Special Bilevel Quadratic Problems for Construction of Worst-Case Feedback Control in Linear-Quadratic Optimal Control Problems under Uncertainties

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Philipps-Universität Marburg

vorgelegt von
Hilke Isabell Stibbe
aus
Arolsen

Juli 2019

Erstgutachter: Prof. Dr. Ekaterina Kostina
Zweitgutachter: Prof. Dr. Thomas Surowiec
Eingereicht: 10. Juli 2019
Tag der Disputation: 30.09.2019
Erscheinungsort: Marburg
Hochschulkennziffer: 1180

Abstract

Almost all mathematical models that describe processes, for instance in industry, engineering or natural sciences, contain uncertainties which arise from different sources. We have to take these uncertainties into account when solving optimal control problems for such processes. There are two popular approaches : On the one hand the so-called closed-loop feedback controls, where the nominal optimal control is updated as soon as the actual state and parameter estimates of the process are available and on the other hand robust optimization, for example worst-case optimization, where it is searched for an optimal solution that is good for all possible realizations of uncertain parameters.

For the optimal control problems of dynamic processes with unknown but bounded uncertainties we are interested in a combination of feedback controls and robust optimization. The computation of such a closed-loop worst-case feedback optimal control is rather difficult because of high dimensional complexity and it is often too expensive or too slow for complex optimal control problems, especially for solving problems in real-time. Another difficulty is that the process trajectory corresponding to the worst-case optimal control might be infeasible. That is why we suggest to solve the problems successively by dividing the time interval and determining intermediate time points, computing the feedback controls of the smaller intervals and allowing to correct controls at these fixed intermediate time points. With this approach we can guarantee that for all admissible uncertainties the terminal state lies in a given prescribed neighborhood of a given state at a given final moment. We can also guarantee that the value of the cost function does not exceed a given estimate.

In this thesis we introduce special bilevel programming problems with solutions of which we may construct the feedback controls. These bilevel problems can be solved explicitly. We present, based on these bilevel problems, efficient methods and approximations for different control policies for the combination of feedback control and robust optimization methods which can be implemented online, compare these approaches and show their application on linear-quadratic control problems.

Zusammenfassung

Fast alle mathematischen Modelle, die Prozesse beschreiben, wie sie zum Beispiel in der Industrie, dem Ingenieurwesen oder den Naturwissenschaften vorkommen, enthalten Unsicherheiten. Diese Unsicherheiten können aus verschiedenen Quellen stammen und müssen bei der Lösung von Optimalsteuerungsproblemen berücksichtigt werden. Hierfür gibt es zwei beliebte Ansätze: Der erste Ansatz ist die so genannte Closed-Loop Feedbacksteuerung, bei der wir die nominale optimale Steuerung immer dann aktualisieren, wenn aktuelle Schätzungen des Zustandes und der Parameter des Prozesses vorhanden sind. Der zweite Ansatz ist die robuste Optimierung, wie zum Beispiel die Worst-Case Optimierung, bei der wir eine optimale Lösung suchen, die für alle Realisierungen der unbekannten Parameter gute Ergebnisse liefert.

Bei Optimalsteuerungsproblemen von dynamischen Systemen mit unbekannten, aber beschränkten Unsicherheiten interessieren wir uns für eine Kombination von Feedbacksteuerung und robuster Optimierung. Eine solche Closed-Loop Worst-Case Feedback Optimalsteuerung ist recht schwierig zu berechnen, da die Probleme hochdimensional komplex sind und damit die Berechnung zu teuer oder zu langsam für komplexe Optimalsteuerungsprobleme ist, gerade wenn wir Probleme in Echtzeit lösen möchten. Eine weitere Schwierigkeit ist, dass die Trajektorie der Steuerung unzulässig ist. Deshalb schlagen wir vor, die Probleme sukzessiv zu lösen, indem wir das Zeitintervall unterteilen, dazwischen liegende Zeitpunkte bestimmen, die Feedbacksteuerungen dieser kleineren Zeitintervalle berechnen und erlauben, die Steuerungen an diesen festen Zwischenzeitpunkten zu korrigieren. Mit diesem Ansatz können wir garantieren, dass für alle zulässigen Störungen der Endzustand im Endzeitpunkt in einer gegebenen vorgeschriebenen Umgebung eines gegebenen Zustands liegt. Außerdem wird dadurch gewährleistet, dass der Wert der Zielfunktion einen gegebenen Wert nicht überschreitet.

In dieser Arbeit führen wir spezielle Bilevel-Optimierungsprobleme ein, mit dessen Lösungen wir die Feedbacksteuerungen konstruieren können. Diese Bilevelprobleme können explizit gelöst werden. Aufbauend auf diesen Bilevelproblemen stellen wir effiziente Methoden und Approximationen verschiedener Steuerungsstrategien für die Kombination von Feedbacksteuerung und robuster Optimierung vor, die online implementiert werden können. Diese Ansätze vergleichen wir und zeigen die Anwendung an linear-quadratischen Problemen.

Acknowledgements

First of all I want to deeply thank my supervisor Prof. Dr. Ekaterina Kostina for giving me the opportunity to write this thesis in her working group. Thank you Ekaterina, for supporting me in an academic and personal way, giving me great advice, encouraging me and having a great working atmosphere in the group.

Thank you to Prof. Dr. Thomas Surowiec for being my second reviewer.

I also want to thank my working group and my colleagues from the department of mathematics and computer science in Marburg and from the faculty of mathematics and computer science in Heidelberg for the great working atmosphere, academical and non-academical discussions, technical support, motivation and many nice lunch and coffee breaks. It was a great time being a part of these groups.

Special thanks go to my mentor Regina Vollmeyer of the University of Frankfurt and my co-mentees Katerina, Nadege, Patricia and Rebecca for great advice, always listening to each other's problems and nice evening meetings with great discussions, tea and cake. Thank you Regina, for always meeting us on your free Saturdays. I also want to thank Mentoring Hessen (formerly SciMento hessenweit) for having me in their program, giving me the opportunity of meeting this great mentor, this wonderful peer-group and the other mentees and for organizing such useful meetings and workshops.

I also owe thanks to Amichai Dreifuss and Marta Sauter for reading parts of this thesis and giving great advice and helpful corrections.

Very special thanks go to my parents, my family and my family in law for always believing in me, supporting me when I needed them and motivating me. Last, I want to thank my husband Nicolas for the understanding, help and love in the last years, for laughing with me and being there for me when I was not in the best mood.

Contents

Abbreviations

CLOCP	**C**losed **L**oop **O**ptimal **C**ontrol **P**roblem
DP	**D**ynamic **P**rogramming
FJ	**F**ritz **J**ohn
KKT	**K**arush **K**uhn **T**ucker
MPC	**M**odel **P**redictive **C**ontrol
MPEC	**M**athematical **P**rogram with **E**quilibrium **C**onstraints
NLP	**N**on-**L**inear **P**rogram
OCP	**O**ptimal **C**ontrol **P**roblem
OLOCP	**O**pen **L**oop **O**ptimal **C**ontrol **P**roblem
RHC	**R**eceding **H**orizon **C**ontrol

Symbols

$\|\cdot\|_2$	l_2 norm $\left(\|x\|_2 = \sqrt{x^T x} = \sqrt{\sum_i x_i^2}\right)$
$\|\cdot\|_A$	norm defined as $\|y\|_S = y^T S y$
$f(x,u)$	right-hand side of dynamical equation
H	Hamiltonian of an optimal control problem
J	cost function of an optimal control problem
$L_2(\cdot)$	Hilbert space
λ	adjoint variable of the Hamiltonian
$\mathbb{N}$	natural numbers: $\{1, 2, 3, \dots\}$
$\mathbb{N}_0$	natural numbers including 0: $\mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\}$
π	control policy
$\bar{\pi}$	approximative control policy
π^0	exact optimal control policy
π^*	approximative control policy
$\pi(\lambda^*)$	approximative control policy
$\pi(x^0)$	approximative control policy
$\pi(\lambda^0, \alpha^0)$	approximative control policy for problems with bounded control

$\pi(\mathcal{T})$	approximative control policy for problems with state constraints
$\pi(\bar{W})$	approximative control policy for systems with bounded disturbances
t	time variable
t_0	initial time moment
t_*	final time moment of time interval
τ	time variable
T	time interval
T_i	time interval from t_{i-1} to t_i
$\mathcal{T}$	time interval
$\mathcal{T}_j$	time interval from τ_{j-1} to τ_j
u	control variable
$u(\cdot, z)$	control law
V_f	terminal penalty
V_N	MPC cost function
V_N^0	MPC optimal value function
$u_\infty^0(\cdot; x)$	optimal value function of an infinite horizon optimal control problem
V_∞	cost function of an infinite horizon optimal control problem
V_∞^0	solution of infinite horizon optimal control problem
$\phi(k; x, u)$	solution of difference equation at time k
w	uncertainty variable
W	class of admissible disturbances
x	state variable of nominal dynamical system
$\dot{x}$	derivative of variable x with respect to the time t ($\dot{x} = \frac{dx}{dt}$)
x_0	initial value of state variable of the nominal dynamical system
x_{k+1}	successor state
z	state variable of actual dynamical system with uncertainties

z_0	initial value of the state variable of the actual dynamical system with uncertainties
$\mathbb{Z}$	integer numbers: $\{\cdots-3,-2,-1,0,1,2,3,\dots\}$

1 Introduction

Mathematical models of real processes which arise in natural science, industry or engineering always contain errors such as statistical uncertainties of state and parameter estimates, model plant mismatch and discretization errors provided by simulation methods. Optimal solutions can be very sensitive to such errors. Furthermore, the realization of optimal scenarios including controls, initial values and design parameters in concrete cases may lead to situations, when the process controlled by nominal optimal control outruns some critical limits and becomes uncontrollable (e.g. explosions in exothermic chemical processes). This is why the application of mathematical optimization to real-life processes demands taking into account uncertainties in the process. To handle the uncertainties we combine in this thesis two mostly used approaches: robust optimization and online/feedback controls. In robust optimization, e.g. “worst-case” optimization, we are looking for an optimal solution which is “good” for all possible realizations of uncertain parameters in a compact set. Online/feedback control is an approach, where the nominal optimal control is updated as soon as the actual state and parameter estimate is available. For the feedback control we consider a current point (τ, x) where x is the current state at time τ. We solve an auxiliary predictive optimal control problem in Model Predictive Control (MPC) scheme (cf. Chapter 4) at some time intervals $[\tau, \tau+T]$ and obtain the solution of this problem as

$$u^0(t; \tau, x), \; t \in [\tau, \tau + T].$$

Then the feedback control at this point is

$$u^*(\tau, x) = u^0(\tau; \tau, x), \; \tau \geq 0, \; x \in \mathbb{R}^n.$$

Using different types of the auxiliary problem at the current time τ, there are various types of MPC schemes, e.g. open-loop optimal control, open-loop min-max optimal control and closed-loop min-max optimal feedback control. In this thesis we follow the *Closed-Loop Min-Max Optimal Feedback Control Approach* (cf. Lee & Yu [55]), which combines the ideas of robust optimization and MPC. This approach leads on the one hand to high computational costs but on the other hand it avoids infeasibility problems resulting from the application of robust optimization. The aim of this thesis is to develop approximative strategies to avoid high computational costs.

In this thesis we consider linear-quadratic optimal control problems with a bounded additive uncertainty. For constructing a guaranteed control strategy for this kind of problem, we correct the control strategy at several intermediate time points, depending on the measurements of the states. As the use of this strategy leads to high computational costs and is practically impossible, we introduce and analyze approximative control strategies, which are suboptimal but still yield a guaranteed policy. The computation of these approximative control policies is equivalent to solving certain convex resp. saddle point optimization problems which can be solved offline. That means we can formulate explicite rules to solve these problems and to construct the corresponding control policy of the original optimal control problem. These rules can be implemented online.

1.1 Contributions

Bilevel Optimization Problem with a Non-Homogeneous Cost Functional

We introduce a special bilevel optimization problem with a non-homogeneous cost functional and a trust-region-type constraint. This bilevel optimization problem is later used to describe practical algorithms for the approximative control policies. For this problem we describe and prove the optimality conditions and show how to solve the problem explicitly. We also formulate and prove a theorem that shows that the special bilevel optimization problem is equivalent to another convex optimization problem in one variable. We also show how to construct the optimal solution of a certain feedback linear quadratic control problem using the solution of the bilevel problem.

Bilevel Optimization Problem with an Additional Quadratic Constraint in the Upper Level Probem

Similar to the previous special bilevel programming problem we introduce this problem, show the optimality conditions, describe an explicite solution and an algorithm for solving and prove the equivalence to a saddle point computed of a convex/concave function.

Approximative Control Policies

Using special bilevel optimization problems, approximative control policies as alternative for the numerical demanding guaranteed optimal control policy can be constructed. We analyze different approximative control policies and compare them theoretically and with a numerical example.

Guaranteed Control for Systems with State Constraints

Using the special bilevel programming problems, we describe generalizations of the control strategies for a problem with state constraints and derive approximative control policies.

Bounded Control

For systems with bounded control we also use the special bilevel optimization problems described before and similarly to the guaranteed control for systems with state constraints the approximative control policies can be derived.

1.2 Outline of the Thesis

This thesis is divided into eight chapters and organized as follows.

- After the introduction we start in Chapter 2 with preliminaries which we later use in the thesis. We shortly introduce convex problems and their optimality conditions in the unconstrained and constrained case. We also give a short overview of solution methods. In Section 2.2 we consider bilevel programming problems and introduce linear and stochastic bilevel problems. We also shortly present an overview of solution methods.

- In Chapter 3 we introduce optimal control problems. We start with continuous optimal control problems, describe a basic problem and formulate Pontryagin's Maximum Principle for the necessary optimality conditions for the control to be optimal. Afterwards we show the application of the Maximum Principle on linear-quadratic optimal control problems. In the fourth part of this chapter, we give a short introduction to discrete optimal control problems and their solution methods. In the last part of this chapter we discuss robust optimization, present types of uncertainties, describe different method types for optimal control problems with uncertainties and give a literature overview of existing methods.

- In Chapter 4 we give a general introduction to Model Predictive Control (MPC). For this, we consider Model Predictive Control in the general case and the Dynamic Programming approach with its solution. Afterwards we present both approaches in the case of uncertainties. We end this chapter with a short section about Feedback Model Predictive Control and the theory of tubes in the case of uncertainties.

- In Chapter 5 we introduce and analyze different types of special bilevel optimization problems which we need to describe the algorithms in Chapters 6 and 7. In Section 5.1 we describe a bilevel optimization problem with a trust-region-type constraint, formulate the optimality conditions, present an explicit solution and show that this problem is equivalent to a certain optimal control problem. Afterwards we similarly describe this also for a bilevel optimization problem with non-homogeneous cost functional and a bilevel optimization problem with an additional quadratic constraint in the upper level problem.

- In Chapter 6 we have a closer look at the optimal control under uncertainties. In Section 6.1 we examine the closed-loop min-max optimal feedback control with one correction moment in more detail. We consider a linear quadratic optimal control problem with an additive unknown but bounded uncertainty. In Section 6.2 we show how to use the bilevel programming problem from Chapter 5 to reformulate this approach and derive a practical algorithm for solving the problem. In Section 6.3 we also construct the robust optimal feedback with more than one intermediate correction point. As the computational costs of this algorithm are high we present and analyze approximative approaches of the closed-loop min max optimal feedback control with intermediate correction moments in Section 6.4 and compare them in Section 6.5. We also formulate practical algorithms for the approximations. At the end of the chapter we show a short numerical example to compare the approximative policies using an application.

- In Chapter 7 we describe three different generalizations of the approximative control policies which we present in Chapter 6. First we present a strategy in which we can control dynamic systems with pointwise state constraints and derive here two different approaches. Then we present the strategy for optimal control problems with bounded controls. In Section 7.4 we discuss a strategy for control problems with uncertainties in which the disturbance is bounded in a norm. We analyze the problems and discuss the corresponding algorithms.

- We end this thesis with a conclusion, in which we summarize the findings from the previous chapters and give an outlook on future developments.

2 Static Optimization

In this chapter we give a short overview about convex optimization and bilevel programming. These are preliminaries which we will need later in this thesis as basic information. In the first part of this chapter we consider convex optimization problems and their optimality conditions in the unconstrained and constrained case. After that we give a short overview of solution methods. In the second part of this chapter we consider different types of bilevel programming problems and give a short overview of solution methods.

2.1 Convex Optimization

We consider the following convex minimization problem

$$\begin{aligned} \min_{x \in D} \;& f(x) \\ \text{s.t. } & g(x) = 0 \\ & h(x) \geq 0 \end{aligned} \tag{2.1}$$

with $x \in D \subset \mathbb{R}^n$, $f : D \subset \mathbb{R}^n \to \mathbb{R}$, $g : D \subset \mathbb{R}^n \to \mathbb{R}^l$ $(l < n)$ and $h : D \subset \mathbb{R}^n \to \mathbb{R}^k$ convex on D and $D \subset \mathbb{R}^n$ a domain. A convex minimization problem has the following three properties:

- Each local minimum is always also a global minimum of the convex problem.
- The set of all minima is convex.
- If the cost functional f is strictly convex, then the minimum is unique.

These and further information can be found in the work of Avriel [3], Boyd & Vandenberghe [17], Ben-Tal & Nemirovski [9], Jarre & Stoer [41], Nocedal & Wright [67].

2.1.1 Optimality Conditions Unconstrained Case

In the case of an unconstrained optimization problem we have $l = k = 0$, which means our convex optimization problem is formulated as follows

$$\min_{x} \; f(x) \tag{2.2}$$

with $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ a convex and sufficiently differentiable function. Since f is differentiable and convex we can formulate the necessary and sufficient condition for a point x^* to be optimal as

$$\nabla f(x^*) = 0. \tag{2.3}$$

That means, for solving (2.2) we have to find a solution of (2.3). In a few special cases we can find this solution by solving (2.3) analytically. But usually we have to use an iterative algorithm to find a solution. We describe methods for solving this problem by an iterative method in Subsection 2.1.3.

2.1.2 Optimality Conditions Constrained Case

With the definition of the Karush-Kuhn-Tucker Condition we obtain first order necessary optimality condition for nonlinear optimization problems.

Definition 2.1. (Karush-Kuhn-Tucker Condition / Karush-Kuhn-Tucker Point)
A point (x^*, λ^*, μ^*) is called **Karush-Kuhn-Tucker Point** (KKT-Point) if the following conditions are satisfied

$$\begin{aligned}
&\nabla f(x^*) + \sum_{i=1}^{l} \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^{k} \mu_j^* h_j(x^*) = 0 \\
&g(x^*) = 0 \\
&h(x^*) \geq 0 \\
&\mu_i^* \geq 0 \\
&\mu_i^* h(x^*) = 0,
\end{aligned} \tag{2.4}$$

with $f(x)$, $g_i(x)$, $h_j(x)$ continuously differentiable. The conditions (2.4) are called **KKT-Conditions**.

The necessary optimality conditions are then formulated as

Theorem 2.2. *(Necessary Optimality Conditions)*
If x^* is a minimum of the convex problem (2.1) and also fulfills certain constraint qualifications, then there exist λ^* and μ^* such that (x^*, λ^*, μ^*) is a KKT-point.

Proof. A proof can be found in Nocedal & Wright [67]. □

Possible Constraint Qualifications are:

- **Linear Independence Constraint Qualification** (LICQ): Requires the linear independence of the gradients of all equality and active inequality constraints.
- **Mangasarian-Fromovitz Constraint Qualification** (MFCQ): MFCQ holds if there exists a vector w such that $\nabla g(x^*)^T w = 0$ for all equality constraints and $\nabla h(x^*) w > 0$ for all active inequality constraints.

A more general formulation of the KKT-Conditions are the Fritz-John Conditions (FJ Conditions) as the constraint qualifications are no longer needed.

Definition 2.3. (Fritz-John Condition / Fritz-John Point)
A point $(z^*, x^*, \lambda^*, \mu^*) \in \mathbb{R}^{1+n+l+k}$ of the optimization problem (2.1) is called **Fritz-John point** (FJ point) if the following conditions are satisfied

$$\begin{aligned} & z^*\nabla f(x^*) + \sum_{i=1}^{l} \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^{k} \mu_j^* \nabla h_j(x^*) = 0 \\ & g(x^*) = 0 \\ & h(x^*) \geq 0 \\ & \mu_i^* \geq 0 \\ & \mu_i^* h(x^*) = 0 \\ & z^* \geq 0. \end{aligned} \tag{2.5}$$

The conditions (2.5) are called **Fritz-John Conditions** (FJ Conditions).

And as before, if x^* is a local minimum of the optimization problem (2.1), there exist λ^*, μ^*, z^* such that $(z^*, x^*, \mu^*, \lambda^*)$ is a FJ-Point and (z^*, μ^*, λ^*) is different from the null vector.

Theorem 2.4. *(Sufficient Optimality Conditions)*
If (x^*, λ^*, μ^*) is a KKT-point, then x^* is a global minimum of the convex problem (2.1).

Proof. A proof can be found in Jarre & Stoer [41]. □

2.1.3 Solution Methods

For unconstrained convex optimization problems we have to find a solution of (2.3). For this we have the initial value $x_0 \in D$, the search direction in the k-th iteration step $d^k \in \mathbb{R}^l$, the step length in the k-th iteration step $t^k \in \mathbb{R}$ and the iteration step $x^{k+1} = x^k + t^k + d^k$.

Definition 2.5. Let f be continuously differentiable and $x \in \mathbb{R}^n$. A vector $d \in \mathbb{R}^n$ is a decent direction of f in x, if there exists a $\bar{t} > 0$ with $f(x + td) < f(x)$ for all $t \in]0, \bar{t}]$.

We also need the following lemma

Lemma 2.6. *Let f be continuously differentiable, $x \in \mathbb{R}^n$ and $d \in \mathbb{R}^n$ with $\nabla^T f(x)d < 0$. Then, d is a descent direction of f in x.*

Proof. Determine the directional derivative

$$f'(x, d) = \lim_{t \to +0} \frac{f(x + td) - f(x)}{t} = \lim_{t \to +0} \frac{\nabla f(\bar{x})d}{t} = \nabla^T f(x)d < 0 \text{ for } \bar{x} \in [x, x + td].$$

Hence, $\frac{f(x+td)-f(x)}{t}x < 0$ for all sufficiently small t. □

With this, we can now formulate the following algorithm for descent methods. We assume to have an initial value x^0 for the iteration step $k = 0$.

1. Check, if x^k satisfies a stop criterion.
2. Determine the descent direction d^k.
3. Determine the step length t^k with $f(x^k + t^k d^k) < f(x^k)$.
4. Set $x^{k+1} = x^k + t^k d^k$ and $k = k + 1$ and and proceed with step 1.

Further information can for instance be found in Boyd & Vandenberghe [17], Avriel [3] and Nocedal & Wright [67].

In the constrained case, in which we have equality and inequality constraints, we can use*Interior-Point Methods*, where we apply Newton's method to a sequence of equality constrained problems or to a sequence of modified versions of the KKT conditions. Possible types of interior point methods are the *Barrier Method* or the *Primal-Dual Interior-Point Method*. Other possible methods can be SQP-methods or augmented Lagrangian methods. Further information to these methods can for instance be found in Boyd & Vandenberghe [17], Nocedal & Wright [67] and Ye [85].

2.2 Bilevel Optimization

A *Bilevel Programming Problem* (BPP) is a hierarchical mathematical optimization problem. The characteristic of a BPP is the formulation of two mathematical programs in one problem. The constraints of the upper level optimization problem are in part defined by a lower level optimization problem. Bilevel optimization problems are nonconvex and nondifferential optimization problem. This section is based on the work of Alizadeh *et al.* [1], Bard [4], Bracken and McGill [18],[19], Colson et al. [25], Dempe [28], [29], [30], Dempe *et al.* [31], Mersha & Dempe [64] Sinha *et al.* [77], Starr & Ho [79] and Wiesemann *et al.* [83].
A general BPP, which is also called *leader's problem*, is stated as follows

$$\begin{aligned} \min_{y \in Y} \;& F(x,y) \\ \text{s.t. } & G(x,y) \leq 0 \\ & H(x,y) = 0 \\ & x \in \Psi(y) \end{aligned} \tag{2.6}$$

with $F : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}$, $G : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_G}$, $H : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_H}$. The function $\Psi(y)$, with $\Psi : \mathbb{R}^{n_y} \to \mathbb{R}^{n_x}$ is the solution set of the *lower level problem*, which is also called *follower's problem* and is formulated as follows

$$\min_{x \in X} \; f(x,y) \tag{2.7}$$

$$\text{s.t. } g(x,y) \leq 0 \tag{2.8}$$

$$h(x,y) = 0 \tag{2.9}$$

with $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}$, $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_g}$, $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_h}$, $x \in \mathbb{R}^{n_x}$, $y \in \mathbb{R}^{n_y}$, $g(x,y) = (g_1(x,y), \ldots, g_{n_g}(x,y))^T$ and $h(x,y) = (h_1(x,y), \ldots, h_{n_h}(x,y))^T$. We assume that the functions F, G, H, f, g and h are sufficiently smooth. We also assume that x is a unique choice for any y. The function F is called the upper level

objective function and the functions G and H are called upper level constraint functions. Typical applications of bilevel optimization can be found in the fields of transportation, engineering or economics.
In the case that the lower level problem has multiple solutions for some of the selections of the upper level decisions maker. Then the leader may assume that the follower can be motivated to choose a best optimal solution in $\Psi(y)$ with respect to the leader's objective function. This is called the *optimistic* formulation of the BPP

$$\begin{aligned} \min_{y} \;\; & \phi_o(y) \\ \text{s.t.} \;\; & G(x,y) \leq 0 \\ & H(x,y) = 0 \\ & y \in Y \end{aligned}$$

where

$$\phi_o(y) = \min_{x}\{F(x,y) : x \in \Psi(y)\}.$$

This might not be possible or even not allowed. Then the leader has to bound the damage that might arise from a choice of the follower. In this case we have a *pessimistic* formulation of the BPP

$$\begin{aligned} \min_{y} \;\; & \phi_p(y) \\ \text{s.t.} \;\; & G(x,y) \leq 0 \\ & H(x,y) = 0 \\ & y \in Y \end{aligned}$$

where

$$\phi_p(y) = \max_{x}\{F(x,y) : x \in \Psi(y)\}.$$

BPPs can be reformulated into usual optimization problems. For instance by using the KKT-conditions of the lower level problem. If the functions $f(x,y)$, $g_i(x,y),\ i = 1,\ldots,n_g$ and $h_j(x,y),\ j = 1,\ldots,n_h$ are differentiable and a regularity condition is satisfied then we can replace the BPP (2.6) by

$$\begin{aligned} \min_{x,y,\lambda,\mu} \;\; & F(x,y) \\ \text{s.t.} \;\; & G(x,y) \leq 0 \\ & H(x,y) = 0 \\ & y \in Y \\ & \nabla_x\{f(x,y) + \lambda^T g(x,y) + \mu^T h(x,y)\} = 0 \\ & \lambda \geq 0,\ \mu \geq 0, \\ & g(x,y) \leq 0,\ \lambda^T g(x,y) = 0 \\ & h(x,y) = 0 \end{aligned}$$

This transformation is only possible, if the lower level problem is convex. This kind of problem is called *mathematical program with complimentarity constraints* (MPCC).

Other possibilities to transform the BPP into an usual optimization problem are to use necessary optimality conditions without Lagrange multipliers (cf. Dempe & Zemkoho [32]) or to use variational inequalities (cf. Kalashnikov & Kalashnikova [43]).

2.2.1 Linear Bilevel Optimization Problem

In the linear case the lower level problem is given as

$$\begin{aligned} \min_x \ & c^T x \\ \text{s.t. } & A_1 x \leq a - A_2 y \\ & x \geq 0 \end{aligned}$$

with $x \in \mathbb{R}^{n_x},\ y \in \mathbb{R}^{n_y},\ a \in \mathbb{R}^p$ and the matrices A_1 and A_2 as well as the vector c are of appropriate dimensions. This linear problem can be obtained by assuming that all functions in problem (2.6) are restricted to be affine. We define $\Psi_L(y)$ by

$$\arg\min_x \{c^T x : A_1 x \leq a - A_2 y,\ x \geq 0\}$$

and hence, we can formulate the linear bilevel optimization problem as

$$\begin{aligned} \min_y \ & d_1^T x + d_2^T y \\ \text{s.t. } & A_3 y = b \\ & y \geq 0 \\ & x \in \Psi_L(y) \end{aligned}$$

with $b \in \mathbb{R}^l$ and all other dimensions are chosen such that they match with the above ones.

2.2.2 Stochastic Bilevel Optimization Problem

A stochastic bilevel optimization problem is formulated as

$$\begin{aligned} \max_{x,y} \ & F_1(x,y) + Q(x,y) \\ \text{s.t. } & G_1(x,y) \leq 0 \\ \min_y \ & f_1(x,y) \\ \text{s.t. } & g_1(x,y) \leq 0 \end{aligned}$$

where $Q(x,y) = E_\xi[\Phi(x,y,\xi)],\ \boldsymbol{\xi} : \Omega \to \mathbb{R}^r$, $\boldsymbol{\xi}$ being a discrete random vector with realizations ξ and support Ξ and $(\Omega, \mathcal{F}, P)$ being a probability space where Ω denotes the set of all random events and $\mathcal{F}$ the set of all subsets of Ω. For any outcome $\xi(\omega) \in \Xi$

$$\begin{aligned} \Phi(x, y, \xi(\omega)) = \max_{x'(\omega)} \;& F_2(x'(\omega), y'(\omega), \omega) \\ \text{s.t.} \;& G_2(x'(\omega), y'(\omega), x, y, \omega) \leq 0 \\ \min_{y'(\omega)} \;& f_2(x'(\omega), y'(\omega), \omega) \\ \text{s.t.} \;& g_2(x'(\omega), y'(\omega), x, \omega) \leq 0 \end{aligned}$$

(cf. Alizadeh *et al.* [1]). In this stochastic problem the first stage decisions are made before the random outcome at the second-stage are observed.

2.2.3 Overview of existing solution methods

There are different possibilities to formulate optimality conditions for bilevel programming problems. One approach is for instance to replace the lower level problem by its Karuh-Kuhn-Tucker conditions and, by applying the optimality conditions for locally Lipschitz continuous problems to the problem, we can get F. John resp. Karush-Kuhn-Tucker type necessary optimality conditions for the bilevel problem which works for general bilevel programming problems. This approach works only in the optimistic case. Another approach, which works for the optimistic as well as the pessimistic case is to denote a region of stability for linear parametric optimization where the optimal basis of the problem remains optimal. From this local optimality of the bilevel problem it follows that the cost function value cannot decrease on each of the regions of stability related to optimal basic solutions of the lower level problem for the parameter under consideration.

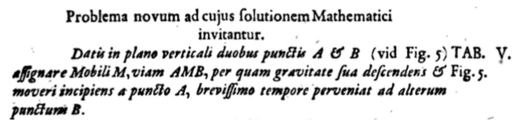
In the literature (eg. Ben-Ayed & Blair [8] and Bard [4]) it often times discussed that solving a bilevel linear programming problem is $\mathcal{NP}$-hard, which colloquially means that the problem cannot in general be solved with a polynomial-time algorithm. As this is not the main topic of this thesis, we will only give a brief overview of some solving algorithms without going into further detail. We can distinguish between enumerative algorithms, descent algorithms and penalty function methods. As bilevel programming problems are nonconvex and nondifferentiable, it is also a task to search for global solutions. As the descent algorithms and the penalty function methods rather search for local optima, we might use enumerative algorithms to search for global solutions. Using an enumerative algorithm we can search within the vertices of the feasible set. This method is based on the optimistic problem and can yield a polynominal algorithm for the linear bilevel programming problem. Further information to this approach can be found e.g. in Bard & Falk [5] and Liu & Spencer [58]. Another approach in the sense of an enumerative algorithm is the search for active inequalities. Here, by formulating additional inequalities with artificial variables for the lower level problem one gets a branch-and-bound algorithm for the search for a global optimistic solution. Further information to this approach can be found in Hansen *et al.* [39]. The descent algorithms are rather used to find local optima for the linear bilevel programming problem. In this approach the necessary optimality conditions for the problem are implemented. This works both for the optimistic and the pessimistic case. To this approach further information can be found in Dempe [27]. Using the penalty function methods a penalty function is added to the problem and the new problem is solved by a decomposition approach. The problem here is that for each value of the penalty parameter the outer nonconvex problem has to be

solved globally. But there are already extensions for this approach where the upper level problem is solved locally but the linear lower level problem is solved globally. Further information to this approach can be found in White & Anandalingam [82].
The branch-and-bound algorithm also works for general bilevel programming problems. Here, the root node of the tree corresponds to the convex lower level problem which is replaced by its KKT-conditions without the complementary constraints. Then lower bounds of can be constructed. Further information to this can be found in eg. Bard & Falk [5]. With penalty functions methods nonlinear bilevel programming problems can be solved. Here, the lower level problem is replaced by a penalized problem and the optimal solution of the new problem converges to the solution of the original bilevel programming problem. Further information to this approach can be found in e.g. Shimizu & Aiyoshi [76]. Another approach to solve nonlinear or linear quadratic bilevel programming problems is the trust-region method. It is an iterative algorithm which is based on the approximation of the original problem by a model around the current iterate. Further information to this method can be found in e.g. Colson *et al.* [26].

3 Optimal Control

Optimal Control is a field of optimization in which dynamic systems are optimized. Optimal Control Problems (OCP) go back to the calculus of variations and the year 1696 (cf. Sussmann & Willems [81]). In June 1696 Johann Bernoulli published the Brachistochrone Problem, which is Greek for the Shortest Time Problem, in Acta Eruditorum as a challenge or an invitation, resp., for all mathematicans. This is the problem as it was formulated originally in 1696 (cf. Acta Eruditorum, June 1696, p. 269): There

Brachistochrone Problem
If two points A and B are given in a vertical plane, assign a path from A to B such that the moving point M from A, under the influence of its own weight, arrives at B in the shortest possible time.


Problema novum ad cujus ſolutionem Mathematici invitantur.

Datis in plano verticali duobus punctis A & B (vid Fig. 5) TAB. V. *aſſignare Mobili M, viam AMB, per quam gravitate ſua deſcendens &* Fig. 5. *moveri incipiens a puncto A, breviſſimo tempore perveniat ad alterum punctum B.*

FIGURE 3.1: The original task that was published in Acta Eruditorum, June 1696 on page 269 by Johann Bernoulli.

were several solutions published by famous mathematicans as Johann Bernoulli himself and his elder brother Jakob Bernoulli, Marquis de L'Hospital and Ehrenfried Walter von Tschirnhausen. Gottfried Wilhelm Leibniz also had a solution but he renounced his publication since his solution was very similar to that of Johann Bernoulli. One of the solutions was published anonymously but Johann Bernoulli identified it later as a solution of Issac Newton.
At this time, Bernoulli already said, that these kinds of problem are not only useful for mechanics, but also for other fields of science. After this challenge a lot of work was done especially by Euler, Lagrange, Hamilton and Weierstrass.

Sussmann & Willems [81] give several reasons, why the Brachistochrone Problem is the birth of the optimal control:

- It is a true minimum time problem of the kind that is studied today in optimal control theory.
- It is the first problem ever that dealt with a dynamical behavior and explicitly asked for the optimal selection of a path.
- A huge part of the calculus of variations, which is based on this problem is the search for the simplest and most general statement of the necessary conditions for optimality. These necessary conditions can nowadays be found in the Maximum Principle of the optimal control theory.

In 1956 L.S. Pontryagin *et al.* [68] formulated the Maximum Principle. It gives us the necessary optimality conditions for the trajectory of an optimal control problem to be optimal and is said to be the birth of the optimal control theory. Since then the

Maximum Principle was improved by several authors using weaker hypotheses, stronger conclusions or generalizations, like Clarke [24] or Sussmann [80].

The aim of discussing *Optimal Control Problems* (OCP) is to determine a control or control law $u(t)$ and an associated state variable $x(t)$ for a dynamical system in order to optimize a given objective functional over a period of time. Nowadays most of the OCPs have too many variables and are too complex to solve them analytically, hence we need numerical methods.

3.1 Continuous Optimal Control

Optimal Control Problems (OCP) are used in different fields of natural science, engineering or economics. The aim of discussing an OCP is to find a control or a control law $u(t)$ and an associated state variable $x(t)$ to optimize a given objective functional $J[x(\cdot), u(\cdot)]$ under certain constraints. The results in this and the following sections are based on the work of Betts [13], Bryson and Ho [20], Gerdts [35], Lewis et al. [56], Kwakernaak and Sivan [52], Rodrigues et al. [71], Schättler and Ledzewicz [73], Sethi and Thompson [75] and further information can be found there.

A simple example for these kinds of problems is a car driver. We consider two cars. One car driver has the aim to reach his destination as fast as possible while another one tries to drive as ecologically as possible. When is the best time to shift? The different aims are formulated in the objective functions and the constraints describe weather, speed limit etc. that may influence.

The general formulation of an OCP is as follows. We have the model equation

$$\dot{x}(t) = f(t, x(t), u(t))$$

where $x : [t_0, t_*] \to \mathbb{R}^{n_x}$ are the states, $u : [t_0, t_*] \to \mathbb{R}^{n_u}$ are the control or control laws and t_0 and t_* can be fixed or free. Then we have the initial value condition

$$x(t_0) = x_0$$

and the boundary conditions

$$\begin{aligned} r(t_0, x(t_0), t_1, x(t_1), \dots, t_*, x(t_*)) &= 0 \\ r(t_0, x(t_0), t_1, x(t_1), \dots, t_*, x(t_*)) &\geq 0. \end{aligned}$$

The control $u(t)$ is bounded in the domain $\Omega \subseteq \mathbb{R}^{n_u}$. There are three different standard types of the cost functional of an OCP. The Mayer type

$$\min_{u(\cdot)} \Phi(t_*, x(t_*)),$$

the Lagrange type

$$\min_{u(\cdot)} \int_{t_0}^{t_*} L(t, x(t), u(t))dt$$

and the Bolza type

$$\min_{u(\cdot)} \Phi(t_*, x(t_*)) + \int_{t_0}^{t_*} L(t, x(t), u(t))dt.$$

Remark 3.1. *All three types of cost functionals are theoretically equivalent, in the sense that hat each type of cost functional can be converted to any other cost functional (cf. Chachuat [22]).*

Altogether we consider the following general OCP

$$\min_{u(\cdot)} \int_{t_0}^{t_*} L(t, x(t), u(t))dt + \Phi(t_*, x(t)) \tag{3.1a}$$

$$\text{s.t. } \dot{x}(t) = f(t, x(t), u(t)) \tag{3.1b}$$

$$u \in \Omega \tag{3.1c}$$

$$x(t_0) = x_0 \tag{3.1d}$$

$$r(t_*, x(t_*)) = 0 \tag{3.1e}$$

with $t \in [t_0, t_*]$.

Definition 3.2. A control u is called feasible, if there exists a x on $[t_0, t_*]$ such that (u, x) satisfies the conditions (3.1b) - (3.1e). The control u^* is called optimal if u^* is feasible and (u^*, x^*) minimizes the cost functional (3.1a).

3.2 Pontryagin's Maximum Principle

As we already stated at the beginning of this chapter in most cases OCPs can not be solved analytically, because they have too many variables, are too complex or even both. Therefore we have to use numerical methods to solve OCPs. We will consider two different types of numerical solution methods. On the one hand there are the direct methods. In this case we first discretize the OCP to reduce it to a nonlinear constrained optimization problem. The nonlinear constrained optimization problem can then be solved for instance with Quasi-Newton methods (e.g. Nocedal & Wright [67]). On the other hand we can use indirect methods, where the solution is based on Pontryagin's Maximum Principle. In the case of indirect methods the problem is reduced to a boundary value problem which can be solved for instance with Multiple Shooting (e.g. Bock [15]). In 1956 L.S. Pontryagin *et al.* [68] developed the *Maximum Principle of Optimal Control* which gives us necessary optimality conditions for the trajectory (x, u) to be optimal (cf. Rodrigues et al. [71] and Schättler [73]).

Theorem 3.3. *(Pontryagin's Maximum Principle)*
Let u^* be an optimal control with corresponding state $x*$. Define

a) Hamiltonian

$$H(t, x, u, \gamma, \lambda) := -\gamma L(t, x, u) + \lambda^T f(t, x, u)$$

b) Extended Cost Functional

$$\Psi(t_*, y(t_*), \alpha) := \gamma\Phi(t_*, y(t_*)) + \alpha^T r(t_*, y(t_*)$$

with $\lambda \in \mathbb{R}^{n_x}$, $\alpha \in \mathbb{R}^{n_r}$, $\gamma \in \mathbb{R}$ and without loss of generality we can assume that $\gamma = 1$ if $\gamma \neq 0$.

Then it holds, that there exist $\gamma \geq 0$, $\alpha \in \mathbb{R}^{n_r}$, $\lambda : [t_0, t_*] \to \mathbb{R}^{n_x}$ with $(\alpha, \lambda(t)) \neq 0$ such that

1. *The adjoint variable $\lambda(t)$ satisfies the adjoint differential equation*

$$\begin{aligned}\dot{\lambda}(t)^T &= \gamma L_x(t, x^*(t), u^*(t)) - \lambda(t)^T f_x(t, x^*(t), u^*(t)) \\ &= -H_x(t, x^*(t), u^*(t), \lambda(t))\end{aligned}$$

2. *Transversality Conditions:*

$$\begin{aligned}\lambda(t_*)^T &= -\Psi_x(t_*, x^*, x^*(t_*), \alpha \\ H(t_*, x^*(t_*), u^*(t_*), \lambda(t_*) &= \Psi_{t_*}(t_*, x^*(t_*), \alpha)\end{aligned}$$

if t_* is free.

3. *The control u^* satisfies the Maximum principle*

$$H(t, x^*(t), u^*(t), \lambda(t)) \geq H(t, x^*(t), \nu(t), \lambda(t))$$

with $\nu \in \Omega$ arbitrary almost everywhere on $[t_0, t_*]$. That means, u^* is the solution of

$$\max_{\nu \in \Omega} H(t, x^*(t), \nu(t), \lambda(t)) = H(t, x^*(t), u^*(t), \lambda(t))$$

almost everywhere on $[t_0, t_*]$.

Proof. First we show the theorem for the following problem

$$\begin{aligned}\min_{u,x} \quad & \int_{t_0}^{t_*} L(t, x(t), u(t))dt + \Phi(t_*, x(t)) =: J(x, u) \\ \text{s.t.} \quad & \dot{x} = f(t, x(t), u(t)), \ x(t_0) = x_0, \ t \in [t_0, t_*] \\ & u \in \Omega,\end{aligned}$$

that means without boundary conditions r. The basic idea is to construct a feasible variation of the optimal solution u^* with the model response x^*:

$$\begin{aligned} \bar{u}(t) &:= u^*(t) + \delta u(t) \\ \bar{x}(t) &= x^*(t) + \Delta x. \end{aligned}$$

We choose δu, get $\bar{u}$ and insert this in the differential equation to get $\bar{x}$. Hence, $\delta x = \bar{x} - x^*$ and δy solves the differential equation

$$\delta \dot{x} = \dot{\bar{x}} - \dot{x}^* = f(t, x^* + \delta x, u^* + \delta u) - f(t, x^*, u^*)$$

with $\delta x(t_0) = 0$. With the stability theorem (cf. Heuser [40]) we get

$$\|\delta x\| = O(\|\delta u\|).$$

For the boundary conditions we consider

$$0 \leq J(\bar{x}, \bar{u}) - J(x^*, u^*) + \int_{t_0}^{t_*} \lambda(t)^T (\dot{\bar{x}} - f(t, \bar{x}, \bar{u}) - x^* + f(t, x^*, u^*))dt + \alpha^T (\underbrace{r(t_*, \bar{x}(t_*)) - r(t_0, x^*(t_*))}_{=0}).$$

Here, u^*, x^* are optimal and $\bar{u} = u^* + \delta u$ and $\bar{x} * x^* + \delta x$ are feasible variations. We use the Taylor expansion

$$\begin{aligned} 0 \leq & \left(\frac{\partial \Phi}{\partial x}(t_*, x^*(t_*) + \lambda(t_*)^T \right) \delta x(t_*) \\ & + \int_{t_0}^{t_*} \left(-\dot{\lambda}(t)^T - \frac{\partial H}{\partial x}(t, x^*, u^*, \lambda) \right) \Delta x(t) dt \\ & + \int_{t_0}^{t_*} H(t, x^*, u^*, \lambda) - H(t, x^*, \bar{u}, \lambda) dt \\ & + \alpha^T \frac{\partial r}{\partial x}(t_*, x^*(t_*)) \delta x(t_*) \\ & + \text{terms of higher order.} \end{aligned}$$

With the transversality condition

$$\lambda(t_*)^T = -\frac{\partial \Psi}{\partial x}(t_*, x(t_*), \alpha) = -\frac{\partial \Phi}{\partial x}(t_*, x(t_*)) - \alpha^T \frac{\partial r}{\partial x}(t_*, x(t_*))$$

and the adjoint differential equation we get the maximum principle.
The proof of the existence of λ and α is not shown here. The proof of the original Maximum Principle can be found in Pontryagin *et al.* [68]. □

3.3 Application: Linear Quadratic Optimal Control Problem

As we show the application of the approaches later in this thesis on linear-quadratic control problems, we consider now the following linear dynamic process

$$\dot{x} = Ax + Bu$$

with $x \in \mathbb{R}^{n_x},\ u \in \mathbb{R}^{n_u}$ and the quadratic cost functional

$$J(t_0) = \frac{1}{2}x^T(t_*)Sx(t_*) + \frac{1}{2}\int_{t_0}^{t_*}(x^TQx + u^TRu)dt$$

with $t \in [t_0, t_*]$, the fixed final time t_*, $S,\ Q$ being symmetric and positive semi-definite and R being symmetric and positive definite. For this problem we have the following Hamiltonian

$$H(t) = \frac{1}{2}(x^TQx + u^tRu) + \lambda^T(Ax + Bu)$$

where $\lambda(t) \in \mathbb{R}^{n_x}$ is an undetermined multiplier. The state and costate equations are

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial \lambda} = Ax + Bu \\ -\dot{\lambda} &= \frac{\partial H}{\partial x} = Qx + A^T\lambda \end{aligned}$$

and the stationary condition is

$$0 = \frac{\partial H}{\partial u} = Ru + B^T\lambda. \tag{3.2}$$

Hence, by solving (3.2) we get the following term for the optimal control

$$u(t) = -R^{-1}B^T\lambda(t).$$

3.4 Discrete Optimal Control

A discrete time optimal control problem is formulated as

$$\begin{aligned} \min_{x,u}\ & \sum_{k=0}^{N} L(x_k, u_k) + \Phi(x_N) \\ \text{s.t.}\ & x_{k+1} = f(x_k, z_k),\ k = 1, \ldots, N-1 \\ & h(x_k, u_k) \leq 0,\ k = 1, \ldots, N-1 \\ & r(x_0, x_N) = 0. \end{aligned}$$

with the controls $u_k \in \mathbb{R}^{n_u}$, the states $x_k \in \mathbb{R}^{n_x}$, the time-invariant dynamical system x_{k+1}, the path constraints $h(x_k, u_k)$ and the boundary constraint function $r(x_0, x_N)$.

There are two important approaches to solve this OCP numerically. With the so-called simultaneous approach we use a structure exploiting nonlinear program (NLP) solver, as we have a large and structured nonlinear program. Possible simultaneous approaches are direct multiple shooting and direct collocation. Another important approach is to reduce the variable space of the NLP by eliminating almost all states by a forward simulation. As result we obtain a reduced problem, which has much less variables. This problem can now be solved by Newton-type methods. This approach is called sequential approach, because the simultation problem and the optimization problem are solved sequentially. These and further information can be found in Betts [13], Bock & Plitt [16].

3.5 Robust Optimal Control

Nowadays, almost every process or dynamic process in industry, engineering, natural science or the like are based on mathematical models which are optimized under certain assumptions. With these models we can simulate processes and gain new information about unknown meachanisms within these processes. Since the real systems are highly complex it is impossible to model all the details and therefore almost every model contains certain approximations, uncertainties, disturbances and assumptions. They can arise during the modeling process due to unknown or unmodeled effects, there might be an additive or multiplicative disturbance that is unknown, or the state of the system may be disturbed. During the measurements of a real process observational errors can occur. And we can obtain discretization errors, which appear while we discretize a continous (i.e. an infinite dimensional) mathematial model into a discrete (which means a finite dimensional) computer model (cf. Rawlings & Mayne [69]). But even though there are uncertainties in these models the goal of robust optimization is to guarantee that certain significant properties are still fulfilled under the influence of uncertainties or inexactness. This means that the models have to consider the uncertainties, disturbances or the like in the model formulation.

3.5.1 Types of Uncertainties

As we just stated systems usually contain uncertainties. In this subsection we will discuss three different types of uncertainties: the unknown additive uncertainty, the not perfectly known system state and the inaccurate model of the system. This subsection is based on the work of Rawlings & Mayne [69].

- **Additive Uncertainty:** In the case of a, usually bounded, additive uncertainty the following equation holds

  $$x_{k+1} = f(x, u) + w$$

  with the uncertainty w. An unbounded uncertainty would make it impossible to guarantee that the state and control constraints can be satisfied. As we will clarify later in this thesis the uncertainty w is in $W \subset \mathbb{R}^n$ which contains the origin. This is the type of uncertainty which we will consider in this thesis.

- **Unknown System State:** If we cannot measure the state directly we have to distinguish between two cases. If we have stochastic optimal control, we have the output $y = Cx + v$ with the measurement noise v. The measurement noise v is assumed to be white noise Gaussian processes, i.e. the values at any pair of times are identically distributed and statistically independent and therefore uncorrelated. Then the state of this optimal control problem is the conditional density of the state x at time k given prior measurements $\{y(0), y(1), \dots, y(k-1)\}$. As the density is very difficult to compute and in the linear case the density can be provided by the Kalman filter, often a suboptimal procedure is used. In this approach, which is called *certainty equivalence* the state x is replaced by an estimate $\hat{x}$ in a control law, which we get by assuming that the state is accessible. In the other case we have a linear function $f(\cdot)$ and we denote the state estimate $\hat{x}$ as

$$\hat{x}_{k+1} = f(\hat{x}, u) + \xi$$

with ξ being the innovation process. The actual state x, which lies in a bounded, possibly time-varying neighborhood of $\hat{x}$ has to satisfy the constraints of the optimal control problem.

- **Inaccurate Model:** If the model of a system, that is supposed to determine the control is uncertain, i.e. we have a parametric uncertainty, we use

  $$x_{k+1} = f(x, u, \theta)$$

  where θ is an unknown parameter which belongs to a compact set Θ.

3.5.2 Approaches for Optimal Control Problems with Uncertainties

In this thesis we will distinguish three different approaches to handle optimal control problems with uncertainties:

1. **Open-loop optimal control** The robustness is analyzed in the case that the uncertainties are neglected in the predictive optimal control problem. This means we just ignore the disturbances. The disadvantage of this method is, that in most cases we just obtain poor results (cf. Rawlings et al. [70] and Scokaert and Mayne [74]).

2. **Open-loop min-max optimal control** In the case of open-loop min-max optimal control we obtain a control as solution for the predictive optimal control problem that has the best value of our cost functional while considering all possible uncertainties that could occur. Hence, we consider the worst-case, that is why we also call this approach open-loop *worst-case* optimal control. This approach does not include any feedback aspects, such that unknown disturbances are not considered during the process. Therefore we have no possibility to correct our control during the process. This means these unknown disturbances might affect the process such that we obtain a feasibility problem (cf. Scokaert and Mayne [74]).

3. **Closed-loop min-max optimal feedback control** The last approach is the closed-loop min-max optimal feedback control. In this procedure we add feedback aspects to the open-loop approach from the previous case. That means, we now have the possibility to correct a control using the information we gain about the state during the process. Hence, we can avoid the feasibility problems of the open-loop approach, but this formulation is very complex. It is very unpractical to

compute it online and therefore we have high computational costs (cf. Lee & Yu [55]).

3.5.3 Methods

There has been a lot of research about robust model predictive control. Mayne *et al.* [63] give a good overview about the beginnings of the research about model predictive control until the year 2000. They consider works about constrained linear and nonlinear dynamic systems and about model predictive control of problems that are difficult to solve, e.g. control of unconstrained nonlinear systems and time varying systems. Good and general introductions to feedback, closed-loop and control policies can be found, e.g., Mayne [62], Kothare *et al.* [49],Lee and Yu [55] or Scokaert and Mayne [74].
Bemporad *et al.*[7] and Magni *et al.* [59] describe methods in their work that combine dynamic and parametric programming approaches for solving discrete min-max optimal control problems under the assumption that the perturbations take values in a polytope. Another approach for solving a predictive optimal control problem can be found in Kerrigan and Maciejowski [42] or in Scokaert and Mayne [74], where single finite dimensional optimization techniques are used. Lee and Yu [55] use dynamic programming by discretizing the state space to solve the predictive optimal control problem in a state feedback form.

In Magni *et al.* [60] a model predictive control algorithm for the solution of a state-feedback robust control problem for discrete-time nonlinear systems is presented. As the decision variable is infinite dimensional these results are not practically usable. For this the constructed control laws are suboptimal and, e.g., in Michalska and Mayne [65] or Chisci *et al.* [23] the authors show a suboptimal routine to avoid the infeasibilty and instability by tightening the constraints.

Limòn *et al.* [57] discuss the stability of constrained nonlinear discrete time systems with bounded additive uncertainty. They use a sequence of nested constraint sets, which yields to a input-to-state stability of nominal model predictive control if the disturbance is sufficiently small. This idea is extended to a procedure which does not need the value function to be continuous and does not require the terminal cost to be a control-Lyapunov function in Grimm *et al.* [38].

A good overview of the theory and computation of minimal and maximal robust invariant sets can be found in Kolmanovsky and Gilbert [45].

Other approaches can be found e.g. in Diehl *et al.* [33]. The authors simplify the robust nonlinear model predictive control problem using a linearization. They also show efficient numerical procedures to determine an approximately optimal control sequence. Nagy and Braatz [66] also use the linearization approach.

Another approach is proposed by Goulart *et al.* [37]. They use a control that is an affine function of the current and past states. The decision variables are the associated parameters.

4 Model Predictive Control and Dynamic Programming

Model Predictive Control (MPC), which is also known as Receding Horizon Control (RHC), is a method of finding a finite control sequence by solving online, at any time moment, at which we can measure the data, a finite horizon optimal control problem. In conventional control methods the control law is precomputed offline. However, MPC also implicitly implements a control law, that can in principle be computed offline. The basic concept of MPC is that at each time step we compute the control by solving an open-loop optimization problem for the prediction horizon. Then we apply only the first value of the control sequence we computed and at the next time step, we measure the system state and recompute. This basic concept is also described in Figure 4.1. In Figure

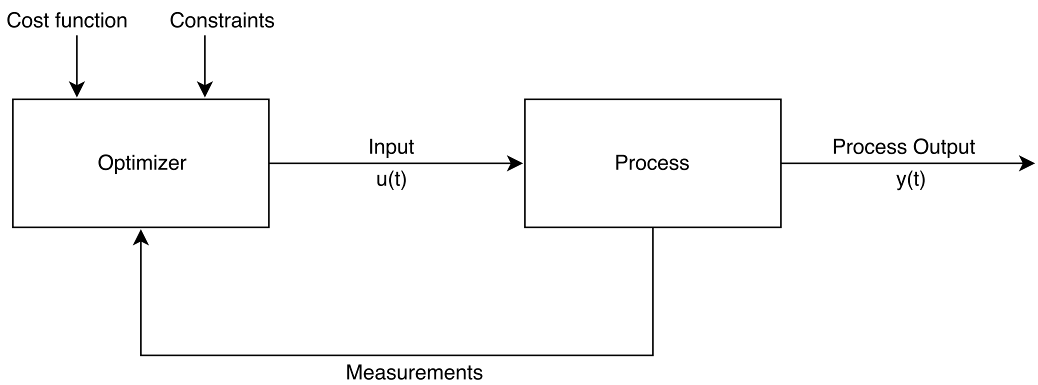


FIGURE 4.1: Feedback Structure of Model Predictive Control

4.2 the past and future control inputs and the measured and predicted future outputs are shown. With Dynamic Programming we also obtain an optimal feedback control

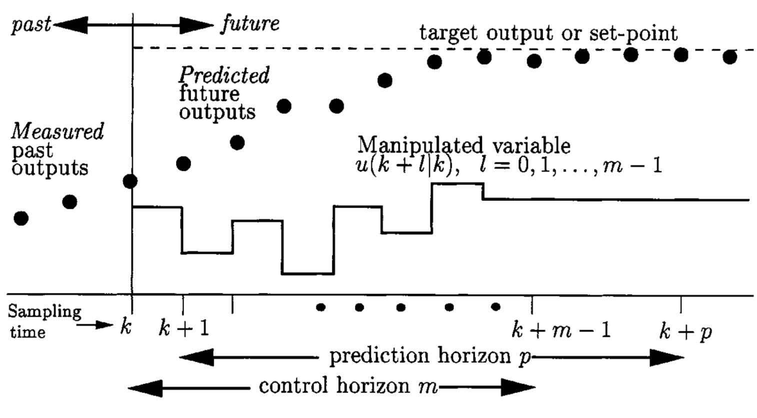


FIGURE 4.2: Basic Presentation of Model Predictive Control (cf. Kohare *et al.* [44])

law, that allows to compute an explicit feedback law offline. The complex problem is divided into subproblems, each of which is only solved once and the results of each subproblem is stored systematically. In this chapter we will first give a short introduction into Model Predictive Control and afterwards we will have a closer look at the Dynamic Programming (DP) approach, formulate Bellman's Principle of Optimality and we will compare the concepts of MPC and DP. Then we will consider both methods in the case that uncertainties are present. We will end the chapter with a short part about Feedback MPC and tubes. This chapter is based on the works of Rawlings & Mayne [69], Bellman & Kalaba [6], Langson *et al.* [53] and Kurzhanski & Varaiya [51].

