}}, \hat{\mathbf{d}}_i)$ is a superposition of these two individual perturbation effects, as can be seen in Figure 3.2. In comparison, the effect of estimation errors on control inputs generated by unconstrained quadratic programming is determined solely by the first effect without the projection component given

that the unconstrained case automatically ensures that $x(t)$ and $\hat{x}_0$ lie within a shared region.

Suppose now that $\tilde{x}(t)$ is such that $x(t) \in \mathrm{R}_i$ and $\hat{x}_0 \in \mathrm{R}_j$ for some pair $i \neq j$. In this situation, the assumption for Corollary 3.1 and (3.9) does not hold. Instead, the outcome of such estimation errors $\tilde{x}(t)$ is a function explicitly depending on both state and state estimation error. This function is affine in the state with parameters depending on the difference between the affine control laws on R_i and R_j and linear in the estimation errors. Corollary 3.1 and (3.9) are a particular case of (3.12) for $i = j$ of the following result.

Corollary 3.2. *If $x(t) \in \mathrm{R}_i$ and $\hat{x}_0 \in \mathrm{R}_j$, then*

$$\tilde{\mathbf{u}}(x(t), \tilde{x}(t)) = \mathbf{u}(x(t)) - \mathbf{u}(\hat{x}_0) = \left(\alpha_i - \alpha_j\right) + \left(\beta_i - \beta_j\right) x(t) + \beta_j \tilde{x}(t). \tag{3.12}$$

This corollary demonstrates that the state estimation error has the capacity to shift the active constraints for $\hat{x}_0$ versus those intended for $x(t)$. This could have significant consequences for the applied control were the state constraints truly important. Indeed, the computed $u(t)$ could be infeasible for the full MPC mpQP at $x(t)$. On the other hand, if the difference between sets of active constraints occurs only for optional or discretionary constraints (included simply because MPC invites us to), there might be little consequence to this violation.

Building on this observation, a question not adressed so far is how the MPC solution and its sensitivity to estimation errors change when we add extra constraints to the initial MPC problem (3.3). In this case, we get a new partition of the feasible region

which is nested in the previously computed partition. That is, each of the new polytopic regions is either identical to or a polytopic subset of a region associated with the initial problem. This results in estimation errors of smaller size being able to change the active region R_i. This is a price paid for enforcing additional constraints in MPC problems.

We next analyze a numerical example in which we demonstrate the effects estimation errors can have on closed-loop performance of a controlled system.

Example 3.2. Consider the system

$$\begin{bmatrix} x(t+1) \\ y(t+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix},$$

and linear quadratic MPC problem

$$\begin{aligned} \min_{u_0, v_0} \ & \left\{ \hat{x}_1^2 + \hat{y}_1^2 + r\left(u_0^2 + v_0^2\right) \right\} \\ \text{s.t. } & \hat{x}_1 - \hat{y}_1 \leq 1, \\ & \hat{x}_1 = \hat{x}_0 + u_0, \\ & \hat{y}_1 = \hat{y}_0 + v_0, \\ & \hat{x}_0 = x(t) - \tilde{x}(t) \\ & \hat{y}_0 = y(t) - \tilde{y}(t). \end{aligned}$$

That is, we aim to achieve $x(t) - y(t) \leq 1$ while minimizing the quadratic objective function. Suppose we choose the control penalty $r > 0$ sufficiently small such that the problem operates on the constraint for all time. Under this assumption, the optimal solution resulting from the above MPC problem is the affine state-estimate feedback

control law

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} -\frac{2+r}{2+2r} & \frac{r}{2+2r} \\ \frac{r}{2+2r} & -\frac{2+r}{2+2r} \end{bmatrix} \begin{bmatrix} \hat{x}_0 \\ \hat{y}_0 \end{bmatrix} + \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix}.$$

This affine control leads to the closed-loop dynamics

$$\begin{bmatrix} x(t+1) \\ y(t+1) \end{bmatrix} = \begin{bmatrix} \frac{r}{2+2r} & \frac{r}{2+2r} \\ \frac{r}{2+2r} & \frac{r}{2+2r} \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} + \begin{bmatrix} \frac{2+r}{2+2r} & -\frac{r}{2+2r} \\ -\frac{r}{2+2r} & \frac{2+r}{2+2r} \end{bmatrix} \begin{bmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{bmatrix}$$

The optimal controls produced by the constrained MPC problem lead to coupling between the state elements $x(t)$ and $y(t)$ in closed loop through the active constraint. The stable closed-loop system matrix has eigenvalues $\lambda_1 = 0$ and $\lambda_2 = r/(1+r)$, with unperturbed steady-state solution $x_\infty = 1/2$ and $y_\infty = -x_\infty = -1/2$. Given white random zero-mean disturbances $\tilde{x}(t)$ and $\tilde{y}(t)$ with covariance matrix

$$W = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

the closed-loop system state has mean $\begin{bmatrix} 1/2 & -1/2 \end{bmatrix}^T$ and motion around the unperturbed steady-state solution. The closed-loop covariance matrix X is determined by the Lyapunov equation

$$\begin{bmatrix} \frac{r}{2+2r} & \frac{r}{2+2r} \\ \frac{r}{2+2r} & \frac{r}{2+2r} \end{bmatrix} X \begin{bmatrix} \frac{r}{2+2r} & \frac{r}{2+2r} \\ \frac{r}{2+2r} & \frac{r}{2+2r} \end{bmatrix} - X + \begin{bmatrix} \frac{2+r}{2+2r} & -\frac{r}{2+2r} \\ -\frac{r}{2+2r} & \frac{2+r}{2+2r} \end{bmatrix} \begin{bmatrix} \frac{2+r}{2+2r} & -\frac{r}{2+2r} \\ -\frac{r}{2+2r} & \frac{2+r}{2+2r} \end{bmatrix} = 0,$$

which for $r = 1$ yields

$$X = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix},$$

with eigenvalues $\lambda_1 = 1/3$ and $\lambda_2 = 1$. Now suppose we replace our original MPC problem by the following problem with modified cost to track the point $\begin{bmatrix} 0 & 4 \end{bmatrix}^T$.

$$\begin{aligned} \min_{u_0, v_0} \ & \left\{ \hat{x}_1^2 + (\hat{y}_1 - 4)^2 + r\left(u_0^2 + v_0^2\right) \right\} \\ \text{s.t. } & \hat{x}_1 - \hat{y}_1 \leq 1, \\ & \hat{x}_1 = \hat{x}_0 + u_0, \\ & \hat{y}_1 = \hat{y}_0 + v_0, \\ & \hat{x}_0 = x(t) - \tilde{x}(t) \\ & \hat{y}_0 = y(t) - \tilde{y}(t). \end{aligned}$$

In this case, the cost function ensures that the constraint will never be active. That is, we compute the optimal control inputs by solving the unconstrained version of a modified problem. The resulting affine optimal control signals are

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{1+r} & 0 \\ 0 & -\frac{1}{1+r} \end{bmatrix} \begin{bmatrix} \hat{x}_0 \\ \hat{y}_0 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{4}{1+r} \end{bmatrix},$$

yielding the decoupled, stable closed-loop dynamics

$$\begin{bmatrix} x(t+1) \\ y(t+1) \end{bmatrix} = \begin{bmatrix} \frac{r}{1+r} & 0 \\ 0 & \frac{r}{1+r} \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{4}{1+r} \end{bmatrix} + \begin{bmatrix} \frac{1}{1+r} & 0 \\ 0 & \frac{1}{1+r} \end{bmatrix} \begin{bmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{bmatrix}.$$

Solving for the closed-loop covariance matrix with $W = I$ and $r = 1$ now yields

$$X = \frac{1}{3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

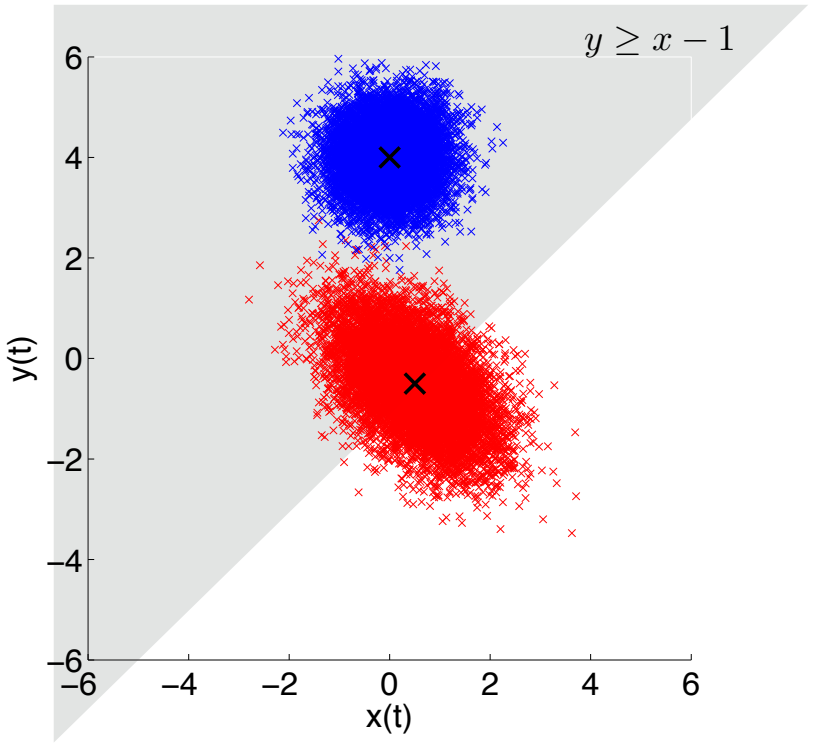

Figure 3.3: Closed-loop simulation of actual state values for Example 3.2, MPC with state estimate feedback. The shaded area is the constraint region $y \geq x - 1$ in the real state variables. Red crosses indicate state values for state estimate feedback MPC with active constraint. Blue crosses are for state estimate feedback without active constraint. Mean values are marked by black xs. Both sets of solutions exhibit the effect of state estimate errors. Note the proportion of infeasible state values for the actively constrained problem. Note also the amplification of the variance of the solutions orthogonal to the active constraint when compared to the unconstrained solution.

which has eigenvalues $\lambda_1 = \lambda_2 = 1/3$. That is, the modified (no longer actively constrained) MPC problem leads to decreased closed-loop variance with respect to the original problem. This demonstrates how the presence of an active constraint in the MPC formulation can potentially cause increased closed-loop sensitivity to estimation errors. Closed-loop behavior for both systems is displayed jointly in Figure 3.3 for zero-mean, normally distributed estimation errors with covariance matrix $W = I$.

3.3 Nonlinear Model Predictive Control

In this section, we discuss briefly how our observations for linear quadratic MPC extend to nonlinear MPC problems, where system, constraints and cost can take more general forms. Suppose our system is described by

$$x(t+1) = f(x(t), u(t))$$

and future states from time t are predicted via

$$\hat{x}_{k+1}(t) = f(\hat{x}_k(t), u_k(t)),$$

$$\hat{x}_0(t) = x(t) - \tilde{x}(t),$$

where $\tilde{x}(t)$ again functions as an additive estimation error produced by an unspecified state estimator. The general nonlinear MPC counterpart to (3.3) takes the form

$$\begin{aligned} &\min_{\mathbf{u}} \left\{ p(\hat{x}_N) + \sum_{k=0}^{N-1} l(\hat{x}_k, u_k) \right\}, \\ &\text{s.t.} \quad g_k(\hat{x}_k, u_k) \leq d, \qquad\qquad k = 0, \ldots, N-1; \\ &\qquad \hat{x}_{k+1} = f(\hat{x}_k(t), u_k(t)), \quad k = 0, \ldots, N-1; \\ &\qquad \hat{x}_0 = x(t) - \tilde{x}(t). \end{aligned} \tag{3.13}$$

Just as in the linear quadratic case, this MPC formulation leads to a multiparametric optimization problem. Using all possible collections of active constraints, we can decompose this optimization problem into a number of local equality-constrained optimization problems with equal parametric cost function. Under mild regularity assumptions on the constraints, the implicit function theorem then states that the cost for each collec-

tion of active constraints can be optimized over a subspace resulting from the equality-constraints, where each non-redundant active constraint reduces the subspace dimension by one. We then arrive at an explicit state-estimate feedback solution over a partition of the feasible region. This partition is generally not polytopic, just as the feasible region cannot be expected to be a polytope. However, the means by which estimation errors can perturb the closed-loop are structurally the same as in the linear quadratic case.

1. State and state estimate may lie within the same region R_i. If $\tilde{x}$ is within a given set including the origin, the estimation error affects the optimal control sequence through the local control law on R_i.

2. State and state estimate may lie within different constraint regions. If $\tilde{x}$ is such that $\hat{x}_0 \in \mathrm{R}_j$ and $x(t) \in \mathrm{R}_i$ with $i \neq j$, the nonlinear MPC problem is additionally perturbed by a change in the set of active equality constraints, resulting in an effect depending on the local control laws on both R_i and R_j.

3.4 Conclusions

In this chapter, we have explored how constraints in model predictive control do not generally come for free. While the commonly employed assumption of full state-feedback masks this drawback, it becomes apparent when dealing with state-estimate feedback control. In this case, larger numbers of active constraints result in extra channels through which estimation errors might be propagated into the control inputs and subsequently perturb the closed-loop dynamics. These observations apply universally

to model predictive control and suggest careful consideration when imposing constraints on optimal control problems.

Acknowledgements

Chapter 3, in full, is a reprint of the material as it appears in: M.A. Sehr, R.R. Bitmead, "Sumptus Cohiberi: The Cost of Constraints in MPC with State Estimates", *Proc. American Control Conference*, 2016, pp 901–906. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Stochastic Model Predictive Control: Output-Feedback, Duality and Performance

4.1 Introduction

MPC, in its original formulation, is a full-state feedback law. This underpins two theoretical limitations of MPC: accommodation of output feedback, and extension to include a cogent robustness theory since the state dimension is fixed. This chapter addresses the first question. There have been a number of approaches, mostly hinging on replacement of the measured true state by a state estimate, which is computed via Kalman filtering [38, 4], moving-horizon estimator [6, 5], tube-based minimax estimators [7], etc. Apart from [6], these designs, often for linear systems, separate the

estimator design from the control design. The control problem may be altered to accommodate the state estimation error by methods such as: constraint tightening [4], chance/probabilistic constraints [8], and so forth.

In this chapter, we first consider Stochastic Model Predictive Control (SMPC), formulated as a variant of Stochastic Optimal Output Feedback Control (SOOFC), without regard to computational tractability restrictions. By taking this route, we establish a formulation of SMPC which possesses central features: accommodation of output feedback and duality/probing; examination of the probabilistic requirements of deterministic and probabilistic constraints; guaranteed performance of the SMPC controller applied to the system. Performance bounds are stated in relation to the infinite-horizon-optimally controlled closed-loop performance.

This chapter does *not* seek to provide a comprehensive survey of the myriad alternative approaches proposed for Stochastic Model Predictive Control (SMPC). For that, we recommend the numerous available references such as [39, 40, 2, 41]. Rather, we present a new algorithm for SMPC based on SOOFC and prove, particularly, performance properties relative to optimality. As a by-product, we acquire a natural treatment of output feedback via the Bayesian Filter and of the associated controller duality required to balance probing for observability enhancement and regulation. The price we pay for general nonlinear systems is the suspension of disbelief in computational tractability. However, the approach delineates a target controller with assured properties. Approximating this intractable controller by a more computationally amenable variant, as opposed to identifying soluble but indirect problems without guarantees, holds the

prospect of approximately attracting the benefits. Such a strategy, using a particle implementation of the Bayesian filter and scenario methods at the cost of losing duality of the control inputs, is discussed in [42].

The structure of the chapter is as follows. Section 4.2 briefly formulates SOOFC, as used in Section 4.3 to present a new SMPC algorithm. After discussing recursive feasibility of this algorithm in Section 4.4, we proceed by establishing conditions for boundedness of the infinite-horizon discounted cost of the SMPC-controlled nonlinear system in Section 4.5. Section 4.6 ties the performance of SMPC to the infinite-horizon SOOFC performance. Section 4.7 provides a brief encapsulation and post-analysis of the results and the set of technical assumptions in the chapter. We conclude the chapter in Section 4.8.

Notation

$\mathbb{R}$ and $\mathbb{R}_+$ are real and non-negative real numbers, respectively. The set of non-negative integers is denoted $\mathbb{N}_0$ and the set of positive integers by $\mathbb{N}_1$. We write sequences as $\mathbf{t}^m \triangleq \{t_0, t_1, \ldots, t_m\}$, where $m \in \mathbb{N}_0$; $\mathbf{t}^\infty$ is an infinite sequence of the same form. pdf(X) denotes the probability density function of random variable X while pdf$(X|Y)$ denotes the conditional probability density function of random variable X given jointly distributed random variable Y. The acronyms a.s., a.e. and i.i.d. stand for *almost sure, almost everywhere* and *independent and identically distributed*, respectively.

4.2 Stochastic Optimal Output-Feedback Control

We consider stochastic optimal control of nonlinear time-invariant dynamics of the form

$$x_{k+1} = f(x_k, u_k, w_k), \quad x_0, \tag{4.1}$$

$$y_k = h(x_k, v_k), \tag{4.2}$$

where $k \in \mathbb{N}_0$, $x_k \in \mathbb{R}^{n_x}$ denotes the state with initial value x_0, $u_k \in \mathbb{R}^{n_u}$ the control input, $y_k \in \mathbb{R}^{n_y}$ the measurement output, $w_k \in \mathbb{R}^{n_w}$ the process noise and $v_k \in \mathbb{R}^{n_v}$ the measurement noise. We denote by

$$\pi_{0|-1} \triangleq \text{pdf}(x_0) \tag{4.3}$$

the known a-priori density of the initial state and by

$$\zeta^k \triangleq \{y_0, u_0, y_1, u_1, \ldots, u_{k-1}, y_k\}, \qquad \zeta^0 \triangleq \{y_0\}$$

the data available at time k. We make the following standing assumptions on the random variables and system dynamics.

Assumption 4.1. The dynamics (4.1-4.2) satisfy

1. $f(\cdot, u, \cdot)$ is differentiable a.e. with full rank Jacobian $\forall\, u \in \mathbb{R}^{n_u}$.

2. $h(\cdot, \cdot)$ is differentiable a.e. with full rank Jacobian.

3. w_k and v_k are i.i.d. sequences with known densities.

4. x_0, w_k, v_l are mutually independent for all $k, l \geq 0$.

Assumption 4.2. The control input u_k at time instant $k \geq 0$ is a function of the data ζ^k and $\pi_{0|-1}$.

