ℓ_{asso}-MPC – Predictive Control with ℓ_1-Regularised Least Squares

Marco Gallieri

Sidney Sussex College

Department of Engineering

University of Cambridge

A thesis submitted for the degree of

Doctor of Philosophy

January 2014

To Mariya

Acknowledgements

I would like to thank my supervisor, Prof. Jan M. Maciejowski for his support, guidance and inspiration throughout my Cambridge years. He has strongly contributed to my formation and patiently helped me to improve the quality of my work.

I would also like to thank my adviser, Prof. Keith Glover for sharing with me his thoughts on my research, and on many other important matters.

I am grateful to the EPSRC grant control for energy and sustainability for funding my research. I would like to express my gratitude to Prof. Colin Jones of the EPFL, for allowing me to visit his group and for sharing his thoughts with me both in Lausanne and Cambridge. I am indebted to Prof. Daniel Limón of the University of Seville, for the time he has kindly dedicated me during ECC13 and for his helpful thoughts. I would also like to thank Prof. James Whidborne, Prof. Eric Kerrigan, Prof. Richard Vinter and Dr. Stefano Longo for the support and interest in my research. For my robust control education, thanks to Prof. Malcolm Smith, Dr. Glenn Vinnicombe, Dr. Jorge Gonçalves, Dr. Ioannis Lestas. I want thank all of the members of the control group, past and present, for support and friendship. I am truly indebted to Dr. Edward N. Hartley, for his exceptional support in these years, and for providing me suggestions and feedback in record time. I'm quite indebted also to Dr. Alison Eele, Dr. Joe Hall, Dr. Rohan Shekar, Dr. Ellie Siva, Dr. Panos Bresas, Dr. Richard Pates, Dr. Dariusz Cieslar, Dr. Amy Koh. Great thanks to Alex Broekhof, Alberto Carignano, Peyman Gifani, David Hayden, Timothy Hughes, Ian McDonnell, Muyiwa Olanrewaju, Kris Parag, Kaoru Yamamoto, Xiaoke Yang. Thanks also to Carlos Andrade-Cabrera, Dr. Alex Darlington, Abhishek Dutta, Dr. Luisa Pires, Dr. Tom Voice, Dr. Xiaochuan Xuan, Dr. Ye Yuan, Dr. Jason Zheng Jiang. Great thanks to my friends Andrea Maffioli, Francesco Battocchio, Antonio D'Ammaro, Alice Cicirello, Olivia Nicoletti, Raphaël Lefevre, Banu Turnaoğlu, who have made my stay in Cambridge truly unforgettable.

Thanks to my family and to Mariya for the support, patience and love.

Abstract

High performance control of large-scale dynamic systems, such as distribution networks or redundantly-actuated marine/air vehicles, requires control policies that are both lean and robust. Whereas control research is quite mature on the latter, the former has been mainly considered in terms of input energy by means of least-squares approaches. These strategies are sometimes not economic for systems with a large number of inputs as all of the control channels are used for most of the time. The use of hybrid controls can lead to improvement at the expense of increased computation.

This thesis presents ℓ_{asso}-MPC, a Model Predictive Control strategy that aims to produce control inputs that are *sparse* in the actuators domain. This is obtained by means of a particular cost function borrowed from the compressed sensing literature. As a result, we obtained an MPC strategy that inherits some of the desirable proprieties of least squares optimal control such as the smoothness of input/output signals as well as the ability to handle uncertainty. In addition, ℓ_{asso}-MPC allows the designer to exploit further the system's actuators by producing control sequences with sparse magnitudes or time-increments. This is obtained by adding an ℓ_1 regularisation penalty to a standard quadratic cost, resulting in a *soft-thresholding*. The computational complexity of ℓ_{asso}-MPC is comparable to that of a standard MPC.

Sufficient conditions are obtained for closed-loop asymptotic stability of a class of non-linear systems under ℓ_{asso}-MPC. This is done by means of a set-theoretic control framework as an extension of standard MPC theory. A set of procedures is reviewed to perform the required offline computation.

A strategy is proposed to compute the regularisation penalty. This allows the designer to specify a set of preferred actuators to be used for most of the time as well as a neighbourhood of the control error origin where the remaining inputs are in stand-by. This concept is demonstrated through a couple of illustrative examples. The first one is the control of an abstract network with primary and secondary links. The second

one is the roll motion control of a linearised aircraft model. In this case, the ailerons and the rudders are used for all of the time while the spoilers are used only when the control error is large.

A contribution is made to the fields of tracking and robust MPC. A novel strategy is proposed for linear systems to provide an upper bound on the magnitude of the additive uncertainty for which ℓ_{asso}-MPC is robustly feasible and locally Input-to-State Stable. The strategy is based on set-theoretic control methods and soft-constraints. The use of the latter is almost inevitable when dealing with uncertainty. In some cases the proposed approach can be much less conservative than standard robust MPC.

The problem of vessel roll reduction in ocean waves by means of rudder and fins is used as the main demonstrator for ℓ_{asso}-MPC. The objective is to reduce the rudder activity by preferring the fins for the reduction of the roll variance. Simulation results on a detailed vessel model illustrate how the proposed ℓ_{asso}-MPC can outperform more common MPC strategies in terms of a smaller roll variance as well as a significant reduction of the rudder activity.

Marco Gallieri

Sidney Sussex College

Cambridge

July 19, 2014.

Declaration

As required by the University Statute, I hereby declare that this dissertation is not substantially the same as any that I have submitted for diploma or degree or any other qualification at any other university. This dissertation is the result of my own work and included nothing which is the outcome of work done in collaboration, except where specified explicitly in the text.

I also declare that the length of this dissertation is less than $65,000$ words and that the number of figures is less than 150.

Marco Gallieri

Sidney Sussex College

Cambridge

July 19, 2014.

CONTENTS

LIST OF FIGURES

NOMENCLATURE

Roman Symbols

A State transition matrix in LTI systems

B Matrix map from input to state in LTI systems

$b(x)$ Right-hand side vector for the constraints in condensed MPC

B_w Matrix map from disturbance to state in LTI systems

C Output matrix in LTI systems

$C(\nu)$ Vessel centripetal and Coriolis matrix

d Vessel down coordinate in the n-frame

$D(\nu)$ Vessel damping matrix

e Vessel East coordinate in the n-frame

$\bar{f}(\cdot)$ Vector field of a nominal dynamic system

$\hat{f}(\cdot)$ Prediction model in MPC

$F(x)$ MPC terminal cost

f Vector field of a perturbed dynamic system

$g(\eta)$ Vessel restoring force

H Hessian matrix in QP and in condensed MPC cost

$J_b^n(\Theta_{nb})$ Vessel velocity and attitude transformation from the b-frame to the Earth-fixed n-frame

K Terminal controller gain matrix

$K_N(x)$ MPC control law

L Inequality coefficient matrix for the state constraints

M Coefficients matrix of state-dependent terms in condensed MPC constraint

m Number of inputs for a dynamic system

M_A Vessel generalised added-mass matrix

M_{RB} Vessel rigid-body generalised mass matrix

$f_K^i(x)$ System evolution at time i in closed-loop with a linear state-feedback control

m_s Number of rows of the MPC matrix S

M_u Matrix parameterisation of steady inputs for LTI systems

M_x Matrix parameterisation of steady states for LTI systems

N Prediction horizon length in MPC

n Number of states of a dynamic system

n Vessel North coordinate in the n-frame

p Vessel roll rate, b-frame

P Quadratic terminal cost matrix in MPC

$\mathbf{Q}$ State predictions penalty coefficients in condensed MPC

q Vessel pitch rate, b-frame

Q State penalty matrix in MPC

q Number of disturbances acting on a dynamic system

$\mathbf{R}$	Input predictions quadratic penalty coefficients in condensed MPC
R	Quadratic input penalty matrix in MPC
r	Vessel yaw rate, b-frame
S	1-norm input penalty matrix in MPC
s_j	Soft-constraint slack variable for j-th step state prediction in MPC
T_{SIM}	Simulation time
u	Vessel surge velocity, b-frame
u	Input vector for a dynamic system
$\underline{\mathrm{u}}$	MPC input prediction sequence
$\hat{u}_j$	MPC input prediction j-steps ahead
v	Vessel sway velocity, b-frame
$V_N(\cdot)$	Cost function in MPC
$V_O(\cdot)$	Steady state offset cost in MPC for tracking
V_{TAS}	Aircraft true airspeed
w	Vessel heave velocity, b-frame
W	Regularisation matrix in LASSO
w	Disturbance vector for a dynamic system
x	State vector for a dynamic system
$\underline{\mathrm{x}}$	MPC state prediction sequence
$\hat{x}_j$	MPC state prediction j-steps ahead
Z	1-norm terminal cost matrix in MPC

Greek Symbols

α_e Vessel fin angle of attack

α_F Vessel fin angle

α_f Vessel fin flow angle

α_R Vessel rudder angle

$\ell(x, u)$ Stage cost in MPC

η Vessel vector of positions and attitude, n-frame

η ℓ_1-regularisation parameter

Γ Linear terms matrix in condensed MPC cost

$\kappa(x)$ A state-feedback control law

$\kappa(x, p)$ A parameter-varying state-feedback control law

Λ Design matrix in LASSO

λ Contraction parameter for a contractive set

$\lambda_i(A)$ The i-th eigenvalue of a symmetric matrix A

$\lambda_{\max}(A)$ Eigenvalue of maximum modulus of a matrix A

$\lambda_{\min}(A)$ Eigenvalue of minimum modulus of a matrix A

χ Vector of decision variables in optimisation

ν Vector of Lagrange multipliers in optimisation

ν Vessel vector of generalised velocities and rates, b-frame

Ω Coefficients matrix of input-dependent terms in condensed MPC constraint

ϕ Vessel roll angle, $n \to b$ frame

ϕ Matrix mapping the input sequence to the terminal state prediction in MPC

ψ Vessel yaw angle, $n \to b$ frame

Ψ State prediction map from initial state in condensed MPC

$\psi_{\mathcal{X}}(x)$ Minkowski function of a set $\mathcal{X}$

σ Vector of slack variables for the 1-norm input penalty

$\sigma_{\max}(S)$ Maximum singular value of a matrix S

$\sigma_{\min}(S)$ Minimum singular value of a matrix S

τ Vessel vector of forces and moments acting on a ship hull

θ Vessel pitch angle, $n \rightarrow b$ frame

Θ State prediction map from input sequence in condensed MPC

θ Parameter defining steady states and inputs in MPC for tracking

Θ_{nb} Vessel Euler angles, $n \rightarrow b$ frame

Υ Permutation matrix for the input predictions in ℓ_{asso}-MPC with prioritised actuators

ζ Vector of slack variables for the Minkowski terminal cost

Superscripts

ii Denotes the elements of a matrix corresponding to the set of auxiliary actuators

i Denotes the elements of a matrix corresponding to the set of preferred actuators

L Denotes the left pseudo-inverse of a matrix

λ Denotes a λ-contractive variant to a previously defined set

o Denotes the optimal MPC cost, or value function

$\star$ Denotes an optimal input or state prediction in MPC

t Denotes whether the considered set is for the regulation or tracking MPC

Subscripts

$\bullet$ Denotes any previously considered subscript

t Denotes a reference steady state, input or output vector

Special Symbols

$\underline{1}$ Vector of ones

$\bullet$ Arbitrary element of appropriate dimension

C-set Convex compact set

$\mathcal{D}_{++}$ Set of positive diagonal matrices

$\mathbb{I}$ Set of integers

$\mathcal{K}$ A function class

$\mathcal{K}(\cdot)$ Controllable set

$\mathcal{KL}$ A function class

$\mathcal{K}_\infty$ A function class

ℓ_{asso}-MPC MPC with an ℓ_1-regularised least squares stage cost

M_θ Matrix parameterisation of steady states and inputs for LTI systems

$\mathcal{O}_\infty$ Maximal admissible set

$\otimes$ Kronecker product of matrices

$\mathbf{part}\mathcal{X}$ A partition of the set $\mathcal{X}$

$\sim$ Pontryagin difference of sets

π Projection of a set

$\mathcal{P}_j$ Region belonging to a partition of the state space

$\mathcal{Q}(\cdot)$ 1-step operator on sets

$\mathbb{R}$ Set of reals

$\mathcal{S}(\cdot)$ Stabilisable set

$\oplus$ Minkowski sum of sets

$\partial f(x)$ Sub-differential of a function f evaluated at x

$\mathbb{U}$ Input constraint set

$\mathbb{W}$ Disturbance set

$\mathbb{X}$ State constraint set

$\mathbb{X}_f$ Terminal constraint set in MPC

$\mathbb{X}_N$ Feasible region and domain of attraction in MPC

$\mathcal{X}_{\text{nom}}$ Set of states where only the preferred actuators are used

Acronyms

RMS Root Mean Squared

AS Aymptotic Stability

BlockDiag(A, B) A block-diagonal composition of matrices A and B

CLF Control Lyapunov Function

DOA Domain of Attraction

FHCOCP Finite-horizon constrained optimal control problem

ISpS Input-to-State practically Stablility

ISS Input-to-State Stability

KKT Karush-Kuhn-Tucker conditions for optimality

LASSO Least Absolute Selection and Shrinking Operator

LDI Linear Differential Inclusion

LMI Linear Matrix Inequality

LP Linear Program

LPV Linear Parameter Varying

LQ-MPC Linear Quadratic MPC

LQR Linear Quadratic Regulator

LS Least Squares

LTI Linear Time Invariant, referred to a dynamic system

MPC Model Predictive Control

mpLP Multi-parametric Linear Program

mpQP Multi-parametric Quadratic Program

PI Positively Invariant set

pS Practical Stability

PWA Piece-Wise Affine function

PWQ Piece-Wise Quadratic function

QCQP Quadratically Constrained Quadratic Program

QP Quadratic Program

RPI Robust Positively Invariant set

SNAME Society of Naval Architects and Marine Engineers

UB Ultimate Boundedness

CHAPTER

ONE

INTRODUCTION

1.1 Foreword

Performance optimisation is the challenge of modern engineering. The control systems engineer is a vital part of the picture. Engineering high performance automatic systems necessarily features the design of smart control algorithms, for scenarios such as power generation or energy saving. In general, an automatic system has to satisfy a particular performance specification, which is tightly related to its particular application field. The paradigm of optimal control allows the designer to formalise project specifications by means of a cost function. During the second half of the last century this paradigm has been significantly extended in order to meet different needs and to cover a larger class of problems. Nowadays more than ever optimal control is a key technology for high performances systems.

Model Predictive Control (MPC) is an optimal control strategy. Being the most successful modern control technology in industry, MPC consists of a numerical implementation of optimal control. At each time step, given the measurements or an estimate of the system state, the future output of the system is predicted over a finite time-window and an optimal control sequence is obtained by minimising a cost function. Then the first element of the sequence is applied to the plant, a new measurement is taken, and the whole procedure is repeated again. Predictions are generated through an approximate model of the system. The main strength of MPC is the intrinsic ability of handling constraints, by means of constrained optimisation. This provides the controller with the ability to safely drive the system towards operating points that could be not achievable by means of other control technologies, potentially improving performance. A further advantage is the possibility of incorporating several layers of control and planning architectures in a single unit, as shown in Figure 1.1, thus reducing the number of interconnections. This is due to the intrinsic ability of MPC to handle constraints on the systems state and the control inputs as well as the potential for using more complex plant models with respect to classic control. The last 20 years have seen a growing enthusiasm among the academic community for MPC, which is now theoretically quite mature.

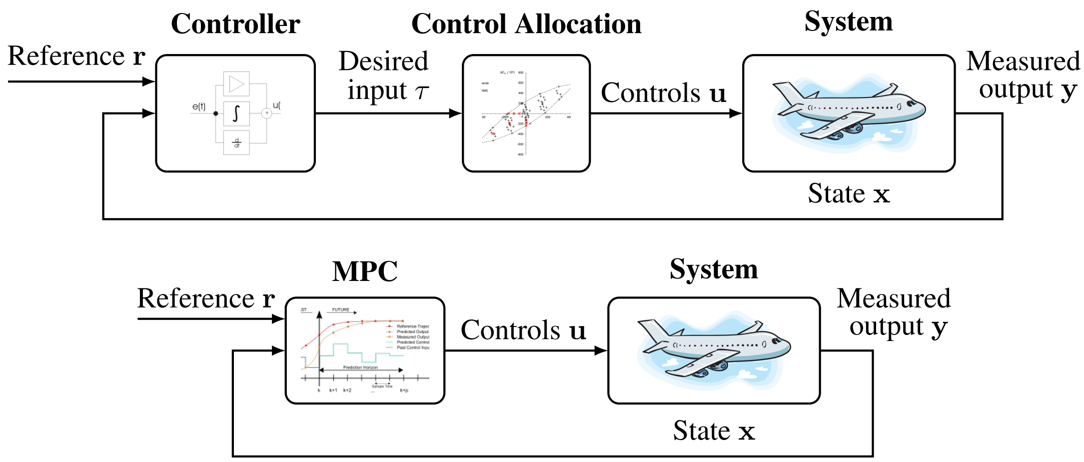

Figure 1.1: Integration of multiple control layers through MPC.

The cost function plays a crucial role in optimal control, being the most significant element of influence for the closed-loop signals behaviour. A wide variety of control problems can be formalised by means of an energy-like quadratic cost. Quadratic cost functions are differentiable. This allows one to compute the minimiser in closed form as in the case of a fundamental *modern-control era* result, the late 1960s *Linear Quadratic Regulator* (LQR). Beside this, a quadratic cost provides other benefits in MPC, such as the well defined conditions for closed-loop stability and a certain level of robustness to uncertainties. For these reasons the most common MPC implementations are still based on a quadratic-input-quadratic-state cost.

Standard quadratic MPC moves *all actuators* for *all of the time*. This is appropriate for most MPC applications, however, in particular cases it might be desirable to move only a subset of the actuators for some particular time period. This is for instance the case of vessel roll reduction and course keeping with rudder and fins. In this application, the fins should be used to reduce the roll motion induced by ocean waves, in other words they should be used most of the time. On the other hand, the rudder affects both the yaw and the roll motion and should be mainly used to steer the ship as well as to help the fins only when the fins have insufficient authority. The use of the latter in fact results in a stronger roll reduction than the former, as well as a smaller amount of additional drag acting on the ship, a reduction of yaw motion interference and lower energy consumption. The vessel configuration is shown in Figure 1.2 together with a picture of the Ansaldo SS Michelangelo, an Italian ocean liner built in the 1960s, where the port fin stabiliser is visible. A similar but larger scale application is the motion control of aircraft, which features nearly

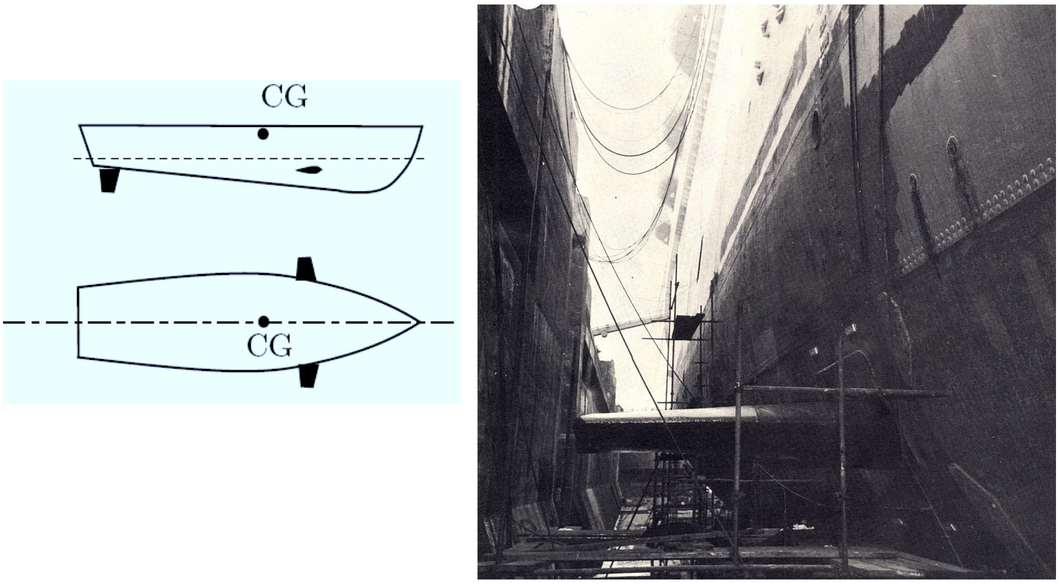

Figure 1.2: Rudder/fins setup for roll reduction based on [Perez & Goodwin, 2008] (left) and a view of the port fin stabiliser on the Ansaldo SS Michelangelo (right).

30 available actuators as shown in Figure 1.3, or a fully connected client-server distribution network with primary and backup links, in Figure 1.4. Further examples of interest include the control of crude oil distillation units, where additional resources are to be used only during the pre-heating stage. For this type of applications the control algorithm should be able to understand, given the current plant state, which actuators are most suitable to perform the given task. This is generally done by dividing the control task into smaller subproblems and by designing small control systems, to be then integrated with minimum interconnection. If the number of these sub-controller could be reduced, then the overall design process could be facilitated. Moreover, achieved performance

may be more robust and economic. Throughout this thesis the above systems are referred to,

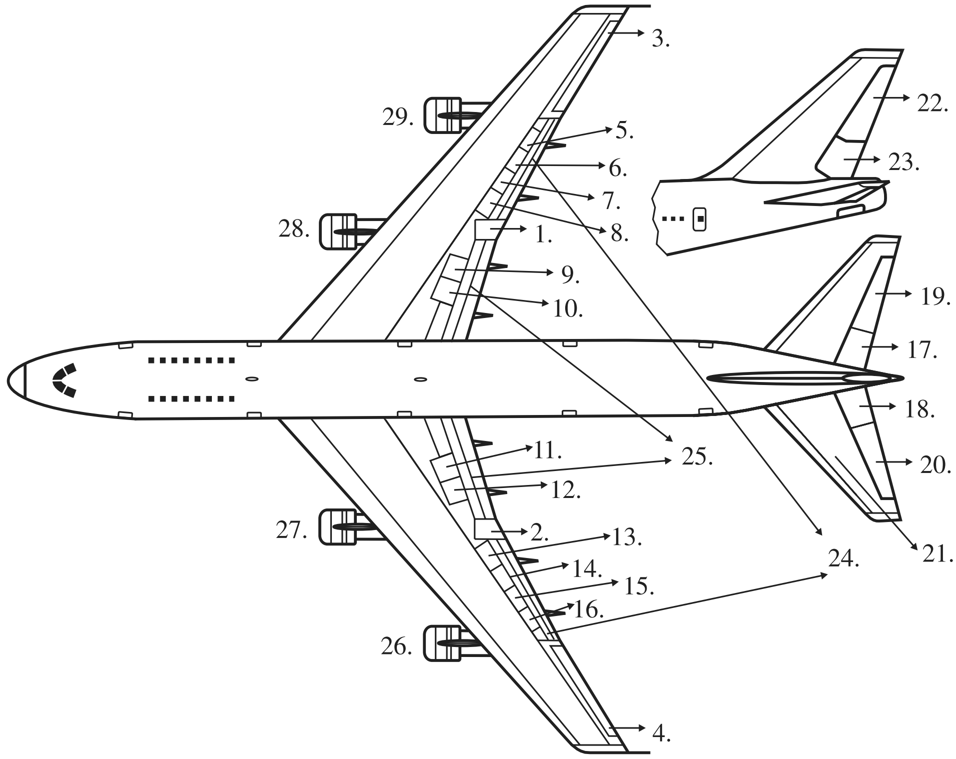

Figure 1.3: Control inputs available on a Boeing 747. Based on [Edwards *et al.* , 2010].

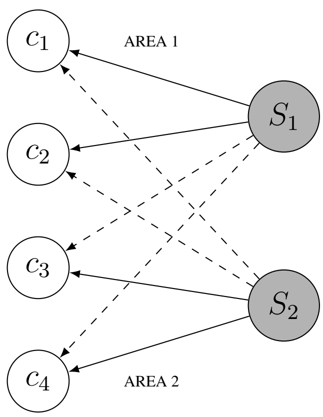

Figure 1.4: Redundant distribution network. Main links (solid), auxiliary links (dashed).

for simplicity, as *redundantly-actuated*. The presence of backup inputs also makes this system capable of an intrinsic level of fault tolerance, the ability of continuing to operate when a part of the system (for instance a certain actuator) fails. MPC in its standard form is known to possess the capabilities for intrinsic fault tolerance [Edwards *et al.* , 2010]. Albeit not explicitly investigated here, the control strategy proposed in this thesis could potentially enhance the fault tolerance ability of standard MPC, the former being able to handle secondary actuators in a smarter way. The motion control of valves, dam and floodgate mechanisms (see Figure 1.5), or wind turbine variable pitch blades also fall in the category of control problems addressed in this thesis, whereas their aim is not to use some actuators more than others, but instead to reduce actuators' wear and tear. From the above considerations it appears to be desirable when applying MPC to such systems to obtain a control policy that is *sparse*, in other words that contains many zero elements instead of small non-zero elements. Sparsity could be desired among the input channels as well as through time.

Figure 1.5: Tokyo floodgates created to protect from typhoon surges. Monitoring and reduction of wear and tear are crucial.

This thesis presents a novel MPC strategy, called ℓ_{asso}-MPC, the cost function of which provides a tradeoff between a standard quadratic criterion and the sum of control signals magnitudes. The approach is inspired by the recent development of ℓ_1-regularised least squares in the field of system identification and signal processing. This strategy, also known as LASSO regression, has been particularly successful in the mentioned fields. The main feature of LASSO regression is *sparsity* of the solution. In the thesis this characteristic will be investigated further through the explicit characterisation of the controller for Linear Time-Invariant (LTI) prediction models. The focus of the thesis is the control of constrained LTI systems, however, closed-loop stability under ℓ_{asso}-MPC is also investigated for the class of non-linear systems with differentiable vector field and the origin as equilibrium. This leads to a set of sufficient conditions that rely on the numerical computation of a control invariant set and a control Lyapunov function. Numerical tools (currently available) are reviewed for the computation of this set. In particular, while these concepts are becoming quite standard for LTI systems, research is still active for the considered non-linear systems class where computational complexity is a strong limitation. Following the stability analysis, ℓ_{asso}-MPC is used for the design of control systems for LTI plants with prioritised and backup actuators. In particular, a procedure is developed so that the designer can specify a nominal operation zone where only a preferred set of actuators is used. Finally, robustness to additive uncertainties is considered together with set-point tracking. Through the development of ℓ_{asso}-MPC general theoretical aspects of MPC have also been reviewed, with the aim of reducing the required offline computation and the conservatism of standard assumptions in MPC theory.

1.2 Thesis Structure

The document is structured as follows

- Chapter 2 is a brief *vademecum* of background material for later reference. Convex sets, functions and convex optimisation are reviewed first as well as the theory of exact penalty functions. The principles for stability and robustness of discrete-time non-linear control systems are then reviewed, in a state-space framework. Set-theoretic control is introduced, being the considered framework for the developments of stability and robustness results of this thesis. Fundamentals of MPC theory follow, including a revised view of conditions for stability and intrinsic robustness, namely the robustness of the system under an MPC designed for the nominal model. The explicit solution of LQ-MPC is also reviewed as it will be used for the analysis of the proposed controller. Finally, robust MPC techniques and MPC for tracking are briefly reviewed.

- Chapter 3 introduces ℓ_{asso}-MPC. First, the LASSO regressor and its common variants are reviewed together with the geometric conditions under which a sparse solution is obtained. Then ℓ_{asso}-MPC for regulation is introduced, in the two different versions considered in the thesis. These versions differ for their terminal costs, upon which stability depends. Initial simulations are performed to compare ℓ_{asso}-MPC with the more common LQ, ℓ_1 and ℓ_∞-MPC, the latter two also being able to obtain sparse solutions at the expense of poorer performance and possible actuator chattering. The geometries of ℓ_1 and ℓ_{asso} problems are analysed for a specific example, to demonstrate how sparsity can be achieved by ℓ_{asso}-MPC without the chattering of ℓ_1-MPC.

- Chapter 4 presents the first version of ℓ_{asso}-MPC, featuring a quadratic terminal cost as in LQ-MPC. In particular, for LTI plants the solution is shown to be piece-wise affine (PWA) in the control error as well as in the weight of the 1-norm input penalty (the regularisation penalty). The solution of ℓ_{asso}-MPC is PWA even for the unconstrained case, in clear contrast with LQ-MPC, and is zero when the control error is in a neighbourhood of the origin. Closed-loop stability is studied for LTI systems, as well as for non-linear systems with differential vector fields and with the origin as a steady-state, using a general regularisation penalty. In particular, it is shown that ultimate boundedness can be achieved using version 1. On the other hand, if only a sub-set of actuators is regularised, then asymptotic stability is proven by using the standard assumptions of LQ-MPC. This case is referred to as partial regularisation, and it leads to the concept of preferred actuators discussed in Chapter 6.

- Chapter 5 presents the second version of ℓ_{asso}-MPC, which provides asymptotic stability for a general regularisation penalty. A set of candidate terminal costs is studied, together with sufficient conditions to satisfy standard assumptions for MPC stabilisation LTI system. In particular, the Minkowski function of the terminal set is proposed as the terminal cost, the computation of which is also reviewed. In contrast to standard MPC theory, in this chapter it is recommended to compute the terminal set independently from the cost function, so that online tuning is facilitated. The considered stability conditions are also extended for the considered class of non-linear systems (the same as for Chapter 4), and the increase in computational complexity is discussed.

- Chapter 6 deals with the control of systems for which there is a clear distinction between preferred and auxiliary actuators, the latter to be used only when the control error is large. Explicit MPC and exact penalty functions are used to show how ℓ_{asso}-MPC can implement this idea. A design procedure is formulated, for each of the two ℓ_{asso}-MPC versions, which allows the designer to impose a certain nominal operation zone, namely, a neighbourhood of the origin in which the auxiliary actuators are never used. Limitations due to constraints and the finite prediction horizon are formalised and demonstrated through an example. While ℓ_{asso}-MPC version 1 can be used to embed an existing LQ-MPC, ℓ_{asso}-MPC version 2 can be used to obtain multiple levels of priority. The paradigm is demonstrated for partial regularisation through the control of a client-server network with prioritised links and of the linearised lateral dynamics of a Boeing 747. In particular, the latter example demonstrates how the use of the spoilers can be confined to situations when the pilot command is larger than a desired threshold.

- Chapter 7 is concerned with the topic of robustly tracking a piece-wise constant reference by means of ℓ_{asso}-MPC with LTI models. Soft constraints are used to recover from infeasibility due to disturbances and a novel terminal set is proposed to obtain robust feasibility, invariance and local Input-to-State Stability (ISS) of the closed loop system. The proposed procedure returns an ∞-norm bound on the allowable additive disturbance. A small example is used to show that the proposed strategy can be less conservative than a standard robust MPC strategy. This results from the relaxation of constraints, which could be anyway inevitable when the assumed uncertainty description is incorrect.

- Chapter 8 investigates the application of ℓ_{asso}-MPC version 1 for vessel roll reduction in ocean waves, by means of rudder and fins. The proposed strategy is compared to the LQ, ℓ_1 and ℓ_∞-MPC, which are outperformed. In particular, the ℓ_{asso}-MPC is shown to provide the highest reduction of the roll variance while making the minimum use of the rudder.

- Chapter 9 provides concluding remarks and suggests future research directions.

1.3 Publications

The work presented in this thesis extends the concepts of the following conference publications:

- Chapters 4, 8: [Gallieri & Maciejowski, 2012].

- Chapter 5: [Gallieri & Maciejowski, 2013a,b].

- Chapter 7: [Gallieri & Maciejowski, 2013a].

During the course of this study, the following journal article has been produced in collaboration with Dr. E. N. Hartley, the content of which is not included in this thesis: [Hartley *et al.* , 2013]. Some of the content of Chapter 6 has been recently submitted as a conference paper.

CHAPTER

TWO

BACKGROUND

2.1 Basic definitions

Denote the set of integers (reals) as $\mathbb{I}$ ($\mathbb{R}$), the set of nonnegative integers (reals) as $\mathbb{I}_{\geq 0}$ ($\mathbb{R}_{\geq 0}$) and the set of positive integers (reals) as $\mathbb{I}_{>0}$ ($\mathbb{R}_{>0}$). The integers from 0 to N are denoted as $\mathbb{I}_{[0,N]}$. Given a square matrix A and the scalar λ_i denoting the i-th eigenvalue of A, then $\lambda_{\max}(A) = \max_i |\lambda_i|$, $\lambda_{\min}(A) = \min_i |\lambda_i|$. The notation BlockDiag$(A, B)$ denotes a block-diagonal matrix, and $\otimes$ denotes the Kronecker product. The identity matrix of size N is denoted by I_N, while $\underline{1}_m$ is a unit column-vector of m elements. A block of zeros in a matrix is simply denoted by 0. Matrix dimensions are omitted if they can be inferred from the context. In the thesis, $A \succ 0$ (respectively $\succeq$) denote that the matrix A is positive (semi-)definite while $\geq$ (respectively $>$) is used for scalar (in)equalities. The scalar "positive-wrapping" operator is defined as

$$(a)_+ = \max(a, 0). \tag{2.1.1}$$

The following definitions are used throughout the thesis:

Definition 2.1.1. $\mathcal{D}_{++}$-matrix
A diagonal matrix S is said to be of class $\mathcal{D}_{++}$ if its entries are strictly positive.

Definition 2.1.2. Affine function
A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *affine* if it has the form linear plus constant, $f(x) = Ax + b$.

Definition 2.1.3. (Generalised) quadratic function

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is a *(generalised) quadratic* if it has the form

$$F(x) = x^T P x + 2d^T x + e. \tag{2.1.2}$$

Definition 2.1.4. A (generalised) quadratic function of the form (2.1.2) is *non-degenerate* iff $P = P^T \succ 0,\ e - d^T P^{-1} d < 0$.

2.2 Convex Sets

Convex sets are generally used to represent convex constraints, the domain of convex functions and level sets. In the following, a set of definition will be given for convex sets and functions, as in [Hindi, 2004]. For an extensive discussion refer to [Boyd & Vandenberghe, 2004; Hindi, 2004].

Definition 2.2.1. Convex set

A set $\mathcal{S} \subseteq \mathbb{R}^n$ is *convex* if it contains the line segment joining any two of its points, that is:

$$x, y \in \mathcal{S}, \quad \theta \in [0, 1] \Rightarrow \theta x + (1 - \theta) y \in \mathcal{S}. \tag{2.2.1}$$

Definition 2.2.2. Convex cone

A set $\mathcal{S} \subseteq \mathbb{R}^n$ is a *convex cone* if it contains all rays passing through its points, which emanate from the origin, as well as all line segments joining any points of those rays, i.e.:

$$x, y \in \mathcal{S}, \quad \alpha, \beta \geq 0 \Rightarrow \alpha x + \beta y \in \mathcal{S}. \tag{2.2.2}$$

Definition 2.2.3. Norm ϵ-ball

A *norm ϵ-ball* is a convex set, $\mathcal{B}_{p,\epsilon}(x_c) = \{x \mid \|x - x_c\|_p \leq \epsilon\}$, $\epsilon > 0$, where $\|\cdot\|_p$ is an ℓ_p-norm.

Definition 2.2.4. Norm cone

$\mathcal{C}_p = \{(x, t) \mid \|x\|_p \leq t\}$, where $\|\cdot\|_p$ is an ℓ_p-norm.

Definition 2.2.5. Closed set

A set $\mathcal{S} \subseteq \mathbb{R}^n$ is *closed* if every point not in $\mathcal{S}$ has a neighborhood disjoint from $\mathcal{S}$, that is:

$$\forall x \notin \mathcal{S}, \quad \exists \epsilon > 0 : \ \mathcal{B}_{p,\epsilon}(x) \cap \mathcal{S} = \emptyset. \tag{2.2.3}$$

Definition 2.2.6. Bounded set

A set $\mathcal{S} \subseteq \mathbb{R}^n$ is *bounded* if it is contained inside some ball, $\mathcal{B}_{p,r}(x)$, of finite radius r, that is:

$$\exists r > 0,\ x \in \mathbb{R}^n :\ \mathcal{S} \subseteq \mathcal{B}_{p,r}(x). \tag{2.2.4}$$

Definition 2.2.7. Compact set

A set $\mathcal{S} \subseteq \mathbb{R}^n$ is *compact* if it is closed and bounded.

Definition 2.2.8. C-set

A set $\mathcal{S} \subset \mathbb{R}^n$ is said to be a *C-set* if it is convex and compact.

2.2.1 Set operators

Proofs of stated lemmas are omitted for brevity. Refer to [Blanchini, 1999].

Definition 2.2.9. Convex Hull

Given a set $\mathcal{S} \subseteq \mathbb{R}^n$, its *convex hull*, $\mathbf{conv}(\mathcal{S}) \subset \mathbb{R}^n$, is the smallest convex set containing $\mathcal{S}$.

Definition 2.2.10. Set union

Given two sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n$, their *union*, $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$, verifies $x \in \mathcal{C} \Leftrightarrow x \in \mathcal{A} \vee x \in \mathcal{B}$.

Definition 2.2.11. Set intersection

Given two sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n$, their *intersection*, $\mathcal{C} = \mathcal{A} \cap \mathcal{B}$, verifies $x \in \mathcal{C} \Leftrightarrow x \in \mathcal{A} \wedge x \in \mathcal{B}$.

Lemma 2.2.1. Convexity is preserved under intersection. The intersection of convex sets of a given type provides a convex set also of the same type.

Definition 2.2.12. Set scaling

Given the convex set $\mathcal{S} \subseteq \mathbb{R}^n$ and the scalar $\alpha \geq 0$, then $\alpha\mathcal{S} = \{x = \alpha y,\ y \in \mathcal{S}\}$.

Definition 2.2.13. Set difference

Given two C-sets $\mathcal{A},\ \mathcal{B} \subset \mathbb{R}^n$, their *set difference* is $\mathcal{A} \backslash \mathcal{B} = \{x \mid x \in \mathcal{A},\ x \notin \mathcal{B}\}$.

Definition 2.2.14. Minkowski sum

Given two sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n$, their *Minkowski sum* is $\mathcal{A} \oplus \mathcal{B} = \{x = a + b \in \mathbb{R}^n \mid a \in \mathcal{A}, b \in \mathcal{B}\}$.

Definition 2.2.15. Set erosion or Pontryagin difference

The *Pontryagin difference* of two sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n$ is $\mathcal{A} \sim \mathcal{B} = \{x \in \mathbb{R}^n \mid x + b \in \mathcal{A}\ \forall b \in \mathcal{B}\}$.

Lemma 2.2.2. The set erosion and Minkowski sum satisfy

$$(\mathcal{A} \sim \mathcal{B}) \oplus \mathcal{B} \subseteq \mathcal{A}.$$

Definition 2.2.16. Projection on a subset

Given two sets $\mathcal{A} \subseteq \mathbb{R}^n$, $\mathcal{B} \subseteq \mathbb{R}^n$, the *projection* of $\mathcal{A}$ on $\mathcal{B}$ is

$$\pi_{\mathcal{B}}(\mathcal{A}) = \{b \in \mathcal{B} : \exists a \in \mathcal{A} \mid a = b + c,\ c^T b = 0\}.$$

Definition 2.2.17. Projection on a lower dimensional space

Given a set $\mathcal{A} \subseteq \mathbb{R}^n$, the *projection* of $\mathcal{A}$ on $\mathbb{R}^m$, with $m < n$, is

$$\pi_m(\mathcal{A}) = \{b \in \mathbb{R}^m : \exists c \in \mathbb{R}^{n-m} \mid (b, c) \in \mathcal{A}\}.$$

Lemma 2.2.3. The operations of *Set scaling*, *Minkowski sum*, *Pontryagin difference* and *Projection on a subset* or *on a subspace* preserve convexity, if the involved sets are convex.

2.2.2 Commonly referred sets

Definition 2.2.18. Halfspace

A closed *halfspace* in $\mathbb{R}^n$ is a convex set, $\mathcal{S} = \{x \mid a^T x \leq b\}$, with $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. If $b = 0$, then $\mathcal{S}$ is also a convex cone.