4.1 Model Predictive Control

This section is based on the work of Rawlings & Mayne [69]. We will consider an infinite horizon optimal control problem. We assume to have the following differential equation

$$\dot{x} = f(x, u)$$

and the cost function

$$V_\infty(x, u(\cdot)) = \int_0^\infty l(x(t), u(t))dt$$

with $x(t)$ and $u(t)$ satisfying $\dot{x} = f(x, u)$, the so-called *stage cost*. Therefore the infinite horizon optimal control problem can be written as

$$\begin{aligned} \min_{u(\cdot)} \;\; & V_\infty(x, u(\cdot)) \\ \text{s.t.} \;\; & \dot{x} = f(x, u) \\ & x(0) = x_0 \\ & u(t) \in U \subseteq \mathbb{R}^{n_u} \\ & x(t) \in X \subseteq \mathbb{R}^{n_x} \end{aligned} \tag{4.1}$$

for all $t \in (0, \infty)$. In the case that $l(\cdot)$ is positive definite the aim of the control is, that the state of the system is directed to the origin. The solution of this problem is denoted by

$$u^0_\infty(\cdot; x)$$

and the optimal value function by

$$V^0_\infty(x).$$

Under this optimal control we can write the closed loop system as

$$\dot{x}(t) = f(x(t), u^0_\infty(t; x)).$$

If we do not have any state constraints and the functions $f(\cdot)$ and $l(\cdot)$ satisfy certain differentiability and growth assumptions then there exists a solution to problem (4.1) for all states x and $V^0_\infty(\cdot)$ is differentiable and satisfies

$$\dot{V}^0_\infty(x) = -l(x, u^0_\infty(0; x)).$$

This means that the origin is an asymptotically stable solution for the closed-loop system. But as we already stated we usually have uncertainties in our systems and hence, we should rather use feedback control than open-loop control, which, however, leads to certain problems. If we only compute online the value of $u^0_\infty(\cdot; x)$ for the current state and not for all x, we still have the problem of having to optimize over a time function $u(\cdot)$ which is infinite dimensional, the time interval $[0, \infty)$ which is semi-infinite and the cost function $V(x, u(\cdot))$ which is usually not convex. Thus we are facing significant optimization difficulties. We can not even guarantee the existence of an optimal control. Therefore we need to approximate our problem (4.1) with a problem which is easier to compute by restricting the system and the control parameterization. For this purpose we will now consider constrained nonlinear time-invariant systems. We can describe the nonlinear system by the following difference equation

$$x_{k+1} = f(x, u) \tag{4.2}$$

with x being the current state, u the current control and x_{k+1} the successor state. The equation (4.2) is the discrete time analog of the continuous time differential equation $\dot{x} = f(x, u)$ and we assume that the function $f(\cdot)$ is continuous and satisfies $f(x(0), u(09) = 0$. For all solutions $x(\cdot)$ of (4.2) with $x(0) = x_0$ and the input control $u(\cdot)$ it holds that

$$x(k+1) = f(x(k), u(k)), \; k = 0, 1, \ldots$$

and

$$x(0) = x_0.$$

Let us introduce the following notations:

- $\phi(k; x, u)$ denotes the solution of (4.2) at time k with the initial state x at time 0 and the control sequence u.
- $\phi(k; (x, i), u)$ denotes the solution of (4.2) at time k if the initial state at time i is x
- (x, i) denotes an event, i.e. the state at time i is x
- $\phi(j - i; x, u)$ denotes the solution of (4.2) at time $j \geq i$ with the initial state x at time i

The function $(x, u) \mapsto \phi(k; x, u)$ is continuous for any k which we will show in the following proposition.

Proposition 4.1. *(cf. Rawlings & Mayne [69])*
Assume that the function $f(\cdot)$ is continuous, then for any integer $k \in \mathbb{Z}$, the function $(x, u) \mapsto \phi(k; x, u)$ is continuous.

Proof. (cf. Rawlings & Mayne [69])
We will prove the proposition by induction. As we know that $\phi(1; x, u(0)) = f(x, u(0))$, the function $(x, u(0)) \mapsto \phi(1; x, u(0))$ is continuous. We assume that the function

$(x, u_{j-1}) \mapsto \phi(j; x, u_{j-1})$ is continuous. With u_j we denote the finite control sequence $\{u(0), u(1), \ldots, u(j)\}$ for any $j \in \mathbb{N}_0$. We consider the function $(x, u_j) \mapsto \phi(j+1; x, u_j)$. Given that

$$\phi(j+1; x, u_j) = f(\phi(j; x, u_{j-1}), u(j))$$

with $f(\cdot)$ and $\phi(j; \cdot)$ being continuous and since $\phi(j+1; \cdot)$ is the composition of the two continuous functions $f(\cdot)$ and $\phi(j; \cdot)$, we can deduce that $\phi(j+1; \cdot)$ is continuous. By induction $\phi(k; \cdot)$ is continuous for any positive integer k. □

In the optimal control problems we usually have constraints which can have the form

$$u(k) \in U, \; x(k) \in X \text{ for all } k \in \mathbb{N}_0 \tag{4.3}$$

We can avoid constraints that involve the control at several times by introducing additional states. In order to do so we can write the common rate constraint $|u(k)-u(k-1)| \leq x$ as $|u(k) - z(k)| \leq c$ with z being an extra state variable which satisfies the difference equation $z^+ = u$ such that $z(k) = u(k-1)$. The constraint $|u-z| \leq c$ is called *mixed constraint* as it includes states and controls. A more general constraint can be formulated as

$$y(k) \in \mathbb{Y}, \text{ for all } k \in \mathbb{N}_0 \tag{4.4}$$

with y satisfying

$$y = h(x, u).$$

As the constraint (4.4) is more general than the constraint (4.3), we can write (4.3) also as $y(k) \in \mathbb{Y}$ with a corresponding choice of the output function $h(\cdot)$ and the output constraint set $\mathbb{Y}$.
The cost of an optimal control problem is usually defined over a finite horizon N to guarantee that the optimal control problem can be solved rapidly enough to permit effective control. We assume x to be the current state and i the current time, then the optimal control problem may be posed as minimizing a cost defined over the interval from time i to time $N+i$. We can write the optimal control problem as follows:

$$\begin{aligned} \min_{x,u} \quad & \sum_{k=i}^{i+N-1} l(x(k), u(k)) + V_f(x(N+1)) \\ \text{s.t.} \quad & x_{k+1} = f(x, u) \\ & x(i) = x \\ & u(k) \in U \\ & x(k) \in X \end{aligned} \tag{4.5}$$

for all $k \in \mathbb{N}_0$ and with V_f the terminal cost and $x := \{x(i), x(i+1), \ldots, x(i+N)\}, u := \{u(i), u(i+1), \ldots, u(i+N-1)\}$. We assume that $l(\cdot)$ is continuous with $l(0,0) = 0$. If we solve the optimal control problem (4.5) we get the following control and state sequences

$$\begin{aligned} u^0(x,i) &= \{u^0(i; (x,i)), u^0(i+1; (x,i)), \ldots, u^0(i+N-1; (x,i))\} \\ x^0(x,i) &= \{x^0(i; (x,i)), x^0(i+1; (x,i)), \ldots, x^0(i+N; (x,i))\} \end{aligned}$$

with $x^0(i;(x,i)) = x$. In the MPC we use the first control action $u^0(i;(x,i))$ in the optimal control sequence $u^0(x,i)$ as control for the process, that means $u(i) = u^0(i;(x,i))$. As we know that the system $x_{k+1} = f(x,u)$, the stage cost $l(\cdot)$ and the terminal cost $V_f(\cdot)$ are all time invariant, we have

$$u^0(x,i) = u^0(x,0)$$
$$x^0(x,i) = x^0(x,0)$$

And therefore from now on we will consider the optimal control problem (4.1) with $i = 0$. For simplicity reasons we will replace $u^0(x,0)$ by $u^0(x)$ and $x^0(x,0)$ by $x^0(x)$. Hence, we can write the optimal control problem as

$$\begin{aligned} \min_{x,u} \quad & \sum_{k=0}^{N-1} l(x(k),u(k)) + V_f(x(N)) \\ \text{s.t.} \quad & x_{k+1} = f(x,u) \\ & x(0) = x \\ & u(k) \in U \\ & x(k) \in X \end{aligned} \tag{4.6}$$

with $u = \{u(0), u(1), \ldots, u(N-1)\}$ and $x = \{x(0), x(1), \ldots, x(N)\}$. In (4.6) we also include the constraint that the state sequence x is a priori a solution of $x_{k+1} = f(x,u)$ to ensure that we can write the problem in the equivalent form of minimizing, with respect to the control sequence u, a cost that is only a function of the initial state x and the control sequence u. Therefore we can rewrite the cost function as

$$V_N(x,u) := \sum_{k=0}^{N-1} l(x(k),u(k)) + V_f(x(N)) \tag{4.7}$$

with $x(k) := \phi(k;x,u)$ for all $k \in \{0,\ldots,N\}$. We also add to the constraints from before an additional terminal constraint on the state

$$x(N) \in X_f$$

with $X_f \subseteq X$ and an additional constraint on the control sequence

$$u \in U_N(x) \tag{4.8}$$

with $U_N(x)$ being the set of control sequences $u := \{u(0), u(1), \ldots, U(N-1)\}$ satisfying the state and control constraints:

$$U_N(x) := \{u | (x,u) \in \mathcal{Z}_N\} \tag{4.9}$$

with the set $\mathcal{Z}_N \subset \mathbb{R}^N \times \mathbb{R}^{N_m}$ defined as

$$\mathcal{Z}_N := \{(x,u) | u(k) \in U,\ \phi(k;x,u) \in X,\ \text{for all } k \in \{0,\ldots,N-1\}, \phi(N;x,u) \in X_f\}.$$

Hence, we can write the optimal control problem as

$$V_N^0(x) := \min_u \{V_N(x,u) | u \in U_N(x)\}. \tag{4.10}$$

The problem (4.10) is a parametric optimization problem with the decision variable u and the cost and the constraint set depending on the parameter x. With $\mathcal{Z}_N$ we denote the set of admissible (x, u), that means the set of (x, u) for which $x \in X$ and for which the constraints of (4.10) are satisfied. We assume X_N to be the set of states in X for which problem (4.10) has a solution, i.e.

$$X_N = \{x \in X | U_N(x) \neq \emptyset\}.$$

This can be rewritten as

$$X_N = \{x \in \mathbb{R}^n | \text{there exists an } u \in \mathbb{R}^{Nm} \text{ such that } (x, u) \in \mathcal{Z}_N\}$$

which is the orthogonal projection of $\mathcal{Z}_N \subset \mathbb{R}^n \times \mathbb{R}^{Nm}$ onto $\mathbb{R}^n$.

With Weierstrass's theorem we can ensure that an optimization problem has a solution if the cost function is continuous (in the decision variable) and if the constraint set is compact. That is what we requested in Proposition 4.1. We will now formulate further assumptions that have to be satisfied in the sequel

Assumption 4.2. (Continuity of system and cost)
The functions $f : X \times U \to \mathbb{R}^n$, $l : X \times U \to \mathbb{R}_{\geq 0}$ and $V_f : X_f \to \mathbb{R}_{\geq 0}$ are continuous and $f(0, 0) = 0$, $l(0, 0) = 0$ and $V_f(0) = 0$.

Assumption 4.3. (Properties of constraint sets)
The sets X and X_f are closed, $X_f \subseteq X$ and U are compact and each set contains the origin.

Proposition 4.4. *(Existence of solution to optimal control problem) (cf. Rawlings & Mayne [69])*
We assume that Assumption 4.2 and Assumption 4.3 hold. Then

a. *The function $V_N(\cdot)$ is continuous in $\mathcal{Z}_N$.*

b. *For each $x \in X_N$, the control constraint set $U_N(x)$ is compact.*

c. *For each $x \in X_N$ a solution to Problem (4.10) exists.*

Proof. The proof can be found in Rawlings & Mayne [69] p. 98. □

The solution of the optimal control problem (4.10) can then be written as

$$u^0(x) = \arg\min_u \{V_N(x, u) | u \in U_N(x)\}.$$

In the case that $u^0(x) = \{u^0(0; x), u^0(1; x), \ldots, u^0(N-1; x)\}$ is unique for each $x \in X_N$, then $u^0 : \mathbb{R}^n \to \mathbb{R}^{Nm}$ is a function. If it is not unique, it is a set-valued function, which means that the values of the function for each x in its domain is a set. Always the first element of $u^0(0; x)$ of the optimal control sequence $u^0(x)$ is applied to the process as control. Afterwards we repeat the whole procedure at the next control point for the successor state. With MPC it is also possible to compute $u(x)$ and therefore also $u(0; x)$ for every x for which the OCP (4.10) is feasible. Then we get the implicit MPC control law $\kappa_N(\cdot)$ which is defined as follows

$$\kappa_N(x) := u^0(0; x), \ x \in X_N.$$

A great advantage of MPC is that we do not have to determine the control law $\kappa_N(\cdot)$ as it is often impossible to determine the control law in the case of constraints or nonlinearities.
If the solution of the OCP (4.10) is not unique at a given state x, then the control law $\kappa_N(\cdot) = u^0(0;\cdot)$ is set valued and the model predictive controller selects one element from the set $\kappa_N(x)$.

4.2 Dynamic Programming

Dynamic Programming (DP) is a method of dividing a complicated problem into smaller and simpler sub-problems and saving of interim results. In the 1950s Richard Bellman introduced this kind of methods with the formulation of Bellman's *Principle of Optimality.* This section is based on the work of Bellman & Kalaba [6], Kurzshanski & Varaiya [51] and Rawlings & Mayne [69].

We consider the discrete time system

$$x_{k+1} = f(x, u) \tag{4.11}$$

as before with $f(\cdot)$ continuous. The constraints of the system can be described by

$$(x, u) \in \mathcal{Z}$$

with $\mathcal{Z}$ being the closed subset of $\mathbb{R}^n \times \mathbb{R}^m$ and by $P_u(\mathcal{Z})$, which is compact, denoting the projection operator $(x, u) \mapsto u$. Often it holds that $\mathcal{Z} = X \times U$ which means that $x \in X$, $u \in U$ and $P_u(\mathcal{Z}) = U$ such that U is compact. The constraint on the terminal state $x(N)$ can be described by

$$x(N) \in X_f.$$

We consider the cost function at the current time i which can be written as

$$V^0(x, i).$$

That means that it is the optimal cost at state x and time i and by $X(i)$ we notate the domain of $V^0(\cdot, i)$. For each time i, $x(i) = \phi(i, (x, 0), u)$ is the solution at time i of (4.11) if the initial state is x at time 0 and the control sequence is u. The cost which belongs to an initial state x at time 0 and a control sequence $u := \{u(0), u(1), \ldots, u(N-1)\}$ can be formulated as

$$V(x, 0, u) = V_f(x(N)) + \sum_{i=1}^{N-1} l(x(i), u(i)). \tag{4.12}$$

The corresponding OCP is then defined by

$$\begin{aligned} V^0(x, 0) = \min_u \ & V(x, 0, u) \\ \text{s.t.} \ & (x(i), u(i)) \in \mathcal{Z}, \ i = 0, 1, \ldots, N-1 \\ & x(N) \in X_f. \end{aligned} \tag{4.13}$$

We can rewrite (4.13) as

$$V^0(x,0) = \min_u \{V(x,0,u) | u \in U(x,0)\} \tag{4.14}$$

with

$$\begin{aligned} u &:= \{u(0), u(1), \dots, u(N-1)\}, \\ U(x,0) &:= \left\{u \in \mathbb{R}^{Nm} | (x(i), u(i)) \in \mathcal{Z},\ i = 0, 1, \dots, N-1; x(N) \in X_f\right\}, \\ x(i) &:= \phi(i; (x,0), u). \end{aligned}$$

By $U(x,0)$ we notate the set of admissible control sequences in the case that the initial state is x at time 0. From $f(\cdot)$ being continuous we can deduce that for all $i \in \{0, 1, \dots, N-1\}$ and all $x \in \mathbb{R}^n$ it holds $u \mapsto \phi(i; (x,0), u)$ is continuous, $u \mapsto V(x,0,u)$ is continuous and $U(x,0)$ is compact. That means that the minimum in (4.14) exists at all $x \in \{x \in \mathbb{R}^n | U(x,0) \neq \emptyset\}$. With DP problem (4.13) for a given state x is contained in a whole family of problems. For each (x,i) we can define these problems by

$$V^0(x,i) = \min_{u^i} \left\{V(x,i,u^i) | u^i \in U(x,i)\right\} \tag{4.15}$$

with

$$u^i := \{u(i), u(i+1), \dots, u(N-1)\},$$

$$V(x,i,u^i) := V_f(x(N)) + \sum_{j=i}^{N-1} l(x(j), u(j)), \tag{4.16}$$

$$U(x,i) := \left\{u^i \in \mathbb{R}^{(N-i)m} | (x(j), u(j)) \in \mathcal{Z},\ j = i, i+1, \dots, N-1, x(N) \in X_f\right\}. \tag{4.17}$$

In (4.16) and (4.17) by $x(j) = \phi(j; (x,i), u^i)$ we denote the solution at time j of the system (4.11) in the case that the initial state is x at time i and the control sequence is u^i. For each i, we denote by $X(i)$ the domain of $V^0(\cdot, i)$ and $U(\cdot, i)$, such that

$$X(i) = \{x \in \mathbb{R}^n | U(x,i) \neq \emptyset\}.$$

For all (x,i) it holds

$$\begin{aligned} V^0(x,i) &= \min_{u^i} \left\{V(x,i,u^i) | u^i \in U(x,i)\right\} \\ &= \min_u \left\{ l(x,u) + \min_{u^{i+1}} V(f(x,u), i+1, u^{i+1}) | \{u, u^{i+1}\} \in U(x,i) \right\} \end{aligned} \tag{4.18}$$

with $u^i = \{u, u(i+1), \dots, u(N-1)\} = \left\{u, u^{i+1}\right\}$. Because of $f(x,u) = x(i+1)$ it also holds that $\left\{u, u^{i+1} \in U(x,i)\right\}$ if and only if $(x,u) \in \mathcal{Z},\ f(x,u) \in X(i+1)$ and $u^{i+1} \in U(f(x,u), i+1)$. With this (4.18) can be rewritten as

$$V^0(x,i) = \min_u \left\{l(x,u) + V^0(f(x,u), i+1) | (x,u) \in \mathcal{Z}, f(x,u) \in X(i+1)\right\} \tag{4.19}$$

for all $x \in X(i)$ with

$$X(i) = \{x \in \mathbb{R}^n | \text{ there exists an } u \text{ such that } (x,u) \in \mathcal{Z} \text{ and } f(x,u) \in X(i+1)\}\,. \tag{4.20}$$

The DP recursion for constrained discrete time optimal control problems is then determined by equations (4.19), (4.20) and the boundary conditions

$$V^0(x,N) = V_f(x) \text{ for all } x \in X(N),\ X(N) = X_f.$$

In the case that there are no state constraints, that means, that $\mathcal{Z} \in \mathbb{R}^n \times U$ with $U \subset \mathbb{R}^m$ is compact, then for all $i \in \{0,1,\dots,N\}$ it holds $X(i) = \mathbb{R}^n$ and the DP equations revert to the familar DP recursion

$$V^0(x,i) = \min_u \left\{ l(x,u) + V^0(f(x,u), i+1) \right\} \text{ for all } x \in \mathbb{R}^n$$

with the boundary condition

$$V^0(x,N) = V_f \text{ for all } x \in \mathbb{R}^n.$$

We can now formulate the Principle of Optimality, which was first done by Richard Bellman in 1957. The original text was stated as follows

Principle of Optimality (cf. Bellman [6])
"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decision must constitute an optimal policy with regard to the state resulting from the first decision."

In the following Lemma, we will formulate and prove the Principle of Optimality in a more mathematical way regarding our problem. Afterwards we will formulate the conditions for the optimal value function and the optimal control law.

Lemma 4.5. *(Principle of Optimality) (cf. Rawlings & Mayne [69])*
Let $x \in X_N$ be arbitrary, let $u := \{u(0), u(1), \dots, u(N-1)\} \in U(x,0)$ denote the solution of (4.14) and let $\{x, x(1), x(2), \dots, x(N)\}$ denote the corresponding optimal state trajectory such that for each i, $x(i) = \phi(i;(x,0),u)$. Then, for any $i \in \{0,1,\dots,N-1\}$, the control sequence $u^i := \{u(i), u(i+1), \dots, u(N-1)\}$ is optimal for (4.15) (any portion of an optimal trajectory is optimal).

Proof. (cf. Rawlings & Mayne [69])
Since u is in $U(x,0)$, the control sequence u^i is in $U(x(i),i)$. We assume that $u^i = \{u(i), u(i+1), \dots, u(N-1)\}$ is not optimal for (4.15), then there exists a control sequence $u' = \{u'(i), u'(i+1), \dots, u'(N-1)\}$ which lies in $U(x(i),i)$ such that $V(x(i),i,u') < V(x(i),u)$. We consider the control sequence $\tilde{u} := \{u(0), u(1), \dots, u(i-1), u'(i), u'(i+1), \dots, u'(N-1)\}$. From that we can deduce that $\tilde{u} \in U(x,0)$ and $V(x,0,\tilde{u}) < V(x,0,u) = V^0(x,0)$, which is a contradiction. Therefore, $u(x(i),i)$ is optimal for (4.15). □

Theorem 4.6. *(Optimal value function & control law from DP) (cf. [69])*
We assume that the function $\psi : \mathbb{R}^n \times \{0, 1, \dots, N\} \to \mathbb{R}$ *for all* $i \in \{1, 2, \dots, N-1\}$ *and all* $x \in X(i)$ *satisfies the DP recursion*

$$\psi(x, i) = \min\{l(x, u) + \psi(f(x, u), i + 1)|(x, u) \in \mathcal{Z}, f(x, u) \in X(i + 1)\}$$
$$X(i) = \{x \in \mathbb{R}^n | \textit{there exists an } u \in \mathbb{R}^m \textit{ such that } (x, u) \in \mathcal{Z},\ f(x, u) \in X(i + 1)\}$$

with the boundary conditions

$$\psi(x, N) = V_f(x) \textit{ for all } x \in X_f,\ X(N) = X_f.$$

Then $\psi(x, i) = V^0(x, i)$ *for all* $(x, i) \in X(i) \times \{0, 1, 2, \dots, N\}$*, that means that the DP recursion yields the optimal value function and the optimal control law.*

Proof. A proof can be found in Rawlings & Mayne [69]. □

In this thesis, we will consider linear quadratic problems and therefore we will now assume the following linear quadratic problem. The system is defined by

$$x_{k+1} = Ax + Bu.$$

We do not have any constraints and the cost function is defined by (4.12) with

$$l(x, u) = \frac{1}{2}x^T Qx + \frac{1}{2}u^T Ru$$

and $V_f(x) = 0$ for all x. The horizon length is N. We suppose that Q is symmetric and positive semidefinite and R is symmetric and positive definite. For this problem we can formulate the DP recursion as

$$V^0(x, i) = \min_u \left\{l(x, u) + V^0(Ax + Bu, i + 1)\right\} \text{ for all } x \in \mathbb{R}^n$$

with the terminal condition

$$V^0(x, N) = 0 \text{ for all } x \in \mathbb{R}^n.$$

We assume that $V(\cdot, i + 1)$ is quadratic and positive semidefinite and therefore we can write it in the form

$$V^0(x, i + 1) = \frac{1}{2}x^T P(i + 1)x$$

with $P(i + 1)$ symmetric and positive semidefinite. Then we can write

$$V^0(x, i) = \frac{1}{2} \min_u \left\{x^T Qx + u^T Ru + (Ax + Bu)^T P(i + 1)(Ax + Bu)\right\}. \tag{4.21}$$

The right hand side of problem (4.21) is a positive definite function of u for all x, such that it is a unique minimizer which is given by

$$\kappa(x, i) = K(i)x$$

with

$$K(i) := -(B^T P(i+1)B * R)^{-1} B^T P(i+1).$$

By setting $u = K(i)x$ in (4.21) we obtain

$$V^0(x,i) = \frac{1}{2} x^T P(i) x$$

with

$$P(i) = Q + A^T P(i+1)A - A^T P(i+1)B(B^T P(i+1)B + R)^{-1} BP(i+1)A.$$

4.3 Dynamic Programming Solution

In contrast to MPC the Dynamic Programming gives us the value function $V_N(\cdot)$ and the implicit MPC control law $\kappa_N(\cdot)$. We consider the OCP (4.10) with the cost function $V_N(\cdot)$(4.7) and the constraints on the control (4.8). The Dynamic Programming approach provides us an optimal policy $\mu^0 = \{\mu_0^0(\cdot), \mu_1^0(\cdot), \dots, \mu_{N-1}^0(\cdot)\}$, that means we obtain a sequence of control laws $\mu_i : X_i \to U,\ i = 0, 1, \dots, N-1$. We will later have a closer look at the domain X_i. Using MPC we have the time-invariant system that satisfies

$$x_{k+1} = f(x, \kappa_N(x)),\ i = 0, 1, \dots, N-1,$$

with $\kappa_N(\cdot) = \mu_0^0(\cdot)$. For DP we consider the time varying controlled system that satisfies

$$x_{k+1} = f(x, \mu_i^0(x)),\ i = 0, 1, \dots, N-1.$$

We define for all $j \in \{0, \dots, N-1\}$

$$\begin{aligned} V_j(x,u) &:= \sum_{k=0}^{j-1} l(x(k), u(k)) + V_f(x(j)), \\ U_j(x) &:= \{u | (x,u) \in \mathcal{Z}_j\} \\ V_j^0(x) &= \min_u \{V_j(x,u) | u \in U_j(x)\} \end{aligned} \tag{4.22}$$

just as in (4.7) and (4.9) with N replaced by j. From before we know that we can solve the following problem

$$V_N^0 = \min_u V_N(x,u) | u \in U_N(x)\} \tag{4.23}$$

for all $x \in X_N$, which is the domain of V_N^0 with DP, but we can also solve problem (4.22) for all $x \in X_j$, which is the domain of V_j^0 for all $j \in \{0, \dots, N-1\}$. Therefore we can write the DP equations as follows

$$V_j^0(x) = \min_{u \in U} \left\{ l(x,u) + V_{j-1}^0(f(x,u)) | f(x,u) \in X_{j-1} \right\},\ \text{for all } x \in X_j \tag{4.24}$$

$$\kappa_j(x) = \arg\min_{u \in U} \left\{ l(x,u) + V_{j-1}^0(f(x,u)) | f(x,u) \in X_{j-1} \right\},\ \text{for all } x \in X_j \tag{4.25}$$

$$X_j = \{x \in X | \text{there exists an } u \in U, \text{ such that } f(x,u) \in X_{j-1}\} \tag{4.26}$$

for $j = 1, 2, \ldots, N$ (j is time to go) with the terminal condition

$$V_0^0(x) = V_f(x) \text{ for all } x \in X_0,\ X_0 = X_f$$

With $V_j^0(x)$ we denote the optimal cost for problem (4.22) for each j if the current state is x, the current time is 0 (or i), the terminal time is j (or $i+j$) and X_j is the domain. By X_j we also describe the set of states in X that can be steered to the terminal set X_f in j steps by an admissible control sequence, which means a control sequence that satisfies the control, state and terminal constraints and therefore it lies in the set $U_j(x)$. Hence, for each j it holds that

$$X_j = \{x \in X | U_j(x) \neq \emptyset\}.$$

Definition 4.7. (Feasible preimage of the state) (cf. [69])
We assume $\mathcal{Z} := X \times U$. The set-valued function $f_{\mathcal{Z}}^{-1} : X \to \mathcal{Z}$ is defined by

$$f_{\mathcal{Z}}^{-1}(x) := f^{-1}(x) \cap \mathcal{Z}$$

with $f^{-1}(x) := \{z \in \mathbb{R}^n \times \mathbb{R}^m | f(z) = x\}$.

For all $j \geq 0$ we define the set $Z_j \subseteq \mathbb{R}^n \times \mathbb{R}^m$ as

$$Z_j := f_{\mathcal{Z}}^{-1}(X_{j-1}) = \{(x, u) | f(x, u) \in X_{j-1}\} \cap \mathcal{Z}.$$

We can then describe the set X_j as

$$X_j = \{x \in \mathbb{R}^n | \text{ there exists an } u \in \mathbb{R}^m \text{ such that } (x, u) \in Z_j\}.$$

With DP we do not only get an optimal control sequence for a given state, but we can also obtain an optimal feedback policy μ^0 or a sequence of control laws with

$$\mu^0 := \{\mu_0(\cdot), \mu_1(\cdot), \ldots, \mu_{N-1}(\cdot)\} = \{\kappa_N(\cdot), \kappa_{N-1}(\cdot), \ldots, \kappa_1(\cdot)\}.$$

At the event (x, i), where the state is x and the time is i, the time to go to the terminal is $N - i$ and the optimal control is

$$\mu_i^0(x) = \kappa_{N-1}(x)$$

which means, that $\mu_i(\cdot)$ is the control law at time i. We consider the initial event $(x, 0)$, which means the state is x at time 0. If the terminal time is N, the optimal control for $(x, 0)$ is $\kappa_N(x)$ and the successor state, i.e. the state at time 1, is

$$x_{k+1} = f(x, \kappa_N(x)).$$

Then at the event $(x_{k+1}, 1)$, the time to go to the terminal is $N - 1$ and the optimal control is $\kappa_{N-1}(x_{k+1}) = \kappa_{N-1}(f(x, \kappa_N(x)))$. If we have a given initial event $(x, 0)$ the optimal policy guarantees to get the optimal state and control trajectories $x^0(x)$ and $u^0(x)$ for which the following difference equations hold

$$\begin{aligned} x(0) &= x & u(0) &= \kappa_N(x) \\ x(i+1) &= f(x(i), u(i)) & u(i) &= \kappa_{N-1}(x(i)) \end{aligned}$$

for $i = 0, 1, \ldots, N-1$. In this case the state and control trajectories are the same as the ones, that we obtain, when we use MPC and solve problem (4.23) directly for the initial event $(x, 0)$ using a mathematical programming algorithm. But the difference is now, that by using DP we can generate a solution for any event (x, i), such that $i \in \{0, \ldots, N-1\}$ and $x \in X_i$. We will now define positive invariant and control invariant sets and afterwards we will formulate properties of the solution to each partial problem (4.22).

Definition 4.8. (Positive and control invariant sets) (cf. [69])

- A set $X \subseteq \mathbb{R}^n$ is positive invariant for $x_{k+1} = f(x)$ if $x \in X$ implies $f(x) \in X$.
- A set $X \subseteq \mathbb{R}^n$ is control invariant for $x_{k+1} = f(x, u),\ u \in U$, if for all $x \in X$ there exists a $u \in U$ such that $f(x, u) \in X$.

Proposition 4.9. *(Existence of solutions to DP recursion) (cf. [69])*
We assume that the Assumptions 4.2 and 4.3 (i.e. $f(\cdot)$, $l(\cdot)$, $V_f(\cdot)$ are continuous and X and X_f are continuous and U is compact and each of the sets contain the origin) hold. Then

1. *For all $j \geq 0$ the cost function $V_j(\cdot)$ is continuous in Z_j and for each $x \in X_j$ the control constraint set $U_j(x)$ is compact and a solution $u^0(x) \in U_j(x)$ to problem (4.22) exists.*
2. *If $X_0 := X_f$ is control invariant for $x_{k+1} = f(x, u),\ u \in U$, then for each $j \in \mathbb{I}_{\geq 0}$, the set X_j is also control invariant and $X_j \supseteq X_{j-1}$ and $0 \in X_j$. Additionally the set X_N is positive invariant for $x_{k+1} = f(x, \kappa_N(x))$.*
3. *For each $j \geq 0$ the set X_j is closed.*

Proof. A proof can be found in Rawlings & Mayne [69]. □

4.4 Model Predictive Control under Uncertainties

An important advantage of the conventional MPC is, that the solution of an open-loop OCP which we solve online is identical to the one that we obtain, when we use DP for the given initial state. We also said, that feedback control is superior to open-loop control in the case of uncertainties in the problem. That means that the OCP that we solve online has to allow feedback to guarantee that its solution is equal to the solution of DP. For the online OCP with horizon N we should rather use a problem, where the decision variable μ is a sequence of control laws, than problem (4.23) where the decision variable u is a sequence of control actions. From now on we will call MPC where the decision variable is a policy *Feedback MPC*. In this approach the policy

$$\mu^0(x) = \{\mu_0^0(\cdot; x), \mu_1^0(\cdot; x), \ldots, \mu_{N-1}^0(\cdot, x)\}$$

is the solution to the OCP. Each of the control laws is a restriction of those which we determine using DP and hence they depend on the initial state x. We only have to determine the value $u^0(x) = \mu_0(x; x)$ of the control law $\mu_0(\cdot; x)$ at the initial state x the following laws we only have to determine of a limited range. Even though Feedback MPC is superior if we have uncertainties but it is enormously more complex than the OCP which is used in the deterministic MPC. Furthermore, the decision variable μ, which is a sequence of control laws is infinite dimensional. And therefore each law or function requires, in general, an infinite dimensional grid to specify it. Feedback MPC is similarly complex as solving the DP equation. That means that MPC, which replaces DP with a solvable open-loop optimization problem in the deterministic case is not easily solved in the case of uncertainties. In the following we will consider the dynamic programming solution under uncertainties.

4.5 Dynamic Programming Solution with Uncertainties

We consider the system

$$x_{k+1} = f(x, u, w) \tag{4.27}$$

with the bounded disturbance input w which represents our uncertainty. The uncertainty w lies in a set $\mathbb{W}$ which is compact convex and contains the origin. We also have state and control constraints as well as terminal constraints as follows

$$x \in X,\ u \in U \text{ and } x(N) \in X_f.$$

We notate the solution of system (4.27) with control and disturbance sequences $u = \{u(0), \ldots, u(N-1)\}$ and $w = \{w(0), \ldots, w(N-1)\}$ at time k as $x(k; x, u, w)$ if the initial state is x at time 0. Analogously the solution at time k with the feedback policy μ and the disturbance sequence w we denote by $x(k; x, \mu, w)$. The value of the cost function is the maximum which is taken over all possible realizations of the disturbance sequence w due to policy μ with the initial state x. We formulate it as

$$V_N(x, \mu) := \max_w \{J_N(x, \mu, w) | w \in W\} \tag{4.28}$$

with $W = \mathbb{W}^n$ the set of admissible disturbance sequences and $J_N(x, \mu, w)$ the cost due to an individual realization w of the disturbance process. This cost is defined by

$$J_N(x, \mu, w) := \sum_{i=0}^{N-1} l(x(i), u(i), w(i)) + V_f(x(N)) \tag{4.29}$$

with $\mu = \{u(0), \mu_1(\cdot), \ldots, \mu_{N-1}(\cdot)\}$, $x(i) = \phi(i; x, \mu, w)$ and $u(i) = \mu_i(x(i))$. With $M(x)$ we denote the set of feedback policies μ that for a given initial state x satisfy the state and control constraints and the terminal constraints for every admissible disturbance sequence $w \in W$. The first element in μ, namely $u(0)$ is a control action and not a control law as the initial state x is known, afterwards the future states are uncertain, therefore the following $\mu_1(\cdot), \ldots, \mu_{N-1}(\cdot)$ are control laws. Hence, we can define $M(x)$ as follows

$$\begin{aligned} M(x) := \{\mu | u(0) \in U,\ \phi(i; x, \mu, w) \in X,\ \mu_i(\phi(i; x, \mu, w)) \in U \text{ for all } i \in \{0, \ldots, N-1\} \\ \text{and } \phi(N; x, \mu, w) \in X_f \text{ for all } w \in W\}. \end{aligned}$$

We denote the robust optimal control problem as

$$\inf_{\mu}\{V_N(x,\mu)|\mu \in M(x)\} \tag{4.30}$$

and if there is a solution to (4.30) it is the policy $\mu^0(x)$ which satisfies

$$\mu^0(x) = \{u^0(0;x), \mu_1^0(\cdot,x), \dots, \mu_{N-1}(\cdot,x)\}$$

and the value function is

$$V_N^0(x) = V_N(x, \mu^0(x)).$$

As in conventional MPC, the control applied to the system state x is $u^0(0;x)$, which is the first element in $\mu^0(x)$ and therefore the implicit model predictive feedback control law is $\kappa_N(\cdot)$. We define it by

$$\kappa_N(x) := u^0(0;x).$$

As we already mentioned it is often impossible to use DP because of the large computational costs. But it is feasible to use it in certain cases such as low dimensional constrained optimal control problems when the system is linear, the constraints are affine and the cost is affine or quadratic. A much better approach is to use min-max DP. We denote the partial value function by $V_i^0(\cdot)$ and the optimal solution to the optimal control problem

$$\inf_{\mu}\{V_i(x,\mu)|\mu \in M(x)\} \tag{4.31}$$

which is defined in (4.30) by replacing N by i with $\kappa_i(\cdot)$ for each $i \in \{0,1,\dots,N\}$. We can then write the DP recursion as

$$\begin{aligned} V_i^0(x) &= \min_{u\in U}\max_{w\in\mathbb{W}}\{l(x,u,w) + V_{i-1}^0(f(x,u,w))|f(x,u,\mathbb{W} \subseteq X_{i-1}\} \\ \kappa_i(x) &= \arg\min_{u\in\mathbb{W}}\max_{w\in\mathbb{W}}\{l(x,u,w) + V_{i-1}^0(f(x,u))|f(x,u,\mathbb{W} \subseteq X_{i-1}\} \\ X_i &= \{x \in X|\text{there exists an } u \in U \text{ such that } f(x,u,\mathbb{W}) \subseteq X_{i-1}\} \end{aligned}$$

with the boundary conditions

$$V_0^0(x) = V_f(x),\ X_0 = X_f.$$

Here, the subscript i is the time to go and for each i the set X_i is the domain of $V_i^0(\cdot)$ and $(\kappa(i))$. And hence, the set X_i is the set of states x for which a solution of problem (4.31) exists. We can also say, that X_i is the set of states that can be robustly steered by state feedback, which means by a policy $\mu \in M(x)$, to X_f in i steps or less and also satisfying all constraints for all disturbance sequences. Therefore we can write

$$V_i^0(x) = \max_{w\in\mathbb{W}}\{l(x,\kappa_i(x),w) + V_{i-1}^0\left(f(x,\kappa_i(x),w)\right)\}. \tag{4.32}$$

As before our aim is to get sufficient conditions to guarantee that the MPC law κ_N is stabilizing. Therefore we have to adapt Assumption 4.2 and Assumption 4.3 to the robust control problem. We will first generalize the Definition 4.8.

Definition 4.10. (Robust control invariance) (cf. [69])
A set $X \subseteq \mathbb{R}^n$ is robust control invariant for the system $x_{k+1} = f(x, u, w),\ w \in \mathbb{W}$, if for every $x \in X$ there exists an $u \in U$ such hat $f(x, u, \mathbb{W}) \subseteq X$.

Definition 4.11. (Robust positive invariance) (cf. [69])
A set X is robust positive invariant for the system $x_{k+1} = f(x, w),\ w \in \mathbb{W}$ if, for every $x \in X,\ f(x, \mathbb{W}) \subseteq X$.

With these definitions we can now generalize the Assumptions 4.2 and 4.3.

Assumption 4.12. (Basic stability assumption in the robust case) (cf. [69])

1. For all $x \in X_f$

$$\min_{u \in U} \max_{w \in \mathbb{W}} \Delta V_f(x, u, w) + l(x, u, w) \leq 0$$

   with $\Delta V_f(x, u, w) = V_f(f(x, u, w)) - V_f(x)$.

2. $X_f \subseteq X$.

This assumption implicitly requires that for each $x \in X_f$ there exists an $u \in U$ such that $f(x, u, \mathbb{W}) \subseteq X_f$. From Assumption 4.12 we can now deduce the following assumption

Assumption 4.13. (Implied stability assumption for the robust case (cf. [69])
The set X_f is robust control invariant for $x_{k+1} = f(x, u, w)$ with $w \in \mathbb{W}$.

From these two assumptions we can deduce the existence of a terminal control law $\kappa_f : X_f \to U$ which has the following properties:

- $\Delta V_f(x, \kappa_f(x), w) + l(x, \kappa_f(x), w) \leq 0$ for all $x \in X_f$ and all $w \in \mathbb{W}$,
- X_f is robust positive invariant for the system $x_{k+1} = f(x, \kappa_f(x), w)$,
- $X_f \subseteq X$ and
- $\kappa_f(X_f) \subseteq U$.

Let us outline some preliminary results in the following theorem that are similar to the ones that we formulated in Section 4.3.

Theorem 4.14. *(Recursive feasibility of control policies) (cf. [69])*
We assume that the Assumptions 4.12 and 4.13 hold. Then

1. $X_N \supseteq X_{N-1} \supseteq \cdots \supseteq X_1 \supseteq X_0 = X_f$.
2. X_i *is robust control invariant for the system* $x_{k+1} = f(x, u, w)$ *for all* $i \in \{0, 1, \ldots, N\}$.
3. X_i *is robust control invariant for the system* $x_{k+1} = f(x, \kappa_i(x), w)$ *for all* $i \in \{0, 1, \ldots, N\}$.
4. $V_i^0(x) \leq V_{i-1}^0(x)$ *for all* $x \in X_{i-1}$ *and for all* $i \in \{0, 1, \ldots, N\}$.
5. $V_N^0(x) \leq V_f(x)$ *for all* $x \in X_f$.

6. $\Delta V_N^0(x, \kappa_N(x), w) + l(x, \kappa_N(x), w) \leq V_N^0(f(x, \kappa_N(x), w) + V_{N-1}^0(f(x, \kappa_N(x), w) \leq 0$ *for all* $(x, w) \in X_N \times \mathbb{W}$.

7. *For any* $x \in X_N$, $\{\kappa_N(x), \kappa_{N-1}(\cdot), \ldots, \kappa_1(\cdot), \kappa - f(\cdot)\}$ *is a feasible policy for problem*

$$\inf_{\mu}\{V_N(x, \mu)|\mu \in M(x)\}$$

and for any $x \in X_{N-1}$ *the set* $\{\kappa_{N-1}(x), \kappa_{N-2}(\cdot), \ldots, \kappa_1(\cdot), \kappa_f(\cdot)\}$ *is a feasible policy for problem (4.30).*

Proof. A proof can be found in Rawlings & Mayne [69]. □

4.6 Feedback MPC and Tubes

Let us again consider the following two systems. First we consider the deterministic nominal system

$$x_{k+1} = f(x, u)$$

with the control variable $u = \{u(0), u(1), \ldots, u(N-1)\}$. The control u is not only one variable but a sequence of control actions. In the case that x_0 is the initial value of the state at time 0 we generate the state sequence $x = \{x(0), x(1), \ldots, x(N)\}$ with $x(0) = x_0$ and $x(i) = \phi(i; x_0, u)$. The second system we consider is the following uncertain system

$$x_{k+1} = f(x, u, w)$$

with the uncertainty w, the control variable $\mu = \{u(0), \mu_1(\cdot), \ldots, \mu_{N-1}(\cdot)\}$. With the initial state x_0 the control policy μ generates a state tube which we denote by

$$\mathcal{X}(x_0, \mu) = \{X(0; x_0), X(1; x_0, \mu), \ldots, X(N; x_0, \mu)\}$$

with $X(0; x_0) = \{x_0\}$ and for all $i \in \mathbb{N}_0$ it holds that

$$X(i; x_0, \mu) = \{\phi(i; x_0, \mu, w)|w \in W\}.$$

Open-loop as well as feedback control both generate a tube of trajectories in the case of uncertainties. The state tube $\mathcal{X}(x, \mu)$ is a bundle of state trajectories with each trajectory corresponding to one realization of an admissible disturbance sequence w. The tube $\mathcal{X}$ represents the solution of the following set difference equation

$$X(i+1) = F(X(i), \mu_i(\cdot)),\ X(0) = \{x_0\}$$

with $F(X, \mu_i(\cdot)) := \{f(x_0, \mu_i(x), w)|x_0 \in X\ \ w \in \mathbb{W}\}$. The MPC problem (4.30) at state x_0 and with $V_N(\cdot)$ and $J_N(\cdot)$ defined in (4.28) and (4.29), respectively, is the same as before

$$\inf_{\mu}\{V_N(x_0, \mu)|\mu \in M(x_0)\}$$

with $M(x_0)$ being the set of feedback policies $\mu = \{u(0), \mu_1(\cdot), \dots, \mu_{N-1}(\cdot)\}$. The state contstraints have to be satisfied by every trajectory in the tube. That means that the control of uncertain systems can rather be seen as the control of tubes than of trajectories, which means that for each initial state a tube is choosen in which all realizations of the state trajectory are bounded. If this choice is suitable we can guarantee that the state and control constraints are satisfied for all possible realizations of the disturbance sequence. But to determine an exact tube is very difficult for linear systems and almost impossible for nonlinear systems. Therefore in Rawlings & Mayne [69] different possibilities for the construction of simple tubes which bound all realizations of the state trajectory are presented. These are approximation of the exact tube which lies inside. Further information about tubes can be found e.g. in Aubin [2] and Bertsekas and Rhodes [11], [12].

5 Special Bilevel Optimization Problems

In Chapter 2 we already discussed general bilevel optimization problems. In this chapter we will describe special bilevel optimization problems with trust region constraints. These problems will be used to describe the algorithms in Chapter 6 and 7.

For completeness we first consider the following bilevel optimization problem, which can be found in Kostyukova & Kostina [47]

$$\min_{\phi} \left\{ \phi^T \mathcal{G} \phi + \max_{\xi} \left\{ \xi^T \mathcal{D} \xi : \ \|\xi - r - \mathcal{G}\phi\|^2_{Q^{-1}} \leq v \right\} \right\} \tag{5.1}$$

with $\phi \in \mathbb{R}^n, \ \xi \in \mathbb{R}^n$ unknown optimization variables, $Q \in \mathbb{R}^{n \times n}, \ \mathcal{G} \in \mathbb{R}^{n \times n}, \ \mathcal{D} \in \mathbb{R}^{n \times n}$ given positive definite matrices, $r \in \mathbb{R}^n$ a given vector and v a given number which will be used to describe the algorithms in Chapter 6. Our aim is to analyze the problem and describe an algorithm to solve it. Using the variable transformation

$$p = \mathcal{M}\xi, \ \psi = \mathcal{M}^{-T}\phi, \ Q^{-1} = \mathcal{M}^T\mathcal{M}$$

we can rewrite problem (5.1) as

$$\min_{\psi} \left(\psi^T G \psi + \max_{p} p^T D p \right), \text{ s.t. } \|p - d - G\psi\|^2 \leq v \tag{5.2}$$

where $G = \mathcal{M}\mathcal{G}\mathcal{M}^T, \ D = \mathcal{M}^{-T}\mathcal{D}\mathcal{M}^{-1}, \ d = \mathcal{M}r$. Let us first consider the lower level problem of problem (5.2):

$$\max_{p} \ p^T D p, \text{ s.t. } \|p - d - G\psi\|^2 \leq v. \tag{5.3}$$

This is a trust region type problem and its solution is described by the following lemma.

Lemma 5.1.
A vector $p \in \mathbb{R}^n$ is optimal in problem (5.3) if and only if there exists $\lambda \geq \lambda_{\max}$ such that p and λ satisfy

$$(-D + \lambda\mathbb{I})p - \lambda(d + G\psi) = 0 \tag{5.4a}$$
$$\|p - d - G\psi\|^2 = v \tag{5.4b}$$
$$\lambda \geq \lambda_{\max}. \tag{5.4c}$$

Here, $\lambda_{\max} > 0$ is the maximal eigenvalue of D.

Proof. For the proof see Sorensen [78]. □

Now let us consider the bilevel problem (5.2). Following the third approach of solving bilevel problems, we can rewrite problem (5.2) as

$$\min f(\psi),\ \psi \in \mathbb{R}^n \tag{5.5}$$

with

$$f(\psi) := \psi^T G\psi + \max_{\|p-d-G\psi\|^2 \le v} p^T Dp,\ \psi \in \mathbb{R}^n. \tag{5.6}$$

Using the solution of problem (5.3) (cf. Lemma 5.1) we can rewrite (5.4a) as:

$$\begin{aligned}
(-D+\lambda\mathbb{I})p &= \lambda(d+G\psi)\\
Dp &= \lambda(p-d-G\psi)\\
(p-(d+G\psi))^T Dp &= \lambda\underbrace{(p-(d+G\psi))^T(p-(d+G\psi))}_{=v}\\
(p-d-G\psi)^T Dp &= \lambda(p-d-G\psi)^T(p-d-G\psi) = \lambda v
\end{aligned}$$

and then

$$p^T Dp = (d+G\psi)^T Dp + \lambda v.$$

and hence

$$\max_{\|p-d-G\psi\|^2 \le v} p^T Dp = \lambda v + (d+G\psi)^T D\lambda(-D+\lambda\mathbb{I})^{-1}(d+G\psi)$$

for some $\lambda \ge \lambda_{\max}$. Thus, there exists a $\lambda \ge \lambda_{\max}$ such that

$$f(\psi) = \psi^T G\psi + \lambda v + (d+G\psi)^T D\lambda(-D+\lambda\mathbb{I})^{-1}(d+G\psi).$$

Function $f(\psi)$ is convex but non-smooth. Let us consider

$$\psi(\lambda) = -\left(-\frac{\mathbb{I}}{\lambda} + D^{-1} + G\right)^{-1} d. \tag{5.7}$$

and denote

$$\psi^* = \psi(\lambda_{\max}) = \left(\frac{\mathbb{I}}{\lambda_{\max}} - D^{-1} - G\right)^{-1} d. \tag{5.8}$$

Let us note that $\left(-\frac{\mathbb{I}}{\lambda} + D^{-1} + G\right)$ is positive definite for $\lambda \ge \lambda_{\max}$. The following two cases can now occur:

$$1)\ \|\psi^*\| < \lambda_{\max}^2 v \tag{5.9}$$
$$2)\ \|\psi^*\| \ge \lambda_{\max}^2 v \tag{5.10}$$

In the first case the vector ψ^* is optimal in (5.5) (see Kostyukova & Kostina [47]).

In the second case, namely $\|\psi^*\| \ge \lambda_{\max}^2 v$, there exists a unique number $\lambda_* \ge \lambda_{\max}$ for which

$$\|\psi(\lambda_*)\|^2 = \lambda_*^2 v \tag{5.11}$$

and $\psi(\lambda_*)$ (5.7) is optimal in (5.5) and corresponding

$$p(\lambda^*) = -D^{-1}\psi(\lambda^*)$$

is optimal in (5.3). In Kostyukova & Kostina [47] it was shown that

Theorem 5.2. *Problem (5.5)*

$$\min f(\psi),\ \ \psi \in \mathbb{R}^n$$

with

$$f(\psi) := \psi^T G\psi + \max_{\|p-d-G\psi\|^2\le v} p^T Dp,\ \ \psi \in \mathbb{R}^n$$

(and consequently Problem (5.2)) is equivalent to the following optimization problem

$$\min_{\lambda} g(\lambda) := d^T\left(D^{-1} + G - \frac{\mathbb{I}}{\lambda}\right)^{-1} d + \lambda v,\ \ s.t.\ \ \lambda \ge \lambda_{\max} \tag{5.12}$$

in the sense that if λ^0 is optimal in (5.12), then the vector $\psi(\lambda^0) = \left(\frac{\mathbb{I}}{\lambda^0} - D^{-1} + G\right)^{-1} d$ solves problem (5.5) and $p(\lambda^0) = -D^{-1}\psi(\lambda^0)$.

Altogether, we can formulate the following algorithm for solving problem (5.5).

Algorithm 5.3. (Solution of Problem (5.5))

`COMPUTE` $\lambda_{\max} = \lambda_{\max}(\mathcal{M}^{-T}\mathcal{D}\mathcal{M}^{-1})$
`COMPUTE` vector $\psi^* = \left(\frac{\mathbb{I}}{\lambda_{\max}} - D^{-1} - G\right)^{-1} d$
  `IF` Case 1) $\|\psi^*\|^2 \le \lambda^2_{\max} v$
    `THEN` solution of Problem (5.5) is given by $\psi^0 = \psi^*$, $p^0 = -D^{-1}\psi^0$
  `ELSE` Case 2) $\|\psi^*\|^2 > \lambda^2_{\max} v$
    `THEN` solve equation $\|\psi(\lambda)\|^2 = \lambda^2 v$ for $\lambda \ge \lambda_{\max}$ (e.g. by Newton's method)
      where $\psi(\lambda) := \left(\frac{\mathbb{I}}{\lambda} - D^{-1} + G\right)^{-1} d$, solution: λ^*
      `WRITE` solution of problem (5.5) as $\psi^0 = \psi(\lambda^*),\ p^0 = -D^{-1}\psi^0$

Using the solution of (5.5) the solution of (5.1) can be computed by using the following variable transformation

$$\phi^0 = \mathcal{M}^T\psi^0,\ \xi^0 = \mathcal{M}^{-1}p^0,$$

where $Q^{-1} = \mathcal{M}^T\mathcal{M}$. With the transformation

$$G = \mathcal{M}\mathcal{G}\mathcal{M}^T,\ \ D = \mathcal{M}^{-T}\mathcal{D}\mathcal{M}^{-1},\ \ d = \mathcal{M}r$$

we can write ψ as

$$\begin{aligned}\psi &= \left(\frac{\mathbb{I}}{\lambda} - D^{-1} - G\right)^{-1} d = \left(\frac{\mathbb{I}}{\lambda} - \mathcal{M}\mathcal{D}^{-1}\mathcal{M}^T - \mathcal{M}\mathcal{G}\mathcal{M}^T\right)^{-1} d\\ &= \left(\frac{\mathcal{M}\mathcal{M}^{-1}\mathcal{M}^{-T}\mathcal{M}^T}{\lambda} - \mathcal{M}\mathcal{D}^{-1}\mathcal{M}^T - \mathcal{M}\mathcal{G}\mathcal{M}^T\right)^{-1} d\end{aligned}$$

$$= \left(\mathcal{M} \left(\frac{\mathcal{M}^{-1}\mathcal{M}^{-T}}{\lambda} - \mathcal{D}^{-1} - \mathcal{G} \right)^{-1} \mathcal{M}^T \right)^{-1} d$$
$$= \mathcal{M}^{-T} \left(\frac{\mathcal{M}^{-1}\mathcal{M}^{-T}}{\lambda} - \mathcal{D}^{-1} - \mathcal{G} \right)^{-1} \mathcal{M}^{-1} d$$
$$= \mathcal{M}^{-T} \left(\frac{Q}{\lambda} - \mathcal{D}^{-1} - \mathcal{G} \right)^{-1} \mathcal{M}^{-1} d$$

and hence, we can write the solution of problem (5.1) as

$$\phi = \mathcal{M}^T \psi = \left(\frac{Q}{\lambda} - \mathcal{D}^{-1} - \mathcal{G} \right)^{-1} r.$$

Therefore we can solve problem (5.1) with the following algorithm.

