Czech Technical University in Prague
Faculty of Electrical Engineering

Master thesis

Stochastic Model Predictive Control

by

Milan Korda

May 2011

Prohlášení

Prohlašuji, že jsem svou diplomovou práci vypracoval samostatně a použil jsem pouze podklady (literaturu, projekty, SW atd.) uvedené v přiloženém seznamu.

V Curychu dne 24. 4. 2011

podpis

Acknowledgment

Foremost, I would like to thank Ing. Jiří Cigler for his excellent leadership and for providing me with the opportunity to complete this thesis at ETH Zurich under the supervision of prof. John Lygeros, whom I am grateful for guidance and the possibility to work in his group. The thesis would not have been the same without the insightful discussions with Dr. Debasish Chatterjee, Dr. Ravi Gondhalekar, Dr. Peter Hokayem, Frauke Oldewurtel and Dr. Federico Ramponi. Finally, my deepest gratitude goes to my family and friends for their unwavering support and encouragement.

Abstract

The setting of this thesis is stochastic optimal control and constrained model predictive control of discrete-time linear systems with additive noise. There are three primary topics.

First, it is the finite horizon stochastic optimal control problem with the expectation of the p-norm as the objective function and jointly Gaussian, although not necessarily independent, disturbances. We develop an approximation strategy that solves the problem in a certain class of nonlinear feedback policies for perfect as well as imperfect state information, while ensuring satisfaction of hard input constraints. A bound on suboptimality of the proposed strategy in the class of aforementioned nonlinear feedback policies is given.

Second, it is the question of mean-square stabilizability of stochastic linear systems with bounded control authority. We provide simplified proofs of existing results on stabilizability of strictly and marginally stable systems, and extend the employed technique to show stability of positive (or negative) parts of the state of marginally unstable systems provided that the control authority is nonzero, but possibly arbitrarily small. We also prove the existence of a mean-square stabilizing Markov policy for marginally stable systems.

Finally, we develop a systematic approach to ensure strong feasibility of stochastic model predictive control problems under affine as well as nonlinear disturbance feedback policies. Two distinct approaches are presented, both of which capitalize on and extend the machinery of (controlled) invariant sets to a stochastic environment. The first approach employs an invariant set as a terminal constraint, whereas the second one constrains the first predicted state. Consequently, the second approach turns out to be completely independent of the policy in question and moreover it produces the largest feasible set amongst all admissible policies. As a result a trade-off between computational complexity and performance can be found without compromising feasibility properties.

Abstrakt

Tato práce se zabývá třemi tématy z oblasti stochastického optimálního a prediktivního řízení linárních systémů s diskrétním časem a aditivním šumem procesu.

Prvně je to problém stochastického optimálního řízení na konečném horizontu s p-normou jako kriteriální funkcí a normálně rozděleným, avšak ne nutně nezávislým, šumem procesu. Je navržena aproximační strategie, která řeší problém v určité třídě nelineárních zpětnovazebních funkcí pro případy úplné i neúplné stavové informace a přitom zajistí dodržení omezení na akční zásah. Horní odhad míry suboptimality navržené strategie v dané třídě nelineárních funkcí je odvozen.

Dále se text zabývá otázkou stabilizovatelnosti linárních stochastických systému ve smyslu omezeného druhého momentu stavu. Uvádíme zjednodušené důkazy některých existujících výsledků pro striktně a marginálně stabilní systémy s tím, že uvedená technika je posléze rozšířena k důkazu stabilizovatelnosti pozitivních (nebo negativních) částí komponent stavu marginálně nestabilních systémů za předpokladu nenulového, ale případně libovolně malého omezení na akční zásah. Dokázána je také existence markovské stabilizující strategie pro marginálně stabilní systémy.

Nakonec je navržen systematický postup pro zajištění rekurzivní proveditelnosti stochastického prediktivního řízení při použítí jak afinních tak nelineárních zpětnovazebních strategií. Představeny jsou dva různé přístupy, z nichž oba staví a rozšiřují nástroje (řízených) invariantních množin do stochastického prostředí. První přístup používá invariatní množinu jako koncové omezení, zatímco druhý omezuje první predikovaný stav. Důsledkem je, že druhý přístup je zcela nezávislý na uvažované zpětnovazební strategii a vede k největší množině řešitelnosti mezi všemi přípustnými strategiemi, což umožnuje najít kompromis mezi výpočetní náročností a kvalitou řízení bez ohledu na velikost množiny řešitelnosti.

České vysoké učení technické v Praze
Fakulta elektrotechnická

Katedra řídicí techniky

ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: **Bc. Milan Korda**

Studijní program: Kybernetika a robotika
Obor: Systémy a řízení

Název tématu: **Stochastické prediktivní řízení**

Pokyny pro vypracování:

1. Seznamte se s stochastickým řízením, analyzujte hlavní výhody a nevýhody používaných metod a jejich potenciálního použití pro stochastické prediktivní řízení (SMPC).
2. Analyzujte problém minimalizace jiných než kvadratických kritérií a pro vybrané realizovatelné případy navrhněte algoritmické řešení.
3. Analyzujte invariantní množiny používané ve stochastickém prediktivním řízení.
4. Implementuje navržené algoritmy a jejich výsledky porovnejte.

Seznam odborné literatury:

Dodá vedoucí práce

Vedoucí: Ing. Jiří Cigler

Platnost zadání: do konce letního semestru 2011/2012

prof. Ing. Michael Šebek, DrSc.
vedoucí katedry

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta elektrotechnická
L.S.
-4-

nz. M. Demlová

prof. Ing. Boris Šimák, CSc.
děkan

V Praze dne 3. 11. 2010

Contents

1. Introduction

Over the last two decades, constrained model predictive control (MPC) has matured substantially. There is now a solid and very general theoretical foundation for stability and feasibility of nominal as well as robust MPC problems [37]. Nevertheless, the connection to another mature field, stochastic optimal control, is still not fully developed although there has been a considerable research effort in this direction over the last years.

A basis for any receding horizon policy is finite horizon cost minimization, which is the first direction of recent research. This problem lies at the heart of stochastic optimal control theory and is known to be extremely difficult with only a handful of problems (e.g. the linear quadratic control) that can be solved optimally. The remainder has to be tackled by various approximation techniques most frequently, but not exclusively, arising from the dynamic programming paradigm [5, 53].

Recent advances in computation and mathematical optimization techniques have, however, opened new ways of dealing with these problems. One of the simplest, yet in most practical applications very effective approach, is the certainty equivalent model predictive control (CE-MPC) [4, 5] that solves a deterministic optimization problem with stochastic disturbances replaced by their estimates based upon the information available at the time, and proceeds in a receding horizon fashion (see Section 2.4). Another popular class of control strategies is the affine disturbance feedback policy which turns out to be equivalent to the affine state-sequence feedback policy via a nonlinear transformation similar to the classical Q-design or Youla-Kučera parametrization. See Section 2.5.2 and also [52, 51].

However convenient the paradigm of affine disturbance feedback may be, its use is prohibitive whenever unbounded stochastic disturbances enter the system in the presence of hard control input bounds since then the linear part necessarily vanishes, which, in effect, renders the policy open loop. One way to overcome this problem is to use a (saturated) nonlinear disturbance feedback as in [22, 29, 51], where this approach was developed for the quadratic cost. The upside of this generalization is the fact that the convexity of the cost function associated with the nonlinear disturbance feedback turns out to be independent of the choice of the nonlinear function (see Section 2.5). This is exploited in Section 3 of this work, which is devoted to developing a tractable extension of this methodology to a general p-norm cost with the additional assumption of the disturbances being jointly Gaussian (but not necessarily independent). Our methodology brings about a significant performance improvement compared to the traditional certainty-equivalent approach while retaining reasonable computational demands compared to sampling or dynamic programming techniques.

Another branch of approximation techniques bounds the disturbances a priori and solves a robust MPC problem, while guaranteeing an open loop probabilistic bound on the performance [7]. This approach, however, tends to be very conservative, and thus the idea of bounding the disturbances a priori based on their distribution appears more often in the context of chance constraints, see e.g. [42]. For different approaches to chance constraints handling see Section 5.1 and also [8, 23, 39].

It is the issue of recursive feasibility of (probabilistic) constraints that has predominantly hampered bridging the gap between stochastic optimal control and constrained model predictive control. The crux of the matter lies in the fact that independent unbounded disturbances additively entering the system cannot give rise to a recursively feasible problem as long as the set of state constraints is compact and control authority bounded. Thus, one has to either develop a backup recovery policy that is triggered when infeasibility occurs or assume compactly supported disturbances. The former was tackled for instance in [21] where an optimal solution (in some sense) was developed using dynamic programming techniques, carrying over the inherent computational burden of dynamic programming techniques, however. The latter was extensively studied in a series of papers [17, 18, 19, 36, 46], where the authors consider various types of constraints and process noise properties, and present multiple techniques to tackle these problems. The common factor of these papers is the use of a perturbed linear state feedback (or pre-stabilization), which necessarily limits the number of degrees of freedom and as a consequence the resulting performance.

In this work, in contrast, the use of affine disturbance feedback, where more degrees of freedom are available, brings about performance improvement but also increased computational effort. This can, however, be overcome by imposing structural constraints on the feedback matrix, allowing to reach a trade-off between performance and computational burden [41]. Furthermore the feasibility of the second of the two approaches presented here is independent of the policy in question and in fact provides the largest feasible set amongst all admissible policies. Our approach takes advantage of the notion of controlled invariance, well established in (robust) constrained model predictive control (see, e.g., [9, 24]), bringing stochastic model predictive control on a sound footing. In fact, we derive results on strong feasibility and least-restrictiveness (see Definitions 5.1 and 5.2) analogous to those of [26, 27, 41] in a stochastic context.

There is a wide range of applications amenable to the methods developed in this thesis. A rich source of such applications is, for instance, building climate control, a typical stochastic environment where p-norm or similar cost functions are ample and constraints have typically probabilistic nature (e.g. the temperature must stay within a certain range with given probability) [42, 43]. Another source of application can be found in Financial Economics, an archetypal example being a portfolio optimization, where, in the simplest setting, one maximizes return under risk constraints of stochastic nature, or minimizes risk under constraints on return (see, e.g., [12, 13, 14, 15]).

Organization

The text starts with a short introduction to stochastic optimal control in Chapter 2 where we set up fundamental problems encountered in this area and outline some existing strategies to deal with them.

The main body of the thesis consists of Chapters 3, 4, 5, each of which is only loosely connected to the others and requires only the basic notions established in Chapter 2. Chapter 3 develops a tractable approximation to the p-norm stochastic optimal control problem of linear systems for perfect (Section 3.2) as well as imperfect (Section 3.2.2) state information. Chapter 4 deals with mean-square stability of linear stochastic systems with bounded control authority. First, some simple connections to another notion of stability are established, and then most of the currently known results on mean-square stabilizability are presented. Chapter 5 deals with recursive feasibility of probabilistic constraints in the presence of bounded disturbances. First we give a brief overview of different types of probabilistic constraints typically encountered in Section 5.1. Main results of this chapter are then collected in Section 5.2, where two distinct approaches are developed, both of which are based on the concept of invariant sets.

Numerical aspects of the p-norm stochastic optimal control approximation are briefly discussed in Section 6.1, whereas numerical examples for all three main chapters are postponed to Section 6.2.

Contribution

The main contribution of the thesis is twofold. First it is the development of a tractable approximation of the p-norm stochastic optimal control problem with nonlinear disturbance feedback for perfect as well as imperfect state information. Second it is the development of a novel approach to enforce strong feasibility of stochastic model predictive control problems based on invariant sets.

Most of the results on mean-square stabilizability have already been established although some of the proof techniques are different and the results on mean-square stabilizability of positive parts of marginally unstable systems and the existence of a stabilizing Markov policy for marginally stable systems are, to the best of our knowledge, new.

2. Stochastic optimal control

2.1. Problem setup

A general stochastic optimal control problem is introduced in this section. This formulation provides a very general framework for a wide variety of problems not only from control engineering, but also from basically any field imaginable since uncertainty (or randomness) in the future is always present.

The goal of a general discrete-time stochastic optimal control is to minimize the cost function

$$J := \mathbf{E}\left\{ l(x_N) + \sum_{k=0}^{N-1} l(x_k, u_k, w_k) \right\}, \tag{2.1}$$

subject to system dynamics

$$x_{k+1} = f_k(x_k, u_k, w_k),$$

where x_k is the state, u_k the control input and w_k is a random vector, over which the expectation is taken, representing a disturbance acting on the system. The initial state x_0 can be random as well although without loss of generality the expectation in (2.1) can be conditioned on x_0 and minimized point-by-point for each value of x_0. The minimization is to be carried out over all Borel measurable causal feedback policies of the form

$$u_k = \phi_k(x_0, x_1, \dots, x_k). \tag{2.2}$$

If the random variables w_k, $k = 0, \dots, N-1$ are independent, an optimal policy (if one exists) turns out to be a state feedback of the form

$$u_k = \tilde{\phi}_k(x_k)$$

for some Borel measurable $\tilde{\phi}$ (see [28] for details).

Note that this is only one of many possible formulations of a stochastic optimal control problem. For instance, a possible generalization is to allow $\phi_k(\cdot)$ to be random given the entire history $(x_0, u_0, \dots, x_{k-1}, u_{k-1}, x_k)$. It turns out, however, that in most practical cases a 'deterministic' policy (2.2) will be optimal in this broader class of randomized policies [28].

The functional nature of stochastic optimal control makes it much more difficult, in fact, except for several cases, intractable, then deterministic optimization in finite dimensional space. Therefore, approximate techniques are usually employed to solve this problem suboptimally.

If full state information were not available, then, at a given time instant, a causal control policy would use all the information available up to this time. Thus the control inputs would be given by

$$u_k = \phi_k(y_0, \dots, y_k), \tag{2.3}$$

where y_k, $k = 0, \dots, N-1$ is the measured output sequence, i.e., a sequence of the form

$$y_k = h_k(x_k, v_k),$$

where v_k, $k = 0, \dots, N-1$ is a random vector representing measurement noise.

Analytically more demanding are infinite horizon problems. Two types of costs are most commonly considered: (i) the discounted infinite horizon cost

$$J := \mathbf{E}\left\{\sum_{k=0}^{\infty} \alpha^k l(x_k, u_k, w_k)\right\}, \tag{2.4}$$

where $0 < \alpha < 1$, and (ii) the long-run average cost

$$J := \limsup_{N\to\infty} \frac{1}{N}\mathbf{E}\left\{\sum_{k=0}^{N-1} l(x_k, u_k, w_k)\right\}, \tag{2.5}$$

where the random sequence $w_0, w_1, \dots$ is typically assumed i.i.d. Both formulations help to improve convergence properties as well as have some practical grounds, e.g., a natural role of the discount factor α in problems arising in economics.

2.2. Dynamic programming

A basic tool for analysis and in some cases for actual solving of stochastic optimization problems is dynamic programming, which is based on the Bellman recursion [5] defined by

$$\begin{aligned} V_N(\mathcal{X}_N) &:= l(x_N), \\ V_k(\mathcal{X}_k) &:= \inf_{u_k}\{l(x_k, u_k) + \mathbf{E}\{V_{k+1}(\mathcal{X}_k, f_k(x_k, u_k, w_k))|\mathcal{X}_k\}\}, \end{aligned} \tag{2.6}$$

where $\mathcal{X}_k = (x_0, \dots, x_k)$ is the state sequence up to time k. It can be shown[1] that if the infimum in (2.6) is attained for all $\mathcal{X}_k$, the optimal value of the problem (2.1) is equal to $J^* = \mathbf{E}V_0(\mathcal{X}_0)$, and an optimal policy is given by

$$\phi_k^*(\mathcal{X}_k) = \arg\min_{u_k}\{l(x_k, u_k) + \mathbf{E}\{V_{k+1}(\mathcal{X}_k, f_k(x_k, u_k, w_k))|\mathcal{X}_k\}\}$$

[1] See [5] for derivation under the assumption of *discrete* random variables w_k being conditionally independent given current state and control, and [6] or [28] for a general treatment.

provided that these functions are Borel-measurable for $k = 0, \dots, N-1$.

The imperfect state information problem can be treated similarly. If the information vector at time k is defined as $\mathcal{I}_k = (y_0, u_0, \dots, y_{k-1}, u_{k-1}, y_k)$, the Bellman recursion becomes

$$\begin{aligned} V_N(\mathcal{I}_N) &:= \mathbf{E}\{l(x_N)|\mathcal{I}_N\}, \\ V_k(\mathcal{I}_k) &:= \min_{u_k}\{\mathbf{E}\{l(x_k,u_k) + V_{k+1}(\mathcal{I}_k, u_k, h_{k+1}(f_k(x_k,u_k,w_k)))|\mathcal{I}_k\}\}, \end{aligned} \tag{2.7}$$

and the optimal control policies

$$\phi_k^*(\mathcal{I}_k) = \arg\min_{u_k}\{\mathbf{E}\{l(x_k,u_k) + V_{k+1}(\mathcal{I}_k, u_k, h_{k+1}(f_k(x_k,u_k,w_k)))|\mathcal{I}_k\}\},$$

which is indeed a function of $y_0, \dots, y_k$ as (2.3) requires since the control inputs $u_0, \dots, u_{k-1}$ contained in $\mathcal{I}_k$ are in turn functions of $y_0, \dots, y_{k-1}$.

Unfortunately, dynamic programming is in its raw form impracticable for most problem instances due to the so-called curse of dimensionality [5], which is even more profound in the stochastic setting, where at time k the expectation of $V_{k+1}(f_k(x_k, u_k, w_k))$ needs to be evaluated for each admissible control input and for all admissible states x_k. There are several problem instances where dynamic programming is actually tractable; the probably best known linear quadratic case is briefly mentioned in Section 2.4.

2.3. Approximate dynamic programming

Despite the lack of applicability described above, dynamic programming is a cornerstone for a whole field of suboptimal control strategies called approximate dynamic programming (ADP). One of the basic approximation techniques is the one-step lookahead control policy defined by

$$\phi_k(\mathcal{X}_k) = \arg\min_{u_k}\{l(x_k,u_k) + \mathbf{E}\{\tilde{V}_{k+1}(\mathcal{X}_k, f_k(x_k,u_k,w_k))|\mathcal{X}_k\}\}, \tag{2.8}$$

where $\tilde{V}_k(\mathcal{X}_k)$, $k = 0, \dots, N-1$, are some approximations of the optimal cost-to-go functions $V_k(\mathcal{X}_k)$. It is not hard to show that if $\tilde{V}_k$ are true cost-to-go functions of any (suboptimal) policy then the resulting one-step lookahead policy will perform no worse [5]. The one-step-lookahead policy with $\tilde{V}_k$ given by a suboptimal policy is sometimes called the rollout algorithm.

Note that the applicability of a particular one-step lookahead policy relies heavily on the ability to fast evaluate the functions $\tilde{V}_k$. This may be prohibitive for the rollout algorithm even if $\tilde{V}_k$ is based on a simple suboptimal policy, such as a certainty equivalent policy, for which there is no analytical expression and evaluation is not fast enough for purposes of minimization in (2.8) where gridding techniques are usually necessary. On the other hand, if $\tilde{V}_k$ has a favourable analytical form, the minimization in (2.8) can sometimes be carried out by means of mathematical programming.

2.3.1. An application of Bellman inequality

A particularly nice application of approximate dynamic programming ideas was given in [60]. The authors were concerned with the infinite horizon discounted problem (2.4), for which they used properties of the Bellman operator[2]

$$(Tg)(x) = \inf_u \{l(x,u) + \alpha \mathbf{E}[g(f(x,u,w_0))]\} \tag{2.9}$$

to construct global underestimators of the optimal cost-to-go function $V(x)$. Two important properties of the Bellman operator, which hold under reasonable conditions [28], are monotonicity, i.e.,

$$g_1 \leq g_2 \Longrightarrow Tg_1 \leq Tg_2,$$

and pointwise convergence to the optimal cost-to-go from any starting function, i.e.,

$$\lim_{k\to\infty} T^k g = V$$

for any[3] function g.

Now it immediately follows from these two properties that if

$$\hat{V} \leq T\hat{V} \tag{2.10}$$

or

$$\hat{V} \leq T^k\hat{V} \tag{2.11}$$

for some $k > 0$ then

$$\hat{V} \leq \lim_{k\to\infty} T^k\hat{V} = V.$$

Hence the Bellman inequality (2.10) and, more generally, the iterated Bellman inequality (2.11) give sufficient conditions for a function $\hat{V}$ to be a global underestimator of V. Moreover, if we restrict ourselves to a finite dimensional subspace spanned by some basis functions

$$V_1, \dots, V_N, \tag{2.12}$$

the Bellman inequality becomes convex in the coefficients of the linear combination $\alpha_1, \dots, \alpha_N$. Indeed, the left-hand side of the Bellman inequality is linear in α_i and the right-hand side is concave since it is the infimum over a family of functions affine in α_i. The iterated Bellman inequality (2.11) is not convex in α_i by itself, but can be approximated by introducing $k-1$ auxiliary functions, which results in an approximation by k Bellman inequalities (see [60] for details).

[2] The random sequence $w_0, w_1, \dots$ is assumed i.i.d., so there could be any other random variable w_k instead of w_0 in the definition of the operator.

[3] From now on we drop the measurability assumptions.

Now it remains to optimize the underestimator in some sense. A viable approach is to maximize the expectation $\mathbf{E}\hat{V}(x)$ for a suitably chosen distribution of x. For instance, if only a global underestimator of V for given distribution of the initial state x_0 were of interest, we would like to solve the problem

$$\begin{array}{ll} \underset{\alpha_1,\ldots,\alpha_N}{\text{maximize}} & \mathbf{E}\hat{V}(x_0) \\ \text{subject to} & \hat{V} \le T\hat{V} \text{ or } \hat{V} \le T^k\hat{V}, \\ & \hat{V} = \alpha_1 V_1 + \ldots + \alpha_N V_N. \end{array}$$

On the other hand if the underestimator were to be used as a cost-to-go approximation for the one-step lookahead policy (2.8), an estimate of the stationary distribution of x might be preferable. In this case it is also beneficial to compute the underestimators for several different distributions of x and take their maximum, which is also an underestimator.

All of this may seem of little practical value since the Bellman inequality is typically difficult to evaluate. However, in the case of convex quadratic cost and basis functions (2.12) and affine control constraints, it is shown in [60] that the (iterated) Bellman inequality can be approximated via the S-procedure to obtain a semidefinite program (SDP) [10]. The program is solved offline only to get a quadratic underestimator, which then results in a receding horizon problem on the horizon of length one if used in the one-step lookahead policy. Maximum of quadratic underestimators yields equally simple receding horizon policy, only a quadratic program (QP) becomes a quadratically-constrained quadratic program (QCQP).

2.4. Certainty equivalent control

One of the simplest suboptimal policy for solving (2.1) is the certainty equivalent control where the random disturbances $w_0, \ldots, w_{N-1}$ are replaced by some fixed values $\hat{w}_0, \ldots, \hat{w}_{N-1}$, which makes the problem finite-dimensional, deterministic. Typical choices for these values are conditional expectations, most likely values or just zeros. The solution of this deterministic optimization problem $\hat{u}_0, \ldots, \hat{u}_{N-1}$ then defines a constant (open-loop) control policy.

A natural way to introduce feedback is to use the certainty equivalent control in the so-called shrinking horizon mode, where new estimates based on all information available are formed at each time $k = 0, \ldots, N-1$, the deterministic problem is resolved with these estimates on the horizon $k, \ldots, N$, and only the first control input is applied. This approach is closely related to the certainty equivalent model predictive control (CE-MPC), the only difference being that with CE-MPC the horizon recedes instead of shrinking, i.e., at each time $k \ge 0$ a deterministic optimization problem is solved on the horizon $k, \ldots, k+N$, which in effect forms a suboptimal *infinite* horizon control policy.

2.4.1. The linear-quadratic case

There are several problems for which the certainty equivalent control policy turns out to be optimal. The best known (at least in control community) is the linear quadratic case where the stage cost is given by

$$l_k(x_k, u_k) = x_k^T Q_k x_k + u_k^T R_k u_k,\ k = 0, \ldots, N-1$$

with the terminal cost $l_N(x_N) = x_N^T Q_N x_N$ and linear dynamics

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

where the disturbance sequence is assumed to be independent and zero-mean. It follows straightforward from the Bellman recursion that the cost-to-go functions are also quadratics of the form

$$\mathbf{E}V_k(x_k) = x_k^T P_k x_k + q_k,$$

where the matrices P_k satisfy the discrete-time Riccati equation [5] and, more importantly, does not depend on the distribution of the disturbance, whose effect is only additive through the term q_k. Thus, the certainty equivalent shrinking horizon control policy is in this case optimal and can actually be expressed as a time-varying linear state feedback

$$u_k = K_k x_k,\ k = 0, \ldots, N-1.$$

It is worth noting that the assumption of the stage cost being quadratic is crucial. If, for instance, a simple input constraints of the form $||u_k||_\infty \leq U_{\max}$ were added, the stage cost would now be

$$l_k(x_k, u_k) = \begin{cases} x_k^T Q_k x_k + u_k^T R_k u_k & ||u_k||_\infty \leq U_{\max} \\ +\infty & \text{otherwise,} \end{cases}$$

which is no longer quadratic, and, indeed, the certainty equivalent control is, in general, no longer optimal. This can be seen from the Bellman recursion where the optimization problem at the stage $N-1$ is a *constrained* quadratic problem, whose optimal value is *piecewise* quadratic in state and hence the cost-to-go, $\mathbf{E}V_{N-1}(x_{N-1})$, is the expectation over a piecewise quadratic function, which is not likely to depend on the disturbance distribution in only additive manner. Thus, the certainty equivalent control will, in general, not be optimal for horizons longer than 2.

The optimality of the certainty equivalent control for quadratic costs translates to the imperfect state information case in the sense that the optimal control policy is given by

$$u_k = K_k \mathbf{E}\{x_k | \mathcal{I}_k\},$$

where K_k is the same state-feedback matrix as in the full state information case. For Gaussian process and measurement noise, the state estimate $\mathbf{E}\{x_k|\mathcal{I}_k\}$ can be obtained by the Kalman filter [54]. The fact that the optimal full state information control law applied to the optimal state estimate leads to the optimal imperfect state information control law is sometimes referred to as the separation principle [5].

2.5. Disturbance feedback

A fairly general class of suboptimal policies solves the infinite dimensional problem (2.1) in a certain finite dimensional subspace of Borel functions. To further analyse the properties of this approach, linear dynamics

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

is assumed. Moreover, since, given the control input sequence up to time k and liner dynamics, there is a one-to-one Borel measurable transformation between the state sequence $x_0, \ldots, x_k$ and the sequence $x_0, w_0, \ldots, w_{k-1}$, the causal disturbance feedback policies of the form

$$u_k = \phi_k(x_0, w_0, \ldots, w_{k-1}) \tag{2.13}$$

can be used instead of the causal state-sequence feedback (2.2).