Definition 2.2.19. Polytope (Polyhedron)

A *polytope* is a *bounded polyhedron*, namely, a bounded intersection of a finite number of halfspaces, $\mathbb{X} = \{x \mid Gx \leq \underline{1}\}$. A polytope can also be represented by its vertices, $v \in \mathcal{V}(\mathbb{X})$.

Definition 2.2.20. Polyhedral partition

The collection of sets $\mathbf{part}\mathcal{P} = \{\mathcal{P}_j\}_{j=1}^N$ with $N \in \mathbb{I}_{>0}$ is called a *polyhedral partition* of a set $\mathcal{P} \subseteq \mathbb{R}^n$ if all $\mathcal{P}_j \subseteq \mathbb{R}^n$ are polyhedra, $\cup_{j=1}^N \mathcal{P}_j = \mathcal{P}$ and $\mathbf{int}\mathcal{P}_i \cap \mathbf{int}\mathcal{P}_j = \emptyset\ \forall i \neq j,\ i, j \in \mathbb{I}_{[1,N]}$.

Definition 2.2.21. Ellipsoid

An *ellipsoid* is a convex set, $\mathcal{E} = \{x \mid (x - x_c)E^{-1}(x - x_c) \leq 1\}$. It can also be expressed as a sublevel set of a non-degenerate generalised quadratic function

$$\mathcal{E} = \{x \mid F(x) \leq 0\},\ F(x) = x^T P x + 2d^T x + e,\ P = P^T \succ 0,\ e - d^T P^{-1} d < 0. \qquad (2.2.5)$$

Definition 2.2.22. First-order cone

A *first-order cone* is the norm cone associated with the ℓ_1-norm, $\mathcal{C}_1 = \{(x, t) \mid \|x\|_1 \leq t\}$.

Definition 2.2.23. Second-order cone

A *second-order cone* is the cone associated with the Euclidean norm, $\mathcal{C}_2 = \{(x, t) \mid \|x\|_2 \leq t\}$.

2.2.2.1 Quadratic constraints and LMIs

A couple of commonly referred Lemmas are stated. Proofs are omitted for brevity. The interested reader can refer to [Boyd & Vandenberghe, 2004; Hindi, 2004].

Lemma 2.2.4. The *second-order cone constraint* (SOCC)

$$\|Ax + b\|_2 \leq c^T x + d, \tag{2.2.6}$$

with $A \in \mathbb{R}^{k \times n}$, is convex. Moreover, it is equivalent to $(Ax + b, c^T x + d) \in \mathcal{S}_c$, where $\mathcal{S}_c$ is the second order cone in $\mathbb{R}^{k+1}$.

Lemma 2.2.5. The solution set of the *Linear Matrix Inequality* (LMI):

$$F(x) = A_0 + x_1 A_1 + \cdots + x_m A_m \succeq 0, \tag{2.2.7}$$

is convex. In particular, any SOCC can be expressed as an LMI, and vice-versa.

2.3 Functions and norms

The following definitions and Lemmas are given according to [Boyd & Vandenberghe, 2004; Lerner, 2010; Robbins, n.d.; Walker, n.d.]:

Definition 2.3.1. Extended function

A function, $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$, is *extended* if it is assumed that $f(x) = \infty, \forall x \notin \mathbf{dom} f$.

Definition 2.3.2. Continuous function

A function $f \; : \; \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ is *continuous on* $\mathcal{X} \subseteq \mathbf{dom} f$ iff

$$\forall x_0 \in \mathcal{X}, \; \forall \epsilon > 0, \; \exists \, \delta(x_0, \epsilon) > 0 : \forall x \in \mathcal{X}, |x - x_0| < \delta(x_0, \epsilon) \Rightarrow |f(x) - f(x_0)| < \epsilon.$$

Definition 2.3.3. Uniformly continuous function

A function $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ is *uniformly continuous on* $\mathcal{X} \subseteq \mathbf{dom} f$ iff

$$\forall \epsilon > 0,\ \exists\, \delta(\epsilon) > 0 : \forall x, x_0 \in \mathcal{X}, |x - x_0| < \delta(\epsilon) \Rightarrow |f(x) - f(x_0)| < \epsilon.$$

Definition 2.3.4. Lipschitz continuous function

A function $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ is *Lipschitz continuous on* $\mathcal{X} \subseteq \mathbf{dom} f$, with *Lipschitz constant* $L \in \mathbb{R}$, iff

$$\forall x, x_0 \in \mathcal{X}, |f(x) - f(x_0)| < L|x - x_0|.$$

Lemma 2.3.1. A Lipschitz continuous function on $\mathcal{X}$ is also uniformly continuous on $\mathcal{X}$.

Lemma 2.3.2. A uniformly continuous function on $\mathcal{X}$ is also continuous on $\mathcal{X}$.

Theorem 2.3.3. (Heine-Cantor)

A function that is continuous on a *compact set*, $\mathcal{X} \subset \mathbb{R}^n$ is uniformly continuous on $\mathcal{X}$.

Definition 2.3.5. Epigraph

The *epigraph* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the set of points above the function, namely:

$$\mathbf{epi} f \ = \ \{(x, \alpha) \in \mathbf{dom} f \times \mathbb{R} \ : \ \alpha \geq f(x)\}. \tag{2.3.1}$$

Different definitions of convexity are given. The purpose of these definitions will be clarified in Section 2.4.1.

Definition 2.3.6. (Strictly) convex function

A function $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ is *(strictly) convex* if its domain, $\mathcal{X} = \mathbf{dom} f$, is convex and if $\forall\, x, y \in \mathcal{X}$, $\forall\, \theta \in [0, 1]$

$$f(\theta x + (1 - \theta) y) \leq (<)\ \theta f(x) + (1 - \theta) f(y). \tag{2.3.2}$$

Lemma 2.3.4. A function $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex iff $\mathbf{epi} f$ is a convex set.

Definition 2.3.7. Strongly convex function

A function $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ is *strongly convex* on the closed convex set $\mathcal{X} \subseteq \mathbf{dom} f$ with respect to the ℓ_p-norm $\|\cdot\|_p$ if $\exists \sigma > 0 : \forall\, x, y \in \mathcal{X}, \forall\, \theta \in [0, 1]$

$$f(\theta x + (1 - \theta) y) \leq \theta f(x) + (1 - \theta) f(y) - \frac{1}{2}\sigma\theta(1 - \theta)\|x - y\|_p^2, \tag{2.3.3}$$

where σ is called the *convexity parameter* of $f(\cdot)$.

Lemma 2.3.5. Given a strongly convex function $f : \ \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$, and a convex function $g : \ \mathbf{dom} g \subseteq \mathbb{R}^n \to \mathbb{R}$, on the set $\mathcal{X}$, then $h(x) = f(x) + g(x)$ is strongly convex on $\mathcal{X}$.

Proof. The result follows trivially from application of the respective definitions. ■

The following definition is based on [Kakade *et al.* , 2012]. Refer to [Vinter, 2000] for a more general discussion, concerning a broader class of functions.

Definition 2.3.8. Sub-differential of a convex function

Given a convex (extended) function, $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, it is *subdifferential at* $x \in \mathcal{X}$ if

$$\partial f(x) = \{\mathbf{v} \in \mathcal{X} \mid \forall z,\ f(x+z) \geq f(x) + \mathbf{v}^T z\}.$$

Moreover, if $x \in \mathcal{X}$, then $\mathbf{v} \in \partial f(x)$ is said to be a *subgradient of* $f(\cdot)$ *at* x.

Definition 2.3.9. Sub-differential of the ℓ_1 and ℓ_∞-norm

The *p-norm subdifferential*, evaluated at u, is denoted $\partial \|u\|_p$. For $p = 1$ any subgradient $\mathbf{v} \in \partial \|u\|_1$ satisfies [Osborne *et al.* , 2000]

$$\mathbf{v}_i = \begin{cases} \operatorname{sign}(u_i) & \text{if } |u_i| > 0, \\ \alpha_i \in [-1,\ 1] & \text{if } u_i = 0. \end{cases} \tag{2.3.4}$$

For $p = \infty$, we have (using the arguments of Section 3.3 of [Boyd & Vandenberghe, n.d.])

$$\partial \|u\|_\infty = \mathbf{conv}(\cup \{\partial |u_i| \mid |u_i| = \|u\|_\infty\}). \tag{2.3.5}$$

Remark 1. It can be shown, for $p = 1, \infty$, that any $\mathbf{v} \in \partial \|u\|_p$ satisfies $\mathbf{v}^T u = \|u\|_p$.

Definition 2.3.10. Differentiable function

An extended function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is *differentiable at* $x \in \mathcal{X}$ iff its subdifferential at x is a *singleton*, namely, $\partial f(x) = \{\nabla f(x)\}$. If this happens $\forall x \in \mathcal{X}$, then f is *differentiable on* $\mathcal{X}$.

Convexity of differentiable functions can be checked by using the following conditions [Boyd & Vandenberghe, 2004; Hindi, 2004]:

Definition 2.3.11. First-order convexity condition

A differentiable function $f : \mathcal{X} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is *convex* if and only if for all $x, x_0 \in \mathcal{X}$:

$$f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0), \tag{2.3.6}$$

where $\partial f(x) = \{\nabla f(x)\}, \forall x \in \mathcal{X}$. Namely, the first order Taylor approximation of f is a *global under-estimator*.

Definition 2.3.12. Second-order (strict) convexity condition

A twice differentiable function $f : \mathcal{X} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is (strictly) *convex* if and only if for all $x \in \mathbf{dom} f$, $\nabla^2 f(x) \succeq 0$ $(\succ 0)$, where $\partial \nabla f(x) = \{\nabla^2 f(x)\}$. In other words, the Hessian of f is positive semi-definite (positive definite) on $\mathcal{X}$.

Definition 2.3.13. Piece-wise affine (quadratic) function

A function $f : \mathcal{P} \rightarrow \mathbb{R}^m$, with $\mathcal{P} \subseteq \mathbb{R}^n$ is piece-wise affine (PWA), respectively, piece-wise quadratic (PWQ) with $m = 1$, if given a partition $\mathbf{part}\mathcal{P}$ it holds that, respectively,

$$\begin{aligned} f(x) &= A_j x + b_j, & \Leftarrow x \in \mathcal{P}_j, \forall \mathcal{P}_j \in \mathbf{part}\mathcal{P}, \\ f(x) &= x^T P_j x + 2 d_j^T x + e_j, & \Leftarrow x \in \mathcal{P}_j, \forall \mathcal{P}_j \in \mathbf{part}\mathcal{P}, \end{aligned}$$

where $A_j \in \mathbb{R}^{m \times n}$, $b_j \in \mathbb{R}^m$, $P_j \in \mathbb{R}^{n \times n}$, $d_j \in \mathbb{R}^n$, $e_j \in \mathbb{R}$, $\forall j \in \mathbb{I}_{[1,N]}$.

Proofs of the following Lemmas are omitted for brevity. The interested reader can refer to [Boyd & Vandenberghe, 2004; Hindi, 2004].

Definition 2.3.14. Functions composition

Given two functions $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^p \rightarrow \mathbb{R}^n$, then $f \circ g(x) = f(g(x))$.

Lemma 2.3.6. The composition of two convex functions is convex.

Lemma 2.3.7. Affine transformation of domain preserves convexity, namely f convex $\Rightarrow f(Ax + b)$ convex.

Definition 2.3.15. Set image under a map

Given a set $\mathcal{A} \subseteq \mathbb{R}^n$ and a map $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, the *image* of $\mathcal{A}$ under f is $f(\mathcal{A}) = \{y = f(x) \in \mathbb{R}^m \mid x \in \mathcal{A}\}$.

Similarly, one can define the set *preimage*, given an invertible map.

Lemma 2.3.8. The image (or preimage) of a convex set under affine transformation is convex. If $\mathcal{S}$, $\mathcal{T}$ are convex, then so are

$$f(\mathcal{T}) = \{Ax + b \mid x \in \mathcal{T}\}, \quad f^{-1}(\mathcal{S}) = \{x \mid Ax + b \in \mathcal{S}\}.$$

Lemma 2.3.9. The minimum over x of a function $f(x, y)$ that convex in both x and y, denoted as $g(y) = \min_x f(x, y)$, is a convex function of y.

The following norm identity can be verified from [Horn & Johnson, 2010], page 314.

Lemma 2.3.10. Given $x \in \mathbb{R}^n$, and $p > r > 0$, then

$$\|x\|_p \leq \|x\|_r \leq n^{(1/r-1/p)}\|x\|_p. \tag{2.3.7}$$

2.4 Convex optimisation

A brief review of convex optimisation is given, based on [Hindi, 2004, 2006] and on [Boyd & Vandenberghe, 2004]. Convex optimisation can be described as a fusion of three disciplines: convex analysis, optimisation, and numerical computation.

Several difficulties may arise in solving a general optimisation problem: presence of local minima, difficult characterisation of feasible domain, stopping criteria, convergence rates, numerical problems. For convex optimisation problems, the first three difficulties are naturally overcome, by definition of convexity. Development of interior point algorithms, first intrduced by [Karmarkar, 1984], provided a drastic improvement in terms of solution time to the point that, at the present time, convex MPC can be applied to many problems in real time [Richter *et al.* , 2013; Wang & Boyd, 2010].

The objectives of this section are: to review some convex problems related to MPC, to provide a brief review of duality and optimality conditions, and to introduce exact penalty functions.

2.4.1 Conic Programming

Conic programs or conic problems are convex problems in which inequalities are specified in terms of affine functions and generalised inequalities. Geometrically, the inequalities are feasible if the range of the affine mappings intersects the cone of the inequality.

A *linear program* (LP) has the form

$$\begin{aligned} \min_{\chi} \ & g^T \chi + d \\ \text{s.t.} \ & \Phi \, \chi = \phi, \\ & \Omega \, \chi \leq b. \end{aligned} \tag{2.4.1}$$

This family of problems can be solved using several algorithms, for instance the *simplex* and its most recent extensions [Boyd & Vandenberghe, 2004].

A *quadratic program* (QP) has the form

$$\begin{aligned} \min_{\chi} \ & \frac{1}{2}\chi^T H \chi + g^T \chi + d \\ \text{s.t.} \ & \Phi \, \chi = \phi, \\ & \Omega \, \chi \leq b. \end{aligned} \tag{2.4.2}$$

where the Hessian is at least positive semi-definite, $H \succeq 0$. More general conic programs include *quadratically constrained quadratic program* (QCQP), *second-order cone program* (SOCP) and *semidefinite program* (SDP). Refer to [Boyd & Vandenberghe, 2004; Hindi, 2004] for details.

2.4.2 Duality

Duality, or *Lagrangian duality*, is a theory used for simplifying and solving optimisation problems. The basic idea in Lagrangian duality is, given a general non-linear program, to augment the objective function in order to account for constraint violation as a weighted extra cost.

Given a general non-linear program

$$\begin{aligned} f^\star = \min_{\chi} \ & f(\chi) \\ \text{s.t.} \ & F(\chi) \leq 0, \\ & h(\chi) = 0. \end{aligned} \tag{2.4.3}$$

where $F \in \mathbb{R}^p$ and $h \in \mathbb{R}^n$ have entries F_i, h_i. Define the *Lagrangian* function, as

$$L(\chi, \lambda, \nu) = f(\chi) + \sum_{i=1}^{m} \nu_i h_i(\chi) + \sum_{i=1}^{p} \lambda_i F_i(\chi). \tag{2.4.4}$$

The *Lagrange dual function* can be defined as

$$\mathcal{L}(\lambda, \nu) = \inf_{\chi \in \mathcal{D}} L(\chi, \lambda, \nu), \tag{2.4.5}$$

where $\mathcal{D}$ is the domain defined by the intersection of inequality and equality constraints.

Definition 2.4.1. Lagrange dual problem

The *Lagrange dual problem* is a convex program, defined as

$$d^\star = \max_{\lambda,\nu} \mathcal{L}(\lambda, \nu) \tag{2.4.6}$$
$$\text{s.t. } \lambda \geq 0. \tag{2.4.7}$$

Note that, if $\lambda < 0$, then $\mathcal{L}(\lambda, \nu) = -\infty$. If $\mathcal{L}(\lambda, \nu) > -\infty$ and the pair (λ, ν) is called *dual feasible* [Boyd & Vandenberghe, 2004; Hindi, 2006]. The optimal solution $(\lambda^\star, \nu^\star)$ of the dual problem is referred to as the *dual optimal*. The original problem is referred to as the *primal* problem.

Definition 2.4.2. Weak duality

Denote $d^\star$ the optimal cost of the Lagrange dual problem, and $f^\star$ the optimal cost of the primal problem. Then, $d^\star$ is the tightest lower bound on $f^\star$. This property is called *weak duality* and it holds even if the primal problem is not convex.

Remark. Weak duality still holds when $f^\star$ and $d^\star$ are infinite. This provides a tool for feasibility inspection: unbounded dual (primal) corresponds to infeasible primal (dual).

The *optimal duality gap*, is defined as:

$$D^\star_{gap} = f^\star - d^\star \geq 0. \tag{2.4.8}$$

Definition 2.4.3. Strong duality

If the equality

$$d^\star = f^\star$$

holds, namely the duality gap is zero, the property of *strong duality* holds.

Strong duality does not hold in the general case. For a convex problem, strong duality can be guaranteed, for instance under *Slater's condition* [Boyd & Vandenberghe, 2004; Hindi, 2006]

Theorem 2.4.1. (Slater's condition)

If a convex (primal) problem is strictly feasible, in other words, if $\exists \bar{\chi} : F(\bar{\chi}) \preceq 0,\ h(\bar{\chi}) = 0,$ then strong duality holds.

Duality can therefore be used to provide bounds for non-convex problems, and stopping criteria for optimisation algorithms. For instance, in convex programming, the duality gap can be evaluated at each iteration, and a stopping criterion can be $D_{gap} \leq \epsilon$.

2.4.3 Optimality conditions

This section provides a set of conditions for optimality. These are then particularised for convex problems.

Definition 2.4.4. First-order optimality (KKT) conditions

Consider the non-linear program (2.4.3). Assume the cost and constraint functions are differentiable and that *strong duality* holds. Define the *Lagrangian* function as (2.4.4). Define $\chi^\star$ as the optimal solution of (2.4.3), $\lambda^\star$ and $\nu^\star$ as the vector of *Lagrange multipliers* at the optimiser for, respectively, inequality and equality constraints. The following first-order conditions, called the *Karush-Kuhn-Tucker conditions*, are necessary (and sufficient, for a convex problem) optimality conditions, [Hindi, 2006]

$$\begin{aligned} F_i(\chi^\star) &\leq 0, \\ h_i(\chi^\star) &= 0, \\ \lambda^\star &\geq 0, \\ \lambda_i^\star F_i(\chi^\star) &= 0, \\ \nabla f(\chi^\star) + (\nabla h(\chi^\star))^T \nu^\star + (\nabla F \chi^\star))^T \lambda^\star &= 0. \end{aligned} \tag{2.4.9}$$

The last condition in (2.4.9) says that the gradient of the Lagrangian must vanish at the optimum, while $\lambda_i^\star F_i(\chi^\star) = 0$, known as *complementary slackness*, is related to inequality constraint satisfaction and dual optimality.

For a QP in the form of (2.4.2), the KKT conditions specialise to the following:

Definition 2.4.5. First-order optimality (KKT) conditions (Quadratic Program)

Given a convex QP in the form of (2.4.2), the necessary and sufficient conditions for global optimality, namely the *KKT conditions*, are [Boyd & Vandenberghe, 2004; Maciejowski, 2002]

$$
\begin{aligned}
-\,\Omega\,\chi^\star - t &= -b, \\
-\Phi\,\chi^\star &= -\phi, \\
\nu^\star &\geq 0, \\
t &\geq 0, \\
t^T\lambda^\star &= 0, \\
H\,\chi^\star + \Phi^T\nu^\star + \Omega^T\lambda^\star &= -g.
\end{aligned} \tag{2.4.10}
$$

An overview of convex optimisation techniques commonly used in MPC can be found in [Boyd & Vandenberghe, 2004; Nocedal & Wright, 1999]. The most commonly used, in the field of MPC, are *active-set* and *interior-point* methods [Karmarkar, 1984]. These algorithms have polynomial complexity[1]. Recent developments provided encouraging results for real-time applications [Richter *et al.* , 2013; Wang & Boyd, 2010].

Most convex optimisation problems presented in this thesis have been solved using a free academic version of the multipurpose solver IBM ILOG CPLEX [IBM, 2013], available online through the IBM Academic Initiative.

2.4.4 Degeneracy and multiple optima

The *active set* of a QP is the set of constraints that are active at the optimum. A QP is said to be *primal degenerate* [Bemporad *et al.* , 2002a] if its active set contains more constraints than the number of decision variables. In this case, the optimal Lagrange multiplier vector may not be unique. A QP is said to be *dual degenerate* [Bemporad *et al.* , 2002a] if its Lagrange dual is primal degenerate. A QP is dual degenerate if its cost function is not strictly positive definite, and its solution may not be unique [Bemporad *et al.* , 2002a].

A result that will often be recalled throughout the thesis is the following:

Lemma 2.4.2. ([Nesterov, 2009], Lemma 6)

Consider a function $f:\ \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$, strongly convex with respect to an ℓ_p-norm, with

[1] in the worst case a complexity of $\mathcal{O}(N^3m^3)$, where N is the MPC prediction horizon length and m is the number of control inputs.

convexity parameter σ. Consider also a closed convex set $\mathcal{X} \subseteq \mathbf{dom} f$. Then, the problem

$$\min_{\chi \in \mathcal{X}} f(\chi) \tag{2.4.11}$$

is solvable and its solution, $\chi^\star$, is *unique*. Moreover, for any $\chi \in \mathcal{X}$ we have

$$f(\chi) \geq f(\chi^\star) + \frac{1}{2}\sigma\|\chi - \chi^\star\|_p^2. \tag{2.4.12}$$

2.4.5 Exact penalty functions

This section is based on the definitions given in Chapter 7 of [Kerrigan, 2000], which follows [Fletcher, 1987]. Consider the non-linear program (2.4.3) and, without loss of generality, assume only the inequality constraints[1]. Consider then the following unconstrained non-smooth program

$$\chi_s^\star = \arg\min_{\chi} f(\chi) + \rho\|F(\chi)^+\|, \tag{2.4.13}$$

where $\|\cdot\|$ can be any norm, $F(\chi)^+$ contains the magnitude of the violation of the equality constraint of (2.4.3) for a given χ, namely, $F(\chi)_j^+ = \max(F(\chi)_j^+, 0)$, and $\rho > 0$ is the *constraint violation penalty weight*.

Definition 2.4.6. Define the *dual norm* of $\|\cdot\|$ as

$$\|u\|_D = \max_{\|v\|\leq 1} u^T v. \tag{2.4.14}$$

Remark 2. It can be shown that $\|\cdot\|_1$ is the dual of $\|\cdot\|_\infty$ and vice versa, and that $\|\cdot\|_2$ is the dual of itself (see A.1.6 of [Boyd & Vandenberghe, 2004]).

Theorem 2.4.3. Exact penalty function ([Fletcher, 1987], Th. 14.3.1)

Denote $\chi^\star$ and $\lambda^\star$ as, respectively, the minimiser and Lagrange multiplier vector of (2.4.3), and $\chi_s^\star$ as the minimiser of (2.4.13). If the penalty weight $\rho > \|\lambda^\star\|_D$ and $F(\chi_s^\star) \leq 0$, then $\chi^\star = \chi_s^\star$. In this case, the constraint violation penalty is said to be an *exact penalty function*.

The above result will be used in Chapters 6 and 7.

[1] Equality constraints can be, for instance, accommodated as back-to-back inequalities, $[h(\chi)^T, \ -h(\chi)^T]^T \leq 0$. This is generally not done in practise, as it makes the optimisation problem primal degenerate and more difficult to solve.

2.5 Constrained control systems in discrete-time domain

The theoretical development in this thesis concerns the control of systems that are time-invariant and that evolve in discrete-time. The focus is on constrained Linear Time-Invariant (LTI) systems, however, closed-loop stability is investigated for the class of non-linear systems with differentiable vector field and with the origin as equilibrium. The numerical tools currently available are reviewed, however, while the concepts upon which stability is based are becoming quite standard for constrained LTI systems, research is still active for the considered non-linear systems class where computational complexity is still a strong limitation.

The considered non-linear systems are of the form:

$$x(k+1) = f(x(k), u(k), w(k)), \quad \text{with } f(0,0,0) = 0, \tag{2.5.1}$$

where $x(k)$ is the system's *state vector*, $u(k)$ is the *input vector*, $w(k)$ is a *disturbance* or *uncertainty vector* at time k. The *vector field* $f : \mathbb{R}^{n \times m \times q} \to \mathbb{R}^n$ is assumed to be continuous and differentiable. The variable $x(k)$ is assumed to be measured at each k. Most theoretical results in this thesis will concern a particular subclass of (2.5.1), the LTI systems:

$$x(k+1) = Ax(k) + Bu(k) + B_w w(k). \tag{2.5.2}$$

Systems (2.5.1–2.5.2) are subject to convex constraints containing the origin of the form:

$$u(k) \in \mathbb{U} \subset \mathbb{R}^m, \quad x(k) \in \mathbb{X} \subset \mathbb{R}^n, \quad w(k) \in \mathbb{W} \subset \mathbb{R}^q, \forall k \in \mathbb{I}_{\geq 0}. \tag{2.5.3}$$

The closed loop expressions of (2.5.1), (2.5.2) under a control law $u(k) = \kappa(x(k))$ are, respectively,

$$x(k+1) = f(x(k), \kappa(x(k)), w(k)), \tag{2.5.4}$$

$$x(k+1) = Ax(k) + B\kappa(x(k)) + B_w w(k). \tag{2.5.5}$$

The nominal dynamics of (2.5.1) and (2.5.2) are, respectively,

$$x(k+1) = \bar{f}(x(k), u(k)) = f(x(k), u(k), 0), \tag{2.5.6}$$

$$x(k+1) = Ax(k) + Bu(k). \tag{2.5.7}$$

Depending on the context, the notation $\bar{x}$ will occasionally be used to denote the nominal evolution of x. The closed loop expressions of (2.5.6), (2.5.7) under a control law $u(k) = \kappa(x(k))$ are, respectively,

$$x(k+1) = \bar{f}(x(k), \kappa(x(k))), \tag{2.5.8}$$

$$x(k+1) = Ax(k) + B\kappa(x(k)). \tag{2.5.9}$$

Definition 2.5.1. Nominal i-th step evolution under a policy

The *i-th step evolution* of the nominal system (2.5.6) from an initial state $x(0)$, under an admissible input sequence $\mathbf{u}_{[i-1]} = \{u(k)\}_0^{i-1} \in \mathbb{U}^i$, is defined as

$$\bar{\phi}(i, x(0), \mathbf{u}_{[i-1]}) = \circ_{k=0}^{i-1} \bar{f}(x(k), u(k)),\ u(k) \in \mathbf{u}_{[i-1]}\ \forall k \in (0,\ i-1), \tag{2.5.10}$$

where $\circ_{k=0}^{i-1}$ represents composition of $\bar{f}$ with itself and $u(k)$ from time 0 to $i-1$, namely,

$$\bar{\phi}(i, x(0), \mathbf{u}_{[i-1]}) = \bar{f}(\bar{\phi}(i-1, x(0), \mathbf{u}_{[i-2]}), u(k)), \quad \bar{\phi}(0, x(0), \cdot) = x(0).$$

2.5.1 Stability definitions

This section concerns fundamental notions for studying stability of the dynamic systems of interest. The results are widely used throughout the thesis.

Remark 3. For the sake of brevity, in the following statements it will be implicitly assumed that for a considered set $\mathcal{X} \subseteq \mathbb{X}$ the considered control law $\kappa(x)$ is *admissible*, that is, $\kappa(x) \in \mathbb{U}, \forall x \in \mathcal{X}$.

Definition 2.5.2. Positively invariant (PI) set

A set $\mathcal{X} \subseteq \mathbb{R}^n$ is said to be *positively invariant* for (2.5.6) under a control law $u(k) = \kappa(x(k))$ if

$$\bar{f}(x, \kappa(x)) \in \mathcal{X},\ \forall x \in \mathcal{X}.$$

Definition 2.5.3. Robustly positively invariant (RPI) set

A set $\mathcal{X} \subseteq \mathbb{R}^n$ is said to be *robustly positively invariant* for (2.5.1) under a control law $u(k) = \kappa(x(k))$ if

$$f(x, \kappa(x), w) \in \mathcal{X},\ \forall x \in \mathcal{X},\ \forall w \in \mathbb{W}.$$

The following definitions are based on [Vidyasagar, 1993]:

Definition 2.5.4. $\mathcal{K}$-function

A continuous function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to be a $\mathcal{K}$-function ($\alpha \in \mathcal{K}$) if it is strictly increasing and if $\alpha(0) = 0$.

Definition 2.5.5. $\mathcal{K}_\infty$-function

A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to be a $\mathcal{K}_\infty$-function ($\alpha \in \mathcal{K}_\infty$) if it is a $\mathcal{K}$-function and if it is radially unbounded, that is $\alpha(r) \to \infty$ as $r \to \infty$.

Definition 2.5.6. $\mathcal{KL}$-function

A continuous function $\beta : \mathbb{R}^2_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to be a $\mathcal{KL}$-function ($\beta \in \mathcal{KL}$) if it is a $\mathcal{K}$-function in the first argument, if it is positive definite and non-increasing in the second argument, and if $\beta(r, t) \to 0$ as $t \to \infty$.

The following statements refer to the system's origin, but they can be extended to arbitrary steady states by a coordinate shift [Limón *et al.* , 2006; Rawlings & Mayne, 2010]:

Definition 2.5.7. Lyapunov stability

Given a PI set $\mathcal{X}$, $0 \in \textbf{int}\mathcal{X}$, the origin of system (2.5.8) is said to be *Lyapunov stable* in $\mathcal{X}$ if

$$\forall \epsilon > 0 \; \exists \delta > 0 : \forall x(0) \in \mathcal{X}, \|x(0)\| \leq \delta \Rightarrow \|x(k)\| < \epsilon, \; \forall k \in \mathbb{I}_{\geq 0}.$$

Definition 2.5.8. Attractivity

Given a PI set $\mathcal{X}$, $0 \in \textbf{int}\mathcal{X}$, the origin of system (2.5.8) is said to be *attractive* in $\mathcal{X}$ if

$$\lim_{k \to \infty} \|x(k)\| = 0, \; \forall x(0) \in \mathcal{X}.$$

Definition 2.5.9. Asymptotic stability (AS)

Given a PI set $\mathcal{X}$, $0 \in \textbf{int}\mathcal{X}$, the origin of system (2.5.8) is said to be *asymptoticaly stable* (AS) in $\mathcal{X}$ if it is Lyapunov stable and attractive in $\mathcal{X}$, equivalently, if there exists a $\mathcal{KL}$-function $\beta(\cdot, \cdot)$ such that

$$\|x(k)\| \leq \beta(\|x(0)\|, k), \; \forall x(0) \in \mathcal{X},$$

the PI set $\mathcal{X}$ is then called a *domain of attraction*.

Definition 2.5.10. Practical stability (pS) Given a PI set $\mathcal{X}$, $0 \in \mathbf{int}\mathcal{X}$, the origin of system (2.5.8) is said to be *practically stable* (pS) in $\mathcal{X}$ if there exists of a $\mathcal{KL}$-function $\beta(\cdot,\cdot)$ and a $\zeta > 0$ such that

$$\|x(k)\| \leq \beta(\|x(0)\|, k) + \zeta,\ \forall x(0) \in \mathcal{X}.$$

Remark 4. Regarding practical stability, a slight abuse of notation will be taken in this thesis by referring to the PI set $\mathcal{X}$ as a domain of attraction.

Definition 2.5.11. (practical) Lyapunov function

Let $\mathcal{X}$ be a compact PI set for system (2.5.8) containing a compact neighbourhood $\mathcal{N}$ of the origin in its interior and let $\alpha_1(\cdot)$, $\alpha_2(\cdot)$, $\alpha_3(\cdot)$ be $\mathcal{K}_\infty$-functions (and $\delta > 0$). A function $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$ with $V(0) = 0$ is a *practical Lyapunov function* in $\mathcal{X}$ if

$$V(x) \geq \alpha_1(\|x\|),\ \forall x \in \mathcal{X}, \tag{2.5.11}$$
$$V(x) \leq \alpha_2(\|x\|),\ \forall x \in \mathcal{N}, \tag{2.5.12}$$
$$V(\bar{f}(x, \kappa(x))) - V(x) \leq -\alpha_3(\|x\|) + \delta,\ \forall x \in \mathcal{X}. \tag{2.5.13}$$

Moreover, if (2.5.13) holds with $\delta = 0$, then $V(\cdot)$ is a *Lyapunov function*.

Theorem 2.5.1. Asymptotic stability

If system (2.5.8) admits a Lyapunov function in $\mathcal{X}$, then the system origin is AS with domain of attraction $\mathcal{X}$.

Proof of the above Theorem can be found, for instance, in [Lazar, 2006; Rawlings & Mayne, 2010].

Theorem 2.5.2. Practical stability

If system (2.5.8) admits a (practical) Lyapunov function in $\mathcal{X}$, then the system origin is pS with domain of attraction $\mathcal{X}$.

Definition 2.5.12. Control Lyapunov Function (CLF)

Let $\mathcal{X}$ be a compact set. A function $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$ with $V(0) = 0$ is a *control Lyapunov function* in $\mathcal{X}$ for system (2.5.6) iff $\exists \kappa(x)$ such that $V(\cdot)$ is a Lyapunov function in $\mathcal{X}$ for (2.5.8).

Remark 5. An important property of (control) Lyapunov functions is that their *level sets* are positively *(control) invariant sets* for the considered system [Blanchini, 1999].

Define the disturbance sequence $\mathbf{w}_{[k]} = \{w_j\}_0^k$. The following definition generalises the concept of stability to uncertain systems [Limón *et al.* , 2006].

Definition 2.5.13. Input-to-state (practical) stability

Given a RPI set $\mathcal{X}$, $0 \in \mathbf{int}\mathcal{X}$, the origin of system (2.5.4) is said to be *(locally) input-to-state practically stable*, or ISpS, in $\mathcal{X}$ if there exists of a $\mathcal{KL}$-function $\beta(\cdot,\cdot)$ and a $\mathcal{K}_\infty$-function γ (and a $\zeta > 0$) such that

$$\|x(k)\| \leq \beta(\|x(0)\|, k) + \gamma(\|\mathbf{w}_{[k-1]}\|) + \zeta,\ \forall x(0) \in \mathcal{X},\ \forall \mathbf{w} : w_j \in \mathbb{W}\ \forall j \in \mathbb{I}_{\geq 0},\ \forall k \in \mathbb{I}_{\geq 0}, \tag{2.5.14}$$

where $\|\mathbf{w}_{[k-1]}\| = \max_{j \in \mathbb{I}_{[0,k-1]}}\{\|w_j\|\}$.

If (2.5.14) holds with $\zeta = 0$, then the origin is *(locally) input-to-state stable*, or ISS.

Definition 2.5.14. IS(p)S Lyapunov function

Let $\mathcal{X}$ be a compact RPI set for system (2.5.4) containing a compact neighbourhood of the origin $\mathcal{N}$ in its interior and let $\alpha_1(\cdot)$, $\alpha_2(\cdot)$, $\alpha_3(\cdot)$ be $\mathcal{K}_\infty$-functions, $\sigma(\cdot)$ be a $\mathcal{K}$-function and $\delta \geq 0$. A function $V : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ with $V(0) = 0$ is an *IS(p)S Lyapunov function* in $\mathcal{X}$ if

$$V(x) \geq \alpha_1(\|x\|),\ \forall x \in \mathcal{X}, \tag{2.5.15}$$

$$V(x) \leq \alpha_2(\|x\|),\ \forall x \in \mathcal{N}, \tag{2.5.16}$$

$$V(f(x, \kappa(x), w)) - V(x) \leq -\alpha_3(\|x\|) + \sigma(\|w\|) + \delta,\ \forall x \in \mathcal{X},\ \forall w \in \mathbb{W}. \tag{2.5.17}$$

If (2.5.17) holds with $\delta = 0$, then $V(\cdot)$ is an *ISS Lyapunov function*.

Theorem 2.5.3. Input-to-state stability

If system (2.5.4) admits an ISS Lyapunov function in $\mathcal{X}$ with respect to $\mathbb{W}$, then the system origin is ISS in $\mathcal{X}$ with respect to $\mathbb{W}$ with *ISS gain* $\sigma(\cdot)$.

Theorem 2.5.4. Input-to-state practical stability

If system (2.5.4) admits an ISpS Lyapunov function in $\mathcal{X}$ with respect to $\mathbb{W}$, then the system origin is ISpS in $\mathcal{X}$, with respect to $\mathbb{W}$, with *ISpS gain* $\sigma(\cdot)$.

The following results are adapted from [Limón *et al.* , 2006]:

Lemma 2.5.5. Uniform continuity and $\mathcal{K}$-functions

Given a uniformly continuous function $V(\cdot)$, there exists a $\mathcal{K}$-function $\sigma(\cdot)$, such that

$$|V(x+w) - V(x)| \leq \sigma(\|w\|),\ \forall x, w \in \mathbf{dom}V. \tag{2.5.18}$$

Theorem 2.5.6. Uniform continuity and ISS

Consider the C-sets $\mathcal{X}$, $\mathbb{W}$. Assume (2.5.4) to be continuous in w, $\forall x \in \mathcal{X},\ \forall w \in \mathbb{W}$[1]. Assume (2.5.8) to be AS in $\mathcal{X}$, then (2.5.4) is ISS in an RPI set $\mathbb{X}^r \subseteq \mathcal{X}$ if $\mathbb{X}^r \neq \emptyset$ and if, either,

1. There exists a continuous Lyapunov function in $\mathcal{X}$ for system (2.5.8),
2. The function $\bar{f}(x, \kappa(x))$, in (2.5.8), is continuous $\forall x \in \mathcal{X}$.

2.5.2 Set-theoretic control

The following definitions and theorems are based on [Blanchini, 1999; Blanchini & Miani, 2008]:

Definition 2.5.15. Control invariant and λ-contractive set

A C-set $\mathcal{X}$ is λ-contractive for (2.5.6) if there exists a $\lambda \in [0, 1)$, and u such that $x \in \mathcal{X} \Rightarrow \bar{f}(x, u) \in \lambda\mathcal{X}$. If this holds for $\lambda = 1$, then the set $\mathcal{X}$ is simply *control invariant*.

If the above condition holds (with $\lambda \in [0, 1)$) for system (2.5.1) $\forall w \in \mathbb{W}$, then the set $\mathcal{X}$ is a λ-contractive set for (2.5.1). If $\lambda = 1$, then the set is simply *robustly control invariant*.

Definition 2.5.16. Minkowski function

Given a C-set $\mathcal{X}$ its *Minkowski function* or gauge function is

$$\psi_{\mathcal{X}}(x) = \min_{\gamma}\{\gamma \mid x \in \gamma\mathcal{X}\}. \tag{2.5.19}$$

Theorem 2.5.7 (Blanchini [1999], Th. 3.3)**.**

The following statements are equivalent:

- The C-set $\mathcal{X}$, with $0 \in \mathbf{int}(\mathcal{X})$, is λ-contractive, for (2.5.6),
- The Minkowski function of $\mathcal{X}$, $\psi_{\mathcal{X}}(x)$, is a CLF in $\mathcal{X}$ for (2.5.7). Hence, $\exists\kappa(x)$ such that (2.5.8) is AS in $\mathcal{X}$.

[1] Uniform continuity is implied by continuity if all the involved sets are bounded, e.g. C-sets, (see Lemma 2.3.3).