Algorithm 5.4. (Solution of Problem (5.1))

`COMPUTE` $\lambda_{\max} = \lambda_{\max}(\mathcal{M}^{-T}\mathcal{D}\mathcal{M}^{-1})$
`COMPUTE` vector $\phi^* = \left(\frac{Q}{\lambda_{\max}} - \mathcal{D}^{-1} - \mathcal{G} \right)^{-1} r$
  `IF` Case 1) $\|\phi^*\|_Q^2 \leq \lambda_{\max}^2 v$
    `THEN` solution of problem (5.1) is given by $\phi^0 = \phi^*$
  `ELSE` Case 2) $\|\phi^*\|_Q^2 > \lambda_{\max}^2 v$
    `THEN` solve equation $\|\phi(\lambda)\|_Q^2 = \lambda^2 v$ for $\lambda \geq \lambda_{\max}$ (e.g. by Newton's method)
      solution: λ^*
      `WRITE` solution of problem (5.1) as $\phi^0 = \phi(\lambda^*)$

Since $f(\psi)$ is convex, with this algorithm we receive a guaranteed global solution for the lower and the upper level problems of the bilevel problem (5.1).

The method can be generalized for the following bilevel problem with linear equality constraints

$$\min_{\phi} \left\{ \phi^T \mathcal{G} \phi + \max_{\xi} \left\{ \xi^T \mathcal{D} \xi : \ \|\xi - r - \mathcal{G}\phi\|_{Q^{-1}}^2 \leq v, \ A\xi = b \right\} \right\} \tag{5.13}$$

where $A \in \mathbb{R}^{p \times n}$ and rank $A = p$. Indeed, let us consider the lower level problem

$$\begin{aligned} \max_{\xi} \ & \xi^T \mathcal{D} \xi \\ \text{s.t.} \ & \|\xi - r - \mathcal{G}\phi\|_{Q^{-1}}^2 \leq v, \ A\xi = b. \end{aligned}$$

We can write

$$\xi = Fz + \hat{\xi}$$

where $F \in \mathbb{R}^{n \times (n-p)}$ is chosen such that $AF = 0$, rank $F = n - p$ and $AFz + A\hat{\xi} = b$, where $\hat{\xi}$ is a particular solution of $A\xi = b$. Then

$$\xi^T \mathcal{D} \xi = \left(z^T F^T + \hat{\xi}^T \right) \mathcal{D} \left(Fz + \hat{\xi} \right) = z^T F^T \mathcal{D} F z + 2\hat{\xi}^T \mathcal{D} F z + \hat{\xi}^T \mathcal{D} \hat{\xi}$$

and the lower level problem can be written as

$$\begin{aligned} &\max_z \; z^T \bar{\mathcal{D}} z + \bar{g}^T z \\ &\text{s.t. } (Fz - \bar{r} - \mathcal{G}\phi)^T Q^{-1} (Fz - \bar{r} - \mathcal{G}\phi) \le v \end{aligned}$$

when $\bar{\mathcal{D}} = F^T \mathcal{D} F$, $\bar{g}^T = 2\hat{\xi}^T \mathcal{D} F$ and $\bar{r} = r - \hat{\xi}$. With a few modifications we can use the methods that we described before.

5.1 Bilevel Optimization Problem with a Trust Region-type Constraint in the Lower Level

Now we consider another bilevel problem, which we will need to describe the algorithms for solving the problems that we will present in Section 7.4

$$\min_p \{p^T G p + \max_\xi \{\xi^T D \xi : \; (d + \xi - p)^T Q^{-1} (d + \xi - p) \le v\}\} \tag{5.14}$$

where $\xi \in \mathbb{R}^n$, $p \in \mathbb{R}^n$, $Q = AA^T$ and A invertible. With the following change of variables

$$y = A^{-1}\xi, \quad \phi = A^{-1} p$$

and the notations

$$\mathcal{G} = A^T G A, \; \mathcal{D} = A^T D A, \; r = A^{-1} d$$

we can rewrite problem (5.14) as

$$\min_\phi \{\phi^T \mathcal{G} \phi + \max_y \{y^T \mathcal{D} y : \; (r + y - \phi)^T (r + y - \phi) \le v\}\} \tag{5.15}$$

with $\phi \in \mathbb{R}^n$, $y \in \mathbb{R}^n$ unknown optimization variables, $\mathcal{G} \in \mathbb{R}^{n \times n}$, $\mathcal{D} \in \mathbb{R}^{n \times n}$ given positive definite matrices, $r \in \mathbb{R}^n$ a given vector and v a given number. Let us first consider the lower level problem

$$\begin{aligned} &\max_y \; y^T \mathcal{D} y \\ &\text{s.t. } (r + y - \phi)^T (r + y - \phi) \le v. \end{aligned} \tag{5.16}$$

We can formulate the following lemma for the necessary and sufficient conditions for a point $y \in \mathbb{R}^n$ to be optimal in problem (5.16).

Lemma 5.5. *A vector $y \in \mathbb{R}^n$ is optimal in problem (5.16) if and only if there exists a number $\lambda \ge \lambda_{\max}$ such that y and λ satisfy*

$$\begin{aligned} \left(-\frac{\mathbb{I}}{\lambda} + \mathcal{D}^{-1}\right) \mathcal{D} y + (r - \phi) &= 0 \\ (r - \phi + y)^T (r - \phi + y) &= v \\ \lambda &\ge \lambda_{\max} \end{aligned} \tag{5.17}$$

where $\lambda_{\max}$ denotes a maximal eigenvalue of $\mathcal{D}$.

Remark 5.6. *Note, that for* $\lambda \geq \lambda_{\max}$ *we have, that the matrix* $(-\mathcal{D} + \lambda\mathbb{I})$ *is positive semidefinite and hence the matrix* $\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)$ *is positive semidefinite as well.*

Proof. The Lagrangian of problem (5.16) can be formulated as

$$\mathcal{L}(y, \lambda) = -y^T \mathcal{D} y - \lambda(v - (r - \phi + y)^T (r - \phi + y))$$

and the corresponding gradient is

$$\nabla_y \mathcal{L}(y, \lambda) = -2\mathcal{D}y + 2\lambda(r - \phi + y) \stackrel{!}{=} 0 \tag{5.18}$$

As it holds that

$$(-\mathcal{D} + \lambda\mathbb{I}) = \lambda \left(-\frac{\mathbb{I}}{\lambda} + \mathcal{D}^{-1}\right) \mathcal{D}$$

we can rewrite (5.18)

$$\left(-\frac{\mathbb{I}}{\lambda} + \mathcal{D}^{-1}\right) \mathcal{D}y = -(r - \phi)$$

and with Lemma 2.4 and Lemma 2.8 from Sorensen [78] Lemma 5.5 is proven. □

Let us consider the bilevel problem (5.15) which can be rewritten as

$$\min_{\phi} \Phi(\phi), \ \phi \in \mathbb{R}^n \tag{5.19}$$

with

$$\Phi(\phi) = \phi^T \mathcal{G} \phi + \max_{\|r+y-\phi\|_2^2 \leq v} y^T \mathcal{D} y.$$

Lemma 5.7. *The function* $\Phi(\phi)$ *in problem (5.19) is convex.*

Proof. Let us consider an arbitrary $\bar{\phi}$ and assume that

$$\bar{\phi} = \lambda\phi' + (1 - \lambda)\phi''$$

with a scalar $\lambda \in]0, 1[$ and some vectors ϕ' and ϕ''. Let $\bar{y} = \bar{y}(\bar{\phi})$ be a vector solving the lower level problem (5.16) for $\phi = \bar{\phi}$, i.e.

$$\bar{y} = \arg \min_{\|r+y-\bar{\phi}\|_2^2 \leq v} y^T \mathcal{D} y$$

and denote

$$\Psi(\bar{\phi}) = \bar{y}^T \mathcal{D} \bar{y}.$$

Compute

$$\bar{y}' = \bar{y} - \left(\bar{\phi} - \phi'\right) \text{ and } \bar{y}'' = \bar{y} - \left(\bar{\phi} - \phi''\right).$$

Then it holds that

$$\begin{aligned} r + \bar{y}' - \phi' &= r + \bar{y} - \left(\bar{\phi} - \phi'\right) - \phi' = r + \bar{y} - \bar{\phi} \\ r + \bar{y}'' - \phi'' &= r + \bar{y} - \bar{\phi} \\ \text{and } \|r + \bar{y}' - \phi'\|_2^2 &= \|r + \bar{y}'' - \phi''\|_2^2 = \|r + \bar{y} - \bar{\phi}\|_2^2 \le v. \end{aligned}$$

Hence $\bar{y}'$ and $\bar{y}''$ are feasible in the lower level problem corresponding to ϕ' and ϕ'' respectively. Since

$$\Psi(\phi) = \max_{\|r+y-\phi\|_2^2 \le v} y^T \mathcal{D} y$$

it follows that

$$\Psi(\phi') \ge (\bar{y}')^T \mathcal{D} \bar{y}' \text{ and } \Psi(\phi'') \ge (\bar{y}'')^T \mathcal{D} \bar{y}''$$

We have

$$\lambda \Psi(\phi') + (1-\lambda)\Psi(\phi'') \ge \lambda (\bar{y}')^T \mathcal{D} \bar{y}' + (1-\lambda)(\bar{y}'')^T \mathcal{D} \bar{y}'' \ge \bar{y}^T \mathcal{D} \bar{y} = \Psi(\bar{\phi}).$$

Indeed, it holds that

$$\begin{aligned} \lambda(\bar{y}')^T \mathcal{D} \bar{y}' &= \lambda(\bar{y} - (\bar{\phi} - \phi'))^T \mathcal{D} (\bar{y} - (\bar{\phi} - \phi')) \\ &= \lambda \bar{y}^T \mathcal{D} \bar{y} + \lambda(\bar{\phi} - \phi')^T \mathcal{D} (\bar{\phi} - \phi') - 2\lambda \bar{y}^T \mathcal{D} (\bar{\phi} - \phi'). \end{aligned}$$

Analogously,

$$(1-\lambda)(\bar{y}'')^T \mathcal{D} \bar{y}'' = (1-\lambda)\bar{y}^T \mathcal{D} \bar{y} + (1-\lambda)(\bar{\phi} - \phi'')^T \mathcal{D}(\bar{\phi} - \phi') - 2(1-\lambda)\bar{y}^T \mathcal{D} (\bar{\phi} - \phi'').$$

Thus, it holds that

$$\begin{aligned} &\lambda(\bar{y}'')^T \mathcal{D} \bar{y}' + (1-\lambda)(\bar{y}'')^T \mathcal{D} \bar{y}'' \\ =&\bar{y}^T \mathcal{D} \bar{y} + \lambda(\bar{\phi} - \phi')^T \mathcal{D}(\bar{\phi} - \phi') + (1-\lambda)(\bar{\phi} - \phi'')^T \mathcal{D} (\bar{\phi} - \phi'') - 2\bar{y}^T \mathcal{D}(\lambda(\bar{\phi} - \phi') + (1-\lambda)(\bar{\phi} - \phi'')) \\ \ge&\bar{y}^T \mathcal{D} \bar{y}, \quad \text{since } \mathcal{D} \text{ is positive definite} \end{aligned}$$

and

$$\lambda(\bar{\phi} - \phi') + (1-\lambda)(\bar{\phi} - \phi'') = \bar{\phi} - (\lambda\phi' + (1-\lambda)\phi'') = \bar{\phi} - \bar{\phi} = 0$$

This means, that $\Psi(\cdot)$ is convex, and hence $\Phi(\cdot)$ is convex.

□

To formulate conditions for a solution ϕ of problem (5.19) we need the following notation

$$\phi(\lambda) = \mathcal{G}^{-1} \left(\mathcal{G}^{-1} + \mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} \right)^{-1} r.$$

Let us note that $\left(\mathcal{G}^{-1} + \mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)$ is positive definite for $\lambda \ge \lambda_{\max}(\mathcal{D})$. It is easy to check that the functions $\phi(\lambda)$ and $y(\lambda) = -\mathcal{D}^{-1}\mathcal{G}\phi(\lambda)$ satisfy

$$(-\mathcal{D} + \lambda \mathbb{I}) y = -\lambda (r - \phi).$$

Furthermore it holds that

$$r - \phi(\lambda) + y(\lambda) = \left(\frac{\mathcal{D}}{\lambda} - \mathbb{I}\right) y(\lambda) + y(\lambda) = \frac{\mathcal{D}}{\lambda} y(\lambda) = -\frac{1}{\lambda}\mathcal{G}\phi(\lambda).$$

Hence,

$$\|r - \phi(\lambda) + y(\lambda)\|_2^2 = \frac{1}{\lambda^2}\|\mathcal{G}\phi(\lambda)\|_2^2.$$

Denote $\phi^* = \phi(\lambda_{\max})$. Similarly to problem (5.5) we distinguish between the following two cases:

$$1)\ \|\mathcal{G}\phi^*\|_2^2 < \lambda_{\max}^2 v \tag{5.20}$$

$$2)\ \|\mathcal{G}\phi^*\|_2^2 \geq \lambda_{\max}^2 v \tag{5.21}$$

Let us first consider Case 1). Let

$$\mathcal{D} = U^T \Lambda U,\ U^T U = \mathbb{I},\ \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

with $0 < \lambda_1 \leq \dots \leq \lambda_n = \lambda_{\max}$ denoting the eigenvalues of $\mathcal{D}$. Consider the vector

$$y_\beta = -\mathcal{D}^{-1}\mathcal{G}\phi^* + \beta u$$

where $\beta \in \mathbb{R}$ and u is an eigenvector corresponding to $\lambda_{\max}$ with $\|u\|_2^2 = 1$. Then

$$\mathcal{D}y_\beta = -\mathcal{G}\phi^* + \lambda_{\max}\beta u.$$

Obviously, for $\phi^* = \mathcal{G}^{-1}\left(\mathcal{G}^{-1} + \mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda_{\max}}\right)^{-1} r$ we have

$$\begin{aligned}(-\mathcal{D} + \lambda_{\max}\mathbb{I})y_\beta &= (-\mathcal{D} + \lambda_{\max}\mathbb{I})(-\mathcal{D}^{-1}\mathcal{G}\phi^*) + \underbrace{(-\mathcal{D}u + \lambda_{\max}u)}_{=0}\beta \\ &= (-\mathcal{D} + \lambda_{\max}\mathbb{I})(-\mathcal{D}^{-1}\mathcal{G}\phi^*) = -\lambda_{\max}(r - \phi_*)\end{aligned}$$

and it holds that

$$\begin{aligned} r - \phi^* + y_\beta &= -\frac{1}{\lambda_{\max}}(-\mathcal{D} + \lambda_{\max}\mathbb{I})y_\beta + y_\beta \\ &= \frac{\mathcal{D}}{\lambda_{\max}} y_\beta = -\frac{\mathcal{G}\phi^*}{\lambda_{\max}} + \beta u \\ &= -\frac{1}{\lambda_{\max}}(\mathcal{G}\phi^* - \lambda_{\max}\beta u).\end{aligned}$$

Since $\|\mathcal{G}\phi^*\|_2^2 < \lambda_{\max}^2 v$ then there exists a number $\beta \neq 0$ that satisfies

$$\|r - \phi^* + y_\beta\|_2^2 = \frac{1}{\lambda_{\max}^2}\|\mathcal{G}\phi^* - \lambda_{\max}\beta u\|_2^2$$

$$= \frac{1}{\lambda_{\max}^2} \left(\|\mathcal{G}\phi^*\|_2^2 - 2\lambda_{\max}\beta \left(u^T\mathcal{G}\phi^*\right) + \lambda_{\max}^2\beta^2 \underbrace{\|u\|_2^2}_{=1} \right)$$
$$= v.$$

Indeed, β should solve the quadratic equation

$$\|\mathcal{G}\phi^*\|_2^2 - 2\lambda_{\max}\beta \left(u^T\mathcal{G}\phi^*\right) + \lambda_{\max}^2\beta^2 = \lambda_{\max}^2 v. \tag{5.22}$$

Obviously, the solutions of (5.22) are the numbers β_1 and β_2

$$\beta_1 = \frac{u^T\mathcal{G}\phi^*}{\lambda_{\max}} + \frac{\sqrt{(u^T\mathcal{G}\phi^*)^2 + \left(\lambda_{\max}^2 v - \|\mathcal{G}\phi^*\|_2^2\right)}}{\lambda_{\max}},$$
$$\beta_2 = \frac{u^T\mathcal{G}\phi^*}{\lambda_{\max}} - \frac{\sqrt{(u^T\mathcal{G}\phi^*)^2 + \left(\lambda_{\max}^2 v - \|\mathcal{G}\phi^*\|_2^2\right)}}{\lambda_{\max}}.$$

For the solution β_1 we have $\beta_1 > 0$ and for the solution β_2 we have $\beta_2 < 0$ since $\lambda_{\max}^2 v - \|\mathcal{G}\phi^*\|_2^2 > 0$ and

$$\sqrt{\left(u^T\mathcal{G}\phi^*\right)^2 + \left(\lambda_{\max}^2 v - \|\mathcal{G}\phi^*\|_2^2\right)} > |u^T\mathcal{G}\phi^*|.$$

Thus we have shown that y_β and ϕ^* satisfy

$$\begin{aligned} &(1)\ \ (-\mathcal{D} + \lambda_{\max}\mathbb{I})\, y_\beta = -\lambda_{\max}(r - \phi^*) \\ &(2)\ \ \|r + y_\beta - \phi^*\|_2^2 = v \end{aligned} \tag{5.23}$$

yielding

$$\max_{\|r+y-\phi^*\|_2^2 \le v} y^T\mathcal{D}y = y_\beta^T\mathcal{D}y_\beta$$

and

$$\Phi(\phi^*) = (\phi^*)^T\mathcal{G}\phi^* + y_\beta^T\mathcal{D}y_\beta.$$

Consider $\bar{\phi} = \phi^* + \Delta\phi$ and $\bar{y} = y_\beta + \Delta y$ with $\Delta y = \Delta\phi$. We have

$$r - \underbrace{(\phi^* + \Delta\phi)}_{\bar{\phi}} + \underbrace{(y_\beta + \Delta y)}_{\bar{y}} = r - \phi^* + y_\beta$$

and hence

$$\|r - \bar{\phi} + \bar{y}\|_2^2 = v$$

and

$$\max_{\|r-\bar{\phi}+y\|_2^2 \le v} y^T\mathcal{D}y \ge \bar{y}^T\mathcal{D}\bar{y}.$$

Compute

$$\begin{aligned}
\Phi(\bar\phi) - \Phi(\phi^*) &= \bar\phi^T \mathcal{G} \bar\phi + \max_{\|r+y-\bar\phi\|_2^2 \le v} y^T \mathcal{D} y - (\phi^*)^T \mathcal{G} \phi^* - y_\beta^T \mathcal{D} y_\beta \\
&\ge \bar\phi^T \mathcal{G} \bar\phi + \bar y^T \mathcal{D} \bar y - (\phi^*)^T \mathcal{G} \phi^* - y_\beta^T \mathcal{D} y_\beta \\
&= (\phi^* + \Delta\phi)^T \mathcal{G} (\phi^* + \Delta\phi) + (y_\beta + \Delta y)^T \mathcal{D} (y_\beta + \Delta y) - (\phi^*)^T \mathcal{G} \phi^* - y_\beta^T \mathcal{D} y_\beta \\
&= \Delta\phi^T \mathcal{G} \Delta\phi + 2\Delta\phi^T \mathcal{G} \phi^* + \underbrace{\Delta y^T}_{\Delta\phi^T} \mathcal{D} \underbrace{\Delta y}_{\Delta\phi} + 2 \underbrace{\Delta y^T}_{\Delta\phi^T} \mathcal{D} y_\beta \\
&= \Delta\phi^T (\mathcal{G} + \mathcal{D}) \Delta\phi + 2\Delta\phi^T (\mathcal{G}\phi^* + \mathcal{D} y_\beta) \\
&= \Delta\phi^T (\mathcal{G} + \mathcal{D}) \Delta\phi + 2\Delta\phi^T \lambda_{\max} \beta u
\end{aligned}$$

for all $\Delta\phi$ and

$$\bar y = y_\beta + \Delta y, \; \Delta y = \Delta\phi, \; \beta = \beta_1 \text{ or } \beta = \beta_2$$

with $\beta_1 > 0$ and $\beta_2 < 0$. Therefore, in the case $\Delta\phi^T u \ge 0$ we choose $\beta = \beta_1$, in case $\Delta\phi^T u < 0$ we choose $\beta = \beta_2$ and get in both cases

$$\Phi(\bar\phi) - \Phi(\phi^*) > 0 \text{ for all } \bar\phi \neq \phi^*$$

since $\mathcal{G}$ and $\mathcal{D}$ are positive definite.
In the case that we have more than one $\lambda_{\max}$. We assume the following set

$$I_* = \{1 \le i \le n : \lambda_i = \lambda_{\max}\}$$

with $u_* = (u_i, \; i \in I_*)$. Then we consider the vector

$$y_\beta = -\mathcal{D}^{-1} \mathcal{G} \phi^* + \sum_{i \in I_*} \beta_i u_i$$

where $\beta_i \in \mathbb{R}$ and u_i are the eigenvectors corresponding to $\lambda_{\max}$ with $\|u_i\|_2^2 = 1$. Then

$$\mathcal{D} y_\beta = -\mathcal{G} \phi^* + \lambda_{\max} \sum_{i \in I_*} \beta_i u_i$$

We have that

$$\begin{aligned}
(-\mathcal{D} + \lambda_{\max} \mathbb{I}) y_\beta &= (-\mathcal{D} + \lambda_{\max} \mathbb{I})(-\mathcal{D}^{-1} \mathcal{G} \phi^* + \sum_{i \in I_*} \beta_i \underbrace{(-\mathcal{D} u_i + \lambda_{\max} u_i}_{=0} \\
&= -\lambda_{\max}(r - \phi *)
\end{aligned}$$

and it holds that

$$\begin{aligned}
r - \phi^* + y_\beta = \frac{\mathcal{D}}{\lambda_{\max}} y_\beta &= -\frac{\mathcal{G}\phi^*}{\lambda_{\max}} + \sum_{i \in I_*} \beta_i u_i \\
&= \frac{1}{\lambda_{\max}} \left(-\mathcal{G}\phi^* + \lambda_{\max} \sum_{i \in I_*} \beta_i u_i \right).
\end{aligned}$$

Since $\|\mathcal{G}\phi^*\| < \lambda_{\max}^2 v$ there exist numbers $\beta_i \neq 0$ that satisfy

$$\|r - \phi^* + y_\beta\|_2^2 = \frac{1}{\lambda_{\max}^2}\left(\|\mathcal{G}\phi^*\|_2^2 - 2\lambda_{\max}\sum_{i\in I_*}\beta_i(u_i^T\mathcal{G}\phi^*) + \lambda_{\max}^2\sum_{i\in I_*}\beta_i^2\|u_i\|_2^2 m\right) \stackrel{!}{=} v. \tag{5.24}$$

We choose

$$\beta_i = \frac{u_i^T\mathcal{G}\phi^*}{\lambda_{\max}}c$$

with $c \in \mathbb{R}$ solving the quadratic equation

$$\|\mathcal{G}\phi^*\|_2^2 - 2\lambda_{\max}c\left(\sum_{i\in I_*}\frac{1}{\lambda_{\max}}(u_i^T\mathcal{G}\phi^*)^2\right) + \lambda_{\max}^2 c^2\left(\sum_{i\in I_*}\frac{1}{\lambda_{\max}^2}(u_i^T\mathcal{G}\phi^*)^2\right) = \lambda_{\max}^2 v. \tag{5.25}$$

Obviously, the solutions of (5.25) are the numbers

$$c_1 = 1 + \frac{\sqrt{\lambda_{\max}^2 v - \|\mathcal{G}\phi^*\|_2^2 + \|u_*^T\mathcal{G}\phi^*\|_2^2}}{\|u_*^T\mathcal{G}\phi^*\|_2^2} > 0$$
$$c_2 = 1 - \frac{\sqrt{\lambda_{\max}^2 v - \|\mathcal{G}\phi^*\|_2^2 + \|u_*^T\mathcal{G}\phi^*\|_2^2}}{\|u_*^T\mathcal{G}\phi^*\|_2^2} < 0.$$

It holds that

$$\beta = \begin{pmatrix}\beta_1\\ \vdots \\ \beta_{|I_*|}\end{pmatrix} \in \mathcal{B}$$

where $\mathcal{B}$ are all vectors that satisfy (5.24). That means, we have shown that $y_\beta = -\mathcal{D}^{-1}\mathcal{G}\phi^* + \sum_{i\in I_*}\beta_i u_i$ and ϕ^* satisfy

$$(1)\ \ (-\mathcal{D} + \lambda_{\max}\mathbb{I})\, y_\beta = -\lambda_{\max}(r - \phi^*)$$
$$(2)\ \ \|r + y_\beta - \phi^*\|_2^2 = v$$

and hence

$$\max_{\|r+y-\phi^*\|_2^2 \le v} y^T\mathcal{D}y = y_\beta^T\mathcal{D}y_\beta.$$

Therefore ϕ^* is optimal. Thus, we have proven the following Lemma:

Lemma 5.8. *If the vector ϕ^* satisfies Case 1) (5.20), then ϕ^* is optimal in the bilevel problem (5.19).*

Let us now consider Case 2). For this case, we can formulate the following lemma.

Lemma 5.9. *If the vector ϕ^* satisfies Case 2), i.e. $\|\mathcal{G}\phi^*\|_2^2 \geq \lambda_{\max}^2 v$, then there exists a unique number $\lambda^* \geq \lambda_{\max}$ for which*

$$\|\mathcal{G}\phi(\lambda^*)\|_2^2 = (\lambda^*)^2 v$$

where

$$\phi(\lambda) = \mathcal{G}^{-1}\left(\mathcal{G}^{-1} + \mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)^{-1} r.$$

Proof. Consider the function $\eta(\lambda) = \|\mathcal{G}\phi\|_2^2 - \lambda^2 v$. We have that $\eta(\lambda_{\max}) \geq 0$ and $\eta(+\infty) = -\infty$. We denote

$$P(\lambda) = \mathcal{G}^{-1} + \mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}.$$

It holds that

$$\begin{aligned}
\frac{d}{d\lambda}\left(\eta(\lambda_{\max})\right) &= \frac{d}{d\lambda}\left(r^T (P(\lambda)^{-1})^T P(\lambda)^{-1} r - \lambda^2 v\right)\\
&= 2r^T P(\lambda)^{-1}\left(-P(\lambda)^{-1}\left(\frac{\mathbb{I}}{\lambda^2}\right)P(\lambda)^{-1} r\right) - 2\lambda v\\
&= -\frac{1}{\lambda^2} r^T P(\lambda)^{-1} P(\lambda)^{-1} P(\lambda)^{-1} r - 2\lambda v\\
&= -\frac{1}{\lambda^2}\|P(\lambda)^{-1} r\|_{P(\lambda)^{-1}}^2 - 2\lambda v \leq 0,
\end{aligned}$$

since $P(\lambda)$ is a positive definite matrix. That means $\eta(\lambda)$, $\lambda \geq \lambda_{\max}$ is monotonically decreasing and there exists a unique number $\lambda^* \geq \lambda_{\max}$ satisfying $\eta(\lambda^*) = 0$. □

If $\|\mathcal{G}\phi^*\| \geq \lambda_{\max}^2 v$, then $y(\lambda^*) = -\mathcal{D}^{-1}\mathcal{G}\phi(\lambda^*)$ and λ^* satisfy the optimality conditions in the lower level problem for $\phi^*(\lambda^*)$ by construction, hence

$$y(\lambda^*)^T \mathcal{D} y(\lambda^*) = \max_{\|r+y-\phi(\lambda^*)\|_2^2 \leq v} y^T \mathcal{D} y.$$

Lemma 5.10. *In case $\|\mathcal{G}\phi(\lambda_{\max})\|_2^2 > \lambda_{\max}^2 v$ the vector $\phi(\lambda^*)$, where $\lambda^* \geq \lambda_{\max}$ is found according to Lemma 5.9, is optimal in the bilevel optimization problem (5.19) (or (5.15)).*

Proof. Consider $\bar{\phi} = \phi(\lambda^*) + \Delta\phi$ and $\bar{y} = y(\lambda^*) + \Delta y$ with $\Delta y = \Delta\phi$. Obviously,

$$\|r + \bar{y} - \bar{y}\|_2^2 = \|r + y(\lambda^*) - \phi^*(\lambda^*)\|_2^2 = v.$$

Compute

$$\begin{aligned}
\Phi(\bar{\phi}) - \Phi(\phi^*) &= \bar{\phi}^T \mathcal{G}\bar{\phi} + \max_{\|r+y-\bar{\phi}\|_2^2 \leq v} y^T \mathcal{D} y - \phi(\lambda^*)^T \mathcal{G}\phi(\lambda^*) - y^T(\lambda^*)\mathcal{D} y(\lambda^*)\\
&\geq \bar{\phi}^T \mathcal{G}\bar{\phi} + \bar{y}^T \mathcal{D}\bar{y} - \phi(\lambda^*)^T \mathcal{G}\phi(\lambda^*) - y^T(\lambda^*)\mathcal{D} y(\lambda^*)\\
&= \Delta\phi^T \mathcal{G}\Delta\phi + 2\Delta\phi^T \mathcal{G}\phi(\lambda^*) + \underbrace{\Delta y^T}_{=\Delta\phi} \mathcal{D} \underbrace{\Delta y}_{=\Delta\phi} + 2\underbrace{\Delta y^T}_{=\Delta\phi} \mathcal{D} y(\lambda^*)
\end{aligned}$$

$$= \Delta\phi^T(\mathcal{G}+\mathcal{D})\Delta\phi + 2\Delta\phi^T \underbrace{(\mathcal{G}\phi(\lambda^*) + \mathcal{D}y(\lambda^*))}_{=0}$$
$$= \Delta\phi^T(\mathcal{G}+\mathcal{D})\Delta\phi > 0 \text{ if } \Delta\phi \neq 0$$

since $\mathcal{G}$ and $\mathcal{D}$ are positive definite. □

The following theorem follows directly from Lemmas 5.9 and 5.10.

Theorem 5.11. *Problem (5.19)*

$$\min_{\phi} \Phi(\phi),\ \ \phi \in \mathbb{R}^n$$

with

$$\Phi(\phi) = \phi^T\mathcal{G}\phi + \max_{\|y-\phi-r\|_2^2 \leq v} y^T\mathcal{D}y$$

(and consequently problem (5.15)) is equivalent to the following optimization problem

$$\min_{\lambda}\ g(\lambda) = \min_{\lambda}\ r^T\left(\mathcal{D}^{-1}+\mathcal{G}^{-1}-\frac{\mathbb{I}}{\lambda}\right)^{-1} r + \lambda v,\ \ s.t.\ \lambda \geq \lambda_{\max}(\mathcal{D}) \tag{5.26}$$

in the sense that if λ^0 is optimal in (5.26), then the vector

$$\phi(\lambda^0) = \mathcal{G}^{-1}\left(\mathcal{D}^{-1}+\mathcal{G}^{-1}-\frac{\mathbb{I}}{\lambda^0}\right)^{-1} r$$

is optimal in problem (5.19).

Proof. The Lagrangian of problem (5.26) is

$$\mathcal{L}(\lambda,\alpha) = g(\lambda) - \alpha(\lambda-\lambda_{\max}),\ \ \alpha \in \mathbb{R},\ \ \alpha \geq 0$$

and the derivative is

$$\frac{d\mathcal{L}(\lambda,\alpha)}{d\lambda} = -r^T\left(\mathcal{D}^{-1}+\mathcal{G}^{-1}-\frac{\mathbb{I}}{\lambda}\right)^{-1}\left(\frac{\mathbb{I}}{\lambda^2}\right)\left(\mathcal{D}^{-1}+\mathcal{G}^{-1}-\frac{\mathbb{I}}{\lambda}\right)^{-1} r + v - \alpha$$
$$= -\frac{1}{\lambda^2}\left\|\left(\mathcal{D}^{-1}+\mathcal{G}^{-1}-\frac{\mathbb{I}}{\lambda}\right) r\right\|_2^2 + v - \alpha.$$

If λ^0 is optimal in (5.26), then there exists a scalar $\alpha \geq 0$ such that it holds that $\frac{d\mathcal{L}(\lambda^0,\alpha)}{d\lambda} = 0$ and $\alpha(\lambda^0 - \lambda_{\max}) = 0$. In the case that $\lambda_{\max} = \lambda^0$ we have $\alpha \geq 0$ and $\|(\mathcal{D}^{-1}+\mathcal{G}^{-1}-\frac{\mathbb{I}}{\lambda_{\max}})^{-1}r\|_2^2 \leq \lambda_{\max}^2 v$ and if $\lambda_{\max} < \lambda^0$ then $\alpha = 0$ and $\|(\mathcal{D}^{-1}+\mathcal{G}^{-1}-\frac{\mathbb{I}}{\lambda^0})^{-1}r\|_2^2 = v$. □

Remark 5.12. *Following Xing et al. [84] the function $g(\lambda)$ can be considered as a canonical dual function of (5.19)*

Lemma 5.13. *Problem (5.26) is convex.*

Proof. The function $g(\lambda)$ is convex according to Lemma A1 in Kostina & Kostyukova [46] and the constraint in (5.26) is linear. □

Remark 5.14. *Having λ^* and $\phi(\lambda^*)$ we can easily restore the solution y^0 of the lower level problem (5.16). If $\|\mathcal{G}\phi^*\|_2^2 \geq \lambda_{\max}^2 v$ the solution is unique and given by $y(\lambda^0) = \mathcal{D}^{-1}\mathcal{G}\phi(x)$. If $\|\mathcal{G}\phi^*\|_2^2 < \lambda_{\max}^2 v$ all vectors y_β with $\beta \in \mathcal{B}$ can by taken as y^0. Note, that y_β is not unique, but global.*

Altogether we can formulate the following algorithm for solving (5.19).

Algorithm 5.15. (Solution of problem (5.19))
`COMPUTE` $\lambda_{\max} = \lambda_{\max}(\mathcal{D})$
`COMPUTE` vector $\phi^* = \mathcal{G}^{-1}(\mathcal{D}^{-1} + \mathcal{G}^{-1} - \frac{\mathbb{I}}{\lambda_{\max}})^{-1}(-r)$
  `IF` Case 1) $\|\mathcal{G}\phi^*\|_2^2 \leq \lambda_{\max}^2 v$
    `THEN` solution of problem (5.19) is given by $\phi^0 = \phi^*$
  `ELSE` Case 2) $\|\mathcal{G}\phi^*\|_2^2 > \lambda_{\max}^2 v$
    `THEN` solve equation $\|\mathcal{G}\phi(\lambda)\|_2^2 = \lambda^2 v$ for $\lambda \geq \lambda_{\max}$,
      where $\phi(\lambda) = \mathcal{G}^{-1}(\mathcal{D} + \mathcal{G}^{-1} - \frac{\mathbb{I}}{\lambda})^{-1} r$ (e.g. by Newton's method)
      solution: λ^*
      `WRITE` solution of problem (5.19) as $\phi^0 = \phi^*(\lambda^*)$

Let us consider the original problem (5.14), namely

$$\min_p \ p^T G p + \max_\xi \xi^T D \xi \tag{5.27}$$

$$\text{s.t. } (d + \xi - p)^T Q^{-1}(d + \xi - p) \leq v. \tag{5.28}$$

We have $\phi = A^{-1}p, y = A^{-1}\xi,\ \mathcal{G} = A^T G A,\ \mathcal{D} = A^T D A,\ r = A^{-1}d$ and hence

$$\mathcal{G}\phi = A^T G p,\ \|\mathcal{G}\phi\|_2^2 = \|Gp\|_Q^2.$$

The two cases (5.20) and (5.21) now read as

$$1)\ \|Gp\|_Q^2 < \lambda_{\max}^2 v \ \text{ and } \ 2)\ \|Gp\|_Q^2 \geq \lambda_{\max}^2 v,$$

where $\lambda_{\max} = \lambda_{\max}(A^T D A)$ and the Lemmas 5.8, 5.16 and 5.10 can be reformulated accordingly. Theorem 5.11 can be reformulated as:

Theorem 5.16. *Problem (5.27)*

$$\min_x \ p^T G p + \max_\xi \xi^T D \xi$$
$$s.t. \ (d + \xi - p)^T Q^{-1}(d + \xi - p) \leq v$$

is equivalent to the following optimization problem

$$\min_\lambda \ d^T \left(D^{-1} + G^{-1} - \frac{Q}{\lambda}\right)^{-1} d + \lambda v,\ \lambda \geq \lambda_{\max}(A^T D A). \tag{5.29}$$

in the sense that if λ^0 is optimal in problem (5.29), then $p^0 = G^{-1}\left(D^{-1} + G^{-1} - \frac{Q}{\lambda^0}\right)^{-1} d$ is optimal in the bilevel problem (5.27).

5.2 Bilevel Optimization Problem with a Non-Homogeneous Cost Functional

Let us consider the following bilevel problem with a non-homogeneous cost function which we will use later to describe the algorithms for solving the problems that we will present in Chapter 7

$$\min_{\phi} \phi^T \mathcal{G}\phi + 2s^T\phi + \max_{\|r+y-\phi\|_2^2 \le v} y^T\mathcal{D}y \tag{5.30}$$

with given symmetric positive definite matrices $\mathcal{G},\ \mathcal{D} \in \mathbb{R}^{n\times n}$, given vectors $s,\ r \in \mathbb{R}^n$ and a given number $v \in \mathbb{R}$. The lower level problem

$$\max_{y}\ y^T\mathcal{D}y \tag{5.31}$$

$$\text{s.t. } \|r+\phi-y\|_2^2 \le v \tag{5.32}$$

is the same as in the previous bilevel problem (5.15) and hence the optimality conditions for the lower level problem remain the same (cf. Lemma 5.5). Let us consider the bilevel problem (5.30) which can be rewritten as

$$\min_{\phi}\ \Phi(\phi),\ \phi \in \mathbb{R}^n \tag{5.33}$$

with

$$\Phi(\phi) = \phi^T\mathcal{G}\phi + 2s^T\phi + \max_{\|r+y-\phi\|_2^2 \le v} y^T\mathcal{D}y.$$

Denote the vector $\phi(\lambda)$ which solves the equation

$$\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)\mathcal{G}\phi = r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s$$

where $\lambda \ge \lambda_{\max}(\mathcal{D})$ and denote the vector $y(\lambda)$ which solves the equation

$$Dy(\lambda) = -\mathcal{G}\phi - s.$$

For $\lambda \ge \lambda_{\max}(\mathcal{D})$ the matrix $\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}$ is positive semidefinite and the matrix $D^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}$ is positive definite. Furthermore $\phi(\lambda)$ and $y(\lambda)$ satisfy

$$r + y - \phi = \frac{1}{\lambda}\mathcal{D}y$$

and

$$\|r+y-\phi\|_2^2 = \frac{1}{\lambda^2}\|\mathcal{D}y\|_2^2.$$

We denote $\phi^* = \phi(\lambda_{\max})$. Analogously to the previous section wen can prove:

1) If $\|\mathcal{G}\phi^* + s\|_2^2 < \lambda_{\max}^2 v$, then there exists an $\beta \in \mathbb{R}$ such that

$$y_\beta = -\mathcal{D}^{-1}(\mathcal{G}\phi - s) + \beta u$$

where u is an eigenvector of $\mathcal{D}$ corresponding to $\lambda_{\max}$, ϕ^* satisfies the optimality conditions of the lower level problem (cf. Lemma 5.8) and ϕ^* solves the bilevel problem (5.33), cf. Lemma 5.9.

2) If $\|\mathcal{G}\phi^* + s\|_2^2 \geq \lambda_{\max}^2 v$, then there exists a $\lambda^* > \lambda_{\max}$ such that

$$\|\mathcal{G}\phi(\lambda^*) + s\|_2^2 = (\lambda^*)^2 v,$$

the vector $\phi(\lambda^*)$ solves the bilevel problem (5.33) and $y(\lambda^*) = \mathcal{D}^{-1}(-\mathcal{G}\phi^* - s)$ satisfies the optimality conditions in the lower level problem for $\phi(\lambda^*)$.

Now we want to proof the following theorem (similar to Theorem 5.11).

Theorem 5.17. *Bilevel problem (5.33)*

$$\min_{\phi} \ \Phi(\phi), \ \phi \in \mathbb{R}^n$$

with

$$\Phi(\phi) = \phi^T \mathcal{G}\phi + 2s^T\phi + \max_{\|r+y-\phi\|_2^2 \leq v} y^T \mathcal{D} y.$$

is equivalent to the following optimization problem

$$\begin{aligned} g(\lambda) = \min_{\lambda \geq \lambda_{\max}(\mathcal{D})} & \left(r - (\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda})s\right)^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \left(r - (\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda})s\right) \\ & - s^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) s + 2r^T s + \lambda v \end{aligned} \tag{5.34}$$

in the sense that if λ^0 is optimal in (5.34) then the vector

$$\phi(\lambda^0) = \mathcal{G}^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) s\right)$$

is optimal in the bilevel problem (5.33) and it holds that $\Phi(\phi(\lambda^0)) = g(\lambda^0)$.

Proof. The Lagrangian of problem (5.34) is

$$\mathcal{L}(\lambda, \nu) = g(\lambda) - \nu(\lambda - \lambda_{\max}), \ \nu \in \mathbb{R}, \ \nu \geq 0$$

and the derivative is

$$\begin{aligned} \frac{d\mathcal{L}(\lambda,\nu)}{d\lambda} &= -\frac{2}{\lambda^2} s^T \mathcal{G}\phi(\lambda) - \frac{1}{\lambda^2}(\mathcal{G}\phi(\lambda))^T \mathcal{G}\phi(\lambda) - \frac{1}{\lambda^2} s^2 + v - \nu \\ &= -\frac{1}{\lambda^2}\|\mathcal{G}\phi(\lambda) + s\|_2^2 + v - \nu, \end{aligned}$$

If λ^0 is optimal in problem (5.34), then there exists a scalar $\nu \geq 0$ such that it holds that $\frac{d\mathcal{L}(\lambda^0,\nu)}{d\lambda} = 0$ and $\nu(\lambda^0 - \lambda_{\max}) = 0$. In the case that $\lambda_{\max} = \lambda^0$ we have $\nu \geq 0$ and $\|\mathcal{G}\phi(\lambda_{\max}) + s\|_2^2 \leq \lambda_{\max}^2 v$ and if $\lambda_{\max} < \lambda^0$ then $\nu = 0$ and $\|\mathcal{G}\phi(\lambda^0) + s\|_2^2 = (\lambda^0)^2 v$. Hence, with 1) and 2) it follows that $\phi(\lambda^0)$ solves problem (5.33).

Let us now show that $g(\lambda^0) = \Phi(\phi(\lambda^0))$. The optimal vectors $y(\lambda)$ and $\phi(\lambda)$ satisfy

$$\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{D}y = \phi - r$$
$$\|r + y - \phi\|_2^2 = v.$$

From $\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{D}y = \phi - r$ it follows that

$$y^T\mathcal{D}y = \lambda v - (r-\phi)^T\mathcal{D}y.$$

With $\mathcal{D}y = -\mathcal{G}\phi - s$ we get

$$y^T\mathcal{D}y = \lambda v + (r-\phi)^T\mathcal{G}\phi + (r-\phi)^T s.$$

Hence, it holds that

$$\begin{aligned}\Phi(\phi) &= \phi^T\mathcal{G}\phi + (r-\phi)^T\mathcal{G}\phi + (r-\phi)^T s + 2s^T\phi + \lambda v\\ &= r^T\mathcal{G}\phi + s^T\phi + r^T s + \lambda v.\end{aligned}$$

Using the formulas

$$\mathcal{G}\phi = \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s\right)$$

we get

$$\begin{aligned}\Phi(\phi) = r^T&\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} r - r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s\\ &+ s^T\mathcal{G}^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s\right) + s^T r + \lambda v.\end{aligned} \tag{5.35}$$

Let us consider the following terms

$$\begin{aligned}&- r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s + r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\mathcal{G}^{-1}s\\ =& r^T s - 2r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s\end{aligned}$$

and

$$\begin{aligned}&- s^T\mathcal{G}^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s\\ =& - s^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s + s^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s.\end{aligned}$$

If we insert these terms now in (5.35) we get

$$\Phi(\phi) = r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} r + r^T s$$

$$
\begin{aligned}
&- 2r^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) s - s^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) s \\
&+ s^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) s + s^T r + \lambda v \\
&= \left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) s\right)^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) s\right) \\
&- s^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) s + \lambda v + 2r^T s.
\end{aligned}
$$

Hence, it holds that $g(\lambda^0) = \Phi(\phi(\lambda^0))$. □

Lemma 5.18. *Problem (5.34) is convex.*

Proof.
- The function $g(\lambda),\ \lambda \in \mathbb{R}$ is convex. To show this we compute the second-order derivative of the function $g(\lambda)$.

$$
\begin{aligned}
\frac{\partial g(\lambda)}{\partial \lambda} &= -\frac{1}{\lambda^2}\|\mathcal{G}\phi + s\|_2^2 + v \\
\frac{\partial^2 g(\lambda)}{\partial \lambda^2} &= \frac{2}{\lambda^3}\|\mathcal{G}\phi + s\|_2^2 - \frac{1}{\lambda^2} 2(\mathcal{G}\phi + s)^T \frac{\partial(\mathcal{G}\phi)}{\partial \lambda}.
\end{aligned}
$$

Since

$$
\begin{aligned}
\frac{\partial(\mathcal{G}\phi)}{\partial \lambda} &= -\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \left(\frac{\mathbb{I}}{\lambda}\right) \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) s\right) \\
&\quad - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \frac{\mathbb{I}}{\lambda^2} s \\
&= -\frac{1}{\lambda^2} \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} (\mathcal{G}\phi + s)
\end{aligned}
$$

we have

$$
\frac{\partial^2 g(\lambda)}{\partial \lambda^2} = \frac{2}{\lambda^3}\|\mathcal{G}\phi + s\|_2^2 + \frac{1}{\lambda^4}(\mathcal{G}\phi + s)^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} (\mathcal{G}\phi + s) \geq 0
$$

for $\lambda \geq 0$ and hence the function $g(\lambda)$ is convex for $\lambda \geq \lambda(\mathcal{D})$.

- The feasible set $\lambda \geq \lambda_{\max}(\mathcal{D})$ is convex.

□

Remark 5.19. *Using some reformulations we can also show that $g(\lambda)$ and $\phi(\lambda)$ in Theorem 5.17 can be reformulated in a different, but equivalent way.*

Theorem 5.20. *Bilevel problem (5.33)*

$$
\min_{\phi}\ \Phi(\phi),\ \ \phi \in \mathbb{R}^n
$$

with

$$
\Phi(\phi) = \phi^T \mathcal{G}\phi + 2s^T\phi + \max_{\|r+y-\phi\|_2^2 \leq v} y^T \mathcal{D} y.
$$

is equivalent to the following optimization problem

$$g(\lambda) = \min_{\lambda \geq \lambda_{\max}} \left(r + \mathcal{G}^{-1}s\right)^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \left(r + \mathcal{G}^{-1}s\right) - s^T\mathcal{G}^{-1}s + \lambda v \qquad (5.36)$$

in the sense that if λ^0 *is optimal in (5.36) then the vector*

$$\phi(\lambda^0) = \mathcal{G}^{-1}s + \mathcal{G}^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} \left(r + \mathcal{G}^{-1}s\right)$$

is optimal in the bilevel problem (5.33) and it holds that $\Phi(\phi(\lambda^0)) = g(\lambda^0)$*.*

Proof. The proof is similar to the proof of Theorem 5.17, but we now consider the following terms of (5.35)

$$\begin{aligned} & -r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s + r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\mathcal{G}^{-1}s \\ & + r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)s \\ = & \, 2r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\mathcal{G}^{-1}s \end{aligned}$$

and

$$\begin{aligned} & -s^T\mathcal{G}^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s \\ = & s^T\mathcal{G}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\mathcal{G}^{-1}s - s^T\mathcal{G}^{-1}s. \end{aligned}$$

If we now insert these terms into (5.35) we get

$$\begin{aligned} \Phi(\phi) = & r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1} r + 2r^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\mathcal{G}^{-1}s \\ & + s^T\mathcal{G}^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\mathcal{G}^{-1}s - s^T\mathcal{G}^{-1}s + \lambda v \\ = & \left(r + \mathcal{G}^{-1}s\right)^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(r + \mathcal{G}^{-1}s\right) - s^T\mathcal{G}^{-1}s + \lambda v. \end{aligned}$$

With the reformulation

$$\begin{aligned} \mathcal{G}\phi + s = & \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)s\right) + \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)s \\ = & \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\left(r + \mathcal{G}^{-1}s\right) \end{aligned}$$

and with the proof of Theorem 5.17 the theorem is proven. □

5.3 Bilevel Optimization Problem with an Additional Quadratic Constraint in the Upper Level Problem

Now, let us consider another bilevel problem which we will use to describe the algorithms for solving the problems that we will present in Chapter 7.3

$$\min_{\phi^T S\phi\leq\delta}(\phi+a)^T\mathcal{G}(\phi+a)+\max_{\|r-\phi+y\|_2^2\leq v} y^T\mathcal{D}y \tag{5.37}$$

with given symmetric positive definite matrices $\mathcal{G},\ \mathcal{D},\ S\in\mathbb{R}^{n\times n}$, given vectors $r,\ a\in\mathbb{R}^n$ and given numbers $\delta,\ v\in\mathbb{R}$. With $\alpha(\delta-\phi^T S\phi)=0,\ \alpha\geq 0,\ \alpha=0$ if $\phi^T S\phi<\delta$ we can reformulate problem (5.37) as

$$\min_{\phi^T S\phi\leq\delta}(\phi+a)^T\mathcal{G}(\phi+a)-\alpha(\delta-\phi^T S\phi)+\max_{\|r-\phi+y\|_2^2\leq v} y^T\mathcal{D}y \tag{5.38}$$

The lower level problem

$$\begin{aligned}&\max_y\ y^T\mathcal{D}y\\ &\text{s.t. }\ \|r+\phi-y\|_2^2\leq v\end{aligned} \tag{5.39}$$

is the same as in bilevel problem (5.15) and hence the optimality conditions for the lower level problem remain the same (cf. Lemma 5.5). Let us consider the bilevel problem (5.38) which can be rewritten as

$$\min_{\|\phi\|_S^2\leq\delta}\ \Phi(\phi),\ \phi\in\mathbb{R}^n \tag{5.40}$$

with

$$\Phi(\phi)=\phi^T(\mathcal{G}+\alpha S)\phi+2a^T\mathcal{G}\phi+a^T\mathcal{G}a-\alpha\delta+\max_{\|r+\phi-y\|_2^2\leq v} y^T\mathcal{D}y.$$

Denote the vector $\phi(\lambda,\alpha)$ which solves the equation

$$\left(\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}\right)+(\mathcal{G}+\alpha S)^{-1}\right)(\mathcal{G}+\alpha S)\phi=r-\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a,$$

where $\lambda\geq\lambda_{\max}(\mathcal{D})$ and $\alpha\geq 0$ and denote the vector $y(\lambda,\alpha)$ which solves the equation

$$\mathcal{D}y(\lambda,\alpha)=-(\mathcal{G}+\alpha S)\phi-\mathcal{G}a.$$

For $\phi(\lambda,\alpha)$ and $y(\lambda,\alpha)$ it holds that

$$\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}\right)\mathcal{D}y=\phi-r.$$

For $\lambda\geq\lambda_{\max}(\mathcal{D})$ we have that the matrix $\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}\right)$ is positive semidefinite and for $\alpha\geq 0$ the matrix $(\mathcal{G}+\alpha S)$ is positive definite and hence the matrix $\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}+(\mathcal{G}+\alpha S)^{-1}\right)$ is positive definite for $\lambda\geq\lambda_{\max}$ and $\alpha\geq 0$. Furthermore $\phi(\lambda,\alpha)$ and $y(\lambda,\alpha)$ satisfy

$$r-\phi+y=\frac{\mathbb{I}}{\lambda}\mathcal{D}y$$

and hence

$$\|r - \phi(\lambda, \alpha) + y(\lambda, \alpha)\|_2^2 = \frac{1}{\lambda^2}\|\mathcal{D}y(\lambda, \alpha)\|_2^2.$$

Consider the case that $\alpha = 0$. Then $\phi(\lambda, 0)$ solves

$$\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)\mathcal{G}\phi(\lambda, 0) = r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a.$$

Denote $\phi^* = \phi(\lambda_{\max}, 0)$. Consider several cases.

Case 1: If $\|\phi^*\|_S^2 < \delta$ and $\|\mathcal{G}(\phi^* + a)\|_2^2 < \lambda_{\max}^2 v$ we can show similarly to (5.23) that there exists a scalar β, such that

$$y_\beta = -\mathcal{D}^{-1}\mathcal{G}(\phi^* + a) + \beta u, \;\; \beta = \beta_1 \text{ or } \beta_2,$$

where u is an eigenvector of $\mathcal{D}$ corresponding to $\lambda_{\max}$ and ϕ^* satisfies the optimality conditions in the lower level problem (cf. Lemma 5.8) and ϕ^* and y_β solves the bilevel problem (5.40), cf. Lemma 5.9.

Case 2: Assume the case $\|\phi^*\|_S^2 < \delta$ and $\|\mathcal{G}(\phi^* + a)\|_2^2 > \lambda_{\max}^2 v$, then we can formulate the following lemma

Lemma 5.21. *If $\|\mathcal{G}(\phi^* + a)\|_2^2 > \lambda_{\max}^2 v$ then there exists a unique number $\lambda^* > \lambda_{\max}$ such that $\phi(\lambda^*, 0)$ satisfies*

$$\|\mathcal{G}(\phi(\lambda^*, 0) + a)\|_2^2 = (\lambda^*)^2 v.$$

Proof. Consider the function $\eta(\lambda) = \|\mathcal{G}(\phi(\lambda, 0) + a)\|_2^2 - \lambda^2 v$. We have that $\eta(\lambda_{\max}) > 0$ and $\eta(+\infty) = -\infty$. It holds that

$$\frac{d\eta(\lambda, 0)}{d\lambda} = -\frac{2}{\lambda^2}\left((\mathcal{G}(\phi + a))^T\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\mathcal{G}(\phi + a)\right) - 2\lambda v < 0.$$

Hence, for $\alpha \equiv 0$ the function $\eta(\lambda, 0)$ is monotonically decreasing if λ is increasing and hence there exists a unique number $\lambda^* > \lambda_{\max}$ such that $\eta(\lambda^*, 0) = 0$. □

Lemma 5.22. *Consider the situation*

$$\|\phi^*\|_S^2 < \delta, \;\; \|\mathcal{G}(\phi^* + a)\|_2^2 > \lambda_{\max}^2 v.$$

Assume the vector $\phi(\lambda^*, 0)$, where $\lambda^* \geq \lambda_{\max}$, is found according to Lemma 5.21, satisfy $\|\phi(\lambda^*, 0)\|_S^2 < \delta$. Then it is optimal in the bilevel optimization problem (5.40).

Proof. Obviously, $y(\lambda^*) = -\mathcal{D}^{-1}\mathcal{G}(\phi(\lambda^*, 0) + a)$ satisfies the optimality conditions in the lower level problem for $\phi(\lambda^*, 0)$ by construction and hence

$$y^T(\lambda^*)\mathcal{D}y(\lambda^*) = \max_{\|r + y - \phi(\lambda^*, 0)\|_2^2 \leq v} y^T\mathcal{D}y.$$

Then following the lines of the proof of Lemma 5.10 we can show that

$$\Phi(\phi(\lambda^*, 0) + \Delta\phi)) - \Phi(\phi(\lambda^*, 0)) > 0$$

for all $\Delta\phi \neq 0$ and $\|\phi(\lambda^*, 0) + \Delta\phi\|_S^2 \leq \delta$. $\square$

Let us note that if $a = 0$ we can formulate the following lemma.

Lemma 5.23. *Assume $a = 0$ and consider the situation*

$$\|\phi^*\|_S^2 < \delta, \ \|\mathcal{G}\phi^*\|_2^2 > \lambda_{\max}^2 v.$$

Let the vector $\phi(\lambda^*, 0)$, where $\lambda^* \geq \lambda_{\max}$, be found according to Lemma 5.21. Then $\phi(\lambda^*, 0)$ satisfies $\|\phi(\lambda^*, 0)\|_S^2 < \delta$ and hence it is optimal in problem (5.40).