Let the finite dimensional subspace at time k be generated by the basis functions

$$\mathcal{E}_k = (e_k^1, \ldots, e_k^{|\mathcal{E}_k|}). \tag{2.14}$$

The disturbance feedback at time k is then given by

$$\phi_k(x_0, w_0, \ldots, w_{k-1}) = \sum_{i=0}^{|\mathcal{E}_k|} \alpha_k^i e_k^i(x_0, w_0, \ldots, w_{k-1}), \tag{2.15}$$

where α_k^i, the basis functions linear combination coefficients, are the optimization variables. The infinite dimensional optimization problem is thus approximated by a finite dimensional one, and, moreover, this problem is convex if the stage cost $l_k(x_k, u_k)$ is convex for all k. This follows straight from convex analysis fundamentals since the cost function is convex for every disturbance realization due to the linear dynamics (since the state and control are then affine in the optimization variables α_k^i), and taking expectation preserves convexity [10].

Note that the finite dimensional approximation can be made arbitrarily accurate by choosing a sufficiently rich set of basis functions (2.14) so this approach provides an appealing alternative to (approximate) dynamic programming. The difference in complexity is most profound for convex cost functions since then powerful convex optimization techniques can be employed to obtain an approximate solution as opposed to largely heuristic techniques of approximate dynamic programming that quickly become intractable in higher dimensions, at least for uncountable state and/or control space.

The only heuristic step involved in the design of a disturbance feedback policy is the choice of a suitable set of basis functions as a trade-off between computational complexity and accuracy. There are two primary factors that contribute to the overall complexity: (i) the size of the basis, i.e., the number of optimization variables, (ii) the speed of evaluation of the cost function with a given set of basis functions. The second factor is a more severe one since in most cases the only way to (approximately) evaluate the cost function (2.1) is by means of Monte Carlo sampling, i.e., to use the approximation

$$\mathbf{E}\left\{ l(x_N) + \sum_{k=0}^{N-1} l(x_k, u_k, w_k) \right\} \approx \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \left\{ l(x_N^{(i)}) + \sum_{k=0}^{N-1} l(x_k^{(i)}, u_k^{(i)}, w_k^{(i)}) \right\}, \tag{2.16}$$

where $w_k^{(i)}$, $k = 0, \ldots, N-1$, $i = 1, \ldots, \tilde{N}$ are samples taken from the joint distribution of $w_0, \ldots, w_{N-1}$, and $u_k^{(i)}$, $x_k^{(i)}$ are the corresponding input and state trajectories generated by these samples.

Even though the convexity is in this case not lost by Monte Carlo sampling, this approach becomes intractable in higher dimensions and for large horizons since the function evaluation and, more importantly, the evaluation of the gradient and Hessian becomes prohibitively demanding due to the sampling if higher accuracy of the cost approximation is required. Note, however, that the size of the optimization problem is determined by the size of the basis and is thus independent of the number of samples $\tilde{N}$. On the other hand, if the problem were formulated with explicit hard constraints, the number of constraint, and possibly associated slack variables, would be proportional the number of samples.

All in all, for most practical applications it is vital that an analytic expression for the cost function in terms of optimization variables α_k^i be found. This is in general impossible for arbitrary combinations of basis and cost functions. Some of the favourable combinations will now be briefly discussed.

2.5.1. Nonlinear feedback, quadratic cost

The most favourable case is the case of the quadratic stage cost

$$l_k(x_k, u_k) = x_k^T Q_k x_k + u_k^T R_k u_k,\ k = 0, \ldots, N-1$$

and linear dynamics since then the cost function (2.1) turns out to be quadratic in the optimization variables α_k^i in (2.15) with arbitrary choice of the nonlinear basis functions (2.14). Indeed, the cost function is then a sum of constant terms and the terms of the form

$$\alpha_k^i \mathbf{E}(e_k^i e_j^l) \alpha_j^l,\ \alpha_k^i \mathbf{E}(e_k^i w_j),\ \alpha_k^i \mathbf{E} e_k^i. \tag{2.17}$$

Hence we only need to evaluate the first and second moments of the nonlinear functions e_k^i, and the moments of the form $\mathbf{E}e_k^i w_j$. This can be done offline, for instance, by means of Monte Carlo simulation. Once these moments are on hand, the problem becomes a quadratic program. Additional constraints can be introduced by a suitable choice of the basis functions e_k^i. For instance, hard control input constraints can be enforced by choosing bounded basis functions and then constraining the coefficients α_k^i (see [22] and also Section 3.2 for details).

The only problem remaining here is the size of the resulting quadratic program, which can be prohibitive if too complex set of basis functions is chosen and fast sampling period is required. For instance, for the affine disturbance policy, described in more detail in the next section, or similar affine-like policies (Section 3.2), the number of variables is approximately $mnN^2/2$. If a general quadratic policies were of interest, i.e., policies such that at each time the policy is a general quadratic function of the past disturbances, the number of variables would be on the order of mn^2N^3. See Section 6.1 for a more detailed discussion.

2.5.2. Affine disturbance feedback

One of the simplest, yet in many cases very effective, closed-loop policy is the causal affine disturbance feedback defined as

$$u_k = \eta_k + \sum_{i=0}^{k-1} K_{k,i} w_i, \ k = 0, \dots, N-1. \tag{2.18}$$

In the case of linear dynamics, the class of all affine disturbance feedback policies coincides with the class of all affine state-sequence policies in the sense that any affine state-sequence policy can be expressed as an affine disturbance feedback policy and vice versa. Indeed, if we denote the corresponding sequences along the horizon as

$$x = \begin{bmatrix} x_0^T & \dots & x_N^T \end{bmatrix}^T, \quad u = \begin{bmatrix} u_0^T & \dots & u_{N-1}^T \end{bmatrix}^T, \quad w = \begin{bmatrix} w_0^T & \dots & w_{N-1}^T \end{bmatrix}^T,$$

and write both policies and the linear state equation in the matrix form

$$u = \eta + Kw, \quad \tilde{u} = \tilde{\eta} + \tilde{K}x,$$
$$x = \mathcal{A}x_0 + \mathcal{B}u + \mathcal{C}w,$$

where

$$\eta = \begin{bmatrix} \eta_0^T & \dots & \eta_{N-1}^T \end{bmatrix}^T, \quad \tilde{\eta} = \begin{bmatrix} \tilde{\eta}_0^T & \dots & \tilde{\eta}_{N-1}^T \end{bmatrix}^T,$$

$$K = \begin{bmatrix} 0 & 0 & \dots & 0 \\ K_{1,1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \\ K_{N-1,1} & \dots & K_{N-1,N-1} & 0 \end{bmatrix}, \quad \tilde{K} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \tilde{K}_{1,1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \\ \tilde{K}_{N-1,1} & \dots & \tilde{K}_{N-1,N} & 0 \end{bmatrix}$$

$$\mathcal{B} = \begin{bmatrix} 0 & & & & \\ B & 0 & & & \\ \vdots & & & & \\ A^{N-2}B & A^{N-3}B & \dots & B & 0 \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} 0 & & & & \\ I & 0 & & & \\ \vdots & & & & \\ A^{N-2} & A^{N-3} & \dots & I & 0 \end{bmatrix},$$

$$\mathcal{A} = \begin{bmatrix} I & A^T & \dots & (A^{N-1})^T \end{bmatrix}^T,$$

we get with the state-sequence feedback that

$$x = \mathcal{A}x_0 + \mathcal{B}\tilde{\eta} + \mathcal{B}\tilde{K}x + \mathcal{C}w.$$

Hence

$$x = (I - \mathcal{B}\tilde{K})^{-1}(\mathcal{A}x_0 + \mathcal{B}\tilde{\eta}) + (I - \mathcal{B}\tilde{K})^{-1}\mathcal{C}w,$$

which is affine in w, and consequently $\tilde{u}$ is also affine in w. The inverse always exists because of the structure of the matrices $\mathcal{B}$ and $\tilde{K}$, which is given by causality. On the other hand, starting with an affine disturbance feedback, we get

$$x = \mathcal{A}x_0 + \mathcal{B}\eta + \mathcal{B}Kw + \mathcal{C}w,$$

so

$$w = (\mathcal{B}K + \mathcal{C})^{\dagger}(x - \mathcal{A}x_0 - \mathcal{B}\eta),$$

which is affine in x, and so is, therefore, u. The last equality follows again from the structure of the matrices and the fact that the first block row of $(x - \mathcal{A}x_0 - \mathcal{B}\eta)$ is zero.

A major advantage of the affine disturbance feedback policy over the state feedback policy is the fact that the resulting optimization problem is affine in the design variables (η, K), which is not the case for the state feedback[4]

The most important feature in the context of this thesis is the fact that the particularly simple form of the policy (2.18) allows for an analytical evaluation for much broader class of cost functions if the disturbances entering the system are Gaussian. In Chapter 3 we show that it is possible to, perhaps after some approximation, arrive at a tractable convex problem for a general ℓ_p cost function even in the case with hard control input bounds, which certainly calls for some modification of the basic affine policy 2.18.

[4]The problem can, however, be transformed to a convex one via a method similar to the classical Q-design procedure [52]. The affine disturbance feedback can actually be thought of as one such transformation.

3. Approximate ℓ_p stochastic optimal control

3.1. Problem statement

This chapter deals with the problem of minimizing the cost function[1]

$$J := \mathbf{E}\left\{ ||Q_N x_N||_p^p + \sum_{k=0}^{N-1} ||Q_k x_k||_p^p + ||R_k u_k||_p^p \right\} \tag{3.1}$$

for $1 \leq p < \infty$ subject to the discrete-time system dynamics

$$x_{k+1} = Ax_k + Bu_k + w_k, \tag{3.2}$$

$x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, and hard input constraints

$$||u_k||_\infty \leq U_{\max},\ k = 0, \dots, N-1, \tag{3.3}$$

where $Q_k \in \mathbb{R}^{n_q \times n}$, $R_k \in \mathbb{R}^{n_r \times m}$ are weighting matrices. All the results derived here generalize with only minor modifications to the case with different bounds on individual control inputs and/or time varying bounds. The disturbance sequence

$$w = [w_0^T, \dots, w_{N-1}^T]^T$$

is assumed to be jointly Gaussian with the covariance matrix Σ_w.

The minimization to be carried out is over all Borel measurable causal disturbance feedback policies

$$u_k = \phi_k(x_0, w_0, \dots, w_{k-1}),\ k = 0, \dots, N-1. \tag{3.4}$$

This problem is, however, in general intractable and various approximation techniques exist, see Chapter 2. Here, we adopt the approach of [22] where the authors propose

[1] For $x = [x_1, \dots, x_n]^T \in \mathbb{R}^n$, the p-norm of x is defined as $||x||_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$ for $p \in [0, \infty)$ and $||x||_\infty = \max_{i=1,\dots,n} |x_i|$.

to search over a class of causal policies affine in certain nonlinear functions of the disturbances, i.e.,

$$u = \eta + Ke(w) = \begin{bmatrix} \eta_0 \\ \vdots \\ \eta_{N-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ K_{1,1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \\ K_{N-1,1} & \dots & K_{N-1,N-1} & 0 \end{bmatrix} e(w), \tag{3.5}$$

where

$$u = [u_0^T, \dots, u_{N-1}^T]^T.$$

The matrix $\eta \in \mathbb{R}^{mN}$ with blocks in $\mathbb{R}^m$, and the strictly lower block triangular matrix $K \in \mathbb{R}^{mN \times nN}$ with blocks in $\mathbb{R}^{m \times n}$ are optimization variables. The choice of the function $e : \mathbb{R}^{nN} \to \mathbb{R}^{nN}$ is discussed later, although it certainly must be bounded should the hard input constraints be satisfied. Hereafter, the bound on $||e(w)||_\infty$ is denoted by ε.

One of the main goals is therefore to solve (at least approximately) the optimization problem

$$\begin{aligned} \underset{\eta,K}{\text{minimize}} \quad & \mathbf{E}\left\{ ||Q_N x_N||_p^p + \sum_{k=0}^{N-1} ||Q_k x_k||_p^p + ||R_k u_k||_p^p \right\} \\ \text{subject to} \quad & u = \eta + Ke(w) \\ & x_{k+1} = Ax_k + Bu_k + w_k \\ & K \text{ is strictly block lower triangular} \\ & \text{constraints on } \eta, K \text{ such that (3.3) is satisfied.} \end{aligned} \tag{3.6}$$

3.2. Tractable solution

The optimization problem (3.6) is, to our knowledge, intractable owing to the p-norm and the nonlinear function $e(w)$ although sampling techniques (see Section 2.5, Eq. (2.16)) can, in principle, be used. We therefore propose to solve a relaxed problem where $u = \eta + Ke(w)$ in (3.6) is replaced with $u = \eta + Kw$ while keeping constraints on η and K such that the hard input constraints are satisfied when the original control policy is used. Suboptimality of this approximation is studied in detail in Section 3.2.3 for $p = 1$. The relaxed problem (as is the original one) must be convex since the objective is convex for each disturbance realization (see [10] and Section 2.5).

In the sequel, we show that the relaxed optimization problem is not only convex but also tractable. To this end, we need an analytical expression for $\mathbf{E}|X|^p$ of a Gaussian random variable X.

3.2.1. Tractability of the proposed approach

Lemma 3.1. *If* $X \sim \mathcal{N}(\mu, \sigma^2)$ *then*

$$g(\mu, \sigma) := \mathbf{E}|X|^p = \frac{2^{p/2}}{\sqrt{\pi}} \sigma^p \Gamma\left(\frac{p+1}{2}\right) \mathrm{M}\left(-\frac{p}{2}, \frac{1}{2}, -\frac{\mu^2}{2\sigma^2}\right) \tag{3.7}$$

and in particular for $p = 1$

$$g(\mu, \sigma) := \mathbf{E}|X| = \sqrt{\frac{2}{\pi}}\, \sigma\, e^{-\frac{\mu^2}{2\sigma^2}} + \mu \,\mathrm{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right), \tag{3.8}$$

where $\Gamma(\cdot)$ *is the Gamma function,* $\mathrm{erf}(\cdot)$ *the error function and* $\mathrm{M}(\cdot,\cdot,\cdot)$ *the Kummer's confluent hypergeometric function (see Appendix A).*

Proof. Follows by a straightforward integration from the definition of the expectation of an absolutely continuous random variable. ■

Now that we have an analytical expression for the (approximate) cost function the gradient and Hessian can be computed by a simple use of vector calculus.

Lemma 3.2. *If* $X \sim \mathcal{N}(\mu, \sigma^2)$ *for* $\sigma > 0$, $X = \mu$ *for* $\sigma = 0$, *and* $\mu(\eta, k) = \mu_0 + b^T\eta$, $\sigma(\eta, k) = ||a + Ck||_2$ *then the function* $f(\eta, k) = (\mathbf{E}|X|^p)(\eta, k)$ *is jointly convex in* (η, k) *and the gradient and Hessian are given by*

$$\nabla f = \frac{\partial f}{\partial \mu}\nabla\mu + \frac{\partial f}{\partial \sigma}\nabla\sigma, \tag{3.9}$$

$$\begin{aligned} \mathrm{Hess}(f) = \nabla\mu &\left[\frac{\partial^2 f}{\partial\mu^2}\nabla\mu + \frac{\partial^2 f}{\partial\mu\partial\sigma}\nabla\sigma\right]^T && (3.10)\\ + \nabla\sigma &\left[\frac{\partial^2 f}{\partial\sigma^2}\nabla\sigma + \frac{\partial^2 f}{\partial\sigma\partial\mu}\nabla\mu\right]^T + \frac{\partial f}{\partial\sigma}\,\mathrm{Jac}(\nabla\sigma), && (3.11)\end{aligned}$$

where

$$\nabla\mu = \begin{bmatrix} b \\ 0 \end{bmatrix}, \qquad \nabla\sigma = \begin{bmatrix} 0 \\ C^T \frac{a+Ck}{\sigma} \end{bmatrix} \tag{3.12}$$

and

$$\mathrm{Jac}(\nabla\sigma) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{\sigma}C^T\left(I - \frac{(a+Ck)(a+Ck)^T}{\sigma^2}\right)C \end{bmatrix}. \tag{3.13}$$

The expressions for partial derivatives are

$$\frac{\partial f}{\partial \mu} = \frac{1}{\sqrt{\pi}} 2^{p/2} \mu p \sigma^{p-2} \gamma M_2, \tag{3.14}$$

$$\frac{\partial f}{\partial \sigma} = \frac{1}{\sqrt{\pi}} 2^{p/2} p \sigma^{p-3} \gamma \left[\sigma^2 M_1 - \mu^2 M_2\right], \tag{3.15}$$

$$\frac{\partial^2 f}{\partial \mu^2} = \frac{1}{3\sqrt{\pi}} 2^{p/2} p \sigma^{p-4} \gamma \left[3\sigma^2 M_2 + \mu^2 (p-2) M_3\right], \tag{3.16}$$

$$\begin{aligned}\frac{\partial^2 f}{\partial \sigma^2} &= \frac{1}{\sqrt{\pi}} 2^{p/2} \sigma^{p-6} \gamma \left\{[\mu^4 + \mu^2(3p-2)\sigma^2\right. \\ &\quad \left. + (p-1)p\sigma^4] M_1 - \mu^2 (1+p) \left(\mu^2 + 2(p-1)\sigma^2\right) M_4\right\},\end{aligned} \tag{3.17}$$

$$\frac{\partial^2 f}{\partial \mu \partial \sigma} = \frac{1}{3\sqrt{\pi}} 2^{p/2} \mu (p-2) p \sigma^{p-5} \gamma \left[\mu^2 M_3 - 3\sigma^2 M_2\right], \tag{3.18}$$

where

$$M_1 = \mathrm{M}\left(-\frac{p}{2}, \frac{1}{2}, -\frac{\mu^2}{2\sigma^2}\right), \qquad M_2 = \mathrm{M}\left(1-\frac{p}{2}, \frac{3}{2}, -\frac{\mu^2}{2\sigma^2}\right), \tag{3.19}$$

$$M_3 = \mathrm{M}\left(2-\frac{p}{2}, \frac{5}{2}, -\frac{\mu^2}{2\sigma^2}\right), \qquad M_4 = \mathrm{M}\left(-\frac{p}{2}, \frac{3}{2}, -\frac{\mu^2}{2\sigma^2}\right)$$

and

$$\gamma = \Gamma\left(\frac{p+1}{2}\right).$$

In particular for $p = 1$ we have

$$\nabla f = \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) \nabla \mu + \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} \nabla \sigma, \tag{3.20}$$

$$\operatorname{Hess}(f) = \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} \left(\frac{1}{\sigma} \begin{bmatrix} b \\ -q\frac{\mu}{\sigma} \end{bmatrix} \begin{bmatrix} b \\ -q\frac{\mu}{\sigma} \end{bmatrix}^T + \operatorname{Jac}(\nabla \sigma)\right), \tag{3.21}$$

where

$$q = C^T \frac{a + Ck}{\sigma}.$$

Proof. Convexity follows from convex calculus fundamentals since

$$f(\eta, k) = \mathbf{E}|\mu_0 + b^T \eta + (a + Ck)^T \tilde{w}|^p$$

for some $\tilde{w} \sim \mathcal{N}(0, I)$, and the right-hand side is convex in (η, k) for every realization of $\tilde{w}$.

The rest is a direct computation, and we carry out in detail only the $p = 1$ case. For the gradient of f we have

$$\nabla f(\mu, \sigma) = \frac{\partial f}{\partial \mu}\nabla\mu + \frac{\partial f}{\partial \sigma}\nabla\sigma = \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\nabla\mu + \sqrt{\frac{2}{\pi}}e^{-\frac{\mu^2}{2\sigma^2}}\nabla\sigma \tag{3.22}$$

with

$$\nabla\mu = \begin{bmatrix} b \\ 0 \end{bmatrix}, \qquad \nabla\sigma = \begin{bmatrix} 0 \\ C^T\frac{a+Ck}{\sigma} \end{bmatrix}. \tag{3.23}$$

The expression for $\nabla\sigma$ follows from the fact that $\nabla||x||_2 = \frac{x}{||x||_2}$ and the chain rule. Now since $\operatorname{Hess}(f) = \operatorname{Jac}(\nabla f)$ and $\operatorname{Jac}(h\tilde{g}) = \tilde{g}(\nabla h)^T + h\operatorname{Jac}(\tilde{g})$ for real-valued function h and multivariate $\tilde{g}$, it follows that

$$\begin{aligned} \operatorname{Hess}(f) = {} & \begin{bmatrix} b \\ 0 \end{bmatrix}\left\{\nabla\operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right\}^T \\ & + \begin{bmatrix} 0 \\ C^T\frac{a+Ck}{\sigma} \end{bmatrix}\left\{\nabla\left(\sqrt{\frac{2}{\pi}}e^{-\frac{\mu^2}{2\sigma^2}}\right)\right\}^T + \sqrt{\frac{2}{\pi}}e^{-\frac{\mu^2}{2\sigma^2}}\operatorname{Jac}(\nabla\sigma) \end{aligned} \tag{3.24}$$

with

$$\operatorname{Jac}(\nabla\sigma) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{||x||_2}C^T\left(I - \frac{xx^T}{||x||_2^2}\right)C \end{bmatrix} \geq 0, \tag{3.25}$$

where $x = a + Ck$ since, again by the chain rule,

$$\operatorname{Jac}_k \nabla\sigma = C^T \operatorname{Jac}\frac{a+Ck}{||a+Ck||_2} = C^T\left[\operatorname{Jac}\left(\frac{y}{||y||_2}\right)\circ(a+Ck)\right]C, \tag{3.26}$$

where $\circ$ denotes the function composition. The remaining two terms in (3.24) are

$$\nabla\operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) = \begin{bmatrix} b \\ 0 \end{bmatrix}\frac{1}{\sigma}\sqrt{\frac{2}{\pi}}e^{-\frac{\mu^2}{2\sigma^2}} - \begin{bmatrix} 0 \\ C^T\frac{a+Ck}{\sigma} \end{bmatrix}\sqrt{\frac{2}{\pi}}\frac{\mu}{\sigma^2}e^{-\frac{\mu^2}{2\sigma^2}}, \tag{3.27}$$

$$\nabla\left(\sqrt{\frac{2}{\pi}}e^{-\frac{\mu^2}{2\sigma^2}}\right) = -\begin{bmatrix} b \\ 0 \end{bmatrix}\sqrt{\frac{2}{\pi}}\frac{\mu}{\sigma^2}e^{-\frac{\mu^2}{2\sigma^2}} + \begin{bmatrix} 0 \\ C^T\frac{a+Ck}{\sigma} \end{bmatrix}\sqrt{\frac{2}{\pi}}\frac{\mu^2}{\sigma^3}e^{-\frac{\mu^2}{2\sigma^2}}. \tag{3.28}$$

Rewriting the Hessian with

$$q := C^T\frac{a+Ck}{\sigma} \tag{3.29}$$

then yields

$$\operatorname{Hess}(f) = \sqrt{\frac{2}{\pi}}e^{-\frac{\mu^2}{2\sigma^2}}\left(\frac{1}{\sigma}\begin{bmatrix} b \\ -q\frac{\mu}{\sigma} \end{bmatrix}\begin{bmatrix} b \\ -q\frac{\mu}{\sigma} \end{bmatrix}^T + \operatorname{Jac}(\nabla\sigma)\right) \geq 0. \tag{3.30}$$

Having computed the Hessian, we can give a direct proof of convexity. The Hessian (3.30) is undefined for $\sigma = 0$, so a little more care is needed. It is easily seen that $f(\eta, k)$ is continuous and that the sequence of smoothed functions

$$f_n(\eta, k) = g\left(\mu(\eta, k), \sqrt{\frac{1}{n} + \sum_i x_i^2}\right),$$

where g is defined in (3.8), converges pointwise to f. The functions f_n are readily shown to be convex by computing their respective Hessians in the same fashion as above, which results in the same expression (3.30) with σ replaced by $\sqrt{\frac{1}{n} + \sum_i x_i^2}$. The function $f(\eta, k)$ is therefore convex since it is a limit of convex functions.

■

Note that, given p, the first two arguments of the hypergeometric functions in (3.19) are constant and the third argument is always negative, which allows for very fast computation of $M_1, \ldots, M_4$, for instance by using methods 1 and 2 of [38] (see Appendix A).

Theorem 3.1. *The optimization problem*

$$\begin{aligned} \underset{\eta, K}{\text{minimize}} \quad & \mathbf{E}\left\{ ||Q_N x_N||_p^p + \sum_{k=0}^{N-1} ||Q_k x_k||_p^p + ||R_k u_k||_p^p \right\} \\ \text{subject to} \quad & u = \eta + Kw \\ & x_{k+1} = Ax_k + Bu_k + w_k \\ & K \text{ is strictly block lower triangular} \\ & |\eta_i| + \varepsilon ||K_i||_\infty \le U_{\max},\ i = 1, \ldots, mN \end{aligned} \tag{3.31}$$

with $w \sim \mathcal{N}(0, \Sigma_w)$ is convex and tractable in the variables (η,K). Furthermore the hard input constraints (3.3) are satisfied under the control policy $u = \eta + Ke(w)$ if $||e(w)||_\infty \le \varepsilon$. Here K_i denotes the i-th row of K and $|| \cdot ||_\infty$ denotes the induced infinity norm[2] *(In particular* not *the maximum absolute value if the matrix is a row vector)*

Proof. The objective function is a sum of terms of the form $\mathbf{E}|q_{jk}^T x_k|^p$ or $\mathbf{E}|r_{jk}^T u_k|^p$, where q_{jk} and r_{jk} denote the j-th rows of Q_k and R_k respectively. Denote also

$$\mathcal{B}_k = [A^{k-1}B, \ldots, B, 0, \ldots, 0], \quad \mathcal{C}_k = [A^{k-1}, \ldots, I, 0, \ldots, 0]F,$$

where $FF^T = \Sigma_w$, and observe that[3]

$$\begin{aligned} q_{jk}^T x_k &= q_{jk}^T (A^k x_0 + \mathcal{B}_k u + \mathcal{C}_k \tilde{w}) \\ &= q_{jk}^T A^k x_0 + q_{jk}^T \mathcal{B}_k \eta + q_{jk}^T (\mathcal{C}_k + \mathcal{B}_k K F)\tilde{w} \end{aligned}$$

[2]The induced p-norm of a matrix A is defined for $1 \le p \le \infty$ as $||A||_p = \sup_{x \ne 0} \frac{||Ax||_p}{||x||_p}$.

