Title
The stochastic robustness of model predictive control and closed-loop scheduling

Author
McAllister, Robert Dakota

Publication Date
2022

Peer reviewed|Thesis/dissertation

University of California
Santa Barbara

The stochastic robustness of model predictive control and closed-loop scheduling

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

Doctor of Philosophy
in
Chemical Engineering

by

Robert D. McAllister

Committee in charge:

Professor James B. Rawlings, Chair
Professor Bassam Bamieh
Professor Phillip Christopher
Professor Michael F. Doherty

September 2022

The Dissertation of Robert D. McAllister is approved.

Professor Bassam Bamieh

Professor Phillip Christopher

Professor Michael F. Doherty

Professor James B. Rawlings, Committee Chair

August 2022

The stochastic robustness of model predictive control and closed-loop scheduling

To my parents.

Acknowledgments

I have had the good fortune to interact with many excellent teachers, peers, and collaborators throughout my academic career. My advisor, James B. Rawlings, has provided incomparable advice and guidance throughout my time in his research group at both UW and UCSB. Jim always challenged me to develop and solve my own research questions, while offering inimitable wisdom and the occasional correction when I found myself too far in the proverbial deep end. The security of having Jim as a conscientious observer of my research, rejecting the particularly ill-conceived ideas while cheering on the good ones, allowed me to pursue the breadth and depth of topics I covered in my research with confidence. For this guidance and the plentiful career advice Jim has bestowed upon me, I am grateful.

Thanks are also due to Christos T. Maravelias of Princeton University for his excellent advice and recommendations on the production scheduling portion of my research. Christos filled a role somewhere between that of an informal co-advisor and collaborator, and our interactions significantly shaped my research for the better. I am grateful to Phillip Christopher for helping me develop as an educator while serving as his TA and co-instructor for the undergraduate process dynamic and control course. Phil allowed me great freedom to modify and improve the course content, a privilege that I am not sure many other faculty would have granted a 2nd or 3rd year graduate student. Phil Christopher also deserves thanks for agreeing to serve on my thesis committee. I would also like to thank Joseph Chada for contributing to my development as an educator and providing ample support for the laboratory projects we added to the process control courses at UCSB. In addition, I thank Mike F. Doherty and Bassam Bamieh for both agreeing to serve on my thesis committee and for their excellent courses in nonlinear dynamics and linear systems, respectively.

The staff at UCSB, and in particular the staff in the chemical engineering department, also deserve thanks for their efforts to support an enjoyable and productive research environment. The move between UW and UCSB was remarkable smooth thanks to the efforts of Erica Diaz, Mike Best, and Rachel Irving at UCSB and Kathy Heinzen at UW. Adrian Cortex was also very helpful in adding new laboratory components to the undergraduate process controls course.

My time in the Rawlings group was also shaped by the exceptional group of students I was privileged to call my peers and colleagues. Michael Risbeck provided advice and an abundance of software that was invaluable during the first few years of my graduate studies. Nishith Patel guided me through my internship at CrossnoKaye and provided me with some practical process control experience. Douglas Allan and Travis Arnold always provided an attentive and helpful sounding board for research questions, as well as camaraderie in our move from Madison to Santa Barbara. Pratyush Kumar deserves special mention as an excellent friend and peer throughout my tenure as a graduate student. The move from Madison to Santa Barbara would have been far more distressful without Pratyush's friendship and good humor. I wish Chris, Steven, Davide, and Titus the best of luck in their research and future endeavors.

During my graduate studies, I enjoyed the unique experience of making new friends in two completely different cities. These friends contributed in a large part to all of the fun I had over these past five years. In Madison, I had the privileged to befriend Michael Jindra, Alec Linot, and Jonathan Sheavly, who I still keep in touch with to this day. During my first year in

Santa Barbara, I found myself in the uncomfortable position of being a second-year graduate student with no social network. The first-years at the time were nonetheless happy to include Pratyush and I in their activities. I am very grateful for the friends I made from those days forward including Patrick Leggieri, Michael Schmithorst, Jordan Finzel, Toby Mazal, Kellie Heom, Sally Jiao, Chris Kuo-Leblanc, Katherine Hurlock, and Candice Macabuhay. The beach days, trivia nights, and dependable Friday night pool match at Pat's place injected some much needed sanity and relaxation into the somewhat hectic graduate school experience. As luck would have it, my Brother, Cooper, and my long-time friends Steven Scheidegg and Kevin Jones also found themselves in California at the same time as me. I thank them and all my friends back on the east coast for their support and the great times we have had over many years.

My wife, Julia, deserves perhaps the most credit for making my experience as a graduate student enjoyable. She moved to Santa Barbara from the Netherlands to join me, even though it was not the optimal move for her career. Graduate school, with its high demands and long hours, can easily strain a relationship, but Julia supported me every step of the way and truly gave me the stability I needed to accomplish this great feat. I only hope that I can provide for her the same environment she provided for me as we move back to the Netherlands for her Ph.D. studies. My time in Santa Barbara was all the better because she was there with me and I am excited for the next chapter of our great adventure together.

To my parents, I am forever grateful for their unconditional love and support. From a young age, they nurtured my curiosity and instilled in me a great respect for knowledge and education. The connection between my parents efforts in raising me and where I stand today is obvious to me. I highly doubt I would have pursued or completed a Ph.D. without their encouragement.

Koty McAllister
Santa Barbara, CA
August 2022

Curriculum Vitæ
Robert D. McAllister

Education

2018-2022	University of California, Santa Barbara (UCSB) – Santa Barbara, CA Ph.D. in Chemical Engineering (Expected)
2017-2018	University of Wisconsin, Madison (UW) – Madison, WI Began studies for a Ph.D. in Chemical Engineering
2013-2017	University of Delaware (UD) – Newark, DE Bachelors of Chemical Engineering

Peer-Reviewed Publications

Under review	R. D. McAllister and J. B. Rawlings. The inherent distributional robusntess of stochastic and nominal model predictive control. IEEE Trans. Auto. Cont., 2022e. Under review
Under review	R. D. McAllister and J. B. Rawlings. The stochastic robustness of nominal and stochastic model predictive control. IEEE Trans. Auto. Cont., 2022d. Under review
Under review	R. D. McAllister and J. B. Rawlings. A suboptimal economic model predictive control algorithm for large and infrequent disturbances. IEEE Trans. Auto. Cont., 2022c. Under review
	R. D. McAllister and J. B. Rawlings. Advances in mixed-integer model predictive control. In *American Control Conference*, pages 364–369, Atlanta, GA, June 8–10, 2022b
	R. D. McAllister and J. B. Rawlings. Nonlinear stochastic model predictive control: Existence, measurability, and stochastic asymptotic stability. *IEEE Trans. Auto. Cont.*, 2022a. doi: https://doi.org/10.1109/TAC.2022.3157131
	R. D. McAllister, J. B. Rawlings, and C. T. Maravelias. The inherent robustness of closed-loop scheduling. *Comput. Chem. Eng.*, 159:107678, 2022. doi: https://doi.org/10.1016/j.compchemeng.2022.107678
	R. D. Mcallister and J. B. Rawlings. Inherent stochastic robustness of model predictive control to large and infrequent disturbances. *IEEE Trans. Auto. Cont.*, 2021. doi: https://10.1109/TAC.2021.3122365
	R. D. McAllister and J. B. Rawlings. Stochastic exponential stability of nonlinear stochastic model predictive control. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 880–885. IEEE, 2021

R. D. McAllister and J. B. Rawlings. Robustness of model predictive control to (large) discrete disturbances. *IFAC-PapersOnLine*, 54(6):64–69, 2021a

R. D. McAllister, J. B. Rawlings, and C. T. Maravelias. Rescheduling penalties for economic model predictive control and closed-loop scheduling. *Ind. Eng. Chem. Res.*, 59(6):2214–2228, 2020

Research Experience

Sep 2017 - Aug 2022	Ph.D. Candidate, University of California, Santa Barbara Mentor: Professor James B. Rawlings
Jun 2021 - Sep 2021	Advanced Controls Intern, CrossnoKaye Mentor: Nishith R. Patel, Ph.D.
Aug 2016 - May 2017	Senior Thesis Research, University of Delaware Mentor: Professor Norman J. Wagner
May 2016 - Aug 2016	Chemical Engineering Intern, W.L. Gore and Associates Mentor: Alexander Agapov, Ph.D.
May 2015 - Aug 2015	Undergraduate Research Fellow, NIST Mentor: Kathleen M. Weigandt, Ph.D.

Teaching Experience

Spring 2020	Teaching Assistant, Process Dynamics and Control, UCSB
Fall 2019	Co-Instructor, Process Dynamics and Control, UCSB
Fall 2018	Teaching Assistant, Process Dynamics and Control, UCSB

Awards

2021	Schlinger Fellowship for Excellence in Chemical Engineering Research (UCSB)
2021	Best Talk in Session, Graduate Student Symposium (UCSB)
2020	CSP Technologies Teacher-Scholar Fellowship (UCSB)
2020	Outstanding Teaching Assistant for Chemical Engineering (UCSB)
2019	Outstanding Teaching Assistant for Chemical Engineering (UCSB)
2016	Chemical Engineering Class of 1950 Scholarship (UD)

Abstract

The stochastic robustness of model predictive control and closed-loop scheduling

by

Robert D. McAllister

Uncertainty is inherent to all science and engineering models. Any algorithm proposed to design, schedule, or control an industrial process must therefore be robust, i.e., the algorithm must be able to withstand and overcome this uncertainty. In this dissertation, we focus specifically on model predictive control (MPC), the advanced control algorithm of choice for chemical process control with a growing list of applications in several other engineering disciplines as well. For deterministic descriptions of this uncertainty, MPC is known to be robust to sufficiently small disturbances. This robustness is afforded by feedback and does not require any characterization of this uncertainty within the control algorithm.

Stochastic descriptions of uncertainty, however, are often better suited to model the behavior of physical systems and have proven highly useful in a variety of science and engineering applications. To that end, we expand the theory of robustness for MPC to address these stochastic descriptions of uncertainty and establish that MPC is robust in this stochastic context. We then apply this theory to the emerging field of stochastic MPC (SMPC), in which a stochastic model of this uncertainty is used directly in the control algorithm. Through the concept of distributional robustness, we further establish that SMPC is robust to uncertainty within even the stochastic model used in the control algorithm. This result in fact unifies the analysis of both MPC and SMPC, thereby allowing a novel comparison of the theoretical properties afforded by these two algorithms. We also demonstrate via suitable examples that including stochastic information in a control algorithm is not always beneficial to the controller's performance.

In the second part of this dissertation, we consider the stochastic robustness of MPC to a new class of large and infrequent disturbances, motivated by recent applications of MPC to production planning and scheduling problems. Using these results and the MPC framework, we then design a closed-loop scheduling algorithm that is robust to the large and infrequent disturbances pertinent to production scheduling problems. This algorithm is further modified to address computational and practical limitations and is therefore suitable for large-scale production scheduling applications.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

To remain competitive in an increasingly dynamic and global market, industries ranging from chemical production and manufacturing to logistic services must continue to implement technologies that improve productivity, minimize environmental impact, and maximize profit. Optimization offers a particularly attractive engineering tool to improve a variety of industrial operations (Grossmann, 2012). By formulating the design, scheduling, or control problem of interest as an optimization problem, the engineer can determine, unrestricted by personal bias and heuristics, the optimal solution. The proper formulation of such an optimization problem also allows the engineer to recompute an optimal solution with little marginal effort if parameters and conditions of the problem statement change. Thus, optimization not only provides an optimal solution to the problem of interest, but also offers an automated method to generate such a solution.

A more direct connection between automation and optimization is apparent when optimization is applied specifically to dynamical systems and process control, as opposed to the optimization of steady-state equations that characterize engineering design problems. In these optimization problems, we typically solve for a *trajectory* of inputs or actions to be implemented at specific times subject to a dynamical model of the system. These systems may include more traditional process control applications (e.g., level or temperature control in reactor) or higher-level problems in the enterprise decision hierarchy such as production

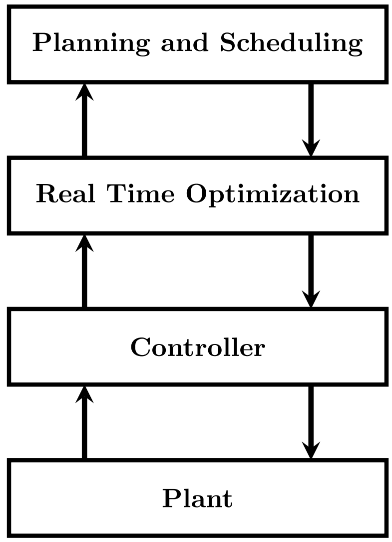

Figure 1.1: Typical automated decision-making hierarchy in a chemical production facility.

planning and scheduling (see Figure 1.1). However, disturbances such as price fluctuations, model inaccuracies, and measurement errors all but ensure that the optimal trajectory computed at one point in time is suboptimal at the next time step, regardless of the level of detail included in the dynamical model. Thus, any algorithm that is proposed to address these process control or scheduling problems should be *robust*, i.e., the algorithm must be able to withstand and overcome relevant disturbances. The topic of robustness is therefore central to the study of control and dynamical systems.

One method to provide robustness to these algorithms is to explicitly include disturbance models in the optimization problem. This method is particularly appealing for problems with significant uncertainty and minimal recourse, e.g., the design of a chemical plant. For process control and dynamical systems, however, the preferred method to provide robustness is through *feedback*. In a feedback or closed-loop method, the algorithm responds to these disturbances in real time as they are observed in the outputs of the plant. The archetypal feedback control structure is depicted in Figure 1.2. Hence, the controller adjusts the inputs to the plant based on these outputs to achieve a specified goal of the control algorithm (e.g.,

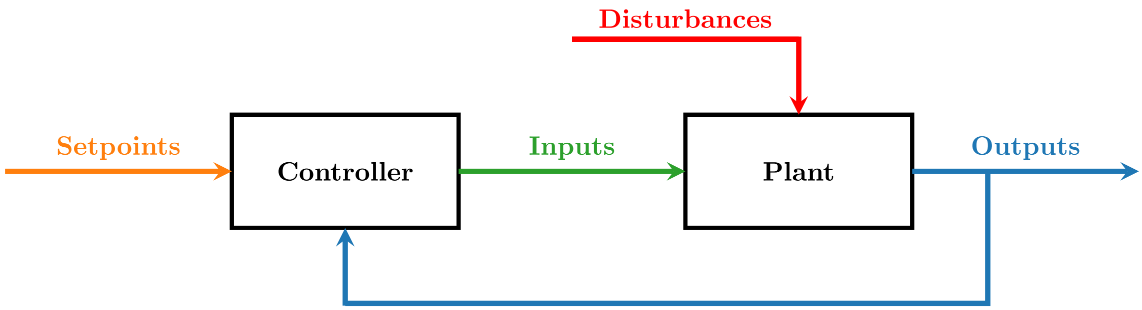

Figure 1.2: A basic feedback control structure.

tracking a setpoint).

The essence of model predictive control (MPC) is precisely this combination of feedback and optimization. By adjusting and resolving the optimization problem at regular and frequent intervals, MPC provides a framework to combine the performance benefits of optimization with the inherent robustness afforded by feedback. Driven by its versatile framework and industrial successes, MPC has become a popular advanced control algorithm among both practitioners and researchers alike (Mayne et al., 2000; Qin and Badgwell, 2003). Although originally developed for process control applications, recent computational and theoretical advances now allow the formulations and theoretical results of MPC to be applied to significantly larger class of problems including higher-level production planning and scheduling problems. We use the term *closed-loop scheduling* to describe the application of MPC specifically to production planning and scheduling problems.

Although powerful, feedback is not a panacea. If applied improperly or naively, feedback can result in nonintuitive, undesirable, and even unstable behavior. Moreover, robustness is not necessarily guaranteed by including feedback in a optimization-based controller. Thus, the design and analysis of these feedback methods is imperative to implementing optimization-based control and automation in practice. While appreciated in process control after many years of research, these caveats of feedback are not always recognized in research fields that are just beginning to incorporate feedback in their algorithms, e.g., closed-loop or online production scheduling.

Under suitable assumptions, nominal MPC is known to be inherently robust to sufficiently small and persistent disturbances, e.g., perturbations, model errors, measurement noise (Grimm et al., 2004; Yu et al., 2014; Allan et al., 2017). Note that we add the term *nominal* to emphasize that no disturbance information is present in the MPC formulation or optimization problem. This definition of robustness considers a deterministic realization of the disturbance and bounds the worst possible performance of the system subject to a given disturbance trajectory. While this property is important and the deterministic description convenient, disturbances in process control applications are usually driven by a stochastic process and not an adversarial opponent. Thus, a stochastic notion of robustness offers an instructive complement to the deterministic definition of robustness typically considered for MPC. While the idea of simulating and analyzing a closed-loop system subject to stochastic disturbances is hardly novel (see Kushner (1965)), theoretical results pertaining to the stochastic robustness of MPC are unavailable. Even for *stochastic* MPC (SMPC) algorithms, which include a stochastic disturbance model explicitly in the optimization problem, theoretical results are often restricted to linear systems. Furthermore, these results assume that the disturbance distribution used to formulate the SMPC algorithm is equivalent to the disturbance distribution of the underlying plant, an assumption that does not hold in any practical setting.

These limitations of the current theoretical results are particularly unfortunate because they render any rigorous comparison of nominal and stochastic MPC impossible. Moreover, there are applications of MPC, such as closed-loop scheduling, in which the most relevant class of disturbances are not adequately addressed with deterministic robustness results. The organizing theme of this dissertation is therefore: the stochastic robustness of model predictive control. Throughout these chapters, we introduce modified and novel mathematical techniques and definitions to analyze and characterize the robustness of nominal and stochastic MPC to stochastic disturbances.

1.1 Outline

We now briefly introduce each of the topics covered in the following chapters and highlight the major contributions in this dissertation.

Robustness and model predictive control

The term robustness is sometimes treated as an abstract or subjective concept in industrial practice. In control theory, however, the term robustness is given a specific mathematical definition that is constructed to characterize the qualitative notions of robustness that are desired in most industrial applications. In general, these qualitative notions can be summarized as: “small implies small,” i.e., an arbitrarily small disturbance produces a similarly small deviation from the goal of the control algorithm. If instead an arbitrarily small error (e.g., rounding error in an computation) can produce significant degradation of the controllers performance, the algorithm is not robust.

While the mathematical form of “small implies small” can be adequately captured with the (relatively) recent advent of comparison functions, the remaining question is how to describe the goal or performance of the controller. Often, this goal is best described as deviation from a specified setpoint or target for the system, but applications of MPC in chemical engineering may frequently define and prioritize performance in terms of *economic* measures represented by the stage cost prescribed to the MPC algorithm. We therefore define robustness in terms of both these potential measures of performance, i.e., distance to the setpoint and value of the stage cost. For stochastic systems and therefore stochastic robustness, we are faced with a similar question: What stochastic property of this performance metric is to be considered? We select *expected value* as the stochastic property to use in these definitions of stochastic robustness given the ease of interpretation and intuitive understanding inherent to this property.

These definitions of robustness are instructive in that they adequately describe *what* closed-loop behavior is desired. As to *how* this closed-loop behavior is to be achieved, we turn to the specific control framework of MPC. In nominal MPC, an approximate and nominal model of the system dynamics is used together with a stage cost characterizing the objective of the control algorithm to formulate an optimization problem. This optimization problem is solved to determine an optimal trajectory of inputs (i.e., decisions) for the system, but only the *first* input in this trajectory is implemented. At the next time step, the state of the system is updated via feedback (e.g., measurements), and the optimization problem is solved again to determine the next input to the system. A graphical depiction of MPC is provided in Figure 1.3, in which the optimal trajectory at each time step, called the *open-loop* trajectory, is shown in the gray region. The *closed-loop* trajectory, including the selected inputs and observed states, is shown in the solid colors. This feedback is the key feature of MPC that provides the control algorithm inherent robustness to disturbances.

In this chapter, we discuss in detail these definitions of deterministic and stochastic robustness for closed-loop systems. We then establish that the closed-loop system generated by nominal MPC is robust in terms of these definitions of deterministic and stochastic robustness, with respect to either distance to the setpoint or value of the stage cost, for sufficiently small (but nonzero) disturbances.

Stochastic model predictive control

The robustness of this nominal MPC algorithm is achieved without incorporating any specific disturbance information in the optimization problem. In recent decades, however, advances in stochastic optimization offer the possibility to incorporate disturbance information directly in the MPC problem. Using the same feedback methodology as MPC, these SMPC formulations also include a stochastic description of the disturbances in the optimiza-

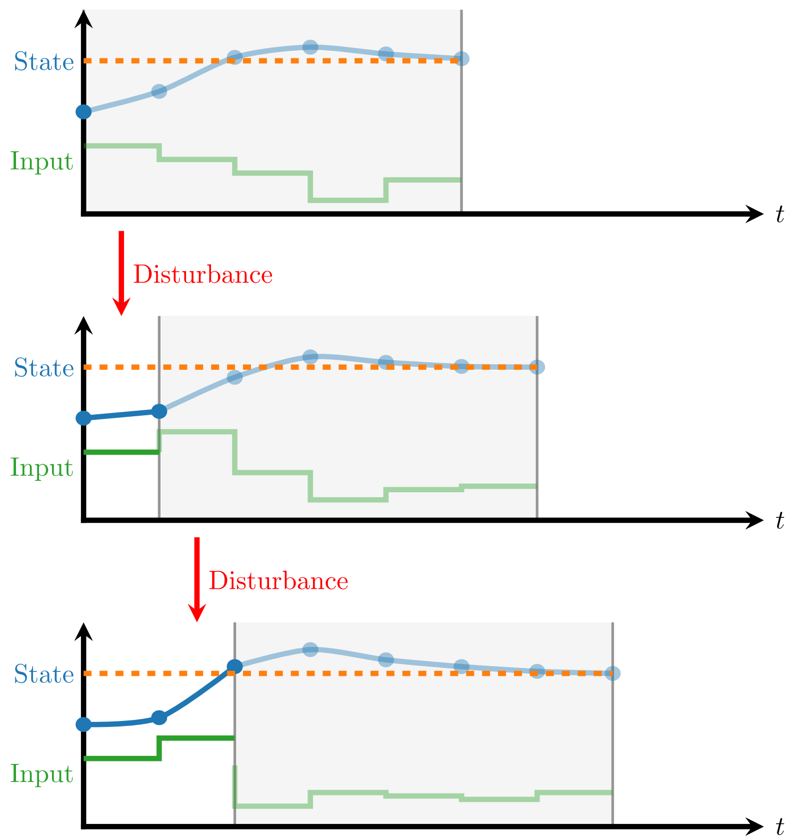

Figure 1.3: Diagram of an MPC controller. The optimal input and state trajectory at each time step is shown in the gray region, while the implemented input and observed state are shown in the solid colors.

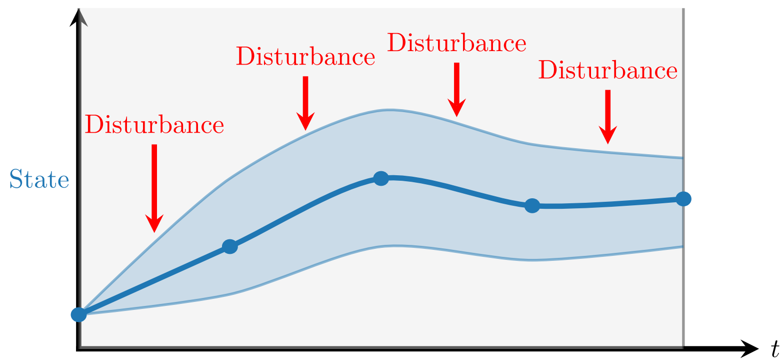

Figure 1.4: Sketch of an optimal state trajectory for an SMPC optimization problem. The faded blue region shows the distribution of states for the optimal solution, while the solid curve shows the expected value of the state.

tion problem, i.e., a probability distribution. Thus, the optimization problem is to minimize the expected value of the cost function for a distribution of states generated by these disturbances. A sketch of an optimal state trajectory and associated distribution are shown in Figure 1.4. Note that the disturbances are added at each time step and the cost function is evaluated based on the distribution of states and not the expected value of the state trajectory.

The goal of SMPC is to produce an algorithm that is more robust than nominal MPC to the specific disturbance and associated distribution of interest. There are, however, numerous limitations to the current theoretical results available for SMPC with most of these results restricted to linear systems. Moreover, the notion of "more robust" is not clearly defined, resulting in an assortment of theoretical results that are not comparable to the theoretical results derived for nominal MPC.

On the topic of SMPC, we provide two main contributions in this chapter. The first is to clearly define stochastic robustness for SMPC in a manner that subsumes the definition of stochastic robustness used for nominal MPC. The definitions are therefore comparable. We then establish that SMPC, under suitable assumptions, renders the closed-loop system robust in this context for the specific disturbance distribution used to formulate the SMPC optimization problem. With these results in hand, the second contribution of this chapter is compare, through theory and examples, the stochastic robustness of nominal and stochastic MPC.

Distributional robustness

One the ubiquitous and yet most impractical assumption made throughout the entirety of theoretical results, analysis, and simulation-studies for SMPC is that the stochastic description of uncertainty used in the optimization problem is equivalent to the stochastic uncertainty in the plant. In practice, these stochastic models are estimated from large volumes of operational data and therefore subject to their own type of uncertainty, often called *distributional* uncertainty. As suggested by the name, distributional uncertainty refers to uncertainty in the probability distribution used as a model for the system.

The somewhat tacit assumption is that feedback provides SMPC some margin of robustness to these distributional uncertainties in the same manner that feedback provides nominal MPC with some margin of robustness to disturbances. To the best of our knowledge, however, this conjecture is not proven, or even explicitly defined, in any SMPC literature. In fact, we were unable to locate any suitable definition of distributional robustness for closed-loop systems in the larger field of stochastic optimal control.

In this chapter, we present a novel definition of distributional robustness for closed-loop systems, which uses the Wasserstein metric to quantify the differences between probability distributions. With suitable modifications to the assumptions typically used for SMPC, we then establish that SMPC is in fact distributionally robust to sufficiently small errors in the stochastic model of the system. In Figure 1.5, we sketch three examples of potential distributional modeling errors covered by this definition of robustness. Note that these errors also include sampling-based approximations, as shown in the third plot of Figure 1.5, which are

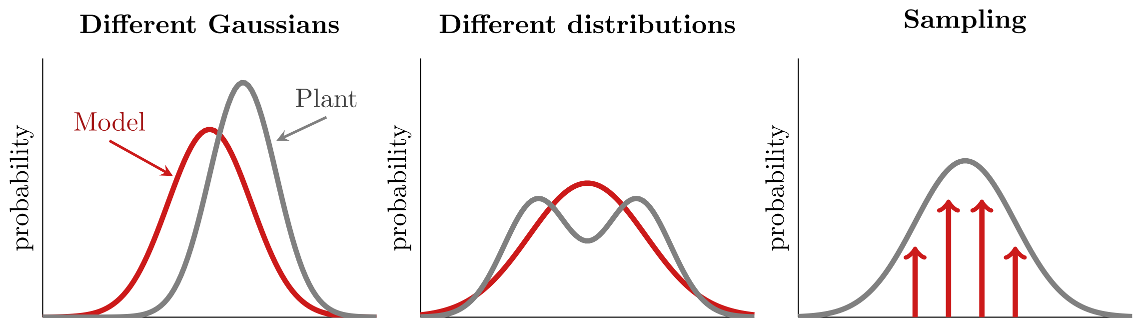

Figure 1.5: Sketches of potential errors between the probability distribution used in the disturbance model (red) and the probability distribution of the disturbances in the plant (gray).

frequently necessary to solve the stochastic optimization problem generated by SMPC. With this result, we address both inaccuracies in the dynamical model of the system as well as the probability distribution of the disturbance. Moreover, we demonstrate that this result unifies the analysis of nominal and stochastic MPC for stochastic closed-loop systems.

Large and infrequent disturbances

These robustness results for nominal and stochastic MPC are strong in that they allow any probability distribution for the disturbance, but weak in that they apply for only *small* disturbances. While the disturbances encountered in most process control applications are well described as small (e.g., model errors and measurement noise), this class of disturbances is no longer sufficient for more recent applications of MPC in higher-level decision making problems, e.g., the production planning and scheduling layer. In these higher-level decision making problems, disturbances are often discrete-valued and best described as *large*. For example, breakdowns or delays in a production scheduling problem can produce significant disruptions in the system and render a manufacturing facility unproductive for hours. In Figure 1.6, we show a comparison of the small, persistent disturbances typical of process control applications and the large, infrequent disturbances encountered in higher-level planning and scheduling problems.

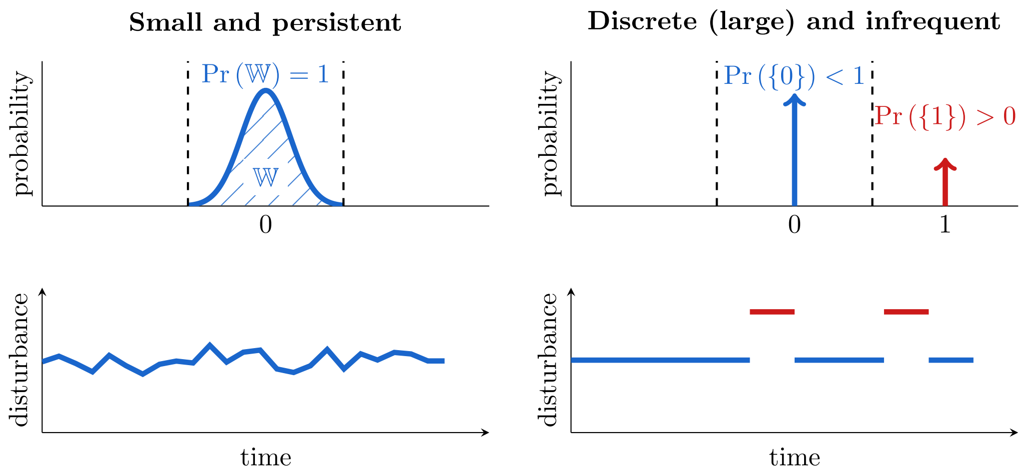

Figure 1.6: Example distributions for a small, persistent disturbance (left) and a discrete and infrequent disturbance (right). Example disturbance trajectories drawn from these distributions are shown in the bottom plots.

These large disturbances, however, are also *infrequent*, in that the probability they occur is small. For this class of large and infrequent disturbances, we therefore provide a different definition of robustness. This definition is strong in that allows large disturbances, but weak in that it applies for only certain probability distributions of the disturbance. We then establish that nominal MPC, under suitable assumptions, provides this property of robustness defined for large and infrequent disturbances. Since the applications of interest in this chapter are often time-varying, we also define nominal MPC and these results for a time-varying system.

Closed-loop scheduling

In a scheduling problem, limited resources are allocated to complete tasks at specific points in time and thereby achieve a specified goal for a manufacturing facility or logistics service. This specified goal is typically a profit maximization or cost minimization objective. An example manufacturing facility and potential schedule are shown in Figure 1.7. Over the last three decades, an extensive body of literature has been developed that focused on formulating these scheduling problems as optimization problems that can then be solved with

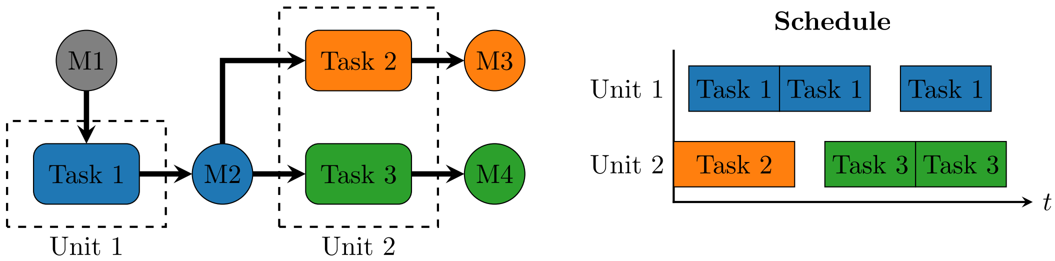

Figure 1.7: Diagram of a manufacturing facility (left) with tasks 1-3 that must be run on units 1 or 2 to consume/produce materials M1-M4 and a corresponding schedule (right) for this manufacturing facility specifying task/unit/time assignment.

available optimization algorithms.

The world, however, seldom accommodates even the best of plans. In practice, disturbances also arise in the production planning and scheduling layer of facility operation in the from of delays, breakdowns, and production yield losses. Note that the most pertinent class of disturbances in production scheduling problems are large and infrequent. In closed-loop scheduling, these disturbances are addressed with the same feedback structure used in MPC, i.e., the schedule is reoptimized with updated facility information at regular intervals. By casting this closed-loop scheduling algorithm in the framework of MPC, we can use the results developed for the stochastic robustness of MPC to define and analyze the robustness of closed-loop scheduling. While there are some simulation-based studies that investigate the performance of closed-loop scheduling for specific case studies, there are no theoretical results that define or establish the robustness of these closed-loop scheduling algorithms.

In this chapter, we use the stochastic robustness results developed throughout this dissertation to construct a closed-loop scheduling algorithm that is guaranteed to be robust to large and infrequent disturbances. Though examples, we demonstrate that this definition of robustness is appropriate for production scheduling and prevents the myopic behavior that is allowed by more naive closed-loop scheduling algorithms. We then extend this algorithm design and associated theoretical results to address practical implementation concerns such as the computational demand of online optimization.

1.2 Notation and basic definitions

We use fairly standard notation, but introduce some of this notation and some basic definitions to ensure clarity in the rest of this dissertation. Let $\mathbb{I}$ and $\mathbb{R}$ denote the integers and reals. Let subscripts denote dimensions and subscripts denote restricts of the integers and reals (e.g., $\mathbb{R}^n_{\geq 0}$ for nonnegative reals and $\mathbb{I}_{0:N}$ for the integers $\{0, 1, \ldots, N\}$). Let $|\cdot|$ denote the Euclidean norm. If not otherwise stated, assume all ambiguously defined sets X are subsets of the reals.

A function $f : X \to \mathbb{R}$ is called lower semicontinuous if the set $\{x \in X : f(x) \leq y\}$ is closed for every $y \in \mathbb{R}$ or equivalently $\liminf_{x \to x_0} f(x) \geq f(x_0)$ for every $x_0 \in X$. A function $f : X \to \mathbb{R}$ is called upper semicontinuous if the set $\{x \in X : f(x) \geq y\}$ is open for all $y \in \mathbb{R}$ or equivalently $\limsup_{x \to x_0} f(x) \leq f(x_0)$ for all $x_0 \in X$. A function that is both lower and upper semicontinuous is continuous. A function $f : X \to Y$ is called Lipschitz continuous if there exists $L \in \mathbb{R}_{\geq 0}$ such that $|f(x_1) - f(x_2)| \leq L|x_1 - x_2|$ for all $x_1, x_2 \in X$. For $X \subseteq \mathbb{R}^n$, $f : X \to Y$ is locally Lipschitz continuous if $f(\cdot)$ is Lipschitz continuous on any compact subset of X.

Let $\mathcal{B}(\Omega)$ denote the Borel algebra of the set Ω, i.e., the collection of all sets that can be formed by countable unions, intersections, and relative complements of all open subsets of Ω. A set $F \subseteq \mathbb{R}^n$ is Borel measurable if $F \in \mathcal{B}(\mathbb{R}^n)$. A function $f : X \to Y$ is Borel measurable if for each open set $O \subseteq Y$, the set $f^{-1}(O) := \{x \in X : f(x) \in O\}$ is Borel measurable, i.e., $f^{-1}(O) \in \mathcal{B}(X)$. A set-valued mapping denoted $S : X \rightrightarrows Y$ is the assignment of each $x \in X$ to a set $S(x) \subseteq Y$. A set-valued mapping $S : X \rightrightarrows Y$ is Borel measurable if for every open set $O \subseteq Y$, the set $S^{-1}(O) := \{x \in X : S(x) \cap O \neq \emptyset\}$ is Borel measurable, i.e., $S^{-1}(O) \in \mathcal{B}(X)$ (Rockafellar and Wets, 1998).

Chapter 2

Robustness and Model Predictive Control

As they form the basis of all subsequent topics discussed in this dissertation, we begin with a thorough introduction to closed-loop stochastic systems, robustness, and model predictive control (MPC). In addition to the introductory nature of this chapter, however, we also introduce and justify a novel definition of stochastic robustness for closed-loop systems. We then establish sufficient conditions for this definition of stochastic robustness and show that this property is in fact implied by a more common definition of deterministic robustness frequently used in MPC analysis. Moreover, we also introduce more general definitions of robustness that depart from the notion of "distance to the origin" that is typical in control theory. These new definitions instead address the robustness of closed-loop (stochastic) systems with respect to some general performance metric defined for the system, e.g., an economic cost function. With these new definitions and results, we then establish that MPC produces a closed-loop system that is robust in both a deterministic and stochastic context and with respect to the stage cost, i.e., performance metric, supplied to the MPC problem formulation.

2.1 State-space dynamical systems

We describe dynamical systems via a state-space representation in which the system, at each instant in time, is fully characterized by the state vector $x \in \mathbb{R}^n$. Information about past history of the system that remains relevant to the future behavior of the system is contained in this state. The state of the system is influenced by the inputs $u \in \mathbb{R}^m$ that must be chosen by some control algorithm. The goal of controller design is to define a control algorithm for the system in order to achieve a predefined objective (e.g., minimizing operation cost or tracking a setpoint). The system may also be perturbed by a disturbance $w \in \mathbb{R}^q$ that is unknown a priori, although we may be able to provide the control algorithm with a stochastic and/or worst-case description of this disturbance.

Engineering models derived from conservation laws (mass, molecules, energy, momentum) typically produce ordinary differential equations (ODEs) of the following form,

$$\frac{dx}{dt} = F(x, u, w) \tag{2.1}$$

in which $F : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^n$ is the continuous time system model. Generally speaking, the state x may be defined to include any variables such that (2.1) holds. The trajectories of the state, input, and disturbance are functions of the time $t \geq 0$ and denoted $x(\cdot)$, $u(\cdot)$, $w(\cdot)$. Given (2.1), the state $x(t)$ at some future time $t \geq 0$ is then fully described by the initial state $x(0)$ and a the trajectories of $u(\cdot)$ and $w(\cdot)$ for $[0, t]$.

We live, however, in the digital age and therefore control actions are rarely implemented in a truly analog (continuous time) fashion. Instead, measurements and state estimates are made at discrete time points with a fixed sampling interval Δ. The input is then defined as a piecewise-constant function, i.e., a zero-order hold, such that the input is constant on the interval $[k\Delta, (k+1)\Delta)$.[1] Thus, discrete time representations of the system are common in

[1]Piecewise-linear inputs, i.e., a first-order hold, or higher-order polynominal functions are also used.

control literature for both theoretical and practical reasons.

In discrete time, the system evolution is given by

$$x(k+1) = f(x(k), u(k), w(k)) \tag{2.2}$$

in which $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^n$ is the discrete time system model. The trajectories of the state, input, and disturbance are now function of the time step $k \in \mathbb{I}_{\geq 0}$ and denoted by the sequences $\mathbf{x} := (x(0), x(1), \dots)$, $\mathbf{u} := (u(0), u(1), \dots)$, and $\mathbf{w} := (w(0), w(1), \dots)$. Note that for a constant sample time and constant input/disturbance, the continuous time system in (2.1), assuming $F(\cdot)$ is a continuous function, can be converted, in theory, to a discrete time system in (2.2) such that the discrete time states $x(k)$ are equivalent to the continuous time states $x(k\Delta)$ for all $k \in \mathbb{I}_{\geq 0}$. In practice, however, continuous time nonlinear ODEs are converted into discrete time systems via approximations such as Runge-Kutta or orthogonal collocation.[2]

Often, the state of the system cannot be directly measured. Instead, the state must be inferred or estimated from the measurements that are available. We denote these measurements or outputs as $y \in \mathbb{R}^p$. These outputs are then taken to be a function of the current state x and input u along with a disturbance $v \in \mathbb{R}^p$ to represent measurement noise inherent to instrumentation (e.g., pressure sensors, thermocouples) as shown in the following equation.

$$y = h(x, u) + v \tag{2.3}$$

The mapping $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$, however, is not typically "one-to-one", i.e., the value of x is not uniquely determined by y and u, even if $v = 0$. Instead, we must estimate the current state of the system from a trajectory of previous y's and u's. With this state-space representation,

[2] By contrast, linear ODEs, i.e. $\frac{dx}{dt} = Ax + Bu + Gw$, can be exactly discreteized (to within machine precision) via matrix exponentials.

controller design is typically broken into two separate problems: Determining this state from past data is termed *state estimation* while the use of this estimated state in determining an input is termed *regulation*. While the topic of state estimation is essential to MPC, we focus this dissertation entirely on the regulation problem and therefore work with the state x, rather than the outputs y.

2.2 Closed-loop stochastic systems

We consider the following discrete time system, written using shorthand notation,

$$x^+ = f(x,u,w) \quad f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^n \tag{2.4}$$

in which $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the controlled input, and $w \in \mathbb{R}^q$ is a disturbance (random variable), and x^+ denotes the successor state. Frequently, the input is subject to constraints (e.g., min/max flow rates) denoted by the set $\mathbb{U} \subseteq \mathbb{R}^m$ and therefore the input is required to satisfy $u \in \mathbb{U}$. We also use $\mathbb{W} \subseteq \mathbb{R}^q$ to denote the set of possible disturbance values, i.e., the disturbance must satisfy $w \in \mathbb{W}$.

We treat the origin ($x = u = 0$) as the steady-state target (setpoint) for the controller, without loss of generality, and consider the following regularity assumption.

Assumption 2.1 (Continuity of system)**.** The system $f : \mathbb{R}^n \times \mathbb{U} \times \mathbb{W} \to \mathbb{R}^n$ is continuous and satisfies $f(0,0,0) = 0$.

Note that this assumption is without loss of generality because any system $\bar{x}^+ = \bar{f}(\bar{x},\bar{u},w)$ and steady-state pair $(\bar{x}_s,\bar{u}_s)$ in the original variable space can be shifted to the origin via deviation variables defined as $x := \bar{x} - \bar{x}_s$ and $u := \bar{u} - \bar{u}_s$. The system is then given by

$$x^+ = f(x,u,w) := \bar{f}(x + \bar{x}_s, u + \bar{u}_s, w) - \bar{x}_s$$

and $(x,u)=(0,0)$ if and only if $(\bar{x},\bar{u})=(\bar{x}_s,\bar{u}_s)$. Moreover, $f(0,0,0)=\bar{f}(\bar{x}_s,\bar{u}_s,0)-\bar{x}_s=0$ because $(\bar{x}_s,\bar{u}_s)$ is a steady state for the system.

We consider w to be a random variable and therefore (2.4) is a stochastic process. In this dissertation, we use a measure-theoretic description of this random variables and corresponding stochastic process. This measure-theoretic framework is particularly useful for the analysis of stochastic MPC in the following two chapters. However, this framework is notably different than the treatment of random variables presented in many engineering courses. We therefore define and discuss these potentially unfamiliar topics in some detail before moving to any discussion of robustness. We begin with the following assumption for the disturbances.

Assumption 2.2 (Disturbances)**.** The disturbances $w\in\mathbb{W}$ are random variables that are independent and identically distributed (i.i.d.) in time and zero mean. The set $\mathbb{W}$ is compact and contains the origin.

Given Assumption 2.2, each random variable w has an equivalent probability measure that we denote $\mu:\mathcal{B}(\mathbb{W})\rightarrow[0,1]$. Recall that $\mathcal{B}(\mathbb{W})$ denotes the Borel algebra of the set $\mathbb{W}$. This probability measure maps an event, which is a Borel measurable subset of $\mathbb{W}$, to a scalar value between zero and one, which is the probability of the event occurring. While seemingly more complicated than the typical treatment of random variables found in engineering courses, a probability measure can be viewed as a generalization of a more familiar notion of probability with the following relation.

$$\mu([a,b])=\Pr\left(a\leq w\leq b\right)$$

in which both sides denote the probability that the random variable $w\in\mathbb{R}$ takes a value between $a\in\mathbb{R}$ and $b\in\mathbb{R}$ with $a\leq b$. The left-hand side, however, also admits a rigorous definition as a mathematical object that we now introduce.

Definition 2.3 (Probability measure)**.** The function $\mu:\mathcal{B}(\mathbb{W})\rightarrow[0,1]$ is a (Borel) probability measure on the set $\mathbb{W}$ if $\mu(\emptyset)=0$, $\mu(\mathbb{W})=1$, and for all countable collections $\{E_i\}_{i\in I}$ of pairwise disjoint sets ($E_i \cap E_j = \emptyset$ if $i \neq j$) satisfying $E_i \in \mathcal{B}(\mathbb{W})$ we have that

$$\mu\left(\bigcup_{i\in I} E_i\right) = \sum_{i\in I} \mu(E_i)$$

Note that this definition is generally in line with our intuition. The probability that nothing happens ($\mu(\emptyset)$) is zero and the probability that something happens ($\mu(\mathbb{W})$) is one. The last condition is in fact a generalization of a more familiar statement for 2 mutually exclusive events:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) \text{ if } A \cap B = \emptyset$$

With this probability measure, we define expected value as the following Lebesgue integral.

$$\mathbb{E}\left[w\right] = \int_{\mathbb{W}} w d\mu(w)$$

which indicates that we are integrating the variable w, over the set $\mathbb{W}$, with the associated measure μ. Thus, $d\mu(w)$ replaces the (continuous) probability density function $p(w)dw$ more commonly found in engineering literature.[3]

One useful feature of the probability measure is that it provides a simple means to combine discrete and continuous probability distributions into a single framework while avoiding the Dirac delta function. Instead, we use the *Dirac measure* that we denote δ_w for some point $w \in \mathbb{W}$ and is defined as

$$\delta_w(S) := \begin{cases} 0, & w \notin S \\ 1, & w \in S \end{cases}$$

for any $S \in \mathcal{B}(\mathbb{W})$. Thus, we can define a discrete probability distribution with $s \in \mathbb{I}_{\geq 1}$

[3]If μ is absolutely continuous with respect to w, we can define $p(w) := \frac{d\mu}{dw}$ in which $\frac{d\mu}{dw}$ is the Radon-Nikodym derivative.

discrete points of equal probability as

$$\mu_d(\cdot) = \frac{1}{s}\sum_{i=1}^{s}\delta_{w_i}(\cdot)$$

in which $\{w_i\}_{i=1}^s$ denotes the sequence of discrete points. All of the subsequent results in this dissertation therefore apply for discrete, continuous, and mixed distributions (e.g., $\mu = (\mu_d + \mu_c)/2$ in which μ_d is a discrete distribution and μ_c is a continuous distribution). The benefit of replacing the Dirac delta function with the Dirac measure is that we can easily establish continuity for functions defined by integrals, regardless of the distribution chosen for μ.

We use $\mathcal{M}(\mathbb{W})$ to denote the collection of all possible probability measures $\mu : \mathcal{B}(\mathbb{W}) \to [0,1]$ that satisfy Assumption 2.2, i.e.,

$$\int_{\mathbb{W}} w d\mu(w) = 0 \quad \forall\, \mu \in \mathcal{M}(\mathbb{W})$$

Since $\mathbb{W}$ is bounded, the second moment of w is finite. For any $\mu \in \mathcal{M}(\mathbb{W})$, we denote the covariance matrix of w as

$$\Sigma := \mathbb{E}\left[(w - \mathbb{E}[w])(w - \mathbb{E}[w])'\right] = \mathbb{E}[ww'] = \int_{\mathbb{W}} ww' d\mu(w)$$

in which the equality holds because $\mathbb{E}[w] = 0$ by Assumption 2.2.

For the i.i.d. random variables $(w(i), w(i+1), \ldots, w(i+N-1))$ and $N \in \mathbb{I}_{\geq 1}$, their joint distribution measure $\mu^N : \mathcal{B}(\mathbb{W}^N) \to [0,1]$ is defined as

$$\mu^N(F) := \mu(F_i)\mu(F_{i+1})\ldots\mu(F_{i+N-1}) \tag{2.5}$$

for all $F = (F_i, F_{i+1}, \ldots, F_{i+N-1}) \in \mathcal{B}(\mathbb{W}^N)$. We also define the sequence of random variables starting from $i = 0$ to time step $k \in \mathbb{I}_{\geq 0}$ as $\mathbf{w}_k := (w(0), w(1), \ldots, w(k-1))$. For this sequence, we define the expected value of a Borel measurable function $g : \mathbb{W}^k \to \mathbb{R}_{\geq 0}$ as the following Lebesgue integral.

$$\mathbb{E}\left[g(\mathbf{w}_k)\right] := \int_{\mathbb{W}^k} g(\mathbf{w}_k) d\mu^k(\mathbf{w}_k)$$

By presenting expected value as a Lebesgue integral, we can establish many useful and fundamental mathematical properties of the stochastic system that may remain unclear with typical probability notation. These properties are particularly relevant in the subsequent chapter on SMPC in which these Lebesgue integrals are evaluated within the optimization problem.

We now introduce the following results for random variables that are used throughout.

Jensen's inequality: Let $\omega \in \Omega$ be a random variable with the probability measure $P : \mathcal{B}(\Omega) \to \mathbb{R}$. Let $g : \Omega \to A$ be a Borel measurable function and $\phi : A \to \mathbb{R}$ be a convex function. Then,

$$\phi\left(\int_{\Omega} g(\omega) dP(\omega)\right) \leq \int_{\Omega} \phi(g(\omega)) dP(\omega) \tag{2.6}$$

and the opposite inequality holds for $\phi(\cdot)$ concave.

Lemma 2.4. *If Assumption 2.2 holds, then* $\mathbb{E}\left[|w|\right] \leq \sqrt{\text{tr}(\Sigma)}$.

Proof. We use Jensen's inequality to show that $\mathbb{E}\left[|w|\right]^2 \leq \mathbb{E}\left[|w|^2\right]$. We then use the fact that $w'w = \text{tr}(w'w) = \text{tr}(ww')$ and note that $\text{tr}(\cdot)$ is a linear operator to give

$$\mathbb{E}\left[|w|^2\right] = \mathbb{E}\left[w'w\right] = \mathbb{E}\left[\text{tr}(w'w)\right] = \mathbb{E}\left[\text{tr}(ww')\right] = \text{tr}(\mathbb{E}\left[ww'\right]) = \text{tr}(\Sigma)$$

Thus, $\mathbb{E}\left[|w|\right]^2 \leq \text{tr}(\Sigma)$ and we take the square root of both sides to complete the proof. □

A state-feedback controller, i.e., control law, is a function that determines the input $u \in \mathbb{U}$ based on the current state of the system x. For MPC, this function is implicitly defined by optimizing a performance metric, i.e., the state cost $\ell(x, u)$, over a finite prediction horizon. Typically, this optimization problem can be solved for only a subset of states $\mathcal{X} \subseteq \mathbb{R}^n$ and the MPC control law is therefore defined for only states in the feasible set $\mathcal{X}$. Thus, we consider a generic control law denoted $\kappa : \mathcal{X} \to \mathbb{U}$ and defined on some subset of the reals, i.e., $\mathcal{X} \subseteq \mathbb{R}^n$. The resulting closed-loop system is then

$$x^+ = f(x, \kappa(x), w) \tag{2.7}$$

In general, the goal of this control law is to drive the state of the system to the origin, i.e., the steady-state target, but other performance metrics and goals can also be considered.

To ensure that the iteration in (2.7) remains well defined, we must also ensure that the state remains within $\mathcal{X}$ for any possible realization of the disturbance. For optimization-based controllers, this property is known as *robust recursive feasibility* of the optimization problem. We characterize this property through positive invariance and robust positive invariance.

Definition 2.5 (Positive invariance)**.** A set $\mathcal{X}$ is positive invariant for the system $x^+ = f(x, \kappa(x), 0)$ if $x^+ \in \mathcal{X}$ for all $x \in \mathcal{X}$.

Definition 2.6 (Robust positive invariance)**.** A set $\mathcal{X}$ is robustly positive invariant (RPI) for the system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ if $x^+ \in \mathcal{X}$ for all $x \in \mathcal{X}$ and $w \in \mathbb{W}$.

If the feasible set of the MPC optimization problem $\mathcal{X}$ is RPI, then the optimization problem is robustly recursively feasible.

By defining this control law and ensuring that $\mathcal{X}$ is RPI, the state at some future time is now fully defined by the initial state of the system and the disturbance sequence. We use the function $\phi(k; x, \mathbf{w}_k)$ to denote the state at a future time $k \in \mathbb{I}_{\geq 0}$, given the initial state $x \in \mathcal{X}$ at $k = 0$, and the disturbance sequence $\mathbf{w}_k \in \mathbb{W}^k$ subject to the iteration in (2.7), i.e.,

$\phi(0; x, \{\}) := x$ and

$$\phi(k+1; x, \mathbf{w}_{k+1}) = f(\phi(k; x, \mathbf{w}_k), \kappa(\phi(k; x, \mathbf{w}_k)), w(k))$$

for all $x \in \mathcal{X}$, $\mathbf{w}_{k+1} \in \mathbb{W}^{k+1}$, and $k \in \mathbb{I}_{\geq 0}$. Note that we do not assume that $\kappa(\cdot)$ is continuous on $\mathcal{X}$ and therefore $f(x, \kappa(x), w)$ and $\phi(k; x, \mathbf{w}_k)$ are not continuous with respect to $x \in \mathcal{X}$. We do, however, require that $\kappa(\cdot)$ is Borel measurable to ensure that $\phi(k; x, \mathbf{w}_k)$ is Borel measurable and therefore all stochastic properties, e.g., expected value, of the closed-loop system are well defined.[4] For any of the subsequent control methods discussed in this dissertation, we verify that $\kappa(\cdot)$ is Borel measurable.

In the subsequent sections (and chapters), we discuss the robustness of closed-loop systems in terms of both deterministic and stochastic metrics of the closed-loop state trajectory $\phi(\cdot)$. We begin, however, by introducing these notions of robustness through a discussion of linear systems and the linear-quadratic regulator.

2.3 The linear case

Before proceeding to the definitions of deterministic and stochastic robustness, we begin with the simplest case: a linear, unconstrained system. Specifically, we consider the system

$$x^+ = f(x, u, w) = Ax + Bu + w \tag{2.8}$$

in which $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$ are matrices and there are no constraints on u, i.e., $\mathbb{U} = \mathbb{R}^m$. We then use the *linear-quadratic regulator* (LQR) to define the control law for this system.

The LQR control law is defined by an *infinite* horizon optimization problem based on the

[4]Note that Borel measurable functions are closed under composition. By induction we establish that $\phi(\cdot)$ is Borel measurable if $\kappa(\cdot)$ and $f(\cdot)$ are Borel measurable (continuous).

nominal evolution of the system ($w = 0$). The cost function is defined as

$$V(x, \mathbf{u}) := \lim_{N\to\infty} \sum_{k=0}^{N-1} \Big(x(k)'Qx(k) + u(k)'Ru(k) \Big) + x(N)'Qx(N)$$

in which $x(k+1) = Ax(k) + Bu(k)$ is the nominal system evolution, $x(0) = x$ is the initial state, and $\mathbf{u} \in \mathbb{U}^\infty$ is an infinite trajectory of control actions. The matrices $Q, R \succ 0$ are tuning parameters chosen to reflect the relative importance of deviations from the origin for different elements of the state and input. The optimization problem with this cost function is written as

$$V^0(x) = \min_{\mathbf{u}} V(x, \mathbf{u})$$

for all $x \in \mathbb{R}^n$ in which $\mathbf{u}^0(x) = (u^0(0), u^0(1), \dots)$ denotes the optimal solution.

If the system $x^+ = Ax + Bu$ is stabilizable, then a solution to this optimization problem exists for all $x \in \mathbb{R}^n$ (Caines and Mayne, 1970; Anderson and Moore, 1981).[5] Moreover, one can use the convenient features of linear systems and quadratic cost functions to solve this infinite horizon optimization problem via dynamic programming (Bertsekas, 1987, p. 58-64). Specifically, the solution to this infinite horizon optimization problem is given by the unique stabilizing solution to the discrete-time algebraic Riccati equation (DARE), i.e., the matrix $P \succ 0$ that solves

$$P = A'PA - (A'PB)(R + B'PB)^{-1}(B'PA) + Q$$

and ensures that $A + BK$ is Schur stable[6] with $K = -(B'PB + R)^{-1}(B'PA)$ (Rawlings et al., 2020, p. 25). The optimal cost is given by $V^0(x) = x'Px$ and the control law, defined by the first input in the optimal solution, is given by $\kappa(x) = u^0(0) = Kx$ for all $x \in \mathbb{R}^n$.

[5]We can also allow $Q \succeq 0$ if (A, Q) is detectable.

[6]All eigenvalues are strictly inside the unit circle on the complex plane.

The nominal closed-loop system is then

$$x^+ = f(x, \kappa(x), 0) = Ax + BKx = A_K x$$

in which $A_K := A + BK$. We can write the nominal closed-loop trajectory as

$$\phi(k; x, \mathbf{0}) = A_K^k x$$

Since the matrix A_K is Schur stable, there exists $\lambda \in (0,1)$ and $\rho \geq 0$ such that $|A_K^k x| \leq \lambda^k \rho |x|$. We can therefore write the following bound on the norm of the nominal closed-loop state trajectory.

$$|\phi(k; x, \mathbf{0})| \leq \lambda^k \rho |x|$$

with $\lambda \in (0,1)$ and $\rho \geq 0$. Thus, the norm of the nominal closed-loop state exponentially convergences to zero with a rate λ and a bound proportional the the initial condition $|x|$ at $k = 0$. A system that admits such a bound is called *exponentially stable*.

We now add the disturbance w back into the system. The resulting closed-loop system is now

$$x^+ = A_K x + w$$

The closed-loop state trajectory can be written using the variation of constants method as

$$\phi(k; x, \mathbf{w}_k) = A_K^k x + \sum_{i=0}^{k-1} A_K^{k-1-i} w(i)$$

Since A_K is Schur stable, we can derive the following bound on the norm of the closed-loop state trajectory.

$$|\phi(k; x, \mathbf{w}_k)| \leq \lambda^k c|x| + c \sum_{i=0}^{k-1} \lambda^{k-1-i} |w(i)| \tag{2.9}$$

in which $\lambda \in (0,1)$ and $c \geq 0$. We then denote $||\mathbf{w}_k|| := \max_{i \in \mathbb{I}_{0:k-1}} |w(i)|$ to give

$$|\phi(k; x, \mathbf{w}_k)| \leq \lambda^k c|x| + \frac{c}{1-\lambda}||\mathbf{w}_k|| \tag{2.10}$$

for all $x \in \mathbb{R}^n$, $\mathbf{w}_k \in \mathbb{W}^k$, and $k \in \mathbb{I}_{\geq 0}$.

For this linear system, the bound in (2.10) characterizes the deterministic robustness of this closed-loop system, i.e., we bound the norm of the closed-loop state given a specific realization of the disturbance sequence $\mathbf{w}_k$. The bound in (2.10) contains two terms: (i) an exponentially decaying term based on the initial condition of the system and (ii) a persistent term that depends on the norm of the disturbances entering the system. In particular, the bound in (2.10) ensures that small values of the disturbance $||\mathbf{w}_k||$ result in similarly small deviations from the origin (proportional to the constant $c/(1-\lambda)$) and for $||\mathbf{w}_k|| = 0$ we recover exponential stability of the nominal system. Furthermore, (2.10) ensures that for a convergent sequence of disturbances, i.e., $w(k) \to 0$ as $k \to \infty$, the state converges to the origin as well, i.e., $\phi(k; x, \mathbf{w}_k) \to 0$.

In addition to this description of deterministic robustness, we may also consider a description of stochastic robustness for this linear system. Specifically, we take the expected value of both sides of (2.9) to give

$$\mathbb{E}\left[|\phi(k; x, \mathbf{w}_k)|\right] \leq \lambda^k c|x| + c\sum_{i=0}^{k-1} \lambda^{k-1-i}\mathbb{E}\left[|w(i)|\right]$$

in which we assume that the initial state x is known exactly (with probability one). We use Lemma 2.4 to give

$$\mathbb{E}\left[|\phi(k; x, \mathbf{w}_k)|\right] \leq \lambda^k c|x| + \frac{c}{1-\lambda}\sqrt{\text{tr}(\Sigma)} \tag{2.11}$$

for all $x \in \mathbb{R}^n$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$.

For the linear case, the bound in (2.11) characterizes the stochastic robustness of the closed-loop system, i.e., we bound a stochastic property of the closed-loop system given a probability distribution for the disturbance. The bound contains two terms: (i) an exponentially decaying term based on the initial condition of the system and (ii) a persistent term that depends on the variance of the disturbance. Analogous to the deterministic bound, (2.11) ensures that small values of $\text{tr}(\Sigma)$ produce similarly small deviations from the origin and for $\Sigma = 0$ we recover exponential stability of the nominal system. Note that the bound in (2.11) in fact holds for any $\mu \in \mathcal{M}(\mathbb{W})$ and corresponding variance Σ.

While analysis of the linear case and the LQR are both instructive and useful for many control problems, we also want to address nonlinear systems and/or systems with input constraints, which are pervasive in chemical engineering problems. In the subsequent sections, we now define nonlinear extensions for these definitions of deterministic and stochastic robustness. We begin with the the definition of deterministic robustness for nonlinear systems.

2.4 Deterministic robustness

The subsequent definitions of robustness throughout this work use *comparison functions*. These functions, popularized only in the last few decades, provide a rigorous means to analyze the stability of nonlinear systems without the ε, δ-arguments that are pervasive in the classical nonlinear systems analysis. Comparison functions include the following classes of functions.

Definition 2.7 ($\mathcal{K}$, $\mathcal{K}_\infty$, and $\mathcal{KL}$ functions)**.** A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is in class $\mathcal{K}$ (written $\alpha(\cdot) \in \mathcal{K}$) if $\alpha(\cdot)$ is continuous, strictly increasing, and $\alpha(0) = 0$. A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is in class $\mathcal{K}_\infty$ if $\alpha(\cdot)$ is in class $\mathcal{K}$ and is also unbounded, i.e., $\lim_{s\to\infty} \alpha(s) = \infty$. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{I}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is in class $\mathcal{KL}$ if for every $k \in \mathbb{I}_{\geq 0}$ the function $\beta(\cdot, k)$ is in class $\mathcal{K}$ and for every $s \in \mathbb{R}_{\geq 0}$ the function $\beta(s, \cdot)$ is nonincreasing and $\lim_{k\to\infty} \beta(s, k) = 0$.

Figure 2.1 provides examples of these comparison functions to illustrate their properties.

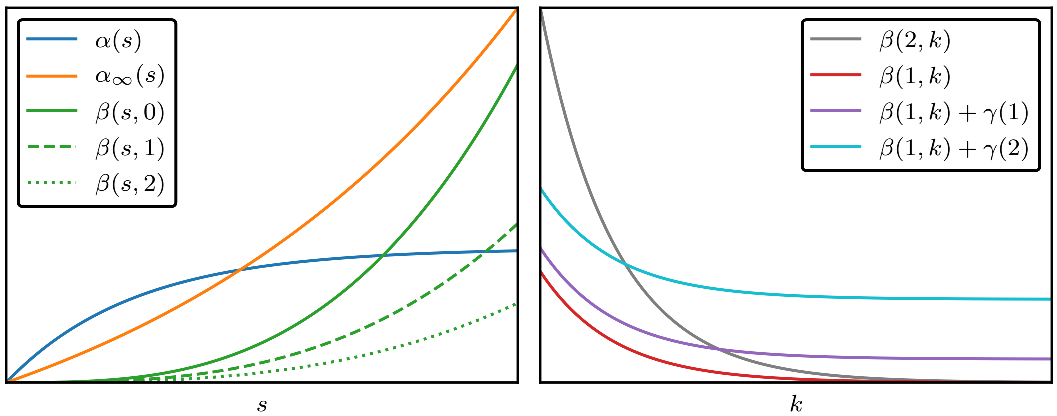

Figure 2.1: Examples of comparison functions ($\gamma(\cdot), \alpha(\cdot) \in \mathcal{K}$, $\alpha_\infty(\cdot) \in \mathcal{K}_\infty$, $\beta(\cdot) \in \mathcal{KL}$) as a function of s (left) and k (right).

Several useful properties of these functions are compiled and established in Khalil (2002, Lemma 4.2) and Kellett (2014). We note a few of these properties here. If $\alpha_1(\cdot) \in \mathcal{K}$ and $\alpha_2(\cdot) \in \mathcal{K}$, then $\alpha_1(\alpha_2(\cdot)) \in \mathcal{K}$. If $\alpha_1(\cdot) \in \mathcal{K}_\infty$, then $\alpha_1^{-1}(\cdot) \in \mathcal{K}_\infty$. If $\alpha_1(\cdot), \alpha_2(\cdot) \in \mathcal{K}$ and $\beta(\cdot) \in \mathcal{KL}$, then $\tilde{\beta}(s,k) := \alpha_1(\beta(\alpha_2(s),k))$ is a $\mathcal{KL}$-function. If $\alpha_1(\cdot) \in \mathcal{K}$, then $\alpha(a+b) \leq \alpha(2a) + \alpha(2b)$.

Before discussing robustness, we begin with a discussion of the *nominal* closed-loop system, i.e., without the disturbances ($\mathbf{w}_k = \mathbf{0}$). Specifically, we consider the following definition of asymptotic stability, which generalizes the definition of exponential stability that is common for linear systems.

Definition 2.8 (Asymptotic stability)**.** The origin is asymptotically stable for the system $x^+ = f(x, \kappa(x), 0)$ in a positive invariant set $\mathcal{X}$ if there exists $\beta(\cdot) \in \mathcal{KL}$ such that

$$|\phi(k; x, \mathbf{0})| \leq \beta(|x|, k)$$

for all $x \in \mathcal{X}$ and $k \in \mathbb{I}_{\geq 0}$.

By the restriction $\beta(\cdot) \in \mathcal{KL}$, this definition of asymptotic stability ensures that the origin is *uniformly stable* for the closed-loop system, i.e., there exists $\alpha(\cdot) \in \mathcal{K}$ such that $|\phi(k; x, \mathbf{0})| \leq$

$\alpha(|x|)$ for all $x \in \mathcal{X}$, and *uniformly attractive*, i.e., $\phi(k; x, \mathbf{0}) \to 0$ uniformly for all $x \in X$ and any compact set $X \subseteq \mathcal{X}$. Note that uniform convergence is a stronger property than pointwise convergence.[7] Exponential stability is then a special case of asymptotic stability in which $\beta(s, k) := \rho\lambda^k s$ with $\rho > 0$ and $\lambda \in (0, 1)$. An example closed-loop trajectory and possible $\mathcal{KL}$ bound are shown in Figure 2.2.

Remark 2.9. Asymptotic stability is sometimes defined as the combination of (non-uniform) local stability and (non-uniform) attractivity of the origin. We refer to this definition as the *classical* definition of asymptotic stability. For continuous nonlinear systems, i.e., $f(x, \kappa(x), 0)$ is a continuous function, the classical definition of asymptotic stability is equivalent to Definition 2.8 (Kellett and Teel, 2004, Prop. 6). Nonlinear MPC, however, can produce a discontinuous control law $\kappa(x)$. For these discontinuous nonlinear systems, the classical definition of asymptotic stability is *weaker* than Definition 2.8 and admits pathological closed-loop dynamics (McAllister and Rawlings, 2021b, Appendix A). Definition 2.8 instead implies a stronger (uniform) version of asymptotic stability and thereby excludes these pathological closed-loop dynamics.

To establish asymptotic stability for a closed-loop system, we use a *Lyapunov function*.

Definition 2.10 (Lyapunov function)**.** A function $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is a Lyapunov function for a system $x^+ = f(x, \kappa(x), 0)$ in a positive invariant set $\mathcal{X}$ if there exist $\alpha_1(\cdot), \alpha_2(\cdot), \alpha_3(\cdot) \in \mathcal{K}_\infty$ such that

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|) \tag{2.12}$$

$$V(f(x, \kappa(x), 0)) \leq V(x) - \alpha_3(|x|) \tag{2.13}$$

[7] We say that $\phi(k; x, \mathbf{0}) \to 0$ *uniformly* for all $x \in X$ if for any $\varepsilon > 0$ there exists $N \in \mathbb{I}_{\geq 0}$ such that $|\phi(k; x, \mathbf{0})| \leq \varepsilon$ for all $k \geq N$ and $x \in X$. *Pointwise* convergence instead considers only a single point $x \in X$, i.e., for any $\varepsilon > 0$ and $x \in X$ there exists $N \in \mathbb{I}_{\geq 0}$ such that $|\phi(k; x, \mathbf{0})| \leq \varepsilon$ for all $k \geq N$.

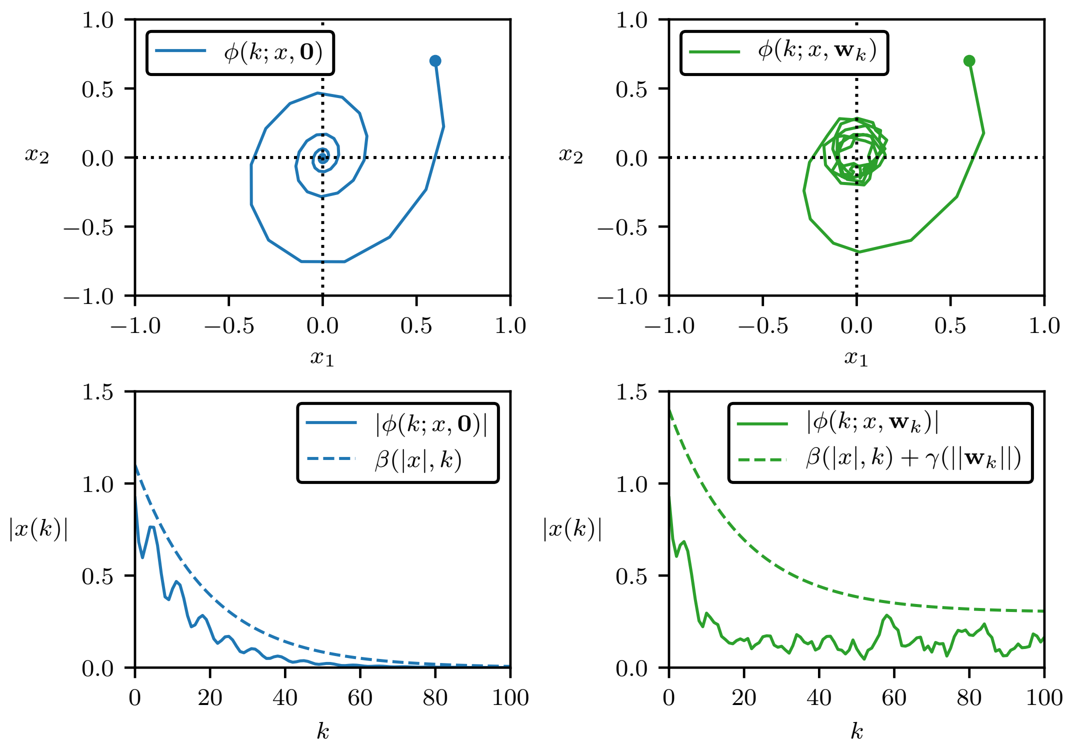

Figure 2.2: Illustration of nominal asymptotic stability (left) and robust asymptotic stability (right) in terms of phase plots (top) and norms of the state trajectories (bottom).

for all $x \in \mathcal{X}$.

If there exists a Lyapunov function for the closed-loop system, then the origin is asymptotically stable. In fact, the existence of a Lyapunov function is a *necessary and sufficient* condition for asymptotic stability. See Rawlings et al. (2020, Thm. B.15, B.17) for a proof of this result.

Proposition 2.11. *The origin is asymptotically stable for a system $x^+ = f(x, \kappa(x), 0)$ in a positive invariant set $\mathcal{X}$ if and only if there exists a Lyapunov function for this system in $\mathcal{X}$.*

Remark 2.12. For the linear system and LQR discussed in the previous section, the Lyapunov function is defined via the solution to the infinite horizon minimization problem, i.e., $V(x) = x'Px$. We have $c_1|x|^2 \leq V(x) \leq c_2|x|^2$ with $c_1, c_2 > 0$ since $P \succ 0$. Since $V(x)$ is the solution to the infinite horizon optimization problem, we also have that

$$V(x^+) - V(x) = -x'Qx - u'Ru \leq -x'Qx \leq -c_3|x|^2$$

in which $c_3 > 0$ because $Q \succ 0$. Note that $\alpha_i(|x|) := c_i|x|^2 \in \mathcal{K}_\infty$ for $i \in \{1, 2, 3\}$ and therefore $V(x)$ is an (exponential) Lyapunov function.

We now reintroduce disturbances in this closed-loop systems. In general, requiring the state of these perturbed systems to converge (uniformly or otherwise) to the origin and remain at this point is unreasonable. Instead, we require the closed-loop system to converge to some neighborhood of the origin. But this requirement is quite mild without additional conditions placed on size of this neighborhood with respect to the disturbance. The key feature of a robust system is that this neighborhood continuously deforms with the size of the disturbance, i.e., an arbitrarily small disturbance produces a similarly small perturbation from the origin. We capture this requirement by using comparison functions in the following definition of robust asymptotic stability (RAS).

Definition 2.13 (Robust asymptotic stability)**.** The origin is robustly asymptotically stable (RAS) for a system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in an RPI set $\mathcal{X}$ if there exist $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$|\phi(k; x, \mathbf{w}_k)| \leq \beta(|x|, k) + \gamma(||\mathbf{w}_k||) \tag{2.14}$$

for all $x \in \mathcal{X}$, $\mathbf{w}_k \in \mathbb{W}^k$, and $k \in \mathbb{I}_{\geq 0}$.

Note the similarities between the bound in (2.14) and (2.10). The exponential convergence is replaced by asymptotic convergence and the linear function of $||\mathbf{w}_k||$ is replaced by a nonlinear function $\gamma(\cdot) \in \mathcal{K}$. This definition of robustness is derived from the more general notion of input-to-state stability (ISS) as defined for continuous time systems in Sontag and Wang (1995) and for discrete time systems in Jiang and Wang (2002). Whereas ISS applies to a general "input" to the system, we use the term RAS to emphasize that we are considering only the disturbance as this input. By requiring $\gamma(\cdot) \in \mathcal{K}$, we ensure that small values of $||\mathbf{w}_k||$ result is similarly small values of $\gamma(||\mathbf{w}_k||)$ and as $||\mathbf{w}_k| \to 0$ we recover nominal asymptotic stability. Furthermore, RAS ensures that for a convergent sequence of disturbances, i.e., $w(k) \to 0$ as $k \to \infty$, the state converges to the origin as well, i.e., $\phi(k; x, \mathbf{w}_k) \to 0$. This convergence is also uniform on compact subsets of $\mathcal{X}$.

If $f(x, \kappa(x), w)$ is a continuous function, we can establish that nominal asymptotic stability and robust positive invariance of the set $\mathcal{X}$ implies robustness (Kellett and Teel, 2004, Thm. 12). For nonlinear MPC, however, $\kappa(x)$ is not necessarily continuous and therefore asymptotic stability of the nominal system does *not* imply RAS. Thus, arbitrarily small disturbances may in fact destabilize a system that is asymptotically stable in the nominal case. To establish RAS for a closed-loop system, we instead use an ISS Lyapunov function.

Definition 2.14 (ISS Lyapunov function)**.** A function $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is an ISS Lyapunov function for the system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in an RPI set $\mathcal{X}$ if there exist $\alpha_1(\cdot), \alpha_2(\cdot), \alpha_3(\cdot) \in$

$\mathcal{K}_\infty$ and $\sigma(\cdot) \in \mathcal{K}$ such that

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|) \tag{2.15}$$

$$V(f(x, \kappa(x), w)) \leq V(x) - \alpha_3(|x|) + \sigma(|w|) \tag{2.16}$$

for all $x \in \mathcal{X}$ and $w \in \mathbb{W}$.

The existence of an ISS Lyapunov function is a necessary and sufficient condition for robust asymptotic stability (Jiang and Wang, 2002; Grüne and Kellett, 2014).

Proposition 2.15. *The origin is RAS for a system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in an RPI set $\mathcal{X}$ if and only if there exists an ISS Lyapunov function for this system in $\mathcal{X}$.*

This framework of Lyapunov and comparison functions is both highly useful and remarkable flexible. As we show in Section 2.7, the cost function defined for MPC provides a natural (ISS) Lyapunov function for the system and therefore these results can be readily applied to MPC. Furthermore, we can extend this ISS Lyapunov function framework to stochastic systems and associated definitions of stochastic robustness.

2.5 Stochastic robustness

Robust asymptotic stability is a strong property in that the bound in (2.14) holds for any probability distribution on the set $\mathbb{W}$. If we however use a stochastic representation for the disturbance, we may instead want a similar bound based on the stochastic properties of the underlying system as shown in (2.11). In this section, we define and discuss a suitable definition of stochastic robustness based on linear version in (2.11).

Many forms of stochastic stability and robustness are already present in control theory. Originating in the 1960's in Kushner (1965, 1967), the notion of stochastic stability for nonlinear systems, i.e., asymptotic stability in probability, was refined more recently in Florchinger (1995). Teel and co-workers constructed and established stronger definitions of *uniform* asymptotic stability in probability and established that stochastic Lyapunov functions ensure uniform convergence (Teel, 2013; Teel et al., 2013, 2014). Analogous to ISS for deterministic systems, stochastic input-to-state stability (SISS) and a corresponding SISS Lyapunov function were defined (Krstic and Deng, 1998; Tang and Basar, 2001; Tsinias, 1998). Over the past decade, this SISS framework has been used in the analysis and control of continuous-time and discrete-time nonlinear stochastic systems (Huang and Mao, 2009; Wu et al., 2016; Zhao et al., 2012; Ding et al., 2015). These works assume that the effect of the stochastic disturbance vanishes once the state of the system reaches the origin (i.e., a multiplicative disturbance) and typically require the closed-loop system to be continuous. In most control applications, however, the stochastic disturbances do not vanish at the origin and MPC may produce a discontinuous control law and therefore closed-loop system. Consequently, the results from stochastic stability theory are not applicable in their current form to the closed-loop systems relevant to MPC. Instead, we define the following notion of stochastic robustness.