Proof. If $a = 0$ it holds that

$$\|\phi(\lambda^*, 0)\|_S^2 < \|\phi^*\|_S^2 < \delta$$

because

$$\frac{\partial \|\phi(\lambda^*, 0)\|_S^2}{\partial \lambda} = -\frac{2}{\lambda^2}\phi^T \underbrace{S\mathcal{G}^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1}\right)^{-1}\mathcal{G}}_{\text{similar to positive definite matrix}}\phi < 0.$$

Hence, if λ is increasing, then $\|\phi(\lambda^*, 0)\|_S^2$ is decreasing and with Lemma 5.22, this lemma is proven. $\square$

If $\|\phi(\lambda^*, 0)\|_2^2 \geq \delta$, where $\lambda^* \geq \lambda_{\max}$ is found according to Lemma 5.21, then the construction of the optimal $\phi(\lambda^*, \alpha^*)$ should be done following Case 4.

Case 3: Assume now the case $\|\phi^*\|_S^2 > \delta$ and $\|\mathcal{G}(\phi^* + a)\| < \lambda_{\max}^2 v$. First of all we find a scalar α^* such that

$$\|\phi(\lambda_{\max}, \alpha^*)\|_S^2 = \delta$$

following the lemma

Lemma 5.24. *If $\|\phi(\lambda_{\max}, 0)\|_S^2 \geq \delta$ then there exists a unique number $\alpha^* > 0$ such that*

$$\|\phi(\lambda_{\max}, \alpha^*)\|_S^2 = \delta.$$

Proof. Consider the function $\rho(\alpha) = \|\phi(\lambda, \alpha)\|_S^2 - \delta$ for fixed $\lambda \geq \lambda_{\max}$. We have that $\rho(0) \geq 0$ and $\rho(+\infty) = -\delta < 0$. It holds that

$$\begin{aligned}\frac{d\rho(\alpha)}{d\alpha} &= -2(S\phi)^T \underbrace{\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\left(\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right) + (\mathcal{G} + \alpha S)^{-1}\right)^{-1}(\mathcal{G} + \alpha S)^{-1}}_{\text{similar to positive definite matrix}}(S\phi) \\ &< 0.\end{aligned}$$

Hence, the function $\rho(\alpha)$ is monotonically decreasing for $\lambda \equiv \lambda_{\max}$ and there exists a unique number $\alpha^* > 0$ such that $\rho(\alpha^*) = 0$. $\square$

Lemma 5.25. *Consider the case $\|\phi^*\|_S^2 \geq \delta$ and let α^* be found according to Lemma 5.24. If $\|(\mathcal{G}+\alpha^* S)\phi(\lambda_{\max},\alpha^*)+\mathcal{G}a\|_2^2 < \lambda_{\max}^2 v$ then there exists a scalar $\beta \in \mathbb{R}$ such that*

$$y_\beta = -\mathcal{D}^{-1}((\mathcal{G}+\alpha^* S)(\phi(\lambda_{\max},\alpha^*)+a)) + \beta u$$

solves the lower level problem for $\phi(\lambda_{\max},\alpha^*)$ (5.39) and $\phi(\lambda_{\max},\alpha^*)$ solves the upper level problem (5.40). Here, as before, u denotes an eigenvector corresponding to the maximal eigenvalue $\lambda_{\max}$ of $\mathcal{D}$.

Proof. It holds that

$$\begin{aligned} y_\beta &= -\mathcal{D}^{-1}\left((\mathcal{G}+\alpha^* S)\phi(\lambda_{\max},\alpha^*)+\mathcal{G}a\right)+\beta u \\ &= -\mathcal{D}^{-1}\left((\mathcal{G}+\alpha^* S)\bar{\phi}-\mathcal{G}a\right)+\beta u \end{aligned}$$

where $\bar{\phi} = \phi(\lambda_{\max},\alpha^*)$. Then it holds that

$$\mathcal{D}y_\beta = -(\mathcal{G}+\alpha^* S)\bar{\phi} - \mathcal{G}a + \lambda_{\max}\beta u.$$

For $\bar{\phi}$ we have

$$\begin{aligned} (-\mathcal{D}+\lambda_{\max}\mathbb{I})y_\beta &= (-\mathcal{D}+\lambda_{\max}\mathbb{I})(-\mathcal{D}^{-1}((\mathcal{G}+\alpha^* S)\bar{\phi}+\mathcal{G}a)+\beta\underbrace{(-\mathcal{D}+\lambda_{\max}\mathbb{I})u}_{=0} \\ &= (\mathbb{I}-\lambda_{\max}\mathcal{D}^{-1})((\mathcal{G}+\alpha^* S)\bar{\phi}+\mathcal{G}a) = -\lambda_{\max}(r-\bar{\phi}) \end{aligned}$$

and it holds that

$$\begin{aligned} r-\bar{\phi}+y_\beta &= -\frac{1}{\lambda_{\max}}(-\mathcal{D}+\lambda_{\max}\mathbb{I})y_\beta + y_\beta \\ &= \frac{\mathcal{D}}{\lambda_{\max}}y_\beta = -\frac{(\mathcal{G}+\alpha^* S)\bar{\phi}+\mathcal{G}a}{\lambda_{\max}} + \beta u. \end{aligned}$$

Since $\|(\mathcal{G}+\alpha^* S)\bar{\phi}+\mathcal{G}a\|_2^2 < \lambda_{\max}^2 v$ there exists a $\beta \neq 0$ that satisfies

$$\|r-\bar{\phi}+y_\beta\|_2^2 = \frac{1}{\lambda_{\max}}\|(\mathcal{G}+\alpha^* S)\bar{\phi}+\mathcal{G}a-\lambda_{\max}\beta u\|_2^2 = v.$$

Thus, y_β and $\bar{\phi}$ satisfy

$$\begin{aligned} &1)\ (-\mathcal{D}+\lambda_{\max}\mathbb{I})y_\beta = -\lambda_{\max}(r-\bar{\phi}) \\ &2)\ \|r-\bar{\phi}+y_\beta\|_2^2 = v \end{aligned}$$

yielding

$$\max_{\|r-\bar{\phi}+y_\beta\|_2^2 \leq v} y^T\mathcal{D}y = y_\beta\mathcal{D}y_\beta$$

and

$$\Phi(\bar{\phi}) = \bar{\phi}^{*T}(\mathcal{G}+\alpha^* S)\bar{\phi} + 2a^T\mathcal{G}\bar{\phi} + a^T\mathcal{G}a - \alpha\delta + y_\beta^T\mathcal{D}y_\beta.$$

Furthermore it holds that $\|\bar{\phi}\|_S^2 = \delta$. Consider $\tilde{\phi} = \bar{\phi} + \Delta\phi$ with $\Delta\phi \neq 0$ such that $\|\tilde{\phi}\|_S^2 \leq \delta$ and $\tilde{y} = y_\beta + \Delta y$ with $\Delta y = \Delta\phi$. Then

$$r - \tilde{\phi} + \tilde{y} = r - (\bar{\phi} + \Delta\phi) + (y_\beta + \Delta y) = r - \bar{\phi} + y_\beta.$$

Hence, $\|r - \tilde{\phi} + \tilde{y}\|_2^2 = v$ and

$$\max_{\|r-\tilde{\phi}+y\|_2^2} y^T \mathcal{D} y \geq \tilde{y}\mathcal{D}\tilde{y}.$$

Similarly to the proof of Lemma 5.8 we have

$$\begin{aligned}
\Phi(\tilde{\phi}) - \Phi(\bar{\phi}) &= \tilde{\phi}^T(\mathcal{G} + \alpha^* S)\tilde{\phi} + 2a^T\mathcal{G}\tilde{\phi} + \max_{\|r+y-\tilde{\phi}\|_2^2 \leq v} y^T\mathcal{D}y - \bar{\phi}^T(\mathcal{G} + \alpha^*)\bar{\phi} - 2a\mathcal{G}\bar{\phi} - y_\beta \mathcal{D} y_\beta \\
&\geq (\bar{\phi} + \Delta\phi)^T(\mathcal{G} + \alpha^* S)(\bar{\phi} + \Delta\phi) + 2a^T\mathcal{G}(\bar{\phi} + \Delta\phi) - \bar{\phi}^T(\mathcal{G} + \alpha^*)\bar{\phi} \\
&\quad - 2a\mathcal{G}\bar{\phi} + \tilde{y}^T\mathcal{D}\tilde{y} - y_\beta^T \mathcal{D} y_\beta \\
&= 2\Delta\phi^T(\mathcal{G} + \alpha^* S)\bar{\phi} + \Delta\phi^T(\mathcal{G} + \alpha^* S)\delta\phi + 2a^T\mathcal{G}\Delta\phi \\
&\quad + \delta y^T \mathcal{D}\Delta y + 2 \underbrace{\Delta y}_{\Delta\phi} \quad \underbrace{\mathcal{D} y_\beta}_{-((\mathcal{G}+\alpha^* S)\bar{\phi}+\mathcal{G}a)+\lambda_{\max}\beta u} \\
&= \Delta\phi^T(\mathcal{G} + \alpha^* S)\Delta\phi + \Delta y^T\mathcal{D}\Delta y + \lambda_{\max}\beta\Delta\phi^T u
\end{aligned}$$

By a proper choice of β we can assure that $\beta\Delta\phi^T u > 0$. Therefore we get, that

$$\Phi(\tilde{\phi}) - \Phi(\bar{\phi}) \geq \Delta\phi^T(\mathcal{G} + \alpha^* S)\Delta\phi + \Delta y^T\mathcal{D}\Delta y > 0$$

for all $\tilde{\phi} \neq \bar{\phi}$. Hence, $\bar{\phi}$ is optimal. □

Lemma 5.26. *Assume $a = 0$ and consider the case $\|\phi^*\|_S^2 \geq \delta$, $\|\mathcal{G}\phi^*\|_2^2 < \lambda_{\max}^2 v$. Let α^* be found according to Lemma 5.24. Then $\phi(\lambda_{\max}, \alpha^*)$ satisfy $\|(\mathcal{G} + \alpha^* S)\phi(\lambda_{\max}, \alpha^*)\|_2^2 < \lambda_{\max}^2 v$ and hence it is optimal in (5.40).*

Proof. Since

$$\frac{\partial\|(\mathcal{G} + \alpha S)\phi(\lambda_{\max}, \alpha)\|}{\partial\alpha} < 0, \ \alpha \geq 0$$

it holds that

$$\|(\mathcal{G} + \alpha^* S)\phi(\lambda_{\max}, \alpha)\|_2^2 < \|\mathcal{G}\phi^*\|_2^2 < \lambda_{\max}^2 v.$$

□

In the case $\|(\mathcal{G} + \alpha S)(\phi(\lambda_{\max}, \alpha^*) + \mathcal{G}a\|_2^2 > \lambda_{\max}^2 v$, where $\alpha^* \geq 0$ is found following Lemma 5.24, then the construction of $\phi(\lambda^*, \alpha^*)$ is performed as in Case 4.
Case 4: Consider the situations

$$\|\phi(\lambda_{\max}, 0)\|_S^2 \geq \delta, \ \|\mathcal{G}(\phi(\lambda_{\max}, 0) + a = \|_2^2 \geq \lambda_{\max}^2 v$$

or

$$\|\mathcal{G}(\phi(\lambda^*, 0) + a)\|_2^2 = (\lambda^*)^2 v, \ \|\phi(\lambda^*, 0)\|_S^2 \geq \delta, \ \lambda^* > \lambda_{\max}$$

or

$$\|\phi(\lambda_{\max}, \alpha^*\|_S^2 = \delta,\ \|(\mathcal{G} + \alpha^* S)(\phi(\lambda_{\max}, \alpha^*) + \mathcal{G}a)\|_2^2 > \lambda_{\max}^2 v,\ \alpha^* > 0.$$

in these situations we search for $\phi^*,\ \alpha^* \geq 0,\ \lambda^* \geq 0$ such that the following system of nonlinear equalities hold:

$$\mathcal{F}(x) = \begin{pmatrix} \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)((\mathcal{G} + \alpha S)\phi + \mathcal{G}a) + \phi - r \\ \|\mathcal{G} + \alpha S)\phi + \mathcal{G}a\|_2^2 - \lambda^2 v \\ \|\phi\|_S^2 - \delta \end{pmatrix} = 0,\ x \in \begin{pmatrix} \phi \\ \alpha \\ \lambda \end{pmatrix}. \tag{5.41}$$

The following lemma gives the conditions where the system (5.41) has a solution.

Lemma 5.27. *Assume that the functions $\mathcal{F} : \mathcal{D} \to \mathbb{R}^{n+2}$ and $\mathcal{J} = \frac{\partial \mathcal{F}}{\partial x} : \mathcal{D} \to \mathbb{R}^{(n+2)\times(n+2)}$ satisfy:*

1. *$\mathcal{J}(\cdot)$ is invertible for all $x \in \mathcal{D}$.*
2. *$\|J^{-1}(y)(J(x + t(y - x)) - J(x))(y - x)\| \leq \omega t\|y - x\|^2$ with $\omega < \infty$ for all $t \in [0, 1]$ and $x, y \in \mathcal{D}$ with $x - y = \mathcal{J}^{-1}(y)\mathcal{F}(x)$.*
3. *Assume that the initial guess $x_0 \in \mathcal{D}$ exists such that $\delta_0 = \frac{\|\mathcal{J}(x_0)^{-1}\mathcal{F}(x_0)\|\omega}{2} < 1$ and $\mathcal{D}_0 := \bar{B}\left(x_0, \frac{\mathcal{J}(x_0)^{-1}\mathcal{F}(x_0)}{1-\delta_0}\right) \subset \mathcal{D}$.*

Denote $\mathcal{D} = \{(\phi, \alpha, \lambda),\ \phi \in \mathbb{R}^n,\ \alpha \in \mathbb{R},\ \lambda \in \mathbb{R},\ 0 \leq \alpha \leq \bar{\alpha},\ 0 < \underline{\lambda} \leq \lambda \leq \bar{\lambda}\}$, where $\bar{\alpha} < \infty$ and $\underline{\lambda} = \lambda_{\max}$, $\bar{\lambda} < \infty$ are some numbers. Then there exist unique $\{\phi^*,\ \alpha^*,\ \lambda^*\} \in \mathcal{D}_0$ such that

$$\mathcal{F}(\phi^*, \alpha^*, \lambda^*) = 0.$$

Proof. The lemma follows form the local contration theorem in Bock [14]. □

Lemma 5.28. *The vector ϕ^* constructed by Lemma 5.27 solves the bilevel programming problem (5.40).*

Proof. Let us introduce the following notations

$$\begin{aligned} \phi(\lambda, \alpha) &= (\mathcal{G} + \alpha S)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + (\mathcal{G} + \alpha S)^{-1}\right)^{-1}\left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a\right) \\ \bar{\phi} &= \phi(\lambda^*, \alpha^*) \\ y(\lambda, \alpha) &= -\mathcal{D}^{-1}((\mathcal{G} + \alpha S)\phi(\lambda, \alpha) + \mathcal{G}a) \\ \bar{y} &= -\mathcal{D}^{-1}\left((\mathcal{G} + \alpha^* S)\bar{\phi} + \mathcal{G}a\right). \end{aligned} \tag{5.42}$$

Since $\phi(\lambda, \alpha)$ satisfies

$$\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + (\mathcal{G} + \alpha S)^{-1}\right)(\mathcal{G} + \alpha S)\phi(\lambda, \alpha) = r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a$$

we have

$$\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)(\mathcal{G} + \alpha S)\phi(\lambda, \alpha) + \phi(\lambda, \alpha) = r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a$$

and hence

$$\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)((\mathcal{G} + \alpha S)\phi(\lambda, \alpha) + \mathcal{G}a) = r - \phi(\lambda, \alpha)$$

Then for $\bar{\phi}$ and $\bar{y}$ we have

$$\begin{aligned}
(-\mathcal{D} + \lambda^*\mathbb{I})\bar{y} &= (-\mathcal{D} + \lambda^*\mathbb{I})(-\mathcal{D}^{-1})\left((\mathcal{G} + \alpha^* S)\bar{\phi} + \mathcal{G}a\right) \\
&= (\mathbb{I} - \lambda^*\mathcal{D}^{-1})\left((\mathcal{G} + \alpha^* S)\bar{\phi} + \mathcal{G}a\right) \\
&= \lambda^*(\frac{\mathbb{I}}{\lambda^*} - \mathcal{D}^{-1})\left((\mathcal{G} + \alpha^* S)\bar{\phi} + \mathcal{G}a\right) \\
&= -\lambda^*(-\bar{\phi} + r)
\end{aligned}$$

and

$$\begin{aligned}
r - \bar{\phi} + \bar{y} &= -\frac{1}{\lambda^*}(-\mathcal{D} + \lambda^*\mathbb{I})\bar{y} + \bar{y} = \frac{\mathcal{D}}{\lambda^*}\bar{y} \\
&= -\frac{1}{\lambda^*}\left((\mathcal{G} + \alpha^* S)\bar{\phi} + \mathcal{G}a\right).
\end{aligned}$$

We have shown that $\bar{\phi}$ and $\bar{y}$ satisfy the optimality conditions in the lower level problem

$$\begin{aligned}
&1)\ (-\mathcal{D} + \lambda^*\mathbb{I})\bar{y} = -\lambda^*(r - \bar{\phi}) \\
&2)\ \|r - \bar{\phi} + \bar{y}\|_2^2 = v.
\end{aligned}$$

Hence, $\bar{y}$ solves the lower level problem for $\bar{\phi}$ and $\Phi(\bar{\phi}) = \bar{\phi}^T(\mathcal{G} + \alpha^* S)\bar{\phi} + 2a^T\mathcal{G}\bar{\phi} + a^T\mathcal{G}a - \alpha\delta + \bar{y}^T\mathcal{D}\bar{y}$. Furthermore it holds that $\|\bar{\phi}\|_S^2 = \delta$. Following the proof of Lemma 5.25 we consider $\tilde{\phi} = \bar{\phi} + \Delta\phi$, $\tilde{y} = y + \Delta y$, $\Delta y = \Delta\phi$ with $\|\tilde{\phi}\|_S^2 \leq \delta$ and can show that

$$\Phi(\tilde{\phi}) - \Phi(\bar{\phi}) > 0$$

for $\tilde{\phi} \neq \bar{\phi}$. □

Theorem 5.29. *The function*

$$\begin{aligned}
g(\lambda, \alpha) = &\left(r - \left(D^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a\right)^T \left(D^{-1} - \frac{\mathbb{I}}{\lambda} + (\mathcal{G} + \alpha S)^{-1}\right)^{-1} \left(r - \left(D^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a\right) \\
&- (\mathcal{G}a)^T\left(D^{-1} - \frac{\mathbb{I}}{\lambda}\right)(\mathcal{G}a) + 2r^T(\mathcal{G}a) + \lambda v - \alpha\delta + a^T\mathcal{G}a
\end{aligned}$$

where $\mathcal{D}^{-1}$, $\mathcal{G}$ and $\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + (\mathcal{G} + \alpha\mathbb{I})^{-1}\right)$ are positive definite for $\alpha \geq 0$ and $\lambda \geq \lambda_{\max}(\mathcal{D})$ is convex in $\lambda \geq 0$ for arbitrary fixed α and concave in $\alpha \geq 0$ for arbitrary fixed λ.

Proof. To show the convexity in $\lambda \geq \lambda_{\max}$ for arbitrary α we show that

$$\frac{\partial^2 g(\lambda, \alpha)}{\partial \lambda^2} \geq 0$$

and to show the concavity in $\alpha \geq 0$ for arbitrary λ we show that

$$\frac{\partial^2 g(\lambda,\alpha)}{\partial \alpha^2} \leq 0.$$

For λ we have

$$\begin{aligned}
\frac{\partial g(\lambda,\alpha)}{\partial \lambda} &= -\frac{1}{\lambda^2}\|(\mathcal{G}+\alpha S)\phi(\lambda,\alpha)+\mathcal{G}a\|_2^2 + v \\
\frac{\partial^2 g(\lambda,\alpha)}{\partial \lambda^2} &= \frac{2}{\lambda^3}\|(\mathcal{G}+\alpha S)\phi(\lambda,\alpha)+\mathcal{G}a\|_2^2 + \frac{1}{\lambda^4}\left((\mathcal{G}+\alpha S)\phi(\lambda,\alpha)+\mathcal{G}a\right)^T \\
&\quad \left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}+\mathcal{G}^{-1}\right)^{-1}\left((\mathcal{G}+\alpha S)\phi(\lambda,\alpha)+\mathcal{G}a\right) \geq 0 \text{ for } \lambda \geq 0
\end{aligned}$$

and for α we have

$$\begin{aligned}
\frac{\partial g(\lambda,\alpha)}{\partial \alpha} &= \|\phi(\lambda,\alpha)\|_S^2 - \delta \\
\frac{\partial^2 g(\lambda,\alpha)}{\partial \alpha^2} &= 2(S\phi(\lambda,\alpha))^T\frac{\partial \phi(\lambda,\alpha)}{\partial \alpha}.
\end{aligned} \tag{5.43}$$

Compute

$$\begin{aligned}
\frac{\partial \phi(\lambda,\alpha)}{\partial \alpha} &= -(\mathcal{G}+\alpha S)^{-1}S(\mathcal{G}+\alpha S)^{-1}\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}+(\mathcal{G}+\alpha S)^{-1}\right)^{-1}\left(r-\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a\right) \\
&\quad + (\mathcal{G}+\alpha S)^{-1}(-1)\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}+\mathcal{G}+\alpha S)^{-1}\right)^{-1}\left(-(\mathcal{G}+\alpha S)^{-1}S(\mathcal{G}+\alpha S)^{-1}\right) \\
&\quad \left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}+(\mathcal{G}+\alpha S)^{-1}\right)\left(r-\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a\right) \\
&= -(\mathcal{G}+\alpha S)^{-1}S\phi(\lambda,\alpha) + (\mathcal{G}+\alpha S)^{-1}\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}+(\mathcal{G}+\alpha S)^{-1}\right)^{-1}(\mathcal{G}+\alpha S)^{-1}S\phi(\lambda,\alpha) \\
&= (\mathcal{G}+\alpha S)^{-1}\left(-\mathbb{I}+\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}+(\mathcal{G}+\alpha S)^{-1}\right)^{-1}(\mathcal{G}+\alpha S)^{-1}\right)S\phi(\lambda,\alpha) \\
&= -(\mathcal{G}+\alpha S)^{-1}\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}+(\mathcal{G}+\alpha S)^{-1}\right)\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}\right)S\phi(\lambda,\alpha).
\end{aligned}$$

If we now insert this into (5.43) we get

$$\frac{\partial^2 g(\lambda,\alpha)}{\partial \alpha^2} = -2(S\phi(\lambda,\alpha))^T\underbrace{(\mathcal{G}+\alpha S)^{-1}\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}+(\mathcal{G}+\alpha S)^{-1}\right)\left(\mathcal{D}^{-1}-\frac{\mathbb{I}}{\lambda}\right)}_{\text{similar to positive definite matrix}}S\phi(\lambda,\alpha) < 0$$

Hence, the function $g(\lambda,\alpha)$ is convex in λ and concave in α. □

Let us analyze the min-max problem

$$\min_{\lambda \geq \lambda_{\max}} \max_{\alpha \geq 0} g(\lambda,\alpha)$$

From Theorem 5.29 we have that $g(\lambda,\alpha)$ is convex in $\lambda \geq \lambda_{\max}$ for arbitrary but fixed α and concave in $\alpha \geq 0$ for arbitrary but fixed λ. The solution $\{\lambda^*,\alpha^*\}$ is a saddle point, which means that

$$g(\lambda^*, \alpha) \leq g(\lambda^*, \alpha^*) \leq g(\lambda, \alpha^*)$$

for all $\lambda \geq \lambda_{\max}$ and for all $\alpha \geq 0$. We define the curves

$$\hat{\lambda}(\alpha) = \arg \min_{\lambda \geq \lambda_{\max}} g(\lambda, \alpha), \;\; \hat{\alpha}(\lambda) = \arg \max_{\alpha \geq 0} g(\lambda, \alpha).$$

Saddle points $\{\lambda^*, \alpha^*\}$ correspond to points where these two curves meet (cf. Bertsekas [10])

(1) Consider $\min\limits_{\lambda \geq \lambda_{\max}} g(\lambda, \alpha)$ for arbitrary fixed α. The optimality conditions read as

$$\begin{aligned} \mathcal{L}_1(\lambda, \nu) &= g(\lambda, \alpha) - \nu(\lambda - \lambda_{\max}) \\ \frac{\partial \mathcal{L}_1(\lambda, \nu)}{\partial \lambda} &= -\frac{1}{\lambda^2} \|(g + \alpha S)\phi + \mathcal{G}a\|_2^2 + v - \nu \overset{!}{=} 0 \end{aligned}$$

with the complimentary conditions

$$\nu(\lambda - \lambda_{\max}) = 0, \;\; \nu \geq 0,$$

where ϕ satisfies (5.35). Analyzing the optimality conditions $\hat{\lambda} = \hat{\lambda}(\alpha)$ satisfies

$$\begin{aligned} &\text{if } \hat{\lambda} > \lambda_{\max} \text{ then } \nu = 0 \text{ and } \|(\mathcal{G} + \alpha S)\phi(\hat{\lambda}, \alpha) + \mathcal{G}a\|_2^2 = \hat{\lambda}^2 v. \\ &\text{if } \hat{\lambda} = \lambda_{\max} \text{ then } \nu \geq 0 \text{ and } \|(\mathcal{G} + \alpha S)\phi(\hat{\lambda}, \alpha) + \mathcal{G}a\|_2^2 \leq \hat{\lambda}^2 v. \end{aligned}$$

(2) Consider $\{\max\limits_{\alpha \geq 0} g(\lambda, \alpha)\}$ for arbitrary fixed λ which is equivalent to $\{-\min\limits_{\alpha \geq 0} - g(\lambda, \alpha)\}$. The optimality conditions read as

$$\begin{aligned} \mathcal{L}_2(\alpha, \mu) &= -g(\lambda, \alpha) - \mu\alpha \\ \frac{\partial \mathcal{L}_2(\alpha, \nu)}{\partial \alpha} &= -\|\phi(\lambda, \alpha)\|_2^2 + \delta - \mu \overset{!}{=} 0 \end{aligned}$$

with the complimentary conditions

$$\mu\alpha = 0, \;\; \mu \geq 0.$$

Analyzing the optimality conditions $\hat{\alpha} = \hat{\alpha}(\lambda)$ satisfies

$$\begin{aligned} &\text{If } \hat{\alpha} > 0 \Rightarrow \mu = 0 \text{ and } \|\phi(\lambda, \hat{\alpha})\|_S^2 = \delta. \\ &\text{If } \hat{\alpha} = 0 \Rightarrow \mu \geq 0 \text{ and } \|\phi(\lambda, 0)\|_S^2 \leq \delta. \end{aligned}$$

Summarizing, we get that the saddle points $\{\lambda^*, \alpha^*\}$ satisfies the following:

1) If $\lambda^* > \lambda_{\max}, \; \alpha^* > 0$, then $\{\lambda^*, \alpha^*\}$ solve the nonlinear system

$$\begin{aligned} \|(\mathcal{G} + \alpha S)\phi(\lambda, \alpha) + \mathcal{G}a\|_2^2 &= \lambda^2 v \\ \|\phi(\lambda, \alpha)\|_S^2 &= \delta. \end{aligned}$$

2) If $\alpha^* = 0$ then $\|\phi(\lambda^*, \alpha^*)\|_S^2 \leq \delta$.

3) If $\lambda^* = \lambda_{\max}$ then $\|(\mathcal{G} + \alpha S)\phi(\lambda_{\max}, \alpha^*) + \mathcal{G}a\|_2^2 \leq \lambda_{\max}^2 v$.

Using this analysis we can formulate the following theorem.

Theorem 5.30. *Bilevel problem (5.38) is equivalent to the following problem*

$$\begin{aligned}\min_{\lambda}\max_{\alpha}\ g(\lambda, \alpha) = & \left(r - \left(D^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a\right)^T \left(D^{-1} - \frac{\mathbb{I}}{\lambda} + (\mathcal{G} + \alpha S)^{-1}\right)^{-1} \left(r - \left(D^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a\right) \\ & - (\mathcal{G}a)^T \left(D^{-1} - \frac{\mathbb{I}}{\lambda}\right)(\mathcal{G}a) + 2r^T(\mathcal{G}a) + \lambda v - \alpha\delta + a^T\mathcal{G}a \\ & s.t.\ \ \alpha \geq 0, \lambda \geq \lambda_{\max}\end{aligned} \tag{5.44}$$

in the sense that if $\{\lambda^*, \alpha^*\}$ solve the problem (5.44) then $\phi(\lambda^*, \alpha^*)$ solves the problem (5.38) and it holds that $g(\lambda^*, \alpha^*) = \Phi(\phi(\lambda^*, \alpha^*))$.

Altogether we can formulate the following algorithm for solving (5.40).

Algorithm 5.31. (Solution of problem (5.40))

`COMPUTE` $\lambda_{\max} = \lambda_{\max}(\mathcal{D})$

`COMPUTE` vector $\phi^* = (\mathcal{G} + \alpha S)^{-1} \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + (\mathcal{G} + \alpha S)^{-1}\right)^{-1} \left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a\right)$

  `IF` Case 1) $\|\phi^*\|_S^2 < \delta$ and $\|\mathcal{G}(\phi^* + a)\|_2^2 < \lambda_{\max}^2 v$
    `THEN` solution of problem (5.40) is given by $\phi^0 = \phi^*$
  `IF` Case 2) $\|\phi^*\|_S^2 < \delta$ and $\|\mathcal{G}(\phi^* + a)\|_2^2 > \lambda_{\max}^2 v$
    `THEN` solve equation $\|\mathcal{G}(\phi(\lambda, 0) + a)\|_2^2 = \lambda^2 v$ for $\lambda \geq \lambda_{\max}$,
      where $\phi(\lambda, 0) = \mathcal{G}^{-1}(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + \mathcal{G}^{-1})^{-1}(r - (\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda})\mathcal{G}a)$
      (e.g. by Newton's method)
      solution: λ^*
      `IF` $\|\phi(\lambda^*, 0)\|_S^2 < \delta$
        `THEN` solution of problem (5.40) as $\phi^0 = \phi^*(\lambda^*, 0)$
      `ELSE` go to Case 4)
  `IF` Case 3) $\|\phi^*\|_S^2 > \delta$ and $\|\mathcal{G}(\phi^* + a)\|_2^2 < \lambda_{\max}^2 v$
    `THEN` solve equation $\|\phi(\lambda_{\max}, \alpha)\|_S^2 = \delta$ for $\alpha \geq 0$,
      where
      $\phi(\lambda_{\max}, \alpha) = (\mathcal{G}+\alpha S)^{-1} \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda_{\max}} + (\mathcal{G} + \alpha S)^{-1}\right)^{-1} \left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda_{\max}}\right)\mathcal{G}a\right)$
      (e.g. by Newton's method)
      solution: α^*
      `IF` $\|\mathcal{G}(\phi(\lambda_{\max}, \alpha^*) + a)\|_2^2 < \lambda_{\max}^2 v$
        `THEN` solution of problem (5.40) as $\phi^0 = \phi^*(\lambda_{\max}, \alpha^*)$
      `ELSE` go to Case 4)
  `ELSE` Case 4)
    `THEN` solve equations
      $\|(\mathcal{G} + \alpha S)\phi(\lambda, \alpha) + \mathcal{G}a\|_2^2 - \lambda^2 v = 0$
      $\|\phi(\lambda, \alpha)\|_S^2 - \delta = 0$
      $\phi(\lambda, \alpha) - (\mathcal{G} + \alpha S)^{-1} \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + (\mathcal{G} + \alpha S)^{-1}\right)^{-1} \left(r - \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda}\right)\mathcal{G}a\right) = 0$
      for $\lambda \geq \lambda_{\max},\ \alpha \geq 0$ (e.g. by Newton's method)
      solution: λ^*, α^*
      `WRITE` solution of problem (5.40) as $\phi^0 = \phi^*(\lambda^*, \alpha^*)$.

6 Closed-Loop Min Max Feedback Controls

Mathematical models of real processes almost always contain uncertainties or disturbances. This is the reason why we will consider optimal control problems with uncertainties or robust optimal control in this chapter. There are several possibilities of sources of uncertainties. One source of uncertainties is for instance when observational errors arise during the measurements. During the modeling process we can obtain errors because of certain modeling assumptions and simplifications, unmodeled effects or even errors in the model. Errors might also arise during the discretization procedures. To consider those uncertainties we will need robust optimization for optimal control under uncertainties, such that the obtained solution of the optimization problem will be a guaranteed good solution for all possible realizations of uncertainties that could occur in the model. For a robust solution we will include state feedback. We assume a predictive optimal control problem on the control interval $[\tau, t]$ and a corresponding optimal control $u^0(t;\tau,x),\ t \in [\tau, t_*]$ depending on the current time τ of the system state x. To obtain a robust solution there is a need of state feedback and this feedback is constructed by the following implicit law

$$u^*(\tau, x) = u^0(\tau;\tau,x),\ \tau \in [0, t_*],\ x \in \mathbb{R}^n. \tag{6.1}$$

As we already stated in the introduction, we can distinguish between three different approaches for robust feedback (6.1): Open-loop optimal control, Open-loop min-max optimal control and Closed-loop min-max optimal feedback control. The third approach with one correction moment is topic of the following section. To present a practical algorithm in Section 6.2 we will reformulate the problem as a bilevel programming problem and discuss an algorithm for its solution. Closed-loop min-max optimal feedback control with several correction points will be discussed in Section 6.3. Section 6.4 presents approximative control policies and a comparison of them. In the last section we will shortly describe a numerical example with which we can compare the different approximative policies.

6.1 Closed Loop Min-Max Optimal Feedback Control Problem with One Correction Point

In this section we will consider a linear-quadratic optimal control problem with uncertainties. We will first formulate linear-quadratic optimal control problems without and with uncertainties and we will present problems that can arise if we ignore the uncertainties. We will then modify the problem formulation by including a feedback aspect, to obtain a guaranteed optimal control and avoid infeasibility problems.

We examine the following optimal control problems. First we consider the optimal control of the nominal system without uncertainties

$$\min_u \; J(u) = \int_T u^2(t)dt \tag{6.2}$$

$$\begin{aligned} \text{s.t. } \dot{x} &= Ax(t) + bu(t) \\ x(0) &= x_0 \\ x(t_*) &= 0 \end{aligned} \tag{6.3}$$

The second optimal control problem takes into account uncertainties in the model dynamics

$$\min \; J(u) = \int_T u^2(t)dt \tag{6.4a}$$

$$\text{s.t. } \dot{z} = Az(t) + bu(t) + gw(t), \; z(0) = z_0 \tag{6.4b}$$

$$\|z(t_*; u(\cdot), w(\cdot))\|_2^2 \le \delta_0^2, \text{ for all } w(\cdot) \in W. \tag{6.4c}$$

In these two problems we have $t \in T = [0, t_*]$, $x(t) \in \mathbb{R}^n$ and $z(t) \in \mathbb{R}^n$ the states of the systems, respectively, $u(t) \in \mathbb{R}$ the control, $w(t) \in \mathbb{R}$ an unknown perturbation and the matrix $A \in \mathbb{R}^{n \times n}$ and the vectors $b \in \mathbb{R}^n, g \in \mathbb{R}^n$ and $x_0 \in \mathbb{R}^n$ are given. The control $u(t)$, the unknown perturbation $w(t)$ and the time $t \in T$ are assumed to belong to $L_2(T)$. Here, $z(t; u(\cdot), w(\cdot))$ with $t \in T$ in (6.4c) is a trajectory of system (6.4b) which is generated by the control $u(\cdot)$ and the perturbation $w(\cdot)$.
We assume that the unknown perturbation $w(\cdot)$ is in the class of all admissible disturbances W, which is defined by:

$$W := \left\{ w(\cdot) = (w(t), t \in T) \in L_2(T) : \int_0^{t_1} w^2(t)dt \le v_1, \int_{t_1}^{t_*} w^2(t)dt \le v_2, \text{for some } t_1 \in T \right\}.$$

Therefore we use uncertainties, that have to satisfy the integral quadratic constraint, which was introduced in the work of V.A. Yakubovich in 1988. The problem of stabilizing a system with uncertain parameters is of interest in the control theory. The methods that treat this kind of problems usually differ in the concept of stability and the form of the uncertainty. A description of the uncertainty sets can be found in the works of Savkin & Petersen [72] and Lee & Zhenghong [55]. In the time-invariant case, there is the ellipsoidal set, which we use in this thesis, and also the axis-aligned polyhedron. The ellipsoidal set has been used by many researchers (eg. Goodwin *et al.* [36], Kosut *et al.* [48], Lau *et al.* [54]) as the parameter estimation under the Gaussian noise assumption naturally yields ellipsoidal bounds. In the case of the axis-aligned polyhedron the Euclidean norm is replaced by the ∞-norm and was also used by many reasearchers(e.g. Campo & Morari [21], Zheng & Morari [86], Genceli and Nikolaou [34]).
The class of bounded disturbances

$$|w(t)| \le \alpha, \; t \in T_1 = [t_0, t_1], \; t \in T_2 = [t_1, t_2], \; t_2 = t_*$$

belongs to the class of admissible disturbances with the following choice of the numbers v_i

$$v_i = \alpha^2(t_i - t_{i-1}),\ i = 1, 2. \tag{6.5}$$

Lemma 6.1. *We consider the two sets*

$$\mathcal{P} = \{p \in \mathbb{R}^n : p = d + \int_{\tau_1}^{\tau_2} s(t)w(t)dt; \int_{\tau_1}^{\tau_2} w^2(t)dt \leq s_0\},$$
$$\bar{\mathcal{P}} = \{p \in \mathbb{R}^n : (p-d)^T S^{-1}(p-d) \leq s_0\}$$

with $s(t),\ t \in [\tau_1, \tau_2]$ a given function, $d \in \mathbb{R}^n$ a given vector and s_0 a given number such that the matrix

$$S := \int_{\tau_1}^{\tau_2} s(t)s^T(t)dt$$

is nonsingular. Then the relation $\mathcal{P} = \bar{\mathcal{P}}$ holds.

Proof. (cf. Kostyukova & Kostina [47]).
We assume that $\bar{p}$ lies on the boundary of set $\mathcal{P}$. Hence,

$$\bar{p} - d = \int_{\tau_1}^{\tau_2} \bar{w}(t)s(t)dt,\ \int_{\tau_1}^{\tau_2} \bar{w}^2(t) = s_0.$$

To show that $\mathcal{P} = \bar{\mathcal{P}}$ we will use a proof by contradiction and therefore we assume that $\bar{p} \notin \bar{\mathcal{P}}$. That means

$$(\bar{p} - d)^T S^{-1}(\bar{p} - d) > s_0.$$

A new perturbation is then constructed by the rules $\tilde{w}(t) = \alpha(\bar{p}-d)^T S^{-1}s(t),\ t \in [\tau_1, \tau_2]$ with $\alpha^2 = \frac{s_0}{\|\bar{p}-d\|^2_{S^{-1}}}$. Hence, the new vector $\tilde{p}$ is given by

$$\tilde{p} := d + \int_{\tau_1}^{\tau_2} \tilde{w}(t)q(t)dt$$

with $q(t) = F(t_*, t)g$ and $F(t, \tau)$ the fundamental matrix of $\dot{x} = Ax$. We use the notations $\bar{\Delta} = \bar{p} - d$ and $\tilde{\Delta} = \tilde{p} - d$ and calculate

$$\begin{aligned} J :&= \|\tilde{\Delta}\|^2_{S^{-1}} - \|\bar{\Delta}\|^2_{S^{-1}} \\ &= \|\tilde{\Delta} - \bar{\Delta}\|^2_{S^{-1}} + 2\bar{\Delta}^T S^{-1}(\tilde{\Delta} - \bar{\Delta}) \\ &= \|\tilde{\Delta} - \bar{\Delta}\|^2_{S^{-1}} + 2\int_{\tau_1}^{\tau_2} \frac{\tilde{w}(t)(\tilde{w}(t) - \bar{w}(t))}{\alpha}. \end{aligned} \tag{6.6}$$

We know that it holds that $\int_{\tau_1}^{\tau_2} \tilde{w}^2(t) = \int_{\tau_1}^{\tau_2} \bar{w}^2(t)dt$, therefore

$$2\int_{\tau_1}^{\tau_2} \tilde{w}(t)(\bar{w}(t) - \tilde{w}(t))dt + \int_{\tau_1}^{\tau_2} (\bar{w}(t) - \tilde{w}(t))^2 dt = 0. \tag{6.7}$$

If we now multiply the upper equation (6.7) with $-\frac{1}{\alpha}$ and add it to (6.6) we obtain

$$J = \|\tilde{\Delta} - \bar{\Delta}\|^2_{S^{-1}} - \frac{1}{\alpha}\int_{\tau_1}^{\tau_2} (\bar{w}(t) - \tilde{w}(t))^2 dt \geq 0$$

with $\alpha < 0$. Hence, $\|\tilde{\Delta}\|^2_{S^{-1}} \geq \|\bar{\Delta}\|^2_{S^{-1}} > s_0$. But it also holds $\tilde{\Delta} = \alpha\bar{\Delta}$ and therefore $\|\tilde{\Delta}\|^2_{S^{-1}} = \alpha^2\|\bar{\Delta}\|^2_{S^{-1}} = s_0$. This means we have a contradiction and it has to be $\bar{p} \in \bar{\mathcal{P}}$.

We still need to show, that $\bar{\mathcal{P}} \subset \mathcal{P}$. We consider $\bar{p} \in \bar{\mathcal{P}}$ and that means it holds that $(\bar{p} - y)^T S^{-1}(\bar{p} - y) \leq s_0$. The new perturbation is now constructed by the rules $\bar{w}(t) = (\bar{p} - y)^T S^{-1} s(t),\ t \in [\tau_1, \tau_2]$. This perturbation satisfies the relations

$$\int_{\tau_1}^{\tau_2} \bar{w}^2(t)dt = \|\bar{p} - y\|^2_{S^{-1}} \leq s_0$$

$$\bar{p} - y - \int_{\tau_1}^{\tau_2} \bar{w}(t)s(t)dt = 0.$$

This means we can deduce $\bar{p} \in \mathcal{P}$ and therefore $\bar{\mathcal{P}} \subset \mathcal{P}$. □

To ensure that our system 6.4 is controllable, we assume that the following relation holds:

$$\text{rank}(b, Ab, \ldots, A^n b) = n. \tag{6.8}$$

Altogether, we can now formulate the stated problem as follows:

OLOCP
Find a control $u(\cdot) = (u(t),\ t \in T)$ that minimizes the quadratic cost functional (6.4a) such that for any perturbation $w(\cdot)$ from the class of admissible disturbances W a trajectory $x(t)$ of the system (6.4b) lies in a δ_0-neighborhood of zero at the final moment $t = t_*$, which is formulated in the inequality (6.4c).

This problem is an open-loop min-max optimal control problem (cf. Section 3.5.2) as we consider the uncertainty but do not have the possibility to correct the control. The open-loop problem can graphically be described as in Figure 6.1. The tube shows the trajectories corresponding to a fixed control $\bar{u}(\cdot)$ that drives the system state to 0 at t_* and a particular realization of the uncertainty. The constraints have to be satisfied by every trajectory $z(\cdot) = z(\cdot; \bar{u}, w)$ in the tube $\|z(t_*; \bar{u}(\cdot), w)\|^2 \leq \delta_2,\ \forall w \in W$(cf. Rawlings

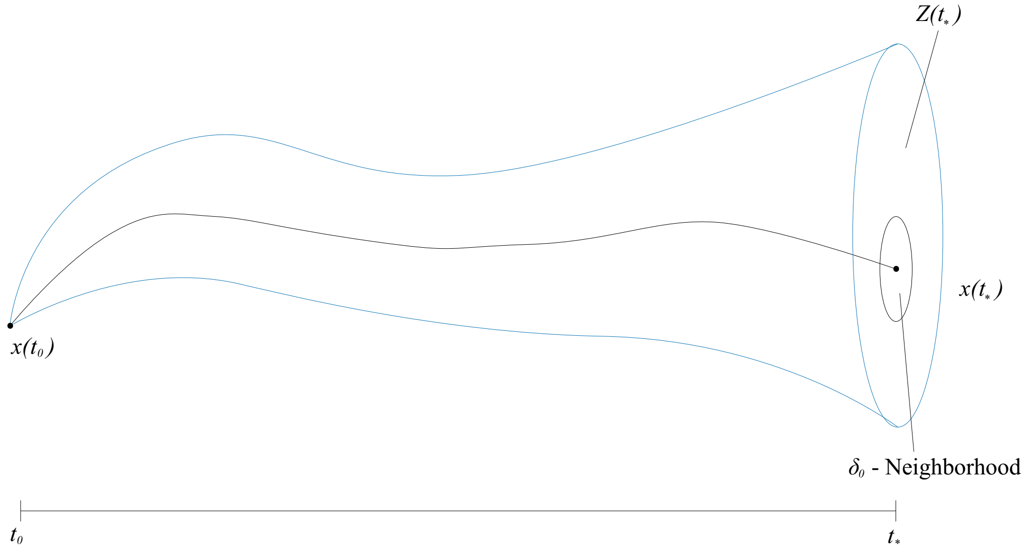


FIGURE 6.1: Trajectories in Open Loop Min Max Optimal Control Problem for some $\bar{u}(\cdot)$

[70]). The ellipsoidal shows the δ_0-neighborhood of zero at the final moment $t = t_*$ inside the set $Z(t_*) = \{z(t_*; \bar{u}(\cdot), w), z(t_0) = x_0, w \in W\}$.

Remark 6.2. *For a more general case, we will later in this thesis use a different formulation for constraint (6.4c) as we will not consider the trajectory to be in the* δ_0*-neighborhood of zero but in the* δ_**-neighborhood of* x_**. We will then use the formulation*

$$\|z(t_*; u(\cdot), w(\cdot)) - x_*\|_2^2 \leq \delta_*^2 \textit{ for all } w(\cdot) \in W.$$

As mentioned in Section 3.5.2 with this problem formulation we could have the problem of not finding a feasible control. To avoid the problem of not finding a control that drives the system in the δ^2-neighborhood of zero at the time moment $t = t_*$, we will modify the problem formulation and include a feedback aspect at one time moment. To show the necessity of the feedback we will use Lemma 2 from Kostyukova & Kostina [47] that shows the problem of not finding a feasible control. For this purpose we consider the following two min-max problems

$$\begin{aligned} \gamma(\tau) := & \min_{u(t), t\in[\tau,t_*], z_*} \; \max_{w(t), t\in[\tau,t_*]} \|z(t_*)\|^2 \\ \text{s.t. } & \dot{z} = Az + bu + gw, \; t \in [\tau, t_*] \\ & z(\tau) = z_* \\ & \int_\tau^{t_*} w^2(t)dt \leq v_2(\tau) \end{aligned} \tag{6.9}$$

and

$$\begin{aligned} \gamma := & \min_{u(\cdot), z_*} \max_{w(\cdot)} \|z(t_*)\|^2 \\ \text{s.t. } & \dot{z} = Az + bu + gw,\ t \in [0, t_*] \\ & z(0) = z_0 \\ & w(\cdot) \in W \end{aligned} \tag{6.10}$$

with the given moment $\tau \in [t_1, t_*]$ and the given value $v_2(\tau) \geq 0$.

Lemma 6.3. *The following relations hold for the values of the cost functionals of the problems (6.9) and (6.10)*

$$\gamma(\tau) = \mu(\tau) v_2(\tau) \tag{6.11a}$$

$$\gamma \geq \bar{\mu} v_1 + \mu(t_1) v_2. \tag{6.11b}$$

Here, $\mu(\tau)$ and $\bar{\mu}$ are the maximal eigenvalues of the matrices

$$\begin{aligned} Q(\tau) &= \int_{\tau}^{t_*} q(t) q^T(t) dt,\ \tau \in [t_1, t_*] \\ \textit{and } \bar{Q} &= \int_0^{t_1} q(t) q^T(t) dt, \end{aligned} \tag{6.12}$$

respectively, with $q(t) = F(t_*, t)g$ and $F(t, \tau)$ the state transition matrix of the system $\dot{x} = Ax$.

The proof of Lemma 6.3 can be found in Kostyukova & Kostina [47].

The lemma shows, that without a feedback aspect it could occur that $\bar{\mu} v_1 + \mu(t_1) v_2 > \delta_0^2$. In this case a control that drives the trajectory of the system (6.4) into the δ_0-neighborhood of zero at the final time moment $t = t_*$ for any perturbation $w(\cdot) \in W$ does not exist. As we already stated, we will now modify the problem formulation by including a feedback aspect to avoid this problem. By including feedback we avoid the infeasibility problem on the one hand but on the other hand the problem may become more complex and unpractical to solve. For this feedback we divide our interval in two smaller intervals $T_1 = [0, t_1]$ and $T_2 = [t_1, t*]$ and choose a new control $\bar{u}(\cdot)$ at the second interval. Therefore at a given time point $t_1 \in\ ]0, t_*]$ we measure the current state and update the control for the second interval $[t_1, t_*]$, considering the new information. Thus, we construct a control like in the closed-loop approach (cf. Section 3.5.2) with one correction moment. This process is graphically shown in Figure 6.2. As in Figure 6.1 we have the tube that shows the trajectories corresponding to a particular realization of the uncertainty. The constraints have to be satisfied by every trajectory in the tube and the red ellipsoidal at the time moment t_1 shows the reachability set $Z(t_1) = \{z(t_1; \bar{u}_1(\cdot), w_1), z(t_0) = x_0; w(\cdot) \in W\}$ where we measure the current state and choose a new control. The black ellipsoidal shows the δ_0-neighborhood around zero in the end point t_* and the red ellipsoidal at t_* shows the set $Z_1(t_*) = \{z(t_*; \bar{u}_2(\cdot), w(\cdot)),\ z(t_1) = x_1 \in Z(t_1),\ w(\cdot) \in W\}$ both inside the set $Z(t_*)$.

It can be shown that in this modified formulation there always exists a control that

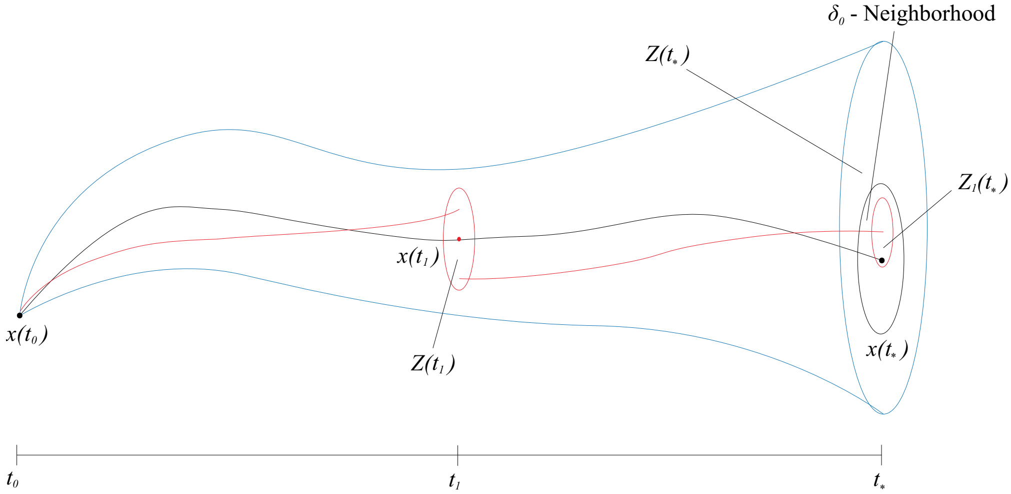


FIGURE 6.2: Trajectories of Open-Loop Control problem with one correction moment for some $\bar{u}_1(\cdot) = \bar{u}(t),\ t \in [t_0, t_1]$ and $\bar{u}_2(\cdot) = \bar{u}(t),\ t \in [t_1, t_*]$

guarantees to drive the trajectory of the system (6.4) in the δ_0-neighborhood of zero at the final moment $t = t_*$ for any perturbation $w(\cdot) \in W$ if for the time moment $t_1 \in]0, t_*]$ the inequality

$$\mu(t_1)v_2 \leq \delta_0^2 \tag{6.13}$$

holds.

The idea:
We have two time intervals $T_1 = [0, t_1]$ and $T_2 = [t_1, t_*]$. We choose an arbitrary control for the first interval and we assume that with this control the trajectory of the system (6.4) is driven to the state $x(t_1) = \xi$ at the time moment $t = t_1$ under the perturbation $w(t) \in T_1$. At this time point we choose a new control for the interval T_2 such that the trajectory of the nominal unperturbed system

$$\dot{x} = Ax + bu,\ t \in T_2,\ x(t_1) = \xi \tag{6.14}$$

takes the zero value at the final time point $t = t_*$.

We know that such a control exists because of the fact that the system is assumed to be controllable. With Lemma 6.3 it follows that the trajectory of the actual perturbed system

$$\dot{z} = Az + bu + gw,\ t \in T_2,\ z(t_1) = \xi$$

appears in the δ_0-neighborhood of zero at the final time moment $t = t_*$ for any perturbation $w(t) \in W,\ t \in T_2$.

This means that we can always find a control that guarantees to drive the trajectory of the system in the δ_0-neighborhood of zero at the final moment $t = t_*$ with this modified approach. This control is called *feasible guaranteed program control*. Without loss of generality we assume that

$$\mu(t_1)v_2 = \delta_0. \tag{6.15}$$

This is possible for the right choice of δ_0. We will now take the cost functional (6.4a) into account. The aim is to find a feasible guaranteed (worst-case) program control which delivers the minimal value to the cost functional (6.4a). The control is constructed by the just described rules above. To solve this problem we start by considering the second interval $T_2 = [t_1, t_*]$ and we assume that at time moment t_1 the state is $x(t_1) = \xi$. We now know with equality (6.15) and Lemma 6.3 that any feasible guaranteed program control drives the trajectory of the unperturbed system from $x(t_1) = \xi$ to $x(t_*) = 0$. To find the best feasible guaranteed program control for the interval T_2 we have to solve the following program

$$\begin{aligned} &\min_u \int_{T_2} u^2(t)dt \\ &\text{s.t. } \dot{x} = Ax + bu \\ &\qquad x(t_1) = \xi, x(t_*) = 0. \end{aligned} \tag{6.16}$$

As it is shown in Pontryagin *et al.* [68] we know that problem (6.16) has a solution. We denote by $(u(\cdot;\xi) = u(t;\xi), t \in T_2)$ an optimal control of problem (6.16) and

$$\int_{T_2} u^2(t;\xi)dt$$

the corresponding optimal value of the cost functional. Now, we consider the first interval $T_1 = [0, t_1]$ and choose the control $u_1(\cdot) = (u_1(t), t \in T_1)$. In the case of the nominal system

$$\dot{x} = Ax + bu_1, \; t \in T_1, \; x(0) = x_0 \tag{6.17}$$

the state at time t_1 takes the value

$$y = y(x_0, u_1) = F(t_1, 0)x_0 + \int_{T_1} f_1(t)u_1(t)dt, \; f_1(t) = F(t_1, t)b.$$

In the case of the actual perturbed system

$$\dot{z} = Az + bu_1 + gw_1, \; t \in T_1, \; z(0) = z_0 \tag{6.18}$$

where $w_1(\cdot) \in W_1(0)$, the state at time t_1 can take any value from the ellipsoid

$$\xi = \xi(z_0, u_1, w_1) \in \left\{ p \in \mathbb{R}^n : p = y(z_0, u_1) + \int_{T_1} q_1(t)w_1(t)dt, \; w_1(\cdot) \in W_1(0) \right\}, \tag{6.19}$$

$$W_1(\tau) = \left\{ w(t) \in L_2([\tau, t_1]) : \int_{\tau}^{t_1} w^2(t)dt \leq v_1(\tau) \right\} \tag{6.20}$$

with $q_1(t) = F(t_1, t)g$ and $v_1(\tau),\ \tau \in T_1$ a given function with $v_1(0) = v_1$. With Lemma 6.1 we can rewrite relation (6.19) as

$$\xi = \xi(z_0, u_1, w_1) \in \{p \in \mathbb{R}^n : (p-y)^T Q^{-1}(p-y) \le v_1\},\ Q = Q_1(0)$$

with

$$Q_1(\tau) = \int_{\tau}^{t_1} q_1(t) q_1^T(t) dt,\ \tau \in T_1. \tag{6.21}$$

This means, that we can write the best guaranteed value of the cost functional (6.4a) for any chosen control $u_1(\cdot)$ as

$$\int_{T_1} u_1^2(t)dt + \max_{\xi \in \mathbb{R}^n, \|x(t_1)-\xi\|^2_{Q^{-1}} \le v_1} \int_{T_2} u^2(t;\xi)dt \tag{6.22}$$

where $x(t),\ t \in T_1$ is a trajectory of the unperturbed system (6.17) and $u(t;\xi),\ t \in T_2$ is an optimal control of problem (6.16).
With this cost functional we can formulate the deterministic problem to find the best control $u_1(\cdot)$ to the corresponding cost functional (6.22):

$$\begin{aligned} &\min_{u_1(\cdot), y} \left(\int_{T_1} u_1^2(t)dt + \max_{\xi} \int_{T_2} u^2(t;\xi)dt \right) \\ &\text{s.t. } \|y - \xi\|^2_{Q^{-1}} \le v_1 \\ &\qquad \dot{x} = Ax + bu_1,\ t \in [0, t_1],\ x(0) = x_0,\ x(t_1) = y \\ &\qquad u(\cdot;\xi) = \operatorname{argmin} \int_{T_2} u^2(t)dt \\ &\qquad \dot{x} = Ax + bu,\ t \in [t_1, t_*],\ x(t_1) = \xi,\ x(t_*) = 0 \end{aligned} \tag{6.23}$$

This is the problem of finding a control with one correction moment only. The more general case is to use a problem formulation, in which we want to find a control law, by using more than one correction moment at fixed intermediate time points $t_i \in T,\ i = 1, \dots, m,\ m > 1$. We will consider this in Section 6.3.