Definition 2.5.17. Ultimate boundedness

Let $\mathcal{S}$ be a neighbourhood of the origin ($0 \in \textbf{int}(\mathcal{S})$). System (2.5.8) is *ultimately bounded* in $\mathcal{S}$ if $\exists\, \mathcal{X} \supset \mathcal{S}$ such that,

$$\forall x(0) \in \mathcal{X}, \exists\, \bar{k} \in \mathbb{I}_{\geq 0} : x(k) \in \mathcal{S}, \forall k \geq \bar{k}.$$

Definition 2.5.18. Lyapunov function outside a set

Let $\mathcal{S}$ be a neighbourhood of the origin. A function $V(\cdot)$ is a *Lyapunov function outside* $\mathcal{S}$ if $\exists\, c : \mathcal{B}_{V,c} = \{x \mid V(x) \leq c\} \subseteq \mathcal{S}$ and if $\exists\, \mathcal{X} \supset \mathcal{S}$ such that $V(\cdot)$ is a Lyapunov function in $\mathcal{X} \backslash \mathcal{B}_{V,c}$.

Theorem 2.5.8. Ultimate boundedness (Blanchini & Miani [2008], Th. 2.24)

If system (2.5.8) admits a Lyapunov function outside $\mathcal{S}$ (from $\mathcal{X}$), then it is ultimately bounded in $\mathcal{S}$ (for all $x(0) \in \mathcal{X}$).

Remark 6. Robust versions of the above statements can be obtained by imposing the corresponding conditions to hold $\forall w \in \mathbb{W}$.

For the case of LTI systems and polytopic constraints, the following holds:

Theorem 2.5.9 (Blanchini & Miani [2008], Cor. 4.43)**.**

A polyhedral C-set $\mathcal{X}$ is λ-contractive for system (2.5.2) iff for each of its vertices $v \in \mathcal{V}(\mathcal{X})$ there exists a control $u_{\{v\}}$ such that, $\forall w \in \mathbb{W}$, we have $Av + Bu_{\{v\}} + B_w w \in \lambda\mathcal{X}$, where $\lambda \in [0, 1)$.

Corollary 2.5.10. Theorem 2.5.9 applies to the nominal system (2.5.7), by taking $\mathbb{W} = \{0\}$.

Set theoretic control makes use of a set of operators. For the study of predictive controllers, the following notions are required [Kerrigan, 2000; Kerrigan & Maciejowski, 2000a]:

Definition 2.5.19. 1-step operator

Given the system (2.5.6), subject to constraints (2.5.3), the 1-*step operator* is defined as

$$\mathcal{Q}(\mathcal{X}) = \{x \in \mathbb{R}^n \;:\; \exists u \in \mathbb{U} \mid \bar{f}(x, u) \in \mathcal{X}\}. \tag{2.5.20}$$

Recalling Definition 2.5.1, we can introduce the following [Kerrigan, 2000; Kerrigan & Maciejowski, 2000a]:

Definition 2.5.20. Controllable set

Given the system (2.5.6), subject to constraints (2.5.3), the *N-step controllable set* from $\mathcal{X} \subseteq \mathbb{R}^n$ to $\mathbb{X}_f \subseteq \mathbb{R}^n$ is defined as

$$\mathcal{K}_N(\mathcal{X}, \mathbb{X}_f) = \{x \in \mathcal{X} : \exists \mathbf{u}_{[N-1]} \in \mathbb{U}^N \mid \hat{x}(j) = \bar{f}^j(x, \mathbf{u}_{[j-1]}) \in \mathcal{X} \, \forall j \in \mathbb{I}_{[0,N]}, \; \hat{x}(N) \in \mathbb{X}_f\}. \tag{2.5.21}$$

Remark 7. As shown in [Kerrigan, 2000; Kerrigan & Maciejowski, 2000a], the N-step controllable set can be computed by recursive application of the 1-step operator. Namely,

$$\mathcal{K}_{i+1}(\mathcal{X}, \mathbb{X}_f) = \mathcal{K}_1(\mathcal{X}, \mathcal{K}_i(\mathcal{X}, \mathbb{X}_f)), \tag{2.5.22}$$

$$\mathcal{K}_1(\mathcal{X}, \mathbb{X}_f) = \mathcal{Q}(\mathbb{X}_f) \cap \mathcal{X}. \tag{2.5.23}$$

In the LTI case, the $\mathcal{Q}$ operator can be implemented using polytopic projections [Kerrigan, 2000].

Definition 2.5.21. Stabilisable set

Given the system (2.5.6), subject to constraints (2.5.3), and given a positively invariant set $\mathbb{X}_f \subseteq \mathbb{X}$, the *N-step stabilisable set* from $\mathcal{X} \subseteq \mathbb{X}$, associated with $\mathbb{X}_f$, is defined as

$$\mathcal{S}_N(\mathcal{X}, \mathbb{X}_f) = \mathcal{K}_N(\mathcal{X}, \mathbb{X}_f). \tag{2.5.24}$$

Note that the above definitions are for the *nominal system*. Robust cases are similarly defined in [Kerrigan, 2000], with the difference that the above conditions must hold $\forall w(k) \in \mathbb{W}$, $\forall k \in \mathbb{I}_{[0,N]}$.

2.6 Model Predictive Control

Model Predictive Control (MPC) is a popular optimal control strategy, in which a constrained finite horizon optimal problem is solved at each iteration on the basis of the available measurement. A plant model is used to compute N future predictions, and an input sequence is generated optimising a performance index. The first control move of the optimal sequence is then applied to the plant. The main strength of predictive control is the ability to handle physical and operational constraints on the plant behaviour as well as on the inputs. MPC is capable of operating the system safely within constraints and can improve performance by allowing operating points to be close to constraint boundaries [Maciejowski, 2002]. Operation points are implicitly determined by the constraints, the prediction horizon length N, and by the performance index or *cost function*.

It is becoming standard, among researchers, to use a *terminal constraint* in MPC [Bemporad *et al.* , 2002a; Chen & Allgöwer, 1998; Chmielewski & Manousiouthakis, 1996; Goodwin *et al.* , 2005; Kerrigan & Maciejowski, 2000a; Maciejowski, 2002; Mayne *et al.* , 2000; Rawlings & Mayne, 2010]. In particular, the state prediction at the end of the horizon is constrained to be into a (nominal or robust) control invariant set (thus called the *terminal set*). The advantage of using this *safe set* is that recursive feasibility is guaranteed for the optimisation problem. Consequently, the feasible region is (nominally or robustly) positively invariant for the closed loop system [Kerrigan & Maciejowski, 2000a]. Additionally, a particular CLF is used to penalise the state prediction (or the control error) at the end of the horizon (thus called the *terminal cost*). These two terminal ingredients are used in most MPC literature to satisfy a set of neat sufficient conditions for (nominal and robust) closed-loop stability [Mayne *et al.* , 2000]. This setting is considered for the majority of the thesis.

2.6.1 Standard MPC formulation

Most MPC implementations are based on a quadratic cost, mainly because quadratic functions are convex and twice differentiable (hence relatively easy to minimise) and partly because of the possibility of inheriting properties from the LQR [Chmielewski & Manousiouthakis, 1996]. The aim of most common MPC strategies is to steer the nominal system (2.5.6) to the origin of the state space (or to a different set-point), while keeping its evolution within constraints and minimising the time summation of N stage costs and of a terminal cost (or cost-to-go). A standard MPC with quadratic cost, or *quadratic MPC*, can be formulated as follows:

Definition 2.6.1. (Quadratic MPC)
Consider the following finite-horizon constrained optimal control problem (FHCOCP)

$$\begin{aligned} V_N^o(x) = \min_{\underline{\mathbf{u}}} \left\{ V_N(x, \underline{\mathbf{u}}) \hat{=} F(\hat{x}_N) + \sum_{j=0}^{N-1} \ell(\hat{x}_j, \hat{u}_j) \right\} \\ s.t. \quad \hat{x}_{j+1} = \hat{f}(\hat{x}_j, \hat{u}_j), \\ \hat{u}_j \in \mathbb{U}, \ \ \forall j \in \mathbb{I}_{[0,\ N-1]},\ \hat{x}_j \in \mathbb{X},\ \forall j \in \mathbb{I}_{[0,\ N]}, \\ \hat{x}_0 = x,\ \hat{x}_N \in \mathbb{X}_f \subseteq \mathbb{X}, \end{aligned} \tag{2.6.1}$$

with stage cost

$$\ell(x, u) = x^T Q x + u^T R u + 2u^T T x, \tag{2.6.2}$$

and terminal cost (cost-to-go)

$$F(x) = x^T P x, \tag{2.6.3}$$

where $\hat{f} : \mathbb{R}^{n \times m} \to \mathbb{R}^n$ is a continuous prediction model, $\underline{\mathrm{u}}^T = \begin{bmatrix} \hat{u}_0^T, & \cdots, & \hat{u}_{N-1}^T \end{bmatrix}$. At each iteration k, the MPC applies to the plant the first move of the optimal policy, $u(k) = \hat{u}_0^\star$, obtained by online solution of (2.6.1-2.6.2), at the current state, $x(k)$. The resulting implicit control law is referred to as $K_N(x) \equiv \hat{u}_0^\star(x)$. The closed-loop evolution of (2.5.1) under quadratic MPC is

$$x(k+1) = f(x(k), K_N(x(k)), w(k)). \tag{2.6.4}$$

The one-step evolution of (2.6.4) for a given x is simply denoted as $x^+ = f(x, K_N(x), w)$.

Formulations based on 1 and ∞-norm costs, $\ell(x,u) = \|x\|_p + \|u\|_p,\ p \in \{1, \infty\}$, have also been investigated in the MPC literature, being the most common alternatives to Quadratic MPC [Maciejowski, 2002; Rao & Rawlings, 2000; Rawlings & Mayne, 2010]. Although they may be useful in some cases, for instance when an LP solver is available instead of a QP solver, formulations based on 1 and ∞ norm often display less desirable performance [Rao & Rawlings, 2000]. Throughout this thesis, quadratic, ℓ_1 and ℓ_∞ MPC will be compared to ℓ_{asso}-MPC.

2.6.2 Considered setup

It is common in the MPC literature to assume that $\hat{f}(\cdot,\cdot) \equiv \bar{f}(\cdot,\cdot)$, namely, that the prediction model corresponds to the nominal plant model. In this thesis, the following is assumed:

Assumption 1. (A1)

(H0) $\hat{f}(\cdot,\cdot)$, $\ell(\cdot,\cdot)$, $F(\cdot)$ are continuous, $\hat{f}$ differentiable, $\hat{f}(0,0) = 0$, $\ell(0,0) = 0$, $F(0) = 0$.

The constraints and the initial state are assumed to satisfy

Assumption 2. (A2)

(H1) $\mathbb{X}$, $\mathbb{X}_f$, and $\mathbb{U}$ are C-sets, containing the origin in their interior,

(H2) $x(0) \in \mathbb{X}_N = \{x \in \mathbb{X} \mid \exists\, \underline{\mathrm{u}} \in \mathbb{U}^N,\ V_N(x, \underline{\mathrm{u}}) < \infty\}$.

In (A2) the MPC cost function is implicitly assumed to be *extended*. The following result is standard in the MPC literature:

Proposition 2.6.1 (Rawlings & Mayne [2010])**.** Assume (A1), (A2). Then, a solution exists for the MPC (2.6.1).

From Proposition 2.6.1, a solution exists for problem (2.6.1), when a differentiable $\hat{f}$ is used. Finding the optimal solution, however, requires to solve a non-convex minimisation, which could be subject to local optima.

Remark 8. As it is common in the MPC literature, in this thesis it is assumed that the global optimum is found at each MPC iteration. The implications of suboptimal solutions for non-convex MPC are discussed, for instance, in [Lazar & Heemels, 2009; Pannocchia *et al.* , 2011; Scokaert *et al.* , 1999]. In particular, under some conditions, suboptimal solutions can be tolerated, with small loss of performance.

Remark 9. Continuity of $\hat{f}$ (or f) is not strictly necessary for using MPC, however, theoretical development for discontinuous systems is beyond the scope of this thesis. The interested reader can refer, for instance, to [Lazar & Heemels, 2009; Lazar *et al.* , 2006] and reference therein.

2.6.3 Recursive feasibility and invariance

In order to safely implement an MPC control law, it is required that the associated optimisation is *feasible for all time*, according to the following [Kerrigan & Maciejowski, 2000a]:

Definition 2.6.2. Feasible for all time

The MPC problem is *feasible for all time* $k \in \mathbb{I}_{\geq 0}$ if and only if **(H2)** holds and all future evolutions of the state belong to the feasible set, namely, $x(k) \in \mathbb{X}_N,\ \forall k \in \mathbb{I}_{\geq 0}$.

Define $C_\infty(\mathbb{X}) = \mathcal{K}_\infty(\mathbb{X}, \mathbb{X})$ as the infinite-time control invariant set in $\mathbb{X}$ [Kerrigan & Maciejowski, 2000a]. The following result provides sufficient conditions for guaranteeing that the MPC problem is feasible for all time:

Theorem 2.6.2. [Kerrigan & Maciejowski, 2000a]

The MPC problem is *feasible for all time* if and only if $x(0) \in \mathbb{X}_N \cap C_\infty(\mathbb{X})$ and $\forall x(k) \in \mathbb{X}_N \cap C_\infty(\mathbb{X})$ the solution to the MPC problem results in $x_1^\star \in \mathbb{X}_N \cap C_\infty(\mathbb{X})$.

If an MPC is feasible for all the time for every feasible initial condition, then we say that it is *recursively feasible*. The following results are from [Kerrigan & Maciejowski, 2000a]:

Definition 2.6.3. Strongly (or recursively) feasible MPC

The MPC problem is *strongly feasible* if and only if for all $x(0) \in \mathbb{X}_N$ the MPC problem is feasible for all time.

Theorem 2.6.3. [Kerrigan & Maciejowski, 2000a]

1. The MPC problem is strongly feasible if and only if the feasible set is a positively invariant set for the (nominal) closed-loop system $x(k+1) = \bar{f}(x(k), K_N(x(k)))$.

2. The MPC problem is strongly feasible only if the feasible set is a control invariant set for the system $x(k+1) = \bar{f}(x(k), u(k))$.

Thanks to Theorem 2.6.3 the conditions for recursive feasibility are translated into set-theoretic control conditions. The next section deals with common MPC assumptions guaranteeing asymptotic stability and, indirectly, the invariance of $\mathbb{X}_N$ and recursive feasibility of the MPC problem when $w(k) = 0, \forall k$.

Remark 10. In the perturbed case ($w(k) \neq 0$ for some k), the notion of *robustly feasible MPC* is introduced, which means that the MPC is strongly feasible $\forall w \in \mathbb{W}$. This (strong) condition is verified if $\mathbb{X}_N$ is robustly control invariant [Kerrigan & Maciejowski, 2000a].

2.6.4 Sufficient conditions for asymptotic stability

A standard set of sufficient assumptions for MPC stability is the following [Rawlings & Mayne, 2010]:

Assumption 3. (A3)

There exist two $\mathcal{K}_\infty$-functions, $\alpha_1,\ \alpha_2$, such that

(H5) $\alpha_1(\|x\|) \leq \ell(x, u),\ \forall x \in \mathbb{X}_N, \forall u \in \mathbb{U}$,

(H6) $F(x) \leq \alpha_2(\|x\|),\ \forall x \in \mathbb{X}_f$,

(H7) $\min_{\{u \in \mathbb{U}\}}\{F \circ f(x, u) + \ell(x, u) \mid f(x, u) \in \mathbb{X}_f\} \leq F(x),\ \forall x \in \mathbb{X}_f$.

It can be noticed that **(H7)** implicitly assumes the control invariance of $\mathbb{X}_f$.

Theorem 2.6.4. [Rawlings & Mayne, 2010] Assume (A1), (A2), (A3). Then, the origin of (2.5.8) is AS in $\mathbb{X}_N$ under the MPC controller, $\kappa(x) = K_N(x)$, with prediction model $\hat{f} \equiv \bar{f}$.

Proof. (Sketch) The result can be obtained, for instance, by means of the *direct method* [Rawlings & Mayne, 2010]. The proof rationale is reported. At time $k+1$, take the previous state and input prediction sequences without their first step. This sequences are still feasible and, since the terminal prediciton at time k was in $\mathbb{X}_f$, one can append to the control sequence the move satisfying (**H7**), obtaining a feasible suboptimal control solution, $\tilde{\underline{\mathrm{u}}}$, with the new terminal state prediction being also in $\mathbb{X}_f$. Thanks to (**H7**), the MPC cost under $\tilde{\underline{\mathrm{u}}}$ is smaller than its value at time k, in particular, $V_N(x(k+1), \tilde{\underline{\mathrm{u}}}) \leq V_N^o(x(k)) - \ell(x(k), K_N(x(k)))$. By optimality, $V_N^o(x(k+1)) \leq V_N(x(k+1), \tilde{\underline{\mathrm{u}}})$, for any feasible $\tilde{\underline{\mathrm{u}}}$. Hence, under (**H5**)–(**H7**), the optimal cost $V_N^o(x)$ is a Lyapunov function in $\mathbb{X}_N$ for the nominal system, with $V_N^o(x(k+1)) - V_N^o(x(k)) \leq -\alpha_1(\|x(k)\|)$. Finally, by Theorem 2.5.1, the system origin is AS with domain of attraction $\mathbb{X}_N$. ∎

2.6.5 Terminal ingredients

In most common formulations, (**H5**) is satisfied by taking $Q \succ 0$. Then, (**H7**) and, consequently, (**H5**) are satisfied by means of a candidate state-feedback *terminal controller*[1]. Given $K \in \mathbb{R}^{m \times n}$, define the $i+1$-step evolution of (2.5.6) under $u = Kx$ as

$$\bar{f}_K^{i+1}(x) = \bar{f}_K^1 \circ \bar{f}_K^i(x), \ \forall i \in \ \mathbb{I}_{\geq 2}, \ \bar{f}_K^1(x) = \bar{f}(x, Kx), \ \bar{f}_K^0(x) = x. \tag{2.6.5}$$

The following assumption is the most common in the MPC literature:

Assumption 4. (A4) There exists a matrix K and a C-set $\mathbb{X}_f$, containing the origin, such that

$$F(\bar{f}_K^i(x)) \leq F(\bar{f}_K^{i-1}(x)) - \ell(\bar{f}_K^{i-1}(x), K\bar{f}_K^{i-1}(x)) \quad \forall x \in \mathbb{X}_f, \ \forall i \in \mathbb{I}_{\geq 0}. \tag{2.6.6}$$

It is common practice, for differentiable systems and a single set point, to linearise the model around the desired equilibrium, then to compute a K that stabilises the linearisation (if stabilisable). Lyapunov theory is then used to determine an admissible invariant set [Chen & Allgöwer, 1998]. This, however, can be very conservative. A review of common approaches follows.

2.6.5.1 Linear Quadratic MPC

For LTI systems and quadratic cost, it is commonly assumed that [Rawlings & Mayne, 2010]:

Assumption 5. (A5) $Q \succeq 0$, $R \succ 0$, (A, B) is stabilisable and $(Q^{1/2}, A)$ detectable.

[1] The terminal controller is never applied to the plant.

In order to satisfy (A4), assume P and K to solve the following DARE

$$(A+BK)^T P(A+BK) - P = -(Q + K^T RK + K^T T + T^T K), \tag{2.6.7}$$

that characterises the discrete-time infinite horizon LQR [Astrom, 1970]. Then, the terminal constraint set can be computed by defining the *input admissible* set, $\bar{X}$, and the *maximal output admissible set*, $\mathcal{O}_\infty$, as [Chmielewski & Manousiouthakis, 1996; Kerrigan & Maciejowski, 2000a]

$$\bar{X} = \{x \in \mathbb{X}\,|Kx \in \mathbb{U}\},\ \mathcal{O}_\infty = \{x \mid (A+BK)^k x \in \bar{X},\ \forall\, k \geq 0\}. \tag{2.6.8}$$

Then, one can take either

$$\mathbb{X}_f = \mathcal{O}_\infty, \tag{2.6.9}$$

or [Bemporad *et al.* , 2002a]

$$\mathbb{X}_f = \{x \mid F(x) \leq c\},\ \text{where } 0 < c < c_m \hat{=} \inf_{x \notin \bar{X}} \{x^T P_\infty x\}. \tag{2.6.10}$$

Using (2.6.9) causes the MPC problem to be a convex QP, while using (2.6.9) gives a more demanding QCQP. Note that both sets from (2.6.9), (2.6.10) are PI under $u = Kx$, however, the $\mathcal{O}_\infty$ set in (2.6.9) is by definition the maximal of all invariant sets. Therefore, the choice of $\mathcal{O}_\infty$ is always less conservative than any other control invariant set, for the same K. Note that, if $(A+BK)$ has eigenvalues strictly inside the unit circle (for instance in the LQR case), or at particular locations on the circle, then $\mathcal{O}_\infty$ is a polytope, and it can be computed in finite time by intersecting a finite number of linear inequalities [Chmielewski & Manousiouthakis, 1996; Kerrigan, 2000]. The locations on the unit circle are characterised in Theorem 4.47 of [Blanchini & Miani, 2008].

2.6.5.2 Non-linear quadratic MPC and differential inclusion

Computation of K and $\mathbb{X}_f$ for (A4), is generally non-trivial. Taking $\mathbb{X}_f = \{0\}$ works theoretically, but it is not recommended in practice, as this could cause numerical difficulties and a small feasible region. Recall that this thesis considers differentiable systems with the origin as equilibrium and take $(\bar{A}, \bar{B})$ from the linearisation of (2.5.6) at the origin. Then, a possible solution is the following (a novel discrete-time version of Lemma 1 of [Chen & Allgöwer, 1998]):

Lemma 2.6.5. Assume $(\bar{A}, \bar{B})$ is stabilisable, $T = 0$. Then there exists a $K \in \mathbb{R}^{m \times n}$, $P \succ 0$, and a scalar, $\alpha > 0$, such that if $\mathbb{X}_f = \{x \in \mathbb{X} \mid F(x) \leq \alpha\}$ then (A4) is satisfied.

Proof. (Lemma 2.6.5) Define $v(x) = \bar{f}(x, Kx) - (\bar{A} + \bar{B}K)x$. Then, (2.5.6) is equivalent, in closed loop with $u = Kx$, to

$$x(k+1) = \tilde{f}(x) \equiv (\bar{A} + \bar{B}K)x + v(x). \tag{2.6.11}$$

From [Lazar *et al.* , 2009] an appropriate matrix gain K can be designed, by means of an LMI, such that

$$\tilde{f}(x)^T P \tilde{f}(x) - x^T P x \leq -x^T \tilde{Q} x + \tau v(x)^T v(x), \tag{2.6.12}$$

for some $P \succ 0$, given $\tilde{Q} \succ 0,\ \tau > 0$. Such $(K,\ P)$ always exists, since $(\bar{A}, \bar{B})$ is stabilisable. Define $\tilde{Q} = Q + K^T R K + \kappa I$, $\kappa > 0$. Define also $\mathcal{B}_{P,\alpha} = \{x \mid x^T P x \leq \alpha\}$, and take α_1 as the maximum constant for which Kx is admissible in $\mathcal{B}_{P,\alpha_1}$. Similarly to [Chen & Allgöwer, 1998], taking $\alpha \in\ (0, \alpha_1]$ such that

$$\max_{x \in \mathcal{B}_{P,\alpha}} \frac{\|v(x)\|_2^2}{\|x\|_2^2} \leq \kappa/\tau, \tag{2.6.13}$$

satisfies (2.6.6). ■

Remark 11. In the proof of Lemma 2.6.5, the gain κ/τ can be seen as a design parameter, that affects the size of $\mathbb{X}_f$.

Remark 12. In Lemma 2.6.5, continuity and differentiability of $\bar{f}$ are required only at the origin.

Alternative methods are available, when $\bar{f}$ is differentiable in a convex set, that potentially enlarge $\mathbb{X}_f$. These methods however *might require a significant increase of offline computation*. In Chapter 5, the terminal ingredients are computed by using Linear Differential Inclusion (LDI) [Blanchini & Miani, 2008; Boyd *et al.* , 1994]. In particular, the following results are considered:

Lemma 2.6.6. LPV embedding or model absorption

Assume $\bar{f}(x, u)$ is differentiable with $\bar{f}(0, 0) = 0$ and consider the artificial Linear Parameter-Varying (LPV) system:

$$\begin{aligned} x(k+1) &= A(r)x(k) + B(r)u(k), \\ A(r) &= \Sigma_i^{n_r} A_i r_i,\ B(r) = \Sigma_i^{n_r} B_i r_i, \\ &\qquad r \in \mathbb{R}^{n_r}_{\geq 0},\ \Sigma_i^{n_r} r_i = 1. \end{aligned} \tag{2.6.14}$$

If there exist a set $\mathcal{S}$ such that

$$\frac{\partial \bar{f}(x,u)}{\partial x, u} \in \mathbf{conv}\left([A_i,\ B_i],\ i \in \mathbb{I}_{[1,n_r]}\right),\ \forall x \in \mathcal{S},\ \forall u \in \mathbb{U}, \tag{2.6.15}$$

then, $\forall x \in \mathcal{S}$, $\forall u \in \mathbb{U}$, there exist a $\bar{r} = r(x,u) \in \mathbb{R}^{n_r}_{\geq 0}$ with $\Sigma_i^{n_r} \bar{r}_i = 1$ such that

$$\bar{f}(x,u) = A(\bar{r})x + B(\bar{r})u. \tag{2.6.16}$$

Then, $\forall x \in \mathcal{S}$, every trajectory of the original system (2.5.6) is also a trajectory of (2.6.14), for an appropriate law $r(k)$. We say that (2.6.14) *embeds* (or *absorbs*) (2.5.6) in $\mathcal{S}$.

Proof. Lemma 2.6.6 follows from the arguments in [Boyd *et al.* , 1994], page 55, sec. 4.3, which make use of the mean-value Theorem [Trench, 2013] Th. 2.3.10, page 83. ■

Theorem 2.6.7. Gain scheduling control

Assume that Lemma 2.6.6 holds. Assume also that, at each time k, a $\bar{r} = r(k)$ is available so that (2.6.16) is satisfied. If there exists a controller $\kappa(x, \bar{r})$ and a C-set $\mathcal{X} \subseteq \mathcal{S}$, positively invariant for (2.6.14) in closed-loop with $\kappa(x, \bar{r})$, then $\mathcal{X}$ is positively invariant also for (2.5.6) in closed-loop with $\kappa(x, \bar{r})$.

Proof. If $\kappa(x, r)$ keeps any trajectory of (2.6.14) inside $\mathcal{X} \subseteq \mathcal{S}$, given x and r, then from Lemma 2.6.6 there exist an appropriate law $\bar{r} = r(k)$ for which every trajectory of the original system (2.5.6) is also a trajectory of system (2.6.14) under $\kappa(x, \bar{r})$. Therefore, every trajectory of (2.5.6) is contained in $\mathcal{X}$, which is then positively invariant. ■

If Lemma 2.6.6 holds, then one can search for a control gain K, or for a more general a gain-scheduling terminal controller $K(r) = \sum_i^{n_r} K_i r_i$, that provides a control invariant set, $\mathbb{X}_f$, for (2.6.14) such that $\mathbb{X}_f \subseteq \mathcal{S}$. Then, from Lemma 2.6.6 and Theorem 2.6.7, the set $\mathbb{X}_f$ is also invariant for (2.5.6). Note that, for what concerns the MPC design, the parameter $\bar{r}$ never needs to be computed.

The matrices $(P,\ K(r))$ can be computed from the solution of the following set of matrix inequalities [Boyd *et al.* , 1994; Kothare *et al.* , 1996] at the parameter space vertices (here assuming $T = 0$):

$$(A_i + B_i K_i)^T P (A_i + B_i K_i) - P \leq -Q - K_i^T R K_i,\ \forall i \in \mathbb{I}_{[0,n_r]}. \tag{2.6.17}$$

If (2.6.17) admits a solution, for some $\mathcal{S}$, then $\mathbb{X}_f$ can be taken as a level set of $x^T P x$, or as any admissible robust invariant set inside $\mathcal{S}$. This will be proven in Chapter 5. A drawback of

LDI methods is that, depending on the system, a solution for (2.6.17) might not exist, for various candidate $\mathcal{S}$. Computation is also intense, and it grows dramatically with the number of parameters.

In Chapter 7 of [Blanchini & Miani, 2008], it is pointed out that, even when (2.6.17) is not solvable (a quadratic CLF cannot be found), the system (2.6.14) could still be robustly stabilisable, for instance with a more general PWA controller (e.g. an MPC). Constraints and dynamics are assumed to be polytopic, causing the largest robustly invariant set for (2.6.14) to be also a polytope. This polytope must satisfy the conditions of Theorem 2.5.9, for all A_i and B_i (see [Blanchini & Miani, 2008], Ch. 7.3). A technique for computing the maximal RPI set for (2.6.14) has been proposed in [Blanchini, 1999]. In particular, λ-contractive sets are constructed by means of polytope projections [Kerrigan, 2000]. This is basically a dynamic programming approach. Then, the resulting set is also λ-contractive for the original system (2.5.6). Polytope projections, required for the technique in [Blanchini, 1999], can be quite demanding. In particular, they can be performed only for relatively small state dimensions. A fundamental step for using the difference inclusion is to determine the set $\mathcal{S}$. A search is required for (relatively large) $\mathcal{S}$, for which the resulting system (2.6.14) is robustly stabilisable, under constraints. This idea is used, for LTI systems, in [Grammatico & Pannocchia, 2013], where $F(x)$ is proportional to the Minkowski function of $\mathbb{X}_f$.

The use of projections can be avoided (at the expense of a smaller terminal set) if, for instance, a stabilising LPV controller is known for system (2.6.14). Then, a λ-contractive set is computed for this closed loop system, as done for $\mathcal{O}_\infty$. Alternatively, one can use a matrix of candidate vertices, X, and then impose the conditions in Theorem 2.5.9 for all A_i and B_i. This is done in Chapter 5, by solving an LP. Similarly to [Grammatico & Pannocchia, 2013], a Minkowski terminal cost is proposed, the difference being the method for computing $\mathbb{X}_f$. In Chapter 5, as well as in [Gallieri & Maciejowski, 2013a,b; Grammatico & Pannocchia, 2013], the terminal set $\mathbb{X}_f$ is computed *a priori*, and it is independent from the cost function. Then, the terminal set is scaled, according to the stage cost. This facilitates fast online tuning, as computing $\mathbb{X}_f$ is effectively the main bottleneck in MPC design. Further alternatives exist, among which are [Bravo *et al.* , 2005; Fiacchini *et al.* , 2012].

2.6.6 Domain of attraction

Recall definitions, 2.5.1, 2.5.19, 2.5.20, 2.5.21. The following is a standard result of MPC theory:

Theorem 2.6.8. Assume (A1), (A2), (A3). Then, the feasible region of problem (2.6.1) is equivalent to the domain of attraction (DOA) of the closed-loop system (2.5.8), and it is given by

$$\mathbb{X}_N = \{x \in \mathbb{X} : \ \exists \underline{\mathbf{u}}_{[N-1]} \in \mathbb{U}^N | \bar{\phi}(i, x, \underline{\mathbf{u}}_{[i-1]}) \in \mathbb{X} \ \forall i \in \mathbb{I}_{[0,\ N]}, \ \bar{\phi}(N, x, \underline{\mathbf{u}}_{[N-1]}) \in \mathcal{O}_\infty\}. \quad (2.6.18)$$

or, equivalently,

$$\mathbb{X}_N = \mathcal{K}_N(\mathbb{X}, \mathbb{X}_f) = \mathcal{S}_N(\mathbb{X}, \mathbb{X}_f), \ \text{and} \ \mathbb{X}_N \supseteq \mathbb{X}_f. \quad (2.6.19)$$

The above results are connected to these in Section 2.6.3. In particular, recursive feasibility of the MPC problem is used to prove stability, and is obtained by means of a control invariant terminal set $\mathbb{X}_f$. The presence of this set, on the other hand, can limit the size of the feasible region and the system operations. The set-theoretic arguments in [Kerrigan & Maciejowski, 2000a] can be used to shown how the choice of $\mathbb{X}_f = \mathcal{O}_\infty$ provides a DOA that is at least as large as when a level set of the terminal cost is used instead. For LTI systems the maximal DOA can be obtained if, for instance, the approach of [Grammatico & Pannocchia, 2013] is used. This requires an increase of computation as discussed in Chapter 5.

Remark 13. The case of $c = 0$ is the trivial completion constraint, $x_N = 0$, which automatically satisfies the conditions for asymptotic stability of the origin for the considered nonlinear systems [Rawlings & Mayne, 2010]. However, for such a constraint one can expect the feasible region (the DOA) to be smaller than for a terminal constraint set, thus restricting the plant operating points. Moreover, since the predictions will reach the origin in finite time such a controller might present an unnecessarily high gain.

2.6.7 Sufficient conditions for intrinsic robustness

It is desirable for a feedback controller to possess a certain degree of robustness to bounded uncertainties. In the nonlinear case, this cannot be guaranteed a priori by a nominally stabilising controller, without performing a robust design [Grimm *et al.* , 2004].

In this thesis, robustness is characterised in terms of robust invariance and IS(p)S. To introduce the discussion, consider the following [Limon *et al.* , 2009]:

Theorem 2.6.9. Assume that Theorem 2.6.4 holds. Moreover, assume that either:

1. $f(x, K_N(x), w)$ is continuous $\forall x \in \mathbb{X}_N$, $\forall w \in \mathbb{W}$,

2. The MPC optimal cost $V_N^o(x)$ is continuous in $\mathbb{X}_f$.

Then, there exists an $r > 0$ such that, if $\mathbb{W} \subseteq \mathcal{B}_{\ell_p,r}(0)$, then the system (2.6.4) admits an RPI set $\mathbb{X}_N^r \subseteq \mathbb{X}_N,\ 0 \in \mathbf{int}\mathbb{X}_N^r$. Moreover, the system is ISS in $\mathbb{X}_N^r$.

Proof. By the Heine-Cantor Theorem 2.3.3, the uniform continuity assumed in Theorem 4 of [Limon *et al.* , 2009] can be relaxed to simple continuity. Then, similarly to Theorem 2.5.6, the existence of a uniformly continuous Lyapunov function $V(x)$ provides the satisfaction of the the ISS property $V(x(k+1)) \leq V(x(k)) - \alpha(\|x(k)\|) + \sigma(\|w\|)$, with $\alpha \in \mathcal{K}_\infty,\ \sigma \in \mathcal{K}$. As shown in the proof of Theorem 4 of [Limon *et al.* , 2009], uniform continuity of the model (the first assumption) provides the existence of such $V(x)$. Under the second assumption, $V(x)$ can be taken as the MPC optimal cost $V_N^o(x)$. The existence of the RPI set $\mathbb{X}_N^r$, given a sufficiently small $r > 0$, is shown by [Limon *et al.* , 2009] to be a result of the ISS property of $V(x)$, being $\mathbb{X}_N^r$ one of its level sets for which $x \in \mathbb{X}_N^r \Rightarrow \alpha(\|x\|) \geq \sigma(r)$. ■

Remark 14. Recall that in this thesis the considered systems models are assumed to be continuous, the cost function is continuous and convex. Therefore, $K_N(x)$ is also continuous and the intrinsic robustness result of Theorem 2.6.9 will apply to any MPC with nominal stability.

The ISS gain of the system is a measure of robust stability and performance. In MPC, this gain depends mainly on the terminal cost, $F(x)$, as will be further discussed in Chapter 7.

2.6.8 Explicit solution of linear quadratic MPC

The term *explicit MPC* [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002a] stands for the closed form representation of an MPC control law, computed offline. For LTI systems this can be done by means of multi-parametric linear or quadratic programming (respectively, mpLP and mpQP). This representation allows for the MPC law to be evaluated without online optimisation, at the cost of additional memory usage [Borrelli *et al.* , 2008, 2009]. Explicit MPC is generally used for applications with fast sampling, including for instance automotive [Di Cairano *et al.* , 2013] and power electronics [Mariéthoz *et al.* , 2012]. This form of MPC can also be useful for post-design analysis, as it provides further insight on the control law and the optimal cost. The material presented in this section will be used, in Chapter 4, to describe the solution of ℓ_{asso}-MPC, as well as in Chapter 6 for the synthesis of a particular type of ℓ_{asso}-MPC.

Procedures for computing the explicit MPC solution are presented for instance in [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002a; Tondel *et al.* , 2003]. The technique is quite consolidated for quadratic, 1 and ∞-norm costs and for systems with piece-wise affine dynamics [Alessio

& Bemporad, 2009]. The basic idea is to partition the feasible region $\mathbb{X}_N$ in a set of N_r *critical regions*, $\mathcal{P}_i$, each one associated with an *active set of constraints* $\mathcal{A}_i$. Complexity of the resulting controller is expressed in terms of N_r and depends (in the worst case exponentially) on the number of variables and constraints, as well as on the cost function parameters [Alessio & Bemporad, 2009; Borrelli *et al.* , 2009]. Several strategies have been proposed for the exploration of the feasible space, required to determine the critical regions, as well as for reducing N_r [Bemporad *et al.* , 2002a; Tondel *et al.* , 2003]. Computational complexity can become an issue for systems with for instance more than $n \geq 5$ states and horizon length $N \geq 5$.

Consider the Quadratic MPC defined by (2.6.1–2.6.3) with a LTI prediction model

$$\hat{f}(x, u) = Ax + Bu,$$

and closed polytopic constraints. The following matrices are widely used throughout the thesis

$$\Psi = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^N \end{bmatrix}, \quad \Theta = \begin{bmatrix} B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix}, \quad \begin{array}{l} \mathbf{Q} = \mathrm{BlockDiag}(I_{N-1} \otimes Q, P), \\ \mathbf{R} = I_N \otimes R, \\ H = (\Theta^T \mathbf{Q} \Theta + \mathbf{R}), \\ \Gamma = \Theta^T \mathbf{Q} \Psi. \end{array} \tag{2.6.20}$$

The Quadratic MPC problem (2.6.1–2.6.3) can be formulated as the following mpQP:

$$\begin{array}{l} V_N^o(x) = x^T Y x + \min_{\underline{u}} (\underline{u}^T H \underline{u} + 2\underline{u}^T \Gamma x) \\ \text{s.t. } \Omega\, \underline{u} \leq \underline{1} - Mx = b(x), \end{array} \tag{2.6.21}$$

where $Y = \Psi^T \mathbf{Q} \Psi + Q$, and the term $x^T Y x$ does not affect the optimal solution. With $R \succ 0$ and polyhedral constraints the following holds:

Theorem 2.6.10. [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002a]

Assume $H \succ 0$, $\begin{bmatrix} H, & \Gamma \\ \Gamma^T & Y \end{bmatrix} \succeq 0$. Then $\mathbb{X}_N$ is a polyhedral set, the value function $V_N^o : \mathbb{X}_N \to \mathbb{R}$ is continuous, convex and piecewise quadratic in x. The optimiser $\underline{u}^\star : \mathbb{X}_N \to \mathbb{U}^N$ is unique, continuous and piecewise affine in x.

Remark 15. Theorem 2.6.10 is not only useful to compute an explicit MPC solution. It provides a theoretical characterisation for $K_N(x)$ and $V_N^o(x)$.