Definition 2.16 (Robust asymptotic stability in expectation)**.** The origin is robustly asymptotically stable in expectation (RASiE) for a system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in an RPI set $\mathcal{X}$ if there exist $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$\mathbb{E}\left[|\phi(k; x, \mathbf{w}_k)|\right] \leq \beta(|x|, k) + \gamma(\text{tr}(\Sigma)) \tag{2.17}$$

for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$.

In contrast to RAS, RASiE bounds the *expected value* of the norm of the closed-loop state based on a *stochastic property* of the disturbance, i.e., $\text{tr}(\Sigma)$. Similar to RAS, however, RASiE ensures that the effect of the initial state $x \in \mathcal{X}$ on the upper bound asymptotically (and uniformly) decreases to zero as $k \to \infty$. The remaining term is independent of k and depends on the distribution of the disturbance w, i.e., $\text{tr}(\Sigma)$ which depends on $\mu \in \mathcal{M}(\mathbb{W})$. By requiring $\gamma(\cdot) \in \mathcal{K}$, we ensure that small values of $\text{tr}(\Sigma)$ result in similarly small deviations from the origin (in terms of expected value) and as $\text{tr}(\Sigma) \to 0$ we recover asymptotic stability of the nominal closed-loop system. We also note that if $\text{tr}(\Sigma) \to 0$, then $\Sigma \to 0$ as well.

To establish RASiE, we use an SISS Lyapunov function similar to the SISS Lyapunov functions used in nonlinear stochastic stability theory. Note that we do not require continuity of $f(x, \kappa(x), w)$ and the bound in (2.19) is based on $\text{tr}(\Sigma)$.

Definition 2.17 (SISS Lyapunov function)**.** The Borel measurable function $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is an SISS Lyapunov function for the system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in an RPI set $\mathcal{X}$ if there exist $\alpha_1(\cdot), \alpha_2(\cdot), \alpha_3(\cdot) \in \mathcal{K}_\infty$ and $\sigma(\cdot) \in \mathcal{K}$ such that

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|) \tag{2.18}$$

$$\int_{\mathbb{W}} V(f(x, \kappa(x), w)) d\mu(w) \leq V(x) - \alpha_3(|x|) + \sigma(\text{tr}(\Sigma)) \tag{2.19}$$

for all $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathbb{W})$.

We can then establish that the existence of an SISS Lyapunov function is a sufficient condition for RASiE if the set $\mathcal{X}$ is *bounded*.[8]

Proposition 2.18. *If a system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ admits an SISS Lyapunov function in an RPI and bounded set $\mathcal{X}$, then the origin is RASiE in $\mathcal{X}$.*

To establish Proposition 2.18, we find the following results, based on Praly and Wang (1996, Lemma 14), useful.

Lemma 2.19. *If $\alpha(\cdot) \in \mathcal{K}$, then for any $b \geq 0$, there exists $\alpha_v(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_v(\cdot)$ is convex and $\alpha_v(s) \leq \alpha(s)$ for all $s \in [0, b]$.*

[8] We did not investigate if this SISS Lyapunov function is also a necessary condition.

Proof. We define

$$\alpha_v(s) := \frac{1}{b}\int_0^s \alpha(r)dr$$

Thus, $\alpha_v(\cdot)$ is strictly increasing and unbounded as $s \to \infty$ since $\alpha(s) > 0$ for all $s > 0$. Since $\alpha(r)$ is continuous, we have the $\alpha_v(s)$ is continuous as well (Rudin, 1976, Thm 6.20). Thus, $\alpha_v(\cdot) \in \mathcal{K}_\infty$. The derivative of $\alpha_v(\cdot)$, i.e., $\frac{d\alpha_v}{ds}(s) = \alpha(s)/b$, is strictly increasing and therefore $\alpha_v(\cdot)$ is a convex function. Furthermore, we have

$$\alpha_v(s) = \frac{1}{b}\int_0^s \alpha(r)dr \leq \frac{1}{b}\int_0^s \alpha(s)dr = \frac{s}{b}\alpha(s) \leq \alpha(s)$$

for all $s \in [0, b]$. □

Corollary 2.20. *If $\alpha(\cdot) \in \mathcal{K}_\infty$, then for any $b \geq 0$, there exists $\alpha_c(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_c(\cdot)$ is concave and $\alpha(s) \leq \alpha_c(s)$ for all $s \in [0, b]$.*

Proof. Note that the inverse $\alpha^{-1}(\cdot)$ exists and $\alpha^{-1}(\cdot) \in \mathcal{K}_\infty$ because $\alpha(\cdot) \in \mathcal{K}_\infty$. We use Lemma 2.19 to construct a convex function $\alpha_v(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_v(r) \leq \alpha^{-1}(r)$ for all $r \in [0, \alpha^{-1}(b)]$. Therefore, $\alpha_c(s) := \alpha_v^{-1}(s) \geq \alpha_2(s)$ for all $s \in [0, b]$. The inverse of a continuous, strictly increasing, and convex function is concave and therefore $\alpha_c(\cdot) \in \mathcal{K}_\infty$ and is a concave function. □

We now prove Proposition 2.18.

Proof of Proposition 2.18. We use the upper bound in (2.18) with 2.19 to give

$$\int_{\mathbb{W}} f(x, \kappa(x), w)d\mu(w) \leq V(x) - \alpha_4(V(x)) + \sigma(\text{tr}(\Sigma))$$

in which $\alpha_4(\cdot) := \alpha_3 \circ \alpha_2^{-1}(\cdot) \in \mathcal{K}_\infty$. Since $\mathcal{X}$ is bounded and $V(x) \leq \alpha_2(|x|)$, there exists $b \geq 0$ such that $V(x) \leq b$ for all $x \in \mathcal{X}$. From Lemma 2.19, we construct $\alpha_v(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_v(\cdot)$ is convex and $\alpha_v(V(x)) \leq \alpha_4(V(x))$ for all $x \in \mathcal{X}$. Therefore, we can replace $\alpha_4(\cdot)$ with $\alpha_v(\cdot)$ to give

$$\int_{\mathbb{W}} V(f(x, \kappa(x), w))d\mu(w) \leq V(x) - \alpha_v(V(x)) + \sigma(\text{tr}(\Sigma))$$

Choose $x \in \mathcal{X}$ and let $x(k) := \phi(k; x, \mathbf{w}_k)$ for all $k \in \mathbb{I}_{\geq 0}$. By the law of total expectation, we have that

$$\mathbb{E}\left[V(x(k+1))\right] \leq \mathbb{E}\Big[V(x(k)) - \alpha_v(V(x(k))) + \sigma(\text{tr}(\Sigma))\Big]$$

for all $k \in \mathbb{I}_{\geq 0}$. We can apply Jensen's inequality with the convexity of $\alpha_v(\cdot)$ to give

$$\mathbb{E}\left[V(x(k+1))\right] \leq \mathbb{E}\left[V(x(k))\right] - \alpha_v\Big(\mathbb{E}\left[V(x(k))\right]\Big) + \sigma(\text{tr}(\Sigma)) \tag{2.20}$$

and note that this inequality holds for all $k \in \mathbb{I}_{\geq 0}$ and any $\mu \in \mathcal{M}(\mathbb{W})$.

Define $\tilde{\gamma}(s) := 2\max\{\alpha_v^{-1}(\sigma(s)), \sigma(s)\}$ and note that $\tilde{\gamma}(\cdot) \in \mathcal{K}$. We now split the bound (2.20) into three separate regions.

If $\mathbb{E}\left[V(x(k))\right] \leq \tilde{\gamma}(\text{tr}(\Sigma))/2$, then

$$\begin{aligned}\mathbb{E}\left[V(x(k+1))\right] &\leq \tilde{\gamma}(\text{tr}(\Sigma))/2 + \sigma(\text{tr}(\Sigma)) \\ &\leq \tilde{\gamma}(\text{tr}(\Sigma))/2 + \tilde{\gamma}(\text{tr}(\Sigma))/2 = \tilde{\gamma}(\text{tr}(\Sigma))\end{aligned}$$

If $\tilde{\gamma}(\text{tr}(\Sigma))/2 \leq \mathbb{E}\left[V(x(k))\right] \leq \tilde{\gamma}(\text{tr}(\Sigma))$, then

$$\begin{aligned}\mathbb{E}\left[V(x(k+1))\right] &\leq \mathbb{E}\left[V(x(k))\right] - \alpha_v(\tilde{\gamma}(\text{tr}(\Sigma))/2) + \sigma(\text{tr}(\Sigma)) \\ &\leq \mathbb{E}\left[V(x(k))\right] \leq \tilde{\gamma}(\text{tr}(\Sigma))\end{aligned}$$

Thus, if $\mathbb{E}\left[V(x(k))\right] \leq \tilde{\gamma}(\text{tr}(\Sigma))$,

$$\mathbb{E}\left[V(x(k+1))\right] \leq \tilde{\gamma}(\text{tr}(\Sigma)) \tag{2.21}$$

If $\mathbb{E}\left[V(x(k))\right] \geq \tilde{\gamma}(\text{tr}(\Sigma))$, then

$$\begin{aligned}\mathbb{E}\left[V(x(k+1))\right] &\leq \mathbb{E}\left[V(x(k))\right] - \sigma_v(\mathbb{E}\left[V(x(k))\right]) + \sigma_v(\mathbb{E}\left[V(x(k))\right]/2) \\ &\leq \lambda_1(\mathbb{E}\left[V(x(k))\right])\end{aligned}$$

in which $\lambda_1(s) := s - \alpha_v(s) + \alpha_v(s/2)$. We have that $\lambda_1(\cdot)$ is continuous, $\lambda_1(0) = 0$, and $\lambda_1(s) < s$ for all $s > 0$. By the same process used in Rawlings et al. (2020, Theorem B.15), we construct $\lambda(\cdot) \in \mathcal{K}_\infty$ such that $\lambda_1(s) \leq \lambda(s) < s$ for all $s > 0$. Thus, we have

$$\mathbb{E}\left[V(x(k+1))\right] \leq \lambda(\mathbb{E}\left[V(x(k))\right]) \tag{2.22}$$

Repeated application of (2.22) and the fact that $\mathbb{E}\left[V(x(0))\right] = V(x)$ gives

$$\mathbb{E}\left[V(x(k))\right] \leq \tilde{\beta}(V(x), k) := \lambda^k(V(x)) \tag{2.23}$$

in which $\lambda^k(\cdot)$ denotes the composition of $\lambda(\cdot)$ with itself k times. Using the same approach as Rawlings et al. (2020, Theorem B.15), we conclude that $\beta(\cdot) \in \mathcal{KL}$.

We combine (2.21) and (2.23) to give,

$$\mathbb{E}\left[V(x(k))\right] \leq \max\{\tilde{\beta}(V(x), k), \tilde{\gamma}(\text{tr}(\Sigma))\} \tag{2.24}$$

Using Lemma 2.19 and the fact that $\mathcal{X}$ is bounded, we can construct a convex function $\alpha_{1,v}(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_{1,v}(|x|) \leq \alpha_1(|x|) \leq V(x)$ for all $x \in \mathcal{X}$. Thus, we apply Jensen's inequality to give

$$\alpha_{1,v}(\mathbb{E}[|x|]) \leq \mathbb{E}[\alpha_{1,v}(|x|)] \leq \mathbb{E}[V(x)]$$

We also have that $V(x) \leq \alpha_2(|x|)$. Therefore,

$$\mathbb{E}[|x(k)|] \leq \max\left\{\alpha_{1,v}^{-1}(\tilde{\beta}(\alpha_2(|x|), k)), \alpha_{1,v}^{-1}(\tilde{\gamma}(\text{tr}(\Sigma)))\right\}$$

We define $\beta(s,k) := \alpha_{1,v}^{-1}(\tilde{\beta}(\alpha_2(s), k))$, $\gamma(s) := \alpha_{1,v}^{-1}(\tilde{\gamma}(s))$, and use the fact that $\max\{a,b\} \leq a + b$ to give

$$\mathbb{E}[|x(k)|] \leq \beta(|x|, k) + \gamma(\text{tr}(\Sigma)) \tag{2.25}$$

Note that $\beta(\cdot) \in \mathcal{KL}$, $\gamma(\cdot) \in \mathcal{K}$ and since the choice of $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$ was arbitrary, (2.25) holds for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$. □

We can also establish the following connection between ISS and SISS Lyapunov functions.

Proposition 2.21. *Let Assumption 2.2 hold. If a Borel measurable function $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is an ISS Lyapunov function for the system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in an RPI set $\mathcal{X}$, then $V(\cdot)$ is also an SISS Lyapunov function for the system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in $\mathcal{X}$.*

Proof. Since $V(\cdot)$ is an ISS Lyapunov function, there exist $\alpha_1(\cdot), \alpha_2(\cdot), \alpha_3(\cdot) \in \mathcal{K}_\infty$ and $\sigma(\cdot) \in \mathcal{K}$ such that equations (2.15) and (2.16) hold. Therefore, (2.18) holds for the same functions $\alpha_1(\cdot)$, $\alpha_2(\cdot)$.

Since $\mathbb{W}$ is compact, there exists $b \geq 0$ such that $|w| \in [0, b]$ for all $w \in \mathbb{W}$. We define a function $\tilde{\sigma}(\cdot) \in \mathcal{K}_\infty$ such that $\sigma(s) \leq \tilde{\sigma}(s)$ for all $s \in \mathbb{R}_{\geq 0}$, e.g., $\tilde{\sigma}(s) := \varepsilon s + \sigma(s)$ with $\varepsilon > 0$. We use Corollary 2.20 to construct a concave function $\sigma_c(\cdot) \in \mathcal{K}_\infty$ such that $\sigma(|w|) \leq \tilde{\sigma}(|w|) \leq \sigma_c(|w|)$ for all $w \in \mathbb{W}$. Choose arbitrary $\mu \in \mathcal{M}(\mathbb{W})$ and integrate both sides of (2.16) with respect to this probability measure. We then use $\sigma_c(\cdot)$ and Jensen's inequality to give

$$\begin{aligned}\int_{\mathbb{W}} V(f(x,\kappa(x),w))d\mu(w) &\leq V(x) - \alpha_3(|x|) + \int_{\mathbb{W}} \sigma(|w|)d\mu(w) \\ &\leq V(x) - \alpha_3(|x|) + \int_{\mathbb{W}} \sigma_c(|w|)d\mu(w) \\ &\leq V(x) - \alpha_3(|x|) + \sigma_c\left(\int_{\mathbb{W}} |w|d\mu(w)\right) \\ &= V(x) - \alpha_3(|x|) + \sigma_c\left(\mathbb{E}[|w|]\right)\end{aligned}$$

From Lemma 2.4, we have that $\mathbb{E}[|w|] \leq \text{tr}(\Sigma)^{1/2}$. We define $\sigma_3(s) := \sigma_c(s^{1/2})$ and note that $\sigma_3(\cdot) \in \mathcal{K}$. Thus, we have that $\sigma_c(\mathbb{E}[|w|]) = \sigma_3(\mathbb{E}[|w|]^2) \leq \sigma_3(\text{tr}(\Sigma))$ and

$$\int_{\mathbb{W}} V(f(x,\kappa(x),w))d\mu(w) \leq V(x) - \alpha_3(|x|) + \sigma_3\left(\text{tr}(\Sigma)\right)$$

Therefore, 2.19 holds for $\alpha_3(\cdot) \in \mathcal{K}_\infty$ and $\sigma_3(\cdot) \in \mathcal{K}$. Note that the choice of μ was arbitrary and therefore (2.19) also holds for all $\mu \in \mathcal{M}(\mathbb{W})$. Thus, $V(\cdot)$ is an SISS Lyapunov function.

□

The converse of Proposition 2.21, however, does not hold. For example, consider the scalar system $x^+ = (0.9+w)x$ for $w \in \mathbb{W} := [-0.2, 0.2]$ distributed such that $\mathbb{E}\left[w\right] = 0$. The system is not ISS because $w = 0.2$ produces an unstable system $x^+ = 1.1x$ and therefore $|x(k)| \to \infty$ for some $w \in \mathbb{W}$. But $V(x) = x^2$ is an SISS Lyapunov function:

$$\begin{aligned}\int_{\mathbb{W}} V(x^+)d\mu(w) &= \int_{\mathbb{W}} \left((0.9+w)x\right)^2 d\mu(w) \\ &\leq 0.81x^2 + 1.8x^2\mathbb{E}[w] + x^2\mathbb{E}[w^2] \\ &\leq 0.81x^2 + 0 + 0.04x^2 \\ &\leq V(x) - \alpha_3(|x|)\end{aligned}$$

with $\alpha_3(s) := 0.15s^2 \in \mathcal{K}_\infty$. In fact, this system admits a stochastic Lyapunov function ($\sigma(\cdot) = 0$) and is mean-squared stable (a stronger property than RASiE).

Proposition 2.21 is particularly important because we can use this result to establish stochastic robustness results for MPC. Specifically, we use Proposition 2.21 to establish that the ISS Lyapunov function typically derived for MPC also confers RASiE. Moreover, Proposition 2.21 leads to the following corollary that establishes a more fundamental connection between RAS and RASiE.

Corollary 2.22. *Let Assumption 2.2 hold. If the origin is RAS for a system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in an RPI and bounded set $\mathcal{X}$, then the origin is also RASiE in $\mathcal{X}$.*

Proof. From Proposition 2.15, we have that RAS implies the existence of an ISS Lyapunov function. We also verify that this ISS Lyapunov function, as constructed in Grüne and Kellett (2014), is Borel measurable. By Proposition 2.21, this ISS Lyapunov function is also an SISS Lyapunov function and by Proposition 2.18 the origin is RASiE. □

2.6 Robustness with respect to stage cost

In an MPC formulation, the control law is implicitly defined by optimizing a performance metric for the system known as the stage cost and denoted by the function $\ell : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}_{\geq 0}$. This stage cost is often chosen as the same quadratic form used in the LQR formulation, i.e., $\ell(x, u) = x'Qx + u'Ru$ with $Q, R \succ 0$. For steady-state tracking applications of MPC, a standard requirement is that the stage cost is lower-bounded by a $\mathcal{K}_\infty$-function of $|x|$, as detailed in the following assumption. We also require that this stage cost is continuous and zero at the origin.

Assumption 2.23 (Stage cost)**.** The state cost $\ell : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}_{\geq 0}$ is continuous and satisfies $\ell(0, 0) = 0$. Moreover, there exists $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_\ell(|x|) \leq \ell(x, u)$ for all $(x, u) \in$

$\mathbb{R}^n \times \mathbb{U}$.

Assumption 2.23 ensures that if $\ell(x, u) \to 0$ then $|x| \to 0$, but allows for significant flexibility in selecting $\ell(\cdot)$. This flexibility is used to "tune" the stage cost to reflect the importance of different state and input variables in the problem of interest. For example, we may place a large penalty on deviations in the concentrations/quality of a product in a reactor but a smaller penalty on deviations in the liquid level of a tank. This flexibility, however, also separates the objective of the MPC problem from the metrics used in the definition of RAS and RASiE. Thus, a reasonable metric for evaluating the robustness of MPC is the one specifically prescribed to the MPC problem formulation: the stage cost $\ell(\cdot)$. Specifically, we investigate the value of this cost along the closed-loop trajectory as defined by the quantity

$$\ell(\phi(k; x, \mathbf{w}_k), \kappa(\phi(k; x, \mathbf{w}_k))) \tag{2.26}$$

For this quantity, we provide definitions of both deterministic and stochastic robustness that we term, respectively, RAS w.r.t. the stage cost $\ell(\cdot)$ and RASiE w.r.t. the stage cost $\ell(\cdot)$. We abbreviate these properties as ℓ-RAS and ℓ-RASiE and define them as follows.

Definition 2.24 (ℓ-RAS)**.** The origin is ℓ-RAS with respect to the stage cost $\ell(x, \kappa(x))$ for a system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in the RPI set $\mathcal{X}$ if there exist $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$\ell(x(k), \kappa(x(k))) \leq \beta(|x|, k) + \gamma(||\mathbf{w}_k||) \tag{2.27}$$

in which $x(k) := \phi(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$, $\mathbf{w}_k \in \mathbb{W}^k$, and $k \in \mathbb{I}_{\geq 0}$.

Definition 2.25 (ℓ-RASiE)**.** The origin is ℓ-RASiE with respect to the stage cost $\ell(x, \kappa(x))$ for a system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ in the RPI set $\mathcal{X}$ if there exist $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$\mathbb{E}\left[\ell(x(k), \kappa(x(k)))\right] \leq \beta(|x|, k) + \gamma(\text{tr}(\Sigma)) \tag{2.28}$$

in which $x(k) := \phi(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$.

Note that these definitions are generalizations of the previous definitions of RAS and RASiE under Assumption 2.23.

We now establish that an ISS (SISS) Lyapunov function that also satisfies $\ell(x, \kappa(x)) \leq V(x)$ ensures that the origin is ℓ-RAS (ℓ-RASiE). Since the Lyapunov function constructed for MPC is (almost) always based on the optimal cost function, the stage cost typically satisfies $\ell(x, \kappa(x)) \leq V(x)$.

Proposition 2.26. *If a system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ admits an ISS Lyapunov function $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$ in an RPI set $\mathcal{X}$ that satisfies $\ell(x, \kappa(x)) \leq V(x)$ for all $x \in \mathcal{X}$, then the origin is ℓ-RAS.*

Proof. Since the system admits an ISS Lyapunov function, we have from Proposition 2.15 that the origin is RAS, i.e., there exists $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$|\phi(k; x, \mathbf{w}_k)| \leq \beta(|x|, k) + \gamma(||\mathbf{w}_k||)$$

for all $x \in \mathcal{X}$, $\mathbf{w}_k \in \mathbb{W}^k$, and $k \in \mathbb{I}_{\geq 0}$. Furthermore, we have that $\ell(x, \kappa(x)) \leq V(x) \leq \alpha_2(|x|)$ and therefore

$$\ell(x(k), \kappa(x(k))) \leq \alpha_2(2\beta(|x|, k)) + \alpha_2(2\gamma(||\mathbf{w}_k||))$$

in which $x(k) := \phi(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$, $\mathbf{w}_k \in \mathbb{W}^k$, and $k \in \mathbb{I}_{\geq 0}$. Note that $\alpha_2(2\beta(\cdot)) \in \mathcal{KL}$ and $\alpha_2(2\gamma(\cdot)) \in \mathcal{K}$ to complete the proof. □

Proposition 2.27. *If a system $x^+ = f(x, \kappa(x), w)$, $w \in \mathbb{W}$ admits an SISS Lyapunov function $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$ in an RPI and bounded set $\mathcal{X}$ that satisfies $\ell(x, \kappa(x)) \leq V(x)$ for all $x \in \mathcal{X}$, then the origin is ℓ-RASiE.*

Proof. From Proposition 2.18, there exists $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$\mathbb{E}\left[|\phi(k; x, \mathbf{w}_k)|\right] \leq \beta(|x|, k) + \gamma(\text{tr}(\Sigma))$$

for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$. Since $\mathcal{X}$ is bounded, we can construct a concave function $\alpha_c(\cdot) \in \mathcal{K}_\infty$ such that $V(x) \leq \alpha_2(|x|) \leq \alpha_c(|x|)$ for all $x \in \mathcal{X}$ via Corollary 2.20. Therefore,

$$\mathbb{E}\left[\ell(x(k), \kappa(x(k)))\right] \leq \mathbb{E}\left[V(x(k))\right] \leq \mathbb{E}\left[\alpha_c(|x(k)|)\right] \leq \alpha_c(\mathbb{E}[|x(k)|])$$

and

$$\mathbb{E}\left[\ell(x(k), \kappa(x(k)))\right] \leq \alpha_c(2\beta(|x|, k)) + \alpha_c(2\gamma(\text{tr}(\Sigma))$$

in which $x(k) := \phi(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$. Note that $\alpha_c(2\beta(\cdot)) \in \mathcal{KL}$ and $\alpha_c(2\gamma(\cdot)) \in \mathcal{K}$ to complete the proof. □

2.7 Model predictive control

2.7.1 Problem formulation and assumptions

For MPC, the disturbance is not considered in the optimization problem. As such, we often refer to this MPC formulation as *nominal* MPC to distinguish it from stochastic or robust MPC, which consider the disturbance explicitly in their problem formulations. Thus, the system model is given by

$$x^+ = f(x, u, 0) \tag{2.29}$$

For a predictive horizon $N \in \mathbb{I}_{\geq 1}$, we use $\hat{\phi}(k; x, \mathbf{u})$ to denote the state trajectory of (2.29) at time $k \in \mathbb{I}_{0:N}$, given the initial condition $x \in \mathbb{R}^n$ and the control trajectory

$$\mathbf{u} = (u(0), u(1), \ldots, u(N-1)) \in \mathbb{U}^N$$

We allow input constraints $u \in \mathbb{U} \subseteq \mathbb{R}^n$, but do not enforce state constraints in the optimization problem. State constraints are sometimes included in MPC formulations and their presence does not disrupt asymptotic stability guarantees for the nominal closed-loop system. These state constraints, however, are problematic once disturbances and nonlinear systems are included in the closed-loop analysis. While input constraints typically represent physical limits of an actuator, e.g., a valve can be set only from 0% to 100% open, state constraints represent desired goals of the controller, e.g., keep the product concentration above the minimum threshold. For a perturbed system, there is no guarantee that these state constraints can be satisfied and including these desired features as constraints can result in infeasible optimization problems. Instead, we assume that these state constraints are converted to exact penalty functions that are included in the stage cost (Zheng and Morari, 1995; Scokaert and Rawlings, 1999; Kerrigan and Maciejowski, 2000). Thus, violation of these state constraints results in a considerable cost increase that is avoided, if possible, in the optimization problem.

In the LQR problem, the control law is derived from an infinite horizon optimization problem. Ideally, we would also solve an infinite horizon optimization problem to determine the control law for nonlinear systems with constraints. This approach, however, is intractable for even linear systems of significant state dimension with input constraints. Instead, MPC uses a *finite* horizon, denoted by $N \in \mathbb{I}_{\geq 1}$, to ensure that the optimization problem is computationally tractable for relevant systems. The drawback of this finite horizon is that optimal control trajectory may be shortsighted and the resulting control law is not necessarily stabilizing or robust. To address this limitation, a standard practice is to include a terminal cost

$V_f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ and terminal constraint $\mathbb{X}_f \subseteq \mathbb{R}^n$ to ensure that the finite horizon optimization problem adequately approximates the infinite horizon problem and therefore avoid myopic control actions. In particular, we construct the terminal cost and constraint to guarantee that the origin is RAS for the closed-loop system. We detail the specific requirements for the terminal cost and constraint through assumptions introduced later in this section.

For MPC with a horizon $N \in \mathbb{I}_{\geq 1}$, we define the set of admissible inputs, feasible initial states, and objective function, respectively, as

$$\mathcal{U}(x) := \{\mathbf{u} \in \mathbb{U}^N : \hat{\phi}(N; x, \mathbf{u}) \in \mathbb{X}_f\}$$
$$\mathcal{X} := \{x \in \mathbb{R}^n : \mathcal{U}(x) \neq \emptyset\}$$
$$V(x, \mathbf{u}) := \sum_{k=0}^{N-1} \ell(\hat{\phi}(k; x, \mathbf{u}), u(k)) + V_f(\hat{\phi}(N; x, \mathbf{u}))$$

The nominal MPC problem for any $x \in \mathcal{X}$ is defined as

$$\mathbb{P}(x) : \ V^0(x) := \min_{\mathbf{u} \in \mathcal{U}(x)} V(x, \mathbf{u}) \tag{2.30}$$

and the optimal solution(s) for a given initial state are denoted $\mathbf{u}^0(x) := \arg\min_{\mathbf{u} \in \mathcal{U}(x)} V(x, \mathbf{u})$. Note that $\mathbf{u}^0(x)$ is a set-valued mapping because there may be multiple solutions to $\mathbb{P}(x)$, i.e., $\mathbf{u}^0(x)$ is defined as a set with more than one element for at least one value of $x \in \mathcal{X}$. To streamline the following presentation, we use a selection rule to define a single-valued control law $\kappa : \mathcal{X} \to \mathbb{U}$ such that

$$\kappa(x) \in u^0(0; x)$$

for all $x \in \mathcal{X}$, in which $u^0(0; x)$ is the set of first inputs in $\mathbf{u}^0(x)$. The resulting closed-loop system is then

$$x^+ = f(x, \kappa(x), w) \tag{2.31}$$

Note that, although $\kappa(x)$ is designed without knowledge of w, the closed-loop system still includes this disturbance.

We consider the following additional assumptions for the MPC problem formulation based on the assumptions in Allan et al. (2017).

Assumption 2.28 (Properties of the constraint sets)**.** The set $\mathbb{U}$ is compact and contains the origin. The set $\mathbb{X}_f$ is defined by $\mathbb{X}_f := \{x \in \mathbb{R}^n : V_f(x) \leq \tau\}$ for some $\tau > 0$.

Assumption 2.29 (Terminal ingredients)**.** The function $V_f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is continuous and $V_f(0) = 0$. There exists a terminal control law $\kappa_f : \mathbb{X}_f \to \mathbb{U}$ such that for all $x \in \mathbb{X}_f$,

$$f(x, \kappa_f(x), 0) \in \mathbb{X}_f \tag{2.32}$$

$$V_f(f(x, \kappa_f(x), 0)) \leq V_f(x) - \ell(x, \kappa_f(x)) \tag{2.33}$$

Note that the condition in (2.33) means that $V_f(\cdot)$ acts as a *local* Lyapunov function in the terminal region for the nominal system $x^+ = f(x, \kappa_f(x), 0)$ and with the specific dissipation rate $\ell(x, \kappa_f(x))$. Thus, the terminal control law must stabilize the origin, but does not need to be optimal for this system and stage cost. While constructing this terminal cost still presents a challenge, we can choose $\tau > 0$ small such that (2.33) needs to hold for only a small region around the origin. Designing a terminal cost and terminal control law to satisfy the requirements of Assumption 2.29 is therefore tractable for many systems of interest.

For example, if the stage cost is quadratic, i.e., $\ell(x, u) = x'Qx + u'Ru$ with $Q, R \succ 0$, we can construct this terminal cost by linearizing the system about the origin (steady state target) and computing the LQR solution for this linearized system (assuming the linearized system is stabilizable). For sufficiently small regions around the origin, this linear model is reasonably accurate and the LQR feedback gain stabilizes the origin. We can therefore construct a terminal cost for this system that satisfies Assumption 2.29 via the procedure in

Rawlings et al. (2020, s 2.5.5). If certain input constraints are active at the origin, we hold these modes of the input constant when computing the linearized model. Provided the linearized model is still stabilizable, we can compute the LQR solution based on this subset of the original input space and use this LQR solution to construct the terminal cost as before.

2.7.2 Existence and measurability

Before proceeding to any closed-loop properties for MPC, we must first establish that a solution to $\mathbb{P}$ exists. Moreover, we need to verify $\kappa(\cdot)$ is Borel measurable to ensure that expected value is well defined for the closed-loop system. Although not a particularly exciting area of analysis[9], verifying these mathematical details is essential to guarantee the veracity of all the subsequent analysis performed with these mathematical objects. Thus, we briefly summarize the results on existence and measurability for MPC with additional details available in McAllister and Rawlings (2021b, Appendix B).

In Rawlings et al. (2020, Prop. 2.4), the authors establish that Assumptions 2.1 and 2.28 combined with continuity of $\ell(\cdot)$ and $V_f(\cdot)$ are sufficient to ensure that a solution to $\mathbb{P}(x)$ exists for all $x \in \mathcal{X}$. Under these same conditions, we can further establish, via Proposition 7.33 in Bertsekas and Shreve (1978), that $\mathcal{X}$ is closed, the optimal cost function $V^0 : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is lower semicontinuous, and that $\mathbf{u}^0 : \mathcal{X} \rightrightarrows \mathbb{U}^N$ is Borel measurable. Furthermore, Bertsekas and Shreve (1978, Lemma 7.18) also establish that for any compact set $\mathbb{U}$, there exists a Borel measurable selection rule for $\mathbf{u}^0(x)$. If we assume that such a selection rule is used to define the control law, we have that $\kappa : \mathcal{X} \to \mathbb{U}$ is well defined and Borel measurable. We summarize these results in the following proposition.

Proposition 2.30. *Let Assumptions 2.1, 2.23, 2.28 and 2.29 hold. Then, $\mathbb{P}(x)$ has a solution for all $x \in \mathcal{X}$, $\mathcal{X}$ is closed, $V^0 : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is lower semicontinuous, and $\mathbf{u}^0 : \mathcal{X} \rightrightarrows \mathbb{U}^N$*

[9] for engineers, at least

is Borel measurable. Furthermore, there exists a Borel measurable selection rule such that the single-valued control law $\kappa : \mathcal{X} \to \mathbb{U}$ *is Borel measurable and satisfies* $\kappa(x) \in u^0(0; x)$ *for all* $x \in \mathcal{X}$.

Remark 2.31. In theory, we could select an exotic selection rule that produces a non-measurable function $\kappa(\cdot)$ from the Borel measurable set-valued mapping $\mathbf{u}^0(\cdot)$ (McAllister and Rawlings, 2021b, Appendix A). We postulate, however, that unintentionally constructing such a selection rule for a system is unlikely as such a construction requires the use of the *axiom of choice* (Solovay, 1970). Thus, we simply assume that a Borel measurable selection rule is chosen. Note assuming that the selection rule is measurable does not supplant the need to verify that $\mathbf{u}^0(\cdot)$ is Borel measurable. If $\mathbf{u}^0(\cdot)$ is not a Borel measurable function, the resulting controller may not be measurable regardless of the selection rule used.

2.7.3 Robustness

While different researches may use slightly modified definitions of robustness, the general notion of robustness for MPC remains the same: For a compact subset of the feasible set ($\mathcal{S} \subseteq \mathcal{X}$), there exists a nonzero margin of robustness ($\delta > 0$) such that $\mathcal{S}$ is RPI, the optimization problem remains feasible, and the origin is RAS for sufficiently small disturbances ($|w| \leq \delta$). For nominal MPC, we refer to this property as *inherent* robustness, as we do not consider the disturbances directly in the optimization problem as done in stochastic or robust MPC formulations. Instead, nominal MPC relies on feedback to address these disturbances and this inherent robustness is often sufficient in industrial applications.

In closed-loop analysis of MPC, the optimal cost function $V^0(x)$ is typically used as a Lyapunov function for the closed-loop system. If $V^0(x)$ is continuous, which is true for a linear model and convex constraints, inherent robustness follows immediately from nominal asymptotic stability (Grimm et al., 2004). Nonlinear MPC with state or terminal constraints, however, can produce potentially discontinuous optimal cost functions. In a particularly consequential paper, Grimm et al. (2004) demonstrated that, due to these discontinuities, nominal asymptotic stability of the origin for nonlinear MPC does not guarantee any margin of robustness. More general conditions on the robustness of MPC without state or terminal constraints are given in Grimm et al. (2007).

In a significant contribution, Yu et al. (2014) established that nonlinear MPC, under suitable assumptions, is inherently robust to sufficiently small additive disturbances. Moreover, this robustness is achieved even if the optimal cost $V^0(x)$ is not continuous. These assumptions exclude hard state constraints but still require a terminal constraint in the optimization problem. The terminal cost and constraint are constructed via the LQR solution of the linearized system at the origin.

Leveraging some of these results, Allan et al. (2017) establish that suboptimal MPC, i.e., an MPC algorithm that does not require optimal solutions to the proposed optimal control problem, is also inherently robust. Given the computational burden and nonconvexity of nonlinear MPC optimization problems, the ability to use potentially suboptimal solutions is particularly important for the online implementation of MPC. To minimize the additional complexity of the subsequent analysis, however, we restrict attention in the following chapters to the optimal MPC problem while noting that extensions of these results to suboptimal MPC are conceivable.

In addition to the results for suboptimal MPC, the analysis in Allan et al. (2017) does not require that $\mathbb{U}$ has an interior, a common assumption in MPC analysis that is used in Yu et al. (2014). By removing this requirement, we allow $\mathbb{U}$ to represent a larger class of input constraints including *integer* constraints, i.e., decision variables that must take strictly integer values.[10] This class of inputs includes binary decisions such as the ON/OFF decisions

[10]By allowing input constraints to be active at the origin, however, the terminal ingredients (Assumption 2.29) must be carefully constructed. See Rao and Rawlings (1999) for a method to construct these terminal ingredients that also applies to problems with integer constraints.

for equipment that are pertinent in higher-level planning and scheduling problems. This fact was articulated in Rawlings and Risbeck (2017) and led to the conclusion: *Any result that holds for standard MPC also holds for MPC with discrete actuators.* In this case, "standard" MPC implies MPC without integer constraints.

The novel contribution of this section is to establish that MPC is also robust in a stochastic context and with respect to the stage cost used in the MPC formulation. While interesting in their own right, these results are particularly important as they facility the direct comparison of nominal MPC, which is almost always characterized by deterministic definitions of robustness, and stochastic MPC, which is almost always characterized by stochastic definitions of robustness. With the assumptions introduced throughout this section, we can establish the following theorem for the robustness of MPC.

Theorem 2.32 (MPC)**.** *Let Assumptions 2.1, 2.2, 2.23, 2.28 and 2.29 hold. For every* $\rho > 0$*, there exists* $\delta > 0$ *such that for any set* $\mathbb{W} \subseteq \{w \in \mathbb{R}^q : |w| \leq \delta\}$*, the closed-loop system* $x^+ = f(x, \kappa(x), w)$*,* $w \in \mathbb{W}$*, and the set* $\mathcal{S} := \{x \in \mathbb{R}^n : V^0(x) \leq \rho\} \cap \mathcal{X}$ *we have that*

(i) The set $\mathcal{S}$ *is RPI.*

(ii) The origin is RAS in the set $\mathcal{S}$*.*

(iii) The origin is RASiE in the set $\mathcal{S}$*.*

(iv) The origin is ℓ*-RAS in the set* $\mathcal{S}$*.*

(v) The origin is ℓ*-RASiE in the set* $\mathcal{S}$*.*

Thus, MPC, for sufficiently small disturbances ($|w| \leq \delta$), satisfies all of the deterministic and stochastic definitions of robustness in the set $\mathcal{S}$. This robustness is achieved without directly considering the disturbance w in the optimization problem and is *inherent* to the MPC formulation through feedback. The first two results are establish in Allan et al. (2017) for suboptimal MPC, but we provide a proof here for the specific case of optimal MPC.

To prove Theorem 2.32, we use the following technical results.

Proposition 2.33. *Let $X \subseteq \mathbb{R}^n$ be closed and suppose a function $g : X \to \mathbb{R}$ is continuous at $x_0 \in X$ and locally bounded on X.*[11] *Then, there exists $\alpha(\cdot) \in \mathcal{K}_\infty$ such that $|g(x) - g(x_0)| \leq \alpha(|x - x_0|)$ for all $x \in X$.*

Proof. See Rawlings and Risbeck (2015, Prop. 14). □

Proposition 2.34. *Let $X \subseteq Y \subseteq \mathbb{R}^n$, with compact X, closed Y, and $g : Y \to \mathbb{R}^n$ continuous. Then there exists $\sigma(\cdot) \in \mathcal{K}_\infty$ such that $|g(x) - g(y)| \leq \sigma(|x - y|)$ for all $x \in X$ and $y \in Y$.*

Proof. See Allan et al. (2017, Prop. 20). □

We then establish the following upper bound for the optimal cost function.

Lemma 2.35. *Let Assumptions 2.1, 2.23, 2.28 and 2.29 hold. Then there exists $\alpha_2(\cdot) \in \mathcal{K}_\infty$ such that $V^0(x) \leq \alpha_2(|x|)$ for all $x \in \mathcal{X}$.*

Proof. Choose $x \in \mathbb{X}_f$ and consider a nominal trajectory generated by repeated application of the terminal control law for the horizon N, i.e., $\mathbf{u}_f := (\kappa_f(x), \kappa_f(f(x, \kappa_f(x), 0)), \dots) \in \mathbb{U}^N$. The set $\mathbb{X}_f$ is RPI for this control law due to Assumption 2.29. Therefore, $\mathbf{u}_f \in \mathcal{U}(x)$ for all $x \in \mathbb{X}_f$. We denote the corresponding state trajectory as $x(k) := \hat{\phi}(k; x, \mathbf{u}_f)$. From Assumption 2.29, we have that

$$V_f(x(k+1)) \leq V_f(x(k)) - \ell(x(k), \kappa_f(x(k)))$$

for all $k \in \mathbb{I}_{0:N-1}$. We sum this inequality from $k = 0$ to $k = N - 1$ to give

$$V_f(x(N)) - V_f(x) \leq -\sum_{k=0}^{N-1} \ell(x(k), \kappa_f(x(k)))$$

[11] *A function $g : X \to \mathbb{R}$ is locally bounded on X if for any compact set $A \subseteq X$, there exists $M > 0$ such that $|g(x)| \leq M$ for all $x \in A$.*

By optimality and the definition of $V(\cdot)$, we have

$$\begin{aligned} V^0(x) &\leq V(x, \mathbf{u}_f) \\ &= \sum_{k=0}^{N-1} \ell(x(k), \kappa_f(x(k))) + V_f(x(N)) \leq V_f(x) \end{aligned}$$

for all $x \in \mathbb{X}_f$. Since $0 \leq V^0(x) \leq V_f(x)$, $V_f(\cdot)$ is continuous, $V^0(0) = V_f(0) = 0$, and $\mathbb{X}_f$ contain the origin in its interior, we have that $V^0(\cdot)$ is continuous at the origin.

We now establish that $V^0(\cdot)$ is locally bounded. Choose any compact set $X \subseteq \mathcal{X}$. Since $V(\cdot)$ is the composition of a finite number of continuous functions, $V(\cdot)$ is continuous and there exists $M \geq 0$ such that $|V(x, \mathbf{u})| \leq M$ for all $X \times \mathbb{U}^N$. Therefore, $|V^0(x)| \leq M$ for all $x \in X$ as well, because $\mathbf{u}^0(x) \subseteq \mathcal{U}(x) \subseteq \mathbb{U}^N$. Thus, $V^0(x)$ is locally bounded. Since $V^0(x)$ is continuous at the origin and is locally bounded, we have from Proposition 2.33 that there exists $\alpha_2(\cdot) \in \mathcal{K}_\infty$ such that

$$V^0(x) = |V^0(x) - V^0(0)| \leq \alpha_2(|x - 0|) = \alpha_2(|x|)$$

for all $x \in \mathcal{X}$. □

We also establish that the origin is asymptotically stable for the nominal system.

Lemma 2.36. *Let Assumptions 2.1, 2.23, 2.28 and 2.29 hold. For every $\rho > 0$, the optimal cost $V^0 : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is a Lyapunov function for the nominal closed-loop system $x^+ = f(x, \kappa(x), 0)$ in the positive invariant set $\mathcal{S} := \{x \in \mathbb{R}^n : V^0(x) \leq \rho\} \cap \mathcal{X}$ and therefore the origin is asymptotically stable in $\mathcal{S}$.*

Proof. For $\rho > 0$, choose $x \in \mathcal{S} \subseteq \mathcal{X}$ and $\mathbf{u}^0 \in \mathbf{u}^0(x)$ such that $\kappa(x) = u^0(0)$. We denote

$\hat{x}^+ = f(x, \kappa(x), 0)$ and $\hat{x}_N := \phi(N; x, \mathbf{u}^0)$. Define the extended input trajectory as

$$\tilde{\mathbf{u}}^+ = (u^0(1), u^0(2), \ldots, u^0(N-1), \kappa_f(\hat{x}_N))$$

Note that $\tilde{\mathbf{u}}^+ \in \mathbb{U}^N$. For this input trajectory, we have that $\hat{x}_N \in \mathbb{X}_f$ because $\mathbf{u}^0 \in \mathcal{U}(x)$ and therefore $f(\hat{x}_N, \kappa_f(\hat{x}_N), 0) \in \mathbb{X}_f$ because of Assumption 2.29. Thus, $\tilde{\mathbf{u}}^+ \in \mathcal{U}(\hat{x}^+)$ and $\hat{x}^+ \in \mathcal{X}$ since $\mathcal{U}(\hat{x}^+) \neq \emptyset$. Furthermore, we have for this extended input trajectory that

$$\begin{aligned} V(\hat{x}^+, \tilde{\mathbf{u}}^+) &= V(x, \mathbf{u}^0) - \ell(x, \kappa(x)) - V_f(\hat{x}_N) + \ell(\hat{x}_N, \kappa(\hat{x}_N)) + V_f(f(\hat{x}_N, \kappa_f(\hat{x}_N))) \\ &\leq V(x, \mathbf{u}^0) - \ell(x, \kappa(x)) \\ &\leq V(x, \mathbf{u}^0) - \alpha_\ell(|x|) \end{aligned} \tag{2.34}$$

in which the first inequality holds because of Assumption 2.29 and the second inequality holds with $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$ because of Assumption 2.23. Therefore,

$$V^0(\hat{x}^+) \leq V(\hat{x}^+, \tilde{\mathbf{u}}^+) \leq V(x, \mathbf{u}^0) = V^0(x) \leq \rho$$

for all $x \in \mathcal{S}$. Since $\hat{x}^+ \in \mathcal{X}$ and $V^0(\hat{x}^+) \leq \rho$, we have that $\hat{x}^+ \in \mathcal{S}$ for all $x \in \mathcal{S}$, i.e., $\mathcal{S}$ is positive invariant for the nominal system.

From Assumption 2.23, we also have $\alpha_\ell(|x|) \leq \ell(x, \kappa(x)) \leq V^0(x)$ with $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$. We use Lemma 2.35 to get the upper bound. The cost decrease condition is shown in (2.34) and therefore $V^0(\cdot)$ is a Lyapunov function. We use Proposition 2.11 to establish that the origin is asymptotically stable. □

With these results in hand, we can now establish Theorem 2.32.

Proof of Theorem 2.32. We use a similar approach as Allan et al. (2017) to establish this result. For $\rho > 0$, choose $x \in \mathcal{S} \subseteq \mathcal{X}$ and $\mathbf{u}^0 \in \mathbf{u}^0(x)$ such that $\kappa(x) = u^0(0)$. For the nominal system, we denote $\hat{x}^+ = f(x, \kappa(x), 0)$, $\hat{x}_N := \phi(N; x, \mathbf{u}^0)$, and define the extended input trajectory as

$$\tilde{\mathbf{u}}^+ = (u^0(1), u^0(2), \ldots, u^0(N-1), \kappa_f(\hat{x}_N))$$

Note that $\tilde{\mathbf{u}}^+ \in \mathcal{U}(\hat{x}^+)$ and $\hat{x}^+ \in \mathcal{S}$ from the proof of Lemma 2.36.

For the perturbed system and any $w \in \mathbb{W}$, we denote $x^+ = f(x, \kappa(x), w)$ and $x_N^+ = \phi(N; x, \tilde{\mathbf{u}}^+)$. We first establish that $x^+ \in \mathcal{X}$. Note that $\mathcal{S}$ is compact since $V^0(\cdot)$ is lower-semicontinuous and $\mathcal{X}$ is closed by Proposition 2.30. Also, the function $V_f(\hat{\phi}(N; x, \mathbf{u}))$ is continuous because it is a composition of a finite number of continuous functions. From Proposition 2.34, there exists $\sigma_f(\cdot) \in \mathcal{K}_\infty$ such that

$$|V_f(\hat{\phi}(N; x^+, \tilde{\mathbf{u}}^+)) - V_f(\hat{\phi}(N; \hat{x}^+, \tilde{\mathbf{u}}^+))| \leq \tilde{\sigma}_f(|x^+ - \hat{x}^+|) \tag{2.35}$$

for all $\hat{x}^+ \in \mathcal{S}$, $\tilde{\mathbf{u}}^+ \in \mathbb{U}^N$, and $x^+ \in \mathbb{R}^n$. Since $f(\cdot)$ is also continuous, there also exists $\sigma_x(\cdot) \in \mathcal{K}_\infty$ such that

$$|x^+ - \hat{x}^+| = |f(x, \kappa(x), w) - f(x, \kappa(x), 0)| \leq \sigma_x(|w|) \tag{2.36}$$

for all $x \in \mathcal{S}$ and $w \in \mathbb{R}^q$. We then combine (2.35) and (2.36) and note that $\hat{\phi}(N; \hat{x}^+, \tilde{\mathbf{u}}^+) = f(\hat{x}_N, \kappa_f(\hat{x}_N), 0)$ to give,

$$V_f(x_N^+) - V_f(f(\hat{x}_N, \kappa_f(\hat{x}_N), 0)) \leq \tilde{\sigma}_f(\sigma_x(|w|)) =: \sigma_f(|w|)$$

in which $\sigma_f(\cdot) \in \mathcal{K}_\infty$. We further use Assumption 2.29 to give

$$\begin{aligned} V_f(x_N^+) &\leq V_f(\hat{x}_N) - \ell(\hat{x}_N, \kappa_f(\hat{x}_N)) + \sigma_f(|w|) \\ &\leq V_f(\hat{x}_N) - \alpha_\ell(|x|) + \sigma_f(|w|) \end{aligned} \tag{2.37}$$

for all $x \in \mathcal{S}$, $w \in \mathbb{W}$. By continuity of $V_f(\cdot)$ and Proposition 2.33, there also exists $\alpha_f(\cdot) \in \mathcal{K}_\infty$ such that

$$V_f(x) = |V_f(x) - V_f(0)| \leq \alpha_f(|x|) \tag{2.38}$$

for all $x \in \mathbb{R}^n$.

Recall that $\mathbb{X}_f := \{x \in \mathbb{R}^n : V_f(x) \leq \tau\}$ and $\hat{x}_N \in \mathbb{X}_f$ because $\mathbf{u}^0 \in \mathcal{U}(x)$. We now define the constant $\delta_1 := \min\{\sigma_f^{-1}(\tau/2), \sigma_f^{-1}(\alpha_\ell(\alpha_f^{-1}(\tau/2)))\}$ and demonstrate that $x_N^+ \in \mathbb{X}_f$ for all $|w| \leq \delta_1$. Note that $\delta_1 > 0$ because $\tau > 0$. If $V_f(\hat{x}_N) \leq \tau/2$, we have from (2.37) that

$$\begin{aligned} V_f(x_N^+) &\leq \tau/2 + \sigma_f(\delta_1) \\ &\leq \tau/2 + \tau/2 = \tau \end{aligned}$$

If $\tau/2 \leq V_f(\hat{x}_N) \leq \tau$, we have that $|x| \geq \alpha_f^{-1}(\tau/2)$. We then substitute this bound into (2.37) to give

$$V_f(x_N^+) \leq \tau - \alpha_\ell(\alpha_f^{-1}(\tau/2)) + \sigma_f(\delta_1) \leq \tau$$

because $\sigma_f(\delta_1) \leq \alpha_\ell(\alpha_f^{-1}(\tau/2))$ by the definition of δ_1. So $V_f(x_N^+) \leq \tau$ and therefore $x_N^+ \in \mathbb{X}_f$. Since $x_N^+ \in \mathbb{X}_f$ and $\tilde{\mathbf{u}}^+ \in \mathbb{U}^N$, we have that $\tilde{\mathbf{u}}^+ \in \mathcal{U}(x^+)$ and therefore $x^+ \in \mathcal{X}$. Thus, the extended input trajectory is feasible for any $w \in \mathbb{W}$.

We now establish that $V^0(x^+) \leq \rho$. Note that $V(\cdot)$ is continuous because it is the composition of a finite number of continuous functions. By Proposition 2.34 there exists $\sigma_\rho(\cdot) \in \mathcal{K}_\infty$ such that

$$|V(x^+, \tilde{\mathbf{u}}^+) - V(\hat{x}^+, \tilde{\mathbf{u}}^+)| \leq \sigma_\rho(|x^+ - \hat{x}^+|)$$

for all $x^+ \in \mathcal{X}$, $\hat{x}^+ \in \mathcal{S}$, and $\tilde{\mathbf{u}}^+ \in \mathbb{U}$. Therefore,

$$V(x^+, \tilde{\mathbf{u}}^+) - V(\hat{x}^+, \tilde{\mathbf{u}}^+) \leq \sigma_\rho(\sigma_x(|w|)) =: \sigma(|w|)$$

and we combine this equation with (2.34) to give

$$V^0(x^+) \leq V(x^+, \tilde{\mathbf{u}}^+) \leq V^0(x) - \alpha_\ell(|x|) + \sigma(|w|) \tag{2.39}$$

For Lemma 2.35, there exists $\alpha_2(\cdot) \in \mathcal{K}_\infty$ such that $V^0(x) \leq \alpha_2(|x|)$ for all $x \in \mathcal{X}$.

We define $\delta_2 := \max\{\sigma^{-1}(\rho/2), \sigma^{-1}(\alpha_\ell(\alpha_2^{-1}(\rho/2)))\}$ and demonstrate that $V^0(x^+) \leq \rho$ for all $|w| \leq \delta_2$. Note that $\delta_2 > 0$ because $\rho > 0$. If $V^0(x) \leq \rho/2$, we have from (2.39) that

$$V^0(x^+) \leq \rho/2 + \sigma(\delta_2) \leq \rho/2 + \rho/2 \leq \rho$$

If $\rho/2 \leq V^0(x) \leq \rho$, we have that $|x| \geq \alpha_2^{-1}(\rho/2)$. We then substitute this bound into (2.39) to give

$$V^0(x^+) \leq \rho - \alpha_\ell(\alpha_2^{-1}(\rho/2)) + \sigma(\delta_2) \leq \rho$$

because $\sigma(\delta_2) \leq \alpha_\ell(\alpha_2^{-1}(\rho/2))$ by the definition of δ_2. So $V^0(x^+) \leq \rho$.

Thus, we define $\delta := \max\{\delta_1, \delta_2\}$ and note that $\delta > 0$ because $\delta_1 > 0$ and $\delta_2 > 0$. So for any $|w| \leq \delta$, we have that $x^+ \in \mathcal{X}$ and $x^+ \in \{x \in \mathbb{R}^n : V^0(x) \leq \rho\}$. Therefore, $x^+ \in \mathcal{S} = \{x \in \mathbb{R}^n : V^0(x) \leq \rho\} \cap \mathcal{X}$ and since the choice of x was arbitrary, we have that $\mathcal{S}$ is RPI for the closed-loop system with $\mathbb{W} \subseteq \{w \in \mathbb{R}^q : |w| \leq \delta\}$, i.e., (i) holds.

In the process of establishing (i), we completed the vast majority of effort to establish the remaining results in Theorem 2.32. From Lemma 2.36, we have that there exist $\alpha_1(\cdot), \alpha_2(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_1(|x|) \leq V^0(x) \leq \alpha_2(|x|)$ for all $x \in \mathcal{S} \subseteq \mathcal{X}$. Furthermore, (2.39) holds for all $x \in \mathcal{S}$ and therefore $V^0(\cdot)$ is an ISS Lyapunov function in the RPI set $\mathcal{S}$. In the proof of Lemma 2.36, we also show that $\ell(x, \kappa(x)) \leq V^0(x)$ for all $x \in \mathcal{S} \subseteq \mathcal{X}$. Thus, we use Proposition 2.15 and Proposition 2.26 to give (ii) and (iv). We then use Proposition 2.21 in combination with Proposition 2.18 and Proposition 2.27 to give (iii) and (v).

□

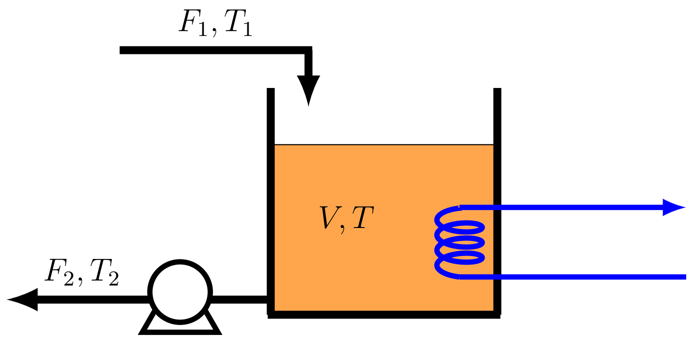

Figure 2.3: Diagram of the heat exchanger example. The process fluid (orange) is cooled by the refrigerant (blue).

2.8 Example

We consider a simple example of MPC based on the process shown in Figure 2.3.[12] The heat exchanger uses a refrigerant (Ammonia, blue) to cool the process fluid (50/50%v Ethylene Glycol/water, orange). The inlet flow-rate F_1 and temperature T_1 are determined by the upstream process and subject to perturbations. The nominal values of these variables are $F_1 = 50$ L/min and $T_1 = 300$ K. We use w_1, w_2 to represent perturbations to these nominal values. We assume, of course, that the tank is well-mixed and therefore $T_2 = T$. The goal is to maintain V and T, and therefore tank height and T_2, at their respective setpoints by manipulating the outlet flow rate F_2 and refrigerant temperature T_r. We assume that the refrigerant temperature can be manipulated "instantly" (at a much faster time-scale than the heat exchanger). With appropriate mass and energy balances, we obtain the following differential equations.

$$\begin{aligned}\frac{dV}{dt} &= 50 - F_2 + w_1 \\ \frac{dT}{dt} &= \frac{50 + w_1}{V}(300 + w_2 - T) + \frac{\alpha}{\rho \hat{C}_p V}(T_r - T)\end{aligned}$$

[12] which is remarkably similar to a certain undergraduate process control project.

in which α is the heat transfer coefficient between the refrigerant and the process fluid and $\hat{C}_p$, ρ are the specific heat capacity and density of the process fluid with $\alpha/(\rho C_p) = 300$ L/min.

The inlet flow rate can take values in the range $F_2 \in [0, 80]$ and the refrigerant temperature may take values in the range $T_r \in [245, 255]$. The target setpoints are $V^s = 1000$ L and $T^s = 255$ K, which leads to the corresponding (nominal) steady state inputs of $F_2^s = 50$ L/min and $T_r^s = 247.5$ K. We define the state, input, and disturbance in deviation variables as $x = [V - V^s, T - T^s]'$, $u = [F_2 - F_2^s, T_r - T_r^s]'$, and $w = [w_1, w_2]'$ with the corresponding (vector-valued) differential equation

$$\frac{dx}{dt} = F(x, u, w)$$

This continuous-time differential equation is then discretized using the Runge-Kutta method (4^{th} order) with one minute time steps ($\Delta = 1$) to give the difference equation

$$x^+ = f(x, u, w)$$

Although the state of the system is defined relative to the steady-state values, we sometimes plot the absolute values of these variables to exemplify the physical meaning of these variables.

For the MPC problem, we define the standard quadratic stage cost as

$$\ell(x, u) = x'Qx + u'Ru$$

with the diagonal matrices

$$Q := \begin{bmatrix} 0.0025 & 0 \\ 0 & 1 \end{bmatrix} \qquad R := \begin{bmatrix} 0.0625 & 0 \\ 0 & 1 \end{bmatrix}$$

To construct the terminal cost and constraints, we linearized the nonlinear difference equation about $(x, u) = (0, 0)$ and compute the LQR solution for this linearized system with the same quadratic stage cost $\ell(\cdot)$ that was just defined. The solution to this problem results in the cost matrix P and feedback gain K that we use to define the terminal cost $V_f(x) := 2x'Px$, terminal constraint $\mathbb{X}_f := \{x : 2x'Px \leq \tau\}$ for $\tau = 1$, and terminal control law $\kappa_f(x) = Kx$. We verify that $V_f(f(x, \kappa_f(x), 0)) \leq V_f(x) - \ell(x, \kappa_f(x))$ for all $x \in \mathbb{X}_f$. We choose a horizon of $N = 20$. Thus, the formulated MPC problem satisfies Assumptions 2.1, 2.23, 2.28 and 2.29.

We solve the MPC problem from the initial state $x(0) = [100, 15]'$, or $V(0) = 1100$ L and $T(0) = 270$ K. In Figure 2.4, we display phase plots of the optimal state and input trajectories. Note that these are not necessarily the closed-loop trajectories for the system.Observe that the input constraints on u_2 are active for the first two inputs in the optimal trajectory and the terminal constraint is easily satisfied.

In Figure 2.5, we plot the nominal closed-loop trajectory ($\mathbf{w} = 0$) from the same initial condition ($x(0) = [100, 15]'$). Both elements of the state (V and T) converge to their respective steady-state targets as t increases, i.e., the target steady state is asymptotically stable.

We now consider disturbances in the closed-loop system. Specifically, we allow fluctuations of ± 10 L/min in F_1 and ± 10 K in T_1 and therefore $\mathbb{W} := \{w : w_1 \in [-10, 10], w_2 \in [-10, 10]\}$, i.e., a square. We assume that both disturbances are independent with uniform distributions on their respective supports. In the top two plots of Figure 2.6, we show the closed-loop state trajectory for 30 different realizations of the disturbance trajectory. Observe that all the trajectories exhibit initial convergence towards a neighborhood of the target

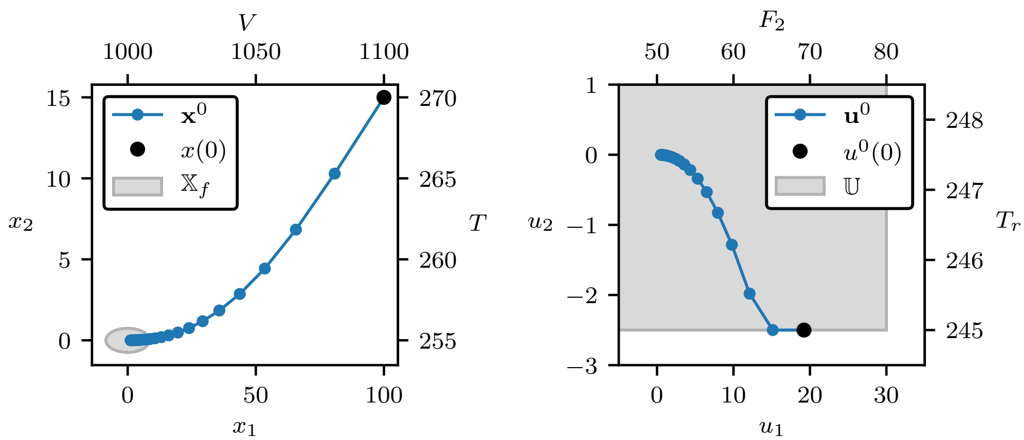

Figure 2.4: Phase plots of the optimal state ($\mathbf{x}^0$) and input trajectories ($\mathbf{u}^0$) for the heat exchanger example with input and terminal state constraint sets shown by the gray shaded regions.

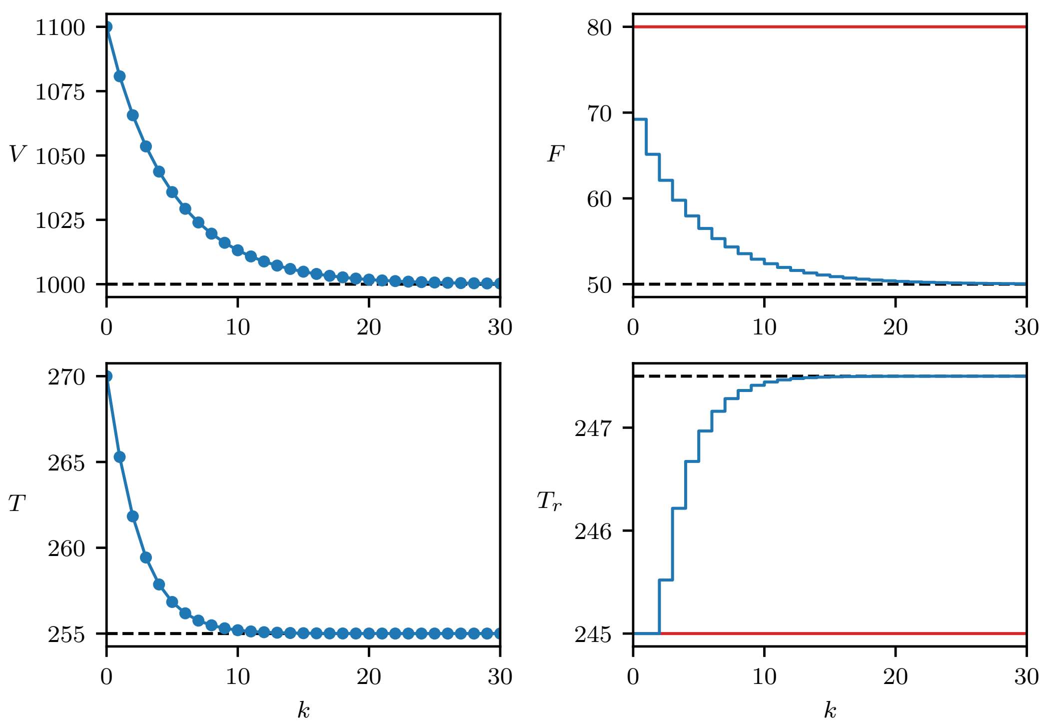

Figure 2.5: The nominal closed-loop state and input trajectories. The target steady-state values are shown with the dashed black lines and input constraints are shown in red.

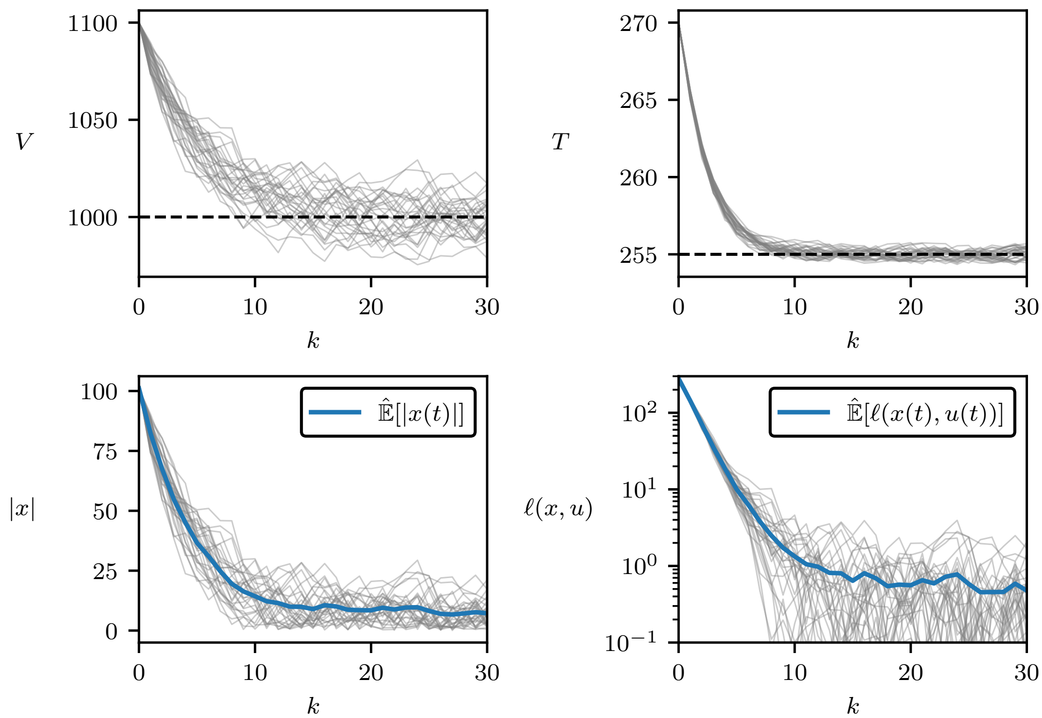

Figure 2.6: Top: The closed-loop state trajectories for 30 realizations of the disturbance trajectory. Bottom: Closed-loop state and stage cost trajectories for these 30 realizations of the disturbance trajectory with their sample averages in blue.

steady state. After this initial convergence, the trajectories remain in this neighborhood for all subsequent t.

In the bottom two plots of Figure 2.6, we show the performance of these trajectories in terms their distance to the target steady state (origin) and stage cost $\ell(x,u)$. Note that all of the trajectories exhibit initial convergence towards the origin and then remain within a neighborhood of the origin. This behavior is therefore consistent with RAS and ℓ-RAS. We then evaluate the sample average of these trajectories, which we denote as $\hat{\mathbb{E}}[\cdot]$, to approximate the expected value used in the definitions of RASiE and ℓ-RASiE. Note that these trajectories are consistent with the definitions of RASiE and ℓ-RASiE. Thus, the behavior shown in Figure 2.6 is consistent with Theorem 2.32 and demonstrates the practical implications of this theorem.

2.9 Summary

Some nonzero margin of robustness to disturbances is essential to deploy a control algorithm in industrial applications. In a deterministic context, robustness implies that arbitrarily small disturbances produce similar small deviations in the performance of the closed-loop system (RAS and ℓ-RAS). In a stochastic context, robustness implies that disturbances with arbitrarily small variances produce similar small deviations in expected value of the performance of the closed-loop system (RASiE and ℓ-RASiE). This "performance" can be characterized by the standard notion of "distance to the target steady state", i.e., $|x|$, or a more general notion in terms of the stage cost $\ell(\cdot)$ specified for the problem of interest (ℓ-RAS and ℓ-RASiE). We established sufficient conditions for these different definitions of robustness via comparison functions and ISS/SISS Lyapunov functions. Moreover, we established that the typical notion of deterministic robustness (RAS) also implies stochastic robustness (RASiE) for bounded RPI sets.

These definitions of deterministic and stochastic robustness provide different strengths and weakness. RASiE and ℓ-RASiE are useful characterizations of the closed-loop system performance for a distribution of disturbance values. RAS and ℓ-RAS, however, are much more instructive properties for *specific* realizations of the disturbance trajectory. In particular, RAS (ℓ-RAS) guarantees that as $|w| \to 0$, we recover asymptotic stability of the origin for that specific trajectory of disturbances, i.e., $\phi(k; x, \mathbf{w}) \to 0$ ($\ell(x(k), u(k)) \to 0$). RASiE (ℓ-RASiE) does not guarantee that we recover asymptotic stability as $|w| \to 0$ for a specific disturbance trajectory because, by definition, RASiE (ℓ-RASiE) considers a distribution of disturbances. For RASiE (ℓ-RASiE), we also require that $\Sigma \to 0$ to recover asymptotic stability of the origin. This fact is significant in comparison nominal and stochastic MPC in the following chapter. In summary, these definitions of stochastic robustness offer better characterizations of closed-loop performance for stochastic systems, but are nonetheless *weaker* properties than the corresponding definitions of deterministic robustness.

The main result in this chapter is Theorem 2.32, which establishes that nominal MPC ensures all of these deterministic and stochastic definitions of robustness for the closed-loop system subject to sufficiently small (but nonzero) disturbances. Nominal MPC achieves this nonzero margin of robustness without any disturbance information explicitly included in the the optimization problem. Instead, this robustness is afforded solely by feedback. As demonstrated in the heat exchanger example, this inherent robustness is often sufficient to address the relevant disturbances for an industrial process.

There may arise, however, applications in which this margin of inherent robustness is insufficient. In these applications, including stochastic information in the problem formulation may prove beneficial. This reasoning and the burgeoning field of stochastic optimization lead to the development of stochastic MPC (SMPC), which we discuss, in extensive and novel detail, in the next two chapters.

Chapter 3

Stochastic MPC

A key requirement of the robustness results established for nominal MPC is that the disturbance must be *sufficiently small* for any of these properties to hold. We can guarantee that this margin of robustness is nonzero (i.e., $\delta > 0$), but drawing any further conclusions for a general nonlinear system of significant dimension is unworkable. The argument made by the proponents of nominal MPC is instead rooted in the industrial success of the control technique. This margin of robustness is typically sufficient in industrial practice and this claim is substantiated by the numerous successful implementations of MPC in industry (Qin and Badgwell, 2003). Alternatively, one may argue that the margin of robustness afforded by nominal MPC, while often satisfactory, is inadequate for systems with exceptional safety concerns, high performance demands, and/or significant uncertainty. In these scenarios, a stochastic or worst-case model of the disturbances should be used directly in the design of the controller to ensure robustness.

This interest led to the development of stochastic MPC (SMPC) in the last two decades. The distinct feature of SMPC is that a disturbance model is included explicitly in the optimization problem with the goal of producing a feedback controller that is "more robust" than nominal MPC to a specific disturbance of interest. Given its early stage of development, however, there are naturally many limitations to the current theory of SMPC, particularly for nonlinear systems. In fact, there is not even a specific definition of stochastic robustness and authors often use a variety of different results in characterizing the properties of SMPC. Moreover, this notion of "more robust," is never clearly defined and, despite the obvious connection between nominal and stochastic MPC (and the claims made by proponents of each approach), there are no rigorous comparisons of the theoretical properties achieved by these different formulations. We attribute this limitation to the fact that closed-loop results for nominal MPC and SMPC are typically deterministic and stochastic, respectively, and are therefore not comparable across these two methods.

By introducing and establishing the stochastic robustness of nominal MPC in the previous chapter, we are uniquely positioned to address this limitation and compare nominal MPC and SMPC across multiple relevant definitions of robustness. We begin by introducing SMPC and establishing sufficient conditions for the stochastic robustness of this controller. SMPC, however, does not ensure robust asymptotic stability (RAS) for nonlinear systems and can thereby produce nonintuitive closed-loop behavior. Motivated by this shortcoming of SMPC, we then propose a constraint-tightened MPC (CMPC) formulation and establish that CMPC satisfies all the definitions of stochastic and deterministic robustness introduced in Chapter 2 for the disturbance used in the problem formulation. Through a few simple examples, we illustrate the implications of these results and demonstrate that, depending on the definition of robustness considered, SMPC is not necessarily more robust than nominal MPC even if the disturbance model is exact.

3.1 Literature review

Given a stochastic description of the uncertainty in the system, the SMPC problem is typically defined as a minimization of the expected value of a sum-of-stage costs (Farina et al., 2016; Mayne, 2016; Mesbah, 2016). The optimization problem is often subject to deterministic constraints that must hold for all realizations of the disturbance as well as probabilistic state and input constraints, sometimes referred to as chance constraints. This stochastic optimization problem, however, is more computationally demanding and further complicates the closed-loop analysis of SMPC relative to nominal MPC. Consequently, the majority of results for SMPC are restricted to linear systems. As with the previous chapter and the rest of this dissertation, we focus on the closed-loop properties of SMPC and do not directly address the topic of stochastic optimization. There is already a significant body of literature devoted to approximating and solving SMPC and stochastic optimal control problems (see Mesbah (2016) for a review).

In one of the first contributions addressing the theoretical properties of SMPC, Primbs and Sung (2009) consider linear systems with multiplicative uncertainty such that the effect of the disturbance vanishes at the origin. By assuming the terminal cost is a *global* stochastic Lyapunov function, the authors establish that the origin is asymptotically stable with probability one. For constrained linear SMPC subject to bounded disturbances, Cannon, Kouvaritakis, and co-authors use a terminal cost and terminal constraint derived from a local Lyapunov function to ensure recursive feasibility and stability in expectation of the closed-loop system (Cannon et al., 2009a,b, 2010; Kouvaritakis et al., 2010). Lorenzen et al. (2016) propose a less restrictive constraint tightening approach and establish that linear SMPC asymptotically stabilizes, with probability one, the minimal robust positive invariant set for the system. Similar results are establish in other subsequent papers for modified SMPC algorithms (Sehr and Bitmead, 2018; Hewing et al., 2020).

For *nonlinear* SMPC, Chatterjee and Lygeros (2014) establish that the expected value of the optimal cost along the closed-loop trajectory is bounded if the terminal cost is a *global* stochastic Lyapunov function. Mayne and Falugi (2019) extend these results to address constrained nonlinear systems subject to bounded, stochastic disturbances and, with terminal constraints, require the terminal cost to be only a *local* stochastic Lyapunov function. Under certain viability and stochastic controllability assumptions, nonlinear SMPC without terminal conditions is also stabilizing, but these assumptions are difficult to verify for nonlinear systems (Lorenzen et al., 2019).

3.2 The stochastic linear-quadratic regulator

Before proceeding to (nonlinear) SMPC, we again begin with the simplest version of the stochastic optimization problem: a linear, unconstrained stochastic system, i.e.,

$$x^+ = f(x, u, w) = Ax + Bu + w$$

in which $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ are matrices and (A, B) is assumed to be stabilizable. We assume that the disturbances w are i.i.d. in time, zero mean, random variables. We even allow $\mathbb{W} = \mathbb{R}^n$, i.e., w is unbounded. We use a quadratic stage cost $\ell(x, u) := x'Qx + u'Ru$ with $Q, R \succ 0$.

Since the stochastic linear-quadratic regulator considers all possible realizations of the (potentially unbouded) disturbance $w \in \mathbb{W}$, we must optimize over a sequence of feedback policies $\Pi := (\pi_0, \pi_1, \dots)$ instead of a single trajectory of control actions $\mathbf{u}$. These feedback policies are functions that map the state of the system at time k to the required input $u(k)$, i.e., $u(k) = \pi_k(x(k))$. Thus, the input at each time k is also a function of the current state.

We define the infinite horizon stochastic cost function as

$$V(x, \Pi) = \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\sum_{k=0}^{N-1}\Big(x(k)'Qx(k) + \pi_k(x(k))'R\pi_k(x(k))\Big) + x(N)'Qx(N)\right]$$

in which $x(k+1) = Ax(k)+B\pi_k(x(k))+w(k)$ is the stochastic system evolution, $x(0) = x$ is the deterministic initial condition, and Π is the infinite trajectory of control policies. Note that we have also normalized this infinite horizon cost by the horizon length $\frac{1}{N}$ so that $V(x, \Pi)$ is finite. The infinite horizon stochastic optimal control problem is then given by

$$V^0(x) = \min_{\Pi} V(x, \Pi)$$

for all $x \in \mathbb{R}^n$ in which the optimal trajectory of policies is denoted $\Pi^0 = (\pi_0^0, \pi_1^0, \dots)$.