6.2 Analysis of problem (6.23) and the Corresponding Bilevel Optimization Problem

In this section we will analyze problem (6.23), which is a bilevel optimization problem. The lower level problem of (6.23) is the problem (6.16). The solution of the optimal control problem (6.16) can be written as

$$u(t;\xi) = -f^T(t)\mathcal{G}_*^{-1}F(t_*, t_1)\xi,\ t \in T_2 \tag{6.24}$$

with $f(t) = F(t_*, t)b,\ \mathcal{G}_* = \mathcal{G}_*(t_1)$ where the matrix $\mathcal{G}_*(\tau)$ is given by

$$\mathcal{G}_*(\tau) = \int_{\tau}^{t_*} f(t) f^T(t) dt. \tag{6.25}$$

The optimal value of the cost functional of (6.16) is equal to

$$\int_{T_2} u^2(t;\xi) dt = \xi^T \mathcal{D} \xi \tag{6.26}$$

where

$$\mathcal{D} := F^T(t_*, t_1) \mathcal{G}_*^{-1} F(t_*, t_1). \tag{6.27}$$

Remark 6.4. *The matrix*

$$Q_1(t_1, t_2) = \int_{t_1}^{t_2} F(t_2, s) g (F(t_2, s) g)^T ds$$

where t_1 is free and t_2 is fixed can directly be computed by

$$\begin{aligned} \dot{p}(t) &= -A p(t),\ p(t_2) = g,\ p \in \mathbb{R}^n \\ \dot{Q}(t) &= p(t) p^T(t),\ Q(t_2) = 0,\ Q \in \mathbb{R}^{n \times n} \end{aligned}$$

with $t \in [t_1, t_2]$ and

$$Q_1(t_1, t_2) = Q(t_1).$$

It holds that

$$Q_0(t_1) := \int_{t_1}^{t_*} F(t_*, s) g (F(t_*, s) g)^T ds = Q_1(t_1, t_*).$$

The matrix

$$\mathcal{D} := F^T(t_*, t_1) \mathcal{G}_*^{-1} F(t_*, t_1)$$

with

$$\mathcal{G}_*(t_1) := \int_{t_1}^{t_*} F(t_*, s) b (F(t_*, s) b)^T ds$$

can directly be computed by

$$\dot{\mathcal{D}} = -A^T \mathcal{D}(t) - \mathcal{D}^T(t) A - \mathcal{D}(t) b (\mathcal{D}(t) b)^T.$$

We can show that problem (6.23) can be rewritten as the following bilevel program

$$\begin{aligned} &\min_{\phi}(\phi^T \mathcal{G}\phi + \max_{\xi} \xi^T \mathcal{D}\xi) \\ &\text{s.t. } \|\xi - r - \mathcal{G}\phi\|^2_{Q^{-1}} \le v_1 \end{aligned} \tag{6.28}$$

with $r \in \mathbb{R}^n$, $\phi, \xi \in \mathbb{R}^n$ unknown given vectors, $\mathcal{G}$ and $\mathcal{D}$ given positive definite matrices and v_1 a given number. In order to show this, remind that for a fixed $y \in \mathbb{R}^n$ the $u_1(\cdot)$, $t \in [0, t_1]$ is a solution of the problem

$$\begin{aligned} &\min_{u_1} \int_{T_1} u_1^2(t)dt \\ &\text{s.t. } \dot{x} = Ax + bu_1, \ t \in [0, t_1] \ x(0) = x_0, \ x(t_1) = y. \end{aligned} \tag{6.29}$$

The optimal control for problem (6.29) can be written

$$u_1(t; y) = \phi^T f_1(t), \ t \in T_1; \ f_1 = F(t_1, t)b \tag{6.30}$$

where ϕ satisfies the following system of linear equations

$$y = F(t_1, 0)x_0 + \mathcal{G}\phi \tag{6.31}$$

and $f_1(t) = F(t_1, t)b$, $\mathcal{G} = \mathcal{G}(0)$. The matrix $\mathcal{G}(\tau)$ is given by

$$\mathcal{G}(\tau) = \int_{\tau}^{t_1} f_1(t) f_1^T(t)dt. \tag{6.32}$$

With the assumption (6.8) the matrix $\mathcal{G}$ is nonsingular. We obtain the following value of the cost functional of problem (6.29)

$$\int_{T_1} u_1^2(t; y)dt = \phi^T \mathcal{G}\phi.$$

With these statements we can rewrite (6.23) as in problem (6.28). Therefore we choose the control for the first interval for problem (6.23) as

$$u_1^0(t) = \left(\phi^0\right)^T f_1(t), \ t \in T_1 \tag{6.33}$$

where $\phi^0 \in \mathbb{R}^n$ is the solution of problem (6.28). For the second interval we choose the control

$$u(t; \xi^*) = -f^T(t)\mathcal{G}_*^{-1} F(t_*, t_1)\xi^*, \ t \in T_2 \tag{6.34}$$

with $\xi^* = x(t_1; u_1^0(\cdot), w_1^*(\cdot))$ the state of the system (6.18) at the time moment $t = t_1$ under the control $u_1(\cdot)$ and the actual perturbation $w_1(\cdot) \in W_1(0)$. Thus the behavior of the dynamic process (6.4) with this constructed guaranteed program control can be formulated as follows

$$\dot{z} = \begin{cases} Az + bu_1^0(t) + gw(t), & t \in T_1 \\ Az + bu(t; x(t_1)) + gw(t), & t \in T_2 \end{cases} \quad z(0) = z_0. \tag{6.35}$$

Thus in problem (6.28)

$$Q(\tau) = \int_{\tau}^{t_*} F(t_*, t)g(F(t_*, t)g)^T dt,\ Q = Q(t_1)$$

$$\mathcal{G} = \mathcal{G}(0) \text{ with } \mathcal{G}(\tau) = \int_{\tau}^{t_1} f_1(t)f_1^T(t)dt,$$

$$\mathcal{D} := F^T(t_*, t_1)\mathcal{G}_*^{-1}F(t_*, t_1)$$

$$\mathcal{G}_* = \mathcal{G}_*(t_1) \text{ with } \mathcal{G}_*(\tau) = \int_{\tau}^{t_*} f(t)f^T(t)dt,$$

are positive definite matrices, $r = F(t_1, 0)x_0$ and v_1 is defined as in (6.5). Using the variable transformation $\psi = \mathcal{M}^{-T}\phi$ where $\mathcal{M}$ is defined by $Q^{-1} = \mathcal{M}^T\mathcal{M}$ we can rewrite problem (6.28) as

$$\min_{\psi}(\psi^T G\psi + \max_{z} z^T Dz), \text{ s.t. } \|z - d - G\psi\|^2 \leq v_1 \tag{6.36}$$

where $G = \mathcal{M}\mathcal{G}\mathcal{M}^T,\ D = \mathcal{M}^{-1^T}\mathcal{D}\mathcal{M}^{-1},\ d = \mathcal{M}r$ and $Q^{-1} = \mathcal{M}^T\mathcal{M}$. This means we now have the same problem as the problem (5.2) in Chapter 5.

Hence, we can directly formulate the following theorem.

Theorem 6.5. *Problem (6.36) is equivalent to the following optimization problem*

$$\min g(\lambda) := d^T \left(D^{-1} + G - \frac{\mathbb{I}}{\lambda} \right)^{-1} d + \lambda v_1,\ s.t.\ \lambda \geq \lambda_{\max}, \tag{6.37}$$

where $\lambda^0_{\max}$ *is the maximal eigenvalue of* D*, in the sense that if* λ^0 *is optimal in (6.37), then the vector* $\psi(\lambda^0) = \left(\frac{I}{\lambda^0} - (D^{-1} + G)\right)^{-1} d$ *solves problem (6.36).*

Altogether, following the first part of Chapter 5 we can formulate the following algorithm for solving problem (6.36):

Algorithm 6.6.

`COMPUTE` vector $\psi^* = \left(\frac{\mathbb{I}}{\lambda_{max}} - D^{-1} - G\right)^{-1} d$
- `IF` Case 1) $\|\psi^*\|^2 \leq \lambda^2_{\max} v_1$
  - `THEN` solution of problem (6.36) is given by $\psi^0 = \psi^*$
- `ELSE` Case 2) $\|\psi^*\|^2 > \lambda^2_{\max} v_1$
  - `THEN` solve equation $\|\psi(\lambda)\|^2 = \lambda^2 v$ for $\lambda \geq \lambda_{\max}$ (e.g. by Newton's method)
    - solution: λ^*
    - `WRITE` solution of problem (6.36) as $\psi^0 = \psi(\lambda^*)$

To obtain a solution of problem (6.28) we use the notations

$$\phi^0 = \mathcal{M}^T\psi^0,\ \xi^* = \mathcal{M}^{-1}p^0$$

where $p^0 = -D^{-1}\psi^*$.

6.2.1 Robust Optimal Feedback

In the last section we discussed a problem reformulation where we include a possibility to correct a control in one point. In this subsection we will discuss the algorithm for the robust optimal feedback using the solution of bilevel optimization problem. We will start with constructing the feedback at the first interval, then we will consider the second interval. We have to consider both intervals separately since the rules of constructing a robust feedback are different at the first and second interval. For constructing the feedback we will use a family of problems

$$\begin{aligned} \min_{u_1(\cdot),y} & \left(\int_{\tau}^{t_1} u^2(t)dt + \max_{\xi} \int_{T_2} u^2(t;\xi)dt \right) \\ \text{s.t. } & \|y-\xi\|^2_{Q_1^{-1}(\tau)} \le v_1(\tau), \\ & \dot{x} = Ax + bu_1,\ x(\tau) = p, x(t_1) = y, \\ & u(\cdot;\xi) = \arg\min \int_{T_2} u^2(t)dt, \\ & \dot{x} = Ax + bu,\ x(t_1) = \xi,\ x(t_*) = 0 \end{aligned} \tag{6.38}$$

with $v_1(\tau) = v_1 - \int_0^{\tau} (w^*(t))^2 dt,\ w^*(t),\ t \in [0,\tau]$ a realized perturbation till a current moment τ. We can rewrite the problem (6.38) analogously to the previous section as

$$\begin{aligned} \min_{\phi}\ & \phi^T \mathcal{G}(\tau)\phi + \max_{\xi} \xi^T D\xi, \\ \text{s.t. } & \|\xi - r(\tau,p) - \mathcal{G}(\tau)\phi\|^2_{Q_1^{-1}(\tau)} \le v_1(\tau) \end{aligned} \tag{6.39}$$

with $r(\tau,p) = F(t_1,\tau)p$, the positive definite matrices D and

$$Q_1(\tau) := \int_{\tau}^{t_1} F(t_1,t)g(F(t_1,t)g)^T$$

and the positive definite matrix $\mathcal{G}(\tau)$. Now we assume $\phi_1^0(\tau,p) \in \mathbb{R}^n$ to be a solution of problem (6.39) $\tau \in [0,t_1],\ p \in \mathbb{R}^n$. We can then construct the control law $u_1^*(t,x),\ t \in T_1$ at the first interval as

$$u_1^*(t,x) = \phi_1^0(t,x)^T f_1(t),\ t \in T_1,\ x \in \mathbb{R}^n. \tag{6.40}$$

We can describe the dynamic process with control law as

$$\dot{x}(t) = Ax(t) + bu_1^*(t,x(t)) + gw(t),\ t \in T_1,\ x(0) = x_0. \tag{6.41}$$

To construct the function

$$\phi_1^0(\tau,p),\ \tau \in T_1,\ p \in \mathbb{R}^n \tag{6.42}$$

we use the described rules from the previous section. With $x^*(\tau)$ we denote a real system state at the current time τ. We are able to construct function (6.42) online only along the realized trajectory $x^*(t),\ t \ge 0$ of system (6.41) generated by the realized perturbation

$w^*(t),\ t \geq 0$. To construct the function

$$\phi^*(\tau) = \phi_1^0(\tau, x^*(\tau)),\ \tau \in T_1 \tag{6.43}$$

we will use the following rules which can be implemented online.
We assume that for the current position $(\tau, x) = (\tau_0, x^*(t_0))$ we know the vector $\phi^*(\tau_0)$ which is constructed as in the previous section. We also assume that the inequality

$$\|\phi^*(\tau_0)\|^2_{Q_1(\tau_0)} < \lambda^2_{\max}(\tau_0)v_1(\tau_0)$$

holds with $\lambda_{\max}(\tau)$ the maximal eigenvalue of the matrix $D(\tau) = \mathcal{M}^{-T}(\tau)\mathcal{D}(\tau)\mathcal{M}^{-1}(\tau)$ where $Q_1^{-1}(\tau) = \mathcal{M}^T(\tau)\mathcal{M}(\tau)$.
For $\tau \in (\tau_0, \tau_1]$ we compute the vector $\phi^*(\tau)$ from

$$\left(\frac{Q_1(\tau)}{\lambda_{\max}(\tau)} - \left(\mathcal{D}^{-1} + \mathcal{G}(\tau)\right)\right)\phi^*(\tau) = F(t_1, \tau)x^*(\tau) \tag{6.44}$$

until the moment τ_1 for which $\tau_1 = t_1$ or $\tau_0 < \tau_1 < t_1$ and $\|\phi^*(\tau_1)\|^2_{Q_1(\tau_1)} = \lambda^2_{\max}(\tau_1)v(\tau_1)$. The matrix functions $\mathcal{G},\ Q_1(\tau),\ F(t_1, \tau)$ and the scalar function $\lambda_{\max}(\tau) \in \mathbb{R},\ \tau \in T_1$ depend only on the elements $A,\ b$ and g of the initial problem and can therefore be computed previously offline.
We now consider the two cases. In case one it holds that $\tau_1 = t_1$ and hence the construction of the feedback at the first interval is finished.
In the second case it holds that $\tau_1 < t_1$. Then we compute the vector $\phi^*(\tau)$ and the number $\lambda(\tau)$ uniquely from

$$\left(\frac{Q_1(\tau)}{\lambda(\tau)} - \left(\mathcal{D}^{-1} + \mathcal{G}(\tau)\right)\right)\phi^*(\tau) = F(t_1, \tau)x^*(\tau), \tag{6.45}$$
$$\|\phi^*(\tau)\|^2_{Q_1(\tau)} = \lambda(\tau)^2 v_1(\tau),\ \lambda(\tau) > \lambda_{\max}(\tau)$$

for $\tau \in (\tau_1, \tau_2]$ with τ_2 such that $\tau_2 = t_1$ or $\tau_1 < \tau_2 < t_1$ and $\lambda(\tau_2) = \lambda_{\max}(\tau_2)$. Again we have two cases. In case one it holds that $\tau_2 = t_1$ and the construction of the feedback at the first interval is finished. In case two it holds that $\tau_1 < \tau_2 < t_1$; then we start the procedure again and compute the vector $\phi^*(\tau)$ for $\tau > \tau_2$ from system (6.44).

Remark 6.7. *1. For the construction of the functions $Q_1(\tau),\ \lambda_{\max}(\tau) \in \mathbb{R},\ \tau \in T_1$ we can use approximations, for example*

$$\bar{Q}_1(\tau) = Q_1(\theta_i),\ \bar{\lambda}_{\max}(\tau) = \lambda_{\max}(\theta_i),\ \tau \in [\theta_i, \theta_{i+1}),\ i = 0, \ldots, N-1,$$

with $N \geq 1$ an integer and $\theta_i,\ i = 0, \ldots, N$ any numbers that satisfy the relations

$$0 = \theta_0 < \theta_1 < \cdots < \theta_{N-1} < \theta_N = t_1.$$

But if we use an approximation, the quality of the control can become worse, because it holds that

$$\{\xi \in \mathbb{R}^n : \|\xi - r(\tau, p) - \mathcal{G}\phi\|^2_{Q_1^{-1}(\tau)} \leq v_1(\tau)\}$$
$$\subset \{\xi \in \mathbb{R}^n : \|\xi - r(\tau, p) - \mathcal{G}\phi\|^2_{\bar{Q}_1^{-1}(\tau)} \leq v_1(\tau)\}$$

2. *It also holds*

$$\mathcal{G}(\tau+\Delta\tau) = \mathcal{G}(\tau) - \int_{\tau}^{\tau+\Delta\tau} f_1(t) f_1^T(t) dt. \qquad (6.46)$$

Hence, we can compute the function $\mathcal{G}(\tau)$ recursively online using path following methods with respect to the parameter $\tau \in [0, t_1]$.

3. *To compute a feedback control $u_1^*(\tau, x)$ at the first interval for any current position (τ, x) we only need to solve system (6.44) or system (6.45).*

After constructing the feedback for the first interval, we will now consider the second interval $T_2 = [t_1, t_*]$. We assume, that at the moment $\tau \in [t_1, t_*)$ the state of the system (6.4) is equal to $x(\tau) = p$. We choose a control $u_2(t)$ at the interval $[\tau, t_*]$. Then at the final moment $t = t_*$ the unperturbed system

$$\dot{x} = Ax + bu_2,\ t \in [\tau, t_*],\ x(\tau) = p$$

appears in the state

$$y(\tau, p, u_2) = F(t_*, \tau)p + \int_{\tau}^{t_*} f(t) u_2(t) dt$$

and the perturbed system

$$\dot{x} = Ax + bu_2 + gw,\ t \in [\tau, t_*],\ x(\tau) = p$$

can appear in any state of the form

$$z = y(\tau, p, u_2) + \int_{\tau}^{t_*} q(t) w(t) dt,\ w(\cdot) \in W_*(\tau)$$

with

$$W_*(\tau) = \{w(t),\ t \in [\tau, t_*] : \int_{\tau}^{t_*} w^2(t) dt \le v_2(\tau) := v_2 - \int_{t_1}^{\tau} (w^*(t))^2 dt\},$$

$w^*(t),\ t \in [t_1, \tau]$ a realized perturbation till the current moment τ. We can deduce from Lemma 6.3 that

$$\max_{w(\cdot) \in W_*(\tau)} \| \int_{\tau}^{t_*} q(t) w(t) dt \|^2 = \mu(\tau) v_2(\tau).$$

Hence, $\|z - y(\tau, p, u_2)\|^2 \le \mu(\tau) v_2(\tau)$, and the control $u_2(t),\ t \in [\tau, t_*]$ is feasible guaranteed if

$$\|y(\tau, p, u_2)\|^2 \le \bar{\delta}^2(\tau) := \delta_0^2 - \mu(\tau) v_2(\tau).$$

We consider a family of problems

$$\min_{u(t),t\in[\tau,t_*]} \int_{\tau}^{t_*} u^2(t)dt \tag{6.47}$$
$$\text{s.t. } \dot{x} = Ax + bu,\ x(\tau) = p,\ \|x(t_*)\|^2 \leq \bar{\delta}^2(\tau).$$

Based on the Maximum Principle (cf. Pontryagin *et al.* [68]) we can formulate the following proposition:

Proposition 6.8. *An optimal control of the problem (6.47) has the form*

$$u(t;\tau,p) = \phi_2^T(\tau,p)f(t),\ t \in [\tau,t_*]$$

with $\phi_2(\tau,p) \in \mathbb{R}^n$ *constructed by the rules*

1. $\phi_2(\tau,p) = 0$ *if* $\|F(t_*,\tau)p\|^2 \leq \bar{\delta}^2(\tau)$
2. $\phi = \phi_2(\tau,p) \in \mathbb{R}^n$ *and the number* $\lambda = \lambda(\tau,p)$ *are uniquely defined by the system*

$$(\mathbb{I} + \lambda\mathcal{G}_*(\tau))\phi = -\lambda F(t_*,\tau)p,\ \|\phi\|^2 = \lambda^2\bar{\delta}^2(\tau),\ \lambda > 0,$$

in the case that $\|F(t_*,\tau)p\|^2 > \bar{\delta}^2(\tau)$. *Here the matrix* $\mathcal{G}_*(\tau)$ *is defined by (6.32).*

We can then compute the feedback control $u_2^*(t,x)$ at the second interval T_2 using the rules

$$u_2^*(t,x) = \phi_2^T(t,x)f(t),\ t \in T_2,\ x \in \mathbb{R}^n, \tag{6.48}$$

and the behavior of the dynamic system (6.4) is described by

$$\dot{x}(t) = Ax(t) + bu_2^*(t,x(t)) + gw(t),\ t \in T_2,\ x(t_1) = z^* \tag{6.49}$$

with z^* the state of the system (6.41) corresponding to the actual perturbation $w^*(t),\ t \in T_1$. Summarizing the construction of the feedback at the first and second interval we can conclude, that the behavior of the dynamic system (6.4) at the whole interval T is described by

$$\dot{x}(t) = Ax(t) + bu^*(t,x(t)) + gw(t),\ t \in T,\ x(0) = x_0,$$

with the feedback

$$u^*(t,x) = \begin{cases} u_1^*(t,x),\ t \in T_1, \\ u_2^*(t,x),\ t \in T_2. \end{cases} \tag{6.50}$$

6.3 Closed-loop Min-Max Optimal Feedback Control Problem with Several Correction Points

In the previous section we allow to correct our control at one time point. In the CLOCP approach we include feedback aspects at fixed intermediate time points $t_i \in T,\ i =$

$1, \dots, m,$ with $m > 1$. We assume that at the control interval T there are given time moments $t_i,\ i = 1, \dots, m$ with

$$0 = t_0 < t_1 < t_2 < \cdots < t_m < t_{m+1} = t_*.$$

We assume now that we estimate the current state $z(t_i) := z(t_i; u_{t_i}(\cdot), w_{t_i})$ of the actual system at each time moment t_i and then correct the control with the new information about the state.

CLOCP
Construct a control policy

$$\pi = (u_i(\cdot; z_{i-1}),\ i = 1, \dots, m+1) \tag{6.51}$$

consisting of control laws

$$u_i(\cdot; z_{i-1}) = (u_i(t; z_{i-1}), t \in T_i), z_{i-1} \in \mathbb{R}^n,\ i = 1, \dots, m+1 \tag{6.52}$$

at each interval $T_i = [t_{i-1}, t_i],\ i = 1, \dots, m+1$, such that for any admissible disturbance $w(\cdot) \in W$ the state

$$z(t) = z(t; \pi, w(\cdot)),\ t \in T$$

of the system

$$\begin{aligned} &\dot{z}(t) = Az(t) + bu_i(t; z(t_{i-1})) + gw(t),\ t \in T_i,\ i = 1, \dots, m+1 \\ &z(0) = z_0 \end{aligned} \tag{6.53}$$

satisfies the relation

$$\|z(t_*; \pi, w(\cdot)) - x_*\|_2^2 \le \delta_*^2 \text{ for all } w(\cdot) \in W \tag{6.54}$$

and the cost functional

$$J(\pi) = \max_{w(\cdot) \in W} J(\pi, w(\cdot))$$

with

$$J(\pi, w(\cdot)) := \sum_{i=1}^{m+1} \int_{T_i} u_i^2(t; z(t_{i-1}; \pi, w_{t_{i-1}}(\cdot)))dt \tag{6.55}$$

takes the minimal value

$$\min_{\pi} J(\pi). \tag{6.56}$$

We assume that the disturbance w is from the class of admissible disturbances:

$$W = \{w(\cdot) \in L_2(T) : \int_{t_{i-1}}^{t_i} w^2(t)dt \le v_i,\ i = 1, \dots m+1\} \tag{6.57}$$

with given numbers $v_i > 0,\ i = 1, \dots, m+1$.
An interpretation of the optimal control problem (predictive problem) as a dynamic game between the disturbance and the control allows us to better illustrate the advantages of the closed-loop prediction. In the open-loop prediction the whole disturbance sequence plays first, while the control sequence is chosen to counteract the worst disturbance realization. In this case, the effect of the uncertainty may grow over the prediction horizon T and may easily lead to infeasiblity of the predictive problem. On the other hand in the closed-loop schemes the disturbance and the control play one move at a time, thus reducing the influence of the disturbance. Let us derive the worst case optimal control policy π (6.51) using Bellman's Principle of Optimality (see Section 4.2) and dynamic programming. As before we distinguish between the actual system and the nominal system. The actual system with disturbances is given by

$$\begin{aligned} \dot{z}(t) &= Az(t) + bu(t) + gw(t) \\ z(0) &= z_0 \end{aligned} \tag{6.58}$$

and the nominal system without disturbances is given by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + bu(t) \\ x(0) &= x_0 \end{aligned} \tag{6.59}$$

with $z(t) \in \mathbb{R}^n$ the state of the actual system, $x(t) \in \mathbb{R}^n$ the state of the nominal system, $A \in \mathbb{R}^{n\times n}$ and $b,\ g \in \mathbb{R}^n$ given matrix and vectors, respectively and $t \in T = [t_0, t_*]$.

We denote the optimal control policy, that solves CLOCP by

$$\begin{aligned} \pi^0 &= (u_i^0(\cdot; z_{i-1}),\ i = 1, \dots, m+1) \\ u_i^0(\cdot; z_{i-1}) &= (u_i^0(t; z_{i-1}), t \in T_i),\ i = 1, \dots, m+1. \end{aligned} \tag{6.60}$$

CLOCP is a closed-loop worst case optimal feedback control problem (cf. Section 3.5.2), as we consider the disturbance and are able to update the control at certain time moments. The control is not a single control as in the open-loop case, but it is a control policy which includes control laws that depend on the current state of each correction moment. The control policy might also be different for a different realization of the disturbance $w(\cdot)$. We note that the policy π^0 and the control laws $u_i^0(\cdot; z_{i-1})$ also depend on the data of the initial problem. Therefore mathematically correct we should write them as

$$\pi^0(z_0, x_*, t_i, i = 0, \dots, m+1) \text{ and } u_i^0(\cdot; z_{i-1}, x_*, t_s, s = i-1, \dots, m+1),$$

but for simplicity reasons, we will neglect the initial information in the cases in which they are fixed. As in relation (6.13) for the existence of a feasible control in the open-loop worst-case problem formulation, we formulate the necessary and sufficient condition for the existence of a feasible policy π that satisfies (6.54) as

$$\mu(t_{m+1}) v_{m+1} \leq \delta_*^2. \tag{6.61}$$

Here, $\mu(t_i)$ denotes the maximal eigenvalue of the matrix

$$Q_i := \int_{t_{i-1}}^{t_i} F(t_i, t) g (F(t_i, t) g)^T dt,\ i = 1, \dots, m+1. \tag{6.62}$$

Therefore relation (6.61) is significantly weaker than relation (6.13).
As in equality (6.15) we assume, without loss of generality, that the parameter $\delta_0 > 0$ is minimal in the case that there exists a feasible control policy if we choose a suitable δ_0. Therefore we can write

$$\mu(t_{m+1})v_{m+1} = \delta_0^2. \tag{6.63}$$

Let us describe the construction of the control policy using dynamic programming. We will first consider the last interval $T_{m+1} = [t_m, t_{m+1}]$. In this interval the aim is to find a control $u(t)$ that minimizes the function

$$\int_{t_m}^{t_{m+1}} u^2(t)dt \tag{6.64}$$

and also satisfies relation (6.54) for the actual system (6.58). Therefore the control $u(t),\ t \in T_{m+1}$ should drive the nominal system (6.59) from the position z_m at t_m to the position x_* at t_{m+1}. With Pontryagin's Maximum Principle (see Theorem 3.3) we can write the optimal control law for this interval T_{m+1} as

$$u(t) = u^0_{m+1}(t; z_m) = (x_* - F_{m+1}z_m)^T G^{-1}_{m+1} F(t_{m+1}, t)b,\ t \in T_{m+1} \tag{6.65}$$

with

$$F_i := F(t_i, t_{i-1}) \tag{6.66}$$

$$G_i = \int_{T_i} F(t_i, t)b(F(t_i, t)b)^T dt,\ i = 1, \ldots, m+1. \tag{6.67}$$

At this control we obtain the following value of the cost functional (6.64)

$$\begin{aligned} J_m(z_m) &= (x_* - F_{m+1}z_m)^T G^{-1}_{m+1}(x_* - F_{m+1}z_m) \\ &= \|x_* - F_{m+1}z_m\|^2_{G^{-1}_{m+1}}. \end{aligned}$$

Therefore the control law $u^0_{m+1}(t; z_m),\ t \in T_{m+1}$, from the optimal policy π^0 is computed by the relation (6.65). We will now proceed with constructing the control law $u^0_m(t; z_{m-1}),\ t \in T_m$ from the optimal policy π^0. We assume that the actual system (6.58) is in some state z_{m-1} at the moment $t = t_{m-1}$. Let a control $u(t),\ t \in T_m$ be found that drives the nominal system (6.59) from the position z_{m-1} at t_m to some position x_m at t_m. Among all such controls we choose the control that minimizes the function

$$\int_{T_m} u^2(t)dt =: \bar{J}(u(\cdot)). \tag{6.68}$$

The control that drives the actual system (6.58) from the position z_{m-1} at t_{m-1} to some position x_m at t_m and minimizes the functional (6.68) is given by

$$u(t; z_{m-1}, x_m) = (x_m - F_m z_{m-1})^T G^{-1}_m F(t_m, t)b,\ t \in T_m \tag{6.69}$$

and

$$\bar{J}_m(u(\cdot; z_{m-1}, x_m)) = (x_m - F_m z_{m-1})^T G_m^{-1}(x_m - F_m z_{m-1}) = \|x_m - F_m z_{m-1}\|^2_{G_m^{-1}}.$$

From the information (6.57) about admissible disturbances at the interval T_m we can show that the actual system (6.58) appears at the moment $t = t_m$ in a state $z_m \in Z_m(x_m)$ with

$$Z_i(x) := \left\{ z \in \mathbb{R}^n : \|z - x\|^2_{Q_i^{-1}} \leq v_i \right\},\ i = 1, \dots, m+1, \tag{6.70}$$

where Q_i is defined by (6.62), cf. Lemma 6.1. That means, if z_{m-1} is a state of the actual system (6.58) at the time moment $t = t_{m-1}$ and if the control laws (6.65) and (6.69) are used at the intervals T_m and T_{m+1} respectively, then the sum of the last two terms of the cost functional (6.55) is equal to

$$\bar{J}_m(u(\cdot; z_{m-1}, x_m)) + \max_{z_m \in Z_m(x_m)} J_m(z_m). \tag{6.71}$$

The control law (6.69) is optimal if the state x_m minimizes the functional (6.71), which means

$$J_{m-1}(z_{m-1}) := \min_{x_m} \left(\|x_m - F_m z_{m-1}\|^2_{G_m^{-1}} + \max_{z_m \in Z_m(x_m)} J_m(z_m) \right). \tag{6.72}$$

We assume that $x_m^0 = x_m^0(z_{m-1})$ solves the problem (6.72). Then, by taking into account (6.69) we can conclude that the control law $u_m^0(t; z),\ t \in T_m$ from the optimal policy π^0 is given by

$$u_m^0(t; z) = (x_m^0(z) - F_m z)^T G_m^{-1} F(t_m, t) b,\ t \in T_m. \tag{6.73}$$

If we use the optimal control laws (6.73) and (6.65) at the intervals T_m and T_{m+1}, respectively under the assumption, that the state of the actual system at $t = t_{m-1}$ is z_{m-1}, we get the result that the cost function takes the value $J_{m-1}(z_{m-1})$, cf. (6.72). Analogously, we suppose that z_{i-1} is the state of the actual system (6.58) at the time moment t_{i-1} then the cost functional is equal to

$$J_{i-1}(z_{i-1}) := \min_{x_i} \left(\|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} + \max_{z_i \in Z_i(x_i)} J_i(z_i) \right) \tag{6.74}$$

if we use the optimal laws

$$u_s^0(t; z_{s-1}),\ t \in T_s,\ z_{s-1} \in \mathbb{R}^n,\ s = i+1, \dots, m+1 \tag{6.75}$$

and the control

$$u(t; z_{i-1}, x_i) = (x_i - F_i z_{i-1})^T G_i^{-1} F(t_i, t) b,\ t \in T_i.$$

We assume that $x_i^0 = x_i^0(z_{i-1})$ solves the problem (6.74). Then the optimal law $u_i^0(t; z),\ t \in T_i$ from the optimal policy π^0 is given by

$$u_i(t; z) = (x_i^0(z) - F_i z)^T G_i^{-1} F(t_i, t) b,\ t \in T_i. \tag{6.76}$$

We apply the previous idea successively for $i = m, \ldots, 1$. As a result we obtain the min-max cost functional (6.56) as

$$J^0 := J(\pi^0) = J_0(z_0) = \min_{x_1} \left(\|x_1 - F_1 z_0\|^2_{G_1^{-1}} + \max_{z_1 \in Z_1(x_1)} J_1(x_1) \right). \tag{6.77}$$

We calculate the optimal control policy π^0 (6.60) using (6.76). The states $x_i^0(z_{i-1}),\ i = m, \ldots, 1$ in (6.76) solves problem (6.74) for $i = m, \ldots, 1$.
We can rewrite problem (6.74) as the following $(m-i)$-level min-max problem

$$J_{i-1}(z_{i-1}) = \min_{x_i} \max_{z_i \in Z_i(x_i)} \ldots \min_{x_m} \max_{z_m \in Z_m(x_m)} \Phi(x_s, z_s,\ s = i, \ldots, m) \tag{6.78}$$

with

$$\Phi(x_s, z_s,\ s = i, \ldots, m) = \sum_{s=i}^{m+1} \|x_s - F_s z_{s-1}\|^2_{G_s^{-1}}$$
$$x_{m+1} = x_*$$

and the decision variables $x_s \in \mathbb{R}^n,\ z_s \in \mathbb{R}^n,\ s = i, \ldots, m$. This multi-level problem consists of nested min-max optimization problems. Figure 6.3 shows the tube where the

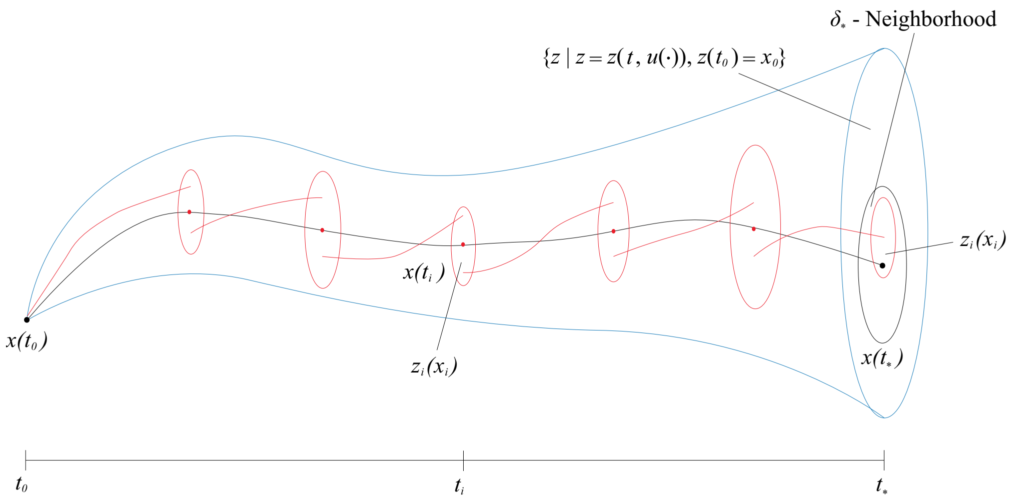


FIGURE 6.3: Optimal feedback control with correction at intermediate time moments for a policy $\pi(\cdot)$.

trajectories corresponding to a particular realization of the uncertainty are inside. The constraints have to be satisfied by every trajectory in the tube. The red ellipsoidals show the reachability sets $Z_i(x_i)$ at the corrrection points. At those points we can measure the current state and correct the control. The black ellipsoidal shows the δ_*^2-neighborhood at our final moment $t = t_*$.

6.4 Approximative Control Policies

As we cannot solve the problem analytically we have to solve each program at each stage numerically which yields high computational costs as we have to first discretize each state to compute and store the min-max cost functional $J_{i-1}(z_{i-1})$ (6.74) and the vectors $x_i^0(z_{i-1})$ for all combinations of the discretized states and for all stages $i = 1, \ldots, m$. Therefore we now present an approximative approach which is suboptimal but it is a more computationally amenable algorithm, which can be implemented online. We formulate and analyze approximative control policies denoted by $\pi(x^0)$, $\pi(\lambda^0)$, $\pi(\lambda^*)$, π^* and $\bar{\pi}$. First we show some results, that are needed to formulate the policies.

We consider the following two optimization problems

$$I_0 := \min_{x} \max_{z \in Z(x)} \left(\|x - F_0 a\|^2_{G_0^{-1}} + \|b - F_* z\|^2_{G_*^{-1}} \right) \tag{6.79}$$

with $Z(x) = \{z \in \mathbb{R}^n : \|z - x\|^2_{\bar{Q}^{-1}} \leq v\}$ and

$$I_* := \min_{\lambda \geq \lambda_*} \left(\|b - \bar{F}_0 a\|^2_{\bar{G}^{-1}(\lambda)} + \lambda v \right) \tag{6.80}$$

with $F_0 \in \mathbb{R}^{n \times n}$ a given matrix, $F_* \in \mathbb{R}^{n \times n}$ a given nonsingular matrix $\lambda \in \mathbb{R}$, $\lambda_* := \lambda_{\max}\left(N^T F_*^T G_*^{-1} F_* N\right)$ where λ_* is the maximal eigenvalue of the matrix $N^T F_*^T G_*^{-1} F_* N$ and where $\bar{Q}^{-1} = (N^{-1})^T N^{-1} \in \mathbb{R}^{n \times n}$ and

$$\begin{aligned} \bar{G}(\lambda) &= F_* G_0 F_*^T + G_* - \frac{F_* Q F_*^T}{\lambda}, \\ \bar{F}_0 &= F_* F_0. \end{aligned}$$

In both problems let the vectors $a, b \in \mathbb{R}^n$, the matrix F_0, the nonsingular matrix $F_* \in \mathbb{R}^{n \times n}$, the positive definite matrices G_0, $G_* \in \mathbb{R}^{n \times n}$, the nonsingular matrix $N \in \mathbb{R}^{n \times n}$ and the positive number $v \in \mathbb{R}$ be given. The matrix $\bar{G}(\lambda)$ is positive definite for all $\lambda \geq \lambda_*$ by construction.

Lemma 6.9. *The equality $I_0 = I_*$ holds. Given an optimal solution $\lambda^0 \in \mathbb{R}$ of the problem (6.80), a solution $x^0 \in \mathbb{R}^n$ of the problem (6.79) is defined by*

$$x^0 = F_0 a - G_0 F_*^T \bar{G}^{-1}(\lambda^0)(\bar{F}_0 a - b).$$

Proof. Using the variable transformation

$$\begin{aligned} p &= F_*(x - F_0 a), \ x = F_*^{-1} p + F_0 a \\ \xi &= b - F_* z, \ z = F_*^{-1}(b - \xi) \end{aligned}$$

and the notations

$$\begin{aligned} G &= (F_* G_0 F_*^T)^{-1} \\ D &= G_*^{-1} \\ Q &= F_* \bar{Q} F_*^T = F_* N (F_* N)^T = A A^T \\ d &= b - F_* F_0 a = b - \bar{F}_0 a \end{aligned}$$

we can apply Theorem 5.16. It is easy to verify that

$$\bar{G}(\lambda) = D^{-1} + G^{-1} - \frac{Q}{\lambda}$$

$$\text{and } \lambda_{\max}((F_*N)^T G_*^{-1}(F_*N)) = \lambda_{\max}(ADA)$$

in problem (5.29). According to Theorem 5.16 it holds that

$$x^0 = F_0 a + F_*^{-1} p^0 = F_0 a + F_*^{-1}(F_* G_0 F_*^T)\bar{G}(\lambda)^{-1}(b - \bar{F}_0 a).$$

□

In what follows we use the following notations

$$Q_i = \int_{t_{i-1}}^{t_i} F(t_i, t) g (F(t_i, t) g)^T dt, \; Q_i = N_i N_i^T \tag{6.81a}$$

$$F_i = F(t_i, t_{i-1}), \; \bar{F}_i = F(t_*, t_{i-1}) \tag{6.81b}$$

$$G_{*i} = \int_{t_i}^{t_*} F(t_*, t) b (F(t_*, t) b)^T dt \tag{6.81c}$$

$$Q_{*i} = \int_{t_{i-1}}^{t_i} F(t_*, t) g (F(t_*, t) g)^T dt \tag{6.81d}$$

$$G_i = \int_{T_i} F(t_i, t) b (F(t_i, t) b)^T dt, \; i = 1, \dots, m+1. \tag{6.81e}$$

$$\tilde{G}_{m+1} = G_{m+1} \tag{6.81f}$$

$$\tilde{G}_i(\lambda_i, \dots, \lambda_m) = G_{*i-1} - \sum_{s=i}^{m} \frac{Q_{*s}}{\lambda_s} \tag{6.81g}$$

$$\bar{G}_i(\lambda_1, \dots, \lambda_{i-1}) = G(t_i, t_0) - \sum_{s=1}^{i-1} \frac{Q_s^i}{\lambda_s} \tag{6.81h}$$

$$A_i = \bar{F}_{i+1} N_i \tag{6.81i}$$

$$\mu_m = \lambda_{\max}(A_m^T G_{m+1} A_m) \tag{6.81j}$$

$$\mu_i(\cdot) = \mu_i(\lambda_{i+1}, \dots, \lambda_m) = \lambda_{\max}\left(A_i^T \tilde{G}_{i+1}^{-1}(\lambda_{i+1}, \dots \lambda_m) A_i\right) \tag{6.81k}$$

$$\mu_m^* = \mu_m \tag{6.81l}$$

$$\mu_i^* = \mu_i(\mu_{i+1}^*, \mu_{i+2}^*, \dots, \mu_m^*), \; i = m-1, m-2, \dots, 1 \tag{6.81m}$$

$$\bar{\mu}_i = \lambda_{\max}(N_i^T F_{i+1}^T G_{i+1}^{-1} F_{i+1} N_i) \tag{6.81n}$$

$$f_i(z, \lambda_i, \dots, \lambda_m) = \left\| x_* - \bar{F}_i z \right\|^2_{\tilde{G}_i^{-1}(\lambda_i, \dots, \lambda_m)} + \sum_{s=i}^{m} \lambda_s v_s, \; i = 1, \dots, m \tag{6.81o}$$

$$f(\lambda_1, \dots, \lambda_m) = \| x_* - F(t_*, t_0) z_0 \|^2_{\tilde{G}_{m+1}^{-1}(\lambda_1, \dots, \lambda_m)} + \sum_{i=1}^{m} \lambda_i v_i \tag{6.81p}$$

$$\Phi(x_s, z_s, \; s = i, \dots, m) = \sum_{s=i}^{m+1} \| x_s - F_s z_{s-1} \|^2_{G_s^{-1}} \tag{6.81q}$$

6.4.1 Problem Formulations

As before we will need the following functions to construct the optimal control policy

$$\begin{aligned} J_m(z) &= \|x_* - F_{m+1}z\|^2_{G^{-1}_{m+1}}, \\ J_{i-1}(z) :&= \min_{x} \max_{z_i \in Z_i(x_i)} \left(\|x_i - F_i z\|^2_{G_i^{-1}} + J_i(z_i) \right),\ i = m, \dots, 1, \\ J_0(z_0) &= \min_{x_1} \left(\|x_1 - F_1 z_0\|^2_{G_1^{-1}} + \max_{z_1 \in Z_1(x_1)} J_1(x_1) \right), \\ J_{i-1}(z_{i-1}) &= \min_{x_i} \max_{z_i \in Z_i(x_i)} \dots \min_{x_m} \max_{z_m \in Z_m(x_m)} \Phi(x_s, z_s,\ s = i, \dots, m). \end{aligned}$$

Further, we define recursively the following problems

$$\begin{aligned} I_m(z) &= J_m(z) \\ P_m(z) :&= J_m(z) \\ I_{i-1}(z) :&= \min_{\lambda_m \geq \mu_m} \min_{\lambda_{m-1} \geq \mu_{m-1}(\cdot)} \dots \min_{\lambda_{i+1} \geq \mu_{i+1}(\cdot)} \min_{x_i} \qquad (6.82) \\ &\qquad \max_{z_i \in Z_i(x_i)} \left(\|x_i - F_i z\|^2_{G_i^{-1}} + f_{i+1}(z_i, \lambda_{i+1}, ..., \lambda_m) \right) \qquad (6.83) \\ P_{i-1}(z) :&= \min_{x_i} \max_{z_i \in Z_i(x_i)} \left(\|x_i - F_i z\|^2_{G_i^{-1}} + I_i(z_i) \right),\ i = m, \dots, 1. \qquad (6.84) \end{aligned}$$

With Lemma 6.9 we can deduce that

$$I_{i-1}(z) := \min_{\lambda_m \geq \mu_m} \min_{\lambda_{m-1} \geq \mu_{m-1}(\cdot)} \dots \min_{\lambda_i \geq \mu_i(\cdot)} f_i(z, \lambda_i, \dots, \lambda_m) \qquad (6.85)$$

Additionally to $J_{i-1}(z)$, $I_{i-1}(z)$ and $P_{i-1}(z)$ we consider also the following problems

$$V^0 = \min_{x_i \in \mathbb{R}^n, i=1,\dots,m} \max_{z_i \in Z_i(x_i), i=1,\dots,m} \Phi(x_i, z_i, i = 1, \dots, m), \qquad (6.86)$$

$$\begin{aligned} W^0 := &\min_{\lambda \in \mathbb{R}^m} f(\lambda) \\ &\text{s.t. } \lambda_i \geq \mu_i(\lambda_{i+1}, \dots, \lambda_m),\ i = 1, \dots, m. \end{aligned} \qquad (6.87)$$

and

$$\begin{aligned} W^* := &\min_{\lambda \in \mathbb{R}^m} f(\lambda) \\ &\text{s.t. } \lambda_i \geq \mu_i^*,\ i = 1, \dots, m. \end{aligned} \qquad (6.88)$$

6.4.2 Analysis of Problem Formulations

Let us now analyze the properties of the problems we formulated in the previous subsection.
First we consider problem $I_{i-1}(z)$ (6.85)

$$I_m(z) = J_m(z)$$

$$I_{i-1}(z) := \min_{\lambda_m \geq \mu_m} \min_{\lambda_{m-1} \geq \mu_{m-1}(\cdot)} \cdots \min_{\lambda_i \geq \mu_i(\cdot)} f_i(z, \lambda_i, \ldots, \lambda_m)$$

In the following propositions we summarize characteristics of the problem

Proposition 6.10. *The set of feasible solutions to problem (6.85) is convex.*

For the proof of Proposition 6.10 we need the following lemma:

Lemma 6.11. *We consider the matrix*

$$S(\lambda) = \left(A - \sum_{i=1}^{k} \frac{B_i}{\lambda_i} \right), \ \lambda = (\lambda_1, \ldots \lambda_k) \in \Lambda,$$

with the matrices A, B_i, $i = 1, \ldots, k$ being positive definite and the set $\Lambda \subset \mathbb{R}^k_+$ being convex. We assume that $S(\lambda)$ is positive definite for all $\lambda \in \Lambda$. Then, the function $\lambda_{\max}(S^{-1}(\lambda))$ is convex at Λ.

Proof. (cf. Kostina & Kostyukova [46])
We can show the convexity of the function $\lambda_{\max}(S^{-1}(\lambda))$ at Λ by proving the concavity of the function $\lambda_{\min}(S(\lambda))$, as it holds that $\lambda_{\max}(S^{-1}(\lambda)) = \frac{1}{\lambda_{\min}(S(\lambda))}$.
We consider the matrix

$$S(\alpha x + (1-\alpha)y) = A - \sum_{i=1}^{k} \frac{B_i}{\alpha x_i + (1-\alpha) y_i} \tag{6.89}$$

with $x, y \in \Lambda$ and $\alpha \in [0,1]$. With the inequality $\frac{1}{(\alpha a + (1-\alpha)b)} \leq \frac{\alpha}{a} + \frac{1-\alpha}{b}$, for all $a > 0, \ b > 0$ and $\alpha \in [0,1]$, we can rewrite matrix (6.89) as

$$\begin{aligned} S(\alpha x + (1-\alpha)y) &= \alpha \left(A - \sum_{i=1}^{k} \frac{B_i}{x_i} \right) + (1-\alpha) \left(A - \sum_{i=1}^{k} \frac{B_i}{y_i} \right) + \sum_{i=1}^{k} B_i \beta_i \\ &= \alpha S(x) + (1-\alpha) S(y) + \bar{B} \end{aligned} \tag{6.90}$$

with $\beta_i \geq 0, \ i = 1, \ldots, k$ some numbers and matrix $\bar{B} = \sum_{i=1}^{k} B_i \beta_i$ positive semidefinite.
With the rewritten matrix (6.90) and the inequalities

$$\begin{aligned} &\lambda_{\min}(S(\alpha x + (1-\alpha)y)) \\ &\quad \geq \lambda_{\min}(\alpha S(x) + (1-\alpha) S(y)) \\ &\quad \geq \alpha \lambda_{\min}(S(x)) + (1-\alpha) \lambda_{\min}(S(y)) \end{aligned}$$

we can deduce the concavity of the function $\lambda_{\min}(S(\lambda))$ (cf. Magnus & Neudecker [61]) and therefore the convexity of the matrix $\lambda_{\max}(S^{-1}(\lambda))$. □

Using Lemma 6.11 we can now prove Proposition 6.10.

Proof. (of Proposition 6.10)
We will use induction to show Proposition 6.10.
For $s = m - 1$ the set $\Lambda_{m-1} := \{\lambda_m \in \mathbb{R} : \mu_m \leq \lambda_m$ is convex and as we constructed the

matrix $\tilde{G}_m(\lambda_m)$ it is positive definite for all $\lambda_m \in \Lambda_{m-1}$.
We now assume that for some index $i < s \leq m-1$ the set

$$\Lambda_s := \{(\lambda_{s+1}, \ldots, \lambda_m) \in \mathbb{R}^{m-i} : \mu_m \leq \lambda_m, \mu_j(\lambda_{j+1}, \ldots, \lambda_m) \leq \lambda_j,\ j = m-1, \ldots s+1\}$$

is convex and the matrix $\tilde{G}_{s+1}(\lambda_{s+1}, \ldots, \lambda_m)$ is positive definite for all $(\lambda_{s+1}, \ldots, \lambda_m) \in \Lambda_{s-1}$.
We now consider the index $s-1$. With Lemma 6.11 for the index $k = m - s$, the matrix $S = A_s \tilde{G}^{-1}_{s+1}(\lambda_{s+1}, \ldots, \lambda_m) A_s$ and the induction assumption we can deduce that the function

$$\mu_s(\lambda_{s+1}, \ldots, \lambda_m) \text{ for } (\lambda_{s+1}, \ldots, \lambda_m) \in \Lambda_s \tag{6.91}$$

is convex. Therefore the set Λ_{s-1} is also convex. We know that the matrix $\tilde{G}_s(\lambda_s, \ldots, \lambda_m)$ is positive definite for all $(\lambda_s, \ldots, \lambda_m) \in \Lambda_{s-1}$ by construction and therefore we can deduce that the sets $\Lambda_s,\ s = m-1, m-2, \ldots, i$ are convex. The set of the feasible solutions to the problem (6.85) coincides with the set Λ_i and therefore the proposition is proven. □

The following proposition follows directly with Proposition 6.10 and Lemma 6.17:

Proposition 6.12. *The cost function of problem (6.85) is convex.*

Remark 6.13. *In the proof of Proposition 6.10 we showed that the functions (6.91), $i = m, \ldots, 1$, are convex. In general case, these functions may not be nondifferentiable at the values $(\bar{\lambda}_{i+1}, \ldots, \bar{\lambda}_m)$ for which the matrix*

$$A_s^T \tilde{G}^{-1}_{s+1}(\bar{\lambda}_{i+1}, \ldots, \bar{\lambda}_m) A_s$$

has several maximal eigenvalues.

In the following lemma we show the relation between the problems $J_i(z_i)$, $P_i(z)$ and $I_i(z)$

Lemma 6.14. *The following inequality holds:*

$$J_i(z) \leq P_i(z) \leq I_i(z),\ i = 0, \ldots, m. \tag{6.92}$$

Proof. The proof is done by induction. By construction we have $J_m(z) = P_m(z) = I_m(z)$ and $J_{m-1}(z) = P_{m-1}(z) = I_{m-1}(z)$. Therefore the inequalities (6.92) hold for $i = m, m-1$. We assume that for some $i \leq m-1$ the inequalities (6.92) hold. We want to show

$$J_{i-1}(z) \leq P_{i-1}(z) \leq I_{i-1}(z). \tag{6.93}$$

With (6.74) and (6.84) it follows that the inequality $J_i(z) \leq I_i(z)$ implies the inequality

$$J_{i-1}(z) \leq P_{i-1}(z). \tag{6.94}$$

With the min-max inequality

$$\max_{y \in Y} \min_{\omega \in W} f(y, \omega) \leq \min_{\omega \in W} \max_{y \in Y} f(y, \omega)$$

where Y and W do not depend on ω and y, respectively, and with (6.83) it holds that

$$\begin{aligned} I_{i-1}(z) &\geq \min_{x_i} \max_{z_i \in Z_i(x_i)} \left(\|x_i - F_i z\|^2_{G_i^{-1}} + \min_{\lambda_m \geq \mu_m} \dots \min_{\lambda_{i+1} \geq \mu_{i+1}(\cdot)} f_{i+1}(z_i, \lambda_{i+1}, \dots, \lambda_m) \right) \\ &= \min_{x_i} \max_{z_i \in Z_i(x_i)} \left(\|x_i - F_i z\|^2_{G_i^{-1}} + I_i(z_i) \right) \\ &= P_{i-1}(z) \end{aligned}$$

With the last inequality and with inequality (6.94) we can deduce inequality (6.93) and therefore the lemma is proven. □

For problem (6.86)

$$V^0 = \min_{x_i \in \mathbb{R}^n, i=1,\dots,m} \max_{z_i \in Z_i(x_i), i=1,\dots,m} \Phi(x_i, z_i, i = 1, \dots, m),$$

we can formulated the following proposition

Proposition 6.15.
The inequality $J^0 \leq V^0$ holds for the optimal values of the cost functionals in the problems (6.78) and (6.86).