[3]Here and hereafter the equality of random elements means the equality of their distributions, not necessarily of the elements itself.

with $\tilde{w} \sim \mathcal{N}(0, I)$. It is clear that $q_{jk}^T x_k$ is Gaussian with the expectation

$$\mu(\eta, k) = \mathbf{E}(q_{jk}^T x_k) = q_{jk}^T A^k x_0 + q_{jk}^T \mathcal{B}_k \eta, \tag{3.32}$$

and standard deviation

$$\sigma(\eta, k) = ||q_{jk}^T (\mathcal{C}_k + \mathcal{B}_k K F)||_2 = ||\mathcal{C}_k^T q_{jk} + (F^T \otimes q_{jk}^T \mathcal{B}_k) S k||_2, \tag{3.33}$$

where $Sk = \mathrm{vec}(K)$ with S being a certain matrix of zeros and ones, and k containing only the nonzero elements of K. Here we employed the equality

$$\mathrm{vec}(ABC) = (C^T \otimes A)\, \mathrm{vec}(B)$$

that relates the vectorization $\mathrm{vec}(\cdot)$ and the Kronecker product $\otimes$ [2]. Similarly

$$r_{jk}^T u_k = r_{jk}^T v_k \eta + r_{jk}^T v_k K F \tilde{w},$$

where v_k is a matrix that selects k-th block row of the size m. Consequently, the expectation and standard deviation become

$$\mu(\eta, k) = \mathbf{E}(r_{jk}^T u_k) = r_{jk}^T v_k \eta, \tag{3.34}$$

$$\sigma(\eta, k) = ||r_{jk}^T v_k K F||_2 = ||(F^T \otimes r_{jk}^T v_k) S k||_2. \tag{3.35}$$

Application of Lemma 3.2, where the expressions for the gradient and Hessian of $(\mathbf{E}|X|^p)(\mu, \sigma)$ were given, now completes the convexity and tractability part of the proof.

Satisfaction of the input constraints follows immediately from the definition of the induced infinity norm and from the assumption that $||e(w)||_\infty \leq \varepsilon$ since

$$|u_i| = ||\eta_i + K_i e(w)||_\infty \leq ||\eta_i||_\infty + ||K_i e(w)||_\infty \leq |\eta_i| + \varepsilon ||K_i||_\infty.$$

■

It may also be of interest to solve the problem with $||\cdot||_p$ instead of $||\cdot||_p^p$, that is, to minimize the cost

$$J := \mathbf{E}\left\{ ||Q_N x_N||_p + \sum_{k=0}^{N-1} ||Q_k x_k||_p + ||R_k u_k||_p \right\} \tag{3.36}$$

instead of (3.1). This indeed can be done with one more approximation (for $p > 1$) based on Jensen's inequality. The individual terms $\mathbf{E}||Q_k x_k||_p$ and $\mathbf{E}||R_k u_k||_p$ can be upper bounded as

$$\mathbf{E}||Q_k x_k||_p \leq \left(\sum_{j=1}^{n_q} \mathbf{E}|q_{jk}^T x_k|^p \right)^{1/p}, \quad \mathbf{E}||R_k u_k||_p \leq \left(\sum_{j=1}^{n_r} \mathbf{E}|r_{jk}^T u_k|^p \right)^{1/p}, \tag{3.37}$$

where q_{jk}^T and r_{jk}^T are, as in the proof of Theorem 3.1, j-th rows of Q_k and R_k.
These upper bounds are convex as is shown in the following Lemma.

Lemma 3.3. *Let $l_i(x, w)$, $i = 1, \ldots, k$, be functions affine in x, and let p be a number greater than or equal to one and w a random variable. Then the function $f(x) = \left(\sum_{i=1}^{k} \mathbf{E}|l_i(x, w)|^p\right)^{1/p}$ is convex.*

Proof. Let $(\Omega, \mathcal{A}, P)$ be the probabilistic space that w is defined on, and let ν be the counting measure on 2^S, where $S = \{1, \ldots, k\}$. Finally define $g(y, x) = l_i(x, w)$ for every $y = (i, w) \in S \times \Omega$ and any x. It is clear that the integral of g over $S \times \Omega$ with respect to the product measure $\mu = \nu \otimes P$ is equal to

$$\int_{S\times\Omega} g \,\mathrm{d}\mu = \sum_{i=1}^{k} \int_{\Omega} l_i(x, w) \,\mathrm{d}P(w) = \sum_{i=1}^{k} \mathbf{E} l_i(x, w),$$

and that, given x, $g(y, x)$ is in L^p if each $l_i(x, w)$ is in L^p. Now the Minkowski inequality can be used to assert the convexity of $f(x)$. Choose $\theta \in [0, 1]$ and x_1, x_2 arbitrary. Then

$$\begin{aligned}
f(\theta x_1 + (1-\theta)x_2) &= \left(\mathbf{E}\sum_{i=1}^{k} |\theta l_i(x_1) + (1-\theta) l_i(x_2, w)|^p\right)^{1/p} \\
&= \left(\int_{S\times\Omega} |\theta g(y, x_1) + (1-\theta) g(y, x_2)|^p \,\mathrm{d}\mu(y)\right)^{1/p} \\
&\leq \left(\int_{S\times\Omega} |\theta g(y, x_1)|^p \,\mathrm{d}\mu(y)\right)^{1/p} + \left(\int_{S\times\Omega} |(1-\theta) g(y, x_2)|^p \,\mathrm{d}\mu(y)\right)^{1/p} \\
&= \theta \left(\mathbf{E}\sum_{i=1}^{k} |l_i(x_1, w)|^p\right)^{1/p} + (1-\theta) \left(\mathbf{E}\sum_{i=1}^{k} |l_i(x_2, w)|^p\right)^{1/p} \\
&= \theta f(x_1) + (1-\theta) f(x_2),
\end{aligned}$$

as desired.
The only inequality employed was the Minkowski inequality [49]. ∎

The proof of Theorem 3.1 now carries over without change for these upper bounds since the last composition with $1/p$ in (3.37) preserves convexity by Lemma 3.3 and the resulting gradient and Hessian are easily computed with the aid of the already computed gradients and Hessians of $\mathbf{E}||Q_k x_k||_p^p$ and $\mathbf{E}||R_k x_k||_p^p$. Indeed, for two functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ we have

$$\nabla(f \circ g) = (\nabla g)(f' \circ g),$$

$$\operatorname{Hess}(f \circ g) = (\nabla g)(\nabla g)^T (f'' \circ g) + (f' \circ g) \operatorname{Hess}(g),$$

which is now applied with

$$g(\eta, k) := \mathbf{E}||Q_k x_k||_p^p \quad \text{or} \quad g(\eta, k) := \mathbf{E}||R_k x_k||_p^p$$

and $f(t) := t^{1/p}$.

3.2.2. Output feedback

Two ways of applying the presented approach to construct an output feedback are presented in this section. The first approach employs Kalman filter innovations feedback [31], whereas the second approach is related to the nonlinear Q-design introduced in [51]. We assume a system in the form

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \\ y_k &= Cx_k + v_k, \end{aligned} \tag{3.38}$$

where the measurement noise v_k and the process noise w_k are, for simplicity, assumed to be zero-mean mutually independent i.i.d. Gaussian random sequences with the covariances $\mathbf{E}v_k v_k^T = \tilde{\Sigma}_v > 0$ and $\mathbf{E}w_k w_k^T = \tilde{\Sigma}_w$, respectively. Corresponding trajectories over the optimization horizon are denoted by

$$y = \begin{bmatrix} y_{t+1}^T & \dots & y_{t+N-1}^T \end{bmatrix}^T, \quad v = \begin{bmatrix} v_{t+1}^T & \dots & v_{t+N-1}^T \end{bmatrix}^T,$$

$$u = \begin{bmatrix} u_t^T & \dots & u_{t+N-1}^T \end{bmatrix}^T, \quad w = \begin{bmatrix} w_t^T & \dots & w_{t+N-1}^T \end{bmatrix}^T.$$

We further assume that at time t, measurements $\mathcal{Y}_t = (y_0, \dots, y_t)$ are on hand, and we want to minimize the cost

$$J_t := \mathbf{E}\left\{ ||Q_N x_{t+N}||_p^p + \sum_{k=0}^{N-1} ||Q_k x_{t+k}||_p^p + ||R_k u_{t+k}||_p^p \;\middle|\; \mathcal{Y}_t \right\} \tag{3.39}$$

over a certain class of output feedback policies.

3.2.2.1. Innovations feedback

First we consider minimization over the policies

$$u = \eta + Ke(\epsilon), \tag{3.40}$$

where $\epsilon = \begin{bmatrix} \epsilon_1^T, \dots, \epsilon_{N-1}^T \end{bmatrix}^T$ is the Kalman filter innovations sequence,

$$\epsilon_k = y_{t+k} - C\hat{x}_{t+k|t+k-1},$$

where

$$\hat{x}_{t+k|t+k-1} = \mathbf{E}(x_{t+k} \mid y_0, \ldots, y_{t+k-1}).$$

Conditionally on $\mathcal{Y}_t$, the innovations sequence ϵ as well as the initial state x_t and the disturbance sequence w are Gaussian random vectors, thus the cost function (3.39) assumes an analytical expression if the affine relaxation of the control law (3.40),

$$u = \eta + K\epsilon, \tag{3.41}$$

is adopted.

First we evaluate the covariance matrix of $[x_t^T, \epsilon^T, w^T]^T$ using the Kalman filter estimate error dynamics. With $\tilde{x}_k = x_{t+k} - \hat{x}_{t+k|t+k-1}$ for $k = 1, \ldots, N-1$ and $\tilde{x}_t = x_t - \hat{x}_{t|t}$ we have

$$\tilde{x}_1 = Ax_t + Bu_t + w_t - A\hat{x}_{t|t} - Bu_t = A\tilde{x}_t + w_t,$$

and

$$\tilde{x}_{k+1} = (A - K_k C)\tilde{x}_k + w_{t+k} - K_k v_{t+k},\ k = 1, \ldots, N-1,$$

where

$$K_k = AP_{k|k-1}C^T(CP_{k|k-1}C^T + \tilde{\Sigma}_v)^{-1}$$

is the Kalman gain associated with the Riccati equation

$$P_{k+1|k} = AP_{k|k-1}A^T + \tilde{\Sigma}_w - K_k(CP_{k|k-1}C^T + \tilde{\Sigma}_v)K_k^T \tag{3.42}$$

with the initial condition $P_{1|0} = AP_{t|t}A^T + \tilde{\Sigma}_w$, where $P_{t|t}$ is the Kalman filter error covariance at time t based on measurements up to this time $\mathcal{Y}_t$. Rewriting the error dynamics in a matrix form gives

$$\tilde{x} = \tilde{\mathcal{A}}\tilde{x}_t + \tilde{\mathcal{B}}w + \tilde{\mathcal{D}}v,$$

where

$$\tilde{x} = \begin{bmatrix} \tilde{x}_{t+1}^T & \ldots & \tilde{x}_{t+N-1}^T \end{bmatrix}^T,$$

and

$$\tilde{\mathcal{A}} = \begin{bmatrix} I \\ A - K_1 C \\ (A - K_2 C)(A - K_1 C) \\ \vdots \\ (A - K_{N-2}C) \cdot \ldots \cdot (A - K_1 C) \end{bmatrix} A, \quad \tilde{\mathcal{B}} = \begin{bmatrix} I & & & & \\ \tilde{B}_{1,1} & I & & & \\ \vdots & & & & \\ \tilde{B}_{N-2,1} & \tilde{B}_{N-2,2} & \ldots & I & 0 \end{bmatrix},$$

$$\tilde{\mathcal{D}} = \begin{bmatrix} 0 & & & & \\ -K_1 & 0 & & & \\ -\tilde{D}_{2,1}K_1 & -K_2 & 0 & & \\ \vdots & & & & \\ -\tilde{D}_{N-2,1}K_1 & -\tilde{D}_{N-2,2}K_2 & \dots & -K_{N-2} & 0 \end{bmatrix}$$

with

$$\tilde{B}_{i,j} = \tilde{D}_{i,j} = (A - K_i C) \cdot \ldots \cdot (A - K_j C).$$

Hence the innovations sequence is given by

$$\epsilon = \tilde{\mathcal{C}}\tilde{x} + v = \tilde{\mathcal{C}}\tilde{\mathcal{A}}\tilde{x}_t + \tilde{\mathcal{C}}\tilde{\mathcal{B}}w + (I + \tilde{\mathcal{C}}\tilde{\mathcal{D}})v,$$

where $\tilde{\mathcal{C}} = \text{bdiag}(C, \dots, C)$. The covariance thus becomes

$$\Sigma = \mathbf{E}\left\{ \begin{bmatrix} x_t \\ \epsilon \\ w \end{bmatrix} \begin{bmatrix} x_t \\ \epsilon \\ w \end{bmatrix}^T \Bigg| \mathcal{Y}_t \right\} = \begin{bmatrix} P_{t|t} & P_{t|t}\tilde{\mathcal{A}}^T\tilde{\mathcal{C}}^T & 0 \\ \tilde{\mathcal{C}}\tilde{\mathcal{A}}P_{t|t} & \Lambda & \tilde{\mathcal{C}}\tilde{\mathcal{B}}\Sigma_w \\ 0 & \Sigma_w\tilde{\mathcal{B}}^T\tilde{\mathcal{C}}^T & \Sigma_w \end{bmatrix}, \tag{3.43}$$

where

$$\Lambda = \tilde{\mathcal{C}}\tilde{\mathcal{A}}P_{t|t}\tilde{\mathcal{A}}^T\tilde{\mathcal{C}}^T + \tilde{\mathcal{C}}\tilde{\mathcal{B}}\Sigma_w\tilde{\mathcal{B}}^T\tilde{\mathcal{C}}^T + (I + \tilde{\mathcal{C}}\tilde{D})\Sigma_v(I + \tilde{\mathcal{D}}^T\tilde{\mathcal{C}}^T), \tag{3.44}$$

and

$$\Sigma_w = \text{bdiag}(\tilde{\Sigma}_w, \dots, \tilde{\Sigma}_w), \quad \Sigma_v = \text{bdiag}(\tilde{\Sigma}_v, \dots, \tilde{\Sigma}_v).$$

Now we are ready to evaluate the individual terms of the cost. Since ϵ and w are zero-mean and, conditionally on $\mathcal{Y}_t$, $x_t \sim \mathcal{N}(\hat{x}_{t|t}, P_{t|t})$, the terms $q_{ik}^T x_{t+k}$ can be expressed as

$$\begin{aligned} q_{ik}^T x_{t+k} &= q_{ik}^T[A^k x_t + \mathcal{B}_k(\eta + K\epsilon) + \mathcal{C}_k w] = q_{ik}^T(\mathcal{B}_k\eta + A^k\hat{x}_{t|t}) + q_{ik}^T[A^k\tilde{x}_t + \mathcal{B}_k K\epsilon + \mathcal{C}_k w] \\ &= q_{ik}^T(\mathcal{B}_k\eta + A_k\hat{x}_{t|t}) + q_{ik}^T[A^k, \mathcal{B}_k K, \mathcal{C}_k]\Sigma^{1/2}\tilde{w} \end{aligned}$$

for $\tilde{w} \sim \mathcal{N}(0, I)$ and

$$\mathcal{B}_k = \left[A^{k-1}B, \dots, B, 0, \dots, 0\right], \quad \mathcal{C}_k = \left[A^{k-1}, \dots, I, 0, \dots, 0\right].$$

Consequently,

$$\mu(\eta, K) = \mathbf{E}(q_{ik}^T x_{t+k}|\mathcal{Y}_y) = q_{ik}^T(A^k\hat{x}_{t|t} + \mathcal{B}_k\eta),$$

and

$$\sigma(\eta, K) = \left|\left| q_{ik}^T A_k \Sigma_{1:n}^{1/2} + q_{ik}^T\mathcal{B}_k K\Sigma_{n+1:l(N-1)+n}^{1/2} + q_{ik}^T\mathcal{C}_k\Sigma_{l(N-1)+n+1:\text{end}}^{1/2} \right|\right|_2,$$

where l is the output dimension and the colon notation selects rows of $\Sigma^{1/2}$.

Similarly for the terms $r_{ik}^T u_k$ we have

$$\mu(\eta, K) = r_{ik}^T v_k \eta, \quad \sigma(\eta, K) = ||r_{ik}^T v_k K \Lambda^{1/2}||_2,$$

where the matrix v_k selects k-th block row of size m and Λ is defined in (3.44).

Thus, we have converted the problem to the framework of Lemma 3.2, which guarantees convexity and tractability.

3.2.2.2. Output error feedback

Instead of taking innovations feedback (3.40), a policy of the form

$$u = \eta + Ke(y - Hu - \mathcal{A}\hat{x}_{t|t}), \tag{3.45}$$

where

$$H = \begin{bmatrix} CB & 0 & & & \\ \vdots & & & & \\ CA^{N-2}B & CA^{N-3}B & \dots & CB & 0 \end{bmatrix},$$

and

$$\mathcal{A} = \begin{bmatrix} (CA)^T & \dots & (CA^{N-1})^T \end{bmatrix}^T$$

also leads to a tractable representation since the feedback term $y - Hu - \mathcal{A}\hat{x}_{t|t}$ can be expressed as

$$y - Hu - \mathcal{A}\hat{x}_{t|t} = \mathcal{A}(x_t - \hat{x}_{t|t}) + \mathcal{D}w + v,$$

where $\mathcal{D}$ is defined in the same way as H with B replaced by I. Thus, according to the previous discussion, the term $y - Hu - \mathcal{A}\hat{x}_{t|t}$ conditioned on $\mathcal{Y}_t$ is a Gaussian random variable with known moments

$$\mathbf{E}(y - Hu - \mathcal{A}\hat{x}_{t|t}) = 0,$$

$$\mathrm{cov}(y - Hu - \mathcal{A}\hat{x}_{t|t}) = \mathcal{A}P_{t|t}\mathcal{A}^T + \mathcal{D}\Sigma_w\mathcal{D}^T + \Sigma_v,$$

where $\Sigma_v = \mathrm{bdiag}(\tilde{\Sigma}_v, \dots, \tilde{\Sigma}_v)$. Lemma 3.2 can now be applied in a straightforward way since if the affine relaxation

$$u = \eta + K(y - Hu - \mathcal{A}\hat{x}_{t|t}) \tag{3.46}$$

is employed instead of (3.45), we get for state-related terms $q_{ik}^T x_k$ that

$$\begin{aligned} q_{ik}^T x_k &= q_{ik}^T[A^k x_t + \mathcal{B}_k(\eta + K(y - Hu - \mathcal{A}\hat{x}_{t|t})) + \mathcal{C}_k\tilde{w}] \\ &= q_{ik}^T[\mathcal{B}_k\eta + A^k x_t + \mathcal{B}_k K\mathcal{A}(x_t - \hat{x}_{t|t}) + (\mathcal{B}_k K\mathcal{D}\Sigma_w^{1/2} + \mathcal{C}_k)\tilde{w} + \mathcal{B}_k K v]. \end{aligned}$$

Consequently

$$\mu(\eta, K) = q_{ik}^T \mathcal{B}_k \eta + q_{ik}^T A^k \hat{x}_{t|t},$$

$$\sigma(\eta, K) = \left|\left| q_{ik}^T \left[A^k P_{t|t}^{1/2}, \mathcal{C}_k, 0 \right] + q_{ik}^T \mathcal{B}_k K \left[\mathcal{A}_k P_{t|t}^{1/2}, \mathcal{D}\Sigma_w^{1/2}, \Sigma_v^{1/2} \right] \right|\right|_2 .$$

Similarly for control terms $r_{ik}^T u_k$ we have

$$r_{ik}^T u_k = r_{ik}^T v_k [\eta + K(y - Hu - \mathcal{A}\hat{x}_{t|t})] = r_{ik}^T v_k [\eta + K\mathcal{A}(x_t - \hat{x}_{t|t}) + K\mathcal{D}w + Kv],$$

and therefore

$$\mu(\eta, K) = r_{ik}^T v_k \eta,$$

$$\sigma(\eta, K) = \left|\left| r_{ik}^T v_k K [\mathcal{A} P_{t|t}^{1/2}, \mathcal{D}\Sigma_w^{1/2}, \Sigma_v] \right|\right|_2 ,$$

which is again exactly in the form required by Lemma 3.2.

A receding horizon implementation of both policies amounts to recursively computing the conditional distribution of x_t by the Kalman filter, minimizing the cost (3.39) every N_c steps over affine innovations or output error feedback policies subject to input constraints, and applying first N_c controls of the corresponding nonlinear policy (3.40) or (3.45). Finally note that in the long run, the steady state values of the Kalman gain and error covariance can be used to compute the joint covariance Σ in (3.43), slightly reducing computational complexity.

3.2.3. Bound on suboptimality

In this section we provide a bound on the suboptimality in (3.6) (with the same constraints on η, K as in (3.31)) of the solution to the relaxed problem (3.31) for $p = 1$. The idea is to bound the difference of the costs under the policies $u = \eta + Kw$ and $u = \eta + Ke(w)$ for given η, K, which in effect bounds the difference of the respective optima. For ease of notation, the result is derived with time-invariant weights, i.e., $Q_k := Q$, $R_k := R$ (and thus $q_{jk} := q_j$, $r_{jk} := r_j$) for all k, but generalizes immediately to the time-varying case.

Lemma 3.4. *The cost J_e incurred under the policy $u = \eta + Ke(w)$ and the cost J_w incurred under the policy $u = \eta + Kw$ differ not more than*

$$(n_q(N+1)||Q||_\infty ||\mathcal{B}_N||_\infty + n_r N ||R||_\infty)\mathbf{E}||e(w) - w||_\infty ||K||_\infty \tag{3.47}$$

Proof. We have

$$|J_e - J_w| \leq \sum_{k=0}^{N} \sum_{j=1}^{n_q} |\mathbf{E}(|q_j^T x_k^e| - |q_j^T x_k^w|)| + \sum_{k=0}^{N-1} \sum_{j}^{n_r} |\mathbf{E}(|r_j^T u_k^e| - |r_j^T u_k^w|)|. \tag{3.48}$$

Next, by Jensen's inequality,

$$\begin{aligned}|\mathbf{E}(|q_j^T x_k^e| - |q_j^T x_k^w|)| &\leq \mathbf{E}\left||q_j^T x_k^e| - |q_j^T x_k^w|\right| \\ &\leq \mathbf{E}(|q_j^T x_k^e - q_j^T x_k^w|) = \mathbf{E}|q_j^T \mathcal{B}_k K(e(w) - w)|,\end{aligned} \tag{3.49}$$

where

$$x_k^e = A^k x_0 + \mathcal{B}_k \eta + \mathcal{B}_k K e(w) + \mathcal{C}_k w, \;\; x_k^w = A^k x_0 + \mathcal{B}_k \eta + \mathcal{B}_k K w + \mathcal{C}_k w.$$

Furthermore

$$\begin{aligned}\mathbf{E}|q_j^T \mathcal{B}_k K(e(w) - w)| &\leq ||q_j^T \mathcal{B}_k K||_\infty \mathbf{E}||e(w) - w||_\infty \\ &\leq ||q_j^T \mathcal{B}_k||_\infty ||K||_\infty \mathbf{E}||e(w) - w||_\infty \\ &\leq ||Q||_\infty ||\mathcal{B}_N||_\infty ||K||_\infty \mathbf{E}||e(w) - w||_\infty.\end{aligned} \tag{3.50}$$

Similar procedure can be carried out for control inputs to yield

$$|\mathbf{E}(|r_j^T u_k^e| - |r_j^T u_k^w|)| \leq ||R||_\infty ||K||_\infty \mathbf{E}||e(w) - w||_\infty.$$

Summing up all terms in (3.48) now leads to the desired result

$$|J_e - J_w| \leq (n_q(N+1)||Q||_\infty ||\mathcal{B}_N||_\infty + n_r N ||R||_\infty)\mathbf{E}||e(w) - w||_\infty ||K||_\infty,$$

which completes the proof. ■

Now it is rather straightforward to derive the suboptimality bound. Denote J_e^* the optimal value of (3.6) and the corresponding minimizer K_e^*, η_e^*. Denote also J_w^* the optimal value of (3.31) and the corresponding optimal solution K_w^*, η_w^*. Finally denote J_e the cost J under the control policy $u = \eta_w^* + K_w^* e(w)$ and J_w the cost J under the policy $u = \eta_e^* + K_e^* w$.

Theorem 3.2. *The solution η_w^*, K_w^* of (3.31) is not more than*

$$\beta := 2(n_q(N+1)||Q||_\infty ||\mathcal{B}_N||_\infty + n_r N ||R||_\infty)\mathbf{E}||e(w) - w||_\infty \frac{U_{\max}}{\varepsilon} \tag{3.51}$$

suboptimal in (3.6).

Proof. It follows from Lemma 3.4 that

$$|J_e - J_w^*| \leq \frac{\beta}{2}, \quad |J_w - J_e^*| \leq \frac{\beta}{2}$$

since $||K_e^*||_\infty \leq U_{\max}/\varepsilon$, $||K_w^*||_\infty \leq U_{\max}/\varepsilon$ because of the constraint on K and η in both optimization problems:

$$|\eta_i| + \varepsilon ||K_i||_\infty \leq U_{\max}, \; i = 1, \ldots, mN$$

implies $||K||_\infty \leq U_{\max}/\varepsilon$.

Now since $J_e^* \leq J_e$ and $J_w^* \leq J_w$ the bound immediately follows

$$0 \leq J_e - J_e^* \leq J_e - J_w^* + J_w - J_e^* = |J_e - J_w^* + J_w - J_e^*| \leq \beta,$$

which completes the proof. ■

The term $\mathbf{E}||e(w) - w||_\infty$ in (3.51) can be computed to virtually arbitrary precision by means of a Monte Carlo simulation. The bound also provides an intuitively obvious guide to selecting the function $e(w)$ in such a way that $e(w)$ and w do not differ very much with high probability. For instance with the choice of $e(w)$ as the elementwise saturation $e_i(w_i) = \mathrm{sat}_r(w_i)$ with $r \gtrsim 4\sqrt{\rho(\Sigma)}$ it is highly likely that the bound will be close to zero and, consequently, the solution to the relaxed problem will be almost optimal in the original one. Note also that this fairly crude bound can be significantly improved by terminating one inequality earlier in (3.50) at the cost of a slightly more complicated expression

$$\tilde{\beta} := 2\mathbf{E}||e(w) - w||_\infty \frac{U_{\max}}{\varepsilon} \left\{ \sum_{k=0}^{N} \sum_{j=1}^{n_q} ||q_j^T \mathcal{B}_k||_\infty + N \sum_{j=1}^{n_r} ||r_j^T||_\infty \right\}. \tag{3.52}$$

4. Stochastic stability of linear systems

In this chapter stochastic stability in the mean-square sense of linear systems with bounded control inputs is discussed, and some connections with a different concept of stability are pointed out.