Theorem 2.6.10 is proved by means of the KKT conditions for (2.6.21)

$$\begin{aligned} 2H\,\underline{\mathrm{u}}^\star(x) + 2\Gamma x + \Omega^T \nu^\star(x) &= 0, \\ -\Omega\,\underline{\mathrm{u}}^\star(x) - s^\star(x) &= -b(x), \\ (s^\star(x))^T \nu^\star(x) &= 0, \\ s^\star(x) &\geq 0, \\ \nu^\star(x) &\geq 0. \end{aligned} \tag{2.6.22}$$

An illustrative proof sketch of Theorem 2.6.10 follows, based upon [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002a; Kerrigan & Maciejowski, 2000b]. Solving (2.6.22) for $\underline{\mathrm{u}}^\star$ provides (since H is invertible)

$$\underline{\mathrm{u}}^\star(x) = -H^{-1}\left(\Gamma x + \frac{1}{2}\Omega^T \nu^\star(x)\right). \tag{2.6.23}$$

If constraints are not active at a particular x, the optimal solution (2.6.23) is linear in x and the LQ MPC will act as a linear controller. Given an active set of constraints, $\mathcal{A}(x)$, denote $\tilde{\nu}$ as the corresponding Lagrange multipliers (the remaining ones being equal to zero). The set of x with the same active set is called a *critical region*. Assume without loss of generality that (2.6.21) is not primal degenerate, namely that Ω has full row rank[1]. Then, substituting (2.6.23) into (2.6.22) provides

$$\tilde{\nu}^\star(x) = -2(\tilde{\Omega}H^{-1}\tilde{\Omega}^T)^{-1}\left(\underline{1} + (\tilde{\Omega}H^{-1}\Gamma - \tilde{M})x\right). \tag{2.6.24}$$

Therefore the optimal Lagrange multipliers are PWA in x. By substituting this back into (2.6.23) it is also possible to show the optimiser to be PWA in $x(k)$. In particular:

$$\underline{\mathrm{u}}^\star(x) = -H^{-1}\Gamma x + \tilde{\Omega}^T(\tilde{\Omega}H^{-1}\tilde{\Omega}^T)^{-1}\left(\underline{1} + (\tilde{\Omega}H^{-1}\Gamma - \tilde{M})x\right). \tag{2.6.25}$$

Continuity of the solution is shown in [Bemporad *et al.* , 2002a] by means of the adjacency of critical regions. These regions form in fact a partition of the feasible set. By similar arguments [Bemporad *et al.* , 2002a] has shown that the value function is continuous, convex (recall Theorem 2.3.7) and PWQ in $x(k)$.

The techniques used for explicit MPC aim to determine the active sets for all $x \in \mathbb{X}_N$ by means of ad-hoc exploration strategies [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002a; Kerrigan & Maciejowski, 2000b]. The parameters describing the first m elements of (2.6.25) are then stored

[1] Redundant inequalities can be removed at a preliminary stage.

in a look-up table. If numerical precision is an issue, for instance when constraints are close to degeneracy, the coefficients of (2.6.24)-(2.6.25) can also be computed with different methods, for instance using QR decomposition [Borrelli *et al.* , 2009; Fletcher, 1987].

The required computation for online evaluation of the control law, as well as the required storage, has been investigated for instance by [Borrelli *et al.* , 2008, 2009]. Define N_c^i the number of half-spaces describing $\mathcal{P}_i$, and $N_C = \sum_{i=o}^{N_r} N_c^i$. The simplest possible implementation, reviewed in [Borrelli *et al.* , 2008] and used in in Chapters 6 of this thesis, requires in the worst case $2nN_C$ floating point operations (flops) and a storage of $(n+1)N_C$ real numbers (including the inequalities for $\mathcal{P}_i$). On the other hand, the algorithm proposed in [Borrelli *et al.* , 2008] (not investigated here) requires, for quadratic cost, $(2n-1)N_r + N_C$ flops and a storage of $(2n-1)N_r + N_C$ reals in the worst case. The control law can also be approximated for further complexity reduction as proposed in [Johansen & Grancharova, 2003; Summers *et al.* , 2011], where stability is guaranteed for a certain level of sub-optimality.

Remark 16. Theorem 2.6.10 can also be used to compute a lower bound on the penalty weight for an exact penalty function [Kerrigan & Maciejowski, 2000b]. This can be used for instance to replace some state constraints with penalties (or *soft constraints*) [Kerrigan & Maciejowski, 2000b], and will be further discussed in Chapters 6 and 7.

2.6.9 Robust MPC and soft-constraints

In real world MPC applications the presence of constraints is twofold. On one hand, control performance can be notably enhanced, as mentioned above. On the other hand, handling constraints can become problematic when recursive feasibility is lost due to excessive disturbances or plant/model mismatch. Uncertainty, feedback control's *raison d'être* [Vinnicombe, 2000], has been object of study within the MPC community, through several approaches. For the purpose of this thesis, the focus is on the more classic deterministic *robust MPC* strategies. A possible choice is the restriction of constraints based on a set-valued uncertainty description. This is done in, among others, [Blanchini & Miani, 2008; Chisci *et al.* , 2001; Goulart *et al.* , 2006; Kuwata *et al.* , 2007; Richards & How, 2006; Shekhar & Maciejowski, 2012] to allow for the existence (or application) of an additional controller, such that the closed-loop trajectory is bounded inside the original constraints. This offers neat but conservative results at the expense of the nominal feasible region.

An indirect way to deal with uncertainty is to use soft-constrained MPC, where slack variables and penalty functions are used instead of the standard "hard" constraints [Kerrigan & Maciejowski, 2000a; Maciejowski, 2002; Zeilinger *et al.* , 2010]. This is generally done to the state constraints. A special case of this is the relaxation of the terminal constraint [Grüne, 2011; Grüne & Pannek, 2011; Limon *et al.* , 2006], that can allow for a larger feasible region. On the other hand, softening the terminal constraint does not always guarantee closed-loop stability from every point of feasible region. The arguments in [Limon *et al.* , 2006] apply to ℓ_{asso}-MPC, but are not addressed here. A further indirect way to deal with uncertainty is to have intermediate targets, as in the recently introduced MPC for tracking [Ferramosca *et al.* , 2008, 2009b; Limon *et al.* , 2008; Zeilinger *et al.* , 2010]. Chapter 7 formulates a synthesis of these ideas, with a novel robust uncertainty bound for soft-constrained MPC for tracking.

2.6.10 MPC for tracking

In the above sections, the use of a terminal constraint has been assumed. This constraint generally depends on the particular set-point we want to regulate to. For changing set-points this means changing the terminal constraint, which is computationally demanding and could also make the problem infeasible. The MPC for tracking, introduced by [Ferramosca *et al.* , 2008, 2009b; Limon *et al.* , 2008] for LTI systems and extended to nonlinear systems in [Ferramosca *et al.* , 2009a], overcomes the limitations of MPC for regularion by means of a an invariant set that is common to a range of feasible steady states. This tracking approach can also benefit from a larger feasible region than standard MPC for regulation. In this thesis MPC for tracking is investigated for LTI systems in Chapter 7. Considering system (2.5.7), the steady states can by characterised by the following:

Lemma 2.6.11. (based on [Ferramosca *et al.* , 2008]) If (A, B) is stabilisable, $\exists M_\theta \in \mathbb{R}^{n+m\times n_\theta}$, $n_\theta \leq n+m$ such that the steady states of the nominal system satisfy $(x_s, u_s) = M_\theta \theta_s$, with $M_\theta^T = [M_x^T \; M_u^T]$, and with $\theta_s = M_x^T x_s + M_u^T u_s$. In particular,

$$M_\theta = \mathbf{null}([A - I, \; B]). \tag{2.6.26}$$

In [Ferramosca *et al.* , 2009b] the case of tracking a piece-wise constant signal, $y_t(k) = Cx_t(k) + Du_t(k)$ is considered. In particular, a virtual set-point is introduced so that the standard MPC cost penalises deviations from this artificial variable. Then a (possibly non-smooth) function is added to the MPC cost, penalising the virtual and the actual set-point mismatch. In [Ferramosca *et al.* , 2009b], the MPC problem (2.6.1) is modified as follows:

1. Add an extra decision variable, θ, and an external reference signal $y_t(k)$,
2. Penalise $\delta\hat{x}_j = \hat{x}_j - M_x\theta$ and $\delta\hat{u}_j = \hat{u}_j - M_u\theta$ instead of $\hat{x}_j,\ \hat{u}_j$,
3. Define a constrained system, Σ_θ, with states $(x(k), \theta(k))$ (where $\theta(k)$ is simply a constant), in closed-loop with a static terminal controller, $u(k) = K(x(k) - M_x\theta(k)) + M_u\theta(k)$,
4. Impose a different terminal constraint set, $(\hat{x}_N, \theta) \in \mathbb{X}_f^t$, with $\mathbb{X}_f^t$ positively invariant for the closed-loop system Σ_θ,
5. Add to the cost function a sub-differentiable penalty, $V_O([C,\ D]M_\theta\theta - y_t(k)) \succ 0$,

2.6.10.1 Stability and convergence

Sufficient conditions for recursive feasibility and asymptotic convergence to a feasible steady state are given in [Ferramosca *et al.* , 2009b], where it it is also shown that convergence occurs to a minimiser of the offset cost V_O if an asymptotically constant $y_t(k)$ is assumed. Results are only sketched here, as they will be further revised in Chapter 7. In particular, in [Ferramosca *et al.* , 2009b] stability of a steady state is proven by the same argument as for Theorem 2.6.4, posed in the coordinates $\delta x, \delta u$. This is obtained through (A4), assumed to hold for Σ_θ in the invariant set $\mathbb{X}_f^t$, with K solving (2.6.7). Then convergence to a minimiser of V_O is proven by contradiction, by showing the sub-differential of the value function (V_N^o) to be non-negative for any feasible steady state that is not a minimiser of V_O.

2.6.10.2 Robustness

The recent development of [Zeilinger *et al.* , 2010] combined the approach of [Ferramosca *et al.* , 2008] with the concept of soft-constraints [Kerrigan & Maciejowski, 2000b; Maciejowski, 2002]. In particular, the approach of [Zeilinger *et al.* , 2010] results in an even larger feasible region than the one of [Ferramosca *et al.* , 2008], and provides a level of intrinsic robustness to bounded uncertainties in the form of Input-to-State Stablility (ISS) [Grimm *et al.* , 2004; Jiang & Wang, 2001; Lazar, 2006; Limon *et al.* , 2009].

2.6.10.3 Proposed contribution

In Chapter 7 a novel set of conditions is given for robust feasibility under bounded additive uncertainties. These conditions require minor modification of the terminal constraint of [Ferramosca *et al.* , 2008] and, if satisfied, they provide an uncertainty bound for which the MPC is always feasible and the feasible region is a RPI set. This region is allowed to be greater than the state constraints and, in most cases, is greater than the feasible region of standard robust MPC. Finally, a local ISS gain, which depends on the system and the cost parameters, is computed[1] .

[1]This includes a corrigendum of the ISS gain provided in Theorem III.7 of [Gallieri & Maciejowski, 2013a].

CHAPTER

THREE

PRINCIPLES OF LASSO MPC

3.1 LASSO regression and ℓ_1-regularised least squares

Least Absolute Selection and Shrinking Operator (LASSO) is a powerful regression technique, which aims to induce sparseness in the solution [Kim *et al.* , 2007a; Ohlsson, 2010; Osborne & Presnell, 2000; Schmidt, 2010; Schuet, 2010; Tibshirani, 1996]. This property is obtained by minimising an ℓ_1-regularised least squares cost function, namely a cost function consisting of a smooth quadratic plus a non-smooth 1-norm term. Common applications of the technique are in the fields of machine learning, feature selection, compressed sensing, and filtering [Kim *et al.* , 2007b; Ohlsson *et al.* , 2010a,b; Osborne & Presnell, 2000; Tibshirani, 1996]. A common regressor used in these fields is also the so-called *Ridge regressor* or ℓ_2-regularised least-squares, that minimises a particular quadratic cost and, from a control engineering perspective, can be related to a deterministic Kalman filter [Willems, 2004]. The latter is not intended to and does not produce sparse estimates (see Chapter 3 of [Hastie *et al.* , 2011]). On the other hand, the solution of a Kalman filter is smooth and unique when the cost is positive definite. In this thesis, a combination of the two is considered, called the *elastic-net* [Hastie *et al.* , 2011; Zou & Hastie, 2005]. A powerful hybrid between the two regressors, the elastic-net allows one to regulate the tradeoff between the sparsity of LASSO and the smoothness of Ridge regression. Elastic-net can also sparsify groups of variables at the same time, depending on the data. This regressor can be represented as a particular case of LASSO [Zou & Hastie, 2005], therefore, throughout the thesis only the term LASSO is used as well as the notation "ℓ_{asso}-MPC". The proposed controller is however based on the elastic-net regression[1].

Noticeably, some other regressors are used in the above fields of statistical learning, the study of which is not in the scope of this thesis. The *sum-of-norms* regularisation (of which LASSO is a special case) uses an ℓ_p-norm penalty to provide sparsity. If the ℓ_2-norm is instead used, then it is possible to set given groups of variables to zero simulatenously [Hastie *et al.* , 2011; Ohlsson *et al.* , 2010a,b, 2012]. Group sparsity is not specifically addressed in this thesis, but it can occur in some particular cases.

3.1.1 Considered regularised regressors

Consider a vector of decision variables, $\underline{u}$, and the observations, $\underline{y}$. The problem

$$\underline{u}^\star = \arg\min_{\underline{u}} \|\underline{y} - \Lambda\underline{u}\|_2^2 + \|W\underline{u}\|_1 \tag{3.1.1}$$

is referred to as an ℓ_1-*regularised least squares* problem [Kim *et al.* , 2007b; Ohlsson *et al.* , 2010a,b; Osborne & Presnell, 2000; Tibshirani, 1996] or LASSO regression in its *unconstrained form*. The convex problem (3.1.1) is non-differentiable at the origin. Non-differentiability causes LASSO to result in, differently from a standard Least Squares (LS) problem, a piecewise affine solution (as a function of $\underline{y}$) with $\|W\underline{u}^\star\|_1 \leq t$, for some $t \geq 0$ [Kim *et al.* , 2007b; Osborne & Presnell, 2000]. Thus, we can also write (3.1.1) in the equivalent (and initial) formulation of [Tibshirani, 1996]:

$$\begin{aligned} \underline{u}^\star &= \arg\min\nolimits_{\underline{u}} \|\underline{y} - \Lambda\underline{u}\|_2^2 \\ &\text{s.t. } \|W\underline{u}\|_1 \leq t, \ \text{for } t \geq 0. \end{aligned} \tag{3.1.2}$$

The ℓ_1-norm constraint in (3.1.2) forces most of the solution's elements to be null, if t is chosen appropriately. This is referred to as *parsimony* or *sparseness*. Define first the ℓ_p-norm on a vector u as $(\sum |u_i|^p)^{1/p}$, for $p > 0$. The ideal way to induce sparseness is to penalise the number of nonzero elements, namely, using the so-called 0-norm [Ljung *et al.* , 2011]:

$$\begin{aligned} \|u\|_0 &= \mathrm{card}(u) = \textstyle\sum_{i=1}^n \varrho(u_i), \\ \varrho(u_i) &= \begin{cases} 1 & (u_i \neq 0) \\ 0 & (u_i = 0). \end{cases} \end{aligned} \tag{3.1.3}$$

[1] This results in the ℓ_{asso}-MPC problem to have a unique solution.

Minimising a least-squares cost with a 0-norm penalty results in a Mixed Integer Program (MIP), an NP-hard problem. On the other hand, as in [Hastie *et al.* , 2011], Figure 3.1 shows that $p = 1$ (featured by LASSO) is the smallest value of p for which an ℓ_p regularised least-squares problem is convex. This is the main reason for the recent success of LASSO [Hastie *et al.* , 2011; Ljung *et al.* , 2011] and it marks a substantial difference between LASSO and other regularised least-squares approaches, for instance the sum-of-squares, or *Ridge regression* [Kim *et al.* , 2007a], for which most of the decision variables will always be non-zero.

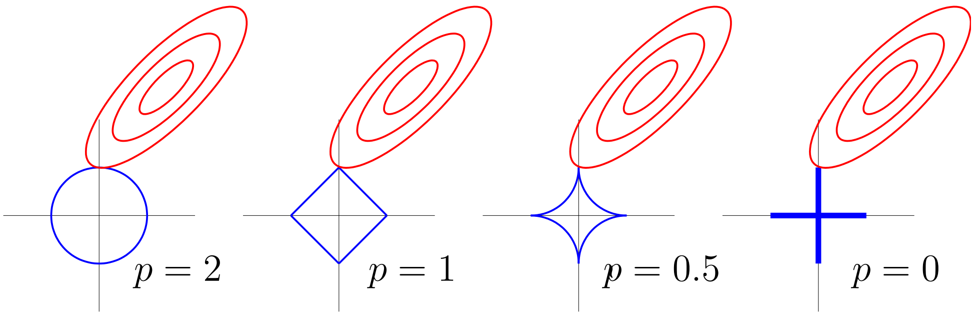

Figure 3.1: Different regularisation problems. Misfit cost (ellipsoids) and penalties (centred at the origin). Based on [Hastie *et al.* , 2011].

As mentioned above, a powerful technique that enjoys the benefits of both LASSO and Ridge is the elastic-net [Zou & Hastie, 2005], that allows for a tradeoff between a sparse (like LASSO) and a smooth unique solution (like Ridge). The elastic-net solves a problem of the form

$$\underline{\mathrm{u}}^{\star} = \arg\min_{\underline{\mathrm{u}}} \|\underline{\mathrm{y}} - \Lambda \underline{\mathrm{u}}\|_2^2 + \alpha \|W\underline{\mathrm{u}}\|_2^2 + (1-\alpha)\|W\underline{\mathrm{u}}\|_1, \ \alpha \in (0,1), \tag{3.1.4}$$

that can easily be expressed in the form of (3.1.1), being effectively a special case of LASSO. Elastic-net has the nice feature of shrinking variables together, if their respective columns of Λ have a high correlation, namely, a high inner product [Zou & Hastie, 2005]. This concept is further elaborated in Chapter 3.4.3 of [Hastie *et al.* , 2011].

3.1.2 Solution characteristics

As shown in [Efron *et al.* , 2004; Osborne & Presnell, 2000], the solution of LASSO for $W = \eta I$, as $\eta \geq 0$ varies, is a continuous piecewise affine function. The set of solutions is called the *regularisation path* [Efron *et al.* , 2004]. In particular, [Osborne & Presnell, 2000] have shown that there exists a finite $\eta_{\max}(\underline{\mathrm{y}})$ such that

$$W = \eta I,\ \eta \geq \eta_{\max}(\underline{\mathrm{y}}) \Rightarrow \underline{\mathrm{u}}^{\star} = 0. \tag{3.1.5}$$

The computation of the regularisation path has been discussed in [Efron *et al.* , 2004]. The extension of the results of [Efron *et al.* , 2004] to the elastic-net, has been presented in [Zou & Hastie, 2005], where the regressor has been also shown to have a PWA regularisation path, and a similar $\eta_{\max}$ is computed for which (3.1.5) holds. Moreover, [Efron *et al.* , 2004; Zou & Hastie, 2005] provide efficient algorithms to compute the path, with the same complexity of a QP. In this thesis, the solution path is computed by means of multi-parametric programming and explicit MPC [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002a]. These techniques are currently applicable to small scale PWA systems (not larger than 5 states/inputs) with horizon length $N \leq 5$.

An interesting property of LASSO (and of elastic-net) is verified when Λ is orthonormal and again if $W = \eta I$, $\eta > 0$ scalar. In this case the solution corresponds to the least squares one (the Ridge one for the elastic-net), passed through the continuous PWA *soft thresholding operator* [Hastie *et al.* , 2011; Osborne & Presnell, 2000; Tibshirani, 1996] with threshold $\eta/2$. Namely,

$$\underline{\mathrm{u}}_i^{\star} = u_{thr} = \mathrm{sign}(\tilde{u}_i)(|\tilde{u}_i| - \eta/2)_+, \tag{3.1.6}$$

where $\tilde{u}$ is the solution for $W = 0$. This dead-zone nonlinearity differs from the discontinuous hard thresholding, as shown in Figure 3.2 (see next page). For MPC applications it will be shown that Λ is not entirely a design parameter, and by construction it cannot be orthonormal. Nevertheless, a few interesting considerations can be made on the behaviour of the solution, as done in the next chapter.

3.1.3 Control-related applications

Recently LASSO-type techniques have been adopted also outside the community of machine learning and compressed sensing. For instance [Ohlsson *et al.* , 2012] provides an extensive review of state filtering for LTI systems in the presence of impulsive disturbances and measurement noise, where appropriate sum-of-norm estimation problems have been formulated. Furthermore, [Schuet, 2010] has shown LASSO-based estimation is effective for fault detection. Control applications of regularisation techniques include, among others, [Lin *et al.* , 2012; Nagahara & Quevedo, 2011; Ohlsson *et al.* , 2010b]. In particular, [Ohlsson *et al.* , 2010b] demonstrated through simulation

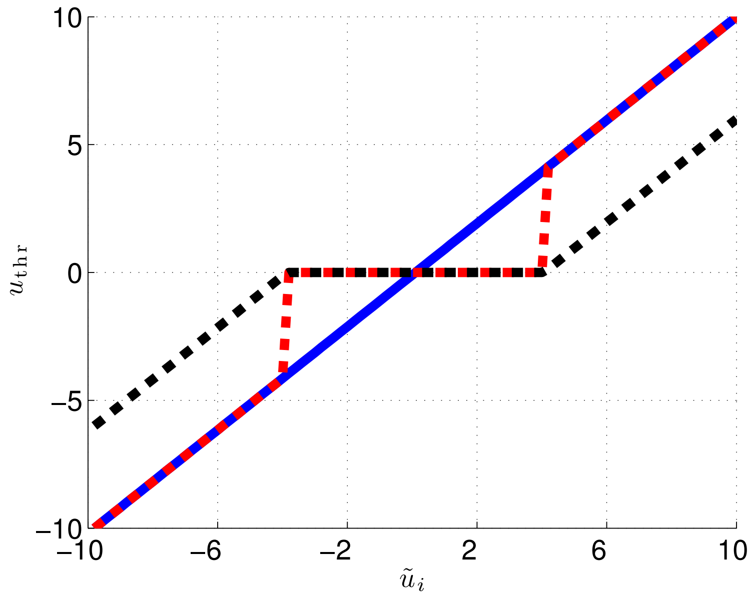

Figure 3.2: Thresholding operators. Original signal (blue), and output of the hard (red) and soft (black) thresholding, with threshold $\eta = 4$.

the use of sum-of-norms for receding horizon path planning with bounded input rates. The authors of [Nagahara & Quevedo, 2011] used LASSO for discrete-time unconstrained receding horizon remote for single input systems. This approach, called *Packetized Predictive Control*, is used to deal with communication losses while transmitting the control sequence. Noticeably, [Nagahara & Quevedo, 2011] showed for the first time that ultimate boundedness (or practical stability) is possible for the nominal and robust cases, inspiring the further developments in the next chapter of this thesis. In [Lin *et al.* , 2012] the elastic-net and other regressors are used for continuos-time LTI state feedback control over networks, with the aim of reducing the number of links while keeping a bounded $\mathcal{H}_2$-norm. In [Gallieri & Maciejowski, 2012] the use of elastic-net has been investigated in MPC for discrete-time constrained over-actuated systems, where a dual-mode approach is used to obtain asymptotic stability. This controller consists of a variable-horizon MPC and is not described in this thesis, since it requires a higher online computation than the proposed ones. In [Gallieri & Maciejowski, 2012] the problem of vessel roll reduction with rudder and fins has been used as a demonstrator for ℓ_{asso}-MPC as well as in the current thesis.

The remainder of the thesis contains only original contribution.

3.2 Formulation of ℓ_{asso}-MPC

LASSO-based MPC consists of solving online a deterministic finite horizon optimal control problem, where the commands are required to have bounded magnitude and to be sparse among the input channels.

Definition 3.2.1. ($\ell_{\mathbf{asso}}$–MPC)
Consider the following finite-horizon constrained optimal control problem (FHCOCP):

$$\begin{aligned} V_N^o(x) = \min_{\underline{u}} \left\{ V_N(x, \underline{u}) \mathrel{\hat{=}} F(\hat{x}_N) + \sum_{j=0}^{N-1} \ell(\hat{x}_j, \hat{u}_j) \right\} \\ s.t. \quad \hat{x}_{j+1} = \hat{f}(\hat{x}_j, \hat{u}_j), \\ \hat{u}_j \in \mathbb{U},\ \forall j \in \mathbb{I}_{[0,\ N-1]},\ \ \hat{x}_j \in \mathbb{X},\ \forall j \in \mathbb{I}_{[0,\ N]}, \\ \hat{x}_0 = x,\ \hat{x}_N \in \mathbb{X}_f \subseteq \mathbb{X}, \end{aligned} \tag{3.2.1}$$

with stage cost

$$\ell(x, u) = x^T Q x + u^T R u + \|Su\|_1, \tag{3.2.2}$$

where the terminal cost F is a strongly convex CLF, and where $\hat{f} : \mathbb{R}^{n \times m} \to \mathbb{R}^n$ is a continuous prediction model, $\underline{u}^T = \left[\begin{array}{ccc} \hat{u}_0^T, & \cdots, & \hat{u}_{N-1}^T \end{array} \right]$, $S \in \mathbb{R}^{m_s \times m}$, $m_s \in \mathbb{I}_{>0}$. At each iteration k, the ℓ_{asso}-MPC applies to the plant the first move of the optimal policy, $u(k) = \hat{u}_0^\star$, obtained by online solution of (3.2.1-3.2.2) at the current state $x(k)$. The generated implicit control law is referred to as $K_N(x) \equiv \hat{u}_0^\star(x)$. The closed-loop evolution of (2.5.1) under ℓ_{asso}-MPC is

$$x(k+1) = f(x(k), K_N(x(k)), w(k)). \tag{3.2.3}$$

The one-step evolution of (3.2.3) for a given x is simply denoted as $x^+ = f(x, K_N(x), w)$.

The controller resulting from receding horizon solution of (3.2.1) is defined as ℓ_{asso}-MPC with (input) magnitude regularisation. Replacing $\hat{u}_j$ by its one step time difference, $\Delta\hat{u}_j = \hat{u}_j - \hat{u}_{j-1}$, problem (3.2.1) will define a ℓ_{asso}-MPC with (input) rate regularisation. The latter can be formulated with simple modifications of the problem (3.2.1) or the system model (for instance including u in the state space), similarly to [Goodwin *et al.* , 2005; Maciejowski, 2002].

Remark 17. The prediction models considered in this thesis are intended to describe the system evolution under the assumption that $w(k) = 0, \forall k$ (nominal predictions), namely, $\hat{f}(x, u) =$

$f(x, u, 0)$.

Remark 18. All results in this thesis are proposed for cost (3.2.2), however, they are also valid for positive definite costs with quadratic cross state-input terms. In this case one can define the extended "state" space[1] $\chi(k) = (x(k), u(k-1))$ with

$$\chi(k+1) = \bar{A}\chi + \bar{B}u(k),\ \bar{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix},\ \bar{B} = \begin{bmatrix} B \\ I \end{bmatrix},\ \bar{Q} = \begin{bmatrix} Q & T^T \\ T & R \end{bmatrix} \succ 0,$$

and then consider $\ell(\chi) = \chi^T \bar{Q}\chi + \|[0,\ S]\chi\|_1$ as the MPC cost function.

3.2.1 Version 1: "Regularised LQR"

Consider an LTI system of the form (2.5.2). In contrast to quadratic MPC, non-differentiability of the 1-norm penalty makes difficult the computation of a closed form solution for the unconstrained ℓ_{asso}-MPC problem, or of the cost-to-go. Therefore, infinite horizon optimality is not guaranteed with the proposed approaches. The first formulation features the quadratic terminal cost

$$F(x) = x^T P x. \tag{3.2.4}$$

If the prediction model $\hat{f}$ is LTI, then this choice of F will result in a *strongly convex* QP. This type of controller is investigated in Chapter 4.

3.2.2 Version 2: General form

This formulation features the terminal cost

$$F(x) = \beta\psi_{\mathbb{X}_f}(x)^2 + \alpha\psi_{\mathbb{X}_f}(x),\ \beta, \alpha \geq 0, \tag{3.2.5}$$

where $\psi_{\mathbb{X}_f}$ is the Minkowski function of the terminal set $\mathbb{X}_f$. This formulation is used to provide an asymptotic stability guarantee. For this purpose, $\mathbb{X}_f$ is taken to be λ-contractive. If the prediction model $\hat{f}$ is LTI, then this choice of F will also result in a *strongly convex* QP. This novel MPC controller is investigated in Chapter 5.

[1]The notion of system state is slightly abused here, since u is memory-less and directly manipulated.

3.3 Initial comparison with quadratic, ℓ_1 and ℓ_∞ MPC

An initial comparison of ℓ_{asso}-MPC with the most common predictive control strategies is performed, to further motivate the theoretical developments presented in the thesis.

3.3.1 Example 1: Sparsity in the actuator space

Consider the LTI system described by

$$A = \begin{bmatrix} 1 & 0.1 \\ -0.1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1.5 & -0.5 \\ -0.4 & 2 \end{bmatrix}, \quad \|u(k)\|_\infty \leq 2, \quad T_s = 0.1. \tag{3.3.1}$$

The system is simulated in closed-loop with 4 different MPC strategies. The controllers solve a problem of the form (3.2.1) with the following parameters:

- **LQ-MPC**: $\ell(x,u) = x^TQx + u^TRu$, $F(x) = x^TPx$, $\mathbb{X}_f = \{0\},\ Q = 0.1 \cdot I,\ R = 5 \cdot I,$
- **LASSO-MPC**: $\ell(x,u) = x^TQx + u^TRu + \|Su\|_1$, $\mathbb{X}_f = \{0\},\ Q = 0.1 \cdot I, R = 0.1 \cdot I,\ S = 5 \cdot I,$
- **ℓ_1-MPC**: $\ell(x,u) = \|Qx\|_1 + \|Ru\|_1$, $\mathbb{X}_f = \{0\},\ Q = 0.1 \cdot I,\ R = 5 \cdot I,$
- **ℓ_∞-MPC**: $\ell(x,u) = \|Qx\|_\infty + \|Ru\|_\infty$, $\mathbb{X}_f = \{0\},\ Q = 0.1 \cdot I,\ R = 5 \cdot I.$

From Figure 3.3, the state evolution is comparable under ℓ_{asso}-MPC and LQ-MPC, even though the proposed tuning provides the former with faster convergence. On the other hand, the input behaviour differs, as in the latter the signals are not zero until the system state is also at zero. For ℓ_{asso}-MPC, the second actuator is instead set to zero before the system converges. The implemented control law and the closed-loop system are therefore non-linear, the control action being *sparse in the actuator domain* for small control errors. For different values of S it is possible to severely reduce the magnitude of the actuator moves, or to make I/O behaviour more similar to LQ-MPC.

The ℓ_1 and ℓ_∞ MPC, in Figure 3.3, generate very different input signals, and seem to offer poorer performance (the ℓ_1 solution is also non-unique). In particular, they have fast convergence, and sparse inputs, however, input signals are very erratic. In particular, using two different implementations (Figure 3.3: left using a condensed LP, right using Yalmip) and the same LP solver, using the ℓ_1-MPC has provided two different solutions. This is due to the fact that the considered LP is degenerate at the initial condition, $x(0) = [10,\ 10]^T$. Multiple solutions are quite undesirable

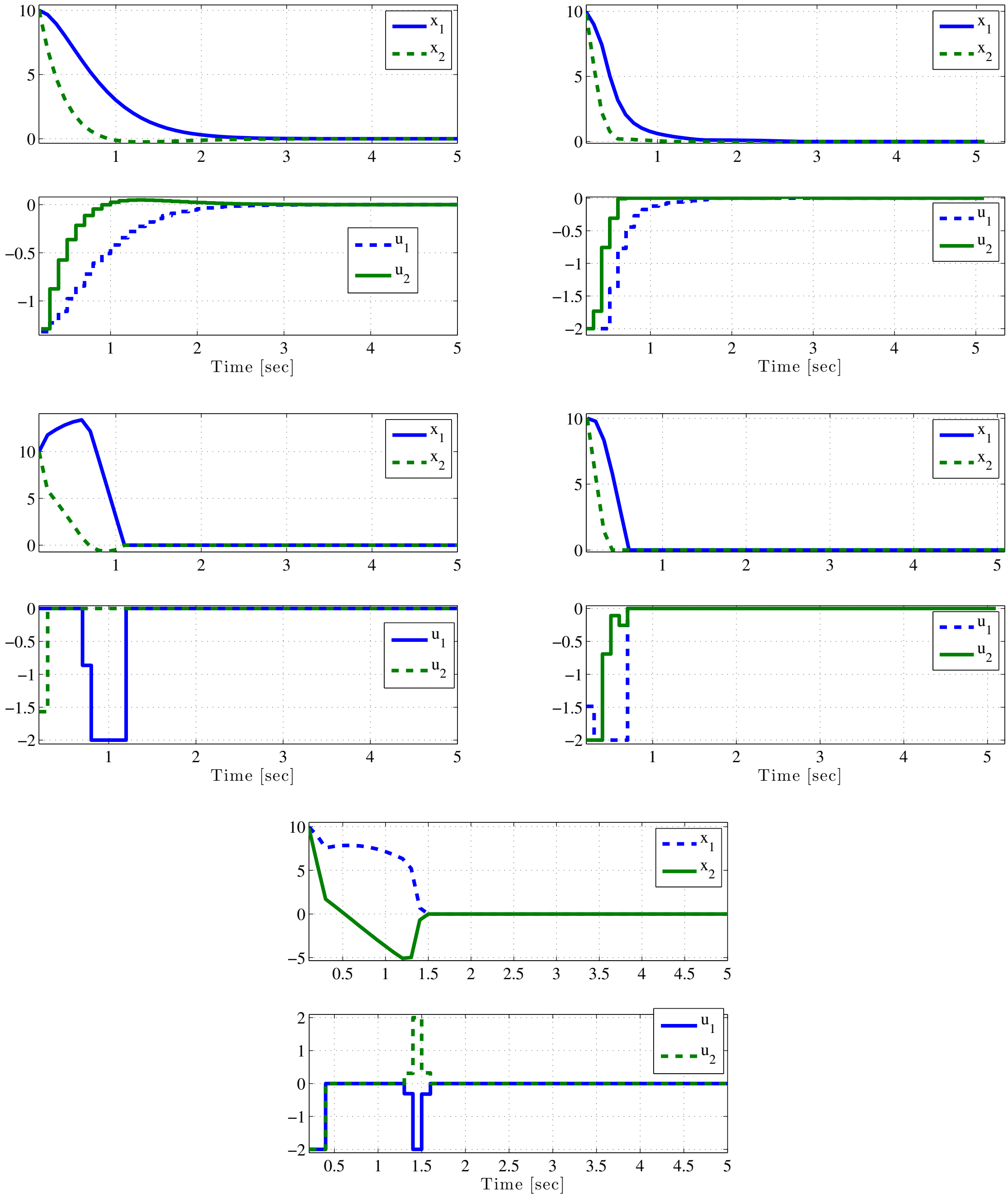

Figure 3.3: LQ-MPC (top left) ℓ_{asso}-MPC (top right), ℓ_1-MPC (center, two different solutions) and ℓ_∞-MPC (bottom) with completion constraint

for an MPC controller, as they may compromise the system's reliability. The ℓ_∞-MPC leaves the system open loop for a short time, and then decides to use differential actuation.

It must be noticed that the cost functions of the above case studies have been tuned with the purpose of obtaining the discussed behaviours. Obviously, different tunings provide different I/O signals. The intention here is not to label one MPC strategy as "good" or "bad" in the absolute sense. Certainly, the presence of undesirable behaviour in ℓ_1 and ℓ_∞-MPC is something to keep in mind, as such behaviour is not likely to be expressed by LQ and ℓ_{asso}-MPC, in the considered setup. The ℓ_1, ℓ_∞ and ℓ_{asso}-MPC can all produce sparse input signals, however, ℓ_{asso}-MPC seems to be the only one to inherit the smooth I/O proprieties of LQ-MPC, as can be expected from the similarity of the cost functions. More considerations about the cost function of the considered MPCs will be given at the end of this chapter. On the plus side, the ℓ_∞-MPC generated controls are also *sparse in time*, which could be of interest in particular applications. This concept will be further considered in the next example.

3.3.2 Example 2: Uncertainty and sparsity in time

All control systems are subject to *uncertainty*, the reduction of which is the only motivation for feedback control [Vinnicombe, 2000].

Consider the following system:

$$x(k+1) = Ax(k) + B(u(k) + w(k)),\ y(k) = x(k) + d(k), \tag{3.3.2}$$

where $w(k) = (w_1(k),\ w_2(k))$, are additive input disturbances, while $d(k)$ is measurement noise. The system matrices and input constraints are

$$A = 1, \quad B = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad \|u(k)\|_\infty \leq 2, \tag{3.3.3}$$

and the sampling time is $T_s = 0.1$ $[sec]$. We wish to determine the control input, $u(k)$, by means of an MPC with state feedback and horizon length $N = 2$. No state constraints are used. Four MPC controllers are considered, with cost functions as in the previous example. The terminal constraint is not included here, and a terminal cost $\|Px\|_p$ is instead used with $P = 0.55$.

3.3.2.1 Nominal case

Figure 3.4 shows the nominal behaviour under the considered controllers. It can be seen that ℓ_{asso}-

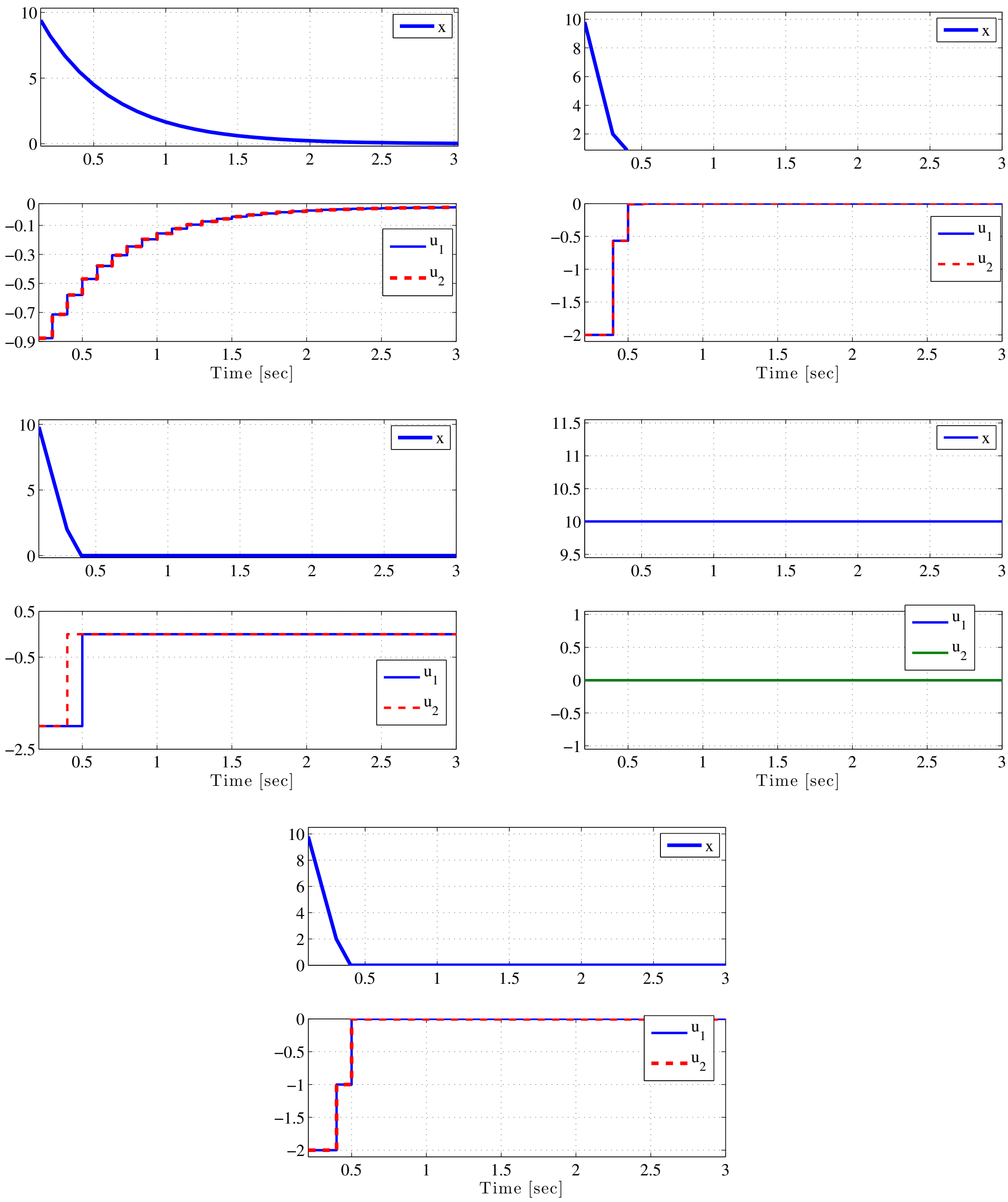

Figure 3.4: Trajectories of LQ-MPC (top letf) ℓ_{asso}-MPC (top right) ℓ_1-MPC (center, two different solutions) and ℓ_∞-MPC (bottom) without terminal constraint

MPC is behaving non-linearly around the origin, where no constraints are active. In particular, the loop is opened when the error is small, and the system converges to $x(\infty) \approx 1.7$, not to the origin. The ℓ_1-MPC presents again 2 different solutions from the initial state $x(0) = 10$. One of them is asymptotically stabilising (implemented using a condensed QP formulation with slack variables), while the other one (implemented using YALMIP for Matlab [Lofberg, 2004]) leaves the system at $x(0)$. As the simulations suggest, nominal closed-loop stability of ℓ_{asso}-MPC version 1 is possible in terms of ultimate boundedness but not always in terms of asymptotic stability. This is discussed in the next chapter.