As with the nominal LQR problem, one can use dynamic programming to show that the solution to this infinite horizon optimization problem is given by the unique stabilizing solution to the discrete-time algebraic Riccati eqution (DARE), i.e., the matrix $P \succ 0$ that solves

$$P = A'PA - (A'PB)(R + B'PB)^{-1}(B'PA) + Q$$

such that $A + BK$ is Schur stable with $K = -(R + B'PB)^{-1}(B'PA)$ (Bertsekas, 2017, s. 3.1). The optimal cost is given by $V^0(x) := \text{tr}(P\Sigma)$ in which $\Sigma \succeq 0$ is the covariance of w. Furthermore, the control law, defined by the first control policy in the optimal solution, is given by $\kappa(x) = \pi_0^0(x) = Kx$.

The implications of this result are remarkable. The control laws derived from the stochastic LQR and nominal LQR are in fact *identical*. We solve the same DARE problem to give the same matrix P, the same feedback gain K, and the same control law $\kappa(x) = Kx$. This property is known as *certainty equivalence*, as first proposed in Simon (1956) and Theil (1957). A discussion of certainty equivalence as it pertains to stochastic optimal control can be found in Van de Water and Willems (1981). More recently, the principle of certainty equivalence has found interest in the machine learning research community as well (Mania et al., 2019).

As it pertains to SMPC, this result means that including stochastic information in the optimization problem produces a different and potentially superior controller only if we consider non-quadratic stage costs, nonlinear systems, or problems in which the input/state constraints are particularly relevant to the control law. By contrast, SMPC problems with approximately linear dynamics, inactive constraints, and quadratic stage costs are liable to produce control laws and closed-loop systems that are very similar, if not the same, as nominal MPC.

3.3 Problem formulation and preliminaries

We consider the same stochastic system introduced in (2.4), but we reintroduce the main features of this system here for convenience. We consider the following discrete-time stochastic system

$$x^+ = f(x,u,w) \quad f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^n \tag{3.1}$$

The input is subject to hard constraints $u \in \mathbb{U} \subseteq \mathbb{R}^m$ and we use $\mathbb{W} \subseteq \mathbb{R}^q$ to denote the set for the disturbances. We make the same assumption for the disturbances used in the previous chapter.

Assumption 3.1 (Disturbances)**.** The disturbances $w \in \mathbb{W}$ are random variables that are i.i.d. in time and zero mean. The set $\mathbb{W}$ is compact and contains the origin.

Given Assumption 3.1, we denote the probability measure of w as $\mu : \mathcal{B}(\mathbb{W}) \to [0,1]$ and we use $\mathcal{M}(\mathbb{W})$ to denote the collection of all possible probability measure $\mu : \mathcal{B}(\mathbb{W}) \to [0,1]$ that satisfy Assumption 3.1, i.e., $\mathbb{E}[w] = 0$ for all $\mu \in \mathcal{M}(\mathbb{W})$. For any $\mu \in \mathcal{M}(\mathbb{W})$, we denote the covariance matrix of w as Σ.

For the i.i.d. random variables $(w(i), w(i+1), \ldots, w(i+N-1))$ and $N \in \mathbb{I}_{\geq 1}$, their joint distribution measure $\mu^N : \mathcal{B}(\mathbb{W}^N) \to [0,1]$ is defined as

$$\mu^N(F) := \mu(F_i)\mu(F_{i+1}) \ldots \mu(F_{i+N-1})$$

for all $F = (F_i, F_{i+1}, \ldots, F_{i+N-1}) \in \mathcal{B}(\mathbb{W}^N)$. We define the sequence of random variables starting from $i = 0$ to $k \in \mathbb{I}_{\geq 0}$ as $\mathbf{w}_k := (w(0), w(1), \ldots, w(k-1))$. For this sequence, we define the expected value of a Borel measurable function $g : \mathbb{W}^i \to \mathbb{R}_{\geq 0}$ as the following Lebesgue integral.

$$\mathbb{E}\left[g(\mathbf{w}_i)\right] := \int_{\mathbb{W}^N} g(\mathbf{w}_i) d\mu^N(\mathbf{w}_i)$$

Since SMPC considers all possible realizations of the disturbance within the optimization problem, selecting a single trajectory of inputs $\mathbf{u}$ for the finite horizon $N \in \mathbb{I}_{\geq 1}$ that satisfies all the required constraints (in particular the terminal constraint) is typically impossible. Instead, we must embed feedback within the optimization problem and optimize over a trajectory of control policies, i.e., a sequence of functions $(\pi_0, \pi_1, \ldots, \pi_{N-1})$ such that the control action at each point in the open-loop optimization problem is given by $u(k) = \pi_k(x(k))$. In practice, however, optimizing over an infinite dimensional object such as a continuous function is intractable. We are only able to compute such a function for the stochastic LQR problem by exploiting the convenient structure of the linear system and quadratics costs. For general nonlinear systems and/or constrains, we loose this capability. To formulate a tractable optimization problem, we instead define a parameterized control policy $\pi : \mathbb{R}^n \times \mathbb{V} \to \mathbb{R}^m$ in which $x \in \mathbb{R}^n$ is the current state of the system and $v \in \mathbb{V} \subseteq \mathbb{R}^l$ are the parameters in the control policy. A common choice for this parameterization is $\pi(x, v) := Kx + v$ in which K is a fixed matrix. Thus, we optimize over a trajectory of parameters $\mathbf{v} := (v(0), v(1), \ldots, v(N-1))$ for this policy and thereby define a trajectory of control policies. As with nominal MPC, we do not implement the entire trajectory of control policies and instead solve for a new trajectory of parameters at each time step.

The resulting system of interest is therefore

$$x^+ = f(x, \pi(x, v), w) \tag{3.2}$$

We use $\hat{\phi}^s(k; x, \mathbf{v}, \mathbf{w})$ to denote the state of the system (3.2) at time $k \in \mathbb{I}_{0:N}$, given the initial condition $x \in \mathbb{R}^n$, the trajectory of control policies $\mathbf{v} \in \mathbb{V}^N$, and the disturbance sequence $\mathbf{w} \in \mathbb{W}^N$. Since we are considering the disturbances directly in the optimization problem, we can consider hard input and state constraints, i.e.,

$$(x, u) \in \mathbb{Z} \subseteq \mathbb{R}^n \times \mathbb{U}$$

Remark 3.2. In some SMPC formulations, probabilistic constraints are also defined, e.g.,

$$\Pr(f(x, u, w) \in \tilde{\mathbb{X}}) \geq 1 - \varepsilon \tag{3.3}$$

for a set $\tilde{\mathbb{X}} \subseteq \mathbb{R}^n$ and constant $\varepsilon \in [0, 1]$. We note, however, that these (one step ahead) probabilistic constraints may be reformulated using the function

$$G(x, u) := 1 - \int_{\mathbb{W}} I_{\tilde{\mathbb{X}}}\left(f(x, u, w)\right) d\mu(w)$$

and the constraint set $\tilde{\mathbb{Z}}_\varepsilon := \{(x, u) : G(x, u) \leq \varepsilon\}$. Thus, (x, u) satisfy (3.3) if and only if $(x, u) \in \tilde{\mathbb{Z}}_\varepsilon$ and we can simply redefine the hard constraints as $\mathbb{Z} \cap \tilde{\mathbb{Z}}_\varepsilon$ to include this probabilistic constraint. Furthermore, we can show that $\tilde{\mathbb{Z}}_\varepsilon$ is in fact closed if $\tilde{\mathbb{X}}$ is closed and $f(\cdot)$ is continuous (McAllister and Rawlings, 2022a, Lemma 1). Thus, we restrict our attention to only hard constraints $(x, u) \in \mathbb{Z}$.

For a horizon $N \in \mathbb{I}_{\geq 1}$ and the terminal constraints $\mathbb{X}_f \subseteq \mathbb{R}^n$, we define the admissible parameter trajectories and feasible initial states as

$$\begin{aligned}
\mathcal{V}(x) := \{\mathbf{v} \in \mathbb{V}^N : (x(k), \pi(x(k), v(k))) \in \mathbb{Z}\ \forall \mathbf{w} \in \mathbb{W}^N,\ k \in \mathbb{I}_{0:N-1}; \\
x(N) \in \mathbb{X}_f\ \forall \mathbf{w} \in \mathbb{W}^N\} \\
\mathcal{X}^s := \{x : \mathcal{V}(x) \neq \emptyset\}
\end{aligned}$$

in which $x(k) = \hat{\phi}^s(k; x, \mathbf{v}, \mathbf{w})$. We use a stage cost $\ell : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ and terminal cost $V_f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ to define the the function

$$J(x, \mathbf{v}, \mathbf{w}) = \sum_{k=0}^{N-1} \ell(x(k), \pi(x(k), v(k))) + V_f(x(N))$$

in which $x(k) = \hat{\phi}^s(k; x, \mathbf{v}, \mathbf{w})$. We define the SMPC cost function as the expected value of $J(\cdot)$, i.e.,

$$V_\mu^s(x, \mathbf{v}) := \int_{\mathbb{W}^N} J(x, \mathbf{v}, \mathbf{w}) d\mu^N(\mathbf{w})$$

The optimization problem for any $x \in \mathcal{X}^s$ is defined as

$$\mathbb{P}_\mu^s(x) : V_\mu^{s0}(x) := \min_{\mathbf{v} \in \mathcal{V}(x)} V_\mu^s(x, \mathbf{v}) \tag{3.4}$$

and the optimal solution for a given initial state are denoted $\mathbf{v}_\mu^{s0}(x) := \arg\min_{\mathbf{v} \in \mathcal{V}(x)} V_\mu^s(x, \mathbf{v})$. Note that $\mathbf{v}_\mu^{s0}$ is a set-valued mapping because there may be multiple solution to $\mathbb{P}_\mu^s(x)$.

As with nominal MPC, we use a selection rule to define a single-valued control law $\kappa : \mathcal{X}^s \to \mathbb{U}$ such that

$$\kappa_\mu^s(x) \in \{\pi(x, v(0)) : \mathbf{v} \in \mathbf{v}_\mu^{s0}(x)\}$$

for all $x \in \mathcal{X}^s$. Note that because we have included the probability measure explicitly in the optimization problem through the cost function, the optimization problem varies with the probability measure $\mu \in \mathcal{M}(\mathbb{W})$. Thus, the optimal cost $V_\mu^{s0}(\cdot)$, optimal solution $\mathbf{v}_\mu^{s0}(\cdot)$, and control law $\kappa_\mu^s(\cdot)$ all vary with μ. The resulting closed-loop system is then

$$x^+ = f(x, \kappa_\mu^s(x), w) \tag{3.5}$$

We use $\phi_\mu^s(k; x, \mathbf{w}_k)$ to denote the state of the system (3.5) at time $k \in \mathbb{I}_{\geq 0}$, given the initial condition $x \in \mathcal{X}^s$, disturbance sequence $\mathbf{w}_k \in \mathbb{W}^k$, and probability measure $\mu \in \mathcal{M}(\mathbb{W})$.

Since the control law varies with μ, so does the closed-loop trajectory $\phi^s_\mu(\cdot)$. This fact becomes relevant when discussing the robustness of SMPC and requires that we extend the definition of RASiE used in the previous chapter to allow for a control law that varies with μ.

For SMPC, we require some of the same assumptions already introduced for nominal MPC that we restate here with a slightly different organization.

Assumption 3.3 (Continuity of system and cost)**.** The system $f : \mathbb{R}^n \times \mathbb{U} \times \mathbb{W} \to \mathbb{R}^n$, stage cost $\ell : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}_{\geq 0}$, and terminal cost $V_f : \mathbb{X}_f \to \mathbb{R}_{\geq 0}$ are continuous and satisfy $f(0,0,0) = 0$, $\ell(0,0) = 0$, and $V_f(0) = 0$.

Assumption 3.4 (Stage cost bound)**.** There exists $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_\ell(|x|) \leq \ell(x,u)$ for all $(x,u) \in \mathbb{R}^n \times \mathbb{U}$.

We also require a few assumptions that differ from the versions used for nominal MPC.

Assumption 3.5 (Properties of the constraint sets; SMPC)**.** The set $\mathbb{Z}$ is closed and contains the origin. The sets $\mathbb{U}$ and $\mathbb{X}_f$ are compact and contain the origin. The set $\mathbb{X}_f$ contains the origin in its interior. The set $\mathcal{X}^s$ is bounded.

Assumption 3.6 (Terminal ingredients; SMPC)**.** There exists a continuous terminal control law $\kappa_f : \mathbb{X}_f \to \mathbb{U}$ such that for all $x \in \mathbb{X}_f$,

$$f(x, \kappa_f(x), w) \in \mathbb{X}_f \quad \forall w \in \mathbb{W} \tag{3.6}$$

$$V_f(f(x, \kappa_f(x), 0)) \leq V_f(x) - \ell(x, \kappa_f(x)) \tag{3.7}$$

Furthermore, $(x, \kappa_f(x)) \in \mathbb{Z}$ and $\pi(x, 0) = \kappa_f(x)$ for all $x \in \mathbb{X}_f$.

Assumption 3.7 (Parameterization)**.** The set $\mathbb{V}$ is compact and contains the origin. The function $\pi : \mathbb{R}^n \times \mathbb{V} \to \mathbb{R}^m$ is continuous.

Assumptions 3.5 and 3.6 are the versions of Assumptions 2.28 and 2.29 adjusted for SMPC. Assumption 3.5 addresses the additional state constraints added to the SMPC problem while dropping the requirement that $\mathbb{X}_f$ is a level set of the terminal cost. Assumption 3.6 requires, in addition to the nominal cost decrease in the terminal region, that the terminal control law renders the terminal set RPI. This requirement ensures that the SMPC algorithm is robustly recursively feasible. Assumption 3.7 sets some basic requirements for the control law parameterization.

An additional requirement in Assumption 3.5 relative to Assumption 2.28 is that $\mathcal{X}^s$ is bounded. This restriction, however, is minor. Most physical systems in engineering applications admit upper and lower bounds on their state (e.g., mole fraction is between zero and one, temperature of a reactor is lower bounded by the temperature of the coolant/inlet stream and upper bounded by an adiabatic limit). Moreover, discretization of continuous ordinary differential equations produces a discrete time system such that $f^{-1}(X) = \{(x,u) \in \mathbb{X} \times \mathbb{U} : f(x,u,0) \in X\}$ is bounded for all bounded X. This fact combined with compact $\mathbb{U}$ and $\mathbb{X}_f$ ensures that $\mathcal{X}$ for nominal MPC and $\mathcal{X}^s$ for SMPC are bounded sets. For a proof of this result and further discussion see Rawlings et al. (2020, Prop. 2.10(d)) and note that $\mathcal{X} \subseteq \mathcal{X}^s$ since $0 \in \mathbb{W}$ (See Lemma 3.23).

Although Assumption 3.6 may seem significantly stronger than Assumption 2.29, the terminal cost and constraint used in nominal MPC is often sufficient for SMPC if the disturbances are again assumed to be sufficiently small (McAllister and Rawlings, 2022d, Lemma 25). The requirements in Assumption 3.6 indicate that the terminal ingredients for SMPC must be compatible with the disturbance of interest. Sufficiently large disturbances may render the construction of a suitable terminal control law and terminal set either difficult of impossible if we consider nonlinear systems and/or input constraints. Assumption 3.6 also ensures that $\mathbb{X}_f \subseteq \mathcal{X}^s$ and therefore $\mathcal{X}^s$ is not empty (See Lemma 3.23).

While not explicitly stated, a key assumption in this chapter and throughout SMPC literature is that the probability measure and set $\mathbb{W}$ used in the SMPC problem formulation are identical to that of the plant. This assumption is, of course, idealized and typically not satisfied for any practical implementation of SMPC. Nonetheless, we proceed, at least initially, under this assumption. The study of idealized SMPC is analogous to that of nominal stability properties for a control algorithm. The merit of this analysis is in establishing the best performance one can expected from SMPC to serve as a baseline. If the performance is not satisfactory under ideal conditions then there is little incentive to study nonideal conditions.

Before proceeding to any closed-loop properties for SMPC, we must first establish that a solution to $\mathbb{P}^s_\mu(x)$ exists for all $x \in \mathcal{X}^s$ and $\mu \in \mathcal{M}(\mathbb{W})$ and verify that $\kappa^s_\mu(\cdot)$ is a Borel measurable function for all $\mu \in \mathcal{M}(\mathbb{W})$. As with nominal MPC, verifying these properties is essential to guarantee the veracity of the subsequent analysis. For nonlinear MPC, however, these topics are seldom discussed. Authors either explicitly assume that the “minimization problem is well-defined” and “the control law is measurable”, or omit this discussion. By contrast, these properties are well established in the larger and related field of stochastic optimal control (Bertsekas and Shreve, 1978).

Fortunately, the assumptions that we already introduced are sufficient to guarantee that a solution to the minimization problem exists. Moreover, we can also establish that $\mathbf{v}^s_\mu(\cdot)$ is a Borel measurable function via the same results in Bertsekas and Shreve (1978) that we used for nominal MPC. We summarize these results in the following proposition and further details can be found in McAllister and Rawlings (2022a).

Proposition 3.8. *Let Assumptions 3.3 and 3.5 to 3.7 hold. Then $\mathbb{P}^s_\mu(x)$ has a solution for all $x \in \mathcal{X}^s$, $\mathcal{X}^s$ is closed, $V^{s0}_\mu : \mathcal{X}^s \to \mathbb{R}_{\geq 0}$ is lower semicontinuous, and $\mathbf{v}^{s0}_\mu : \mathcal{X}^s \rightrightarrows \mathbb{V}^N$ is a Borel measurable function for all $\mu \in \mathcal{M}(\mathbb{W})$. Furthermore, there exists a Borel measurable selection rule such that the single-valued control law $\kappa^s_\mu : \mathcal{X}^s \to \mathbb{U}$ is Borel measurable and satisfies $\kappa^s_\mu(x) \in \{\pi(x, v(0)) : \mathbf{v} \in \mathbf{v}^{s0}_\mu(x)\}$ for all $x \in \mathcal{X}^s$.*

We can now discuss the robustness of the closed-loop trajectory.

3.4 Extended definitions of stochastic robustness

The definition of RASiE provided in the previous chapter (Definition 2.16) assumes a fixed control law for all $\mu \in \mathcal{M}(\mathbb{W})$. For nominal MPC this characterization is suitable, but for SMPC the control law is not fixed and we must consider this fact in the definition of stochastic robustness. Thus, we redefine RASiE as follows while allowing the control law $\kappa_\mu : \mathcal{X} \to \mathbb{U}$ and therefore closed-loop trajectory $\phi_\mu(k; x, \mathbf{w}_k)$ to vary with $\mu \in \mathcal{M}(\mathbb{W})$. Note that since the constraints in SMPC do not vary with μ, the feasible set $\mathcal{X}^s$ remains fixed for all $\mu \in \mathcal{M}(\mathbb{W})$ and this fact is reflected in the subsequent definitions.[1]

Definition 3.9 (RASiE)**.** The origin is robustly asymptotically stable in expectation (RASiE) for a system $x^+ = f(x, \kappa_\mu(x), w)$, $w \in \mathbb{W}$ in an RPI set $\mathcal{X}$ if there exist $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$\mathbb{E}\left[|\phi_\mu(k; x, \mathbf{w}_k)|\right] \leq \beta(|x|, k) + \gamma(\text{tr}(\Sigma)) \tag{3.8}$$

for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$.

Note that the upper bound depends on the probability measure through only the argument $\text{tr}(\Sigma)$. The functions $\beta(\cdot)$ and $\gamma(\cdot)$ are the same for all $\mu \in \mathcal{M}(\mathbb{W})$. We can similarly extend the definition of ℓ-RASiE to allow for control laws and closed-loop trajectories that vary with $\mu \in \mathcal{M}(\mathbb{W})$.

Definition 3.10 (ℓ-RASiE)**.** The origin is ℓ-RASiE with respect to the stage cost $\ell(x, \kappa_\mu(x))$ for a system $x^+ = f(x, \kappa_\mu(x), w)$, $w \in \mathbb{W}$ in an RPI set $\mathcal{X}$ if there exist $\beta(\cdot) \in \mathcal{KL}$ and

[1] If probabilistic constraints are included, however, the set $\mathcal{X}^s$ may vary with μ.

$\gamma(\cdot) \in \mathcal{K}$ such that

$$\mathbb{E}\left[\ell(x(k), \kappa_\mu(x(k)))\right] \leq \beta(|x|, k) + \gamma(\text{tr}(\Sigma)) \tag{3.9}$$

in which $x(k) := \phi_\mu(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$.

In both cases, these definitions are generalizations of their previous versions in Definitions 2.16 and 2.25 and are equivalent to their previous versions for control laws that do not vary with μ, i.e., $\kappa_\mu(\cdot) = \kappa(\cdot)$.

We also use an SISS Lyapunov function specifically tailored for SMPC. Since we intend to use the optimal cost function for SMPC as the SISS Lyapunov function, we must allow this Lyapunov function to vary with μ. Moreover, the optimal cost function for SMPC may not satisfy $V_\mu^{s0}(0) = 0$. Thus, we need to allow the upper bound for the SISS Lyapunov function to also vary with μ and the corresponding Σ. We redefine the SISS Lyapunov function as follows.

Definition 3.11 (SISS Lyapunov function)**.** The Borel measurable function $V_\mu : \mathcal{X} \to \mathbb{R}_{\geq 0}$, defined for all $\mu \in \mathcal{M}(\mathbb{W})$, is an SISS Lyapunov function for the system $x^+ = f(x, \kappa_\mu(x), w)$, $w \in \mathbb{W}$ in an RPI set $\mathcal{X}$ if there exist $\alpha_1(\cdot), \alpha_2(\cdot), \alpha_3(\cdot) \in \mathcal{K}_\infty$ and $\sigma_2(\cdot), \sigma_3(\cdot) \in \mathcal{K}$ such that

$$\alpha_1(|x|) \leq V_\mu(x) \leq \alpha_2(|x|) + \sigma_2(\text{tr}(\Sigma)) \tag{3.10}$$

$$\int_{\mathbb{W}} V_\mu(f(x, \kappa_\mu(x), w)) d\mu(w) \leq V_\mu(x) - \alpha_3(|x|) + \sigma_3(\text{tr}(\Sigma)) \tag{3.11}$$

for all $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathbb{W})$.

We note that this definition is a generalization of Definition 2.17 for SISS Lyapunov functions. For a control law and Lyapunov function that do not vary with μ, an SISS Lyapunov function according to Definition 2.17 is also an SISS Lyapunov function according to Definition 3.11, i.e., $\kappa_\mu(\cdot) = \kappa(\cdot)$ and $V_\mu(\cdot) = V(\cdot)$ with $\sigma_3(\cdot) = \sigma(\cdot)$ and any $\sigma_2(\cdot) \in \mathcal{K}$. We now establish that the existence of an SISS Lyapunov function in terms of Definition 3.11 is a sufficient condition for RASiE in terms of Definition 3.9. Unless otherwise stated, we use these new definitions in lieu of the previous ones in Chapter 2 for the rest of this chapter, e.g., we take RASiE to mean Definition 3.9 and not Definition 2.16.

Proposition 3.12. *If a system $x^+ = f(x, \kappa_\mu(x), w)$, $w \in \mathbb{W}$, with $\mathbb{W}$ bounded, admits an SISS Lyapunov function in an RPI and bounded set $\mathcal{X}$, then the origin is RASiE in $\mathcal{X}$.*

Proof. We assume without loss of generality that $\alpha_3(s) \leq \alpha_2(s)$ for all $s \in \mathbb{R}_{\geq 0}$.[2] Define $\alpha_4(s) := \alpha_3\left(\alpha_2^{-1}(s/2)\right)$ and note that $\alpha_4(\cdot) \in \mathcal{K}_\infty$ and $\alpha_4(s) \leq s$ for all $s \in \mathbb{R}_{\geq 0}$ because $\alpha_3(s) \leq \alpha_2(s)$. We have the following inequality.

$$\begin{aligned}
\alpha_4(V_\mu(x)) &\leq \alpha_4\left(\alpha_2(|x|) + \sigma_2(\text{tr}(\Sigma))\right) \\
&\leq \alpha_4\left(2\alpha_2(|x|)\right) + \alpha_4\left(2\sigma_2(\text{tr}(\Sigma))\right) \\
&= \alpha_3(|x|) + \alpha_4\left(2\sigma_2(\text{tr}(\Sigma))\right)
\end{aligned}$$

By rearranging, we have $-\alpha_3(|x|) \leq -\alpha_4(V_\mu(x)) + \alpha_4(2\sigma_2(\text{tr}(\Sigma)))$ and therefore,

$$\int_{\mathbb{W}} V_\mu(f(x, \kappa_\mu(x), w))d\mu(w) \leq V_\mu(x) - \alpha_4(V_\mu(x)) + \sigma(\text{tr}(\Sigma))$$

in which $\sigma(s) := \sigma_3(s) + \alpha_4\left(2\sigma_2(s)\right)$ and $\sigma(\cdot) \in \mathcal{K}$. Since $\mathbb{W}$ is bounded, $\text{tr}(\Sigma)$ is bounded. Since $\mathcal{X}$ is also bounded, there exists $b \geq 0$ such that $V_\mu(x) \leq \alpha_2(|x|) + \sigma_2(\text{tr}(\Sigma)) \leq b$ for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and corresponding Σ. From Lemma 2.19, we construct $\alpha_v(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_v(\cdot)$ is convex and $\alpha_v(V_\mu(x)) \leq \alpha_4(V_\mu(x))$ for all $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathbb{W})$. Therefore, we have

$$\int_{\mathbb{W}} V_\mu(f(x, \kappa_\mu(x), w))d\mu(w) \leq V_\mu(x) - \alpha_v(V_\mu(x)) + \sigma(\text{tr}(\Sigma)) \tag{3.12}$$

[2] If this inequality does not hold, we can construct a new $\alpha_2(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_3(s) \leq \alpha_2(s)$ for all $s \in \mathbb{R}_{\geq 0}$ and (3.10) still holds.

For arbitrary $\mu \in \mathcal{M}(\mathbb{W})$ and $x \in \mathcal{X}$, let $x(k) := \phi_\mu(k; x, \mathbf{w}_k)$ for all $k \in \mathbb{I}_{\geq 0}$. We then proceed based on the Proof of Proposition 2.18 to arrive at (2.24), i.e.,

$$\mathbb{E}\left[V_\mu(x(k))\right] \leq \max\{\tilde{\beta}(V_\mu(x), k), \tilde{\gamma}(\text{tr}(\Sigma))\} \tag{3.13}$$

in which $\tilde{\beta}(\cdot) \in \mathcal{KL}$ and $\tilde{\gamma}(\cdot) \in \mathcal{K}$. Using Lemma 2.19 and the fact that $\mathcal{X}$ is bounded, we can construct a convex function $\alpha_{1,v}(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_{1,v}(|x|) \leq \alpha_1(|x|) \leq V_\mu(x)$ for all $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathbb{W})$. Thus, we apply Jensen's inequality to give

$$\alpha_{1,v}(\mathbb{E}\left[|x|\right]) \leq \mathbb{E}\left[\alpha_{1,v}(|x|)\right] \leq \mathbb{E}\left[V_\mu(x)\right]$$

and therefore

$$\begin{aligned}
\mathbb{E}\left[|x(k)|\right] &\leq \max\left\{\alpha_{1,v}^{-1}\left(\tilde{\beta}(V_\mu(x), k)\right), \alpha_{1,v}^{-1}\left(\tilde{\gamma}(\text{tr}(\Sigma))\right)\right\} \\
&\leq \beta_1(V_\mu(x), k) + \gamma_1(\text{tr}(\Sigma))
\end{aligned} \tag{3.14}$$

with $\beta_1(s, k) := \alpha_{1,v}^{-1}\left(\tilde{\beta}(s, k)\right) \in \mathcal{KL}$ and $\gamma_1(s) := \alpha_{1,v}^{-1}\left(\tilde{\gamma}(s)\right) \in \mathcal{K}$. We then use the upper bound for $V_\mu(\cdot)$ to give

$$\begin{aligned}
\mathbb{E}\left[|x(k)|\right] &\leq \beta_1(\alpha_2(|x|) + \sigma_2(\text{tr}(\Sigma)), k) + \gamma_1(\text{tr}(\Sigma)) \\
&\leq \beta_1(2\alpha_2(|x|), k) + \beta_1(2\sigma_2(\text{tr}(\Sigma)), 0) + \gamma_1(\text{tr}(\Sigma)) \\
&= \beta(|x|, k) + \gamma(\text{tr}(\Sigma))
\end{aligned}$$

in which $\beta(s, k) := \beta_1(2\alpha_2(s), k) \in \mathcal{KL}$ and $\gamma(s) := \beta_1(2\sigma_2(s), 0) + \gamma_1(s) \in \mathcal{K}$.

Note that the choice of $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathbb{W})$ was arbitrary and the functions $\beta(\cdot)$ and $\gamma(\cdot)$ are constructed independently of μ. Therefore, these same functions $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ satisfy (3.8) for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$. $\square$

We can also establish the following sufficient condition for ℓ-RASiE.

Proposition 3.13. *If a system $x^+ = f(x, \kappa_\mu(x), w)$, $w \in \mathbb{W}$, with $\mathbb{W}$ bounded, admits an SISS Lyapunov function $V_\mu : \mathcal{X} \to \mathbb{R}_{\geq 0}$ in an RPI and bounded set $\mathcal{X}$ that satisfies $\ell(x, \kappa_\mu(x)) \leq V_\mu(x)$ for all $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathbb{W})$, then the origin is ℓ-RASiE in $\mathcal{X}$.*

Proof. From Proposition 3.12, there exists $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ such that (3.8) holds for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$. Since $\mathcal{X}$ is bounded, we can construct a concave function $\alpha_c(\cdot) \in \mathcal{K}_\infty$ that satisfies $\alpha_2(|x|) \leq \alpha_c(|x|)$ for all $x \in \mathcal{X}$ via Corollary 2.20. Therefore,

$$
\begin{aligned}
\mathbb{E}\left[\ell(x(k), \kappa_\mu(x(k)))\right] &\leq \mathbb{E}\left[V_\mu(x(k))\right] \\
&\leq \mathbb{E}\Big[\alpha_c(|x(k)|) + \sigma_2(\text{tr}(\Sigma))\Big] \\
&\leq \alpha_c(\mathbb{E}[|x(k)|]) + \sigma_2(\text{tr}(\Sigma)) \\
&\leq \alpha_c(2\beta(|x|, k)) + \alpha_c(2\gamma(\text{tr}(\Sigma))) + \sigma_2(\text{tr}(\Sigma))
\end{aligned}
$$

in which $x(k) := \phi_\mu(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$. Define $\tilde{\beta}(s,k) := \alpha_c(2\beta(s,k))$ and $\tilde{\gamma}(s) := \alpha_c(2\gamma(s)) + \sigma_2(s)$ and note that $\tilde{\beta}(\cdot) \in \mathcal{KL}$ and $\tilde{\gamma}(\cdot) \in \mathcal{K}$ to complete the proof. □

3.5 Stochastic robustness of SMPC

With these definitions of stochastic stability extended to address $\kappa_\mu(\cdot)$, we now establish that SMPC is both RASiE and ℓ-RASiE. Analogous to nominal MPC, we establish these results by showing that the optimal cost of the SMPC problem $V_\mu^{s0}(\cdot)$ is an SISS Lyapunov function. We begin by establishing some important properties of SMPC through the following lemmata. First, we establish that the terminal cost and constraint ensure a stochastic cost decrease condition within $\mathbb{X}_f$.

Lemma 3.14. *Let Assumptions 3.1, 3.3 and 3.5 to 3.7 hold. Then there exists $\sigma(\cdot) \in \mathcal{K}$ such that*

$$\int_{\mathbb{W}} V_f(f(x, \kappa_f(x), w))d\mu(w) \leq V_f(x) - \ell(x, \kappa_f(x)) + \sigma(\text{tr}(\Sigma))$$

for all $x \in \mathbb{X}_f$ and $\mu \in \mathcal{M}(\mathbb{W})$.

Proof. Since $V_f(\cdot)$, $f(\cdot)$, and $\kappa_f(\cdot)$ are continuous and $\mathbb{X}_f$ and $\mathbb{W}$ are bounded, we have from Proposition 2.34 that there exists $\alpha(\cdot) \in \mathcal{K}_\infty$ such that

$$|V_f(f(x, \kappa_f(x), w)) - V_f(f(x, \kappa_f(x), w))| \leq \alpha(|w|)$$

for all $x \in \mathbb{X}_f$ and $w \in \mathbb{W}$. We combine this inequality with (3.7) to give

$$V_f(f(x, \kappa_f(x), w)) \leq V_f(x) - \ell(x, \kappa_f(x)) + \alpha(|w|)$$

Then we apply Corollary 2.20 to construct a concave function $\alpha_c(\cdot) \in \mathcal{K}_\infty$ such that $\alpha(|w|) \leq \alpha_c(|w|)$ for all $w \in \mathbb{W}$ since $\mathbb{W}$ is bounded. We evaluate the Lebesgue integral of both sides of the inequality with respect to $\mu \in \mathcal{M}(\mathbb{W})$ and apply Jensen's inequality to give

$$\int_{\mathbb{W}} V_f(f(x, \kappa_f(x), w))d\mu(w) \leq V_f(x) - \ell(x, \kappa_f(x)) + \alpha_c(\mathbb{E}[|w|])$$

We apply Lemma 2.4, define $\sigma(s) := \alpha_c(s^{1/2})$, and note that $\sigma(\cdot) \in \mathcal{K}$ to complete the proof. □

Note that the construction of $\sigma(\cdot)$ is independent of μ since the functions $V_f(\cdot)$, $f(\cdot)$, and $\kappa_f(\cdot)$ are independent of μ. As an important special case, we can show that if the disturbance is additive and the terminal cost quadratic, the $\mathcal{K}$-function takes a specific form more reminiscent of the stochastic LQR problem.

Lemma 3.15. *Let Assumptions 3.1, 3.3 and 3.5 to 3.7 hold with* $f(x,u,w) := g(x,u) + w$ *and* $V_f(x) := x'Px$ *for positive semidefinite* P*. Then we have that*

$$\int_{\mathbb{W}} V_f(f(x,\kappa_f(x),w))d\mu(w) \leq V_f(x) - \ell(x,\kappa_f(x)) + \text{tr}(P\Sigma) \tag{3.15}$$

for all $x \in \mathbb{X}_f$ *and* $\mu \in \mathcal{M}(\mathbb{W})$*.*

Proof. Substitute these specific equations for $f(\cdot)$ and $V_f(\cdot)$ into the following integral to give

$$\int_{\mathbb{W}} V_f(f(x,\kappa_f(x),w))d\mu(w) \\ = \int_{\mathbb{W}} \left(g(x,\kappa_f(x))'Pg(x,\kappa_f(x)) + 2g(x,\kappa_f(x))'Pw + w'Pw\right) d\mu(w)$$

Note that $g(x,\kappa_f(x))$ is constant w.r.t. w and $\mathbb{E}[w] = 0$ to give

$$\int_{\mathbb{W}} V_f(f(x,\kappa_f(x),w))d\mu(w) = g(x,\kappa_f(x))'Pg(x,\kappa_f(x)) + \int_{\mathbb{W}} w'Pwd\mu(w)$$

Then note that $f(x,\kappa_f(x),0) = g(x,\kappa_f(x))$ and

$$\mathbb{E}[w'Pw] = \text{tr}(\mathbb{E}[w'Pw]) = \mathbb{E}\left[\text{tr}(w'Pw)\right] = \mathbb{E}\left[\text{tr}(Pw'w)\right] = \text{tr}(P\mathbb{E}[w'w]) = \text{tr}(P\Sigma)$$

to give

$$\int_{\mathbb{W}} V_f(f(x,\kappa_f(x),w))d\mu(w) = V_f(f(x,\kappa_f(x),0)) + \text{tr}(P\Sigma)$$

for all $x \in \mathbb{X}_f$ and $\mu \in \mathcal{M}(\mathbb{W})$. □

Again, we note that P is fixed for all μ. We can replace $\text{tr}(P\Sigma)$ with a $\mathcal{K}$-function of $\text{tr}(\Sigma)$ in (3.15) by noting that $\text{tr}(P\Sigma) \leq \bar{\lambda}_P\text{tr}(\Sigma)$ in which $\bar{\lambda}_P \geq 0$ is the maximum eigenvalue of the positive semidefinite matrix P. We can now extend the cost decrease in Lemma 3.14 to the optimal cost and the entire feasible set $\mathcal{X}^s$.

Lemma 3.16. *Let Assumptions 3.1, 3.3 and 3.5 to 3.7 hold. Then the set $\mathcal{X}^s$ is RPI for the system $x^+ = f(x, \kappa_\mu^s(x), w)$, $w \in \mathbb{W}$ for any $\mu \in \mathcal{M}(\mathbb{W})$ and there exists $\sigma(\cdot) \in \mathcal{K}$ such that*

$$\int_{\mathbb{W}} V_\mu^{s0}(f(x, \kappa_\mu^s(x), w)) d\mu(w) \leq V_\mu^{s0}(x) - \ell(x, \kappa_\mu^s(x)) + \sigma(\mathbf{tr}(\Sigma))$$

for all $x \in \mathcal{X}^s$ and $\mu \in \mathcal{M}(\mathbb{W})$.

Proof. Choose $x \in \mathcal{X}^s$, $\mu \in \mathcal{M}(\mathbb{W})$, and $\mathbf{v}^0 \in \mathbf{v}_\mu^{s0}(x)$ such that $\kappa_\mu^s(x) = \pi(x, v^0(0))$. For all $\mathbf{w} = (w(0), w(1), \dots, w(N-1)) \in \mathbb{W}^N$, we have that

$$(\hat{\phi}^s(k; x, \mathbf{v}^0, \mathbf{w}), \pi(\hat{\phi}^s(k; x, \mathbf{v}^0, \mathbf{w}), v^0(k))) \in \mathbb{Z}$$

Also, $x(N, \mathbf{w}) := \hat{\phi}^s(N; x, \mathbf{v}^0, \mathbf{w}) \in \mathbb{X}_f$ and therefore

$$f(x(N, \mathbf{w}), \kappa_f(x(N, \mathbf{w})), w(N)) \in \mathbb{X}_f$$

for all $w(N) \in \mathbb{W}$ by Assumption 3.6. Thus, the candidate trajectory

$$\tilde{\mathbf{v}}^+ := (v^0(1), v^0(2), \dots, v^0(N-1), 0)$$

satisfies $\tilde{\mathbf{v}}^+ \in \mathcal{V}(x^+)$ for $x^+ = f(x, \kappa_\mu^s(x), w(0))$ and all $w(0) \in \mathbb{W}$. So $\mathcal{V}(x^+)$ is nonempty and $x^+ \in \mathcal{X}^s$. Since the choice of $x \in \mathcal{X}^s$ and $\mu \in \mathcal{M}(\mathbb{W})$ was arbitrary, $\mathcal{X}^s$ is RPI for any $\mu \in \mathcal{M}(\mathbb{W})$.

We define

$$\tilde{\mathbf{w}}^+ = (w(1), w(2), \dots, w(N-1), w(N))$$

and using the definition of $J(\cdot)$ we obtain

$$J(x^+, \tilde{\mathbf{v}}^+, \tilde{\mathbf{w}}^+) = J(x, \mathbf{v}^0, \mathbf{w}) - \ell(x, \kappa_\mu^s(x)) + \eta(x(N, \mathbf{w}), w(N)) \tag{3.16}$$

in which

$$\eta(x,w) := -V_f(x) + \ell(x,\kappa_f(x)) + V_f(f(x,\kappa_f(x),w))$$

From Lemma 3.14 and the fact that $x(N,\mathbf{w}) \in \mathbb{X}_f$, there exists $\sigma(\cdot) \in \mathcal{K}$ such that

$$\int_{\mathbb{W}^{N+1}} \eta(x(N,\mathbf{w}),w(N))d\mu^N(\mathbf{w})d\mu(w(N)) \le \sigma(\text{tr}(\Sigma))$$

We also have the following equality.

$$\int_{\mathbb{W}^{N+1}} J(x,\mathbf{v}^0,\mathbf{w})d\mu^N(\mathbf{w})d\mu(w(N)) = V_\mu^{s0}(x)$$

And by optimality we have that

$$V_\mu^{s0}(x^+) \le \int_{\mathbb{W}^N} J(x^+,\tilde{\mathbf{v}}^+,\tilde{\mathbf{w}}^+)d\mu(w(1))\dots d\mu(w(N))$$

We combine these inequalities with (3.16) to give

$$\begin{aligned}
&\int_{\mathbb{W}} V_\mu^{s0}(x^+)d\mu(w) \\
&\quad \le \int_{\mathbb{W}^{N+1}} J(x^+,\tilde{\mathbf{v}}^+,\tilde{\mathbf{w}}^+)d\mu^N(\mathbf{w})d\mu(w(N)) \\
&\quad \le \int_{\mathbb{W}^{N+1}} \Big(J(x,\mathbf{v}^0,\mathbf{w}) - \ell(x,\kappa_\mu^s(x)) + \eta(x(N,\mathbf{w}),w(N))\Big)d\mu^N(\mathbf{w})d\mu(w(N)) \\
&\quad \le V_\mu^{s0}(x) - \ell(x,\kappa_\mu^s(x)) + \sigma(\text{tr}(\Sigma))
\end{aligned}$$

Substitute $x^+ = f(x,\kappa_\mu^s(x),w)$ and let $w = w(0)$. Note that since the choice of $x \in \mathcal{X}^s$, $\mu \in \mathcal{M}(\mathbb{W})$ was arbitrary and $\sigma(\cdot)$ is constructed independently of $\mu \in \mathcal{M}(\mathbb{W})$ via Lemma 3.14, this inequality holds for all $x \in \mathcal{X}^s$ and $\mu \in \mathcal{M}(\mathbb{W})$. □

Note that the function $\sigma(\cdot) \in \mathcal{K}$ is the same in Lemma 3.14 and Lemma 3.17. We now establish an upper bound for the optimal cost function.

Lemma 3.17. *Let Assumptions 3.1, 3.3 and 3.5 to 3.7 hold. Then there exist* $\alpha_2(\cdot) \in \mathcal{K}_\infty$ *and* $\sigma_2(\cdot) \in \mathcal{K}$ *such that* $V_\mu^{s0}(x) \leq \alpha_2(|x|) + \sigma_2(\mathrm{tr}(\Sigma))$ *for all* $x \in \mathcal{X}^s$ *and* $\mu \in \mathcal{M}(\mathbb{W})$*.*

Proof. Choose $x \in \mathbb{X}_f$ and $\mu \in \mathcal{M}(\mathbb{W})$ and consider a trajectory generated by repeated application of the terminal control law, i.e., $x(k) := \hat{\phi}(k; x, \mathbf{0}, \mathbf{w})$ since $\pi(x, 0) = \kappa_f(x)$. The set $\mathbb{X}_f$ is RPI for this control law due to Assumption 3.6 and therefore $x(k) \in \mathbb{X}_f$ for all $k \in \mathbb{I}_{0:N}$. From Assumption 3.6 and Lemma 3.14, we have that

$$\int_{\mathbb{W}^N} \left(V_f(x(k+1)) - V_f(x(k))\right) d\mu^N(\mathbf{w}) \leq -\int_{\mathbb{W}^N} \ell(x(k), \kappa_f(x(k))) d\mu^N(\mathbf{w}) + \sigma(\mathrm{tr}(\Sigma))$$

We sum both sides of the inequality from $k = 0$ to $k = N - 1$ to give

$$\int_{\mathbb{W}^N} \left(V_f(x(N)) - V_f(x(0))\right) d\mu^N(\mathbf{w}) \leq -\int_{\mathbb{W}^N} \sum_{k=0}^{N-1} \ell(x(k), \kappa_f(x(k))) d\mu^N(\mathbf{w}) + N\sigma(\mathrm{tr}(\Sigma))$$

By rearranging and substituting in the definition of $J(\cdot)$ and $x = x(0)$, we have

$$\int_{\mathbb{W}^N} J(x, \mathbf{0}, \mathbf{w}) d\mu^N(\mathbf{w}) \leq V_f(x) + N\sigma(\mathrm{tr}(\Sigma))$$

By optimality, we have that

$$V_\mu^{s0}(x) \leq V_f(x) + N\sigma(\mathrm{tr}(\Sigma))$$

Since the choice of $x \in \mathbb{X}_f$ and $\mu \in \mathcal{M}(\mathbb{W})$ was arbitrary, this inequality must hold for all $x \in \mathbb{X}_f$ and $\mu \in \mathcal{M}(\mathbb{W})$. Furthermore, $V_f(\cdot)$ and $\sigma(\cdot)$ are constructed independently of μ.

We now define

$$H(x) := \max \left\{ \sup_{\mu \in \mathcal{M}(\mathbb{W})} \left(V_\mu^{s0}(x) - N\sigma(\text{tr}(\Sigma)) \right), 0 \right\}$$

and note that $0 \leq H(x) \leq V_f(x)$ for all $x \in \mathbb{X}_f$. Since $V_f(\cdot)$ is continuous, $H(0) = V_f(0) = 0$, and $\mathbb{X}_f$ contains the origin in its interior, we know that $H(x)$ is continuous at the origin. We also establish that $H(x)$ is locally bounded. Let X be a compact subset of $\mathcal{X}^s$. The function $J : \mathbb{R}^n \times \mathbb{V}^N \times \mathbb{W}^N \rightarrow \mathbb{R}_{\geq 0}$ is a composition of a finite number of continuous functions and is therefore continuous. Thus, $J(\cdot)$ has an upper bound on the compact set $X \times \mathbb{V}^N \times \mathbb{W}^N$. Since $\mathcal{V}(x) \subseteq \mathbb{V}^N$ for all $x \in \mathcal{X}^s$, $V_\mu^{s0} : \mathcal{X}^s \rightarrow \mathbb{R}_{\geq 0}$ must satisfy the same upper bound for all $\mu \in \mathcal{M}(\mathbb{W})$. Thus, $H(x)$ must satisfy this same upper bound because $H(x) \leq \sup_{\mu \in \mathcal{M}(\mathbb{W})} V_\mu^{s0}(x)$. Since $0 \leq H(x)$ as well and the choice of X was arbitrary, $H(x)$ is locally bounded on $\mathcal{X}^s$.

Since $H(x)$ is locally bounded, satisfies $H(0) = 0$, and is continuous at $x = 0$, we can apply Proposition 2.33 to construct $\alpha_2(\cdot) \in \mathcal{K}_\infty$ such that $H(x) \leq \alpha_2(|x|)$ for all $x \in \mathcal{X}^s$. Furthermore, we have that

$$V_\mu^{s0}(x) - N\sigma(\text{tr}(\Sigma)) \leq H(x) \leq \alpha_2(|x|)$$

for all $x \in \mathcal{X}^s$ and $\mu \in \mathcal{M}(\mathbb{W})$. □

We observe that the function $\sigma_2(\cdot)$ constructed in Lemma 3.17 increases with increasing horizon length N. As an alternative to Lemma 3.17, we may instead assume that $V_\mu^{s0}(\cdot)$ is continuous at the origin. With this alternate assumption, we can find $\alpha_2(\cdot) \in \mathcal{K}_\infty$ such that

$$V_\mu^{s0}(x) \leq V_\mu^{s0}(0) + \alpha_2(|x|) \quad \forall x \in \mathcal{X}^s$$

Note, however, that $V_\mu^{s0}(0)$ is typically not zero for SMPC if the stage cost satisfies Assumption 3.4. Only is specific situations, e.g., multiplicative disturbance models in which the effect of the disturbance vanishes at the origin, is $V_\mu^{s0}(0) = 0$. Furthermore, we also expect the value of $V_\mu^{s0}(0)$ to increase with increasing N similar to the bound derived in Lemma 3.17. We propose, therefore, that this increase with horizon length is not a weakness of the analysis approach, but an underlying characteristic of SMPC, particularly for nonlinear systems.

With these results, we can now establish the following theorem for SMPC.

Theorem 3.18 (SMPC)**.** *Let Assumptions 3.1 and 3.3 to 3.7 hold with* $\mathbb{W}$ *and* $\mu \in \mathcal{M}(\mathbb{W})$ *known exactly. For the system* $x^+ = f(x, \kappa_\mu^s, w)$*,* $w \in \mathbb{W}$*, we have that:*

(i) The set $\mathcal{X}^s$ *is RPI.*

(ii) The origin is RASiE in $\mathcal{X}^s$*.*

(iii) The origin is ℓ*-RASiE in the set* $\mathcal{X}^s$*.*

Proof. From Assumption 3.4 and the definition of $V_\mu^{s0}(\cdot)$, we have that

$$\alpha_\ell(|x|) \leq \ell(x, \kappa_\mu^s(x)) \leq V_\mu^{s0}(x)$$

for all $x \in \mathcal{X}^s$ and $\mu \in \mathcal{M}(\mathbb{W})$. Note that $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$ is independent of μ. We then use Lemma 3.16 and Assumption 3.4 to give

$$\int_{\mathbb{W}} V_\mu^{s0}(f(x, \kappa_\mu^s(x), w))d\mu(w) \leq V_\mu^{s0}(x) - \alpha_\ell(|x|) + \sigma(\text{tr}(\Sigma))$$

for all $x \in \mathcal{X}^s$ and $\mu \in \mathcal{M}(\mathbb{W})$. Lastly, we use Lemma 3.17 to obtain the upper bound for $V_\mu^{s0}(\cdot)$. Thus, $V_\mu^{s0}(\cdot)$ is an SISS Lyapunov function with $\alpha_1(\cdot) = \alpha_3(\cdot) = \alpha_\ell(\cdot) \in \mathcal{K}_\infty$ from Assumption 3.4, $\sigma_3(\cdot) = \sigma(\cdot) \in \mathcal{K}$ from Lemma 3.16, and $\alpha_2(\cdot) \in \mathcal{K}_\infty$ and $\sigma_2(\cdot) \in \mathcal{K}$ from Lemma 3.17. We apply Proposition 3.12 and Proposition 3.13 to complete the proof. □

Note that, unlike nominal MPC, the origin is *not* guaranteed to be RAS for SMPC. For linear systems, quadratic costs, and specifically chosen terminal costs, (Goulart and Kerrigan, 2008) establish that SMPC is RAS, but this results relies on properties, such as convexity of the optimal cost, that do not extend to the nonlinear SMPC problem. We discuss the practical implications of this fact in Section 3.7.

3.6 Constraint-tightened MPC

A key strength of SMPC is that the disturbances considered in the problem formulation provide a natural means to tighten the state and input constraints to ensure robust constraint satisfaction for the closed-loop trajectory, i.e.,

$$(\phi^s_\mu(k; x, \mathbf{w}_k), \kappa^s_\mu(\phi^s_\mu(k; x, \mathbf{w}_k))) \in \mathbb{Z}$$

for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, $\mathbf{w}_k \in \mathbb{W}^k$, and $k \in \mathbb{I}_{\geq 0}$. In certain control problems, SMPC is used primarily for this purpose, and the stochastic objective function is not essential to the design goal.

Tube-based MPC is particularly suited for these problems as it provides middle ground between nominal and stochastic MPC. By using the stochastic MPC framework to (conservatively) tighten constraints offline, tube-based MPC ensures robust constraint satisfaction while retaining the computational and theoretical convenience afforded by a nominal objective function. These tube-based methods were first proposed for linear systems subject to worst-case disturbances (Chisci et al., 2001; Mayne et al., 2005) and then extended to consider nonlinear systems (Limón Marruedo et al., 2002; Cannon et al., 2011; Mayne et al., 2011). Stochastic descriptions of these disturbances can also be used to construct tubes that satisfy probabilistic (chance) constraints for the system (Cannon et al., 2010). The notion of incremental stabilizability can also be used to tighten constraints without requiring complicated offline computations (Köhler et al., 2020; Santos et al., 2019).

Since this dissertation focuses on the closed-loop properties of these MPC algorithms, we consider a somewhat different problem than the typical tube-based MPC formulation. Specifically, we propose using the same control parameterization, disturbance set, and therefore set of admissible control parameterizations ($\mathcal{V}(x)$) as SMPC, but consider an objective function evaluated for only the nominal trajectory. We call this formulation constraint-tightened MPC (CMPC). This formulation, unlike tube-based MPC, does not lend itself to offline computation of the set $\mathcal{V}(\cdot)$ and therefore does offer the same computational efficiencies as tube-based MPC. Instead, this formulation serves as an idealized version of tube-based MPC. That is, we tighten the constraints no more than necessary to ensure robust constraint satisfaction. Conversely, tube-based MPC formulations can be viewed as methods to conservatively approximate $\mathcal{V}(\cdot)$ offline.

We define the CMPC optimization problem as

$$\mathbb{P}^c(x): \ V^{c0}(x) := \min_{\mathbf{v}\in\mathcal{V}(x)} J(x, \mathbf{v}, \mathbf{0}) \tag{3.17}$$

for any $x \in \mathcal{X}^s$ and the optimal solutions are denoted $\mathbf{v}^{c0}(x) := \arg\min_{\mathbf{v}\in\mathcal{V}(x)} J(x, \mathbf{v}, \mathbf{0})$. Thus, we are using the disturbance set $\mathbb{W}$ to construct $\mathcal{V}(x)$, but we do not use the probability measure μ in the optimization problem. Note that $V^{c0}(\cdot)$ and $\mathbf{v}^{c0}(\cdot)$ do not vary with μ. We use a state feedback parameterization ($\pi(x(k), v(k))$), but disturbance feedback parameterizations are also used in tube-based MPC ($\tilde{\pi}(w(k-1), v(k))$). These two parameterizations produce an equivalent control law for linear systems (Goulart et al., 2006).

We use a Borel measurable selection rule to define the single-valued control law $\kappa^c : \mathcal{X}^s \rightarrow \mathbb{U}$ such that $\kappa^c(x) \in \{\pi(x, v(0)) : \mathbf{v} \in \mathbf{v}^{c0}(x)\}$ for all $x \in \mathcal{X}^s$. The resulting closed loop system is then

$$x^+ = f(x, \kappa^c(x), w) \tag{3.18}$$

We use $\phi^c(k; x, \mathbf{w}_k)$ to denote state of the system (3.18) at time $k \in \mathbb{I}_{\geq 0}$, given the initial condition $x \in \mathcal{X}^s$ and disturbance sequence $\mathbf{w}_k \in \mathbb{W}^k$. Note that the control law and therefore closed-loop trajectory do not depend on the probability measure μ. Since $\mathbb{P}^c(x) = \mathbb{P}^s_\mu(x)$ for all $x \in \mathcal{X}^s$ and $\mu(\{0\}) = 1$, Proposition 3.8 applies to CMPC, i.e., $\mathbb{P}^c(x)$ has a solution for all $x \in \mathcal{X}^s$, $V^{c0} : \mathcal{X}^s \to \mathbb{R}_{\geq 0}$ is lower semicontinuous, and $\mathbf{v}^{c0} : \mathcal{X}^s \rightrightarrows \mathbb{V}^N$ is a Borel measurable function. Furthermore, there exists a Borel measurable selection rule such that $\kappa^c(\cdot)$ is a Borel measurable function. Thus, all relevant mathematical objects are well defined and Borel measurable.

By using the disturbance set $\mathbb{W}$ to construct the constraints, like SMPC, we ensure that $\mathcal{X}^s$ is RPI for the closed-loop system. By using a nominal objective function, like nominal MPC, we ensure that the system satisfies all of the deterministic and stochastic definitions of robustness considered in Chapter 2. Thus, we have the following theorem for CMPC.

Theorem 3.19 (CMPC). *Let Assumptions 3.1 and 3.3 to 3.7 hold with $\mathbb{W}$ known exactly. For the system $x^+ = f(x, \kappa^s_\mu, w)$, $w \in \mathbb{W}$, we have that:*

1. *The set $\mathcal{X}^s$ is RPI.*
2. *The origin is RAS in the set $\mathcal{X}^s$.*
3. *The origin is RASiE in the set $\mathcal{X}^s$.*
4. *The origin is ℓ-RAS in the set $\mathcal{X}^s$.*
5. *The origin is ℓ-RASiE in the set $\mathcal{X}^s$.*

Proof. We use the same approach as the proof of Lemma 3.16 to establish that $\mathcal{X}^s$ is RPI. Choose $x \in \mathcal{X}^s$ and $\mathbf{v}^0 \in \mathbf{v}^{c0}(x)$ such that $\kappa^c(x) = \pi(x, v(0))$. For all $\mathbf{w} = (w(0), w(1), \dots, w(N-$

$1)) \in \mathbb{W}^N$, we have that

$$(\hat{\phi}^s(k; x, \mathbf{v}^0, \mathbf{w}), \pi(\hat{\phi}^s(k; x, \mathbf{v}^0, \mathbf{w}), v^0(k))) \in \mathbb{Z}$$

Also, $x(N, \mathbf{w}) := \hat{\phi}^s(N; x, \mathbf{v}^0, \mathbf{w}) \in \mathbb{X}_f$ and therefore

$$f(x(N, \mathbf{w}), \kappa_f(x(N, \mathbf{w})), w(N)) \in \mathbb{X}_f$$

for all $w(N) \in \mathbb{W}$ by Assumption 3.6. Thus, the candidate trajectory

$$\tilde{\mathbf{v}}^+ := (v^0(1), v^0(2), \ldots, v^0(N-1), 0)$$

satisfies $\tilde{\mathbf{v}}^+ \in \mathcal{V}(x^+)$ for $x^+ = f(x, \kappa_\mu^s(x), w(0))$ and all $w(0) \in \mathbb{W}$. Thus, $\mathcal{V}(x^+)$ is nonempty and $x^+ \in \mathcal{X}^s$. Since the choice of $x \in \mathcal{X}^s$ was arbitrary, $\mathcal{X}^s$ is RPI and (i) holds.

We now establish that $V^{c0}(\cdot)$ is an ISS Lyapunov function. From (3.16) we have that

$$J(x^+, \tilde{\mathbf{v}}^+, \mathbf{0}) = J(x, \mathbf{v}^0, \mathbf{0}) - \ell(x, \kappa_\mu^c(x)) + \eta(x(N, \mathbf{0}), 0)) \tag{3.19}$$

in which

$$\eta(x, w) := -V_f(x) + \ell(x, \kappa_f(x)) + V_f(f(x, \kappa_f(x), w))$$

From Assumptions 3.4 and 3.6, we then have

$$J(f(x, \kappa^c(x), 0), \tilde{\mathbf{v}}^+, \mathbf{0}) \leq J(x, \mathbf{v}^0, \mathbf{0}) - \alpha_\ell(|x|)$$

for all $x \in \mathcal{X}^s$ in which $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$.

We have that $J(\cdot)$ is continuous and $\mathcal{X}^s$, $\mathbb{U}$, $\mathbb{V}$ are compact. By Proposition 2.34, there exists $\sigma(\cdot) \in \mathcal{K}$ such that

$$|J(f(x,u,w),\mathbf{v},\mathbf{0}) - J(f(x,u,0),\mathbf{v},\mathbf{0})| \leq \sigma(|w|)$$

for all $(x,u,w) \in \mathcal{X}^s \times \mathbb{U} \times \mathbb{W}$ and $\mathbf{v} \in \mathbb{V}^N$. Thus, we have for all $x \in \mathcal{X}^s$ and $w \in \mathbb{W}$ that

$$\begin{aligned} V^{c0}(x^+) &\leq J(f(x,\kappa^c(x),w),\tilde{\mathbf{v}}^+,\mathbf{0}) \\ &\leq J(f(x,\kappa^c(x),0),\tilde{\mathbf{v}}^+,\mathbf{0}) + \sigma(|w|) \\ &\leq J(x,\mathbf{v}^0,\mathbf{0}) - \alpha_\ell(|x|) + \sigma(|w|) \end{aligned}$$

From Assumption 3.4, we also have that $\alpha_\ell(|x|) \leq \ell(x,\kappa^c(x)) \leq V^{c0}(x)$ for all $x \in \mathcal{X}^s$.

For the upper bound, we note that $V^{c0}(x) = V^{s0}_\mu(x)$ for $\mu(\{0\}) = 1$ and $\Sigma = 0$. We therefore can use Lemma 3.17 with $\Sigma = 0$ to construct $\alpha_2(\cdot) \in \mathcal{K}_\infty$ such that

$$V^{c0}(x) = V^{s0}_\mu(x) \leq \alpha_2(|x|)$$

for all $x \in \mathcal{X}^s$.

Thus, $V^{c0}(\cdot)$ is an ISS Lyapunov function on the RPI set $\mathcal{X}^s$ with $\ell(x,\kappa^c(x)) \leq V^{c0}(x)$ for all $x \in \mathcal{X}^s$. We use Proposition 2.15 and Proposition 2.26 to establish (ii) and (iv). We also use Proposition 2.21 to establish that $V^{c0}(\cdot)$ is also an SISS Lyapunov function. We then apply Proposition 2.18 and Proposition 2.27 to give (iii) and (v). □

3.7 Examples and comparisons

With Theorems 2.32, 3.18 and 3.19 in hand, we make the following observation: nominal MPC, SMPC, and CMPC satisfy the same definitions of stochastic robustness. Thus, including stochastic information in the formulation of SMPC has not afforded SMPC some unique property unobtainable by either nominal MPC or CMPC. This observation, however, does not mean that the quantitative performance of these methods is the same. In certain cases, nominal MPC and SMPC may produce significantly different closed-loop trajectories. Hence, the next question to answer is: Which method is *more* robust? Based on the definitions of stochastic robustness presented in this chapter, we consider three specific conjectures that characterize the notion that SMPC is more robust than nominal MPC. While we restrict these conjectures to comparing nominal MPC and SMPC to streamline their presentation, we also perform simulations of CMPC and include this method in the subsequent discussion. Several of these examples are adapted from McAllister and Rawlings (2022d).

Conjecture 3.20. *Let $f(\cdot)$, $\ell(\cdot)$, $V_f(\cdot)$, $\mathbb{U}$, $\mathbb{X}_f$, μ, $\mathbb{W}$ be the same for nominal MPC and SMPC. Then, the feasible set for SMPC, $\mathcal{X}^s$, is larger than the RPI set for nominal MPC for the same disturbance set $\mathbb{W}$, i.e., $\mathcal{S} \subseteq \mathcal{X}^s$.*

Conjecture 3.21. *Let $f(\cdot)$, $\ell(\cdot)$, $V_f(\cdot)$, $\mathbb{U}$, $\mathbb{X}_f$, μ, $\mathbb{W}$ be the same for nominal MPC and SMPC. For any $x \in \mathcal{X}^s$,*

$$\lim_{k\to\infty} \mathbb{E}\left[|\phi^s_\mu(k; x, \mathbf{w}_k)|\right] \leq \lim_{k\to\infty} \mathbb{E}\left[|\phi(k; x, \mathbf{w}_k)|\right]$$

if these limits exist, i.e., SMPC is better than nominal MPC in terms of the expected norm of the closed-loop state (RASiE).

Conjecture 3.22. *Let $f(\cdot)$, $\ell(\cdot)$, $V_f(\cdot)$, $\mathbb{U}$, $\mathbb{X}_f$, μ, $\mathbb{W}$ be the same for nominal MPC and SMPC. For any $x \in \mathcal{X}^s$,*

$$\lim_{k\to\infty} \mathbb{E}\left[\ell(x^s(k), \kappa^s_\mu(x(k)))\right] \leq \lim_{k\to\infty} \mathbb{E}\left[\ell(x(k), \kappa(x(k)))\right]$$

if these limits exist in which $x^s(k) := \phi^s_\mu(k; x, \mathbf{w}_k)$ and $x := \phi(k; x, \mathbf{w}_k)$, i.e., SMPC is better than nominal MPC in terms of the expected value of the closed-loop stage cost (ℓ-RASiE).

In the following subsections, we use a few simple examples to investigate these conjectures and compare the strengths, weakness, and closed-loop behavior of nominal MPC, SMPC, and CMPC.

3.7.1 RPI sets

Conjecture 3.20 frames the discussion of robustness based on the respective RPI sets for nominal MPC and SMPC. A larger RPI set means that the controller is robustly recursively feasible for a larger set of initial states and is therefore "more robust". Recall that CMPC and SMPC share the same RPI set $\mathcal{X}^s$. Thus, any conclusion for $\mathcal{X}^s$ also holds for CMPC. We begin with a comparison of three important sets for each of these problems.

Lemma 3.23. *Let Assumptions 3.1 and 3.3 to 3.7 hold. Let $f(\cdot)$, $\mathbb{U}$, $\mathbb{X}_f$, $\mathbb{W}$ be the same for nominal MPC and SMPC. Then $\mathbb{X}_f \subseteq \mathcal{X}^s \subseteq \mathcal{X}$.*

Proof. For any $x \in \mathbb{X}_f$, we have that $\mathbf{0} \in \mathcal{V}(x)$ because Assumption 3.6 ensures that $\mathbb{X}_f$ is RPI for the system $x^+ = f(x, \kappa_f(x), w)$ and $(x, \kappa_f(x)) \in \mathbb{Z}$. So for any $x \in \mathbb{X}_f$, $\mathcal{V}(x) \neq \emptyset$ and therefore $x \in \mathcal{X}^s$ as well. Thus, $\mathbb{X}_f \subseteq \mathcal{X}^s$.

For any $x \in \mathcal{X}^s$ and $\mathbf{v} \in \mathcal{V}(x)$ we know that

$$\pi(\hat{\phi}^s(k; x, \mathbf{v}, \mathbf{0}), v(k)) \in \mathbb{U}$$

for all $k \in \mathbb{I}_{0:N-1}$ and $\hat{\phi}^s(N; x, \mathbf{v}, \mathbf{0}) \in \mathbb{X}_f$ because $0 \in \mathbb{W}$ by Assumption 3.1. Thus, we can define $\mathbf{u} = (u(0), \dots, u(N-1))$ such that

$$u(k) := \pi(\hat{\phi}^s(k; x, \mathbf{v}, \mathbf{0}), v(k))$$

and we have that $\hat{\phi}(k; x, \mathbf{u}) = \hat{\phi}^s(k; x, \mathbf{v}, \mathbf{0})$. Therefore $\mathbf{u} \in \mathbb{U}^N$, $\hat{\phi}(N; x, \mathbf{u}) \in \mathbb{X}_f$, and $\mathbf{u} \in \mathcal{U}(x)$. So for any $x \in \mathcal{X}^s, \mathcal{U}(x) \neq \emptyset$ and therefore $x \in \mathcal{X}$ as well. Thus, $\mathcal{X}^s \subseteq \mathcal{X}$. □

Conjecture 3.20, however, compares the set $\mathcal{X}^s$ to the RPI set for nominal MPC from Theorem 2.32, i.e., the set $\mathcal{S}$. By definition, $\mathcal{S} \subseteq \mathcal{X}$, but establishing the relative sizes of $\mathcal{S}$ and either $\mathbb{X}_f$ or $\mathcal{X}^s$ for a nonlinear system is difficult.

Instead, we demonstrate a counter example to Conjecture 3.20. Consider the scalar system

$$x^+ = x + u + w$$

with the input constraint $\mathbb{U} = [-2, 2]$ and the disturbance set $\mathbb{W} = [-1, 1]$. Choose the stage cost $\ell(x, u) = x^2 + u^2$, terminal cost $V_f(x) = 2x^2$, terminal constraint $\mathbb{X}_f = [-2, 2]$, and control law parameterization $\pi(x, v) = -x + v$. We have that $\mathcal{X} = \{x : |x| \leq 2 + 2N\}$ and $\mathcal{X}^s = \{x : |x| \leq 2 + N\}$ since SMPC must address the potential for a disturbance of $|w| = 1$ at each time step in the optimization problem while still satisfying the terminal constraint. Thus, we have that $\mathbb{X}_f \subset \mathcal{X}^s \subset \mathcal{X}$ for all $N \geq 1$, in which these are strict subsets. For the disturbance set of interest, however, the entire feasible set $\mathcal{X}$ is RPI for the nominal MPC controller. Thus, we have

$$\mathcal{X}^s \subset \mathcal{X} = \mathcal{S}$$

in which $\mathcal{X}^s$ is a strict subset of $\mathcal{S}$ and Conjecture 3.20 does not hold. Moreover, the control laws for nominal MPC, CMPC, and SMPC are identical for this problem for all $x \in \mathcal{X}^s$. The only difference between these three MPC formulations for this problem is that nominal MPC has a larger feasible set and a larger RPI set.

In Figure 3.1, we provide an illustration of these sets for a general nonlinear system informed by Lemma 3.23 and this counter example to Conjecture 3.20. Note that $\mathcal{S}$ is drawn such that it is neither a subset or superset of $\mathcal{X}^s$, since either result is possible for a general nonlinear system. Next, we consider a simple liquid level control problem that serves as a counter example to Conjecture 3.21.

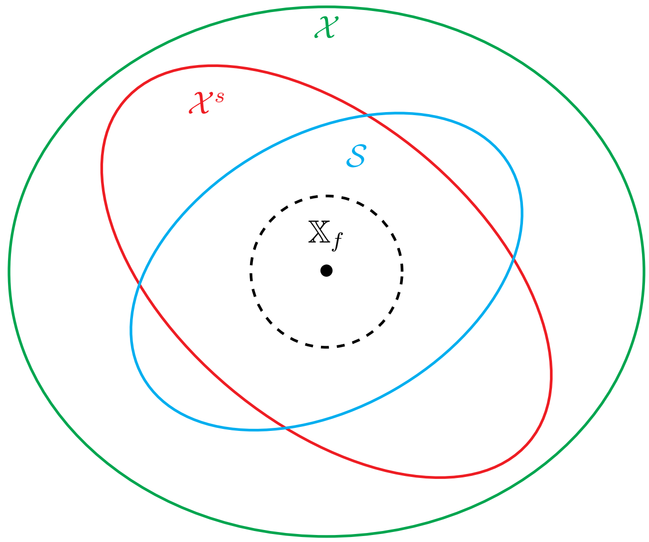

Figure 3.1: Illustration of the relevant sets for nominal MPC and SMPC. Note that $\mathcal{S}$ is not a subset or superset of $\mathcal{X}^s$.

3.7.2 Liquid level control

We now consider a simple liquid level control example with the two tanks shown in Figure 3.2. The goal is to control the height of liquid in each tank, denoted h_1 and h_2 for tanks 1 and 2, respectively. Both tanks have an area of 1. We can adjust the inlet flow rate into tank 1 $F_1 \in [0, 2]$, and the effluent flow rate from tank 2 $F_2 \in [0, 2]$. Tank 1 drains into tank 2 by gravity at a rate proportional to the height of tank 1. We assume, however, that this proportionality constant is subject to uncertainty, i.e., the flow rate from tank 1 to to tank 2 is given by $(1 + w)h_1$ in which w may take values in the finite set $\mathbb{W} := \{-0.3, 0, 0.3\}$. The variable w is distributed according to the probability measure $\mu(\{0.3\}) = \mu(\{-0.3\}) = 0.35$ and $\mu(\{0\}) = 0.3$. The target steady state is $h_1^s = h_2^s = F_1^s = F_2^s = 1$ and we define the state and input in terms of deviation variables: $x = [h_1 - h_1^s, h_2 - h_2^s]'$ and $u = [F_1 - F_1^s, F_1 - F_1^s]'$. With appropriate mass balances on each tank, we have the following differential equations

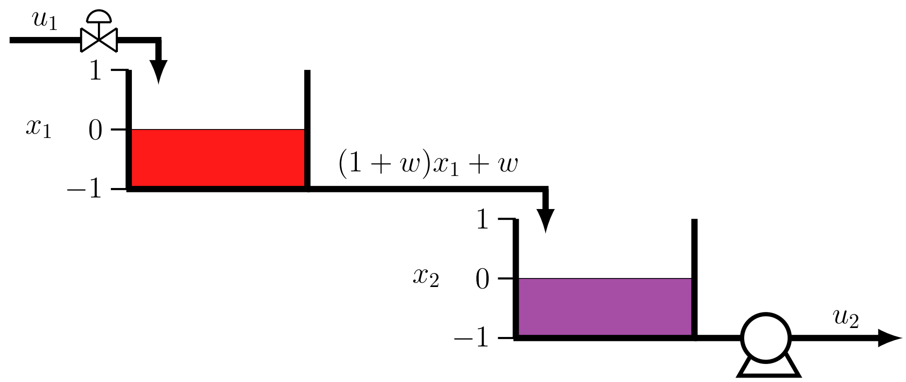

Figure 3.2: Illustration of the liquid level control example. There are two tanks with gravity driven flow from tank 1 to tank 2.

for the system in deviation variables.

$$\frac{dx_1}{dt} = -(1 + w)x_1 + u_1 + w$$
$$\frac{dx_2}{dt} = (1 + w)x_1 - u_2 + w$$

The nominal system ($w = 0$) is linear, but the effect of the disturbance results in a nonlinear differential equation. Although the uncertainty is parametric, i.e., the value of w represents uncertainty in the proportionality constant, and therefore multiplicative in the original system, the system in terms of deviation variables includes both multiplicative and additive effects from this disturbance. Thus, the effect of this parametric uncertainty does not vanish for $(x, u) = 0$. Since the support is finite and the nominal system is linear, we can in fact discretize this differential equation (assuming a zero-order hold on the inputs and disturbance) exactly for all $w \in \mathbb{W}$. Furthermore, we can evaluate the expected value of the cost function in the SMPC optimization problem by enumerating all possible disturbance trajectories. We similarly evaluate the expected value of the closed-loop state and stage cost by simulating all possible disturbance trajectories.

The constraints on F_1, F_2 produce the input constraints $u_1, u_2 \in [-1, 1]$. We define the stage cost $\ell(x,u) = x'Qx + u'Ru$ with

$$Q := \begin{bmatrix} 0.1 & 0 \\ 0 & 20 \end{bmatrix} \qquad R := \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$$

These penalties are chosen to strongly discourage (penalize) any deviations in the height of the second tank. Nonetheless, this stage cost satisfies Assumption 3.4.

We use the LQR cost P and gain K for the nominal system ($w = 0$) to define the terminal cost $V_f(x) = x'Px$ and control law parameterization $\pi(x,v) := Kx + v$. We define the terminal constraint as $\mathbb{X}_f := \{x : |x_1| \leq 0.4, \ |x_2| \leq 0.4\}$ and verify that this terminal constraint satisfies Assumption 3.6 with the terminal control law $\kappa_f(x) := Kx$. Although we have not chosen $\mathbb{X}_f$ as a level set of $V_f(\cdot)$, the nominal system is linear, the constraints are convex, and therefore nominal MPC is nonetheless robustly asymptotically stable (Grimm et al., 2004, Cor. 13). The control law parameterization for SMPC and CMPC is $\pi(x,v) = Kx + v$ with $\mathbb{V}$ chosen such that for any $(x,u) \in \mathcal{X}^s \times \mathbb{U}$ there exists $v \in \mathbb{V}$ that satisfies $\pi(x,v) = u$.

In Figure 3.3, we plot the closed-loop trajectories for each realization of the disturbance and the expected values of these trajectories for MPC and SMPC, respectively, with a horizon of $N = 3$. Since deviations in x_2 are assigned a large cost, the SMPC controller decides to decrease the height of the first tank to minimize the effect of the disturbance on x_2. While there are clear benefits to this approach in terms of the expected stage cost of the system, the behavior is nonintuitive in terms of a typical tracking control problem. Indeed, SMPC drives the state *away* from the origin and outside of the terminal set as well ($|x_1| > 0.4$). The closed-loop trajectory for CMPC ($N = 3$) is identical to nominal MPC ($N = 3$) and therefore omitted.

In Figure 3.4, we plot the expected value of the norm of the state and stage cost for the

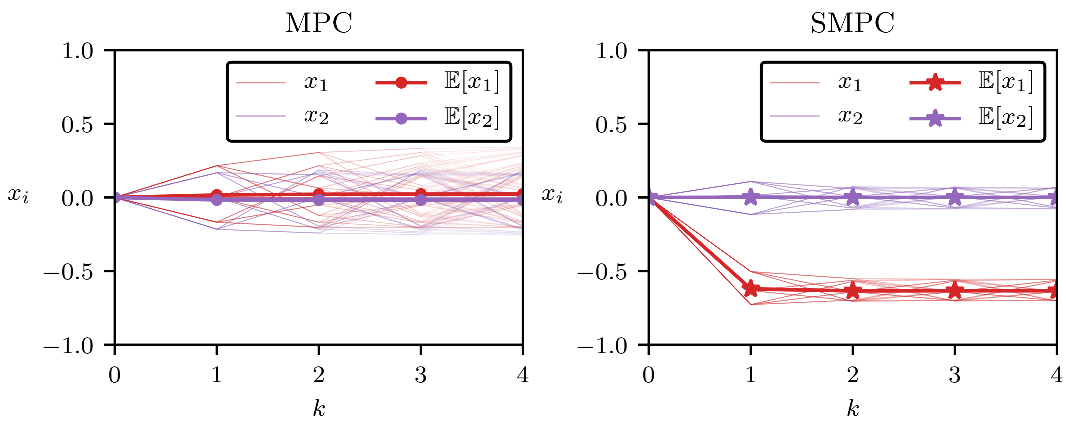

Figure 3.3: Closed-loop trajectories for MPC (left) and SMPC (right) for the liquid level control problem with $N = 3$.