Throughout this chapter we consider the linear system

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

with the disturbance sequence $\{w_k,\, k \in \mathbb{N}_0\}$ independent. There will be no specific assumptions on the disturbance distribution, only on the boundedness of some of its moments.

Definition 4.1. *A discrete-time stochastic system is said to be mean-square stable if*

$$\sup_{k\geq 0} \mathbf{E}||x_k||_2^2 < \infty.$$

This is a convenient concept of stability to work with although it gives no protection against large deviations occurring over a long period of time. Indeed, a sequence of i.i.d. Gaussian random variables $\{x_k,\, k \in \mathbb{N}_0\}$ is certainly mean-square bounded and still

$$P\left(\limsup_{k\to\infty} x_k = \infty\right) = \lim_{N\to\infty} P\left(\limsup_{k\to\infty} x_k > N\right) = 1$$

since

$$P\left(\limsup_{k\to\infty} x_k > N\right) = 1 - P\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}[x_k \leq N]\right) = 1$$

as $P\left(\bigcap_{k=n}^{\infty}[x_k \leq N]\right) = 0$ by the i.i.d. assumption and the fact that $P(x_k \leq N) < 1$ for a Gaussian random variable x_k.

Another plausible concept of stability of a stochastic process is the requirement of a zero probability of divergent trajectories. In this case, the mean-square stability (or in fact boundedness of any moment) does guarantee zero probability of a divergent trajectory as shows the following Lemma.

Lemma 4.1. *Let a stochastic process $\{x_k,\, k \in \mathbb{N}_0\}$ have a bounded p-th moment, i.e., $\sup_{k\geq 0} \mathbf{E}||x_k||_2^p < \infty$ for some $p > 0$. Then there is a zero probability of a divergent trajectory, i.e.,*

$$P\left(\lim_{k\to\infty} ||x_k||_2 = \infty\right) = 0.$$

Proof. Assume that $P(\lim_{k\to\infty} ||x_k||_2 = \infty) = \delta > 0$. Then

$$P\left(\bigcup_{n=0}^{\infty}\bigcap_{k=n}^{\infty}[||x_k||_2 \geq t]\right) = \lim_{n\to\infty}\left(\bigcap_{k=n}^{\infty}[||x_k||_2 \geq t]\right) \geq \delta$$

for every t. Hence there exists $0 < \delta' < \delta$ such that for every t there is an n such that $P(||x_n||_2 \geq t) \geq \delta'$, which is a contradiction since

$$P(||x_n||_2 \geq t) \leq \frac{\mathbf{E}||x_n||_2^p}{t^p} \leq \frac{\sup_{k\geq 0}\mathbf{E}||x_k||_2^p}{t^p}$$

by Markov's inequality. ∎

On the other hand, zero probability of divergent trajectories does not imply mean-square boundedness. Consider a sequence of independent random variables $\{x_k, k \in \mathbb{N}_0\}$ having a Laplace distribution with densities $f_{x_k}(x) = \frac{1}{2}ke^{-|x|/k}$. Then

$$\mathbf{E}|x_k|^2 = 2k^2 \to \infty,$$

but

$$P\left(\bigcap_{k=n}^{\infty}[|x_k| > t]\right) = \prod_{k=n}^{\infty} P(|x_k| > t) = \prod_{k=n}^{\infty} e^{-\frac{t}{k}} = 0$$

for any t and any n. The last equality follows by taking a logarithm, which gives a divergent $1/k$ series. Hence, using the same argument as in the proof of Lemma 4.1

$$P\left(\lim_{k\to\infty}|x_k| = \infty\right) = 0.$$

4.1. Conditions for stabilizability

The current state of knowledge on mean-square stabilizability of linear systems with bounded control inputs is as follows:

A **is strictly stable** Mean-square stability is guaranteed with *any* bounded control input. See Theorem 4.1.

A **is marginally stable** The state is mean-square stabilizable provided that the 4-th moment of the noise sequence is bounded and there is sufficient control authority. See Lemma 4.2 and Theorem 4.2 for a proof with slightly stronger assumptions on the noise sequence.

A **is marginally unstable** Open problem. See Section 4.4.

A **is strictly unstable** The state is not mean-square stabilizable whenever the control authority is bounded.

4.2. Strictly stable case

Now we provide a slight generalization and an elementary proof of a result on mean-square boundedness of Schur stable ($\rho(A) < 1$) systems that already appeared in [22].

Theorem 4.1. *Let u_k, w_k be two stochastic processes defined on the same probabilistic space with $||u_k||_\infty \le U_{\max}$ a.s. and $\sup_{i,j} ||\mathbf{E}\{w_i w_j^T\}|| < \infty$. The state of the system $x_{k+1} = Ax_k + Bu_k + w_k$ then stays mean-square bounded (i.e., $\sup_k \mathbf{E}||x_k||_2^2 < \infty$) provided that $\mathbf{E}||x_0||_2^2 < \infty$ and $\rho(A) < 1$.*

Proof. $\mathbf{E}||x_k||_2^2 = \mathrm{tr}(\mathbf{E}\{x_k x_k^T\})$ and consequently it suffices to show that $\mathbf{E}\{x_k x_k^T\}$ is bounded in any norm because of the norm equivalence on finite dimensional vector spaces and the fact that $\mathrm{tr}(\cdot)$ coincides with the nuclear norm on the space of positive semidefinite matrices. The proof proceeds by direct evaluation:

$$\begin{aligned}\mathbf{E}(x_k x_k^T) &= \mathbf{E}\{(A^k x_0 + \mathcal{B}_k U_k + \mathcal{C}_k W_k)(A^k x_0 + \mathcal{B}_k U_k + \mathcal{C}_k W_k)^T\} \\ &= A^k P_0 (A^k)^T + A^k \mathbf{E}\{x_0 U_k^T\}\mathcal{B}_k^T + \mathcal{B}_k \mathbf{E}\{U_k x_0^T\}(A^k)^T \\ &+ \mathcal{B}_k \mathbf{E}\{U_k U_k^T\}\mathcal{B}_k^T + \mathcal{B}_k \mathbf{E}\{U_k W_k^T\}\mathcal{C}_k^T + \mathcal{C}_k \mathbf{E}\{W_k U_k^T\}\mathcal{B}_k^T \\ &+ \mathcal{C}_k \mathbf{E}\{W_k W_k^T\}\mathcal{C}_k^T,\end{aligned} \tag{4.1}$$

where

$$\begin{aligned} U_k &= [u_0^T, \ldots, u_{k-1}^T]^T, \quad W_k = [w_0^T, \ldots, w_{k-1}^T]^T, \\ \mathcal{B}_k &= \left[A^{k-1}B, \ldots, B\right], \quad \mathcal{C}_k = \left[A^{k-1}, \ldots, I\right].\end{aligned}$$

The boundedness of the first term is obvious, the boundedness of the second and third terms follows from the fact that $||\mathbf{E}\{x_0 U_k^T\}||_2 \le U_{\max}\sqrt{mk\mathbf{E}||x_0||_2^2}$ (this follows directly by Jensen's and Cauchy-Schwarz inequalities). The boundedness of $\mathcal{B}_k$ is obvious by the assumption that $\rho(A) < 1$, and therefore the second and third terms actually go to zero.

Consider now any family of matrices M_{rq} such that $||M_{rq}|| \le \hat{\Delta}$ for all r, q and some $\hat{\Delta} < \infty$. For such a family and any submultiplicative norm $||\cdot||$ we have

$$\left|\left|\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} A^i M_{rq} A^j\right|\right| \le \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} ||A^i|| ||M_{rq}|| ||A^j|| \le \hat{\Delta}\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} ||A^i|| ||A^j||. \tag{4.2}$$

The first term in (4.2) is therefore bounded since the last series is convergent by the assumption that $\rho(A) < 1$.

The theorem then follows since the last four terms in (4.1) can be casted in the stated form with $r = k - i - 1$, $q = k - j - 1$ and M_{rq} componentwise bounded (by Cauchy-Schwarz inequality and the assumptions on u_k, w_k) and hence $||\cdot||$ bounded due to the norm equivalence. ∎

Corollary 4.1. *The receding horizon implementation of the control policy defined by solving the optimization problem (3.31) every $N_c \leq N$ steps and applying the first N_c control inputs generated by the policy $u = \eta + Ke(w)$ renders the state x_k mean-square bounded provided that $\rho(A) < 1$.*

Proof. Follows directly from Theorem 4.1 since the constraints in (3.31) ensure that the inputs stay bounded. ■

4.3. Marginally stable case

Now we turn to the more interesting question of mean-square stabilizability of marginally stable ($\rho(A) = 1$, Lyapunov stable) linear systems with bounded control inputs. The question was answered to satisfactory extent in [30] where the authors employed a deeper result about boundedness of moments of a sequence with a negative drift proven in [44]. At full strength, this result is outside the scope of this work, but there is a simpler argument given stronger hypotheses. In either case there is the crucial assumption of a bounded p-th moment of the process noise. The strongest result available ensures that there is a causal feedback policy that renders the system mean-square bounded provided that the deterministic part of the system is stabilizable, there is sufficiently large control authority and the *fourth* moment of the process noise is bounded. Under the hypothesis of the p-th moment of the process noise being bounded for some $p > 6$, a more direct argument that blends the ideas of [30] and [44] can be devised.

First recall the notation

$$\mathcal{B}_n = \begin{bmatrix} A^{n-1}B & A^{n-2}B & \dots & B \end{bmatrix}, \quad \mathcal{C}_n = \begin{bmatrix} A^{n-1} & A^{n-2}, \dots, I \end{bmatrix}$$

and denote q the first integer such that $\mathcal{B}_q$ has full rank. In the following we will assume that the pair (A, B) is reachable, so that $q \leq n$. The technical part of the argument is given in the following Lemma where we assume the system matrix being orthogonal, i.e., $||Ax||_2 = ||x||_2$ for all x. We also denote

$$\operatorname*{sat}_{\gamma}(x) := \begin{cases} \gamma \frac{x}{||x||_2} & ||x||_2 > \gamma \\ x & \text{otherwise.} \end{cases}$$

Lemma 4.2. *Let w_k be an i.i.d random sequence with $\mathbf{E}||w_0||_2^p < \infty$ for some $p > 2$. Let further be the pair (A, B) reachable, and the matrix A orthogonal. Then there exists a causal feedback policy with $||u_k||_\infty < U_{\max}$ such that the state of the system $x_{k+1} = Ax_k + Bu_k + w_k$ satisfies $\sup_{k \geq n} \mathbf{E}||x_k||_2^r < \infty$ for every $0 \leq r < (p-2)/2$ provided that $U_{\max} > q\mathbf{E}||w_0||_2 ||\mathcal{B}_q^\dagger A^q||_\infty$.*

Proof. We define a subsampled process of the state norms by

$$X_n := ||x_{nq}||_2, \ n \geq 0.$$

Thus we would like to prove that

$$\sup_{n \geq 0} \mathbf{E} X_n^r < \infty,$$

and the Lemma will follow by a routine use of triangle inequalities for the original process.

Since

$$\mathbf{E} X_n^r = \int_0^\infty P(X_n^r \geq t) \,\mathrm{d}t = \int_0^\infty t^{r-1} P(X_n \geq t) \,\mathrm{d}t,$$

we can proceed by bounding $P(X_n \geq t)$ by decomposing according to the last time before n that the sequence was less than or equal to some $\mathcal{J} > q\mu$, where $\mu := \mathbf{E}||w_0||_2$. For $t > \mathcal{J}$ we have

$$\begin{aligned}
P(X_n \geq t) &= \sum_{k=0}^{n-1} P\left(X_n \geq t, X_k \leq \mathcal{J}, X_i > \mathcal{J}, i \in \mathbb{N}_{k+1}^{n-1}\right) \\
&\leq \sum_{k=0}^{n-1} P\left(X_n \geq t \mid X_k \leq \mathcal{J}, X_i > \mathcal{J}, i \in \mathbb{N}_{k+1}^{n-1}\right) \\
&= \sum_{k=0}^{n-1} P\left(X_k + X_{k+1} - X_k + \ldots + X_n - X_{n-1} \geq t \mid X_k \leq \mathcal{J}, X_i > \mathcal{J}, i \in \mathbb{N}_{k+1}^{n-1}\right).
\end{aligned}$$

Now define the control input outside $\mathcal{J}$ as

$$u_{qt:q(t+1)} = -\mathcal{B}_q^\dagger A^q \operatorname*{sat}_\gamma(x_t), \ ||x_t||_2 > \mathcal{J}, \tag{4.3}$$

and arbitrarily otherwise with

$$q\mu < \gamma < \min\left\{\mathcal{J}, \frac{U_{\max}}{||\mathcal{B}_q^\dagger A^q||_\infty}\right\}.$$

Such a γ exists due to the assumption on $U_{\max}$. This choice of γ guarantees that $||u_{qt:q(t+1)}||_\infty \leq U_{\max}$ since

$$||u_{qt:q(t+1)}||_\infty = ||B_q^\dagger A^q||_\infty || \operatorname*{sat}_\gamma(x_t)||_\infty \leq ||B_q^\dagger A^q||_\infty || \operatorname*{sat}_\gamma(x_t)||_2 \leq ||B_q^\dagger A^q||_\infty \gamma \leq U_{\max}.$$

With this control input we get for $i \geq 1$

$$\begin{aligned}
X_{k+i+1} - X_{k+i} &= ||A^q x_{q(k+i)} - \mathcal{B}_q \mathcal{B}_q^\dagger A^q \operatorname*{sat}_\gamma(x_{q(k+i)}) + \mathcal{C}_q w_{q(k+i):q(k+i+1)}||_2 - ||x_{q(k+i)}||_2 \\
&\leq ||x_{q(k+i)} - \operatorname*{sat}_\gamma(x_{q(k+i)})||_2 + ||\mathcal{C}_q w_{q(k+i):q(k+i+1)}||_2 - ||x_{q(k+i)}||_2 \\
&\leq ||\mathcal{C}_q w_{q(k+i):q(k+i+1)}||_2 - \gamma,
\end{aligned} \tag{4.4}$$

where in the first inequality we used the orthogonality of A and the fact that $\mathcal{B}_q\mathcal{B}_q^\dagger = I$ since $\mathcal{B}_q$ has full row rank. The second inequality follows from the definition of the $\mathrm{sat}_\gamma(\cdot)$ function. Next, by a similar argument, we have

$$X_{k+1} - X_k \leq ||\mathcal{B}_q u_{qk:q(k+1)}||_2 + ||\mathcal{C}_q w_{qk:q(k+1)}||_2 \leq \sqrt{m}||\mathcal{B}_q||_2 U_{\max} + ||\mathcal{C}_q w_{qk:q(k+1)}||_2,$$

and, of course, $X_k \leq \mathcal{J}$. Finally, again by orthogonality of A,

$$||\mathcal{C}_q w_{q(k+i):q(k+1+i)}||_2 = ||\sum_{j=0}^{q-1} A^{q-j-1} w_{q(k+i)+j}||_2 \leq \sum_{j=0}^{q-1} ||w_{q(k+i)+j}||_2, \tag{4.5}$$

for $i \geq 0$.

Putting it all together using the i.i.d. assumption and denoting $\alpha := \sqrt{m}||\mathcal{B}_q||_2 U_{\max}$ we get

$$\begin{aligned} P(X_n \geq t) &\leq \sum_{k=0}^{n-1} P\left(\sum_{j=1}^{(n-k)q} ||w_j||_2 \geq t - \mathcal{J} - \alpha + (n-k-1)\gamma\right) \\ &\leq \sum_{k=0}^{n-1} \frac{\mathbf{E}\left|\sum_{j=1}^{(n-k)q} \{||w_j||_2 - \mu\}\right|^p}{[t - \mathcal{J} - \alpha + (n-k-1)\gamma - (n-k)q\mu]^p} \end{aligned}$$

by Markov's inequality.

The Martingale $M_n := \sum_{i=0}^{n}(||w_i||_2 - \mu)$ can be bounded by Burkholder's inequality [11, 32] as follows. First, by Minkowski inequality,

$$\mathbf{E}|M_n|^p = \mathbf{E}|M_n - M_0 + M_0|^p \leq 2^p\mathbf{E}|M_n - M_0|^p + 2^p\mathbf{E}|M_0|^p = 2^p\mathbf{E}|M_n - M_0|^p + 2^p C_p',$$

for $C_p' := \mathbf{E}\,|||w_0||_2 - \mu|^p$. The process $M_n - M_0$ is now a Martingale null at zero, so that by Burkholder's and Minkowski inequalities

$$\begin{aligned} \mathbf{E}|M_n - M_0|^p &\leq c_b\mathbf{E}\left(\sum_{i=1}^{n}(M_i - M_{i-1})^2\right)^{p/2} \leq c_b\left(\sum_{i=1}^{n}(\mathbf{E}|M_i - M_{i-1}|^p)^{2/p}\right)^{p/2} \\ &\leq c_b n^{p/2} \max_{1\leq i\leq n} \mathbf{E}|M_i - M_{i-1}|^p. \end{aligned}$$

Since $\mathbf{E}|M_i - M_{i-1}| = \mathbf{E}\,|||w_i||_2 - \mu|^p = C_p'$ for all i, we arrive at the bound

$$\begin{aligned} P(X_n \geq t) &\leq \sum_{k=0}^{n-1} \frac{((n-k)q)^{p/2}2^p C_p' + 2^p C_p'}{[t - \mathcal{J} - \alpha + (n-k-1)\gamma - (n-k)q\mu]^p} \\ &\leq \sum_{l=0}^{n-1} \frac{(lq)^{p/2}2^p C_p' + 2^p C_p'}{[t - \mathcal{J} - \alpha - \gamma + l(\gamma - q\mu)]^p} \leq \sum_{l=0}^{\infty} \frac{(lq)^{p/2}2^p C_p' + 2^p C_p'}{[t - \mathcal{J} - \alpha - \gamma + l(\gamma - q\mu)]^p} \end{aligned}$$

for $t > \mathcal{J} + \alpha + \gamma$. Since γ was chosen such that $\gamma - q\mu > 0$ we get

$$\begin{aligned}\int_0^\infty t^{r-1} P(X_n \geq t)\,\mathrm{d}t &\leq b + \sum_{l=0}^{\infty} \int_{\mathcal{J}+\alpha+\gamma+1}^{\infty} t^{r-1} \frac{(lq)^{p/2} 2^p C_p' + 2^p C_p'}{[t - \mathcal{J} - \alpha - \gamma + l(\gamma - q\mu)]^p}\,\mathrm{d}t \\ &\approx \sum_{l=0}^{\infty} l^{(2r-p)/2},\end{aligned}$$

where $b = (\mathcal{J} + \alpha + \gamma + 1)^r / r$. This finishes the proof since the right-hand side is independent of n and finite for $r < (p-2)/2$. ■

Now we are ready to prove the main stability theorem by transforming the system matrix to the real Jordan form as in [30].

Theorem 4.2. *Let w_k be an i.i.d random sequence. Let further be the pair (A, B) stabilizable, the matrix A Lyapunov stable and $U_{\max} > q\mathbf{E}||w_0||_2 ||\mathcal{B}_q^\dagger A^q||_\infty$. Then there exists a causal feedback policy with $||u_k||_\infty < U_{\max}$ such that for the system $x_{k+1} = Ax_k + Bu_k + w_k$ holds*

I. the state x_k is mean-square bounded if $\mathbf{E}||w_0||_2^p < \infty$ for some $p > 6$,

II. there is a zero probability of a divergent trajectory if $\mathbf{E}||w_0||_2^p < \infty$ for some $p > 2$.

Proof. There exists a basis such that the system matrix is in the real Jordan form, and hence, by the stabilizability and Lyapunov stability assumption,

$$A = \begin{bmatrix} A_s & 0 \\ 0 & A_o \end{bmatrix}, \quad B = \begin{bmatrix} B_s \\ B_o \end{bmatrix},$$

where the matrix A_s is strictly stable, i.e., $\rho(A) < 1$, the matrix A_o is orthogonal, and the pair (A_o, B_o) is reachable. Thus the part of the state corresponding to the matrices (A_s, B_s) is mean-square bounded by Theorem 4.1, and the part of the state corresponding to the pair (A_o, B_o) can be stabilized by the previous Lemma 4.2 with $r = 2$. The zero probability of divergent trajectories follows from the same Lemma and Lemma 4.1. ■

Remark 4.1. *Results of [30] show that the first claim of Theorem 4.2 holds with $p = 4$. The argument is built on a fundamental result of [44], which is far beyond the scope of this text.*

4.3.1. Receding horizon stabilization

A closer look at the proof of the Lemma 4.2 reveals that it is not the particular form of the policy 4.3, but rather a negative drift condition that it ensures. The negative drift condition can be seen by taking the expectation on both sides of (4.4), which in view of (4.5) yields $\mathbf{E}(X_{k+i+1} - X_{k+i} \mid X_{k+i}) \leq q\mu - \gamma < 0$ on the event $[X_{k+i} > \mathcal{J}]$. This condition can be readily incorporated into an online optimization procedure even if a nonlinear disturbance feedback of the form $\eta + Ke(w)$ (see Eq. (3.5)) is employed since the preceding results guarantees feasibility with $K = 0$. To derive a stabilizing constraint observe that

$$\begin{aligned}\mathbf{E}(||x_{k+q}|| - ||x_k|| \mid x_k) &= \mathbf{E}(||A^q x_k + \mathcal{B}_q u_{k:k+q-1} + C_q w_{k:k+q-1}|| - ||x_k|| \mid x_k) \\ &\leq \mathbf{E}(||A^q x_k + \mathcal{B}_q u_{k:k+q-1}||_2 \mid x_k) + \mathbf{E}||C_q w_{k:k+q-1}||_2 - ||x_k||.\end{aligned}$$

Furthermore[1]

$$\begin{aligned}\mathbf{E}(||A^q x_k + \mathcal{B}_q u_{k:k+q-1}||_2 \mid x_k) &= \mathbf{E}||A^q x_k + B_q \eta_{1:qm} + K_{1:qm} e(w_{k:k+N-1})||_2 \\ &\leq ||A^q x_k + B_q \eta_{1:qm}||_2 + \mathbf{E}||B_q K_{1:qm} e(w_{k:k+N-1})||_2 \\ &\leq ||A^q x_k + B_q \eta_{1:qm}||_2 + \sqrt{n}||B_q K_{1:qm}||_\infty \mathbf{E}||e(w_{k:k+N-1})_{1:qn}||_\infty,\end{aligned}$$

where the term $\tilde{\varepsilon} := \mathbf{E}||e(w_{k:k+N-1})_{1:qn}||_\infty$ can be computed by means of Monte Carlo or upper bounded by $\varepsilon = ||e(w)||_\infty < \infty$. To guarantee mean-square stability we require that $\mathbf{E}(||x_{k+q}|| - ||x_k|| \mid x_k) \leq -\delta$ for some $\delta > 0$. Thus, considering that $\mathbf{E}||C_q w_{k:k+q-1}||_2 \leq q\mathbf{E}||w_0||_2$, the stabilizing constraint is

$$||A^q x_k + B_q \eta_{1:qm}||_2 + \sqrt{n}\tilde{\varepsilon}||B_q K_{1:qm}||_\infty - ||x_k|| \leq -q\mathbf{E}||w_0||_2 - \delta \tag{4.6}$$

whenever $||x_k||_2 > \mathcal{J}$, which is a second-order cone constraint.

A stabilizing receding horizon policy then employs a control horizon $N_c = q$ with the understanding that the stabilizing constraint (4.6) is triggered only if $||x_k||_2 > \mathcal{J}$.

4.3.2. Output feedback stabilization

The preceding approach immediately generalizes to output feedback stabilization if Gaussian disturbances are assumed for the process and measurement noise.

Assuming the linear stochastic system (3.38) and assumptions of Section 3.2.2, the conditional expectation is given by the Kalman filter equation

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + L_k \epsilon_{k|k-1}, \tag{4.7}$$

where $\hat{x}_{k|k-1} = \mathbf{E}(x_k \mid y_0, \dots, y_{k-1})$, $\epsilon_{k|k-1} = y_k - C\hat{x}_{k|k-1}$ and L_k is the Kalman gain. The innovations sequence ϵ_k is Gaussian and furthermore has bounded variance (and hence all moments) under the assumption of detectability of the pair (A, C), which also ensures that the estimate error $x_k - \hat{x}_{k|k-1}$ is mean-square bounded with variance given

[1] Here we abuse the notation and denote by $w_{k:k+q-1}$ the vector $[w_k^T, \dots, w_{k+q-1}^T]^T$ and similarly for $u_{k:k+q-1}$, but subscripts of the matrices K, η, $e(w)$ denote row selections.

by the limiting solution of the Riccati equation (3.42). Thus, Theorem 4.2 guarantees existence of a mean-square stabilizing policy for the estimator (4.7) whose state is accessible. This, however, ensures existence of a stabilizing feedback policy for the original system since

$$\begin{aligned}\mathbf{E}||x_k||_2^2 &= \mathbf{E}||x_k - \hat{x}_{k|k-1} + \hat{x}_{k|k-1}||_2^2 &\quad (4.8)\\ &\leq \mathbf{E}||x_k - \hat{x}_{k|k-1}||_2^2 + 2\mathbf{E}(||\hat{x}_{k|k-1}||_2||x_k - \hat{x}_{k|k-1}||_2) + \mathbf{E}||\hat{x}_{k|k-1}||_2^2 &\quad (4.9)\end{aligned}$$

with the first and third terms bounded according to the preceding discussion. Boundedness of the second term then follows by the Cauchy-Schwarz inequality as

$$2\mathbf{E}(||\hat{x}_{k|k-1}||_2||x_k - \hat{x}_{k|k-1}||_2) \leq 2\sqrt{\mathbf{E}||\hat{x}_{k|k-1}||_2^2}\sqrt{\mathbf{E}||x_k - \hat{x}_{k|k-1}||_2^2}.$$

Discussion on receding horizon stabilization of Section 4.3.1 now applies to both output feedback policies of Section 3.2.2.

4.3.3. Existence of a stabilizing Markov policy

The mean-square stabilizing policies arising from Theorem 4.2 are not Markov (i.e., require knowledge not only of the current but also of some of the preceding states) if the reachability index q is greater than one. A natural question to ask is whether there exists a stabilizing Markov control policy, that is, a policy for which the control action at time k depends only on the measured state x_k, and not on the previous states $x_{k-1}, x_{k-2}, \ldots, x_0$. First, we establish a general result of the existence of a stabilizing *randomized* Markov policy given a stabilizing non-Markov policy, and then we exploit the particular form of the policy from Theorem 4.2 to construct a *deterministic* Markov policy. Both policies constructed are, however, time-varying.