As there is no direct feedthrough from u_k to y_k, Assumptions 4.1 and 4.2 assure that system (4.1-4.2) is a *controlled Markov process* [43]. Assumption 4.1 further ensures that f and h enjoy the *Ponomarev 0-property* [44] and hence that x_k and y_k possess joint and marginal densities.

4.2.1 Information State & Bayesian Filter

Definition 4.1. The conditional density of state x given data ζ^k,

$$\pi_k \triangleq \mathrm{pdf}\left(x_k \mid \zeta^k\right), \qquad k \in \mathbb{N}_0, \tag{4.4}$$

is the *information state* of system (4.1-4.2).

For a Markov system such as (4.1-4.2), the information state is propagated via the *Bayesian Filter* (e.g. [45, 46]):

$$\pi_k = \frac{\mathrm{pdf}(y_k \mid x_k)\, \pi_{k|k-1}}{\int \mathrm{pdf}(y_k \mid x_k)\, \pi_{k|k-1}\, dx_k}, \tag{4.5}$$

$$\pi_{k+1|k} \triangleq \int \mathrm{pdf}(x_{k+1} \mid x_k, u_k)\, \pi_k\, dx_k, \tag{4.6}$$

for $k \in \mathbb{N}_0$ and density $\pi_{0|-1}$ as in (4.3). For linear dynamics and Gaussian noise, the recursion (4.5-4.6) yields the Kalman Filter.

Definition 4.2. The recursion (4.5-4.6) defines the mapping

$$\pi_{k+1} = T\left(\pi_k, y_{k+1}, u_k\right), \qquad k \in \mathbb{N}_0. \tag{4.7}$$

4.2.2 Cost and Constraints

Definition 4.3. $\mathbb{E}_k[\,\cdot\,]$ and $\mathbb{P}_k[\,\cdot\,]$ are expected value and probability with respect to state x_k – with conditional density π_k – and i.i.d. random variables $\{(w_j, v_{j+1}) : j \geq k\}$.

Given the available data ζ^0, we aim to select non-anticipatory (i.e. subject to Assumption 4.2) control inputs u_k to minimize

$$J_N(\pi_0, \mathbf{u}^{N-1}) \triangleq \mathbb{E}_0\left[\sum_{j=0}^{N-1} \alpha^j c(x_j, u_j) + \alpha^N c_N(x_N)\right], \tag{4.8}$$

where N is the control horizon, $c : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}_+$ the stage cost, $c_N : \mathbb{R}^{n_x} \to \mathbb{R}_+$ the terminal cost and $\alpha \in \mathbb{R}_+$ a discount factor. Drawing from the literature (e.g. [47, 43]), optimal controls in (4.8) must inherently be *separated* feedback policies. That is, control input u_k depends on data ζ^k and initial density $\pi_{0|-1}$ solely through the current information state π_k, leading to the closed-loop architecture displayed in Figure 4.1. Optimality thus requires propagating π_k and policies g_k, where

$$u_k = g_k(\pi_k). \tag{4.9}$$

Cost (4.8) then reads

$$J_N(\pi_0, \mathbf{g}^{N-1}) = \mathbb{E}_0\left[\sum_{j=0}^{N-1} \alpha^j c(x_j, g_j(\pi_j)) + \alpha^N c_N(x_N)\right]. \tag{4.10}$$

Extending stochastic optimal control problems with cost (4.10) to the *infinite horizon* (see [47, 48]) typically requires $\alpha < 1$ and omitting the terminal cost term $c_N(\cdot)$, leading to

$$J_\infty(\pi_0, \mathbf{g}^\infty) \triangleq \mathbb{E}_0\left[\sum_{j=0}^{\infty} \alpha^j c(x_j, g(\pi_j))\right]. \tag{4.11}$$

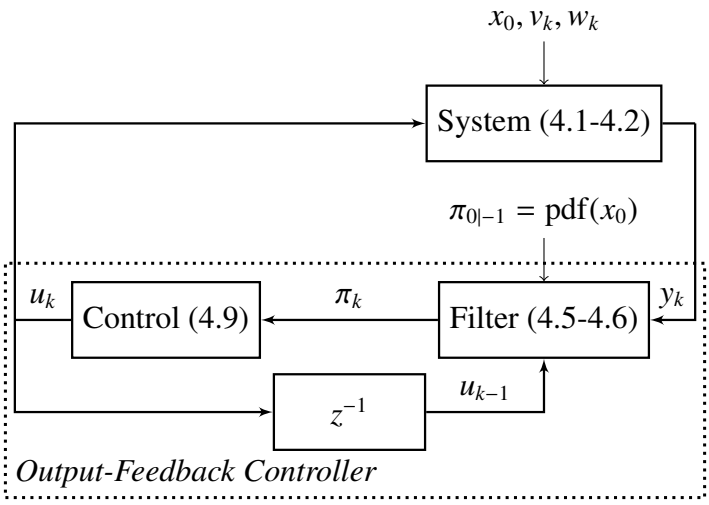

Figure 4.1: Closed-loop system architecture for stochastic optimal output-feedback control based on information state π_k.

In addition to minimizing the expected value cost (4.10), we impose probabilistic state constraints of the form

$$\mathbb{P}_k [x_k \in \mathcal{X}_k] \geq 1 - \epsilon_k, \qquad k \in \mathbb{N}_1 \tag{4.12}$$

for $\epsilon_k \in [0, 1)$. That is, we enforce constraints with respect to the known distributions of the future noise variables and the conditional density of the current state x_k, captured by the information state π_k. Moreover, we consider input constraints of the form

$$u_k = g_k(\pi_k) \in \mathcal{U}_k, \qquad k \in \mathbb{N}_0. \tag{4.13}$$

When discussing infinite-horizon optimal control with cost (4.11), we replace the state constraints (4.12) by the stationary probabilistic state constraints

$$\mathbb{P}_k [x_k \in \mathcal{X}_\infty] \geq 1 - \epsilon_\infty, \qquad k \in \mathbb{N}_1 \tag{4.14}$$

for $\epsilon_\infty \in [0, 1)$ and the input constraints (4.13) by

$$u_k = g_k(\pi_k) \in \mathcal{U}_\infty, \qquad\qquad k \in \mathbb{N}_0.$$

Definition 4.4. Denote by $\mathcal{D}$ the set of all densities on $\mathbb{R}^{n_x}$. Further define $C_k \subseteq \mathcal{D}, k \in \mathbb{N}_1$, to be the set of all π_k of x_k satisfying the probabilistic constraint (4.12). Define C_∞ likewise for (4.14).

4.2.3 Stochastic Optimal Control

Definition 4.5. Given dynamics (4.1-4.2), $\alpha \in \mathbb{R}_+$ and horizon $N \in \mathbb{N}_1$, define the *finite-horizon stochastic optimal control problem*

$$\mathcal{P}_N(\pi_0) : \begin{cases} \inf_{\mathbf{g}^{N-1}} & J_N(\pi_0, \mathbf{g}^{N-1}) \\ \text{s.t.} & \mathbb{P}_j\left[x_j \in \mathcal{X}_j\right] \geq 1 - \epsilon_j,\ j = 1, \ldots, N. \\ & g_j(\pi_j) \in \mathcal{U}_j,\ j = 0, \ldots, N-1. \end{cases}$$

Definition 4.6. Given dynamics (4.1-4.2) and $\alpha \in \mathbb{R}_+$, define the *infinite-horizon stochastic optimal control problem*

$$\mathcal{P}_\infty(\pi_0) : \begin{cases} \inf_{\mathbf{g}^\infty} & J_\infty(\pi_0, \mathbf{g}^\infty) \\ \text{s.t.} & \mathbb{P}_j\left[x_j \in \mathcal{X}_\infty\right] \geq 1 - \epsilon_\infty,\ j \in \mathbb{N}_1. \\ & g_j(\pi_j) \in \mathcal{U}_\infty,\ j \in \mathbb{N}_0. \end{cases}$$

Definition 4.7. π_0 is feasible for $\mathcal{P}_N(\cdot)$ if there exists a sequence of policies $\mathbf{g}^{N-1}$ such that, $\{w_j, v_{j+1}\}_{j\geq 0}$-a.s., $u_j = g_j(\pi_j)$ satisfy the constraints and $J_N(\pi_0, \mathbf{g}^{N-1})$ is finite. Define feasibility likewise for $\mathcal{P}_\infty(\pi_0)$.

In Stochastic Optimal Control, feasibility entails the existence of policies $g_k(\cdot)$ such that for any $\pi_k \in C_k$, $g_k(\pi_k) \in \mathcal{U}_k$ and

$$\pi_{k+1} = T(\pi_k, y_{k+1}, g_k(\pi_k)) \in C_{k+1}, \qquad (w_k, v_{k+1}) - \text{a.s.}$$

Even though the state constraints (4.12) are probabilistic, this condition results in an equivalent almost sure constraint on the conditional state densities. The stochastic optimal feedback policies in $\mathcal{P}_N(\pi_0)$ may now be computed in principle by solving the Stochastic Dynamic Programming Equation (SDPE),

$$\begin{aligned} V_k(\pi_k) \triangleq \inf_{g_k(\cdot)} & \mathbb{E}_k \left[c(x_k, g_k(\pi_k)) + \alpha V_{k+1}(\pi_{k+1}) \right], \\ \text{s.t. } & \pi_{k+1} \in C_{k+1}, \ (w_k, v_{k+1}) - \text{a.s.} \\ & g_k(\pi_k) \in \mathcal{U}_k \end{aligned} \tag{4.15}$$

for $k = 0, \ldots, N-1$ and $\pi_k \in C_k$. The equation is solved backwards in time, from its terminal value

$$V_N(\pi_N) \triangleq \mathbb{E}_N \left[c_N(x_N) \right], \qquad \pi_N \in C_N. \tag{4.16}$$

Solution of the SDPE is the primary source of the restrictive computational demands in Stochastic Optimal Control. The reason for this difficulty lies in the dependence of the future information state in each step of (4.15-4.16) on the current and future control inputs. While the dependence on future control inputs is limiting even in deterministic control, the computational burden is drastically worsened in the stochastic case because of the complexity of the operator T_k in (4.7). On the other hand, optimality via the SDPE leads to a control law of *dual* nature. *Dual optimal control* connotes the

compromise in optimal control between the control signal's function to reveal the state and its function to regulate that state. These dual actions are typically antagonistic [49]. The duality of stochastic optimal control is a generic feature, although there exist some problems – called *neutral* – where the probing nature of the control evanesces, linear Gaussian control being one such case.

Notice that, while the Bayesian Filter (4.5-4.6) can be approximated to arbitrary accuracy using a Particle Filter [46], the SDPE cannot be easily simplified without loss of optimal probing in the control inputs. While control laws generated without solution of the SDPE can be modified artificially to include certain excitation properties, as discussed for instance in [50, 11], such approaches are suboptimal and do not generally enjoy the theoretical guarantees discussed below. For the stochastic optimal control problems considered here, excitation of the control signal is incorporated automatically and as necessary through the optimization. The optimal control policies, $g_j^\star(\cdot)$, will inherently inject excitation into the control signal depending on the quality of state knowledge embodied in π_k.

4.3 Stochastic Model Predictive Control

Algorithm 1 *

(Dual Optimal) SMPC

Given: $\pi_{0|-1} \in \mathcal{D}$ and $\alpha \in \mathbb{R}_+$

1: **Offline:** Solve $\mathcal{P}_N(\cdot)$ for $g_0^\star(\cdot)$ via (4.15-4.16)

2: **Online:**

3: **for** $k \in \mathbb{N}_0$ **do**

4: Measure y_k

5: Compute π_k

6: Apply first optimal control policy, $u_k = g_0^\star(\pi_k)$

7: Compute $\pi_{k+1|k}$

8: **end for**

Notice how this algorithm differs from common practice in SMPC [51, 41] in that we explicitly use the information states $\pi_k \in \mathcal{D}$. Throughout the literature, these information states – conditional densities – are replaced by best available, or certainty-equivalent state estimates in $\mathbb{R}^{n_x}$. While this makes the problem more tractable, one no longer solves the underlying stochastic optimal control problem. As we shall demonstrate in this chapter, using information state π_k and optimal policy $g_0^\star(\cdot)$ resulting from solution of Problem $\mathcal{P}_N(\pi_k)$ at each time instance leads to a number of results regarding closed-loop performance on the infinite horizon.

4.4 Recursive Feasibility

Assumption 4.3. $\pi_{0|-1}$ yields π_0 feasible for $\mathcal{P}_N(\cdot)$, v_0-a.s.

Assumption 4.4. The constraints in $\mathcal{P}_N(\cdot)$ and $\mathcal{P}_\infty(\cdot)$, for $j = 1, \ldots, N-1$, satisfy

$$C_{j+1} \subseteq C_j \subseteq C_\infty, \qquad \mathcal{U}_j \subseteq \mathcal{U}_{j-1} \subseteq \mathcal{U}_\infty.$$

Assumption 4.5. For all densities $\pi_k \in C_N$, there exists a policy $\tilde{g}(\pi_k)$ satisfying

$$\tilde{g}(\pi_k) \in \mathcal{U}_{N-1},$$

$$T(\pi_k, y_{k+1}, \tilde{g}(\pi_k)) \in C_N, \ (w_k, v_{k+1}) - \text{a.s.},$$

$$c(x_k, \tilde{g}(\pi_k)) < \infty.$$

Theorem 4.1. *Given Assumptions 4.3-4.5, SMPC yields π_k feasible for $\mathcal{P}_N(\cdot)$, $\{w_j, v_{j+1}\}_{j\geq 0}$-a.s., for all $k \in \mathbb{N}_1$.*

Proof. It is sufficient to verify that, provided feasible initial information state π_0, information state π_1 remains feasible for $\mathcal{P}_N(\cdot)$. By feasibility of π_0 (Assumption 4.3), there exists a sequence of policies $\mathbf{g}^{N-1}$ such that

$$\mathbb{P}_{j+1}\left[x_{j+1} \in \mathcal{X}_{j+1}\right] \geq 1 - \epsilon_{j+1},$$

$$u_j = g_j(\pi_j) \in \mathcal{U}_j,$$

$$J_N(\pi_0, \mathbf{g}^{N-1}) < \infty,$$

a.s. for $j = 0, \ldots, N-1$. Now consider information state π_1, propagated via (4.5-4.6), and feedback policies

$$\hat{\mathbf{g}}^{N-1} \triangleq \{g_1, g_2, \ldots, g_{N-1}, \tilde{g}\}.$$

By Assumptions 4.4 and 4.5, we have

$$\mathbb{P}_{j+2}\left[x_{j+2} \in \mathcal{X}_{j+1}\right] \geq 1 - \epsilon_{j+1},$$

$$u_{j+1} = \hat{g}_j(\pi_{j+1}) \in \mathcal{U}_j,$$

$$J_N(\pi_1, \hat{\mathbf{g}}^{N-1}) < \infty,$$

a.s. for $j = 0, \ldots, N-1$, confirming feasibility of π_1 for $\mathcal{P}_N(\cdot)$. □

The above proof follows directly as a stochastic version of the corresponding result in deterministic MPC, e.g. [27]. Notice that recursive feasibility and compact $\mathcal{X}_1$ immediately implies a stability result independent of the cost (4.10), i.e.

$$\mathbb{P}_k[x_k \in \mathcal{X}_1] \geq 1 - \epsilon_1, \qquad \{w_j, v_{j+1}\}_{j\geq 0} - \text{a.s.}, \tag{4.17}$$

for $k \in \mathbb{N}_1$.

4.5 Convergence and Stability

Assumption 4.6. For a given $\alpha \in \mathbb{R}_+$, the terminal feedback policy $\tilde{g}(\pi)$ specified in Assumption 4.5 satisfies

$$\alpha\, \mathbb{E}_\pi\left[c_N(f(x, \tilde{g}(\pi), w))\right] - c_N(x) \overset{\text{a.s.}}{\leq} -c(x, \tilde{g}(\pi)) \tag{4.18}$$

for all densities π of x with $\pi \in C_N$. The expectation $\mathbb{E}_\pi[\cdot]$ is with respect to state x – with density π – and w.

For $\alpha \geq 1$, Assumption 4.6 can be interpreted as the existence of a stochastic Lyapunov function on the terminal set of densities, C_N. If (4.18) holds for $\alpha \geq 1$, it naturally holds for all $\alpha \in (0, 1]$.