3.3.2.2 Gaussian input disturbance

Figure 3.5 shows the closed-loop behaviour when the system is under the input disturbance $w(k) = (w_1(k),\ w_2(k))$, the elements of which are different realisations of a zero-mean Gaussian noise with variance 2. It can be noticed that LQ-MPC behaves similarly to an LQR controller, with small and non-zero control signals. On the other hand, ℓ_{asso}-MPC has produced control signals that are *sparse in time*, and overall "less expensive" (in the ℓ_1 sense). Moreover, ℓ_{asso}-MPC seems to have visibly reduced the effect of the noise, with respect to the LQ-MPC, with basically the same actuators usage. At the same time the ℓ_∞-MPC moves the actuators much more, and achieves just the same noise reduction. Noticeably, the two actuators behave in the same way (having the same model) for all MPC strategies, except for the ℓ_1-MPC. The latter in fact uses u_1 most of the time (u_1 is chattering) without any design specification, and produces an extremely sparse u_2 signal. This chattering behaviour is never possible for LQ and ℓ_{asso}-MPC, provided that $R \succ 0$. This will also be proven in the next chapter.

3.3.2.3 Gaussian measurement noise

Figure 3.6 shows the closed-loop behaviour under a measurement noise, $d(k)$, that is a realisation of a zero-mean Gaussian noise with variance 0.5. Once again, the ℓ_{asso}-MPC produces a sparse in time input sequence, unlike LQ and ℓ_∞-MPC. Moreover, this time the noise reduction is much higher than for all of the other strategies. In particular, the ℓ_1 and ℓ_∞-MPC are in this case not able to match the noise reduction of the LQ and ℓ_{asso}-MPC, their control signals being excessively lively.

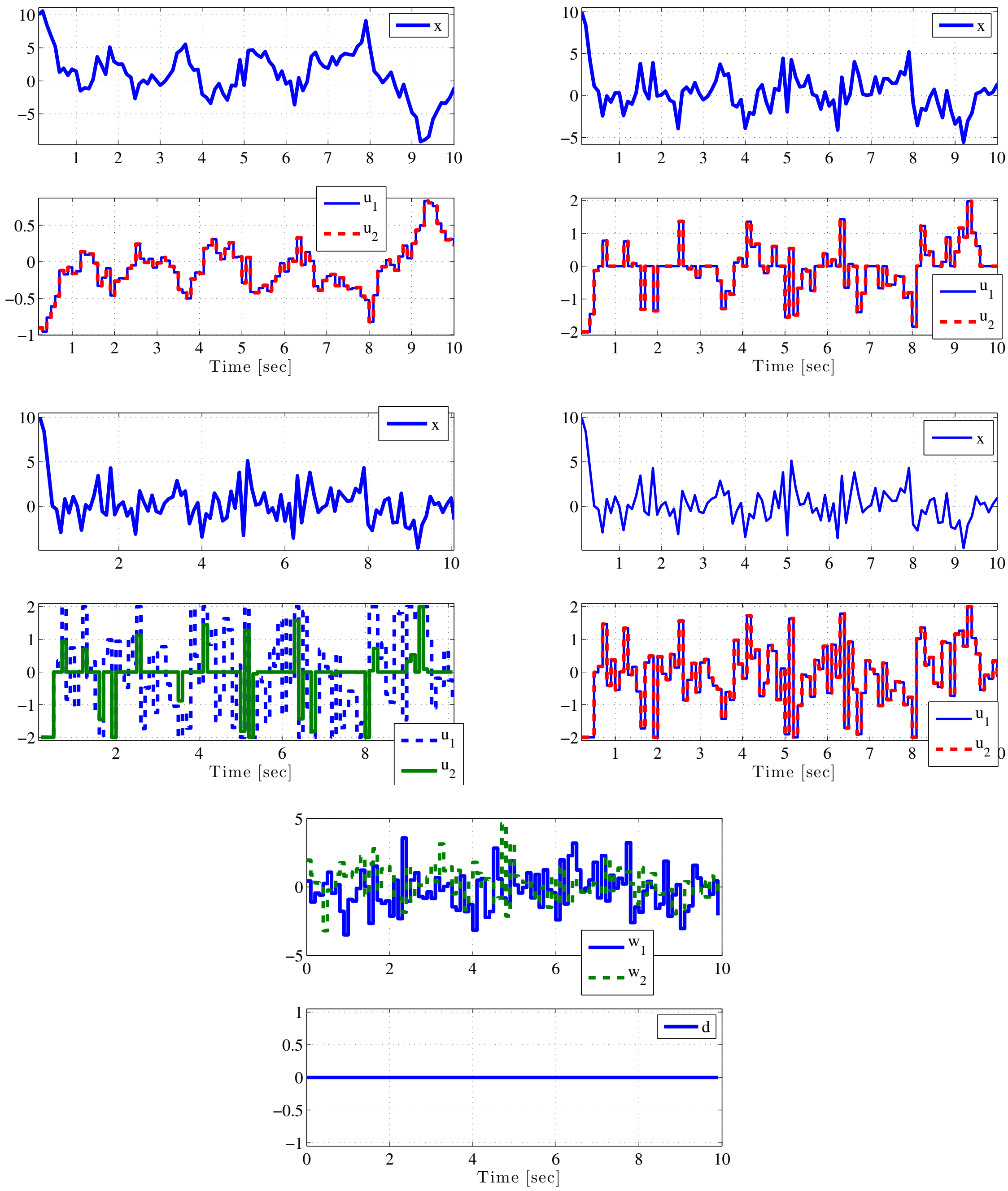

Figure 3.5: Trajectories of LQ-MPC (top left) ℓ_{asso}-MPC (top right) ℓ_1-MPC (middle letf) and ℓ_∞-MPC (middle right) under two zero-mean Gaussian input disturbances (bottom) with variance 2.

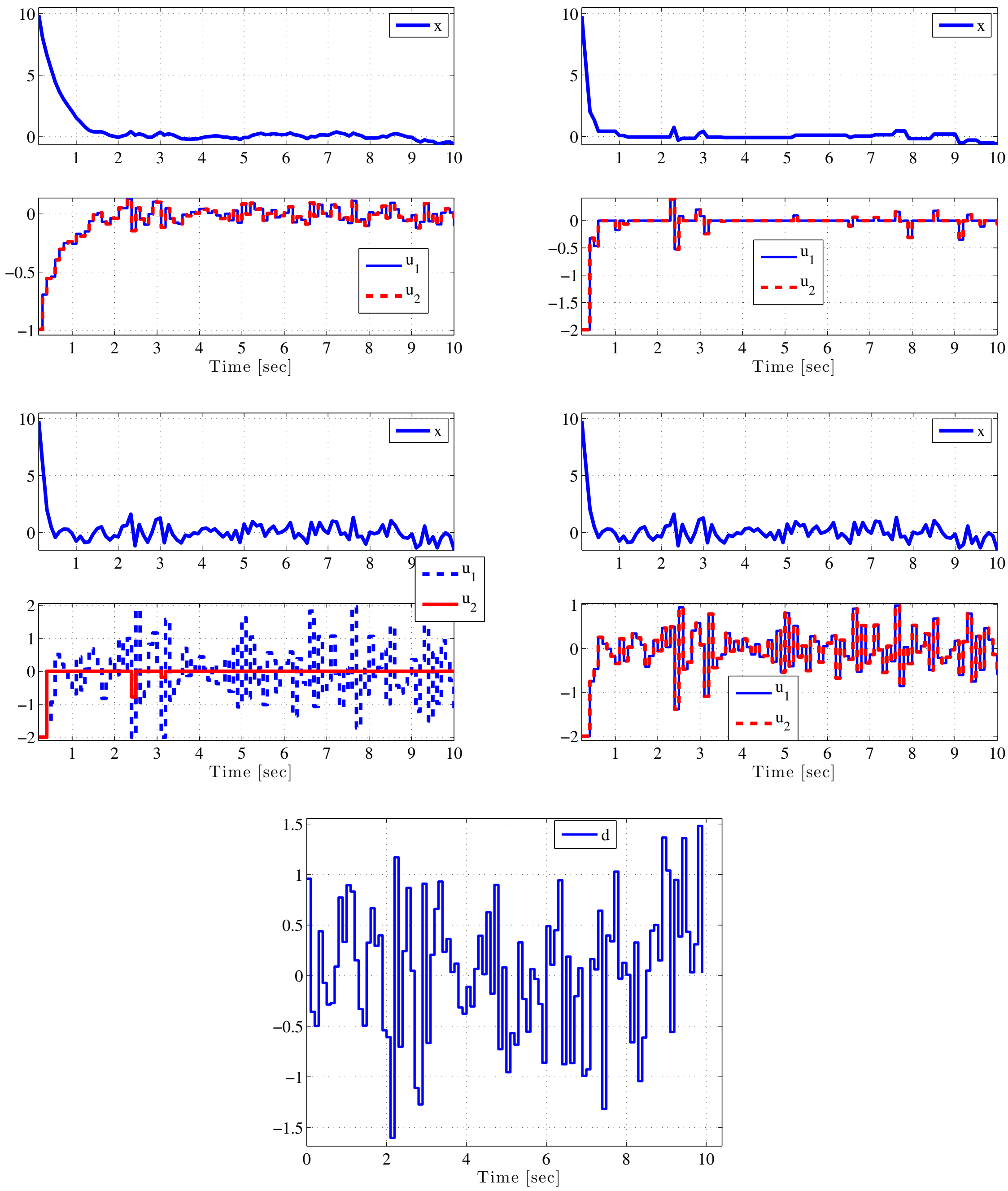

Figure 3.6: Trajectories of LQ-MPC (top left) ℓ_{asso}-MPC (top right) ℓ_1-MPC (middle letf) and ℓ_∞-MPC (middle right) under a zero-mean Gaussian measurement noise disturbance (bottom) with variance 0.5.

3.3.2.4 Comments

Similarly to LQ-MPC, in the above simulations ℓ_{asso}-MPC has been able, with the same Q, R and P, to handle relatively well both input and measurement uncertainties. At the same time, ℓ_{asso}-MPC has produced sparse input sequences, that could be suitable for applications where the control action is expensive. On the other hand, MPC based on ℓ_1 and ℓ_∞ costs have been able to perform well in only one of the two cases (for this tuning the input or state noise, even tough other tunings are possible). These controllers have produced sparse yet unsatisfactory input behaviour. In particular, the ℓ_1-MPC suffers from strong chattering. An interpretation of these different actuators behaviour is given in the next Section, where the role of the cost function and the constraints is discussed.

3.3.3 Geometric interpretation of the solution behaviour

In order to understand further the implications of using a LASSO cost, rather than for instance a 1-norm cost, consider an LTI system with $A = I_{2\times 2}$, $B = I_{2\times 2}$, and an MPC controller with $N = 1$, $F(x) = \ell(x, 0)$, $\mathbb{X} = \mathbb{R}^2$, $\mathbb{U} = \mathbb{R}^2$, and with terminal constraint $\mathbb{X}_f = \{x : L_f x \leq \underline{1}\}$ (in Figure 3.7) where

$$L_f = \begin{bmatrix} 0.25 & 1 \\ -0.25 & 1 \\ -1 & -0.25 \\ -1 & 0.25 \\ -0.25 & -1 \\ 0.2500 & -1 \\ 1 & -0.25 \\ 1 & 0.25 \\ -0.7289 & -0.7289 \\ -0.7289 & 0.7289 \\ 0.7289 & 0.7289 \\ 0.7289 & -0.7289 \end{bmatrix}. \tag{3.3.4}$$

Denote $u = (u_1, u_2)$, $x = (x_1, x_2)$ and consider the following cost functions (the system state is not penalised for illustrative purpose, and $F(x) = 0$)

- `Case 1`: $\ell(x, u) = \|u\|_1$

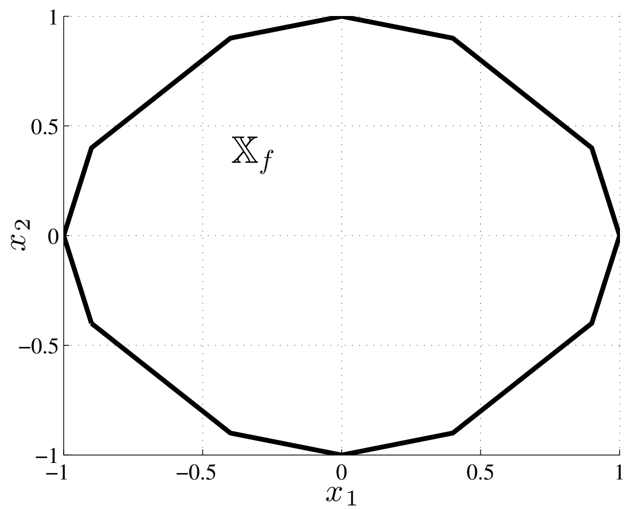

Figure 3.7: Considered terminal constraint

- `Case 2`: $\ell(x,u) = \|u\|_1 + u^T u$
- `Case 3`: $\ell(x,u) = \|u\|_1 + u^T u + 10(u_2)^2$
- `Case 4`: $\ell(x,u) = \|u\|_1 + u^T u + 3|u_2|$

Figure 3.8 shows some level sets of the 1-norm cost function for `Case 1`, as well as the constraint $\mathbb{X}_f$, expressed as a function of u for the initial state $x(0) = (2,\ 0.8)$. It can be noticed that, since for this particular value of x the level set boundary of the cost is parallel to one of the constraint facets, the problem has multiple optimal solutions (the facet boundary of the constraint is highlighted in purple). This is the reason for the chattering seen in the previous section. It can be seen, in particular, that one of the optimal solutions features $u_2 = 0$ and could potentially lead to a sparse control signal. From Figure 3.8 it is also possible to speculate on how uncertainty could affect the solution, as the optimal u could dramatically change if either the constraint or cost function are slightly perturbed (recall from Chapter 2.6.8 that the MPC cost generally depends on the state). The (sparsity-optimal) ℓ_0 solution, in Figure 3.8, is clearly the one with only $u_2 \neq 0$. On the other hand, Figure 3.9 shows that `Case 2` has a unique optimum. This is however not sparse, for this particular tuning and initial state. The cost function level sets for `Case 2` are described by curves that are overall smoother than for `Case 1`, thanks to the quadratic terms in the cost. This in fact is what makes the solution unique for ℓ_{asso}-MPC, as it will be proven in the next chapter. Smoothness also means that, similarly to LQ but in contrast with ℓ_1-MPC, small perturbations of the cost and constraints have a smooth effect on the solution of ℓ_{asso}-MPC. Moreover, ℓ_{asso}-MPC can achieve the same noise rejection of LQ-MPC that is superior to the ℓ_1 and ℓ_∞ case. Noticeably, the curves in Figure 3.9 still feature the non-smooth vertices of the ℓ_1-norm. These are key for ℓ_{asso}-MPC to have a sparse solution.

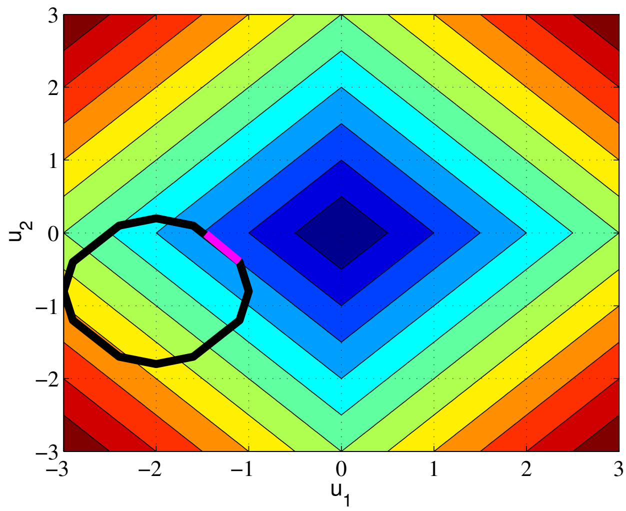

Figure 3.8: `Case 1`: Constrained 1-norm MPC with multiple optima (constraints facet in purple)

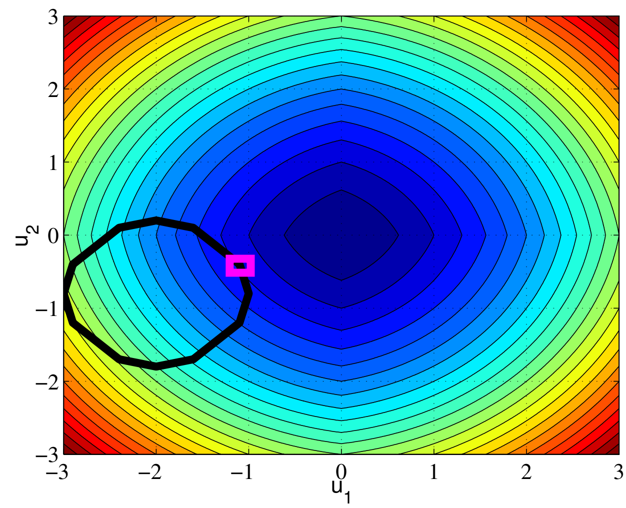

Figure 3.9: `Case 2`: Constrained LASSO-type MPC. Solution is unique but not sparse

In order to search for a unique and sparse solution, for instance preferring u_1 over u_2, the following modifications are made to the cost function of `Case 2`. In `Case 3` the weight on the square of u_2 is increased by 10 while, in `Case 4` the corresponding 1-norm weight is increased by 3. Figure 3.10 shows the result for `Case 3`. In particular, the level set boundaries are smoother than in the previous cases and the solution is not sparse. On the other hand, curves for `Case 4`, in Figure 3.11, have more pronounced peaks, on the u_2 direction, thus providing the desired sparse solution with $u_2 = 0$. In conclusion, it can be noticed that a tradeoff exists between the smoothness of the cost and the sparsity of the solution. One possible way to regulate this tradeoff is to use an elastic-net cost as done in this thesis.

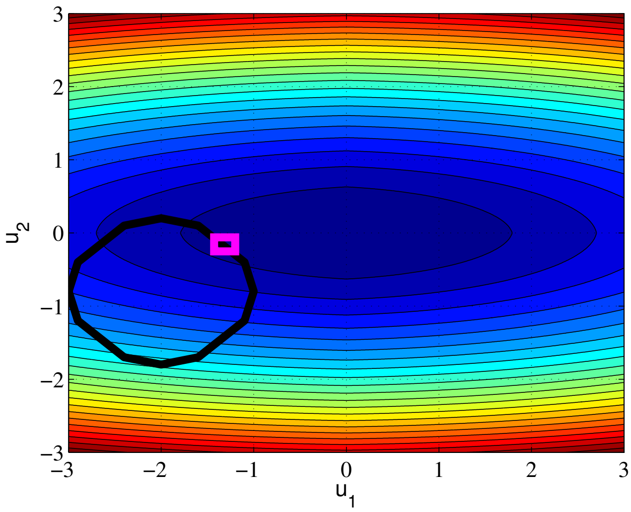

Figure 3.10: `Case 3`: Constrained LASSO-type MPC with increased quadratic penalty on u_2. Solution is unique but not sparse

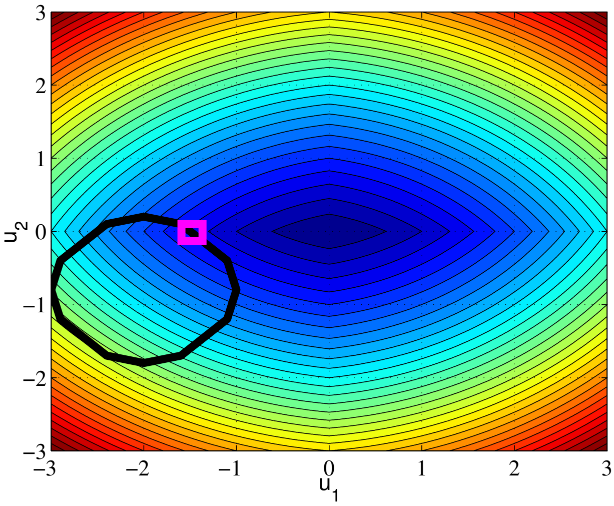

Figure 3.11: `Case 4`: Constrained LASSO-type MPC with increased ℓ_1 penalty on u_2. Solution is unique and sparse ($u_2^\star = 0$)

3.4 Conclusions

This chapter has introduced the two ℓ_{asso}-MPC formulations investigated throughout the thesis. First, a review of ℓ_1-regularisation and LASSO regression has been given, focusing on the solution's behaviour. In particular, the solution of LASSO regression and its variants can be sparse, namely many elements of it can be zero, at the expense of a bounded approximation error (or bias). This translates into the first of the proposed ℓ_{asso}-MPC versions, which is based on the *elastic-net* regression. Initial simulations have shown that ℓ_{asso}-MPC version 1 is able to produce input sequences that are sparse both through time and through the actuator channels, however, the former can occur only in particular cases. Sparsity in time can occur for instance if the system is subject to noise but can hardly be guaranteed a priori, since open-loop and closed-loop MPC predictions are generally not the same. This is not the case for sparsity through actuators, since only the first control move is involved. This motivates the further investigation of actuators sparsity carried out in the thesis. Temporal sparsity is on the other hand not further investigated.

A simple LTI example has been used to illustrate the geometric conditions for sparsity through actuators as well for having a unique solution. As it will be proven in the next chapter, these characteristics are possessed at the same time by ℓ_{asso}-MPC, in contrast with the more common ℓ_1, ℓ_∞ and LQ-MPC. In particular, ℓ_1 and ℓ_∞ MPCs can produce sparse solutions but may suffer from poor control performance and actuator chattering. The LQ-MPC provides good performance but it cannot provide a sparse control solution. At the same time, the output behaviour obtained with ℓ_{asso}-MPC can closely resemble the one obtained with LQ-MPC.

CHAPTER

FOUR

VERSION 1: ℓ_1-INPUT REGULARISED QUADRATIC MPC

4.1 Introduction

This chapter introduces the first and simpler of the proposed ℓ_{asso}-MPC versions. The cost function is the same as for a standard quadratic MPC, with the addition of a 1-norm penalty on the input channels. Similarly to LASSO regression and the elastic-net, it is shown that this cost function can produce (for LTI systems) sparse input predictions under certain conditions. This characteristic is then shown to result in a control law that is non-linear, even without constraints. This is in clear contrast with LQ-MPC, the unconstrained solution of which is a constant gain. The explicit solution of ℓ_{asso}-MPC for LTI systems is shown to be unique, continuous and piecewise affine in the control error. The main feature of the control law is an *implicit input dead-zone*, namely, the regularised actuators being exactly at zero when the control error belongs to a certain neighbourhood of the origin. The shape and extent of this region depend on the cost function as well as the constraints. The explicit solution is computed for a small example to illustrate these properties. The solution path of ℓ_{asso}-MPC is then shown to be also piecewise affine, and it is computed by means of a single multi-parametric quadratic program.

The consequence of using this cost is an *implicit input dead-zone*, a better characterisation of ℓ_{asso}-MPC than the general term *sparsity* inherited from compressed sensing. In fact, because of the receding horizon nature of ℓ_{asso}-MPC, closed loop signals are not necessarily sparse even if their open loop predictions are so. For nominal systems sparsity is more likely to be obtained *among actuator channels* than through time. This will be further investigated towards the end of the chapter, leading to the concept of *preferred* and *auxiliary actuators* in Chapter 6.

Closed-loop nominal stability under ℓ_{asso}-MPC version 1 is investigated for non-linear systems with differentiable vector field and the origin as equilibrium. This is done using the same terminal ingredients as for a quadratic MPC, computed in Chapter 2. In particular, if the system is open-loop unstable, then regularising all actuators invalidates the standard assumptions for proving MPC stability. In this case, it is shown that there exist an upper bound on the 1-norm penalty for which the closed-loop trajectory is ultimately bounded in a neighbourhood of the origin. This is similar to a result from [Nagahara & Quevedo, 2011], with the difference that the constrained case is considered in this chapter. Secondly, the concept of partial regularisation is introduced, for which only some actuators are included in the 1-norm penalty. The strategy is shown to provide closed-loop asymptotic stability under certain conditions. As mentioned above, this leads to the concepts discussed in Chapter 6.

4.2 Regularised input behaviour

In this chapter, the terminal cost for the ℓ_{asso}-MPC problem (3.2.2) is assumed be

$$F(x) = x^T P x, \ P \succ 0. \tag{4.2.1}$$

The input behaviour of the unconstrained ℓ_{asso}-MPC is investigated first. The focus is on LTI models, in order to understand what the proposed controller has in common with LASSO regression and the elastic-net. Recall the definitions (2.6.20). Define also

$$W = I_N \otimes S, \tag{4.2.2}$$

and

$$\begin{aligned} \Lambda &= -\begin{bmatrix} S_Q \Theta \\ S_R \end{bmatrix}, \quad H = \Lambda^T \Lambda, \quad S_Q^T S_Q = \mathbf{Q}, \\ \underline{\mathrm{y}} &= \begin{bmatrix} S_Q \Psi x \\ 0 \end{bmatrix}, \quad \Gamma x = -\Lambda^T \underline{\mathrm{y}}, \quad S_R^T S_R = \mathbf{R}. \end{aligned} \tag{4.2.3}$$

The following is obtained:

Lemma 4.2.1. For the prediction model (2.5.7), problem (3.2.1) is an ℓ_1-regularized LS problem, subject to constraints.

Proof. (Lemma 4.2.1) Substitute (4.2.3) in (3.1.1) to obtain

$$(\underline{\mathrm{y}} - \Lambda\underline{\mathrm{u}})^2 + \|W\underline{\mathrm{u}}\|_1 = \underline{\mathrm{u}}^T H\underline{\mathrm{u}} + 2\underline{\mathrm{u}}^T\Gamma x + \|W\underline{\mathrm{u}}\|_1 + \underline{\mathrm{y}}^T\underline{\mathrm{y}}, \tag{4.2.4}$$

that is the cost function of ℓ_{asso}-MPC as well as of LASSO regression. ■

From Lemma 4.2.1, we can expect ℓ_{asso}-MPC version 1 to inherit some of the properties of LASSO. Assume, for the moment, that the constraints can be represented in the condensed form:

$$\Omega\underline{\mathrm{u}} \leq b(x), \tag{4.2.5}$$

where $b(x)$ is affine in x. From (4.2.4)–(4.2.5), problem (3.1.1) has the following Lagrangian:

$$L(\underline{\mathrm{u}}, v) = (\underline{\mathrm{y}} - \Lambda\underline{\mathrm{u}})^2 + \|W\underline{\mathrm{u}}\|_1 + (\Omega\underline{\mathrm{u}} - b(x))^T v, \tag{4.2.6}$$

where v is a vector of Lagrange multipliers. By sub-differentiating (4.2.6), with respect to $\underline{\mathrm{u}}$, the following KKT optimality conditions are obtained for (3.1.1) subject to (4.2.5):

$$2\Lambda^T(\underline{\mathrm{y}} - \Lambda\underline{\mathrm{u}}^\star) + \Omega^T v^\star \in W^T\,\partial\|W\underline{\mathrm{u}}^\star\|_1, \tag{4.2.7}$$

$$\Omega\underline{\mathrm{u}}^\star = b(x) - s^\star, \tag{4.2.8}$$

$$s^\star \cdot v^\star = 0, \tag{4.2.9}$$

$$s^\star \geq 0,\ \ v^\star \geq 0. \tag{4.2.10}$$

Condition (4.2.7) corresponds to [Kim *et al.* , 2007b; Osborne & Presnell, 2000]

$$2\Lambda^T(\underline{\mathrm{y}} - \Lambda\underline{\mathrm{u}}^\star) + \Omega^T v^\star = W^T\mathbf{v}, \tag{4.2.11}$$

where $\mathbf{v}$ can be any subgradient of $\|\cdot\|_1$ at $W\underline{\mathrm{u}}^\star$, namely, $\mathbf{v} \in \partial\|W\underline{\mathrm{u}}^\star\|_1$ and $\|\mathbf{v}\|_\infty = 1$.

4.2.1 Unconstrained solution & input dead-zone

Substituting (4.2.3) in (4.2.7), the unconstrained solution ($v^\star = 0$) of ℓ_{asso}-MPC for the LTI system (2.5.7) is (when $R \succ 0$ is assumed)

$$\underline{\mathrm{u}}^\star = -H^{-1}\left(\Gamma x(k) + \frac{1}{2}W^T\mathbf{v}\right). \tag{4.2.12}$$

In equation (4.2.12), the first element of the right-hand side corresponds to the unconstrained solution of a quadratic MPC problem, that is (3.2.2) for $S = 0$. The second is an additional piecewise-constant element. The following is obtained, the proof of which will be given in Section 4.4:

Theorem 4.2.2. Assume $R \succ 0$ and $S \in \mathcal{D}_{++}$. Then, for LTI systems, the solution of the unconstrained ℓ_{asso}-MPC version 1 is a unique, continuous PWA function of $x(k)$ and S.

Theorem 4.2.2 marks the fundamental difference between ℓ_{asso} and LQ-MPC: the unconstrained solution of the latter is a linear function of x, while the former is piece-wise affine (this will be further discussed in the later in chapter). On the other hand, the solution of ℓ_{asso}-MPC is unique as for an LQ-MPC, which in a sense certifies the input signals to be *well behaved*. This concept will be further discussed later in the chapter. The solution path for different values of S is also PWA as for the LASSO and the elastic-net regression (see Chapter 3.1.2).

Assume, for simplicity, $S \in \mathcal{D}_{++}$. Then, recalling (2.3.4), condition (4.2.11) becomes

$$-2(\Gamma x + H\underline{\mathbf{u}}^{\star})_i \in \begin{cases} \{+W_{ii}\} & \text{if } \underline{\mathbf{u}}_i^{\star} > 0, \\ \{-W_{ii}\} & \text{if } \underline{\mathbf{u}}_i^{\star} < 0, \\ [-W_{ii},\ W_{ii}] & \text{if } \underline{\mathbf{u}}_i^{\star} = 0. \end{cases} \tag{4.2.13}$$

Conditions (4.2.13) denote an interesting propriety of ℓ_{asso}-MPC, inherited from LASSO regression. In particular, the solution's behaviour is related to the soft-thresholding operator. The relation is in the following result:

Theorem 4.2.3. Assume $R \succ 0,\ S \in \mathcal{D}_{++}$ and no constraints. If $\underline{\mathbf{u}}_j^{\star} \neq 0,\ \forall j \neq i$, then

$$2\Big|(\Theta^T \mathbf{Q}\Psi x)_i + \sum_{j\neq i} H_{ij}\underline{\mathbf{u}}_j^{\star}\Big| < W_{ii} \Rightarrow \underline{\mathbf{u}}_i^{\star} = 0. \tag{4.2.14}$$

Proof. (Theorem 4.2.3) Substitute (4.2.3) into (4.2.13), to have $-2(\Theta^T \mathbf{Q}\Psi x + H\underline{\mathbf{u}}^{\star}) = W^T\mathbf{v}$. Assume $\underline{\mathbf{u}}_j^{\star} \neq 0,\ \forall j \neq i$. Since $W = \mathrm{Diag}(W_{ii})$ and $W_{ii} > 0$, the i-th row of $W^T\mathbf{v}$ is given by $W_{ii}\mathbf{v}_i$. From (4.2.13)-(4.2.14), and from Definition 2.3.9, it follows that $\mathbf{v}_i \in (-1,\ 1)$, giving (by Definition 2.3.9) $W_{ii}\underline{\mathbf{u}}_i^{\star} = 0$ and $\underline{\mathbf{u}}_i^{\star} = 0$. Finally, since $\underline{\mathbf{u}}^{\star}$ is PWA in x (see Theorem 4.4.2), then (4.2.14) defines a convex region of the state space, in which $\underline{\mathbf{u}}_i^{\star} = 0$. ∎

Remark 19. Theorem 4.2.3 implies that for a certain set of states some decision variables are zero. This set is convex since the $\underline{\mathbf{u}}^{\star}$ is PWA in x. It will be soon shown that this set contains the origin. Therefore, the unconstrained ℓ_{asso}-MPC version 1 has an implicit "open-loop zone".

Consider now problem (3.1.2), with $W = I$. Interestingly, the optimal Lagrange multiplier for (3.1.2) is

$$\lambda^\star(x, \underline{\mathrm{u}}^\star) = 2\|\Theta^T \mathbf{Q}(\Psi x + \Theta\, \underline{\mathrm{u}}^\star) + \mathbf{R}\, \underline{\mathrm{u}}^\star\|_\infty. \tag{4.2.15}$$

This suggests the following result, that stems directly from [Osborne & Presnell, 2000]:

Theorem 4.2.4. Assume $S \in \mathcal{D}_{++}$. Then, $\underline{\mathrm{u}}^\star = 0$ is the solution of the unconstrained ℓ_{asso}-MPC, iff

$$\min_i S_{ii} \geq \lambda^\star(x, 0) = 2\|\Theta^T \mathbf{Q}\Psi x\|_\infty. \tag{4.2.16}$$

Proof. (Theorem 4.2.4) From Theorem 4.2.3, the "if" condition, with inequality, follows by taking the minimum over i of the right-hand side of inequality (4.2.14) and the maximum of the left hand-side, for $\underline{\mathrm{u}} = 0$. Similarly, the strict equality and the "only-if" follows from the necessity and sufficiency of conditions (4.2.7)–(4.2.10). ■

As a consequence of Theorem 4.2.4, the unconstrained ℓ_{asso}-MPC sets all input signals to zero (system is open-loop) if the control error is in a certain convex neighbourhood of the origin. The coming sections will investigate how the input behaviour is modified by input and state constraints.

4.3 Implementation for LTI models

This section discusses two possible implementations of ℓ_{asso}-MPC version 1. These are helpful for the analysis of the explicit control law and do not necessarily provide the most efficient online implementations.

Suppose to have the polytopic constraints, $\mathbb{X} = \{x \mid Lx \leq \underline{1}\}$, $\mathbb{U} = \{u \mid Eu \leq \underline{1}\}$, and $\mathbb{X}_f = \{x \mid L_f\, x \leq \underline{1}\}$. Then, a constrained ℓ_{asso}-MPC for LTI systems can be implemented by solving, for instance, the following strongly convex non-smooth problem:

$$\begin{aligned} &\underline{\mathrm{u}}^\star = \arg\min_{\underline{\mathrm{u}}}\ \underline{\mathrm{u}}^T H \underline{\mathrm{u}} + 2\underline{\mathrm{u}}^T \Gamma x + \|W\underline{\mathrm{u}}\|_1 \\ &\text{s.t. } \Omega\, \underline{\mathrm{u}} \leq b(x),\ x = x(k), \end{aligned} \tag{4.3.1}$$

where

$$\Omega = \begin{bmatrix} (I_{(N\times N)} \otimes L)\Theta \\ L_f \phi \\ I_{(N\times N)} \otimes E \end{bmatrix}, \quad b(x) = \underline{1} - Mx, \quad M = \begin{bmatrix} (I_{(N\times N)} \otimes L)\Phi \\ L_f A^N \\ 0 \end{bmatrix}, \tag{4.3.2}$$

and where $\Phi = P\phi$. In the following, it will be shown how to formulate (4.3.1) as a smooth QP.

4.3.1 Using slack variables

Problem (4.3.1) can be re-formulated, for instance, by introducing mN slack variables, σ, and $2mN$ slack constraints, as the constrained semi-definite mpQP

$$\begin{aligned}&\chi^\star = \arg\min_\chi \tfrac{1}{2}\,\chi^T \bar{H}\chi + \chi^T \bar{G}(x)\\ &\text{s.t. } \bar{\Omega}\,\chi \leq \bar{b}(x),\; x = x(k),\end{aligned} \tag{4.3.3}$$

where

$$\chi = \begin{bmatrix} \underline{\mathrm{u}} \\ \sigma \end{bmatrix}, \quad \bar{H} = \begin{bmatrix} 2H & 0 \\ 0 & 0 \end{bmatrix}, \quad \bar{G}(x) = \begin{bmatrix} 2\Gamma x \\ \underline{1} \end{bmatrix}. \tag{4.3.4}$$

Constraint matrices for (4.3.3) are

$$\bar{\Omega} = \begin{bmatrix} \Omega \quad 0 \\ \Pi \end{bmatrix}, \quad \bar{b}(x) = \begin{bmatrix} b(x) \\ 0 \end{bmatrix}, \quad \Pi = \begin{bmatrix} W & -I_{mN} \\ -W & -I_{mN} \end{bmatrix}. \tag{4.3.5}$$

At each time k, the signal applied to the plant is given by

$$K_N(x(k)) = [I_m\; 0\; \ldots\; 0]\text{BlockDiag}(I_{mN},\, 0)\chi^\star. \tag{4.3.6}$$

Note that to implement (4.3.3) (either online or offline) it is necessary to use a solver that can handle dual degeneracy. This is however not necessary for the formulation proposed in the next section, which results in a positive definite mpQP. As a consequence of this, in the coming sections the optimal solution of ℓ_{asso}-MPC will be shown to be unique when $R \succ 0$, as already done for the unconstrained case in Theorem 4.2.2.

4.3.2 Splitting actuators for strict convexity

An alternative formulation is proposed in order to obtain a strictly convex QP. This, for instance, allows one to choose among a larger variety of online QP solvers as well as offline multi-parametric solvers [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002b]. In the next section, the formulation presented here will be used to explicitly characterise the solution of ℓ_{asso}-MPC.

A positive definite Hessian is needed. As noticed by [Tibshirani, 1996], (for $S \in \mathcal{D}_{++}$) the decision variables can be taken as $\underline{u}_+,\ \underline{u}_-$, with $\underline{u} = \underline{u}_+ - \underline{u}_-$. Then in [Tibshirani, 1996] the additional constraints $\underline{u}_+ \geq 0\ \underline{u}_- \geq 0$ are included in the formulation and $\bar{\Lambda} = [\Lambda,\ -\Lambda]$ can be used to define H. For the general case, [Tibshirani, 1996] has noted that this formulation only provides a positive semi-definite H. However, as noticed in [Hartley *et al.* , 2013], the choice of $R \in \mathcal{D}_{++}$ and $S \in \mathcal{D}_{++}$ for ℓ_{asso}-MPC allows one to re-formulate the problem with a positive definite Hessian. A more general result is proposed here, for any $R \succ 0$.

Consider

$$R \succ 0,\ S = \bar{S}R^{1/2},\ \bar{S} \in \mathcal{D}_{++} \subset \mathbb{R}^m_{\geq 0} \times \mathbb{R}^m_{\geq 0}. \tag{4.3.7}$$

In this case ℓ_{asso}-MPC version 1 can be implemented by using, for instance, the following condensed formulation:

$$\begin{array}{l} \chi^\star = \arg\min_\chi \frac{1}{2}\, \chi^T \bar{H} \chi + \chi^T \bar{G}(x) \\ \text{s.t. } \bar{\Omega}\, \chi \leq b(x),\ x = x(k). \end{array} \tag{4.3.8}$$

The above formulation is obtained thorugh the following steps:

1. Define the new coordinates $z = R^{1/2}u$.
2. Split variables in $(z_+,\ z_-)$ with $z = z_+ - z_-$.
3. Define $\chi = (\underline{z}_+,\ \underline{z}_-)$, namely, the predictions of z_+ and z_-, subject to a non-negativity constraint.