Proof. We apply the min-max inequality

$$\max_{y \in Y} \min_{w \in W} f(y, w) \leq \min_{w \in W} \max_{y \in Y} f(y, w),$$

where Y and W do now depend on w and y respectively and get

$$\begin{aligned} J^0 &= \min_{x_1} \max_{z_1 \in Z_1(x_1)} \dots \min_{x_{m-1}} \max_{z_{m-1} \in Z_{m-1}(x_{m-1}} \min_{x_m} \max_{z_m \in Z_m(x_m)} \Phi(x_i, z_i, i = 1, \dots, m) \\ &\leq \min_{x_1} \max_{z_1 \in Z_1(x_1)} \dots \min_{x_{m-1}} \min_{x_m} \max_{z_{m-1} \in Z_{m-1}(x_{m-1})} \max_{z_m \in Z_m(x_m)} \Phi(x_i, z_i, i = 1, \dots, m) \\ &\leq \dots \leq \min_{x_i \in \mathbb{R}^n, i=1,\dots,m} \max_{z_i \in Z_i(x_i), i=1,\dots,m} \Phi(x_i, z_i, i = 1, \dots, m) = V^0 \end{aligned}$$

□

The problem (6.86) is a two-level optimization problem in the variables $x_i \in \mathbb{R}^n,\ i = 1, \dots, m$ (at the upper level) and $z_i \in \mathbb{R}^n,\ i = 1, \dots, m$ (at the lower level). In general, such problems are non-convex and non-continuous and are assumed to be very difficult for computations. However, special properties and structures of problem (6.86) allow us to derive an effective method for its solution.
To apply Lemma 6.9 to problem (6.86) we rewrite the problem in the following form

$$V^0 = \min_{x_m} \dots \min_{x_2} \left(S_1(x_2) + \max_{z_2 \in Z_2(x_2)} \dots \max_{z_m \in Z_m(x_m)} \sum_{i=3}^{m+1} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} \right) \qquad (6.95)$$

with

$$S_1(x_2) = \min_{x_1} \max_{z_1 \in Z_1(x_1)} \left(\|x_1 - F_1 z_0\|^2_{G_1^{-1}} + \|x_2 - F_2 z_1\|^2_{G_2^{-1}} \right).$$

If we now apply Lemma 6.9 we get

$$S_1(x_2) = \min_{\lambda_1 \geq \bar{\mu}_1} \left(\|x_2 - F_2 a_1\|^2_{\bar{G}_2^{-1}(\lambda_1)} + \lambda_1 v_1 \right) \tag{6.96}$$

with

$$\begin{aligned} \bar{\mu}_1 :=& \lambda_{\max}\left(N_1^T F_2^T G_2^{-1} F_2 N_1\right), \; a_1 := F_1 z_0, \\ \bar{G}_2(\lambda_1) =& F_2 G_1 F_2^T + G_2 - \frac{F_2 Q_1 F_2^T}{\lambda_1} = G(t_2, t_0) - \frac{Q_1^2}{\lambda_1} \\ Q_1 =& \int_{t_0}^{t_1} F(t_1, t) g (F(t_1, t) g)^T dt = N_1 N_1^T \end{aligned} \tag{6.97}$$

$$G(t_j, t_i) = \int_{t_i}^{t_j} F(t_j, t) b (F(t_j, t) b)^T dt, \; Q_i^j := \int_{t_{i-1}}^{t_i} F(t_j, t) g (F(t_j, t) g)^T dt. \tag{6.98}$$

If we now substitute (6.96) into problem (6.95) we get

$$V^0 = \min_{\lambda_1 \geq \bar{\mu}_1} \min_{x_m} \dots \min_{x_3} \left(S_2(x_3, \lambda_1) + \lambda_1 v_1 + \max_{z_3 \in Z_3(x_3)} \dots \max_{z_m \in Z_m(x_m)} \sum_{i=4}^{m+1} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} \right) \tag{6.99}$$

with

$$S_2(x_3, \lambda_1) := \min_{x_2} \max_{z_2 \in Z_2(x_2)} \left(\|x_2 - F_2 a_1\|^2_{\bar{G}_2^{-1}(\lambda_1)} + \|x_3 - F_3 z_2\|^2_{G_3^{-1}} \right).$$

Applying Lemma 6.9 to the latter problem yields

$$S_2(x_3, \lambda_1) = \min_{\lambda_2 \geq \bar{\mu}_2} \left(\|x_3 - F_3 a_2\|^2_{\bar{G}_3^{-1}(\lambda_1, \lambda_2)} + \lambda_2 v_2 \right) \tag{6.100}$$

with

$$\begin{aligned} \bar{\mu}_i :=& \lambda_{\max}\left(N_i^T F_{i+1}^T G_{i+1}^{-1} F_{i+1} N_i\right), \; a_i := F_i a_{i-1}, \\ \bar{G}_i(\lambda_1, \dots, \lambda_{i-1}) :=& F_i \bar{G}_{i-1}(\lambda_1, \dots, \lambda_{i-2}) F_i^T + G_i - \frac{F_i Q_{i-1} F_i^T}{\lambda_{i-1}} \\ =& G(t_i, t_0) - \frac{Q_1^i}{\lambda_1} - \frac{Q_2^i}{\lambda_2} - \dots - \frac{Q_{i-1}^i}{\lambda_{i-1}}, \\ Q_i =& \int_{t_{i-1}}^{t_i} F(t_i, t) g (F(t_i, t) g)^T dt = N_i N_i^T. \end{aligned} \tag{6.101}$$

$G(t_j, t_0)$ and Q_i^j are defined by (6.98). If we now substitute (6.100) into (6.99) and perform recursively $m-1$ times the described application of Lemma 6.9 we get the following formulation of V^0

$$V^0 = \min_{\lambda_1 \geq \bar{\mu}_1} \min_{\lambda_2 \geq \bar{\mu}_2} \dots \min_{\lambda_m \geq \bar{\mu}_m} f(\lambda_1, \lambda_2, \dots, \lambda_m)$$

with

$$f(\lambda_1, \lambda_2, \ldots, \lambda_m) = \|x_{m+1} - F_{m+1}a_m\|^2_{\bar{G}^{-1}_{m+1}(\lambda_1,\lambda_2,\ldots,\lambda_m)} + \sum_{i=1}^{m} \lambda_i v_i,$$

$$a_m = F(t_m, t_0)z_0, \; x_{m+1} = x_*,$$

$$\bar{G}_{m+1}(\lambda_1, \lambda_2, \ldots, \lambda_m) = G(t_*, t_0) - \sum_{i=1}^{m} \frac{Q_i^{m+1}}{\lambda_i}.$$

This means, that we transformed the two-level min-max problem (6.86) in the variables $x_i \in \mathbb{R}^n, i = 1, \ldots, m$ at the upper level and $z_i \in \mathbb{R}^n, i = 1, \ldots, m$ in the lower level to the one-level minimization problem in the m variables $\lambda = (\lambda_1, \ldots, \lambda_m)^T$. We can write this problem as

$$V^0 = \min_{\lambda \in \mathbb{R}^m} f(\lambda) \qquad (6.102)$$

$$\text{s.t. } \lambda \geq \bar{\mu}$$

with $f(\lambda) := d^T \bar{G}^{-1}_{m+1}(\lambda)d + v^T\lambda$, the vectors $\bar{\mu} = (\bar{\mu}_1, \ldots, \bar{\mu}_m)^T$, $v = (v_1, \ldots, v_m)^T$, $d = d(z_0, x_*) = x_* - F(t_*, t_0)z_0$ being given.

Proposition 6.16.
The function $f(\lambda)$ is continuous and convex at $\Lambda = \{\lambda \in \mathbb{R}^n : \lambda \geq \bar{\mu}\}$.

The proposition follows directly from the following lemma:

Lemma 6.17. *We consider the function $f(\lambda) = d^T S(\lambda)d$, for $\lambda \in \Lambda$, with the matrix S being defined through $S = S(\lambda) = A - \sum_{i=1}^{m} \frac{B_i}{\lambda_i}$, and with $d \in \mathbb{R}^n$ a given vector, $A \in \mathbb{R}^{n\times n}$ a given matrix, $B_i \in \mathbb{R}^{n\times n}$, $i = 1, \ldots, m$ given positive definite matrices and $\Lambda \subset \mathbb{R}^m_+$ a convex set. We also assume that the matrix $S(\lambda)$ is positive definite for each $\lambda \in \Lambda$. Then the function $f(\lambda)$, $\lambda \in \Lambda$ is convex.*

Proof. To show the lemma we will show that the Hessian of $f(\lambda)$

$$\frac{\partial^2 f(\lambda)}{\partial \lambda^2} = 2(D + diag(\alpha_j, j = 1, \ldots, m))$$

with

$$\alpha_j = \frac{d^T S^{-1} B_j S^{-1} d}{\lambda_j^3}, \; j = 1, \ldots, m$$

and

$$D = \left(\frac{B_1 S^{-1} d}{\lambda_1^2}, \ldots \frac{B_m S^{-1} d}{\lambda_m^2}\right)^T S^{-1} \left(\frac{B_1 S^{-1} d}{\lambda_1^2}, \ldots \frac{B_m S^{-1} d}{\lambda_m^2}\right)$$

is positive definite for all $\lambda \in \Lambda$. The matrix D is positive definite since the matrix $S^{-1}(\lambda)$ is positive definite for all $\lambda \in \Lambda$ and it holds that $\alpha_j \geq 0$, $j = 1, \ldots, m$ because of the inequalities $\lambda_j > 0$, $j = 1, \ldots, m$. Therefore the Hessian is positive definite. □

Problem (6.102) has a solution that can be found by standard methods of nonlinear programming.

For problem (6.87)

$$\begin{aligned} W^0 := &\min_{\lambda \in \mathbb{R}^m} \; f(\lambda) \\ &\text{s.t. } \lambda_i \geq \mu_i(\lambda_{i+1}, \dots, \lambda_m), \; i = 1, \dots, m \end{aligned}$$

we need the following reformulations for J^0 and $P_{i-1}(z_{i-1})$

$$J^0 = \min_{x_1} \max_{z_1 \in Z_1(x_1)} \dots \min_{x_{m-1}} \max_{z_{m-1} \in Z_{m-1}(x_{m-1})} \left(\sum_{i=1}^{m-1} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} + P_{m-1}(z_{m-1}) \right) \tag{6.103}$$

where

$$P_{m-1}(z_{m-1}) = \min_{x_m} \max_{z_m \in Z_m(x_m)} \left(\|x_m - F_m z_{m-1}\|^2_{G_m^{-1}} + \|x_* - F_{m+1} z_m\|^2_{G_{m+1}^{-1}} \right).$$

Using Lemma 6.9 we can write

$$P_{m-1}(z_{m-1}) = \min_{\lambda_m \geq \mu_m} \left(\|x_* - \bar{F}_m z_{m-1}\|^2_{\tilde{G}_m^{-1}} + \lambda_m v_m \right) \tag{6.104}$$

where $\bar{F}$ and $\tilde{G}(\lambda_m)$ are defined in (6.81). If we now substitute (6.104) into (6.103) we obtain an expression for J^0

$$J^0 = \min_{x_1} \max_{z_1 \in Z_1(x_1)} \dots \min_{x_{m-2}} \max_{z_{m-2} \in Z_{m-2}(x_{m-2})} \left(\sum_{i=1}^{m-2} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} + P_{m-2}(z_{m-2}) \right), \tag{6.105}$$

where

$$\begin{aligned} P_{m-2}(z_{m-2}) := &\min_{x_{m-1}} \max_{z_{m-1} \in Z_{m-1}(x_{m-1})} \min_{\lambda_m \geq \mu_m} \\ &\left(\|x_{m-1} - F_{m-1} z_{m-2}\|^2_{G_{m-1}^{-1}} + \|x_* - \bar{F}_m z_{m-1}\|^2_{\tilde{G}_m^{-1}(\lambda)} + \lambda_m v_m \right) \end{aligned}$$

With Lemma 6.9 we can conclude that

$$\begin{aligned} &P_{m-2}(z_{m-2}) \leq I_{m-2}(z_{m-2}) := \min_{\lambda \geq \mu_m} \min_{x_{m-1}} \max_{z_{m-1} \in Z_{m-1}(x_{m-1})} \\ &\left(\|x_{m-1} - F_{m-1} z_{m-2}\|^2_{G_{m-1}^{-1}} + \|x_* - \bar{F}_m z_{m-1}\|^2_{\tilde{G}_m^{-1}(\lambda)} + \lambda_m v_m \right) \\ &= \min_{\lambda_m \geq \mu_m} \min_{\lambda_{m-1} \geq \mu_{m-1}(\lambda_m)} \left(\|x_* - \bar{F}_{m-1} z_{m-2}\|^2_{\tilde{G}_{m-1}^{-1}(\lambda_{m-1}, \lambda_m)} + \sum_{s=m-1}^{m} \lambda_s v_s \right). \end{aligned} \tag{6.106}$$

Here, the matrices $\bar{F}_i$ and $\tilde{G}(\lambda)$ are defined as in (6.81). Using (6.105) and (6.106) we get

$$J^0 \le \min_{x_1} \max_{z_1 \in Z_1(x_1)} \dots \min_{x_{m-3}} \max_{z_{m-3} \in Z_{m-3}(x_{m-3})} \left(\sum_{i=1}^{m-3} \|x_i - F_i z_{i-1}\|{G_i^{-1}}^2 + P_{m-3}(z_{m-3}) \right) \tag{6.107}$$

where

$$P_{m-3}(z_{m-3}) := \min_{x_{m-2}} \max_{z_{m-2} \in Z_{m-2}(x_{m-2})} \min_{\lambda_m \ge \mu_m} \min_{\lambda_{m-1} \ge \mu_{m-1}(\cdot)} \left(\|x_{m-2} - F_{m-2} z_{m-3}\|^2_{G^{-1}_{m-2}} + \|x_* - \bar{F}_{m-1} z_{m-2}\|_{\tilde{G}^{-1}_{m-1}(\lambda_{m-1},\lambda_m)} + \sum_{s=m-1}^{m} \lambda_s v_s \right).$$

If we again use Lemma 6.9 we obtain

$$P_{m-3}(z_{m-3}) \le I_{m-3}(z_{m-3}) := \min_{\lambda_m \ge \mu_m} \min_{\lambda_{m-1} \ge \mu_{m-1}(\cdot)} \min_{\lambda_{m-2} \ge \mu_{m-2}(\cdot)} \left(\|x_* - \bar{F}_{m-2} z_{m-3}\|_{\tilde{G}^{-1}_{m-2}(\lambda_{m-2},\lambda_{m-1},\lambda_m)} + \sum_{s=m-2}^{m} \lambda_s v_s \right). \tag{6.108}$$

By substituting (6.108) into (6.107) and continuing the described operations recursively we get the inequality

$$J^0 \le P_0(z_0)$$

with

$$P_0(z_0) = \min_{\lambda_m \ge \mu_m} \min_{\lambda_{m-1} \ge \mu_{m-1}} \dots \min_{\lambda_1 \ge \mu_1(\cdot)} \left(\|x_* - \bar{F}_1 z_0\|_{\tilde{G}_1^{-1}(\lambda_1,\dots,\lambda_m)} + \sum_{s=1}^{m} \lambda_s v_s \right). \tag{6.109}$$

This means that the optimal value of the cost functional in problem (6.109) may be taken as an approximation for the optimal value of the cost function in the original problem (6.77). By analyzing the relations (6.81) we can easily verify that

$$\|x_* - \bar{F}_1 z_0\|_{\tilde{G}_1^{-1}(\lambda_1,\dots,\lambda_m)} + \sum_{s=1}^{m} \lambda_s v_s = f(\lambda)$$

where $f(\lambda)$ is defined as in (6.102). Therefore we can rewrite problem (6.109) as W^0 (6.87)

$$\begin{aligned} W^0 := &\min_{\lambda \in \mathbb{R}^m} \; f(\lambda) \\ &\text{s.t. } \lambda_i \ge \mu_i(\lambda_{i+1}, \dots, \lambda_m), \; i = 1, \dots, m. \end{aligned} \tag{6.110}$$

Problem (6.88)

$$W^* := \min_{\lambda \in \mathbb{R}^m} f(\lambda)$$

$$\text{s.t. } \lambda_i \geq \mu_i^*, \ i = 1, \ldots, m$$

is a problem of convex function minimization subject to bounds on variables. We will now analyze the relations between the problems (6.102) and (6.88). For this we will need the following lemmas:

Lemma 6.18. *Let the matrices* $\mathcal{A}_* = \mathcal{A}(\lambda_m^*, \ldots, \lambda_i^*)$ *and* $\bar{\mathcal{A}} = \mathcal{A}(\bar{\lambda}_m, \ldots, \bar{\lambda}_i)$ *be defined by*

$$\mathcal{A}(\lambda_m, \ldots, \lambda_i) = M - \sum_{s=i}^{m} \frac{B_s}{\lambda_s}$$

with the matrices $M, \ B_s, \ s = i, \ldots, m$ *being positive definite,* $\bar{\lambda}_s = \lambda_s^* + \Delta\lambda_s, \ \lambda_s^* > 0, \ \Delta\lambda_s \geq 0, \ s = i, \ldots, m.$ *We assume further that the matrix* $\mathcal{A}_*$ *is positive definite. Then it holds that*

$$\lambda_{\max}(\mathcal{A}^{-1}) \geq \lambda_{\max}(\bar{\mathcal{A}}^{-1}).$$

Proof. It is easy to verify, that $\bar{\mathcal{A}} = \mathcal{A}_* + \mathcal{B}$ with the matrix $\mathcal{B} = \sum_{s=i}^{m} \frac{B_s \Delta\lambda_s}{\lambda^* \lambda_s}$ being positive definite. Then it holds that

$$\lambda_{\min}(\mathcal{A}_*) \leq \lambda_{\min}(\bar{\mathcal{A}}).$$

□

Lemma 6.19. *The maximal eigenvalues of the matrices*

$$\tilde{M} := A_{i-1}^T \tilde{G}_i^{-1}(\lambda_i^*, \ldots, \lambda_m^*) A_{i-1} \text{ and } M := N_{i-1} F_i^T G_i^{-1} F_i N_{i-1}$$

with the matrices A_{i-1} *and* $\tilde{G}_i^{-1}(\lambda_i^*, \ldots, \lambda_m^*)$ *being defined by* (6.81) satisfy the inequality

$$\lambda_{\max}(M) \geq \lambda_{\max}(\tilde{M}).$$

Proof. To prove the lemma we have to show that

$$\lambda_{\min}(M^{-1}) \leq \lambda_{\min}(\tilde{M}^{-1}).$$

By (6.81) it holds that $\tilde{M} = M + B$ with the matrices $\tilde{M}, \ M, \ B$ being positive definite. And with Magnus & Neudecker [61] we can deduce that $\lambda_{\max}(M) \geq \lambda_{\max}(\tilde{M})$. □

The Lemmas 6.18 and 6.19 yield the inequalities

$$\bar{\mu}_i \geq \mu_i^*, \ i = 1, \ldots, m$$

and therefore any feasible solution in problem (6.88) is feasible in problem $I_0(z_0)$. Therefore the following inequalities hold

$$V^0 \geq W^* \geq W_0 \geq J_0.$$

This means, that we can use the optimal value of the cost functional in (6.88) as an approximation of the cost functional of the original problem (6.78).

6.4.3 Approximative Control Policy $\pi(x^0)$

We consider the control policy $\pi = \pi(x_1, \ldots, x_m)$ for a fixed set of system states $x_1, x_2, \ldots, x_m$ with the following control laws

$$u_i(t; z) = (x_i - F_i z)^T G_i^{-1} F(t_i, t) b, \; t \in T_i, \; i = 1, \ldots, m+1. \tag{6.111}$$

We can interpret this policy as follows. We assume z_{i-1} to be the state of the actual system at the time moment $t = t_{i-1}$. We choose a control at the interval $T_i = [t_{i-1}, t_i]$ which drives the nominal system (6.59) from the actual position (t_{i-1}, z_{i-1}) into the given position (t_i, x_i) by minimizing the function $\int\limits_{T_i} u^2(t)dt$.

We can now define an optimal control problem with a cost functional $J(\pi)$ that depends on the policy π but does not depend on the disturbance $w(\cdot) \in W$. A conventional choice is

$$J(\pi) = J(\pi(x_1, \ldots, x_m)) = V(x_1, \ldots, x_m) = \max_{z_i \in Z_i(x_i), i=1,\ldots,m} \Phi(x_i, z_i, i = 1, \ldots, m). \tag{6.112}$$

Here, we take into account that for any admissible disturbance $w(t), \; t \in T_i$ the state of the actual system z_i and the state of the nominal system x_i are related through $z_i \in Z_i(x_i)$, with $Z_i(x)$ defined in (6.70). We then choose among all policies π with control law (6.111) a policy that minimizes the cost functional (6.112). The problem can be formulated as V^0 (cf. 6.86)

$$V^0 = \min_{x_i \in \mathbb{R}^n, i=1,\ldots,m} \; \max_{z_i \in Z_i(x_i), i=1,\ldots,m} \Phi(x_i, z_i, i = 1, \ldots, m).$$

The optimal value of the cost functional of problem (6.86) satisfies Proposition 6.15 which means that we can consider the problem (6.86) as an approximation of problem (6.78) and a policy $\pi = \pi(x_1^0, \ldots, x_m^0)$, which is constructed by a solution $x_1^0, \ldots, x_m^0 \in \mathbb{R}^n$ of problem (6.86) can be considered as an approximation of an optimal policy π^0 with the control laws (6.76).

Algorithm 6.20. (Determination of control laws of an approximative policy $\pi(x^0)$)

`GIVEN` Initial state $z_0 = z(t_0)$

`COMPUTE` the matrices F_i, G_i, Q_i, μ_i for $i = 1, \ldots, m+1$

`FORMULATE and SOLVE` (offline) the problem

$$V^0 = \min_{\lambda \in \mathbb{R}^m} f(\lambda)$$
$$\text{s.t. } \lambda \geq \bar{\mu}$$

`SOLUTION:` $\lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0)^T$

`FOR` $i = m, \ldots, 1$ (offline)

`START and SAVE` $x_{m+1}^0 = x_*$
`COMPUTE`

$\bar{G}_i(\lambda_1^0, \ldots, \lambda_{i-1}^0) = G(t_i, t_0) - \sum_{s=1}^{i-1} \frac{Q_s^i}{\lambda_s}$

$d_i = x_{i+1}^0 - F(t_{i+1}, t_0)z_0$

$\psi_i^0 = F_{i+1}^T \bar{G}_{i+1}^{-1}(\lambda_1^0, \ldots, \lambda_i^0)d_i$

$x_i^0 = \bar{G}_i(\lambda_1^0, \ldots, \lambda_{i-1}^0)\psi_i^0 + F(t_i, t_0)z_0$

`SAVE` x_i^0

`END`

`FOR` $i = 1, \ldots, m+1$

`GIVEN` state z_{i-1}

`DETERMINE` the control law at $t \in T_i$

$u_i(t; z_i) = (x_i^0 - F_i z_{i-1})^T G_i^{-1} F(t_i, t)b,\ t \in T_i$

`END`

`DENOTE` the control policy composed of the control laws by $\pi(x^0)$

We want to underline that the described policy $\pi(x_1^0, \ldots, x_m^0)$ is an optimal guaranteed policy at the set of policies of type (6.111) with respect to the worst-case performance (6.55). The policy $\pi(x_1^0, \ldots, x_m^0)$ guarantees for any admissible disturbance $w(\cdot) \in W$

- that the initial state z_0 of the actual system (6.58) is steered into the δ_*-neighborhood of the terminal state x_* in m steps,
- and the value of the cost functional at the realized control does not exceed V_0.

For the cost functional, the estimate V^0 is exact in the class of policies (6.111), namely

$$J\left(\pi(x_1^0, \ldots, x_m^0)\right) = V^0 \leq J(\pi(x_1, \ldots, x_m)).$$

6.4.4 Approximative Control Policy $\pi(\lambda^0)$

First we will use an optimal solution $\lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0)^T$ of problem (6.87) to construct an approximative control policy $\pi = \pi(\lambda^0)$ which guarantees that for all admissible $w(\cdot) \in W$

- the terminal state of the actual system (6.58) remains in the δ_*- neighborhood of the terminal state x_* at the time $t = t_*$,
- and the value of the cost functional at the realized control does not exceed W^0.

Given the solution $x_1^0 = x_1^0(z_0)$ of the problem

$$
\begin{aligned}
W_0(z_0) :=& \sum_{i=2}^{m} \lambda_i^0 v_i \\
&+ \min_{x_1} \Big(\|x_1 - F_1 z_0\|^2_{G_1^{-1}} + \max_{z_1 \in Z_1(x_1)} \|x_* - \bar{F}_2 z_1\|_{\tilde{G}_2^{-1}(\lambda_2^0,\dots,\lambda_m^0)} \Big)
\end{aligned}
\tag{6.113}
$$

we construct the control law $u(\cdot; z_0)$ at the first interval by the rule

$$
u_1(t; z_0) = (x_1^0(z_0) - F_1 z_0)^T G_1^{-1} F(t_1, t) b,\ t \in T_1. \tag{6.114}
$$

We will describe the rules of solving problem (6.113) later. Using Lemma 6.9 and the previous computations it yields

$$
W_0(z_0) = \min_{\lambda_1 \ge \mu_1^0} \Big(\|x_* - F_1 z_0\|^2_{\tilde{G}_1^{-1}(\lambda_2^0,\dots,\lambda_m^0)} + \lambda_1 v_1 \Big) + \sum_{i=2}^{m} \lambda_i v_i
$$

where $\mu_1^0 = \mu_1(\lambda_2^0, \dots, \lambda_m^0)$. The control (6.114) at the interval T_1 has the cost

$$
\int_{T_1} u_1^2(t; z_0) dt = \|x_1^0(z_0) - F_1 z_0\|^2_{G_1^{-1}} =: \Delta W(z_0).
$$

Using (6.113) we may estimate

$$
\begin{aligned}
W_0(z_0) \ge& \|x_1^0(z_0) - F_1 z_0\|^2_{G_1^{-1}} + \|x_* - \bar{F}_2 z_1\|^2_{\tilde{G}_2^{-1}(\lambda_2^0,\lambda_3^0,\dots,\lambda_m^0)} + \sum_{i=2}^{m} \lambda_i^0 v_i \\
\ge& \Delta W_1(z_0) + \sum_{i=3}^{m} \lambda_i^0 v_i \\
&+ \min_{\lambda_2 \ge \mu_2^0} \Big(\|x_* - \bar{F}_2 z_1\|^2_{\tilde{G}_2^{-1}(\lambda_2^0,\lambda_3^0,\dots,\lambda_m^0)} + \lambda_2 v_2 \Big) \\
=& \Delta W_1(z_0) + W_1(z_1) \text{ for all } z_1 \in Z_1(x_1^0(z_0)).
\end{aligned}
\tag{6.115}
$$

Here and in the following μ_i^0 denotes $\mu_i^0 := \mu_i(\lambda_{i+1}^0, \dots, \lambda_m^0),\ i = 2, \dots, m$ and $W_{i-1}(z_{i-1})$ denotes

$$
\begin{aligned}
W_{i-1}(z_{i-1}) :=& \min_{x_i} \Big(\|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} \\
&+ \max_{z_i \in Z_i(x_i)} \|x_* - \bar{F}_{i+1} z_i\|^2_{\tilde{G}_{i+1}^{-1}(\lambda_{i+1}^0,\dots,\lambda_m^0)} \Big) + \sum_{s=i+1}^{m} \lambda_s^0 v_s.
\end{aligned}
\tag{6.116}
$$

By construction the control (6.114) drives the nominal system from the position z_0 at $t = 0$ into the position $x_1^0(z_0)$ at $t = t_1$. Hence, the actual system (6.58) at the moment $t = t_1$ appears at some state $z_1 \in Z_1(x_1^0(z_0))$. Thus, the inequality (6.115) holds for any state z_1 of the actual system (6.58), generated by the control (6.114) and any admissible disturbance $w(t),\ t \in T_1$.

Let us now consider the case for $i \ge 2$. Depending on the state z_{i-1} of the actual system

(6.58) at the moment $t = t_{i-1}$ we define the control law at the interval T_i by the rule

$$u_i(t; z_{i-1}) = (x_i^0(z_{i-1}) - F_i z_{i-1})^T G_i^{-1} F(t_i, t)b,\ t \in T_i, \tag{6.117}$$

where $x_i^0 = x_i^0(z_{i-1})$ solves problem (6.116). The control (6.117) at the interval T_i has the cost

$$\int_{T_i} u_i^2(t; z_0)dt = \|x_i^0(z_{i-1}) - F_i z_{i-1}\|^2_{G_i^{-1}} =: \Delta W_i(z_{i-1}).$$

Using (6.116) and Lemma 6.9 we may estimate

$$\begin{aligned} W_{i-1}(z_{i-1}) \geq & \Delta W_i(z_{i-1}) + \sum_{s=i+2}^{m} \lambda_s^0 v_s \\ & + \min_{\lambda_{i+1} \geq \mu_{i+1}^0} \left(\|x_* - \bar{F}_{i+1} z_i\|^2_{\tilde{G}_{i+1}^{-1}(\lambda_{i+1}, \lambda_{i+2}^0, \dots, \lambda_m^0)} + \lambda_{i+1} v_{i+1} \right) \\ = & \Delta W_i(z_{i-1}) + W_i(z_i) \text{ for all } z_i \in Z_i(x_i^0(z_{i-1})). \end{aligned} \tag{6.118}$$

Repeating recursively the arguments for $i = 3, \dots, m$ we obtain the control policy $\pi(\lambda^0)$ with the control laws (6.117) for each stage $i = 1, \dots, m$. This means, that for any admissible disturbance $w(\cdot) \in W$ the inequalities (6.118) hold at all stages $i = 1, \dots, m+1$. Hence, for all $w(\cdot) \in W$ we have

$$z_{m+1} \in Z_{m+1}(x_{m+1}^0(z_m)) = Z_{m+1}(x_*),\ \sum_{i=1}^{m+1} \Delta W_i(z_{i-1}) \leq W_0(z_0) = W^0.$$

Lemma 6.21. *It holds that*

$$\max_{z_i \in Z_i(x_i^0(z_{i-1})), i=1,\dots,m} \sum_{i=1}^{m+1} \Delta W_i(z_i) = W^0$$

or in the other notations (see (6.55))

$$J(\pi(\lambda^0)) = W^0.$$

The latter expression means, that for the policy $\pi(\lambda^0)$ the estimate W^0 of the cost functional is exact, which means that there exists an admissible disturbance $w(\cdot) \in W$ under which realization and under the policy $\pi(\lambda^0)$ the cost functional at the resulting control has the value W^0.

Let us now discuss the rules of constructing solutions $x_i^0(z_{i-1})$ of problem (6.113) and (6.116). According to Lemma 6.9 we have to compute a solution $\lambda_i^* = \xi_i(\lambda^0, z_{i-1})$ to the problem

$$\min_{\lambda_i \geq \mu_i^0} \left(\|x_* - \bar{F}_i z_{i-1}\|_{\tilde{G}_i^{-1}(\lambda_i, \lambda_{i+1}^0, \dots, \lambda_m^0)} + \lambda_i v_i \right). \tag{6.119}$$

Problem (6.119) is a convex one-dimensional optimization problem at the subspace $\lambda_i \geq \mu_i^0$. We can construct its solution by the following algorithm

Algorithm 6.22. (Determination of $\xi_i(\lambda, z_{i-1})$)

`INPUT` $\tilde{S}(\mu_i, \lambda_{i+1}, \ldots, \lambda_m)$, $\bar{F}_i$, G_{*i}, Q_{*i}, N_i, μ_i, λ_i for $i = 1, \ldots, m+1$

`FOR` $i = 1, \ldots, m$

`IF` $\|N_i^T \bar{F}_{i+1}^T \tilde{S}_i^{-1}(\mu_i, \lambda_{i+1}, \ldots, \lambda_m)(x_* - \bar{F}_i z_{i-1})\|^2 \leq (\mu_i)^2 v_i$

`THEN` $\xi_i(\lambda, z_{i-1}) = \mu_i$,

`ELSE`

`COMPUTE` root of

$\|N_i^T \bar{F}_{i+1}^T \tilde{S}_i^{-1}(\xi_i, \lambda_{i+1}, \ldots, \lambda_m)(x_* - \bar{F}_i z_{i-1})\|^2 = (\xi_i)^2 v_i.$

`THEN` $\xi_i(\lambda, z_{i-1}) = \xi_i > \mu_i$

`DENOTE` the solution by $\xi_i(\lambda, z_{i-1})$

`END`

Remark 6.23. *Given the solution $\xi_i(\lambda^0, z_{i-1})$ we construct a solution $x_i^0(\lambda^0, z_{i-1})$ of problem (6.116) by*

$$\begin{aligned} \Psi_i^0(\lambda^0, z_{i-1}) &:= \bar{F}_{i+1}^T \tilde{G}_i^{-1}(\xi_i(\lambda^0, z_{i-1}), \lambda_{i+1}^0, \ldots, \lambda_m^0)(x_* - \bar{F}_i z_{i-1}) \\ x_i^0(\lambda^0, z_{i-1}) &= F_i z_i + G_i \Psi_i^0(\lambda^0, z_{i-1}). \end{aligned}$$

By including x_i^0 into the control law (6.117) it yields

$$\begin{aligned} u_i(t; z_{i-1}) = (x_* - \bar{F}_i z_{i-1})^T \tilde{G}_i^{-1}(\xi_i(\lambda^0, z_{i-1}), \lambda_{i+1}^0, \ldots, \lambda_m^0) F(t_*, t) b \\ t \in T_i, \ i = 1, \ldots, m+1. \end{aligned}$$

Altogether we can formulate the following algorithms for constructing optimal control laws which define the control policy $\pi(\lambda^0)$.

Algorithm 6.24. (Determination of control laws of an approximative policy $\pi(\lambda^0)$)

`GIVEN` Initial state $z_0 = z(t_0)$

`COMPUTE` the matrices F_i, $\bar{F}_i$, G_i, G_{*i}, Q_{*i}, N_i, $\tilde{G}_1(\lambda_1, \ldots, \lambda_m)$ for $i = 1, \ldots, m+1$

`FORMULATE and SOLVE` (offline) the problem

$$\begin{aligned} I_0(z_0) = \min_{\lambda_i,\ i=1,\ldots,m} f_1(z_0, \lambda_1, \ldots, \lambda_m) = \|x_* - \bar{F}_1 z_0\|^2_{\tilde{G}_1^{-1}(\lambda_1,\ldots,\lambda_m)} + \sum_{s=1}^m \lambda_s v_s \\ \text{s.t. } \lambda_i \geq \mu_i(\lambda_{i+1}, \ldots, \lambda_m), \ i = 1, \ldots, m \end{aligned} \tag{6.120}$$

`SOLUTION` of (6.120): $\lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0)^T$

`FOR` $i = 1, \dots, m$

`GIVEN` state z_{i-1}

`COMPUTE` $\tilde{G}_i(\lambda_i^0, \dots, \lambda_m^0) = G_{*i-1} - \sum\limits_{s=i}^{m} \frac{Q_{*s}}{\lambda_s^0}$

`COMPUTE` $\xi_i(\lambda^0, z_{i-1})$ using Algorithm 6.22 with

$$\tilde{S}(\mu_i, \lambda_{i+1}, \dots, \lambda_m) = \tilde{G}_i(\lambda_i^0, \dots, \lambda_m^0),\ \lambda_i = \lambda_i^0 \text{ and } \mu_i = \mu_i^0$$

`DETERMINE` the control law for $t \in T_i$

$$u_i(t; z_{i-1}) = (x_* - \bar{F}_i z_{i-1})^T \tilde{G}_i^{-1}(\xi_i(\lambda^0, z_{i-1}), \lambda_{i+1}^0, \dots, \lambda_m^0) F(t_*, t) b$$

`END`

`DENOTE` the control policy composed of the control laws by $\pi(\lambda^0)$

6.4.5 Approximative Control Policies $\pi(\lambda^*)$ and $\pi(\tilde{\lambda})$

As we already mentioned in Remark 6.13 the functions $\mu_i(\lambda_{i+1}, \dots, \lambda_m),\ i = 1, \dots, m$ in the constraints of problem (6.120) are convex, but in general, can be nondifferentiable. The nondifferentiability can cause numerical difficulties to the solution of problem (6.120). Hence, along problem (6.120) we consider now problem (6.88). With the solution $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$ of problem (6.88) we define the policy $\pi(\lambda^*)$ following the rules that we described in the previous section where the upper index 0 is changed to $*$, e.g. $\lambda_i^0 \to \lambda_i^*$. The policy $\pi(\lambda^*)$ drives the actual system into the δ_*-neighborhood of the terminal state x_* for any admissible disturbance with the cost which does not exceed the value W^*. While the bound W^0 is exact for the policy $\pi(\lambda^0)$ the bound W^* might be inexact for the policy $\pi(\lambda^*)$. In general, we have the following inequalities

$$\max_{z_i \in Z_i(y_i^*(z_{i-1})), i=1,\dots,m} \sum_{i=1}^{m} \Delta W_i^*(z_i) \le W^* \text{ or } J(\pi(\lambda^*)) \le W^*. \tag{6.121}$$

An inequality in (6.121) can appear since an inequality

$$\mu_i^0 := \mu_i(\lambda_{i+1}^*, \dots, \lambda_m^*) < \mu_i^* := \mu_i(\mu_{i+1}^*, \dots, \mu_m^*) \tag{6.122}$$

can hold in the problem of the type (6.119).

That means we can formulate the following algorithm for constructing the control policy $\pi(\lambda^*)$

Algorithm 6.25. (Determination of control laws of an approximative policy $\pi(\lambda^*)$).

`GIVEN` Initial state $z_0 = z(t_0)$

`COMPUTE` the matrices F_i, $\bar{F}_i$, G_i, G_{*i}, Q_{*i}, N_i, μ_i^* for $i = 1, \dots, m+1$

`FORMULATE and SOLVE` (offline) the problem

$$W^* := \min_{\lambda \in \mathbb{R}^m} f(\lambda)$$
$$\text{s.t. } \lambda_i \geq \mu_i^*, \; i = 1, \ldots, m$$

`SOLUTION` $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)$

`FOR` $i = 1, \ldots, m$

`GIVEN` state z_{i-1}

`COMPUTE` $\tilde{G}_i(\lambda_i^*, \ldots, \lambda_m^*) = G_{*i-1} - \sum\limits_{s=i}^{m} \frac{Q_{*s}}{\lambda_s^*}$

`COMPUTE` $\xi_i(\lambda^*, z_{i-1})$ using Algorithm 6.22 with

$$\tilde{S}(\mu_i, \lambda_{i+1}, \ldots, \lambda_m) = \tilde{G}_i(\lambda_i^*, \ldots, \lambda_m^*), \; \lambda_i = \lambda_i^* \text{ and } \mu_i = \mu_i^*$$

`DETERMINE` the control law for $t \in T_i$

$$u_i(t; z_{i-1}) = (x_* - \bar{F}_i z_{i-1})^T \tilde{G}_i^{-1}(\xi_i(\lambda^*, z_{i-1}), \lambda_{i+1}^*, \ldots, \lambda_m^*) F(t_*, t) b$$

`END`

`DENOTE` the control policy composed of the control laws by $\pi(\lambda^*)$

Remark 6.26. *With the solution $\xi_i(\lambda^*, z_{i-1})$ we can restore the solution $x_i^0(\lambda^*, z_{i-1})$ of problem W^* with*

$$\Psi_i^0(\lambda^*, z_{i-1}) := \bar{F}_{i+1}^T \tilde{G}_i^{-1}(\xi_i(\lambda^*, z_{i-1}), \lambda_{i+1}^*, \ldots, \lambda_m^*)(x_* - \bar{F}_i z_{i-1})$$
$$x_i^0(\lambda^*, z_{i-1}) = F_i z_{i-1} + G_i \Psi_i^0(\lambda^*, z_{i-1}).$$

Remark 6.27. *Let us note, that for any feasible solution $\tilde{\lambda} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m)$ of the problem (6.120) the policy $\pi(\tilde{\lambda})$ with the guaranteed cost $\tilde{W} := f(\tilde{\lambda})$ can be defined by the rules from the previous section with λ_i^0 being replaced by $\tilde{\lambda}_i, \; i = 1, \ldots m$. The estimate $\tilde{W}$ is exact for the policy $\pi(\tilde{\lambda})$ if and only if the following property holds.*

Property 6.28. At each stage $i, \; 1 \leq i \leq m$, the number $\tilde{\lambda}_i$ solves the problem of type (6.119), with z_{i-1} being chosen by $z_{i-1} = \tilde{z}_{i-1}(\tilde{y}_{i-1}(\tilde{z}_{i-2}))$. Here $\tilde{x}_{i-1}(z_{i-2})$ and $\tilde{z}_{i-1}(\tilde{x}_{i-1}(\tilde{z}_{i-2}))$ are optimal solutions of the outer and the inner problems in (6.116) respectively, with i being replaced by $i-1$ and λ_s^0 being replaced by $\tilde{\lambda}_s, \; i = 1, \ldots, m$.

The relation of type (6.122) is necessary for Property 6.28. We will now describe a simple method of computing a feasible (possibly non-optimal) solution in the problem (6.120) that possesses the Property 1 and satisfies $f(\tilde{\lambda}) = \tilde{W} \leq W^*$. For this we solve problem (6.88). Let λ^* be a solution of (6.88). We set $\tilde{\lambda}_m = \lambda_m^*, \; \tilde{\mu}_m = \mu_m^*, \; i = m-1$ and perform recursively the following steps:

1. We assume, that we already know at stage $i,\ 0 \le i \le m-1$ the numbers $\tilde{\lambda}_s,\ s = m, m-1, \ldots, i+1$ that satisfy the inequalities

$$\tilde{\lambda}_s \le \tilde{\mu}_s := \mu_s\left(\tilde{\lambda}_{s+1}, \ldots, \tilde{\lambda}_m\right),\ s = m, m-1, \ldots, i+1.$$

2. Form and solve the convex programming problem with simple bounds on variables

$$\begin{aligned} &\min_{\lambda \in \mathbb{R}^m} \ f(\lambda) \\ &\quad \text{s.t. } \lambda_s = \tilde{\lambda}_s,\ s = i+1, \ldots, m,\ \lambda_s \ge \hat{\mu}_s,\ s = 1, \ldots, i \end{aligned} \tag{6.123}$$

   with

$$\hat{\mu}_s = \mu_s\left(\hat{\mu}_{s+1}, \ldots, \hat{\mu}_i, \tilde{\lambda}_{i+1}, \ldots, \tilde{\lambda}_m\right),\ s = i, i-1, \ldots, 1.$$

3. We assume, that $\left(\hat{\lambda}_1, \ldots, \hat{\lambda}_i, \tilde{\lambda}_{i+1}, \ldots, \tilde{\lambda}_m\right)$ is an optimal solution in (6.123). Then we put $\tilde{\lambda}_i := \hat{\lambda}_i$, replace i with $i-1$ and repeat the operations 1.-3.

In the described procedure there is no necessity to solve the problem (6.123) at each stage. Indeed, assume that at the stage j the optimal solution

$$\left(\hat{\lambda}_1, \ldots, \hat{\lambda}_j, \tilde{\lambda}_{j+1}, \ldots, \tilde{\lambda}_m\right) \tag{6.124}$$

of the corresponding problem (6.123) satisfies the relations

$$\hat{\lambda}_{j-1} > \hat{\mu}_{j-1} \text{ or } \hat{\lambda}_{j-1} = \hat{\mu}_{j-1} \text{ and } \hat{\lambda}_j = \hat{\mu}_j.$$

Then the optimal solution of problem (6.123) corresponding to the next stage $j-1$ coincides with the solution (6.124) from the previous stage j. This means we can formulate the following algorithm for constructing the control policy $\pi(\tilde{\lambda})$.

Algorithm 6.29. (Determination of control laws of an approximative policy $\pi(\tilde{\lambda})$)

`GIVEN` Initial state $z_0 = z(t_0)$

`COMPUTE` the matrices $F_i,\ \bar{F}_i,\ G_i,\ G_{*i},\ Q_{*i},\ N_i,\ \mu_i^*$ for $i = 1, \ldots, m+1$

`FORMULATE and SOLVE` (offline) the problem

$$\begin{aligned} W^* := &\min_{\lambda \in \mathbb{R}^m} f(\lambda) \\ &\text{s.t. } \lambda_i \ge \mu_i^*,\ i = 1, \ldots, m \end{aligned}$$

`SOLUTION` $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)$

`SET` $\tilde{\lambda}_m = \lambda_m^*,\ \tilde{\mu}_m = \mu_m^*$

`FOR` $i = m-1, \ldots, 1$

`FORMULATE and SOLVE` problem

$$\begin{aligned} &\min_{\lambda \in \mathbb{R}^m} \; f(\lambda) \\ &\quad \text{s.t. } \lambda_s = \tilde{\lambda}_s, \; s = i+1, \dots, m \; \lambda_s \geq \hat{\mu}_s, \; s = 1, \dots, i \end{aligned} \tag{6.125}$$

`SOLUTION` $(\hat{\lambda}_1, \dots \hat{\lambda}_j, \tilde{\lambda}_{j+1}, \dots, \tilde{\lambda}_m)$

`SET` $\tilde{\lambda}_i := \hat{\lambda}_i$

`END`

`FOR` $i = 1, \dots, m$

`GIVEN` state z_{i-1}

`COMPUTE` $\tilde{G}_i(\tilde{\lambda}_i, \dots, \tilde{\lambda}_m) = G_{*i-1} - \sum_{s=i}^{m} Q_{*s}/\tilde{\lambda}_s$

`COMPUTE` $\xi_i(\tilde{\lambda}, z_{i-1})$ using Algorithm 6.22 with

$$\tilde{S}(\mu_i, \lambda_{i+1}, \dots, \lambda_m) = \tilde{G}_i(\tilde{\lambda}_i, \dots, \tilde{\lambda}_m), \; \lambda_i = \tilde{\lambda}_i^* \text{ and } \mu_i = \tilde{\mu}_i$$

`DETERMINE` the control law for $t \in T_i$

$$u_i(t; z_{i-1}) = (x_* - \bar{F}_i z_{i-1})^T \tilde{G}_i^{-1}(\xi_i(\tilde{\lambda}, z_{i-1}), \tilde{\lambda}_{i+1}, \dots, \tilde{\lambda}_m) F(t_*, t) b$$

`END`

`DENOTE` the control policy composed of the control laws by $\pi(\tilde{\lambda})$

Remark 6.30. *With the solution $\xi_i(\tilde{\lambda}, z_{i-1})$ we can restore the solution $x_i^0(\tilde{\lambda}, z_{i-1})$ of problem (6.125) with*

$$\begin{aligned} \Psi_i^0(\tilde{\lambda}, z_{i-1}) &:= \bar{F}_{i+1}^T \tilde{G}_i^{-1}(\xi_i(\tilde{\lambda}, z_{i-1}), \tilde{\lambda}_{i+1}, \dots, \tilde{\lambda}_m)(x_* - \bar{F}_i z_{i-1}) \\ x_i^0(\tilde{\lambda}, z_{i-1}) &= F_i z_{i-1} + G_i \Psi_i^0(\tilde{\lambda}, z_{i-1}). \end{aligned}$$

6.4.6 Approximative Control Policy π^*

We can introduce another approximative control police π^* which is described by the control laws

$$u_i(t; z_{i-1}) = (\kappa_i^0(z_{i-1}) - F_i z_{i-1})^T G_i^{-1} F(t_i, t) b \tag{6.126}$$

with $t \in T_i, \; z_{i-1} \in \mathbb{R}^n, \; i = 1, \dots, m+1$ and $x_i^0 = \kappa_i^0(z)$ solves problem (6.83). It is supposed to better approximate the optimal control $\pi(\lambda^0)$ because of the inequality

$$J_{i-1}(z) \leq I_{i-1}(z) \leq W_{i-1}(z) \tag{6.127}$$

which we can deduce from (6.93) and (6.116). If we use this control policy π^* instead of the control policy (6.117) we have to solve online the problem (6.83) at every time moment $t = t_{i-1}$. Because of Lemma 6.9 this is equivalent to the convex programming problem $I_{i-1}(z_{i-1})$ (6.85) depending on the state z_{i-1} of the actual system (6.58) with respect to the decision variables $\lambda_m, \dots, \lambda_i$.
Altogether we can formulate the following algorithm for constructing the control policy π^*.

Algorithm 6.31. (Determination of control laws of an approximative policy π^*)

`GIVEN` Initial state $z_0 = z(t_0)$

`COMPUTE` the matrices F_i, $\bar{F}_i$, G_i, G_{*i}, Q_{*i}, N_i for $i = 1, \dots, m+1$

`FOR` $i = 1, \dots, m$

`GIVEN` state z_{i-1}

`FORMULATE and SOLVE` (offline) the problem

$$\begin{aligned} I_{i-1}(z_{i-1}) &= \min_{\lambda_i} f_i(z_0, \lambda_i, \dots, \lambda_m) \\ \text{s.t. } & \lambda_s \geq \mu_s(\lambda_{s+1}, \dots, \lambda_m), \;\; s = i, \dots, m \end{aligned}$$

`SOLUTION` $\lambda^0 = (\lambda_i^0, \dots, \lambda_m^0)^T$

`COMPUTE` $\tilde{G}_i(\lambda_i^0, \dots, \lambda_m^0) = G_{*i-1} - \sum\limits_{s=i}^{m} \frac{Q_{*s}}{\lambda_s^0}$

`COMPUTE` $\xi(\lambda^0, z_{i-1})$ using Algorithm 6.22 with

$$\tilde{S}(\mu_i, \lambda_{i+1}, \dots, \lambda_m) = \tilde{G}_i(\lambda_i^0, \dots, \lambda_m^0), \;\; \lambda_i = \lambda_i^0 \text{ and } \mu_i = \mu_i^0$$

`DETERMINE` the control law for $t \in T_i$

$$u_i(t; z_{i-1}) = (x_* - \bar{F}_i z_{i-1})^T \tilde{G}_i^{-1}(\xi_i(\lambda^0, z_{i-1}), \lambda_{i+1}^0, \dots, \lambda_m^0) F(t_*, t) b$$

`END`

`DENOTE` the control policy composed of the control laws by $\pi(\lambda^0)$

Remark 6.32. *With the solution $\xi_i(\lambda^0, z_{i-1})$ we can restore the solution $x_i^0(\lambda^*, z_{i-1})$ of problem (6.120) with*

$$\begin{aligned} \Psi_i^0(\lambda^0, z_{i-1}) &:= \bar{F}_{i+1}^T \tilde{G}_i^{-1}(\xi_i(\lambda^0, z_{i-1}), \lambda_{i+1}^0, \dots, \lambda_m^0)(x_* - \bar{F}_i z_{i-1}) \\ x_i^0(\lambda^0, z_{i-1}) &= F_i z_{i-1} + G_i \Psi_i^0(\lambda^0, z_{i-1}). \end{aligned}$$

6.4.7 Approximative Control Policy $\bar{\pi}$

Using the idea of classical feedback, we can introduce another approximative control policy, which we will denote by $\bar{\pi}$. For this control policy, we assume that the actual state of the system at a time t_{i-1} is z_{i-1}. We compute a control such that it drives the state of the nominal system (6.59) from the position z_{i-1} at $t = t_{i-1}$ to the given position x_* at $t = t_*$ and the control minimizes the cost functional

$$\int_{t_{i-1}}^{t_*} u^2(t)dt.$$

Now, this optimal control can be written as

$$u(t; t_{i-1}, z_{i-1}) = (x_* - F(t_*, t_{i-1})z_{i-1})^T G_{*t-1} F(t_*, t)b, \ t \in [t_{i-1}, t_*] \tag{6.128}$$

with G_{*t-1} as in the definition (6.81c). We use this control for the actual system at the interval $[t_{i-1}, t_i]$ to drive this actual system to a state z_i at the moment t_i. At the new position (t_i, z_i) we update the control by the rule (6.128) using this new position. We then continue the process. With these rules we obtain a control policy

$$\bar{\pi} = (\bar{u}_i(\cdot; z_{i-1}), \ i = 1, \ldots, m+1)$$

which is determined by the control laws

$$\bar{u}_i(t, z_{i-1}) = (x_* - F(t_*, t_{i-1})z_{i-1})^T G_{*i-1}^{-1} F(t_*, t)b$$

with $t \in T_i, \ z_{i-1} \in \mathbb{R}^n, \ i = 1, \ldots, m+1$ as in (6.60). The guaranteed value of the cost functional (6.55) of the control policy $\bar{\pi}$ is given by

$$J(\bar{\pi}) = \max_{w(\cdot) \in W} J(\bar{\pi}, w(\cdot))$$

with

$$J(\bar{\pi}, w(\cdot)) = \sum_{i=1}^{m+1} \int_{T_i} \bar{u}_i^2 \left(t; z\left(t_{i-1}; \bar{\pi}, w_{t_{i-1}}(\cdot)\right)\right) dt.$$

Here, $z(t; \bar{\pi}, w_t(\cdot))$ is the state of the system (6.53) at the moment t which is generated by the control policy $\bar{\pi}$ and the disturbance $w_t(\cdot)$. Altogether we can formulate the following algorithm for constructing the control policy $\bar{\pi}$.

Algorithm 6.33. (Determination of control laws of an approximative policy $\bar{\pi}$)

`GIVEN` Initial state $z_0 = z(t_0)$

`COMPUTE` the matrices $\bar{F}_i$ and G_{*i} for $i = 1, \ldots, m+1$

`FOR` $i = 1, \ldots, m$

`DETERMINE` the control law for $t \in T_i$

$$u_i(t; z_{i-1}) = (x_* - \bar{F}_i z_{i-1})^T G_{*i} F(t_*, t) b$$

`END`

`DENOTE` the control policy composed of the control laws by $\bar{\pi}$

6.5 Comparison of Different Control Policies

In this section we will first summarize and then compare the different policies we presented in this chapter. We considered the following policies:

1. The optimal policy π^0 with the control law

$$u_i^0(t; z) = (x_i^0(z) - F_i z)^T G_i^{-1} F(t_i, t) b, \ t \in T_i, \ i = 1, \dots, m+1,$$

   where $x_i^0(z)$ solves the problem

$$J_{i-1}(z) := \min_{x_i} \left(\|x_i - F_i z\|^2_{G_i^{-1}} + \max_{z_i \in Z_i(x_i)} J_i(z_i) \right).$$

2. The approximative policy $\pi(x^0)$ with the control laws

$$u_i^0(t; z) = (x_i^0 - F_i z)^T G_i^{-1} F(t_i, t) b, \ t \in T_i, \ i = 1, \dots, m+1,$$

   where $x_i^0, \ i = 1, \dots, m$ solves the problem

$$V^0 = \min_{x_i \in \mathbb{R}^n, i=1,\dots,m} \ \max_{z_i \in Z_i(x_i), i=1,\dots,m} \Phi(x_i, z_i, i = 1, \dots, m). \tag{6.129}$$

3. The approximative policy $\pi(\lambda^0)$ with the control law

$$u_i^0(t; z) = (y_i^0(z) - F_i z)^T G_i^{-1} F(t_i, t) b, \ t \in T_i, \ i = 1, \dots, m+1,$$

   with $y_i(z)^0$ solving the problem

$$W_{i-1}(z) = \min_{y_i} \left(\|y_i - F_i z\|^2_{G_i^{-1}} + \max_{z_i \in Z_i(y_i)} \|x_* - \bar{F}_{i+1} z_i\|_{G_{i+1}^{-1}(\lambda_{i+1}^0, \dots, \lambda_m^0)} \right) + \sum_{s=i+1}^{m} \lambda_s^0 v_s. \tag{6.130}$$

4. The approximative policy π^* with the control law

$$u_i^0(t; z) = (x_i^0 - F_i z)^T G_i^{-1} F(t_i, t) b, \ t \in T_i, \ i = 1, \dots, m+1,$$

   where $x_i^0, \ i = 1, \dots, m$ solves the problem

$$I_{i-1}(z) = \min_{\lambda_m \geq \mu_m} \ \min_{\lambda_{m-1} \geq \mu_{m-1}(\cdot)} \cdots \min_{\lambda_{i+1} \geq \mu_{i+1}(\cdot)} \ \min_{x_i} \ \max_{z_i \in Z_i(x_i)} \left(\|x_i - F_i z\|^2_{G_i^{-1}} + f_{i+1}(z_i, \lambda_{i+1}, \dots, \lambda_m) \right)$$

$$= \min_{\lambda_m \geq \mu_m} \min_{\lambda_{m-1} \geq \mu_{m-1}(\cdot)} \cdots \min_{\lambda_{i+1} \geq \mu_{i+1}(\cdot)} f_i(z, \lambda_i, \ldots, \lambda_m).$$

5. The approximative policy $\bar{\pi}$ with the control law

$$u_i(t; z_{i-1}) = (x_* - \bar{F}_i z_{i-1})^T G_{*i} F(t_*, t) b,$$

which is constructed based on the classical feedback.