Theorem 4.2. *Given Assumptions 4.3-4.6, SMPC yields*

$$\lim_{M\to\infty} \sum_{k=0}^{M} \alpha^k c(x_k, g_0^\star(\pi_k)) \overset{a.s.}{<} \infty. \tag{4.19}$$

Proof. Denote by M_k the discounted $\mathcal{P}_N$-cost-to-go,

$$\begin{aligned} M_k &\triangleq \sum_{j=k}^{k+N-1} \alpha^j c(x_j, g_{j-k}^\star(\pi_j)) + \alpha^{k+N} c_N(x_{k+N}) \\ &= \alpha^k \left(\sum_{j=0}^{N-1} \alpha^j c(x_{k+j}, g_j^\star(\pi_{k+j})) + \alpha^N c_N(x_{k+N}) \right), \end{aligned}$$

where $g_j^\star(\cdot)$, $j = 0, \ldots, N-1$, are the optimal feedback policies in Problem $\mathcal{P}_N(\cdot)$. Moreover, define $\mathcal{F}_k$ as the σ-algebra generated by the initial state x_0 with density $\pi_{0|-1}$ and the i.i.d. noise sequences w_j and v_j for $j = 0, \ldots, k+N-1$. Then M_k is $\mathcal{F}_k$-measurable and $M_k \geq 0$ by non-negativity of stage and terminal cost. Then,

$$\mathbb{E}_0 \left[M_{k+1} \mid \mathcal{F}_k \right] = \alpha^{k+1} \mathbb{E}_0 \left[\sum_{j=0}^{N-1} \alpha^j c(x_{j+k+1}, g_j^\star(\pi_{j+k+1})) + \alpha^N c_N(x_{k+N+1}) \mid \mathcal{F}_k \right],$$

and, by optimality of the policies $g_j^\star(\cdot)$ in $\mathcal{P}_N(\cdot)$,

$$\begin{aligned} \mathbb{E}_0[M_{k+1} \mid \mathcal{F}_k] \overset{\text{a.s.}}{\leq} \mathbb{E}_0 \Big[& M_k - \alpha^k c(x_k, g_0^\star(\pi_k)) - \alpha^{k+N} c_N(x_{k+N}) + \\ & \alpha^{k+N} c(x_{k+N}, \tilde{g}(\pi_{k+N})) + \alpha^{k+N+1} c_N(f(x_{k+N}, \tilde{g}(\pi_{k+N}), w_{k+N})) \mid \mathcal{F}_k \Big], \end{aligned}$$

where $\tilde{g}(\cdot)$ denotes the terminal feedback policy, specified by Assumptions 4.5 and 4.6, and feasibility follows as in the proof of Theorem 4.1. Given that

$$M_k - \alpha^k c(x_k, g_0^\star(\pi_k)) - \alpha^{k+N} c_N(x_{k+N}) + \alpha^{k+N} c(x_{k+N}, \tilde{g}(\pi_{k+N}))$$

is $\mathcal{F}_k$-measurable, we then have

$$\mathbb{E}_0[M_{k+1} \mid \mathcal{F}_k] \overset{\text{a.s.}}{\leq} M_k - \alpha^k c(x_k, g_0^\star(\pi_k)) - \alpha^{k+N} c_N(x_{k+N}) + \alpha^{k+N} c(x_{k+N}, \tilde{g}(\pi_{k+N})) + \alpha^{k+N+1} \mathbb{E}_0[c_N(f(x_{k+N}, \tilde{g}(\pi_{k+N}), w_{k+N})) \mid \mathcal{F}_k].$$

By Assumption 4.6, this yields

$$\mathbb{E}_0 [M_{k+1} \mid \mathcal{F}_k] \overset{\text{a.s.}}{\leq} M_k - \alpha^k c(x_k, g_0^\star(\pi_k)). \tag{4.20}$$

Taking expectations in (4.20) further gives

$$\mathbb{E}_0 [M_{k+1}] \leq \mathbb{E}_0 \left[M_k - \alpha^k c(x_k, g_0^\star(\pi_k)) \right],$$

where $\mathbb{E}_0 [M_0] < \infty$ via feasibility of π_0 for $\mathcal{P}(\cdot)$. By positivity of the stage cost, this yields

$$\sup_{k \in \mathbb{N}_0} \mathbb{E}_0 [|M_k|] < \infty. \tag{4.21}$$

Inequalities (4.20) and (4.21) with non-negativity of the stage cost show that M_k is a non-negative L^1-supermartingale on its filtration $\mathcal{F}_k$ and thus, by Doob's Martingale Convergence Theorem (see [52]), converges almost surely to a finite random variable,

$$M_k \overset{\text{a.s.}}{\to} M_\infty < \infty, \text{ as } k \to \infty. \tag{4.22}$$

Now define Z_k to be the discounted sample $\mathcal{P}_N$ cost-to-go plus the achieved MPC cost at time k,

$$Z_k \triangleq M_k + \sum_{j=0}^{k-1} \alpha^j c(x_j, g_0^\star(\pi_j)) \geq 0.$$

Then,

$$\mathbb{E}_0\left[Z_{k+1} \mid \mathcal{F}_k\right] \overset{\text{a.s.}}{\leq} M_k - \alpha^k c(x_k, g_0^\star(\pi_k)) + \sum_{j=0}^{k} \alpha^j c(x_j, g_0^\star(\pi_j)) = Z_k.$$

That is, recognizing that $Z_0 = M_0$ so that $\mathbb{E}_0[|Z_0|] < \infty$, Z_k also is a non-negative L^1-supermartingale and converges almost surely to a finite random variable

$$Z_k \overset{\text{a.s.}}{\to} Z_\infty < \infty, \text{ as } k \to \infty.$$

However, by definition of Z_k and (4.22), this implies (4.19). □

While the discount factor α may not seem to play a major role in this result, notice that small values of α may be required to satisfy Assumption 4.6. For $\alpha \geq 1$, (4.19) implies almost sure convergence to 0 of the achieved stage cost.

Assumption 4.7. State x is detectable via the stage cost:

$$c(x_k, u_k) \overset{\text{a.s.}}{\to} 0 \text{ as } k \to \infty \qquad \Longrightarrow \qquad x_k \overset{\text{a.s.}}{\to} \mathcal{X} \text{ as } k \to \infty.$$

Theorem 4.3. *Given Assumptions 4.3-4.7, SMPC with $\alpha \geq 1$ yields*

$$\lim_{M\to\infty} \sum_{k=0}^{M} c(x_k, g_0^\star(\pi_k)) \overset{a.s.}{<} \infty$$

and

$$x_k \overset{a.s.}{\to} \mathcal{X}, \text{ as } k \to \infty. \tag{4.23}$$

Proof. First proceed as in the proof of Theorem 4.2. By Doob's Decomposition Theorem (see [53]) on (4.22), there exists a martingale $\mathcal{M}_k$ and a decreasing sequence $\mathcal{A}_k$

such that $M_k = \mathcal{M}_k + \mathcal{A}_k$, where $A_k \to A_\infty$ a.s. by (4.22). Using this decomposition, (4.20) yields

$$c(x_k, g_0^\star(\pi_k)) \leq \alpha^k c(x_k, g_0^\star(\pi_k)) \overset{\text{a.s.}}{\leq}$$

$$M_k - \mathbb{E}_0\left[M_{k+1} \mid \mathcal{F}_k\right] = \mathcal{A}_k - \mathbb{E}_0\left[\mathcal{A}_{k+1} \mid \mathcal{F}_k\right] \overset{\text{a.s.}}{\leq} \mathcal{A}_k - \mathbb{E}_0\left[\mathcal{A}_\infty \mid \mathcal{F}_k\right].$$

Taking limits as $k \to \infty$ and re-invoking non-negativity of the stage cost then leads to $c(x_k, g_0^\star(\pi_k)) \to 0$ a.s., which by the detectability condition on the stage cost (Assumption 4.7) verifies (4.23). □

While (4.23) holds only for $\alpha \geq 1$, notice that SMPC for $\alpha \in [0, 1)$ with recursive feasibility possesses the default stability property (4.17). For zero terminal cost $c_N(x) \equiv 0$, Assumption 4.8 replaces Assumption 4.6 to guarantee (4.19), a finite discounted infinite-horizon SMPC cost.

Assumption 4.8. The terminal feedback policy $\tilde{g}(\pi)$ specified in Assumption 4.5 satisfies

$$c(x, \tilde{g}(\pi)) \overset{\text{a.s.}}{=} 0$$

for all densities π of x with $\pi \in C_N$.

Corollary 4.1. *Given Assumptions 4.3-4.5 and 4.8, SMPC with zero terminal cost* $c_N(x) \equiv 0$ *yields*

$$\lim_{M\to\infty} \sum_{k=0}^{M} \alpha^k c(x_k, g_0^\star(\pi_k)) \overset{a.s.}{<} \infty.$$

Moreover, if $\alpha = 1$ and Assumption 4.7 is added, we have

$$x_k \overset{a.s.}{\to} \mathcal{X}, \text{ as } k \to \infty.$$

4.6 Infinite-Horizon Performance Bounds

In the following, we establish performance bounds for SMPC, implemented on the infinite horizon as a proxy to solving the infinite-horizon stochastic optimal control problem $\mathcal{P}_\infty(\pi)$. These bounds are in the spirit of previously established bounds reported for deterministic MPC in [26] and the stochastic full state-feedback case in [28].

Assumption 4.9. There exist $\gamma \in [0, 1]$ and $\eta \in \mathbb{R}_+$ such that

$$\mathbb{E}_0 \left[V_0(T(\pi_0, y_1, g_0^\star(\pi_0))) - V_1(T(\pi_0, y_1, g_0^\star(\pi_0))) \right] \leq \gamma \, \mathbb{E}_0 \left[c(x_0, g_0^\star(\pi_0)) \right] + \eta \tag{4.24}$$

for all densities π_0 of x_0 which are feasible in $\mathcal{P}_N(\cdot)$.

Definition 4.8. Denote by $\mathbf{g}^{MPC}$ the SMPC implementation of policy $g_0^\star(\cdot)$ on the infinite horizon, i.e.

$$\mathbf{g}^{MPC} \triangleq \{g_0^\star, g_0^\star, g_0^\star, \ldots\}.$$

Similarly, $\mathbf{g}^{\star^{N-1}}$ and $\mathbf{g}^{\star^\infty}$ are the optimal sequences of policies in Problems $\mathcal{P}_N(\cdot)$ and $\mathcal{P}_\infty(\cdot)$, respectively.

Theorem 4.4. *Given Assumptions 4.3-4.5 and 4.9, SMPC with $\alpha \in [0, 1)$ yields*

$$(1 - \alpha\gamma) \, J_\infty(\pi_0, \mathbf{g}^{\star^\infty}) \leq (1 - \alpha\gamma) \, J_\infty(\pi_0, \mathbf{g}^{MPC}) \leq J_N(\pi_0, \mathbf{g}^{\star^{N-1}}) + \frac{\alpha}{1 - \alpha}\eta. \tag{4.25}$$

Proof. The optimal value function in the SDPE (4.15) satisfies $V_0(\pi_0) = J_N(\pi_0, \mathbf{g}^{\star^{N-1}})$, so that optimality of policy $g_0^\star(\cdot)$ in Problem $\mathcal{P}_N(\pi_0)$ implies

$$\begin{aligned} V_0(\pi_0) = \mathbb{E}_0 \left[c(x_0, g_0^\star(\pi_0)) + \alpha V_1(T(\pi_0, y_1, g_0^\star(\pi_0))) \right] \\ + \alpha \mathbb{E}_0 \left[V_0(T(\pi_0, y_1, g_0^\star(\pi_0))) \right] \\ - \alpha \mathbb{E}_0 \left[V_0(T(\pi_0, y_1, g_0^\star(\pi_0))) \right], \end{aligned}$$

which by Assumption 4.9 yields

$$(1 - \alpha\gamma)\mathbb{E}_0 \left[c(x_0, g_0^\star(\pi_0)) \right] \leq V_0(\pi_0) - \alpha \mathbb{E}_0 \left[V_0(T(\pi_0, y_1, g_0^\star(\pi_0))) \right] + \alpha\eta. \tag{4.26}$$

Now denote by $J_\infty^M(\pi_0, \mathbf{g}^{MPC})$ the first $M \in \mathbb{N}_1$ terms of the infinite-horizon cost $J_\infty(\pi_0, \mathbf{g}^{MPC})$ subject to the SMPC implementation of policy $g_0^*(\cdot)$. By (4.26), we have

$$\begin{aligned} (1 - \alpha\gamma) J_\infty^M(\pi_0, \mathbf{g}^{MPC}) = (1 - \alpha\gamma)\mathbb{E}_0 \left[\sum_{k=0}^{M-1} \alpha^k c(x_k, g_0^\star(\pi_k)) \right] \leq \\ \mathbb{E}_0 \Big[V_0(\pi_0) - \alpha V_0(\pi_1) + \alpha\eta + \alpha V_0(\pi_1) - \alpha^2 V_0(\pi_2) + \\ \alpha^2\eta + \ldots + \alpha^{M-1} V_0(\pi_{M-1}) - \alpha^M V_0(\pi_M) + \alpha^M \eta \Big], \end{aligned}$$

such that

$$(1 - \alpha\gamma) J_\infty^M(\pi_0, \mathbf{g}^{MPC}) \leq J_N(\pi_0, \mathbf{g}^{\star^{N-1}}) - \alpha^M \mathbb{E}_0 \left[J_N(\pi_M, \mathbf{g}^{\star^{N-1}}) \right] + \left(\alpha + \ldots + \alpha^M \right) \eta,$$

which by non-negativity of the stage cost confirms the right-hand inequality in (4.25) in the limit as $M \to \infty$. The left-hand inequality follows directly from optimality. □

In the special case $\gamma = 0$, we impose the following assumption on the terminal cost to obtain an insightful corollary to Theorem 4.4.

Assumption 4.10. For $\alpha \in [0, 1)$, there exists $\eta \in \mathbb{R}_+$ such that the terminal policy $\tilde{g}(\cdot)$ specified in Assumption 4.5 satisfies

$$\mathbb{E}_\pi \left[\alpha\, c_N(f(x, \tilde{g}(\pi), w)) - c_N(x)\right] \leq -\mathbb{E}_\pi \left[c(x, \tilde{g}(\pi))\right] + \frac{\eta}{\alpha^{N-1}},$$

for all densities π of x with $\pi \in C_N$. The expectation $\mathbb{E}_\pi[\cdot]$ is with respect to state x – with density π – and w.

Corollary 4.2. *Given Assumptions 4.3-4.5 and 4.10, SMPC with $\alpha \in [0, 1)$ yields*

$$J_\infty(\pi_0, \mathbf{g}^{\star^\infty}) \leq J_\infty(\pi_0, \mathbf{g}^{MPC}) \leq J_N(\pi_0, \mathbf{g}^{\star^{N-1}}) + \frac{\alpha}{1-\alpha}\eta.$$

Proof. For conditional densities π_1 of x_1 such that $\pi_1 \in C_1$, use optimality and subsequently Assumption 4.10 to conclude

$$\begin{aligned} V_0(\pi_1) - V_1(\pi_1) &= \mathbb{E}_1\Bigg[\left(\sum_{k=0}^{N-1} \alpha^k c(x_{k+1}, g_k^\star(\pi_{k+1})) + \alpha^N c_N(x_{N+1})\right) \\ &\qquad - \left(\sum_{k=0}^{N-2} \alpha^k c(x_{k+1}, g_{k+1}^\star(\pi_{k+1})) + \alpha^{N-1} c_N(x_N)\right)\Bigg] \\ &\leq \mathbb{E}_1[\alpha^{N-1} c(x_N, \tilde{g}(\pi_N)) + \alpha^N c_N(f(x_N, \tilde{g}(\pi_N), w_N)) - \alpha^{N-1} c_N(x_N)] \\ &\leq \eta, \end{aligned}$$

which by (4.17) implies $V_0(\pi_k) - V_1(\pi_k) \leq \eta$ for $k \in \mathbb{N}_1$. However, this means Assumption 4.9 is satisfied with $\gamma = 0$ and thus completes the proof by Theorem 4.4. □

This Corollary relates the following quantities: *design cost*, $J_N(\pi_0, \mathbf{g}^{*^{N-1}})$, which is known as part of the SMPC calculation, *optimal cost* $J_\infty(\pi_0, \mathbf{g}^{*^\infty})$ which is unknown (otherwise we would use $\mathbf{g}^{*^\infty}$), and unknown infinite-horizon SMPC *achieved cost* $J_\infty(\pi_0, \mathbf{g}^{MPC})$.

4.7 Discussion and Remarks

4.7.1 Interpretation of Results

SMPC is cast as a variant of stochastic optimal output feedback control (SOFC). This is a closed-loop calculation which does not involve the computation of an open-loop: control sequence, state sequence, or state density sequence. SMPC consists of two distinct pieces: the Bayesian filter recursion (4.5-4.6) yielding the conditional state density π_k and the optimal feedback policy, $g_0^\star(\cdot)$. This is reminiscent of an explicit implementation of MPC [35, 33], $u_k = \kappa(x_k)$, with two central distinctions. The (finite-dimensional) full-state measurement is replaced by the (infinite-dimensional) density π_k, and the explicit feedback law is computed as a closed-loop quantity via the SDPE (4.15) in place of open-loop criterion minimization. Conventional MPC is a memoryless nonlinear state-feedback controller, while SMPC has memory via π_k and, therefore, may include probing for state observability. This dual action of the control signal captured in $g_0^\star(\cdot)$ is inherent to stochastic optimal control. The presented SMPC algorithm is new, as are its performance and stability guarantees.

To our knowledge, the feasibility analysis of the stochastic optimal control problem via consideration of density classes is new. This permits treatment of recursive feasibility for SMPC. Notably, constraints – whether probabilistic or almost sure – require almost sure satisfaction with respect to the propagated information states. Performance guarantees of SMPC are made in comparison to performance of the infinite-horizon stochastically optimal controlled system and are presented in Theorem 4.4 and Corol-

lary 4.2. These results extend those of [28], which pertain to full state stochastic optimal feedback. Other examples of stochastic performance bounds are mostly restricted to linear systems and, while computable, do not relate to the optimal constrained control.

Linear gaussian systems collapse the SMPC algorithm into standard structure. The Bayesian filter is identical to the Kalman filter since the conditional state density is gaussian and therefore described by its conditional mean and covariance alone. The associated optimal control problem exhibits the Separation Principle of Wonham [54, 55], which establishes that, for these systems with not-necessarily quadratic cost functions, the optimal control is a function of the conditional mean of the information state. Further, with quadratic costs, the optimal feedback control function is linear feedback [56]. Separation in this case means that the optimal output feedback controller comprises the optimal (mean-square) estimator together with the optimal state feedback law. More generally, the Bayesian filter possesses no optimality property itself just as the feedback policy $g_0^\star$ in isolation is not *optimal*.

4.7.2 Analysis of Assumptions

The sequence of assumptions becomes more inscrutable as our study progresses. However, they deviate only slightly from standard assumptions in MPC, suitably tweaked for stochastic applications. Assumptions 4.1 and 4.2 are regularity conditions permitting the development of the Bayesian filter via densities and restricting the controls to causal policies. Assumptions 4.3 and 4.4 limit the constraint sets and initial state density to admit treatment of recursive feasibility.

Assumptions 4.5, 4.6, 4.8 and 4.10 each concerns a putative terminal control policy, $\tilde{g}(\cdot)$. Assumption 4.5 implies positive invariance of the terminal constraint set under $\tilde{g}$. Using the martingale analysis of the proof of Theorem 4.3, Assumption 4.6 ensures that the extant $\tilde{g}$ achieves finite cost-to-go on the terminal set. The cost-detectability Assumption 4.7 is familiar in Optimal Control to make the implication that finite cost forces state convergence. Assumption 4.8 temporarily replaces Assumption 4.6 only to consider the zero terminal cost case. Assumptions 4.9 and 4.10 presume monotonicity of the finite-horizon cost with increasing horizon, firstly for the optimal policy $g_0^\star$ and then for the putative terminal policy, $\tilde{g}$ on the terminal set. These monotonicity assumptions mirror those of, for example, [26] for deterministic MPC and [28] for full-state stochastic MPC. They underpin the deterministic Lyapunov analysis and the stochastic Martingale analysis based on the cost-to-go.

4.7.3 Duality in Optimal Control

Apart from linear systems, state observability properties depend on the control inputs [57, 58]. The following system is a distillation of a network congestion control problem studied in [58].

$$
\begin{aligned}
x_{k+1}^1 &= w_k^1 + u_k, & w_k^1 &\sim \mathcal{U}[0,1], \\
x_{k+1}^2 &= w_k^2, & w_t^2 &\sim \mathcal{U}[5,6], \\
y_k &= \max\left[x_k^1, x_k^2\right] + v_k.
\end{aligned}
$$

Evidently, there are control sequences, $\{u_t\}$, which render x_t^1 or x_t^2 unobservable from $\{u_t, y_t\}$. Control $u_t \equiv 0$ is one such sequence. Duality is a property of output feedback control with or without parameter estimation.