In the new coordinates, χ, problem (4.3.1) is equivalent to (4.3.8), where

$$\bar{H} = 2[\bar{\Theta},\ -\bar{\Theta}]^T \mathbf{Q} [\bar{\Theta},\ -\bar{\Theta}] + 2I,\quad \bar{W} = I_{2N} \otimes \bar{S},\quad \bar{G}(x) = 2[\bar{\Theta},\ -\bar{\Theta}]^T \mathbf{Q}\Psi x + \bar{W}\underline{1}\ , \tag{4.3.9}$$

and where

$$\bar{\Theta} = \Theta R^{-1/2} = \begin{bmatrix} BR^{-1/2} & 0 & \cdots & 0 \\ ABR^{-1/2} & BR^{-1/2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}BR^{-1/2} & \cdots & & BR^{-1/2} \end{bmatrix}. \tag{4.3.10}$$

The constraint matrices for (4.3.8) are

$$\bar{\Omega} = \begin{bmatrix} \Omega R^{-1/2} & -\Omega R^{-1/2} \\ & -I_{2mN} \end{bmatrix},\quad b(x) = \begin{bmatrix} \underline{1} - Mx \\ 0_{(2mN\times 1)} \end{bmatrix}. \tag{4.3.11}$$

At each time k, the signal applied to the plant is given by $u(k) = K_N(x(k))$, with

$$K_N(x(k)) = R^{1/2}[I\ 0\ \dots\ 0][I_{mN},\ -I_{mN}]\chi^{\star}. \tag{4.3.12}$$

Notice that in (4.3.8) we have $\bar{H} \succ 0$, namely, (4.3.8) is strictly convex. To obtain this in (4.3.9) the cross product elements of the Hessian, $-z_+ \cdot z_-$, have been neglected, and only the identity is kept in $\bar{H}$. This is possible because, at the optimum, it always holds that $z_+ \cdot z_- = 0$. To verify this condition, assume that $z^\star = z_+ - z_-$ with $z_+ \cdot z_- \neq 0$. The cost of this choice is the same for the state penalty, if $z^\star$ stays the same. On the other hand, the input cost is $z^\star \cdot z^\star + \underline{1}^T(Sz_+ + Sz_-)$ that is greater than $z^\star \cdot z^\star + \underline{1}^T S z^\star$, unless $-z_+ \cdot z_- = 0$. Therefore the solution of (4.3.8) and (4.3.1) are the same.

4.4 Explicit solution for LTI models

Uniqueness and continuity of the solution are a valuable feature for MPC as they imply that the actuators will not chatter. It is known, as verified in the previous chapter, that MPC based on ℓ_1 and ℓ_∞ costs can have multiple solutions [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002b], potentially leading to actuator chattering (for instance if two combinations of actuators have a similar effect on the plant) and to excessive wear and tear. For ℓ_{asso}-MPC version 1, we have the following result:

Theorem 4.4.1. Assume that $\mathbb{X}, \mathbb{U}, \mathbb{X}_f$ are C-sets and that $R \succ 0$. Then the solution of ℓ_{asso}-MPC version 1 is unique.

Proof. (Theorem 4.4.1) This is the same as the Proof of Theorem 4.2.2. In particular, since $R \succ 0$, the cost function of (4.3.1) is strongly convex. Then the result follows from Lemma 2.4.2 since the constraints are C-sets. ■

The explicit solution of ℓ_{asso}-MPC version 1 for constrained LTI system is characterised by the following, that follows from [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002a]:

Theorem 4.4.2. Assume $\mathbb{X}, \mathbb{U}$ are polytopic C-sets, and $\mathbb{X}_f = \mathcal{O}_\infty$, for a given stabilising K. Assume also that $R \succ 0$, and either $S = \bar{S}$ or $S = \bar{S}R^{1/2}$ hold for $\bar{S} \in \mathcal{D}_{++} \subset \mathbb{R}^m_{\geq 0} \times \mathbb{R}^m_{\geq 0}$. Then for ℓ_{asso}-MPC version 1 the following hold:

1. The feasible region, $\mathbb{X}_N$, is a polyhedron. Assume a fixed $S \in \mathbb{R}^m_{\geq 0} \times \mathbb{R}^m_{\geq 0}$. Then $\forall x \in \mathbb{X}_N$ the value function, $V_N^o(x)$, is continuous, convex, and piece-wise quadratic (PWQ). Moreover, the solution $K_N(x)$ is unique, piece-wise affine (PWA) and continuous, $\forall x \in \mathbb{X}_N$.

2. For every $x \in \mathbb{X}_N$ there exists a closed and bounded polytope, $\mathcal{S}_x \subset \mathbb{R}^m_{\geq 0}$, such that the solution path of ℓ_{asso}-MPC version 1 at x is the solution of a strictly convex mpQP, with parameter $\bar{s} = \mathrm{Diag}(\bar{S})$, and $\bar{s} \in \mathcal{S}_x$. The solution path at x is unique, continuous and PWA in $\bar{s}$.

3. There exists a closed and bounded polytope, $\mathcal{S} \subset \mathbb{R}^m_{\geq 0}$, such that, $\forall x \in \mathbb{X}_N$, the solution of ℓ_{asso}-MPC version 1 is unique, continuous and PWA in $(x, \bar{s})$ for $\bar{s} \in \mathcal{S}$.

Proof. (Theorem 4.4.2)

1. From the given assumptions, problem (4.3.1) can be formulated as the smooth positive definite mpQP (4.3.8), with polytopic constraints. Since the state appears only in the linear term of the mpQP, the claim follows from Theorem 1 of [Alessio & Bemporad, 2009].

2. Notice that, in (4.3.8), S also appears only in the linear term. Since we are enforcing constraints, from (4.2.7) it can be deduced that there exists an upper bound on the magnitude of $\mathcal{S}$, above which the solution will not change. Therefore, the claim follows again from Theorem 1 of [Alessio & Bemporad, 2009].

3. This is a consequence of points 1 and 2. In particular, both x and S appear only in the linear terms of the mpQP, that can be therefore parameterised in terms of both.

$\blacksquare$

Remark 20. Extensions of Theorem 4.4.2 are possible, for stage costs featuring additional cross state-input terms, and for different structures of S. In particular, the multi-parametric programming techniques in [Alessio & Bemporad, 2009; Bemporad *et al.* , 2002a; Tondel *et al.* , 2003] can be used to express $K_N(x)$ as a unique PWA function of x, as outlined in Section 2.6.8. This is possible even when R is positive semi-definite, in which case the algorithm in [Tondel *et al.* , 2003] can be used to obtain a continuous solution. This is generally done for MPC based on 1-norm and ∞-norm costs to avoid chattering. Chattering cannot occur for ℓ_{asso}-MPC, when $R \succ 0$.

We are now ready to prove Theorem 4.2.2:

Proof. (Theorem 4.2.2) The result is a special case of Theorem 4.4.2. In particular, without state and input constraint problem (4.3.8) still features the non-negativity constraints. Therefore, each unbounded domain $\mathcal{X} \subset \mathbb{R}^n$ can be divided into a finite number of critical regions, and the unique solution can be computed as the PWA function in [Bemporad *et al.* , 2002a] as reviewed in Chapter 2.6.8. In particular, from Theorem 4.2.4, the solution path can be computed at x by restricting the domain of the parameter, $\bar{s} = \mathrm{Diag}(S)$, to $\mathcal{S}_x = \{\bar{s} \mid \bar{s}_i \in (0, \lambda^\star(x, 0)], \forall i\}$. ■

4.4.1 Example 1: Soft-thresholding and dead-zone

A scalar LTI system is considered, with $A = B = Q = R = 1$, $N = 3$, $S = 100$, $|x| \leq 5$, $|u| \leq 1$, and $P = 1.6180$ solving the DARE. The terminal constraint is $|x_N| \leq 1.6180$. The explicit solution of ℓ_{asso}-MPC is computed using the Matlab Multi-Parametric Toolbox [Kvasnica *et al.* , 2004]. In order to do so, the input is again divided in two positive components, $u = u_+ - u_-$, providing $\bar{H} \succ 0$ (required by the MPT). The control law, in Figure 4.1, is a PWA function of x with a *dead-zone*. In particular, when x is in the terminal set, the plant is left open-loop. Notice that every x in the feasible region is steered to the terminal set[1]. It should be noticed that, without constraints, Theorem 4.2.4 would provide $\underline{\mathrm{u}}^\star = 0$, for all $|x| \leq 13.9$. The PWQ optimal cost, $V_N^o(x)$, is shown

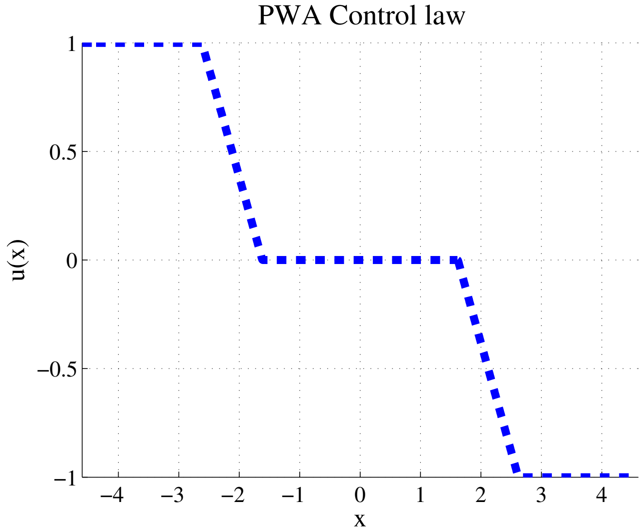

Figure 4.1: Explicit ℓ_{asso}-MPC for a scalar system: Control law is PWA with a *dead-zone*.

[1]This happens despite the assumption (A7) for ultimate boundedness (proposed in the next section) not being satisfied.

in Figure 4.2.

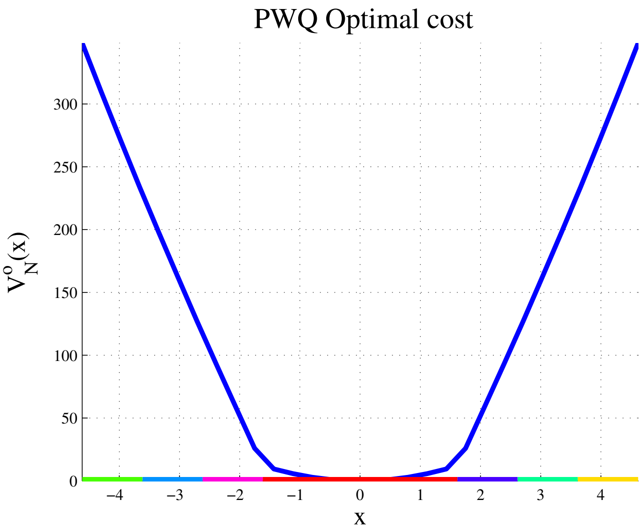

Figure 4.2: Explicit ℓ_{asso}-MPC for a scalar system: Optimal cost is PWQ. Coloured areas indicate state space partitions.

4.4.2 Example 2: Solution path for all states

The solution path is computed for $S \in (0, 30]$. The control law, in Figure 4.3, 4.4, is a PWA in x, S. It can be noticed, in Figure 4.3, how initially the *dead-zone* varies linearly with S, to then stop for approximately $S \geq 11$. This is due to the terminal constraint. The PWQ optimal cost shown is in Figure 4.5. The considered example is very simple. More complex behaviour can be expected in more complex examples, depending on the system model and on the constraints.

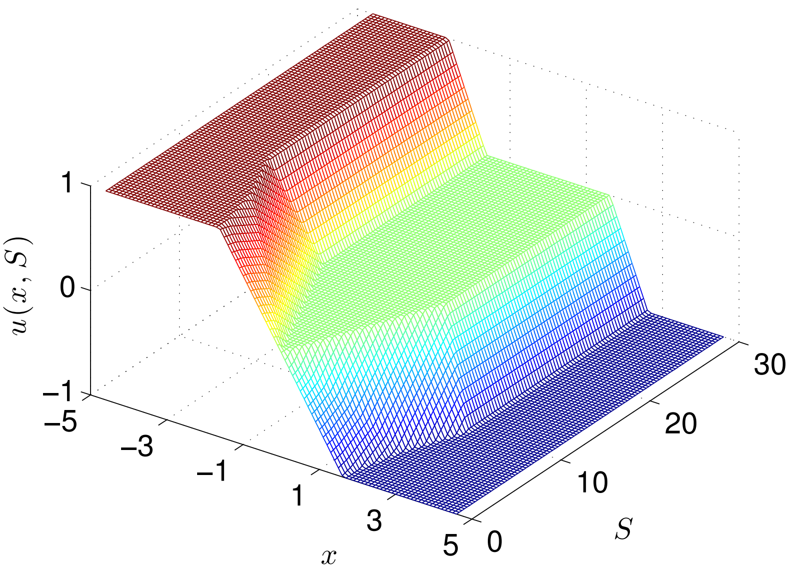

Figure 4.3: ℓ_{asso}-MPC solution path for a scalar system: The dead-zone varies with S.

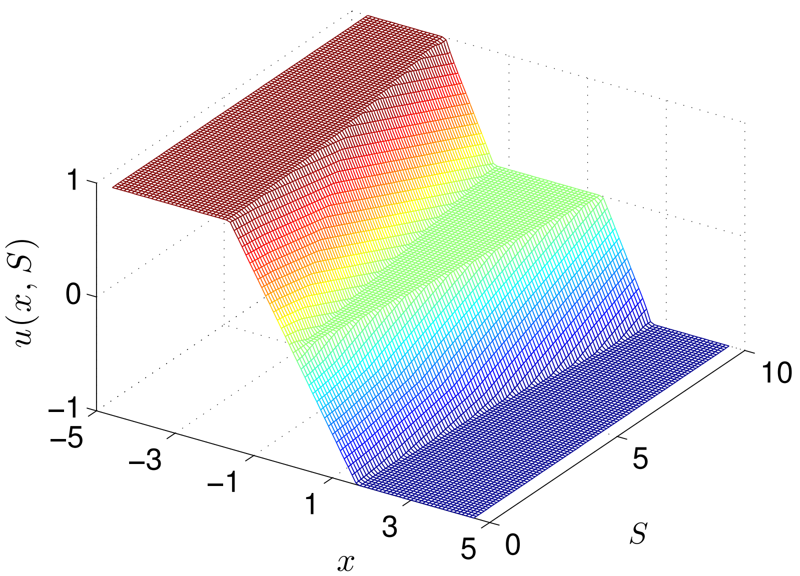

Figure 4.4: ℓ_{asso}-MPC solution path for a scalar system (zoom): The dead-zone varies with S.

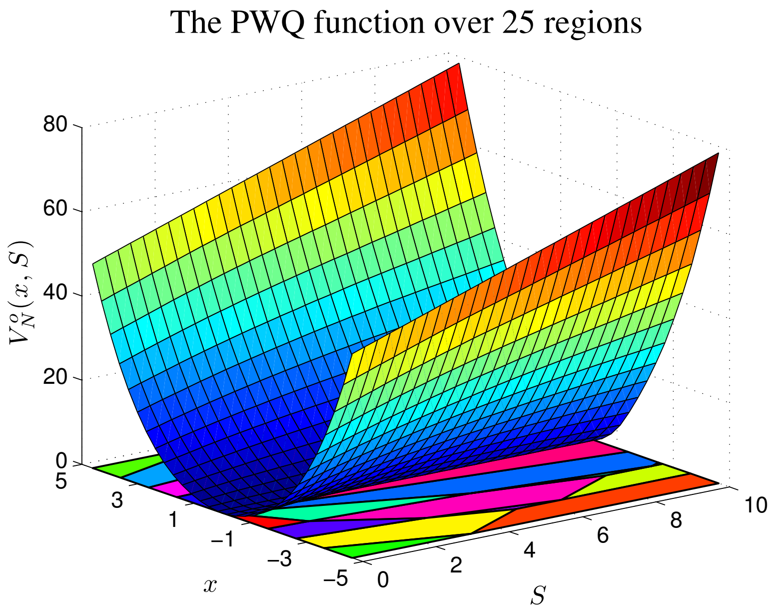

Figure 4.5: Explicit ℓ_{asso}-MPC for a scalar system: Optimal cost is PWQ in $x,\ S$. Coloured areas indicate the x, S space partitions.

4.5 Closed-loop stability analysis

This section concerns the study of closed-loop nominal stability under ℓ_{asso}-MPC version 1. From the previous sections, it seems to be clear that asymptotic stability of the origin is not possible for an arbitrary matrix S. This will be first formalised in terms of ultimate boundedness of the solution, that is obtained under certain assumptions on S. Then, it will be shown that asymptotic stability can be achieved in the redundantly-actuated case if only a subset of the actuators is regularised. The focus is on non-linear systems with a differentiable vector field and the origin as an equilibrium. The proposed results rely on the standard terminal ingredients for MPC.

4.5.1 General regularisation penalty and ultimate boundedness

Recall, from (2.6.5), the definition of the i-step closed-loop evolution from x under a linear controller $u = Kx$, denoted $\bar{f}_K^i(x)$. Since the terminal cost is quadratic and the stage cost features a 1 norm, the assumption (A4) cannot be satisfied by means of algebraic manipulations or scaling. For instance, suppose that (A4) is satisfied for the quadratic part of the cost. Since $F(x)$ is a CLF in $\mathbb{X}_f$, with exponential stability, there exists a $\lambda \in [0, 1)$ such that $F(\bar{f}_K^{i+1}(x)) - F(\bar{f}_K^i(x)) \leq (\lambda - 1)F(\bar{f}_K^{i+1}(x))$. Take $i = 0$, for simplicity, and try imposing $\alpha(\lambda-1)\|P^{1/2}x\|_2^2 \leq -\|SKx\|_1$, from some $\alpha > 0$, in a neighbourhood of the origin. This, unless when $\|SKx\|_1 = 0$, is impossible for any finite α, since $\|SKx\|_1/\|P^{1/2}x\|_2^2$ diverges to ∞ as $\|x\| \to 0$. This shows that standard MPC assumptions are difficult to be met by means of this terminal cost, unless $\|SKx\|_1 = 0, \ \forall x$.

Let's first assume to have a quadratic MPC, to which we wish to add a general regularisation penalty $\|Su\|_1$. Consider the following:

Assumption 6. (A6)

Assume $Q \succ 0, \ R \succ 0$, and that $(\mathbf{H1})$ from (A2) holds. Assume also that the set $\mathbb{X}_f$ is PI under $u = Kx$, with

$$F(\bar{f}_K^{i+1}(x)) - F(\bar{f}_K^i(x)) \leq -\ell(\bar{f}_K^i(x), K\bar{f}_K^i(x)) + \|SK\bar{f}_K^i(x)\|_1, \ \forall x \in \mathbb{X}_f, \ \forall i \in \mathbb{I}_{\geq 0}. \quad (4.5.1)$$

If the quadratic MPC satisfies (A2) and (A4), then it can be easily verified that ℓ_{asso}-MPC satisfies **(A6)**. Define

$$\mathrm{cond}(P) = \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}. \quad (4.5.2)$$

Note that from **(A6)** we have that $\mathrm{cond}(P) \in [1, \ \infty)$.

To aim for practical stability and ultimate boundedness of the solution, the following is needed:

Lemma 4.5.1. Assume (A6). Then, there exists two $\mathcal{K}_\infty$-functions, $\alpha_1(r),\ \alpha_2(r)$, such that

$$\alpha_1(\|x\|_2) \le V_N^o(x) \le \alpha_2(\|x\|_2),\ \forall x \in \mathbb{X}_N, \tag{4.5.3}$$

where, $\alpha_1(r) = \lambda_{\min}(Q)r^2$, and

$$\begin{aligned}\alpha_2(r) = \epsilon\tilde{\alpha}_2(r),\ \epsilon = \max\left(1, \bar{V}_N/\tilde{\alpha}_2(\bar{r})\right),\ \bar{V}_N \ge V_N^o(x)\ \forall x \in \mathbb{X}_N,\ \bar{r} > 0,\\ \tilde{\alpha}_2(r) = \lambda_{\max}(P)r^2 + \sqrt{m_s}N\,\mathrm{cond}(P)\ \|SK\|_2 r.\end{aligned} \tag{4.5.4}$$

Proof. (Lemma 4.5.1) Trivially, $\alpha_1(\|x\|_2) = \lambda_{\min}(Q)\|x\|_2^2$, follows from $\ell(x,u) \ge \lambda_{\min}(Q)\|x\|_2^2, \forall u$, and $V_N^o(x) \ge \ell(x, K_N(x)), \forall x$. To prove the existence of $\alpha_2(\|x\|_2)$, define the sequence $\tilde{u} = \{Kf^i(x),\ i = 0,\ldots,N-1\}$, that is a feasible solution for $x \in \mathbb{X}_f$. By optimality and by recursive application of (4.5.1) it follows that $\forall x \in \mathbb{X}_f$:

$$\begin{aligned} V_N^o(x) \le V_N(x,\tilde{u}) = F(x) + \sum_{i=0}^{N-1} \|SK\bar{f}_K^i(x)\|_1 \le \\ F(x) + \sum_{i=0}^{N-1} \sqrt{m_s}\|SK\bar{f}_K^i(x)\|_2 \le \\ F(x) + \sum_{i=0}^{N-1} \sqrt{m_s}\|SK\|_2\ \|\bar{f}_K^i(x)\|_2 \le \\ F(x) + \sqrt{m_s}N\|SK\|_2\,\mathrm{cond}(P)\|x\|_2 = \tilde{\alpha}(\|x\|_2), \end{aligned} \tag{4.5.5}$$

where in the last inequality we have used the fact that, from **(A6)** , $\|P^{1/2}\bar{f}_K^i(x)\|_2 \le \|P^{1/2}x\|_2\ \forall i \ge 0$. The above $\tilde{\alpha}(r)$ is a $\mathcal{K}_\infty$ function. Since $\mathbb{X}$ and $\mathbb{U}$ are C-sets then so is $\mathbb{X}_f$. Therefore the existence of such $\tilde{\alpha}(\|x\|_2)$ for all $x \in \mathbb{X}_f$ implies [Limón *et al.* , 2006; Rawlings & Mayne, 2010] the existence of another $\mathcal{K}_\infty$ function, $\alpha_2(\|x\|_2) \ge V_N^o(x), \forall x \in \mathbb{X}_N$, of the desired form. Compactness of $\mathbb{X}$ can also be relaxed by considering a C-set of initial conditions as done by [Limón *et al.* , 2006]. ∎

A limitation on S is now imposed. Define:

$$\begin{aligned}\delta_1 = \max\nolimits_{x\in\mathbb{X}_N} \|x\|_2^2,\quad \delta_2 = \sqrt{m_s}\cdot\max\nolimits_{x\in\mathbb{X}_f}\|SKx\|_2,\\ \delta_3 = \sqrt{m_s}\cdot\max\nolimits_{x\in\mathbb{X}_f}\|Kx\|_2.\end{aligned} \tag{4.5.6}$$

Define also the C-set

$$\mathcal{B} = \{x \mid \|x\|_2^2 \le \gamma\}, \;\; \gamma = \delta_2/\sigma_{\min}(Q) > 0. \tag{4.5.7}$$

The following relation is assumed for Q and S:

Assumption 7. (A7)
$\sigma_{\max}(S) < \lambda_{\min}(Q)\delta_1/\delta_3$.

A set of lemmas is now presented, that leads to the first main result:

Lemma 4.5.2. Assume (A6), (A7). Then, $\mathbb{X}_N \backslash \mathcal{B}$ is not empty.

Proof. (Lemma 4.5.2) From (A7), it follows that

$$\gamma = \delta_2/\lambda_{\min}(Q) \le \delta_3\sigma_{\max}(S)/\lambda_{\min}(Q) < \delta_1, \tag{4.5.8}$$

where $\delta_1 = \max_{x\in\mathbb{X}_N} \|x\|_2^2$. From the above result we have:

$$\mathcal{T} = \{x \mid \gamma < \|x\|_2^2 \le \delta_1\} \ne \emptyset, \;\; \mathcal{T} \subseteq \mathbb{X}_N, \;\; \mathcal{B} \cap \mathcal{T} = \emptyset,$$

that is $\mathcal{T} \subseteq \mathbb{X}_N \backslash \mathcal{B}$, implying that $\mathbb{X}_N \backslash \mathcal{B}$ is not empty. ∎

Lemma 4.5.3. Assume (A6), (A7). Then,

$$\ell(x, K_N(x)) > \|SKx_N\|_1, \; \forall x_N \in \mathbb{X}_f, \; \forall x \in \mathbb{X}_N \backslash \mathcal{B}. \tag{4.5.9}$$

Proof. (Lemma 4.5.3) Assume (A7). Then, from Lemma 4.5.2, $\mathbb{X}_N \backslash \mathcal{B}$ is not empty. Consider any $x \in \mathbb{X}_N \backslash \mathcal{B}$. Then, it follows that

$$\begin{aligned} \|x\|_2^2 > \gamma = \delta_2/\lambda_{\min}(Q) \Rightarrow \\ \lambda_{\min}(Q)\|x\|_2^2 > \sqrt{m_s} \max_{x\in\mathbb{X}_f} \|SKx\|_2 \ge \\ \ge \max_{x\in\mathbb{X}_f} \|SKx\|_1. \end{aligned} \tag{4.5.10}$$

Since, from (A6) and Lemma 4.5.1, $\lambda_{\min}(Q)\|x\|_2^2$ is a lower-bound of $\ell(x, u)$ for all u, then we have:

$$\ell(x, K_N(x)) > \|SKx_N\|_1, \; \forall x_N \in \mathbb{X}_f, \; \forall x \in \mathbb{X}_N \backslash \mathcal{B},$$

as claimed. ∎

Define $\mathcal{T} = \{x \mid V_N^o(x) \leq b\}$, $b = \max_{x \in \mathbb{X}_N \setminus \mathcal{B}}\{V_N^o(x)\}$. We are now ready to state the first stability result:

Theorem 4.5.4. Assume (A6), (A7). Then under the ℓ_{asso}-MPC version 1 it follows that for all $x \in \mathbb{X}_N$:

1. The trajectory of (2.5.8) is ultimately bounded (UB) in $\mathcal{T}$, that is: $x(k) \to \mathcal{T}$ as $k \to \infty$, $\forall x(0) \in \mathbb{X}_N$ and $\mathcal{T}$ is PI (so is $\mathbb{X}_N$),

2. System (2.5.8) is practically stable, namely:

$$\begin{aligned} \|x(k)\|_2 \leq \beta(\|x(0)\|_2, k) + \zeta, \\ \beta(r,k) = \sqrt{\frac{\rho^k}{\lambda_{min}(Q)}\alpha_2(r)},\ \rho = 1 - \frac{\lambda_{min}(Q)}{\alpha_2(1)}, \\ \zeta = \sqrt{\frac{1}{1-\rho}\left(\gamma + \frac{1}{4}\right)}. \end{aligned} \tag{4.5.11}$$

Notice that $\rho \in [0,\ 1)$ and hence β is a $\mathcal{KL}$-function.

Proof. (Theorem 4.5.4) The proof relies on the direct method in [Mayne *et al.* , 2000; Rawlings & Mayne, 2010]. The one-step evolution of (2.5.8), from a given x, is denoted as x^+. From Lemma 4.5.1 the optimal cost function, $V_N^o(x)$, is positive definite with $V_N^o(0) = 0$. Given, at time k, the optimal sequence, $\underline{\mathrm{u}}^\star = \{\hat{u}^\star_{0|k}, \ldots, \hat{u}^\star_{N-1|k}\}$, then the sequence, $\tilde{\underline{\mathrm{u}}} = \{\hat{u}^\star_{1|k}, \ldots, \hat{u}^\star_{N-1|k}, K\hat{x}^\star_{N|k}\}$, is also admissible at time $k+1$. Hence

$$\begin{aligned} V_N^o(x^+) \leq V_N(x, \tilde{\underline{\mathrm{u}}}) = V_N^o(x) - \ell(x, K_N(x)) \\ -F(\hat{x}^\star_{N|k}) + \ell(\hat{x}^\star_{N|k}, K\hat{x}^\star_{N|k}) + F(\bar{f}_K^1(\hat{x}^\star_{N|k})). \end{aligned} \tag{4.5.12}$$

Then, from (A6), we have:

$$V_N^o(x^+) \leq V_N^o(x) - \ell(x, K_N(x)) + \|SK\hat{x}^\star_{N|k}\|_1. \tag{4.5.13}$$

Assume now (A7). Then, from Lemma 4.5.3, it follows that

$$\ell(x, K_N(x)) > \|SKx_N\|_1,\ \forall x_N \in \mathbb{X}_f,\ \forall x \in \mathbb{X}_N \setminus \mathcal{B}. \tag{4.5.14}$$

1. Combine (4.5.13) with (4.5.14) to have $V_N^o(x^+) - V_N^o(x) < 0,\ \forall x \in \mathbb{X}_N \backslash \mathcal{B}$. This, together with Lemma 4.5.1, results in $V_N^o(x)$ being a Lyapunov function in $\mathbb{X}_N \backslash \mathcal{B}$. Hence, according to Theorem 2.5.8 (namely to Theorem 2.24 of [Blanchini & Miani, 2008]), the trajectory of (2.5.8) is UB in $\mathcal{T}$ for all $x \in \mathbb{X}_N$.

2. The proof proceeds similarly to [Nagahara & Quevedo, 2011]. Recall (4.5.12) and Lemma 4.5.1, to have:

$$
\begin{aligned}
V_N^o(x^+) &\leq V_N^o(x) - \lambda_{\min}(Q)\|x\|_2^2 + \|SKx_N\|_1, \\
&\leq V_N^o(x) - \lambda_{\min}(Q)\|x\|_2^2 + \delta_2, \\
&\leq \left(1 - \frac{\lambda_{\min}(Q)\|x\|_2^2}{V_N^o(x)}\right) V_N^o(x) + \delta_2, \\
&\leq \left(1 - \frac{\lambda_{\min}(Q)\|x\|_2^2}{\alpha_2(\|x\|_2)}\right) V_N^o(x) + \delta_2.
\end{aligned}
$$

If $\|x\|_2^2 \leq \|x\|_2$, it follows that

$$
\begin{aligned}
V_N^o(x^+) \leq{}& \left(1 - \frac{\lambda_{\min}(Q)\|x\|_2}{\alpha_2(\|x\|_2)}\right) V_N^o(x) + \delta_2 \\
& + \lambda_{\min}(Q)(\|x\|_2 - \|x\|_2^2), \qquad (4.5.15)\\
\leq{}& \left(1 - \frac{\lambda_{\min}(Q)}{\alpha_2(1)}\right) V_N^o(x) + \delta_2 + \frac{\lambda_{\min}(Q)}{4}. \qquad (4.5.16)
\end{aligned}
$$

Recalling (4.5.2) we have:

$$
\frac{\lambda_{\min}(Q)}{\alpha_2(1)} = \frac{\lambda_{\min}(Q)}{\epsilon(\lambda_{\max}(P) + \sqrt{m_s}N\ \mathrm{cond}(P)\ \|SK\|_2)} < 1,
$$

since $\epsilon \geq 1$, $\lambda_{\max}(P) \geq \lambda_{\min}(Q)$, $\|SK\|_2 > 0$ and $\mathrm{cond}(P) \geq 1$.

For $\|x\|_2 \leq \|x\|_2^2$, we have:

$$
\begin{aligned}
V_N^o(x^+) &\leq \left(1 - \frac{\lambda_{\min}(Q)\|x\|_2}{\alpha_2(\|x\|_2)}\right) V_N^o(x) + \delta_2 \\
&\leq \left(1 - \frac{\lambda_{\min}(Q)}{\alpha_2(1)}\right) V_N^o(x) + \delta_2 + \frac{\lambda_{\min}(Q)}{4}. \qquad (4.5.17)
\end{aligned}
$$

Iterating the above inequality over time we have:

$$V_N^o(x(k)) \leq \rho^k V_N^o(x(0)) + \sum_{i=0}^{k-1} \rho^i \left(\delta_2 + \frac{\lambda_{\min}(Q)}{4}\right).$$

From the inequalities of Lemma 4.5.1, and from the fact that $\rho \in [0, 1)$, it follows that

$$\begin{aligned}\|x(k)\|_2^2 &\leq \frac{\rho^k}{\lambda_{\min}(Q)}\alpha_2(\|x(0)\|_2) + \frac{1}{1-\rho}\left(\frac{\delta_2}{\lambda_{\min}(Q)} + \frac{1}{4}\right)\\ &= \frac{\rho^k}{\lambda_{\min}(Q)}\alpha_2(\|x(0)\|_2) + \frac{1}{1-\rho}\left(\gamma + \frac{1}{4}\right),\end{aligned}$$

hence the claim follows.

■

Theorem 4.5.4 stated that the system's closed-loop trajectory under an ℓ_1-regularised quadratic MPC might not converge to the origin but will asymptotically stay in neighbourhood of it, the size of which depends on S. In the next section it will be shown how, under less general assumptions, closed-loop asymptotic stability is achieved.

Remark 21. For LTI systems, if $Q \succeq 0$, then Theorem 4.5.4 is valid for $y(k) = Q^{1/2}x(k)$. Assume $(Q^{1/2}, A)$ is detectable. Then it can be shown that $x(k)$ is also ultimately bounded.

Remark 22. The proposed results are sufficient. The bound $\mathcal{T}$ can be conservative as well as (A7). The feasible region $\mathbb{X}_N$ does not change with S, and it is PI. The δ_1 in the proof can be considered as an abstract quantity, the computation of which is not necessary for using the controller.

Remark 23. The results of [Nagahara & Quevedo, 2011] have been inspirational for Theorem 4.5.4. However, the proof presented here differs from [Nagahara & Quevedo, 2011] for several reasons. The first being that here the terminal cost is not modified but a limitation on S is imposed instead. The two ideas could be in principle equivalent. A stronger difference is that the proposed results consider non-linear systems with constraints and multiple inputs, while [Nagahara & Quevedo, 2011] focused on single inputs LTI systems (as well as on robustness to communication uncertainties in the case of remote control, not addressed in this thesis). In the proposed results constraints are dealt by means of (A7) and the assumptions on S are based on Q, not directly to P.

4.5.2 Partial regularisation and asymptotic stability

Assume now that only some actuators are regularised, namely that the 1-norm penalty in (3.2.2) involves only a subset of inputs labelled $u^{\{\mathrm{ii}\}}$. The remaining set of actuators (non-regularised) are instead labelled $u^{\{\mathrm{i}\}}$. This means, for instance, that

$$S = \begin{bmatrix} 0 & S_{\{\mathrm{ii}\}} \end{bmatrix}, \tag{4.5.18}$$

for some matrix, $S_{\{\mathrm{ii}\}}$. Note that, in (4.5.18), the actuators have been grouped as $u = (u^{\{\mathrm{i}\}},\ u^{\{\mathrm{ii}\}})$. This approach, that we shall refer to as *partial regularisation*, can be used in applications where there is a clear set of *main* or *preferred* actuators (namely $u^{\{\mathrm{i}\}}$) and an *auxiliary* or *backup actuators* set ($u^{\{\mathrm{ii}\}}$). This type of scenarios will be further discussed in Chapter 6.

The scope of this section is to show how the use of an S matrix in the form of (4.5.18) can provide asymptotic stability. In order to do this, let us introduce the notation $\Sigma_{\{\mathrm{i}\}}$ (respectively $\Sigma_{(\mathrm{ii})}$), referring to system (2.5.6) when only $u^{\{\mathrm{i}\}}$ (respectively $u^{\{\mathrm{ii}\}}$) is used. The following is obtained:

Theorem 4.5.5. Assume $Q \succ 0,\ R \succ 0$, and (A2). Assume also that (A4) holds for system $\Sigma_{\{\mathrm{i}\}}$. Then, under the ℓ_{asso}-MPC version 1 with S given by (4.5.18), it follows that the origin of (2.5.8) is AS $\forall\, x \in \mathbb{X}_N$.

Proof. By assumption, there exists a suboptimal and stabilising linear controller, $u^{\{\mathrm{i}\}} = Kx$, $u^{\{\mathrm{ii}\}} = 0$, in $\mathbb{X}_f$. The overall stabilising input vector is then $u = [K^T\ 0]^T x$. Since (A4) is satisfied for $\Sigma_{\{}\mathrm{i}\}$, for the overall system we have that (A6) is again satisfied, with the difference that, from (4.5.18) , we have:

$$S\begin{bmatrix} K \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & S_{\{\mathrm{ii}\}} \end{bmatrix}\begin{bmatrix} K \\ 0 \end{bmatrix} = 0, \tag{4.5.19}$$

thus satisfying (A4) also for the full system. Then the result is obtained by application of Theorem 2.24, p. 123 of [Rawlings & Mayne, 2010], for instance by means of the direct method argument of [Mayne *et al.* , 2000] reviewed in Chapter 2.6.4 of this thesis (Theorem 2.6.4). ■

Remark 24. For the LTI case, it is assumed that $\Sigma_{\{\mathrm{i}\}}$ is stabilisable. If $Q \succeq 0$, then Theorem 4.5.5 is valid for $y(k) = Q^{1/2}x(k)$. Assume $(Q^{1/2}, A)$ is detectable. Then the origin of the state space is also AS.

Theorem 4.5.5 opens to an interesting application of ℓ_{asso}-MPC version 1: given a nominal quadratic MPC design it is possible to include new actuators in the loop, and to limit their behaviour through regularisation. From what seen in this chapter, we can expect the regularised actuator not to be in operation for the whole time, but to be set into operation only in particular regions of the state space. This will be further discussed in Chapter 6.

4.5.3 Computational aspects for LTI and non-linear systems

The proposed stability results rely on the computation of a control invariant terminal set and of a control Lyapunov function for the considered system. Providing that these ingredients can be computed then the system stability can be determined by Theorem 4.5.4 or Theorem 4.5.5. For LTI systems it is common to consider the LQR controller and its value function, as shown in Chapter 2.6.5.1. Even when more intensive methods are used (for instance to obtain a larger terminal set and a larger MPC feasible region) the required ingredients can be computed by means of relatively standard tools, providing the system is detectable, the state and input dimensions are contained (approximately less than 6) and the constraints contain the origin in their interior. While the required offline complexity still remains a bottleneck in the application of terminal constraints in MPC, even for LTI models, the computation is generally contained in this case.

Differently from the linear case, considering non-linear models introduces additional complexity. It is clear that online complexity is increased due to the loss of convexity in the MPC cost. The offline complexity is also significantly increased. In particular, using LDIs as in Chapter 2.6.5.2, requires first the computation of (at least) one LPV system that locally contains the vector field gradient (assuming the system to be differentiable and the origin to be an equilibrium). This is not a trivial task. Moreover, the selected LPV system may not be controllable by means of a piecewise linear state feedback or, even when this is possible, the required LMI (2.6.17) may not have a solution. In this case it might be necessary to restrict the domain of interest and compute another local LPV system. This procedure may be iterated several times without success, for instance if the system is not stabilisable by means of a quadratic Lyapunov function [Blanchini, 1999; Blanchini & Miani, 2008]. The next chapter will deal with the more general polytopic Lyapunov functions, however, the computation of the terminal ingredients remains the main bottleneck for the use of MPC with terminal constraint for non-linear models. A possible alternative is to use a local LQR controller, as done in Lemma 2.6.5, to obtain the terminal ingredients. This could however result in a very small terminal set and feasible region (recall that this is also the domain of attraction).