For the values of the policies we have

$$J_{i-1}(z) \leq I_{i-1}(z)$$

and from (6.130) we can deduce that it also holds that

$$I_{i-1}(z) \leq W_{i-1}(z)$$

and therefore

$$J_{i-1}(z) \leq I_{i-1}(z) \leq W_{i-1}(z). \tag{6.131}$$

Due to the inequalities (6.131) the policy π^* should better approximate the optimal policy, but the disadvantage of this policy is that at each time moment $t = t_{i-1}$ we have to solve the following nonlinear programming problem depending on the state z_{i-1} of the actual system

$$\begin{aligned} I_{i-1}(z_{i-1}) = \min_{\lambda_s, s=i,\ldots,m} & \left(d^T(z_{i-1}) \tilde{G}_i^{-1}(\lambda_i, \ldots, \lambda_m) d(z_{i-1}) + \sum_{s=i}^{m} \lambda_i v_i \right) \\ \text{s.t. } & \lambda_s \geq \mu_s(\lambda_{s+1}, \ldots, \lambda_s), \; s = m, m-1, \ldots, i. \end{aligned} \tag{6.132}$$

with $d(z_{i-1}) = x_* - \bar{F}_i z_{i-1}$ we denote the vector known at the moment $t = t_i$. This problem is a convex programming problem in $(m-i+1)$ variables $\lambda_i, \ldots, \lambda_m$. In comparison to this policy, using the policy $\pi(\lambda^0)$ we only have to solve off-line the convex programming problem (6.120) instead of solving the convex programming problem (6.132) online at each time $t = t_{i-1}$ if we use the policy π^*. This means that the policy π^* is computationally impractical even thought it might yield better results.
Altogether we can conclude the following:

- **Control Laws and States:** All policies $\pi^0, \pi(x^0),\ \pi(\lambda^0)$ and π^* have the same expression for the control laws but differ in the states at the time t_i, to which the control law $u_i(t; z),\ t \in T_i$ drives the nominal system from the actual position (t_{i-1}, z). In policy $\pi(x^0)$ the state x_i^0 solves problem (6.129) which is fixed through the process, that means that it does not depend on an actual position (t_{i-1}, z). In the policies $\pi^0,\ \pi(\lambda^0)$ and π^* the state always depends on the actual position and solves one of the problems $J_{i-1}(z),\ W_{i-1}(z)$ or $I_{i-1}(z)$, respectively.

- **Cost Functions:** The optimal values of the cost functionals are related through the inequality (6.131).

- **Transformation:** Problem $J_{i-1}(z)$ can be transformed into an $(m-i+1)-$level min-max problem in the variables $x_s \in \mathbb{R}^n,\ z_2 \in \mathbb{R}^n,\ s = i, \ldots, m$ (see for instance the problem (6.78)). Problem $W_{i-1}(z)$ can be transformed into an one-dimensional convex problem with a bound on a variable (see problem (6.116)). Problem $I_{i-1}(z)$ can be transformed into a constrained convex programming problem in $(m-i+1)$ variables $\lambda_s \in \mathbb{R},\ s = i, \dots, m$ with $(m-i+1)$ constraints (see problem (6.132)).

- **Nonlinearity:** The control laws of policy $\pi(x_i^0)$ are time-dependent but linear in the state z. The control laws of the policies $\pi^0,\ \pi(\lambda^0)$ and π^* are time-dependent and nonlinear in the state z.

6.6 Numerical Example

Example 6.1. ***Simple Pendulum***
We consider a simple pendulum which is a idealized string pendulum. It is a simple model to describe pendulum swings. In the case of a simple pendulum the following simplification assumptions are made. We assume that the cord on which the bob swings is massless, inextensible and always remains taut. Furthermore we assume that the bob is a point mass and that motion only occurs in two dimensions. We also neglect any friction or air resistance. The gravitational field is uniform and the support does not move. Then we can write the motion of a simple pendulum as

$$\ddot{\theta} = -\left(\frac{g}{l}\right)\sin\theta + \frac{u(t)}{l}.$$

with g the acceleration due to gravity, l the length of the pendulum and θ the angular displacement. For linearizing the system we use the small-angle approximation, which means, we consider

$$\sin\theta = \theta$$

Therefore we can transform the upper system in a system of first order ordinary differential equations and add a control u(t) as follows:

$$\begin{aligned} \dot{x}_1(t) &= x_2 \\ \dot{x}_2(t) &= \left(-\frac{g}{l}\right)x_1 + u(t). \end{aligned}$$

For our purposes, we also add an additive uncertainty $w(t)$. We assume $g = 9.81m/s^2$ and $l = 5cm$. This means we consider the following problem

$$\begin{aligned} \dot{z}(t) &= Az(t) + bu(t) + gw(t) \\ z(0) &= z_0 \end{aligned}$$

with

$$A = \begin{pmatrix} 0 & 1 \\ -\frac{g}{l} & 0 \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad g = \begin{pmatrix} 10 \\ 10 \end{pmatrix}$$

and

$$z_0 = \begin{pmatrix} 1.57 \\ 0 \end{pmatrix}, \quad x^* = \begin{pmatrix} 0.94 \\ -17.5 \end{pmatrix}, t^* = 5.$$

The policy was tested on the correction points

$$T = \{t_1 = 1,\ t_2 = 2,\ t_3 = 3,\ t_4 = 4\}.$$

In Table 6.1 the results of the tests of the different policies are presented that confirms the theoretical analysis. We compared the policy $\pi(\lambda^0)$ *with the classical feedback* $\bar{\pi}$ *and with the control method, in which we do not take into account any uncertainty (nu). We can see that the values of the cost functionals of the classical feedback and of the one without any consideration of the uncertainty are always better than the values of the cost functionals of* $\pi(\lambda^0)$*. But on the other hand, the values of* x_2^* *and* x_3^* *differ from the actual* x^**. As we already stated, the computation of the policy* π^* *is not practical. In the case of* $w(t) = 0.32(\cos(t) + \sin(t))$ *in the upper example we obtain the cost function values*

$$J(\pi^*, w(\cdot)) = 120.15 \quad J(\pi(\lambda^0)) = 120.85,$$

which means, that the policy π^* *is, in this case, slightly better than* $\pi(\lambda^0)$ *but it is not efficient at all, as we have to solve a minimization problem in every iteration.*

TABLE 6.1: Values of the cost functionals $J_1 = J(\pi(\lambda^0)),\ J_2 = J(\bar{\pi})$ and J_3, in which we do not take into account any uncertainties, the corresponding values of $x_1^*,\ x_2^*$ and x_3^*, respectively on different $w(\cdot)$.

$w(t)$		0	$\frac{1}{3}$	$-\frac{1}{3}$	$\cos(t)/3$	$-\cos(t)/3$	$\cos(\sqrt{t})/3$	$-\cos(\sqrt{t})/3$
$\pi(\lambda^0)$	J_1	0.006	135.2	135.3	83.44	83.62	42.90	43.0
	x_1^*	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$
$\bar{\pi}$	J_2	0.002	38.70	38.82	82.88	83.53	16.50	16.83
	x_2^*	$(0.94\ \ -17.5)^T$	$(1.19\ \ -20.17)^T$	$(0.69\ \ -14.83)^T$	$(0.78\ \ -15.76)^T$	$(1.10\ \ -19.24)^T$	$(0.84\ \ -16.39)^T$	$(1.04\ \ -18.61)^T$
nu	J_3	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	x_3^*	$(0.94\ \ -17.5)^T$	$(1.13\ \ -18.66)^T$	$(0.74\ \ -16.36)^T$	$(1.14\ \ -18.66)^T$	$(0.74\ \ -16.35)^T$	$(1.14\ \ -18.66)^T$	$(0.74\ \ -16.36)^T$

$w(t)$		$\frac{1}{4}$	$-\frac{1}{4}$	$\sin(t)/3$	$-\sin(t)/3$	$\sin(\sqrt{t})/3$	$-\sin(\sqrt{t})/3$	$0.32(\cos(t)+\sin(t))$
$\pi(\lambda^0)$	J_1	76.03	76.11	49.72	49.83	91.12	91.16	120.85
	x_1^*	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$
$\bar{\pi}$	J_2	21.76	21.85	26.06	26.51	50.94	51.01	86.75
	x_2^*	$(1.13\ \ -19.5)^T$	$(0.75\ \ -15.5)^T$	$(0.75\ \ -15.48)^T$	$(1.13\ \ -19.52)^T$	$(1.17\ \ -19.93)^T$	$(0.71\ \ -15.07)^T$	$(0.60\ \ -13.90)^T$
nu	J_3	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	x_3^*	$(1.09\ \ -18.37)^T$	$(0.79\ \ -16.64)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(0.94\ \ -17.5)^T$	$(1.13\ \ -18.61)^T$

7 Generalizations

In this chapter we generalize the approximative policies which we considered in the previous chapters. In the first part of this chapter we consider a policy with which we can control dynamic systems with pointwise state constraints. After that in the second part we generalize the policy by adding additional constraints to the control and in the last section we consider a disturbance which is bounded in a norm. We also formulate practical algorithms for determining the control policies.

7.1 Guaranteed Control for Systems with State Constraints

Again we assume to have the actual dynamic system as before

$$\begin{aligned} &\dot{z}(t) = Az(t) + bu(t) + gw(t) \\ &z(0) = z_0 \\ &\operatorname{rank}(b, Ab, \dots, A^{n-1}b) = n. \end{aligned} \tag{7.1}$$

and the nominal system

$$\begin{aligned} &\dot{z}(t) = Az(t) + bu(t) \\ &z(0) = z_0 \\ &\operatorname{rank}(b, Ab, \dots, A^{n-1}b) = n. \end{aligned} \tag{7.2}$$

with the time interval $\mathcal{T} = [0, \tau_*]$ and we have given time moments $\tau_j \in \mathcal{T}$ and $t_{ji} \in \mathcal{T}_j = [\tau_{j-1}, \tau_j],\ i = 0, \dots, m,\ j = 1, \dots, N$ with

$$\begin{gathered} 0 = \tau_0 < \tau_1 < \cdots < \tau_N = \tau_* \\ \tau_{j-1} = t_{j0} < t_{j1} < \cdots < t_{jm} < t_{jm+1} = \tau_j. \end{gathered}$$

The disturbance w is from the class of admissible disturbances which is defined as

$$\bar{W} = \{w(\cdot) \in L_2[\mathcal{T}] : \int_{t_{ji-1}}^{t_{ji}} w^2(t)dt \le v_{ji},\ i = 1, \dots, m+1,\ j = 1, \dots, N\} \tag{7.3}$$

with v_{ji} given. As before we can correct the control at the time moment t_{ji} using the information about the current state $z_{ji} = z(t; u_t(\cdot), w_t(\cdot))\Big|_{t=t_{ji}}$ of the actual dynamic system (7.1) with the control $u(t),\ t \in [0, t_{ji})$ and the actual disturbance $w(t),\ t \in [0, t_{ji})$.

The problem can then be described as

Construct a control policy

$$\pi(\mathcal{T}) = (\pi_j(\mathcal{T}_j), j = 1, \dots, N)$$
$$\pi_j(\mathcal{T}_j) = (u_{ji}(\cdot; z_{ji-1}), i = 1, \dots, m+1)$$

consisting of control laws

$$u_{ji}(\cdot; z_{ji-1}) = (u_{ji}(t; z_{ji-1}), t \in \mathcal{T}_j),\ z_{ji-1} \in \mathbb{R}^n,\ i = 1, \dots, m+1,\ j = 1, \dots, N$$

for each interval $\mathcal{T}_{ji},\ j = 1, \dots, N,\ i = 1, \dots, m+1$, such that for any admissible disturbance $w(\cdot)$ the trajectory

$$z(t) = z(t; \pi, w(\cdot)),\ t \in \mathcal{T}_j$$

of the system

$$\dot{z}(t) = Az(t) + bu_{ji}(t; z(\tau_{ji-1})) + gw(t),\ t \in \mathcal{T}_j \tag{7.4}$$

satisfies

$$\|z(\tau_j; u_{\tau_j}(\cdot), w_{\tau_j}(\cdot)) - x^*_{\tau j}\|_2^2 \leq \delta_j^2$$

and the cost functional

$$J(\pi) = \max_{w(\cdot) \in \bar{W}} \sum_{j=1}^{N} \int\limits_{\mathcal{T}_j} u_j^2(t; x(\tau_{j-1}; \pi, w_{j-1}(\cdot))d\tau$$

takes the minimal value

$$\min_{\pi} J(\pi). \tag{7.5}$$

We can now correct the control at every time moment t_{ji} using the information of the current state. Therefore the set of feasible guaranteed policies is much wider than the set of feasible guaranteed program controls.

Theorem 7.1. *(Existence of a guaranteed problem control)*
There always exists a control that guarantees to drive the trajectory of the system (7.4) in the δ_0-neighborhood of zero at the final moment $t = t_{jm+1}$ of each interval $\mathcal{T}_j$ for any perturbation $w(\cdot) \in \bar{W}$ if and only if the inequality

$$\mu(t_j^*) v_{m+1} \leq \delta_*^2$$

holds. Here, $\mu(t_{ji})$ is the maximal eigenvalue of

$$Q_{ji} := \int\limits_{\mathcal{T}_{ji}} F(t_{ji}, t) g (F(t_{ji}, t) g)^T dt,\ i = 1, \dots, m+1.$$

The Theorem follows directly from the lemmas 6.1 and 6.3.

As in equality (6.15) we assume, without loss of generality, that the parameter $\delta_* > 0$ is minimal in the case that there exists a feasible control policy if we choose a suitable δ_*. Therefore we can write

$$\mu(t_j^*)v_{m+1} = \delta_*^2. \tag{7.6}$$

Let us first consider the last interval $\mathcal{T}_{jm+1}$ on the interval $\mathcal{T}_j,\ j = 1, \dots, N$. In this interval the aim is to find a control $u(t)$ that minimizes the function

$$\int_{t_{jm}}^{t_{jm+1}} u^2(t)dt. \tag{7.7}$$

The optimal control law for the interval $\mathcal{T}_{jm+1}$ can be written as

$$u(t) = u^0_{jm+1}(t; z_{jm}) = (x_j^* - F_{jm+1}z_{jm})^T G^{-1}_{jm+1} F(t_{jm+1}, t)b,\ t \in \mathcal{T}_{jm+1} \tag{7.8}$$

with

$$\begin{aligned} F_{ji} &:= F(t_{ji}, t_{ji-1}) \\ G_{ji} &:= \int_{\mathcal{T}_j} F(t_{ji}, t_j)b(F(t_{ji}, t_j)b)^T dt_j. \end{aligned}$$

Hence, at this control we obtain the following value of the cost functional (7.7)

$$\begin{aligned} J_{jm}(z_{jm}) &= (x_j^* - F_{jm+1}z_{jm})^T G^{-1}_{jm+1}(x_j^* - F_{jm+1}z_{jm}) \\ &= \|x_j^* - F_{jm+1}z_{jm}\|^2_{G^{-1}_{jm+1}} \end{aligned}$$

Therefore the control law $u^0_{jm+1}(t; z_{jm}),\ t \in \mathcal{T}_{jm+1}$ from the optimal control policy $\pi_j(\mathcal{T}_j)$ is computed by the relation (7.8). Let us now consider the previous time interval and construct the control law $u^0_{jm}(t; z_{jm-1}),\ t \in \mathcal{T}_{jm}$. We assume that the actual system (7.1) is in some state z_{jm-1} at the moment $t = t_{jm-1}$. Let a control $u(t),\ t \in \mathcal{T}_{jm}$ be found that drives the nominal system (7.2) from the position z_{jm-1} at $t = t_{jm-1}$ to some position x_{jm} at t_{jm}. Among all such controls we choose the control that minimizes the function

$$\int_{t_{jm-1}}^{t_{jm}} u^2(t)dt =: \bar{J}(u(\cdot)). \tag{7.9}$$

The control that drives the actual system (7.1) from the position z_{jm-1} at $t = t_{jm-1}$ to some position z_{jm} at t_{jm} and minimizes the cost functional (7.9) is given by

$$u(t; z_{jm-1}, x_{jm}) = (x_{jm} - F_{jm}z_{jm-1})^T G^{-1}_{jm} F(t_{jm}, t)b,\ t \in \mathcal{T}_{jm} \tag{7.10}$$

and

$$\begin{aligned} \bar{J}_{jm}(u(\cdot; z_{jm-1}, x_{jm})) &= (x_{jm} - F_{jm}z_{jm-1})^T G^{-1}_{jm}(x_{jm} - F_{jm}z_{jm-1}) \\ &= \|x_{jm} - F_{jm}z_{jm-1}\|^2_{G^{-1}_{jm}}. \end{aligned}$$

Using the admissible disturbance (7.3) at the interval $\mathcal{T}_{jm}$ we can show that the actual system (7.1) appears at the moment $t = t_{jm}$ in a state $z_{jm} \in Z_{jm}(x_{jm})$ with

$$Z_{ji} := \left\{ z \in \mathbb{R}^n : \|z - x\|^2_{Q^{-1}_{ji}} \leq v_{ji} \right\},\ i = 1, \dots, m+1.$$

That means, if z_{jm-1} is a state of the actual system (7.1) at the time moment $t = t_{jm-1}$ and if the control laws (7.8) and (7.10) are used at the intervals $\mathcal{T}_{jm}$ and $\mathcal{T}_{jm+1}$ respectively, then the sum of the last two terms of the cost functional is equal to

$$\bar{J}_{jm}(u(\cdot; z_{jm-1}, x_{jm})) + \max_{z_{jm} \in Z_{jm}(x_{jm})} J_{jm}(z_{jm}). \tag{7.11}$$

The control law (7.10) is optimal if the state x_{jm} minimizes the functional (7.11), which means

$$J_{jm-1} := \min_{x_{jm}} \left(\|x_{jm} - F_{jm} z_{jm-1}\|^2_{G^{-1}_{jm}} + \max_{z_{jm} \in Z_{jm}(x_{jm})} J_{jm}(z_{jm}) \right). \tag{7.12}$$

We assume that $x^0_{jm} = x^0_{jm}(z_{jm-1})$ solves the problem (7.12). Then, by taking into account (7.10) we can conclude that the control law $u^0_{jm}(t; z),\ t \in \mathcal{T}_{jm}$ from the optimal policy $\pi_j(\mathcal{T}_j)$ is given by

$$u^0_{jm}(t; z) = (x^0_{jm}(z) - F_{jm} z)^T G^{-1}_{jm} F(t_{jm}, t) b,\ t \in \mathcal{T}_{jm}. \tag{7.13}$$

If we use the optimal control laws (7.13) and (7.8) at the intervals $\mathcal{T}_{jm}$ and $\mathcal{T}_{jm+1}$ respectively under the assumption that the state of the actual system at $t = t_{jm-1}$ is z_{jm-1} we get the result that the cost function takes the value $J_{jm-1}(z_{jm-1})$ as in (7.14). Analogously, we suppose that z_{ji-1} is the state of the actual system (7.1) at the time moment t_{ji-1} then the cost functional is equal to

$$J_{ji-1}(z_{ji-1}) := \min_{x_{ji}} \left(\|x_{ji} - F_{ji} z_{ji-1}\|^2_{G^{-1}_{ji}} + \max_{z_{ji} \in Z_{ji}(x_{ji})} J_{ji}(z_{ji}) \right) \tag{7.14}$$

if we use the optimal control laws

$$u^0_{js}(t; z_{js-1}),\ t \in \mathcal{T}_{js},\ z_{js-1} \in \mathbb{R}^n,\ s = i+1, \dots, m+1$$

and the control

$$u(t; z_{ji-1}, x_{ji}) = (x_{ji} - F_{ji} z_{ji-1})^T G^{-1}_{ji} F(t_{ji}, t),\ t \in \mathcal{T}_{ji}.$$

We assume that $x^0_{ji} = x^0_{ji}(z_{ji-1})$ solves the problem (7.14). Then the optimal control law $u^0_{ji}(t; z),\ t \in \mathcal{T}_{ji}$ from the optimal policy $\pi_j(\mathcal{T}_j)$ is given by

$$u_{ji}(t; z) = (x^0_{ji}(z) - F_{ji} z)^T G^{-1}_{ji} F(t_i, t) b,\ t \in \mathcal{T}_{ji}.$$

We use the following approximative policy

$$\pi(\mathcal{T}) = (\pi_j(z, x^*_j, t_{ji}, i = 0, \dots, m+1), j = 1, \dots, N)$$

as it is practically impossible to construct the policy numerically. To describe the rules for constructing each j-th control policy

$$\pi_j(z, x_j^*, t_{ji}, i = 0, \dots, m+1)$$

we will need the following definitions

$$G_{ji}^* = \int_{t_{ji}}^{t_j*} F(t_j^*, t_j) b (F(t_j^*, t_j) b)^T dt_j \tag{7.15a}$$

$$Q_{ji}^* := \int_{t_{ji-1}}^{t_{ji}} F(t_j^*, t_j) g (F(t_j^*, t_j) g)^T dt_j \tag{7.15b}$$

$$\tilde{G}_{m+1} = G_{jm+1} \tag{7.15c}$$

$$\tilde{G}_{ji}(\lambda_{ji}, \dots \lambda_{jm}) = G_{ji-1} - \sum_{s=i}^{m} Q_{js}^* / \lambda_s \tag{7.15d}$$

$$\bar{F}_{ji} = F(t_j^*, t_{ji-1}) \tag{7.15e}$$

$$\mu_{ji}(\cdot) = \mu_{ji}(\lambda_{ji+1}, \dots, \lambda_{jm}) = \lambda_{\max}(M_{ji}^T \tilde{G}_{ji+1}(\lambda_{ji+1}, \dots, \lambda_{jm}) M_{ji}) \tag{7.15f}$$

$$M_{ji} = \bar{F}_{ji+1} N_{ji} \tag{7.15g}$$

$$f_{ji}(z, \lambda_{ji}, \dots, \lambda_{jm}) = \|x_j^* - \bar{F}_{ji} z\|_{\tilde{G}_j i^{-1}(\lambda_{ji}, \dots, \lambda_{jm})} + \sum_{s=i}^{m} \lambda_{js} v_{js}, \; i = 1, \dots, m. \tag{7.15h}$$

In each j-th step we need the following functions

$$\begin{aligned} J_{jm}(z) &= \|x_j^* - F_{jm+1} z\|^2_{G_{jm+1}^{-1}} \\ J_{ji-1}(z) &:= \min_{x_{ji}} \max_{z_{ji} \in Z_{ji}(x_{ji})} (\|x_{ji} - F_{jm+1} z\|^2_{G_{ji}^{-1}} + J_{ji}(z_{ji})) \\ i &= m, \dots, 1 \end{aligned} \tag{7.16}$$

and let us also introduce the following functions

$$\begin{aligned} I_{jm}(z) &= J_{jm}(z) \\ P_{jm}(z) &:= J_{jm}(z) \\ I_{ji-1}(z) &:= \min_{\lambda_{jm} \geq \mu_{jm}} \min_{\lambda_{jm-1} \geq \mu_{jm-1}(\cdot)} \dots \\ & \min_{\lambda_{ji+1} \geq \mu_{ji+1}(\cdot)} \min_{x_{ji}} \max_{z_{ji} \in Z_{ji}(x_{ji})} \left(\|x_{ji} - F_{ji} z\|^2_{G_{ji}^{-1}} + f_{ji+1}(z_{ji}, \lambda_{ji+1}, \dots \lambda_{jm}) \right) \end{aligned} \tag{7.17}$$

With Lemma 6.9 we can equivalently write

$$I_{ji-1}(z) := \min_{\lambda_{jm} \geq \mu_{jm}} \min_{\lambda_{jm-1} \geq \mu_{jm-1}(\cdot)} \dots \min_{\lambda_{ji} \geq \mu_{ji}(\cdot)} f_{ji}(z, \lambda_{ji}, \dots, \lambda_{jm})$$

$$P_{ji-1}(z) := \min_{x_{ji}} \max_{z_{ji} \in Z_{ji}(x_{ji})} \left(\|x_{ji} - F_{ji} z\|^2_{G_{ji}^{-1}} + I_{ji}(z_{ji}) \right) \tag{7.18}$$

$$i = m, \dots, 1. \tag{7.19}$$

And we can formulate the following lemma

Lemma 7.2. *The following inequality holds:*

$$J_{ji}(z) \le P_{ji}(z) \le I_{ji}(z), \;\; i = 0, \ldots, m. \tag{7.20}$$

Proof. As it is constructed it holds that

$$J_{jm}(z) = P_{jm}(z) = I_{jm}(z)$$

and

$$J_{jm-1}(z) = P_{jm-1}(z) = I_{jm-1}(z).$$

Therefore the inequalities (7.20) hold for $i = m, m-1$. We assume that for some $i \le m-1$ the inequalities (7.20) hold. We want to show

$$J_{ji-1} \le P_{ji-1} \le I_{ji-1}. \tag{7.21}$$

With (7.16) and (7.18) it follows that the inequality $J_{ji}(z) \le I_{ji}(z)$ implies the inequality

$$J_{ji-1}(z) \le P_{ji-1}(z). \tag{7.22}$$

With the general min-max inequality

$$\max_{y \in Y} \min_{\omega \in W} f(y, \omega) \le \min_{\omega \in W} \max_{y \in Y} f(y, \omega)$$

where Y and W do not depend on ω and y, respectively, and with (7.17) it holds that

$$\begin{aligned}
&I_{ji-1}(z) \\
&\quad \ge \max_{x_{ji}} \min_{z_{ji} \in Z_{ji}(x_{ji})} \left(\|x_{ji} - F_{ji} z\|^2_{G^{-1}_{ji}} + \min_{\lambda_{jm} \ge \mu_{jm}} \ldots \min_{\lambda_{ji+1} \ge \mu_{ji+1}(\cdot)} f_{ji+1}(z_{ji}, \lambda_{ji+1}, \ldots, \lambda_{jm}) \right) \\
&\quad = \max_{x_{ji}} \min_{z_{ji} \in Z_{ji}(x_{ji})} \left(\|x_{ji} - F_{ji} z\|^2_{G^{-1}_{ji}} + I_{ji}(z_{ji}) \right) \\
&\quad = P_{ji-1}(z).
\end{aligned}$$

With the last inequality and with inequality (7.22) we can deduce inequality (7.21) and therefore the lemma is proven. □

With $i = 0$ it holds that

$$J_{j0}(z_{j0}) \le I_{j0}(z_{j0})$$

where

$$I_{j0}(z_{j0}) = \min_{\lambda_{jm} \ge \mu_{jm}} \min_{\lambda_{jm-1} \ge \mu_{jm-q}(\cdot)} \cdots \min_{\lambda_{j1} \ge \mu_{j1}(\cdot)} f_{j1}(z_{j0}, \lambda_{j1}, \ldots, \lambda_{jm}) \tag{7.23}$$

and with $f_{j1}(z, \lambda_{j1}, \ldots, \lambda_{jm})$ and $\mu_{ji}(\cdot)$, $i = 1, \ldots, m$ being defined in (7.15). This means, that we can use the optimal value of the problem (7.23) as an approximation for the optimal value of the cost functional of the original problem. Therefore, we can rewrite problem (7.23) as

$$\begin{aligned} I_{j0}(z_{j0}) &= \min_{\lambda_{ji}} f_{j1}(z_{j0}, \lambda_{j1}, \ldots, \lambda_{jm}) \\ \text{s.t. } \lambda_{ji} &\geq \mu_{ji}(\lambda_{ji+1}, \ldots, \lambda_{jm}), \ i = 1, \ldots, m \end{aligned} \tag{7.24}$$

We assume that we have found an optimal solution $\lambda_j^0 = (\lambda_{j1}^0, \ldots, \lambda_{jm}^0)^T$ of problem 7.24 for a given z_{j0} and we will now construct the policy $\pi_j(\mathcal{T}_j)$.
By x_j^* we denote $x_{jm+1}^0(\lambda_j^0, z_j)$ and by $x_{ji}^0(\lambda_j^0, z_j)$ we denote the solution of the problem

$$\begin{aligned} W_{ji-1} :=& \sum_{s=i+1}^{m} \lambda_{js}^0 v_{js} \\ &+ \min_{x_{ji}} \left(\|x_{ji} - F_{ji} z_j\|^2_{G_{ji}^{-1}} + \max_{z_{ji} \in Z_{ji}(x_{ji})} \|x_j^* - \bar{F}_{ji+1} z_j i\|^2_{\tilde{G}_{i+1}^{-1}(\lambda_{ji+1}^0, \ldots, \lambda_{jm}^0)} \right) \end{aligned} \tag{7.25}$$

For the interval $\mathcal{T}_j$ we define the control law by

$$\begin{aligned} u_{ji}(t_j; z_{ji-1}) = (x_{ji}^0(\lambda_j^0, z_{ji-1}) - F_{ji} z_{ji-1})^T G_{ji}^{-1} F_j(t_j i, t_j) b \\ t_j \in \mathcal{T}_j, \ z_{ji-1} \in \mathbb{R}^n, \ i = 1, \ldots, m+1 \end{aligned} \tag{7.26}$$

The control policy, which is composed of control laws (7.26) is denoted by $\pi_j(\mathcal{T}_j)$. Let us note that the solution λ_j^0 of problem (7.24) depends on the data $(z_{j0}, x_j^*, t_{ji}, i = 0, \ldots, m+1)$ of the initial problem and therefore it would be correct to write $\lambda_j^0 = \lambda_j^0(z_{j0}, x_j^*, t_{ji}, i = 0, \ldots, m+1)$ and also $\pi_j(\mathcal{T}_j) = \pi_j(\mathcal{T}_j(z_{j0}, x_j^*, t_{ji}, i = 0, \ldots, m+1))$.

Theorem 7.3. *For a given initial state $z_j(0) = z_{j0}$, the constructed policy $\pi_j(\mathcal{T}_j)$ guarantees for all admissible $w(\cdot) \in \bar{W}$ that*

- *the terminal state of the actual system remains in the δ_*-neighborhood of the terminal state x_j^* at the time $t = t_j^*$.*
- *the value of the cost functional at the realized control does not exceed the number $I_{j0}(z_{j0}) = W_{j0}(z_{j0})$.*

For the policy $\pi_j(\mathcal{T}_j)$ the estimate $W_{j0}(z_{j0})$ is exact.

Proof. We assume $i = 1$ and we let z_{ji-1} be an actual system at the time moment $t = t_{ji-1}$. We consider the problem $W_{ji-1}(z_{ji-1})$ (cf. (7.25)). With the solution $x_{ji}^0 = x_{ji}^0(\lambda^0, z_{ji-1})$ of the problem $W_{ji-1}(z_{ji-1})$ we can construct the control law $u_{ji}(\cdot; z_{ji-1})$ at the interval $\mathcal{T}_j$ as in (7.26). The corresponding cost functional for the control $u_{ji}(t; z_{ji-1})$ at the interval $\mathcal{T}_j$ is then

$$\begin{aligned} \int\limits_{T_{ji}} u_{ji}^2(t; z_{ji-1}) dt &= \|x_{ji}^0(\lambda^0, z_{ji-1}) - F_{ji} z_{ji-1}\|^2_{G_{ji}^{-1}} \\ &=: \Delta W_{ji}(z_{ji-1}). \end{aligned}$$

With (7.25) and Lemma 6.9 it holds that

$$
\begin{aligned}
&W_{ji-1}(z_{ji-1}) \\
&\quad \geq \|x^0_{ji}(\lambda^0, z_{ji-1}) - F_{ji}z_{ji-1}\|^2_{G^{-1}_{ji}} + \|x^*_j - \bar{F}_{ji+1}z_{ji}\|^2_{\tilde{G}^{-1}_{ji+1}(\lambda^0_{ji+1},\dots,\lambda^0_{jm})} + \sum_{s=i+1}^{m} \lambda^0_{js}v_{js} \\
&\quad \geq \Delta W_{ji}(z_{ji-1}) + \sum_{s=i+2}^{m} \lambda^0_{js}v_{js} \\
&\qquad + \min_{\lambda_{ji+1}\geq\mu^0_{ji+1}} \left(\|x^*_j - \bar{F}_{ji+1}z_{ji}\|^2_{\tilde{G}^{-1}_{ji+1}(\lambda_{ji+1},\lambda^0_{ji+2},\dots,\lambda^0_{jm})} + \lambda_{ji+1}v_{ji+1} \right) \\
&\quad = \Delta W_{ji}(z_{ji-1}) + W_{ji}(z_{ji})
\end{aligned}
\tag{7.27}
$$

for all $z_{ji} \in Z_{ji}(x^0_{ji}(\lambda^0_j, z_{ji-1}))$.

The control law (7.26) drives the nominal system from the position (t_{ji-1}, z_{ji-1}) into the position $(t_{ji}, x^0_{ji}(\lambda^0_j, z_{ji-1}))$. Therefore we know that the actual system (7.1) is at some state $z_{ji} \in Z_{ji}(x^0_{ji}(\lambda^0_j, z_{ji-1}))$ at the time moment $t = t_{ji}$. That means that the inequality (7.27) holds for any state z_{ji} of the actual system (7.1) which is generated by the control $u_{ji}(t; z_{ji-1})$ at any admissible disturbance $w(t),\ t \in \mathcal{T}_j$.
We repeat this procedure recursively for the arguments $i = 2, \dots, m$. We get the result that for any admissible disturbance $w(\cdot) \in \bar{W}$ the inequalities (7.27) hold at all stages $i = 1, \dots, m+1$. Therefore, for all $w(\cdot) \in \bar{W}$ it holds that

$$
\begin{aligned}
z_{jm+1} &= z(t^*_j; \pi(\lambda^0_j), w(\cdot)) \in Z_{jm+1}(x^0_{jm+1}(\lambda^0_j, z_{jm})) \\
&= Z_{jm+1}(x^*_j)
\end{aligned}
$$

and

$$
\sum_{i=1}^{m+1} \Delta W_{ji}(z_{ji-1}) \leq W_{j0}(z_{j0}).
$$

Also, the following equality holds

$$
\max_{z_{ji}\in Z_{ji}(x^0_{ji}(\lambda^0_j, z_{ji-1})),\ i=1,\dots,m} \sum_{i=1}^{m+1} \Delta W_{ji}(z_{ji-1}) = W_{j0}(z_{j0})
$$

which can be rewritten as

$$
J(\pi_j(\mathcal{T}_j)) = W_{j0}(z_{j0}).
$$

Hence, the estimate $W_{j0}(z_{j0})$ of the cost functional is exact for the control policy $\pi_j(\mathcal{T}_j)$. This means that there is an admissible disturbance $w(\cdot) \in W$ such that under the policy $\pi_j(\mathcal{T}_j)$ the cost functional at the resulting control has the value $W_{j0}(z_{j0})$. □

We will now construct the solution of the subsidiary problem (7.25). From Lemma 7.2

it follows that we need to compute a solution $\lambda^*_{ji} = \xi_{ji}(\lambda^0_j, z_{ji-1})$ of the problem

$$\min_{\lambda_{ji}\geq\mu^0_{ji}} \|x^*_j - \bar{F}_{ji}z_{ji-1}\|^2_{\tilde{G}^{-1}_{ji}(\lambda_{ji},\lambda^0_{ji+1},\ldots,\lambda^0_{jm})} + \lambda_{ji}v_{ji}. \tag{7.28}$$

This problem is a convex one-dimensional optimization problem at the subspace $\lambda_{ji} \geq \mu^0_{ji}$. We can construct the solution by the following rule: If it holds that

$$\|N^T_{ji}\bar{F}^T_{ji+1}\tilde{G}^{-1}_{ji}(\mu^0_{ji}, \lambda^0_{ji+1}, \ldots, \lambda^0_{jm})(x^*_j - \bar{F}_{ji}z_{ji-1})\|^2 \leq (\mu^0_{ji})^2 v_{ji}$$

then

$$\xi_{ji}(\lambda^0_j, z_{ji-1}) = \mu^0_{ji}$$

else $\xi_{ji}(\lambda^0_j, z_{ji-1}) = \xi_{ji} > \mu^0_{ji}$ is a unique root of the following equation

$$\|N^T_{ji}\bar{F}^T_{ji+1}\tilde{G}^{-1}_{ji}(\xi_{ji}, \lambda^0_{ji+1}, \ldots, \lambda^0_{jm})(x^*_j - \bar{F}_{ji}z_{ji-1})\|^2 - (\xi_{ji})^2 v_{ji} = 0.$$

With the solution $\xi_{ji}(\lambda^0_j, z_{ji-1})$ of the problem (7.28) we can construct a solution $x^0_{ji}(\lambda^0_j, z_{ji-1})$ of problem (7.25) as

$$\begin{aligned}
\psi^0_{ji}(\lambda^0_j, z_{ji-1}) &:= \bar{F}^T_{ji+1}\tilde{G}^T_{ji}(\xi_{ji}(\lambda^0_j, z_{ji-1}), \lambda^0_{ji+1}, \ldots, \lambda^0_{j,})(x^*_j - \bar{F}_{ji}z_{ji-1}) \\
x^0_{ji}(\lambda^0_j, z_{ji-1}) &= F_{ji}z_{ji-1} + G_{ji}\psi^0_{ji}(\lambda^0_j, z_{j-1})
\end{aligned}$$

Using this the control law (7.26) can be written as

$$u_{ji}(t; z_{ji-1}) = (x^*_j - \bar{F}_{ji}z_{ji-1})^T\tilde{G}^{-1}_{ji}(\xi_{ji}(\lambda^0_j, z_{ji-1}), \lambda^0_{ji+1}, \ldots, \lambda^0_{jm})F(t^*_j, t_j)b$$

with $t_j \in \mathcal{T}_j,\ i = 1, \ldots, m$. With these rules we can construct the control of the actual system (7.1) as follows for the first two time intervals:

1. Determine the policy

   $$\pi_1(z(0), x^*_1, t_{1i}, i = 0, \ldots, m+1)$$

   by constructing the control laws as in e.g. (7.26) using the information about the current state $z_0 = z(0)$ of the actual system.

2. Use the control laws of the policy to control the actual system (7.1) at the interval $\mathcal{T}_1$.

3. At the moment $t = \tau_1$ the actual system is driven to some state $z(\tau_1)$ satisfying $\|z(\tau_1) - x^*_1\|^2_2 \leq \delta^2_1$.

4. Use the policy $\pi_2\left(z(\tau_1), x^*_2, t_{2i}, i = 0, \ldots, m+1\right)$ with the state $z(\tau_1)$ of the actual system at the moment $t = \tau_1$ and a given vector x^*_2 to control the system at the interval $\mathcal{T}_2$.

We repeat this strategy for the rest of the time intervals. Altogether the policy

$$\pi(\mathcal{T}) = (\pi_j(\mathcal{T}_j),\ j = 1, \ldots, N)$$

is then constructed by the control policies

$$\pi_j(\mathcal{T}_j) = (u_{ji}(\cdot; Z_{ji-1}), i = 1, \dots, m+1)$$

for each interval $\mathcal{T}_{ji} = [t_{ji-1}, t_{ji}],\ j = 1, \dots, N,\ i = 1, \dots, m+1$. And the control policies $\pi_j(\mathcal{T}_j)$ consist of the control laws

$$u_{ji}(\cdot; z_{ji-1}) = (u_{ji}(t; z_{ji-1}),\ t \in \mathcal{T}_{ji}$$
$$z_{ji-1} \in \mathbb{R}^n,\ i = 1, \dots, m+1,\ j = 1, \dots, N.$$

With these information we have shown that there exists a control policy $\pi(\mathcal{T})$ for controlling the system (7.1).

Theorem 7.4. *(Existence of the control policy $\pi(\mathcal{T})$)*
For a given initial state $z(0) = z_0$ the constructed policy $\pi(\mathcal{T})$ guarantees for all admissible $w(\cdot) \in \bar{W}$ that

- *the terminal state of the actual system remains in the δ_*-neighborhood of the terminal state x_* at the time $t = \tau_*$.*
- *the value of the cost functional at the realized control does not exceed the number $I_0(z_0) = W_0(z_0)$.*

For the policy $\pi(\mathcal{T})$ the estimate $W_0(z_0)$ of the cost functional is exact.

Remark 7.5. *The value of the cost functional at each interval $\mathcal{T}_j$ is equal to β_j for the described strategy:*

$$\beta_j = \max_{z \in \bar{Z}_j(x^*_{j-1})} I_0\left(z, x^*_j, t_{ji}, i = 0, \dots, m+1\right)$$

with

$$\bar{Z}_j(x) = \{z \in \mathbb{R}^n : \exists f \in \mathbb{R}^r, z = x + N_{j-1,m+1}f, f^T f \leq v_{j-1,m+1}\}$$

and

$$I_0(z, x^*_j, t_{ji}, i = 0, \dots, m+1)$$

is the optimal value in problem 7.24.

We can also use this strategy for the case that the constraint satisfies a given accuracy δ_* at the moments $\tau_j := j\Delta\tau,\ \Delta\tau > 0,\ j = 1, 2, \dots$ over the infinite horizon. The value of the cost functional

$$\int_{\mathcal{T}_j} u^2(t)dt$$

with $\mathcal{T}_j = [\tau_{j-1}, \tau_j]$ is then guaranteed to be smaller than some number β_j that is computed before.

Algorithm 7.6. (Determination of control laws of an approximative policy $\pi(\mathcal{T})$)

`FOR` $j = 1, \dots, N$

`GIVEN` Initial state $z_{j0} = z(t_{j0})$

`COMPUTE` the matrices F_{ji}, $\bar{F}_{ji}$, G_{ji}, G_{*ji}, Q_{*ji}, N_{ji} for $i = 1, \ldots, m+1$

`FORMULATE and SOLVE` (offline) the problem

$$\begin{aligned} I_{j0}(z_{j0}) &= \min_{\lambda_{ji}} f_1(z_{j0}, \lambda_{j1}, \ldots, \lambda_{jm}) \\ \text{s.t. } \lambda_{ji} &\geq \mu_{ji}(\lambda_{ji+1}, \ldots, \lambda_{jm}),\ i = 1, \ldots, m \end{aligned} \tag{7.29}$$

`SOLUTION` of problem (7.29): $\lambda_j^0 = (\lambda_{j1}^0, \ldots, \lambda_{jm}^0)^T$

`FOR` $i = 1, \ldots, m$

`GIVEN` z_{ji-1}

`COMPUTE` $\tilde{G}_{ji}(\lambda_{ji}^0, \ldots, \lambda_{jm}^0) = G_{*ji-1} - \sum\limits_{s=i}^{m} Q_{*js}/\lambda_{js}^0$

`COMPUTE` $\xi(\lambda_j^0, z_{ji-1})$ using Algorithm 6.22 with

$\tilde{S}(\mu_i, \lambda_{i+1}, \ldots, \lambda_m) = \tilde{G}_{ji}(\lambda_{ji}^0, \ldots, \lambda_{jm}^0)$, $\lambda_i = \lambda_{ji}^0$ and $\mu_i = \mu_{ji}^0$

`DETERMINE` the control law

$$u_{ji}(t_j; z_{ji-1}) = (x_j^* - \bar{F}_{ji} z_{ji-1})^T \tilde{G}_{ji}^{-1}(\xi_{ij}(\lambda_j^0, z_{ji-1}), \lambda_{ji+1}^0, \ldots, \lambda_{jm}^0) F(t_j^*, t_j) b$$

`END`

`DENOTE` the control policy composed of the control laws by $\pi_j(\mathcal{T}_j)$

`END`

`DENOTE` the control policy composed of the control policies $\pi_j(\mathcal{T}_j)$ by $\pi(\mathcal{T})$

Remark 7.7. *Given the solution $\xi(\lambda_j^0, z_{ji-1})$ we can restore the solution of problem (7.29) by*

$$\begin{aligned} \Psi_{ji}^0(\lambda_j^0, z_{ji-1}) &= \bar{F}_{ji+1}^T \tilde{G}_{ji}^{-1}(\xi_{ji}(\lambda_j^0, z_{ji-1}), \lambda_{ji+1}^0, \ldots, \lambda_{jm}^0)(x_j^* - \bar{F}_{ji} z_{ji-1}) \\ x_{ji}^0(\lambda_j^0, z_{ji-1}) &= F_{ji} z_{ji-1} + G_{ji} \Psi_{ji}^0(\lambda_j^0, z_{ji-1}). \end{aligned}$$

7.2 Alternative Approach to Guaranteed Control for Systems with State Constraints

Another possibility to solve such problems is the following approach. We assume to have the same actual dynamic system (7.1) with the time interval $T = [0, t_{m+1}]$ and the given disturbance (7.3) as before. The optimal control law at the interval $T_{m+1} = [t_m, t_{m+1}]$

is given by

$$u(t) = (x^*_{m+1} - F_{m+1}z_m)^T G^{-1}_{m+1} F(t_{m+1}, t)b,\ t \in T_{m+1} \tag{7.30}$$

and the cost functional at the control (7.30) gets the value

$$J_m(z_m) = \|x^*_{m+1} - F_{m+1}z_m\|^2_{G^{-1}_{m+1}}. \tag{7.31}$$

Consider the moment t_{m-1} and assume that the system is at the position z_{m-1} at $t = t_{m-1}$. The control that drives us from the position z_{m-1} at $t = t_{m-1}$ to some position x_m at $t = t_m$ is

$$u(t; z_{m-1}, x_m) = (x_m - F_m z_{m-1})^T G^{-1}_m F(t_m, t)b \tag{7.32}$$

and the corresponding cost function can be written as

$$\bar{J}_m(z(\cdot; z_{m-1}, x_m)) = \|x_m - F_m z_{m-1}\|^2_{G^{-1}_m}.$$

The actual state at this moment is $z_m \in Z_m(x_m)$ with

$$Z_m(x_m) = \{z \in \mathbb{R}^n | \|z - x_m\|^2_{Q^{-1}_m} \leq v_m\}.$$

Any state $z_m \in Z_m(x_m)$ should additionally satisfy

$$\|z_m - x^*_m\|^2_{S_m} \leq \delta_m$$

where S_m and δ_m are given. That is why x_m should satisfy

$$x_m \in B_m = \{x | \|x - x^*_m\|^2_{S_m} \leq \bar{\delta}_m\}$$

where $\bar{\delta}_m$ is such that

$$\|z_m - x^*_m\|^2_{S_m} \leq \delta_m, \quad \text{for all } z_m \in Z_m(x_m).$$

If z_{m-1} is the state of the actual system at the time $t = t_{m-1}$ and if the control laws (7.30) and (7.32) are used at the intervals T_{m+1} and T_m, respectively then the sum of the last terms in the cost functional is equal to

$$\bar{J}_m(u(\cdot; z_{m-1}, x_m)) + \max_{z_m \in Z_m(x_m)} J_m(z_m). \tag{7.33}$$

Clearly, for the position z_{m-1} at t_{m-1} the control law (7.32) is optimal if the state x_m minimizes the functional (7.33), that is

$$J_{m-1}(z_{m-1}) := \min_{x_m \in B_m} \left(\|x_m - F_m z_{m-1}\|^2_{G^{-1}_m} + \max_{z_m \in Z_m(x_m)} \|x^*_{m+1} - F_{m+1}z_m\|^2_{G^{-1}_{m+1}} \right). \tag{7.34}$$

We assume that $x^0_m = x^0_m(z_{m-1})$ solves the problem (7.34). Then, by taking into account (7.32) we can conclude that the control law $u^0_m(t; z),\ t \in T_m$ from the optimal policy $\pi(\mathcal{T})$ is given by

$$u^0_m(t; z) = (x^0_m(z) - F_m z)^T G^{-1}_m F(t_m, t)b,\ t \in T_m. \tag{7.35}$$

If we use the optimal control laws (7.35) and (7.30) at the intervals T_m and T_{m+1}, respectively under the assumption, that the state of the actual system at $t = t_{m-1}$ is z_{m-1}, we get the result that the cost function takes the value $J_{m-1}(z_{m-1})$, cf. (7.34). Analogously, we suppose that z_{i-1} is the state of the actual system (7.1) at the time moment t_{i-1} then the cost functional is equal to

$$J_{i-1}(z_{i-1}) := \min_{x_i \in B_i} \left(\|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} + \max_{z_i \in Z_i(x_i)} J_i(z_i) \right) \tag{7.36}$$

where $B_i = \{x | \|x - x_i^*\| \leq \bar{\delta}_i\}$ where $\bar{\delta}_i$ is chosen such that

$$\|z_i - x_i^*\|^2_{S_i} \leq \delta_i \text{ for all } z_i \in Z_i(x_i),$$

if we use the optimal laws

$$u_s^0(t; z_{s-1}),\ t \in T_s,\ z_{s-1} \in \mathbb{R}^n,\ s = i+1, \ldots, m+1 \tag{7.37}$$

and the control

$$u(t; z_{i-1}, x_i) = (x_i - F_i z_{i-1})^T G_i^{-1} F(t_i, t) b,\ t \in T_i.$$

We assume that $x_i^0 = x_i^0(z_{i-1})$ solves the problem (7.36). Then the optimal law $u_i^0(t; z),\ t \in T_i$ from the optimal policy $\pi(\mathcal{T})$ is given by

$$u_i(t; z) = (x_i^0(z) - F_i z)^T G_i^{-1} F(t_i, t) b,\ t \in T_i. \tag{7.38}$$

We apply the previous idea successively for $i = m, \ldots, 1$. As a result we obtain the min-max cost functional (7.5) as

$$J^0 := J(\pi(\mathcal{T})) = J_0(z_0) = \min_{x_1 \in B_1} \left(\|x_1 - F_1 z_0\|^2_{G_1^{-1}} + \max_{z_1 \in Z_1(x_1)} J_1(x_1) \right), \tag{7.39}$$

where $B_1 = \{x | \|x - x_1^*\| \leq \bar{\delta}_1\}$. We calculate the optimal control policy $\pi(\mathcal{T})$ using (7.38). The states $x_i^0(z_{i-1}),\ i = m, \ldots, 1$ in (7.38) solves problem (7.36) for $i = m, \ldots, 1$.
We can rewrite problem (7.36) as the following $(m-i)$-level min-max problem

$$J_{i-1}(z_{i-1}) = \min_{x_i \in B_i} \max_{z_i \in Z_i(x_i)} \cdots \min_{x_m \in B_m} \max_{z_m \in Z_m(x_m)} \Phi(x_s, z_s,\ s = i, \ldots, m) \tag{7.40}$$

with

$$\Phi(x_s, z_s,\ s = i, \ldots, m) = \sum_{s=i}^{m+1} \|x_s - F_s z_{s-1}\|^2_{G_s^{-1}}$$
$$x_{m+1} = x_{m+1}^*$$

and the decision variables $x_s \in \mathbb{R}^n,\ z_s \in \mathbb{R}^n,\ s = i, \ldots, m$. This multi-level problem consists of nested min-max optimization problems.
Let us consider the following problem

$$J(z_{m-1}) = \min_{x_m \in B_m} \max_{z_m \in Z_m(x_m)} \left(\|x_m - F_m z_{m-1}\|^2_{G_m^{-1}} + \|x_{m+1}^* - F_{m+1} z_m\|^2_{G_{m+1}^{-1}} \right) \tag{7.41}$$

where

$$\begin{aligned} B_m &= \{x \| \|x - x_m^*\|_{S_m}^2 \leq \bar{\delta}_m\} \\ Z_m(x_m) &= \{z \| \|z - x_m\|_{Q_m^{-1}}^2 \leq v_m\} \end{aligned}$$

and let us introduce the following variable transformation and notations

$$\begin{aligned} p &= F_{m+1}(x_m - F_m z_{m-1}) \\ x_m &= F_{m+1}^{-1} p + F_m z_{m-1} \\ \xi &= -x_{m+1}^* + F_{m+1} z_m \\ z_m &= F_{m+1}^{-1}(\xi + x_{m+1}^*) \\ G &= \left(F_{m+1} G_m F_{m+1}^T\right)^{-1} \\ D &= G_{m+1}^{-1} \\ Q &= F_{m+1} Q_m F_{m+1}^T \\ d &= x_{m+1}^* - F_{m+1} F_m z_{m-1} \\ \delta &= \delta_m, \ v = v_m \\ \bar{S} &= (F_{m+1}^T)^{-1} S_m F_{m+1}^{-1} \\ \bar{a} &= -F_{m+1} F_m z_{m-1} + F_{m+1} x_m^* = -F_{m+1}(F_m z_{m-1} - x_m^*). \end{aligned}$$

With these transformations and notations we can rewrite problem (7.41) as

$$\begin{aligned} \min_p \ & p^T G p + \max_\xi \ \xi^T D \xi \\ \text{s.t.} \ & (d + \xi - p)^T Q^{-1} (d + \xi - p) \leq v \\ & (p - \bar{a})^T \bar{S} (p - \bar{a}) \leq \delta. \end{aligned} \tag{7.42}$$

And with the introduction of one more variable transformation

$$\begin{aligned} & Q = AA^T, \ y = A^{-1}\xi, \ \phi = A^{-1}(p - \bar{a}) \\ & \mathcal{G} = A^T G A, \ \mathcal{D} = A^T D A, \ r = A^{-1}(d - \bar{a}) \\ & a = A^{-1}\bar{a}, \ S = A^T \bar{S} A \end{aligned}$$

we can rewrite problem (7.42) and hence problem (7.41) as

$$\begin{aligned} \min_\phi \ & (\phi + a)^T \mathcal{G} (\phi + a) + \max_y y^T \mathcal{D} y \\ \text{s.t.} \ & (r + y - \phi)^T (r + y - \phi) \leq v \\ & \phi^T S \phi \leq \delta. \end{aligned} \tag{7.43}$$

Here $\mathcal{G}, \ \mathcal{D}, \ S \in \mathbb{R}^{n \times n}$ symmetric positive definite, $a, \ r \in \mathbb{R}^n, \ v, \ \delta \in \mathbb{R}$ are given matrices, vectors and scalars. We make use of Theorem 5.30. Then problem (7.43) is equivalent to the following problem

$$\begin{aligned}\min_{\lambda}\max_{\alpha}\ & \left(r + (\mathcal{G}+\alpha S)^{-1}\mathcal{G}a\right)^T \left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda} + (\mathcal{G}+\alpha S)^{-1}\right)^{-1} \left(r + (\mathcal{G}+\alpha S)^{-1}\mathcal{G}a\right) \\ & - (\mathcal{G}a)^T(\mathcal{G}+\alpha S)^{-1}\mathcal{G}a + \lambda v - \alpha\delta \\ & \text{s.t. } \alpha \geq 0, \lambda \geq \lambda_{\max}(\mathcal{D})\end{aligned} \tag{7.44}$$

in the sense that if α and λ solve problem (7.44), then

$$\phi(\lambda,\alpha) = (\mathcal{G}+\alpha S)^{-1}\left(\mathcal{D}^{-1} - \frac{\mathbb{I}}{\lambda^2} + (\mathcal{G}+\alpha S)^{-1}\right)^{-1}\left(r + (\mathcal{G}+\alpha S)^{-1}\mathcal{G}a\right) - (\mathcal{G}+\alpha S)^{-1}\mathcal{G}a$$

solves problem (7.43). For the solution of (7.44) we get

$$\begin{aligned}x_m^0 =& x_m^* - \left(G_m^{-1}+\alpha S_m\right)^{-1} F_{m+1}^T\Bigg(\left(F_{m+1}^{-1}\right)^T G_m^{-1}(F_m z_{m-1} - x_m^*) + \Bigg(G_{m+1} - \frac{F_{m+1}Q_m F_{m+1}^T}{\lambda} \\ & + F_{m+1}(G_m^{-1}+\alpha S_m)^{-1}F_{m+1}^T\Bigg)^{-1}\left(x_{m+1}^* - F_{m+1}x_m^* - F_{m+1}(G_m^{-1}+\alpha S_m)^{-1}G_m^{-1}(F_m z_{m-1} - x_m^*)\right)\Bigg).\end{aligned} \tag{7.45}$$

Using (7.45) we get

Lemma 7.8. *The original problem (7.41) is equivalent to*

$$\begin{aligned}\min_{\lambda}\max_{\alpha}\ & \left\|x_{m+1}^* - F_{m+1}\left(x_m^* - (G_m^{-1}+\alpha S_m)^{-1}G_m^{-1}(F_m z_{m-1} - x_m^*)\right)\right\|^2_{\bar{G}^{-1}(\lambda,\alpha)} \\ & - \|G_m^{-1}(F_m z_{m-1} - x_m^*)\|^2_{\left(G_m^{-1}+\alpha S_m\right)^{-1}} + \lambda v_m - \alpha\bar{\delta}_m \\ s.t.\ & \lambda \geq \lambda_{\max}(N_m^T F_{m+1}^T G_{m+1}^{-1} F_{m+1} N_m) \\ & \alpha \geq 0\end{aligned}$$

where $Q_m = N_m N_m^T$ *and* $\bar{G}(\lambda,\alpha) = G_{m+1} - \frac{F_{m+1}Q_m F_{m+1}^T}{\lambda} + F_{m+1}(G_m^{-1}+\alpha S_m)^{-1}F_{m+1}^T$.