Dual optimal control connotes the compromise in optimal control between the control signal's function to reveal the state and its function to regulate the state. The dependence of the information state on the control signal is evident from (4.6). These dual actions are typically antagonistic [49], leading Florentin [59], Jacobs and Patchell [60], Bar-Shalom and Tse [61, 62] to introduce the notions of caution and probing in stochastic optimal control. The duality of stochastic optimal control is a generic feature, but there exist some circumstances where the probing nature of the control evanesces due to the specific case at hand. Fel'dbaum [63, 64] classifies such problems as *neutral*. The linear Gaussian stochastic optimal control problem is one instance given that in this case the conditional mean and variance of the state provide a sufficient statistic with respect to the conditional state density.

When system identification is combined with control, even in linear systems, the *state* includes both the plant state and the parameters. Identifiability of the system parameters depends on the signal excitation properties. A number of papers [65, 50, 11] seek to impose artificially excitation conditions on the system input signals as part of the constraints of MPC design. These are not optimal controls. For the stochastic optimal control problems considered here, excitation of the control signal is incorporated into the control, if necessary, as part of the optimization. The optimal control policy functions, $g_j^\star(\cdot)$, will inherently inject excitation into the control signal depending on the quality of

state knowledge as embodied in π_k. So these policies are complicated mappings of the information state, which explains the difficulty of solving the SDPE.

4.7.4 Other Variants of Stochastic MPC

In their recent book [40], Kouvaritakis and Cannon provide a thorough analysis of approaches to stochastic MPC, particularly in regard to linear systems. Their survey includes Tube MPC for probabilistic constraints and uncertain systems with specified state dimension. For stochastic output feedback control they adopt a *certainty equivalence* control approach where a (finite-dimensional) state estimate, $\hat{x}_k$, replaces the actual state in the MPC calculations. Traditionally, the conditional mean, $\mathbb{E}_{\pi_k}(x)$, is used.

Scenario methods deal with optimization of difficult, non-convex problems in which the initial task is recast as a parametrized collection of simpler convex problems. Robust control can be formulated in this fashion. Random sampling of the parameter is performed and the resulting collection or union of optimizations is solved. This sampling can occur by drawing from a known distribution or by repeated experiments. In the MPC context, this has been proposed by a number of authors including [66, 67, 68, 69, 70] with emphasis on stochastic dynamics and robustness with the random parameter varying with time. The focus has been on full state feedback for systems with linear dynamics and with probabilistic state constraints. The technical construction is to take a sufficient number of samples (scenarios) to provide an adequate reconstruction of future state distributions for design. These distribution are, however, inherently open-loop constructions.

From the perspective of this chapter, the scenario approach is best connected with a possible numerical implementation of the prediction step of the Bayesian filter (4.6). [The measurement update stage (4.5) can be approximated by a resampling process with the combined calculation yielding the Particle Filter.] It is interesting to note that scenarios here occur in the state conditional density calculation, while the scenario approach to MPC deals exclusively with full state feedback and uses sampling to compute control signals, not control policies. The material presented here is not captured by the scenario MPC literature because, in the latter approach, the densities of future (randomized) parameters depend in no way on the control signals, while this control duality is at the core of SMPC as in the present chapter. The focus on linear systems in the scenario literature is stronger evidence of this distinction.

A better comparison to the present work than the scenario approach to MPC are tube-based methods as discussed for instance in [9, 40, 29, 7]. These methods treat perturbed linear systems with noise-corrupted output measurements. The disturbance signals are typically bounded and may be either deterministic or stochastic in nature. The forward propagation of the system dynamics with disturbance yields tubes in the state space within which the actual states lie. Feedback control is then designed to operate across all the tubes in a receding horizon fashion. This results in a robustness to the actual disturbances. The achieved closed-loop performance of such tube MPC controlled systems can be calculated (or bounded) from the system information. The results extend to the stochastic case, especially with quadratic criteria.

These authors treat state estimate feedback via certainty equivalence for systems

with linear dynamics. Under certainty equivalence, the control signal is a function of a finite-dimensional statistic, usually the conditional mean, of the state, as opposed to maintaining a copy of the entire conditional density. This state estimate can be derived from a number of possible sources: Kalman filter, Luenberger observer, MHE, etc. Once one restricts attention to systems with linear dynamics, duality drops from the picture. Closed-loop output-feedback optimality of these methods is not considered.

4.8 Conclusions

The central contribution of the chapter is the presentation of an SMPC algorithm based on SOOFC. This yields a number of theoretical properties of the controlled system, some of which are simply recognized as the stochastic variants of results from deterministic full-state feedback MPC with their attendant assumptions, including for instance Theorem 4.1 for recursive feasibility. Theorem 4.2 is the main stability result in establishing the finiteness of the discounted cost of the SMPC-controlled system. Theorem 4.3 and Corollary 4.1 deal with consequent convergence of the state in special cases.

Performance guarantees of SMPC are made in comparison to performance of the infinite-horizon stochastically optimally controlled system and are presented in Theorem 4.4 and Corollary 4.2. These results extend those of [28], which pertain to full-state feedback stochastic optimal control and which therefore do not accommodate duality. Other examples of stochastic performance bounds are mostly restricted to linear systems

and, while computable, do not relate to the optimal constrained control. While the formal stochastic results are traceable to deterministic predecessors, the divergence from earlier work is also notable. This concentrates on the use of the information state to accommodate measurements and the exploration of control policy functionals stemming from the Stochastic Dynamic Programming Equation. The resulting output feedback control possesses duality and optimality properties which are either artificially imposed in or absent from earlier approaches.

We further suggest two potential strategies to ameliorate the computational intractability of the Bayesian filter and SDPE, famous for its *curse of dimensionality*. Firstly, one may use the Particle filter implementation of the Bayesian filter, which has many examples of fast execution for small state dimensions, which with a loss of duality can be combined with scenario methods. This approach is discussed in [42] as an approximation of the algorithm in this chapter. Secondly, we point out that our algorithm becomes computationally tractable for the special case of Partially Observable Markov Decision Processes (POMDPs), which may be used either to approximate a nonlinear model or to model a given system in the first place. This strategy inherits the dual nature of our SMPC algorithm for general nonlinear systems.

Acknowledgements

Chapter 4, in part, has been submitted for publication of the material as it may appear in: M.A. Sehr, R.R. Bitmead, “Stochastic Model Predictive Control: Output-

Feedback, Duality and Performance", *Automatica*. The dissertation author was the primary investigator and author of this paper.

Chapter 5

Particle Model Predictive Control: Tractable Stochastic Nonlinear Output-Feedback MPC

5.1 Introduction

Model Predictive Control (MPC), in its original formulation, is a full-state feedback law (see [1, 2, 3]). This underpins two theoretical limitations of MPC: accommodation of output-feedback, and extension to include a compelling robustness theory given the state dimension is fixed. This chapter addresses the first of these issues in a rather general, though practical setup.

There has been a number of approaches to output-feedback MPC, mostly hinging on the replacement of the measured true state by a state estimate, which is computed

via Kalman filtering (e.g. [38, 4]), moving-horizon estimator (e.g. [6, 5]), tube-based minimax estimators (e.g. [7]), etc. Apart from [6], these designs, often for linear systems, separate the estimator design from the control design. The control problem may be altered to accommodate the state estimation error by methods such as: constraint tightening as in [4], chance/probabilistic constraints as in [9] or [8], and so forth. Likewise, for nonlinear problems, where the state estimation behavior is affected by control signal properties, the control may be modified to enhance the excitation properties of the estimator, as suggested in [10, 11]. Each of these aspects of accommodation is made in an isolated fashion.

The stochastic nonlinear output-feedback MPC algorithm presented in this chapter is motivated by the structure of Stochastic Model Predictive Control (SMPC) via finite-horizon stochastic optimal control. The latter method requires propagating conditional state densities using a Bayesian Filter (BF) and solution of the Stochastic Dynamic Programming Equation (SDPE). By virtue of implementing a truly optimal finite-horizon control law in a receding horizon fashion, one can deduce a number of properties of the closed-loop dynamics, including recursive feasibility of the SMPC controller, stochastic stability and bounds characterizing closed-loop infinite-horizon performance, as discussed in [71].

Unfortunately, solving for the stochastic optimal output-feedback controller, even on the finite horizon, is computationally intractable except for special cases such as linear quadratic Gaussian MPC because of the need to solve the SDPE, which incorporates the duality of the optimal control law in its effect on state observability. While the

BF, required to propagate the conditional state densities, is readily approximated using a Particle Filter (PF), open-loop solution of the SDPE results in the loss of the duality of the optimal control. While not discussed in this chapter, this effect can be mitigated sub-optimally by imposing excitation requirements as in [10, 11].

Approximately propagating the conditional state densities by means of the PF naturally invites combination with the more recent advances in Scenario Model Predictive Control (SCMPC), as discussed for instance by [72, 66, 67, 68, 69, 70]. Scenario methods deal with optimization of difficult, non-convex problems in which the initial task is recast as a parametrized collection of simpler, generally convex problems. Random sampling of uncertain signals and parameters is performed and the resulting collection of deterministic problem instances is solved. The focus has been on full state feedback for systems with linear dynamics and probabilistic state constraints. The technical construction is to take a sufficient number of samples (scenarios) to provide an adequate reconstruction of future controlled state densities for design.

In contrast to solving the SDPE underlying the stochastic optimal control problem, the future controlled state densities in SCMPC are open-loop constructions. However, they present a natural fit combined with the particle-based conditional density approximations generated by the PF, where individual particles can be interpreted as scenarios from an estimation perspective. Moreover, while SCMPC is typically formulated in the linear case, the basic idea extends to the nonlinear case, albeit with the loss of many computation-saving features. In this chapter, we propose and discuss this output-feedback version of SCMPC combined with the PF, which we call Particle Model

Predictive Control (PMPC). Compared with the stochastic optimal output-feedback controller (computed via BF and SDPE), the PMPC controller is suboptimal in not accommodating future measurement updates and thereby losing both exact constraint violation probabilities along the horizon and the probing requirement inherent to stochastic optimal control. On the other hand, PMPC enables a generally applicable and, at least for small state dimensions, computationally tractable alternative for nonlinear stochastic output-feedback control.

The structure of the chapter is as follows. We briefly introduce the problem setup in Section 5.2 and SMPC in Section 5.3 and proceed by introducing the PMPC control algorithm based on its individual components and parameters in Section 5.4. After describing the algorithm and its correspondence to SMPC, we use a challenging scalar nonlinear example to demonstrate computational tractability and dependence of the proposed PMPC closed-loop behavior on a number of parameters in Section 5.5. The example features nonlinear state and measurement equations and probabilistic state constraints under significant measurement noise. Finally, we conclude with Section 5.6.

5.2 Stochastic Optimal Control – Setup

We consider receding horizon output-feedback control for nonlinear stochastic systems of the form

$$x_{t+1} = f(x_t, u_t, w_t), \quad x_0 \in \mathbb{R}^n, \tag{5.1}$$

$$y_t = h(x_t, v_t), \tag{5.2}$$

starting from known initial state probability density function, $\pi_{0|-1} = \text{pdf}(x_0)$. To this end, we denote the data available at time t by

$$\zeta^t \triangleq \{y_0, u_0, y_1, u_1, \ldots, u_{t-1}, y_t\}, \qquad \zeta^0 \triangleq \{y_0\}.$$

The *information state,* denoted π_t, is the conditional density of state x_t given data ζ^t.

$$\pi_t \triangleq \text{pdf}\left(x_t \mid \zeta^t\right). \tag{5.3}$$

We further impose the following standing assumption on the random variables and control inputs.

Assumption 5.1. The signals in (5.1-5.2) satisfy:

1. $\{w_t\}$ and $\{v_t\}$ are sequences of independent and identically distributed random variables.

2. x_0, w_t, v_l are mutually independent for all $t, l \geq 0$.

3. The control input u_t at time instant $t \geq 0$ is a function of the data ζ^t and given initial state density $\pi_{0|-1}$.

Denote by $\mathbb{E}_t[\,\cdot\,]$ and $\mathbb{P}_t[\,\cdot\,]$ the conditional expected value and probability with respect to state x_t – with conditional density π_t – and random variables $\{(w_k, v_{k+1}) : k \geq t\}$, respectively, and by ϵ_k the constraint violation level of constraint $x_k \in \mathbb{X}_k$. Our goal is to solve the *finite-horizon stochastic optimal control problem* (FHSOCP)

$$\mathcal{P}_N(\pi_t) : \begin{cases} \inf_{u_t,\ldots,u_{t+N-1}} & \mathbb{E}_t\left[\sum_{k=t}^{t+N-1} c(x_k, u_k) + c_N(x_{t+N})\right], \\ \quad \text{s.t.} & x_{k+1} = f(x_k, u_k, w_k), \\ & x_t \sim \pi_t, \\ & \mathbb{P}_{k+1}\left[x_{k+1} \in \mathbb{X}_{k+1}\right] \geq 1 - \epsilon_{k+1}, \\ & u_k \in \mathbb{U}_k, \\ & k = t, \ldots, t+N-1. \end{cases}$$

In theory, solving the FHSOCP at each time t and subsequently implementing the first control in a receding horizon fashion leads to a number of desirable closed-loop properties, as discussed in [71]. However, solving the FHSOCP is computationally intractable in practice, a fact that has led to a number of approaches in MPC for nonlinear stochastic dynamics. We propose a novel strategy that is oriented at the structure of SMPC based on the FHSOCP, but numerically tractable at least for low state dimensions.

As a result of the Markovian state equation (5.1) and measurement equation (5.2), the optimal control inputs in the FHSOCP must inherently be *separated* feedback policies (e.g. [47, 43]). That is, control input u_t depends on the available data ζ^t and initial density $\pi_{0|-1}$ solely through the current information state, π_t. Optimality thus

requires propagating π_t and policies g_t, where

$$u_t = g_t(\pi_t). \tag{5.4}$$

Motivated by this two-component separated structure of stochastic optimal output-feedback control, we propose an extension of the SCMPC approach to nonlinear systems, merged with a numerical approximation of the information state update via particle filtering. Before proceeding with this novel approach, we briefly revisit the two components of SMPC via solution of the FHSOCP.

5.3 Stochastic Model Predictive Control

The information state is propagated via the *Bayesian Filter* (see e.g. [45, 46]):

$$\pi_t = \frac{\text{pdf}(y_t \mid x_t)\, \pi_{t|t-1}}{\int \text{pdf}(y_t \mid x_t)\, \pi_{t|t-1}\, dx_t}, \tag{5.5}$$

$$\pi_{t+1|t} \triangleq \int \text{pdf}(x_{t+1} \mid x_t, u_t)\, \pi_t\, dx_t, \tag{5.6}$$

for $t \in \{0, 1, 2, \ldots\}$ and initial density $\pi_{0|-1}$. The recursion (5.5-5.6) has the following features:

- The *measurement update* (5.5) combines the *a priori* conditional density, $\pi_{t|t-1}$, and $\text{pdf}(y_t \mid x_t)$, derived from (5.2) using knowledge of: the function $h(\cdot, \cdot)$, the density of v_t, and the value of y_t.

- The *time update* (5.6) combines π_t and $\text{pdf}(x_{t+1}|x_t, u_t)$, derived from (5.1) using knowledge of: control input u_t, function $f(\cdot, \cdot, \cdot)$, and the density of w_t.

- For linear Gaussian systems, the filter recursion (5.5-5.6) reduces to the well-known Kalman Filter.

Combined with solution of the FHSOCP, this leads to the following SMPC algorithm, as discussed in [71].

Algorithm 2 Stochastic Model Predictive Control

1: **Offline:**
2: Solve $\mathcal{P}_N(\cdot)$ for the first optimal policy, $g_0^\star(\cdot)$.
3: **Online:**
4: **for** $t = 0, 1, 2, \ldots$ **do**
5: Measure y_t
6: Compute π_t
7: Apply first optimal control policy, $u_t = g_0^\star(\pi_t)$
8: Compute $\pi_{t+1|t}$
9: **end for**

Notice how this algorithm differs from common practice in stochastic model predictive control in that it explicitly uses the information states π_t. Throughout the literature, these information states – conditional densities – are commonly replaced by state estimates. While this makes the problem more tractable, one no longer solves the underlying stochastic optimal control problem.The central divergence however lies in Step 2 of the algorithm, in which the SDPE is presumed solved offline for the optimal feedback policies, $g_t(\pi_t)$, from (5.4). This is an extraordinarily difficult proposition in many cases but captures the optimality, and hence duality, as a closed-loop feedback

control law. The complexity of this step lies not only in computing a vector functional but also in the internal propagation of the information state within the SDPE.

5.4 Tractable Nonlinear Output-Feedback Model Predictive Control

In this section, we motivate a novel approach to output-feedback MPC that maintains the separated structure of SMPC while being numerically tractable for modest problem size.

5.4.1 Approximate Information State & Particle Filter

The BF (5.5-5.6) propagates the information state π_t to implement a necessarily separated stochastic optimal output-feedback control law. While implementing this recursion precisely is possible only in special cases such as linear Gaussian systems, where the densities can be finitely parametrized, the BF can be implemented approximately by means of the Particle Filter, with the approximation improving with the number of particles, as described for instance in [46]. In parallel with the BF, the PF consists of two parts: the forward propagation of the state density, and the resampling of the density using the next measurement.

The following algorithm describes a version of the PF amenable to PMPC in the context of this chapter. This is a slightly modified version of the filter design described

by [46].

Algorithm 3 Particle Filter (PF)

1: Sample N_p particles, $\{x^-_{0,p},\ p = 1, \ldots, N_p\}$, from density $\pi_{0|-1}$.

2: **for** $t = 0, 1, 2, \ldots$ **do**

3: Measure y_k.

4: Compute the relative likelihood q_p of each particle $x^-_{t,p}$ conditioned on the measurement y_t by evaluating $\text{pdf}(y_t \mid x^-_{t,p})$ based on (5.2) and $\text{pdf}(v_t)$.

5: Normalize $q_p \to q_p / \sum_{p=1}^{N_p} q_p$.

6: Sample N_p particles, $x^+_{t,p}$, via *resampling* based on the relative likelihoods q_p.

7: Given u_t, propagate $x^-_{t+1,p} = f(x^+_{t,p}, u_t, w_{t,p})$, where $w_{t,p}$ is generated based on $\text{pdf}(w_t)$.

8: **end for**

While a number of variations – such as roughening of the particles and differing *resampling* strategies, including importance sampling – of this basic algorithm may be sensible depending on the system at hand, this basic algorithm suffices in presenting a numerical method of approximating the Bayesian Filter to arbitrary degree of accuracy with increasing number of particles N_p (see e.g. [73]). For a more detailed discussion on the PF for use in state-estimate feedback control, see [74].