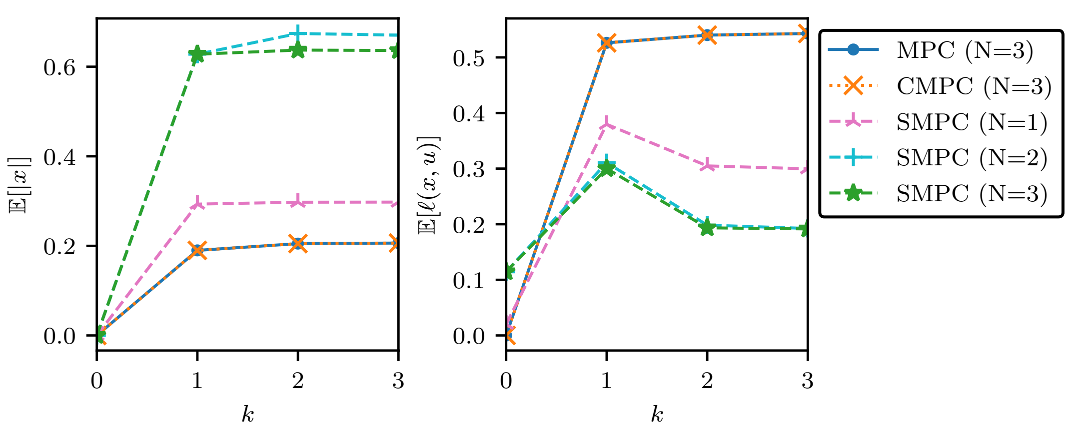

Figure 3.4: The expected value of the norm of the state and stage cost for the closed-loop trajectory of different controllers in the liquid level control problem.

closed-loop trajectory of each controller. As we may anticipate, SMPC achieves a lower expected stage cost as $k \to \infty$ than nominal MPC or CMPC. The value of $\mathbb{E}[|x(k)|]$, however, is larger for SMPC than nominal MPC for an otherwise equivalent problem. By the end of the simulation at $k = 3$, the value of $\mathbb{E}[|x(k)|]$ appear to be constant and we presume that the limit of $\mathbb{E}[|x(k)|]$ is approximately equal to the value at $k = 3$. Thus, Conjecture 3.21 does not hold.

We also observe in Figure 3.4 that the values of $\mathbb{E}[|x(k)|]$ are larger for SMPC with $N = 3$ than for SMPC with $N = 1$. Thus, the dependence of $\sigma_2(\cdot)$ in Lemma 3.17 on the horizon

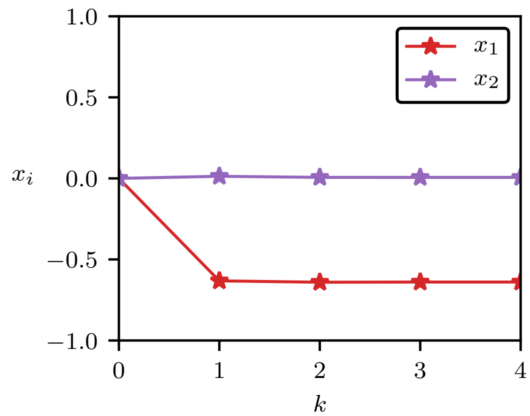

Figure 3.5: The closed-loop trajectory for SMPC ($N = 3$) subject to $\mathbf{w}_k = \mathbf{0}$ for the liquid level control problem.

length N appears to indicate an underlying characteristic of nonlinear SMPC and is not necessarily a shortcoming of the analysis approach used in this work.

As noted in the discussion after Theorem 3.18, one significant distinction between nominal MPC and SMPC (for nonlinear systems) is that SMPC does not guarantee that the origin is RAS for the closed-loop system. We demonstrate the implications of this fact by considering a nominal realization of the disturbance, i.e., $\mathbf{w}_k = \mathbf{0}$. We plot the nominal closed-loop trajectory for SMPC in Figure 3.5. Despite that fact that no disturbance occurs, SMPC drives the state of the system away from the origin and is therefore not RAS or nominally asymptotically stable. By contrast, nominal MPC and CMPC keep the state of the system at the origin for a nominal realization of the disturbance.

3.7.3 State constraints

An important capability of SMPC and CMPC is the systematic constraint tightening procedure inherent to these problem formulations. By explicitly considering the disturbance realizations in the constraints of the optimization problem, we can guarantee robust constraint satisfaction of the closed-loop system for all $w \in \mathbb{W}$. We now consider a two state linear system to illustrate the benefits of this systematic constraint tightening procedure. The system is described by

$$x^+ = Ax + Bu + Gw$$
$$A = \begin{bmatrix} 1 & 0.1 \\ -0.1 & 0.95 \end{bmatrix} B = \begin{bmatrix} 5 \\ 0.1 \end{bmatrix} G = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

with $u \in [-1, 1]$. We assume a finite support for w with $\mu(\{0.05\}) = \mu(\{-0.05\}) = 0.35$ and $\mu(\{0\}) = 0.3$. We also consider the state constraint $|x_1| \leq 1$.

We use a quadratic stage cost $\ell(x, u) := x'Qx + u'Ru$ with

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix} \qquad R = 0.1$$

We use the LQR cost P and gain K for the unconstrained nominal system to define the terminal cost $V_f(x) := x'Px$, terminal constraint $\mathbb{X}_f := \{x : x'Px \leq 1\}$, and terminal control law $\kappa_f(x) := Kx$. The control law parameterization for SMPC and CMPC is $\pi(x, v) = Kx + v$ with $\mathbb{V}$ chosen such that for any $(x, u) \in \mathcal{X}^s \times \mathbb{U}$ there exists $v \in \mathbb{V}$ that satisfies $\pi(x, v) = u$. We verify that this design satisfies all the required assumptions for nominal MPC, SMPC, and CMPC. We choose a horizon of $N = 4$.

Since SMPC and CMPC can guarantee robust state constraint satisfaction (if $\mathbb{W}$ is exact), we include the state constraint as a hard constraint in these optimization problems. For nominal MPC, we instead convert this state constraint to a large violation penalty. With this penalty, we ensure that the constraint is satisfied for the nominal system (if possible) while retaining robust recursive feasibility of the optimization problem. Specifically, we redefine the stage cost as

$$\ell(x, u) := x'Qx + u'Ru + \lambda |x_1|_{[-1,1]}$$

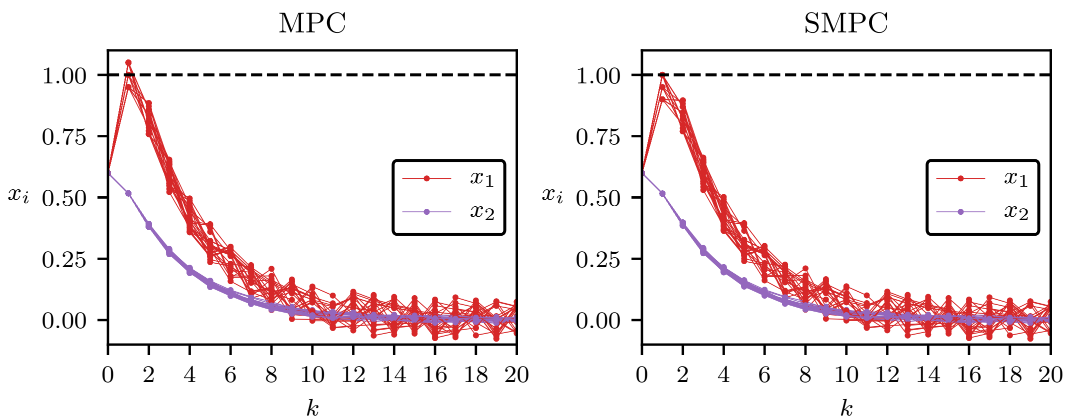

Figure 3.6: Closed-loop trajectories for nominal MPC (left) and SMPC (right) subject to 30 different realizations of the disturbance sequence for the state constraint example.

in which $\lambda > 0$ is the large violation penalty and $|x_1|_{[-1,1]} = \min\{|x_1 - y| : y \in [-1, 1]\}$. We find that $\lambda = 100$ is sufficient to ensure constraint satisfaction in the nominal optimization problem when possible. We use this stage cost in the subsequent statistics for the closed-loop systems of nominal MPC, CMPC, and SMPC.

In Figure 3.6, we plot the closed-loop trajectories for 30 realizations of the disturbance (drawn from the specified distribution) for MPC and SMPC. Both controllers initially drive x_1 away from the origin to minimize the value of x_2 and therefore the stage cost. The main difference between these methods is the value of $x_1(1)$. Nominal MPC drives the nominal value of $x_1(1)$ to the limit of state constraint in the interest of minimizing the nominal stage cost. Thus, any value of $w(0) > 0$ results in violation of the constraint $|x_1(1)| \leq 1$. By contrast, SMPC and CMPC are aware of the potential for a disturbance and therefore leave a buffer for the nominal value of $x_1(1)$ to accommodate this disturbance. Thus, SMPC and CMPC guarantee robust constraint satisfaction. For $k \geq 2$, however, all three methods produce similar state trajectories.

In Figure 3.7, we plot the sample average performance of each method in terms of $\hat{\mathbb{E}}[|x|]$ and $\hat{\mathbb{E}}[\ell(x, u)]$. Note that CMPC produces nearly identical performance to SMPC without using a stochastic cost function. At $k = 1$, MPC violates the state constraint and therefore

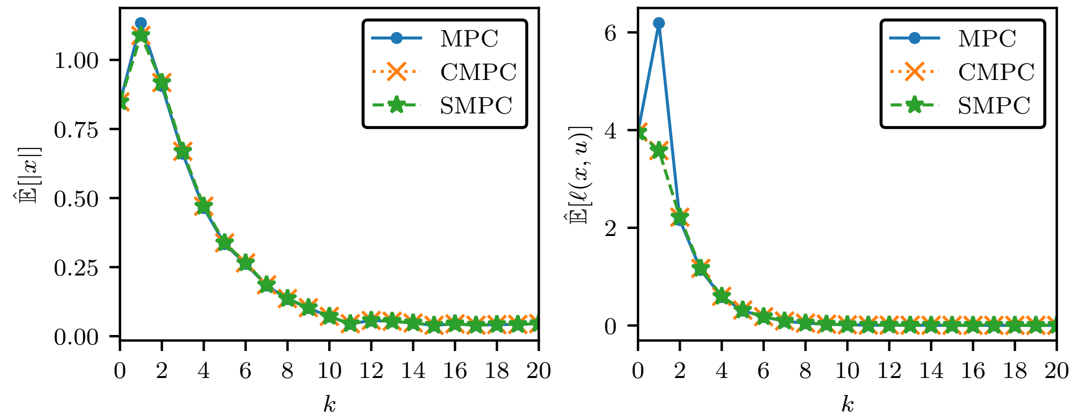

Figure 3.7: Sample average of the norm of the closed-loop state and closed-loop stage cost for the state constraint example.

the closed-loop performance, particularly in terms of the state cost, is inferior to SMPC and CMPC. But for $k \geq 2$, all of these controllers produce nearly equivalent performance.

We observe that once the state of this system is inside the terminal region and state/input constraints are not active, the optimal controller for both the nominal and stochastic linear system is in fact the LQR feedback gain used to construct the control law parameterization, i.e., $\kappa_\mu(x) = \pi(x, 0) = Kx$ for all $x \in \mathbb{X}_f$. Thus, nominal MPC, CMPC, and SMPC all use the same control law within the terminal and achieve the same closed-loop performance in this region.

3.7.4 Inventory control

The previous two examples focused on set point tracking and therefore use quadratic and positive definite stage costs. One of the key advantages of the MPC framework, however, is that the stage cost can be chosen to directly represent the economic costs of the problem of interest. We now consider a simple inventory control problem to demonstrate this capability. We describe the dynamics of this inventory control problem with the following scalar difference equation.

$$x^+ = x + u + w$$

in which $x \in \mathbb{R}$ is the current inventory level, $u \in [-1, 1]$ is the amount of inventory we choose to add or remove, and w is a disturbance. The disturbance w is a random variable with the probability distribution $\mu(\{-0.5\}) = \mu(\{0.5\}) = \varepsilon/2$ and $\mu(\{0\}) = 1 - \varepsilon$ for some $\varepsilon \in [0, 1]$. The economic stage cost consists of a storage cost for $x > 0$ and a larger backlog cost if $x < 0$. The stage cost is defined as follows.

$$\ell(x, u) := \max\{0, x\} + 3 \max\{0, -x\}$$

Note that the stage cost is not quadratic and is *asymmetric* about the origin. Nonetheless, this stage cost still satisfies Assumption 3.4 as $|x| \leq \ell(x, u)$ for all $x, u \in \mathbb{R}^2$.

We define the terminal set $\mathbb{X}_f := [-1, 1]$, terminal cost $V_f(x) = \ell(x, 0)$, and control law parameterization $\pi(x, v) = -x + v$. We choose $N = 3$ and $\mathbb{V} = [-5, 5]$ such that for any $(x, u) \in \mathcal{X}^s \times \mathbb{U}$ there exists $v \in \mathbb{V}$ that satisfies $\pi(x, v) = u$. We again solve the SMPC problem and calculate the expected value of the closed-loop state and stage cost by simulating all possible disturbance trajectories.

Since this problem is one dimensional, we can easily compute and plot the control law for SMPC for different values of ε and therefore different values of the variance $\text{tr}(\Sigma) = \varepsilon/4$. We plot these control laws in Figure 3.8 for all $x \in [-2, 2]$. Observe that for $\varepsilon < 0.5$, the control law is identical to that of nominal MPC and CMPC, i.e., deadbeat control that drives the state of the system to the origin as aggressively as possible. For these small values of ε, the disturbance is not significant enough to alter the control law. For $\varepsilon > 0.5$, however, the controller drives the state of the system to $x = 0.5$ instead of $x = 0$. By maintaining this extra inventory, we avoid the large backlog penalty that can occur if $w = -0.5$.

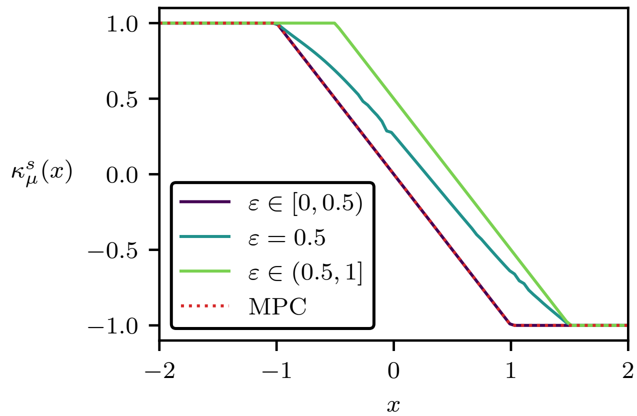

Figure 3.8: Control laws $\kappa_\mu^s(\cdot)$ for the inventory control problem with varying values of $\varepsilon \in [0, 1]$.

At exactly $\varepsilon = 0.5$, any value between these two curves, such as the curve shown in blue, is an optimal control action. Based on these observations, we conclude that the control law for SMPC in this simple inventory control problem is in fact *discontinuous* with respect to the probability distribution. By discontinuous, we mean that for a fixed $x \in \mathcal{X}^s$, arbitrarily small changes in ε can produce jumps in the optimal control action.

For $\varepsilon = 0.6$ and $x(0) = 2$, we simulate the closed-loop trajectory for nominal MPC and SMPC. For this value of ε, the SMPC control law maintains extra inventory to avoid the large backlog penalty and is therefore distinct from the nominal MPC control law. In Figure 3.9, we plot the expected value of the norm of the state and stage cost for the closed-loop trajectories of nominal MPC and SMPC. The closed-loop trajectory for CMPC is identical to nominal MPC and therefore omitted from Figure 3.9. Similar to the liquid level control problem, we observe that SMPC produces a larger value of $\mathbb{E}\left[|x(k)|\right]$ for all $k \in \mathbb{I}_{\geq 0}$, but a smaller value of $\mathbb{E}\left[\ell(x(k), u(k))\right]$. Unlike the liquid level control problem, however, the stage cost is the most important metric of robustness for this economic problem. Thus, the benefit of SMPC is clear for this economic applications of MPC.

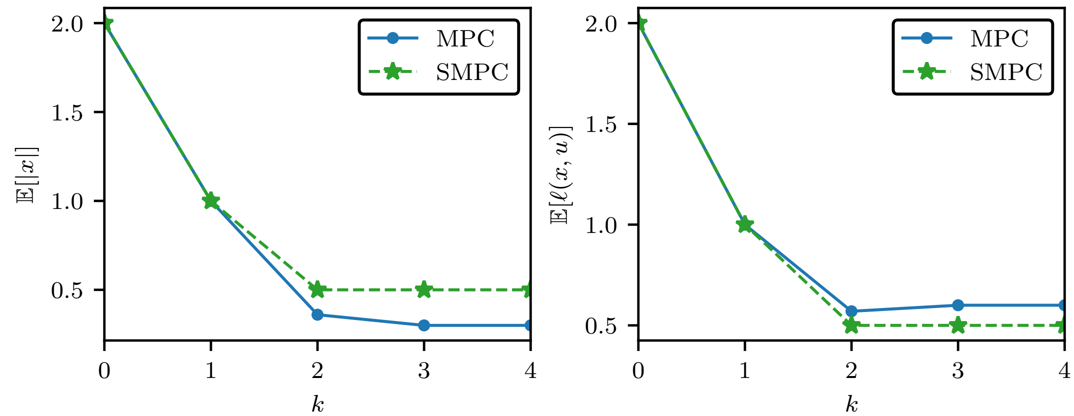

Figure 3.9: The expected value of the norm of the closed-loop state and closed-loop stage cost for nominal MPC and SMPC in the inventory control example.

3.8 Discussion and Summary

In this chapter, we introduced SMPC and established several novel results for this problem formulation. These results included fundamental mathematical properties, such as existence of optimal solutions and measurability of the control law (Proposition 3.8), as well as characterizations of the closed-loop performance via the definitions of RASiE and ℓ-RASiE (Theorem 3.18). One key shortcoming of SMPC, for nonlinear systems, is that the closed-loop system is not robust in a deterministic context. Consequently, there is no guarantee that the closed-loop state converges to, or remains at, the origin for a nominal realization of the disturbance ($\mathbf{w}_k = \mathbf{0}$). Motivated by this caveat of SMPC, we proposed CMPC as a middle ground between nominal MPC and SMPC. By using disturbance information to tighten constraints, but optimizing over a nominal cost function, CMPC retains all the stochastic and deterministic robustness properties of nominal MPC, but with the same RPI set as SMPC.

Informed by this collection of results, we then compared these three MPC formulations through a few conjectures and several examples. In these examples, we observe that SMPC can often find a superior operating point or trajectory than nominal MPC, in terms of the expected stage cost. Thus, Conjecture 3.22 is well motivated and is supported by all of these examples. Nonetheless, we were unable to prove this conjecture. Conjectures 3.20 and 3.21, however, do not hold. As shown in Section 3.7.1, nominal MPC may produce a larger RPI set than SMPC for an equivalent problem formulation. In Sections 3.7.2 and 3.7.4, we demonstrate counter examples to Conjecture 3.21. Thus, the claim that SMPC is necessarily " more robust" than nominal MPC should be qualified. There are reasonable definitions of robustness for which nominal MPC can outperform SMPC.

For control applications that prioritize economic performance, i.e., stage cost minimization, over stability of a target steady state, SMPC appears to offer a clear benefit over nominal MPC or CMPC, e.g, Section 3.7.4. If instead stability of a target steady state is prioritized, the benefits of SMPC are less clear, e.g., Sections 3.7.1 and 3.7.2. Feedback is sufficient to ensure that MPC, without knowledge of the disturbance model or distribution, achieves the same type of stochastic robustness afforded by SMPC, i.e., RASiE and ℓ-RASiE, for sufficiently small disturbances ($|w| \leq \delta$). If this margin of robustness is too small and/or robust constraint satisfaction is required for safety-critical applications, CMPC can be employed to achieve both stochastic and deterministic robustness without a stochastic objective function. Perhaps the most important critique of SMPC is that the origin is not guaranteed to be RAS, a property often seen as essential for a control algorithm.

One of the key assumptions made throughout this chapter is that the dynamical model and disturbance distribution used in the SMPC problem formulation are equivalent to that of the underlying plant. In practice, however, we do not have access to an exact model and distribution. At best, we can construct a reasonable approximation or estimate of these components. Thus, all the results discussed in this chapter are for *idealized* SMPC and the performance of SMPC in practice may significantly degrade relative to the idealized case. As an example, we revisit the control law in Figure 3.8. Note that this control law is discontinuous with respect to the probability distribution and, therefore, arbitrarily small errors in determining ε may produce significant changes in the selected control action. This observation raises the following important question: what happens to these performance guarantees if the system model and distribution used in the SMPC problem are not exact? In the next chapter, we address this question and show that, indeed, SMPC is *robust* to sufficiently small errors in the system model and disturbance distribution.

Chapter 4

Distributional Robustness

The most tenuous assumption made in SMPC, and indeed much of stochastic optimal control in general, is that the stochastic description of uncertainty, i.e., the dynamical model and probability distribution of the disturbance, is exact and comprehensive. This assumption, however, does not hold in any practical setting. The disturbance model and distribution, which are typically identified from data, are not exact and there may be relevant disturbances that are entirely absent from these models. This assumption is, of course, a reasonable and important starting point. The study of stochastic properties for idealized SMPC is analogous to the study of nominal properties for nominal MPC. The merit of this analysis is in establishing the best performance one can expect for SMPC. If the performance of idealized SMPC is insufficient, then there is little incentive to study nonideal conditions.

With the results established for idealized SMPC in the previous chapter, we now posses a suitable foundation to remove this idealized assumption of exact disturbance models and distributions. Specifically, we address an open question in the field of SMPC that is of significant practical concern: What, if any, robustness does SMPC confer for errors in the probability distribution used in the problem formulation, i.e., the *distributional robustness* of SMPC? The unwritten hypothesis is that feedback, similar to nominal MPC, provides some margin of inherent distributional robustness to SMPC and thereby addresses small discrepancies in the dynamical model and disturbance distribution. Again, we describe this robustness as *inherent* because we do not specifically design for this distributional robustness via, for example, the emerging field of distributionally robust optimization. This hypothesis, however, has never been formally stated or established for SMPC.

In this chapter, we address this question and establish that SMPC, under suitable assumptions, is inherently distributionally robust to errors in the probability distribution of the disturbance. We define this notion of distributional robustness via the Wasserstein metric, a metric that quantifies the distance between two probability measures and is introduced in Section 4.2. This result addresses incorrectly or unmodeled disturbances that enter the closed-loop system and also demonstrates the efficacy of scenario optimization as a means to approximate and solve the SMPC problem. Moreover, we can treat CMPC and nominal MPC as special cases of SMPC within this framework and thereby unify the descriptions of stochastic robustness across these different MPC formulations.

4.1 Problem formulation and preliminaries

To properly formulate and analyze the closed-loop properties of non-idealized SMPC, we must introduce both the underlying stochastic system of the plant and the stochastic model used in the SMPC optimization problem. Thus, we reintroduce the stochastic system and SMPC problem formulation in this chapter with additional notation to distinguish properties of the underlying stochastic system and the stochastic model. All quantities (e.g., sets and probability measures) of the stochastic model used in the SMPC problem formulation are indicated with a “hat” placed on the notation for that quantity used in the previous chapter (e.g., $\hat{\mathbb{W}}$ and $\hat{\mu}$). We also drop the superscript “s” for quantities defined for SMPC, e.g., we use $\mathcal{X}$ instead of $\mathcal{X}^s$ in this chapter.

4.1.1 The stochastic system(s)

We consider the following discrete-time stochastic system

$$x^+ = f(x, u, w) \quad f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^n$$

in which $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the controlled input, $w \in \mathbb{W} \subseteq \mathbb{R}^q$ is the stochastic disturbance, and x^+ is the successor state. We let $\mathbf{w}_i := (w(0), w(1), \ldots, w(i-1))$ denote a sequence of disturbances. We consider the following assumption.

Assumption 4.1 (Disturbances)**.** The disturbances $w \in \mathbb{W}$ are random variables that are i.i.d. in time. The set $\mathbb{W}$ is compact and contains the origin.

Given Assumption 4.1, we denote the probability measure for w as $\mu : \mathcal{B}(\mathbb{W}) \to [0, 1]$. Let $\mathcal{M}(\mathbb{W})$ denote the set of all probability measures on the measurable space $(\mathbb{W}, \mathcal{B}(\mathbb{W}))$. We define expected value of a Borel measurable function $g : \mathbb{W}^i \to \mathbb{R}$ as the Lebesgue integral

$$\mathbb{E}\left[g(\mathbf{w}_i)\right] := \int_{\mathbb{W}^i} g(\mathbf{w}_i) d\mu(w(0)) d\mu(w(1)) \ldots d\mu(w(i-1))$$

with respect to any $\mu \in \mathcal{M}(\mathbb{W})$. Note that we no longer require w to be zero mean and therefore expand the collection of probability measures included in $\mathcal{M}(\mathbb{W})$.

In this chapter, we do not assume that the set $\mathbb{W}$ or measure μ is known. Instead, we have access to only a *model* of the set $\mathbb{W}$ and the probability measure μ that we denote $\hat{\mathbb{W}}$ and $\hat{\mu}$, respectively. In the SMPC problem formulation, the stochastic system evolves according to the following stochastic model and without knowledge of $\mathbb{W}$ or μ.

$$x^+ = f(x, u, \hat{w}) \quad \hat{w} \in \hat{\mathbb{W}} \tag{4.1}$$

in which $\hat{w}$ is distributed according to the measure $\hat{\mu}$.

We may assume that $\hat{\mathbb{W}} \subseteq \mathbb{W}$ without loss of generality because we can increase the size of $\mathbb{W}$ to fit $\hat{\mathbb{W}}$ and assign these additional values measure zero with μ. We may also, without loss of generality, define $\hat{\mu}$ on the domain $\mathcal{B}(\mathbb{W})$ by assigning measure zero to all the points in $\mathbb{W}$ that are not in $\hat{\mathbb{W}}$, i.e., $\hat{\mu} : \mathcal{B}(\mathbb{W}) \rightarrow [0,1]$ such that $\hat{\mu}(\mathbb{W} \setminus \hat{\mathbb{W}}) = 0$. Specifically, we have that

$$\int_{\mathbb{W}} g(\hat{w}) d\hat{\mu}(\hat{w}) = \int_{\hat{\mathbb{W}}} g(\hat{w}) d\hat{\mu}(\hat{w})$$

for all measurable functions $g(\cdot)$. We ensure that μ and $\hat{\mu}$ are defined on the same domain to facilitate the comparison of these two measures via the Wasserstein metric. We formalize these requirements of the stochastic model through the following assumption.

Assumption 4.2 (Disturbance model)**.** The random variables $\hat{w}(i)$ are i.i.d. in time, with a known probability measure $\hat{\mu} : \mathcal{B}(\mathbb{W}) \rightarrow [0,1]$. The set $\hat{\mathbb{W}}$ is compact and contains the origin. The probability distribution satisfies $\hat{\mu}(\hat{\mathbb{W}}) = 1$.

We assume that there is a fixed set $\hat{\mathbb{W}}$ used in the SMPC algorithm, as also done in the previous chapter for idealized SMPC. Note that the subsequent assumptions are based on a single set $\hat{\mathbb{W}}$. We do, however, allow for different $\hat{\mu}$ and derive bounds that apply for any $\hat{\mu}$ that satisfies Assumption 4.2. We use $\hat{\mathcal{M}}(\mathbb{W})$ to denote the set of all probability measures on the measurable space $(\mathbb{W}, \mathcal{B}(\mathbb{W}))$ that satisfy Assumption 4.2, i.e., $\hat{\mu}(\hat{\mathbb{W}}) = 1$ for all $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$. Note that $\hat{\mathcal{M}}(\mathbb{W}) \subseteq \mathcal{M}(\mathbb{W})$. We illustrate these sets in Figure 4.1.

We use the Hausdorff distance, a standard measure of distance between sets, to characterize the distance between $\mathbb{W}$ and $\hat{\mathbb{W}}$. The Hausdorff distance between two sets $X, Y \subseteq \mathbb{R}^n$ is defined as

$$d_H(X,Y) := \max \left\{ \sup_{x \in X} |x|_Y, \ \sup_{y \in Y} |y|_X \right\}$$

in which $|x|_Y$ denotes the point-to-set distance from x to Y, i.e., $|x|_Y = \inf_{y \in Y} |x - y|$. Note that since $\hat{\mathbb{W}} \subseteq \mathbb{W}$ and both sets are compact, we have that $d_H(\mathbb{W}, \hat{\mathbb{W}}) = \max_{w \in \mathbb{W}} |w|_{\hat{\mathbb{W}}}$. We

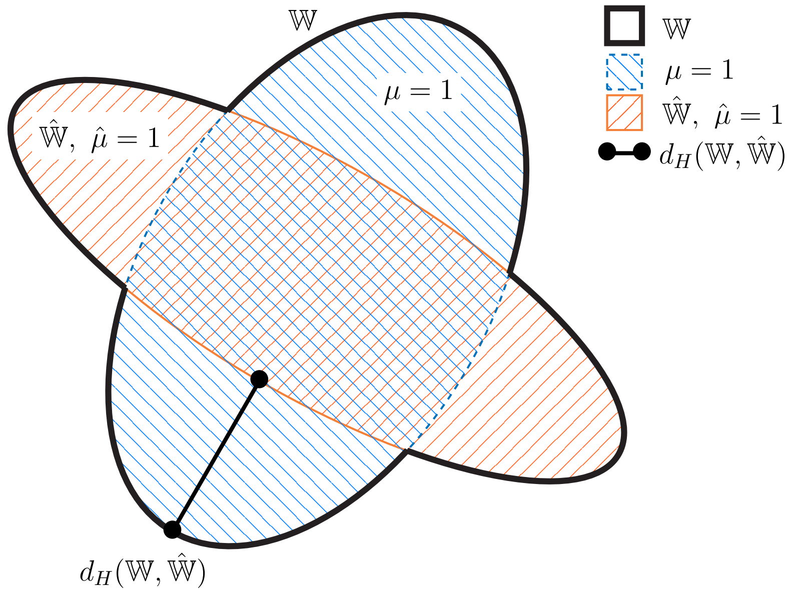

Figure 4.1: Illustration of the sets $\mathbb{W}$, $\hat{\mathbb{W}}$ and the probability measures μ, $\hat{\mu}$. We also show the Hausdorff distance $d_H(\cdot)$ between these two sets.

show an example of this distance in Figure 4.1.

We emphasize that the stochastic model $(\hat{\mathbb{W}}, \hat{\mu})$ and stochastic system $(\mathbb{W}, \mu)$ are not necessarily equivalent. Furthermore, we allow $\hat{\mathbb{W}}$ to be a finite set (e.g., $\hat{\mathbb{W}} = \{0, 1\}$) even if $\mathbb{W}$ is uncountable (e.g., $\mathbb{W} = [0, 1]$) and can therefore address scenario-based (empirical) approximations of $\mathbb{W}$ and μ within this framework. This framework is also general enough to address disturbances that are entirely absent from the SMPC model and optimization problem. For example, we can consider nominal MPC by defining $\hat{\mathbb{W}} = \{0\}$ and $\hat{\mu}(\{0\}) = 1$. We can also consider "over-modeling" of the disturbance in SMPC. For example, we use $\hat{\mathbb{W}} = [-1, 1]$ in the SMPC problem, but $\mu([-0.5, 0.5]) = 1$, i.e., the probability that the disturbance takes a value outside of $[-0.5, 0.5]$ is zero. Thus, we can represent incorrectly modeled ($\hat{\mu} \neq \mu$), unmodeled ($\hat{\mathbb{W}} \neq \mathbb{W}$), or out-of-sample ($\hat{\mathbb{W}}$ is finite) disturbances.

For any $i \in \mathbb{I}_{\geq 0}$, $N \in \mathbb{I}_{\geq 1}$, the sequence of i.i.d. random variables $\hat{\mathbf{w}} := (\hat{w}(i), \hat{w}(i +$

$1), \ldots, \hat{w}(i+N-1))$ has a joint probability measure $\hat{\mu}^N : \mathcal{B}(\mathbb{W}^N) \to [0,1]$ defined as

$$\hat{\mu}^N(F) = \hat{\mu}(F_i)\hat{\mu}(F_{i+1})\ldots\hat{\mu}(F_{i+N-1})$$

for all $F = (F_i, F_{i+1}, \ldots, F_{i+N-1}) \in \mathcal{B}(\mathbb{W}^N)$. For any Borel measurable function $g : \mathbb{W}^N \to \mathbb{R}$, we define expected value with respect to $\hat{\mu}$ as

$$\hat{\mathbb{E}}\left[g(\hat{\mathbf{w}})\right] = \int_{\hat{\mathbb{W}}^N} g(\hat{\mathbf{w}}) d\hat{\mu}^N(\hat{\mathbf{w}})$$

Note that we use $\hat{\mathbb{E}}[\cdot]$ to indicate expected value with respect to $\hat{\mu}$ instead of μ. We frequently use the expected value of $|\hat{w}|$ in the following analysis and note the following inequality

$$\hat{\mathbb{E}}\left[|\hat{w}|\right] = \int_{\hat{\mathbb{W}}} |\hat{w}| d\hat{\mu}(\hat{w}) \leq \sqrt{\text{tr}\left(\hat{\Sigma}\right) + \left(\hat{\mathbb{E}}[\hat{w}]\right)^2}$$

in which $\hat{\Sigma} \succeq 0$ is the covariance matrix of $\hat{w}$.

Remark 4.3. Unlike the previous chapter, we do not assume that the disturbance in the SMPC optimization problem is zero mean, i.e., we allow $\hat{\mathbb{E}}[\hat{w}] \neq 0$. Thus, the covariance matrix $\hat{\Sigma}$ that appears in the results introduced in Chapters 2 and 3 does not appear in the subsequent results in this chapter. While the assumption $\hat{\mathbb{E}}[\hat{w}] = 0$ is typically used without loss of generality if a continuous distribution is considered (e.g., a uniform distribution), scenario-based approximations of μ do not produce empirical distributions with zero mean. Thus, we do not restrict $\hat{\mu}$ to only measures of zero mean. In the special case that $\hat{\mathbb{E}}[\hat{w}] = 0$, we may use the upper bound $\hat{\mathbb{E}}[|\hat{w}|] \leq \hat{\Sigma}^{1/2}$ to introduce $\hat{\Sigma}$ into the following results.

4.1.2 SMPC problem formulation

We again define a parametrized control policy $\pi : \mathbb{R}^n \times \mathbb{V} \to \mathbb{R}^m$ in which $x \in \mathbb{R}^n$ is the current state of the system and $v \in \mathbb{V} \subseteq \mathbb{R}^l$ are the parameters in the control policy. The resulting system is

$$x^+ = f(x, \pi(x, v), \hat{w}) \quad \hat{w} \in \hat{\mathbb{W}} \tag{4.2}$$

in which $\hat{w}$ is distributed according to $\hat{\mu}$. Let $\hat{\phi}(k; x, \mathbf{v}, \hat{\mathbf{w}})$ denote the solution of (4.2) at time $k \in \mathbb{I}_{0:N}$, given the initial state $x \in \mathbb{R}^n$, the trajectory of control policy parameters $\mathbf{v} \in \mathbb{V}^N$, and disturbances trajectory $\hat{\mathbf{w}} \in \hat{\mathbb{W}}^N$.

We consider hard input constraints, i.e., $u \in \mathbb{U} \subseteq \mathbb{R}^m$, but do not allow hard or probabilistic constraints on the state. Since we do not assume that the disturbance set is exact, a disturbance that is not considered in the set $\hat{\mathbb{W}}$ may cause the closed-loop system to violate these constraints. Similar to nominal MPC, we assume that all state constraints (except the terminal constraint) are converted to penalty functions in the stage cost.

For a horizon $N \in \mathbb{I}_{\geq 1}$ and terminal constraints $\mathbb{X}_f \subseteq \mathbb{R}^n$, we define

$$\begin{aligned}\mathcal{V}(x) := \{\mathbf{v} \in \mathbb{V}^N : \pi(\hat{\phi}(k; x, \mathbf{v}, \hat{\mathbf{w}}), v(k)) \in \mathbb{U}\ \forall \hat{\mathbf{w}} \in \hat{\mathbb{W}}^N,\ k \in \mathbb{I}_{0:N-1}; \\ \hat{\phi}(N; x, \mathbf{v}, \hat{\mathbf{w}}) \in \mathbb{X}_f\}\end{aligned}$$

The set of all feasible initial states is denoted

$$\mathcal{X} := \{x \in \mathbb{R}^n : \mathcal{V}(x) \neq \emptyset\}$$

The remainder of the SMPC problem formulation is nearly identical to that in Chapter 3, except that we use the disturbance set $\hat{\mathbb{W}}$ and probability measure $\hat{\mu}$. Note, however, that we drop the subscript s to indicate SMPC, as we do not consider nominal MPC and CMPC as separate formulations in this chapter.

For the stage cost $\ell : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ and terminal cost $V_f : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$, we define the function

$$J(x, \mathbf{v}, \hat{\mathbf{w}}) := \sum_{k=0}^{N-1} \ell(x(k), \pi(x(k), v(k))) + V_f(x(N))$$

in which $x(k) = \hat{\phi}(k; x, \mathbf{v}, \hat{\mathbf{w}})$. We define the SMPC cost function as

$$V_{\hat{\mu}}(x, \mathbf{v}) := \hat{\mathbb{E}}\left[J(x, \mathbf{v}, \hat{\mathbf{w}})\right] = \int_{\hat{\mathbb{W}}^N} J(x, \mathbf{v}, \hat{\mathbf{w}}) d\hat{\mu}^N(\hat{\mathbf{w}})$$

The optimization problem is defined as

$$\mathbb{P}_{\hat{\mu}}(x): \; V^0_{\hat{\mu}}(x) := \min_{\mathbf{v} \in \mathcal{V}(x)} V_{\hat{\mu}}(x, \mathbf{v}) \tag{4.3}$$

and the optimal solution(s) for a given distribution $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$ are defined by the set-valued mapping $\mathbf{v}^0_{\hat{\mu}} : \mathcal{X} \rightrightarrows \mathbb{V}^N$ such that

$$\mathbf{v}^0_{\hat{\mu}}(x) := \arg \min_{\mathbf{v} \in \mathcal{V}(x)} V_{\hat{\mu}}(x, \mathbf{v})$$

We use a Borel measurable selection rule to define a single-valued control law $\kappa_{\hat{\mu}} : \mathcal{X} \rightarrow \mathbb{U}$ such that

$$\kappa_{\hat{\mu}}(x) \in \{\pi(x, v(0)) : \mathbf{v} \in \mathbf{v}^0_{\hat{\mu}}(x)\}$$

for all $x \in \mathcal{X}$. Note that the control law depends on the disturbance model $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$.

We now reintroduce the underlying disturbance set and distribution in the closed-loop system

$$x^+ = f(x, \kappa_{\hat{\mu}}(x), w) \quad w \in \mathbb{W} \tag{4.4}$$

in which w is distributed according to $\mu \in \mathcal{M}(\mathbb{W})$ but the control law is defined based on

$\hat{\mu}$. We use $\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)$ to denote the solution to (4.4) at time $k \in \mathbb{I}_{\geq 0}$, given the initial condition $x \in \mathcal{X}$ and disturbance sequence $\mathbf{w}_k \in \mathbb{W}^k$. The deterministic value of the closed-loop trajectory $\phi_{\hat{\mu}}(\cdot)$ depends on the disturbance model $\hat{\mu}$ because $\kappa_{\hat{\mu}}(\cdot)$ depends on $\hat{\mu}$. But the disturbance takes values $w \in \mathbb{W}$ and the expected value of the closed-loop system is evaluated based on probability measure $\mu \in \mathcal{M}(\mathbb{W})$. Thus, in the subsequent analysis, we discuss quantities such as

$$\mathbb{E}\left[|\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)|\right] = \int_{\mathbb{W}^k} |\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)| d\mu^k(\mathbf{w}_k)$$

4.1.3 Assumptions for SMPC

To establish the distributional robustness of SMPC, we require stronger assumptions than for idealized SMPC that are similar to the assumptions required for nominal MPC in Chapter 2.

Assumption 4.4 (Continuity of system and cost)**.** The model $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^n$, control parameterization $\pi : \mathbb{R}^n \to \mathbb{V} \to \mathbb{R}^m$, stage cost $\ell : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, and terminal cost $V_f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ are locally Lipschitz continuous. Furthermore, $f(0, 0, 0) = 0$, $\ell(0, 0) = 0$, and $V_f(0) = 0$.

Note that we have strengthened the usual assumption of continuity to local Lipschitz continuity. This additional restriction, however, is minor because most physical systems are described by local Lipschitz continuous functions.[1] Notably, quadratic costs and linear systems satisfy this requirement. Furthermore, locally Lipschitz continuous functions are required if we intend to use gradient-based, nonlinear optimization solvers to solve these optimal control problems. This fact is true for nominal MPC problems as well. This assumption of local Lipschitz continuity, however, does not imply that the optimal cost function $V^0_{\hat{\mu}}(\cdot)$ or control law $\kappa_{\hat{\mu}}(\cdot)$ are continuous for $x \in \mathcal{X}$.

[1] For example, all of the example problems discussed in this dissertation consider locally Lipschitz continuous system models.

Assumption 4.5 (Properties of constraint sets)**.** The sets $\mathbb{U}$ and $\mathbb{V}$ are compact and contain the origin and $\mathbb{X}_f := \{x \in \mathbb{R}^n : V_f(x) \leq \tau\}$ for some $\tau > 0$. The set $\mathcal{X}$ is bounded. The control law parameterization satisfies $\pi(x, v) \in \mathbb{U}$ for all $x \in \mathbb{R}^n$ and $v \in \mathbb{V}$.

The final requirement of Assumption 4.5 means that $\pi(x, v) = Kx + v$ may not be a valid control law parameterization. Instead, we can use $\pi(x, v) = \text{sat}_{\mathbb{U}}(Kx + v)$ in which $u = \text{sat}_{\mathbb{U}}(s)$ maps s to the nearest input that satisfies $u \in \mathbb{U}$, i.e., $\text{sat}_{\mathbb{U}}(s) = \arg\min_{u\in\mathbb{U}} |u - s|$.

Assumption 4.6. There exists a locally Lipschitz continuous terminal control law $\kappa_f : \mathbb{X}_f \to \mathbb{U}$ and $\tilde{\tau} \leq \tau$ such that for all $x \in \mathbb{X}_f$,

$$f(x, \kappa_f(x), \hat{w}) \in \{x : V_f(x) \leq \tilde{\tau}\} \subset \mathbb{X}_f, \ \forall \hat{w} \in \hat{\mathbb{W}} \tag{4.5}$$

$$V_f(f(x, \kappa_f(x), 0)) \leq V_f(x) - \ell(x, \kappa_f(x)) \tag{4.6}$$

Furthermore, $\pi(x, 0) = \kappa_f(x)$ for all $x \in \mathbb{X}_f$.

Thus, the terminal control law must drive $x \in \mathbb{X}_f$ to the *interior* of $\mathbb{X}_f$ for all $\hat{w} \in \hat{\mathbb{W}}$. This assumption is stronger than the analogous assumption typically used in SMPC, i.e., Assumption 3.6, that requires only robust positive invariance of the terminal set. By strengthening this assumption, we allow for some nonzero error in approximating $\mathbb{W}$ with $\hat{\mathbb{W}}$.

For tracking problems, we also require the usual lower bound on the stage cost.

Assumption 4.7. There exists a function $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$ such that

$$\ell(x, u) \geq \alpha_\ell(|x|)$$

for all $(x, u) \in \mathbb{R}^n \times \mathbb{U}$.

These assumptions address the construction of the SMPC optimization problem and the set $\hat{\mathbb{W}}$, but do not place any restrictions on the disturbance set $\mathbb{W}$ and probability measure

μ. We also allow any $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$ for the chosen set $\hat{\mathbb{W}}$. We also note that Proposition 3.8 still holds for this SMPC formulation and therefore all relevant stochastic properties are well defined.

4.2 Distance and convergence for probability measures

We now have two probability measures: the disturbance distribution μ and a model of the disturbance distribution $\hat{\mu}$ that is used in the SMPC optimization problem. The goal is to show that small differences between μ and $\hat{\mu}$ produce similarly small deviations in the closed-loop performance bounds relative to idealized SMPC. Moreover, we wish to show that as $\hat{\mu} \to \mu$, we recover the idealized SMPC guarantees established in Theorem 3.18. This goal requires that we first address an important mathematical question: How do we define this difference or distance between probability measures? And furthermore, how do we define the convergence $\hat{\mu} \to \mu$?

4.2.1 Wasserstein metric

The most intuitive concept of a distance or metric is the Euclidean distance $|x - y|$ between two points $x, y \in \mathbb{R}^n$. The concept of distance, however, can be generalized[2] to address functions, sets, and indeed measures. For measures, there are several notions of distance available, but we find the Wasserstein metric most suitable for the task at hand. While this metric initially found interest in the field of optimal transport (Villani, 2009), there are several recent applications of the Wasserstein metric in machine learning, state estimation, and optimal control. We note a few of these contributions here.

In a particularly significant contribution, Arjovsky et al. (2017) use the Wasserstein metric as a loss function in training generative adversarial networks (GANs). The authors showed

[2]like any good mathematical concept

several notably improvements in algorithmic stability and performance compared to the Jensen-Shannon divergence loss function typically used to train GANs. The Wasserstein metric has also been used as a loss function in multi-label learning problems (Frogner et al., 2015).

The Wasserstein metric can also be used to define an ambiguity set, i.e., a set of potential probability measures within some radius (defined by the Wasserstein metric) of a central probability measure. This ambiguity set is then used to formulate a min-max optimization problem in which the goal is to minimize the cost of the stochastic objective function subject to the worst possible probability measure within this ambiguity set. This problem formulation is known as distributionally robust optimization (DRO) (Goh and Sim, 2010). Shafieezadeh Abadeh et al. (2018) use this concept of distributionally robust optimization to formulate a new Kalman filtering algorithm and based on a tractable optimization problem. Yang (2020) propose a DRO formulation for nonlinear stochastic optimal control based on sample average approximation (SAA). We note that in both of these cases, the specific problem formulation is discussed in great detail, but the analysis of the closed-loop system is not addressed. Thus, it remains unknown if these algorithms confer any of this distributional robustness included in the optimization problem to the actual closed-loop system. In contrast to these approaches, we do not use the Wasserstein metric in the formulation of SMPC problem. Instead, we use the Wasserstein metric only as a means to quantify the distance between μ and $\hat{\mu}$. We then characterize the behavior of the closed-loop system via this distance in the definition of distributional robustness introduced in the following section.

We consider the type-1 version of the Wasserstein metric, sometimes known as the Kantorovich-Rubinstein metric, defined as follows. Recall that we use $\mathcal{M}(\mathbb{W})$ to denote the space of all Borel probability measures on the compact set $\mathbb{W} \subseteq \mathbb{R}^q$.

Definition 4.8 (Wasserstein metric)**.** The Wasserstein metric (type-1) is denoted $W : \mathcal{M}(\mathbb{W}) \times$

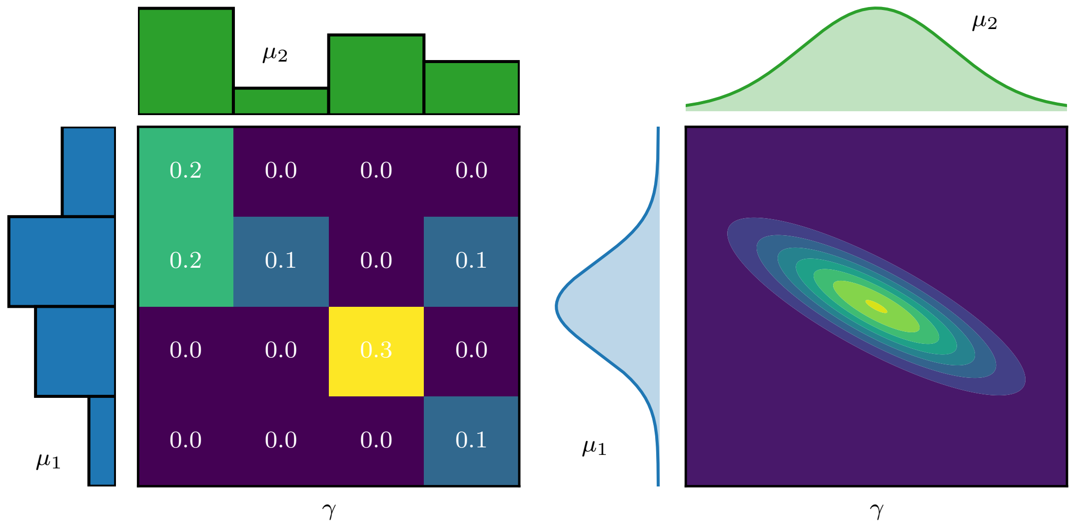

Figure 4.2: Potential values for $\gamma \in \Gamma(\mu_1, \mu_2)$ for measures with finite support (left) and $\mathbb{W} = \mathbb{R}$ (right).

$\mathcal{M}(\mathbb{W}) \to \mathbb{R}_{\geq 0}$ and defined as

$$W(\mu_1, \mu_2) := \inf_{\gamma \in \Gamma(\mu_1, \mu_2)} \int_{\mathbb{W} \times \mathbb{W}} |w_1 - w_2| d\gamma(w_1, w_2)$$

for all $\mu_1, \mu_2 \in \mathcal{M}(\mathbb{W})$, in which $\Gamma(\mu_1, \mu_2)$ denotes the collection of all measures on $\mathbb{W} \times \mathbb{W}$ with marginals μ_1 and μ_2, i.e., $\gamma \in \Gamma(\mu_1, \mu_2)$ must satisfy

$$\mu_1(\cdot) = \int_{\mathbb{W}} \gamma(\cdot, w_2) dw_2 \qquad \mu_2(\cdot) = \int_{\mathbb{W}} \gamma(w_1, \cdot) dw_1$$

The Wasserstein metric has received significant attention in the field of optimal transport (Villani, 2009) and is most convenient to describe according to this application. The measure $\gamma(\cdot) \in \Gamma(\mu_1, \mu_2)$ can be viewed as a *transport plan* for relocating density or "earth" from a distribution described by μ_1 to another distribution described by μ_2. For this reason, the Wasserstein metric for finite sets $\mathbb{W}$ is often referred to as the "earth mover's" distance. We plot potential values of $\gamma(\cdot)$ for one dimensional distributions in Figure 4.2.

Thus, determining the Wasserstein metric amounts to solving for the optimal transport plan in which the cost is given by the Euclidean distance $|\cdot|$ multiplied by the density of "earth" to be moved. Note that the Wasserstein metric satisfies all the axioms of a distance on $\mathcal{M}(\mathbb{W})$ for compact $\mathbb{W}$, i.e., the metric is finite, symmetric, $W(\mu_1, \mu_2) = 0$ if and only if $\mu_1 = \mu_2$, and satisfies triangle inequality

$$W(\mu_1, \mu_3) \leq W(\mu_1, \mu_2) + W(\mu_2, \mu_3)$$

for all $\mu_1, \mu_2, \mu_3 \in \mathcal{M}(\mathbb{W})$.

To better illustrate this Wasserstein metric, we consider one dimensional distributions μ_1 and μ_2 on the set $\mathbb{W} := [-4, 4]$. The probability measure μ_1 is a truncated normal distribution and μ_2 is a discrete distribution consisting of three events with equal probability, i.e., $\mu_2(\{-1\}) = \mu_2(\{0\}) = \mu_2(\{1\}) = 1/3$. In Figure 4.3, we plot the probability density function $p_i(w) = \frac{d\mu_i}{dw}(w)$ and cumulative density function $F_i(w) := \mu_i([-4, w])$ for these two probability measures. We use arrows to represent the delta functions in $p_2(w)$. For one dimensional systems, the Wasserstein metric admits a convenient simplification in terms of cumulative distribution functions:

$$W(\mu_1, \mu_2) = \int_{\mathbb{W}} |F_1(w) - F_2(w)| dw$$

In the bottom plot of Figure 4.3, we show $|F_1(w) - F_2(w)|$ for these two probability measures. The area under this curve is the value of the Wasserstein distance between these two probability measures, with $W(\mu_1, \mu_2) = 0.343$ for this example.

As a particular simple example, we consider the Wasserstein distance between two point masses. Recall that we defined δ_m as the Dirac measure for some point $m \in \mathbb{R}^q$, i.e., $\delta_m(S) = 1$ if $m \in S$ and zero otherwise. We consider two measures $\mu_1 = \delta_{m_1}$ and $\mu_2 = \delta_{m_2}$ for some points $m_1, m_2 \in \mathbb{R}^q$. For this example, $W(\mu_1, \mu_2) = |m_1 - m_2|$ thus preserving the intuitive

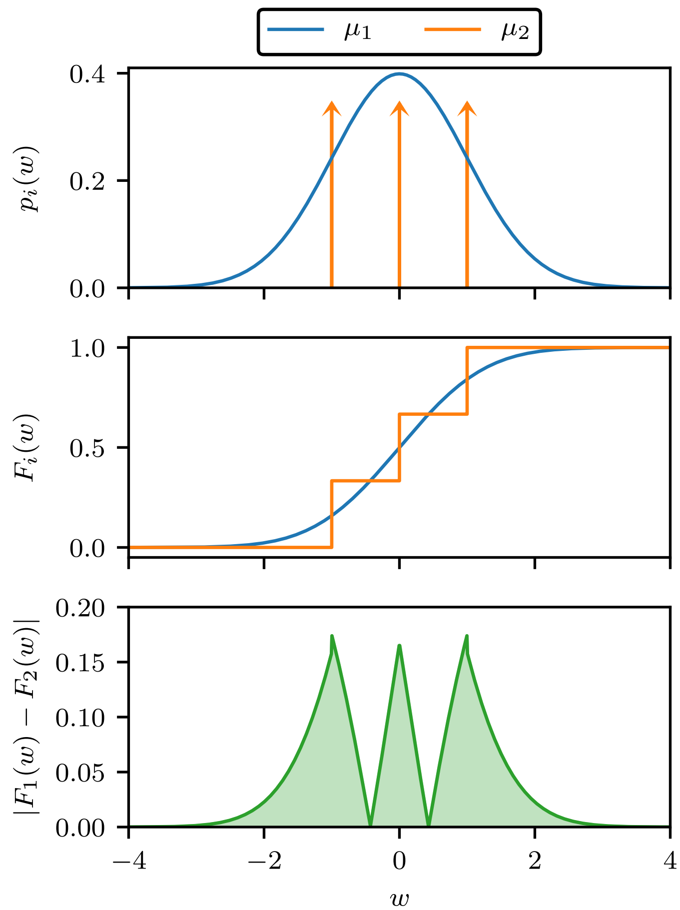

Figure 4.3: The probability density $p(w)$ and cumulative distribution $F(w)$ for two probability measures μ_1, μ_2 on the set $\mathbb{W} := [-4, 4]$. The arrows indicate delta functions, each with a weight of 1/3. The bottom plot is the absolute difference between $F_1(w)$ and $F_2(w)$ for all $w \in [-4, 4]$.

notion of Euclidean distance between these points.

Another useful relation is the Wasserstein distance between any probability measure and a point mass, i.e., $\mu_1 \in \mathcal{M}(\mathbb{W})$ and $\mu_2 = \delta_{m_2}$ for $m_2 \in \mathbb{W}$. For these two measures, one can show that the only available $\gamma \in \Gamma(\mu_1, \mu_2)$ is given by the product $\gamma(S_1, S_2) = \mu_1(S_1)\delta_{m_2}(S_2)$ for all $S_1, S_2 \in \mathcal{B}(\mathbb{W})$. Thus, the Wasserstein distance between these two distributions reduces to

$$W(\mu_1, \delta_{m_2}) = \int_{\mathbb{W}} |w_1 - m_2| d\mu_1(w_1)$$

If $m_2 = 0$, this further reduces to $\int_{\mathbb{W}} |w_1| d\mu_1(w_1)$, i.e., the expected value of the norm of w_1.

A particularly important result is the dual representation of the Wasserstein metric shown in the following theorem.

Theorem 4.9 (Kantorovich-Rubinstein)**.** *For any probability measures $\mu_1, \mu_2 \in \mathcal{M}(\mathbb{W})$, we have that*

$$W(\mu_1, \mu_2) = \sup_{g \in \mathcal{L}} \left\{ \int_{\mathbb{W}} g(w_1) d\mu_1(w_1) - \int_{\mathbb{W}} g(w_2) d\mu_2(w_2) \right\}$$

in which $\mathcal{L}$ denotes the space of all Lipschitz continuous function with $|g(w_1) - g(w_2)| \leq |w_1 - w_2|$ for all $w_1, w_2 \in \mathbb{W}$.

Further discussion of this result can be found in Villani (2009, Remark 6.5). Thus, for any Lipschitz continuous function $g(\cdot)$ with Lipschitz constant $L \geq 0$ on $\mathbb{W}$, we have that

$$\int_{\mathbb{W}} g(w_1) d\mu_1(w_1) - \int_{\mathbb{W}} g(w_2) d\mu_2(w_2) \leq L W(\mu_1, \mu_2)$$

for all $\mu_1, \mu_2 \in \mathcal{M}(\mathbb{W})$. This inequality is essential to establish the following results in this chapter and is the reason we restrict the SMPC problem formulation to locally Lipschitz continuous functions in Assumption 4.4.

4.2.2 Weak convergence of measures

In addition to the Wasserstein metric, we also introduce a notion of convergence for probability measures in $\mathcal{M}(\mathbb{W})$. We first recall the concept of convergence for sequences of points. We say that the sequence of points $\{w_i\}_{i=1}^{\infty}$ in $\mathbb{R}^q$ converges to a point $w \in \mathbb{R}^q$ if

$$\lim_{i\to\infty} |w_i - w| = 0$$

i.e., for any $\varepsilon > 0$, there exists $k \in \mathbb{I}_{\geq 0}$ such that $|w_i - w| \leq \varepsilon$ for all $i \geq k$. We use the notation $w_i \to w$ to indicate this convergence.

For a sequence of probability measures $\{\mu_i\}_{i=1}^{\infty}$ in $\mathcal{M}(\mathbb{W})$, we say that μ_i converges *weakly* to $\mu \in \mathcal{M}(\mathbb{W})$ if

$$\lim_{i\to\infty} \left| \int_{\mathbb{W}} g(w) d\mu_i(w) - \int_{\mathbb{W}} g(w) d\mu(w) \right| = 0$$

for all continuous functions $g : \mathbb{W} \to [-1, 1]$. We use the notation $\mu_i \to \mu$ to denote weak convergence.

Remark 4.10. We refer to this convergence as *weak* because we consider only *continuous* functions in the definition. Total variation convergence instead considers all *measurable* functions $g : \mathbb{W} \to [-1, 1]$ and is therefore a stronger notion of convergence. Total variation convergence, however, is often too strong for many relevant problems. For example, sampling-based empirical distributions may never converge in total variation to the continuous distribution from which they are constructed.

The notion of weak convergence for probability measures is particularly important for sampling-based empirical approximations of the probability measure. For example, consider a probability measure $\mu \in \mathcal{M}(\mathbb{W})$. We draw $s \in \mathbb{I}_{\geq 1}$ random samples from μ that we denote

$\{\hat{\omega}_i\}_{i=1}^s$ and define the empirical probability measure as

$$\hat{\mu}_s(\cdot) := \frac{1}{s}\sum_{i=1}^{s}\delta_{\omega_i}(\cdot)$$

For $\mathbb{W} \subseteq \mathbb{R}^q$, one can show that $\hat{\mu}_s \to \mu$ (with probability one) as $s \to \infty$ via the strong law of large numbers and the fact that $\mathbb{W}$ is separable (Varadarajan, 1958).

In the context of weak convergence of probability measures, the choice of Wasserstein metric is particularly clear.

Theorem 4.11 (W metrizes $\mathcal{M}$)**.** *The Wasserstein metric metrizes convergence in* $\mathcal{M}(\mathbb{W})$ *for compact* $\mathbb{W} \subseteq \mathbb{R}^q$*, i.e., for the sequence* $\{\mu_i\}_{i=1}^{\infty}$*,* $\mu_i \to \mu$ *if and only if* $W(\mu_i, \mu) \to 0$*.*

A more general version of this result is given in Villani (2009, Theorem 6.9). Thus, weak convergence and convergence of the Wasserstein metric are equivalent on compact subsets of the reals. Moreover, the Wasserstein metric may be used to quantify weak convergence (Fournier and Guillin, 2015). In particular, this result demonstrates that the choice of Wasserstein metric for the subsequent analysis is appropriate.

4.3 Distributional robustness of closed-loop systems

We now use the Wasserstein metric to define distribution robustness for closed-loop nonlinear systems. First, we recall the definition of robustness positive invariance and modify this definition slightly to account for the new notation used in the SMPC formulation.

Definition 4.12 (Robust positive invariance)**.** The set $\mathcal{X}$ is robustly positive invariant (RPI) for the system $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$ if $x \in \mathcal{X}$ implies that $x^+ \in \mathcal{X}$ for all $w \in \mathbb{W}$ and $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$.

Thus, we consider all possible $w \in \mathbb{W}$ in the true system and all probability measures that may be used in the SMPC formulation $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$ to define the control law $\kappa_{\hat{\mu}}(\cdot)$. We now define distributional robustness for closed-loop nonlinear systems.

Definition 4.13 (Distributionally robust asymptotic stability in expectation)**.** The origin is distributionally robustly asymptotically stable in expectation (DRASiE) for the system $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$ in the RPI set $\mathcal{X}$ if there exist $\beta(\cdot) \in \mathcal{KL}$ and $\gamma_1(\cdot), \gamma_2(\cdot) \in \mathcal{K}$ such that

$$\mathbb{E}\left[|\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)|\right] \leq \beta(|x|, k) + \gamma_1\left(\hat{\mathbb{E}}[|\hat{w}|]\right) + \gamma_2\left(W(\mu, \hat{\mu})\right) \tag{4.7}$$

for all $x \in \mathcal{X}$, $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$.

Note that the closed-loop trajectory on the left-hand side of (4.7) is a function of $\hat{\mu}$ via the control law defined by the SMPC problem and μ via the definition of expected value. As with every other definition of the robustness in this dissertation, the $\mathcal{KL}$ function in the upper bound ensures that the effect off the initial condition $x \in \mathcal{X}$ asymptotically vanishes as $k \to \infty$. However, we now have two persistent terms in the upper bound.

The function $\gamma_1(\hat{\mathbb{E}}[|\hat{w}|])$ accounts for the effect of the modeled disturbance ($\hat{w}$) in the control law design and the ideal system (if $\mu = \hat{\mu}$). Note that $\hat{\mathbb{E}}[|\hat{w}|]$ is a function of $\hat{\mu}$. If $\hat{\mathbb{E}}[\hat{w}] = 0$, i.e., the model of the disturbance distribution is zero mean, we can replace $\hat{\mathbb{E}}[|\hat{w}|]$ with the upper bound $\text{tr}(\hat{\Sigma})^{1/2}$ as per Remark 4.3.

The function $\gamma_2(W(\mu, \hat{\mu}))$ accounts for the discrepancy between the probability measure $\hat{\mu}$ used in the SMPC optimization problem and the probability measure of the underlying closed-loop system μ. If $\mu = \hat{\mu}$, then $\gamma_2(W(\mu, \hat{\mu})) = 0$ and we recover the usual bound for idealized SMPC shown in Definition 3.9 for zero mean disturbance distributions. The bound in (4.7) ensures that arbitrarily small differences between $\hat{\mu}$ and μ, in terms of the Wasserstein distance, produce similarly small deviations from the closed-loop bound derived for idealized SMPC. We further discuss the implications of DRASiE for SMPC in Section 4.5.

We can also define DRASiE with respect to the stage cost $\ell(\cdot)$ as follows.

Definition 4.14 (ℓ-DRASiE)**.** The origin is ℓ-DRASiE with respect to the stage cost $\ell(x, \kappa_{\hat{\mu}}(x))$ for the system $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$ in the RPI set $\mathcal{X}$ if there exist $\tilde{\beta}(\cdot) \in \mathcal{KL}$ and $\tilde{\gamma}_1(\cdot), \tilde{\gamma}_2(\cdot) \in \mathcal{K}$ such that

$$\mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] \leq \tilde{\beta}(|x|, k) + \tilde{\gamma}_1\left(\hat{\mathbb{E}}[|\hat{w}|]\right) + \tilde{\gamma}_2\left(W(\mu, \hat{\mu})\right) \tag{4.8}$$

in which $x(k) := \phi_{\hat{\mu}}(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$, $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$.

We now further modify the SISS Lyapunov function in Definition 3.11 to serve as a sufficient condition for DRASiE and ℓ-DRASiE.

Definition 4.15 (SISS Lyapunov function)**.** The measurable function $V_{\hat{\mu}} : \mathcal{X} \to \mathbb{R}_{\geq 0}$, defined for all $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, is an SISS Lyapunov function for the system $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$ in the RPI set $\mathcal{X}$ if there exist $\alpha_1(\cdot), \alpha_2(\cdot), \alpha_3(\cdot) \in \mathcal{K}_\infty$ and $\sigma_2(\cdot), \sigma_3(\cdot), \sigma_4(\cdot) \in \mathcal{K}$ such that

$$\alpha_1(|x|) \leq V_{\hat{\mu}}(x) \leq \alpha_2(|x|) + \sigma_2\left(\hat{\mathbb{E}}[|\hat{w}|]\right) \tag{4.9}$$

$$\int_{\mathbb{W}} V_{\hat{\mu}}(f(x, \kappa_{\hat{\mu}}(x), w)) d\mu(w) \leq V_{\hat{\mu}}(x) - \alpha_3(|x|) + \sigma_3\left(\hat{\mathbb{E}}[|\hat{w}|]\right) + \sigma_4(W(\mu, \hat{\mu})) \tag{4.10}$$

for all $x \in \mathcal{X}$, $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, and $\mu \in \mathcal{M}(\mathbb{W})$.

We then use this SISS Lyapunov function to establish DRASiE for a closed-loop system.

Proposition 4.16. *If a system $x^+ = f(x, \kappa_{\hat{\mu}}, w)$, $w \in \mathbb{W}$ admits an SISS Lyapunov function in the RPI and bounded set $\mathcal{X}$, then the origin is DRASiE.*

Since the proof is similar to Proposition 3.12, we provide only an outline of the proof in this chapter. The full proof is available in the Appendix of McAllister and Rawlings (2022e).

Outline of proof. The proof proceeds in a similar manner to the proof of Proposition 3.12 except we also carry the term $\sigma_4(W(\mu, \hat{\mu}))$ through the same operations as $\sigma_3(\hat{\mathbb{E}}[|\hat{w}|])$. Specifically, we use the same first steps as the proof of Proposition 3.12 to show

$$\int_{\mathbb{W}} V_{\hat{\mu}}(f(x, \kappa_{\hat{\mu}}(x), w)) d\mu(w) \leq V_{\hat{\mu}}(x) - \alpha_v(V_{\hat{\mu}}(x)) + \tilde{\sigma}_3\left(\hat{\mathbb{E}}[|\hat{w}|]\right) + \sigma_4(W(\mu, \hat{\mu}))$$

in which $\alpha_v(\cdot) \in \mathcal{K}_\infty$ is convex and $\tilde{\sigma}_3(s) := \sigma_3(s) + \alpha_3(\alpha_2^{-1}(\sigma_2(s)))$ (compare with (3.12)). We then treat the quantity $c(\mu, \hat{\mu}) := \tilde{\sigma}_3\left(\hat{\mathbb{E}}[|\hat{w}|]\right) + \sigma_4(W(\mu, \hat{\mu}))$ with all the same operations performed on $\sigma_3(\text{tr}(\Sigma))$ in the proof of Proposition 3.12 to give

$$\mathbb{E}[|\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)|] \leq \beta_1(V_{\hat{\mu}}(x), k) + \tilde{\gamma}_2(c(\mu, \hat{\mu}))$$

with $\beta_1(\cdot) \in \mathcal{KL}$ and $\tilde{\gamma}_2(\cdot) \in \mathcal{K}$ (compare with (3.14)). We use the upper bound on $V_{\hat{\mu}}(x)$ to give

$$\mathbb{E}[|\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)|] \leq \beta(|x|, k) + \tilde{\gamma}_1(\hat{\mathbb{E}}[|\hat{w}|]) + \tilde{\gamma}_2(c(\mu, \hat{\mu}))$$

in which $\beta(s, k) := \beta_1(2\alpha_2(s), k) \in \mathcal{KL}$ and $\tilde{\gamma}_1(s) := \beta_1(2\sigma_2(s), 0) \in \mathcal{K}$. To complete the proof, we unpack $c(\mu, \hat{\mu})$ to give

$$\mathbb{E}[|\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)|] \leq \beta(|x|, k) + \gamma_1\left(\hat{\mathbb{E}}[|\hat{w}|]\right) + \gamma_2\left(W(\mu, \hat{\mu})\right)$$

in which $\gamma_1(s) := \tilde{\gamma}_1(s) + \tilde{\gamma}_2(2\tilde{\sigma}_3(s)) \in \mathcal{K}$ and $\gamma_2(s) := \tilde{\gamma}_2(2\sigma_4(s)) \in \mathcal{K}$. □

We can further use the SISS Lyapunov function to establish ℓ-DRASiE.

Proposition 4.17. *If a system $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$ admits an SISS Lyapunov function in the RPI and bounded $\mathcal{X}$ that satisfies $\ell(x, \kappa_{\hat{\mu}}(x)) \leq V_{\hat{\mu}}(x)$ for all $x \in \mathcal{X}$ and $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, then the origin is ℓ-DRASiE.*

Proof. We use the same approach as the proof of Proposition 3.13. From Proposition 4.16, there exist $\beta(\cdot) \in \mathcal{KL}$ and $\gamma_1(\cdot), \gamma_2(\cdot) \in \mathcal{K}$ such that (4.7) holds for all $x \in \mathcal{X}$, $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$,

$\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$. By Corollary 2.20, we construct concave $\alpha_c(\cdot) \in \mathcal{K}_\infty$ such that $\alpha_2(|x|) \leq \alpha_c(|x|)$ for all $x \in \mathcal{X}$. Thus, we have

$$\begin{aligned}
\mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] &\leq \mathbb{E}\left[V_{\hat{\mu}}(x(k))\right] \\
&\leq \mathbb{E}\left[\alpha_c(|x(k)|) + \sigma_2\left(\hat{\mathbb{E}}[|\hat{w}|]\right)\right] \\
&\leq \alpha_c(\mathbb{E}[|x(k)|]) + \sigma_2\left(\hat{\mathbb{E}}[|\hat{w}|]\right) \\
&\leq \alpha_c(2\beta(|x|, k)) + \alpha_c\left(2\gamma_1\left(\hat{\mathbb{E}}[|\hat{w}|]\right)\right) + \sigma_2\left(\hat{\mathbb{E}}[|\hat{w}|]\right) \\
&\qquad + \alpha_c\left(2\gamma_2\left(W(\mu, \hat{\mu})\right)\right)
\end{aligned}$$

in which $x(k) = \phi_{\hat{\mu}}(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$, $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$. Define $\tilde{\beta}(\cdot) := \alpha_c(2\beta(\cdot))$, $\tilde{\gamma}_1(\cdot) := \alpha_c(2\gamma_1(\cdot)) + \sigma_2(\cdot)$, and $\tilde{\gamma}_2(s) := \alpha_c(2\gamma_2(\cdot))$. Note that $\tilde{\beta}(\cdot) \in \mathcal{KL}$ and $\tilde{\gamma}_1(\cdot), \tilde{\gamma}_2(\cdot) \in \mathcal{K}$ to complete the proof. □

Thus, the ISS/SISS framework is able to accommodate this new definition of distributional robustness with minimal modifications, once again demonstrating the flexibility of this framework for nonlinear system analysis. We can now apply these results to SMPC.

4.4 Inherent distributional robustness of SMPC

The main result of this chapter is now stated.

Theorem 4.18. *Let Assumptions 4.1, 4.2 and 4.4 to 4.7 hold for fixed* $\hat{\mathbb{W}} \subseteq \mathbb{R}^q$. *Then there exists* $\delta > 0$ *such that for any set* $\mathbb{W} \subseteq \mathbb{R}^q$ *satisfying* $d_H(\mathbb{W}, , \hat{\mathbb{W}}) \leq \delta$ *and the system* $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$ *we have that:*

(i) The set $\mathcal{X}$ *is RPI.*

(ii) The origin is DRASiE in the set $\mathcal{X}$.

(iii) The origin is ℓ-DRASiE in the set $\mathcal{X}$.

Thus, for a sufficiently small difference between $\mathbb{W}$ and $\hat{\mathbb{W}}$ (in terms of the Hausdorff distance), the set $\mathcal{X}$ is RPI, i.e., the SMPC problem remains robustly recursively feasible, and the origin is both DRASiE and ℓ-DRASiE. To establish this result, we require the following intermediate lemmata. This first lemma is a general mathematical result for integrals of Lipschitz continuous functions.

Lemma 4.19. *For a Lipschitz continuous function $g : X \times S \to \mathbb{R}$ with $X \subseteq \mathbb{R}^n$ and $S \subseteq \mathbb{R}^q$, $G(x) := \int_S g(x,s)d\mu(s)$ is also a Lipschitz continuous function with the same Lipschitz constant for all $\mu \in \mathcal{M}(S)$.*

Proof. Let $L \geq 0$ denote the Lipschitz constant for $g(\cdot)$. For $x_1, x_2 \in X$, we have

$$\begin{aligned}|G(x_1) - G(x_2)| &= \left| \int_S \Big(g(x_1,s) - g(x_2,s) \Big) d\mu(s) \right| \\ &\leq \int_S |g(x_1,s) - g(x_2,s)|\, d\mu(s) \\ &\leq \int_S L|x_1 - x_2| d\mu(w) = L|x_1 - x_2|\end{aligned}$$

in which the last equality holds because $\mu \in \mathcal{M}(S)$ is a probability measure, i.e., $\int_S d\mu(s) = 1$. Thus, $G(\cdot)$ is a Lipschitz continuous function with the same Lipschitz constant as $g(\cdot)$ for all $\mu \in \mathcal{M}(S)$. □

The second lemma provides an upper bound in the terminal region. This results leverages the Lipschitz continuity of the system model, terminal control law, and terminal cost to construct a bound similar to Lemma 3.14.

Lemma 4.20. *Let Assumptions 4.2 and 4.4 to 4.6 hold. There exists exists $L_f \geq 0$ such that*

$$\int_{\hat{\mathbb{W}}} V_f(f(x, \kappa_f(x), \hat{w}))d\hat{\mu}(\hat{w}) \leq V_f(x) - \ell(x, \kappa_f(x)) + L_f \mathbb{E}\left[|\hat{w}|\right] \tag{4.11}$$

for all $x \in \mathbb{X}_f$ *and* $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$.

Proof. Since $V_f(\cdot)$, $f(\cdot)$, $\kappa_f(\cdot)$ are locally Lipschitz continuous and $\mathbb{X}_f$, $\hat{\mathbb{W}}$ are bounded, there exists $L_f \geq 0$ such that

$$|V_f(f(x, \kappa_f(x), \hat{w})) - V_f(f(x, \kappa_f(x), 0))| \leq L_f |\hat{w}|$$

for all $x \in \mathbb{X}_f$ and $\hat{w} \in \hat{\mathbb{W}}$. Therefore,

$$\int_{\hat{\mathbb{W}}} V_f(f(x, \kappa_f(x), \hat{w})) d\hat{\mu}(\hat{w}) \leq V_f(f(x, \kappa_f(x), 0)) + L_f \hat{\mathbb{E}}[|\hat{w}|]$$

We combine this bound with (4.6) to give (4.11). □

We then establish an upper bound for the optimal cost function.

Lemma 4.21. *Let Assumptions 4.2 and 4.4 to 4.6 hold. Then there exist* $\alpha_2(\cdot) \in \mathcal{K}_\infty$ *and* $\sigma_2(\cdot) \in \mathcal{K}$ *such that* $V^0_{\hat{\mu}}(x) \leq \alpha_2(|x|) + \sigma_2(\hat{\mathbb{E}}[|\hat{w}|])$ *for all* $x \in \mathcal{X}$ *and* $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$.

The proof of this result is nearly identical to the proof of Lemma 3.17 and therefore omitted. The full proof can be found in the Appendix of McAllister and Rawlings (2022e, Lemma 20).

We can now establish Theorem 4.18.

Proof of Theorem 4.18. We first establish that there exists $\delta > 0$ such that $\mathcal{X}$ is RPI. Since $f(\cdot)$, $\pi(\cdot)$ are locally Lipschitz continuous and $\mathcal{X}$ is bounded, there exists $L_x > 0$ such that

$$|f(x, \pi(x, v), w) - f(x, \pi(x, v), \hat{w})| \leq L_x |w - \hat{w}|$$

for all $x \in \mathcal{X}$, $v \in \mathbb{V}$, and $w, \hat{w} \in \mathbb{W}$. For $x \in \mathcal{X}$ and $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, choose $\mathbf{v}^0 \in \mathbf{v}^0_{\hat{\mu}}(x)$ such that $\kappa_{\hat{\mu}}(x) = \pi(x, v^0(0))$, $\hat{\mathbf{w}} \in \hat{\mathbb{W}}^N$, and define

$$\tilde{\mathbf{v}}^+ := (v^0(1), v^0(2), \dots, v^0(N-1), 0)$$
$$\tilde{\mathbf{w}}^+ := (\hat{w}(1), \hat{w}(2), \dots, \hat{w}(N-1), \hat{w}(N))$$

for some $\hat{w}(N) \in \hat{\mathbb{W}}$. We denote $x^+(w) = f(x, \kappa_{\hat{\mu}}(x), w)$, $x(N) = \hat{\phi}(N, x, \mathbf{v}^0, \hat{\mathbf{w}})$, and $x^+(N; w) = \hat{\phi}(N; x^+(w), \tilde{\mathbf{v}}^+, \tilde{\mathbf{w}}^+)$. Note that $x \in \mathcal{X}$, $w \in \mathbb{W}$, $\tilde{\mathbf{v}}^+ \in \mathbb{V}^N$, and $\tilde{\mathbf{w}}^+ \in \hat{\mathbb{W}}^N$ are all bounded.