4.3.3.1. Randomized Markov policy

In this section, we will establish a general result for Markov control processes that, for any policy π, guarantees the existence of a Markov policy $\tilde{\pi}$ such that

$$P^{\tilde{\pi}}(x_{k+1} \in B \mid x_k) = P^{\pi}(x_{k+1} \in B \mid x_k), \quad k \in \mathbb{N}_0, \tag{4.10}$$

where the $P^{\tilde{\pi}}$ and P^{π} denote the probability measures under the policies $\tilde{\pi}$ and π, respectively.

We assume a Markov control process with state-space $\mathcal{X}$, action space $\mathcal{U}$ and transition kernel Q, that is, $P(x_{k+1} \in B|x_k, u_k) = Q(B \mid x_k, u_k)$ for any Borel set $\mathcal{B} \subset \mathcal{X}$, $x_k \in \mathcal{X}$ and $u_k \in \mathcal{U}$. A stochastic linear system with *independent* additive noise is clearly an example of a Markov control process. See [28] for a detailed treatment of optimal control of Markov control processes. History up to time t is denoted by

$h_t = (x_0, u_0, \dots, x_{t-1}, u_{t-1}, x_t)$, and the corresponding product space that h_t takes values in by $H_t = \mathcal{X} \times \mathcal{U} \times \dots \times \mathcal{X} \times \mathcal{U} \times \mathcal{X}$. Let further $\pi = \{\pi_t\}_{t=0}^{\infty}$ denote a general (possibly randomized and non-markov) control policy, that is, $\pi_t(\cdot|\cdot)$ is a stochastic kernel[2] on $\mathcal{U}$ given H_t, and, for a Borel set $C \subset \mathcal{U}$, $\pi(C|h_t) = P^{\pi}(u_t \in C|h_t)$. Marginal distribution (or law) of a random variable X is denote by $P_X(\cdot)$, and the conditional distribution of a random variable X given Y is denoted by $P_{X|Y}(\cdot|\cdot)$. All spaces considered are assumed to be Borel, so that regular conditional distributions exist and the measures $P^{\tilde{\pi}}$ and P^{π} in (4.10) can be constructed (see [28, 32] for details).

First observe that, given an initial distribution $P^{\pi}_{x_0}(B) = P^{\tilde{\pi}}_{x_0}(B) := P(x_0 \in B)$, equality (4.10) implies that the process $\{x_t\}_{t=0}^{\infty}$ has the same one-dimensional marginals under both policies π and $\tilde{\pi}$ (and hence if π is mean-square stabilizing, so is $\tilde{\pi}$). Indeed, the initial distributions are the same, and, assuming $P^{\pi}_{x_k} = P^{\tilde{\pi}}_{x_k}$ for some $k \in \mathbb{N}$,

$$\begin{aligned} P^{\pi}(x_{k+1} \in B) &= \int_{\mathcal{X}} P^{\pi}(x_{k+1} \in B|x_k = x)\, P^{\pi}_{x_k}(dx) \\ &= \int_{\mathcal{X}} P^{\tilde{\pi}}(x_{k+1} \in B|x_k = x)\, P^{\tilde{\pi}}_{x_k}(dx) = P^{\tilde{\pi}}(x_{k+1} \in B). \end{aligned}$$

Now we construct one particular $\tilde{\pi}$ satisfying (4.10). Note that the constructed policy may and in most cases will be randomized (i.e., $\tilde{\pi}_k(\cdot|x_k)$ may not be a Dirac measure). The policy is defined in such a way that

$$P^{\tilde{\pi}}(u_k \in C \mid x_k) = P^{\pi}(u_k \in C \mid x_k)$$

for every Borel set $C \subset \mathcal{U}$. This is satisfied if the stochastic kernels $\tilde{\pi}_k$, $k \in \mathbb{N}$ constituting π are defined as

$$\tilde{\pi}_k(C \mid x_k) := P^{\pi}(u_k \in C \mid x_k) = \mathbf{E}^{\pi}[\pi_k(C \mid h_k) \mid x_k] = \int_{H_k} \pi_k(C \mid h_k)\, P^{\pi}_{h_k|x_t}(dh_k|x_k). \tag{4.11}$$

If defined in this way, $\tilde{\pi}_t$ is clearly a stochastic kernel on $\mathcal{U}$ given $\mathcal{X}$, and consequently $\{\tilde{\pi}_t\}_{t=0}^{\infty}$ is a Markov control policy. It remains to show that (4.10) holds. This turns out to be a direct consequence of the following lemma.

Lemma 4.3. *Let $K(\cdot|\cdot)$ be a stochastic kernel on $\mathcal{Z}$ given $\mathcal{Y}$. Let further μ be a measure on $\mathcal{Y}$ and $f : \mathcal{Z} \to \mathbb{R}$ a nonnegative measurable function. Then*

$$\int_{\mathcal{Y}} \int_{\mathcal{Z}} f(z) K(dz|y)\mu(dy) = \int_{\mathcal{Z}} f(z)\nu(dz),$$

[2]A stochastic kernel $K(\cdot|\cdot)$ on $\mathcal{X}$ given $\mathcal{Y}$ satisfies the following two properties: $K(\cdot|y)$ is a probability measure on $\mathcal{X}$ for each $y \in \mathcal{Y}$ and $K(B|\cdot)$ is a measurable function on $\mathcal{Y}$ for every measurable $B \subset \mathcal{X}$. Note that the transition kernel Q is also a stochastic kernel.

where

$$\nu(C) = \int_{\mathcal{Y}} K(C|y)\mu(dy)$$

for any Borel set $C \subset \mathcal{Z}$.

Proof. The proof is standard, starting with indicators and then using approximation by simple functions. Thus, suppose first that $f = \mathbb{1}_A$ for a Borel set $A \subset \mathcal{Z}$. Then

$$\int_{\mathcal{Y}} \int_{\mathcal{Z}} \mathbb{1}_A K(dz|y)\mu(dy) = \int_{\mathcal{Y}} K(A|y)\mu(dy),$$

and

$$\int_{\mathcal{Z}} \mathbb{1}_A \nu(dz) = \nu(A) = \int_{\mathcal{Y}} K(A|y)\mu(dy),$$

which shows that the equality holds for indicator functions of measurable sets. The equality then extends by linearity to any simple function. Finally, approximation of f by simple functions and the monotone convergence theorem finish the proof. ∎

The lemma now immediately gives equality (4.10) as

$$\begin{aligned} P^{\pi}(x_{k+1} \in B|x_k) &= \int_{H_k} \int_{\mathcal{U}} Q(B|x_k, u_k)\pi_k(du_k|h_k)P^{\pi}_{h_k|x_k}(dh_k|x_k) \\ &= \int_{\mathcal{U}} Q(B|x_k, u_k)\tilde{\pi}_k(du_k|x_k) = P^{\tilde{\pi}}(x_{k+1} \in B|x_k), \end{aligned}$$

where the second equality follows from Lemma 4.3 individually for each value of x_k.

This establishes the existence of a mean-square stabilizing Markov control policy $\tilde{\pi}$ given *any* mean-square stabilizing policy π. Note that this result is purely theoretical since practical implementation of this policy requires the ability to sample from the distribution $P_{h_t|x_t}$, which is difficult even for the special case of a q-periodic policy arising from Theorem 4.2.

4.3.3.2. Deterministic Markov policy

The q-periodicity of the policy arising from Theorem 4.2 can be exploited to construct a deterministic (i.e., not randomized) Markov policy. A stabilizing q-periodic policy computes at times nq, $n \in \mathbb{N}_0$, a sequence of control inputs $u_{nq}, u_{nq+1}\ldots, u_{q(n+1)-1}$ ensuring that the negative drift condition

$$\mathbf{E}\{||x_{(n+1)q}||_2 \mid x_{nq}\} - ||x_{nq}||_2 \leq a \tag{4.12}$$

is satisfied for some $a > 0$ whenever $||x_{nq}||_2 > \mathcal{J}$. If $||x_{nq}||_2 \leq \mathcal{J}$, the control sequence can be arbitrary. The sequence is then applied in open loop to the system. This suggest that a simple q-periodic Markov policy can be constructed by solving at each time a stochastic optimal control problem minimizing the expectation of the 2-norm of the state at the next integer multiple of q over all open loop policies. Thus, define functions $\phi_0, \ldots, \phi_{q-1}$ as

$$\phi_k(x) := u_0^k(x), \quad k = 0, \ldots, q-1,$$

where $u_0^k(x)$ comes from the optimal solution $u_0^k(x), \ldots, u_{q-1-k}^k(x)$ to the problem

$$\begin{aligned} \underset{u_0^k, \ldots, u_{q-1-k}^k}{\text{minimize}} \quad & \mathbf{E}\{||x_{q-k}||_2 \mid x_0 = x\} \\ \text{subject to} \quad & x_{i+1} = Ax_i + Bu_i + w_i \\ & u_i \in \mathcal{U}, \; i = 0, \ldots, q-1-k \end{aligned} \tag{4.13}$$

A q-periodic deterministic Markov policy $\pi^d = \{\pi_k^d\}_{k=0}^{\infty}$ is then defined by successive concatenation of $\phi_0, \ldots, \phi_{q-1}$, that is,

$$\pi_k^d(x_k) := \phi_{k \bmod q}(x_k).$$

It is easy to see that this policy ensures no smaller negative drift than the original non-Markov one since the controlled process is Markov and, at each time $nq, \ldots, (n+1)q-1$, the remainder of the open-loop sequence that ensures (4.12) is feasible in (4.13). Consequently, the deterministic Markov policy π^d is mean-square stabilizing whenever the original non-Markov policy is.

The problem (4.13) may be relatively difficult to solve since there is typically no closed-form expression for the cost function. The problem is, however, small scale (the horizon length is at most equal to the reachability index), and therefore sampling techniques can be used to obtain good enough approximation of the cost (see Eq. (2.16)).

To demonstrate the deterministic Markov policy, we consider a 3-dimensional orthogonal system and an i.i.d. Gaussian disturbance sequence given by the matrices

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}, \quad B = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{E} w_k w_k^T = \begin{bmatrix} 0.0042 & 0 & 0 \\ 0 & 0.0042 & 0 \\ 0 & 0 & 0.0085 \end{bmatrix},$$

which is (up to a coordinate transformation and an asymptotically stable part) the same setting as in [47]. The control authority is bounded by $U_{\max} = 1$, and the saturation level is $\gamma = U_{\max}/||\mathcal{B}_3^\dagger A^3||_\infty = 0.1213$. The initial state is $x_0 = [-0.5, -0.5, -2]^T$. Simulation results over 1000 Monte Carlo runs shown in Figure 4.1 suggest that both the non-Markov policy arising form Theorem 4.2 and the deterministic Markov policy are indeed stabilizing, and that the latter performs slightly better in terms of the mean-square of the state.

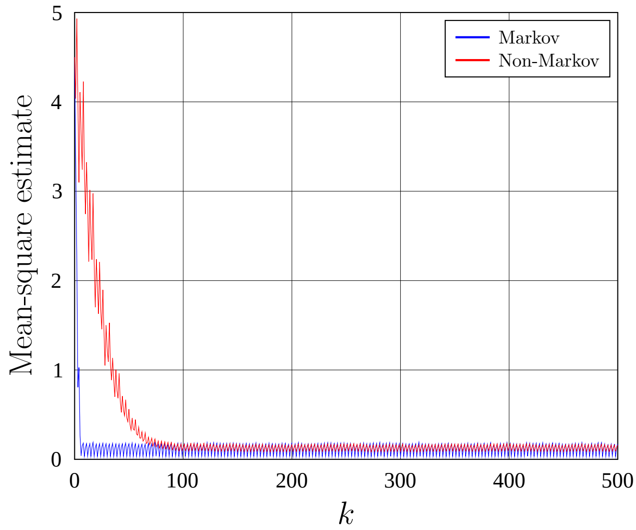

Figure 4.1.: Mean-square stabilization by Markov and non-Markov policies.

4.4. Marginally unstable case

Now we briefly discuss the marginally unstable case. There are no conclusive results even for the simplest case of a discrete double integrator although simulation results suggest that at least this system should be mean-square stabilizable. We derive at least a partial result on stabilizability of positive or negative part of both states of this system with arbitrarily small but nonzero control authority. We consider the system

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = A \begin{bmatrix} x_k \\ y_k \end{bmatrix} + Bu_k + \begin{bmatrix} w_k \\ v_k \end{bmatrix}, \tag{4.14}$$

where

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

The disturbance sequence $[v_k, w_k]$ is assumed to be independent with respect to time, and to have bounded p-th moment as $C_p = \sup_{k \geq 0} \max\{E|v_k|^p, E|w_k|^p\} < \infty$.

To derive a bound on positive parts of the state we define τ_n as the last time before n when both x_k and y_k were below some $\mathcal{J} \geq \max\{x_0, y_0\}$[3], i.e.,

$$\tau_n = \max\{k \leq n \mid \max\{x_k, y_k\} \leq \mathcal{J}\}.$$

[3] Here we could have chosen any $\mathcal{J} \in \mathbb{R}$ independently of x_0 and y_0. This would only allow for the possibility of $\tau = -\infty$, which is, however, not pathological in any way, only increases notational burden. All of the calculation below would actually remain valid with one more term in the decomposition according to the values of τ_n and with $\mathcal{J}$ replaced by $\max\{\mathcal{J}, x_0, y_0\}$.

Finally define the control policy as

$$u_k = \begin{cases} -U_{\max} & \max\{x_k, y_k\} \geq \mathcal{J} \\ 0 & \text{otherwise.} \end{cases} \tag{4.15}$$

Then for $t > \mathcal{J}$ we have

$$\begin{aligned}
&P(x_n \geq t) = \sum_{k=0}^{n-1} P(x_n \geq t, \tau_n = k) \leq \sum_{k=0}^{n-1} P(x_n > t \mid \tau_n = k) \\
&= \sum_{k=0}^{n-1} P\left(x_k + (n-k)y_k - U_{\max}\sum_{i=0}^{n-k-2} i + \sum_{i=0}^{n-k-1} iv_{n-i-1} + \sum_{i=0}^{n-k-1} w_{n-i-1} \geq t \mid \tau_n = k\right) \\
&\leq \sum_{k=0}^{n-1} P\left((n-k+1)\mathcal{J} - 0.5(n-k-1)(n-k-2)U_{\max} + \sum_{i=0}^{n-k-1}(iv_{n-i-1} + w_{n-i-1}) \geq t\right) \\
&\leq \sum_{k=0}^{n-1} \frac{\mathbf{E}\left|\sum_{i=0}^{n-k-1}(iv_{n-i-1} + w_{n-i-1})\right|^p}{(t - (n-k+1)\mathcal{J} + 0.5(n-k-1)(n-k-2)U_{\max})^p}
\end{aligned}$$

for $t \geq t_0$, where t_0 is chosen such that $t_0 - (n-k+1)\mathcal{J} + 0.5(n-k-1)(n-k-2)U_{\max})^p$ is positive for all k and n. Such a t_0 clearly exists for any $U_{\max} > 0$ and any $\mathcal{J} \in \mathbb{R}$.

Now observe that $\sum_{i=0}^{n-k-1}(iv_{n-i-1} + w_{n-i-1})$ is a sum of independent zero-mean random variables and hence a martingale. Thus by Burkohlder's and Minkowski inequalities we have

$$\begin{aligned}
\mathbf{E}\left|\sum_{i=0}^{n-k-1}(iv_{n-i-1} + w_{n-i-1})\right|^p &\leq c_p(n-k)^{p/2} \max_{0 \leq i \leq n-k-1} \mathbf{E}|iv_{n-i-1} + w_{n-i-1}|^p \\
&\leq 2^p c_p (n-k)^{p/2} \max_{0 \leq i \leq n-k-1} \{\max\{i^p \mathbf{E}|v_{n-i-1}|^p, \mathbf{E}|w_{n-i-1}|^p\}\} \\
&\leq c_p 2^p (n-k)^{3p/2} C_p,
\end{aligned}$$

where $c_p < \infty$ comes from Burkohlder's inequality. Thus we have arrived at the bound

$$\begin{aligned}
\mathbf{E}(x_n^+)^r &= \int_0^\infty P(x_n^r \geq t)\,\mathrm{d}t \leq t_0^r/r + \int_{t_0}^\infty t^{r-1} P(x_n \geq t)\,\mathrm{d}t \\
&\leq t_0^r/r + \int_{t_0}^\infty t^{r-1} \sum_{k=0}^{n-1} \frac{c_p 2^p (n-k)^{3p/2} C_p}{(t - (n-k+1)\mathcal{J} + 0.5(n-k-1)(n-k-2)U_{\max})^p}\,\mathrm{d}t \\
&\leq t_0^r/r + \sum_{l=1}^\infty \int_{t_0}^\infty t^{r-1} \frac{c_p 2^p C_p l^{3p/2}}{(t - (l+1)\mathcal{J} + 0.5(l-1)(l-2)U_{\max})^p}\,\mathrm{d}t \approx \sum_{l=0}^\infty \frac{l^{3p/2}}{l^{2(p-r)}},
\end{aligned}$$

which is finite for $p > 4r+2$. In particular, mean-square stability is ensured for $p > 10$. Mean-square boundedness of y_k^+ follows from Theorem 4.2 for any $p > 6$ or directly by an analogous computation.

This technique can be immediately extended to a Jordan block of arbitrary size and any matrix B such that the last row of B is nonzero.

On the other hand the problem of mean-square stabilizability of the state norm, not only positive parts is more difficult. In particular, a proof that an extension of the policy (4.15) also stabilizes the negative parts of the state seems to be far from trivial. There are at least some obvious modifications of the policy that look stabilizing in computer simulations. For example for the policy

$$u_k = \begin{cases} -U_{\max} & x_k > \mathcal{J}, y_k \geq -\mathcal{J} \\ U_{\max} & x_k < -\mathcal{J}, y_k \leq \mathcal{J} \\ 0 & \text{otherwise} \end{cases} \tag{4.16}$$

we have no proof of stability of either positive or negative parts, but computer simulations in Figures 4.2, 4.3 suggest that the policy is indeed stabilizing. Figure 4.4, by contrast, shows unstable behaviour when the $A(2,2)$ element of the matrix A is perturbed from 1 to 1.1.

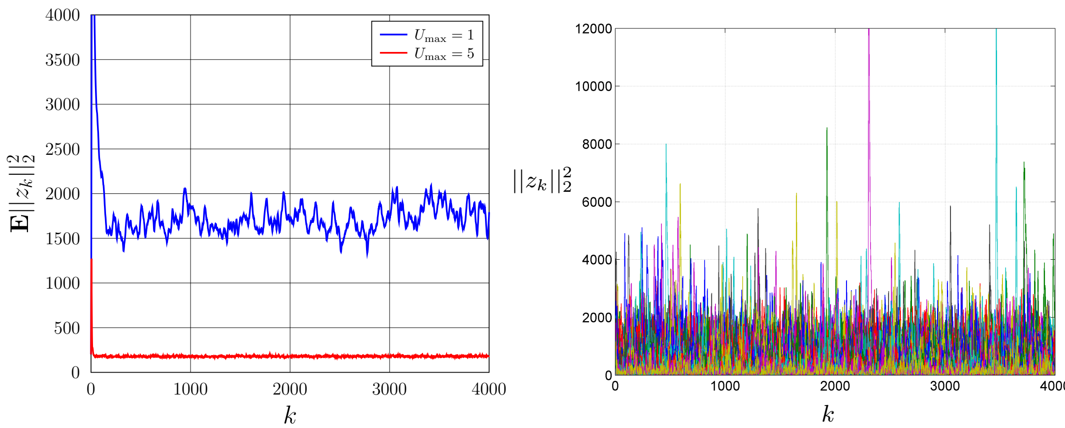

Figure 4.2.: Simulation over 2000 Monte Carlo runs using policy (4.16) with $\mathcal{J} = 10$ and $z_0 = [x_0, y_0]^T = [10, 10]^T$. Left: mean-square estimate for two values of $U_{\max}$. Right: individual trajectories for $U_{\max} = 5$.

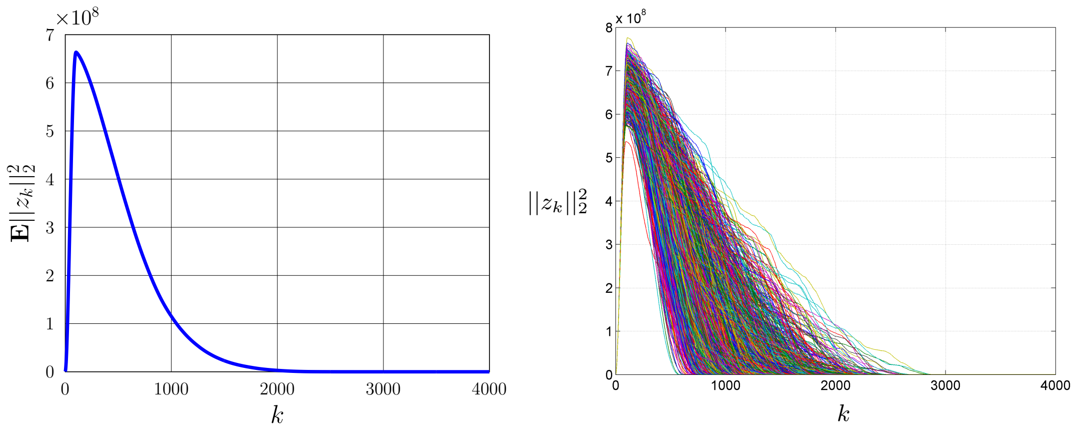

Figure 4.3.: Simulation results over 1000 Monte Carlo runs using policy (4.16) with $\mathcal{J} = 10$, $U_{\max} = 5$ and $z_0 = [x_0, y_0]^T = [500, 500]^T$. Left: mean-square estimate. Right: individual trajectories.

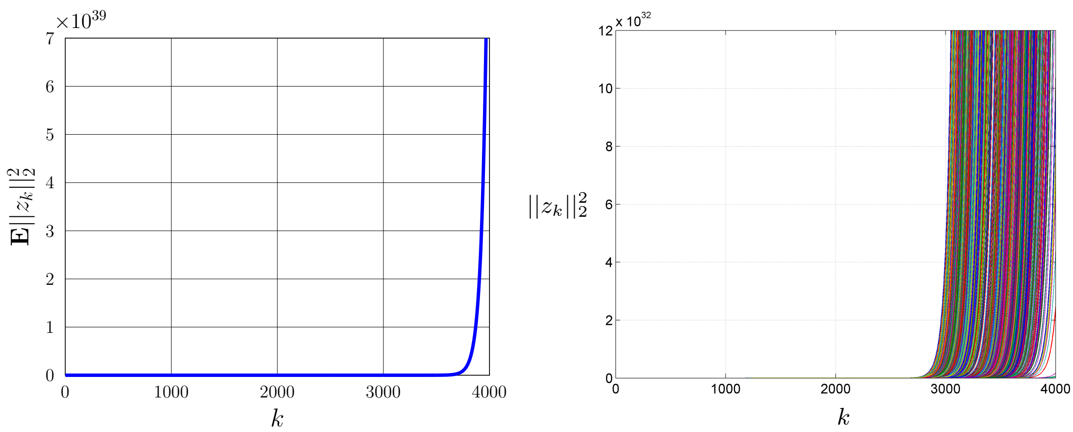

Figure 4.4.: Simulation results over 1000 Monte Carlo runs using policy (4.16) with $\mathcal{J} = 10$, $U_{\max} = 5$, $z_0 = [x_0, y_0]^T = [500, 500]^T$ and perturbed matrix A with $A(2,2) = 1.01$. Left: mean-square estimate. Right: individual trajectories.

5. Recursive feasibility via invariant sets

This chapter is concerned with developing techniques to ensure recursive feasibility of probabilistically-constrained stochastic model predictive control problems. First we give a brief overview of probabilistic (chance) constraints typically encountered in the context of stochastic model predictive control in Section 5.1, and then we move on to the development of recursively feasible algorithms in Section 5.2.

5.1. Chance constraints

In the presence of unbounded additive disturbances it is no longer possible[1] to satisfy hard state constraints of the form $f(x_k) \le 0$. Therefore, some kind of soft constraint has to be introduced, a natural and probably the most widely used choice being a probabilistic (or chance) constraint of the form

$$P(f(x_k) \le 0) \ge 1 - \alpha. \tag{5.1}$$

This is the sole type of constraints that will be addressed in this section, even though there are several other possibilities for hard constraints softening. For example constraints of the type

$$\mathbf{E} f(x_k) \le 0$$

or integrated chance constraints, where satisfaction of

$$\int_0^\infty P(f(x_k) > s)\,\mathrm{d}s \le \alpha$$

is required, often lead to a convex representation (see [23] for a survey).

For further analysis it is convenient to rewrite the chance constraint (5.1) in the form

$$P(f(w, \theta) \le 0) \ge 1 - \alpha, \tag{5.2}$$

where w is a random vector and θ is a vector of optimization variables. For a Gaussian random variable, this constraint turns out to be convex if the function f is affine individually in w and θ (i.e., f is bilinear). In particular, the constraint translates to

[1] Some mild technical assumptions are needed here. An obvious set of assumptions might require that the disturbance sequence is i.i.d. with $P(w_k \in G) > 0$ for every set open set G, the pair (A, F), where $\mathbf{E} w_k w_k^T = FF^T$, is reachable, and the set $\{x \mid f(x) > 0\}$ has nonempty interior.

an affine constraint if it is affine jointly in (w, θ) and a second-order cone constraint if it is affine individually in w and θ, but not jointly in (w, θ), which is shown in the following Lemma.

Lemma 5.1. *If* $w \sim \mathcal{N}(0, FF^T)$ *then*

$$P(a^T\theta + b^T w + \theta^T C w + d \le 0) \ge 1 - \alpha \tag{5.3}$$

is equivalent to

$$a^T\theta + d + \Phi^{-1}(1-\alpha)||F^T(b + C^T\theta)||_2 \le 0. \tag{5.4}$$

Proof.

$$\begin{aligned} P(a^T\theta + b^T w + \theta^T C w + d \le 0) &= P((b^T + \theta^T C)F\tilde{w} \le -a^T\theta - d) \\ &= \Phi\left(\frac{-a^T\theta - d}{||F^T(b + C^T\theta)||_2}\right), \end{aligned}$$

where $\tilde{w} \sim \mathcal{N}(0, I)$. The last equality follows from the fact that $(b^T + \theta^T C)F\tilde{w}$ is a zero-mean Gaussian random variable with the variance $||F^T(b + C^T\theta)||_2^2$.

The result now immediately follows. ∎

Note that the constraint (5.4) is a second-order cone constraint for $C \neq 0$ (and $\alpha < 0.5$) and an affine constraint for $C = 0$ in the optimization variable θ, which is in accordance with the preceding discussion.