5.4.2 Scenario MPC and Particle Model Predictive Control

The Scenario Approach to MPC (e.g. [66, 67, 68, 69, 70]) commences from state x_t or state estimate, $\hat{x}_{t|t}$. It propagates, i.e. simulates, an open-loop controlled stochastic

system with sampled process noise density pdf(w_t). These propagated samples are then used to evaluate controls for constraint satisfaction and for open-loop optimality with probabilities tied to the sampled w_t densities. In many regards, this is congruent to repeated forward propagation of the PF via (5.6) without measurement update (5.5) and commencing from a singular density at x_t or $\hat{x}_{t|t}$. Particle MPC simply replaces the starting point, $\hat{x}_{t|t}$, by the collection of particles $\{x^+_{t,p}, p = 1, \ldots .N_p\}$ distributed as π_t, as illustrated in Figure 5.1.

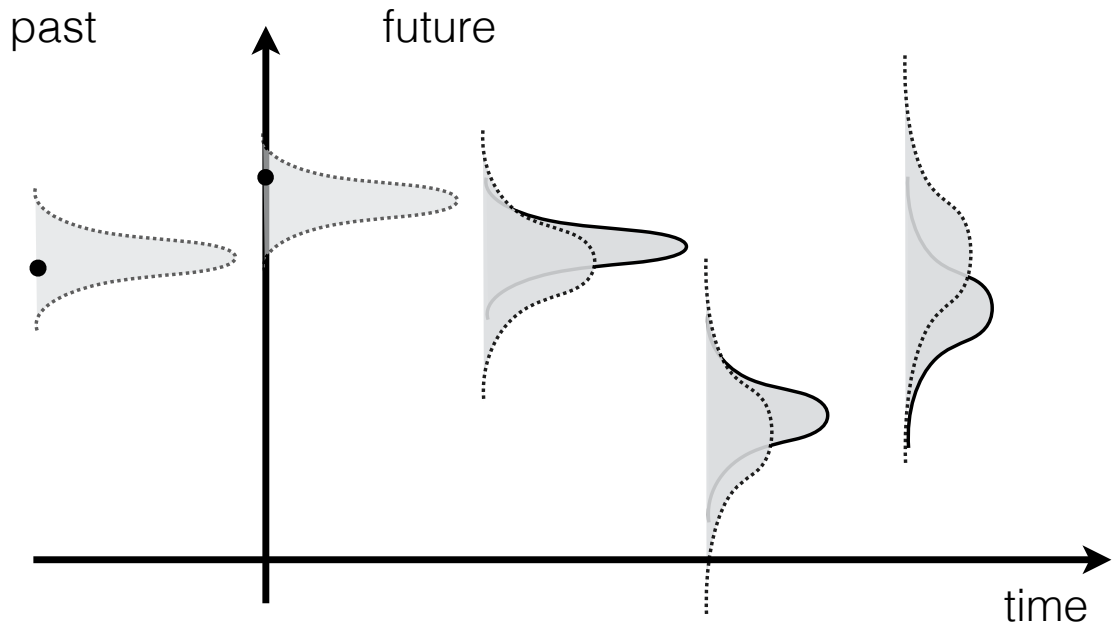

Figure 5.1: State density evolution in: Scenario MPC calculations (dots and solid outlines) and, Particle MPC (dashed outlines), for three steps into the future.

Before introducing the PMPC algorithm, we define a sampled, particle version

of the FHSOCP, with N_s scenarios and N_p available *a posteriori* particles at time t,

$$\tilde{\mathcal{P}}_N(\{x_{t,p}^+, p = 1, \ldots, N_p\}) : \begin{cases} \inf_{u_t, \ldots, u_{t+N-1}} & \sum_{s=0}^{N_s} \left(\sum_{k=t}^{t+N-1} c(x_{k,s}, u_k) + c_N(x_{t+N,s}) \right), \\ \text{s.t.} & x_{k+1,s} = f(x_{k,s}, u_k, w_{k,s}), \\ & x_{t,s} \in \{x_{t,p}^+, p = 1, \ldots, N_p\}, \\ & \tilde{\mathbb{P}}_{k+1} \left[x_{k+1} \in \mathbb{X}_{k+1} \right] \geq 1 - \epsilon_{k+1}, \\ & u_k \in \mathbb{U}_k, \\ & s = 1, \ldots, N_s, \quad k = t, \ldots, t + N - 1, \end{cases}$$

where statement

$$\tilde{\mathbb{P}}_{k+1} \left[x_{k+1} \in \mathbb{X}_{k+1} \right] \geq 1 - \epsilon_{k+1}$$

means that $x_{k+1,s} \in \mathbb{X}_{k+1}$ for at least $(1 - \epsilon_{k+1})N_s$ scenarios. Following the approach in [75], one may also choose to replace this constraint by $x_{k+1} \in \mathbb{X}_{k+1}$ and select the number of scenarios N_s according to the desired constraint violation levels ϵ_{k+1}. We are now in position to formulate the PMPC algorithm following the schematic in Figure 5.1.

Algorithm 4 Particle Model Predictive Control (PMPC)

1: Generate N_p *a priori* particles, $x^-_{0,p}$, based on $\pi_{0|-1}$.

2: **for** $t = 0, 1, 2, \ldots$ **do**

3: Measure y_t.

4: Compute the relative likelihood q_p of each particle $x^-_{t,p}$ conditioned on the measurement y_t by evaluating $\text{pdf}(y_t \mid x^-_{t,p})$ based on (5.2) and $\text{pdf}(v_t)$.

5: Normalize $q_p \to q_p / \sum_{p=1}^{N_p} q_p$.

6: Generate N_p *a posteriori* particles, $x^+_{t,p}$, via *resampling* based on the relative likelihoods q_p.

7: Solve $\tilde{\mathcal{P}}_N(\{x^+_{t,p}, p = 1, \ldots, N_p\})$ for the optimal scenario control values $u^\star_t, \ldots, u^\star_{t+N-1}$.

8: Given $u^\star_t$, propagate $x^-_{t+1,p} = f(x^+_{t,p}, u^\star_t, w_{t,p})$, where $w_{t,p}$ is generated based on $\text{pdf}(w_t)$.

9: **end for**

5.4.3 Computational Demand

Computational tractability of PMPC deteriorates with increasing: number of particles; number of scenarios; system dimensions; control signal grid spacing; MPC horizon. While the number of particles required for satisfactory performance of the PF grows exponentially with the state dimension (e.g. [76]), it is unclear how to select an appropriate number of scenarios in the nonlinear case. Suppose the state and input dimensions are n and m and that the numbers of particles and scenarios are chosen as

$N_p = P^n$ and $N_s = S^n$ for positive integers P and S, respectively, and that the MPC horizon is N. Further assuming a grid of U^m points in the control space and brute-force evaluation of all possible sequences, the order of growth for PMPC is approximately $O(P^n + S^n U^{mN})$.

Notice that the computational demand associated with the conditional density approximation in PMPC is additive in terms of the overall computational demand. This indicates that, provided the PF is computationally tractable for given state dimensions, tractability of PMPC is roughly equivalent to tractability of standard state-feedback SCMPC. In the example below, we found that scenario optimization tends to be the computational bottleneck at least for low system dimensions. Clearly, this observation holds only when the scenario optimization is performed by explicit enumeration of all feasible sequences over a grid in the control space, which may be avoided for particular problem instances. But the experience also confirms that in the nonlinear case the open- or closed-loop control calculation dominates the computational burden in comparison to state estimation.

5.5 Numerical Example

Consider the scalar, nominally unstable nonlinear system

$$x_{t+1} = 1.5\, x_t + \operatorname{atan}\left((x_t - 1)^2\right) u_t + w_t,$$

$$y_t = x_t^3 - x_t + v_t,$$

where x_0, w_t and v_l are mutually independent random variables for all $t, l \geq 0$ and

$$x_0 \sim \mathcal{U}(1,2), \qquad w_t \sim \mathcal{U}(-2,2), \qquad v_t \sim \mathcal{N}(0,5),$$

for all $t \geq 0$. We aim to minimize the quadratic cost function

$$J_N(\pi_t, u_t, \ldots, u_{t+N-1}) = \mathbb{E}_t\left[\sum_{k=t}^{t+N-1}\left(100\, x_k^2 + u_k^2\right) + 100\, x_{t+N}^2\right],$$

while satisfying the constraints

$$\mathbb{P}_{k+1}[x_{k+1} \geq 1] \geq 0.9, \qquad -5 \leq u_k \leq 5,$$

along the control horizon N, that is $k \in \{t, \ldots, t+N-1\}$ for $t \geq 0$. Notice how this system has both limited observability and controllability close to the constraint but infeasible unconstrained optimal states. In combination with the very noisy measurements, this is a challenging control problem. To implement PMPC as described in Section 5.4 for this nonlinear stochastic output-feedback control problem, we further restrict the control inputs to integer values, such that $u_t \in \{-5, -4, \ldots, 4, 5\}$. Figures 5.2- 5.5 display simulated closed-loop state trajectories, control values and measurement values for four PMPC controllers with differing parameters subject to the same realizations of process and measurement noise, respectively.

Figure 5.2 displays closed-loop simulation results under PMPC with horizon $N = 3$, $N_p = 5,000$ particles and $N_s = 1,000$ scenarios. While the poor observability properties of the system show close to the probabilistic constraint, it is satisfied at all times in this simulation. This is still the case when decreasing the number of particles to $N_p = 100$ in the simulation displayed Figure 5.3. However, we see how in this case,

the decreased accuracy of the PF leads to larger state-values in closed-loop. Similar behavior is observed in Figure 5.4 when reducing the number of scenarios to $N_s = 50$. Additionally, the controller violates the probabilistic constraint 3 times in this case. This trend continues when reducing the horizon to $N = 2$, as displayed in Figure 5.5.

5.6 Conclusion

We presented PMPC as a novel approach to output-feedback control of stochastic nonlinear systems. Generating scenarios not only from the distribution of the process noise but also from the particles of the Particle Filter, PMPC combines the benefits of the Particle Filter and Scenario MPC in a natural fit, allowing for a numerically tractable version of stochastic MPC with general nonlinear dynamics, cost and probabilistic constraints. Given a particular system instance, the algorithm and its properties may be adapted to exploit specific problem structure. Such extensions include: sub-optimal probing via additional constraints; scenario removal; provable closed-loop properties such as constraint satisfaction with specified confidence levels; optimization over parametrized policies.

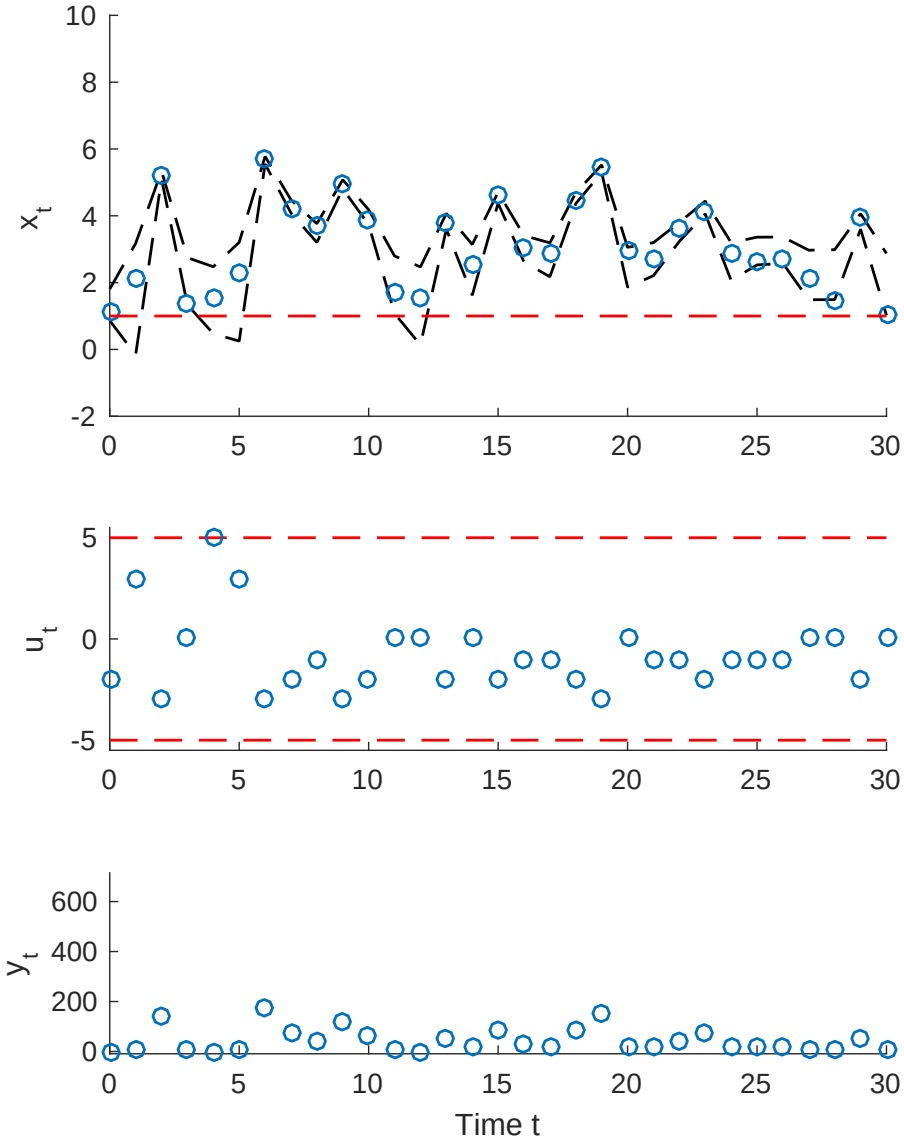

Figure 5.2: Simulation data for example in Section 5.5 over 30 samples, running PMPC with control horizon $N = 3$, number of particles $N_p = 5,000$ and number of scenarios $N_s = 1,000$. State, control and measurement values (blue), probabilistic and hard constraints (red), 95% confidence interval of PF (black).

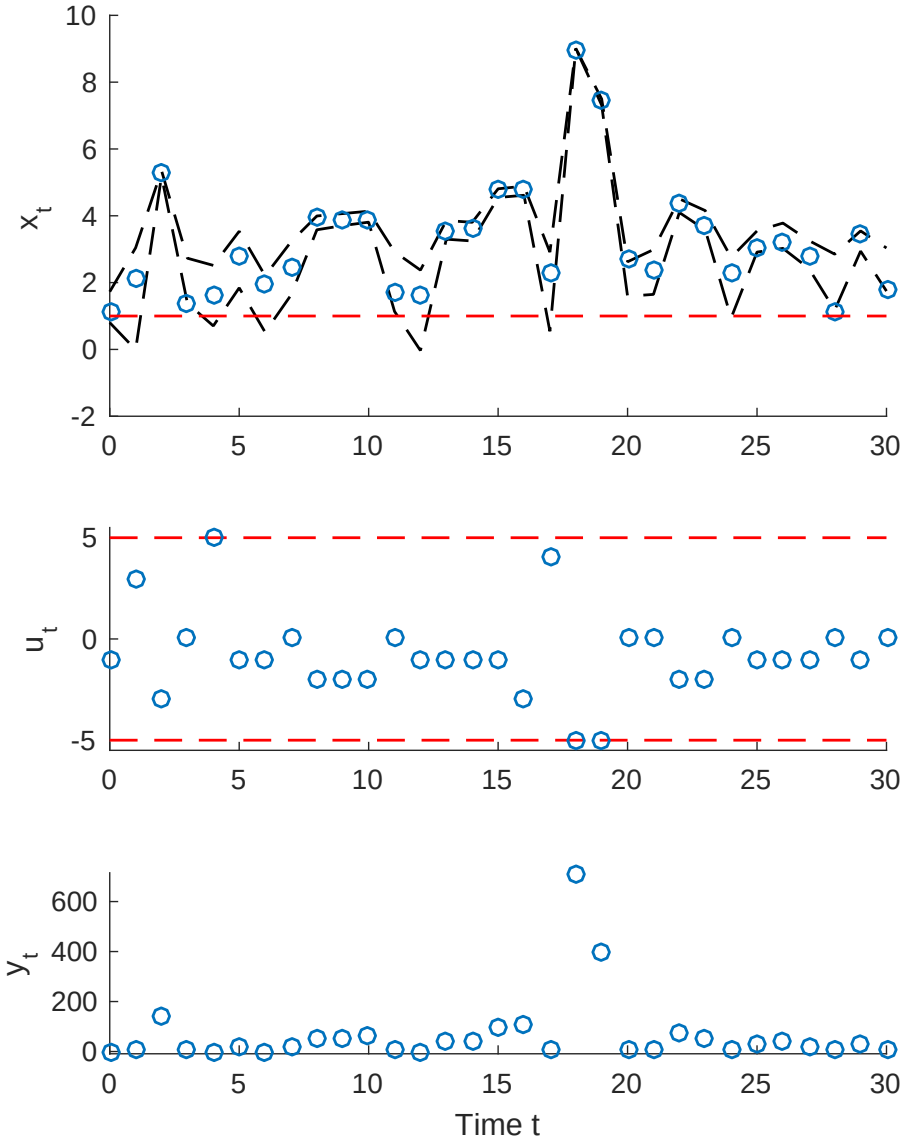

Figure 5.3: Simulation data for example in Section 5.5 over 30 samples, running PMPC with control horizon $N = 3$, number of particles $N_p = 100$ and number of scenarios $N_s = 1,000$. State, control and measurement values (blue), probabilistic and hard constraints (red), 95% confidence interval of PF (black).