The function $\hat{\phi}(N; \cdot)$ is locally Lipschitz continuous since it is a composition of a finite number of locally Lipschitz continuous functions. Therefore, $V_f(\hat{\phi}(N; \cdot))$ is also locally Lipschitz continuous and there exists $\tilde{L}_f > 0$ such that

$$\begin{aligned} V_f(x^+(N; w)) - V_f(x^+(N; \hat{w})) &\leq |V_f(x^+(N; w)) - V_f(x^+(N; \hat{w}))| \\ &\leq \tilde{L}_f |x^+(w) - x^+(\hat{w})| \\ &\leq \tilde{L}_f L_x |w - \hat{w}| \end{aligned}$$

for all $x \in \mathcal{X}$, $w, \hat{w} \in \mathbb{W}$, $\tilde{\mathbf{v}}^+ \in \mathbb{V}^N$, and $\tilde{\mathbf{w}}^+ \in \hat{\mathbb{W}}^N$. Since $x(N) \in \mathbb{X}_f$, we have from Assumption 4.6 that $V_f(x^+(N; \hat{w})) \leq \tilde{\tau}$ for all $\hat{w} \in \hat{\mathbb{W}}$ and therefore,

$$V_f(x^+(N; w)) \leq \tilde{\tau} + \tilde{L}_f L_x |w - \hat{w}|$$

for all $w \in \mathbb{W}$ and $\hat{w} \in \hat{\mathbb{W}}$. Thus, for any $w \in \mathbb{W}$, we can choose $\hat{w} \in \hat{\mathbb{W}}$ to minimize the value of $|w - \hat{w}|$ and we have that

$$V_f(x^+(N; w)) \leq \tilde{\tau} + \tilde{L}_f L_x |w|_{\hat{\mathbb{W}}}$$

for all $w \in \mathbb{W}$. We define

$$\delta := \frac{\tau - \tilde{\tau}}{\tilde{L}_f L_x} \tag{4.12}$$

and note that $\delta > 0$. Thus, for all sets $\mathbb{W}$ such that $d_H(\mathbb{W}, \hat{\mathbb{W}}) \leq \delta$, we have that $|w|_{\hat{\mathbb{W}}} \leq \delta$ and therefore $V_f(x^+(N; w)) \leq \tau$. Hence, $x \in \mathcal{X}$ implies that $x^+(N; x) \in \mathbb{X}_f$, $\tilde{\mathbf{v}}^+ \in \mathcal{V}(x^+(w))$, and $x^+(w) \in \mathcal{X}$. Recall that $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$ and $x \in \mathcal{X}$ were chosen arbitrarily. Thus, $\mathcal{X}$ is RPI for the closed-loop system $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$ for any $\mathbb{W} \subseteq \mathbb{R}^q$ such that $d_H(\mathbb{W}, \hat{\mathbb{W}}) \leq \delta$ and we have established (i).

We now establish an expected cost decrease condition for the probability measure $\hat{\mu}$ similar to Lemma 3.16. Using the definition of $J(\cdot)$, we have

$$J(x^+(\hat{w}), \tilde{\mathbf{v}}^+, \tilde{\mathbf{w}}^+) = J(x, \mathbf{v}^0, \tilde{\mathbf{w}}) - \ell(x, \kappa_{\hat{\mu}}(x)) + \eta(x(N), \hat{w}(N)) \tag{4.13}$$

in which

$$\eta(x, w) := -V_f(x) + \ell(x, \kappa_f(x)) + V_f(f(x, \kappa_f(x), w))$$

From Lemma 4.20 and the fact that $x(N) \in \mathbb{X}_f$, there exists $L_f > 0$ such that

$$\int_{\hat{\mathbb{W}}^{N+1}} \eta(x(N), \hat{w}(N)) d\hat{\mu}^N(\hat{\mathbf{w}}) d\hat{\mu}(\hat{w}(N)) \leq L_f \hat{\mathbb{E}}[|\hat{w}|]$$

We also have the equality

$$V^0_{\hat{\mu}}(x) = \int_{\hat{\mathbb{W}}^{N+1}} J(x, \mathbf{v}^0, \hat{\mathbf{w}}) d\hat{\mu}^N(\hat{\mathbf{w}}) d\hat{\mu}(\hat{w}(N))$$

and therefore

$$\int_{\mathbb{W}} V_{\hat{\mu}}(x^+(\hat{w}), \tilde{\mathbf{v}}^+) d\hat{\mu}(\hat{w}) = \int_{\hat{\mathbb{W}}^{N+1}} J(x^+(\hat{w}), \tilde{\mathbf{v}}^+, \tilde{\mathbf{w}}^+) d\hat{\mu}^N(\hat{\mathbf{w}}) d\hat{\mu}(\hat{w}(N))$$
$$\leq V^0_{\hat{\mu}}(x) - \ell(x, \kappa_{\hat{\mu}}(x)) + L_f \hat{\mathbb{E}}[|\hat{w}|] \tag{4.14}$$

in which we can exchange $\hat{\mathbb{W}}$ with $\mathbb{W}$ for the domain of integration since Assumption 4.2 ensures that $\hat{\mu}(\mathbb{W} \setminus \hat{\mathbb{W}}) = 0$.

Next, we use Theorem 4.9 to exchange $\hat{\mu}$ for μ. The function $J(x, \mathbf{v}, \mathbf{w})$ is the composition of a finite number of locally Lipschitz continuous functions and is therefore locally Lipschitz continuous. Thus, $J(x, \mathbf{v}, \mathbf{w})$ is Lipschitz continuous on the compact set $\mathcal{X} \times \mathbb{V}^N \times \hat{\mathbb{W}}^N$. From Lemma 4.19, we have the $V_{\hat{\mu}}(x, \mathbf{v})$ is also Lipschitz continuous with the same Lipschitz constant for all $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$. Thus, there exists $L_J > 0$ such that

$$|V_{\hat{\mu}}(f(x, u, w), \tilde{\mathbf{v}}^+) - V_{\hat{\mu}}(f(x, u, \hat{w}), \tilde{\mathbf{v}}^+)| \leq L_J |w - \hat{w}|$$

for all $w, \hat{w} \in \mathbb{W}$, $x \in \mathcal{X}$, $u \in \mathbb{U}$, $\tilde{\mathbf{v}}^+ \in \mathbb{V}^N$, and $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$. We choose arbitrary $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$ and use Theorem 4.9 to give

$$\int_{\mathbb{W}} V_{\hat{\mu}}(x^+(w), \tilde{\mathbf{v}}^+) d\mu(w) \leq \int_{\mathbb{W}} V_{\hat{\mu}}(x^+(\hat{w}), \tilde{\mathbf{v}}^+) d\hat{\mu}(\hat{w}) + L_J W(\mu, \hat{\mu}) \tag{4.15}$$

for all $x \in \mathcal{X}$, $\tilde{\mathbf{v}}^+ \in \mathbb{V}^N$, $\mu \in \mathcal{M}(\mathbb{W})$. Note that the choice of $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$ was arbitrary and (4.15) holds for all $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$ with the same value of $L_J > 0$. We combine (4.14) and (4.15) and by optimality we have

$$\int_{\mathbb{W}} V^0_{\hat{\mu}}(x^+(w)) d\mu(w) \leq V^0_{\hat{\mu}}(x) - \ell(x, \kappa_{\hat{\mu}}(x)) + L_f \hat{\mathbb{E}}[|\hat{w}|] + L_J W(\mu, \hat{\mu}) \tag{4.16}$$

for all $x \in \mathcal{X}$, $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, and $\mu \in \mathcal{M}(\mathbb{W})$.

We now establish that $V^0_{\hat{\mu}}(\cdot)$ is an SISS Lyapunov function according to Definition 4.15. From Assumption 4.7, there exists $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$ such that $-\ell(x,\kappa_{\hat{\mu}}(x)) \leq -\alpha_\ell(|x|)$ and we substitute this inequality in to (4.16). Therefore, (4.10) holds with $V_{\hat{\mu}}(\cdot) := V^0_{\hat{\mu}}(\cdot)$, $\alpha_3(\cdot) := \alpha_\ell(\cdot)$, $\sigma_3(\cdot) := L_f s$, and $\sigma_4(s) := L_J s$. We also use Assumption 4.7 to show that $\alpha_1(|x|) := \alpha_\ell(|x|) \leq \ell(x,\kappa_{\hat{\mu}}(x)) \leq V^0_{\hat{\mu}}(x)$ for all $x \in \mathcal{X}$. With Lemma 4.21, we can construct the desired upper bound for $V^0_{\hat{\mu}}(\cdot)$. Thus, $V^0_{\hat{\mu}}(\cdot)$ satisfies the requirements in Definition 4.15 for an SISS Lyapunov function. By Propositions 4.16 and 4.17, the origin is DRASiE and ℓ-DRASiE. □

As noted in the previous chapter, an important class of applications for SMPC are economic problems in which the stage cost is chosen to directly represent an economic, environmental, or performance metric for the process (e.g., production cost, carbon production). If this cost satisfies Assumption 4.7, then Theorem 4.18 still applies and ℓ-DRASiE is particularly relevant. In Chapters 2 and 3, we exclusively considered MPC and SMPC formulations that satisfy this requirement for the stage cost. Unfortunately, this requirement restricts the space of economic cost functions that we may consider with SMPC and can exclude many relevant problems. Thus, in economic applications of MPC, i.e., economic MPC, we often drop Assumption 4.7 and instead analyze the closed-loop system without this assumption. By removing this assumption, we obtain a weaker, but still instructive, result for economic applications of SMPC.

Theorem 4.22. *Let Assumptions 4.1, 4.2 and 4.4 to 4.6 hold. Then there exists $\delta > 0$ such that for any $\mathbb{W} \subseteq \mathbb{R}^q$ satisfying $d_H(\mathbb{W},\hat{\mathbb{W}}) \leq \delta$ and the system $x^+ = f(x,\kappa_{\hat{\mu}}(x),w)$, $w \in \mathbb{W}$ we have that $\mathcal{X}$ is RPI. Furthermore, there exist $L_f, L_J > 0$ such that the closed-loop trajectory satisfies*

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} \mathbb{E}\left[\ell(x(k),\kappa_{\hat{\mu}}(x(k)))\right] \leq L_f \hat{\mathbb{E}}[|\hat{w}|] + L_J W(\mu,\hat{\mu}) \tag{4.17}$$

in which $x(k) = \phi_{\hat{\mu}}(k;x,\mathbf{w}_k)$ for all $x \in \mathcal{X}$, $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, and $\mu \in \mathcal{M}(\mathbb{W})$.

Proof. In the proof of Theorem 4.18, we establish that $\mathcal{X}$ is RPI and that the bound in (4.16) holds without using Assumption 4.7. We choose $x \in \mathcal{X}$, $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$ and denote the closed-loop trajectory $x(k) = \phi_{\hat{\mu}}(k; x, \mathbf{w}_k)$. We then apply the law of total expectation to (4.16) and rearrange to give,

$$\mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] \leq \mathbb{E}\left[V_{\hat{\mu}}^0(x(k))\right] - \mathbb{E}\left[V_{\hat{\mu}}^0(x(k+1))\right] + L_f\hat{\mathbb{E}}[|\hat{w}|] + L_J W(\mu, \hat{\mu})$$

We sum both sides of this inequality from $k = 0$ to $T - 1$, cancel terms, and divide by T to give,

$$\frac{1}{T}\sum_{k=0}^{T-1}\mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] \leq \frac{V_{\hat{\mu}}^0(x) - \mathbb{E}\left[V_{\hat{\mu}}^0(x(T))\right]}{T} + L_f\hat{\mathbb{E}}[|\hat{w}|] + L_J W(\mu, \hat{\mu})$$

Since $V_{\hat{\mu}}(x, \mathbf{v})$ is Lipschitz continuous on the bounded set $\mathcal{X} \times \mathbb{V}^N$ for all $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$, we have that $V_{\hat{\mu}}^0(x)$ is bounded for all $x \in \mathcal{X}$ and $\hat{\mu} \in \hat{\mathcal{M}}(\mathbb{W})$. Therefore, $\mathbb{E}[V_{\hat{\mu}}^0(x(T))]$ is also bounded since $x(T) \in \mathcal{X}$. We take the limit $T \to \infty$ and note that $V_{\hat{\mu}}^0(x)/T$ and $\mathbb{E}\left[V_{\hat{\mu}}^0(x(T))\right]/T$ to give (4.17). □

The inequality in (4.17) ensures that as $T \to \infty$, the time-averaged expected value of the stage cost is upper bounded by a constant proportional to $\hat{\mathbb{E}}[|\hat{w}|]$ and $W(\mu, \hat{\mu})$. Note that we use the limit supremum ($\limsup$) because the limit of left-hand side of this equation may not exist. Analogous to ℓ-DRASiE, the effect of the initial condition vanishes as $T \to \infty$, but (4.17) does not enforce any properties on the initial transient behavior of the closed-loop trajectory (other than the fact that $\mathcal{X}$ is RPI). By contrast, the bound in (4.8) must hold for all $k \in \mathbb{I}_{\geq 0}$. The bound in (4.17) reduces to a standard result for idealized SMPC that was first derived for nonlinear systems by Chatterjee and Lygeros (2014). This result for idealized SMPC has become a somewhat standard closed-loop property and is often reestablished for new SMPC formulations to ensure "stability" of the proposed formulation. Similar results are also available for (idealized) robust MPC formulations (Bayer et al., 2016).

4.5 Discussion

We now discuss several insights derived from these results. This discussion addresses (i) the robustness of SMPC to incorrectly or unmodeled disturbances, (ii) scenario optimization as a means to approximate and solve SMPC optimization problems, and (iii) the unification of stochastic robustness results across different MPC formulations.

4.5.1 SMPC

Idealized SMPC For idealized SMPC, we have that $\mu = \hat{\mu}$ and $\mathbb{W} = \hat{\mathbb{W}}$. Under these conditions, $W(\mu, \hat{\mu}) = 0$, $d_H(\mathbb{W}, \hat{\mathbb{W}}) = 0$, and (4.7), (4.8) and (4.17) reduce to their idealized SMPC counterparts, i.e., we have

$$\mathbb{E}\left[|\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)|\right] \leq \beta(|x|, k) + \gamma_1\left(\mathbb{E}[|w|]\right) \quad \forall k \in \mathbb{I}_{\geq 0} \tag{4.18}$$

$$\mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] \leq \tilde{\beta}(|x|, k) + \tilde{\gamma}_1\left(\mathbb{E}[|w|]\right) \quad \forall k \in \mathbb{I}_{\geq 0} \tag{4.19}$$

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] \leq L_f \mathbb{E}[|w|] \tag{4.20}$$

for all $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathbb{W})$.

Incorrectly modeled disturbances We assume that $\mathbb{W} = \hat{\mathbb{W}}$, but the distribution is incorrect, i.e., $\mu \neq \hat{\mu}$. In this case, $d_H(\mathbb{W}, \hat{\mathbb{W}}) = 0$ and robust recursive feasibility is guaranteed ($\mathcal{X}$ is RPI). The closed-loop performance bounds with respect to $\mathbb{E}[|x(k)|]$ or $\mathbb{E}[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))]$ degrade relative to the idealized case with respect to the distance between $\hat{\mu}$ and μ, i.e, $W(\mu, \hat{\mu})$. Thus, as $\hat{\mu} \to \mu$ we recover the idealized SMPC bounds. Furthermore, arbitrarily small differences between $\hat{\mu}$ and μ, in terms of the Wasserstein metric, produce similarly small deviations in these bounds. Note that selecting $\hat{\mathbb{W}}$ larger than the support of μ, i.e., $\mu(\hat{\mathbb{W}}) = 1$, still ensures that $\hat{\mathbb{W}} = \mathbb{W}$ because we can increase the size of $\mathbb{W}$ and assign these values zero measure. Thus, designing for a larger set of disturbances than the system encounters does not affect feasibility. This additional conservatism, however, may reduce performance since $W(\hat{\mu}, \mu) > 0$ for $\hat{\mu} \neq \mu$ and can also reduce the size of the feasible set $\mathcal{X}$. Furthermore, there is an implicit relationship between the terminal constraint and $\hat{\mathbb{W}}$ via Assumption 4.6. For a sufficiently large set $\hat{\mathbb{W}}$, there may not exist any $\kappa_f(\cdot)$, $\mathbb{X}_f$, and $V_f(\cdot)$ that satisfy Assumption 4.6.

Unmodeled disturbances In this case, we have that $\mathbb{W} \neq \hat{\mathbb{W}}$ and $\mu \neq \hat{\mu}$. This case represents disturbances that were "undermodeled" in that $\mathbb{W}$ is larger than $\hat{\mathbb{W}}$, e.g., $\hat{\mathbb{W}} = \{w \in \mathbb{R}^q : |w| \leq 1\}$, but $\mathbb{W} = \{w \in \mathbb{R}^q : |w| \leq 2\}$ and $\mu(\hat{\mathbb{W}}) < 1$. This case also covers elements or directions of $\mathbb{W}$ that are entirely absent in $\hat{\mathbb{W}}$, e.g., $\hat{\mathbb{W}} = \{w \in \mathbb{R}^q : |w| \leq 1 \text{ and } w_1 = 0\}$ or $\hat{\mathbb{W}} = \{0\}$, but $\mathbb{W} := \{w \in \mathbb{R}^q : |w| \leq 1\}$ with $\mu(\hat{\mathbb{W}}) < 1$. In fact, this representation can also account for an error in the model $f(\cdot)$. For example, we consider the model $\hat{f}(x, u, w_1)$, but the actual system evolves according to $f(x, u, w) = \hat{f}(x, u + w_2, 0) + w_3$ in which $w = (w_1, w_2, w_3)$. Nonetheless, we may still define $\hat{\mathbb{W}} = \{w \in \mathbb{R}^3 : |w_1| \leq 1,\ w_2 = 0,\ w_3 = 0\}$ and $\mathbb{W} = \{w \in \mathbb{R}^3 : |w_1| \leq 1,\ |w_2| \leq 1,\ |w_3| \leq 1\}$. Thus, the SMPC optimization problem considers only disturbances in w_1, while $w_2 = w_3 = 0$. Provided $\delta \geq \sqrt{2}$ for this problem formulation, we have that $d_H(\mathbb{W}, \hat{\mathbb{W}}) = \sqrt{2} \leq \delta$ and therefore $\mathcal{X}$ is RPI and (4.7), (4.8) and (4.17) hold despite the significant mismodeling of the disturbance in this example.

4.5.2 Scenario optimization

For nonlinear systems, the stochastic optimization problem proposed in (4.3) is typically intractable for continuous probability measures, e.g., a truncated normal distribution. Instead, scenario optimization methods are often used to approximate and solve the stochastic optimization problem. By selecting a finite set of possible scenarios from the underlying disturbance set and probability measure, the expected cost of the stochastic optimization problem is approximated by the average cost of these scenarios. The constraints are then required to hold for all scenarios considered. Performance bounds for the optimal solution/cost of stochastic optimization problems using this scenario optimization method is a topic that has generated much interest, with applications beyond SMPC. For scenario approximations to robust and chance-constrained optimization problems, one can show probabilistic guarantees, e.g., the solution to the approximate problem is within some confidence interval of the solution to the origin problem (Calafiore and Campi, 2006; Esfahani et al., 2014). This confidence interval shrinks as the number of samples increases, but still allows for a nonzero probability that the constraints of the original problem are violated.

The quality of this approximation, however, is irrelevant for SMPC if near exact approximations still produce poor controllers. Thus, the contribution in this chapter is novel because we are discussing the performance of the *closed-loop system* generated by repeated solutions to this approximated optimization problem and not the quality of each individual approximation. We can thereby bound the performance of the controller subject to a scenario-based approximation of the stochastic optimization problem.

We proceed by defining the stochastic optimization problem in (4.3) with an empirical distribution generated via a scenario-based approximation of the original stochastic optimal control problem. Thus, we can analyze the scenario approximation error as an additional error in representing the disturbance set and probability measure for the plant. Specifically, we construct this scenario optimization problem by drawing $s \in \mathbb{I}_{\geq 1}$ samples $\hat{\omega}_i$ from the set $\hat{\mathbb{W}}$ and probability measure $\hat{\mu}$. We then define the set $\hat{\mathbb{W}}_s := \{\hat{\omega}_1, \hat{\omega}_2, \ldots, \hat{\omega}_s\}$ and the empirical probability measure

$$\hat{\mu}_s(\cdot) := \frac{1}{s} \sum_{i=1}^{s} \delta_{\hat{\omega}_i}(\cdot)$$

Note that $\hat{\mathbb{W}}_s$ and $\hat{\mu}_s(\cdot)$ satisfy the requirements in Assumption 4.2 for all $s \in \mathbb{I}_{\geq 1}$. Moreover, if Assumption 4.6 holds for $\hat{\mathbb{W}}$, Assumption 4.6 also holds for $\hat{\mathbb{W}}_s \subseteq \hat{\mathbb{W}}$. Thus, we may use $\hat{\mathbb{W}}_s$ and $\hat{\mu}_s(\cdot)$ in place of $\hat{\mathbb{W}}$ and $\hat{\mu}$ for all algorithms and results in this work. In particular, we have that Theorems 4.18 and 4.22 hold for $\hat{\mathbb{W}}_s$ and $\hat{\mu}_s(\cdot)$ in place of $\hat{\mathbb{W}}$ and $\hat{\mu}$.

We can therefore draw several important conclusions for scenario optimization as a means to approximate the SMPC optimization problem. First, if $\hat{\mathbb{W}}$ is sufficiently close to $\mathbb{W}$ ($d_H(\mathbb{W}, \hat{\mathbb{W}}) \leq \delta/2$) and the sampling of $\hat{\mathbb{W}}$ is sufficiently dense ($d_H(\hat{\mathbb{W}}, \hat{\mathbb{W}}_s) \leq \delta/2$), then $\mathcal{X}$ is RPI ($d_H(\mathbb{W}, \hat{\mathbb{W}}_s) \leq \delta$). This observation suggests that sampling-based approximations can ensure robust recursive feasibility of SMPC if a sufficient number of samples are used. In fact, deliberate construction of $\hat{\mathbb{W}}_s$ to ensure that the approximation is sufficiently dense on $\hat{\mathbb{W}}$ may be preferable to constructing $\hat{\mathbb{W}}_s$ via random sampling. If $\hat{\mathbb{W}}_s$ is constructed with random sampling, $d_H(\hat{\mathbb{W}}, \hat{\mathbb{W}}_s)$ may take larger values with some small probability. This result is a significant departure from the usual results for scenario-based approximations because it implies that there is a finite number of samples, if chosen deliberately, that ensure a deterministic property of the closed-loop system, i.e., $\mathcal{X}$ is RPI. Recall that recursive feasibility is essential for industrial implementation of any optimization-based controller.

Second, the performance is bounded by the distance between $\hat{\mu}_s$ and μ. Since the Wasserstein metric is a proper metric, we can use the triangle inequality to show that

$$W(\mu, \hat{\mu}_s) \leq W(\mu, \hat{\mu}) + W(\hat{\mu}, \hat{\mu}_s)$$

As the number of samples increases, i.e., $s \to \infty$, we can further establish that

$$W(\mu, \hat{\mu}_s) \to W(\mu, \hat{\mu})$$

Thus, the performance bounds in (4.7), (4.8) and (4.17) converge to their values for the original stochastic optimization problem as the number of samples increases. Moreover, we can characterize this convergence of the Wasserstein metric in terms of the number of samples and the dimension of w via results available in Fournier and Guillin (2015). For example, we can show that for $q \geq 3$, there exists $C > 0$ such that

$$\mathbb{E}_s \left[W(\hat{\mu}, \hat{\mu}_s) \right] \leq C(s^{-1/q} + s^{-1})$$

in which $\mathbb{E}_s$ is the expected value evaluated over the random samples taken from $\hat{\mu}$ since $\hat{\mu}_s$ is constructed from s random variables $\hat{\omega}_i$.

The key conclusion from this discussion is that scenario-based approximations of SMPC optimization problems also benefit from the inherent distributional robustness afforded by feedback. For computational reasons, these scenario-based approximations are already and frequently used to solve nonlinear SMPC problems. The results presented here, however, provide a novel theoretical justification for this approximation in the context of closed-loop systems.

4.5.3 CMPC

The conclusions of Theorem 4.18 and Theorem 4.22 are particularly interesting given their ability to unify the notions of stochastic robustness across all SMPC, CMPC, and nominal MPC. We can treat both CMPC and nominal MPC as special cases of SMPC and therefore use these two theorems to draw conclusions about all of these MPC formulations. We discuss CMPC in this subsection and nominal MPC in the next subsection.

For $\hat{\mu}(\{0\}) = 1$, the SMPC formulation in (4.3) reduces to CMPC, in which we optimize a nominal objective function subject to tightened constraints, i.e.,

$$\min_{\mathbf{v} \in \mathcal{V}(x)} J(x, \mathbf{v}, \mathbf{0})$$

We therefore have $\hat{\mathbb{E}}[|\hat{w}|] = 0$ and the Wasserstein metric reduces to

$$W(\mu, \hat{\mu}) = \mathbb{E}\left[|w|\right]$$

and therefore (4.7), (4.8) and (4.17) can be simplified.

Corollary 4.23 (CMPC). *Let Assumptions 4.1, 4.2 and 4.4 to 4.6 hold with $\hat{\mu}(\{0\}) = 1$. Then there exists $\delta > 0$ such that for any set $\mathbb{W} \subseteq \mathbb{R}^q$ satisfying $d_H(\mathbb{W}, \hat{\mathbb{W}}) \leq \delta$, the set $\mathcal{X}$ is RPI for the system $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$. Furthermore, there exists $L_J > 0$ such that the closed-loop trajectory satisfies*

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] \leq L_J \mathbb{E}[|w|] \tag{4.21}$$

in which $x(k) = \phi_{\hat{\mu}}(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathbb{W})$. If Assumption 4.7 also holds, there exist $\beta(\cdot), \tilde{\beta}(\cdot) \in \mathcal{KL}$ and $\gamma_2(\cdot), \tilde{\gamma}_2 \in \mathcal{K}$ such that

$$\mathbb{E}\left[|\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)|\right] \leq \beta(|x|, k) + \gamma_2\left(\mathbb{E}[|w|]\right) \tag{4.22}$$

$$\mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] \leq \tilde{\beta}(|x|, k) + \tilde{\gamma}_2\left(\mathbb{E}[|w|]\right) \tag{4.23}$$

for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, and $k \in \mathbb{I}_{\geq 0}$.

Note that $\gamma_2(\cdot)$, $\tilde{\gamma}_2(\cdot)$, and L_J appear in (4.21) to (4.23), but $\gamma_1(\cdot)$, $\tilde{\gamma}_1(\cdot)$, and L_f appear in (4.18) to (4.20). This observation suggests, as discussed in Chapter 3, that the performance of CMPC and SMPC may differ quantitatively. The qualitative behavior, however, is likely similar for an otherwise equivalent problem, i.e., increases in $\mathbb{E}\left[|w|\right]$ produce increases in each of these upper bounds.

We now consider an interesting question: is the relationship between the constants L_f and L_J general under specific assumptions? For example, can we claim that $L_f \leq L_J$ for a specific class of problems? We note, however, that L_f and L_J are constructed via different routes. The constant L_f depends entirely on the terminal region, while L_J also includes effects from the stage cost and control law parameterization over the horizon N. Furthermore, these bounds are often too conservative to produce useful quantitative information, much like the $\mathcal{K}$ functions used in the other bounds for the closed-loop system. We therefore do not recommend calculating these bounds a priori as a means to assess the potential benefits of implementing SMPC over a nominal objective function. Simulation studies remain the best means to evaluate the quantitative benefits of SMPC for specific problems of interest. See, for example, Kumar et al. (2019)

4.5.4 Nominal MPC

For $\hat{\mu}(\{0\}) = 1$ and $\hat{\mathbb{W}} = \{0\}$, the SMPC problem reduces to nominal MPC in which we have the feedback law $\pi(x, v)$ embedded in the optimization problem. This type of parameterization, however, is not uncommon for nominal MPC formulations. By parameterizing the feedback law used in the nominal MPC problem, we can "pre-stabilize" the open-loop system and thereby ensure that the MPC optimization problem is well conditioned (Jerez et al., 2011; Rossiter et al., 1998). Moreover, if we choose $\pi(x, v) = v$ and $\mathbb{V} = \mathbb{U}$, the nominal MPC problem introduced in Chapter 2 is equivalent to this simplified SMPC problem. For this choice of $\pi(x, v) = v$ and $\mathbb{V} = \mathbb{U}$, we have that

$$\mathcal{V}(x) = \mathcal{U}(x) = \{u \in \mathbb{U}^N : \hat{\phi}(N; x, \mathbf{u}, \mathbf{0}) \in \mathbb{X}_f\}$$

and the optimization problem becomes

$$\min_{\mathbf{v} \in \mathcal{V}(x)} J(x, \mathbf{v}, \mathbf{0}) = \min_{u \in \mathcal{U}(x)} J(x, \mathbf{u}, \mathbf{0})$$

For this choice of $\pi(\cdot)$ and $\mathbb{V}$, Assumptions 4.4 and 4.5 reduces to their nominal MPC counterparts, Assumptions 2.1 and 2.28, with the requirement of local Lipschitz continuity added to Assumption 2.1 and bounded $\mathcal{X}$ added to Assumption 2.28. For Assumption 4.6, the requirement in (4.5) becomes

$$f(x, \kappa_f(x), 0) \in \{x \in \mathbb{R}^n : V_f(x) \leq \tilde{\tau}\} \tag{4.24}$$

Thus, the terminal control law must drive the subsequent state for the nominal system to the interior of $\mathbb{X}_f$ for all $x \in \mathbb{X}_f$. If Assumption 4.7 also holds, the nominal cost decrease condition in (4.6) combined with the definition of $\mathbb{X}_f := \{x \in \mathbb{R}^n : V_f(x) \leq \tau\}$ is sufficient to guarantee that (4.24) also holds for some $\tilde{\tau} < \tau$.

Lemma 4.24. *Let Assumptions 4.4, 4.5 and 4.7 hold. If there exists a terminal control law* $\kappa_f : \mathbb{X}_f \to \mathbb{U}$ *such that*

$$V_f(f(x, \kappa_f(x), 0)) \leq V_f(x) - \ell(x, \kappa_f(x))$$

for all $x \in \mathbb{X}_f$*, then there exists* $\tilde{\tau} < \tau$ *such that* (4.24) *holds for all* $x \in \mathbb{X}_f$*.*

Proof. We have that

$$V_f(f(x, \kappa_f(x), 0)) \leq V_f(x) - \alpha_\ell(|x|)$$

for all $x \in \mathbb{X}_f$. Since $V_f(\cdot)$ is Lipschitz continuous on the compact set $\mathbb{X}_f$ and $V_f(0) = 0$, there exists $L > 0$ such that $V_f(x) \leq L|x|$ for all $x \in \mathbb{X}_f$. We define

$$\tilde{\tau} := \tau - \min\{\tau/2, \alpha_\ell(\tau/(2L))\}$$

Note that $\tilde{\tau} < \tau$. If $V_f(x) \leq \tau/2$, we have that $V_f(f(x, \kappa_f(x), 0)) \leq \tau/2 \leq \tilde{\tau}$. If $\tau/2 \leq$

$V_f(x) \leq \tau$, then we have

$$V_f(f(x, \kappa_f(x), 0)) \leq \tau - \alpha_\ell(\tau/(2L)) \leq \tilde{\tau}$$

Thus, $V_f(f(x, \kappa_f(x), 0)) \leq \tilde{\tau}$ and (4.24) holds for all $x \in \mathbb{X}_f$. □

If Assumption 4.7 does not hold, however, the requirement in (4.24) is notably different than the requirements typically considered or used for economic MPC. As we discuss later in this subsection, this requirement allows us to derive new results for the robustness of nominal economic MPC.

With this problem formulation, we have that $d_H(\mathbb{W}, \{0\}) = \sup_{w\in\mathbb{W}} |w|$ and the bounds in (4.7), (4.8) and (4.17) reduce to the equations in (4.21) to (4.23), i.e., we obtain the same closed-loop bounds as CMPC. Thus, we can use Theorems 4.18 and 4.22 to establish the stochastic robustness of nominal MPC to sufficiently small disturbances.

Corollary 4.25 (Nominal MPC)**.** *Let Assumptions 4.1, 4.2 and 4.4 to 4.6 hold with $\hat{\mu}(\{0\}) = 1$ and $\hat{\mathbb{W}} = \{0\}$. Then there exists $\delta > 0$ such that for any set $\mathbb{W} \subseteq \{w \in \mathbb{R}^q : |w| \leq \delta\}$, the set $\mathcal{X}$ is RPI for the system $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$. Furthermore, there exists $L_J > 0$ such that the closed-loop trajectory satisfies*

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} \mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] \leq L_J \mathbb{E}[|w|] \tag{4.25}$$

in which $x(k) = \phi_{\hat{\mu}}(k; x, \mathbf{w}_k)$ for all $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathbb{W})$. If Assumption 4.7 also holds, there exist $\beta(\cdot), \tilde{\beta}(\cdot) \in \mathcal{KL}$ and $\gamma_2(\cdot), \tilde{\gamma}_2 \in \mathcal{K}$ such that

$$\mathbb{E}\left[|\phi_{\hat{\mu}}(k; x, \mathbf{w}_k)|\right] \leq \beta(|x|, k) + \gamma_2\left(\mathbb{E}[|w|]\right) \tag{4.26}$$

$$\mathbb{E}\left[\ell(x(k), \kappa_{\hat{\mu}}(x(k)))\right] \leq \tilde{\beta}(|x|, k) + \tilde{\gamma}_2\left(\mathbb{E}[|w|]\right) \tag{4.27}$$

for all $x \in \mathcal{X}$, $\mu \in \mathcal{M}(\mathbb{W})$, *and* $k \in \mathbb{I}_{\geq 0}$.

We now compare Corollary 4.25 to Theorem 2.32 for nominal MPC. An obvious difference is that Theorem 2.32 defines stochastic robustness in terms of the covariance of the disturbance (Σ) instead of the the expected norm of the disturbance ($\mathbb{E}[|w|]$). As noted in Remark 4.3, we can use $\text{tr}(\Sigma)$ in place of $\mathbb{E}[|w|]$ for all of the results in this chapter if the disturbance is zero mean. Alternatively, we can also establish the results from Chapters 2 and 3 with $\mathbb{E}[|w|]$ in place of $\text{tr}(\Sigma)$.

A more significant difference between these results is the RPI sets of interest. In Theorem 2.32, we allow for the choice of a compact set $\mathcal{S} := \mathcal{X} \cap \{x \in \mathbb{R} : V^0(x) \leq \rho\}$ for some $\rho > 0$ that then determines the margin of robustness $\delta > 0$ for the system. In Corollary 4.25, we instead assume that $\mathcal{X}$ is bounded, which is usually the case, and consider only the set $\mathcal{S} = \mathcal{X}$ and its associated margin of robustness $\delta > 0$. Thus, Theorem 2.32 is somewhat more informative, owing to the fact that the result is derived specifically for nominal MPC. Nonetheless, both results arrive at the same important conclusion: Nominal MPC provides some margin of stochastic robustness to disturbances.

Theorems 4.18 and 4.22 also seem to suggest that the value of $\delta > 0$ is constant across all three MPC formulations, but this conclusion is incorrect. The value of $\delta > 0$ in Corollary 4.25 is *not* necessarily the same as in Corollary 4.23 and Theorems 4.18 and 4.22. For an otherwise equivalent problem, the value of $\tilde{\tau} > 0$ in Assumption 4.6 may be significantly smaller (but not larger) for $\hat{\mathbb{W}} = \{0\}$ than for a set $\hat{\mathbb{W}}$ that includes more than the origin. Note that in the proof of Theorem 4.18, the value of δ is defined in (4.12) as

$$\delta := \frac{\tau - \tilde{\tau}}{\tilde{L}_f L_x}$$

Thus, a decrease in $\tilde{\tau}$ increases the value of δ. However, the feasible set $\mathcal{X}$ may also be larger (but not smaller) for $\hat{\mathbb{W}} = \{0\}$ than for a set $\hat{\mathbb{W}}$ that includes more than the origin. Thus, the Lipschitz constants $\tilde{L}_f, L_x > 0$ may also increase for nominal MPC relative to SMPC or CMPC. The value of $\delta > 0$ can therefore be larger or smaller for nominal MPC compared to SMPC or CMPC depending on the relative changes in these constants. As discussed in Section 3.7.1, general conclusions about which approach is the "most" robust in terms of RPI sets remain elusive.

One of the unintended results revealed by Corollary 4.25 is that Assumption 4.7 (or an analogous dissipativity assumption) is not required to ensure that $\mathcal{X}$ is RPI for sufficiently small disturbances. Instead, we rely on the requirement in (4.24) to establish that $\mathcal{X}$ is RPI and therefore ensure that the optimization problem is robustly recursively feasible. By ensuring that $\mathcal{X}$ is RPI, we can also derive the performance bound in (4.25) that is analogous to the bound derived for idealized stochastic MPC. Thus, nominal economic MPC, i.e., nominal MPC without Assumption 4.7, is robust to sufficiently small disturbances in an economic context. This result is, to the best of our knowledge, new for nominal MPC. However, terminal control laws that satisfy (4.24) and (4.6) may be difficult to construct for relevant systems and stage costs.

4.6 Numerical example

To demonstrate some of the implications of these technical results, we consider a simple numerical example based on the inventory control problem discussed in Section 3.7.4 and adapted from McAllister and Rawlings (2022e). In this example, we have inventory levels at two separate warehouses (x_1, x_2) and we need to determine the amount of inventory to add or remove for each facility at each time step (u_1, u_2). Negative values of x_1, x_2 represent backlog. We also consider fluctuations in demand (w_1, w_2), which gives the following discrete-time dynamics.

$$x_1^+ = x_1 + u_1 + w_1$$

$$x_2^+ = x_2 + u_2 + w_2$$

Note that the dynamical equations are decoupled, but the input is subject to the min/max constraints $u_1, u_2 \in [-2, 2]$ and the total product flow constraint $u_1 + u_2 \in [-2, 2]$. We plot the set $\mathbb{U}$ in Figure 4.4.

For the SMPC controller, we define $\hat{w} = (\hat{w}_1, \hat{w}_2) \in \hat{\mathbb{W}} := \{-0.5, 0, 0.5\} \times \{0\}$, i.e, $\hat{w}_1 \in \{-0.5, 0, 0.5\}$ and $\hat{w}_2 = 0$. Note that this SMPC formulation is equivalent to leaving $\hat{w}_2$ out of the dynamical equations used in the SMPC optimization problem. For some $\hat{\varepsilon}_1 \in [0, 1]$, we define $\hat{\mu}$ such that

$$\hat{\mu}(\{(-0.5, 0)\}) = \hat{\mu}(\{(0.5, 0)\}) = \hat{\varepsilon}_1/2 \quad \text{and} \quad \hat{\mu}(\{(0, 0)\}) = 1 - \hat{\varepsilon}_1$$

We use the stage cost

$$\ell(x, u) = \max\{x_1, 0\} + 0.5 \max\{x_2, 0\} + 5 \max\{-x_1, 0\} + 2.5 \max\{-x_2, 0\} \\ + 0.5|u_1| + 0.5|u_2|$$

in which we assign larger penalties to negative values of x_1, x_2 (backlog) than positive values of x_1, x_2 (extra inventory). We also penalize increasing/decreasing the inventory levels with u_1, u_2. To better illustrate this cost function, we provide a contour plot in Figure 4.4 of $\ell(x, 0)$. We now construct the terminal constraint and cost for this problem. We choose the terminal

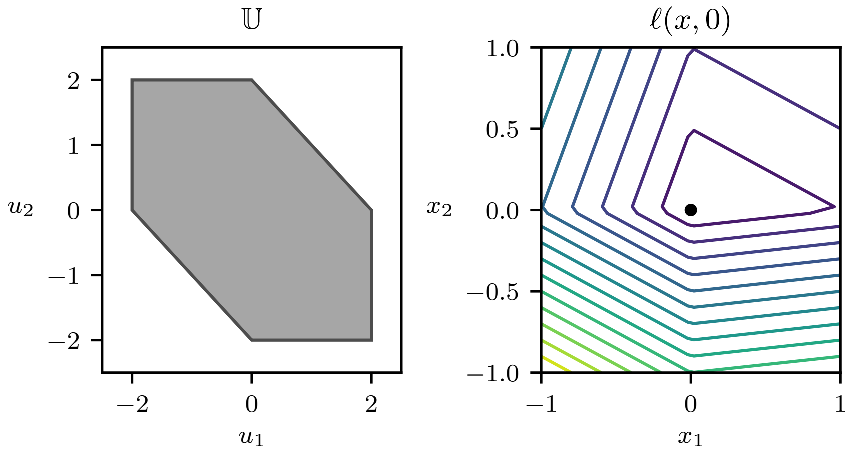

Figure 4.4: Input constraint set $\mathbb{U}$ (left) and contour plot of the stage cost $\ell(x, 0)$ (right) for the numerical example.

cost $V_f(x) = 6|x_1| + 6|x_2|$ and $\tau = 6$ to give

$$\mathbb{X}_f = \{x \in \mathbb{R}^2 : V_f(x) \le 6\} = \{x \in \mathbb{R}^2 : |x_1| + |x_2| \le 1\}$$

We use the terminal control law $\kappa_f(x) = -x$ and verify that

$$V_f(f(x, \kappa_f(x), 0)) = 0 \le 6|x_1| + 6|x_2| - \ell(x, -x)$$

for all $x \in \mathbb{X}_f$. Moreover,

$$V_f(f(x, \kappa_f(x), \hat{w})) \le 6|\hat{w}_1| \le 3$$

for all $x \in \mathbb{X}_f$, $\hat{w} \in \hat{\mathbb{W}}$ and therefore (4.5) holds with $\tilde{\tau} = 3$. We use the parametrization $\pi(x, v) = \text{sat}_{\mathbb{U}}(-x + v)$ and let $\mathbb{V} = \{v \in \mathbb{R}^2 : |v_1| \le 10,\ |v_1| \le 10\}$. This formulation satisfies Assumptions 4.4 to 4.7.

As noted in Section 3.7.4, SMPC formulations can produce control laws that are *discontinuous* with respect to the probability distribution. To illustrate changes in the control law with respect to changes in $\hat{\mu}$, we fix $x_2 = 0$ and calculate the values of $\kappa_{\hat{\mu}}(x)$ for $x_1 \in [-3, 3]$

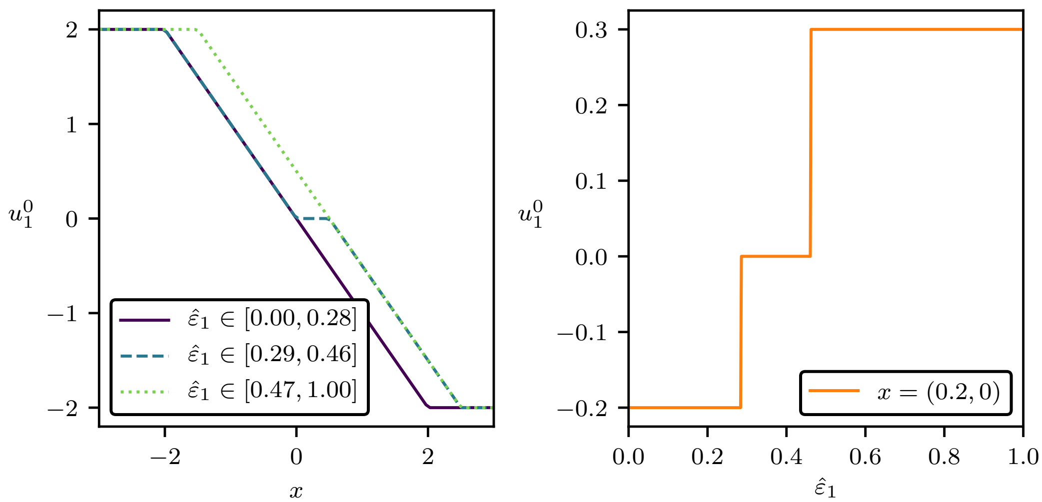

Figure 4.5: The first element of $\kappa_{\hat{\mu}}(x)$, denoted u_1^0, for all $x_1 \in [-3, 3]$ and different regions of $\hat{\varepsilon}_1$ (left). Also, u_1^0 for all $\hat{\varepsilon}_1$ and fixed $x = (0.2, 0)$ (right).

for all $\hat{\varepsilon}_1 \in [0, 1]$. There are three different regions of $\hat{\varepsilon}_1$ that produce a different control law in each region. We plot the first element of $\kappa_{\hat{\mu}}(x)$, denoted u_1^0, on the left side of Figure 4.5 for each of these regions. Note that the second element $\kappa_{\hat{\mu}}(x)$ is zero for all $\hat{\varepsilon}_1 \in [0, 1]$ and $x_2 = 0$.

For $\hat{\varepsilon}_1 \leq 0.28$, the optimal control action is to drive x_1 to zero as aggressively as permitted by the constraints. For $\hat{\varepsilon}_1 \geq 0.47$, we instead drive x_1 to 0.5 as aggressively as permitted by the constraints. In contrast to the example in Section 3.7.4, we include a penalty for $|u_1|$ and therefore produce an intermediate control law for $\hat{\varepsilon}_1 \in [0.29, 0.46]$. In this region, we drive x_1 to zero for $x_1 \leq 0$, drive x_1 to 0.5 for $x_1 \geq 0.5$, and take no action for $x_1 \in [0, 0.5]$.

The transitions between these three regions appear to be discontinuous with respect to changes in $\hat{\varepsilon}_1$. To demonstrate these discontinuities, we fix $x = (0.2, 0)$ and plot u_1^0 for all $\hat{\varepsilon}_1 \in [0, 1]$ on the right side of Figure 4.5. These discontinuities occur at $\hat{\varepsilon}_1 \approx 0.284$ and $\hat{\varepsilon}_1 \approx 0.461$. For the exact values of $\hat{\varepsilon}_1$ at which these discontinuities occur, we observe that all values of $u_1 \in [-0.2, 0]$ and $u \in [0, 0.25]$, respectively, are optimal for $x = (0.2, 0)$. Thus, arbitrarily small changes in $\hat{\mu}$ can produce significantly different controllers.

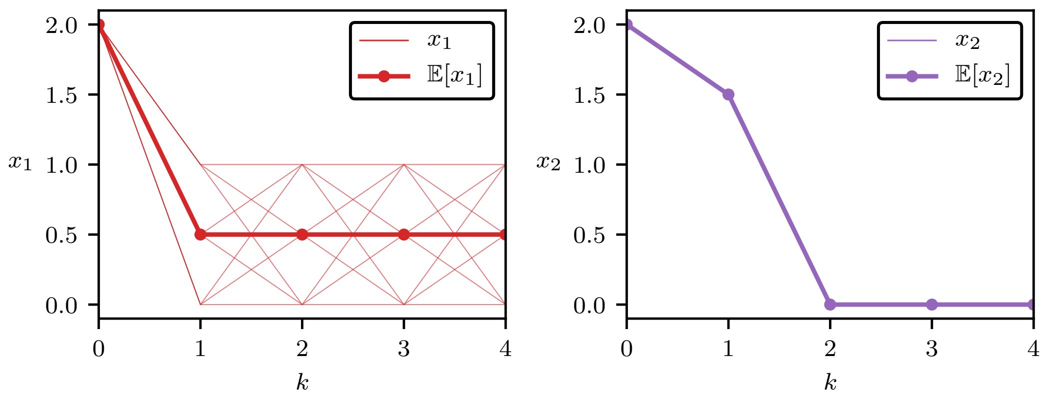

Figure 4.6: Closed-loop trajectories for x_1, x_2 for idealized SMPC with their expected values.

For this example, we now consider the closed-loop trajectory for idealized SMPC. We choose $\hat{\varepsilon}_1 = 2/3$, start the system at $x_0 = (2, 2)$, and simulate the closed-loop trajectory with $\mathbb{W} = \hat{\mathbb{W}}$ and $\mu = \hat{\mu}$. We plot all possible trajectories and their associated expected value for x_1, x_2 up to $k = 4$ in Figure 4.6. Similar to the example in Section 3.7.4, we observe that SMPC ensures that x_1 remains positive for all realizations of the disturbance and thereby avoids the large penalty for backlog. Note, however, that x_2 is driven to zero since $\hat{w}_2 = 0$.

We now introduce the unmodeled disturbance w_2. Let $\mathbb{W} := \{-0.5, 0, 0.5\}^2$, i.e., $w_1, w_2 \in \{-0.5, 0, 0.5\}$, and $\mu(\{(w_1, w_2)\}) = \mu_1(\{w_1\})\mu_2(\{w_2\})$ in which

$$\mu_i(\{-0.5\}) = \mu(\{0.5\}) = \hat{\varepsilon}_i/2 \quad \text{and} \quad \mu_i(\{0\}) = 1 - \hat{\varepsilon}_i$$

and $\varepsilon_i \in [0, 1]$ for $i = 1, 2$. Thus w_1 and w_2 are independent and zero mean. We assume that w_1 is modeled correctly and therefore $\varepsilon_1 = \hat{\varepsilon}_1 = 2/3$, but we also include the unmodeled disturbance w_2 via ε_2. For these probability measures, the Wasserstein metric is given by $W(\mu, \hat{\mu}) = \varepsilon_2/2$.

Starting from $x_0 = (2, 2)$, we simulate the closed-loop trajectories for SMPC subject to this unmodeled disturbance. We plot the expected value of the norm of the closed-loop

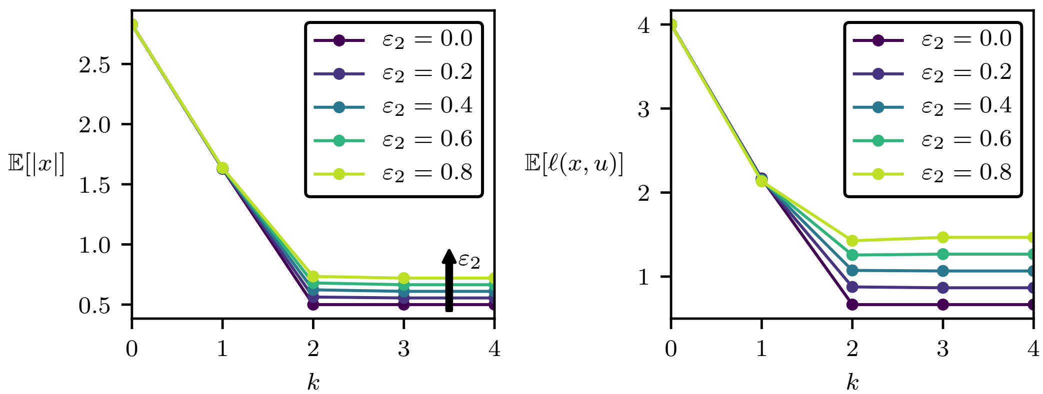

Figure 4.7: Expected value of the norm of the closed-loop state and closed-loop stage cost for multiple values of ε_2.

state trajectory, denoted $\mathbb{E}[|x(k)|] = \mathbb{E}[|\phi_{\hat{\mu}}(k; x_0, \mathbf{w}_k)|]$, and closed-loop stage cost, denoted $\mathbb{E}[\ell(x(k), u(k))]$, for multiple values of ε_2 in Figure 4.7. For $\varepsilon_2 = 0$, we have idealized SMPC and we observe that $\mathbb{E}[|x(k)|]$ and $\mathbb{E}[\ell(x,u)]$ initially decrease with increasing k until leveling off, consistent with the theoretical results for idealized SMPC. As we increase ε_2 and thereby increase $W(\mu, \hat{\mu})$, the values of $\mathbb{E}[|x(k)|]$ and $\mathbb{E}[\ell(x(k), u(k))]$ increase for each $k \geq 2$. This behavior is consistent with the DRASiE, ℓ-DRASiE, and therefore Theorem 4.18, i.e., the upper bounds on these performance metrics of the closed-loop trajectory increase with increasing $W(\mu, \hat{\mu})$.

In this closed-loop simulation, we observe that the optimization problem remains feasible for all realizations of $w \in \mathbb{W}$, despite the fact that $\mathbb{W} \neq \hat{\mathbb{W}}$. Thus, $\mathcal{X}$ is RPI for the the closed-loop system with $w \in \mathbb{W}$. To illustrate the robustness of SMPC to differences in $\mathbb{W}$ and $\hat{\mathbb{W}}$, we first compute $\mathcal{X}$ for the set $\hat{\mathbb{W}}$ used the SMPC formulation. We then sample this feasible region and simulate one step of the closed-loop trajectory from these sampled points to construct the region

$$\mathcal{X}^+ := \{f(x, \kappa_{\hat{\mu}}(x), \hat{w}) : x \in \mathcal{X},\ \hat{w} \in \hat{\mathbb{W}}\}$$

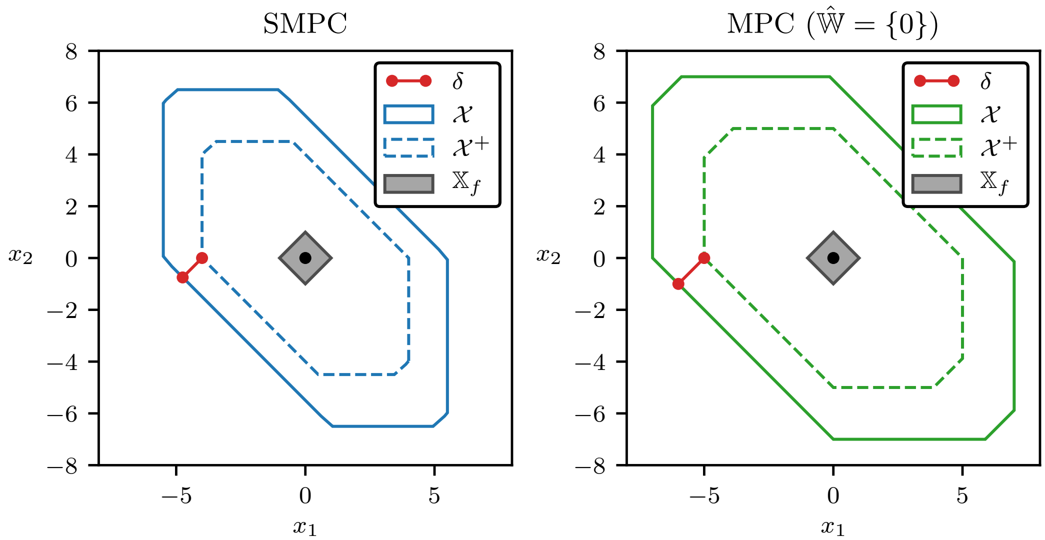

Figure 4.8: The feasible set for SMPC (left) and MPC (right) $\mathcal{X}$ with $\hat{\mathbb{W}}$ and the set of states after one application of the control law subject to $\hat{w} \in \hat{\mathbb{W}}$, denoted $\mathcal{X}^+$.

On the left side of Figure 4.8, we plot these two sets for SMPC. Note that $\mathcal{X}^+$ is a strict subset of $\mathcal{X}$. Thus, there exists some set of nonzero disturbances, in addition to $\hat{\mathbb{W}}$, that can be added to the closed-loop system such that $\mathcal{X}$ remains RPI. For example, we see that adding ± 0.5 to the x_2 direction of any $x \in \mathcal{X}^+$ does not force the state outside of $\mathcal{X}$, confirming the observations of robust recursive feasibility noted in the previous paragraph.

With this information, we can in fact compute the largest $\delta > 0$ such that $\mathcal{X}$ is RPI for the closed-loop system $x^+ = f(x, \kappa_{\hat{\mu}}(x), w)$, $w \in \mathbb{W}$ with any $\mathbb{W} \in \mathbb{R}^2$ that satisfies $d_H(\mathbb{W}, \hat{\mathbb{W}}) \leq \delta$, i.e.,

$$\delta := \max \left\{ \delta : f(x, \kappa_{\hat{\mu}}(x), w) \in \mathcal{X} \ \forall w \in \mathbb{W} \text{ and } d_H(\mathbb{W}, \hat{\mathbb{W}}) \leq \delta \right\}$$

For this example, the value of δ is equal to the minimum distance between the boundary of $\mathcal{X}$ and $\mathcal{X}^+$, as shown in left plot of Figure 4.8. The exact value of this distance is $\delta = 3\sqrt{2}/4 \approx 1.06$ for SMPC. For nominal MPC, we set $\hat{\mathbb{W}} = \{0\}$ and perform the same calculations and plot the resulting sets on the right side of Figure 4.8. Again, we can calculate the margin of robustness exactly and find that $\delta_{mpc} = \sqrt{2} \approx 1.41$.

Note that $\delta_{mpc} > \delta$, corroborating the discussion that followed Corollary 4.25. This suggests that SMPC is trading some of its "general" robustness to disturbances, i.e., robustness to any $|w| \leq \delta$, for "specific" robustness to disturbances in the set $\hat{\mathbb{W}}$. This trade, however, is not one-to-one and SMPC still maintains a reasonable margin of robustness to unmodeled ("general") disturbances in this example.

4.7 Summary

Nominal MPC, as shown in Chapter 2, is robust to sufficiently small errors in the *deterministic* dynamical model. In this chapter, we established that SMPC is robust to sufficiently small errors in the *stochastic* dynamical model. Feedback is the key component of both algorithms that facilitates this robustness. These errors in the stochastic dynamical model may include incorrectly modeled disturbances, for which the distribution is incorrect, and unmodeled disturbances, for which the disturbance is entirely absent from the stochastic model. The only requirement is that these unmodeled disturbances are i.i.d. in time and enter the state via a continuous function $f(x, u, w)$. Moreover, these errors can include discrepancies in the stochastic dynamical model introduced by scenario-based approximations of the stochastic optimization problem.

In addition to the implications for SMPC, Theorems 4.18 and 4.22 also allow us to characterize the stochastic robustness of other MPC formulations, i.e., CMPC and nominal MPC, as special cases of SMPC. Corollaries 4.23 and 4.25 unify the analysis of these three different problem formulations and largely subsume the results discussed in Chapter 2 for the stochastic robustness of nominal MPC.[3]. In fact, this approach reveled a novel requirement for the

[3]With the exception that we require bounded $\mathcal{X}$ in the current chapter.

terminal constraint of nominal economic MPC that ensures the feasible set remains RPI for sufficiently small disturbances.

The definitions of distributional robustness and the SISS Lyapunov framework introduced in this chapter are not restricted to SMPC or the MPC formulation in general. Definitions 4.13 and 4.15 apply to any stochastic dynamical system given by $f(x, \kappa_{\hat{\mu}}(x), w)$ with some control law that depends on a probability distribution $\hat{\mu}$. Indeed, these definitions are therefore applicable to the larger field of stochastic optimal control and, potentially, the developing field of distributionally robust control.

Although these results offer significant insight into the robustness and efficacy of SMPC in practice, these results do not significantly change the conclusions in Chapter 3. The benefits of implementing SMPC compared to nominal MPC remain unclear in general. Furthermore, a rigorous means to determine, prior to a simulation study, if such an investment in complexity is worthwhile for a specific problem of interest remains unavailable. Similar questions can be proposed for distributionally robust control as well, e.g., is the inherent distributional robustness of SMPC, or nominal MPC for that matter, sufficient? Or, is there a clear and significant benefit to explicitly designing for this distributional robustness via distributionally robust optimization (DRO)? As the field of DRO and any application in MPC is new, there is no clear answer to these questions and, as with the comparison between nomianl MPC and SMPC, a rigorous answer to this question may remain elusive.

These robustness results all require that the disturbance or model error is sufficiently *small*. Even for SMPC, the requirements detailed in Assumption 4.6 for the terminal constraint and cost curtail the size of disturbances that can be considered in the SMPC formulation. While this characterization is reasonable for many traditional process control applications, there are some MPC applications, such as production scheduling problems, in which the relevant disturbances are not sufficient small. These disturbance are instead *large*, but infrequent in that they occur with small probability. In the subsequent chapter, we consider the robustness of MPC to this class of large and infrequent disturbances.

Chapter 5

Large and Infrequent Disturbances

In the previous three chapters, a key characteristic of these robustness results is that the disturbance or model error can be treated as sufficiently *small.* For most process control applications, this characterization is reasonable. The model errors, measurement noise, and perturbations anticipated in many relevant applications are well described as small disturbances. The scope of MPC and this dissertation, however, are not limited to traditional process control applications.

Aided by both theoretical and computational advances, MPC is now being applied to higher-level planning and scheduling problems with both the usual continuous-valued input decisions (e.g, "how much do I buy/sell?") as well as integer/binary-valued decisions (e.g., "do I turn this unit ON/OFF?") that are not typically considered in MPC applications. As first recognized in Rawlings and Risbeck (2017), the compact set $\mathbb{U}$ is already general enough to enforce integer constraints on the subset of the inputs. Thus, the theoretical results derived for traditional MPC applications with continuous decisions also hold MPC applications with discrete decisions. To emphasize this point, we note that *all* of the results in this thesis already include systems with both continuous and discrete input decisions.

With discrete decisions and scheduling problems, we must now consider discrete disturbances such as task delays or breakdowns in equipment. While the transition from continuous to discrete disturbances does not invalidate the analysis in the previous chapters, the argument that a discrete disturbances is "sufficiently small" is dubious at best. Breakdowns in equipment that can occur in a production scheduling problem are not, and should not be considered, small disturbances. Thus, the theory developed so far for sufficiently small disturbances is now insufficient. We instead characterize these disturbances as "large." In practice, however, these "large" disturbances are also typically infrequent, e.g., a delay occurs with some small probability. We therefore refer to this class of disturbances as *large*, because the disturbances are bounded away from zero and cannot be considered sufficiently small, and *infrequent*, because the probability that these disturbances occur is small. This description also applies to a variety of disturbances that can be encountered for even tradition process control systems such as faults, missing measurements, and communication failures.

For this class of disturbances, any deterministic bound must consider the worst deterministic performance possible for the system, e.g., the entire production line is broken indefinitely. Consequently, a deterministic bound is far too conservative and offers little insight for a closed-loop system subject to this class of disturbances. We instead exploit the infrequent nature of these disturbances and propose a stochastic definition of robustness for large and infrequent disturbances that is distinct from, but similar to, the previous stochastic robustness results derived for sufficiently small disturbances. We then establish conditions that ensure nominal MPC is inherently robust to large and infrequent disturbances in this stochastic context. In this chapter, we also introduce time-varying MPC to address the time-varying dynamical models, costs, and constraints that are common in higher-level planning and scheduling problems. These time-varying systems subject to large and infrequent disturbances are the primary focus for the remainder of this dissertation.

5.1 Time-varying MPC

We a consider time-varying system of the form

$$x^+ = f(x,u,w,t) \quad f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{I}_{\geq 0} \to \mathbb{R}^n$$

in which $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^n$ is the controller input, $w \in \mathbb{W} \subseteq \mathbb{R}^q$ is the stochastic disturbance, at the discrete time index $t \in \mathbb{I}_{\geq 0}$. For a given initial time $t \in \mathbb{I}_{\geq 0}$ and $k \in \mathbb{I}_{\geq t}$, let

$$\mathbf{w}_{t:k} := (w(t), w(t+1), \dots, w(k-1))$$

denote a sequence of disturbances from the initial time t until time k. We again consider the usual assumption for these stochastic disturbances.

Assumption 5.1 (Disturbances)**.** The disturbances $w \in \mathbb{W}$ are random variables that are i.i.d. in time. The set $\mathbb{W}$ is compact and contains the origin.

Given Assumption 5.1, let $\mu : \mathcal{B}(\mathbb{W}) \to [0,1]$ denote the probability measure for w and $\mathcal{M}(\mathbb{W})$ denote the set of all probability measures on the measurable space $(\mathbb{W}, \mathcal{B}(\mathbb{W}))$. We define expected value of a Borel measurable function $g : \mathbb{W}^{k-t} \to \mathbb{R}$ as the Lebesgue integral

$$\mathbb{E}[g(\mathbf{w}_{t:k})] := \int_{\mathbb{W}^{k-t}} g(\mathbf{w}_{t:k}) d\mu(w(t)) d\mu(w(t+1)) \dots d\mu(w(k-1))$$

with respect to any $\mu \in \mathcal{M}(\mathbb{W})$. We similarly define the probability that $g(\mathbf{w}_{t:k}) \in S$ for some Borel measurable function $g : \mathbb{W}^{k-t} \to \mathbb{R}^n$ and set $S \subseteq \mathbb{R}^n$ as

$$\Pr\big(g(\mathbf{w}_{t:k}) \in S\big) := \int_{\mathbb{W}^{k-t}} I_S\big(g(\mathbf{w}_{t:k})\big) d\mu(w(t)) d\mu(w(t+1)) \dots d\mu(w(k-1))$$

in which $I_S(\cdot)$ is the indicator function, i.e., $I_S(x) = 1$ if $x \in S$ and zero otherwise.

5.1.1 MPC formulation

We consider time-varying nominal MPC in which the system is described by

$$x^+ = f(x, u, 0, t) \tag{5.1}$$

For a horizon $N \in \mathbb{I}_{\geq 0}$, let $\hat{\phi}(k; x, \mathbf{u}, t)$ denote the open-loop state solution to the nominal system in (5.1) at time $k \in \mathbb{I}_{[t,N]}$, given the state $x \in \mathbb{R}^n$ at time $t \in \mathbb{I}_{\geq 0}$ and the input sequence

$$\mathbf{u} := (u(t), u(t+1), \ldots, u(t+N-1))$$

Unlike Chapter 2, we allow state constraints in the MPC formulation. We often need these state constraints in scheduling and planning problems to enforce requirements that are important for the realism of the system, e.g., inventory must be nonnegative. State constraints that represent nonphysical restrictions and/or desired goals of the system should still be avoided to ensure robust recursive feasibility. We also allow these constraints to be time-varying such that at time $t \in \mathbb{I}_{\geq 0}$ we require that

$$(x, u) \in \mathbb{Z}(t) \subseteq \mathbb{R}^n \times \mathbb{U}$$

for $\mathbb{U} \subseteq \mathbb{R}^m$ and the sequence of sets $(\mathbb{Z}(t))_{t=0}^{\infty}$. We also use a time-varying terminal constraint given by the sequence $(\mathbb{X}_f(t))_{t=0}^{\infty}$.

For these constraints, we define the set of admissible inputs and feasible initial states as follows.

$$\begin{aligned}\mathcal{U}(x,t) &:= \{\mathbf{u} \in \mathbb{U}^N : (\hat{\phi}(k;x,\mathbf{u},t), u(k)) \in \mathbb{Z}(k)\ \forall k \in \mathbb{I}_{t:t+N-1} \\ &\qquad \text{and } \hat{\phi}(N;x,\mathbf{u},t) \in \mathbb{X}_f(N)\} \\ \mathcal{X}(t) &= \{x \in \mathbb{R}^n : \mathcal{U}(x,t) \neq \emptyset\}\end{aligned}$$

Note that these sets also depend on the time $t \in \mathbb{I}_{\geq 0}$. We define the time-varying stage cost $\ell : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{I}_{\geq 0} \to \mathbb{R}$ and terminal cost $V_f : \mathbb{R}^n \times \mathbb{I}_{\geq 0} \to \mathbb{R}_{\geq 0}$. The MPC cost function is then defined as

$$V(x,\mathbf{u},t) := \sum_{k=i}^{t+N-1} \ell(x(k),u(k),k) + V_f(x(t+N),t+N)$$

in which $x(k) := \hat{\phi}(k;x,\mathbf{u},t)$.

For any $t \in \mathbb{I}_{\geq 0}$ and $x \in \mathcal{X}(t)$, the nominal MPC problem is defined as

$$\mathbb{P}(x,t) := V^0(x,t) := \min_{\mathbf{u} \in \mathcal{U}(x,t)} V(x,\mathbf{u},t)$$

and the optimal solution(s) are denoted $\mathbf{u}^0(x,t) := \arg\min_{\mathbf{u} \in \mathcal{U}(x,t)} V(x,\mathbf{u},t)$. Note that $\mathbf{u}^0(x,t)$ is again a set-valued mapping, and we use a Borel measurable selection rule to define the single-valued control law $\kappa(\cdot,t) : \mathcal{X}(t) \to \mathbb{U}$ such that $\kappa(x,t) \in \{u(0) : \mathbf{u} \in \mathbf{u}^0(x,t)\}$ for all $x \in \mathcal{X}(t)$ and $t \in \mathbb{I}_{\geq 0}$. The closed-loop system is then given by

$$x^+ = f(x,\kappa(x,t),w,t) \tag{5.2}$$

in which both the underlying system and control law vary with time $t \in \mathbb{I}_{\geq 0}$. Let $\phi(k;x,\mathbf{w}_{t:k},t)$ denote the solution to (5.2) at time $k \in \mathbb{I}_{\geq t}$, given the initial condition $x \in \mathcal{X}(t)$, initial time $t \in \mathbb{I}_{\geq 0}$, and disturbance sequence $\mathbf{w}_{t:k} \in \mathbb{W}^{k-t}$.

5.1.2 Assumptions

We require the usual assumptions of continuity, closed constraint sets, and bounded inputs. Note that we do not require the terminal set $\mathbb{X}_f(t)$ or the feasible set $\mathcal{X}(t)$ to be bounded.

Assumption 5.2 (Continuity of sysem and cost)**.** The model $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{I}_{\geq 0} \rightarrow \mathbb{R}^n$, stage cost $\ell : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{I}_{\geq 0} \rightarrow \mathbb{R}$, and terminal cost $V_f : \mathbb{R}^n \times \mathbb{I}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ are continuous. The function $\ell(\cdot)$ is bounded from below. Also, $f(0,0,0,t) = 0$, $\ell(0,0,t) = 0$, and $V_f(0,t) = 0$ for all $t \in \mathbb{I}_{\geq 0}$.

Assumption 5.3 (Properties of constrain set)**.** For each $t \in \mathbb{I}_{\geq 0}$, the sets $\mathbb{Z}(t)$ and $\mathbb{X}_f(t)$ are closed and contain the origin. The set $\mathbb{U}$ is compact and contains the origin.

With these assumptions, we are again assuming that the system has been shifted such that the target setpoint is at the origin. This setpoint, however, may also vary with time along with the system. For example, let $\bar{x}$ and $\bar{u}$ denote the original state and input variables for the system dynamics

$$\bar{x}^+ = \bar{f}(\bar{x}, \bar{u}, w, t) \qquad \bar{f} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{I}_{\geq 0} \rightarrow \mathbb{R}^n$$

We also have time-varying constraints for the original system such that

$$(\bar{x}, \bar{u}) \in \bar{\mathbb{Z}}(t) \subseteq \mathbb{R}^n \times \bar{\mathbb{U}}$$

We then consider a time-varying reference trajectory given by the state and input sequences $\bar{\mathbf{x}}_r$ and $\bar{\mathbf{u}}_r$ defined as

$$\bar{\mathbf{x}}_r := (\bar{x}_r(0), \bar{x}_r(1), \dots) \qquad \bar{\mathbf{u}}_r := (\bar{u}_r(0), \bar{u}_r(1), \dots)$$

We require that this trajectory satisfies

$$\bar{x}_r(t+1) = \bar{f}(\bar{x}_r(t), \bar{u}_r(t), 0, t)$$

and the constraints $(\bar{x}_r(t), \bar{u}_r(t)) \in \bar{\mathbb{Z}}(t)$ for all $t \in \mathbb{I}_{\geq 0}$, i.e., the reference trajectory is a feasible trajectory for the nominal system. To shift this system to the origin, we define the deviation variables

$$x(t) := \bar{x}(t) - \bar{x}_r(t) \qquad u(t) := \bar{u}(t) - \bar{u}_r(t)$$

the dynamics

$$x^+ = f(x, u, w, t) := \bar{f}(x + \bar{x}_r(t), u + \bar{u}_r(t), w, t) - \bar{x}_r(t+1)$$

and constraints

$$\mathbb{Z}(t) := \{(\bar{x} - \bar{x}_r(t), \bar{u} - \bar{u}_r(t)) : (\bar{x}, \bar{u}) \in \bar{\mathbb{Z}}(t)\}$$

Note that the requirements for $\mathbf{x}_r$ and $\mathbf{u}_r$ ensure that $f(0,0,0,t) = 0$ and $(0,0) \in \mathbb{Z}(t)$. We can similarly shift the stage cost such that $\ell(0,0,t) = 0$. An example state trajectory is shown in original and deviation variables in Figure 5.1. This reference trajectory may be constant, periodic, or any time-varying sequence that is a valid trajectory for the nominal closed-loop system. We can also use a time-varying reference trajectory as the target for a system that is *time-invariant* in the original variables, but the system in terms of the deviation variables x and u is ultimately time-varying.

While shifting the system, constraints, and cost to the origin is convenient for mathematical analysis, the MPC optimization problem can still be solved in the original variables. The control law for the original MPC problem is equivalent to the shifted MPC problem plus the

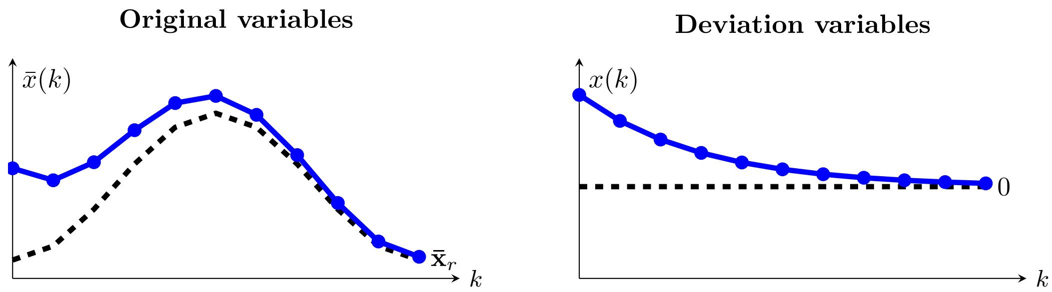

Figure 5.1: The same state trajectory shown in original ($\bar{x}$) and deviation (x) variables with respect to the reference trajectory ($\bar{\mathbf{x}}_r$).

reference input, i.e., $\bar{\kappa}(\bar{x}, t) = \kappa(x, t) + \bar{u}_r(t)$ in which $\bar{\kappa}(\bar{x}, t)$ is the control law defined by the original MPC problem. Given this equivalence, we discuss the theoretical properties of only the shifted MPC problem in this chapter.

The time-varying terminal ingredients for the shifted MPC problem must satisfy the following restrictions.

Assumption 5.4 (Terminal control law)**.** There exists a terminal control law $\kappa_f : \mathbb{X}_f \times \mathbb{I}_{\geq 0} \to \mathbb{U}$ such that

$$f(x, \kappa_f(x, i), 0, i) \in \mathbb{X}_f(i + 1) \tag{5.3}$$

$$V_f(f(x, \kappa_f(x, i), 0, i), i + 1) \leq V_f(x, i) - \ell(x, \kappa_f(x, i), i) \tag{5.4}$$

for all $x \in \mathbb{X}_f(i)$ and $i \in \mathbb{I}_{\geq 0}$. Furthermore, $(x, \kappa_f(x, i)) \in \mathbb{Z}(i)$ for all $x \in \mathbb{X}_f$ and $i \in \mathbb{I}_{\geq 0}$.

If asymptotic stability of the origin (reference trajectory) is the primary goal for the controller, then we also require that the stage cost satisfies the following assumption.

Assumption 5.5 (Stage cost bound)**.** There exist $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$ such that

$$\alpha_\ell(|x|) \leq \ell(x, u, i)$$

for all $(x, u) \in \mathbb{Z}(i)$ and $i \in \mathbb{I}_{\geq 0}$.

In previous chapters, we constructed the terminal set via a level set of the terminal cost function, i.e., $\mathbb{X}_f := \{x \in \mathbb{R}^n : V_f(x) \leq \tau\}$ for some $\tau > 0$. Thus, the origin is necessarily contained in the interior of the terminal set. In this chapter, we instead use a more general condition called *weak controllability*.

Assumption 5.6 (Weak controllability)**.** There exists $\alpha_2(\cdot) \in \mathcal{K}_\infty$ such that

$$V^0(x,i) \leq \alpha_2(|x|)$$

for all $x \in \mathcal{X}(i)$ and $i \in \mathbb{I}_{\geq 0}$.

See Rawlings et al. (2020, Prop. 2.38) for a variety of other conditions for which Assumption 5.6 assumption holds, including the condition that $\mathbb{X}_f(i)$ contains the origin in its interior for all $i \in \mathbb{I}_{\geq 0}$.

Since the system and feasible set $\mathcal{X}(t)$ are now time-varying, we redefine positive invariance and robust positive invariant as follows.

Definition 5.7 (Positive invariance)**.** A sequence of sets $(\mathcal{X}(t))_{i=t}^{\infty}$ is positive invariant for the nominal system $x^+ = f(x, \kappa(x,t), 0, t)$ if $x^+ \in \mathcal{X}(t+1)$ for all $x \in \mathcal{X}(t)$ and $t \in \mathbb{I}_{\geq 0}$.

Definition 5.8 (Robust positive invariance)**.** A sequence of sets $(\mathcal{X}(t))_{i=t}^{\infty}$ is robustly positive invariant (RPI) for the system $x^+ = f(x, \kappa(x,t), w, t)$, $w \in \mathbb{W}$ if $x^+ \in \mathcal{X}(t+1)$ for all $x \in \mathcal{X}(t)$, $w \in \mathbb{W}$, and $t \in \mathbb{I}_{\geq 0}$.

5.1.3 Nominal properties

With these assumptions, we can establish some nominal closed-loop properties of time-varying MPC that are generalizations of properties previously established for time-invariant MPC. These results and corresponding proofs are available in either Rawlings et al. (2020, s.

2.4.5) or Risbeck and Rawlings (2019). We begin with a key cost decrease condition for the nominal system.

Lemma 5.9. *If Assumptions 5.2 to 5.4 hold, then the sequence of sets* $(\mathcal{X}(t))_{t=0}^{\infty}$ *is positive invariant for the nominal system* $x^+ = f(x, \kappa(x,t), 0, t)$ *and*

$$V^0(f(x, \kappa(x,t), 0, t), t+1) \leq V^0(x,t) - \ell(x, \kappa(x,t), t)$$

for all $x \in \mathcal{X}(t)$ *and* $t \in \mathbb{I}_{\geq 0}$.

From this nominal cost decrease condition we can establish the following result for economic applications in which Assumptions 5.5 and 5.6 are not required.

Theorem 5.10 (Nominal performance)**.** *Let Assumptions 5.2 to 5.4 hold. Then the sequence of sets* $(\mathcal{X}(t))_{t=0}^{\infty}$ *is positive invariant for the nominal closed-loop system* $x^+ = f(x, \kappa(x,t), 0, t)$. *Furthermore, the nominal closed-loop trajectory satisfies*

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{k=t}^{t+T-1} \ell(x(k), \kappa(x(k), k), k) \leq 0$$

in which $x(k) = \phi(k; x, \mathbf{0}, t)$ *for all* $x \in \mathcal{X}$ *and* $t \in \mathbb{I}_{\geq 0}$.