If the distribution of w is not Gaussian, it is still possible to obtain a representation of the constraint (5.3) as

$$a^T\theta + d + \gamma \le 0, \tag{5.5}$$

where γ is the smallest value such that $P((b^T + \theta^T C)w \le \gamma) \ge 1 - \alpha$, i.e, $\gamma = F^{-1}_{(b^T+\theta^T C)w}(1-\alpha)$, where $F^{-1}_{(b^T+\theta^T C)w}(\cdot)$ is the left quantile function of $(b^T + \theta^T C)w$. The quantile γ can be obtained, for instance, by a Monte Carlo simulation, which leads to a tractable affine representation for $C = 0$ since then γ is independent of the optimization variable θ. This observation can be used to construct a recursively feasible stochastic MPC algorithm for a broad class of control policies, which is carried out in detail in Section 5.2.

5.1.1. Joint chance constraints

So far only individual chance constraints, i.e., constraints with $f(w, \theta)$ scalar, have been discussed. In practice, however, joint chance constraints, where $f(w, \theta)$ is a vector, are of equal importance. Unfortunately in this case an exact tractable representation does usually not exist even if the constraint is convex, since mere evaluation of the constraint requires a computation of a multivariate integral, which becomes prohibitive in higher dimensions.

The simplest way to approximate a joint chance constraint of the form (5.2), where $f(w,\theta) = \big[f_1(w,\theta),\dots,f_m(w,\theta)\big]^T$, is by employing the Boole's inequality to get an approximation by individual constraints as

$$P(f(w,\theta) \nleq 0) = P\left(\bigcup_{i=1}^{m} f_i(w,\theta) > 0\right) \leq \sum_{i=1}^{m} P(f_i(w,\theta) > 0). \tag{5.6}$$

Now if α_i, $i = 1,\dots,m$ are chosen such that

$$P(f_i(w,\theta) > 0) \leq \alpha_i, \quad i = 1,\dots,m \tag{5.7}$$

and $\sum_{i=1}^{m} \alpha_i = \alpha$, equations (5.7) clearly give a conservative approximation of the original constraint (5.2).

A typical choice for α_i's is $\alpha_i = \alpha/m$, $i = 1,\dots,m$. However, conservatism of this approximation can be significantly reduced if the α_i's are included in the optimization problem as free variables with the only constraint that they be nonnegative and sum to α, which is the idea of risk allocation introduced in [8]. In the setting of Lemma 5.1 this leads to a convex representation if $C = 0$ and $\alpha \leq 0.5$, since $\Phi^{-1}(x)$ is convex on $[0.5, 1]$.

The requirement that α does not exceed 0.5 is typically not a practical limitation, since for $\alpha > 0.5$ the expectation of $f(w,\theta)$ is allowed to be positive, which is a rare situation.

On the other hand, the loss of convexity for $C \neq 0$ pose a real problem and has to be tackled. In practice, the optimization variable θ can usually be partitioned as $\theta = [\theta_1^T, \theta_2^T]^T$ and the constraint (5.3) written as

$$P(a^T\theta_1 + b^T w + \theta_2^T C w + d \leq 0) \geq 1 - \alpha. \tag{5.8}$$

A simple alternating convex optimization scheme can now be adopted to approximately solve an optimization problem with a constraint of the form (5.8). First the problem is solved in variables (θ_1, θ_2) with α_i fixed to α/m, then solved with variables $(\theta_1, \alpha_1, \dots, \alpha_m)$ and θ_2 fixed to the corresponding part of the optimal solution of the previous problem, and then again solved in variables (θ_1, θ_2) with α_i fixed from the optimal solution of the previous problem, and so on. The constraint is now convex in each iteration of this scheme, and only one or two iterations are usually enough to get a solid improvement.

5.2. Strongly feasible stochastic MPC

In this section we develop a systematic approach to ensure recursive feasibility of stochastic model predictive control problems using invariant set techniques. The recursive feasibility will be enforced only through constraints, and consequently the resulting stochastic MPC problems will actually be strongly feasible (see Definition 5.1). Two approaches are developed; the first one employs a terminal constraint, whereas the second one constrains the first predicted state. The second approach turns out to be completely independent of the policy in question and moreover it is least-restrictive in the sense that it produces the largest feasible set amongst all admissible policies (see Definition 5.2).

We consider the linear time-invariant stochastic dynamic system

$$x_{k+1} = Ax_k + Bu_k + w_k,\ k \in \mathbb{N}_0$$

with the state $x_k \in \mathbb{R}^n$, the control $u_k \in \mathbb{R}^m$, and the i.i.d disturbance sequence $w_k \in \mathbb{R}^n$. It is assumed that the state x_k is known at time k for all $k \in \mathbb{N}_0$, and that the pair (A, B) is stabilizable.

Our aim is to develop a systematic approach to ensure that the closed-loop state trajectory satisfies the probabilistic constraints

$$P(g_j^T x_k \leq h_j) \geq 1 - \alpha_j,\ k \in \mathbb{N},\ j \in \mathbb{N}_1^r, \tag{5.9}$$

while minimizing some cost function and satisfying hard input constraints

$$u_k \in \mathcal{U} := \{u \in \mathbb{R}^m \mid ||u||_\infty \leq U_{\max}\},\ k \in \mathbb{N}_0. \tag{5.10}$$

The allowed probability of violation $\alpha_j \in [0, 1]$ typically comes directly from application requirements, but it can also be viewed as a tuning parameter tracing a trade-off curve between constraint violation and incurred cost.

The polyhedral intersection of the individual constraints $g_j^T x \leq h_j$ is referred to as the constraint set and denoted by

$$\mathcal{X} := \{x \in \mathbb{R}^n \mid Gx \leq h\}, \tag{5.11}$$

where g_j^T and h_j form the rows of $G \in \mathbb{R}^{r\times n}$ and $h \in \mathbb{R}^r$ respectively.

To ensure satisfaction of (5.9), it is sufficient to guarantee that

$$P(g_j^T x_{k+1} \leq h_j \mid x_k) \geq 1 - \alpha_j,\ j \in \mathbb{N}_1^r, \tag{5.12}$$

for all $k \in \mathbb{N}_0$. In the sequel, we will focus on developing techniques to render the constraint recursively feasible, that is, to guarantee its feasibility under a given control policy at each time $k \in \mathbb{N}_0$.

Remark 5.1. *The presented approach exhibits a certain degree of conservatism since satisfaction of (5.12) for all $k \in \mathbb{N}_0$ is only sufficient for (5.9). However, (5.12) offers a tractable and straightforwardly implementable condition, in contrast to (5.9), which generally cannot be exactly accommodated in a cost-minimization procedure.*

We let

$$\boldsymbol{u} := [u_0^T, \dots, u_{N-1}^T]^T, \quad \boldsymbol{w} := [w_0^T, \dots, w_{N-1}^T]^T$$

denote the predicted input sequence and the disturbance sequence along the horizon N, respectively.

The recursive feasibility is considered with respect to the affine disturbance feedback policy

$$\boldsymbol{u} = \eta + K\boldsymbol{w} = \begin{bmatrix} \eta_0 \\ \eta_1 \\ \vdots \\ \eta_{N-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ K_{1,1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \\ K_{N-1,1} & \dots & K_{N-1,N-1} & 0 \end{bmatrix} \boldsymbol{w}, \tag{5.13}$$

applied in a receding horizon fashion.

In the presence of independent unbounded disturbances additively entering the system, it is impossible to ensure recursive feasibility at all times. Hence we assume that the common distribution of the disturbance sequence w_k, $k \in \mathbb{N}_0$ is supported on the compact set

$$\mathcal{W} = \{w \in \mathbb{R}^n \mid ||w||_\infty \le \Delta < \infty\}, \tag{5.14}$$

and we let w denote any random variable having the same distribution as w_k, $k \in \mathbb{N}_0$.

Remark 5.2. *The presented approach can immediately be generalized to polytopically or quadratically bounded disturbances, still giving rise to tractable convex optimization problems. We, however do not, consider more general disturbance specifications for the sake of brevity.*

In what follows we consider the optimization problem

$$\begin{aligned} \underset{\eta,K}{\text{minimize}} \quad & J(\eta, K) := \mathbf{E}\left\{ ||Q_N x_N||_p^p + \sum_{k=0}^{N-1} ||Q_k x_k||_p^p + ||R_k u_k||_p^p \right\} \\ \text{subject to} \quad & \boldsymbol{u} = \eta + K\boldsymbol{w} \text{ structured as in (5.13)} \\ & P(g_j^T x_{k+1} \le h_j) \ge 1 - \alpha_j, \ j \in \mathbb{N}_1^r \\ & \boldsymbol{u} \in \mathcal{U}^N \\ & x_{k+1} = Ax_k + Bu_k + w_k \end{aligned} \tag{5.15}$$

Remark 5.3. *The particular form of the cost function J does not affect the theoretical discussion of this section, because here we are interested only in the feasibility properties of (5.15) which are independent of J. We, however, employ the above cost function in the example of Section 6.2.*

A receding horizon application of (5.15) gives rise to a time-invariant state-feedback control policy (in fact a whole family of them) $\pi(x) = \eta_0(x)$, where $\eta_0(x)$ can come from any feasible point of (5.15), (η, K), with the initial state $x_0 = x$. The corresponding closed-loop trajectory

$$x_{k+1} = Ax_k + B\pi(x_k) + w_k \tag{5.16}$$

then satisfies (5.12) for all $k \in \mathbb{N}_0$ as long as the problem (5.15) stays feasible at all times.

Remark 5.4. *Note that only the first-step ($k = 0$) probabilistic constraint in (5.15) is important for closed-loop satisfaction of (5.12). This is exploited in Section 5.3.2 to derive a least-restrictive formulation (see Definition 5.2).*

Our primary goal is to ensure strong feasibility of the problem (5.15), or its version with only the first-step probabilistic constraint.

Definition 5.1 (Strong feasibility [34])**.** *A stochastic model predictive control problem is said to be strongly feasible if for every feasible initial state the closed-loop trajectory remains feasible due to any admissible disturbance realization and any sequence of feasible control inputs generated in a receding horizon fashion.*

Our secondary goal is to derive a problem that is least-restrictive.

Definition 5.2 (Least restrictiveness)**.** *A stochastic model predictive control problem is said to be least-restrictive if it is strongly feasible and there is no initial state x_0 outside its feasible set and no policy satisfying the input constraints such that the closed-loop trajectory generated by the policy, starting from x_0, satisfies the probabilistic constraint (5.12) for all $k \in \mathbb{N}_0$ and for all admissible disturbance realizations $\{w_k \in \mathcal{W}\}_{k=0}^{\infty}$.*

5.3. Main results

Two approaches to assert recursive feasibility are presented, both of which are based on robust controlled invariant sets that have become a standard tool in receding horizon control [9]. We begin with the definition of the feasibility region of the constraint (5.12), which plays a crucial role in what follows.

Definition 5.3 (Stochastic feasibility set)**.** *The stochastic feasibility set of the constraint (5.12) is*

$$\mathcal{X}_f := \{x \mid \exists u \in \mathcal{U} \text{ s.t. } P(g_j^T(Ax + Bu + w) \le h_j) \ge 1 - \alpha_j \ \forall\, j \in \mathbb{N}_1^r\}.$$

Being a projection of a polyhedron, this set is polyhedral as well. Indeed, we have

$$\begin{aligned}\mathcal{X}_f &= \{x \mid \exists u \in \mathcal{U} \text{ s.t. } F_{g_j^T w}(h_j - g_j^T(Ax + Bu)) \ge 1 - \alpha_j \ \forall\, j \in \mathbb{N}_1^r\} \\ &= \{x \mid \exists u \in \mathcal{U} \text{ s.t. } g_j^T(Ax + Bu) \le h_j - F_{g_j^T w}^{-1}(1 - \alpha_j) \ \forall\, j \in \mathbb{N}_1^r\},\end{aligned} \tag{5.17}$$

where $F_{g_j^T w}(\cdot)$ and $F^{-1}_{g_j^T w}(\cdot)$ are respectively the cumulative distribution and left quantile function of $g_j^T w$.

Remark 5.5. *The quantiles $F^{-1}_{g_j^T w}(1-\alpha_j)$, $j \in \mathbb{N}_1^r$ are the only quantities that need to be computed before standard algorithms for the construction of invariant sets can be employed. They can be computed offline to virtually arbitrary precision for any reasonable distribution of w, for instance, by means of Monte Carlo techniques.*

Note also that the stochastic feasible set $\mathcal{X}_f$ is, in general, neither a subset nor a superset of the constraint set $\mathcal{X}$ (see numerical example).

5.3.1. Terminal constraint

First we adopt a dual mode paradigm where the affine disturbance feedback policy (5.13) is used for predictions in mode 1, that is, at times $k = 0, \dots, N-1$, and any stabilizing state feedback in mode 2, that is, at times $k \geq N$ [37].

In mode 1 we have, given x_k,

$$P(g_j^T x_{k+1} \leq h_j \mid x_k) = P(g_j^T (Ax_k + Bu_k + w_k) \leq h_j \mid x_k),\ j \in \mathbb{N}_1^r.$$

Thus to ensure satisfaction of (5.12) we require that

$$g_j^T (Ax_k + Bu_k) \leq h_j - F^{-1}_{g_j^T w}(1-\alpha_j),\ j \in \mathbb{N}_1^r,\ k \in \mathbb{N}_0^{N-1}$$

for all possible states x_k reachable at time k by the disturbance sequence up to this time, $w_0^{k-1} := (w_0, \dots, w_{k-1})$, under a given policy in mode 1. Employing the definition of the affine disturbance feedback (5.13) on the left-hand side, we get

$$\begin{aligned} g_j^T (Ax_k + Bu_k) &= g_j^T [A(A^k x_0 + \mathcal{B}_k(\eta + K\boldsymbol{w}) + C_k \boldsymbol{w}) + B(\eta_k + K_k \boldsymbol{w})] \\ &= g_j^T (A^{k+1} x_0 + \mathcal{B}_{k+1}\eta) + g_j^T (\mathcal{B}_{k+1} K + A\mathcal{C}_k)\boldsymbol{w}, \end{aligned}$$

where

$$\mathcal{B}_k = [A^{k-1}B, \dots, B, 0, \dots, 0], \quad \mathcal{C}_k = [A^{k-1}, \dots, I, 0, \dots, 0],$$

and η_k and K_k denote k-th block rows of size m. Thus, considering the worst-case value of the uncertain term over all disturbances,

$$\max_{\boldsymbol{w} \in \mathcal{W}^N} g_j^T (\mathcal{B}_{k+1} K + A\mathcal{C}_k)\boldsymbol{w} = ||g_j^T (\mathcal{B}_{k+1} K + A\mathcal{C}_k)||_\infty \Delta, \tag{5.18}$$

we get a sufficient condition for recursive feasibility in mode 1

$$g_j^T (A^{k+1} x_0 + \mathcal{B}_{k+1}\eta) \leq h_j - ||g_j^T (\mathcal{B}_{k+1} K + A\mathcal{C}_k)||_\infty \Delta - F^{-1}_{g_j^T w}(1-\alpha_j), \tag{5.19}$$

for all $k \in \mathbb{N}_0^{N-1}$ and $j \in \mathbb{N}_1^r$.

Remark 5.6. *Although there is the disturbance sequence over the whole horizon* $\boldsymbol{w}$ *in the above computation, only the disturbances* w_0^{k-1} *contribute to the worst-case value due to the structure of the matrices* $\mathcal{B}_k$ *and* $\mathcal{C}_k$*.*

In mode 2 we use a stabilizing state feedback $u_k = K_s x_k$ with the corresponding strictly stable feedback dynamics matrix $\hat{A} = A + BK_s$. One step predictions in mode 2 now read

$$g_j^T x_{N+i+1} = g_j^T(\hat{A}^{i+1}\hat{x}_N + \hat{A}^{i+1}(\mathcal{B}_N K + \mathcal{C}_N)w_0^{N-1} + \hat{A}\hat{\mathcal{C}}_k w_N^{i-1} + w_{N+i}),$$

where $\hat{x}_N = A^N x_0 + \mathcal{B}_N \eta$ and $\hat{\mathcal{C}}_i = [\hat{A}^{i-1}, \hat{A}^{i-2}, \dots, I]$. Thus considering worst case values over w_0^{N+i-1} we get a sufficient condition

$$\begin{aligned} g_j^T \hat{A}^{i+1}(A^N x_0 + \mathcal{B}_N \eta) \leq h_j - ||g_j^T \hat{A}^{i+1}(\mathcal{B}_N K + \mathcal{C}_N)||_\infty \Delta \\ - ||g_j^T \hat{A}\hat{\mathcal{C}}_i||_\infty \Delta - F_{g_j^T w}^{-1}(1-\alpha_j) \end{aligned} \tag{5.20}$$

for all $j \in \mathbb{N}_1^r$ and $i \in \mathbb{N}_0$.

This infinite number of constraints can be expressed in terms of the maximum robust invariant subset of the feasibility region with respect to the closed-loop dynamics $x_{k+1} = (A + BK_s)x_k + w_k$, hard input constraints $||K_s x_k|| \in \mathcal{U}$ and the chance constraint $P(g_j^T(A + BK_s)x_k + w_k \leq h_j) \geq 1 - \alpha_j$. In other words, we a employ a set $\mathcal{X}_r^{K_s} \subset \mathcal{X}_f$ such that for all $x \in \mathcal{X}_r^{K_s}$

$$\begin{aligned} (A + BK_s)x + w &\in \mathcal{X}_r^{K_s}, \\ ||K_s x||_\infty &\leq U_{\max}, \\ g_j^T(A + BK_s)x &\leq h_j - F_{g_j^T w}^{-1}(1-\alpha_j) \ \forall w \in \mathcal{W} \ \forall j \in \mathbb{N}_1^r. \end{aligned} \tag{5.21}$$

It is assumed that the set is polyhedral and nonempty in the form $\mathcal{X}_r^{K_s} = \{x \in \mathbb{R}^n \mid Sx \leq z\}$. If the set were nonempty but not polyhedral, an inner approximation that is polyhedral can always be constructed. See Appendix B and [9, 27] for some algorithms to construct (controlled) invariant sets or their polyhedral approximations.

Recursive feasibility is now ensured by the requirement that the state x_N lands robustly inside $\mathcal{X}_r^{K_s}$, that is,

$$A^N x_0 + \mathcal{B}_N \eta + (\mathcal{B}_N K + \mathcal{C}_N)\boldsymbol{w} \in \mathcal{X}_r^{K_s} \quad \forall \boldsymbol{w} \in \mathcal{W}^N,$$

which is equivalent to

$$s_j^T(A^N x_0 + \mathcal{B}_N \eta) \leq z_j - ||s_j^T(\mathcal{B}_N K + \mathcal{C}_N)||_\infty \Delta \quad \forall j \in \mathbb{N}_1^{r'}, \tag{5.22}$$

where s_j^T and z_j are the rows of the matrices $S \in \mathbb{R}^{r' \times n}$ and $z \in \mathbb{R}^{r'}$ defining $\mathcal{X}_r^{K_s}$.

Hard input constraints are enforced explicitly in mode 1 as

$$|\eta_i| + \Delta||K_i||_\infty \leq U_{\max},\ i = 1, \dots, mN, \tag{5.23}$$

and implicitly in mode 2 through the relation (5.21). Here the subscript i denotes i-th row (not block row) of the corresponding matrix.

Hence we have arrived at the following theorem.

Theorem 5.1. *For the stochastic model predictive control problem*

$$\begin{aligned} \underset{\eta, K}{\text{minimize}} \quad & J(\eta, K) \\ \text{subject to} \quad & \boldsymbol{u} = \eta + K\boldsymbol{w} \text{ structured as in (5.13)} \\ & x_{k+1} = Ax_k + Bu_k + w_k,\ k \in \mathbb{N}_0^{N-1} \\ & (5.19),\ (5.22),\ (5.23) \end{aligned} \tag{5.24}$$

the following holds:

I. The problem is strongly feasible.

II. The closed-loop state trajectory satisfies (5.9) and (5.12).

Proof. I. Given any feasible solution (η, K) (structured as in (5.13)) at time zero, we are guaranteed to have a feasible point $(\tilde{\eta}, \tilde{K})$ at time one with $(\tilde{\eta}, \tilde{K})$ constructed as

$$\tilde{\eta} = \begin{bmatrix} \eta_1 + K_{1,1}w_0 \\ \eta_2 + K_{2,1}w_0 \\ \vdots \\ \eta_{N-1} + K_{N-1,1}w_0 \\ \eta_L \end{bmatrix}, \quad \tilde{K} = \begin{bmatrix} 0 & 0 \\ \hat{K} & 0 \\ K_L & 0 \end{bmatrix},$$

where

$$\hat{K} = \begin{bmatrix} K_{2,2} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \\ K_{N-1,2} & \dots & K_{N-1,N-1} & 0 \end{bmatrix}.$$

The last block rows η_L and K_L can be determined from the fact that $K_s x_N$ defines a feasible input at time N provided that the previous inputs were generated by the policy $\boldsymbol{u} = \eta + K\boldsymbol{w}$. Thus for the last block rows we have

$$\eta_L = K_s(A^N x_0 + \mathcal{B}_N \eta) + K_s[(\mathcal{B}_N K + \mathcal{C}_N)]_{1:n} \cdot w_0$$

$$K_L = K_s[\mathcal{B}_N K + \mathcal{C}_N]_{n+1:nN},$$

where $[A]_{p:q}$ denotes the sub-matrix of a matrix A consisting of columns p through q. Strong feasibility now follows by induction.

II. Satisfaction of probabilistic constraints on the state at the next time instant using any feasible input is ensured by (5.19) with $k = 0$. Hence the constraints (5.9) and (5.12) are satisfied if the problem is strongly feasible, which is guaranteed by I.

■

5.3.2. First-step constraint

An alternative approach to assert strong feasibility is to constrain at each time step only the predicted state at the very next time instant to a certain invariant set, in our case the maximum stochastic robust controlled invariant set (see Definition 5.5). This type of technique was recently introduced in the context of nominal as well as robust model predictive control [26, 41].

Definition 5.4. *A set $\mathcal{X}_{rc} \subset \mathbb{R}^n$ is a stochastic robust controlled invariant set if it satisfies the following condition:*

$$\begin{aligned} \forall\, x \in \mathcal{X}_{rc}\ \exists\, u \in \mathcal{U} \text{ s.t.: }\ & Ax + Bu + w \in \mathcal{X}_{rc} \quad \forall\, w \in \mathcal{W}, \\ & P(g_j^T(Ax + Bu + w) \leq h_j) \geq 1 - \alpha_j\ \forall\, j \in \mathbb{N}_1^r. \end{aligned} \tag{5.25}$$

Definition 5.5 (MSRCI set)**.** *The maximum stochastic robust controlled invariant set (MSRCI) is the largest set $\mathcal{X}_{rc}^* \subset \mathbb{R}^n$ that is stochastic robust controlled invariant according to Definition 5.4. The MSRCI set can be explicitly defined as*

$$\begin{aligned} \mathcal{X}_{rc}^* = \Big\{ x_0 \in \mathcal{X}_f \mid \exists\, \phi : \mathbb{R}^n \to \mathbb{R}^m \text{ s.t. } & x_{k+1} = Ax_k + B\phi(x_k) + w_k, \\ & P(g_j^T Ax_k + B\phi(x_k) + w \leq h_j) \geq 1 - \alpha_j, \\ & \phi(x_k) \in \mathcal{U} \quad \forall\, j \in \mathbb{N}_1^r\ \forall\, k \in \mathbb{N}_0\ \forall \{w_k \in \mathcal{W}\}_{k=0}^{\infty} \Big\}. \end{aligned} \tag{5.26}$$

Remark 5.7. *It is clear that the MSRCI set $\mathcal{X}_{rc}^*$ is a superset of the maximum robust controlled invariant subset of $\mathcal{X}$ for any choice of $\alpha_j \in [0, 1]$. However, it is not true in general that the two sets coincide when $\alpha_j = 0$ for all j, but rather $\mathcal{X}_{rc}^*$ is then equal to the set from which the maximum robust controlled invariant subset of $\mathcal{X}$ can be reached in one step. This implies that those states in $\mathcal{X}_{rc}^*$ (for $\alpha_j = 0$) that are not in the maximum robust controlled invariant subset of $\mathcal{X}$ must be outside $\mathcal{X}$. See the numerical example in Section 6.2.*

Using the same argument as with the stochastic feasibility set (5.17), the stochastic robust controlled invariance condition (5.25) can be expressed as

$$\begin{aligned} \forall\, x \in \mathcal{X}_{rc}\ \exists\, u \in \mathcal{U} \text{ s.t.: }\ & Ax + Bu + w \in \mathcal{X}_{rc} \quad \forall\, w \in \mathcal{W}, \\ & g_j^T(Ax + Bu) \leq h_j - F_{g_j^T w}^{-1}(1 - \alpha_j)\ \forall\, j \in \mathbb{N}_1^r, \end{aligned}$$

which shows that these sets can be determined by standard algorithms for construction of (maximum) robust controlled invariant sets. Consequently, all of the results for maximum robust controlled invariant sets hold. In particular the set can be expressed as an intersection of possibly infinite number of polyhedra. Hence, the set is convex, and moreover if $\mathcal{X}_f$ is compact, so is $\mathcal{X}^*_{rc}$ [9].

Again, it is assumed that the MSRCI set $\mathcal{X}^*_{rc}$ is polyhedral and nonempty in the form $\mathcal{X}^*_{rc} = \{x \mid \tilde{S}x \leq \tilde{z}\}$. If the set were nonempty but not polyhedral, an inner approximation that is stochastic robust controlled invariant and polyhedral can always be constructed (see Appendix B and [27]). This approximation is no longer maximum, rendering the problem more restrictive than the original one, yet still strongly feasible.

Thus, given the initial state x_0, the only necessary constraints to ensure strong feasibility and satisfaction of the probabilistic constraints are

$$
\begin{gathered}
Ax_0 + B\eta_0 + w \in \mathcal{X}^*_{rc} \;\; \forall\, w \in \mathcal{W}, \\
P(g_j^T(Ax_0 + B\eta_0 + w) \leq h_j) \geq 1 - \alpha_j \; \forall\, j \in \mathbb{N}_1^r, \\
\eta_0 \in \mathcal{U},
\end{gathered}
$$

which translates to

$$\tilde{s}_j^T(Ax_0 + B\eta_0) \leq \tilde{z}_j - ||\tilde{s}_j^T||_\infty \Delta \;\; \forall\, j \in \mathbb{N}_1^{\tilde{r}}, \tag{5.27}$$

$$g_j^T(Ax_0 + B\eta_0) \leq h_j - F^{-1}_{g_j^T w}(1 - \alpha_j) \;\; \forall\, j \in \mathbb{N}_1^r, \tag{5.28}$$

$$||\eta_0||_\infty \leq U_{\max}, \tag{5.29}$$

where $\tilde{s}_j^T$ and $\tilde{z}_j$ are the rows of the matrices $\tilde{S} \in \mathbb{R}^{\tilde{r} \times n}$ and $\tilde{z} \in \mathbb{R}^{\tilde{r}}$ defining $\mathcal{X}^*_{rc}$.