4.6 Conclusions

This chapter has analysed the behaviour of ℓ_{asso}-MPC version 1, an ℓ_1-regularised MPC with quadratic terminal cost. For LTI systems the explicit control solution is a unique, continuous piece-wise affine functions of the system state. Moreover, the regularised control signals are equal to zero in a neighbourhood of the origin. In other words, ℓ_{asso}-MPC version 1 features an *implicit input dead-zone* or *open-loop zone*. Differently from ℓ_1 and ℓ_∞-MPC, which can also leave the system in open loop when close to the origin, the uniqueness and continuity of the solution of ℓ_{asso}-MPC version 1 provide the advantage that actuator chattering is not possible for LTI models. Furthermore, it has been shown that the set of ℓ_{asso}-MPC solutions for a continuous range of regularisation penalties is also continuous and can be computed by means of a single multi-parametric quadratic program.

Closed-loop ultimate boundedness has been proven for a class of non-linear systems, under a particular assumption on the regularisation penalty. For the special case where only some actuators are regularised, it is has also been shown that the closed-loop system origin can be made asymptotically stable. This requires the system to be stabilisable by means of the non-regularised actuators.

The concept of implicit input dead-zone introduced in this chapter can help making realistic speculations on the input signals behaviour, being more specific than the term sparsity. The results from this chapter suggest a particular application of ℓ_{asso}-MPC version 1: given a nominal quadratic MPC, it is possible to include new actuators in the loop, and to limit their behaviour through regularisation. This will be further discussed in Chapter 6.

CHAPTER

FIVE

VERSION 2: LASSO MPC WITH STABILISING TERMINAL COST

5.1 Chapter outline

This chapter presents a modified version of ℓ_{asso}-MPC that guarantees closed-loop asymptotic stability for arbitrary 1-norm input penalties. The approach is based on a new terminal cost and terminal constraint. Appropriate scalings are computed to adjust the proposed former according to the stage cost matrices. In particular, two strategies are presented to compute the required ingredients. The first one makes use of norm inequalities and linear matrix inequalities. The required assumptions can, however, be difficult to meet as they might require the solution of a nonlinear program. A second approach is then proposed, which makes use of set-theoretic control principles. In particular, the proposed terminal cost is based on the Minkowski function associated with the terminal set, which is taken to be λ-contractive.

Computation of the terminal set is one of the main bottlenecks in the design of MPC, if standard assumptions are to be fulfilled. This could limit real world applications of most common MPC theory. In this chapter the terminal set is computed offline, independently of the stage cost. This is in contrast to most standard MPC theory, where the terminal set depends on the stage cost and must therefore be re-computed for tuning. The proposed approach aims in fact to facilitate online tuning, thus reducing the gap between theory and applications. A technique is reviewed for the computation of a relatively large terminal set, which makes use of linear programming. This can potentially increase the domain of attraction of the controller. A simple example is given to illustrate the proposed techniques.

For LTI models the proposed ℓ_{asso}-MPC version 2 requires the solution of a strongly convex quadratic program, similar to version 1. This enjoys a unique piece-wise affine optimiser. The conditions for having sparse solutions in terms of actuator channels are however slightly different from ℓ_{asso}-MPC version 1. In particular, the presence of a norm in the terminal cost makes sparsity more difficult to be obtained. A set of conditions and arrangements are discussed that provide sparse solutions.

To the author's knowledge [Gallieri & Maciejowski, 2013a, 2012, 2013b] are the first and only references to provide asymptotic stability conditions for ℓ_{asso}-MPC or any other regularised MPCs. In Section 5.2, the approach of [Gallieri & Maciejowski, 2013b] is reviewed and extended to non-linear systems through LPV embedding. The approach based on Minkowski functions, presented in Section 5.3 as well as in [Gallieri & Maciejowski, 2013a], is computationally less demanding than the ones in [Gallieri & Maciejowski, 2013b].

5.2 Terminal cost using LMIs and norm inequalities

Assume the terminal cost to be[1]:

$$F(x) = \|Zx\|_1 + x^T P x, \text{ with } P \succeq 0. \tag{5.2.1}$$

In order to have closed loop stability, it is sufficient for $F(x)$ to satisfy (A3), given an invariant set $\mathbb{X}_f$. The LTI case is considered, with $u = Kx$ and K constant. An extension for a class of non-linear systems is discussed in Section 5.7.1. The computation of Z is based on some of the concepts in [Lazar *et al.* , 2006; Raković & Lazar, 2012].

5.2.1 ℓ_1-terminal cost

Assume $P = 0$. Define $\tilde{Q} = Q^{1/2}$, $\tilde{R} = R^{1/2}$. Consider the LTI case (2.5.7). For (A3) to hold, we need:

$$\|Z(A+BK)x\|_1 - \|Zx\|_1 + \|\tilde{Q}x\|_2^2 + \|\tilde{R}Kx\|_2^2 + \|SKx\|_1 \leq 0, \; \forall x \in \mathbb{X}_f. \tag{5.2.2}$$

[1]Under this assumption, it can be easily shown that problem (3.2.1–3.2.2) is still an ℓ_1-regularised LS problem.

Assume $Z = \alpha\hat{Z}$, where $\alpha > 0$ and that $\hat{Z}$ has full column rank, with left pseudo-inverse $\hat{Z}^L = (\hat{Z}^T\hat{Z})^{-1}\hat{Z}^T$. Define:

$$\beta = 1 - \|\hat{Z}(A+BK)\hat{Z}^L\|_1, \quad \sigma = \|\tilde{Q}\hat{Z}^L\|_2^2 + \|\tilde{R}K\hat{Z}^L\|_2^2, \quad \zeta = \|SK\hat{Z}^L\|_1. \tag{5.2.3}$$

The following result is obtained:

Lemma 5.2.1. Consider system (2.5.7), an asymptotically stabilising K, and the associated PI set $\mathbb{X}_f$. Then, condition (5.2.2) and (A3) are satisfied by $F(x) = \|Zx\|_1$, if $\beta > 0$ and if

$$\alpha \geq \frac{1}{\beta}\left(\max_{x\in\mathbb{X}_f}\|\hat{Z}x\|_1\sigma + \zeta\right). \tag{5.2.4}$$

Proof. (Lemma 5.2.1) Assume $x \in \mathbb{X}_f$. Substituting (5.2.3) in (5.2.4), we have:

$$\begin{aligned} 1 &\geq \|\hat{Z}(A+BK)\hat{Z}^L\|_1 + \tfrac{1}{\alpha}\|SK\hat{Z}^L\|_1 + \tfrac{1}{\alpha}\max_{(x\in\mathbb{X}_f)}\|\hat{Z}x\|_1\left(\|\tilde{Q}\hat{Z}^L\|_2^2 + \|\tilde{R}K\hat{Z}^L\|_2^2\right) \\ &\geq \|\hat{Z}^L(A+BK)\hat{Z}^L\|_1 + \tfrac{1}{\alpha}\|SK\hat{Z}^L\|_1 + \tfrac{1}{\alpha}\|\hat{Z}x\|_1\left(\|\tilde{Q}\hat{Z}^L\|_2^2 + \|\tilde{R}K\hat{Z}^L\|_2^2\right). \end{aligned} \tag{5.2.5}$$

Recalling that $Z = \alpha\hat{Z}$, the last part of the inequality (5.2.5) is equal to

$$\begin{aligned} &\|Z(A+BK)Z^L\|_1 + \|SKZ^L\|_1 + \|Zx\|_1\left(\|\tilde{Q}Z^L\|_2^2 + \|\tilde{R}KZ^L\|_2^2\right) \\ &= \|Z(A+BK)Z^L\|_1 + \|SKZ^L\|_1 + \|Zx\|_1^2\left(\|\tilde{Q}Z^L\|_2^2 + \|\tilde{R}KZ^L\|_2^2\right)/\|Zx\|_1, \end{aligned} \tag{5.2.6}$$

since $\|x\|_2 \leq \|x\|_1$, this is an upper bound of

$$\|Z(A+BK)Z^L\|_1 + \|SKZ^L\|_1 + \|Zx\|_2^2\left(\|\tilde{Q}Z^L\|_2^2 + \|\tilde{R}KZ^L\|_2^2\right)/\|Zx\|_1. \tag{5.2.7}$$

From the above statements it follows that:

$$1 \geq \|Z(A+BK)Z^L\|_1 + \|SKZ^L\|_1 + \left(\|\tilde{Q}Z^L\|_2^2\|Zx\|_2^2 + \|\tilde{R}KZ^L\|_2^2\|Zx\|_2^2\right)/\|Zx\|_1. \tag{5.2.8}$$

Multiply both sides of (5.2.8) by $\|Zx\|_1$ to have:

$$\begin{aligned} \|Zx\|_1 \geq\ & \|Z(A+BK)\hat{Z}^L\|_1\|Zx\|_1 + \|SK\hat{Z}^L\|_1\|Zx\|_1 + \|\tilde{Q}\hat{Z}^L\|_2^2\|Zx\|_2^2 + \\ & \|\tilde{R}K\hat{Z}^L\|_2^2\|Zx\|_2^2 \geq \|Z(A+BK)x\|_1 + \|SKx\|_1 + \|\tilde{Q}x\|_2^2 + \|\tilde{R}Kx\|_2^2, \end{aligned} \tag{5.2.9}$$

from which condition (5.2.2) follows. ■

It should be noticed that a maximisation over a potentially large set is required by (5.2.4). This could lead to a very large α and an aggressive controller. To avoid the maximisation a more general result is provided in the coming section.

5.2.2 Mixed ℓ_1/ℓ_2^2-terminal cost

The aim of this section is to avoid the maximisation in (5.2.4). Assume that $P \succ 0$ and, again, that $Z = \alpha\hat{Z} \succ 0$, where $\hat{Z}$ has full column rank. The following result is obtained:

Lemma 5.2.2. Consider system (2.5.7) and an asymptotically stabilising K. Then there exists a PI set $\mathbb{X}_f$ for (2.5.7) in closed-loop with $u = Kx$. Moreover, (A3) is satisfied by $F(x) = x^T Px + \|Zx\|_1$ if P solves

$$(A + BK)^T P(A + BK) - P \leq -(Q + K^T RK), \tag{5.2.10}$$

and if

$$\alpha \geq \frac{\zeta}{\beta},\ \beta > 0, \tag{5.2.11}$$

where ζ and β are defined in (5.2.3).

Proof. (Lemma 5.2.2) Asymptotic stability means that $(A+BK)$ has all its eigenvalues inside the unit circle. This implies the existence of an invariant polytope, $\mathbb{X}_f = \mathcal{O}_\infty$, as discussed in Chapter 2.6.5.1.

Consider now (5.2.10). Since P solves (5.2.10), it follows that

$$x^T(A + BK)^T P(A + BK)x - x^T Px \leq -x^T(Q + K^T RK)x,\ \forall x \in \mathcal{O}_\infty. \tag{5.2.12}$$

and, consequently:

$$F((A + BK)x) - F(x) + \ell(x, Kx) \leq \|Z(A + BK)x\|_1 - \|Zx\|_1 + \|SKx\|_1,\ \forall x \in \mathcal{O}_\infty. \tag{5.2.13}$$

In order to verify $F((A + BK)x) - F(x) \leq -\ell(x, Kx)$, we want to have:

$$\|Z(A + BK)x\|_1 - \|Zx\|_1 + \|SKx\|_1 \leq 0,\ \forall x \in \mathcal{O}_\infty. \tag{5.2.14}$$

Notice that

$$\|Z(A+BK)x\|_1 - \|Zx\|_1 + \|SKx\|_1 \leq \|\hat{Z}(A+BK)\hat{Z}^L\|_1\|Zx\|_1 - \|Zx\|_1 + \frac{1}{\alpha}\|SK\hat{Z}^L\|_1\|Zx\|_1 \leq 0, \tag{5.2.15}$$

is verified if

$$\frac{1}{\alpha}\underbrace{\|SK\hat{Z}^L\|_1}_{\zeta} \leq \underbrace{1 - \|\hat{Z}(A+BK)\hat{Z}^L\|_1}_{\beta}, \tag{5.2.16}$$

from which the required inequality follows. ∎

5.2.3 Computational issues

The proposed method presents certain computational issues. The main difficulty, for LTI system, is to find a K, or a $\hat{Z}$, such that $\|\hat{Z}(A+BK)\hat{Z}^L\|_1 \leq 1$ is satisfied. In principle, if we fix $\hat{Z}$ this could be done by solving an LP for K. However, since $\hat{Z}$ has full column rank we need to achieve at best $\|A+BK\|_2 < \frac{1}{\sigma_{\min}(\hat{Z})\sigma_{\min}(\hat{Z}^L)}$ or in the worst case $\|A+BK\|_2 \leq \frac{1}{\sqrt{n_Z}\sigma_{\max}(\hat{Z})\sigma_{\max}(\hat{Z}^L)}$. If $\hat{Z}$ is orthonormal, then we still need $\|A+BK\|_1 \leq 1$. This requires all the elements of A to be moved by means of K. To do so, the matrix B must have have full row rank, n. In this case, one can then fix $\hat{Z}$ and proceed with an LP. However, if the system has less actuators than states, then the conditions for Lemma 5.2.1 and Lemma 5.2.2 cannot be satisfied for an arbitrary $\hat{Z}$. One could, for instance, try to solve a non-convex programme in K and $\hat{Z}$, with increased computation and no solution guarantees. A possibility is to have a deadbeat K, so that $\hat{Z}$ could be in the left null-space of $A+BK$. However, an excessive gain would notably reduce the size of the admissible set, potentially resulting in an even smaller invariant set, and domain of attraction.

The discussed limitations apply for instance to the most common industrial cases, where the inputs time differences, Δu, and not their magnitude are penalised. In this case, one can regularise Δu, to obtain piece-wise constant signals, and for instance to reduce actuator wear and tear. To do so the terminal cost in Lemmas 5.2.1 and 5.2.2 is not the most suitable. In fact, it is common for Δu formulations to include u into an extended state vector $z = (x, u) \in \mathbb{R}^{n+m}$, and to build an artificial system that has $\Delta u \in \mathbb{R}^m$ as inputs. For this system MPC stability can be imposed again by using a terminal cost and a terminal constraint. However, the number of elements of z is at least as large as the number of elements of Δu, hence, a non-linear optimisation is again needed to meet Lemma 5.2.1 and Lemma 5.2.2, with no solution guarantee. This is indeed a very strong limitation of the above Lemmas, that can basically be used only for systems with more actuators than states and with regularisation penalty on u.

5.3 Terminal cost using Minkowski functions

This formulation allows one to compute the stabilising ingredients for ℓ_{asso}-MPC version 2 by means of convex linear programming. The terminal cost is taken as

$$F(x) = \beta\psi^2_{\mathbb{X}_f}(x) + \alpha\psi_{\mathbb{X}_f}(x), \text{ with } \beta, \alpha > 0, \tag{5.3.1}$$

where $\psi_{\mathbb{X}_f}$ is the Minkowski function of the terminal set $\mathbb{X}_f$ [Blanchini & Miani, 2008] and $\psi^2_{\mathbb{X}_f}$ is the function squared. In particular, the scalars $\beta, \alpha > 0$ and the set $\mathbb{X}_f$ will be computed to satisfy (A3). For this purpose, $\mathbb{X}_f$ will be a λ-contractive set.

Terminal costs using Minkowski functions have been considered, for instance, in [Raković & Lazar, 2012] for MPCs with 1 and ∞-norm costs and in [Grammatico & Pannocchia, 2013] for LQ-MPC. In this thesis, a combination of $\psi_{\mathbb{X}_f}$ and its $\psi^2_{\mathbb{X}_f}$ is used for guaranteeing ℓ_{asso}-MPC stability.

5.3.1 Minkowski functions

If the constraints are polytopic C-sets, then $\mathbb{X}_f$ can also be taken as a polytope C-set, with the equivalent irreducible representations [Blanchini & Miani, 2008]:

$$\mathbb{X}_f = \{x \mid Gx \leq \underline{1}\}, \tag{5.3.2}$$

$$\mathbb{X}_f = \{x = Xz \mid z \geq 0,\ \underline{1}^T z \leq 1\}, \tag{5.3.3}$$

where G has full column rank, and the columns of X are the vertices of $\mathbb{X}_f$. Define G_i as the i-th row of G. The Minkowski function of $\mathbb{X}_f$ takes the form [Blanchini & Miani, 2008]:

$$\psi_{\mathbb{X}_f}(x) = \max_i\{G_i x\} = \min_z\{\underline{1}^T z : x = Xz,\ z \geq 0\}. \tag{5.3.4}$$

Remark 25. A Minkowski function $\psi_{\mathbb{X}_f}(x)$ is a *semi-norm* [Horn & Johnson, 2010], and it enjoys several properties. In particular, it is continuous, convex, positive definite, homogeneous of order 1 (namely $\psi(cx) = c\psi(x)$ where c is a scalar) and sub-additive (the triangle inequality holds) [Blanchini & Miani, 2008]. Since we assume that $0 \in \text{int}(\mathbb{X}_f)$, then we also have that $\psi_{\mathbb{X}_f}(x) =$

0 iff $x = 0$, and then $\psi_{\mathbb{X}_f}(x)$ is a *vector norm* [Horn & Johnson, 2010]. Moreover, (5.3.4) is the solution of an LP therefore $\psi_{\mathbb{X}_f}(x)$ can be represented by a PWA function of x [Alessio & Bemporad, 2009; Spjtvold *et al.* , 2007].

In the case of $\mathbb{X}_f$ symmetric with respect to the origin we have, for different G and X:

$$\mathbb{X}_f = \{x \mid \|Gx\|_\infty \leq 1\}, \tag{5.3.5}$$

$$\mathbb{X}_f = \{x = Xz \mid \|z\|_1 \leq 1\}, \tag{5.3.6}$$

$$\psi_{\mathbb{X}_f}(x) = \|Gx\|_\infty = \min_z\{\|z\|_1 : x = Xz\}. \tag{5.3.7}$$

5.3.2 Computation of the terminal cost

Assume $\mathbb{X} = \{x \mid Lx \leq \underline{1}\}$, and $\mathbb{U} = \{u \mid Eu \leq \underline{1}\}$. Consider a polytopic C-set, $\mathcal{S} \subseteq \mathbb{X}$, with vertices $V = \mathcal{V}(\mathcal{S})$. Then the feedback gain K can be taken, for instance, as the solution of the following LP (similar to [Blanchini & Miani, 2008], Ch. 7.3):

$$\begin{array}{rll} (KY) = & \arg\min_{\bar{K},\bar{Y}} & \|\bar{K}\|_1 \\ & \text{s.t.:} & (A + B\bar{K})V = V\bar{Y}, \\ & & \underline{1}^T\bar{Y} \leq \lambda\underline{1}^T, \\ & & E\bar{K}V \leq [\underline{1}, \ \ldots, \ \underline{1}], \ \lambda \in (0, 1). \end{array} \tag{5.3.8}$$

The, following is instrumental to compute the terminal cost scalings:

Lemma 5.3.1. If problem (5.3.8) is solvable then:

1. The set $\mathcal{S} = \mathbf{conv}(V)$ is λ-contractive for (2.5.7) and $\psi_{\mathbb{X}_f}(x)$ is a CLF in $\mathbb{X}_f = \mathcal{S}$.

2. For any p-norm, $\exists\ \underline{\mathrm{a}}_p,\ \bar{\mathrm{a}}_p,\ \underline{\mathrm{b}}_p,\ \bar{\mathrm{b}}_p > 0$ such that, $\forall x \in \mathbb{X}_f$:

$$\underline{\mathrm{a}}_p\|x\|_p \leq \psi_{\mathbb{X}_f}(x) \leq \bar{\mathrm{a}}_p\|x\|_p, \tag{5.3.9}$$

$$\psi_{\mathbb{X}_f}(x^+) - \psi_{\mathbb{X}_f}(x) \leq (\lambda - 1)\underline{\mathrm{a}}_p\|x\|_p, \tag{5.3.10}$$

$$\underline{\mathrm{b}}_p\|x\|_p^2 \leq \psi^2_{\mathbb{X}_f}(x) \leq \bar{\mathrm{b}}_p\|x\|_p^2, \tag{5.3.11}$$

$$\psi^2_{\mathbb{X}_f}(x^+) - \psi^2_{\mathbb{X}_f}(x) \leq (\lambda^2 - 1)\underline{\mathrm{b}}_p\|x\|_p^2, \tag{5.3.12}$$

   for $x^+ = (A + BK)x$.

Proof. (Lemma 5.3.1) In the same order:

1. The result is a direct application of Proposition 7.34, page 254 of [Blanchini & Miani, 2008].

2. The norm identities are obtained as a consequence of $\mathcal{S}$ being λ-contractive. First, (5.3.10) and (5.3.12) are a consequence of contractivity. Then, the existence of the lower bounds in (5.3.9) and (5.3.11) follows from the fact, in Remark 25, that $\psi_{\mathcal{S}}(x) = 0$ iff $x = 0$, hence $\psi_{\mathcal{S}}(x)$ is a vector norm, and it can be upper and lower bounded by other norms and $\mathcal{K}_\infty$-functions, for instance with inequalities similar to the ones in Lemma 2.3.10. Since $\psi_{\mathcal{S}}(x) \le \|Gx\|_\infty$ and, from Theorem 2.5.7, $\|x\|_\infty \le \|x\|_p$ for any p, then the upper bounds in (5.3.9) and (5.3.11) can be taken as, $\bar{a}_p = \|G\|_\infty$ and $\bar{b}_p = \|G\|_\infty^2$, for any p-norm.

■

Contractivity of $\mathcal{S}$ is used to satisfy (A3), thus providing closed loop stability of ℓ_{asso}-MPC version 2. Assume that Lemma 5.3.1 holds. Then, it is possible to take $\mathbb{X}_f = \mathcal{S}$ and the stabilising terminal cost can be computed as follows:

Lemma 5.3.2. If Lemma 5.3.1 holds, then (A3) is satisfied for system (2.5.7) in closed loop with $u = Kx$, if $\mathbb{X}_f = \mathcal{S}$ and if $F(x)$ is given by (5.3.1) with

$$\beta \ge \frac{\lambda_{\max}(Q) + \lambda_{\max}(K^T RK)}{\underline{b}_2(1-\lambda^2)}, \quad \alpha \ge \frac{\|SK\|_1}{\underline{a}_1(1-\lambda)}. \tag{5.3.13}$$

Proof. (Lemma 5.3.2) Assume that Lemma 5.3.1 holds. Then, $\mathbb{X}_f = \mathcal{S}$ is λ-contractive and also control invariant. Take β as in (5.3.13). It follows that

$$0 \ge \beta\underline{b}_2(\lambda^2 - 1)\|x\|_2^2 + (\lambda_{\max}(Q) + \lambda_{\max}(K^T RK))\|x\|_2^2. \tag{5.3.14}$$

However, from Lemma 5.3.1 part 2, we have that $\forall x \in \mathbb{X}_f$

$$\beta\underline{b}_2(\lambda^2 - 1)\|x\|_2^2 \ge \beta(\lambda^2 - 1)\psi^2_{\mathbb{X}_f}(x) \ge \beta(\psi^2_{\mathbb{X}_f}(x^+) - \psi^2_{\mathbb{X}_f}(x)). \tag{5.3.15}$$

At the same time we have:

$$(\lambda_{\max}(Q) + \lambda_{\max}(K^T RK))\|x\|_2^2 \ge x^T(Q + K^T RK)x. \tag{5.3.16}$$

Combine (5.3.15) and (5.3.16) to have, $\forall x \in \mathbb{X}_f$, that

$$\beta(\psi^2_{\mathbb{X}_f}(x^+) - \psi^2_{\mathbb{X}_f}(x)) + x^T(Q + K^T RK)x \le 0. \tag{5.3.17}$$

From (5.3.13) we have that

$$\alpha\underline{a}_1(\lambda-1)+\|SK\|_1\leq 0. \tag{5.3.18}$$

From Lemma 5.3.1 we have, $\forall x\in\mathbb{X}_f$, that

$$\begin{aligned} 0 &\geq (\alpha\underline{a}_1(\lambda-1)+\|SK\|_1)\|x\|_1 \geq \alpha(\lambda-1)\psi_{\mathbb{X}_f}(x)+\|SKx\|_1, && (5.3.19)\\ &\geq \alpha(\psi_{\mathbb{X}_f}(x^+)-\psi_{\mathbb{X}_f}(x))+\|SKx\|_1. && (5.3.20)\end{aligned}$$

Add the terms of (5.3.17) and (5.3.20) to obtain the inequality

$$\beta(\psi^2_{\mathbb{X}_f}(x^+)-\psi^2_{\mathbb{X}_f}(x))+\alpha(\psi_{\mathbb{X}_f}(x^+)-\psi_{\mathbb{X}_f}(x))\leq -x^T(Q+K^TRK)x-\|SKx\|_1,\ \forall x\in\mathbb{X}_f, \tag{5.3.21}$$

required by (A3). ∎

Remark 26. Since, $\psi_{\mathbb{S}}$ is continuous, convex, positive definite and PWA in x, with $\psi_{\mathbb{S}}(0)=0$, it also follows that $\psi^2_{\mathbb{S}}(x)$ is continuous, convex and PWQ, with $\psi^2_{\mathbb{S}}(0)=0$. Then, the quantities $\underline{a}_p$ and $\underline{b}_p$ can be computed, for instance, from the explicit solution of one of the two mpLPs on the right-hand side of (5.3.4).

5.3.2.1 Symmetric constraints

In the particular case of $\mathbb{X}$ and $\mathbb{U}$ symmetric with respect to the origin (0-symmetric) then, any invariant $\mathbb{S}$ is also 0-symmetric. Then, the Minkowski function is given by (5.3.7) can be computed directly, as well as some (conservative) bounds for Lemma 5.3.1. For instance, for $p=\infty$ and $p=2$, we can take

$$\underline{a}_p=\frac{1}{\sqrt{n_g}}\sigma_{\min}(G),\ \underline{b}_p=\frac{1}{n_g}\sigma_{\min}(G)^2, \tag{5.3.22}$$

where n_g is the number of rows of G, namely the number of inequalities for $\mathbb{S}$. This follows from the fact that $(\|Gx\|_\infty)^2\geq\frac{1}{n_g}\|Gx\|_2^2\geq\frac{1}{n_g}\|x\|_2^2/\|(G^TG)^{-1}\|_2$ (note that G has full column rank), and that $\|x\|_2\geq\|x\|_\infty$. Notice that, if the number of inequalities in n_g is large, then taking $\underline{a}_p$ and $\underline{b}_p$ as in (5.3.22) could give quite large values of α and β. However, if the set is large, then $\sigma_{\min}(G)$ can be expected to be quite small. Alternatively, one can compute the scalings as follows:

Lemma 5.3.3. Assume $\mathbb{S}$ is symmetric. Then, in Lemma 5.3.2, it is possible to take

$$\beta \geq \frac{n\left(\|Q^{1/2}G^L\|_2^2 + \|(G^L)^T K^T R K G^L\|_2\right)}{1-\lambda^2}, \quad \alpha \geq \frac{n_s\|SKG^L\|_\infty}{1-\lambda}. \tag{5.3.23}$$

Proof. (Lemma 5.3.3) The results are obtained through similar steps as taken for Lemma 5.3.2, with the difference that here the stage cost elements are upper bounded by ∞-norms, and $\psi_{\mathbb{S}}$ is left as it is. The left pseudo-inverse of G, G^L, is used to collect $\|Gx\|_\infty$ and to remove x from the inequalities. ■

Remark 27. The above lower bounds could be conservative, as they hold globally, not only in $\mathbb{X}_f$. A less conservative scaling can be found by means of an LP with constraints $\underline{a}_p\|v\|_p \leq \psi_{\mathbb{X}_f}(v),\ \forall v \in \mathcal{V}(\mathbb{X}_f)$, for $p = 1, \infty$.

Remark 28. It can be shown that ℓ_{asso}-MPC version 2 can be formulated as a strictly convex QP, for instance when a Minkowski function cost is used. Thus the explicit solution is unique and PWA, by the same arguments as in Chapter 4.4. This will be further discussed in the next section.

5.4 Terminal set and domain of attraction

The previous sections provided a way to compute a stabilising terminal cost, given a λ-contractive terminal set. Differently from more classic MPC approaches (see Chapter 2.6.5.1), let's assume that $\mathbb{X}_f = \mathbb{S}$, as computed in the previous Section. As mentioned above, the set $\mathbb{S}$ does not depend on the stage cost, and it can be computed once, prior to cost function tuning. Recalling Chapter 2.6.6, it is clear that $\mathbb{X}_f$ affects, to a certain extent, the size of the the DOA, $\mathbb{X}_N$, (together with the horizon length N). In particular, from the definitions in Chapter 2.6.6 it can be noticed that the DOA of any considered stabilising MPC satisfies

$$\mathbb{X} \supseteq \mathbb{X}_N \supseteq \mathbb{X}_f. \tag{5.4.1}$$

Therefore, maximising $\mathbb{X}_f$ may be beneficial to the DOA. Since it is proposed to compute $\mathbb{X}_f$ online, then the size of $\mathbb{X}_N$ will remain unchanged, for a change in the stage cost. This is a further advantage with respect to the approaches of Chapter 2.6.5.1.

5.4.1 The considered method: control at the vertices

Consider Algorithm 1, based on Proposition 7.34, p. 254 of [Blanchini & Miani, 2008].

Algorithm 1 Terminal set computation (based on Proposition 7.34 of [Blanchini & Miani, 2008])

---

1. Choose a matrix V containing a set of candidate vertices, and $\lambda \in (0,1)$, $\epsilon \geq 0$, $\gamma_{\max} > 0$.

2. Solve the LP

$$\begin{array}{ll} \max_{\chi} & \gamma - \epsilon \|\bar{K}\|_1 \\ \text{s.t.:} & A\gamma V + B\bar{K}V = V\bar{Y}, \\ & \underline{1}^T \bar{Y} \leq \gamma \lambda \underline{1}^T, \\ & E\bar{K}V \leq [\underline{1}, \dots, \underline{1}], \\ & 0 < \gamma \leq \gamma_{\max}, \end{array} \tag{5.4.2}$$

   where $\chi = (\gamma,\ \bar{K},\ \bar{Y})$, and the entries of $\bar{Y}$ are all non-negative.

3. **IF** (5.4.2) is feasible **THEN** $K \leftarrow \frac{1}{\gamma^\star}\bar{K}^\star$ **ELSE** goto 1.

4. $\mathbb{X}_f \leftarrow \text{conv}(\gamma^\star V)$.

Algorithm 1 is an improved version of the technique proposed in [Gallieri & Maciejowski, 2013b]. In particular, Algorithm 1 requires at best the solution of a single The parameter ϵ can be used to regulate the tradeoff between the size of the terminal set, and the sparsity of K. A sparse K could be useful, for instance, for the prioritised actuators scenarios of Chapter 6. On the other hand, $\epsilon = 0$ provides the largest set for the given V and the given LPV system. Since from (5.4.1) the DOA is contained in $\mathbb{X}$, a possible choice is to take V as the vertices of $\mathbb{X}$. In this case, $\gamma^\star = 1$ provides the largest DOA.

Remark 29. If, in Algorithm 1, $V = [I,\ -I]$, then $\mathbb{X}_f = \{x \mid \|x\|_1 \leq \gamma^\star\}$.

The weakness of Algorithm 1 is the requirement of a matrix of vertices V, such that problem (5.4.2) is solvable. The search for V can be non-trivial. A possible way to proceed is to partition the state space in cones, and then take a vertex on each ray (accounting for possible symmetry). The idea of conic partitions has been considered in [Lazar, 2010], where an algorithm is given for computing ∞-norm contractive sets. Several contributions exist, that aim to solve for V. In [Cannon *et al.* , 2003], a CLF $\|Gx\|_\infty$ is computed by assuming K to be constant and G invertible. No structure is imposed on K.

5.4.2 Alternative methods

Two alternative formulations are discussed, that rely on different types of invariant sets. These are more general than Algorithm 1, however, more computation is required.

5.4.2.1 Using the maximal admissible set

A different formulation can be used if a stabilising controller K for (2.5.7) is known. In this case, one can and compute the maximal admissible set inside $\mathcal{S}$ for the closed-loop system [Blanchini & Miani, 2008; Kerrigan, 2000]. This is formalised by the following:

Lemma 5.4.1. If an asymptotically stabilising controller, K, is known for the unconstrained system (2.5.7) together with a set $\mathcal{S}$ such that $Kx \in \mathbb{U}$, $\forall x \in \mathcal{S}$ then, problem (5.3.8) has a solution for $V = \mathcal{V}(\mathcal{O}_\infty^\lambda)$, where $\mathcal{O}_\infty^\lambda$ is the maximal admissible set for the system:

$$\begin{aligned} x(k+1) &= \frac{1}{\lambda}(A+BK)x, \\ x &\in \mathcal{S},\ Kx \in \mathbb{U}. \end{aligned} \tag{5.4.3}$$

Therefore, Lemma 5.3.1 is satisfied. The set $\mathcal{O}_\infty^\lambda$ exists and it can be computed in finite time, if $\lambda \in (\lambda_{\max}(A+BK),\ 1)$.

Proof. (Lemma 5.4.1) By definition of maximal admissible set, it follows that $x(0) \in \mathcal{O}_\infty^\lambda \Rightarrow x(k) \Rightarrow \mathcal{O}_\infty^\lambda, \forall k \geq 0$ for system (5.4.3). Therefore, for the system (2.5.7) we have $x(0) \in \mathcal{O}_\infty^\lambda \Rightarrow x(k) \in \lambda^k \mathcal{O}_\infty^\lambda, \forall k \geq 0$, namely, the set is λ-contractive. The existence of the set $\mathcal{O}_\infty^\lambda$ follows from the fact that, if $\lambda \in (\lambda_{\max}(A+BK),\ 1)$, then $\frac{1}{\lambda}(A+BK)$ has all eigenvalues strictly in the unit circle (see Theorem 5.17, p. 180 of [Blanchini & Miani, 2008]). Moreover, $\mathcal{O}_\infty^\lambda$ consists of a finite number of inequalities (it is finitely determined), as in Theorem 5.17 of [Blanchini & Miani, 2008]. ■

5.4.2.2 Using a robustly control invariant set

A second alternative to Algorithm 1 has been proposed, for LQ-MPC, in [Grammatico & Pannocchia, 2013]. In particular, a λ-contractive set is computed without restricting the terminal controller structure. This is computed with the algorithm proposed in [Blanchini, 1994]. Moreover, in [Grammatico & Pannocchia, 2013] a scaling is computed for the quadratic terminal cost, taken as the square of the Minkowski function of $\mathbb{X}_f$. This has inspired the choice of terminal cost made in this thesis, which differs from the one in [Gallieri & Maciejowski, 2013a], where the Minkowski function was used together with a different quadratic terminal cost. The approach of [Grammatico & Pannocchia, 2013] can provide a the maximal λ-contractive set (up to a numerical tolerance), however, it relies on the use of polytopic projections.

5.5 Input sparsity and dead-zone

In this section, it is shown that the dead-zone capability, discussed in the previous chapter for ℓ_{asso}-MPC version 1, can be inherited by ℓ_{asso}-MPC version 2. Consider a terminal cost of the type:

$$F(x) = x^T P x + \|Zx\|_p, \text{ for } p = 1, \infty. \tag{5.5.1}$$

This choice is representative, for instance, of the case of symmetric constraints.

Define

$$\Phi = Z[A^{N-1}B, A^{N-2}B, \cdots, B]. \tag{5.5.2}$$

Similar to the previous chapter, the constrained ℓ_{asso}-MPC version 2 with polytopic constraints is based on the online solution of the QP

$$\begin{aligned} &\min_{\underline{\mathrm{u}}} \ \underline{\mathrm{u}}^T H \underline{\mathrm{u}} + 2\underline{\mathrm{u}}^T \Gamma x + \|W\underline{\mathrm{u}}\|_1 + \|ZA^N x + \Phi \underline{\mathrm{u}}\|_p \\ &\text{s.t. } \Omega\, \underline{\mathrm{u}} \leq \underline{1} - Mx, \ x = x(k). \end{aligned} \tag{5.5.3}$$

The Lagrangian of (5.5.3) is given by

$$L(\underline{\mathrm{u}}, \nu) = \underline{\mathrm{u}}^T H \underline{\mathrm{u}} + 2\underline{\mathrm{u}}^T \Gamma x - (\underline{\mathrm{u}}^T \Omega^T - \underline{1}^T)\nu + \|W\underline{\mathrm{u}}\|_1 + \|ZA^N x + \Phi \underline{\mathrm{u}}\|_p. \tag{5.5.4}$$

Then the KKT conditions for (5.5.3) are:

$$\begin{aligned} 2H\,\underline{\mathbf{u}}^\star + 2\Gamma x + \Omega^T \nu^\star + \Phi^T \mathbf{r}^\star &= -W^T \mathbf{v}^\star \\ -\Omega\,\underline{\mathbf{u}}^\star - s^\star &= -(\underline{1} - Mx) \\ (s^\star)^T \nu^\star &= 0 \\ s^\star &\geq 0 \\ \nu^\star &\geq 0 \\ \mathbf{v}^\star &\in \partial \|W\underline{\mathbf{u}}^\star\|_1 \\ \mathbf{r}^\star &\in \partial \|ZA^N x + \Phi \underline{\mathbf{u}}^\star\|_p, \\ x &= x(k). \end{aligned} \tag{5.5.5}$$

In the unconstrained case the KKT conditions (5.5.5) become

$$-2(\Gamma x + H\underline{\mathbf{u}}^\star) - \Phi^T \mathbf{r}^\star = \lambda W^T \mathbf{v}^\star, \tag{5.5.6}$$

$$\mathbf{v}^\star \in\ W^T \partial \|W\underline{\mathbf{u}}^\star\|_1, \tag{5.5.7}$$

$$\mathbf{r}^\star \in \partial \|PA^N x + \Phi \underline{\mathbf{u}}^\star\|_p, \tag{5.5.8}$$

$$x = x(k). \tag{5.5.9}$$

Theorem 5.5.1. The unconstrained ℓ_{asso}-MPC version 2, with $F(x)$ as in (5.5.1), has an implicit input thresholding capability. In particular, assume $W \in \mathcal{D}_{++}$. Then it follows that

$$2|\Gamma_i\, x(k) + \sum_{j \neq i} H_{ij} \underline{\mathbf{u}}_j^\star| + \sum_j |\Phi_{ji}|. < W_{ii} \Rightarrow \underline{\mathbf{u}}_i^\star = 0. \tag{5.5.10}$$

Proof. Given a particular value of x and the corresponding solution $\underline{\mathbf{u}}^\star$, taking W_{ii} from the first part of (5.5.10) gives:

$$\begin{aligned} W_{ii} &> 2|\Gamma_i\, x(k) + \sum_{j \neq i} H_{ij} \underline{\mathbf{u}}_j^\star| + \max_{\mathbf{r}, \|\mathbf{r}\|_\infty \leq 1} |(\Phi)_i^T \mathbf{r}| \\ &\geq \max_{\mathbf{r}, \|\mathbf{r}\|_\infty \leq 1} \{|2\Gamma_i\, x(k) + 2\sum_{j \neq i} H_{ij} \underline{\mathbf{u}}_j^\star + (\Phi)_i^T \mathbf{r}|\} \\ &\geq |2\Gamma_i\, x(k) + 2\sum_{j \neq i} H_{ij} \underline{\mathbf{u}}_j^\star + (\Phi)_i^T \mathbf{r}^\star|, \end{aligned} \tag{5.5.11}$$

where we have used the fact that (for $p = 1, \infty$) $\mathbf{r}^T v = \|v\|_p$ by Definition 2.3.9, from which

$\|\mathbf{r}^\star\|_\infty \leq 1$. Given $W \in \mathcal{D}_{++}$, (5.5.11) provides the KKT condition (5.5.6) to be satisfied with $\mathbf{v}_i \in (-1,\ 1)$, implying that $W_{ii}\underline{\mathbf{u}}_i^\star = 0$ hence $\underline{\mathbf{u}}_i^\star = 0$. ■

Noticeably, condition (5.5.10) has an additional term on the left hand side, with respect to (4.2.14). From Theorem 5.5.1 it is clear that using ℓ_{asso}-MPC version 2 is more difficult to have some inputs at zero for most of the time. In particular, assume $S = \eta I$, for some $\eta > 0$. From (4.2.2) and (5.7.20) we have $W_{ii} = \eta,\ \forall i$ and $\Phi = \eta \bar{Z}[A^{N-1}B,\ \ldots,\ B]$, for some $\bar{Z}$ with full column rank. This has direct influence on sparsity, as the penalty in (5.5.3) is smaller for the channels with $\|A^{N-1}B_{\bullet i}\|_1$ small. In other words thresholding is facilitated for actuators that are less relevant to stabilisation. This is illustrated in the next example.