Theorem 5.10 was first established for time-invariant economic MPC in Angeli et al. (2012) and later extended to time-varying economic MPC in Risbeck and Rawlings (2019, Thm. 1). Thus, the average performance of the closed-loop system is no worse than the origin (reference trajectory), in terms of the stage cost.

We can also define asymptotic stability of the origin (reference trajectory).

Definition 5.11 (Asymptotic stability)**.** The origin is asymptotically stable for the nominal closed-loop system $x^+ = f(x, \kappa(x,t), 0, t)$ in a positive invariant sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$

if there exists $\beta(\cdot) \in \mathcal{KL}$ such that

$$|\phi(k; x, \mathbf{0}, t)| \leq \beta(|x|, k - t)$$

for all $x \in \mathcal{X}$, $t \in \mathbb{I}_{\geq 0}$, and $k \in \mathbb{I}_{\geq t}$.

For time-varying systems and optimal cost functions, we require a time-varying Lyapunov function. We can then use this Lyapunov function as a sufficient condition of asymptotic stability.

Definition 5.12 (Lyapunov function)**.** A function $V(\cdot, t) : \mathcal{X}(t) \to \mathbb{R}_{\geq 0}$ is a Lyapunov function for a system $x^+ = f(x, \kappa(x, t), 0, t)$ in a positive invariant sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$ if there exist $\alpha_1(\cdot), \alpha_2(\cdot), \alpha_3(\cdot) \in \mathcal{K}_\infty$ such that

$$\alpha_1(|x|) \leq V(x, t) \leq \alpha_2(|x|)$$
$$V(f(x, \kappa(x, t), 0, t)) \leq V(x) - \alpha_3(|x|)$$

for all $x \in \mathcal{X}(t)$ and $t \in \mathbb{I}_{\geq 0}$.

Proposition 5.13. *If a system $x^+ = f(x, \kappa(x, t), 0, t)$ admits a Lyapunov function in the positive invariant sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$, then the origin is asymptotically stable for the nominal closed-loop system.*

The proof of Proposition 5.13 is similar to the time-invariant result and is available in Rawlings et al. (2020, Thm. B.24). With the addition of Assumption 5.5 and Assumption 5.6, we can establish that the reference trajectory is asymptotically stable for the nominal closed-loop system.

Theorem 5.14 (Asymptotic stability)**.** *Let Assumptions 5.2 to 5.6 hold. Then the nominal system $x^+ = f(x, \kappa(x, t), 0, t)$ admits a Lyapunov function in the positive invariant sequence of sets*

$(\mathcal{X}(t))_{t=0}^{\infty}$ *and the origin is asymptotically stable.*

While the extension from time-invariant to time-varying MPC is somewhat straightforward, this result is a recent development (Risbeck and Rawlings, 2019, Thm. 3). A more detailed proof is available in Rawlings et al. (2020, Thm 2.39). By extending the MPC formulation and associated assumptions to address the more general class of time-varying problems, we are able to establish time-varying versions of the same nominal performance and asymptotic stability results that are typically established for time-invariant MPC. It stands to reason that all of the properties established in this thesis for time-invariant MPC can be extended to time-varying MPC with suitable modifications to the assumptions. Since these extensions add additional notation to the analysis problem without any notably insights, we have thus far restricted this thesis to the time-invariant case. We depart from this approach only in these later chapters because the primary use of the theory developed for large and infrequent disturbances is to address closed-loop scheduling, an application that requires the time-varying MPC formulation introduced here.

5.2 Large and infrequent disturbances

We introduce this class of large and infrequent disturbance by discussing them in contrast to the small and persistent disturbances addressed in the previous chapters. In Theorem 2.32, we establish that there exists some nonzero margin of robustness, that we denote $\delta_0 > 0$ in the current chapter, for nominal MPC. For disturbances within this margin of robustness ($|w| \leq \delta_0$), nominal MPC is robust in both a deterministic and stochastic context. Thus, these small disturbances are contained in the set $\mathbb{W}_0 := \{w \in \mathbb{R}^q : |w| \leq \delta_0\}$.

In this chapter, we consider the following question. If the disturbance w is *not small*, i.e., $w \notin \mathbb{W}_0$, what happens to the robustness of nominal MPC? To answer this question, we introduce the set $\mathbb{W}_1$ to represent the *large* disturbances not included in $\mathbb{W}_0$ and consider the

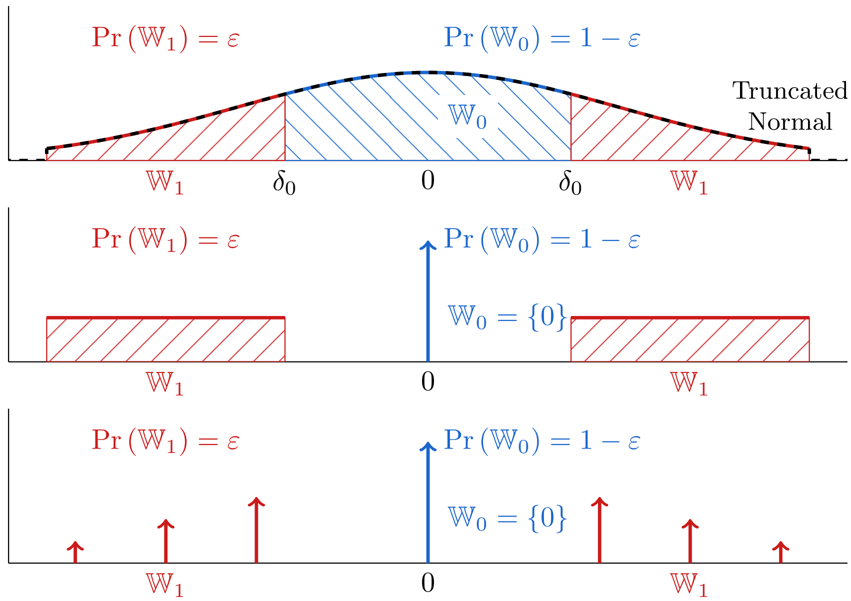

Figure 5.2: Three probability distributions for the disturbance w depicting the small, persistent disturbances ($\mathbb{W}_0$) and large, infrequent disturbances ($\mathbb{W}_1$).

probability

$$\mu(\mathbb{W}_1) = \Pr(w \in \mathbb{W}_1)$$

Let this set satisfy $\mathbb{W}_1 \cap \mathbb{W}_0 = \emptyset$ such that $\inf_{w \in \mathbb{W}_1} |w| > 0$, i.e., the large disturbances are "bounded away from zero." Note that $\mathbb{W}_1$ includes discrete-valued disturbances that may not be included in $\mathbb{W}_0$.

For example, consider the truncated normal distribution in the top plot of Figure 5.2. The disturbance w may take values that exceed δ_0, but these events are infrequent in that $\mu(\mathbb{W}_1)$ is small. This description also applies to the other distributions shown in Figure 5.2. In particular, we note the bottom plot in Figure 5.2 in which the disturbance takes only discrete values.

For the disturbances in $\mathbb{W}_1$, nominal MPC may not ensure deterministic robustness. In fact, there may not exist *any* controller that can ensure deterministic robustness for the system $f(\cdot)$ and disturbances considered in $\mathbb{W}_1$. As established in Corollary 2.22, deterministic robustness implies stochastic robustness, but the converse does not hold. Thus, there may exist a range of disturbances such that the system is robust in a stochastic context, but not in a deterministic context. We therefore investigate what conditions are required to guarantee that $\mathbb{W}_1$ is in that range.

Under suitable assumptions, we show that nominal MPC is robust to these large disturbances in a stochastic context provided these disturbances are sufficiently *infrequent*, i.e., $\mu(\mathbb{W}_1)$ is sufficiently small. This characterization includes disturbances such as faults, missing measurements, communication failures, breakdowns, large delays, and large price/demand spike in economic applications. If these large disturbances are sufficiently frequent, indeed no controller is expected to be robust in any sense against them.

Remark 5.15. Through Assumption 5.1, we require that the disturbances are i.i.d. and therefore $\mu(\mathbb{W}_1)$ is not time-varying. We can, however, extend the subsequent results to time-varying (but independent) probability distributions by allowing the distribution $\mu(\cdot)$ to vary with time, i.e., we consider $\mu(\mathbb{W}_1, k)$.

For clarity, we restrict attention to the case $\mathbb{W}_0 = \{0\}$ (e.g., the middle and bottom plots of Figure 5.2), such that there are only two possibilities: (i) the nominal behavior ($w = 0$) occurs with probability $1 - \mu(\mathbb{W}_1)$ or (ii) a large disturbance ($w \in \mathbb{W}_1$) occurs with probability $\mu(\mathbb{W}_1)$. Specifically, we require the following assumption.

Assumption 5.16 (Only large disturbances)**.** The disturbance set satisfies $\mathbb{W} = \mathbb{W}_0 \cup \mathbb{W}_1$ with $\mathbb{W}_0 = \{0\}$.

5.2.1 Motivating example

Consider the scalar system

$$x^+ = x + u + 2w$$

with $x \in \mathbb{R}$ and $u \in [-1, 1]$. We consider the controller $\kappa(x) := \text{sat}_{[-1,1]}(-3x/5)$ and use $x(k) = \phi(k; x, \mathbf{w}_k)$ to denote the closed-loop state trajectory. Note that this controller is defined for all $x \in \mathbb{R}$.

For this system, we have the Lyapunov function $V(x) := x^2$ such that the function $\alpha_3(\cdot) \in \mathcal{K}_\infty$ defined as

$$\alpha_3(|x|) := \begin{cases} \frac{21}{25}|x|^2 & |x| < \frac{5}{3} \\ 2|x| - 1 & |x| \geq 5/3 \end{cases}$$

satisfies

$$V(f(x, \kappa(x), 0)) \leq V(x) - \alpha_3(|x|)$$

for all $x \in \mathbb{R}$. Thus, the origin of the closed-loop system is asymptotically stable.

If w is assumed to be a continuous-valued random variable with $|w| < 0.5$, then the closed-loop system is robustly asymptotically stable (RAS), i.e., $\delta_0 < 0.5$. If however $w \in \mathbb{W} := \{0, 1\}$, i.e., $\mathbb{W}_1 = \{1\}$ is a large (discrete-valued) disturbance, then there exists a worst-case scenario in which $w = 1$ at every time and the system moves further away from $x = 0$ at each step. The system is not robust to this discrete-valued disturbance in the usual deterministic sense (RAS).

We now consider that this large disturbance is also infrequent in that

$$\varepsilon = \Pr(w = 1) = \mu(\mathbb{W}_1)$$

for some $\varepsilon \in (0, 1)$. We simulate the closed-loop system subject to this stochastic disturbance from $x(0) = 30$ and $\varepsilon = 0.4$. The results of 50 realizations of the disturbance trajectory are plotted in Figure 5.3. Note that these realizations do not admit a deterministic upper bound for $|x(k)|$ as $k \to \infty$. Given a sufficient number of time steps, the probability that $|x(k)|$ exceeds any finite upper bound is nonzero.

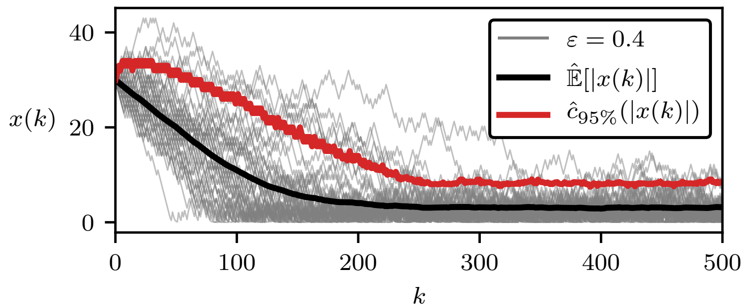

Figure 5.3: Closed-loop state for 50 realizations of the disturbance trajectory with $\varepsilon = 0.4$. The sample average and 95% confidence bound for the norm of the closed-loop state are evaluated for 1000 realizations of the disturbance trajectory.

We then calculate and plot the sample average of the norm of the closed-loop state trajectory at each k, denoted $\hat{\mathbb{E}}[|x(k)|]$, in Figure 5.3 for 1000 realizations of the disturbance trajectory. Note that $\hat{\mathbb{E}}[|x(k)|]$ converges to a constant value as $k \to \infty$ and appears to admit a finite upper bound for all k, consistent with the definition of RASiE considered in previous chapters. In addition to the sample average, we also calculate a confidence interval for this trajectory. Specifically, we define the sample 95%-confidence trajectory $\hat{c}_{95\%}(|x(k)|)$ at each k as the minimum bound for $|x(k)|$ that holds for at least 95% of the disturbance trajectory realizations considered (950 disturbance realizations in this example). We also plot this trajectory in Figure 5.3. Observe that this 95%-confidence trajectory converges to a constant as $k \to \infty$ and appears to admit a finite upper bound.

In Figure 5.4, we consider multiple values of $\varepsilon = \Pr(w = 1)$ and plot the sample average and 95%-confidence trajectories for these different values of ε. For $\varepsilon < 0.5$, both the sample average and 95%-confidence trajectories exhibit behavior similar to $\varepsilon = 0.4$ and converge to a constant value as $k \to \infty$. We construct an upper bound for this constant value by defining

$$\hat{\gamma}(\varepsilon) := \max\left\{\hat{\mathbb{E}}[|x(k)|] : k \in \mathbb{I}_{400:500}\right\}$$

$$\hat{\gamma}_{95\%}(\varepsilon) := \max\left\{\hat{c}_{95\%}(|x(k)|) : k \in \mathbb{I}_{400:500}\right\}$$

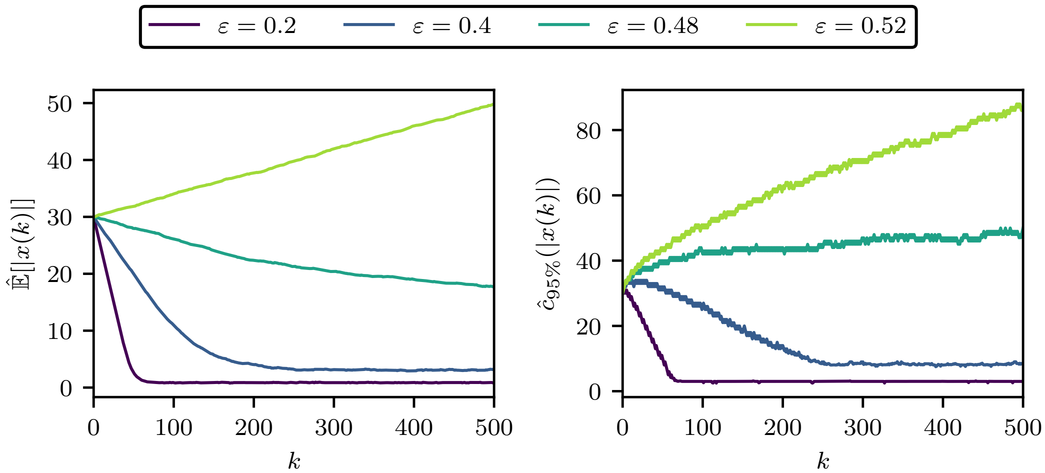

Figure 5.4: The expected value, denoted $\mathbb{E}[|x(k)|]$, and 95% confidence bound, denoted $\hat{c}_{95\%}(|x(k)|$, for the norm of the closed-loop trajectory for different values of ε.

We plot these values in Figure 5.5 across a range of $\varepsilon \in (0, 0.5)$. Note that both $\hat{\gamma}(\varepsilon)$ and $\hat{\gamma}_{95\%}(\varepsilon)$ exhibit behavior that is consistent with $\mathcal{K}$-functions, i.e., these functions increase with increasing ε and converge to zero as $\varepsilon \to 0$.

These results suggest that the closed-loop system admits an ISS-type bound for both the expected value $\mathbb{E}[|x(k)|]$ and 95%-confidence bound $c_{95\%}(|x(k)|)$. Note that we remove the "hat" notation to indicate that these are exact quantities and not sampling based approximation. Specifically, we claim that there exists $\beta(\cdot), \beta_{95\%}(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot), \gamma_{95\%}(\cdot) \in \mathcal{K}$ such that

$$\mathbb{E}[|x(k)|] \leq \beta(|x(0)|, k) + \gamma(\varepsilon) \quad (5.5)$$

$$c_{95\%}(|x(k)|) \leq \beta_{95\%}(|x(0)|, k) + \gamma_{95\%}(\varepsilon) \quad (5.6)$$

for all $x(0) \in \mathbb{R}$, $k \in \mathbb{I}_{\geq 0}$, and $\varepsilon < 0.5$. We can equivalently write (5.6) as

$$\Pr\Big(|x(k)| \leq \beta(|x(0)|, k) + \gamma(\varepsilon)\Big) \geq 0.95 \quad (5.7)$$

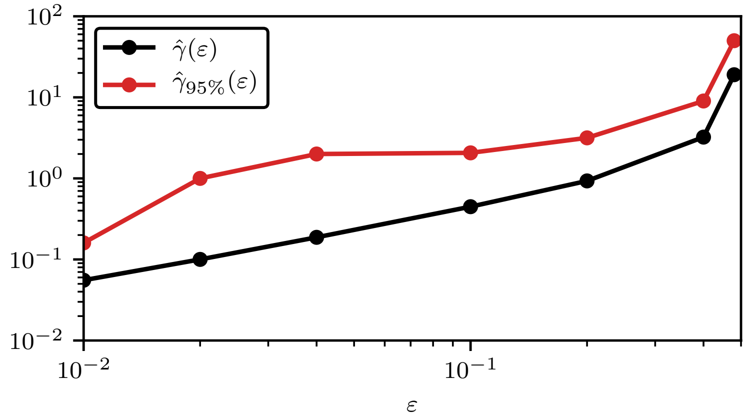

Figure 5.5: Plot of the maximum value of $\mathbb{E}[|x(k)|]$ for $k \in [400, 500]$, denoted $\hat{\gamma}(\varepsilon)$, and the maximum value of $\hat{c}_{95\%}(|x(k)|$ for $k \in [400, 500]$, denoted $\hat{\gamma}_{95\%}(\varepsilon)$.

i.e., for each $k \in \mathbb{I}_{\geq 0}$ the closed-loop system satisfies this bound with a probability of at least 95%.

We do not, however, expect this bound to hold for disturbances that occur with high probability. For $\varepsilon = 0.52$, we see in Figure 5.4 that the sample average and 95%-confidence trajectories diverge as $k \to \infty$ and do not admit any ISS-type bound. Thus, there exists some $\delta > 0$ such that for any $\varepsilon \leq \delta$, we obtain the bounds in (5.5) and (5.7). If $\varepsilon > \delta$, these bounds do not hold. This value of δ now describes the margin of robustness for the system to these large disturbances. For this example, we observe that $\delta < 0.5$.

5.2.2 Definitions of robustness

For large and infrequent disturbances, we consider two stochastic definitions of robustness. These definitions use bounds that depend on a stochastic property of the disturbance. In Chapters 2 and 3, this stochastic property is the covariance of the disturbance Σ, in Chapter 4 we use $\mathbb{E}[|w|]$, and in this chapter we consider $\mu(\mathbb{W}_1) = \Pr(w \in \mathbb{W}_1)$. We note that $\mu(\mathbb{W}_1)$ is related to $\mathbb{E}[|w|]$ via the following bounds.

$$\frac{\mathbb{E}[|w|]}{\sup_{w\in\mathbb{W}_1}|w|} \leq \mu(\mathbb{W}_1) \leq \frac{\mathbb{E}[|w|]}{\inf_{w\in\mathbb{W}_1}|w|}$$

Recall that $\inf_{w\in\mathbb{W}_1}|w| > 0$ in the definition of $\mathbb{W}_1$ and $\mathbb{W}$ is compact.

In previous chapters, we constrained the maximum size of the disturbances included in the robustness analysis, but allowed for any probability distribution on this set of potential disturbances. For large and infrequent disturbances, we instead consider a fixed set $\mathbb{W}_1$ and constrain the probability distribution for the disturbance through the value of $\mu(\mathbb{W}_1) = \Pr(w \in \mathbb{W}_1)$. For some $\delta \in [0,1]$, let $\mathcal{M}(\mathbb{W},\delta)$ denote the set of probability measures on the domain $\mathcal{B}(\mathbb{W})$ such that $\mu(\mathbb{W}_1) \in [0,\delta]$ for all $\mu \in \mathcal{M}(\mathbb{W},\delta)$. We define robust asymptotic stability in expectation (RASiE) for large and infrequent disturbances as follows.

Definition 5.17 (RASiE for large, infrequent disturbances)**.** The origin is RASiE for large, infrequent disturbances for a system $x^+ = f(x,\kappa(x,t),w,t)$, $w \in \mathbb{W}$ and $\delta > 0$ in an RPI sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$ if there exist $\beta(\cdot) \in \mathcal{KL}$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$\mathbb{E}[|\phi(k;x,\mathbf{w}_{t:k},t)|] \leq \beta(|x|,k-t) + \gamma(\mu(\mathbb{W}_1)) \tag{5.8}$$

for all $x \in \mathcal{X}(t)$, $\mu \in \mathcal{M}(\mathbb{W},\delta)$, $t \in \mathbb{I}_{\geq 0}$, and $k \in \mathbb{I}_{\geq t}$.

Definition 5.17 ensures that for infrequent disturbances ($\mu(\mathbb{W}_1) \leq \delta$) we have the bound in (5.8). This definition is consistent with the observations from the motivating example. For $\mu(\mathbb{W}_1) < 0.5$, the closed-loop system in the motivating example can, on average, recover from the disturbance before another disturbance occurs and therefore satisfies the bound in (5.8). For $\mu(\mathbb{W}_1) > 0.5$, however, the value of $\mathbb{E}[|\phi(k;x,\mathbf{w}_{t:k},t)|]$ diverges as $k \to \infty$. We can similarly define an SISS Lyapunov function for systems subject to these large and infrequent disturbances.

Definition 5.18 (SISS Lyapunov function for large, infrequent disturbances)**.** A function $V(\cdot,t) : \mathcal{X}(t) \to \mathbb{R}_{\geq 0}$ is an SISS Lyapunov function for large, infrequent disturbances for a system $x^+ = f(x,\kappa(x,t),w,t)$, $w \in \mathbb{W}$ and $\delta > 0$ in an RPI sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$ if there exist $\alpha_1(\cdot), \alpha_2(\cdot), \alpha_3(\cdot) \in \mathcal{K}_\infty$ and $\sigma(\cdot) \in \mathcal{K}$ such that

$$\alpha_1(|x|) \leq V(x,t) \leq \alpha_2(|x|)$$
$$\int_{\mathbb{W}} V(f(x,\kappa(x,t),w,t))d\mu(w) \leq V(x) - \alpha_3(|x|) + \sigma(\mu(\mathbb{W}_1))$$

for all $x \in \mathcal{X}(t)$, $\mu \in \mathcal{M}(\mathbb{W},\delta)$, and $t \in \mathbb{I}_{\geq 0}$.

In Proposition 2.18, we established that an SISS Lyapunov function is a sufficient condition for RASiE if the state of the system is *bounded*. The proof of Proposition 2.18 relied on the fact that $\mathcal{X}$ is bounded to construct convex lower bounds for the required $\mathcal{K}_\infty$-functions and apply Jensen's inequality. For many physical systems, $\mathcal{X}$ is bounded, particularly if the terminal set is also bounded. For the time-varying case, we say that the sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$ is bounded if the set $\cup_{t=0}^{\infty}\mathcal{X}(t)$ is bounded. We therefore have the following result as a straightforward extension of Proposition 2.18 to time-varying systems and large disturbances.

Proposition 5.19. *If a system $x^+ = f(x,\kappa(x,t),w,t)$, $w \in \mathbb{W}$ admits an SISS Lyapunov function for large, infrequent disturbances in the RPI and bounded sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$, then the origin is RASiE for large, infrequent disturbance in $(\mathcal{X}(t))_{t=0}^{\infty}$.*

Production planning and scheduling problems, however, are the exception to this rule. These formulations often consider variables such as backlog, representing unmet order volume, that are technically unbounded. In practice, a facility cannot accumulate infinite backlog[1], but there is no specific upper bound for this state. Moreover, simulation studies of

[1]Without first going out of business

these production scheduling problems should allow for the possibility of diverging backlog as it indicates insufficient production capacity when subject to disturbances. Thus, requiring bounded $\mathcal{X}(t)$ is undesirable for the analysis of production planning and scheduling applications of MPC.

Without the assumption of bounded $\mathcal{X}(t)$, the SISS Lyapunov function in Definition 5.18 is not sufficient to establish RASiE. To see why unbounded $\mathcal{X}(t)$ presents such a challenge, we note that seemingly benign distributions can nonetheless produce infinite expected value on unbounded supports. For example, we consider a Pareto distribution given by the cumulative distribution function $F(x) = 1 - \frac{1}{x}$ for all $x \geq 1$. This type of distribution, originally used to describe the allocation of wealth in society, has been used to model a variety of sociological and physical phenomenon. Confidence intervals for this distribution are well defined, but the expected value of a random variable described by this cumulative distribution is *infinite*.

To address unbounded $\mathcal{X}(t)$, we weaken the definition of stochastic robustness used in this section to consider confidence intervals instead of expected value of the closed-loop trajectory. Specifically, we define robust asymptotic stability in *probability* (RASiP) for large, infrequent disturbances.

Definition 5.20 (RASiP for large, infrequent disturbances)**.** The origin is RASiP for large, infrequent disturbances for a system $x^+ = f(x, \kappa(x,t), w, t)$, $w \in \mathbb{W}$ and $\delta > 0$ in a robustly positive invariant sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$ if for each $p \in (0,1)$ there exist $\beta_p(\cdot) \in \mathcal{KL}$ and $\gamma_p(\cdot) \in \mathcal{K}$ such that

$$\Pr\Big(|\phi(k; x, \mathbf{w}_{t:k}, t)| \leq \beta_p(|x|, k-t) + \gamma_p(\mu(\mathbb{W}_1))\Big) \geq p \tag{5.9}$$

for all $x \in \mathcal{X}(t)$, $\mu \in \mathcal{M}(\mathbb{W}, \delta)$, $t \in \mathbb{I}_{\geq 0}$, and $k \in \mathbb{I}_{\geq t}$.

Definition 5.20 ensures that for sufficiently infrequent disturbances, $\mu(\mathbb{W}_1) \leq \delta$, we can construct the bound in (5.9) for any confidence level $p \in (0,1)$. In the motivating example, we considered $p = 0.95$ and the results in this example are again consistent with Definition 5.20. We can show that RASiE implies RASiP via Markov's inequality, but the converse is not true in general. This observation supports the claim that RASiP is a weaker property that RASiE. Furthermore, the SISS Lyapunov function in Definition 5.18 is a sufficient condition for RASiP without bounded $\mathcal{X}(t)$.

Proposition 5.21. *If a system $x^+ = f(x, \kappa(x,t), w, t)$, $w \in \mathbb{W}$ admits an SISS Lyapunov function for large, infrequent disturbances in the RPI sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$, then the origin is RASiP for large, infrequent disturbance in $(\mathcal{X}(t))_{t=0}^{\infty}$.*

The proof Proposition 5.21 uses a similar approach to the proof of Proposition 2.18, but requires a significant amount of additional mathematical details to construct the function $\beta_p(\cdot) \in \mathcal{KL}$. In particular, this result relies on the work of Teel and co-workers in a series of publications that detail necessary and sufficient conditions for global asymptotic stability in probability (Teel, 2013; Teel et al., 2013, 2014). Some useful results from these works are summarized in McAllister and Rawlings (2020) and a proof of Proposition 5.21 is given in Mcallister and Rawlings (2021, Prop. 3).

5.3 Robustness of MPC to large, infrequent disturbances

5.3.1 Additional assumptions

Although we refer to these disturbances as large, we do not allow disturbances of arbitrary size. In particular, we require that the MPC remains feasible for the closed-loop trajectory subject to these large disturbances.

Assumption 5.22 (Robust recursive feasibility)**.** The sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$ is robustly positive invariant for the closed-loop system $x^+ = f(x, \kappa(x,t), w, t)$, $w \in \mathbb{W}$.

Note that assuming recursive feasibility for MPC is sometimes inappropriate. In previous chapters, we established robust recursive feasibility by removing state constraints, enforcing stronger restrictions on the terminal set, and requiring the disturbances to be sufficiently small. For large disturbances, we no longer have this ability and must instead choose the constraints in the optimization problem with care. In general, the MPC problem must be recursively feasible *by design* for the specific problem and large disturbance of interest via careful choices for the state and terminal constraints. For example, we still cannot enforce state constraints that represent desired goals for the system as these constraint may still produce infeasible optimization problems.

For production scheduling applications of MPC, Assumption 5.22 is often reasonable. The input and state constraints in the problem formulation are used only to enforce realistic decisions, e.g., nonnegative inventory, and disturbances cannot force violations of these constraints. Sufficiently long horizons and reasonable terminal constraints ensure that any possible state for the facility can be driven to the terminal region with N inputs. We discuss these production scheduling formulations in more detail in the subsequent chapter.

MPC implementations without state or terminal constraints easily satisfy Assumption 5.22 since the feasible set is then $\mathcal{X}(t) = \mathbb{R}^n$ for all $t \in \mathbb{I}_{\geq 0}$. These formulations can also ensure nominal (practical) asymptotic stability with suitable dissipativity assumptions and terminal costs (Grüne and Stieler, 2014; Limon et al., 2006; Zanon and Faulwasser, 2018). Unfortunately, these dissipativity assumptions are difficult to verify for nonlinear systems and often these MPC formulations are limited to practical asymptotic stability, in which the nominal system is guaranteed to converge to only some region around the origin. Nonetheless, there exists a significant class of MPC implementations that satisfy Assumption 5.22, particularly for the higher level planning and scheduling problems.

In addition to feasibility, we also require a bound on the cost function subject to these large disturbances.

Assumption 5.23 (Maximum cost increase)**.** There exist $b_1, b_2 \geq 0$ such that

$$V^0(f(x, \kappa(x,t), w, t), t+1) \leq V^0(x,t) + b_1 \ell(x, \kappa(x,t), t) + b_2 \tag{5.10}$$

for all $x \in \mathcal{X}(t)$, $w \in \mathbb{W}_1$, and $t \in \mathbb{I}_{\geq 0}$.

The bound in (5.10) is notably weaker than the bound required for an ISS Lyapunov function. The constant in front of the stage cost is *positive* in (5.10) and the optimal cost may therefore increase at a rate proportional to the current stage cost. If the set $\cup_{t=0}^{\infty}\mathcal{X}(t)$ is bounded, Assumption 5.23 is satisfied for $b_1 = 0$ and some large, finite $b_2 > 0$. However, we are also interested in applications for which $\cup_{t=0}^{\infty}\mathcal{X}(t)$ is not bounded. For certain important cases, we can show that Assumption 5.23 does hold without bounded $\cup_{t=0}^{\infty}\mathcal{X}(t)$. For example, we have the following result for MPC formulations with exponential cost bounds.

Lemma 5.24. *Let Assumptions 5.1 to 5.4, 5.16 and 5.22 with* $a, c_1, c_2 > 0$ *such that*

$$c_1|x|^a \leq \ell(x,u,t) \qquad V^0(x,t) \leq c_2|x|^2$$

for all $(x,u) \in \mathbb{Z}(t)$, $x \in \mathcal{X}(t)$, *and* $t \in \mathbb{I}_{\geq 0}$. *If there exist* $e_1, e_2 \geq 0$ *such that*

$$|f(x,u,w,t) - f(x,u,0,t)| \leq e_1|x| + e_2$$

for all $(x,u) \in \mathbb{Z}(t)$, $w \in \mathbb{W}_1$, *and* $t \in \mathbb{I}_{\geq 0}$, *then Assumption 5.23 holds.*

Proof. Choose $x \in \mathcal{X}(t)$ and $t \in \mathbb{I}_{\geq 0}$. We use Lemma 5.9 and the fact that $c_1|x| \leq \ell(x, \kappa(x,t), t) \leq$

$V^0(x,t)$ to give

$$\begin{aligned} c_1|f(x,\kappa(x,t),0,t)|^a &\leq V^0(f(x,\kappa(x,t),0,t)) \\ &\leq V^0(x) - \ell(x,\kappa(x,t),t) \\ &\leq (c_2 - c_1)|x|^a \end{aligned} \tag{5.11}$$

We have from the upper bound on $V^0(\cdot)$ that

$$\begin{aligned} V^0(f(x,\kappa(x,t),w,t),t+1) &\leq c_2|f(x,\kappa(x,t),w,t)|^a \\ &\leq c_2\left(|f(x,\kappa(x,t),0,t)| + e_1|x| + e_2\right)^a \\ &\leq 2^a c_2|f(x,\kappa(x,t),0,t)|^a + (4e_1)^a|x|^a + (4e_2)^a \end{aligned}$$

We substitute in (5.11) to give

$$V^0(f(x,\kappa(x,t),w,t),t+1) \leq \tilde{b}_1|x|^a + b_2$$

in which $\tilde{b}_1 := 2^a c_2((c_2 - c_1)/c_1 + 2^a e_1^a)$ and $b_2 = c_2(4e_2)^a$. We again use the fact that $c_1|x| \leq \ell(x,\kappa(x,t),t) \leq V^0(x,t)$ to complete the proof with $b_1 := \tilde{b}_1/c_1 - 1$. ☐

The exponential bounds on the stage cost and optimal cost in Lemma 5.24 are stronger than the bounds required by Assumptions 5.5 and 5.6. In fact, an MPC problem that satisfies these exponential bounds produces a nominal closed-loop system that is *exponentially* stable. MPC formulations with suitable terminal conditions and quadratic costs satisfy these bounds Rawlings et al. (2020, s. 2.4), but constructing a terminal constraint that satisfies both Assumption 5.4 and Assumption 5.22 for the large disturbance of interest may prove difficult. For the specific case of stable linear systems without state constraints, we can construct a global quadratic Lyapunov function for the terminal cost and thereby remove the terminal constraint citet[s. 2.5.3]rawlings:mayne:diehl:2020. Another method to verify Assumption 5.23 that does not require a nonnegative stage cost is available in Mcallister and Rawlings (2021, Lemma 7).

5.3.2 Main results

For economic applications of nominal MPC, i.e., without Assumptions 5.5 and 5.6, we can establish the following result for the robustness of MPC to large and infrequent disturbances.

Theorem 5.25 (Robust performance for large, infrequent disturbances)**.** *Let Assumptions 5.1 to 5.4, 5.16, 5.22 and 5.23 hold. Then there exist* $\delta > 0$ *and* $\bar{\gamma}(\cdot) \in \mathcal{K}$ *such that*

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{k=t}^{t+T-1} \mathbb{E}[\ell(x(k), \kappa(x(k), k), k)] \leq \bar{\gamma}(\mu(\mathbb{W}_1)) \tag{5.12}$$

in which $x(k) = \phi(k; x, \mathbf{w}_{t:k}, t)$ *for all* $x \in \mathcal{X}(t)$*,* $\mu \in \mathcal{M}(\mathbb{W}, \delta)$*, and* $t \in \mathbb{I}_{\geq 0}$*.*

Proof. Choose $x \in \mathcal{X}(t)$ and $t \in \mathbb{I}_{\geq 0}$. If $w = 0$, we have from Lemma 5.24 that

$$V^0(f(x, \kappa(x,t), 0, t), t+1) \leq V^0(x,t) - \ell(x, \kappa(x,t), t)$$

We then combine this bound with Assumption 5.23 using the indicator function for $\mathbb{W}_1$, i.e.,

$$\begin{aligned} V^0(f(x, \kappa(x,t), w, t), t+1) \leq V^0(x,t) - (1 - I_{\mathbb{W}_1}(w))\ell(x, \kappa(x,t), t) \\ + I_{\mathbb{W}_1}(w)\left(b_1 \ell(x, \kappa(x,t), t) + b_2\right) \end{aligned}$$

Taking the expected value and combining terms give

$$\mathbb{E}[V^0(f(x, \kappa(x,t), w, t), t+1)] \leq V^0(x,t) - (1 - \mu(\mathbb{W}_1) - b_1\mu(\mathbb{W}_1))\ell(x, \kappa(x,t), t) + b_2\mu(\mathbb{W}_1)$$

We choose $0 < \delta < 1/(1+b_1)$ which gives

$$\mathbb{E}[V^0(f(x,\kappa(x,t),w,t),t+1)] \leq V^0(x,t) - b_3\ell(x,\kappa(x,t),t) + b_2\mu(\mathbb{W}_1) \tag{5.13}$$

in which $b_3 := (1-(1+b_1)\delta) > 0$. Since the choice of x and t was arbitrary, (5.13) holds for any $x \in \mathcal{X}(t)$ and $t \in \mathbb{I}_{\geq 0}$.

For any initial $x \in \mathcal{X}(t)$ and $t \in \mathbb{I}_{\geq 0}$, we denote the closed-loop trajectory $x(k) = \phi(k; x, \mathbf{w}_{t:k}, t)$ and the input $u(k) = \kappa(x(k), k)$ for all $k \in \mathbb{I}_{\geq t}$. From (5.13) and the law of total expectation we may write

$$\mathbb{E}[V^0(x(k+1),k+1)] \leq \mathbb{E}[V^0(x(k),k)] - b_3\mathbb{E}[\ell(x(k),u(k),k)] + b_2\mu(\mathbb{W}_1)$$

for all $k \in \mathbb{I}_{\geq t}$. We sum both sides of this inequality from t to $t+T-1$ with $T \in \mathbb{I}_{\geq 1}$, divide by T, and rearrange to give

$$\frac{b_3}{T}\sum_{k=t}^{t+T-1}\mathbb{E}[\ell(x(k),u(k),k)] \leq \frac{V^0(x(t),t) - \mathbb{E}[V^0(x(T),T)]}{T} + b_2\mu(\mathbb{W}_1)$$

Since $\ell(\cdot)$ is bounded from below (Assumption 5.2) and $V_f(\cdot) \geq 0$, there exists finite $M \in \mathbb{R}$ such that $V^0(x(T),T) \geq M$ and therefore

$$\frac{b_3}{T}\sum_{k=t}^{t+T-1}\mathbb{E}[\ell(x(k),u(k),k)] \leq \frac{V^0(x(t),t) - M}{b_3 T} + \bar{\gamma}(\mu(\mathbb{W}_1))$$

in which $\bar{\gamma}(\mu(\mathbb{W}_1)) := (b_2/b_3)\mu(\mathbb{W}_1) \in \mathcal{K}$. We take the limit supremum as $T \to \infty$ such that the initial cost and M vanish to give (5.12) □

We note that (5.12) holds for only $\mu(\mathbb{W}_1) \leq \delta$ for some small $\delta > 0$, i.e., for only sufficiently infrequent disturbances. Within this range of sufficiently infrequent disturbances, the upper bound (5.12) increases with increasing $\mu(\mathbb{W}_1)$. Conversely, as $\mu(\mathbb{W}_1) \to 0$ we recover the nominal guarantee for the system in Theorem 5.10. If we also include Assumptions 5.5 and 5.6, nominal MPC renders the origin RASiP to large and infrequent disturbances.

Theorem 5.26 (Robustness for large, infrequent disturbances)**.** *Let Assumptions 5.1 to 5.6, 5.16, 5.22 and 5.23 hold. Then there exists $\delta > 0$ such that the origin is RASiP for the system $x^+ = f(x, \kappa(x,t), w, t)$, $w \in \mathbb{W}$ in the sequence of robustly positive invariant sets $(\mathcal{X}(t))_{t=0}^{\infty}$ for all $\mu \in \mathcal{M}(\mathbb{W}, \delta)$. If the set $\cup_{t=0}^{\infty}\mathcal{X}(t)$ is bounded, the origin is also RASiE.*

Proof. By applying the lower bound in Assumption 5.5 to (5.13), there exists $\alpha_\ell(\cdot) \in \mathcal{K}_\infty$ and $\sigma(s) = b_2 s \in \mathcal{K}$ such that

$$\mathbb{E}[V^0(f(x, \kappa(x,t), w, t), t+1)] \leq V^0(x,t) - b_3\alpha_\ell(|x|) + \sigma(\mu(\mathbb{W}_1))$$

for all $\mu(\mathbb{W}_1) \leq \delta$ and some $\delta > 0$ that satisfies $\delta < 1/(1+b_1)$. There also exists $\alpha_2(\cdot) \in \mathcal{K}_\infty$ from Assumption 5.6 such that

$$\alpha_\ell(|x|) \leq \ell(x, \kappa(x,t), t) \leq V^0(x,t) \leq \alpha_2(|x|)$$

Thus, $V^0(\cdot)$ is an SISS Lyapunov function for large, infrequent disturbances and the origin is RASiP by Proposition 5.21. □

Theorem 5.26 ensure that there exists some margin of robustness $\delta > 0$ for these large and infrequent disturbances. Thus, MPC ensures that the system can recover from these large disturbances provided the probability of them occurring is sufficiently small. For sufficiently frequent disturbances ($\mu(\mathbb{W}_1) > \delta$), however, the origin may not be RASiP.

The value of δ in Theorems 5.25 and 5.26 is defined as $\delta < 1/(1+b_1)$ in which b_1 is defined in Assumption 5.23. As $b_1 \to \infty$, we have that $\delta \to 0$ and the system is no longer robust to large disturbances with nonzero probability. Conversely, as $b_1 \to 0$, we have that $\delta \to 1$ and the system is robust to large disturbances that occur with any probability less than one, i.e., $\mu(\mathbb{W}_1) < 1$.

If $\cup_{t=0}^{\infty}\mathcal{X}(t)$ is bounded and $(\mathcal{X}(t))_{t=0}^{\infty}$ satisfy Assumption 5.22, we can choose $b_1 = 0$ and b_2 sufficiently large such that Assumption 5.23 holds. Moreover, we can strength RASiP to RASiE with Proposition 5.19 and obtain a result that is remarkable similar to stochastic robustness of small and persistent disturbances. Specifically, we have the following corollary to Theorem 5.26 in which RASiE holds for any $\delta \in (0,1)$.

Corollary 5.27. *Let Assumptions 5.1 to 5.6, 5.16 and 5.22 hold and assume* $\cup_{t=0}^{\infty}\mathcal{X}(t)$ *is bounded. Then for any* $\delta \in (0,1)$*, the origin is RASiE for the system* $x^+ = f(x, \kappa(x,t), w, t)$*,* $w \in \mathbb{W}$ *in the robustly positive sequence of sets* $(\mathcal{X}(t))_{t=0}^{\infty}$ *for all* $\mu \in \mathcal{M}(\mathbb{W}, \delta)$*.*

5.4 Examples

5.4.1 Time-invariant reactor scheduling

We first consider a time-invariant example of product blending adapted from McAllister and Rawlings (2021a) and similar to an example in Rawlings and Risbeck (2017). Two batch reactors with different specifications deliver product to a single blending tank. Product is withdrawn from this tank at the start of each hour. The goal of the controller is to maintain the total volume and concentration of the product species in the blending tank while delivering a specific volume downstream every hour. A diagram of this example problem is shown in Figure 5.6.

Let v_T denote the volume of fluid in the tank and m_T denote the mass of product species in the tank. Each reactor can be set to produce a different volume of fluid and mass of product species each hour, within some available range specific to each reactor. The volume of the

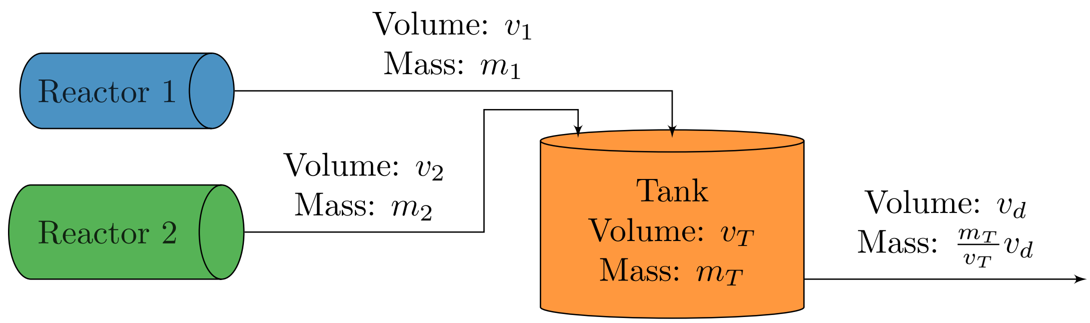

Figure 5.6: Diagram of time-invariant reactor scheduling problem.

fluid and mass of product species delivered from reactors 1 and 2 to the blending tank each hour is denoted v_1, m_1 and v_2, m_2, respectively. We plot the constraints for these reactors and the blending tank in Figure 5.7. To enforce minimum capacity requirements of each reactor we use the binary variables z_1 and z_2 to indicate if the reactors are 'on' or 'off'. We also allow the controller to select the quantity of volume withdrawn from the tank, which we denote v_d. The discrete time model of the facility is

$$v_T^+ = v_T + v_1(1 - d_1) + v_2(1 - d_2) - v_d$$
$$m_T^+ = m_T + m_1(1 - d_1) + m_2(1 - d_2) - \frac{m_T}{v_T} v_d$$

In addition to these variables, we also include the binary disturbance variables d_1 and d_2 that represent breakdowns or unplanned maintenance of reactors 1 and 2. We also require that $v_d \in [0, 10]$ and $v_d \leq v_T$ to ensure that we do not withdraw more material from the blending tank than is available at the start of each hour. Thus, we have the constraints $(x, u) \in \mathbb{Z} \subseteq \mathbb{X} \times \mathbb{U}$.

The system has two states $\bar{x} = (v_T, m_T)$, seven inputs $\bar{u} = (z_1, v_1, m_1, z_2, v_2, m_2, v_d)$, and two *binary* disturbances $w = (d_1, d_2)$. The steady-state target is $\bar{x}_r = (15, 4.5)$ and $\bar{u}_r = (1, 2, 0.35, 1, 5, 1.75, 7)$. This steady-state target is shown with black dots in Figure 5.7. Note that this steady-state target is equivalent to a time-invariant reference trajectory. We

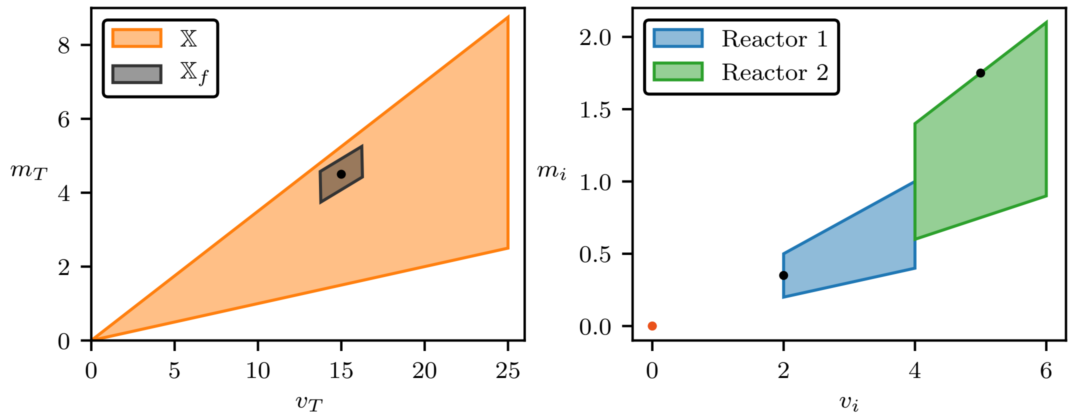

Figure 5.7: State and input constraints for reactor scheduling example. Steady-state values are indicated with black dots.

denote the shifted state and input as $x := \bar{x} - \bar{x}_r$ and $u := \bar{u} - \bar{u}_r$. We define the time-invariant stage cost as $\ell(x, u) := x'Qx + u'Ru$ with diagonal $Q := \text{diag}([1, 1])$ and $R := \text{diag}([0, 0.25, 0.5, 0, 0.25, 0.5, 2])$.

To construct the terminal cost and constraint, we linearize the system and consider only m_1 and v_2 as free inputs. We also choose $m_2 = \rho_{max} v_2$ such that the input constraint for m_2 remains active and fix all other inputs to their steady-state values. We then determine the LQR solution P and corresponding state-feedback gain K for the reduced system with only two free inputs. This procedure produces the terminal control law $\kappa_f(x) := Kx$ and terminal cost $V_f(x) := x'Px$. Since the terminal control law and input constraints (for the free inputs m_1 and v_2) are linear, we can construct the terminal set

$$\mathbb{X}_f := \{x \in \mathbb{X} : \kappa_f(x) \in \mathbb{U}\}$$

as shown in Figure 5.7. We verify that this choice of terminal cost and constraint satisfy Assumption 5.4. Since Q is positive definite and $\mathbb{X}_f$ contains the origin in its interior, Assumptions 5.5 and 5.6 hold. We choose a horizon of $N = 6$ such that $\mathcal{X}$ is RPI for this system. We also note that the system satisfies the conditions of Lemma 5.24 and therefore Assump-

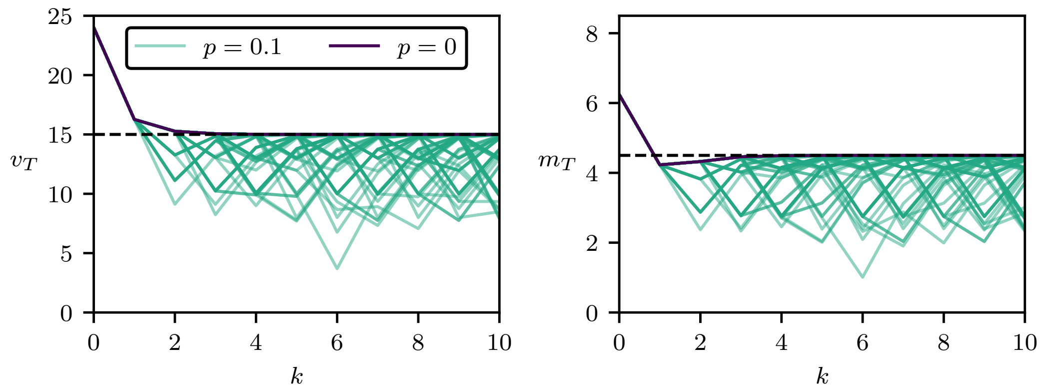

Figure 5.8: Closed-loop trajectories for 50 realizations of the disturbance with $p = 0.1$ for the reactor scheduling example. The black line indicates the nominal trajectory ($p = 0$) for comparison.

tion 5.23 holds as well.[2] Since $\mathcal{X}$ is bounded, Theorem 5.26 ensures that the origin is RASiE as well as RASiP to large, infrequent disturbances for this example.

We simulate the closed-loop trajectory for this system subject to breakdowns that occur with some probability p for each reactor, i.e., $\Pr(d_1 = 1) = \Pr(d_2 = 1) = p$. Thus, $\mu(\mathbb{W}_1) = p^2 + 2p(1-p)$. We initialize the system at $x(0) = (24, 6.24)$ and consider 100 realizations of the disturbance trajectory. In Figure 5.8, we plot these trajectories for $p = 0.1$. We then calculate the sample mean of $|x|$ for 100 realizations of the closed-loop trajectory with different values of p and plot the results in Figure 5.9. We observe an initial decrease in the sample average as the effect of the initial condition vanishes. After this initial decrease, the sample average trajectory satisfies a finite upper bound that increases with increasing p ($\mu(\mathbb{W}_1)$) and decays to zero as $p \to 0$. These results are consistent with RASiE and therefore Theorem 5.26.

5.4.2 Production scheduling

We now consider a simple scheduling example that also serves as a light introduction to the topics in the subsequent chapter. The goal is to meet the demand for material 1 (M1) by

[2] We can also verify that Assumption 5.23 holds by noting that $\mathcal{X}$ is bounded.

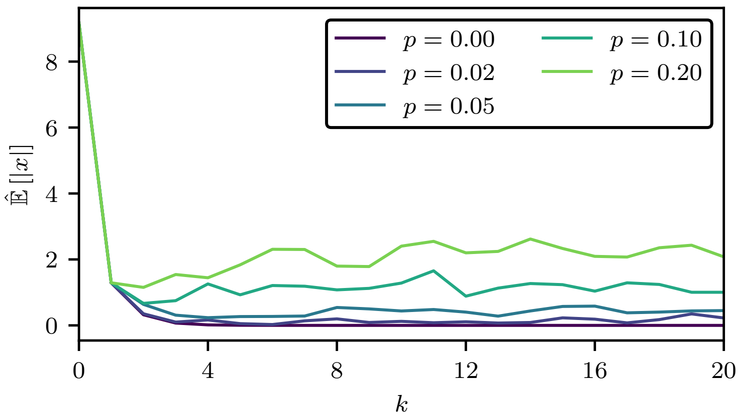

Figure 5.9: Sample mean of the norm of the closed-loop state trajectory for the reactor scheduling example with different values of p.

converting raw material (assumed to be in abundant supply) to M1 through task 1 (T1). Task 1 can produce between 5 and 16 units of M1 and has a nominal processing time of 2 time steps. The demand for M1 is 4 units per time step. If demand is not met, the facility accumulates backlog that must be offset at later times. The penalty for storing M1 is 10 (per unit per time step) and the penalty for maintaining backlog is 100 (per unit per time step).

To model this system, we use a state-space scheduling model (Subramanian et al., 2012). The binary decision variable W is unity if T1 starts at the current time step. We also define the continuous decision variable B that specifies the batch size, i.e., between 5 and 16 units, for T1. To track these decisions in the state of the system, we *lift* these variables via the state variables $\bar{W}_n, \bar{B}_n$ for $n \in \{0, 1, 2\}$. The integer n represents the number of time steps that task has progressed (e.g., if $\bar{W}_1 = 1$ then T1 has been running for 1 time step). We also consider a 1 time step delay in the progress of T1 as a potential disturbance. We use $Y \in \{0, 1\}$ to represent this disturbance an note that Y is inherently discrete-valued (large) due to the discrete-time grid. The dynamics for the task progress are given by

$$\begin{aligned}
\bar{W}_0^+ &= (W + \bar{W}_0)Y \\
\bar{W}_1^+ &= (W + \bar{W}_0)(1 - Y) + \bar{W}_1 Y \\
\bar{W}_2^+ &= W_1(1 - Y) \\
\bar{B}_0^+ &= (B + \bar{B}_0)Y \\
\bar{B}_1^+ &= (B + \bar{B}_0)(1 - Y) + \bar{B}_1 Y \\
\bar{B}_2^+ &= B_1(1 - Y)
\end{aligned}$$

Note that if $Y = 1$, the progress of the task does not advance.

Inventory and backlog (unmet demand) of M1 are denoted S and U, respectively. These states are integrators influenced by the batch size of T1 ending ($\bar{B}_2$), shipments to meet demand (H), and demand (4 units of M1 per time step). We also allow up to 1 unit M1 demand to be outsourced each time step and therefore removed from the facilities backlog. We use the decision variable C to represent this outsourcing/canceling of backlog. To discourage using this action, we assign a large penalty of 800 per unit to this decision. Thus, the dynamics for inventory and backlog are as follows.

$$\begin{aligned}
S^+ &= S + \bar{B}_2 - H \\
U^+ &= U + 4 - C - H
\end{aligned}$$

Next, we require constraints to enforce realism of the scheduling model. We require that $U \geq 0$, $S \in [0, 20]$, $W \in \{0, 1\}$, $C \in [0, 1]$, $H \in [0, 20]$, and the constraint

$$\bar{W}_0 + \bar{W}_1 \leq 1$$

to ensure that only one active task is running at each time step. We also require B to satisfy the min/max batch size constraints if T1 is starting.

$$5W \leq B \leq 16W$$

If T1 is not starting ($W = 0$), this constraint also requires that $B = 0$.

We now have a discrete-time, state-space representation of the system with

$$\begin{aligned} \bar{x} &= \left(\bar{W}_0, \bar{W}_1, \bar{W}_2, \bar{B}_0, \bar{B}_1, \bar{B}_2, S, U\right) \\ \bar{U} &= (W, B, H, C) \end{aligned}$$

and $w = (Y)$. The discrete-time system dynamics can be written as

$$\bar{x} = \bar{f}(\bar{x}, \bar{u}, w)$$

with the constraints $(\bar{x}, \bar{u}) \in \bar{\mathbb{Z}}$ and the stage cost

$$\bar{\ell}(\bar{x}, \bar{u}) = 10S + 100U + 800C$$

In the nominal case, the facility can meet demand while operating at 50% capacity. The optimal periodic solution is to start a new T1 every two time steps at a batch size of 8 units. Demand is met every time step and 4 units of material is held in storage every other time step. By repeating this periodic solution we can construct an infinite horizon reference trajectory. This reference trajectory is shown in Figure 5.10 and denoted $(\bar{x}_r(t))_{t=0}^{\infty}$ and $(\bar{u}_r(t))_{t=0}^{\infty}$.

Although the dynamics of the original variables are time-invariant[3], the reference trajectory is time-varying and the shifted problem is therefore time-varying. Thus, the shifted

[3]Demand and therefore the original dynamical model are often time-varying for scheduling problems.

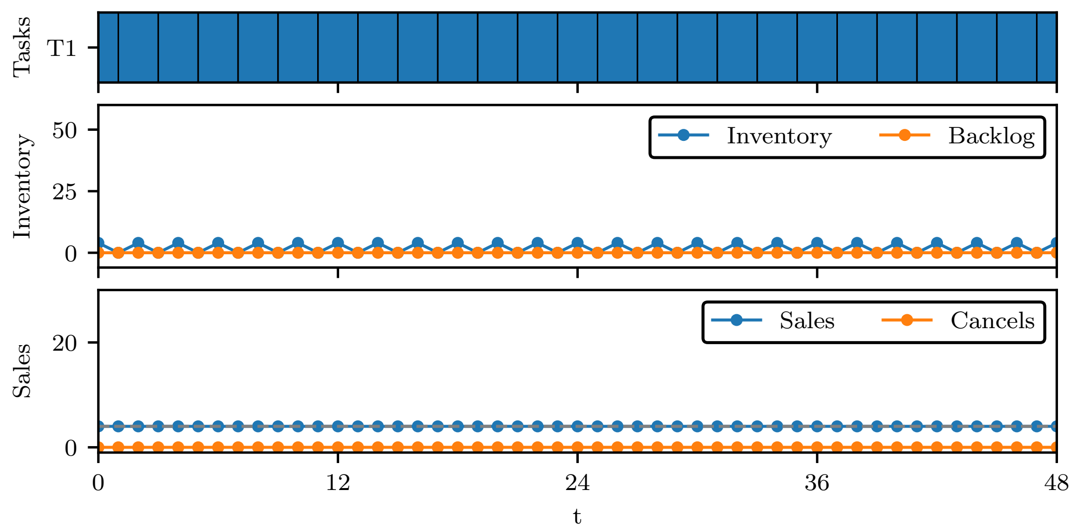

Figure 5.10: Periodic optimal solution and reference trajectory for the production scheduling example. The top plot is a Gantt chart with blue blocks representing executions of T1. The inventory, backlog, sales, and canceled backlog are shown in the middle and bottom plots.

dynamics, constraints, and stage cost are defined as

$$x^+ = f(x, u, w, t) := \bar{f}(x + \bar{x}_r(t), u + \bar{u}_r(t), w) - \bar{x}_r(t+1)$$

$$\mathbb{Z}(t) := \{(\bar{x} - \bar{x}_r(t), \bar{u} - \bar{u}_r(t)) : (\bar{x}, \bar{u}) \in \bar{\mathbb{Z}}(t)\}$$

$$\ell(x, u, t) := \bar{\ell}(x + \bar{x}_r(t), u + \bar{u}_r(t)) - \bar{\ell}(\bar{x}_r(t), \bar{u}_r(t))$$

With the terminal constraint, we require that all elements of the state, except backlog, terminate in phase with the reference trajectory. For the shifted state, this means that all of these elements, except the element of the state corresponding to backlog, must be zero. For backlog, the terminal region includes any nonnegative real number. Thus, we have

$$\mathbb{X}_f(t) := \{(0, 0, 0, 0, 0, 0, 0)\} \times \mathbb{R}_{\geq 0}$$

and we use the terminal cost

$$V_f(x,t) = 900U + 100U^2$$

By allowing backlog to take any nonnegative value, the sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$ is now robustly positive invariant for the disturbance of interest and Assumption 5.22 is satisfied. Assumption 5.4 is satisfied by the terminal control law $\kappa_f(x,t) := (0,0,0,\min\{U,1\})$. We can further establish that Assumption 5.23 holds for this problem by exploiting the integrator dynamics of the unbounded backlog state. We provide more details on this procedure in the subsequent chapter for a general class of scheduling problems.

If a delay occurs every time step, no M1 is ever produced and backlog continues to accumulate. Instead, a more realistic scenario includes task delays that occur infrequently with $\varepsilon := \Pr(Y = 1)$. We simulate 100 realizations of the closed-loop trajectory with an initial backlog of 40 units for multiple values of ε.

An example closed-loop trajectory is shown in Figure 5.11 for $\varepsilon = 0.3$. We observe that the backlog decreases from the initial value at 40 units despite a few task delays. Once the backlog nears zero, infrequent task delays force the backlog to increase but, on average, the facility can recover before another delay occurs. Thus, the sample average performance of the system (in terms of stage cost) is bounded, even though specific realizations of the disturbance trajectory may produce diverging backlog and cost as $t \to \infty$.

Note that this problem formulation does not satisfy Assumption 5.5 and the system is not necessarily dissipative with respect to this stage cost. Thus, the closed-loop scheduling formulation does not guarantee RASiE or RASiP. Instead, the main concern in closed-loop scheduling is the economic performance of the system that we define via

$$\Delta(t) := \frac{1}{t}\sum_{k=0}^{t-1} \ell(x(k), u(k), k)$$

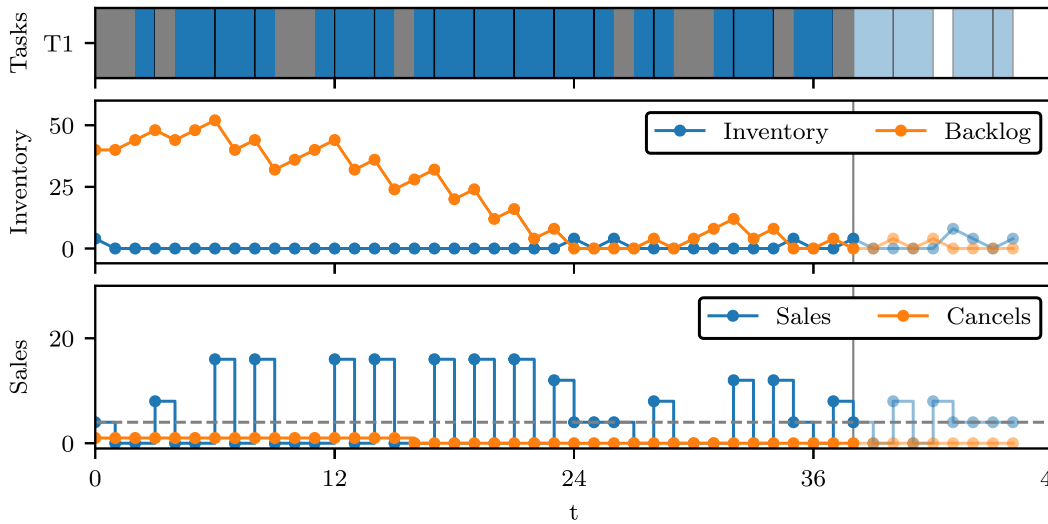

Figure 5.11: Closed-loop trajectory for production scheduling example. The top plot is a Gantt chart with blue blocks representing executions of T1. The inventory, backlog, sales, and outsourced backlog are shown in the middle and bottom plots. The open-loop trajectory, or schedule, for the next 8 time steps is shown in faded colors.

in which $x(k) := \phi(k; x, \mathbf{w}_k, 0)$ and $u(k) := \kappa(x(k), k)$ are the closed-loop state and input trajectories. Note that $\Delta(t)$ is equivalent to the closed-loop performance metric considered in Theorem 5.25. In Figure 5.12 we plot the sample average of $\Delta(t)$ for different values of ε. As $t \to \infty$, the sample average of $\Delta(t)$ decays towards some nonnegative constant specific to each value of ε for all $\varepsilon \leq 0.45$. As ε increases this constant increases for $\varepsilon \leq 0.45$. At $\varepsilon = 0.5$, the sample average instead diverges. Thus, the system exhibits behavior consistent with Theorem 5.25 and we presume that $0.45 < \delta < 0.5$.

5.5 Summary

Discrete-valued and large disturbances are a key feature of production planning and scheduling applications of MPC. While these disturbances are often too large to admit the deterministic robustness results discussed in the previous chapters, the infrequent nature of these large disturbances permits a stochastic description of robustness. Under suitable assump-

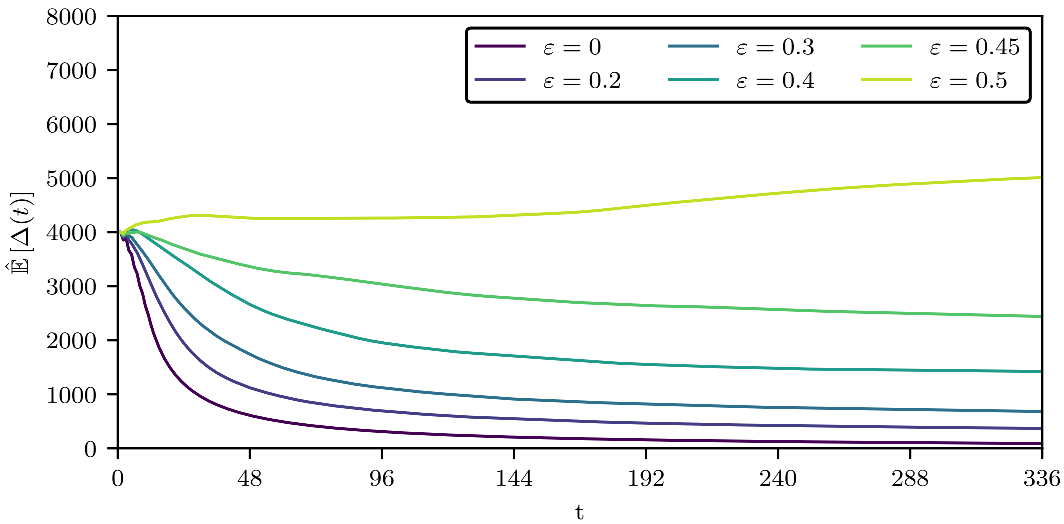

Figure 5.12: Sample average of $\Delta(t)$ for 100 simulations with different values of ε.

tions, nominal MPC is robust in this stochastic context, provided these large disturbances are sufficiently *infrequent*. This characterization can also include large disturbances such as faults and missing measurements, which are encountered in traditional process control applications. We again emphasize that this robustness is provided by feedback and without explicitly including the disturbance in the optimization problem. We do, however, need to design the MPC problem such that $\mathcal{X}$ is RPI for the disturbance of interest.

Given the stochastic nature of these disturbances, SMPC may seem like a desirable alternative to address these large and infrequent disturbances. We note, however, that SMPC requires a terminal region that is RPI for the disturbance of interest. For small disturbances, this design procedure is relatively straightforward. For large disturbances, constructing such a terminal region may prove difficult. To demonstrate this difficultly, we revisit the production scheduling example in Section 5.4.2. The terminal region used in this example is suitable for nominal MPC, but is not sufficient for SMPC. To accommodate the task delay in an SMPC formulation and retain the desired closed-loop properties, we must expand the terminal region to include any possible configuration of the facility and design a corresponding terminal cost that satisfies the required cost decrease condition for the entire state space. Typically, constructing such a terminal cost for the entire state space is intractable. Furthermore, any improvement gained from this SMPC formulation is likely to be small since these disturbances occur with small probability. Thus, we postulate that addressing these large and infrequent disturbances with SMPC, particularly for production scheduling problems, offers a minimal potential for improvement while incurring a significant increase in the effort required to formulate and solve these SMPC problems. We therefore do not formulate or discuss an SMPC formulation for large and infrequent disturbances.

While these results are general and can be applied to any system that satisfies the assumptions herein, the main application driving the development of these results, as suggested by the discussion thus far, is production scheduling. In the next chapter, we address this application in detail. Specifically, we develop a closed-loop scheduling algorithm that is inherently robust to large and infrequent disturbances.

Chapter 6

Closed-Loop Scheduling

Due to the different time scales, problem formulations, and computational requirements, process control and production scheduling problems are often viewed as two separate subfields within the larger discipline of process systems engineering. While this separation still persists for historical reasons, there is an emerging class of *online* scheduling algorithms that are beginning the blur the line between traditional process control and scheduling problems. These closed-loop or online scheduling algorithms address uncertainty in plant operations in real-time through consistent and frequent reoptimization and rescheduling, i.e., these algorithms respond to *feedback* from the system. The goal of these algorithms, whether explicitly stated or otherwise implied, is to provide the scheduling algorithm with *robustness* to relevant disturbances. This robustness has been explored via specific case studies, but there are no theoretical results available that characterize or establish this property of robustness for closed-loop scheduling algorithms and the corresponding closed-loop system.

In this chapter, we present and justify an appropriate definition of robustness for closed-loop scheduling. Since the main class of disturbances relevant to production scheduling applications are large and infrequent, we use the results in Chapter 5 to construct this definition of robustness. We already introduced a simple scheduling problem in Section 5.4.2 and designed an MPC formulation that renders the closed-loop system robust to large and infrequent disturbances for this specific example. We now seek to generalize this result. For this general class of production scheduling problems with an available reference trajectory, we construct a closed-loop scheduling algorithm that is inherently robust to large and infrequent disturbances. We then apply this algorithm to a more complicated production scheduling example to demonstrate the implications of these results. We also discuss extensions to this algorithm designed to reduce the computational burden and avoid "schedule nervousness" for large-scale industrial applications of closed-loop scheduling.

6.1 Introduction and literature review

Scheduling is a crucial activity for a variety of manufacturing facilities including chemical production, pharmaceuticals, food processing, metal fabrication, and logistic services (Maravelias, 2012; Harjunkoski et al., 2014). The task of generating reasonable schedules for a manufacturing facility is traditionally accomplished by means of human intuition, heuristics, and experience. Over the past few decades, however, optimization has found a significant role in generating high-quality schedules. Through mathematical abstractions, such as the state-task network (STN), a scheduling problem is formulated as mixed integer linear program (MILP). The objective function is based on some performance metric of the facility (e.g., cost) and the constraints are constructed to reflect the psychical constraints of the production facility (e.g., production times, capacities) (Kondili et al., 1993; Pantelides, 1994).

Since the advent of these mathematical models, expanding the scope of these models and improving the optimization algorithms for these specific problem formulations has received considerable attention. Generating a single schedule, however, is not sufficient in practice. Disturbances (e.g., demand variations, delays, and breakdowns) ensure that the optimal schedule at one time is inferior or infeasible a short time later. One method to accommodate these disturbances is to generate a single schedule that accounts for this uncertainty a priori. This approach can include robust optimization in which the schedule is designed to remain feasible for some set of possible realizations of the disturbance. The objective function can be based on the worst case realization of the disturbance (Lin et al., 2004) or the nominal objective function (Vin and Ierapetritou, 2001; Li and Ierapetritou, 2008c). Stochastic optimization formulations for scheduling problems were also proposed (Sand and Engell, 2004; Bonfill et al., 2004; Balasubramanian and Grossmann, 2004), in which the expected value of the cost function is optimized via a stochastic description of the uncertainty. These stochastic and robust optimization approaches are typically limited to demand uncertainty, sometimes called "left-hand side" uncertainty in the resulting MILP problem formulation. Two-stage adaptive robust optimization (ARO) can also be used to generate a set of possible schedules for the system (Shi and You, 2016). Depending on the realization of uncertainty in the first stage, the schedule is then updated. In particular, Lappas and Gounaris (2016); Lappas et al. (2019) use a multi-stage ARO that includes uncertainty in processing times in addition to demand variations. The multi-stage ARO formulations, in contrast to a standard feedback method, generates the set of possible recourse actions a priori. Thus, there are no available recourse actions for a disturbance that is not included in the initial ARO problem formulation.

For systems with high uncertainty and minimal recourse, accounting for disturbances a priori in the optimization problem is beneficial. Unfortunately, characterizing this uncertainty is difficult and often these schedules are overly conservative. In other words, these methods sacrifice nominal performance to accommodate disturbances that may or may not occur. Furthermore, a disturbance that is not included in the robust or stochastic optimization problem may render even the robust schedules infeasible. Further discussion on the merits of robust and stochastic optimization for production scheduling can be found in Li and Ierapetritou (2008a); Harjunkoski et al. (2014)

The alternative to accounting for uncertainty a priori is to react in real time to the realization of uncertainty, i.e., a *feedback* method. Initially called reactive rescheduling, these scheduling algorithms used heuristics and logical rules to handle disturbances and reschedule

(Cott and Macchietto, 1989; Kanakamedala et al., 1994; Huercio et al., 1995; Rodrigues et al., 1996; Elkamel and Mohindra, 1999). As the speed of MILP solvers improved, online reoptimization of the schedule in real time became feasible. Upon observing a disturbance, part of the schedule is fixed and updated with the disturbance, while the remaining part of schedule is reoptimized (Vin and Ierapetritou, 2000). Initially, these algorithms focused on minimizing schedule alterations and the size of the optimization problem by freezing variables that were deemed unrelated to the current disturbance and only rescheduling after a disturbance is observed (Mendez and Cerdá, 2004; Janak et al., 2006; Ferrer-Nadal et al., 2007; Novas and Henning, 2010; Chu and You, 2014). Two stage stochastic optimization can also be used with this reactive scheduling approach (Cui and Engell, 2010). Multi-parametric programming can be used to reduce online computation time, but these approaches suffer from the curse of dimensionallity as the size of the scheduling problem increases (Li and Ierapetritou, 2008b; Kopanos and Pistikopoulos, 2014). None of these algorithms, however, offer any theoretical guarantees for the performance or robustness of the closed-loop system.

Instead of rescheduling only when a disturbance occurs, a natural extension of these reactive scheduling methods is to reoptimize the schedule at fixed sampling intervals regardless of whether a disturbance occurs (Subramanian et al., 2012; Gupta and Maravelias, 2016; Gupta et al., 2016). These algorithms are sometimes called online or rolling horizon scheduling, but we use the title *closed-loop* scheduling in this work to emphasize the parallels between this approach to scheduling and closed-loop process control. As noted throughout this dissertation, feedback alone is not sufficient to guarantee robustness and may produce poor closed-loop performance, even in the nominal case, if the closed-loop scheduling algorithm is not carefully designed (Subramanian et al., 2012; Gupta and Maravelias, 2016; Risbeck et al., 2019). A graphical depiction of closed-loop scheduling is provided in Figure 6.1.

Given the presentation provided in this dissertation, the similarities between closed-loop scheduling and closed-loop control may appear both intuitive and obvious. Even a decade ago,

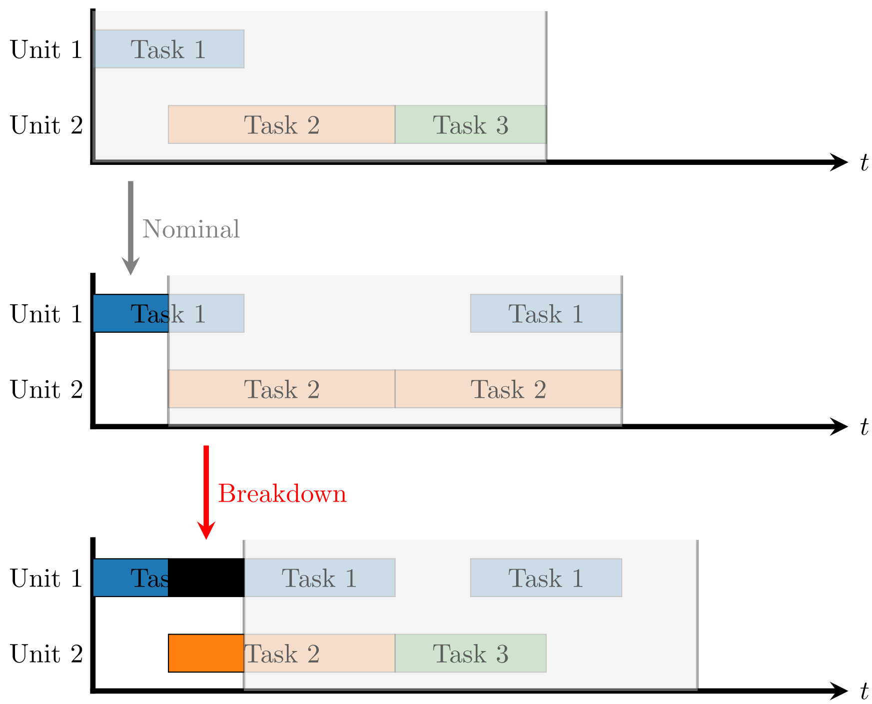

Figure 6.1: A diagram of closed-loop scheduling. The optimal open-loop schedule is shown in the gray region, while the implement (closed-loop) schedule is shown in the solid colors.

however, these parallels were neither intuitive nor obvious to the separate research communities of production scheduling and process control. The key limitation was that production scheduling problem formulations were seldom written as dynamical systems in the same manner as the dynamical systems encountered in process control problems. This approach made the resulting optimization problems more compact and easier to solve, but also made them incompatible with the typical analysis tools used in process control. The work of Subramanian et al. (2012) removed this limitation. The authors demonstrated that the typical STN formulation used for production scheduling problems can be converted to a dynamical state-space model. Thus, we can treat closed-loop scheduling as an MPC problem in which there are both discrete-valued inputs and economic cost functions. Moreover, we can use the theoretical results developed throughout this dissertation, specifically Chapter 5, to analyze the closed-loop behavior of closed-loop scheduling.