The following theorem now summarizes these observations.

Theorem 5.2. *For the stochastic model predictive control problem*

$$
\begin{aligned}
\underset{\eta, K}{\text{minimize}} \quad & J(\eta, K) \\
\text{subject to} \quad & \boldsymbol{u} = \eta + K\boldsymbol{w} \textit{ structured as in (5.13)} \\
& x_{k+1} = Ax_k + Bu_k + w_k, \; k \in \mathbb{N}_0^{N-1} \\
& (5.27),\; (5.28),\; (5.29)
\end{aligned} \tag{5.30}
$$

the following holds:

I. The problem is strongly feasible.

II. The problem is least-restrictive with the feasibility set equal to the associated MSRCI set $\mathcal{X}^*_{rc}$.

III. The closed-loop state trajectory satisfies the probabilistic constraints (5.9) and (5.12).

Proof. I. Given constraints (5.27), (5.28), (5.29), strong feasibility follows immediately by construction of the MSRCI set as follows. Given initial state $x_0 \in \mathcal{X}_{rc}^*$ and *any* feasible point (η, K) at time zero, the constraint (5.27) guarantees that the state at the next time instant stays robustly in $\mathcal{X}_{rc}^*$ after application of the first control move η_0. The result now follows by induction.

II. The least-restrictivness follows from the equivalent characterization of $\mathcal{X}_{rc}^*$ (5.26). The fact that $\mathcal{X}_{rc}^*$ is the feasible set of the problem is clear from the maximality of the MSRCI and the problem constraints.

III. Satisfaction of the probabilistic constraint (5.12) (and hence (5.9)) along the horizon follows from the constraint (5.28) and from strong feasibility.

∎

5.3.2.1. Mode 1 constraints

Theorem 5.2 tells us that if the stochastic *maximum* robust controlled invariant set is employed, the problem is feasible at time zero (and then by induction at all times) if and only if $x_0 \in \mathcal{X}_{rc}^*$. Even though constraints (5.27), (5.28) and (5.29) are sufficient, it may be beneficial for the sake of cost minimization to also include the mode 1 constraints (5.19) and (5.23). This can, however, unnecessarily reduce the size of the feasible set. Indeed, the additional constraints employ explicitly the affine disturbance feedback policy, whereas $\mathcal{X}_{rc}^*$ is maximum with respect to all policies. A remedy proposed in [41, 26] is to relax the additional constraints in a minimal way such that the feasible set remains unchanged. This amounts to replacing (5.19) and (5.23) with

$$g_j^T(A^{k+1}x_0 + \mathcal{B}_{k+1}\eta) \le h_j - ||g_j^T A(\mathcal{B}_k K + \mathcal{C}_k)||_\infty \Delta - F^{-1}_{g_j^T w_k}(1-\alpha_j) + \zeta_j, \tag{5.31}$$

and

$$|\eta_i| + \Delta ||K_i||_\infty \le U_{\max} + \xi_i \tag{5.32}$$

where $\zeta = [\zeta_1, \dots, \zeta_r]^T$ and $\xi = [\xi_1, \dots, \xi_{mN}]^T$ are minimal in some sense (e.g., in the 2-norm) such that the feasible set does not shrink. It is shown in [26] that computation of such a minimal relaxation gives rise to a convex problem where enumeration of all vertices of $\mathcal{X}_{rc}^*$ is necessary, which can quickly become prohibitive in larger dimensions. If this is the case, one can, however, always resort to a soft relaxation, that is, to keep ζ and ξ as optimization variables, and add regularization terms to the cost. If the 2-norms of ζ and ξ are of interest, this approach leads to the cost of the form

$$\tilde{J}(\eta, K) = J(\eta, K) + \gamma_1 ||\zeta||_2^2 + \gamma_2 ||\xi||_2^2 \tag{5.33}$$

with some positive γ_1 and γ_2.

5.3.2.2. Structural constraints

The second approach is particularly useful when additional structure is imposed on the matrix K and/or η in order to reduce the number of decision variables, and consequently the computational burden. A typical structure of the matrix K might be block-banded, i.e., allowing only a limited recourse via the disturbance sequence in the sense that $K_{i,j} = 0$ for $j < i - j_0$ for some fixed $j_0 \geq 0$. Another viable structure is a diagonal one for which $K_{i,j} = K_{i+1,j+1}$. See [41] for a comparison of various blocking strategies in the context of robust model predictive control.

It can be seen from the proof of Theorem 5.1 that this additional structural constraint cannot be easily accommodated within the first approach. On the other hand, the MSRCI set in the second approach, and hence its feasibility properties, remain completely unaffected as long as the first control move is free. This is a major advantage of the second approach since it allows for a trade-off between performance and complexity of the resulting problem while retaining (least-restrictive) strong feasibility. This is in fact one of the motivations behind the results of [26, 27] in the context of the standard move-blocking strategies widely employed in receding horizon control.

5.3.3. Nonlinear feedback

In this section we outline how to extend both of the presented approaches to the case of a nonlinear disturbance feedback of the form

$$\boldsymbol{u} = \eta + Ke(\boldsymbol{w}), \; ||e(\boldsymbol{w})||_\infty \leq \varepsilon. \tag{5.34}$$

First note that the second approach is completely independent of the policy in question as long as the additional mode 1 constraints (5.19) and (5.23) are not used. In fact, the MSRCI set depends only on the probabilistic constraints (5.9), the disturbance set $\mathcal{W}$ and the set of admissible controls $\mathcal{U}$.

When the nonlinear disturbance feedback is employed in the framework of the first approach or when the mode 1 constraints are included in the second approach, all of the calculations in Section 5.3.1 remain valid with Δ replaced by ε in (5.23) and (5.27), and an appropriate change in (5.18) and (5.22) according to the kind of nonlinear function

$e(\boldsymbol{w})$ employed. The worst-case value in (5.18) can be upper-bounded as

$$\max_{\boldsymbol{w}\in\mathcal{W}^N} g_j^T(\mathcal{B}_{k+1}Ke(\boldsymbol{w}) + A\mathcal{C}_k(\boldsymbol{w})) \leq \tag{5.35}$$

$$\leq \max_{\boldsymbol{w}\in\mathcal{W}^N} |g_j^T(\mathcal{B}_{k+1}Ke(\boldsymbol{w}) - \mathcal{B}_{k+1}K\boldsymbol{w})| + \max_{\boldsymbol{w}\in\mathcal{W}^N} |g_j^T(\mathcal{B}_{k+1}K + A\mathcal{C}_k)\boldsymbol{w}|$$

$$\leq ||g_j^T\mathcal{B}_{k+1}K||_\infty \sup_{\boldsymbol{w}\in\mathcal{W}^N} ||\boldsymbol{w} - e(\boldsymbol{w})||_\infty + ||g_j^T(\mathcal{B}_{k+1}K + A\mathcal{C}_k)||_\infty\Delta. \tag{5.36}$$

The only unknown term in the above expression is the supremum, which can, however, be easily evaluated exactly or upper-bounded for the most common choices of the nonlinear function $e(\boldsymbol{w})$. For instance, if $e(\boldsymbol{w})$ is a componentwise saturation with $\varepsilon = ||e(\boldsymbol{w})||_\infty < \Delta$ we get

$$\sup_{\boldsymbol{w}\in\mathcal{W}^N} ||\boldsymbol{w} - e(\boldsymbol{w})||_\infty = \Delta - \varepsilon.$$

Thus, both approaches readily extend to the case of a nonlinear disturbance feedback.

5.3.4. Discussion

Theorem 5.2 states that the feasible set of the second approach is maximal amongst all admissible policies. Thus the feasible set of the first formulation is necessarily a subset of, or equal to, the feasible set of the second one. On the other hand, construction of the robust invariant set with respect to the linear state feedback in the first approach is substantially less computationally demanding than the construction of the maximum robust controlled invariant set in the second problem. At this point it should, however, be emphasized that both sets are computed offline.

In conclusion we note that the recursively feasible chance constraint (5.9) translates to *affine* constraints on η and K regardless of the disturbance distribution, which is in sharp contrast to the traditional 'open-loop' chance constraints that lead to second-order-cone constraints for Gaussian disturbances (with the affine disturbance feedback) and have typically no exact representation otherwise [39, 1]. This is not completely unexpected since in (5.12) the stochastic nature of the problem comes into play at the last step only (from k to $k + 1$), whereas all of the previous disturbances have to be treated robustly.

6. Numerical implementation and examples

6.1. Numerical implementation

There are several aspects of the algorithms presented in Chapter 3 that may require special care in a practical implementation. First of all, the affine-like policies the Chapter is dealing with may lead to considerably larger optimization problems than the traditional certainty-equivalent model predictive control gives rise to, limiting the class of potential applications. This can, however, be ameliorated by exploiting the problem structure and by a sensible choice of optimization tools. A closely related problem is the fact that although convex, the problems encountered are not amenable to be solved by standard cone solvers (e.g. Sedumi [55]) due to the special functions involved. The last hindrance is the potential nondifferentiability of the cost function, which arises from the problem formulation itself since the p-norm is nondifferentiable. However, this is a place where the stochastic nature of the problem is actually helpful, rendering the problem smooth through the action of the expectation operator.

The problem of large size is briefly discussed first. Throughout the thesis the affine-like policies of the form

$$u = \eta + Ke(w) = \begin{bmatrix} \eta_0 \\ \vdots \\ \eta_{N-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ K_{1,1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \\ K_{N-1,1} & \dots & K_{N-1,N-1} & 0 \end{bmatrix} e(w), \tag{6.1}$$

were of the main concern. The number of optimization variables (not including slack variables associated with input constraints) is determined by the number of nonzero elements in the matrices η and K. If, n, m and N denote, as usual, the state-space dimension, the number of inputs and the optimization horizon length respectively, the number of optimization variables is

$$mn\frac{N(N-1)}{2} \approx mn\frac{N^2}{2}.$$

Furthermore, the input constraint

$$|\eta_i| + \varepsilon||K_i||_\infty \leq U_{\max},\ i = 1, \dots, mN$$

or equivalently

$$|\eta_i| + \varepsilon||K_i^T||_1 \leq U_{\max},\ i = 1, \ldots, mN$$

introduces another $mn\frac{N^2}{2}$ slack variables when represented as

$$\begin{aligned}
\eta_i + \varepsilon \mathbf{1}^T t_i &\leq U_{\max}, \quad i = 1, \ldots, mN \\
-\eta_i + \varepsilon \mathbf{1}^T t_i &\leq U_{\max}, \quad i = 1, \ldots, mN \\
-t_i \leq K_i^T &\leq t_i, \quad i = 1, \ldots, mN,
\end{aligned}$$

where t_i, $i = 1, \ldots, mN$, are vectors of slack variables, $\mathbf{1}$ denotes a vector of ones and the last inequality is componentwise. The total number of variables is thus approximately mnN^2. This number can grow quite rapidly, quickly giving rise to problems intractable for solvers that need to evaluate an exact Newton step at some point [10]. Indeed, a moderate-size problem with $n = 4$, $m = 2$ and a longer horizon $N = 48$ leads to a problem with approximately 18 500 variables and $9\,000 \times 9\,000$ dense block of the Hessian, which, although still solvable, results in prohibitively long computation time of the Newton step, let alone memory consumption on the order of several hundred of megabytes.

Thus it is necessary to resort to an approximate computation of the Newton step, preferable one where there is no need to actually form the Hessian. Although there are other possibilities (see [40]), here we chose the conjugate gradient method as a tool to solve a large system of equations with a positive definite matrix [40]. The key feature of the method is the fact that it only needs to evaluate the product of the Hessian and a vector, allowing for a dramatic speed-up in computation of the Newton step while, when tuned properly, returning excellent search direction compared to the first order gradient method. A downside of the method is higher problem dependence and the need for more careful tuning and/or preconditioning.

Using the conjugate gradient method, the key point of the implementation becomes a fast evaluation of the product of the Hessian and a vector. Recall the expressions for the Hessian derived in Lemma 3.2:

$$\begin{aligned}
\mathrm{Hess}(f) = \nabla\mu &\left[\frac{\partial^2 f}{\partial \mu^2}\nabla\mu + \frac{\partial^2 f}{\partial\mu\partial\sigma}\nabla\sigma\right]^T \\
&+ \nabla\sigma\left[\frac{\partial^2 f}{\partial \sigma^2}\nabla\sigma + \frac{\partial^2 f}{\partial\sigma\partial\mu}\nabla\mu\right]^T + \frac{\partial f}{\partial\sigma}\,\mathrm{Jac}(\nabla\sigma),
\end{aligned}$$

with

$$\mathrm{Jac}(\nabla\sigma) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{\sigma}C^T\left(I - \frac{(a+Ck)(a+Ck)^T}{\sigma^2}\right)C \end{bmatrix}.$$

This is a very convenient form for a multiplication by a vector since the first two terms translate to a vector-vector multiplication and the only matrix-vector multiplication is the one with $\mathrm{Jac}(\nabla\sigma)$, which can again be very fast if carried out properly. From Eq. (3.33) or (3.35) we see that the matrix C in the above expression has dimension $nN \times \frac{1}{2}mnN(N-1)$. Thus, denoting the vector to multiply with by v, and letting $\tilde{v} := Cv$, evaluation in the order

$$C^T\left(\tilde{v} - \frac{1}{\sigma^2}(a+Ck)[(a+Ck)\tilde{v}]\right)$$

will essentially amount to two multiplications of C and one multiplication of C^T by a vector, all of them being fast since the first dimension of C is modest.

As far as the problem of special functions present in the optimization problem, probably the only solution is to use a general (nonlinear) convex solver (e.g., IPOPT [57]). In our case we got by with a handwritten Matlab implementation of a primal-dual solver [20] with the exact Newton step replaced by the conjugate gradient one.

The problem of nondifferentiability is probably the least severe one as it can occur only on the rare occasion of zero variance of one of the terms of the cost. In fact it follows from Eq. (3.33) that, as a result of causality, it cannot happen for the $q_{ik}^T x_k$ terms if the noise is not degenerate (i.e., the noise covariance matrix has full rank). On the other hand it could occur for $r_{ik}^T u_k$ terms if the optimal solution were such that the k-th block row of K was zero, i.e., there was no recourse at time k. There are several ways to deal with this situation. For instance, one can fix the troublesome input to zero and optimize only over the remaining nonzero block-rows of K. This may, however, seem a little impractical since the situation must be detected and adjusted for online, leading to potential bursts in computation time. A more practical heuristic approach adopted here is to smooth the cost function as at the end of the proof of Lemma 3.2. This can be done just by replacing the 2-norm in the expression for the standard deviation $\sigma(\eta, k)$ in (3.33) or (3.35) by a smoothed version

$$||x||_2 \approx \sqrt{\frac{1}{j} + \sum_i x_i^2}$$

for some appropriately large j.

6.2. Numerical examples

Four numerical examples are presented to illustrate the methods developed in the thesis. Numerical simulations for some results on mean-square stability can be found at the end of Sections 4.3 and 4.4.

6.2.1. Finite horizon

First we consider a fixed horizon stochastic control optimal problem. For the system matrices and the noise covariance matrix we chose

$$A = \begin{bmatrix} 1 & -0.4 \\ 0.1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}, \quad \Sigma_w = I \otimes \begin{bmatrix} 8 & 5 \\ 5 & 6 \end{bmatrix}$$

with w_k zero-mean jointly Gaussian. We set the weighting matrices to $Q = I$ and $R = 0.1$, the input constraints to $U_{\max} = 30$, the initial state to $x_0 = [1, -1]^T$ and the optimization horizon to $T = 12$. The function $e(w)$ was chosen as the elementwise saturation that saturates the disturbances at $4\sqrt{\max(\text{diag}(\Sigma_w))} = 11.31$. We compared our method (with $N_c = N = T$) with the standard certainty equivalent MPC ($N_c = 1$, $N = T$) and with the shrinking horizon CE-MPC ($N_c = 1$, $N(k) = T - k$, $k = 0, \ldots, T-1$). Furthermore, we tried out the proposed method with $K = 0$ against the certainty equivalent open loop control (i.e., CE-MPC with $N_c = N = T$). For the sake of completeness we tried out our method in the shrinking horizon mode with $N_c = 2, N(k) = T - k$ as well. The respective objective functions were evaluated using 1000 Monte Carlo runs. The results for the p-norm minimization using the upper-bounds (3.37) are summarized in Table 6.1, which shows that our method (without shrinking) outperforms the others by a significant margin except perhaps for SH-MPC where the difference is smaller and, naturally, our method in the shrinking horizon mode. On the other hand, unlike with MPC strategies, there is no need for online optimization with our method in this setting. It is also worth noting that our method with $K = 0$ (i.e., an open loop policy) slightly outperforms the certainty equivalent open loop control, which is in contrast with the quadratic cost case where this strategy is optimal in the class of open loop policies. Finally, we evaluate the improved version of the bound (3.52), which yields $\tilde{\beta} = 4.5 \cdot 10^{-5}$ showing that the solution found by (3.31) is in this case basically optimal in (3.6) for $p = 1$. For $p > 0$ no conclusions about the suboptimality of the solution can be made, although the superior performance compared to the CE-MPC suggests that it be should very small.

Table 6.1.: Comparison of control policies over the optimization horizon $T = 12$.

p	SH-(3.31)	(3.31)	SH-MPC	MPC	$(3.31), K = 0$	OL
1	85.9	90.3	98.3	120.0	140.5	144.2
1.5	73.3	76.8	85.0	102.7	115.2	118.0
2	70.5	74.1	80.1	96.0	110.9	113.7

6.2.2. Receding horizon

Our second example compares the proposed method with the certainty equivalent MPC in a receding horizon regime. In this example we consider the respective matrices

$$A = \begin{bmatrix} 1 & 1 \\ -0.5 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{E}\{w_i w_j^T\} = \begin{bmatrix} 8 & 5 \\ 5 & 6 \end{bmatrix} \delta_{ij}$$

with w_k zero-mean Gaussian and independent, where δ_{ij} is the Kronecker delta. The weighting matrices were set to $Q = I$ and $R = 0.1I$, the input constraints to $U_{\max} = 10$, and the initial state to $x_0 = [1, -1]^T$. We compared our control policy with $N = 12$, $N_c = 4$ against CE-MPC with $N = 12$, $N_c = 1$ in a receding horizon mode over the simulation time $T = 100$. Again, we used the 4-sigma rule to get $\varepsilon = 13.9$. Figure 6.1 shows the accumulation of the cost over the simulation time, while Figure 6.2 depicts the evolution of the state's 2-norm-square expectation suggesting its boundedness, which was to be expected since $\rho(A) = \sqrt{2}/2 < 1$. 100 Monte Carlo runs were used to evaluate the expectations in the costs.

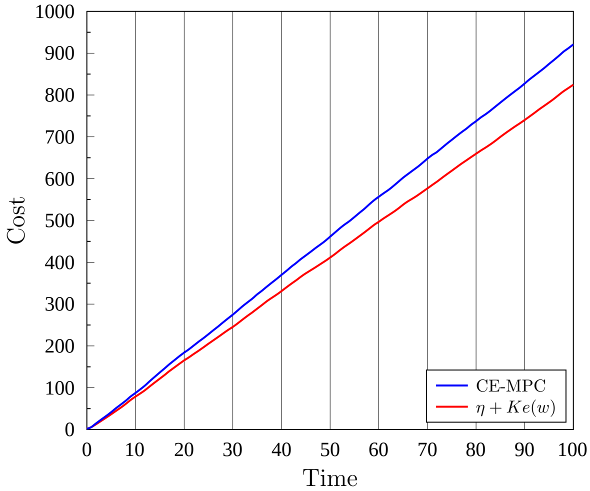

Figure 6.1.: Comparison of costs over the simulation time $T = 100$ in a receding horizon regime with $N = 12$ and $N_c = 4$ for our policy and $N_c = 1$ for CE-MPC. Final costs are 824.7 for our control scheme and 921.1 for CE-MPC.

6.2.3. Recursive feasibility

6.2.4. Invariant set demonstration

Our third example examines the strongly feasible model predictive control algorithm of Section 5.2. We compared both the terminal (AD-T) and the first-step (without mode 1

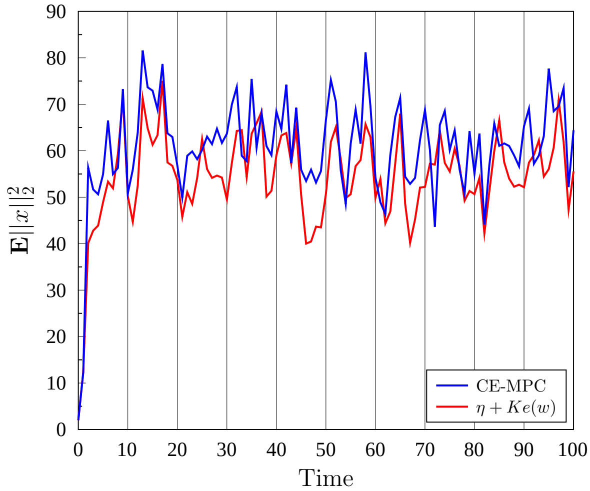

Figure 6.2.: Evolution of $\mathbf{E}||x||_2^2$ under our receding horizon control policy with $N = 12$, $N_c = 4$ and CE-MPC control policy with $N = 12$, $N_c = 1$.

constraints) (AD-F) affine disturbance feedback polices against the perturbed linear state feedback stochastic MPC (P-SMPC) of [36] and the robust affine disturbance feedback (AD-R). The additional parameters for the P-SMPC policy are $\hat{N} = 40$ and $n^* = 1$ (see [36] for the meaning of the parameters). As the first step constraint for AD-R we used the MSRCI set with a zero probability of violation (i.e., $\alpha_j = 0$ for all j), $\mathcal{X}'_{\text{rob}}$, which is in general *not* the maximum robust controlled invariant subset of $\mathcal{X}$, $\mathcal{X}_{\text{rob}}$ (see Remark 5.7 and Figure 6.3).

We consider the system given by the matrices

$$A = \begin{bmatrix} 1.25 & -0.15 \\ 0.25 & 1.02 \end{bmatrix}, \quad B = \begin{bmatrix} 0.14 \\ 0.12 \end{bmatrix},$$

where w_k is an i.i.d. sequence obtained by truncating the standard normal distribution at $\Delta = ||w_k||_\infty \leq 3$. The weighting matrices were set to $Q = I$ and $R = \sqrt{1.1}$ and the input constraints to $U_{\max} = 250$. We chose a quadratic ($p = 2$) cost function J, which can be evaluated exactly for all of the policies considered. The constraint set $\mathcal{X}$ is given by two constraints $g_1^T x \leq h_1$ and $g_2^T x \leq h_2$ with $g_1 = [-0.41, 1]^T$, $h_1 = 31$ and $g_2 = [-0.7593, 1]^T$, $h_2 = 43.494$, and the allowed probability of violation $\alpha_1 = \alpha_2 = 0.1$. All of the policies were applied in a receding horizon fashion with the prediction horizon $N = 8$. We chose $K_s = [1.7303, \ -13.1020]$ as the mode 2 controller for AD-T as well as the base policy for P-SMPC. Note that the LQ optimal state feedback cannot be used in this case since then $X_r^{K_s}$ turns out to be empty and as a consequence both policies are globally infeasible. For the sake of comparison we also included the LQ-optimal policy itself. The initial state $x_0 = [13.34, \ 42.46]^T$ was chosen to lie on the boundary of $\mathcal{X}'_{\text{rob}}$.

All of the (invariant) sets considered and the initial state are depicted in Figure 6.3.

Performance and constraint violation was evaluated over 500 Monte Carlo runs. The accumulation of the cost over the simulation horizon $T = 20$ is depicted in Figure 6.4 and the final costs are listed in Table 6.2. The two proposed strongly feasible MPC formulations outperform the P-SMPC and AD-R policies and, naturally, perform worse than the LQ-optimal policy. We also observe tight satisfaction of the chance constraint at the time $k = 1$ with our policies; the constraint violation is 9.6 % for both, which is close, but within, the prescribed 10 % limit. The P-SMPC and AD-R policies are more conservative here, exhibiting zero violation. The LQ-optimal control, in contrast, violated the constraint in 89.0 % of the 500 runs performed. Violations at other time steps were zero or negligible for all investigated policies.

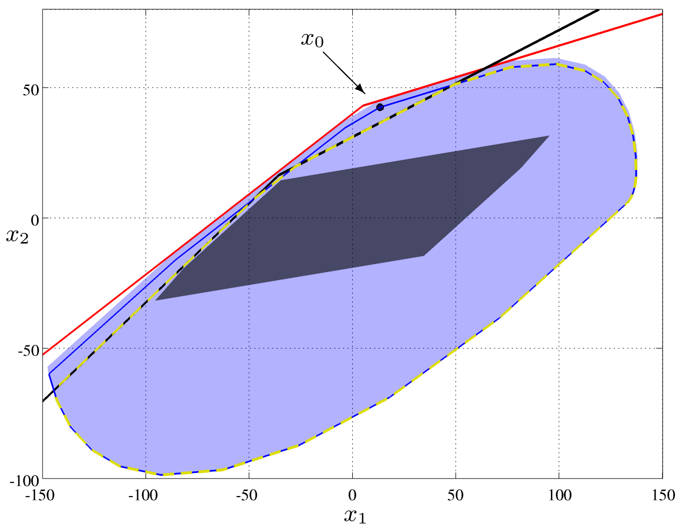

Figure 6.3.: Constraint set $\mathcal{X}$ (below solid black line), feasible set $\mathcal{X}_f$ (below solid red line), MSRCI set $\mathcal{X}^*_{rc}$ (light blue interior), linear state-feedback positively invariant set $\mathcal{X}_r^{K_s}$ (dark blue interior), maximum robust controlled invariant set $\mathcal{X}_{\text{rob}}$ (dashed yellow boundary), zero violation ($\alpha_j = 0$) MSRCI set $\mathcal{X}'_{\text{rob}}$ (blue solid boundary). Note that $\mathcal{X}^*_{rc} \not\subset \mathcal{X}$ and $\mathcal{X}_{\text{rob}} \neq \mathcal{X}'_{\text{rob}}$, but $\mathcal{X}_{\text{rob}} \subset \mathcal{X}'_{\text{rob}} \subset \mathcal{X}^*_{rc} \subset \mathcal{X}_f$ and $\mathcal{X}_{\text{rob}} \subset \mathcal{X}$.