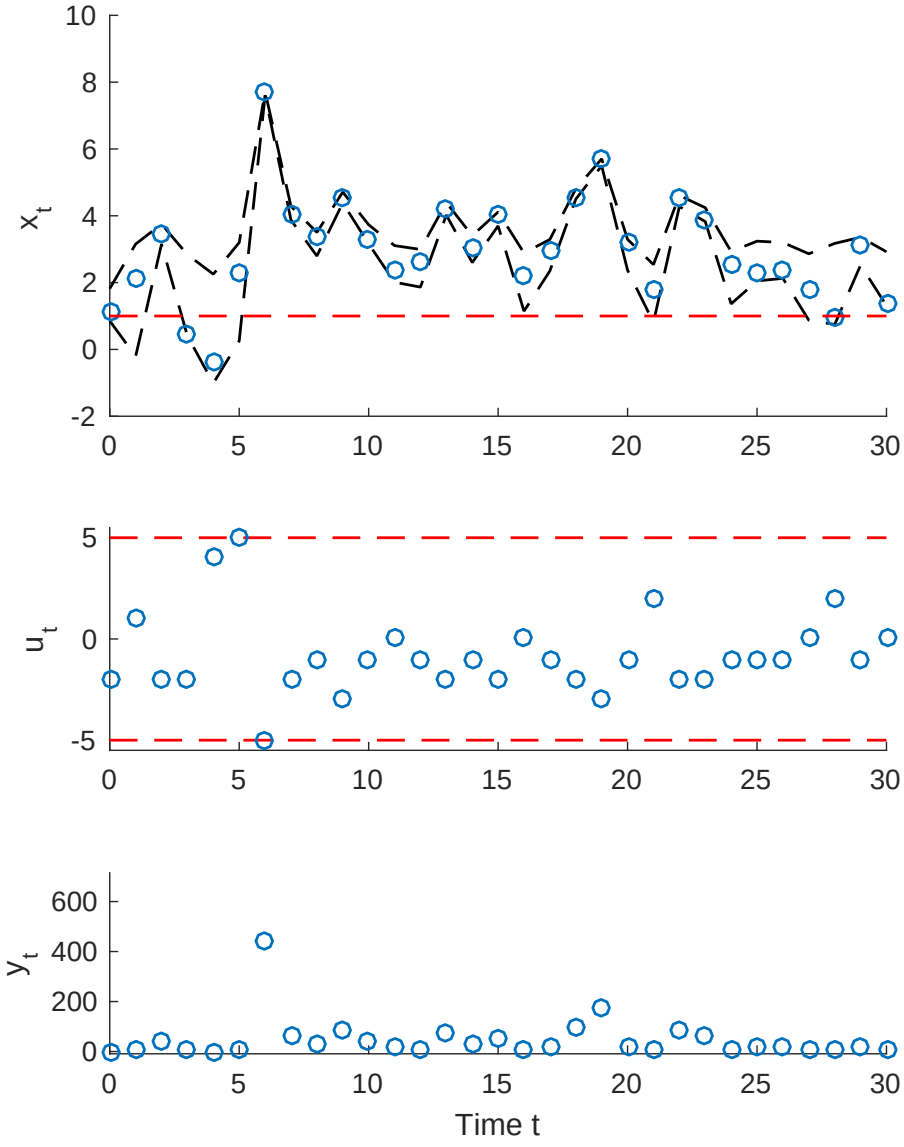

Figure 5.4: Simulation data for example in Section 5.5 over 30 samples, running PMPC with control horizon $N = 3$, number of particles $N_p = 5,000$ and number of scenarios $N_s = 50$. State, control and measurement values (blue), probabilistic and hard constraints (red), 95% confidence interval of PF (black).

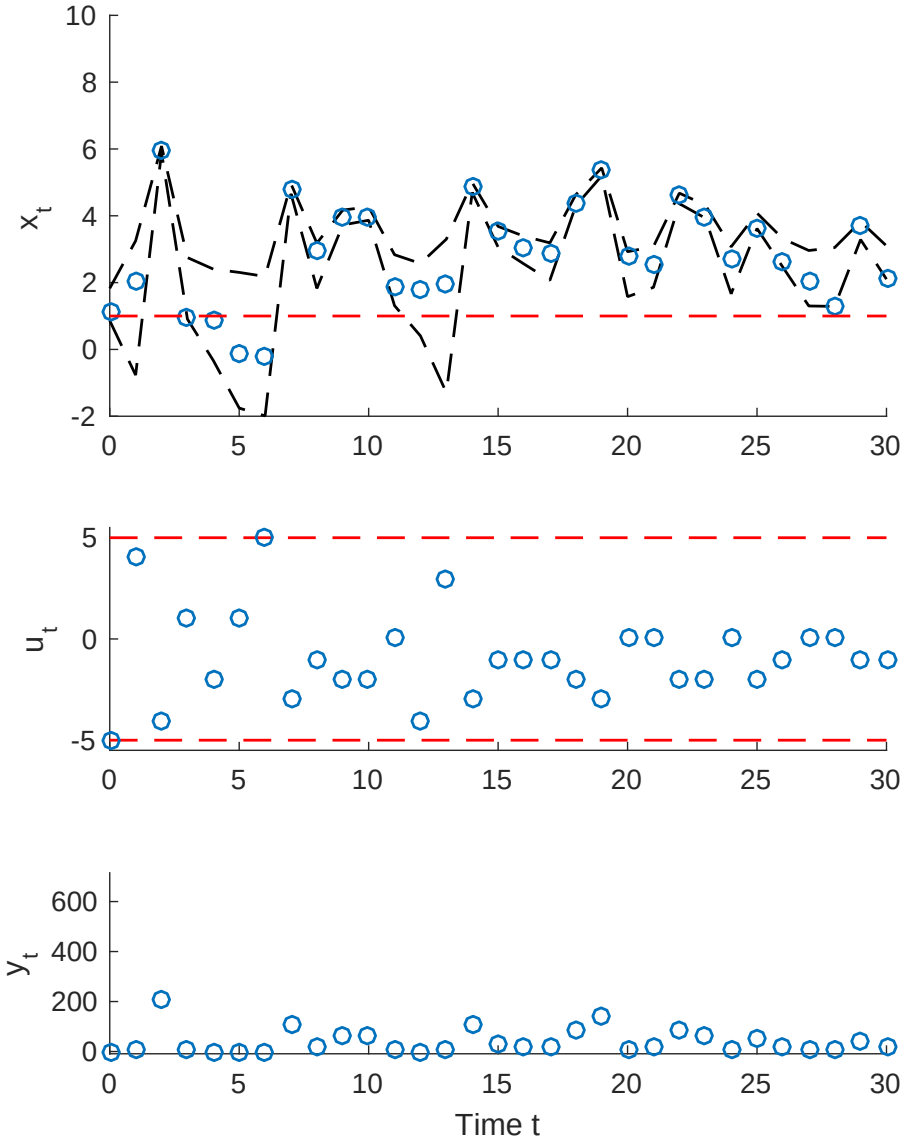

Figure 5.5: Simulation data for example in Section 5.5 over 30 samples, running PMPC with control horizon $N = 2$, number of particles $N_p = 5,000$ and number of scenarios $N_s = 1,000$. State, control and measurement values (blue), probabilistic and hard constraints (red), 95% confidence interval of PF (black).

Acknowledgements

Chapter 5, in full, is a reprint of the material as it will appear in: M.A. Sehr, R.R. Bitmead, "Particle Model Predictive Control: Tractable Stochastic Nonlinear Output-Feedback MPC", *Proc. IFAC World Congress*, 2017. The dissertation author was the primary investigator and author of this paper.

Chapter 6

Performance of Model Predictive Control of POMDPs

6.1 Introduction

Model Predictive Control (MPC) is well applied and popular because of its capacity to handle constraints and its simple formulation as an open-loop finite-horizon optimization problem evaluated on the receding horizon [1, 2]. There are a few areas in which MPC is wanting for more complete results, notably in the area of output feedback control and the associated requirement to manage the duality of the control signal in stochastic MPC (SMPC) problems. When SMPC is developed as a logical extension of finite-horizon Stochastic Optimal Control, which demands computation of closed-loop policies, it inherits the computational intractability of this latter subject via the inclusion of the Bayesian filter, required to propagate the conditional state densities,

and the stochastic dynamic programming equation.

Results exist relating the infinite-horizon performance of MPC to both the optimal performance and the performance computed as part of the finite-horizon optimization. These performance bounds are available in both the deterministic [26] and the stochastic [71] settings, were one ever able to solve the underlying finite-horizon stochastic problem computationally. While approximation of SMPC based on Stochastic Optimal Control via more tractable surrogate problems is possible, such as for instance in [5, 7, 72, 42], one generally loses the associated closed-loop guarantees, in particular regarding infinite-horizon performance of the generated control laws.

In this chapter, we derive new performance results for SMPC of systems described by Partially Observable Markov Decision Processes (POMDPs, see e.g. [77, 78]). POMDP system models of small to moderate dimensions admit tractable computation of finite-horizon stochastic optimal control laws while preserving the control signal duality, and so are attractive propositions with which to approach implementable SMPC [79, 80]. In deriving perfomance bounds for this specific class of problems, we examine their relation to the deterministic and stochastic continuous-state results, highlighting the role of value function monotonicity with horizon. All theorems discussed in this chapter exhibit the same conceptual structure displayed in Figure 6.1.

While the capability of handling constraints is a raison-d'être for MPC, constraints complicate this analysis and add little to the discussion about closed-loop cost. Thus, as in most of [26], we omit the explicit consideration in this chapter and point out that constraints may be reinserted subject to recursive feasibility assumptions.

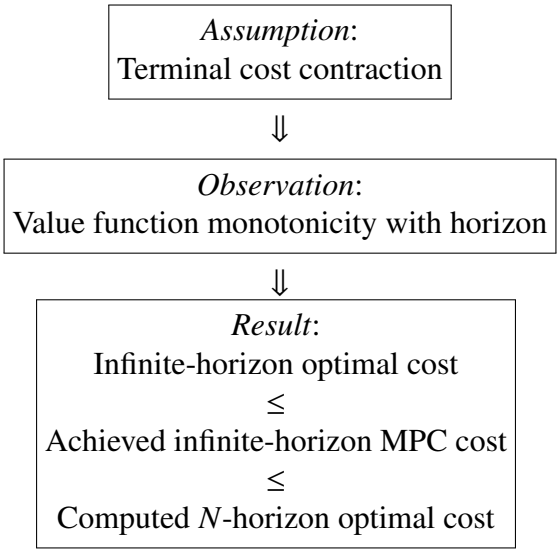

Figure 6.1: Conceptual structure of the results in Chapter 6.

The chapter is organized as follows. We revisit a particular infinite-horizon performance result from [26] in Section 6.2. We then proceed by reviewing a stochastic counterpart to this result, derived in [71], which we extend to receding horizon control of POMDPs in Section 6.4. A specific POMDP example from healthcare is studied in Section 6.5 to demonstrate numerically the satisfaction of assumptions, interpret control duality, and evaluate performance bounds on the infinite control horizon. The example, introduced in [80], displays in particular the dual nature of SMPC based on Stochastic Optimal Control.

6.2 Deterministic Model Predictive Control

This section revisits a performance result for deterministic MPC from [26], which we extend to SMPC for nonlinear systems (see also [71]) and POMDPs below.

Consider the nonlinear dynamic system

$$x_{t+1} = f(x_t, u_t),$$

where $x_t \in X$ and $u_t \in U$ for $t \in \mathbb{N}_0 \triangleq \{0, 1, 2, \ldots\}$ and metric spaces X, U. Further define the space of control sequences $u : \mathbb{N}_0 \to U$ as $\mathcal{U}$. In principle, we aim to find control policy $\mu : X \to U$ that minimizes the infinite-horizon cost functional

$$J_\infty(x_0, u) \triangleq \sum_{k=0}^{\infty} c(x_k, u_k), \tag{6.1}$$

where $c : X \times U \to \mathbb{R}_+$ is the stage cost. We define the optimal value function associated with cost (6.1) as

$$J_\infty^\star(x_0) \triangleq \inf_u J_\infty(x_0, u)$$

Given that solution of this infinite-horizon optimal control problem, even in the deterministic case, is usually intractable, a popular approach is to replace (6.1) by a finite-horizon optimal control problem over horizon $N \in \mathbb{N}_0$, with cost functional

$$J_N(x_0, u) \triangleq \sum_{k=0}^{N-1} c(x_k, u_k) + c_N(x_N), \tag{6.2}$$

where $c_N : X \to \mathbb{R}_+$ denotes an optional terminal cost term. The optimal value function corresponding to (6.2) is defined by

$$J_N^\star(x_0) \triangleq \inf_u J_N(x_0, u). \tag{6.3}$$

We further denote the sequence of optimal control policies in this finite-horizon problem by μ^N, with first control policy $\mu_0^N : X \to U$, which is implemented repeatedly in an

MPC law, denoted by

$$\mu^N_{\text{MPC}} \triangleq \{\mu^N_0, \mu^N_0, \ldots\}.$$

We now aim to provide computational estimates of the infinite-horizon achieved MPC cost $J_\infty(x_0, \mu^N_{\text{MPC}})$ in relation to the computed finite-horizon optimal cost $J_N(x, \mu^N)$. This goal can be achieved, for instance, by using the following assumption.

Assumption 6.1. For all $x \in X$, there exists $u \in U$ such that

$$c_N(f(x, u)) \leq c_N(x) - c(x, u).$$

This assumption on the terminal cost c_N in (6.2) then leads to the following performance guarantee.

Theorem 6.1 (Performance of deterministic MPC [26])**.** *Given Assumption 6.1, the inequality*

$$J^\star_\infty(x) \leq J_\infty(x, \mu^N_{MPC}) \leq J^\star_N(x)$$

holds for all $x \in X$.

This result, which is a special case of Theorem 6.2 in [26], allows us to provide bounds on the achieved infinite-horizon performance of the closed-loop system when choosing the terminal cost, c_N, as a Lyapunov function. This result is particularly useful because we compute the upper bound implicitly when generating our MPC control law, μ^N_{MPC}. Theorem 6.1 follows given that Assumption 6.1 implies that the underlying finite-horizon optimal value function $J^\star_N(x)$ is monotonically non-increasing with increasing

control horizon N. Notice, further, how this result not only provides infinite- but also finite-horizon closed-loop performance guarantees. This follows simply by

$$J_\infty^M(x, \mu_{\mathrm{MPC}}^N) \leq J_\infty(x, \mu_{\mathrm{MPC}}^N),$$

for all $M \in \mathbb{N}_0$ and $x \in X$, where

$$J_\infty^M(x_0, u) \triangleq \sum_{k=0}^{M} c(x_k, u_k).$$

The stochastic extension of this observation is of interest in particular for applications such as the healthcare example provided in Section 6.5 below, where infinite-horizon performance may not be of particular interest given the inherent finite-horizon nature of the control problem. We next provide results of similar quality to Theorem 6.1 for SMPC and in particular SMPC applied to POMDPs in the following sections.

6.3 Stochastic Model Predictive Control

We next discuss closed-loop performance of SMPC as in [71]. Committing a slight abuse of notation, we shall recycle most of the symbols used previously in Section 6.2 above. Consider nonlinear stochastic systems of the form

$$x_{t+1} = f(x_t, u_t, w_t), \tag{6.4}$$

$$y_t = h(x_t, v_t), \tag{6.5}$$

where $x_t \in X$, $u_t \in U$, $y_t \in Y$ for $t \in \mathbb{N}_0$ and metric spaces X, U, Y, respectively. Starting from known initial state density $\pi_{0|-1} = \mathrm{pdf}(x_0)$, we denote the data available at time t

by

$$\zeta^t \triangleq \{y_0, u_0, y_1, u_1, \ldots, u_{t-1}, y_t\}, \qquad \zeta^0 \triangleq \{y_0\}.$$

We further impose the following standing assumption on the random variables and control inputs.

Assumption 6.2. The signals in (6.4-6.5) satisfy:

1. w_t and v_t are i.i.d. sequences with known densities.
2. x_0, w_t, v_l are mutually independent for all $t, l \in \mathbb{N}_0$.
3. The control input u_t at time instant $t \in \mathbb{N}_0$ is a function of the data ζ^t and given initial state density $\pi_{0|-1}$.

The *information state,* denoted π_t, is the conditional probability density function of state x_t given data ζ^t,

$$\pi_t \triangleq \text{pdf}\left(x_t \mid \zeta^t\right).$$

As a result of the Markovian dynamics (6.4-6.5), optimal control inputs must inherently be *separated* feedback policies (e.g. [47, 43]). That is, optimal control input u_t depends on the data ζ^t and initial density $\pi_{0|-1}$ solely through the current information state, π_t. Optimality thus requires propagating π_t and policies g_t, where

$$u_t = g_t(\pi_t).$$

Definition 6.1. $\mathbb{E}_t[\,\cdot\,]$ and $\mathbb{P}_t[\,\cdot\,]$ are expected value and probability with respect to state x_t – with conditional density π_t – and i.i.d. random variables $\{(w_k, v_{k+1}) : k \geq t\}$.

Notice that stochastic optimal control on the infinite horizon (see [47, 48]) typically requires a discount factor $\alpha < 1$, casting the stochastic version of (6.1) as

$$J_\infty(\pi_0, g) \triangleq \mathbb{E}_0\left[\sum_{k=0}^{\infty} \alpha^k c(x_k, g_k(\pi_k))\right], \tag{6.6}$$

with corresponding finite-horizon cost

$$J_N(\pi_0, g) \triangleq \mathbb{E}_0\left[\sum_{k=0}^{N-1} \alpha^k c(x_k, g_k(\pi_k)) + \alpha^N c_N(x_N)\right]. \tag{6.7}$$

Defining the optimal value function $J_N^\star(\pi_0)$ as in (6.3),

$$J_N^\star(\pi_0) \triangleq \inf_{g_k(\cdot)} J_N(\pi_0, g),$$

finite-horizon stochastic optimal feedback policies may be computed, in principle, by solving the stochastic dynamic programming equation,

$$J_{N-k}^\star(\pi_k) \triangleq \inf_{g_k(\cdot)} \mathbb{E}_k\left[c(x_k, g_k(\pi_k)) + \alpha J_{N-k-1}^\star(\pi_{k+1})\right], \tag{6.8}$$

for $k = 0, \ldots, N-1$. The equation is solved backwards in time, from its terminal value,

$$J_0^\star(\pi_N) \triangleq \mathbb{E}_N\left[c_N(x_N)\right]. \tag{6.9}$$

Similarly to Section 6.2, we denote by: $J_\infty^\star(\pi)$ the infinite-horizon optimal value function; μ^N the sequence of optimal policies in (6.8-6.9); μ_0^N the first element of this sequence; $\mu_{\text{MPC}}^N \triangleq \{\mu_0^N, \mu_0^N, \ldots\}$ the receding horizon implementation of this sequence. We next impose the following stochastic counterpart to Assumption 6.1 to discuss the infinite horizon cost of the SMPC law μ_{MPC}^N.

Assumption 6.3. For $\alpha \in [0, 1)$, there exist $\eta \in \mathbb{R}_+$ and a policy $\tilde{g}(\cdot)$ such that

$$\mathbb{E}_\pi \left[\alpha\, c_N(f(x, \tilde{g}(\pi), w))\right] \leq \mathbb{E}_\pi \left[c_N(x) - c(x, \tilde{g}(\pi))\right] + \frac{\eta}{\alpha^{N-1}},$$

for all densities π of $x \in X$. The expectation $\mathbb{E}_\pi[\cdot]$ is with respect to state x – with conditional density π – and w.

This assumption then leads to the following extension of Theorem 6.1 to SMPC of system (6.4-6.5).