Using this framework and an available reference trajectory, Risbeck et al. (2019) propose a closed-loop scheduling algorithm that ensures nominal closed-loop performance via terminal equality constraints. This nominal guarantee is an important first step as it precludes particularly poor closed-loop performance, but nominal performance does not guarantee robustness to disturbances. Simulation frameworks that can be used to study the empirical robustness of specific case studies date back to Honkomp et al. (1999), with a more recent study available in Gupta and Maravelias (2020). While these studies allow the authors to draw useful insights for the design of closed-loop scheduling algorithms, these conclusions are inherently limited to specific case studies. In contrast to the quantitative and empirical robustness results provided by these simulation studies, we instead focus on defining and establishing the property of robustness for a general class of closed-loop scheduling problems. We then design a closed-loop scheduling algorithm that is guaranteed to be robust in this context, under a set of reasonable assumptions.

6.2 Problem formulation

6.2.1 State-space scheduling model

We consider a general production scheduling problem for batch processes with a discrete-time grid. To represent the manufacturing facility we use the STN representation. The facility consists of tasks $i \in \mathbf{I}$, units $j \in \mathbf{J}$, and materials $k \in \mathbf{K}$. The subset of tasks i than can be run on unit j is denoted $\mathbf{I}_j$. Materials for which we have demand or that can be sold for profit are called products and are denoted by the subset $\mathbf{K}^P \subseteq \mathbf{K}$. We denote the intermediate and feedstock materials as $\mathbf{K}^I = \mathbf{K} \setminus \mathbf{K}^P$. In general, the goal of these scheduling problems is to convert feedstock or raw materials to final products through a series of tasks and intermediate materials. We note that these intermediate materials can include "renewable" materials that account for thermal, electrical, or labor demands of a task. For example, the facility generates steam to provide heat to certain processes. This resource must be allocated across the facility such that we do not consume more than the maximum rate of steam generation for the facility. Steam capacity is consumed when a task starts and replenished when the task ends. Thus, we call this steam capacity a renewable material for the facility.

The parameter τ_{ij} denotes the processing time of task i on unit j. These tasks produce and consume materials when they start and finish. The parameters β_{ij}^{min} and β_{ij}^{max} denote the minimum and maximum batch size for task i on unit j. The parameters $\rho_{ik}/\bar{\rho}_{ik}$ denote the ratio of material k produced (> 0 for production, < 0 for consumed) by starting/completing, respectively, task i relative to the batch size of the task.

We now construct a state-space scheduling model similar to Subramanian et al. (2012); Gupta and Maravelias (2016). We define the binary decision variables W_{ij} to be unity if task i starts on unit j at the current time. We also define the continuous decision variable B_{ij} to denote the batch size assignment for the same task. To track these decisions in the state of the system, we "lift" W_{ij} and B_{ij} with the state variables $\bar{W}_{ij}^n$ and $\bar{B}_{ij}^n$ for $n \in \{0, \dots, \tau_{ij}\}$. The index n is the progress of the task, e.g., $\bar{W}_{ij}^n = 1$ indicates that task i on unit j is n/τ_{ij} complete at the current time. We include state variables for inventory of intermediates and feedstocks denoted $\tilde{S}_k$ for $k \in \mathbf{K}^I$. The state variables S_k and U_k are the inventory and backlog (unmet demand), respectively, of the products $k \in \mathbf{K}^P$. Inventory of each material must not exceed a maximum inventory capacity denoted $\tilde{\psi}_k$ for each $k \in \mathbf{K}^I$ and ψ_k for each $k \in \mathbf{K}^P$. We consider incoming deliveries $\zeta_k(t)$ for each $k \in \mathbf{K}^I$ and outgoing demand $\xi_k(t)$ for each $k \in \mathbf{K}^P$ that vary with the time index $t \in \mathbb{I}_{\geq 0}$. The decision variable H_k is the amount of material $k \in \mathbf{K}^P$ shipped to meet demand with a maximum throughput $\eta_k > 0$, i.e., $H_k \leq \eta_k$.

We also add an option to "hold" material in a processing unit after a task is complete. For each task $i \in \mathbf{I}_j$, we can also include a task $i' \in \mathbf{I}_j$ that holds the material in the processing unit. This task can be used only after task i is completed and consumes and produces the same set of materials produced by task i with a processing time of one time step. We denote the subset of these holding tasks as $\mathbf{I}_j^h \subseteq \mathbf{I}_j$ and the mapping of each hold task i' to the corresponding production task $i \in \mathbf{I}_j \setminus \mathbf{I}_j^h$ as $h(i') = i$. This option, originally proposed by (Kondili et al., 1993), is often *necessary* to ensure that scheduling algorithms satisfy maximum inventory constraints for intermediate materials when subject to delays and demand variations. For example, if we experience a delay in downstream production, intermediate material may not be consumed at the time specified by the previous schedule. If a task producing this intermediate material is completed at the same time, we may not have sufficient inventory capacity. By allowing the scheduler the option to leave this material in the unit, we can avoid violating the maximum inventory constraints for the facility and therefore recover a feasible schedule. (Avadiappan and Maravelias, 2021) make a similar observation and propose adding delays as optimization variables to ensure feasibility.

The decision variable V_k is the amount of material purchased/sold in excess of demand (> 0 for purchased, < 0 for sold). The parameters ν_k^P and ν_k^S denote the maximum amount of material that can be purchased or sold at each time step for each material $k \in \mathbf{K}$ in excess of demand or shipments. We often have $\nu_k^P = \nu_k^S = 0$ such that extra material cannot be purchased or sold. Thus, we cover cost minimization problems (meet demand at minimum cost), profit maximization problems (sell as much product as possible), and any combination of these two problem types.

We also allow for a material disposal action. In practice, this action includes a wider range of options than the usual interpretation of disposal, i.e., waste. Instead, we can ship the material to a long-term storage facility or move the material to a separate storage unit on-site. The key characteristic of this action is that we pay a single (large) cost that removes the material from the inventory state variable and therefore the cost function. The decision variable D_k is the amount of inventory disposed at a given time step for material $k \in \mathbf{K}^P$.

We use the parameter μ_k to denote the maximum amount of inventory that may be disposed for each material $k \in \mathbf{K}^P$ at each time step.

We now consider the disturbances in the scheduling model. Let the binary variable Y_j be unity if unit j is delayed by one time step. Let the binary variable Z_j be unity if unit j experiences a breakdown or total loss during $[t, t+1)$. We define fractional yield loss of batch size/material with the variable L_j which takes values in $[0, 1]$, e.g., $L_j = 0.25$ is a 25% loss of batch size on unit j. Note that individual units are not permitted to run more than one task and we can therefore specify these disturbances by only the unit affected.

We now define the dynamic evolution of the state variables. All the variables on the right-hand side of these equations are at time t and the left-hand side at time $t + 1$, denoted by $^+$. For all $i \in \mathbf{I}_j$ and $j \in \mathbf{J}$ with $\tau_{ij} \geq 2$, we have

$$
\begin{aligned}
(\bar{W}_{ij}^0)^+ &= (\bar{W}_{ij}^0 + W_{ij})Y_j(1 - Z_j) \\
(\bar{W}_{ij}^1)^+ &= \left((\bar{W}_{ij}^0 + W_{ij})(1 - Y_j) + \bar{W}_{ij}^1 Y_j\right)(1 - Z_j) \\
(\bar{W}_{ij}^{\tau_{ij}})^+ &= \bar{W}_{ij}^{\tau_{ij}-1}(1 - Y_j)(1 - Z_j) \\
(\bar{B}_{ij}^0)^+ &= (\bar{B}_{ij}^0 + B_{ij})Y_j(1 - Z_j)(1 - L_j) \\
(\bar{B}_{ij}^1)^+ &= \left((\bar{B}_{ij}^0 + B_{ij})(1 - Y_j) + \bar{B}_{ij}^1 Y_j\right)(1 - Z_j)(1 - L_j) \\
(\bar{B}_{ij}^{\tau_{ij}})^+ &= \bar{B}_{ij}^{\tau_{ij}-1}(1 - Y_j)(1 - Z_j)(1 - L_j)
\end{aligned}
$$

and

$$
\begin{aligned}
(\bar{W}_{ij}^n)^+ &= (W_{ij}^{n-1}(1 - Y_j) + W_{ij}^n Y_j)(1 - Z_j) \\
(\bar{B}_{ij}^n)^+ &= (B_{ij}^{n-1}(1 - Y_j) + B_{ij}^n Y_j)(1 - Z_j)(1 - L_j)
\end{aligned}
$$

for all $n \in \{1, \ldots, \tau_{ij} - 1\}$. If $\tau_{ij} = 1$, we have instead

$$\begin{aligned}
(\bar{W}_{ij}^0)^+ &= (\bar{W}_{ij}^0 + W_{ij})Y_j(1 - Z_j) \\
(\bar{W}_{ij}^1)^+ &= (\bar{W}_{ij}^0 + W_{ij})(1 - Y_j)(1 - Z_j) \\
(\bar{B}_{ij}^0)^+ &= (\bar{B}_{ij}^0 + B_{ij})Y_j(1 - Z_j)(1 - L_j) \\
(\bar{B}_{ij}^1)^+ &= (\bar{B}_{ij}^0 + B_{ij})(1 - Y_j)(1 - Z_j)(1 - L_j)
\end{aligned}$$

For intermediate and feedstock materials ($k \in \mathbf{K}^I$), we model the inventory dynamics as

$$\tilde{S}_k^+ = \tilde{S}_k + \sum_{j \in \mathbf{J}} \sum_{i \in \mathbf{I}_j} \left(\bar{\rho}_{ik} \bar{B}_{ij}^{\tau_{ij}} + \rho_{ik} B_{ij}\right) + V_k + \zeta_k$$

For products ($k \in \mathbf{K}^P$), we model the discrete-time evolution of inventory and backlog as

$$\begin{aligned}
S_k^+ &= S_k + \sum_{j \in \mathbf{J}} \sum_{i \in \mathbf{I}_j} \left(\bar{\rho}_{ik} \bar{B}_{ij}^{\tau_{ij}} + \rho_{ik} B_{ij}\right) + V_k - H_k - D_k \\
U_k^+ &= U_k - H_k + \xi_k
\end{aligned}$$

To streamline notation, we now define each variable or parameter without subscripts to indicate a column vector containing the variable at each subscript, e.g.,

$$\bar{B} := [(\bar{B}_{ij}^n \ \forall j \in \mathbf{J}, i \in \mathbf{I}_j, n \in \mathbb{I}_{0:\tau_{ij}})]' \qquad V := [(V_k \ \forall k \in \mathbf{K})]' \qquad \pi := [(\pi_k \ \forall k \in \mathbf{K})]'$$

The state, input, and disturbance for the system are then defined as

$$x := \begin{bmatrix} \bar{W} \\ \bar{B} \\ \tilde{S} \\ S \\ U \end{bmatrix} \qquad u := \begin{bmatrix} W \\ B \\ V \\ H \\ D \end{bmatrix} \qquad w := \begin{bmatrix} Y \\ Z \\ L \end{bmatrix} \tag{6.1}$$

The dynamic equations introduced in the previous paragraph can be represented in state-space form as follows.

$$x^+ = f(x, u, w, t) \tag{6.2}$$

with $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, and $w \in \mathbb{R}^q$. Note that the system is time-varying because shipments $\zeta_k(t)$ and demand $\xi_k(t)$ are time-varying.

While the variables Y_j, Z_j, L_j enter the model as bi-linear terms, these variables are not considered in the nominal optimization problem ($w = 0$). Thus, we have a linear affine description of the nominal system,

$$f(x, u, 0, t) = Ax + Bu + c(t) \tag{6.3}$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $c(t) \in \mathbb{R}^n$.

Remark 6.1. In certain cases, we may have accurate predictions of these disturbance variables at future times (e.g., we know that there is maintenance on unit j at time t'). In these cases, we could include these disturbances as parameters in the nominal optimization problem to produce a superior schedule. From a process control perspective, we may call this *feedforward* control, i.e., we use information about upcoming disturbances to better determine the control action. To streamline the following presentation and discussion, we exclude these disturbance predictions from the closed-loop scheduling algorithm. Although this additional information may improve the performance of the algorithm, we find that this information is not necessary to demonstrate the robustness of closed-loop scheduling. Observing and reacting to disturbances after they occur is sufficient to guarantee some margin of robustness for closed-loop scheduling.

In addition to the state-space dynamics in (6.2), we impose the following constrains on the state and input at each time step to enforce one-task-per-unit and batch size requirements.

$$\sum_{i \in \mathbf{I}_j} \sum_{n=0}^{\tau_{ij}} \bar{W}_{ij}^n \leq 1 \quad \forall j \in \mathbf{J} \tag{6.4}$$

$$\beta_{ij}^{min} W_{ij} \leq B_{ij} \leq \beta_{ij}^{max} W_{ij} \quad \forall i \in \mathbf{I}_j,\ j \in \mathbf{J} \tag{6.5}$$

For the hold tasks, we require that these tasks are only available after the corresponding production task is complete, i.e.,

$$W_{ij} \leq \bar{W}_{h(i)j}^{\tau_{h(i)j}} + \bar{W}_{ij}^1 \quad \forall i \in \mathbf{I}_j^h \tag{6.6}$$

We also enforce the appropriate ranges for each element of the state and input.

$$\begin{gathered} W_{ij}, \bar{W}_{ij}^n, X_j \in \{0,1\} \quad B_{ij}, \bar{B}_{ij}^n \in [0, \beta_{ij}^{max}] \quad \forall i \in \mathbf{I}_j,\ j \in \mathbf{J} \\ \tilde{S}_k \in [0, \tilde{\psi}_k] \quad \forall k \in \mathbf{K}^I \\ S_k \in [0, \psi_k] \quad U_k \geq 0 \quad \forall k \in \mathbf{K}^P \\ V_k \in [-\nu_k^S, \nu_k^P] \quad H_k \in [0, \eta_k] \quad D_k \in [0, \mu_k] \quad \forall k \in \mathbf{K} \end{gathered} \tag{6.7}$$

We rewrite the constraints in (6.4), (6.5) and (6.7) as $x \in \mathbb{X}$ and $u \in \mathbb{U}$ for the sets $\mathbb{X} \subseteq \mathbb{R}^n$ and $\mathbb{U} \subseteq \mathbb{R}^m$. We add (6.6) to these constraints by defining the set $\mathbb{Z} \subseteq \mathbb{X} \times \mathbb{U}$ and requiring that $(x, u) \in \mathbb{Z}$. We also have that $w \in \mathbb{W}$ with

$$\mathbb{W} := \{w : Y_j, Z_j \in \{0,1\},\ L_j \in [0,1] \quad \forall j \in \mathbf{J}\}$$

We now introduce the cost function for this system. Let α_{ij}^F and α_{ij}^P denote the fixed and proportional costs of task i on unit j. Let π_k denote the sales prices of material $k \in \mathbf{K}$. Let π_k^S and π_k^U denote the inventory and backlog cost, respectively, per unit of product $k \in \mathbf{K}^P$. Inventory costs for $k \in \mathbf{K}^I$ are zero. We use the parameter π_k^D to denote the cost of disposing material. Note that all of these costs are proportional to elements of state and input and we can therefore write the economic cost as

$$\bar{\ell}(x, u, t) = q'x + r'u \tag{6.8}$$

in which

$$q := \begin{bmatrix} 0 \\ 0 \\ 0 \\ \pi^S \\ \pi^U \end{bmatrix} \qquad r := \begin{bmatrix} \alpha^F \\ \alpha^P \\ \pi \\ 0 \\ \pi^D \end{bmatrix}$$

Thus, the scheduling problem can now be written as a time-varying MPC problem.

6.2.2 MPC problem and assumptions

To properly formulate the MPC problem and to benchmark the closed-loop performance, we first require a reference trajectory for the nominal system. This reference trajectory is comprised of a sequence of states $\mathbf{x}_r$ and inputs $\mathbf{u}_r$ that form a feasible trajectory for the nominal system, i.e., the reference trajectory satisfies the following assumption.

Assumption 6.2 (Reference trajectory)**.** The reference trajectory $(\mathbf{x}_r, \mathbf{u}_r)$ satisfies

$$x_r(t+1) = f(x_r(t), u_r(t), t)$$

with $(x_r(t), u_r(t)) \in \mathbb{Z}$ for all $t \in \mathbb{I}_{\geq 0}$.

For systems with periodic demand profiles, an optimal periodic schedule (computed a priori for the system) can serve as the reference trajectory. Heuristic methods may also be used to construct this reference trajectory. We emphasize, however, that the economic performance of this trajectory is important as this reference trajectory forms the benchmark for all subsequent results. Improvements in the reference trajectory are therefore reflected in the following performance guarantees.

In contrast to Chapter 5, we do not shift the system such that the reference trajectory is at the origin. While shifting the system is useful for analysis, the original system is often easier to interpret and implement in the optimization problem. Since this chapter focuses on algorithm development as well as theoretical results for closed-loop scheduling, we therefore leave the state, input, and system in the original variables and present all definitions, algorithms, assumptions, and results in these original variables. We do, however, shift the stage cost as follows.

$$\ell(x, u, t) := \bar{\ell}(x, u, t) - \bar{\ell}(x_r(t), u_r(t), t) = q'(x - x_r(t)) + r'(u - u_r(t)) \tag{6.9}$$

such that $\ell(x_r(t), u_r(t), t) = 0$. Note that the stage cost is time-varying even if $\bar{\ell}()$ is not time-varying because the reference trajectory is time-varying.

We now briefly reintroduce time-varying MPC. The nominal system is described by

$$x^+ = f(x, u, 0, t) \tag{6.10}$$

For a horizon $N \in \mathbb{I}_{\geq 1}$, let $\hat{\phi}(k; x, \mathbf{u}, t)$ denote the open-loop state solution to the nominal system in (6.10) at time $k \in \mathbb{I}_{t:N}$, given the initial state $x \in \mathbb{X}$ at time $t \in \mathbb{I}_{\geq 0}$ and the input trajectory $\mathbf{u} \in \mathbb{U}^N$. We also consider a sequence of terminal constraints $(\mathbb{X}_f(t))_{t=0}^{\infty}$ and terminal cost $V_f : \mathbb{R}^n \times \mathbb{I}_{\geq 0} \to \mathbb{R}_{\geq 0}$, to be defined later. We define the set of admissible inputs, admissible initial conditions, and objective function as follows.

$$\begin{aligned}
\mathcal{U}(x,t) &:= \{\mathbf{u} \in \mathbb{U}^N : (\hat{\phi}(k; x, \mathbf{u}, t), u(k)) \in \mathbb{Z}(k)\ \forall k \in \mathbb{I}_{t:t+N-1} \\
&\qquad \text{and } \hat{\phi}(N; x, \mathbf{u}, t) \in \mathbb{X}_f(N)\} \\
\mathcal{X}(t) &:= \{x \in \mathbb{X} : \mathcal{U}(x,t) \neq \emptyset\} \\
V(x, \mathbf{u}, t) &:= \sum_{k=i}^{t+N-1} \ell(\hat{\phi}(k; x, \mathbf{u}, t), u(k), k) + V_f(\hat{\phi}(t+N; x, \mathbf{u}, t), t+N)
\end{aligned}$$

The MPC problem is then

$$\mathbb{P}(x,t):\ V^0(x,t) := \min_{\mathbf{u} \in \mathcal{U}(x,t)} V(x, \mathbf{u}, t)$$

and the optimal solution(s) are denoted $\mathbf{u}^0(x,t)$. We assume a Borel measurable selection rule is applied to define the single-valued control law $\kappa(\cdot, t) : \mathcal{X}(t) \to \mathbb{U}$ such that $\kappa(x,t) \in \{u(0) : \mathbf{u} \in \mathbf{u}^0(x,t)\}$ for all $x \in \mathcal{X}(t)$ and $t \in \mathbb{I}_{\geq 0}$. The closed-loop system is therefore

$$x^+ = f(x, \kappa(x,t), w, t) \tag{6.11}$$

and we let $\phi(k; x, \mathbf{w}_{t:k}, t)$ denote the solution to (6.11) at time $k \in \mathbb{I}_{\geq t}$, given the initial state

$x \in \mathcal{X}(t)$ at time $t \in \mathbb{I}_{\geq 0}$ and the disturbance sequence $\mathbf{w}_{t:k} \in \mathbb{W}^{k-t}$. In the context for production scheduling problems, $\mathbf{u}^0(x,t)$ defines the open-loop schedule for a given state x and time t while $\phi(\cdot, x, \mathbf{w}_k, t)$ and the corresponding input trajectory are the closed-loop or implemented schedule for the facility.

Remark 6.3. Since this shifted stage cost $\ell(\cdot)$ produces a constant shift in the cost function relative to the original economic cost $\bar{\ell}(\cdot)$, either cost function produces the same optimal solution $\mathbf{u}^0(\cdot)$ and therefore control law. Thus, we can also formulate the MPC problem in terms of $\bar{\ell}(\cdot)$ without altering the following results.

We require the following standard assumptions for the system, stage cost, and constraints.

Assumption 6.4 (System, stage cost, and constraints)**.** The model $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{I}_{\geq 0} \to \mathbb{R}^n$ and stage cost $\ell : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{I}_{\geq 0} \to \mathbb{R}$ are continuous. The function $\ell(\cdot)$ is bounded from below. The set $\mathbb{Z}$ is closed and $\mathbb{U}$ is compact.

We then require the following assumption for the terminal cost and constraint in the MPC problem.

Assumption 6.5 (Terminal conditions)**.** The set $\mathbb{X}_f(t)$ is closed for all $t \in \mathbb{I}_{\geq 0}$. The terminal cost $V_f(\cdot)$ is continuous and lower bounded with $V_f(x_r(t), t) = 0$. There exists a terminal control law $\kappa_f : \mathbb{X}_f \times \mathbb{I}_{\geq 0} \to \mathbb{U}$ such that

$$f(x, \kappa_f(x,t), 0, t) \in \mathbb{X}_f(t+1) \tag{6.12}$$

$$V_f(f(x, \kappa_f(x,t), 0, t), t+1) \leq V_f(x,t) - \ell(x, \kappa_f(x), t) \tag{6.13}$$

for all $x \in \mathbb{X}_f(t)$ and $t \in \mathbb{I}_{\geq 0}$. Furthermore, $(x, \kappa_f(x,t)) \in \mathbb{Z}(t)$ for all $x \in \mathbb{X}_f(t)$ and $t \in \mathbb{I}_{\geq 0}$.

With these assumptions, we can establish Theorem 5.10, which we restate as follows.

Theorem 6.6 (Nominal performance)**.** *Let Assumptions 6.2, 6.4 and 6.5 hold. Then the sequence of sets* $(\mathcal{X}(t))_{t=0}^{\infty}$ *is positive invariant for the nominal closed-loop system* $x^+ = f(x, \kappa(x,t), 0, t)$*. Furthermore, the nominal closed-loop trajectory satisfies*

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{k=t}^{t+T-1} \ell(x(k), \kappa(x(k), k), k) \leq 0$$

in which $x(k) = \phi(k; x, \mathbf{0}, t)$ *for all* $x \in \mathcal{X}$ *and* $t \in \mathbb{I}_{\geq 0}$.

6.3 Disturbances and inherent robustness

The class of disturbances most relevant to production scheduling applications are typically large and infrequent, such as breakdowns and delays. We therefore restrict our attention to systems with *only* large and infrequent disturbances, i.e., $w \in \mathbb{W}_1$, via the following assumption.

Assumption 6.7 (Only large disturbances)**.** The disturbance set satisfies $\mathbb{W} = \mathbb{W}_0 \cup \mathbb{W}_1$ with $\mathbb{W}_0 = \{0\}$.

We also consider the usual assumption of independence for these disturbances.

Assumption 6.8 (Disturbances)**.** The disturbances $w \in \mathbb{W}$ are random variables that are i.i.d. in time. The set $\mathbb{W}$ is compact and contains the origin.

We use $\mu : \mathcal{B}(\mathbb{W}) \to [0,1]$ to denote the probability measure for w. For $\delta \in [0,1]$, we define $\mathcal{M}(\mathbb{W}, \delta)$ as the set of all probability measures on the measurable space $(\mathbb{W}, \mathcal{B}(\mathbb{W}))$ such that $\mu(\mathbb{W}_1) \in [0, \delta]$ for all $\mu \in \mathcal{M}(\mathbb{W}, \delta)$.

For production scheduling applications, economic performance of the closed-loop system is more important than stability of a specific reference trajectory. Thus, Theorem 5.25

provides the most relevant notion of robustness for closed-loop scheduling. Based on this theorem, we define economic robustness to large, infrequent disturbances as follows.

Definition 6.9 (Economically robust to large, infrequent disturbances)**.** The system $x^+ = f(x, \kappa(x,t), w, t)$, $w \in \mathbb{W}$ is economically robust to large, infrequent disturbances with respect to the stage cost $\ell(\cdot)$ and reference trajectory $(\mathbf{x}_r, \mathbf{u}_r)$ in an RPI sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$ if there exist $\delta > 0$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{k=t}^{t+T-1} \mathbb{E}\left[\ell(x(k), \kappa(x(k), k), k)\right] \leq \gamma(\mu(\mathbb{W}_1)) \tag{6.14}$$

in which $x(k) = \phi(k; x, \mathbf{w}_{t:k}, t)$ for all $x \in \mathcal{X}(t)$, $\mu \in \mathcal{M}(\mathbb{W}, \delta)$, and $t \in \mathbb{I}_{\geq 0}$.

As discussed in Chapter 5, this results ensure that for sufficient infrequent disturbances ($\mu(\mathbb{W}_1) \leq \delta$ for some $\delta > 0$), (6.14) holds for the closed-loop system. The key feature of this bound is that $\gamma(\cdot)$ is a $\mathcal{K}$-function. Recall that the reference trajectory is constructed for the nominal system and the stage cost $\ell(\cdot)$ is defined relative to the economic cost of this reference trajectory. Thus, disturbances that occur with arbitrarily small probability produce, at most, similar small deviations in the economic performance of the system relative to reference trajectory. Moreover, the guarantee in (6.14) ensures that as the reliability of the plant increases, i.e., $\varepsilon \to 0$, we approach the nominal performance of the system. We again note that $\gamma(\cdot)$ is often too conservative to be a useful quantitative bound, but the fact that this bound exists is significant.

To establish this property of robustness for the closed-loop system, we require the same assumptions used in Chapter 5.

Assumption 6.10 (Robust recursive feasibility)**.** The sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$ is robustly positive invariant for the closed-loop system $x^+ = f(x, \kappa(x,t), w, t)$, $w \in \mathbb{W}$.

Assumption 6.11 (Maximum cost increase)**.** There exist $b_1, b_2 \geq 0$ such that

$$V^0(f(x, \kappa(x,t), w, t), t+1) \leq V^0(x,t) + b_1\ell(x, \kappa(x,t), t) + b_2 \tag{6.15}$$

for all $x \in \mathcal{X}(t)$, $w \in \mathbb{W}_1$, and $t \in \mathbb{I}_{\geq 0}$.

We also restate Theorem 5.25, modified to use the definition of robustness in Definition 6.9.

Theorem 6.12. *Let Assumptions 6.2, 6.4, 6.5, 6.7, 6.8, 6.10 and 6.11 hold. Then the system* $x^+ = f(x, \kappa(x,t), w, t)$*,* $w \in \mathbb{W}$ *is economically robust to large, infrequent disturbances with respect to the stage cost* $\ell(\cdot)$ *and reference trajectory* $(\mathbf{x}_r, \mathbf{u}_r)$ *in the RPI sequence of sets* $(\mathcal{X}(t))_{t=0}^{\infty}$.

6.3.1 Motivating example

As the practical meaning and utility of this definition of robustness may not be clear yet, we now consider a motivating example to illustrate three important points:

1. Without careful construction of the terminal cost and constraints, closed-loop scheduling algorithms do not have this property of robustness.

2. Suitable nominal performance of a closed-loop scheduling algorithm does not imply that the closed-loop system is robust to disturbances.

3. Without this property of robustness, the closed-loop scheduling algorithm may produce myopic and undesirable behavior when subject to disturbances.

We consider a simple scheduling problem, adapted from McAllister et al. (2022), with a single unit and two tasks, T1 and T2, that produce the product M1. Raw materials are assumed to be abundant are therefore ignored in the scheduling problem. Each task requires 2 time steps to complete ($\tau_{ij} = 2$) and there is demand of 1 unit of M1 every 2 time steps ($\xi(t) = 1$ if t is even). T1 produces up to 1 unit of M1 ($\beta_{11}^{max} = 1$) at a cost of 60 ($\alpha_{11}^{F} = 60$) while

T2 produces up to 1.2 units of M1 ($\beta_{21}^{max} = 1.2$) at a cost of 90 ($\alpha_{21}^{F} = 90$). The cost for maintaining inventory and backlog is 1 and 10, respectively ($\pi^S = 1$ and $\pi^U = 10$). We also allow up to 1 unit of disposal of M1 ($\mu_1 = 1$) each time step at a cost of 10 ($\pi^D = 10$). This disposal action, however, is unused in all of the subsequent closed-loop simulations. For the nominal system, the optimal periodic schedule is to run T1 at maximum capacity and in phase with demand. We treat this optimal periodic schedule as the reference trajectory for the subsequent discussion and definition of the stage cost.

We first consider a closed-loop scheduling algorithm without a terminal constraint and cost, i.e., $\mathbb{X}_f = \mathbb{R}^n$ and $V_f(x, t) = 0$. This approach is representative of most online or closed-loop scheduling algorithms proposed in literature, in which a terminal cost and constraint are not included. Nonetheless, the nominal performance of these algorithms in empirical studies is often satisfactory. For example, if we initialize this motivating example in phase with the optimal periodic schedule with a horizon of $N = 24$, the closed-loop trajectory is identical to the optimal periodic schedule. This trajectory is shown in Figure 6.2. Thus, the nominal performance of this scheduling algorithm without a terminal constraint or cost appears to be acceptable. Nominal performance however does not imply robustness.

We now consider a delay in the completion of T1 at time $t = 2$, i.e., a disturbance. We use a horizon of $N = 24$ and we plot the closed-loop trajectory in Figure 6.3. After this delay, the algorithm continues to run T1 every two time steps while seemingly indifferent to the backlog penalty incurred at every other time step. The consequence of this choice are severe. The total cost continues to increase relative to the optimal periodic schedule for as long as we run the simulation. We emphasize that this behavior is a choice of scheduling *algorithm* and not caused by an inherent limitation of the underlying system. The system has a means to address backlog: run T2 instead of T1, produce extra M1 to offset this backlog, and return to the optimal periodic schedule. The algorithm, however, chooses to not run T2. Thus, we subjected the system to a single disturbance and the resulting closed-loop performance

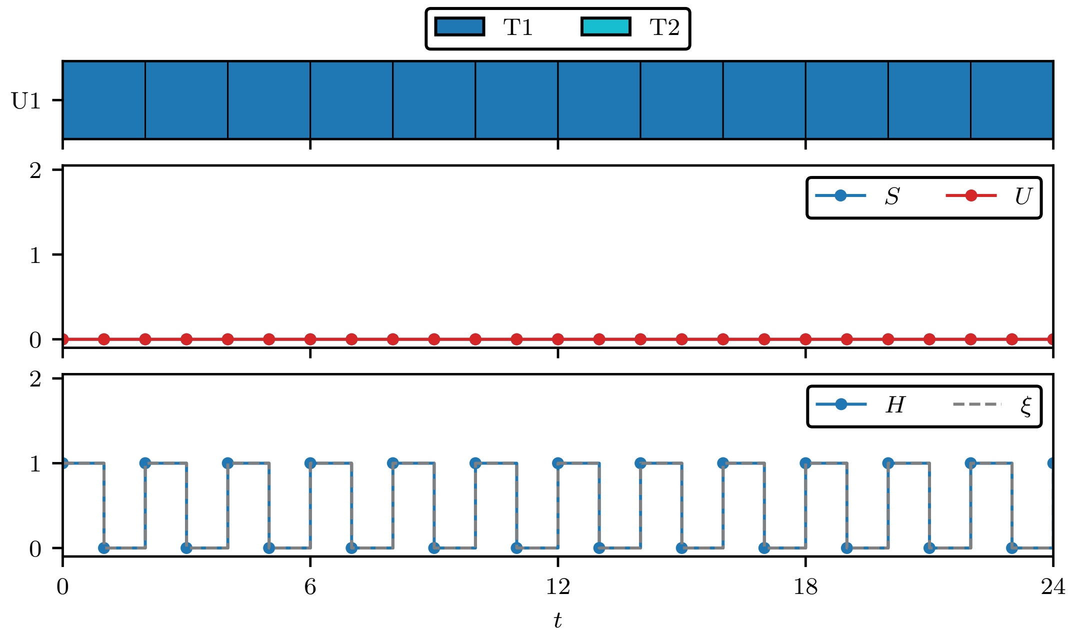

Figure 6.2: Closed-loop schedule for the motivating example without terminal ingredients.

degrades significantly and permanently. We note that this behavior occurs after a single delay regardless of when this delay occurs.

In terms of Definition 6.9, the algorithm and resulting closed-loop system in this example are not robust. To justify this conclusion, we first assume that this delay occurs with arbitrarily small probability that we denote $\varepsilon := \Pr(Y = 1) > 0$. Nonetheless, as $T \to \infty$, this disturbance must occur at least once and the behavior in Figure 6.3 eventually occurs. Therefore, the following inequality holds for all $\varepsilon > 0$.

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}[\ell(x(k), \kappa(x(k), k), k)] \geq 5 \tag{6.16}$$

Thus, the motivating example without a proper terminal cost and constraint violates Definition 6.9 and is therefore not robust. Conversely, a system that has the property of robustness as defined in Definition 6.9 cannot produce the behavior shown in Figure 6.3.

We now further investigate the reasons for this myopic behavior. The scheduling algo-

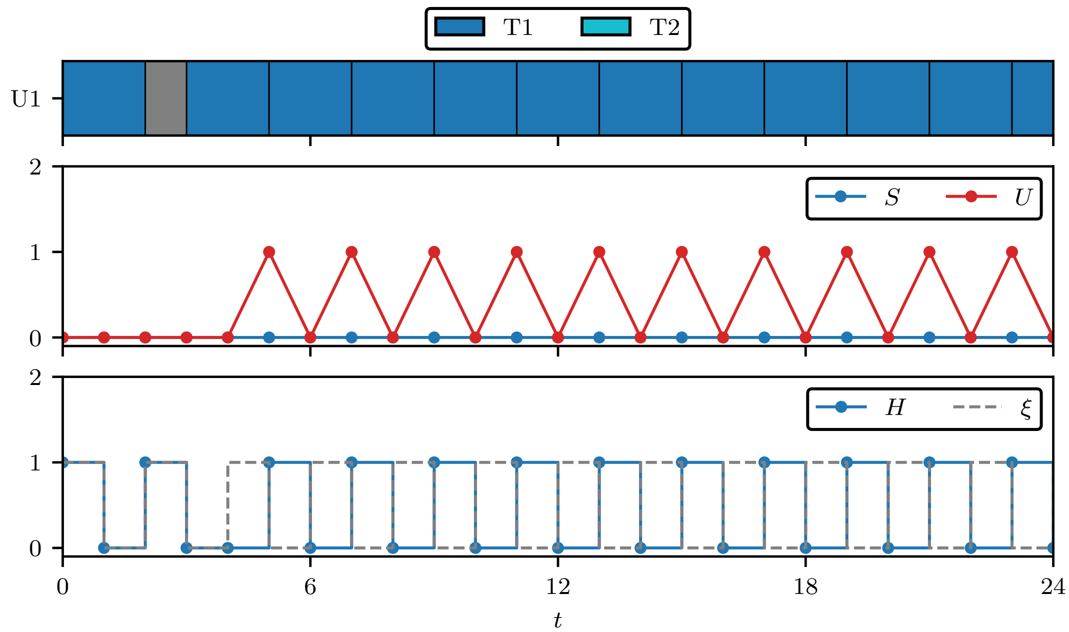

Figure 6.3: Closed-loop schedule for the motivating example without terminal ingredients and with a delay at $t = 2$.

rithm neglects T2 because the cost to address backlog by running T2 is in fact larger than the cost to retain backlog for 24 time steps of operation. Hence, the optimal choice for this horizon length is to never deal with the backlog. Without an appropriate terminal cost/constraint, the optimization problem is unaware of the consequences of any choices that are not realized within the chosen horizon length. In this example, the algorithm is unaware of the persistent cost of backlog and therefore selects a shortsighted approach, i.e., suffering the cost of backlog to avoid the added cost of running T2.

One approach to address this problem is to use a longer horizon length to attempt to better account for these persistent costs and therefore avoid myopic decisions. While this approach is sometimes used in industrial implementation of MPC, the horizon length necessary to ensure nominal performance and robustness of the closed-loop system is typically unknown. Moreover, the horizon length require to avoid this behavior may be so large that the optimization problem is intractable. In this motivating example, a horizon length of 24 time steps

(12 times the processing time and period of the demand cycle) was insufficient. If we scale this observation to large facilities with more complex pathways from raw materials to products, the required horizon length to avoid myopic behavior may be prohibitively large to address with current optimization solvers. Without a terminal cost and constraint, we can test the system's robustness through simulation studies, but we can not *guarantee* that the algorithm is robust without enumerating and testing ever possible state and disturbance that the manufacturing facility may encounter. Indeed, there are even systems for which increasing the horizon length does not ensure nominal performance or stability (Risbeck et al., 2019).

We instead propose a more scalable and generalizable solution to construct a robust closed-loop scheduling algorithm via appropriate an terminal constraint and cost. By designing a terminal constraint and cost that satisfy Assumptions 6.5, 6.10 and 6.11, we can guarantee that the closed-loop scheduling algorithm is robust in terms of Definition 6.9 without the computational demands of long horizon lengths and exhaustive testing of the algorithm.

Using the procedure discussed in the subsequent section, we design terminal ingredients that satisfy Assumptions 6.5, 6.10 and 6.11 for this motivating example. We then decrease the horizon length to only *8 time steps* and plot the resulting closed-loop trajectory in Figure 6.4. With this terminal cost and constraint, the closed-loop scheduling algorithm now recognizes that the cost of maintaining backlog outweighs the cost of running T2 in the long term and therefore chooses to run T2 after the disturbance occurs. After paying down the backlog, the system returns to the optimal periodic schedule. Thus, the system recovers after a single disturbances with only a temporary increase in the economic operating cost. After a sufficient amount of time, the time-average cost of the closed-loop system with a single delay is no different than the time-average cost of the nominal system, i.e.,

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} \mathbb{E}[\ell(x(k),\kappa(x(k),k),k)] \to 0$$

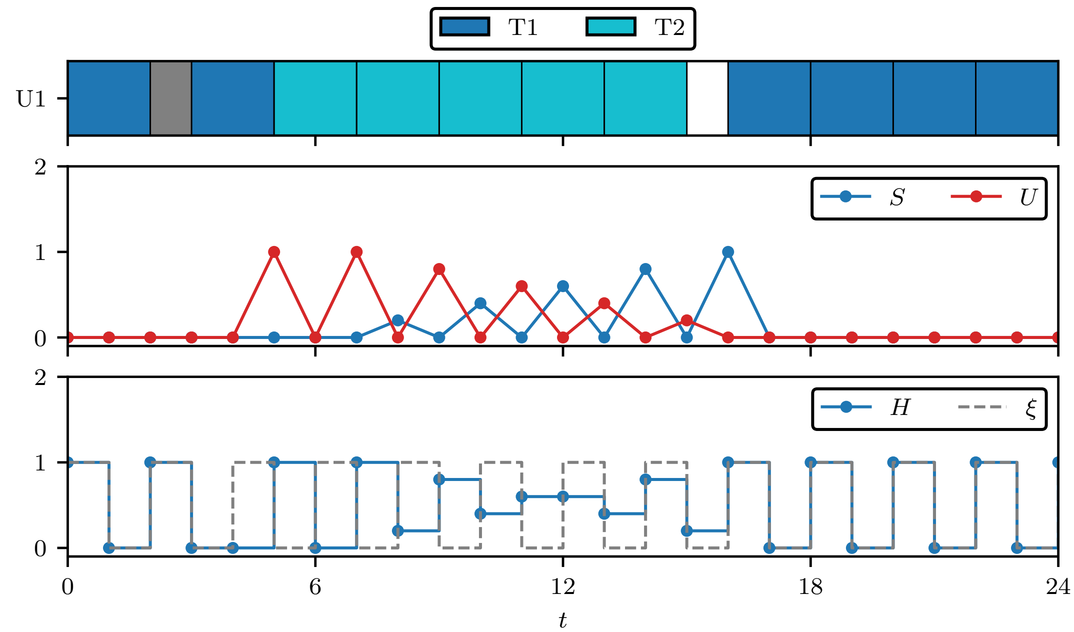

Figure 6.4: Closed-loop schedule for the motivating example with terminal ingredients and subject to a delay at $t = 2$.

as $\Pr(Y = 1) \to 0$. This behavior is an example of robustness, as defined in Definition 6.9 and guaranteed by the terminal cost and constraint chosen for this algorithm.

6.4 Robustness of closed-loop scheduling

In this section, we design a terminal cost and constraint such that the closed-loop scheduling algorithm is robust. We first note that the model, stage cost, and constraints defined for the general production scheduling problem satisfy Assumption 6.4. A reference trajectory that satisfies Assumption 6.2 can be generated by a periodic optimization problem for the nominal system or any heuristic scheduling method.

6.4.1 The terminal constraint

In selecting the terminal constraint for any MPC problem, we must balance two competing goals: to ensure feasibility of the optimization problem and to admit a terminal cost that can satisfy Assumption 6.5. To ensure feasibility, we want to select as large of a terminal region as possible, i.e., choose $\mathbb{X}_f$ as close to $\mathbb{X}$ as possible. To construct a terminal cost that satisfies Assumption 6.5, we want to select as small of a terminal region as possible.

One option for the terminal constraint and cost is to choose $\mathbb{X}_f(t) := \{x_r(t)\}$ and $V_f(\cdot) = 0$. We thereby guarantee that Assumption 6.5 holds with $\kappa_f(x_r(t)) = u_r(t)$. Unfortunately, the terminal equality constraint defined by $\mathbb{X}_f(t)$ significantly reduces the size of the feasible set $\mathcal{X}(t)$ and may render the optimization problem infeasible for relevant disturbances. Thus, we begin by expanding this terminal region to ensure robust recursive feasibility of the sequence of sets $(\mathcal{X}(t))_{t=0}^{\infty}$.

For the subsequent discussion, the reference trajectory is given by

$$x_r(t) = \begin{bmatrix} \bar{W}_r(t) \\ \bar{B}_r(t) \\ \tilde{S}_r(t) \\ S_r(t) \\ U_r(t) \end{bmatrix}$$

For sufficiently long horizons, requiring $\bar{W}$ and $\bar{B}$ to terminate in-phase with the reference trajectory is unlikely to affect feasibility of the optimization problem. Similarly, requiring inventory $\tilde{S}$ and S to meet or exceed their corresponding values in the reference trajectory is unlikely to restrict the optimization problem. We do not, however, allow any amount of inventory that exceeds the reference trajectory in the terminal constraint. Instead, we define the parameters

$$\tilde{\omega}_k = \tilde{\psi}_k - \max_{t \in \mathbb{I}_{\geq 0}} \tilde{S}_{r,k}(t)$$

$$\omega_k = \psi_k - \max_{t \in \mathbb{I}_{\geq 0}} S_{r,k}(t)$$

for all $k \in \mathbf{K}^I$ and $k \in \mathbf{K}^P$, respectively. These parameters represent the maximum amount of inventory that the facility can retain in excess of the reference trajectory without violating maximum inventory constraints for all $t \in \mathbb{I}_{\geq 0}$. In the terminal constraint, we do not allow inventories that exceed the reference trajectory by more than $\tilde{\omega}$ or ω, as appropriate.

We can make these arguments for these elements of the state primarily because these variables are bounded by physical limitations of the system. Thus, a sufficiently long horizon allows the scheduling algorithm to drive the system to the exact configuration required ($\bar{W} = \bar{W}_r(t)$, $\bar{B} = \bar{B}_r(t)$) with sufficient inventories ($\tilde{S} \geq \tilde{S}_r(t)$, $S \geq S_r(t)$). Backlog, however, is not bounded. For any horizon length, there also exists a sufficiently large initial backlog such that the terminal state cannot reach the reference trajectory. Thus, we relax the terminal constraint to include any value of backlog that exceeds the reference trajectory ($U \geq U_r(t)$).

The terminal constraint is therefore

$$\begin{aligned}\mathbb{X}_f(t) := \{x \in \mathbb{X} : \bar{W} = \bar{W}_r(t),\ \bar{B} = \bar{B}_r(t),\\ 0 \leq \tilde{S} - \tilde{S}_r(t) \leq \tilde{\omega},\ 0 \leq S - S_r(t) \leq \omega,\ U \geq U_r(t)\}\end{aligned} \tag{6.17}$$

in which the inequalities are element-wise comparisons. Note that $\mathbb{X}_f(t)$ is closed for all $t \in \mathbb{I}_{\geq 0}$. This relaxed terminal constraint ensures robust recursive feasibility for a much larger class of scheduling problems than a terminal equality constraint. Verifying Assumption 6.10, however, remains difficult for scheduling problems of industrial scale.

With terminal constraints, we still need the horizon length to exceed a potentially unknown minimum to satisfy Assumption 6.10. This requirement is very similar to that of using sufficiently long horizons to guarantee robustness without terminal costs and constraints. There is, however, an important practical difference between these methods that is worth noting. If a horizon length is insufficient to satisfy Assumption 6.10 with terminal constraints, the performance and robustness of the closed-loop system remains satisfactory until an infeasible optimization problem occurs. This infeasible optimization problem is easy to observe and address by increasing the horizon length. If we instead remove the terminal constraint and cost, an insufficient horizon length is not easily recognized and symptoms of this shortcoming may not be obvious until the algorithm is run for a long time. Thus, we may experience poor closed-loop performance until the problem is identified. Furthermore, terminal constraints are prudent because they enable the design of suboptimal closed-loop scheduling algorithms, which we discuss in Section 6.6. We therefore develop an algorithm for closed-loop scheduling that relies on terminal constraints.

6.4.2 The terminal cost

To construct a terminal cost that satisfies Assumption 6.5, one typically begins by determining a terminal control law for the system that renders the reference trajectory (origin) asymptotically stable. The terminal cost is then defined as the infinite horizon cost of the nominal system subject to this terminal control law extending from any $x \in \mathbb{X}_f(t)$, or some suitable approximation of this infinite horizon cost. A key requirement is that the terminal cost is an analytic function such that we can evaluate this cost in the MPC optimization problem.

In previous chapters, we often used a quadratic stage cost and therefore constructed the terminal control law and terminal cost via the LQR solution of the nominal linearized system centered at the target steady-state (or reference trajectory). The terminal constraint can then include any state such that the input constraints are not active. This approach produces a quadratic function for the terminal cost that is easy to implement in the optimization problem.

Closed-loop scheduling problems, however, include discrete-valued input constraints that are always active and use an economic cost function that is neither quadratic nor positive definite with respect to the reference trajectory. By requiring $\bar{W}$ and $\bar{B}$ to terminate exactly in phase with the reference trajectory via the terminal constraint, we need to construct a terminal control law and cost for only inventory and backlog. Due to the many state and input constraints in the scheduling problem, we rely on the additional flexibility afforded by the inventory disposal action to construct this terminal control law and terminal cost. We also require that the reference trajectory *overproduces* each of the final products by some margin $\sigma > 0$. By requiring the reference trajectory to overproduce each product, we ensure that some recourse is available to the system along the reference trajectory, i.e., we can use the excess production to address backlog. We specify these requirements in the following assumption.

Assumption 6.13 (Disposal and overproduction)**.** We can dispose of inventory for all products, i.e., $\mu_k > 0$ for all $k \in \mathbf{K}^P$. The reference trajectory overproduces and disposes of some small amount of each product at every time step, i.e., there exists $\sigma_k > 0$ such that $D_{r,k}(t) \in [\sigma_k, \mu_k/2]$ for all $t \in \mathbb{I}_{\geq 0}$ and $k \in \mathbf{K}^P$. Also, $H_{r,k}(t) + \sigma_k \leq \eta_k$ for all $k \in \mathbf{K}^P$.

The ability to dispose of extra product and the requirement that we do so in the reference trajectory may, at first, seem undesirable. Why would we dispose of material that we may need in the future? The reason is based on the balance between the persistent cost of inventory and the fixed cost of production. For example, if we produce one unit of product material at a cost of 10, but must store the material for 100 time steps at a cost of 1 per time step, the cost of retaining the material in storage far outweighs the cost to just reproduce the material again at a later time. Moreover, if demand at future time steps is subject to significant uncertainty, holding on to potentially useless material may not be an economically beneficial strategy. Recall that this "disposal" action can also include shipping the material to a long-term storage facility or any other action that removes the material from the cost function of the scheduling problem.

We also note that most robust and stochastic optimization approaches to scheduling lead to overproduction of final products relative to the nominal demand pattern. To our knowledge, these methods do not discuss how this excess inventory is to be handled if the realized demand profile does not require this extra production. In contrast to these robust and stochastic optimization approaches to scheduling that are used in an open-loop fashion, we instead require overproduction only in the reference trajectory used to construct the terminal constraint and cost for the scheduling algorithm. Thus, we do not necessarily overproduce or dispose of these products in the closed-loop trajectory, as we demonstrate in subsequent examples.

To construct a reference trajectory that satisfies Assumption 6.13, we can solve a finite horizon, periodic optimization problem with the requirement that $D_k \geq \sigma_k$ for all time steps and some small margin $\sigma_k > 0$ for each $k \in \mathbf{K}^P$. While this reference trajectory results in a higher cost relative to a periodic solution that allows any $D_k \geq 0$, we can choose a small value of σ_k such that there difference between these two periodic solutions is small. For example, we selected $\sigma = 0.01$ in the motivating example. Note that the guarantees in Theorem 6.6 and Theorem 6.12 are now relative to this new reference trajectory. Nonetheless, the closed-loop performance of this scheduling algorithm can, and often does, outperform these bounds.

With this reference trajectory, we define the terminal control law as

$$\kappa(x,t) := \begin{bmatrix} W_r(t) \\ B_r(t) \\ V_r(t) \\ H_r(t) + \min\{\Delta U, \sigma\} \\ D_r(t) + \min\{\Delta S, \mu/2\} - \min\{\Delta U, \sigma\} \end{bmatrix} \tag{6.18}$$

in which $\Delta S := S - S_r(t)$ and $\Delta U := U - U_r(t)$. The first step in verifying Assumption 6.5 is to show that this terminal control law is a feasible input and that $f(x, \kappa(x,t), 0, t) \in \mathbb{X}_f(t+1)$ for all $x \in \mathbb{X}_f(t)$ and $t \in \mathbb{I}_{\geq 0}$. To establish these properties of $\kappa_f(\cdot)$, we exploit the nominal dynamics of a scheduling problem.

We note that the structure of the nominal dynamical equation for a scheduling problem is special in that all the eigenvalues of A are either zero or one. The lifted states are defined by deadbeat dynamics, i.e., eigenvalues of zero, while elements of inventory and backlog are all integrators, i.e., eigenvalues of one. Scheduling problems therefore satisfy the following assumption.

Assumption 6.14 (Integrating dynamics)**.** If $x - x_r(t) = [0\ 0\ \Delta\tilde{S}'\ \Delta S'\ \Delta U']'$ and $u - u_r(t) = [0\ 0\ 0\ \Delta H'\ \Delta D']'$, then $x^+ = f(x, u, 0, t)$ satisfies

$$x^+ - x_r(t+1) = \begin{bmatrix} 0 \\ 0 \\ \Delta\tilde{S}' \\ \Delta S - \Delta H - \Delta D \\ \Delta U - \Delta H \end{bmatrix} \tag{6.19}$$

With this additional assumption, we can establish the following result.

Lemma 6.15. *Consider the system in* (6.2)*, constraint* $\mathbb{Z}$*, and stage cost in* (6.9) *defined for production scheduling. Let Assumptions 6.2, 6.13 and 6.14 hold and let the terminal constraint be defined by* (6.17)*. If* $x \in \mathbb{X}_f(t)$*, then* $(x, \kappa_f(x,t)) \in \mathbb{Z}$ *and* $f(x, \kappa_f(x,t), 0, t) \in \mathbb{X}_f(t+1)$ *for the terminal control law defined in* (6.18).

Proof. From Assumption 6.2, $\kappa_f(\cdot)$ satisfies the required constraints for W, B, and V for all $x \in \mathbb{X}_f(t)$. For $x \in \mathbb{X}_f(t)$, we have that

$$H_r(t) + \min\{\Delta U, \sigma\} \leq H_r(t) + \sigma \leq \eta$$
$$D_r(t) + \min\{\Delta S, \mu/2\} - \min\{\Delta U, \sigma\} \leq \mu$$

from Assumption 6.13. Furthermore, $D_r(t) \geq \sigma$, so $D \geq 0$ as well. Thus, $(x, \kappa_f(x,t)) \in \mathbb{Z}$ for all $x \in \mathbb{X}_f(t)$ and $t \in \mathbb{I}_{\geq 0}$.

Next, we consider $x^+ = f(x, \kappa_f(x,t), 0, t)$. From Assumption 6.14, we have (6.19) with $\Delta\tilde{S}, \Delta S, \Delta U \geq 0$. From the definition of $\kappa_f(\cdot)$, we have that $\Delta H = \min\{\Delta U, \sigma\}$ and $\Delta D = \min\{\Delta S, \mu/2\} - \min\{\Delta U, \sigma\}$ in (6.19). Thus, $\Delta U - \Delta H \geq 0$. We also have that

$$\Delta S - \Delta H - \Delta D = \Delta S - \min\{\Delta S, \mu/2\}.$$

and therefore $0 \leq \Delta S - \Delta H - \Delta D \leq \Delta S$. Since $x \in \mathbb{X}_f(t)$, we have that $0 \leq \Delta S \leq \omega$ and $0 \leq \Delta\tilde{S} \leq \tilde{\omega}$. Thus, we have that $x^+ \in \mathbb{X}_f(t+1)$. □

We now use this terminal control law to construct the following infinite horizon cost function.

$$V_\infty^{\kappa_f}(x,t) := \sum_{k=t}^{\infty} \ell(x(k), \kappa_f(x(k),k), k)$$

in which $x(k+1) = f(x(k), \kappa_f(x(k),k), 0, k)$ and $x(t) = x$. This infinite horizon cost function can in fact be written as an analytic function via the following transformation.

Given the terminal constraint in (6.17), terminal control law in (6.18), and the fact that $\tilde{S}$ does not affect the stage cost, we can reduce the system of interest in this cost function to a lower dimension without loss of information. Specifically, we define the variable

$$z := R_x(x - x_r(t)) = \begin{bmatrix} \Delta S \\ \Delta U \end{bmatrix}$$

for all $x \in \mathbb{X}_f(t)$ and $t \in \mathbb{I}_{\geq 0}$ using an appropriate transformation matrix $R_x \in \mathbb{R}^{\bar{n} \times n}$. From Assumption 6.14 and letting $u = \kappa_f(x, t)$, we have the reduced system dynamics

$$z^+ = z - \min\{z, m\} = \begin{bmatrix} \Delta S - \min\{\Delta S, \mu/2\} \\ \Delta U - \min\{\Delta U, \sigma\} \end{bmatrix}$$

and stage cost

$$\ell(x, u, t) = \tilde{q}' z + \tilde{r}' \min\{z, m\}$$

in which

$$m = \begin{bmatrix} \mu/2 \\ \sigma \end{bmatrix} \qquad \tilde{q} = \begin{bmatrix} \pi^S \\ \pi^U \end{bmatrix} \qquad \tilde{r} = \begin{bmatrix} \pi^D \\ -\pi^D \end{bmatrix}$$

Each element of z_i is nonnegative and given by the equation

$$z_i(k) = \min\{z_i(t) - (k - t)m_i, 0\}$$

for future times $k \in \mathbb{I}_{\geq t}$. Since the stage cost is linear and each element of z is decoupled, we can write the infinite horizon cost as a sum of functions of each element z_i as follows.

$$V_\infty^{\kappa_f}(x, t) = \sum_{i=1}^{\bar{n}} V_{\infty,i}^{\kappa_f}(z_i, t)$$

These individual cost functions for each element of z_i are given by

$$\begin{aligned} V_{\infty,i}^{\kappa_f}(z_i,t) &= \sum_{k=t}^{\infty} \tilde{q}_i z_i(k) + \tilde{r}_i \min\{z_i(k), m_i\} \\ &= \sum_{k=0}^{N_i(z_i)} \tilde{q}_i(z_i - km_i) + \tilde{r}_i z_i \\ &= \tilde{q}_i\left(N_i(z_i)+1\right)\left(z_i - km_i N_i(z_i)/2\right) + \tilde{r}_i z_i \end{aligned}$$

in which $N_i(z_i) = \lfloor z_i/m_i \rfloor$ and $\lfloor \cdot \rfloor$ denotes rounding down to the nearest integer. This infinite horizon cost function is both analytic can be used as a terminal cost because

$$V_{\infty}^{\kappa_f}(f(x,\kappa_f(x,t),0,t),t+1) - V_{\infty}^{\kappa_f}(x,t) = \ell(x,\kappa_f(x,t))$$

for all $x \in \mathbb{X}_f(t)$ and $t \in \mathbb{I}_{\geq 0}$ by the definition of an infinite horizon cost.

We can potentially include this cost function in an optimization problem by replacing $N_i(z_i)$ with the integer decision variable N_i for each element of z. We enforce the constraint $N_i \geq z_i/m_i - 1$ and since $V_{\infty,i}^{\kappa_f}(\cdot)$ increases with increasing N_i, the optimizer selects the smallest integer N_i that satisfies the constraint $N_i \geq z_i/m_i - 1$, i.e., the largest integer less than z_i/m_i. Moreover, we can verify that $V_{\infty,i}^{\kappa_f}(\cdot)$ and therefore $V_{\infty}^{\kappa_f}(\cdot)$ are continuous on $\mathbb{X}_f(t)$.

Alternatively, we can construct a more convenient approximation of this infinite horizon cost function to use as the terminal cost. If we use the approximation $M(z_i) \approx z_i/m_i$, we have

$$V_{\infty,i}^{\kappa_f}(z_i,t) = \frac{\tilde{q}_i}{2m_i} z_i^2 + \left(\frac{\tilde{q}_i}{2} + \tilde{r}_i\right) z_i$$

While this approximation does not satisfy Assumption 6.5, we require only a minor modification to construct the following terminal cost function that is used for the rest of this chapter.

$$V_f(x,t) := \sum_{i=1}^{\bar{n}} \left(\frac{\tilde{q}_i}{2m_i} z_i^2 + (\tilde{q}_i + \tilde{r}_i) z_i \right) \tag{6.20}$$

We also can write (6.20) in the more general form

$$V_f(x,t) = (x - x_r(t))' P (x - x_r(t)) + p'(x - x_r(t)) \tag{6.21}$$

in which $P \in \mathbb{R}^{n \times n}$ and $p \in \mathbb{R}^n$ are defined as

$$P := \text{diag}([0\ 0\ 0\ (\pi^S/\mu)'\ (\pi^U/2\sigma)'])$$
$$p' := [0\ 0\ 0\ (\pi^S + \pi^D)'\ (\pi^U - \pi^D)']$$

To ensure that (6.20) is a valid terminal cost, we require the following minor restrictions on inventory and backlog costs. We also use this assumption in the next subsection to verify that Assumption 6.11 holds.

Assumption 6.16. All final products incur positive inventory and backlog cost, i.e., there exists $c_1 > 0$ such that $\pi_k^S, \pi_k^U \geq c_1$ for all $k \in \mathbf{K}^P$. Inventories for intermediates and feedstock materials are bounded, i.e., $\tilde{\psi}_k < \infty$ for all $k \in \mathbf{K}^I$.

Proposition 6.17. *Consider the system in* (6.2)*, constraint* $\mathbb{Z}$*, and stage cost in* (6.9) *defined for production scheduling. Let Assumptions 6.2, 6.13, 6.14 and 6.16 hold. Then Assumption 6.5 is satisfied for the terminal constraint in* (6.17)*, terminal control law in* (6.18)*, and terminal cost in* (6.20).

Proof. From Lemma 6.15, we have that $(x, \kappa_f(x,t)) \in \mathbb{Z}$ and $x^+ = f(x, \kappa_f(x,t), 0, t) \in \mathbb{X}_f(t+1)$ for all $x \in \mathbb{X}_f(t)$ and $t \in \mathbb{I}_{\geq 0}$. By construction, we have that $V_f(x_r(t), t) = 0$ because $z = 0$ if $x = x_r(t)$ for all $t \in \mathbb{I}_{\geq 0}$. From Assumptions 6.13 and 6.16, we have that $\tilde{q}_i/(2m_i)$ is strictly positive and therefore $V_f(\cdot)$ is bounded from below for all $x \in \mathbb{X}_f$ and

$t \in \mathbb{I}_{\geq 0}$.

Next, we verify that $\kappa_f(x,t)$ satisfies

$$V_f(x^+, t+1) - V_f(x,t) + \ell(x, \kappa_f(x,t), t) \leq 0 \tag{6.22}$$

for all $x \in \mathbb{X}_f(t)$ and $t \in \mathbb{I}_{\geq 0}$. Given the definition of $V_f(\cdot)$ we can equivalently write (6.22) as

$$\sum_{i=1}^{\bar{n}} \Big(V_{f,i}(z_i^+, t+1) - V_{f,i}(z_i, t) + \tilde{q}_i z_i + \tilde{r}_i v_i \Big) \leq 0 \tag{6.23}$$

in which $v_i = \min\{z_i, m_i\}$. We then have that

$$\begin{aligned} V_{f,i}(z_i^+, t+1) &= \frac{\tilde{q}_i}{2m_i}(z_i - v_i)^2 + (\tilde{q}_i + \tilde{r}_i)(z_i - v_i) \\ &= \frac{\tilde{q}_i}{2m_i} z_i^2 - \frac{\tilde{q}_i}{m_i} z_i v_i + \frac{\tilde{q}_i}{2m_i} v_i^2 + (\tilde{q}_i + \tilde{r}_i) z_i - (\tilde{q}_i + \tilde{r}_i) v_i \\ &= V_{f,i}(z_i, t) - \frac{\tilde{q}_i}{m_i} z_i v_i + \frac{\tilde{q}_i}{2m_i} v_i^2 - (\tilde{q}_i + \tilde{r}_i) v_i \end{aligned}$$

We rearrange this equation, add $\tilde{q}_i z_i + \tilde{r}_i v_i$ to both sides, and note that $\tilde{q}_i \geq 0$ from Assumption 6.16 to give

$$V_{f,i}(z_i^+, t+1) - V_{f,i}(z_i, t) + \tilde{q}_i z_i + \tilde{r}_i v_i = \tilde{q}_i \left(\frac{1}{2m_i} v_i^2 - \frac{1}{m_i} z_i v_i + z_i - v_i \right)$$

If $z_i \leq m_i$, we have that $v_i = \min\{z_i, m_i\} = z_i$ and therefore

$$\frac{1}{2m_i} v_i^2 - \frac{1}{m_i} z_i v_i + z_i - v_i = -\frac{1}{2m_i} z_i^2 \leq 0$$

If instead $z_i > m_i$, then $v_i = \min\{z_i, m_i\} = m_i$ and we have

$$\frac{1}{2m_i} v_i^2 - \frac{1}{m_i} z_i v_i + z_i - v_i = -m_i/2 \leq 0$$

Thus, we have that

$$V_{f,i}(z_i^+, t+1) - V_{f,i}(z_i, t) + \tilde{q}_i z_i + \tilde{r}_i v_i \leq 0$$

and each term of the summation in (6.23) is nonpositive. Therefore, $\kappa_f(\cdot)$ satisfies (6.22) for all $x \in \mathbb{X}_f(t)$ and $t \in \mathbb{I}_{\geq 0}$ and Assumption 6.5 holds. □

Note that the terminal cost is quadratic and the resulting optimization problem is therefore a mixed-integer quadratic program (MIQP). Fortunately, most mixed-integer optimization solvers that are frequently used to solve scheduling problems, such as Gurobi, can also handle MIQPs. The proposed optimization problem is therefore well within the capabilities of current optimization software.

Remark 6.18. If quadratic terms must be avoided in the formulation, we can instead choose $P = 0$ and use

$$p' := [0\ 0\ 0\ (b\pi^S/2\mu + \pi^D)'\ (b\pi^U/\sigma - \pi^D)']$$

in (6.21). For any $b \geq 0$, this terminal cost satisfies Assumption 6.5 if $z_i \leq b$ for all i. This observation suggests that a large linear penalty assessed on excess inventory and backlog (relative to the reference trajectory) may be sufficient to ensure robustness in practice.

6.4.3 Verifying Assumption 6.11

With this terminal cost and constraint, we now verify that Assumption 6.11 holds for this problem formulation. First, we note that the production scheduling model and disturbances considered in this chapter satisfy the following assumption.

Assumption 6.19. There exists $e_3 \geq 0$ such that

$$|f(x,u,w,t) - f(x,u,0,t)| \leq e_3$$

for all $(x, u) \in \mathbb{Z}$, $w \in \mathbb{W}$, and $t \in \mathbb{I}_{\geq 0}$.

Furthermore, the nominal system and stage cost are linear and therefore satisfy the properties given in the following two lemmata.

Lemma 6.20. *Let Assumption 6.4 hold. For the nominal system in* (6.3) *and fixed* $N \in \mathbb{I}_{\geq 0}$*, there exist* $e_1, e_2 > 0$ *satisfying*

$$|\hat{\phi}(k; x_1, \mathbf{u}_1, t) - \hat{\phi}(k; x_2, \mathbf{u}_2, t)| \leq e_1|x_1 - x_2| + e_2 \tag{6.24}$$

for all $x_1, x_2 \in \mathbb{X}$*,* $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^N$*,* $t \in \mathbb{I}_{\geq 0}$ *and* $k \in \mathbb{I}_{t:t+N}$*.*

Lemma 6.21. *Let Assumptions 6.4 and 6.16 hold. For the stage cost in* (6.9)*, there exist* $c_1, c_2 > 0$ *and* $d_1, d_2 \geq 0$ *satisfying*

$$c_1|x_1 - x_2| - d_1 \leq |\ell(x_1, u_1, t) - \ell(x_2, u_2, t)| \leq c_2|x_1 - x_2| + d_2$$

for all $x_1, x_2 \in \mathbb{X}$*,* $u_1, u_2 \in \mathbb{U}$*, and* $t \in \mathbb{I}_{\geq 0}$*.*

Using these results, we now establish that Assumption 6.11 holds for this formulation.

Proposition 6.22. *Consider the system in* (6.2)*, constraint* $\mathbb{Z}$*, and stage cost in* (6.9) *defined for production scheduling. Let the terminal constraint and terminal cost be defined by* (6.17) *and* (6.21)*. Let Assumptions 6.2, 6.4, 6.5, 6.10, 6.16 and 6.19 hold. Then we have that Assumption 6.11 holds.*

Proof. Choose $x \in \mathcal{X}(t)$, $t \in \mathbb{I}_{\geq 0}$, and define $x^+ = f(x, \kappa(x, t), w, t)$ for some $w \in \mathbb{W}$ and $\hat{x}^+ = f(x, \kappa(x, t), 0, t)$. Let $\mathbf{u}^+ = \mathbf{u}^0(x^+, t+1)$ and $\hat{\mathbf{u}}^+ = \mathbf{u}^0(\hat{x}^+, t+1)$. We denote $x^+(k) = \hat{\phi}(k; x^+, \mathbf{u}^+, t+1)$, $\hat{x}^+(k) = \hat{\phi}(k; \hat{x}^+, \hat{\mathbf{u}}^+, t+1)$ for all $k \in \mathbb{I}_{t+1:t+1+N}$ and denote $x_f^+ = x^+(t+N+1)$, $\hat{x}_f^+ = \hat{x}^+(t+N+1)$. From Lemma 6.20 and Assumption 6.19, we have that

$$|x^+(k) - \hat{x}^+(k)| \leq e_1 e_3 + e_2 =: e_4$$

We use this bound with Lemma 6.21 to give

$$\begin{aligned}\ell(x^+(k), u^+(k), k) - \ell(\hat{x}^+(k), \hat{u}^+(k), k) &\leq c_2|x^+(k) - \hat{x}^+(k)| + d_2 \\ &\leq c_2 e_4 + d_2\end{aligned}$$

Therefore, the difference between optimal costs is upper bounded by a constant plus the difference between terminal costs.

$$\begin{aligned}V^0(x^+, t+1) - V^0(\hat{x}^+, t+1) \leq& \sum_{k=t+1}^{t+N} \Big(\ell(x^+(k), u^+(k), k) - \ell(\hat{x}^+(k), \hat{u}^+(k), k)\Big) \\ &+ V_f(x_f^+, t+N+1) - V_f(\hat{x}_f^+, t+N+1) \\ \leq& \tilde{b}_1 + V_f(x_f^+, t+N+1) - V_f(\hat{x}_f^+, t+N+1)\end{aligned}$$

in which $\tilde{b}_1 := N(c_2 e_4 + d_2)$.

Next, we construct a bound for the difference between terminal costs. Let $y_1 = x_f^+ - x_r(t+N+1)$, $y_2 = \hat{x}_f^+ - x_r(t+N+1)$, and note that $|y_1 - y_2| = |x_f^+ - \hat{x}_f^+| \leq e_4$. By the definition of $V_f(\cdot)$, we have

$$\begin{aligned}V_f(x_f^+, t) &= y_1' P y_1 + p' y_1 \\ &= (y_1 - y_2 + y_2)' P (y_1 - y_2 + y_2) + p'(y_1 - y_2 + y_2) \\ &= (y_1 - y_2)' P (y_1 - y_2) + (y_1 - y_2)' P y_2 + y_2' P y_2 + p'(y_1 - y_2) + p' y_2 \\ &\leq V_f(\hat{x}_f^+, t) + 2e_4|P||y_2| + |P|e_4^2 + |p|e_4 \\ &\leq V_f(\hat{x}_f^+, t) + 2e_4|P||y_2| + \tilde{b}_2\end{aligned}$$

in which $\tilde{b}_2 := |P|e_4^2 + |p|e_4$. We also have that

$$|y_2| = |\hat{x}_f^+ - x_r(t+N+1)| \leq e_1|\hat{x}^+ - x_r(t+1)| + e_2$$

Since $\hat{x}^+$ is the nominal evolution of the system, we have that

$$\begin{aligned}|\hat{x}^+ - x_r(t+1)| &\leq e_1|x - x_r(t)| + e_2 \\ &\leq (e_1/c_1)|\ell(x,\kappa(x,t),t)| + (e_1d_1/c_1) + e_2\end{aligned}$$

and therefore

$$2e_4|P||y_2| \leq b_1|\ell(x,\kappa(x,t),t)| + \tilde{b}_3$$

in which $b_1 := 2e_4|P|e_1^2/c_1$ and $\tilde{b}_3 := 2e_4|P|(e_1d_1/c_1 + e_1e_2 + e_2)$. Therefore, the optimal cost difference has the following upper bound.

$$V^0(x^+,t+1) - V^0(\hat{x}^+,t+1) \leq b_1|\ell(x,\kappa(x,t),t)| + \tilde{b}_1 + \tilde{b}_2 + \tilde{b}_3 \tag{6.25}$$

From Assumption 6.4, there exists $m \in \mathbb{R}_{\leq 0}$ such that $\ell(x,u,t) \geq m$ for all $(x,u,t) \in \mathbb{Z} \times \mathbb{I}_{\geq 0}$. Therefore, we have that

$$V^0(x^+,t+1) - V^0(\hat{x}^+,t+1) \leq b_1\ell(x,\kappa(x,t),t) + \tilde{b}_1 + \tilde{b}_2 + \tilde{b}_3 + 2b_1m \tag{6.26}$$

With the nominal cost decrease ensured by Assumption 6.5, we have that

$$V^0(\hat{x}^+,t+1) - V^0(x,t) \leq -\ell(x,\kappa(x,t),t) \leq -m$$

We combine this nominal cost decrease with (6.26) to give

$$V^0(x^+, t+1) - V^0(x, t) \leq (b_1 - 1)\ell(x, \kappa(x, t), t) + b_2$$

in which $b_2 \geq \tilde{b}_1 + \tilde{b}_2 + \tilde{b}_3 + (2b_1 - 1)m$. Therefore, Assumption 6.11 holds. □

6.4.4 The main result

With the terminal constraint, terminal cost, and assumptions introduced in this section, we have the following theorem.

Theorem 6.23. *Consider the system in* (6.2)*, constraint* $\mathbb{Z}$*, and stage cost in* (6.9)*, defined for production scheduling. Let the terminal constraint and terminal cost be defined by* (6.17) *and* (6.21)*. Let Assumptions 6.2, 6.4, 6.7, 6.8, 6.10, 6.13, 6.14, 6.16 and 6.19 hold. Then the system* $x^+ = f(x, \kappa(x, t), w, t)$, $w \in \mathbb{W}$ *is economically robust to large, infrequent disturbances with respect to the stage cost* $\ell(\cdot)$ *and the reference trajectory* $(\mathbf{x}_r, \mathbf{u}_r)$ *in the RPI sequence of sets* $(\mathcal{X}(t))_{t=0}^{\infty}$*, i.e., the closed-loop scheduling algorithm is inherently robust.*

Proof. Apply Propositions 6.17 and 6.22 to establish that Assumptions 6.5 and 6.11 hold. Then apply Theorem 6.12 to complete the proof. □

We now review the assumptions used in Theorem 6.23. Assumption 6.2 requires a valid reference trajectory, e.g., a nominal periodic schedule. Assumption 6.4 holds for the model, constraint, and stage cost introduced in this chapter. For production scheduling algorithms, the most pertinent class of disturbances are large and we therefore use Assumption 6.7. Furthermore, we require these disturbances to be independent and identically distributed via Assumption 6.8. By choosing a sufficiently long horizon and defining the terminal constraint to allow a range of values for each material's inventory and an unbounded amount of backlog for each material, we ensure that Assumption 6.10 is satisfied for a general class of production scheduling problems. Assumption 6.13 requires that the reference trajectory overproduces each final product by some small margin σ and that there is an option to dispose of product material. Assumption 6.14 holds because of the structure of the optimization problem and Assumption 6.16 requires that we choose positive costs for maintaining inventory and backlog of product materials. Assumption 6.19 holds for the disturbances considered in this scheduling problem.

We note that Assumptions 6.10 and 6.19 are the only two assumptions that address the behavior of the system subject to disturbances. In the future, we may want to include additional large and infrequent disturbances in this problem formulation that are currently absent. For these additional disturbances, we need to check only that Assumptions 6.10 and 6.19 still hold, since the remaining assumptions are unaffected.

6.5 Example

We consider a scheduling problem adapted from McAllister et al. (2022) and depicted in Figure 6.5. Unit 1 (U1) can run task 1 (T1) to produce the intermediate material M1. Unit 2 (U2) can run task 2 or 3 (T2 or T3) that consume M1 to produce either M2 or M3, respectively. We also allow for a hold task (T4) that can take place after T1. This hold task is important to retain robust recursive feasibility of the scheduling algorithm due to the maximum inventory capacity for M1. There is demand for 45 units of M2 every 6 hours. If we are unable to meet this demand, we accumulate backlog. While there is no specific demand for M3, we may sell up to 5 units of M3 each hour. An optimal schedule therefore meets the demand for M2 while maximizing the production and sale of M3. The parameters for this scheduling problem are specified in Tables 6.1 and 6.2.

For this scheduling problem, we plot the optimal periodic schedule for a 48 hour horizon and 1 hour time steps with the requirement that we overproduce M2 by 0.05 units per hour,

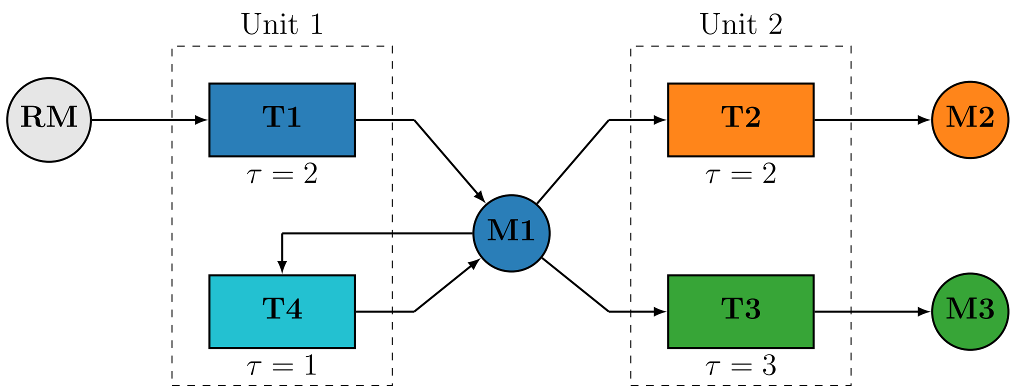

Figure 6.5: Diagram of manufacturing pathways in the example problem.

Table 6.1: Task and unit parameters for the scheduling example problem.

Task (Unit)	β_{ij}^{min}	β_{ij}^{max}	τ_{ij}
T1 (U1)	5	20	2
T2 (U2)	10	20	2
T3 (U2)	10	20	3
T4 (U1)	5	20	1

Table 6.2: Material parameters for the scheduling example problem.