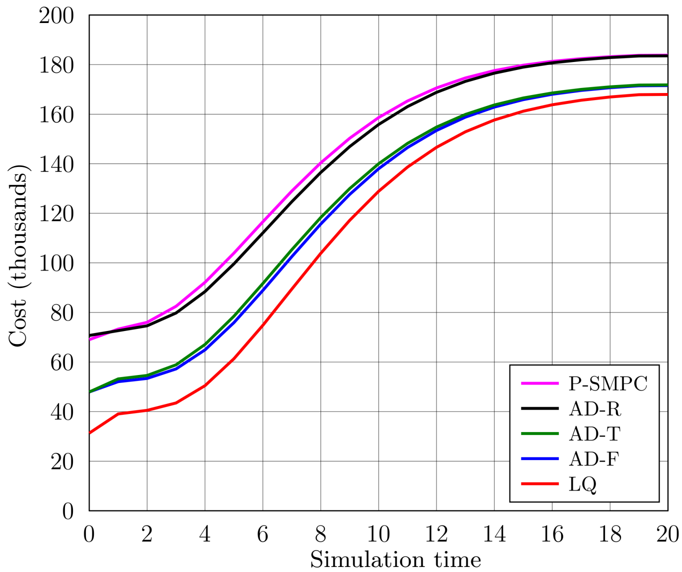

Figure 6.4.: Comparison of costs over the simulation time $T = 20$ for the five investigated strategies.

Table 6.2.: Comparison of control policies over the optimization horizon $T = 20$. First row: percentage increase over the LQ-optimal policy. Second row: the probability of violating the state constraints at time one. The final cost of the LQ-optimal policy, J_{LQ}, is $1.6795 \cdot 10^5$.

Policy	LQ	AD-F	AD-T	AD-R	P-SMPC
$100\frac{J-J_{\mathrm{LQ}}}{J_{\mathrm{LQ}}}$	0	2.132	2.298	9.277	9.431
$P(g_1^T x_1 > h_1)$	0.89	0.096	0.096	0	0

6.2.5. Long-run constraint violations

In the previous example, there are no constraint violations when stationarity is reached. The next example shows that it is possible to achieve repeated constraint violations in stationarity, and thus to obtain significant performance improvement compared to the robust MPC by fully exploiting the probabilistic nature of constraints over a long period of time. We consider the system

$$x_{k+1} = Ax_k + Bu_k + w_k = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u_k + w_k,$$

with the noise sequence w_k having the standard normal distribution truncated at $\Delta = 3$. The initial state is $x_0 = [5, 5]^T$. The only constraint on the state is $P(x_2 \geq 0) \geq 1 - \alpha$, whereas the control authority is bounded by $U_{\max} = 12$. Simulations were carried out for four values of the allowed probability of constraint violation: $\alpha = 0.1$, $\alpha = 0.2$, $\alpha = 0.3$ and $\alpha = 0.4$. We compared the first-step affine disturbance feedback (AD-F) with the robust affine disturbance feedback policy (AD-R) and the LQ optimal policy. The prediction and control horizons were $N = 8$ and $N_c = 1$ for both disturbance feedback policies.

Instead of Monte Carlo analysis, we examined constraint violations over a single, but very long, trajectory. Simulation results are depicted in Figures 6.5, 6.6 and 6.7. Table 6.3 then summarizes the results. For all four values of α, the closed-loop trajectory under the first-step affine disturbance feedback tightly satisfies the probabilistic constraint, and as a result achieves a significant performance improvement over the robust affine disturbance feedback policy. The LQ optimal policy, which is oblivious to all constraints, naturally outperforms both policies, but violates the probabilistic constraint substantially.

Table 6.3.: Comparison of control policies over the optimization horizon $T = 10000$

policy	LQ	$\alpha = 0.4$	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$	Robust
J/J_{LQ}	1	1.07	1.29	1.73	2.68	10.23
#violations	4920	3916	2942	1983	992	0

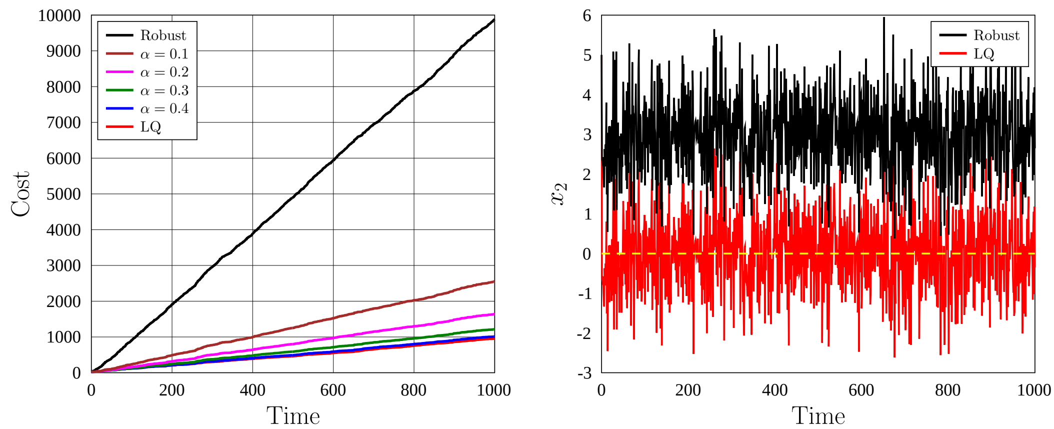

Figure 6.5.: Simulation results over a time period $T = 1000$. Left: cost functions. Right: sample path of the x_2 for LQ and robust control policies.

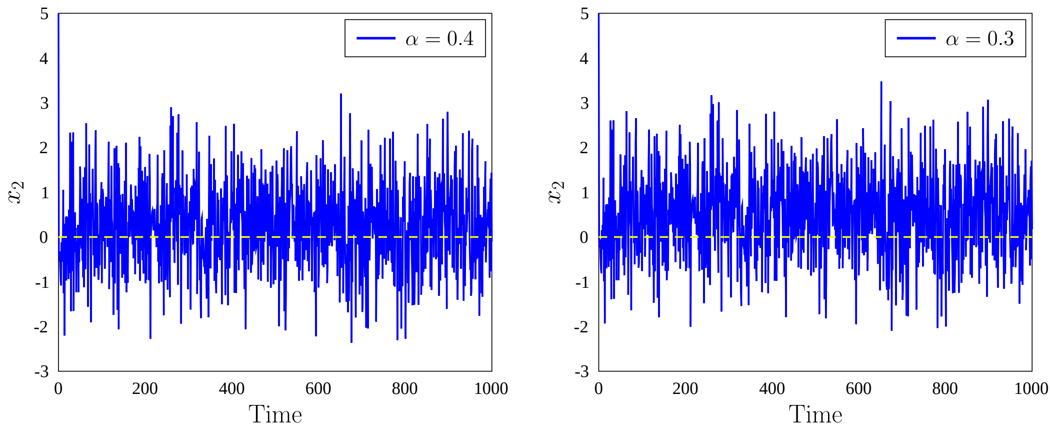

Figure 6.6.: Simulation results over a time period $T = 1000$. Left: sample path of x_2 for $\alpha = 0.4$. Right: sample path of x_2 for $\alpha = 0.3$.

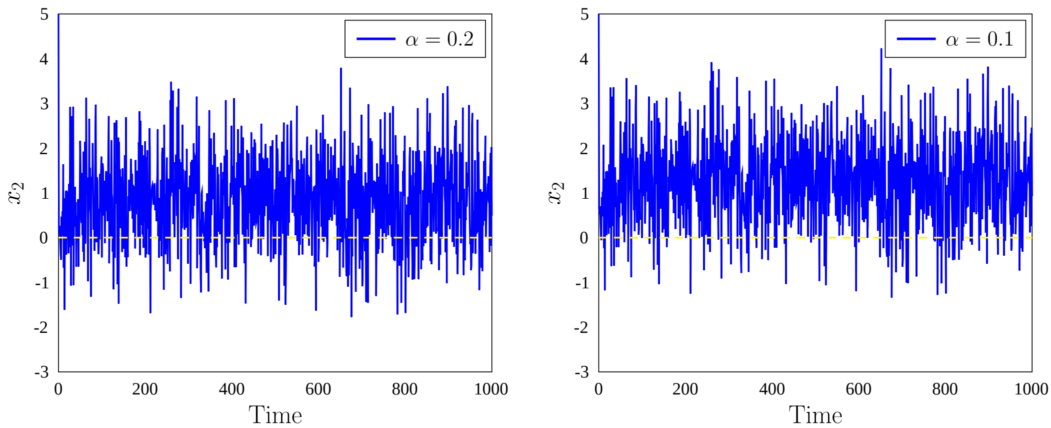

Figure 6.7.: Simulation results over a time period $T = 1000$. Left: sample path of x_2 for $\alpha = 0.2$. Right: sample path of x_2 for $\alpha = 0.1$.

7. Conclusion

In this thesis we investigated three areas of stochastic optimal control and constrained model predictive control of linear systems with additive noise and bounded control inputs.

ℓ_p stochastic optimal control First, we dealt with the expectation of the p-norm stochastic optimal control problem with perfect as well as imperfect state information. For both problems we developed an approximate solution technique ensuring bounded control inputs in the presence of Gaussian disturbances. The approximation technique leads to a tractable convex optimization problem without resorting to sampling techniques, thus keeping acceptable computational requirements. Moreover, we constructed a suboptimality bound of our method in a certain class of nonlinear feedback control policies.

Stability of linear stochastic systems Second, we discussed mean-square stabilizability of linear stochastic systems with bounded control authority. To a large extent, the problem has already been solved in recent years, and the remaining open problems of stabilization of marginally unstable systems or the existence of a stabilizing static state-feedback policy for the marginally stable case seems to be far from trivial. Nevertheless, we developed simplified proofs of existing results, and proved mean-square stabilizability of positive (or negative) parts of marginally unstable systems. We also proved the existence of a mean-square stabilizing Markov policy for marginally stable systems. The question of mean-square stabilizability of marginally unstable systems remains open, and the answer (whether positive or negative) will most likely require different techniques than those of this text, at least in dimensions greater than two.

Recursive feasibility We developed a systematic approach to ensure recursive feasibility of a set of probabilistic constraints under affine as well as nonlinear disturbance feedback policies. The first approach employs the well established notion of terminal constraints, whereas the second one takes advantage of the more recently developed first-step constraint. Both of them turn out to have direct analogies in a stochastic environment carrying over their advantages and disadvantages. In particular the first approach is policy-dependent and hence not amenable to imposing additional structural constraints on the feedback matrix. In contrast, the second approach is policy-independent and results in the largest feasible set amongst all admissible policies. There is still a certain degree of conservatism since the recursive feasibility is enforced via one-step conditional probabilities. Future research should focus on reducing this conservatism by allowing for larger excursions outside the constraint set with predefined (low) probability.

List of Symbols

2^S	Power set of a set S, page 22
$\mathrm{bdiag}(\cdot, \dots, \cdot)$	Block diagonal matrix composed of the arguments, page 25
$\mathbb{N}_i^j$	Set of consecutive integers $\{i, \dots, j\}$, page 35
$\mathbf{E}$	Expectation of a random variable, page 5
$\mathbf{E}(\cdot \mid \mathcal{Y})$	Conditional expectation given $\mathcal{Y}$, page 6
$\mathbf{E}^\pi$	Expectation under a policy π, page 40
ε	Bound on the nonlinear function $e(\cdot)$, page 16
$\mathrm{Hess}(\cdot)$	Hessian matrix of a twice differentiable function, page 17
$\mathbb{1}_A$	Indicator function of a set A, page 41
$\mathrm{Jac}(\cdot)$	Jacobian matrix of a differentiable function, page 17
$\lVert\cdot\rVert_p$	p-norm of a column vector or induced p-norm of a matrix, page 20
$\mathbb{N}_0$	The set of nonnegative integers, page 31
$\mathbb{R}$	The set of real numbers, page 15
$\mathcal{N}(\mu, \Sigma)$	Normal distribution with the expectation μ and covariance matrix Σ, page 17
$\mu \otimes \nu$	Product measure of measures μ and ν, page 22
$\nabla(\cdot)$	Gradient of a differentiable function, page 17
$\mathbf{1}$	Column vector of ones, page 62
Σ_v	Joint covariance matrix of the measurement noise along the prediction horizon, page 25
Σ_w	Joint covariance matrix of the process noise along the prediction horizon, page 15

$\mathrm{tr}(\cdot)$	Trace of a matrix, page 33
$\mathrm{vec}(\cdot)$	Vectorization of a matrix, page 21
$A \otimes B$	Kronecker product of matrices A and B, page 21
l	Output dimension, page 25
m	Input dimension, page 15
N	Optimization or prediction horizon, page 15
n	State-space dimension, page 15
P	Probability measure, page 22
P^{π}	Probability measure under a policy π, page 39
x^{+}	Positive part of a real number x, page 45
mod	Modulo operation, page 42

Bibliography

[1] M. Agarwal, E. Cinquemani, D. Chatterjee, and J. Lygeros. On convexity of stochastic optimization problems with constraints. In *European Control Conference*, pages 2827–2832, 2009.

[2] D. S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton University Press, second edition, 2009.

[3] D. P. Bertsekas. *Dynamic Programming: Deterministic and Stochastic Models*. Prentice Hall, first edition, 1987.

[4] D. P. Bertsekas. Dynamic programming and suboptimal control: A survey from ADP to MPC. In *CDC Proceedings*, 2005.

[5] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, third edition, 2007.

[6] D. P. Bertsekas and E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific, 2007.

[7] D. Bertsimas and D.B. Brown. Constrained stochastic LQC: a tractable approach. *IEEE Transactions on Automatic Control*, 52(10):1826–1841, 2007.

[8] L. Blackmore and M. Ono. Convex chance constrained predictive control without sampling. In *Proceedings of the AIAA Guidance, Navigation and Control Conference*, 2009.

[9] F. Blanchini. Set invariance in control. *Automatica*, 35(11):1747–1767, 1999.

[10] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[11] D. L. Burkholder. Distribution function inequalities for martingales. *The Annals of Probability*, 1(1):19–42, 1973.

[12] G. Calafiore and M. C Campi. On two-stage portfolio allocation problems with affine recourse. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, pages 8042–8047, 2005.

[13] G. C Calafiore. A linear reaction technique for dynamic asset allocation in the presence of transaction costs. In *47th IEEE Conference on Decision and Control*, pages 1765–1770.

[14] G. C Calafiore. Multi-period portfolio optimization with linear control policies. *Automatica*, 44(10):2463–2473, 2008.

[15] Giuseppe Carlo Calafiore. An affine control method for optimal dynamic asset allocation with transaction costs. *SIAM Journal on Control and Optimization*, 48(4):22–54, 2009.

[16] M. Cannon, P. Couchman, and B. Kouvaritakis. Mpc for stochastic systems. *Assessment and Future Directions of Nonlinear Model Predictive Control*, pages 255–268, 2007.

[17] M. Cannon, B. Kouvaritakis, and D. Ng. Probabilistic tubes in linear stochastic model predictive control. *Systems & Control Letters*, 58(10-11):747–753, 2009.

[18] M. Cannon, B. Kouvaritakis, S. V Raković, and Q. Cheng. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Transactions on Automatic Control*, 56(1):194–200, 2011.

[19] M. Cannon, B. Kouvaritakis, and X. Wu. Model predictive control for systems with stochastic multiplicative uncertainty and probabilistic constraints. *Automatica*, 45(1):167–172, 2009.

[20] P. Carbonetto. Matlab implementation of a primal-dual interior-point solver for convex programs with constraints. `http://www.cs.ubc.ca/~pcarbo/convexprog.html`, 2008.

[21] D. Chatterjee, E. Cinquemani, G. Chaloulos, and J. Lygeros. On Stochastic Control up to a Hitting Time. In *IEEE Conference on Decision and Control*, pages 531–536, 2009.

[22] D. Chatterjee, P. Hokayem, and J. Lygeros. Stochastic receding horizon control with bounded control inputs: a vector space approach. *IEEE Transactions on Automatic Control*, 2009.

[23] E. Cinquemani, M. Agarwal, D. Chatterjee, and J. Lygeros. On convex problems in chance-constrained stochastic model predictive control. *Arxiv preprint arXiv:0905.3447*, 2009.

[24] C. E. T. Dórea and J. C. Hennet. (A, B)-invariant polyhedral sets of linear discrete-time systems. *Journal of optimization theory and applications*, 103(3):521–542, 1999.

[25] E. G. Gilbert and K. T. Tan. Linear systems with state and control constraints: The theory and application of maximal output admissible sets. *IEEE Transactions on Automatic Control*, 36(9):1008–1020, 1991.

[26] R. Gondhalekar and J. Imura. Least-restrictive move-blocking model predictive control. *Automatica*, 46(7):1234–1240, 2010.

[27] R. Gondhalekar, J. Imura, and K. Kashima. Controlled invariant feasibility – A general approach to enforcing strong feasibility in MPC applied to move-blocking. *Automatica*, 45(12):2869–2875, 2009.

[28] O. Hernandez-Lerma and J. B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, first edition, 1996.

[29] P. Hokayem, D. Chatterjee, and J. Lygeros. On stochastic model predictive control with bounded control inputs. pages 6359–6364, 2009.

[30] P. Hokayem, D. Chatterjee, F. Ramponi, G. Chaloulos, and J. Lygeros. Stable stochastic receding horizon control of linear systems with bounded control inputs. In *Mathematical Theory of Networks and Systems*, pages 31–36, 2010.

[31] P. Hokayem, E. Cinquemani, D. Chatterjee, and J. Lygeros. Stochastic MPC with Imperfect State Information and Bounded Controls. In *UKACC International Conference on CONTROL*, Coverntry, UK, 2010.

[32] O. Kallenberg. *Foundations of Modern Probability*. Springer New York, second edition, 2010.

[33] E. C Kerrigan and J. M Maciejowski. Robust feasibility in model predictive control: Necessary and sufficient conditions. In *IEEE Conference on Decision and Control*, pages 728–733, 2001.

[34] E.C. Kerrigan. *Robust Constraint Satisfaction: Invariant Sets and Predictive Control*. PhD thesis, University of Cambridge, UK, 2000.

[35] M. Korda and J. Cigler. A tractable approximation of the 1-norm stochastic mpc with bounded control inputs. In *American Control Conference*, 2011. To appear.

[36] 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.

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

[38] K. E. Muller. Computing the confluent hypergeometric function, M(a,b,x). *Numerical Mathematics*, (90):179–196.

[39] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2007.

[40] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, second edition, 2006.

[41] F. Oldewurtel, R. Gondhalekar, C. N Jones, and M. Morari. Blocking parameterizations for improving the computational tractability of affine disturbance feedback MPC problems. In *48th IEEE Conference on Decision and Control*, pages 7381–7386, 2009.

[42] F. Oldewurtel, C. N Jones, and M. Morari. A tractable approximation of chance constrained stochastic MPC based on affine disturbance feedback. In *47th IEEE Conference on Decision and Control*, pages 4731–4736, 2008.

[43] F. Oldewurtel, A. Parisio, C. N Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth. Energy efficient building climate control using stochastic model predictive control and weather predictions. In *American Control Conference*, pages 5100–5105, 2010.

[44] R. Pemantle and J. S Rosenthal. Moment conditions for a sequence with negative drift to be uniformly bounded in lr. *Stochastic Processes and their Applications*, 82(1):143–155, 1999.

[45] T. Perez, H. Haimovich, and G. C Goodwin. On optimal control of constrained linear systems with imperfect state information and stochastic disturbances. *International Journal of Robust and Nonlinear Control*, 14(4):379–393, 2004.

[46] James A. Primbs and Chang Hwan Sung. Stochastic receding horizon control of constrained linear systems with state and control multiplicative noise. *IEEE Transactions on Automatic Control*, 54(2):221–230, 2009.

[47] Federico Ramponi, Debasish Chatterjee, Andreas Milias-Argeitis, Peter Hokayem, and John Lygeros. Attaining mean square boundedness of a marginally stable stochastic linear system with a bounded control input. *IEEE Transactions on Automatic Control*, 55(10):2414–2418, 2010.

[48] James B. Rawlings and David Q. Mayne. *Model Predictive Control: Theory and Design*. Nob Hill Publishing, first edition, 2009.

[49] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, third edition, 1986.

[50] pages = 41-54 S. Raković, year = 2009. Set theoretic methods in model predictive control. *Nonlinear Model Predictive Control*.

[51] J. Skaf and S. Boyd. Nonlinear Q-Design for convex stochastic control. *IEEE Transactions on Automatic Control*, 54(10):2426–2430, 2009.

[52] J. Skaf and S. Boyd. Design of affine controllers via convex optimization. *Submitted to IEEE Transactions on Automatic Control*, 2010.

[53] J. Skaf, S. Boyd, and A. Zeevi. Shrinking-horizon dynamic programming. *To appear in International Journal of Robust and Nonlinear Control*, 2010.

[54] R. F. Stengel. *Optimal Control and Estimation.* Dover Publications, 2004.

[55] J. Sturm. Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11:625–653, 1999.

[56] D.H. van Hessem and O.H. Bosgra. A full solution to the constrained stochastic closed-loop MPC problem via state and innovations feedback and its receding horizon implementation. In *42nd IEEE Conference on Decision and Control*, pages 4731–4736, 2003.

[57] A. Wächter and L. T. Biegler. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57, 2006.

[58] Y. Wang and S. Boyd. Performance bounds for linear stochastic control. *Systems & Control Letters*, 58(3):178–182, 2009.

[59] Y. Wang and S. Boyd. Approximate dynamic programming via iterated bellman inequalities. Manuscript, 2010.

[60] Y. Wang and Stephen Boyd. Approximate dynamic programming via iterated bellman inequalities. `http://www.stanford.edu/~boyd/papers/pdf/adp_iter_bellman.pdf`, June 2010.

[61] B. Williams, M. Ono, B. C Williams, M. Embedded, et al. Efficient motion planning algorithm for stochastic dynamic systems with constraints on probability of failure. *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence*, 2008.

[62] D. Williams. *Probability with Martingales.* Cambridge University Press, first edition, 1991.

A. Kummer's confluent hypergeometric function

A brief description of some of the Kummer's confluent hypergeometric function computational methods is given here. Only the methods 1 and 2 of the article [38] are outlined since only these were employed in the framework of this thesis. Kummer's confluent hypergeometric function is defined as

$$M(a, b, x) = \sum_{i=0}^{\infty} \frac{(a)_i}{(b)_i} \frac{x^i}{i!}, \tag{A.1}$$

where the rising factorial $(a)_i$ is given by

$$(a)_i = \prod_{k=0}^{i-1} (a + k).$$

In the algorithms used in the thesis, the first two arguments a and b are constant and small in magnitude, and x is always negative. Then the analysis of the article [38] shows that it is effective to use the (truncated) defining series (A.1) for small x (approx. $|x| < 20$) and the asymptotic expansion

$$M(a, b, x) \approx \frac{\Gamma(b)(-x)^{-a}}{\Gamma(b-a)} \sum_{i=0}^{N} \frac{(1 + a - b)_i (a)_i}{i!(-x)^i} \tag{A.2}$$

for $x < -20$.

Both of these approximations can be evaluated recursively, usually requiring only several terms of the respective sums to get a good enough approximation. Furthermore, depending on p only, the values of $\Gamma(b)$ and $\Gamma(b-a)$ can be precomputed. The maximum approximation error for $M_1, \ldots, M_4$ in (3.19) is typically on the order of 10^{-8} around $x = -20$ and much smaller otherwise. The computation time proved to be negligible in the context of the overall algorithm.

B. Construction of controlled invariant sets

A brief description of most common algorithms for construction of controlled invariant sets is given here. Given a system

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

and a polyhedral set $\mathcal{X}$, we seek the maximum robust controlled invariant subset, i.e., the largest set $\mathcal{X}_{rc}^* \subset \mathcal{X}$ such that

$$\forall x \in \mathcal{X}_{rc}^* \; \exists\, u \in \mathcal{U} \text{ s.t.}: \; Ax + Bu + w \in \mathcal{X}_{rc}^* \;\; \forall w \in \mathcal{W}, \tag{B.1}$$

where $\mathcal{U}$ is a polyhedral input constraint set.

The first algorithm constructs a nested sequence of supersets of $\mathcal{X}_{rc}^*$ as follows [9]. Set $\mathcal{X}_0 := \mathcal{X}$ and define recursively for $i \geq 1$

$$\mathcal{X}_i := \{x \in \mathcal{X}_{i-1} \mid \exists\, u \in \mathcal{U} \text{ s.t.}\, Ax + Bu + w \in \mathcal{X}_{i-1} \;\; \forall w \in \mathcal{W}\}.$$

Then clearly $\mathcal{X}_{i+1} \subset \mathcal{X}_i$ for all $i \geq 0$ and

$$\mathcal{X}_{rc}^* = \bigcap_{i=0}^{\infty} \mathcal{X}_i.$$

In particular $\mathcal{X}_{rc}^* = \mathcal{X}_i$ whenever $\mathcal{X}_i = \mathcal{X}_{i+1}$, in which case $\mathcal{X}_{rc}^*$ is polyhedral.

The second algorithm on the other hand starts from any robust controlled invariant subset of $\mathcal{X}$ and constructs an increasing sequence of subsets of $\mathcal{X}_{rc}^*$ [27]. If possible, the initial invariant set can be taken as the invariant set with respect to a stabilizing linear state feedback. Thus we assume that a robust controlled invariant subset $\mathcal{Y}_0 \subset \mathcal{X}$ is on hand and define recursively for $i \geq 1$

$$\mathcal{Y}_i := \{x \in \mathcal{X} \mid \exists\, u \in \mathcal{U} \text{ s.t.}\, Ax + Bu + w \in \mathcal{Y}_{i-1} \;\; \forall w \in \mathcal{W}\}.$$

Due to the invariance of $\mathcal{Y}_0$, each set $\mathcal{Y}_i$ is also invariant and $\mathcal{Y}_i \subset \mathcal{Y}_{i+1}$ for all $i \geq 0$. Hence[1]

$$\mathcal{X}_{rc}^* = \bigcup_{i=0}^{\infty} \mathcal{Y}_i,$$

[1] To be precise, the closure of the union must be taken to obtain $\mathcal{X}_{rc}^*$ if $\mathcal{Y}_{i+1} \neq \mathcal{Y}_i$ for all i.

and $\mathcal{X}_{rc}^* = \mathcal{Y}_i$ whenever $\mathcal{Y}_i = \mathcal{Y}_{i+1}$.

Main advantage of the second algorithm is the fact that it produces a robust controlled invariant set at each step, which is not the case for the first algorithm.

When the maximum set is not finitely determined or its complexity is prohibitive, the first algorithm can serve as a stopping criterion for the second one. For all i we have

$$\mathcal{Y}_i \subset \mathcal{X}_{rc}^* \subset \mathcal{X}_i,$$

and thus, for instance, the Hausdorff distance between $\mathcal{X}_i$ and $\mathcal{Y}_i$

$$\rho_H(\mathcal{X}_i, \mathcal{Y}_i) = \sup_{x \in \mathcal{X}_i} \inf_{y \in \mathcal{Y}_i} ||x - y||_2$$

can be used as a basis for a stopping criterion [27].