Theorem 6.2 (Performance of stochastic MPC [71])**.** *Given Assumption 6.3, SMPC with* $\alpha \in [0, 1)$ *yields*

$$J^\star_\infty(\pi) \leq J_\infty(\pi, \mu^N_{MPC}) \leq J^\star_N(\pi) + \frac{\alpha}{1-\alpha}\eta,$$

for all densities π *of* $x \in X$*.*

This result relates the following quantities in SMPC: *design cost*, $J^\star_N(\pi)$, which is evaluated as part of the SMPC computation; *optimal cost*, $J^\star_\infty(\pi)$, which is unknown (otherwise we would use the infinite-horizon optimal policy); and, unknown infinite-horizon SMPC *achieved cost* $J_\infty(\pi, \mu^N_{\text{MPC}})$. The result, which must exhibit duality and satisfaction of the stochastic programming equation (6.8-6.9), is special in that SMPC approaches relying on approximation of the finite horizon Stochastic Optimal Control problem, as commonly found in the literature, do not generally yield statements regarding performance of the implemented control laws on the infinite horizon. This fact is linked inherently to the loss of the dual optimal nature of the control inputs when avoiding solution of (6.8-6.9).

As in Section 6.2 and Theorem 6.1, the proof of Theorem 6.2 via Assumption 6.3 relies on verifying monotonicity of the underlying optimal value function $J_N^\star(\pi)$. We next proceed by extending this result and its proof to dual optimal receding horizon control of POMDPs.

6.4 Stochastic MPC for POMDPs

POMDPs are characterized by probabilistic dynamics on a finite state space $X = \{1, \ldots, n\}$, finite action space $U = \{1, \ldots, m\}$, and finite observation space $Y = \{1, \ldots, o\}$. POMDP dynamics are defined by the conditional state transition and observation probabilities

$$\mathbb{P}\left(x_{t+1} = j \mid x_t = i, u_t = a\right) = p_{ij}^a, \tag{6.10}$$

$$\mathbb{P}\left(y_{t+1} = \theta \mid x_{t+1} = j, u_t = a\right) = r_{j\theta}^a, \tag{6.11}$$

where $t \in \mathbb{N}_0$, $i, j \in X$, $a \in U$, $\theta \in Y$. The state transition dynamics (6.10) correspond to a conventional Markov Decision Process (MDP, e.g. [81]). However, the control actions u_t are to chosen based on the known initial state distribution $\pi_0 = \text{pdf}(x_0)$ and the sequences of observations, $\{y_1, \ldots, y_t\}$, and controls $\{u_0, \ldots, u_{t-1}\}$, respectively. That is, we are choosing our control actions in a Hidden Markov Model (HMM, e.g. [82]) setup. Notice that, while POMDPs conventionally do not have an initial observation y_0 in (6.11), as is commonly assumed in nonlinear system models of the form (6.4-6.5), one can easily modify this basic setup without altering the discussion below.

Given control action $u_t = a$ and measured output $y_{t+1} = \theta$, the information state π_t in a POMDP is updated via

$$\pi_{t+1,j} = \frac{\sum_{i \in X} \pi_{t,j} p_{ij}^a r_{j\theta}^a}{\sum_{i,j \in X} \pi_{t,j} p_{ij}^a r_{j\theta}^a},$$

where $\pi_{t,j}$ denotes the j^{th} entry of the row vector π_t. To specify the cost functionals (6.6) and (6.7) in the POMDP setup, we write the stage cost as $c(x_t, u_t) = c_i^a$ if $x_t = i \in X$ and $u_t = a \in U$, summarized in the column vectors $c(a)$ of the same dimension as row vectors π_k. Similarly, the terminal cost terms are $c_N(x_t) = c_{i,N}$ if $x_N = i \in X$, summarized in the column vector c_N. The infinite horizon cost functional defined in Section 6.3 then follows as

$$J_\infty(\pi_0, g) = \mathbb{E}_0 \left[\sum_{k=0}^{\infty} \alpha^k \pi_k c(g_k(\pi_k)) \right],$$

with corresponding finite-horizon variant

$$J_N(\pi_0, g) = \mathbb{E}_0 \left[\sum_{k=0}^{N-1} \alpha^k \pi_k c(g_k(\pi_k)) + \alpha^N \pi_N c_N \right].$$

Extending (6.8-6.9), optimal control decisions may then be computed via

$$J_{N-k}^\star(\pi_k) = \min_{g_k(\cdot)} \left\{ \pi_k c(g_k(\pi_k)) + \alpha \sum_{\theta \in Y} \mathbb{P}\left(y_{k+1} = \theta \mid \pi_k,\, g_k(\pi_k)\right) J_{N-k-1}^\star(\pi_{k+1}) \right\}, \tag{6.12}$$

for $k = 0, \ldots, N-1$, from terminal value function

$$J_0^\star(\pi_N) = \pi_N c_N. \tag{6.13}$$

Using the notation for optimal finite- and infinite-horizon value functions as well as MPC policies introduced in Section 6.3, we next prove the following auxiliary result before extending the performance guarantees in Theorem 6.2 to SMPC on POMDPs.

Lemma 6.1. *If there exist* $\gamma \in [0, 1]$ *and* $\eta \in \mathbb{R}_+$ *such that*

$$\mathbb{E}_0 \left[J_N^\star(\pi_1) - J_{N-1}^\star(\pi_1) \right] \leq \gamma \mathbb{E}_0 \left[\pi_0 c(\mu_0^N(\pi_0)) \right] + \eta, \tag{6.14}$$

for all densities π_0 *of* $x_0 \in X$*, then SMPC with discount factor* $\alpha \in [0, 1)$ *yields*

$$(1 - \alpha\gamma)\, J_\infty^\star(\pi_0) \leq (1 - \alpha\gamma)\, J_\infty(\pi_0, \mu_{MPC}^N) \leq J_N^\star(\pi_0) + \frac{\alpha}{1 - \alpha}\eta. \tag{6.15}$$

Proof. Optimality of the initial policy $\mu_0^N(\cdot)$ implies

$$J_N^\star(\pi_0) = \mathbb{E}_0 \left[\pi_0 c(\mu_0^N) + \alpha J_{N-1}^\star(\pi_1) \right] + \alpha \mathbb{E}_0 \left[J_N^\star(\pi_1) - J_N^\star(\pi_1) \right],$$

which by (6.14) yields

$$(1 - \alpha\gamma)\mathbb{E}_0 \left[\pi_0 c(\mu_0^N(\pi_0)) \right] \leq J_N^\star(\pi_0) - \alpha \mathbb{E}_0 \left[J_N^\star(\pi_1) \right] + \alpha\eta. \tag{6.16}$$

Now denote by $J_\infty^M(\pi_0, \mu_{\text{MPC}}^N)$ the first $M \in \mathbb{N}_1$ terms of the achieved infinite-horizon cost $J_\infty(\pi_0, \mu_{\text{MPC}}^N)$ subject to the SMPC implementation of policy $\mu_0^N(\cdot)$. By (6.16), we have

$$\begin{aligned}(1 - \alpha\gamma) J_\infty^M(\pi_0, \mu_{\text{MPC}}^N) = (1 - \alpha\gamma)\mathbb{E}_0 \left[\sum_{k=0}^{M-1} \alpha^k \pi_k c(\mu_0^N(\pi_k)) \right] \leq \\ \mathbb{E}_0 \Big[J_N^\star(\pi_0) - \alpha J_N^\star(\pi_1) + \alpha\eta + \alpha J_N^\star(\pi_1) - \alpha^2 J_N^\star(\pi_2) + \\ \alpha^2\eta + \ldots + \alpha^{M-1} J_N^\star(\pi_{M-1}) - \alpha^M J_N^\star(\pi_M) + \alpha^M \eta \Big],\end{aligned}$$

such that

$$(1 - \alpha\gamma) J_\infty^M(\pi_0, \mu_{\text{MPC}}^N) \leq J_N^\star(\pi_0) - \alpha^M \mathbb{E}_0 \left[J_N^\star(\pi_M) \right] + \left(\alpha + \ldots + \alpha^M \right) \eta,$$

which confirms the right-hand inequality in (6.15) in the limit as $M \to \infty$. The left-hand inequality follows directly from optimality on the infinite horizon. □

This lemma then leads to the following assumption and subsequent performance result in the spirit of Theorems 6.1-6.2.

Assumption 6.4. For $\alpha \in [0, 1)$, there exist $\eta \in \mathbb{R}_+$ and a policy $\tilde{g}(\cdot)$ such that

$$\mathbb{E}_0 \left[\alpha \pi_1 c_N\right] \leq \mathbb{E}_0 \left[\pi_0 c_N - \pi_0 c(\tilde{g}(\pi_0))\right] + \frac{\eta}{\alpha^{N-1}}, \tag{6.17}$$

for all densities π_0 of $x_0 \in X$.

Theorem 6.3. *[Performance of SMPC for POMDPs] Given Assumption 6.4, SMPC for POMDPs with $\alpha \in [0, 1)$ yields*

$$J^\star_\infty(\pi) \leq J_\infty(\pi, \mu^N_{MPC}) \leq J^\star_N(\pi) + \frac{\alpha}{1-\alpha}\eta,$$

for all densities π of $x \in X$.

Proof. Use optimality and Assumption 6.4 to conclude

$$\begin{aligned} J^\star_N(\pi_1) - J^\star_{N-1}(\pi_1) = \mathbb{E}_0 \Bigg[& \left(\sum_{k=0}^{N-1} \alpha^k \pi_{k+1} c(\mu^N_k(\pi_{k+1})) + \alpha^N \pi_{N+1} c_N \right) \\ & - \left(\sum_{k=0}^{N-2} \alpha^k \pi_{k+1} c(\mu^N_{k+1}(\pi_{k+1})) + \alpha^{N-1} \pi_N c_N \right) \Bigg] \\ \leq \mathbb{E}_0 & \left[\alpha^{N-1} \pi_N c(\tilde{g}(\pi_N)) + \alpha^N \pi_{N+1} c_N - \alpha^{N-1} \pi_N c_N \right] \\ \leq \eta, & \end{aligned}$$

which implies (6.14) with $\gamma = 0$ and thus completes the proof by Lemma 6.1. □

6.5 Numerical Example in Healthcare

6.5.1 Problem Setup

The remainder of this chapter discusses a particular numerical example of decisions on treatment and diagnosis in healthcare, displaying specifically the use of dual control in SMPC applied to a POMDP. Consider a patient treated for a specific disease which can be managed but not cured. For simplicity, we assume that the patient does not die under treatment. While this transition would have to be added in practice, it results in a time-varying model, which we avoid in order to keep the following discussion compact.

The example, introduced in [80], is set up as follows. The disease encompasses three stages with severity increasing from Stage 1 through Stage 2 to Stage 3, transitions between which are governed by a Markov chain with transition probability matrix

$$P = \begin{bmatrix} 0.8 & 0.2 & 0.0 \\ 0.0 & 0.9 & 0.1 \\ 0.0 & 0.0 & 1.0 \end{bmatrix},$$

where P is the matrix with values p_{ij} at row i and column j. All transition and observation probability matrices below are defined similarly. Once our patient enters Stage 3, Stages 1 and 2 are inaccessible for all future times. However, Stage 3 can only be entered through Stage 2, a transition from which to Stage 1 is possible only under costly treatment. The same treatment inhibits transitions from Stage 2 to Stage 3. We have access to the patient state only through tests, which will result in one of three possible

values, each of which is representative of one of the three disease stages. However, these tests are imperfect, with non-zero probability of returning an incorrect disease stage. All possible state transitions and observations are illustrated in Figure 6.2.

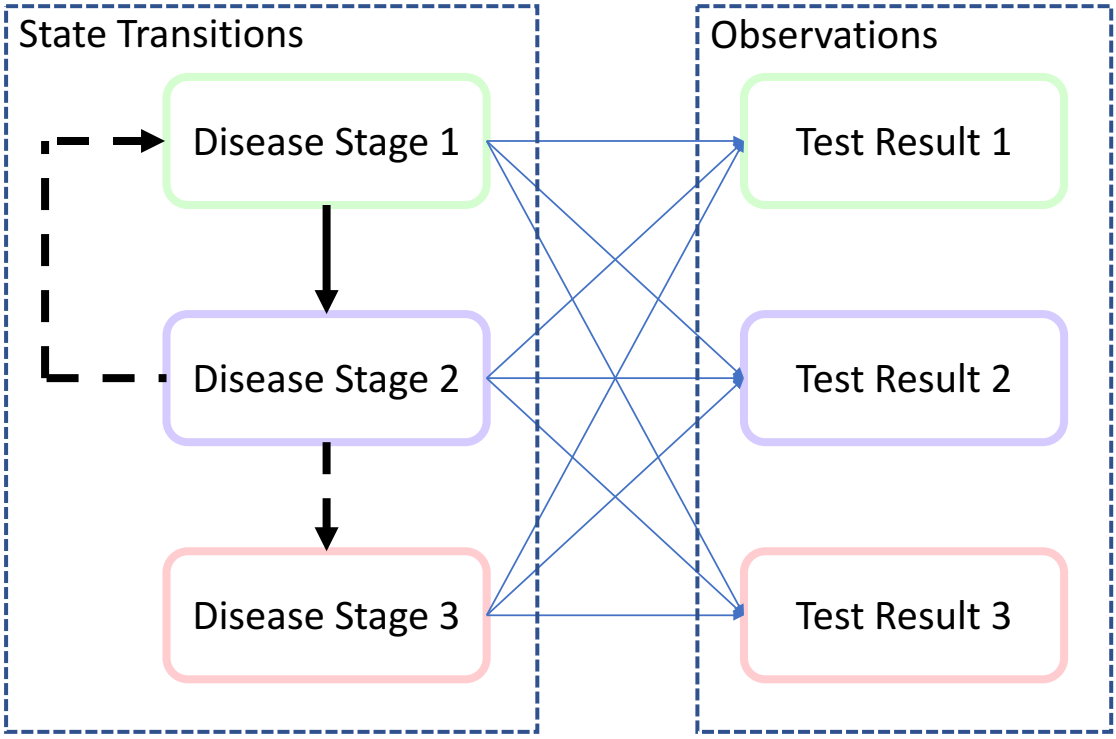

Figure 6.2: Feasible state transitions and possible test results in healthcare example. Solid arrows for feasible state transitions and observations. Dashed arrows for transitions conditional on treatment and diagnosis decisions.

At each point in time, the current information state π_t is available to make one of four possible decisions:

1. Skip next appointment slot
2. Schedule new appointment
3. Order rapid diagnostic test

4. Apply available treatment

Skipping an appointment slot results in the patient progressing through the Markov chain describing the transition probabilities of the disease without medical intervention, without new information being available after the current decision epoch. Scheduling an appointment does not alter the patient transition probabilities but provides a low-quality assessment of the current disease stage, which is used to refine the next information state. The third option, ordering a rapid diagnostic test, allows for a high-quality assessment of the patient's state, leading to a more reliable refinement of the next information state than possible when choosing the previous decision option. The results from this diagnostic test are considered available sufficiently fast so that the patient state remains unchanged under this decision. The remaining option entails medical intervention, allowing transition from Stage 2 to Stage 1 while preventing transition from Stage 2 to Stage 3. Transition probabilities $P(a)$, observation probabilities $R(a)$, and stage cost vectors $c(a)$ for each decision are summarized in Table 6.1. Additionally, we impose the terminal cost

$$c_N = \begin{bmatrix} 0 & 4 & 30 \end{bmatrix}^T.$$

6.5.2 Rationale for Duality

Intuitively, we expect an efficient policy for this problem to attempt avoiding transitions to Stage 3 while managing the resources required to schedule appointments, order tests, or apply medical intervention. This may, in principle, be achieved by a

policy akin to the following structure:

1. Skip appointments when Stages 2 and 3 are unlikely.

2. Schedule appointments when Stages 2 and 3 are likely but the probability for Stage 2 is below some threshold.

3. Order diagnostic test if the probability of Stage 2 lies in a specific range.

4. Proceed with medical intervention if the probability of Stage 2 is high.

While the optimal policy may be somewhat more intricate, this simple decision structure could be acceptable in practice. However, even this simple structure includes duality in the decisions, demonstrated by including the diagnostic test even though it does not alter the patient state. That is, this decision improves the quality of available information at a cost, also called *exploration*. This improvement in the available information allows us to apply medical intervention at appropriate times, which is called *exploitation*.

6.5.3 Computational Results

The trade-off between these two principal decision categories is precisely what is encompassed by duality, which we can include in an optimal sense by solving (6.12-6.13) and applying the resulting initial policy in receding horizon fashion. This is demonstrated in Figure 6.3, which shows simulation results for SMPC with control horizon $N = 5$ and discount factor $\alpha = 0.98$. As anticipated, the stochastic optimal receding horizon policy shows a structure not drastically different from the decision structure

motivated above. In particular, diagnostic tests are used effectively to decide on medical intervention.

In order to apply Theorem 6.3 to this particular example, we choose the policy $\tilde{g}(\cdot)$ in Assumption 6.4 always to apply medical intervention. Using the worst-case scenario for the expectations in (6.17), which entails transition from Stage 1 to Stage 2 under treatment, we can satisfy Assumption 6.4 with $\eta = 7.92$. The computed cost in our simulation is $J_N^\star(\pi_0) \approx 11.36$. Combined with the discount factor $\alpha = 0.98$, we thus have the upper bound

$$J_\infty(\pi_0, \mu_{\mathrm{MPC}}^N) \leq J_N^\star(\pi_0) + \frac{\alpha}{1-\alpha}\eta \approx 400$$

via application of Theorem 6.3. Denoting by e_j the row-vector with entry 1 in element j and zeros elsewhere, the observed (finite-horizon) cost corresponding with Figure 6.3 is

$$J_\infty^{\mathrm{obs}} = \sum_{k=0}^{29} e_{x_k} c(\mu_0^N(\pi_k)) \approx 38.53 < 400.$$

While this bound is not particularly tight, one may modify the discount factor α or the terminal cost c_N to achieve a tighter estimate of the achieved MPC cost.