Material	π_k^S	π_k^U	π_k	π_k^D	ν_k^S	μ_k	ψ_k or $\tilde{\psi}_k$
M1	-	-	-	-	0	-	40
M2	1	10	-	12	0	1	100
M3	1	-	10	12	5	1	20

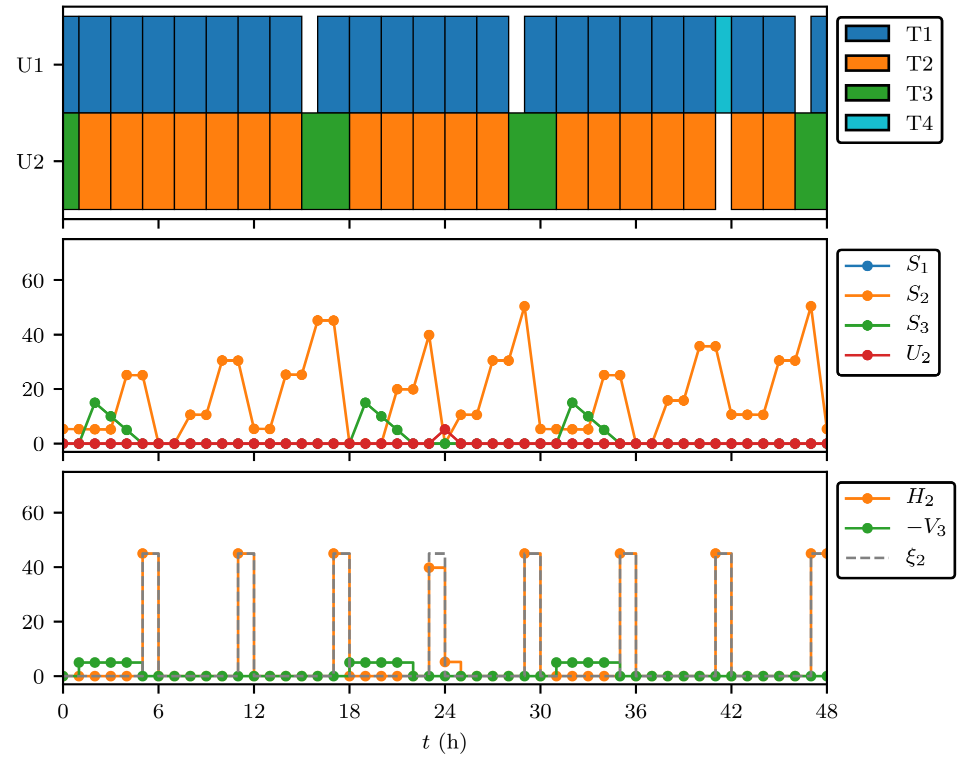

Figure 6.6: Optimal periodic schedule for a horizon length of 48 hours with $\sigma_2 = 0.05$.

i.e., $\sigma_2 = 0.05$ and we constrain $D_2 \geq \sigma_2$ in the periodic optimization problem. We plot this periodic schedule in Figure 6.6 and use this periodic schedule as the reference trajectory in all subsequent simulations. Note that we omit D_2 from Figure 6.6, since the values of D_2 are too small to distinguish from zero on the chosen y-axis scales.

For the closed-loop system, we consider the potential for breakdowns (Z), 1 hour delays (Y), and 20% yield losses (L) for both U1 and U2 at each hour. Let $\varepsilon = \Pr(w \neq 0) = \Pr(w \in \mathbb{W}_1)$, and split the probability of a disturbance occurring equally between each disturbance type, i.e.,

$$\Pr(Y_j = 1) = \Pr(Z_j = 1) = \Pr(L_j = 0.2) = 1 - (1 - \varepsilon)^{1/6}$$

for both $j = 1, 2$. We use a horizon of $N = 12$ for the closed-loop scheduling algorithm with a time step of 1 hour and we use terminal constraints and costs constructed from the periodic

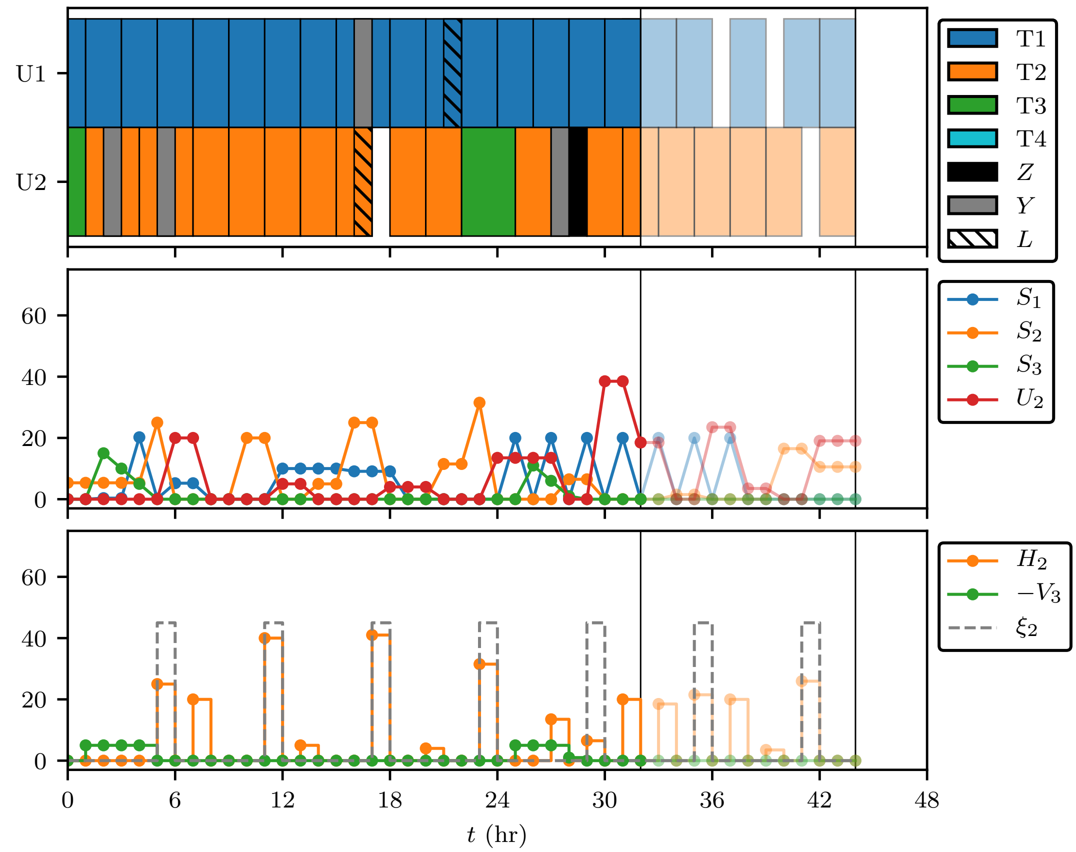

Figure 6.7: An example closed-loop trajectory (solid) and open-loop schedule (faded) for $\varepsilon = 0.2$.

reference trajectory according to (6.17) and (6.21). Since there is no specific demand for M3, we set $U_3 = 0$ and ignored this element of the state in the terminal constraint. We choose the initial state of the facility to be the state of the periodic optimal schedule at $t = 0$. Recall that in this closed-loop scheduling algorithm, we update the state and resolve the optimization problem every hour to compute a new schedule and control action.

In Figure 6.7, we plot an example of the closed-loop trajectory (solid) and current open-loop schedule (faded) with $\varepsilon = 0.1$. Note that backlog in the open-loop schedule is not required to terminate in phase with the reference trajectory, i.e., we allow $U_2(44) \geq 0$. Thus, the optimization problem remains feasible. No inventory is disposed in this closed-loop trajectory or any of the simulations in this section.

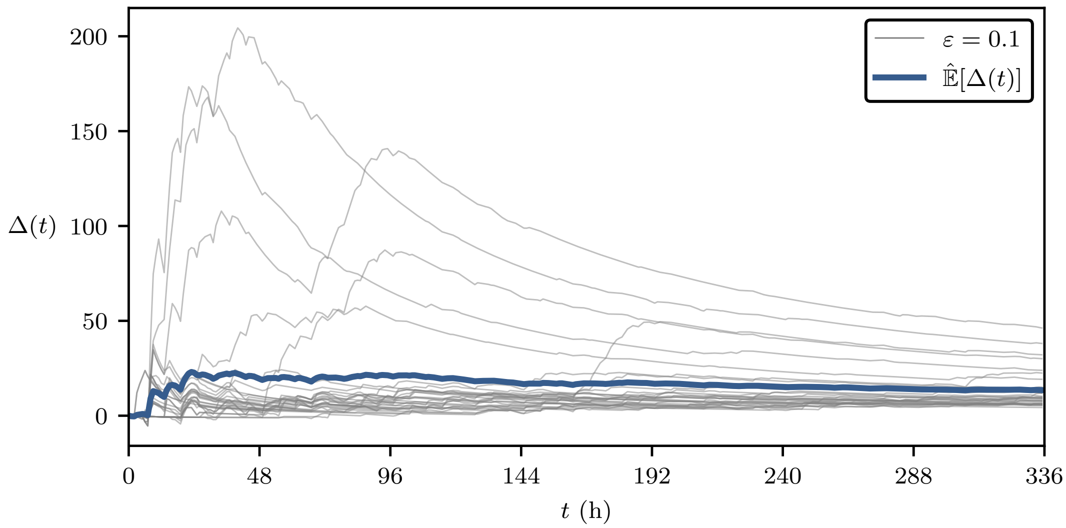

Figure 6.8: Trajectories of $\Delta(t)$ and their sample average for 30 simulations of the closed-loop system with $\varepsilon = 0.1$.

To characterize the economic performance of the closed-loop trajectory, we define

$$\Delta(t) := \frac{1}{t}\sum_{k=0}^{t-1} \ell(x(k), u(k), k)$$

in which $x(k) = \phi(k; x, \mathbf{w}_k, 0)$ is the closed-loop state trajectory and $u(k) = \kappa(x(k), k)$ is the corresponding input trajectory. In other words, $\Delta(t)$ is the time-average cost of the closed-loop trajectory relative to the reference trajectory up time t. Note that this performance metric is used in Definition 6.9 to define robustness. We also used this economic performance metric in Section 5.4.2.

We simulate the closed-loop trajectory for 336 hours (2 weeks) for 30 different realizations of the disturbance trajectory with $\varepsilon = 0.1$. We plot the values of $\Delta(t)$ for each of these closed-loop trajectories in Figure 6.8. Although the individual simulations present a range of trajectories and performance, the sample average of these trajectories appears to converge to a constant value as $t \to \infty$.

In Figure 6.9, we plot the sample average of $\Delta(t)$ for multiple values of ε. We observe that

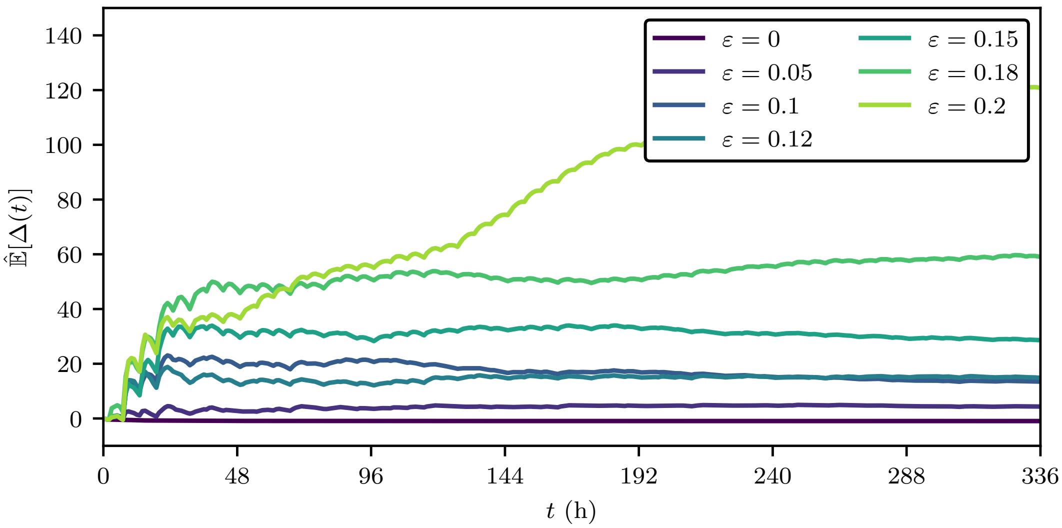

Figure 6.9: Sample average $\hat{\mathbb{E}}[\Delta(t)]$ for 30 simulations of the closed-loop trajectory and multiple values of ε.

for $\varepsilon \leq 0.15$, the trajectories of $\hat{\mathbb{E}}[\Delta(t)]$ appear to converge to a constant value for each ε. Thus, we presume that $\delta \geq 0.15$ for this example in Definition 6.9.

In addition to the terminal constraint and cost proposed in (6.17) and (6.21) (denoted the *LQ* algorithm), we also consider the performance of two other choices of terminal conditions. In the *NTC* algorithm, we omit the terminal constraint and set the terminal cost to zero, i.e., $\mathbb{X}_f = \mathbb{R}^n$ and $V_f(\cdot) = 0$. In the *linear* algorithm, we use the terminal constraint in (6.17) with the linear approximation of (6.21) discussed in Remark 6.18.[1] For each of these algorithms, we observe that $\hat{\mathbb{E}}[\Delta(t)]$ diverges for $\varepsilon > 0.15$. Since divergence of $\hat{\mathbb{E}}[\Delta(t)]$ is primarily the result of physical limitations for the facility (e.g., max production capacity), this observation is reasonable.

For each of these algorithms, we define $\hat{\gamma}(\varepsilon) := \hat{\mathbb{E}}[\Delta(336)]$ as an approximation of the infinite limit in (6.14) for $\varepsilon \leq 0.15$. We plot the values of $\hat{\gamma}(\varepsilon)$ for these three algorithms in Figure 6.10. For the LQ and linear algorithms, the behavior of $\hat{\gamma}(\varepsilon)$ is consistent with Definition 6.9. For the NTC algorithm, however, the value of $\hat{\gamma}(0)$ is greater than zero, indicating

[1]We use the constant $b = 100$ in the approximation detailed in Remark 6.18

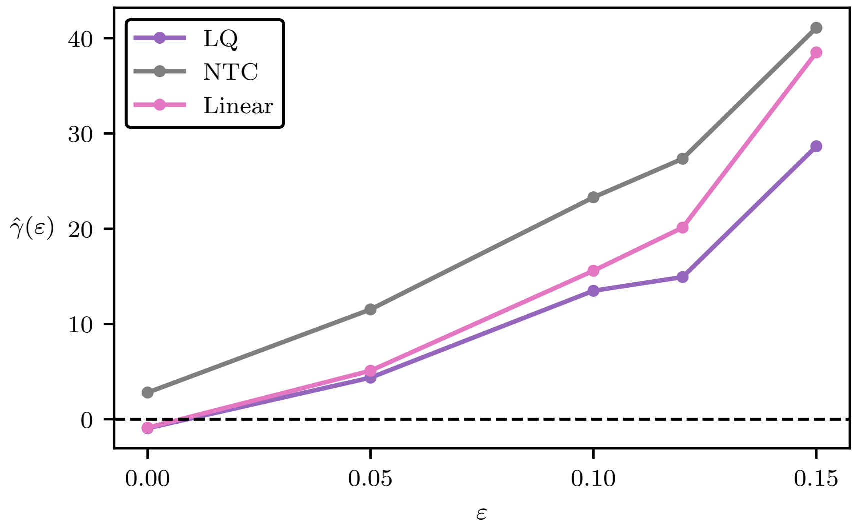

Figure 6.10: The value of $\hat{\gamma}(\varepsilon)$ for multiple values of ε and three production scheduling algorithms with different terminal constraints and costs.

that the nominal closed-loop performance of the system is worse than the reference trajectory. Thus, the NTC algorithm is not robust in terms of Definition 6.9. In fact, the performance of the NTC algorithm is worse than both the LQ and linear algorithm for all values of ε tested. The LQ and linear algorithm both outperform the reference trajectory and do not dispose of any product material. For smaller values of ε, the performance of the linear algorithm is similar to the LQ algorithm.

Note that for all of these simulations with various disturbance trajectories, we did not encounter a single infeasible optimization problem for the LQ or linear closed-loop scheduling algorithm with the terminal constraint in (6.17). Thus, we conclude that a horizon of N=12 is sufficient for Assumption 6.10 to hold with this terminal constraint. If we instead used a terminal equality constraint ($\mathbb{X}_f(t) = \{x_r(t)\}$), we frequently encounter infeasible optimization problems that render the algorithm useless in practice.

6.6 Extensions for large-scale scheduling problems

We now discuss a few extensions to the proposed closed-loop scheduling algorithm that improve the utility of this algorithm in practice. Specifically, we introduce a closed-loop scheduling algorithm that includes a rescheduling penalty, which discourages frequent alterations to the open-loop schedule for the facility, and can operate with *suboptimal* solutions to the proposed optimization problem. It turns out that these two seemingly distinct topics are in fact highly related and therefore merit a combined discussion. We then establish that this new algorithm is also robust to large and infrequent disturbances.

By allowing reoptimization of the entire schedule at each time step, the schedule can change frequently and significantly. In most process control applications of MPC, changes in the open-loop control trajectory are less relevant. For a scheduling problem in a manufacturing facility, however, human operators often observe and execute many components of the schedule computed by these algorithms. Frequent and nonintuitive alterations to the schedule, particular in the near future, may frustrate these operators and lead to so-called schedule nervousness (Kopanos et al., 2008). While no direct economic metric is typically available for this effect, schedule nervousness can nonetheless lead to significant performance loss, increased delays, and even safety concerns in a manufacturing facility (Kopanos et al., 2008). Thus, we want to incorporate a means to balance the economic benefits of rapid rescheduling with the potential schedule nervousness that these rescheduling events may produce in a manufacturing facility. Specifically, we introduce a generalized *rescheduling penalty* that is added to the economic cost function of the previous closed-loop scheduling algorithm.

In addition to schedule nervousness, we also address the computational limitations of the closed-loop scheduling algorithm. Results such as Theorem 6.12, assume that an optimal solution to the MPC problem is available at each time step. For large-scale and industrially relevant production scheduling problems, however, solving these mixed-integer optimization problems to optimality is often intractable. Instead, large-scale MILPs or MIQPs are typically solved to within some global optimality gap, i.e., the cost of the reported solution is within some fixed range of the optimal cost (often reported as a percentage). We therefore want algorithms for closed-loop scheduling that retain these guarantees of nominal and robust economic performance, but do not require optimal solutions to the stated optimization problem. These limitations also exist in other applications of MPC, particular for nonlinear systems. Thus, the design and analysis of suboptimal MPC algorithms, i.e., algorithms that operate without optimal solutions to the MPC problem, are important topics of research.

The approaches to suboptimal MPC are divided into two main categories: warm-start suboptimal MPC and optimality-gap suboptimal MPC. In warm-start suboptimal MPC, a feasible initial control sequence (warm start), based on the previous open-loop trajectory, is used to initialize the optimization algorithm (Scokaert et al., 1999). The optimizer must then compute a control sequence that is no worse than the warm start in terms of the cost function. This algorithm also requires that the terminal control law is available to compute this warm-start input trajectory. With this algorithm, the origin is nominally asymptotically stable and inherently robust to sufficiently small disturbances (Allan et al., 2017; Pannocchia et al., 2011). Moreover, warm-start economic MPC provides the same nominal performance guarantee established in Theorem 6.6 (Risbeck, 2018, Thm. 3.15). A critical requirement to establish the robustness of this warm-start algorithm is that the disturbances are sufficiently small such that that the warm-start input trajectory is feasible for the perturbed state. The large disturbances encountered in production scheduling applications, however, often render the previously computed schedule infeasible, e.g., a task delay may require all of the subsequent tasks in the schedule to be similarly delayed.

In optimality-gap suboptimal MPC, the cost of the computed solution must be within a specific gap of the global optimum (Lazar and Heemels, 2009). Robustness to disturbances is also established for this algorithm, but requires a positive definite stage cost and does not extend to the nominal performance guarantee in Theorem 6.6 (Lazar and Heemels, 2009; Picasso et al., 2012).

Thus, neither of these algorithms meet the requirements of the desired suboptimal closed-loop scheduling algorithm. We instead propose a hybrid algorithm that combines the strengths of these two suboptimal MPC algorithms. This algorithm is specifically designed to address the large disturbances relevant to production scheduling algorithms. Moreover, we include the rescheduling penalty seamlessly in this hybrid algorithm. We now introduce this hybrid algorithm, starting with the rescheduling penalty.

6.6.1 Rescheduling penalties

To define a rescheduling penalty for the system, we first define a warm-start input trajectory for the system by extending the previous schedule $\mathbf{u}$ with the terminal control law, i.e.,

$$\tilde{\mathbf{u}}^{+}(x, \mathbf{u}, t) := (u(t+1), \ldots, u(t+N-1), \kappa(x(t+N), t+N))$$

in which $x(t+N) = \hat{\phi}(t+N; x, \mathbf{u}, t)$. This input trajectory can also be treated as the *incumbent* schedule, i.e., the trajectory in which no tasks are rescheduled. We then augment the economic cost function in the previous closed-loop scheduling algorithm with a rescheduling penalty $R : \mathbb{U}^N \times \mathbb{U}^N \to \mathbb{R}_{\geq 0}$ such that the new cost function becomes

$$V_R(x, \tilde{\mathbf{u}}, \mathbf{u}, t) = V(x, \mathbf{u}, t) + R(\tilde{\mathbf{u}}, \mathbf{u})$$

With this new cost function, the optimization problem is now defined as

$$\mathbb{P}_R(x, \tilde{\mathbf{u}}, t) : \; V_R^0(x, \tilde{\mathbf{u}}, t) := \min_{\mathbf{u} \in \mathcal{U}(x,t)} V_R(x, \tilde{\mathbf{u}}, \mathbf{u}, t) \tag{6.27}$$

Note that the constraint set $\mathcal{U}(x,t)$ does not change and therefore the feasible set $\mathcal{X}(t)$ for this optimization problem is unchanged. We require that this rescheduling penalty is continuous and is zero if the schedule is unchanged.

Assumption 6.24 (Rescheduling penalty)**.** The rescheduling penalty $R : \mathbb{U}^N \times \mathbb{U}^N \to \mathbb{R}_{\geq 0}$ is continuous and satisfies $R(\mathbf{u},\mathbf{u}) = 0$ for all $\mathbf{u} \in \mathbb{U}^N$.

For example, we can define the rescheduling penalty as

$$R(\tilde{\mathbf{u}},\mathbf{u}) = \sum_{k=t}^{t+N-1} |\tilde{u}(k) - u(k)|$$

or we can implement more complicated and asymmetric penalties tailored to the application of interest. For more details, discussion, and additional penalty forms see McAllister et al. (2020).

6.6.2 A hybrid suboptimal MPC algorithm

We now introduce the hybrid suboptimal MPC algorithm that combines the features of both warm-start and optimality-gap suboptimal MPC. This algorithm requires that the computed solution achieves an optimality gap of less that some specified constant $\rho \geq 0$. If the warm start is feasible, we also require the computed solution to be no worse than the warm start in terms of the cost function. We define this algorithm in the following paragraph.

The set of admissible input trajectories for warm-start suboptimal MPC is

$$\check{\mathcal{U}}^w(x,\tilde{\mathbf{u}},t) := \{\mathbf{u} : \mathbf{u} \in \mathcal{U}(x,t),\ V_R(x,\tilde{\mathbf{u}},\mathbf{u},t) \leq V_R(x,\tilde{\mathbf{u}},\tilde{\mathbf{u}},t)\}$$

Note that if $\tilde{\mathbf{u}} \notin \mathcal{U}(x,t)$, i.e., the warm start is infeasible, then $\check{\mathcal{U}}^w(x,\tilde{\mathbf{u}},t)$ may be empty. We define the set of admissible input trajectories with an optimality gap less than $\rho \geq 0$ as

$$\check{\mathcal{U}}^g(x, \tilde{\mathbf{u}}, t) := \{\mathbf{u} : \mathbf{u} \in \mathcal{U}(x, t),\ V_R(x, \tilde{\mathbf{u}}, \mathbf{u}, t) \leq V_R^0(x, \tilde{\mathbf{u}}, t) + \rho\}$$

The hybrid algorithm is defined by the following combination of these sets.

$$\check{\mathcal{U}}(x, \tilde{\mathbf{u}}, t) := \begin{cases} \check{\mathcal{U}}^w(x, \tilde{\mathbf{u}}, t) \cap \check{\mathcal{U}}^g(x, \tilde{\mathbf{u}}, t) & ;\ \tilde{\mathbf{u}} \in \mathcal{U}(x, t) \\ \check{\mathcal{U}}^g(x, \tilde{\mathbf{u}}, t) & ;\ \tilde{\mathbf{u}} \notin \mathcal{U}(x, t) \end{cases}$$

We summarize this suboptimal MPC algorithm as follows.

Algorithm 6.25 (Hybrid suboptimal MPC)**.** Obtain the initial state $x \in \mathcal{X}(t)$ and any initial warm start $\tilde{\mathbf{u}} \in \mathbb{U}^N$. Then repeat

1. Obtain the current state x and warm start $\tilde{\mathbf{u}}$.
2. Compute any $\mathbf{u} \in \check{\mathcal{U}}(x, \tilde{\mathbf{u}}, t)$.
   (a) If $\tilde{\mathbf{u}}$ is feasible, compute $\mathbf{u} \in \check{\mathcal{U}}^w(x, \tilde{\mathbf{u}}, t) \cap \check{\mathcal{U}}^g(x, \tilde{\mathbf{u}}, t)$.
   (b) If $\tilde{\mathbf{u}}$ is not feasible, compute $\mathbf{u} \in \check{\mathcal{U}}^g(x, \tilde{\mathbf{u}}, t)$.
   (c) Inject the first element of the input sequence $\mathbf{u}$.
   (d) Compute the next warm start $\tilde{\mathbf{u}}^+(x, \mathbf{u}, t)$.

Since the control action is now a function of the warm start, we find it convenient to discuss the extended state $z := (x, \tilde{\mathbf{u}})$ for the system. The dynamics of this extended state are given by

$$z^+ \in H(z, w, t) := \left\{ \begin{pmatrix} f(x, u, w, t) \\ \tilde{\mathbf{u}}^+(x, \mathbf{u}, t) \end{pmatrix} :\ \mathbf{u} \in \check{\mathcal{U}}(z, t) \right\},\ w \in \mathbb{W} \tag{6.28}$$

in which u is the first element of $\mathbf{u}$. We also define the sets

$$\mathcal{Z}(t) := \mathcal{X}(t) \times \mathbb{U}^N$$

for all $t \in \mathbb{I}_{\geq 0}$. We use $\phi_z(k; z, \mathbf{w}_k, t)$ to denote any solution of (6.28) with the initial extended state $z \in \mathcal{X}(t) \times \mathbb{U}^N$ at time $t \in \mathbb{I}_{\geq 0}$, given the disturbance sequence $\mathbf{w}_{t:k} \in \mathbb{W}^{k-t}$. This trajectory is therefore a *selection* from the set of potential solutions for the closed-loop system defined by (6.28). All results are then established for any such selection. For this trajectory, we use $\phi_x(k; z, \mathbf{w}_k, t)$ and $\phi_u(k; z, \mathbf{w}_k, t)$ to denote the corresponding x and u trajectory, respectively. We also use $\phi_{\mathbf{u}}(k; z, \mathbf{w}_k, t)$ to denote the computed open-loop control trajectory $\mathbf{u}$ at each time step.

6.6.3 Robustness to large and infrequent disturbances

Since we are now addressing the extended state of the system, we need to modify the definition of economic robustness in Definition 6.9 to suit the system in (6.28).

Definition 6.26 (Economically robust to large, infrequent disturbances)**.** The system in (6.28) is economically robust to large, infrequent disturbances with respect to the stage cost $\ell(\cdot)$ and the reference trajectory $(\mathbf{x}_r, \mathbf{u}_r)$ in an RPI sequence of sets $(\mathcal{Z}(t))_{t=0}^{\infty}$ if there exist $\delta > 0$ and $\gamma(\cdot) \in \mathcal{K}$ such that

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{k=t}^{t+T-1} \mathbb{E}[\ell(x(k), u(k), k)] \leq \gamma(\mu(\mathbb{W}_1)) \tag{6.29}$$

in which $x(k) = \phi_x(k; z, \mathbf{w}_{t:k}, t)$ and $u(k) = \phi_u(k; z, \mathbf{w}_{t:k}, t)$ for all $z \in \mathcal{Z}(t)$, $\mu \in \mathcal{M}(\mathbb{W}, \delta)$, and $t \in \mathbb{I}_{\geq 0}$.

We also modify Assumption 6.10 as follows.

Assumption 6.27 (Robust recursive feasibility)**.** The sequence of sets $(\mathcal{Z}(t))_{t=0}^{\infty}$ is robustly positive invariant for the system in (6.28), i.e., $H(z,w,t) \subseteq \mathcal{Z}(t+1)$ for all $z \in \mathcal{Z}(t)$, $w \in \mathbb{W}$, and $t \in \mathbb{I}_{\geq 0}$.

We require one additional assumption for the stage cost of the system. Note that this assumption is satisfied by the linear stage costs used in the production scheduling problem formulation.

Assumption 6.28 (Uniformly continuous stage cost)**.** There exists $d \geq 0$ such that

$$|\ell(x,u_1,t) - \ell(x,u_2,t)| \leq d$$

for all $(x,u_1) \in \mathbb{Z}$, $(x,u_2) \in \mathbb{Z}$, and $t \in \mathbb{I}_{\geq 0}$.

We can now establish the following result.

Theorem 6.29. *Let Assumptions 6.2, 6.4, 6.5, 6.7, 6.8, 6.11, 6.24, 6.27 and 6.28 hold. Then the system in* (6.28) *is economically robust to large, infrequent disturbances with respect to the stage cost $\ell(\cdot)$ and reference trajectory $(\mathbf{x}_r, \mathbf{u}_r)$ in the RPI sequence of sets $(\mathcal{Z}(t))_{t=0}^{\infty}$.*

Proof. Choose any $z = (x, \tilde{\mathbf{u}}) \in \mathcal{Z}(t)$, $t \in \mathbb{I}_{\geq 0}$, and an input trajectory that satisfies $\mathbf{u} \in \check{\mathcal{U}}(z,t)$. Denote the first input of this trajectory as u. We denote the subsequent extended state as $z^+ = (x^+, \tilde{\mathbf{u}}^+) \in H(z,w,t)$.

If $w = 0$, we have that $\tilde{\mathbf{u}}^+ \in \mathcal{U}(x^+, t+1)$ by Assumption 6.5 and the subsequent input trajectory must satisfy $\mathbf{u}^+ \in \check{\mathcal{U}}^w(z, t+1)$. Let $x_f := \hat{\phi}(N; x, \mathbf{u}, t)$, $u_f := \kappa(x_f, t+N)$, and

$x_f^+ = f(x_f, u_f, 0, t)$. From Assumption 6.5 and Assumption 6.24, we have that

$$\begin{aligned}
V_R(z^+, \mathbf{u}^+, t+1) &\leq V_R(z^+, \tilde{\mathbf{u}}^+, t+1) \\
&= V(z^+, t+1) \\
&= V(x, \mathbf{u}, t) - \ell(x, u, t) \\
&\qquad - V_f(x_f, t+N) + \ell(x_f, u_r, t+N) + V_f(x_f^+, t+N+1) \\
&\leq V(x, \mathbf{u}, t) - \ell(x, u, t) \\
&\leq V_R(z, \mathbf{u}, t) - \ell(x, u, t)
\end{aligned}$$

If $w \in \mathbb{W}_1$, we have $x^+ = f(x, u, w, t)$ and note that $\tilde{\mathbf{u}}^+$ is not necessarily a feasible warm start for x^+. Instead, we know that $\mathbf{u}^+ \in \check{\mathcal{U}}^g(z, t+1)$ and therefore

$$V_R(z^+, \mathbf{u}^+, t+1) \leq V_R^0(z^+, t+1) + \rho$$

Since $R(\cdot)$ is continuous and $\mathbb{U}$ is compact, there exists $r \geq 0$ such that $R(\tilde{\mathbf{u}}, \mathbf{u}) \leq r$ for all $\tilde{\mathbf{u}}, \mathbf{u} \in \mathbb{U}^N$. From Assumption 6.11, we therefore have

$$\begin{aligned}
V_R(z^+, \mathbf{u}^+, t+1) &\leq V_R^0(z^+, t+1) + \rho \\
&\leq V^0(x^+, t+1) + r + \rho \\
&\leq V^0(x, t) + b_1 \ell(x, \kappa(x, t), t) + b_2 + r + \rho \\
&\leq V_R(z, \mathbf{u}, t) + b_1 \ell(x, \kappa(x, t), t) + b_2 + r + \rho
\end{aligned}$$

We now apply Assumption 6.28 to give

$$V_R(z^+, \mathbf{u}^+, t+1) \leq V_R(z, \mathbf{u}, t) + b_1 \ell(x, u, t) + b_3$$

in which $b_3 := b_1 d + b_2 + r + \rho$.

To streamline notation, we now define $y := (z, \mathbf{u})$ and $y^+ = (z^+, \mathbf{u}^+)$. We then combine the bounds with and without the disturbance as follows.

$$V_R(y^+, t+1) \leq V_R(y, t) - (1 - I_{\mathbb{W}_1}(w))\ell(x, u, t) + I_{\mathbb{W}_1}(w)(b_1 \ell(x, u, t) + b_3)$$

Taking the expected value and combining terms gives,

$$\mathbb{E}\left[V_R(y^+, t+1)\right] - \mathbb{E}\left[V_R(y, t)\right] \leq -(1 - \mu(\mathbb{W}_1) - b_1 \mu(\mathbb{W}_1))\mathbb{E}\left[\ell(x, u, t)\right] + b_3 \mu(\mathbb{W}_1)$$

We choose $0 < \delta < 1/(1 + b_1)$ such that

$$\mathbb{E}\left[V_R(y^+, t+1)\right] - \mathbb{E}\left[V_R(y, t)\right] \leq -b_4 \mathbb{E}\left[\ell(x, u, t)\right] + b_3 \mu(\mathbb{W}_1) \tag{6.30}$$

with $b_4 := 1 - (1 + b_1)\delta$. Note that $b_4 > 0$.

From an initial $z \in \mathcal{Z}(t)$ and $t \in \mathbb{I}_{\geq 0}$, we denote the closed-loop trajectories as $z(k) = \phi_z(k; z, \mathbf{w}_{t:k}, t)$, $x(k) = \phi_x(k; z, \mathbf{w}_{t:k}, t)$, $u(k) = \phi_u(k; z, \mathbf{w}_{t:k}, t)$, and $\mathbf{u}(k) = \phi_{\mathbf{u}}(k; z, \mathbf{w}_{t:k}, t)$. Furthermore, we define $y(k) = (z(k), \mathbf{u}(k))$. With (6.30) and the properties of iterated expectations, we have

$$\mathbb{E}\left[V_R(y(k+1), k+1)\right] - \mathbb{E}\left[V_R(y(k), k)\right] \leq -b_4 \mathbb{E}\left[\ell(x(k), u(k), k)\right] + b_3 \mu(\mathbb{W}_1)$$

for all $k \in \mathbb{I}_{\geq t}$. We sum both sides of this inequality from t to $t + T - 1$, divide by T, and rearrange to give

$$\frac{b_4}{T} \sum_{k=t}^{t+T-1} \mathbb{E}\left[\ell(x(k), u(k), k)\right] \leq \frac{V_R(y(t), t) - \mathbb{E}\left[V_R(y(t+T), t+T)\right]}{T} + b_3 \mu(\mathbb{W}_1)$$

By Assumption 6.4 and Assumption 6.24, there exists $M \in \mathbb{R}$ such that $V_R(z, \mathbf{u}, k) \geq M$ for all $z \in \mathcal{Z}(k)$, $\mathbf{u} \in \mathbb{U}^N$, and $k \in \mathbb{I}_{\geq 0}$. Therefore, we have

$$\frac{1}{T}\sum_{k=t}^{t+T-1} \mathbb{E}\left[\ell(x(k), u(k), k)\right] \leq \frac{V_R(y(t), t) - M}{b_4 T} + \gamma(\mu(\mathbb{W}_1))$$

in which $\gamma(s) := (b_3/b_4)s$. We take the limit supremum as $T \to \infty$ so that the initial cost and M vanish to give (6.29). Note that $\gamma(\cdot) \in \mathcal{K}$ to complete the proof. □

Thus, the proposed suboptimal MPC algorithm with rescheduling penalties is economically robust to large, infrequent disturbances. We can now apply this algorithm to closed-loop scheduling and establish the following corollary of Theorem 6.29.

Corollary 6.30. *Consider the system in* (6.2)*, constraint* $\mathbb{Z}$*, and stage cost in* (6.9) *defined for production scheduling. Let the terminal constraint and terminal cost be defined by* (6.17) *and* (6.21)*. Let Assumptions 6.2, 6.4, 6.7, 6.8, 6.13, 6.14, 6.16, 6.19 and 6.27 hold. Then the system in* (6.28) *is economically robust to large, infrequent disturbances with respect to the stage cost* $\ell(\cdot)$ *and reference trajectory* $(\mathbf{x}_r, \mathbf{u}_r)$ *in the RPI sequence of sets* $(\mathcal{Z}(t))_{t=0}^{\infty}$*, i.e., the suboptimal closed-loop scheduling algorithm with a rescheduling penalty is inherently robust.*

6.7 Summary

In this chapter, we introduced a general class of production scheduling problems and demonstrated that closed-loop or online scheduling algorithms fit within the framework of MPC. We then defined the term robustness for the class of large and infrequent disturbances prevalent in production scheduling applications and justified this definition with a small motivating example. With suitable assumptions and a carefully constructed terminal constraint and cost, we established that the proposed closed-loop scheduling algorithm is inherently robust to large and infrequent disturbances such as breakdowns, delays, and yield losses. To address the high computational burden and potential for "schedule nervousness" inherent to closed-loop scheduling, we further developed a suboptimal MPC algorithm that includes a rescheduling penalty. Moreover, we established that this novel suboptimal MPC algorithm is also robust to large and infrequent disturbances. The proposed algorithm is therefore suitable for large-scale industrial applications of closed-loop scheduling.

We recognize, however, that constructing the terminal constraint and cost as detailed in (6.17) and (6.21) may be inconvenient in engineering practice for many systems of interest. We therefore conclude this chapter with a selection of well motivated suggestions to construct a more practical approximation of the algorithm proposed in this chapter:

1. Determine a suitable reference trajectory for the nominal manufacturing facility based on heuristics, optimization, past experience, or some combination thereof.

2. Use this reference trajectory as a terminal equality constraint for all state variables aside from inventory and backlog.

3. Allow the terminal region to include any values of inventory and backlog that exceed the reference trajectory.

4. Define the terminal cost to asses a large linear penalty on deviations of inventory and backlog from their corresponding values in the reference trajectory.

5. Use the incumbent schedule as a warm start for the MILP produced by this MPC problem and set a minimum optimality gap.

6. If the optimization problem becomes infeasible, increase the horizon length.

7. If schedule nervousness is a concern in the manufacturing facility, implement a rescheduling penalty that places a 1-norm penalty on changes to the decision variable W in the input trajectory, i.e., $R(\tilde{\mathbf{u}}, \mathbf{u}) = \sum_{k=t}^{t+N-1} |\tilde{W}(k) - W(k)|_1$.

While this approximate algorithm does not exactly satisfy the required assumptions of Theorem 6.12 or Theorem 6.29, these suggestions may significantly improve the performance and robustness of closed-loop scheduling algorithms compared to algorithms without any terminal costs or constraints.

Chapter 7

Conclusions

A key question in using models to design, schedule, and control chemical processes is how to handle uncertainty. Do we attempt to characterize and include this uncertainty directly in the problem formulation? Respond to this uncertainty as it is observed? Or use some combination of these two methodologies? In process control, the preferred method to address uncertainty is through feedback from the system. Even for cutting edge control algorithms based on stochastic optimization, feedback still remains an integral part of the control algorithm. Both algorithmically simple to implement and intuitively similar to the manner in which humans naturally address uncertainty, feedback has proven time and again to be a remarkably powerful tool in engineering practice.

Although powerful, feedback methods also introduce their own set of complications to the underlying system analysis. The dynamics of the closed-loop system emerging from this feedback control structure are complicated and sometimes nonintutive. Analyzing and characterizing the behavior of new control algorithms and the resulting closed-loop systems are therefore essential components of control theory.

One can, and often does, use a deterministic description of this uncertainty to characterize the behavior and robustness of these closed-loop systems. While this deterministic description produces instructive and important results, stochastic descriptions of uncertainty are often better suited to model the behavior of physical systems. We therefore pursued stochastic descriptions of uncertainty and robustness throughout this dissertation. In particular, we developed a extensive theory for the stochastic robustness of MPC, the advanced control method of choice for chemical process control with applications in many other engineering disciplines as well.

In Chapter 2, we discussed and defined stochastic robustness for a closed-loop system. We then established that nominal MPC is robust in this stochastic context for sufficiently small disturbances. While significant in its own right, this result is particularly important because we can now compare the theoretical properties of nominal and stochastic MPC in terms of this definition of stochastic robustness.

In Chapter 3, we introduced SMPC and established some fundamental mathematical properties for the problem formulation that are often neglected in SMPC literature. We then established that idealized SMPC admits the same definition of stochastic robustness that was established for nominal MPC. This result suggests that nominal and stochastic MPC provide the same qualitative closed-loop behavior in the presence of stochastic uncertainty in the dynamical model. The comparison of robustness across these two MPC formulations is therefore a quantitative inquiry, best explored through simulation studies.

A significant limitation of the analysis in Chapter 3 and throughout much of stochastic optimal control is that these results assume the stochastic description of uncertainty used in the problem formulation is equivalent to the stochastic uncertainty encountered in the plant. While convenient for the analysis of these stochastic optimal control methods, this assumption does not hold in practice. Stochastic descriptions of uncertainty are subject to their own distributional uncertainty. In Chapter 4, we addressed this limitation and established, under suitable assumptions, that SMPC is distributionally robust. Sufficiently small errors in the stochastic dynamical model therefore produce similarly small deviations in the closed-loop performance of SMPC. Scenario-based approximations of the SMPC optimization problem are covered by this result as well. This result is also important because it unifies the seemingly disparate results established for nominal and stochastic MPC in Chapters 2 and 3. In fact, Theorem 4.18 serves as a single summarizing result of the analysis developed throughout Chapters 2 to 4.

With this analysis, we intentionally raise a particularly contentious issue in control theory and algorithm design that is best stated by Wonham (1969):

> Since the mathematical model is usually greatly complicated by explicitly including stochastic features, it is always to be asked whether the extra effort is worthwhile, i.e. whether it leads to a control markedly superior in performance to one designed on the assumption that stochastic disturbances are absent.

While we restricted the discussion in this dissertation to a comparison of nominal and stochastic MPC, the debate between nominal and stochastic optimal control methods is broader and older than the research field of MPC. At the time, Wonham presented a conclusion that was very critical of stochastic optimal control.

> In the case of feedback controls the general conclusion is that only marginal improvements can be obtained unless the disturbance level is very high, in which case the fractional improvement from stochastic optimization may be large, but the system is useless anyway, That is, efforts to counter disturbances by simply warping the velocity field in state space are generally misplaced.

In recent decades, the effort required to include these stochastic features in controller design and optimization problems has been greatly diminished by algorithmic and computational advances. Nonetheless, the benefits of including this stochastic information remain unclear. Nominal MPC, for example, offers similar qualitative behavior to that of SMPC. Through several examples in Section 3.7, we further demonstrated that SMPC is not necessarily more robust than nominal MPC in terms of relevant quantitative metrics of robustness. In particular, SMPC can produce nominal closed-loop systems that are not asymptotically stable, resulting in nonintuitive and potentially undesirable closed-loop behavior. If we consider sufficiently large disturbances such that nominal MPC is no longer robust, we may not be able to design an SMPC algorithm for this class of disturbances anyway. Thus, the conclusion of Wonham remains remarkably relevant and accurate today for setpoint tracking applications of MPC.

There is perhaps one update that should be made to Wonham's conclusion as it pertains to *economic* applications of MPC. In these economic applications, performance in terms of the stage cost is more important than the stability of a target steady state (or reference trajectory) for the closed-loop system. For economic applications, the goal of SMPC is not only to reject disturbances by "warping the velocity field", but also to select operating trajectories that boast superior expected performance than the potentially suboptimal steady state or reference trajectory supplied to the MPC problem. These economic costs are frequently "asymmetric" in that the cost of deviating from the setpoint or reference trajectory is smaller in one direction than the opposite direction. Exploiting these asymmetries in the presence of uncertainty can prove beneficial. The inventory control problem in Section 3.7.4 provides a simple example of this asymmetry and the potential benefit of SMPC.

The first part of this dissertation, spanning Chapters 2 to 4, addressed the question of stochastic robustness for the class of sufficiently small disturbances that best characterizes the uncertainty observed in most process control applications. But as we expand the purview of MPC to include higher-level planning and scheduling problems, we must also expand the class of disturbances considered in these robustness results. In Chapter 5, we considered a class of large and infrequent disturbances that best characterize the uncertainty observed in these planning and scheduling applications of MPC. We then established that nominal MPC is robust to this class of large disturbances, in a stochastic context, provided that these disturbances are also sufficiently infrequent. This result requires that the nominal MPC algorithm is robustly recursively feasible *by design*, but this assumption is often satisfied for these planning and scheduling problems through careful design of the constraints.

In Chapter 6, we applied these results for large and infrequent disturbances to design a closed-loop scheduling algorithm that is inherently robust relative to a prescribed reference trajectory for the system. In particular, we developed a method to construct a terminal constraint and cost for a general production scheduling problem that satisfy the required assumptions detailed in Chapter 5. We can also add rescheduling penalties to this problem formulation to discourage alterations to the open-loop schedule and thereby improve operator acceptance of the new algorithm. We also detailed a closed-loop scheduling algorithm that can operate with suboptimal solutions to the proposed optimization problem. We further established that this suboptimal algorithm remains robust to large and infrequent disturbances. The proposed algorithm is therefore computationally efficient and can be applied to large-scale applications of closed-loop scheduling.

The conceptual development of defining robustness for closed-loop scheduling is perhaps a more important contribution in Chapter 6 than the development of a new closed-loop scheduling algorithm. This definition of robustness characterizes the behavior of the closed-loop system and is therefore not restricted to MPC. The property is general, and therefore applicable to any closed-loop scheduling algorithm.

The definitions of robustness for the control of dynamical systems have evolved considerably over time. The ISS-framework pioneered by Sontag and Wang (1995) has proven both highly useful and flexible for the study of both the stochastic and distributional robustness of nominal and stochastic MPC. Many of the developments in this dissertation apply to closed-loop systems in general and may therefore find utility in the larger field of stochastic optimal control. With that in mind, we close with some suggestions for future directions of research.

7.1 Future directions

We discuss several potential directions for future research based on the results in this dissertation. Some of these topics are quite promising and indicated as such. Others are subject to numerous limitations that may hinder any further development of the idea, which we also note in the subsequent discussion.

7.1.1 Economic MPC

In Corollary 4.25, we established that economic MPC, without the requirement of a positive definite or dissipative stage cost, is robust in terms of the average economic performance. This fortuitous result requires an interesting and new restriction in (4.24) for the terminal control law and terminal cost. Since this result was not the original focus of this research direction, the implications of this result were not further explored. Thus, a potential direction of future research is to investigate this result further. In particular, we want to determine if the there exist terminal costs and control laws that satisfy the nominal cost decrease condition in (4.6) as well as (4.24). If so, is there a general procedure to construct this terminal cost?

7.1.2 SMPC

One of the significant shortcomings of SMPC for nonlinear systems is that the closed-loop system is not necessarily RAS. Thus, the closed-loop system may not stabilize the specified setpoint or reference trajectory in the nominal case. While this behavior may be acceptable in economic application, in which stability of a specific setpoint is not required, this behavior may be nonintuitive or undesirable for setpoint tracking applications of MPC.

For linear systems and quadratic costs, RAS is guaranteed for SMPC. This result is established by exploiting the convexity and continuity of the optimal cost function. For nonlinear systems and more general cost functions, are there additional conditions on the SMPC problem that guarantee RAS? If so, we can use these conditions to determine if SMPC is liable to produce undesirable closed-loop behavior for setpoint tracking control problems.

Establishing Conjecture 3.22, or a similar result, may also provide some clarity to the debate between nominal and stochastic MPC. Conjecture 3.22 is well motivated and supported by the examples in this dissertation. Larger-scale studies of SMPC also suggest that Conjecture 3.22 may hold (Kumar et al., 2019). Unfortunately, we do not see a clear path forward at this time. The results in Theorem 4.22 and Corollary 4.25 offer some insights into this conjecture, but nonetheless produce two different constants for these two formulations that are, to the best of our knowledge, not comparable for a general nonlinear system. With additional restrictions on the stage cost and system we may be able to establish a (weaker) version of Conjecture 3.22 that provides some insight. For example, this result may indicate what conditions for the system and stage cost are required for SMPC to produce significant gains compared to nominal MPC. At this time, however, simulation studies remain the only means to compare the performance of nominal and stochastic MPC for a problem of interest.

7.1.3 Distributional robustness

In Chapter 4, we established that SMPC is distributionally robust to sufficiently small errors in the stochastic model of uncertainty in the SMPC formulation. We can, in theory, further extend the SMPC problem formulation to include a description of this distributional uncertainty directly in the problem formulation via the emerging field of distributionally robust optimization (DRO). DRO can also be used to address the inherent limitations of scenario-based approximations of SMPC. While this proposed formulation is interesting, solving this DRO problem is even more computationally demanding than a stochastic optimization problem. Algorithmic advances in the field of DRO are continuing to improve the efficacy of this method, but the approach remains intractable for nonlinear systems and the online implementation required of MPC.

The definition of distributional robustness and the SISS framework proposed in Chapter 4 are general in that they apply to any closed-loop stochastic system. These results can therefore be applied to the larger field of stochastic optimal control. For example, this framework may be useful in investigating the distributional robustness of infinite horizon stochastic optimal control or stochastic dynamic programming.

In particular, the notion of distributional robustness defined in this dissertation is able to address scenario-based approximations of stochastic optimal control problems. This capability provides a potentially useful connection to data-driven methods for controller design. These data-driven or reinforcement learning methods effectively sample a stochastic dynamical system to determine a control law for that system. Characterizing the convergence of these data-driven methods with increased samples is therefore a very important topic of research. While the framework developed in Chapter 4 does not address this question directly, the results in this chapter may provide useful inspiration for future work on the topic of data-driven control and reinforcement learning.

7.1.4 Closed-loop scheduling

The scheduling problem in Chapter 6 includes only batch processes and produces a linear dynamical model for the nominal system. This state-space scheduling model can also be extended to include continuous processes, i.e., production occurs at each time step and we are scheduling transitions between different operating points in the unit (e.g., different product qualities in a reactor) (Risbeck et al., 2019). We may also want to consider *blending* problems similar to the example in Section 5.4.1. These blending problems frequently occur in chemical manufacturing, in which different feedstocks must be blended to produce the target combination of key properties for the products. Thus, a potential direction for future work is to extend the results in Chapter 6 to include these additional scheduling models.

The extension to include the scheduling of continuous processes should be relatively simple as the nominal dynamical model remains linear. Moreover, the terminal constraint can still require all of the lifted elements of the state to terminate in phase with the reference trajectory. Thus, we expect that the modifications to the terminal constraint and cost required to accommodate this new scheduling problem are minor.

The extension to include blending problems is likely more difficult since the dynamical system is now *nonlinear*. Nonetheless, the majority of results in Chapter 6 are not restricted to linear systems. Assumption 6.14 still applies because the inventory and backlog elements of the state are still integrators. Lemma 6.20 uses the fact that the nominal model is linear, but the desired bound in (6.24) is not unique to linear systems. Thus, a path to extend these results to blending problems is available.

Another important topic in production planning and scheduling is the so called "short-term" scheduling problem. In these short-term scheduling problems, there is no high-quality reference trajectory available for the system because there is no regular pattern of demand. Instead, we must react to new orders as they occur. The robustness results in Chapter 6 are nonetheless still relevant to this problem.

In effect, these new orders are disturbances entering the system. The goal of the scheduling algorithm is therefore to address and reject these disturbances as they occur. In this case, the nominal demand for each product is zero and the optimal nominal trajectory is an idle facility. We can therefore use this idle facility as the reference trajectory in the definition of robustness for these short-term scheduling problems. While an idle facility is not a "high-quality" reference trajectory for the system, the bound in Definition 6.9 is still meaningful. A short-term scheduling algorithm that admits this definition of robustness still ensures that small amounts of backlog are not ignored and that unnecessary inventory is not maintained for products we do not need. If there are only a few products for the facility, constructing a terminal cost and constraint based on the procedure in Chapter 6, may still be possible. If, however, there are dozens (or even hundreds) of products for the facility, constructing a terminal cost and constraint according to Chapter 6 may be unwise. Since these short-term scheduling problems are common in some manufacturing facilities, further research into closed-loop scheduling algorithms for this class of problems is worthwhile.

Bibliography

D. A. Allan, C. N. Bates, M. J. Risbeck, and J. B. Rawlings. On the inherent robustness of optimal and suboptimal nonlinear MPC. *Sys. Cont. Let.*, 106:68 – 78, 2017. ISSN 0167-6911. doi: 10.1016/j.sysconle.2017.03.005.

B. D. O. Anderson and J. B. Moore. Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM J. Cont. Opt.*, 19(1):20–32, 1981.

D. Angeli, R. Amrit, and J. B. Rawlings. On average performance and stability of economic model predictive control. *IEEE Trans. Auto. Cont.*, 57(7):1615–1626, 2012.

M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein generative adversarial networks. In *International conference on machine learning*, pages 214–223. PMLR, 2017.

V. Avadiappan and C. T. Maravelias. State estimation in online batch production scheduling: concepts, definitions, algorithms and optimization models. *Comput. Chem. Eng.*, 146: 107209, 2021.

J. Balasubramanian and I. Grossmann. Approximation to multistage stochastic optimization in multiperiod batch plant scheduling under demand uncertainty. *Ind. Eng. Chem. Res.*, 43 (14):3695–3713, 2004.

F. A. Bayer, M. Lorenzen, M. A. Müller, and F. Allgöwer. Robust economic model predictive control using stochastic information. *Automatica*, 74:151–161, 2016.

D. Bertsekas. *Dynamic programming and optimal control: Volume I*, volume 1. Athena scientific, 2017.

D. P. Bertsekas. *Dynamic Programming*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1987.

D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete Time Case*. Academic Press, New York, 1978.

A. Bonfill, M. Bagajewicz, A. Espuña, and L. Puigjaner. Risk management in the scheduling of batch plants under uncertain market demand. *Ind. Eng. Chem. Res.*, 43(3):741–750, 2004.

P. E. Caines and D. Q. Mayne. On the discrete time matrix Riccati equation of optimal control. *Int. J. Control*, 12(5):785–794, 1970.

G. C. Calafiore and M. C. Campi. The scenario approach to robust control design. *IEEE Trans. Auto. Cont.*, 51(5):742–753, 2006.

M. Cannon, B. Kouvaritakis, and D. Ng. Probabilistic tubes in linear stochastic model predictive control. *Sys. Cont. Let.*, 58(10-11):747–753, 2009a.

M. Cannon, B. Kouvaritakis, and X. Wu. Probabilistic constrained mpc for multiplicative and additive stochastic uncertainty. *IEEE Trans. Auto. Cont.*, 54(7):1626–1632, 2009b.

M. Cannon, B. Kouvaritakis, S. V. Rakovic, and Q. Cheng. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Trans. Auto. Cont.*, 56(1):194–200, 2010.

M. Cannon, J. Buerger, B. Kouvaritakis, and S. Rakovic. Robust tubes in nonlinear model predictive control. *IEEE Trans. Auto. Cont.*, 56(8):1942–1947, 2011.

D. Chatterjee and J. Lygeros. On stability and performance of stochastic predictive control techniques. *IEEE Trans. Auto. Cont.*, 60(2):509–514, 2014.

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

Y. Chu and F. You. Moving horizon approach of integrating scheduling and control for sequential batch processes. *AIChE J.*, 60(5):1654–1671, 2014.

B. Cott and S. Macchietto. Minimizing the effects of batch process variability using online schedule modification. *Comput. Chem. Eng.*, 13(1-2):105–113, 1989.

J. Cui and S. Engell. Medium-term planning of a multiproduct batch plant under evolving multi-period multi-uncertainty by means of a moving horizon strategy. *Comput. Chem. Eng.*, 34(5):598–619, 2010.

D. Ding, Z. Wang, B. Shen, and G. Wei. Event-triggered consensus control for discrete-time stochastic multi-agent systems: The input-to-state stability in probability. *Automatica*, 62: 284–291, 2015.

A. Elkamel and A. Mohindra. A rolling horizon heuristic for reactive scheduling of batch process operations. *Eng. Optim.*, 31(6):763–792, 1999.

P. M. Esfahani, T. Sutter, and J. Lygeros. Performance bounds for the scenario approach and an extension to a class of non-convex programs. *IEEE Trans. Auto. Cont.*, 60(1):46–58, 2014.

M. Farina, L. Giulioni, and R. Scattolini. Stochastic linear model predictive control with chance constraints–a review. *J. Proc. Cont.*, 44:53–67, 2016.

S. Ferrer-Nadal, C. A. Méndez, M. Graells, and L. Puigjaner. Optimal reactive scheduling of manufacturing plants with flexible batch recipes. *Ind. Eng. Chem. Res.*, 46(19):6273–6283, 2007.

P. Florchinger. Lyapunov-like techniques for stochastic stability. *SIAM J. Cont. Opt.*, 33(4): 1151–1169, 1995.

N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory and Relat. Fields*, 162(3):707–738, 2015.

C. Frogner, C. Zhang, H. Mobahi, M. Araya, and T. A. Poggio. Learning with a Wasserstein loss. *Adv. Neural Inf. Process. Syst.*, 28, 2015.

J. Goh and M. Sim. Distributionally robust optimization and its tractable approximations. *Oper. Res.*, 58(4-part-1):902–917, 2010.

P. J. Goulart and E. C. Kerrigan. Input-to-state stability of robust receding horizon control with an expected value cost. *Automatica*, 44(4):1171–1174, 2008.

P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski. Optimization over state feedback policies for robust control with constraints. *Automatica*, 42:523–533, 2006.

L. Grüne and M. Stieler. Asymptotic stability and transient optimality of economic MPC without terminal conditions. *J. Proc. Cont.*, 24(8):1187–1196, 2014.

G. Grimm, M. J. Messina, S. E. Tuna, and A. R. Teel. Examples when nonlinear model predictive control is nonrobust. *Automatica*, 40:1729–1738, 2004.

G. Grimm, M. J. Messina, S. E. Tuna, and A. R. Teel. Nominally robust model predictive control with state constraints. *IEEE Trans. Auto. Cont.*, 52(10):1856–1870, Oct 2007.

I. E. Grossmann. Advances in mathematical programming models for enterprise-wide optimization. *Comput. Chem. Eng.*, 47:2–18, 2012. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2012.06.038. URL http://www.sciencedirect.com/science/article/pii/S0098135412002220.

L. Grüne and C. M. Kellett. ISS-Lyapunov functions for discontinuous discrete-time systems. *IEEE Trans. Auto. Cont.*, 59(11):3098–3103, Nov 2014. ISSN 0018-9286. doi: 10.1109/TAC.2014.2321667.

D. Gupta and C. T. Maravelias. On deterministic online scheduling: Major considerations, paradoxes and remedies. *Comput. Chem. Eng.*, 94:312–330, 2016.

D. Gupta and C. T. Maravelias. Framework for studying online production scheduling under endogenous uncertainty. *Comput. Chem. Eng.*, 135:106670, 2020.

D. Gupta, C. T. Maravelias, and J. M. Wassick. From rescheduling to online scheduling. *Chem. Eng. Res. Des.*, 116:83–97, 2016.

I. Harjunkoski, C. T. Maravelias, P. Bongers, P. M. Castro, S. Engell, I. E. Grossmann, J. Hooker, C. Méndez, G. Sand, and J. Wassick. Scope for industrial applications of production scheduling models and solution methods. *Comput. Chem. Eng.*, 62:161 – 193, 2014. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2013.12.001.

L. Hewing, K. P. Wabersich, and M. N. Zeilinger. Recursively feasible stochastic model predictive control using indirect feedback. *Automatica*, 119:109095, 2020.

S. Honkomp, L. Mockus, and G. Reklaitis. A framework for schedule evaluation with processing uncertainty. *Comput. Chem. Eng.*, 23(4-5):595–609, 1999.

L. Huang and X. Mao. On input-to-state stability of stochastic retarded systems with Markovian switching. *IEEE Trans. Auto. Cont.*, 54(8):1898–1902, 2009.

A. Huercio, A. Espuna, and L. Puigjaner. Incorporating on-line scheduling strategies in integrated batch production control. *Comput. Chem. Eng.*, 19:609–614, 1995.

S. Janak, C. Floudas, J. Kallrath, and N. Vormbrock. Production scheduling of a large-scale industrial batch plant. II. reactive scheduling. *Ind. Eng. Chem. Res.*, 45(25):8253–8269, 2006.

J. L. Jerez, E. C. Kerrigan, and G. A. Constantinides. A condensed and sparse QP formulation for predictive control. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 5217–5222. IEEE, 2011.

Z.-P. Jiang and Y. Wang. A converse Lyapunov theorem for discrete-time systems with disturbances. *Sys. Cont. Let.*, 45:49–58, 2002.

K. B. Kanakamedala, G. V. Reklaitis, and V. Venkatasubramanian. Reactive schedule modification in multipurpose batch chemical plants. *Ind. Eng. Chem. Res.*, 33(1):77–90, 1994.

C. M. Kellett. A compendium of comparison function results. *Math. Contr. Sign. Syst.*, 26(3): 339–374, 2014. doi: 10.1007/s00498-014-0128-8.

C. M. Kellett and A. R. Teel. Smooth Lyapunov functions and robustness of stability for difference inclusions. *Sys. Cont. Let.*, 52:395–405, 2004.

E. C. Kerrigan and J. M. Maciejowski. Soft constraints and exact penalty functions in model predictive control. In *Control 2000 Conference, Cambridge*, pages 2319–2327, 2000.

H. K. Khalil. *Nonlinear Systems.* Prentice-Hall, Upper Saddle River, NJ, third edition, 2002.

J. Köhler, R. Soloperto, M. A. Müller, and F. Allgöwer. A computationally efficient robust model predictive control framework for uncertain nonlinear systems. *IEEE Trans. Auto. Cont.*, 66(2):794–801, 2020.

E. Kondili, C. C. Pantelides, and R. Sargent. A general algorithm for short term scheduling of batch operations–I. MILP formulation. *Comput. Chem. Eng.*, 17:211–227, 1993.

G. M. Kopanos and E. N. Pistikopoulos. Reactive scheduling by a multiparametric programming rolling horizon framework: A case of a network of combined heat and power units. *Ind. Eng. Chem. Res.*, 53(11):4366–4386, 2014.

G. M. Kopanos, E. Capón-García, A. Espuña, and L. Puigjaner. Costs for rescheduling actions: a critical issue for reducing the gap between scheduling theory and practice. *Ind. Eng. Chem. Res.*, 47(22):8785–8795, 2008.

B. Kouvaritakis, M. Cannon, S. V. Raković, and Q. Cheng. Explicit use of probabilistic distributions in linear predictive control. *Automatica*, 46(10):1719 – 1724, 2010. ISSN 0005-1098. doi: https://doi.org/10.1016/j.automatica.2010.06.034.

M. Krstic and H. Deng. *Stabilization of nonlinear uncertain systems.* Springer-Verlag, 1998.

R. Kumar, J. Jalving, M. J. Wenzel, M. J. Ellis, M. N. ElBsat, K. H. Drees, and V. M. Zavala. Benchmarking stochastic and deterministic MPC: A case study in stationary battery systems. *AIChE J.*, 65(7):e16551, 2019.

H. J. Kushner. On the stability of stochastic dynamical systems. *Proceedings of the National Academy of Sciences of the United States of America*, 53(1):8, 1965.

H. J. Kushner. *Stochastic Stability and Control*, volume 33 of *Mathematics in Science and Engineering*. Academic Press, New York, 1967.

N. H. Lappas and C. E. Gounaris. Multi-stage adjustable robust optimization for process scheduling under uncertainty. *AIChE J.*, 62(5):1646–1667, 2016.

N. H. Lappas, L. A. Ricardez-Sandoval, R. Fukasawa, and C. E. Gounaris. Adjustable robust optimization for multi-tasking scheduling with reprocessing due to imperfect tasks. *Optim. Eng.*, 20(4):1117–1159, 2019.

M. Lazar and W. P. M. H. Heemels. Predictive control of hybrid systems: Input-to-state stability results for sub-optimal solutions. *Automatica*, 45(1):180–185, 2009.

Z. Li and M. Ierapetritou. Process scheduling under uncertainty: Review and challenges. *Comput. Chem. Eng.*, 32(4):715–727, 2008a.

Z. Li and M. G. Ierapetritou. Reactive scheduling using parametric programming. *AIChE J.*, 54(10):2610–2623, 2008b. ISSN 1547-5905. doi: 10.1002/aic.11593.

Z. Li and M. G. Ierapetritou. Robust optimization for process scheduling under uncertainty. *Ind. Eng. Chem. Res.*, 47(12):4148–4157, 2008c.

D. Limon, T. Alamo, F. Salas, and E. F. Camacho. On the stability of MPC without terminal constraint. *IEEE Trans. Auto. Cont.*, 51(5):832–836, May 2006.

D. Limón Marruedo, T. Álamo, and E. F. Camacho. Stability analysis of systems with bounded additive uncertainties based on invariant sets: stability and feasibility of MPC. In *Proceedings of the American Control Conference*, pages 364–369, Anchorage, Alaska, May 2002.

X. Lin, S. L. Janak, and C. A. Floudas. A new robust optimization approach for scheduling under uncertainty: I. bounded uncertainty. *Comput. Chem. Eng.*, 28(6-7):1069–1085, 2004.

M. Lorenzen, F. Dabbene, R. Tempo, and F. Allgöwer. Constraint-tightening and stability in stochastic model predictive control. *IEEE Trans. Auto. Cont.*, 62(7):3165–3177, 2016.

M. Lorenzen, M. A. Müller, and F. Allgöwer. Stochastic model predictive control without terminal constraints. *Int. J. Robust and Nonlinear Control*, 29(15):4987–5001, 2019.

H. Mania, S. Tu, and B. Recht. Certainty equivalence is efficient for linear quadratic control. *Adv. Neural Inf. Process. Syst.*, 32, 2019.

C. Maravelias. General framework and modeling approach classification for chemical production scheduling. *AIChE J.*, 58(6):1812–1828, 2012.

D. Q. Mayne. Robust and stochastic model predictive control: Are we going in the right direction? *Annual Rev. Control*, 41:184 – 192, 2016. ISSN 1367-5788. doi: 10.1016/j.arcontrol.2016.04.006.

D. Q. Mayne and P. Falugi. Stabilizing conditions for model predictive control. *Int. J. Robust and Nonlinear Control*, 29(4):894–903, 2019.

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

D. Q. Mayne, M. M. Seron, and S. V. Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219–224, Feb 2005.

D. Q. Mayne, E. C. Kerrigan, E. J. van Wyk, and P. Falugi. Tube-based robust nonlinear model predictive control. *International Journal of Robust and Nonlinear Control*, 21(11):1341–1353, 2011. ISSN 1099-1239. doi: 10.1002/rnc.1758. URL `http://dx.doi.org/10.1002/rnc.1758`.

R. D. McAllister and J. B. Rawlings. Stochastic Lyapunov functions and asymptotic stability in probability. Technical Report 2020–02, TWCCC Technical Report, August 2020. URL `https://sites.engineering.ucsb.edu/~jbraw/jbrweb-archives/tech-reports/twccc-2020-02.pdf`.

R. D. McAllister and J. B. Rawlings. Robustness of model predictive control to (large) discrete disturbances. *IFAC-PapersOnLine*, 54(6):64–69, 2021a.

R. D. McAllister and J. B. Rawlings. Stochastic model predictive control: Existence and measurability. Technical Report 2021–01, TWCCC Technical Report, March 2021b. URL https://sites.engineering.ucsb.edu/~jbraw/jbrweb-archives/tech-reports/twccc-2021-01.pdf.

R. D. Mcallister and J. B. Rawlings. Inherent stochastic robustness of model predictive control to large and infrequent disturbances. *IEEE Trans. Auto. Cont.*, 2021. doi: https://10.1109/TAC.2021.3122365.

R. D. McAllister and J. B. Rawlings. Stochastic exponential stability of nonlinear stochastic model predictive control. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 880–885. IEEE, 2021.

R. D. McAllister and J. B. Rawlings. Nonlinear stochastic model predictive control: Existence, measurability, and stochastic asymptotic stability. *IEEE Trans. Auto. Cont.*, 2022a. doi: https://doi.org/10.1109/TAC.2022.3157131.

R. D. McAllister and J. B. Rawlings. Advances in mixed-integer model predictive control. In *American Control Conference*, pages 364–369, Atlanta, GA, June 8–10, 2022b.

R. D. McAllister and J. B. Rawlings. A suboptimal economic model predictive control algorithm for large and infrequent disturbances. IEEE Trans. Auto. Cont., 2022c. Under review.

R. D. McAllister and J. B. Rawlings. The stochastic robustness of nominal and stochastic model predictive control. IEEE Trans. Auto. Cont., 2022d. Under review.

R. D. McAllister and J. B. Rawlings. The inherent distributional robusntess of stochastic and nominal model predictive control. IEEE Trans. Auto. Cont., 2022e. Under review.

R. D. McAllister, J. B. Rawlings, and C. T. Maravelias. Rescheduling penalties for economic model predictive control and closed-loop scheduling. *Ind. Eng. Chem. Res.*, 59(6):2214–2228, 2020.

R. D. McAllister, J. B. Rawlings, and C. T. Maravelias. The inherent robustness of closed-loop scheduling. *Comput. Chem. Eng.*, 159:107678, 2022. doi: https://doi.org/10.1016/j.compchemeng.2022.107678.

C. A. Mendez and J. Cerdá. An MILP framework for batch reactive scheduling with limited discrete resources. *Comput. Chem. Eng.*, 28(6-7):1059–1068, 2004.

A. Mesbah. Stochastic model predictive control. *IEEE Ctl. Sys. Mag.*, pages 30–44, Dec 2016.

J. Novas and G. Henning. Reactive scheduling framework based on domain knowledge and constraint programming. *Comput. Chem. Eng.*, 34(12):2129–2148, 2010.

G. Pannocchia, J. B. Rawlings, and S. J. Wright. Conditions under which suboptimal nonlinear MPC is inherently robust. *Sys. Cont. Let.*, 60:747–755, 2011.

C. Pantelides. Unified frameworks for optimal process planning and scheduling. In *Proceedings on the second conference on foundations of computer aided operations*, pages 253–274. Cache Publications New York, 1994.

B. Picasso, D. Desiderio, and R. Scattolini. Robust stability analysis of nonlinear discrete-time systems with application to MPC. *IEEE Trans. Auto. Cont.*, 57(1):185–191, Jan 2012. ISSN 0018-9286. doi: 10.1109/TAC.2011.2163363.

L. Praly and Y. Wang. Stabilization in spite of matched unmodeled dynamics and an equivalent definition of input-to-state stability. *Mathematics of Control, Signals and Systems*, 9(1):1–33, 1996.

J. A. Primbs and C. H. Sung. Stochastic receding horizon control of constrained linear systems with state and control multiplicative noise. *IEEE Trans. Auto. Cont.*, 54(2):221–230, 2009.

S. J. Qin and T. A. Badgwell. A survey of industrial model predictive control technology. *Control Eng. Pract.*, 11(7):733–764, 2003.

C. V. Rao and J. B. Rawlings. Steady states and constraints in model predictive control. *AIChE J.*, 45(6):1266–1278, 1999.

J. B. Rawlings and M. J. Risbeck. On the equivalence between statements with epsilon-delta and K-functions. Technical Report 2015–01, TWCCC Technical Report, December 2015. URL `http://jbrwww.che.wisc.edu/tech-reports/twccc-2015-01.pdf`.

J. B. Rawlings and M. J. Risbeck. Model predictive control with discrete actuators: Theory and application. *Automatica*, 78:258–265, 2017.

J. B. Rawlings, D. Q. Mayne, and M. M. Diehl. *Model Predictive Control: Theory, Design, and Computation.* Nob Hill Publishing, Santa Barbara, CA, 2nd, paperback edition, 2020. 770 pages, ISBN 978-0-9759377-5-4.

M. J. Risbeck. *Mixed-Integer Model Predictive Control with Applications to Building Energy Systems.* PhD thesis, University of Wisconsin–Madison, July 2018.

M. J. Risbeck and J. B. Rawlings. Economic MPC for time-varying cost and peak demand charge optimization. *IEEE Trans. Auto. Cont.*, 65(7):2957–2968, Jul 2019. doi: 10.1109/TAC.2019.2939633.

M. J. Risbeck, C. T. Maravelias, and J. B. Rawlings. Unification of closed-loop scheduling and control: State-space formulations, terminal constraints, and nominal theoretical properties. *Comput. Chem. Eng.*, 129:106496, 2019.

R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis.* Springer-Verlag, 1998.

M. Rodrigues, L. Gimeno, C. Passos, and M. Campos. Reactive scheduling approach for multipurpose chemical batch plants. *Comput. Chem. Eng.*, 20:S1215–S1220, 1996.

J. A. Rossiter, B. Kouvaritakis, and M. J. Rice. A numerically robust state-space approach to stable-predictive control strategies. *Automatica*, 34(1):65–73, 1998.

W. Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, New York, third edition, 1976.

G. Sand and S. Engell. Modeling and solving real-time scheduling problems by stochastic integer programming. *Comput. Chem. Eng.*, 28(6):1087–1103, 2004.

T. L. Santos, A. D. Bonzanini, T. A. N. Heirung, and A. Mesbah. A constraint-tightening approach to nonlinear model predictive control with chance constraints for stochastic systems. In *2019 American Control Conference (ACC)*, pages 1641–1647. IEEE, 2019.

P. O. M. Scokaert and J. B. Rawlings. Feasibility issues in linear model predictive control. *AIChE J.*, 45(8):1649–1659, Aug 1999.

P. O. M. Scokaert, D. Q. Mayne, and J. B. Rawlings. Suboptimal model predictive control (feasibility implies stability). *IEEE Trans. Auto. Cont.*, 44(3):648–654, Mar 1999.

M. A. Sehr and R. R. Bitmead. Stochastic output-feedback model predictive control. *Automatica*, 94:315–323, 2018.

S. Shafieezadeh Abadeh, V. A. Nguyen, D. Kuhn, and P. M. Mohajerin Esfahani. Wasserstein distributionally robust Kalman filtering. *Adv. Neural Inf. Process. Syst.*, 31, 2018.

H. Shi and F. You. A computational framework and solution algorithms for two-stage adaptive robust scheduling of batch manufacturing processes under uncertainty. *AIChE J.*, 62(3): 687–703, 2016.

H. A. Simon. Dynamic programming under uncertainty with a quadratic criterion function. *Econometrica*, pages 74–81, 1956.

R. M. Solovay. A model of set-theory in which every set of reals is Lebesgue measurable. *Ann. Math.*, pages 1–56, 1970.

E. D. Sontag and Y. Wang. On the characterization of the input to state stability property. *Sys. Cont. Let.*, 24:351–359, 1995.

K. Subramanian, C. T. Maravelias, and J. B. Rawlings. A state-space model for chemical production scheduling. *Comput. Chem. Eng.*, 47:97–110, Dec 2012.

C. Tang and T. Basar. Stochastic stability of singularly perturbed nonlinear systems. In *Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No. 01CH37228)*, volume 1, pages 399–404. IEEE, 2001.

A. R. Teel. A Matrosov theorem for adversarial Markov decision processes. *IEEE Trans. Auto. Cont.*, 58(8):2142–2148, 2013.

A. R. Teel, J. P. Hespanha, and A. Subbaraman. Equivalent characterizations of input-to-state stability for stochastic discrete-time systems. *IEEE Trans. Auto. Cont.*, 59(2):516–522, 2013.

A. R. Teel, J. P. Hespanha, and A. Subbaraman. A converse Lyapunov theorem and robustness for asymptotic stability in probability. *IEEE Trans. Auto. Cont.*, 59(9):2426–2441, 2014.

H. Theil. A note on certainty equivalence in dynamic planning. *Econometrica*, pages 346–349, 1957.

J. Tsinias. Stochastic input-to-state stability and applications to global feedback stabilization. *Int. J. Control*, 71(5):907–930, 1998.

H. Van de Water and J. Willems. The certainty equivalence property in stochastic control theory. *IEEE Trans. Auto. Cont.*, 26(5):1080–1087, 1981.

V. S. Varadarajan. On the convergence of sample probability distributions. *Sankhyā*, 19(1/2): 23–26, 1958.

C. Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

J. P. Vin and M. G. Ierapetritou. A new approach for efficient rescheduling of multiproduct batch plants. *Ind. Eng. Chem. Res.*, 39(11):4228–4238, 2000.

J. P. Vin and M. G. Ierapetritou. Robust short-term scheduling of multiproduct batch plants under demand uncertainty. *Ind. Eng. Chem. Res.*, 40(21):4543–4554, 2001.

W. M. Wonham. Optimal stochastic control. *Automatica*, 5:113–118, 1969.

X. Wu, Y. Tang, and W. Zhang. Input-to-state stability of impulsive stochastic delayed systems under linear assumptions. *Automatica*, 66:195–204, 2016.

I. Yang. Wasserstein distributionally robust stochastic control: A data-driven approach. *IEEE Trans. Auto. Cont.*, 66(8):3863–3870, 2020.

S. Yu, M. Reble, H. Chen, and F. Allgöwer. Inherent robustness properties of quasi-infinite horizon nonlinear model predictive control. *Automatica*, 50(9):2269 – 2280, 2014. ISSN 0005-1098. doi: 10.1016/j.automatica.2014.07.014.

M. Zanon and T. Faulwasser. Economic MPC without terminal constraints: Gradient-correcting end penalties enforce asymptotic stability. *J. Proc. Cont.*, 63:1–14, 2018.

P. Zhao, W. Feng, and Y. Kang. Stochastic input-to-state stability of switched stochastic non-linear systems. *Automatica*, 48(10):2569–2576, 2012.

A. Zheng and M. Morari. Control of linear unstable systems with constraints. In *American Control Conference, Seattle, Washington*, pages 3704–3708, 1995.