We want to investigate approximative policies. For simplicity of notations and better readability we assume that $\bar{a} = 0$. Then the problem of Lemma 7.8 reads as

$$\begin{aligned}\min_{\lambda}\max_{\alpha}\ & \|x_{m+1}^* - F_{m+1}x_m^*\|^2_{\bar{G}^{-1}(\lambda,\alpha)} + \lambda v_m - \alpha\bar{\delta}_m \\ \text{s.t. } & \lambda \geq \lambda_{\max}(N_m^T F_{m+1}^T G_{m+1}^{-1} F_{m+1} N_m) \\ & \alpha \geq 0\end{aligned}$$

Altogether we formulate recursively the following problem for the original control policy

$$\begin{aligned}J_m(z) &= \|x_{m+1}^* - F_{m+1}z\|^2_{G_{m+1}^{-1}} \\ J_{m-1}(z) &= \min_{x_m\in B_m}\max_{z_m\in Z_m(x_m)}\left(\|x_m - F_m z\|^2_{G_m^{-1}} + J_m(z)\right) \\ J_{i-1}(z) &= \min_{x_i\in B_i}\max_{z_i\in Z_i(x_i)}\left(\|x_i - F_i z\|^2_{G_m^{-1}} + J_i(z_i)\right)\end{aligned}$$

$$J_0(z) = \min_{x_1 \in B_1} \max_{z_1 \in Z_1(x_1)} \dots \min_{x_m \in B_m} \max_{z_m \in Z_m(x_m)} \Phi(z, x_1, z_1, \dots, x_m, z_m) \tag{7.46}$$

where

$$\Phi(z, x_1, z_1, \dots, x_m, z_m) = \sum_{i=1}^{m+1} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}}$$
$$x_{m+1} = x^*_{m+1}; z_0 = z.$$

As in Chapter 6 we cannot solve this problem analytically. This yields to high computational costs and hence we will now introduce approximative control policies. To construct these policies we formulate the problems I_0, V_0, W^0 and W^* similar to Chapter 6.

Problem I_0
Define for $i = m, \dots, 0$:

$$I_m(z) = J_m(z) = \|x^*_{m+1} - F_{m+1} z\|^2_{G^{-1}_{m+1}}$$
$$I_{m-1}(z) = \min_{x_m \in B_m} \max_{z_m \in Z_m(x_m)} \left(\|x_m - F_m z\|^2_{G_m^{-1}} + \|x^*_{m+1} - F_{m+1} z\|^2_{G^{-1}_{m+1}} \right) \tag{7.47}$$
$$I_0(z) = \min_{x \in B} \max_{z \in Z(x)} \left(\|x - F_0 z\|^2_{G_0^{-1}} + \|x^*_{m+1} - F_{m+1} z\|^2_{G^{-1}_{m+1}} \right) \tag{7.48}$$

With Lemma 7.8 we can rewrite problem (7.48) as

$$I_{m-1}(z) = \min_{\lambda \geq \lambda_{\max}(N_m^T F^T_{m+1} G^{-1}_{m+1} F_{m+1} N_m)} \max_{\alpha \geq 0} \|x^*_{m+1} - \underbrace{F_{m+1} F_m}_{\bar{F}_m} z\|^2_{G_m^{-1}(\lambda_m, \alpha_m)}$$
$$= \min_{\lambda \geq \mu} \max_{\alpha \geq 0} f_m(z, \lambda_m, \alpha_m),$$

where $f_m(z, \lambda_m, \alpha_m)$ is constructed as follows. We define the matrices

$$\tilde{G}_{m+1} = G_{m+1}$$
$$\tilde{G}_m(\lambda_m, \alpha_m) = G_{m+1} - \frac{F_{m+1} Q_m F^T_{m+1}}{\lambda_m} + F_{m+1}(G_m^{-1} + \alpha_m S_m)^{-1} F^T_{m+1}.$$

Using the Sherman-Morrison formula we can rewrite

$$(G_m^{-1} + \alpha_m S_m)^{-1} = G_m - \alpha_m G_m (S_m^{-1} + \alpha_m G_m)^{-1} G_m$$

and then we define

$$G_{m+1} + F_{m+1} G_m F^T_{m+1}$$
$$= \int_{t_m}^{t_{m+1}} (F(t_{m+1}, t)b)(F(t_{m+1}, t)b)^T dt + F(t_{m+1}, t_m) \int_{t_{m-1}}^{t_m} (F(t_m, t)b)(F(t_m, t)b)^T dt \; F^T(t_{m+1}, t_m)$$
$$= \int_{t_{m-1}}^{t_{m+1}} (F(t_{m+1}, t)b)(F(t_{m+1}, t)b)^T dt$$
$$= G_{*m-1}.$$

Analogously we can write for Q_m

$$\begin{aligned} F_{m+1}Q_mF_{m+1}^T &= F(t_*,t_m)\int_{t_{m-1}}^{t_m} F(t_m,t)g(F(t_m,t)g)^T dt F(t_*,t_m)^T \\ &= \int_{t_{m-1}}^{t_m} F(t_*,t)g(F(t_*,t)g)^T dt \\ &= Q_{*m} \end{aligned}$$

Hence we can rewrite $\tilde{G}_m(\lambda_m,\alpha_m)$ as

$$\tilde{G}_m(\lambda_m,\alpha_m) = G_{*m-1} - \frac{Q_{*m}}{\lambda_m} - \alpha_m F_{m+1}G_m(S_m^{-1}+\alpha_m G_m)^{-1}G_m F_{m+1}^T.$$

For the simplicity of notation we denote $H_m = H_m(\alpha_m) := G_m(S_m^{-1}+\alpha_m G_m)^{-1}G_m$. With

$$\begin{aligned} A_m &= \bar{F}_{m+1}N_m \\ \mu_m &= \lambda_{\max}(A_m^T G_{m+1}^{-1} A_m) \end{aligned}$$

we get

$$f_m(z,\lambda_m,\alpha_m) = \|x^*_{m+1} - \bar{F}_m z\|^2_{\tilde{G}_m^{-1}(\lambda_m,\alpha_m)} + \lambda_m v - \alpha_m \delta.$$

To find an expression for $I_{i-1}(z)$ we successively compute

$$\begin{aligned} I_{m-2}(z) &= \min_{\lambda_m\ge\mu_m}\max_{\alpha_m\ge 0}\min_{x_{m-1}\in B_{m-1}}\max_{z_{m-1}\in Z_{m-1}(x_{m-1})} \|x_{m-1}-F_{m-1}z\|^2_{G_{m-1}^{-1}} + f_m(z_{m-1},\lambda_m,\alpha_m) \\ &= \min_{\lambda_m\ge\mu_m}\max_{\alpha_m\ge 0}\Big(\lambda_m v - \alpha_m\delta \\ &+ \min_{x_{m-1}\in B_{m-1}}\max_{z_{m-1}\in Z_{m-1}(x_{m-1})} \|x_{m-1}-F_{m-1}z\|^2_{G_{m-1}^{-1}} + \|x^*_{m+1}-\bar{F}_m z_{m-1}\|^2_{\tilde{G}_m^{-1}(\lambda_m,\alpha_m)}\Big) \\ &= \min_{\lambda_m\ge\mu_m}\max_{\alpha_m\ge 0}\Big(\lambda_m v_m - \alpha_m\delta_m \\ &+ \min_{\lambda_{m-1}\ge\mu_{m-1}}\max_{\alpha_{m-1}\ge 0}\|x^*_{m+1}-\bar{F}_{m-1}z_{m-1}\|^2_{\tilde{G}_{m-1}^{-1}(\lambda_{m-1},\alpha_{m-1},\lambda_m,\alpha_m)} + \lambda_{m-1}v_{m-1} - \alpha_{m-1}\delta_{m-1}\Big), \end{aligned}$$

where

$$\mu_{m-1} = \mu_{m-1}(\lambda_m,\alpha_m) = \lambda_{\max}\big(\underbrace{N_{m-1}^T F_m^T}_{A_{m-1}^T}\tilde{G}_m^{-1}(\lambda_m,\alpha_m)\underbrace{F_m^T N_{m-1}}_{A_{m-1}}\big)$$

$$\tilde{G}_{m-1}(\lambda_{m-1},\alpha_{m-1},\lambda_m,\alpha_m) = G_{*m-2} - \sum_{s=m-1}^{m}\frac{Q_{*s}}{\lambda_s} - \sum_{s=m-1}^{m}\alpha_s F_{s+1}H_s(\alpha_s)F_{s+1}^T$$

$$f_{m-1}(z,\lambda_{m-1},\alpha_{m-1},\lambda_m,\alpha_m) = \|x^*_{m+1}-\bar{F}_{m-1}z\|^2_{\tilde{G}_{m-1}^{-1}(\lambda_{m-1},\alpha_{m-1},\lambda_m,\alpha_m)} + \sum_{s=m-1}^{m}\lambda_s v_s - \sum_{s=m-1}^{m}\alpha_s\delta_s. \tag{7.49}$$

Analogously we can write

$$I_{i-1}(z) = \min_{\lambda_m \geq \mu_m} \max_{\alpha_m \geq 0} \ldots \min_{\lambda_{i+1} \geq \mu_{i+1}} \max_{\alpha_{i+1} \geq 0} \min_{x_i \in B_i} \max_{z_i \in Z_i(x_i)} \Big(\|x_i - F_i z\|^2_{G_i^{-1}} + f_{i+1}(z_i, \lambda_{i+1}, \ldots, \lambda_m, \alpha_{i+1}, \ldots, \alpha_m) \Big). \tag{7.50}$$

Problem V^0

Problem V^0 is formulated as

$$V^0(z_0) = \min_{\substack{x_i \in B_i \\ i=1,\ldots,m}} \max_{\substack{z_i \in Z_i(x_i) \\ i=1,\ldots,m}} \sum_{s=1}^{m} \|x_s - F_s z_{s-1}\|^2_{G_s^{-1}} \tag{7.51}$$

To apply Lemma 7.8 to problem (7.51) we rewrite the problem in the following form

$$V^0(z_0) = \min_{x_m \in B_m} \ldots \min_{x_2 \in B_2} \left(S_1(x_2) + \max_{z_2 \in Z_2(x_2)} \ldots \max_{z_m \in Z_m(x_m)} \sum_{i=3}^{m+2} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} \right), \tag{7.52}$$

with

$$S_1(x_2) = \min_{x_1 \in B_1} \max_{z_1 \in Z_1(x_1)} \|x_1 - F_1 z_0\|^2_{G_1^{-1}} + \|x_2 - F_2 z_1\|^2_{G_2^{-1}}. \tag{7.53}$$

If we now apply Lemma 7.8 (where $\bar{a} = 0$) we get

$$S_1(x_2) = \min_{\lambda_1 \geq \bar{\mu}_1} \max_{\alpha_1 \geq 0} \|x_2 - F_2 a_1\|_{\bar{G}_2^{-1}(\lambda_1, \alpha_1)} + \lambda_1 v_1 - \alpha_1 \delta_1$$

where

$$\begin{aligned} \bar{\mu}_1 &= \lambda_{\max}(N_1^T F_2^T G_2^{-1} F_2 N_1),\ a_1 = F_1 z_0 \\ \bar{G}_2(\lambda_1, \alpha_1) &= G_2 - \frac{F_2 Q_1 F_2^T}{\lambda_1} + F_2 (G_1^{-1} + \alpha_1 S_1)^{-1} F_2^T \\ &= G(t_2, t_0) - \frac{Q_1^2}{\lambda_1} - \alpha_1 F_2 H_1(\alpha_1) F_2^T. \end{aligned}$$

Hence, if we now substitute problem (7.53) into problem (7.52) we get

$$V^0 = \min_{\lambda_1 \geq \mu_1} \max_{\alpha_1 \geq 0} \min_{x_m \in B_m} \ldots \min_{x_3 \in B_3} \Bigg(S_2(x_3, \lambda_1, \alpha_1) + \lambda_1 v_1 - \alpha_1 \delta_1 \tag{7.54}$$

$$+ \max_{z_3 \in Z_x(x_3)} \ldots \max_{z_m \in Z_m(x_m)} \sum_{i=4}^{m+1} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} \Bigg) \tag{7.55}$$

with

$$S_2(x_3, \lambda_1, \alpha_1) = \min_{x_2 \in B_2} \max_{z_2 \in Z_2(x_2)} \|x_2 - F_2 a_1\|_{\bar{G}_2^{-1}(\lambda_1, \alpha_1)} + \|x_3 - F_3 z_2\|^2_{G_3^{-1}}$$

Applying Lemma 7.8 to the latter problems yields

$$S_2(x_3,\lambda_1,\alpha_1) = \min_{\lambda_2\geq\bar{\mu}_2}\max_{\alpha_2\geq 0}\|x_3 - F_3a_2\|^2_{\bar{G}_2^{-1}(\lambda_1,\lambda_2,\alpha_1,\alpha_2)} + \lambda_2 v_2 - \alpha_2\delta_2, \tag{7.56}$$

where

$$\begin{aligned}\bar{\mu}_2 &= \lambda_{\max}(N_2^TF_3^TG_3^{-1}F_3N_2),\ a_2 = F_2a_1\\ \bar{G}_3^{-1}(\lambda_1,\lambda_2,\alpha_1,\alpha_2) &= G_3 - \frac{F_3Q_2F_3^T}{\lambda_2} + F_2(\bar{G}_2^{-1}(\lambda_1,\alpha_1)+\alpha_2S_2)^{-1}F_3^T\\ &= G(t_3,t_0) - \frac{Q_1^3}{\lambda_1} - \alpha_1F_2H_1F_2^T - \frac{Q_2^3}{\lambda_2} - \alpha_2F_3H_2F_3^T.\end{aligned}$$

If we now substitute (7.56) into (7.54) and perform recursively $m-1$ times the described application of Lemma 7.8 we get the following formulation of V^0

$$V^0 = \min_{\lambda_1\geq\bar{\mu}_1}\max_{\alpha_1\geq 0}\min_{\lambda_2\geq\bar{\mu}_2}\max_{\alpha_2\geq 0}\ldots\min_{\lambda_m\geq\bar{\mu}_m}\max_{\alpha_m\geq 0} f(\lambda_1,\ldots,\lambda_m,\alpha_1,\ldots\alpha_m)$$

with

$$f(\lambda_1,\ldots,\lambda_m,\alpha_1,\ldots,\alpha_m) = \|x^*_{m+1} - F_{m+1}a_m\|^2_{\bar{G}^{-1}_{m+1}(\lambda_1,\ldots,\lambda_m,\alpha_1,\ldots,\alpha_m)} + \sum_{i=1}^m \lambda_i v_i - \sum_{i=1}^m \alpha_i\delta_i$$

$$a_m = F(t_m,t_0)z_0,$$

$$\bar{G}_{m+1}(\lambda_1,\ldots,\lambda_m,\alpha_1,\ldots,\alpha_m) = G(t_*,t_0) - \sum_{i=1}^m\frac{Q_i^{m+1}}{\lambda_i} - \sum_{i=1}^m\alpha_iF_{i+1}H_iF_{i+1}^T.$$

This means, that we transformed the two-level min-max problem (7.51) in the variables $x_i\in\mathbb{R}^n,\ i=1,\ldots,m$ at the upper level and $z_i\in\mathbb{R}^n, i=1,\ldots,m$ in the lower level to the min-max problem in the m variables $\lambda=(\lambda_1,\ldots,\lambda_m)^T$ and $\alpha=(\alpha_1,\ldots,\alpha_m)^T$. We can write this problem as

$$V^0 = \min_{\lambda\geq\bar{\mu}}\max_{\alpha\geq 0} f(\lambda,\alpha) \tag{7.57}$$

with $f(\lambda,\alpha) := d^T\bar{G}^{-1}_{m+1}(\lambda,\alpha)d + v^T\lambda - \delta^T\alpha$, the vectors $\bar{\mu}=(\bar{\mu}_1,\ldots,\bar{\mu}_m)^T$, $v=(v_1,\ldots,v_m)^T$,
$\delta=(\delta_1,\ldots,\delta_m)^T$, $d=d(z_0,x^*_{m+1}) = x^*_{m+1} - F(t_*,t_0)z_0$ being given.

Problem W^0

Let us consider the following problem

$$\begin{aligned} W^0 := \min_\lambda\max_\alpha\ & f(\lambda,\alpha)\\ \text{s.t. } & \lambda_i \geq \mu_i(\lambda_{i+1},\ldots,\lambda_m,\alpha_{i+1},\ldots,\alpha_m),\ i=1,\ldots,m\\ & \alpha\geq 0.\end{aligned} \tag{7.58}$$

To analyze this problem we need the following reformulations

$$J^0 = \min_{x_1 \in B_1} \max_{z_1 \in Z_1(x_1)} \dots \min_{x_{m-1} \in B_{m-1}} \max_{z_{m-1} \in Z_{m-1}(x_{m-1})} \left(\sum_{i=1}^{m-1} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} + P_{m-1}(z_{m-1}) \right) \tag{7.59}$$

where

$$P_{m-1}(z_{m-1}) = \min_{x_m \in B_m} \max_{z_m \in Z_m(x_m)} \left(\|x_m - F_m z_{m-1}\|^2_{G_m^{-1}} + \|x^*_{m+1} - F_{m+1} z_m\|^2_{G_{m+1}^{-1}} \right).$$

Using Lemma 7.8 we can write

$$P_{m-1}(z_{m-1}) = \min_{\lambda_m \geq \mu_m} \max_{\alpha \geq 0} \left(\|x^* - \bar{F}_m z_{m-1}\|^2_{G_m^{-1}(\lambda_m, \alpha_m)} + \lambda_m v_m - \alpha_m \delta_m \right) \tag{7.60}$$

where $G_m(\lambda_m, \alpha_m) = \tilde{G}_m = \tilde{G}_m(\lambda_m, \alpha_m)$ If we now substitute (7.60) into (7.59)we get an expression for J^0

$$J^0 = \min_{x_1 \in B_1} \max_{z_1 \in Z_1(x_1)} \dots \min_{x_{m-2} \in B_{m-2}} \max_{z_{m-2} \in Z_{m-2}(x_{m-2})} \left(\sum_{i=1}^{m-2} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} + P_{m-2}(z_{m-2}) \right), \tag{7.61}$$

where

$$P_{m-2}(z_{m-2}) := \min_{x_{m-1} \in B_{m-1}} \max_{z_{m-1} \in Z_{m-1}(x_{m-1})} \min_{\lambda_m \geq \mu_m} \max_{\alpha_m \geq 0}$$
$$\left(\|x_{m-1} - F_{m-1} z_{m-2}\|^2_{G_{m-1}^{-1}} + \|x^*_{m+1} - \bar{F}_m z_{m-1}\|^2_{\tilde{G}_m^{-1}(\lambda, \alpha)} + \lambda_m v_m - \alpha_m \delta_m \right)$$

With Lemma 7.8 we can conclude that

$$\begin{aligned} P_{m-2}(z_{m-2}) \leq I_{m-2}(z_{m-2}) &:= \min_{\lambda_m \geq \mu_m} \max_{\alpha_m \geq 0} \min_{x_{m-1} \in B_{m-1}} \max_{z_{m-1} \in Z_{m-1}(x_{m-1})} \\ &\left(\|x_{m-1} - F_{m-1} z_{m-2}\|^2_{G_{m-1}^{-1}} + \|x^*_{m+1} - \bar{F}_m z_{m-1}\|^2_{\tilde{G}_m^{-1}(\lambda, \alpha)} + \lambda_m v_m - \alpha_m \delta_m \right) \\ &=: I_{m-2}(z_{m-2}) \qquad (7.62) \\ &= \min_{\lambda_m \geq \mu_m} \max_{\alpha_m \geq 0} \min_{\lambda_{m-1} \geq \mu_{m-1}(\lambda_m)} \max_{\alpha_{m-1} \geq 0} \\ &\left(\|x^*_{m+1} - \bar{F}_{m-1} z_{m-2}\|^2_{\tilde{G}_{m-1}^{-1}(\lambda_{m-1}, \lambda_m, \alpha_{m-1}, \alpha_m)} + \sum_{s=m-1}^{m} \lambda_s v_s - \sum_{s=m-1}^{m} \alpha_s \delta_s \right), \qquad (7.63) \end{aligned}$$

where μ_{m-1} and $\tilde{G}_{m-1}$ are defined in (7.49). Using (7.61) and (7.63) we get

$$\begin{aligned} J^0 \leq \min_{x_1 \in B_1} \max_{z_1 \in Z_1(x_1)} &\dots \min_{x_{m-3} \in B_{m-3}} \max_{z_{m-3} \in Z_{m-3}(x_{m-3}} \\ &\left(\sum_{i=1}^{m-3} \|x_i - F_i z_{i-1}\|^2_{G_i^{-1}} + P_{m-3}(z_{m-3}) \right) \end{aligned} \tag{7.64}$$

where

$$P_{m-3}(z_{m-3}) := \min_{x_{m-2} \in B_{m-2}} \max_{z_{m-2} \in Z_{m-2}(x_{m-2})} \min_{\lambda_m \geq \mu_m} \max_{\alpha_m \geq 0} \min_{\lambda_{m-1} \geq \mu_{m-1}(\cdot)} \max_{\alpha_{m-1} \geq 0}$$

$$\Bigg(\|x_{m-2} - F_{m-2}z_{m-3}\|^2_{G^{-1}_{m-2}} + \|x^*_{m+1} - \bar{F}_{m-1}z_{m-2}\|_{\tilde{G}^{-1}_{m-1}(\lambda_{m-1},\lambda_m,\alpha_{m-1},\alpha_m)} + \sum_{s=m-1}^{m} \lambda_s v_s - \sum_{s=m-1}^{m} \alpha_s \delta_s\Bigg).$$

If we again use Lemma 7.8 we obtain

$$P_{m-3}(z_{m-3}) \le I_{m-3}(z_{m-3}) := \min_{\lambda_m \ge \mu_m} \max_{\alpha_m} \min_{\lambda_{m-1}\ge\mu_{m-1}(\cdot)} \max_{\alpha_{m-1}} \min_{\lambda_{m-2}\ge\mu_{m-2}(\cdot)} \max_{\alpha_{m-2}\ge 0} \Bigg(\|x^*_{m+1} - \bar{F}_{m-2}z_{m-3}\|_{\tilde{G}^{-1}_{m-2}(\lambda_{m-2},\lambda_{m-1},\lambda_m,\alpha_{m-2},\alpha_{m-1},\alpha_m)} + \sum_{s=m-2}^{m} \lambda_s v_s - \sum_{s=m-2}^{m} \alpha_s\delta_s\Bigg). \tag{7.65}$$

By substituting (7.65) into (7.64) and continuing the described operations recursively we get the inequality

$$J^0 \le P_0(z_0)$$

with

$$P_0(z_0) = \min_{\lambda_m\ge\mu_m} \max_{\alpha_m\ge 0} \min_{\lambda_{m-1}\ge\mu_{m-1}} \max_{\alpha_{m-1}\ge 0} \dots \min_{\lambda_1\ge\mu_1(\cdot)} \max_{\alpha_1\ge 0} \Bigg(\|x^*_{m+1} - \bar{F}_1 z_0\|_{\tilde{G}^{-1}_1(\lambda_1,\dots,\lambda_m,\alpha_1,\dots,\alpha_m)} + \sum_{s=1}^{m}\lambda_s v_s - \sum_{s=1}^{m}\alpha_s\delta_s\Bigg). \tag{7.66}$$

This means that the optimal value of the cost functional in problem (7.66) may be taken as an approximation for the optimal value of the cost function in the original problem (7.46). By analyzing the relations (7.49) we can easily verify that

$$\|x^*_{m+1} - \bar{F}_1 z_0\|_{\tilde{G}^{-1}_1(\lambda_1,\dots,\lambda_m,\alpha_1,\dots,\alpha_m)} + \sum_{s=1}^{m}\lambda_s v_s - \sum_{s=1}^{m}\alpha_s\delta_s = f(\lambda,\alpha)$$

where $f(\lambda, \alpha)$ is defined as in (7.57). Therefore we can rewrite problem (7.66) as W^0 (7.58).

Problem W^*

Problem W^* is formulated as

$$\begin{aligned} W^* := \min_{\lambda\in\mathbb{R}^m} \max_{\alpha\ge 0} & f(\lambda,\alpha) \\ \text{s.t. } & \lambda_i \ge \mu^*_i, \\ & \alpha_i \ge 0,\ i = 1,\dots,m. \end{aligned} \tag{7.67}$$

where

$$\begin{aligned}
&\mu_m^* = \mu_m = \lambda_{\max}(A_m^T G_{m+1} A_m) \\
&\mu_i^* = \mu_i(\mu_{i+1}^*, \ldots, \mu_m^*, 0, \ldots, 0) = \lambda_{\max}(A_i^T \tilde{G}_{i+1}^{-1}(\mu_{i+1}^*, \ldots, \mu_m^*, 0, \ldots,)A_i) \\
&\tilde{G}_{m+1} = G_{m+1} \\
&\tilde{G}_i(\lambda_i, \ldots, \lambda_m, 0, \ldots, 0) = G_{i-1} - \sum_{s=i}^{m} \frac{Q_s}{\lambda_s}.
\end{aligned} \tag{7.68}$$

Following the rules of Section 6.4 we can now construct the different approximative control policies based on the solution of the problems presented in this section.

For example we can formulate the following algorithm for the approximative control policy $\pi(\lambda^0, \alpha^0)$.

Algorithm 7.9. (Determination of control laws of approximative policy $\pi(\lambda^0, \alpha^0)$)

`GIVEN` Initial state $z_0 = z(t_0)$

`COMPUTE` the matrices F_i, $\bar{F}_i$, G_i, G_{*i}, Q_{*i}, N_i for $i = 1, \ldots, m+1$

`FORMULATE and SOLVE` (offline) the problem

$$\begin{aligned}
\min_{\lambda \in \mathbb{R}^m} \max_{\alpha \in \mathbb{R}^m} \;& f(z_0, \lambda, \alpha) \\
\text{s.t. } \;& \lambda_i \geq \mu_i(\lambda_{i+1}, \ldots, \lambda_m, \alpha_{i+1}, \ldots, \alpha_m) \\
& \alpha \geq 0, \; i = 1, \ldots, m
\end{aligned} \tag{7.69}$$

`SOLUTION` of problem (7.69): $\lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0)^T$, $\alpha^0 = (\alpha_1^0, \ldots, \alpha_m^0)^T$

`FOR` $i = 1, \ldots, m$

`GIVEN` z_{i-1}

`SOLVE` problem

$$\min_{\lambda_i \geq \mu_i} \max_{\alpha_i \geq 0} \; \|x_{m+1}^* - \bar{F}_i z_{i-1}\|_{\tilde{G}_{m+1}^{-1}(\lambda_i, \lambda_{i+1}^0, \ldots, \lambda_m^0, \alpha_i, \lambda_{i+1}^0, \ldots, \alpha_m^0)} + \lambda_i v_i - \alpha_i \delta_i \tag{7.70}$$

`SOLUTION` of problem (7.70): λ_i^0, α_i^0

`DETERMINE` the control law

$$u_i(t; z_{i-1}) = (x_* - \bar{F}_i z_{i-1})^T \tilde{G}_i^{-1}(\lambda_i^0, \lambda_{i+1}^0, \ldots, \lambda_m^0, \alpha_i^0, \alpha_{i+1}^0, \ldots, \alpha_m^0) F(t_*, t) b$$

`END`

`DENOTE` the control policy composed of the control laws by $\pi(\lambda^0, \alpha^0)$

Remark 7.10. *The solution of problem (7.70) can be computed using Algorithm 5.31.*

7.3 Bounded Control

As we stated at the beginning of this chapter, we can also generalize the strategies from Section 6.4 for problems with additional constraints on the control. For this we now assume the following problem

Construct a control policy

$$\pi = (u_i(\cdot; z_{i-1}),\ i = 1, \dots, m+1)$$

consisting of control laws

$$u_i(\cdot; z_{i-1}) = (u_i(t; z_{i-1}), t \in T_i),\ z_{i-1} \in \mathbb{R}^n,\ i = 1, \dots, m+1$$

for each interval $T_i,\ i = 1, \dots m+1$, such that for any admissible disturbance $w(\cdot)$ the trajectory

$$z(t) = z(t; \pi, w(\cdot)),\ t \in T$$

of the system

$$\dot{z}(t) = Az(t) + bu_i(t; z(t_{i-1})) + gw(t),\ t \in T_i,\ i = 1, \dots, m+1,\ z(0) = z_0$$

and the control laws satisfy

$$\int\limits_{T_i} u_i^2(t; z(t_{i-1}; \pi, w(\cdot)))dt \le r_i,\ i = 1, \dots, m+1$$

and the cost functional

$$J(\pi) = \max_{w(\cdot) \in \bar{W}} \sum_{i=1}^{m+1} (z(t_i; \pi, w(\cdot)) - x_i^*)^T D_i (z(t_i; \pi, w(\cdot)) - x_i^*)$$

takes the minimal value

$$\min_{\pi} J(\pi)$$

with given matrices D_i, given vectors $x_i^* \in \mathbb{R}^n$ and given numbers $r_i \in \mathbb{R}$, $i = 1, \dots, m+1$.

The policies and problems $J^0,\ I_0,\ P_0,\ V^0,\ W^0$ and W^* can be formulated in a similar way as in Section 7.2 with the difference that B_i is defined as

$$B_i(z) = \{x \in \mathbb{R}^n : \|x - F_i z\|^2_{G_i^{-1}} \le r_i\}.$$

7.4 Disturbances Bounded in Norm

In this section we will consider another generalization. We will analyze a system in which the disturbance is bounded in a norm. Before, we used the class of admissible disturbances (6.57). In the case of class (6.57) the number of correction points and their values $t_i,\ i = 1, \ldots, m$ are determined by the class (6.57). Now, we will construct a class of admissible disturbances with which we are able to choose the number of correction points arbitrarily. The set

$$\bar{W} = \{\bar{w}(\cdot) \in L_1(t) : |\bar{w}(t)| \le \alpha,\ t \in T\} \tag{7.71}$$

with α being a given number is a subset of the class of admissible disturbances (6.57) with an arbitrary choice of moments $t_i,\ i = 1, \ldots,$ and the choice of numbers v_i as following

$$v_i = \alpha^2(t_i - t_{i-1}),\ i = 1, \ldots, m+1.$$

With the choice of (7.71) as class of admissible disturbances we now have the possibility to arbitrarily choose the correction points $t_i,\ i = 1, \ldots, m$. The larger the number of correction points, which means the smaller $\max(t_{i+1} - t_i),\ i = 0, \ldots, m$, the better the original class of admissible disturbances (6.57) approximates the set (7.71). Therefore we can apply the results from the previous sections. From Kurzhanski and Varaiya [50] we can deduce that for the reachability set

$$X_i := \left\{ x \in \mathbb{R}^n | x = \int\limits_{t_{i-1}}^{t_i} F(t_i, t) g \bar{w}(t) dt, |\bar{w}(t)| \le \alpha \right\}$$

there exists a matrix $\bar{N}_i \in \mathbb{R}^{n \times r}$ and a number $\bar{v}_i > 0$ such that

$$X_i \subset M_i := \{x \in \mathbb{R}^n : x = \bar{N}_i f, f^T f \le \bar{v}_i\}$$

and the ellipsoidal is minimal. The minimality of M_i and constructive rules of computing the matrix $\bar{N}_i$ and $\bar{v}_i$ are described in Kurzhanski & Varaiya [50]. Hence, we may choose arbitrary correction points $t_i,\ i = 1, \ldots, m$ and compute the matrices $\bar{N}_i$ and the numbers $\bar{v}_i$, $i = 1, \ldots, m+1$ by the rules in [50]. Then, we use the strategies described before in Chapter 6 where we change $Z_i(x),\ i = 1, \ldots m+1$ (6.70) by

$$\bar{Z}_i(x) := \{\text{there exist } f_i; z = x + \bar{N}_i f_i,\ f_i^T f_i \le \bar{v}_i\},\ i = 1, \ldots m+1.$$

Let us show how the problem $I_0(z)$ changes in this case. Consider the problem

$$\min_{x_m} \|x_m - F_m z_{m-1}\|^2_{G_m^{-1}} + \max_{z_m \in Z_m(x_m)} \|x^*_{m+1} - F_{m+1} z_m\|^2_{G_{m+1}} \tag{7.72}$$

where $Z_m(x_m) = \{z | z = x_m + N_m f_m,\ f_m^T f_m \le v_m\}$ and $N_m \in \mathbb{R}^{n \times r}$ and let us introduce the following variable transformation and notations

$$\begin{aligned}
p &= F_{m+1}(x_m - F_m z_{m+1}) \\
x_m &= F_{m+1}^{-1} p + F_m z_{m+1} \\
\xi &= -x^*_{m+1} + F_{m+1} z_m \\
z_m &= F_{m+1}^{-1}(\xi + x^*_{m+1} \\
G &= (F_{m+1} G_m F_{m+1}^T)^{-1} \\
D &= G_{m+1}^{-1}
\end{aligned}$$

$$
\begin{aligned}
d &= x^*_{m+1} - F_{m+1}F_m z_{m-1}\\
\delta &= \delta_m,\ v = v_m\\
A &= F_{m+1}N_m,\ f = f_m.
\end{aligned}
$$

With these notations we can rewrite problem (7.72) to the following bilevel optimization problem

$$
\begin{aligned}
&\min_p\ p^TGp + \max_{\xi,f}\ \xi^TD\xi\\
&\text{s.t. } d + \xi - p - Af = 0,\ f^Tf \le v.
\end{aligned}
$$

Let us first consider the lower level problem

$$
\begin{aligned}
&\max_{\xi,f}\ \xi^TD\xi\\
&\text{s.t. } d + \xi - p - Af = 0,\ f^Tf \le v.
\end{aligned}
\tag{7.73}
$$

Then we can formulate the following lemma

Lemma 7.11. *The vectors $\xi,\ f \in \mathbb{R}^n$ are optimal in problem (7.73) if and only if there exists a number $\lambda \ge \lambda_{\max}$ such that $\lambda,\ \xi,\ f$ satisfy*

$$
\begin{aligned}
(-A^TDA + \lambda\mathbb{I})f - A^TD(p-d) &= 0\\
f^Tf &= v\\
d + \xi - p - Af &= 0
\end{aligned}
$$

where $\lambda_{\max}$ denotes the maximal eigenvalue of the matrix A^TDA.

Denote $p(\lambda) = G^{-1}\left(D^{-1} + G^{-1} - \frac{AA^T}{\lambda}\right)^{-1} d$. Similarly to Chapter 5 we can prove the following lemmas.

Lemma 7.12. *If $p^* = p(\lambda_{\max})$ satisfies the condition*

$$\|Gp^*\|^2_{AA^T} \le \lambda^2_{\max}v$$

then p^* is optimal in the bilevel optimization problem (7.72).

Lemma 7.13. *If the vector p^* satisfies $\|Gp^*\|^2_{AA^T} \ge \lambda^2_{\max}v$ then there exists a unique number $\lambda^* \ge \lambda_{\max}$ such that it holds that*

$$\|Gp(\lambda)\|^2_{AA^T} = \lambda^2_{\max}v$$

and the vector $p(\lambda^*)$ solves the bilevel optimization problem (7.72).

The following theorem follows directly from Lemmas 7.12 and 7.13.

Theorem 7.14. *Problem 7.72 is equivalent to the following optimization problem*

$$
\begin{aligned}
&g(\lambda) = \min_\lambda\ d^T\left(D^{-1} + G^{-1} - \frac{AA^T}{\lambda}\right)^{-1} d + \lambda v\\
&s.t.\ \lambda \ge \lambda_{\max}(A^TDA)
\end{aligned}
\tag{7.74}
$$

in the sense that if λ^0 *is optimal in problem (7.74) then*

$$p^0 = G^{-1}\left(D^{-1} + G^{-1} - \frac{AA^T}{\lambda^0}\right)^{-1} d$$

is optimal in problem (7.72).

8 Conclusion

In this thesis we considered optimal control problems of dynamic systems with unknown but bounded uncertainties. For solving these problems the uncertainties had to be taken into account and for this, usually two different approaches are suggested. On the one hand feedback controls, where the nominal optimal control is updated as soon as the actual state and parameter estimates are available and on the other hand robust optimization, for example worst-case optimization, where it is searched for an optimal solution that is good for all possible realizations of uncertain parameters. We were interested in using a combination of feedback control and robust optimization.

Below, we will summarize the major contents of this thesis and give an outlook on future work.

After a general introduction to the motivation and contributions of this thesis in Chapter 1, we started in Chapter 2 with an overview of convex and bilevel optimization problems. We discussed the optimality conditions of convex optimization problems and presented a summary of solution methods. In Section 2.2, we described different types of bilevel optimization problems and also gave an overview of solution methods. In Chapter 3 we introduced the basic ideas of optimal control problems. For the continuous optimal control problems we formulated Pontryagin's Maximum Principle and showed the application of it on linear-quadratic optimal control problems. We also shortly introduced discrete optimal control problems and corresponding solution methods. In Section 3.5 we discussed robust optimization. We stated types of uncertainties and three different approaches to handle optimal control problems under uncertainties. Here, the third approach, the closed-loop min-max optimal feedback control approach, was very important, since this was the basic approach, we used throughout this thesis. We also presented a literature overview of robust optimal control. These two chapters were the basic preliminaries for the rest of the thesis.

In Chapter 4 we introduced the theory of Model Predictive Control. We also discussed the concepts of Dynamic Programming and in this context we formulated Bellmann's Principle of Optimality. We compared both approaches and also presented them in the case of uncertainties. In the case of uncertainties we also discussed the theory of Feedback Model Predictive Control and the theory of tubes.

For describing approximative control policies for optimal control that are described in Chapter 6 we needed special bilevel optimization problems. For this we introduced and analyzed three different types of special bilevel optimization problems in Chapter 5. For all three problems we formulated and proved the optimality conditions, presented an explicit solution and showed how to construct the optimal solution of a certain feedback linear quadratic optimal control problem using the solution of the bilevel optimization problem. These bilevel problems can be solved explicitly and offline. We also presented practical and efficient algorithms to solve these problems.

In Chapter 6 we considered the closed-loop min-max optimal feedback control. As the computation of such a control is rather difficult and slow and the costs are often too expensive we suggested to solve this kind of problem successively by dividing the time interval and computing the controls of these smaller time intervals at intermediate time points. We started by discussing this strategy with one correction moment in more detail. As application problem we used a linear quadratic optimal control problem with an additive unknown but bounded uncertainty. We used the special bilevel programming problems of Chapter 5 to rewrite this optimal control problem and formulated and discussed a practical algorithm for solving it. We also presented closed-loop min-max optimal feedback control with more than one correction points. We were able to show that we can guarantee for all admissible disturbances that the terminal state lies in a given prescribed neighborhood of a given state at the given final moment and that the value of the cost functional does not exceed a given estimate. As the computational costs of this algorithm are really high we formulated different approximative approaches which are suboptimal but can be implemented online. We compared these different approaches theoretically and in a numerical example.

In Chapter 7 we discussed three different generalizations of the approximative policies that we presented in Chapter 6 using the newly introduced special bilevel optimization problems of Chapter 5. The first generalization was a strategy in which we can control dynamic systems with pointwise state constraints. For this generalization we derived two different approaches. Afterwards we introduced a strategy for optimal control problems with bounded controls and the last generalization was a strategy in which the disturbances was bounded in a norm. For these generalizations we analyzed the problems and discussed corresponding algorithms.

In future work it is desirable to develop the approximative control policies that were shown in this thesis for nonlinear systems and to include efficient methods for solving the convex/saddle point problem and the problem with bounded control.

List of Figures

List of Tables

Bibliography

[1] Alizadeh, S.M., Marcotte, P., Savard, G.: *Two-stage stochastic bilevel programming over a transportation network*. Transportation Research Part B: Methodological, Elsevier, vol. 58(C), pages 92-105, 2013.

[2] Aubin, J.P.: *Viability theory*. Systems & Control: Foundations & Applications. Birkhäuser, Boston, Basel, Berlin, 1991.

[3] Avriel, M.: *Nonlinear Programming: Analysis and Methods*. Dover Publications, Inc., Mineola, New York, 2003.

[4] Bard, J.F.: *Some Properties of the Bilevel Programming Problem*. Technical Note, Journal of Optimization Theory and Applications. Vol. 68, No. 2, February 1991.

[5] Bard, J.F., Falk, J.E.: *An Explicit Solution to the Multi-Level Programming Problem*. Comput. & Ops. Res. Vol 9, No. 1, pp. 77-100, 1982.

[6] Bellman, R., Kalaba, R.: *Dynamic Programming and Modern Control Theory*. Academic Press Inc., New York and London, 1965.

[7] Bemporad, A., Borrelli, F., Morari, M.: *Min-Max Control of Constrained Uncertain Discrete-Time Linear Systems*. Technical Notes and Correspondence, IEEE Transactions on Automatic Control, Vol. 48, No.9, 2003.

[8] Ben-Ayed, O., Blair, C.E.: *Computational Difficulties of Bilevel Linear Programming*. Operations Research, Vol. 38, No. 3, pp. 556-560, 1990.

[9] Ben-Tal, A., Nemirovski, A.: *Lectures on Modern Convex Optimization: Analysis, Algorithms, Engineering Applications*. MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.

[10] Bertsekas, D.P.: *Convex Optimizataion Theory*. Athena Scientific, 2009.

[11] Bertsekas, D.P., Rhodes, I.B.: *Recursive State Estimation for a Set-Membership Description of Uncertainty*. IEEE Transactions on Automatic Control, Vol. AC-16, No. 2, pp.117-128, 1971.

[12] Bertsekas, D.P., Rhodes, I.B.: *On the Minimax Reachability of Target Sets and Target Tubes*. Automatica, Vol 7, No. 2, pp.233-247, 1971.

[13] Betts, J.T.: *Practical Methods for Optimal Control Using Nonlinear Programming*. Society for Industrial and Applied Mathematics, 2001.

[14] Bock, H.G.: *Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen*. Bonner Mathematische Schriften 183 (1987).

[15] Bock, H.G.: *Numerical Solution of Nonlinear Multipoint Boundary Value Problems with Application to Optimal Control.* Zeitschrift für Angewandte Mathematik und Mechanik Vol. 58, p.407, 1978

[16] Bock, H.G., Plitt, K.J.: *A multiple shooting algorithm for direct solution of optimal control problems.* Proceedings of the 9th IFAC World Congress, Budapest, 1984.

[17] Boyd, S., Vandenberghe, L.: *Convex Optimization.* Cambridge University Press, Cambridge 2004. Seventh printing with corrections 2009.

[18] Bracken, J., McGill, J.M.: *Mathematical Programs with Optimization Problems in the Constraints.* Operations Research 21, 37-44, 1973

[19] Bracken, J., McGill, J.M.: *A Method for Solving Mathematical Programs with Nonlinear Programs in the Constraints.* Operations Research 22, 1097-1101

[20] Bryson, A.E., Ho, Y.-C.: *Applied Optimal Control.* Blaisdell Publishing Company, 1969.

[21] Campo, P.J., Morari, M.: *Robust Model Predictive Control.* 1987 American Control Conference, Minneapolis, MN, USA, pp. 1021-1026, 1987. DOI: 10.23919/ACC.1987.4789462

[22] Chachuat, B.C.: *Nonlinear and Dynamic Optimization: From Theory to Practice* Laboratoire d'Automatique, Ecole Polytechnique Federale de Lausanne, 2007.

[23] Chisci, L., Rossiter, J.A., Zappa, G.: *Systems with persistent disturbances: predictive control with restricted constraints.* Automatica, Vol. 37, No. 7, pp. 1019-1028, 2001.

[24] Clarke, F.H.: *The Maximum Principle under Minimal Hypotheses.* SIAM J. Control and Optimization, Vol. 14, No. 6, 1976.

[25] Colson, B., Marcotte, P., Savard, G.: *An overview of bilevel optimization.* Springer Science + Business Media, Inc. 2007

[26] Colson, B., Marcotte, P., Savard, G.: *A Trust-Region Method for Nonlinear Bilevel Programming: Algorithm and Computational Experience.* in: Computational Optimization and Applications, 30, pp. 211-227, Springer Science + Business Media, Inc. 2005.

[27] Dempe, S.: *A Simple Algorithm for the Linear Bilevel Programming Problem.* Optimization, Vol. 18 No.3, pp. 373-385, 1987.

[28] Dempe, S.: *An Implicit Function Approach to Bilevel Programming Problems.* In: Migdalas A., Pardalos P.M., Värbrand P. (eds) Multilevel Optimization: Algorithms and Applications. Nonconvex Optimization and Its Applications, vol 20. Springer, Boston, MA, 1998.

[29] Dempe,S.: *Foundations of Bilevel Programming.* Nonconvex Optimization and Its Applications, Volume 61, Kluwer Academic Publishers, 2002.

[30] Dempe, S.: *Bilevel Programming - A Survey.* Preprint TU Bergakademie Freiberg No. 2003-11, Faculty of Mathematics and Computer Science, 2003.

[31] Dempe, S., Luo, G., Franke, S.: *Pessimistic bilevel linear optimization*. TU Bergakademie Freiberg, Fakultät für Mathematik und Informatik, Preprint 2016-05.

[32] Dempe, S., Zemkoho, A.B.: *The bilevel programming problem: reformulations, constraint qualifications and optimality conditions*. Mathematical Programming, Series A (2013) 138:447-473.

[33] Diehl, M., Bock, H.G., Kostina, E. *An approximation technique for robust nonlinear optimization*. Math. Program., Ser. B 107, pp. 213-230, 2006.

[34] Genceli, H., Nikolaou, M.: *Robust stabilty analysis of constrained l_1-norm model predictive control*. AIChE Journal, Vol. 39, Issue 12, pp. 1954-1965, 1993.

[35] Gerdts, M.: *Solving mixed-integer optimal control problems by branch&bound: a case study from automobile test-driving with gear shift*. Optimal Control Applications and Methods 26: 1-18, 2005. Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/oca.751.

[36] Goodwin, G.C., Ninness, B., Salgado, M.E.: *Quantification of Uncertainty in Estimation*. 1990 American Control Conference, San Diego, CA, USA, pp. 2400-2405, 1990. DOI: 10.23919/ACC.1990.4791157

[37] Goulart, P.J., Kerrigan, E.C., Maciejowski, J.M.: *Optimization over state feedback policies for robust control with constraints*. Automatica, Vol. 42, No. 4, pp. 523-533, 2006.

[38] Grimm, G., Messina, M.J., Tuna, S.E., Teel, A.R.: *Nominally Robust Model Predictive Control With State Constraints*. IEEE Transactions on Automatic Control, Vol. 52, No. 10, pp. 1856-1870, 2007.

[39] Hansen, P., Jaumard, B., Savard, G.: *New Branch-and-Bound Rules for Linear Bilevel Programming*. SIAM J. Sc. Stat. Comput. Vol 13, No. 5, pp- 1194-1217, 1992.

[40] Heuser, H.: *Gewöhnliche Differentialgleichungen: Einführung in Lehre und Gebrauch*. B.G. Teubner, Stuttgart, 1989.

[41] Jarre, F., Stoer, J.: *Optimierung*. Springer-Verlag, Berlin Heidelberg 2004.

[42] Kerrigan, E.C., Maciejowski, J.M.: *Feedback min-max model predictive control using a single linear program: robust stability and the explicit solution*. International Journal of Robust and Nonlinear Control, Vol. 14, pp. 395-413, 2004.

[43] Kalashnikov, V.V., Kalashnikova, N.I.: *Solving Two-Level Variational Inequality*. Journal of Global Optimization, Vol. 8, pp. 289-294, Kluwer Academic Publishers, 1996.

[44] Kohare, M.V., Mettler, B., Morari, M., Bendotti, P., Falinower, C.-M.: *Level Control in the Steam Generator of a Nuclear Power Plant*. IEEE Transactions on Control Systems Technology, Vol. 8, p. 55-69, 2000.

[45] Kolmanovsky, I., Gilbert, E.G.: *Maximal Output Admissible Sets for Discrete-Time Systems with Disturbance Inputs*. Proceedings of the American Control Conference, Washington, pp. 1995-1999, 1995.

[46] Kostina, E., Kostyukova, O. *Worst-case control policies for (terminal) linear-quadratic control problems under disturbances.* International Journal of Robust and Nonlinear Control 19: 1940-1958. Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/oca.1417, 27 March 2009.

[47] Kostyukova, O., Kostina, E.: *Robust optimal feedback for terminal linear-quadratic control problems under disturbances.* Math. Program., Ser. B 107, 131-153, 2006.

[48] Kosut, R.L., Lau, M., Boyd, S.: *Identification of Sytems with Parametric and Nonparametric Uncertainty.* Proceedings of the 1990 American Control Conference, San Diege, Ca, USA, pp. 2412-2417, 1990.

[49] Kothare, M.V., Balakrishnan, V., Morari, M.: *Robust Constrained Model Predictive Control using Linear Matrix Inequalities.* Automatica, Vol. 32, No. 10, pp. 1361-1379, 1996.

[50] Kurzhanski, A.B., Varaiya, P.: *On ellipsoidal techniques for reachability analsis - Part I. External approximations.* Optimization Methods and Software 2002; 17(N2): 177-206.

[51] Kurzshanski, A.B., Varaiya, P.: *Dynamics and Control of Trajectory Tubes - Theory and Computation.* Birkhäuser Science, Springer International Publishing Switzerland, 2014.

[52] Kwakernaak, H., Sivan, R.: *Linear Optimal Control Systems.* John Wiley & Sons, Inc., 1972.

[53] Langson W., Chryssochoos, I., Rakovic, S.V., Mayne, D.Q.: *Robust model predictive control using tubes.* Automatica, Vol. 40 pp. 125-133, 2004.

[54] Lau, M.K., Boyd, S., Kosut, R.L., Franklin, G.F.: *A Robust Control Design for FIR Plants with Parameter Set Uncertainty.* Proceedings of the 30th IEEE Conference on Decision and Control, Boston, Vol. 1 Dec. pp. 83-88, 1991.

[55] Lee, J.H., Yu, Z.: *Worst-case Formulations of Model Predictive Control Systems with Bounded Parameters.* Automatica, Vol. 33. No. 5. pp 763-781, 1997.

[56] Lewis, F.L., Vrabie, D., Syrmos, V.L.: *Optimal Control.* John Wiley & Sons, Inc., 2012.

[57] Limòn Marruedo, D., Álamo, T., Camacho, E.F.: *Stability analysis of systems with bounded additive uncertainties based on invariant sets: Stability and feasibility of MPC.* Proceedings of the American Control Conference Anchorage, AK, pp- 364-369, 2002.

[58] Liu, Y.-H., Spencer, T. H.: *Solving a bilevel linear program when the inner decision maker controls few variables.* European Journal of Operational Research Vol. 81, pp. 644-651, 1995.

[59] Magni, L., De Nicolao, G., Scattolini, R., Allgöwer: *Robust Receding Horizon Control for Nonlinear Discrete-Time Systems.* 15th Triennial World Congress, Barcelona, Spain, 2002.

[60] Magni, L., De Nicolao, G., Scattolini, R., Allgöwer: *Robust model predictive control for nonlinear discrete-time systems*. International Journal of Robust and Nonlinear Control, 13, pp. 229-246 (DOI: 10.1002/rnc.815), 2003.

[61] Magnus, J.R., Neudecker, H.: *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons Ltd. Third Edition, pp.85-86, 2007.

[62] Mayne, D.Q.: *Optimization in model based control*. In Proceedings of the 4th IFAC Symposium Dynamics and Control of Chemical Reactors, Distillation Columns, and Batch Processes, editor: James B. Rawlings, pp 229-242, Copenhagen, Denmark, 1995.

[63] Mayne, D.Q., Rawlings, J.B., Rao, C.V, Scokaert, P.O.M: *Constrained model predictive control: Stability and optimality*. Automatica Vol. 36. No. 6, pp 789-814, 2000.

[64] Mersha, A.G., Dempe, S.. *Linear bilevel programming with upper level constraints depending on the lower level solution*. Applied Mathematics and Computation, Vol. 180, Issue 1, pp 247-254, 2006.

[65] Michalska, H., Mayne, D.Q. *Robust Receding Horizon Control of Constrained Nonlinear Systems*. IEEE Transactions on Automatic Control, Vol. 38, No. 11, pp. 1623-1633, 1993.

[66] Nagy, Z.K., Braatz, R.D.: *Open-loop and closed-loop robust optimal control of batch processes using distributional and worst-case analysis*. Journal of Process Control 14, pp. 411-422, 2004.

[67] Nocedal, J., Wright, S.J. *Numerical Optimization* Springer Science+Business Media, LLC 2006.

[68] Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mishchenko, E.F.: *The Mathematical Theory of Optimal Processes*. John Wiley & Sons, Inc. Third Printing, September 1965.

[69] Rawlings, J.B., Mayne, D.Q.: *Model Predictive Control: Theory and Design*. Nob Hill Publishing, LLC, Electronic Download (5th), 2015.

[70] Rawlings, J.B., Meadows, E.S., Muske, K.R.: *Nonlinear Model Predictive Control: A tutorial and survey*. Proc. Int. Symp. Adv. Control of Chemical Processes ADCHEM, Kyoto, Japan, 1994.

[71] Rodrigues, H.S., Monteiro, M.T.T., Torres, D.F.M.: *Optimal Control and Numerical Software: An Overview*. Preprint, *Systems Theory: Perspectives, Applications and Developments*, Nova Science Publishers, Editor: Francisco Miranda, Jan. 28th 2014

[72] Savkin, A.V., Petersen, I.R.: *A connection between H^∞ control and the absolute stabilizability of uncertain systems*. Elsevier Science B.V. 1994.

[73] Schättler, H., Ledzewicz, U.: *Geometric Optimal Control: Theory, Methods and Examples*. Interdisciplinary Applied Mathematics 38, Springer Science+Business Media, 2012.

[74] Scokaert, P.O.M., Mayne, D.Q. *Min-Max Feedback Model Predictive Control for Constrained Linear Systems.* IEEE Trans. Automat. Cont. 43 (8), 1136-1142, 1998.

[75] Sethi, S.P., Thompson, G.L.: *Optimal Control Theory: Applications to Management Science and Economics.* Springer Science + Business Meida, Inc. 2nd Edition 2000

[76] Shimizu, K., Aiyoshi, E.: *A new computational method for Stackelberg and min-max problems by use of a penalty method.* IEEE Transactions on Automatic Control, 26, pp. 460-466, 1981.

[77] Sinha, A., Malo, P., Kalyanmoy, D.: *A Review on Bilevel Optimization: From Classical to Evolutionary Approaches and Applications.* IEEE Transactions on Evolutionary Computation, vol. 22, no. 2, pp. 276-295, April 2018. DOI: 10.1109/TEVC.2017.2712906

[78] Sorensen, D.C.: *Newton's Method with a Model Trust Region Modification.* SIAM Journal Numerical Analysis, Vol 19, No. 2, pp. 409-426, April 1982.

[79] Starr, A.W., Ho, Y.C.: *Nonzero-Sum Differential Games.* Journal of Optimization Theory and Applications: Vol. 3, No. 3, 1969.

[80] Sussmann, H.J.: *A strong version of the Maximum Principle under weak hypotheses.* Proceedings of the 33rd Conference on Decision and Control lake Buena Vista, FL, 1994.

[81] Sussmann, H.J., Willems, J.C.: *300 Years of Optimal Control: From The Brachystochrone to the Maximum Principle.* IEEE Control Systems Magazine 17 (3): pp 32-44, June 1997.

[82] White, D.J., Anandalingam, G.: *A Penalty Function Approach for Solving Bi-Level Linear Programs.* Journal of Global Optimization Vol. 3, pp. 397-419, 1993.

[83] Wiesemann, W., Tsoukalas, A., Kleniati, P.-M., Rustem, B.: *Pessimistic Bilevel Optimization.* Siam J. Optim., Vol. 23, No 1., pp. 353-380, 2013.

[84] Xing, W., Fang, S.-C., Sheu, R.-L., Zhang, L.: *Canonical Dual Solutions to Quadratic Optimization over One Quadratic Constraint..* Asia-Pacific Journal of Operational Research 32, 2015.

[85] Ye, Y.: *Interior Point Algorithms: Theory and Analysis.* John Wiley & Sons, Inc., 1997.

[86] Zheng, Z.Q., Morari, M.: *Robust Stability of Constrained Model Predictive Control.* 1993 American Control Conference, San Francisco, CA, USA, pp. 379-383, 1993. DOI: 10.23919/ACC.1993.4792879