6.6 Conclusions

We extended closed-loop achieved performance guarantees well-known in deterministic MPC to SMPC and in particular receding horizon control of POMDPs, which allow tractable solution of the underlying Stochastic Optimal Control problems and thus duality of the control inputs in an optimal sense. The basic formulations in this chapter

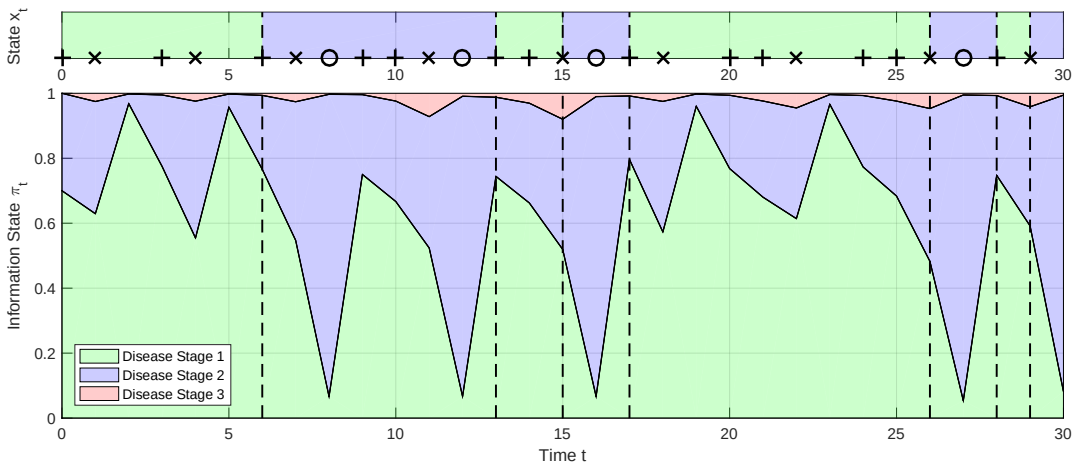

Figure 6.3: Simulation results for SMPC with horizon $N = 5$ and discount factor $\alpha = 0.98$. Top plot displays patient state and transitions, with optimal SMPC decisions based on current information state: appointment (pluses); diagnosis (crosses); treatment (circles). Bottom plot shows information state evolution. Dashed vertical lines mark time instances of state transitions.

can be modified, for instance, by introducing state and input constraint sets or time-varying (monotonic) stage costs. While this requires additional assumptions to maintain recursive feasibility of the MPC and SMPC inputs, the cost discussion is rather similar. We demonstrated use of the novel results using a particular POMDP instance in healthcare decision making, demanding the use of probing control inputs in order to adequately decide upon the proper and cost-effective use of medical intervention.

Table 6.1: Problem data for healthcare decision making example.

Decision a	Transition Probabilities $P(a)$	Observation Probabilities $R(a)$	Cost $c(a)$
1: Skip next appointment slot	$\begin{bmatrix} 0.80 & 0.20 & 0.00 \\ 0.00 & 0.90 & 0.10 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 5 \\ 5 \end{bmatrix}$
2: Schedule new appointment	$\begin{bmatrix} 0.80 & 0.20 & 0.00 \\ 0.00 & 0.90 & 0.10 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.40 & 0.30 & 0.30 \\ 0.30 & 0.40 & 0.30 \\ 0.30 & 0.30 & 0.40 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
3: Order rapid diagnostic test	$\begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.90 & 0.05 & 0.05 \\ 0.05 & 0.90 & 0.05 \\ 0.05 & 0.05 & 0.90 \end{bmatrix}$	$\begin{bmatrix} 4 \\ 3 \\ 4 \end{bmatrix}$
4: Apply available treatment	$\begin{bmatrix} 0.80 & 0.20 & 0.00 \\ 0.75 & 0.25 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.40 & 0.30 & 0.30 \\ 0.30 & 0.40 & 0.30 \\ 0.30 & 0.30 & 0.40 \end{bmatrix}$	$\begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix}$

Acknowledgements

Chapter 6, in part, has been submitted for publication of the material as it may appear in: M.A. Sehr, R.R. Bitmead, "Performance of Model Predictive Control of POMDPs", *Proc. 56th IEEE Conference on Decision and Control*, 2017. The dissertation author was the primary investigator and author of this paper.

Conclusions

This dissertation has broached the subject of duality in Stochastic Optimal Control and, thereby, its presence in formulations of stochastic MPC based on Stochastic Optimal Control. Our aim has been to highlight two central aspects: the reliance of optimality on probing to manage future information states along the horizon, and the attendant computational intractability. The benefit accrued is infinite-horizon stochastic MPC performance quantitatively comparable to truly stochastically optimal.

A central contribution of this dissertation is the presentation of an SMPC algorithm based on Stochastic Optimal Control in Chapter 4. This yields a number of theoretical properties of the controlled system. Performance guarantees of SMPC are made in comparison to performance of the infinite-horizon stochastically optimally controlled system and are presented in Theorem 4.4 and Corollary 4.2. These results extend those of [28], which pertain to full-state feedback Stochastic Optimal Control and which therefore do not accommodate duality. Other examples of stochastic performance bounds are mostly restricted to linear systems and, while computable, do not relate to the optimal constrained control. While the formal stochastic results are traceable to deterministic

predecessors, the divergence from earlier work is also notable. This concentrates on the use of the information state to accommodate measurements and the exploration of control policy functionals stemming from the Stochastic Dynamic Programming Equation. The resulting output feedback control possesses duality and optimality properties which are either artificially imposed in or absent from earlier approaches.

We further suggested two potential strategies to ameliorate the computational intractability of the Bayesian filter and the Stochastic Dynamic Programming Equation, famous for its *curse of dimensionality*. Firstly, one may use the Particle filter implementation of the Bayesian filter, which with a loss of duality can be combined with scenario methods, as outlined in Chapter 5. Secondly, we point out that our dual optimal SMPC algorithm becomes computationally tractable for the special case of Partially Observable Markov Decision Processes, which may be used either to approximate a nonlinear model or to model a given system in the first place. This strategy inherits the dual nature of our SMPC algorithm for general nonlinear systems, as discussed specifically in Chapter 6.

Bibliography

[1] D. Q. Mayne, J. B. Rawlings, C. V. Rao, P. O. M. Scokaert, Constrained model predictive control: Stability and optimality, Automatica 36 (6) (2000) 789–814.

[2] D. Q. Mayne, Model predictive control: Recent developments and future promise, Automatica 50 (12) (2014) 2967–2986.

[3] J. M. Maciejowski, Predictive Control with Constraints, Prentice Hall, Englewood Cliffs, NJ, 2002.

[4] J. Yan, R. R. Bitmead, Incorporating state estimation into model predictive control and its application to network traffic control, Automatica 41 (4) (2005) 595–604.

[5] D. Sui, L. Feng, M. Hovd, Robust output feedback model predictive control for linear systems via moving horizon estimation, in: American Control Conference, Seattle, WA, 2008, pp. 453–458.

[6] D. A. Copp, J. P. Hespanha, Nonlinear output-feedback model predictive control with moving horizon estimation, in: 53rd IEEE Conference on Decision and Control, Los Angeles, CA, 2014, pp. 3511–3517.

[7] D. Q. Mayne, S. V. Raković, R. Findeisen, F. Allgöwer, Robust output feedback model predictive control of constrained linear systems: Time varying case, Automatica 45 (9) (2009) 2082–2087.

[8] A. T. Schwarm, M. Nikolaou, Chance-constrained model predictive control, AIChE Journal 45 (8) (1999) 1743–1752.

[9] M. Cannon, Q. Cheng, B. Kouvaritakis, S. V. Raković, Stochastic tube MPC with state estimation, Automatica 48 (3) (2012) 536–541.

[10] L. Chisci, J. A. Rossiter, G. Zappa, Systems with persistent disturbances: predictive control with restricted constraints, Automatica 37 (7) (2001) 1019–1028.

[11] G. Marafioti, R. R. Bitmead, M. Hovd, Persistently exciting model predictive control, International Journal of Adaptive Control and Signal Processing 28 (6) (2014) 536–552.

[12] P. Lebowitz, Schedule the short procedure first to improve OR efficiency, AORN journal 78 (4) (2003) 651–659.

[13] E. N. Weiss, Models for determining estimated start times and case orderings in hospital operating rooms, IIE transactions 22 (2) (1990) 143–150.

[14] S. P. Meyn, Control techniques for complex networks, Cambridge University Press, 2008.

[15] H. J. Kushner, Heavy traffic analysis of controlled queueing and communication networks, Vol. 47, Springer, 2001.

[16] D. Swaroop, J. Hedrick, String stability of interconnected systems, IEEE Transactions on Automatic Control 41 (3) (1996) 349–357.

[17] P. Seiler, A. Pant, J. K. Hedrick, Disturbance propagation in vehicle strings, IEEE Transactions on Automatic Control 49 (10) (2004) 1835–1842.

[18] M. I. Reiman, L. M. Wein, Dynamic scheduling of a two-class queue with setups, Operations Research 46 (4) (1998) 532–547.

[19] T. Cayirli, E. Veral, Outpatient scheduling in health care: a review of literature, Production and Operations Management 12 (4) (2003) 519–549.

[20] D. Gupta, B. Denton, Appointment scheduling in health care: Challenges and opportunities, IIE transactions 40 (9) (2008) 800–819.

[21] B. Cardoen, E. Demeulemeester, J. Beliën, Operating room planning and scheduling: A literature review, European Journal of Operational Research 201 (3) (2010) 921–932.

[22] D. Gupta, Surgical suites' operations management, Production and Operations Management 16 (6) (2007) 689–700.

[23] M. A. Sehr, R. R. Bitmead, J. Fontanesi, Multi-class appointments in individualized healthcare: Analysis for scheduling rules, in: Proc. European Control Conference, Linz, Austria, 2015, pp. 1219–1224.

[24] A. Stuart, K. Ord, Kendall's Advanced Theory of Statistics, 6th Edition, Vol. I, Hodder Arnold Publication, London, 1994.

[25] B. Jansson, Choosing a good appointment system—a study of queues of the type (D, M, 1), Operations Research 14 (2) (1966) 292–312.

[26] L. Grüne, A. Rantzer, On the infinite horizon performance of receding horizon controllers, IEEE Transactions on Automatic Control 53 (9) (2008) 2100–2111.

[27] L. Grüne, J. Pannek, Nonlinear model predictive control, Springer, London, 2011.

[28] D. J. Riggs, R. R. Bitmead, MPC under the hood/sous le capot/unter der haube, in: 4th IFAC Nonlinear Model Predictive Control Conference, 2012, pp. 363–368.

[29] D. Q. Mayne, S. V. Raković, R. Findeisen, F. Allgöwer, Robust output feedback model predictive control of constrained linear systems, Automatica 42 (7) (2006) 1217–1222.

[30] F. Allgöwer, T. A. Badgwell, J. S. Qin, J. B. Rawlings, S. J. Wright, Nonlinear predictive control and moving horizon estimation—an introductory overview, in: Advances in control, Springer, 1999, pp. 391–449.

[31] C. Løvaas, M. M. Seron, G. C. Goodwin, Robust output-feedback model predictive control for systems with unstructured uncertainty, Automatica 44 (8) (2008) 1933–1943.

[32] M. M. Seron, J. De Dona, G. C. Goodwin, Global analytical model predictive control with input constraints, in: Decision and Control, 2000. Proceedings of the 39th IEEE Conference on, Vol. 1, IEEE, 2000, pp. 154–159.

[33] A. Bemporad, M. Morari, V. Dua, E. N. Pistikopoulos, The explicit linear quadratic regulator for constrained systems, Automatica 38 (1) (2002) 3–20.

[34] A. Bemporad, F. Borrelli, M. Morari, Model predictive control based on linear programming—the explicit solution, IEEE Transactions on Automatic Control 47 (12) (2002) 1974–1985.

[35] A. Alessio, A. Bemporad, A survey on explicit model predictive control, in: D. Raimondo, L. Magni, F. Allgöwer (Eds.), Nonlinear model predictive control, Springer, Berlin, 2009, pp. 345–369.

[36] C. N. Jones, M. Barić, M. Morari, Multiparametric linear programming with applications to control, European Journal of Control 13 (2) (2007) 152–170.

[37] A. G. Hadigheh, O. Romanko, T. Terlaky, Sensitivity analysis in convex quadratic optimization: simultaneous perturbation of the objective and right-hand-side vectors, Algorithmic Operations Research 2 (2) (2007) 94.

[38] M. A. Sehr, R. R. Bitmead, Sumptus cohiberi: The cost of constraints in MPC with state estimates, in: American Control Conference, Boston, MA, 2016, pp. 901–906.

[39] G. C. Goodwin, H. Kong, G. Mirzaeva, M. M. Seron, Robust model predictive control: reflections and opportunities, Journal of Control and Decision 1 (2) (2014) 115–148.

[40] B. Kouvaritakis, M. Cannon, Model Predictive Control, Springer, Switzerland, 2016.

[41] A. Mesbah, Stochastic model predictive control: An overview and perspectives for future research, IEEE Control Systems Magazine, Accepted.

[42] M. A. Sehr, R. R. Bitmead, Particle model predictive control: Tractable stochastic nonlinear output-feedback MPC, arXiv:1612.00505. To appear in proc. IFAC World Congress, Toulouse, France (2017).

[43] P. R. Kumar, P. Varaiya, Stochastic Systems: Estimation, Identification, and Adaptive Control, Prentice-Hall, Englewood Cliffs, NJ, 1986.

[44] S. P. Ponomarev, Submersions and preimages of sets of measure zero, Siberian Mathematical Journal 28 (1) (1987) 153–163.

[45] Z. Chen, Bayesian filtering: From Kalman filters to particle filters, and beyond, Statistics 182 (1) (2003) 1–69.

[46] D. Simon, Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches, John Wiley & Sons, New York, NY, 2006.

[47] D. P. Bertsekas, Dynamic programming and optimal control, Athena Scientific, Belmont, MA, 1995.

[48] D. P. Bertsekas, S. E. Shreve, Stochastic optimal control: The discrete time case, Vol. 23, Academic Press, New York, NY, 1978.

[49] A. A. Fel'dbaum, Optimal control systems, Academic Press, New York, NY, 1965.

[50] H. Genceli, M. Nikolaou, New approach to constrained predictive control with simultaneous model identification, AIChE journal 42 (10) (1996) 2857–2868.

[51] B. Kouvaritakis, M. Cannon, Stochastic model predictive control, in: J. Baillieul, T. Samad (Eds.), Encyclopedia of Systems and Control, Springer, London, 2015, pp. 1350–1357.

[52] J. L. Doob, Classical Potential Theory and Its Probabilistic Counterpart, Springer, Berlin, 1984.

[53] J. L. Doob, Stochastic processes, John Wiley & Sons, New York, NY, 1953.

[54] W. M. Wonham, On the separation theorem of stochastic control, SIAM Journal on Control 6 (2) (1968) 312–326.

[55] T. T. Georgiou, A. Lindquist, The separation principle in stochastic control, redux, IEEE Transactions on Automatic Control 58 (10) (2013) 2481–2494.

[56] M. Athans, P. L. Falb, Optimal Control: An Introduction to the Theory and its Applications, McGraw-Hill, New York, NY, 1966.

[57] R. Hermann, A. J. Krener, Nonlinear controllability and observability, IEEE Transactions on Automatic Control 22 (5) (1977) 728–740.

[58] A. R. Liu, R. R. Bitmead, Stochastic observability in network state estimation and control, Automatica 47 (2011) 65–78.

[59] J. J. Florentin, Optimal, probing, adaptive control of a simple Bayesian system, International Journal of Electronics 13 (2) (1962) 165–177.

[60] O. L. R. Jacobs, J. W. Patchell, Caution and probing in stochastic control, International Journal of Control 16 (1) (1972) 189–199.

[61] Y. Bar-Shalom, E. Tse, Dual effect, certainty equivalence, and separation in stochastic control, IEEE Transactions on Automatic Control 19 (5) (1974) 494–500.

[62] Y. Bar-Shalom, E. Tse, Caution, probing, and the value of information in the control of uncertain systems, Annals of Economic and Social Measurement 5 (3) (1976) 323–337.

[63] A. A. Fel'dbaum, Dual-control theory, part I, Automation & Remote Control 21 (9) (1961) 874–880.

[64] A. A. Fel'dbaum, Dual-control theory, part II, Automation & Remote Control 21 (11) (1961) 1033–1039.

[65] N. M. Filatov, H. Unbehauen, Survey of adaptive dual control methods, IEE Proceedings - Control Theory and Applications 147 (1) (2000) 118–128.

[66] G. C. Calafiore, L. Fagiano, Stochastic model predictive control of LPV systems via scenario optimization, Automatica 49 (6) (2013) 1861–1866.

[67] S. Grammatico, X. Zhang, K. Margellos, P. Goulart, J. Lygeros, A scenario approach for non-convex control design, IEEE Transactions on Automatic Control 61 (2) (2016) 334–345.

[68] J. H. Lee, From robust model predictive control to stochastic optimal control and approximate dynamic programming: A perspective gained from a personal journey, Computers & Chemical Engineering 70 (2014) 114–121.

[69] A. Mesbah, S. Streif, R. Findeisen, R. D. Braatz, Stochastic nonlinear model predictive control with probabilistic constraints, in: American Control Conference, Portland, OR, 2014, pp. 2413–2419.

[70] G. Schildbach, L. Fagiano, C. Frei, M. Morari, The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations, Automatica 50 (12) (2014) 3009–3018.

[71] M. A. Sehr, R. R. Bitmead, Stochastic model predictive control: Output-feedback, duality and guaranteed performance, submitted to Automatica (2016).

[72] L. Blackmore, M. Ono, A. Bektassov, B. C. Williams, A probabilistic particle-control approximation of chance-constrained stochastic predictive control, IEEE Transactions on Robotics 26 (3) (2010) 502–517.

[73] A. F. M. Smith, A. E. Gelfand, Bayesian statistics without tears: a sampling–resampling perspective, The American Statistician 46 (2) (1992) 84–88.

[74] J. B. Rawlings, D. Q. Mayne, Model predictive control: Theory and design, Nob Hill Publishing, Madison, WI, 2009.

[75] G. Schildbach, L. Fagiano, M. Morari, Randomized solutions to convex programs with multiple chance constraints, SIAM Journal on Optimization 23 (4) (2013) 2479–2501.

[76] C. Snyder, T. Bengtsson, P. Bickel, J. Anderson, Obstacles to high-dimensional particle filtering, Monthly Weather Review 136 (12) (2008) 4629–4640.

[77] R. D. Smallwood, E. J. Sondik, The optimal control of partially observable markov processes over a finite horizon, Operations Research 21 (5) (1973) 1071–1088.

[78] L. P. Kaelbling, M. L. Littman, A. R. Cassandra, Planning and acting in partially observable stochastic domains, Artificial intelligence 101 (1) (1998) 99–134.

[79] Z. Sunberg, S. Chakravorty, R. S. Erwin, Information space receding horizon control, IEEE Transactions on Cybernetics 43 (6) (2013) 2255–2260.

[80] M. A. Sehr, R. R. Bitmead, Tractable dual optimal stochastic model predictive control: An example in healthcare, arXiv:1704.07770. Submitted for publication to Proc. 1st IEEE Conference on Control Technology and Applications (2017).

[81] M. L. Puterman, Markov decision processes: discrete stochastic dynamic programming, John Wiley & Sons, 2014.

[82] R. J. Elliott, L. Aggoun, J. B. Moore, Hidden Markov models: estimation and control, Vol. 29, Springer Science & Business Media, 2008.