5.6 Example: Maximum DOA and enhanced optimality

Consider again the LTI system with matrices

$$\begin{aligned} A &= \begin{bmatrix} 0.15 & 0.1 \\ 0 & 1.1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \\ Q &= \begin{bmatrix} 20 & 0 \\ 0 & 60 \end{bmatrix}, \quad R = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, \end{aligned} \tag{5.6.1}$$

with state and input magnitude constraints $\|x\|_\infty \leq 30$, $\|u\|_\infty \leq 5$, $N = 3$, and with $S = \eta \cdot I$. The proposed strategies are compared, in terms of the following criteria:

- The average closed-loop cost.
- The size of the DOA.
- The control error.
- The input signals.

5.6.1 The DOA / optimality tradeoff

Consider, as a benchmark, the solution of an Infinite Horizon Optimal Control Problem (IHOCP) with stage cost (3.2.2). This ideal problem has clearly the maximal DOA and is optimal with respect to the stage cost $\ell(x, u)$ both in open-loop and closed-loop. When using a finite horizon

MPC we would also like to have a large DOA, as well as good closed-loop performance with respect to $\ell(x,u)$. These can both be improved, for instance, by increasing the horizon length N, when computationally acceptable. Consider, however, a fixed value of N. Since we wish the MPC solution to be optimal with respect to $\ell(x,u)$, we would like the terminal cost $F(x)$ to not have a strong influence on the total MPC cost, while guaranteeing stability. For instance, we would like the scaling α, β to be not too large. On the other hand, we wish to have a relatively large terminal set, $\mathbb{X}_f$, so that the DOA, $\mathbb{X}_N$, could be large even for a relatively short N. Recall Problem (5.4.2) and Lemma 5.3.3, to observe the existence of a tradeoff between the values of α, β and the maximisation of the size of $\mathbb{X}_f$. In particular, from (5.4.2), the smallest values of α, β, are obtained from the smallest γ.

Regulating the terminal cost/constraint tradeoff could be, in principle, a way to reduce sub-optimality with respect to an ideal infinite horizon problem. Moreover, numerical difficulties could occur if α is quite large or $\mathbb{X}_f$ quite small (for example, when an interior-point method is used [Kim *et al.* , 2007b]). In order to develop a measure of the DOA volume for the coming example, consider the approximation of the volume of a polytope, $\mathcal{S} = \{x \mid Lx \leq \underline{1}\}$, by means of the associated ellipsoid $\varepsilon_{\mathcal{S}} = \{x \mid x^T \Sigma_L x \leq 1\}$, where $\Sigma_L = \frac{1}{q} L^T L$, and where q is the number of rows of L. Define $\tilde{L} = \Sigma_L^{1/2}$. Then, the volume of $\varepsilon_{\mathcal{S}}$ is proportional to $\mathbf{d}_{\mathcal{S}} = \det(\tilde{L}^{-1})$. Hence, it is possible to select the γ providing the best $\mathbf{d}_{\mathbb{X}_N} \mid \alpha$ tradeoff. This requires the computation of $\mathbb{X}_N$, that can be done using polytope projections. However, if one wishes not to compute $\mathbb{X}_N$, the value of γ itself can be used, to compare different choices of $\mathbb{X}_f$. It must be pointed out that this might not be the best way to approximate the size of the terminal set, and it is only used for illustrative purpose.

5.6.2 Simulation results

The following cases are investigated:

1. `con-sets`: Similar to [Limon *et al.* , 2005]. A sequence of 4 terminal sets is used. These are the i-step controllable sets to the origin for $i = 4, 3, 2, 1$. The sequence is then terminated by the origin. At each time step the next element is selected.
2. `1-norm`: ℓ_1-terminal cost. K places all eigenvalues at 0.001, $\mathbb{X}_f = \mathcal{O}_\infty$.
3. `1-norm` (γ): ℓ_1-terminal cost. $\gamma = 1$, $\mathbb{X}_f = \mathbb{X} = \mathbb{X}_N$.
4. `1-norm` $(\alpha|\gamma)$: ℓ_1-terminal cost, $\gamma = 0.5$, $\mathbb{X}_f = \mathcal{O}_\infty$.
5. `mixed`: mixed terminal cost. K places all eigenvalues at 0.001, $\mathbb{X}_f = \mathcal{O}_\infty$.
6. `mixed` (γ): mixed terminal cost. $\gamma = 1$, $\mathbb{X}_f = \mathbb{X} = \mathbb{X}_N$.

7. `mixed` $(\alpha|\gamma)$: mixed terminal cost, $\gamma = 0.5$, $\mathbb{X}_f = \mathcal{O}_\infty$.
8. `Mink-fun` (γ): Minkowski terminal cost (∞-norm), using Lemma 5.3.3 and Algorithm 1, with $\epsilon = 1$, $\gamma = 1$, $\mathbb{X}_f = \mathbb{X} = \mathbb{X}_N$.
9. `inf-hor`: Open loop optimal control with $N = 200$.

All terminal controllers are computed with Algorithm 1, with the additional constraint (5.2.11) for cases 2,4,5,6. The resulting DOAs and state trajectories are shown in Figures 5.1–5.4 (`con-sets`, `1-norm`, `1-norm` (γ) and `mixed` $(\alpha|\gamma)$), for $\eta = 300$, where the x (y) axis represents the first (second) state. In Figure 5.3–5.4, it can be noticed that the proposed strategies achieve the maximum DOA.

Table 5.1 shows the results for $\eta = 300$, where MSE is the mean squared control error, and $\mathrm{E}[V_N^o] = \frac{1}{200}\sum_{i=0}^{200} \ell(x(i), u(i))$ is the expected closed-loop cost, both averaged over 50 simulations of the duration of 200 samples from different initial conditions. For all of the considered cases, the Mean Absolute Input (MAI) value is approximately 0.047, and 100% of control comes from input 2.

Table 5.1: Control results per ℓ_{asso}-MPC strategy ($\lambda = 300$)

Strategy	MSE	β	α	γ	$\mathbf{d}_{\mathbb{X}_N}$	$\mathrm{E}[V_N^o]$
`con-sets`	1.767	62	-	-	1460	184.68
`1-norm`	1.772	-	2710	-	951	184.64
`1-norm` (γ)	1.772	-	10^5	1	1800	184.64
`1-norm` $(\alpha\|\gamma)$	1.772	-	19260	0.5	1800	184.64
`mixed`	1.772	60	630	-	951	184.64
`mixed` (γ)	1.772	915	1500	1	1800	184.64
`mixed` $(\alpha\|\gamma)$	1.772	277	857	0.5	1800	184.64
`Mink-fun` (γ)	1.772	$1.2\ 10^5$	$6.1\ 10^4$	1	1800	184.64
`inf-hor`	1.769	-	-	-	1800	184.62

Figures 5.5–5.7 compare the I/O behaviour for three of the approaches, with $x(0) = [15,\ 15]^T$, with $\eta = 300$ and $\eta = 5$. The different strategies show very similar input/output behaviour. In particular, when $\eta = 5$ both states are rapidly stabilised to the origin by means of both actuators. On the other hand, for $\eta = 300$ only one actuator is used. This is exactly the sort of control allocation for which ℓ_{asso}-MPC is recommended.

From Table 5.1 it can be noticed that all the proposed approaches are "more optimal" than `con-sets`, with respect to $\mathrm{E}[V_N^o]$. On the other hand, the proposed strategies have all similar

$\mathrm{E}[V_N^o]$, which is slightly higher than for `inf-hor`. It can also be noticed that `con-sets` is slightly more aggressive, having a smaller MSE than `inf-hor`, and a smoother input 2 behaviour than all the others. `Mink-fun` (γ) behaves similarly to the other strategies, with I/O signals quite similar to Figure 5.7. However, the values of α and β obtained with `Mink-fun` (γ) are conservative with respect to `mixed` (γ), where $\beta = \|P\|_2$ and the quadratic part of $F(x)$ requires the solution of a Lyapunov equation or of a DARE. Conservativeness originates mainly from the use of a global lower bound for the Minkowski function.

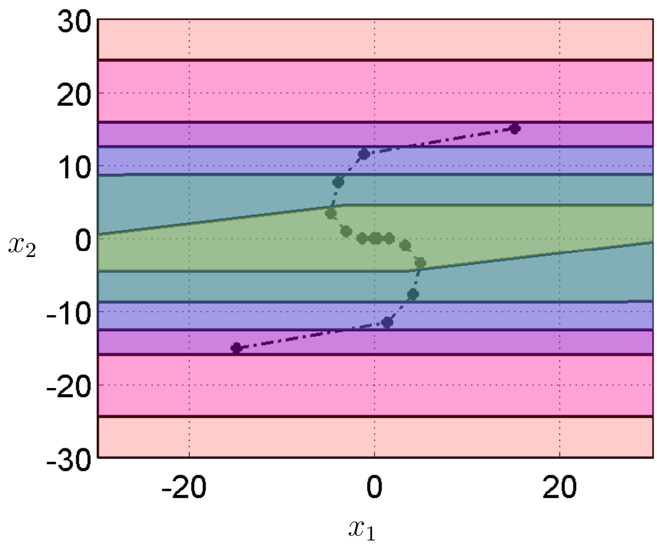

Figure 5.1: `con-sets`: 4 invariant sets, $\mathbb{X}_N$, $\mathbb{X}$, and trajectories ($\eta = 300$).

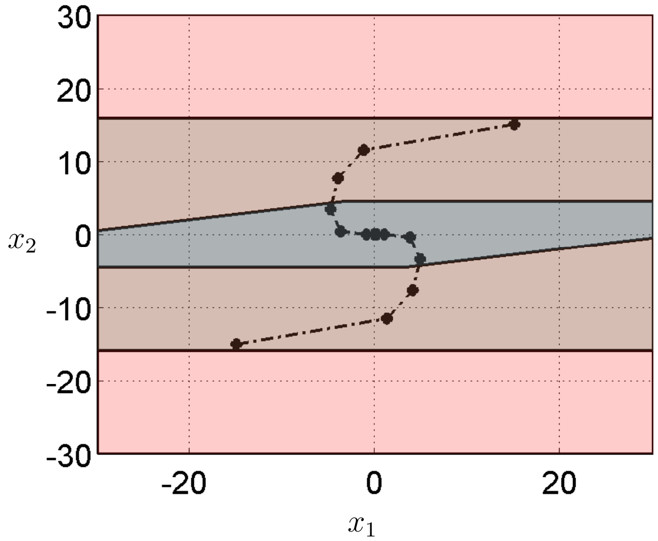

Figure 5.2: `1-norm`: Sets, $\mathcal{O}_\infty$, $\mathbb{X}_N$, $\mathbb{X}$, and trajectories ($\eta = 300$).

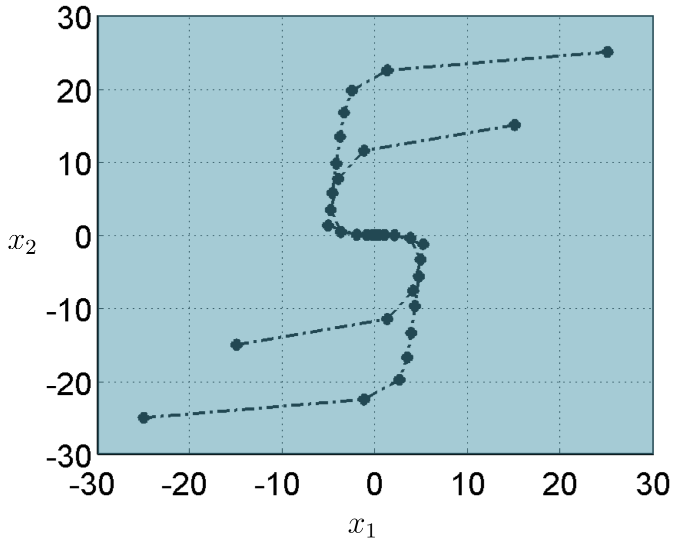

Figure 5.3: `1-norm` (γ): Sets, $\mathcal{O}_\infty \equiv \mathbb{X}_N \equiv \mathbb{X}$, and trajectories ($\eta = 300$).

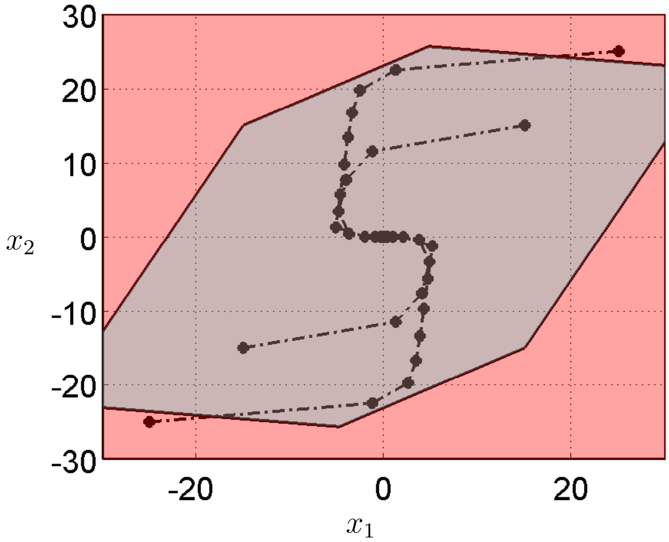

Figure 5.4: `mixed` $(\alpha|\gamma)$: Sets, $\mathcal{O}_\infty$, $\mathbb{X}_N \equiv \mathbb{X}$, and trajectories ($\eta = 300$).

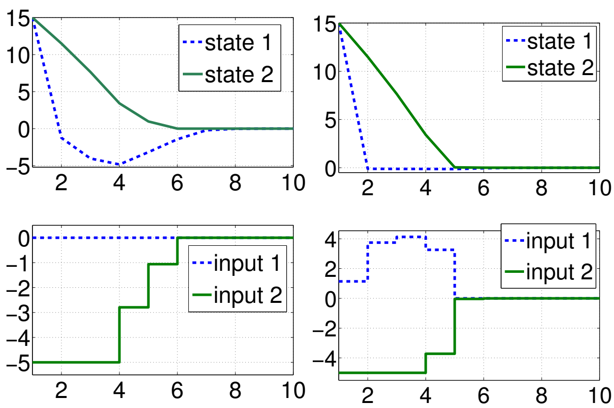

Figure 5.5: `con-sets`: I/O trajectories for $\eta = 300$ (left), $\eta = 5$ (right).

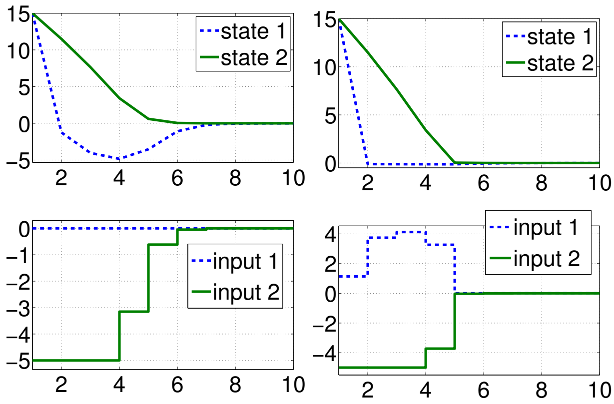

Figure 5.6: `1-norm`: I/O trajectories for $\eta = 300$ (left), $\eta = 5$ (right).

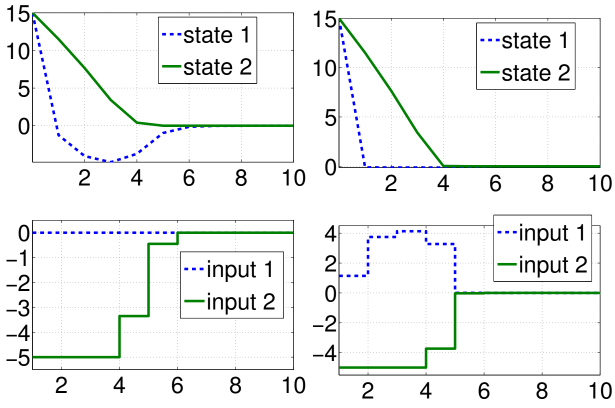

Figure 5.7: `mixed` $(\alpha|\gamma)$: I/O trajectories for $\eta = 300$ (left), $\eta = 5$ (right).

5.7 Terminal ingredients for non-linear systems

The results of the previous sections are now extended to the considered class of non-linear systems with differentiable vector field and the origin as equilibrium. The proposed solutions make use of LPV embedding (similar to Chapter 2.6.5.2). In particular, the candidate terminal controller is taken as $u = K(r)x$, where r is the scheduling parameter for the LPV embedding system (2.6.14), with $r \in \mathbb{R}^{n_r}_{\geq 0}$, $\Sigma r_i = 1$. This comports an increase in computational complexity from the LTI case, which is discussed throughout the section.

5.7.1 Mixed ℓ_1/ℓ_2^2-terminal cost

Assume that $P \succ 0$ and, again, that $Z = \alpha \hat{Z} \succ 0$, where $\hat{Z}$ has full column rank. Consider system (2.5.6). The following result extend the ones of Section 5.2 by means of local LPV model absorption and of the principle of control at the vertices, as discussed in Chapter 2.6.5.2. For convenience, condition (2.6.17) is restated here:

$$(A_i + B_i K_i)^T P (A_i + B_i K_i) - P \leq -Q - K_i^T R K_i, \ \forall i \in \mathbb{I}_{[0,n_r]}.$$

Lemma 5.7.1. Assume the existence of a C-set $\mathbb{S} \subseteq \mathbb{X}$ such that Lemma 2.6.6 holds. Then, $\forall x(0) \in \mathbb{S}$, $\forall u(k) \in \mathbb{U}$, the trajectory of (2.5.6) is in the convex hull of the ones of the LPV system (2.6.14) for all r. Assume also that P satisfies (2.6.17), for an appropriate set of matrices $K_i,\ i \in \mathbb{I}_{[1,n_r]}$. If for some $\alpha > 0$

$$\max_i \alpha(\|\hat{Z}(A_i + B_i K_i)\hat{Z}^L\|_1 - 1) + \|SK_i\hat{Z}^L\|_1 \leq 0, \tag{5.7.1}$$

then there exists a set $\mathbb{X}_f \subseteq \mathbb{S}$ that is PI for (2.6.14), as well as for (2.5.6), in closed-loop with the controller $\kappa(x, r) = \Sigma_i^{n_r} K_i r_i x$. Moreover, (A3) is satisfied for (2.5.6).

Proof. (Lemma 5.7.1) The quadratic part of the cost is addressed first. We want to obtain that, $\forall r \in \mathbb{I}_{\geq 0}, \Sigma_i r_i = 1$

$$(A(r) + B(r)K(r))^T P (A(r) + B(r)K(r)) - P + Q + K(r)^T R K(r) \preceq 0. \tag{5.7.2}$$

Similar to [Blanchini & Miani, 2008; Boyd *et al.* , 1994], define $G = P^{-1}$, $Y(r) = K(r)G$, and pre and post-multiply (5.7.2) by G, to have:

$$G \succeq 0,\ G - (GA^T(r) + Y^T(r)B^T(r))^T G^{-1}(A(r)G + B(r)Y(r)) - GQG - Y(r)^T RY(r) \succeq 0. \tag{5.7.3}$$

Similar to [Boyd *et al.* , 1994; Kothare *et al.* , 1996; Lazar *et al.* , 2006], by Schur complementing, condition (5.7.3) is equivalent to the following LMI:

$$M(p) = \begin{bmatrix} G & GA^T(r) + Y^T(r)B^T(r) & (R^{1/2}Y(r))^T & (Q^{1/2}G)^T \\ A(r)G + B(r)Y(r) & G & 0 & 0 \\ R^{1/2}Y(r) & 0 & I & 0 \\ Q^{1/2}G & 0 & 0 & I \end{bmatrix} \succeq 0. \tag{5.7.4}$$

Since $A(r) = \Sigma_i^{n_r} A_i r_i$, and similarly for $B(r),\ K(r)$, then $M(r) = \Sigma_i^{n_r} M_i r_i$, where M_i is $M(r)$ evaluated at $r_i = 1,\ r_j = 0\ \forall j \neq i$. Since $r_i \geq 0\ \forall i$, then $M(r) \succeq 0 \Leftrightarrow M_i \geq 0, \forall i$. Finally, again by Schur complementing, the condition on all M_i translates into the necessary and suffient (2.6.17). Since $x^T P x$ is a Lyapunov function, $\mathbb{X}_f$ can be taken, for instance, as one of its admissible level sets, inside $\mathcal{S}$.

For the 1-norm part, we want, for all r and for some $\alpha > 0$, to have:

$$\max_r \alpha(\|\hat{Z}(A(r) + B(r)K(r))\hat{Z}^L\|_1 - 1) + \|SK(r)\hat{Z}^L\|_1 \leq 0. \tag{5.7.5}$$

For a fixed $\alpha > 0$, the inequality (5.7.5) is satisfied if and only if it is satisfied at each vertex $\nu_i = (A_i, B_i, K_i)$. Therefore, condition (5.7.1) is necessary and sufficient for (5.7.5) and, together with (2.6.17), it provides a sufficient condition for (A3) to be satisfied for system (2.6.14). Since, by Lemma 2.6.6, the convex hull of the trajectories of (2.6.14) for all r contains the one of (2.5.6), $\forall x \in \mathcal{S} \supseteq \mathbb{X}_f$, $\forall u \in \mathbb{U}$, then by Theorem 2.6.7 it folows that $\mathbb{X}_f$ is positively invariant for (2.5.6) in closed-loop with $u = \kappa(x, r)$. Therefore (A3) is satisfied. ■

Remark 30. In Lemma 5.7.1, it is also possible to require (2.5.6) to be embedded in (2.6.14) only when in closed loop with $u(k) = \kappa(x(k), r(k))$, for some $r(k)$, instead of for all $u(k) \in \mathbb{U}$.

The computational complexity of this approach, discussed in Section 5.2.3 for the LTI case, is further increased in the non-linear case by the requirement of solving a set of LMIs. These might in fact not be solvable for the candidate $\mathcal{S}$. These limitations are overcome in the next section, once again by means of Minkowski functions.

5.7.2 Minkowski function terminal cost

Theoretical results from Section 5.3 are now extended for the considered class of non-linear systems, together with a review of existing techniques for computing the terminal set.

Consider a polytopic C-set, $\mathcal{S} \subseteq \mathbb{X}$, with vertices $V = \mathcal{V}(\mathcal{S})$. Then, for the i-th vertex of system (2.6.14), the feedback gains K_i can be taken, for instance, as the solution of the following LP (similar to [Blanchini & Miani, 2008], Ch. 7.3):

$$\begin{array}{lll} (K_i, Y_i) = & \arg\min_{\bar{K}_i, \bar{Y}_i} & \|\bar{K}_i\|_1 \\ & \text{s.t.:} & (A_i + B_i \bar{K}_i)V = V\bar{Y}_i, \\ & & \underline{1}^T \bar{Y}_i \leq \lambda \underline{1}^T, \\ & & E\bar{K}_i V \leq [\underline{1}, \ \ldots, \ \underline{1}], \ \lambda \in (0,1). \end{array} \tag{5.7.6}$$

The following is instrumental to compute the terminal cost scalings:

Lemma 5.7.2. If problem (5.7.6) is solvable for all $i \in \mathbb{I}_{[1,n_r]}$, then:

1. The set $\mathcal{S} = \mathbf{conv}(V)$ is λ-contractive for (2.6.14), for all $r \in \mathbb{R}^{n_r}_{\geq 0}$, $\underline{1}^T r = 1$.

2. If Lemma 2.6.6 holds in $\mathcal{S}$, then $\mathcal{S}$ is λ-contractive for (2.5.6), and $\psi_{\mathbb{X}_f}(x)$ is a CLF in $\mathbb{X}_f = \mathcal{S}$.

3. For any p-norm, $\exists\ \underline{\mathbf{a}}_p,\ \bar{\mathbf{a}}_p,\ \underline{\mathbf{b}}_p,\ \bar{\mathbf{b}}_p > 0$ such that, $\forall x \in \mathbb{X}_f$:

$$\underline{\mathbf{a}}_p \|x\|_p \leq \psi_{\mathbb{X}_f}(x) \leq \bar{\mathbf{a}}_p \|x\|_p, \tag{5.7.7}$$

$$\psi_{\mathbb{X}_f}(x^+) - \psi_{\mathbb{X}_f}(x) \leq (\lambda - 1)\underline{\mathbf{a}}_p \|x\|_p, \tag{5.7.8}$$

$$\underline{\mathbf{b}}_p \|x\|_p^2 \leq \psi^2_{\mathbb{X}_f}(x) \leq \bar{\mathbf{b}}_p \|x\|_p^2, \tag{5.7.9}$$

$$\psi^2_{\mathbb{X}_f}(x^+) - \psi^2_{\mathbb{X}_f}(x) \leq (\lambda^2 - 1)\underline{\mathbf{b}}_p \|x\|_p^2, \tag{5.7.10}$$

   for $x^+ = (A(r) + B(r)K(r))x$ and, if part 2) holds, then also for $x^+ = \bar{f}(x, K(r)x)$.

Proof. (Lemma 5.7.2) In the same order:

1. The result is a direct application of Proposition 7.34, page 254 of [Blanchini & Miani, 2008].

2. If Lemma 2.6.6 holds, then $\forall x \in \mathcal{S}$, the trajectory of (2.5.6) is contained in the set of trajectories of (2.6.14). Then Theorem 2.6.7 also holds from which there exist an appropriate time-varying $r(k)$ so that $\mathcal{S}$ is also λ-contractive for (2.6.14) in closed-loop with $u(k) = K(r(k))x(k)$. Finally, by Theorem 2.5.7 the Minkowski function of $\mathcal{S}$ is a CLF for (2.5.6).

3. The statement follows from the the same arguments as for part 2 of Lemma 5.3.1.

■

Refer to the set $\mathcal{S}$, in Lemma 5.7.2, as the LPV *embedding set*. Contractivity of $\mathcal{S}$ is used to satisfy (A3), thus providing closed loop stability of ℓ_{asso}-MPC version 2. Assume that Lemma 5.7.2 holds. Then, it is possible to take $\mathbb{X}_f = \mathcal{S}$ and the stabilising terminal cost can be computed as follows:

Lemma 5.7.3. If Lemma 5.7.2 holds, then (A3) is satisfied for system (2.5.6) in closed loop with $u = K(r)x$, if $\mathbb{X}_f = \mathcal{S}$ and if $F(x)$ is given by (5.3.1) with

$$\beta \geq \frac{\lambda_{\max}(Q) + \lambda_{\max}(K_i^T R K_i)}{\underline{\mathrm{b}}_2(1-\lambda^2)}, \ \alpha \geq \frac{\|SK_i\|_1}{\underline{\mathrm{a}}_1(1-\lambda)}, \ \forall i \in \mathbb{I}_{[1,n_r]}. \tag{5.7.11}$$

Proof. (Lemma 5.7.3) Assume that Lemma 5.7.2 holds. Then, $\mathbb{X}_f = \mathcal{S}$ is λ-contractive and also control invariant. Take β as in (5.7.11). For a fixed i, it follows that

$$0 \geq \beta\underline{\mathrm{b}}_2(\lambda^2 - 1)\|x\|_2^2 + (\lambda_{\max}(Q) + \lambda_{\max}(K_i^T R K_i))\|x\|_2^2. \tag{5.7.12}$$

However, from Lemma 5.7.2 part 3, we have that $\forall x \in \mathbb{X}_f$

$$\beta\underline{\mathrm{b}}_2(\lambda^2 - 1)\|x\|_2^2 \geq \beta(\lambda^2 - 1)\psi^2_{\mathbb{X}_f}(x) \geq \beta(\psi^2_{\mathbb{X}_f}(x^+) - \psi^2_{\mathbb{X}_f}(x)). \tag{5.7.13}$$

At the same time we have:

$$(\lambda_{\max}(Q) + \lambda_{\max}(K_i^T R K_i))\|x\|_2^2 \geq x^T(Q + K_i^T R K_i)x. \tag{5.7.14}$$

Combine (5.7.13) and (5.7.14) to have, $\forall x \in \mathbb{X}_f$, that

$$\beta(\psi^2_{\mathbb{X}_f}(x^+) - \psi^2_{\mathbb{X}_f}(x)) + x^T(Q + K_i^T R K_i)x \leq 0, \ \forall i \ \in \mathbb{I}_{[1,\, n_r]}. \tag{5.7.15}$$

Then, thanks to the arguments of Lemma 5.7.1, the inequality (5.7.15) holds for the LPV system (2.6.14) for any r.

From (5.7.11) we have that

$$\alpha \underline{a}_1(\lambda - 1) + \|SK_i\|_1 \leq 0, \; \forall i \; \in \mathbb{I}_{[1,\, n_r]}. \tag{5.7.16}$$

For a fixed i, from Lemma 5.7.2 we have, $\forall x \in \mathbb{X}_f$, that

$$0 \geq (\alpha \underline{a}_1(\lambda - 1) + \|SK_i\|_1)\|x\|_1 \geq \alpha(\lambda - 1)\psi_{\mathbb{X}_f}(x) + \|SK_i x\|_1, \tag{5.7.17}$$

$$\geq \alpha(\psi_{\mathbb{X}_f}(x^+) - \psi_{\mathbb{X}_f}(x)) + \|SK_i x\|_1. \tag{5.7.18}$$

The above linear inequalities also hold $\forall r$, as they hold for all vertices. Add the terms of (5.7.15) and (5.7.18) to obtain the inequality

$$\beta(\psi^2_{\mathbb{X}_f}(x^+) - \psi^2_{\mathbb{X}_f}(x)) + \alpha(\psi_{\mathbb{X}_f}(x^+) - \psi_{\mathbb{X}_f}(x)) \leq -x^T(Q + K_i^T R K_i)x - \|SK_i x\|_1, \; \forall x \in \mathbb{X}_f, \tag{5.7.19}$$

required by (A3). ∎

Remark 31. As in Remark 26, the function $\psi_{\mathbb{S}}$ is continuous, convex, positive definite and PWA in x, with $\psi_{\mathbb{S}}(0) = 0$. As a consequence of this $\psi^2_{\mathbb{S}}(x)$ is also continuous, convex and it is PWQ, with $\psi^2_{\mathbb{S}}(0) = 0$. Therefore, the quantities $\underline{a}_p$ and $\underline{b}_p$ can be computed, for instance, from the explicit solution of one of the two mpLPs on the right-hand side of (5.3.4).

Remark 32. Notice that, if the terminal controller gain is constant and independent of r, then the terminal ingredients computation is faster. However, for non-linear systems a single stabilising K might not exist or, in other cases, the terminal set could be small if a constant K is used.

5.7.2.1 Symmetric constraints

In the particular case of $\mathbb{X}$ and $\mathbb{U}$ symmetric with respect to the origin (0-symmetric) then, any invariant $\mathbb{S}$ is also 0-symmetric. Then, the Minkowski function is given by (5.3.7). When the terminal set is set is large, one can compute the scalings as follows:

Lemma 5.7.4. Assume $\mathbb{S}$ is symmetric. Then, in Lemma 5.7.3, it is possible to take

$$\beta \geq \frac{n\left(\|Q^{1/2}G^L\|_2^2 + \|(G^L)^T K_i^T R K_i G^L\|_2\right)}{1 - \lambda^2}, \; \alpha \geq \frac{n_s \|SK_i G^L\|_\infty}{1 - \lambda}, \; \forall i \in \mathbb{I}_{[1,\, n_r]}. \tag{5.7.20}$$

Proof. (Lemma 5.7.4) The results are obtained through similar steps as taken for Lemma 5.7.3, with the difference that here the stage cost elements are upper bounded by ∞-norms, and $\psi_{\mathbb{S}}$ is left as it is. The left pseudo-inverse of G, G^L, is used to collect $\|Gx\|_\infty$ and to remove x from the inequalities. ■

5.7.3 Terminal set and domain of attraction

Computation of invariant and contractive sets for constrained non-linear systems is a non-trivial problem, and it is beyond the scope of this thesis. This section illustrates a possible approach that is based on the solution of (at best) one constrained LP. This requires the knowledge for a set of vectors for which the LP is feasible. The search for these vertices could however be non-trivial. A set of alternative methodologies are also reviewed from the MPC literature, that have more solution guarantees, at the expense of computation increase.

5.7.3.1 The considered method: control at the vertices

Consider Algorithm 2, based on Proposition 7.34, p. 254 of [Blanchini & Miani, 2008].

Algorithm 2 Terminal set computation (based on Proposition 7.34 of [Blanchini & Miani, 2008])

1. Choose a matrix V containing a set of candidate vertices, and $\lambda \in (0, 1)$, $\epsilon \geq 0$, $\gamma_{\max} > 0$.
2. Solve the LP
$$\begin{array}{ll} \max_\chi & \gamma - \epsilon \Sigma_{i=1}^{n_r} \|\bar{K}_i\|_1 \\ \text{s.t.:} & A_i \gamma V + B_i \bar{K}_i V = V \bar{Y}_i, \\ & \underline{1}^T \bar{Y}_i \leq \gamma \lambda \underline{1}^T, \\ & E \bar{K}_i V \leq [\underline{1}, \dots, \underline{1}], \ \forall i \in \mathbb{I}_{[1,\, n_r]}, \\ & 0 < \gamma \leq \gamma_{\max}, \end{array} \tag{5.7.21}$$
where $\chi = (\gamma,\ \bar{K}_1,\ \dots,\ \bar{K}_{n_r},\ \bar{Y}_1,\ \dots,\ \bar{Y}_{n_r})$, and the entries of $\bar{Y}_i$ are all non-negative.
3. **IF** (5.7.21) is feasible **THEN** $K_i \leftarrow \frac{1}{\gamma^\star} \bar{K}_i^\star$, $\forall i \in \mathbb{I}_{[1,\, n_r]}$ **ELSE** goto 1.
4. $\mathbb{X}_f \leftarrow \mathrm{conv}(\gamma^\star V)$.

Algorithm 2 is the extension of Algorithm 1 for LPV systems. Similar to Algorithm 1, Algorithm 2 algorithm 2 requires at best the solution of a single LP, the number of inequality constraints of which scales linearly with n_r. The parameter ϵ can be used to regulate the tradeoff between the size of the terminal set, and the sparsity of the K_is. Sparse K_is could be useful, for instance, for the prioritised actuators scenarios of Chapter 6. On the other hand, $\epsilon = 0$ provides the largest set for the given V and the given LPV system. Since from (5.4.1) the DOA is contained in $\mathbb{X}$, a possible choice is to take V as the vertices of $\mathbb{X}$. In this case, $\gamma^\star = 1$ provides the largest DOA. The weakness of Algorithm 2 is the requirement of a matrix of vertices V, such that problem (5.4.2) is solvable. The search for V can be non-trivial as discussed in Section 5.4.1. The reader is referred to the methods in [Cannon *et al.* , 2003; Fiacchini *et al.* , 2012; Lazar & Jokic, 2010].

Once a candidate terminal set (with its terminal controller) is obtained, from Algorithm 2 or the ones in [Cannon *et al.* , 2003; Fiacchini *et al.* , 2012; Lazar & Jokic, 2010], one could then obtain a less conservative set for instance by using the technique in Section 5.7.3.2. This allows for a relatively large set and potentially for a sparse terminal controller.

5.7.3.2 Using the maximal admissible set

A different formulation can be used if an appropriate LPV embedding set $\mathcal{S}$ and a stabilising controller $K(r)$ for (2.6.14) are known. This has higher complexity than the technique of the previous section, since if requires the intersection of a number of inequalities that is combinatorial in n_r (see p. 180 of [Blanchini & Miani, 2008])) in contrast to the linear complexity of (5.7.21). On the other hand, the following strategy is capable to provide the least conservative terminal constraint for (2.6.14), given $K(r)$.

The idea is compute the maximal admissible set inside $\mathcal{S}$ for the closed-loop system [Blanchini & Miani, 2008; Kerrigan, 2000]. This is formalised by the following:

Lemma 5.7.5. If an asymptotically stabilising controller, $K(r)$, is known for the unconstrained system (2.6.14) together with a set $\mathcal{S}$ such that $K(r)x \in \mathbb{U},\ \forall x \in \mathcal{S}$ then, problem (5.7.6) has a solution for $V = \mathcal{V}(\mathcal{O}_\infty^\lambda)$, where $\mathcal{O}_\infty^\lambda$ is the maximal admissible set for the system:

$$\begin{aligned} x(k+1) &= \frac{1}{\lambda}(A(r) + B(r)K(r))x, \\ x &\in \mathcal{S},\ K_i x \in \mathbb{U}, \forall i \in \mathbb{I}_{[1,\ n_r]}. \end{aligned} \tag{5.7.22}$$

Therefore, Lemma 5.7.2 is satisfied. The set $\mathcal{O}_\infty^\lambda$ exists and it can be computed in finite time, if $\lambda \in (\max_i \lambda_{\max}(A_i + B_i K_i),\ 1)$.

Proof. (Lemma 5.7.5) By definition of maximal admissible set, it follows that $x(0) \in \mathcal{O}_\infty^\lambda \Rightarrow x(k) \Rightarrow \mathcal{O}_\infty^\lambda, \forall k \geq 0$ for system (5.7.22). Therefore, for the system (2.6.14) we have $x(0) \in$

$\mathcal{O}_\infty^\lambda \Rightarrow x(k) \in \lambda^k \mathcal{O}_\infty^\lambda, \forall k \geq 0$, namely, the set is λ-contractive. The existence of the set $\mathcal{O}_\infty^\lambda$ follows from the fact that, if $\lambda \in (\max_i \lambda_{\max}(A_i + B_i K_i),\ 1)$, then $\frac{1}{\lambda}(A(r) + B(r)K(r))$ has all eigenvalues strictly in the unit circle (see Theorem 5.17, p. 180 of [Blanchini & Miani, 2008]). Moreover, $\mathcal{O}_\infty^\lambda$ consists of a finite number of inequalities as in Theorem 5.17 of [Blanchini & Miani, 2008]. ■

Remark 33. Using Lemma 5.7.5 can be computationally demanding, for LPV systems with many parameters. In fact, the computation $\mathcal{O}_\infty^\lambda$ has combinatorial complexity, depending on the number of parameters n_r (see Chapter 5.4, p. 180 of [Blanchini & Miani, 2008]). On the other hand, for LTI systems the computation is much less intensive and the use of the maximal admissible set is quite common in the MPC literature [Kerrigan, 2000; Rawlings & Mayne, 2010].

5.7.3.3 Using a robustly control invariant set

Similar to the LTI case, in Section 5.4.2.2, a further alternative to Algorithm 2 can be derived by means of the results in [Blanchini, 1994; Grammatico & Pannocchia, 2013]. This however relies on polytope projections, the complexity of which can increase significantly with n_r (since the number of half-spaces increases). A more tractable alternative to this formulation is given in [Blanchini & Miani, 2003], where it is shown how a dynamic LPV compensator can be computed with small loss of generality.

5.8 Conclusions

This chapter has presented two formulations ℓ_{asso}-MPC that provide closed-loop asymptotic stability for LTI systems as well as their extension for a class of non-linear systems. These strategies make use of a terminal constraint set and differ for the particular choice of the terminal cost. In particular, the first approach features a quadratic function plus a scaled p-norm. The implementation of the proposed stability conditions for this particular cost can, however, require the solution of a non-linear optimisation even for LTI systems. The second strategy, on the other hand, uses a weighted sum of the terminal set Minkowski function and its square and requires less computation. This new approach is referred to as "ℓ_{asso}-MPC version 2" and assumes the terminal set to be λ-contractive.

A set of techniques has been reviewed for the computation of the required candidate terminal controller and terminal set, with the aim of obtaining a large domain of attraction. By means of the Minkowski function terminal cost the required offline complexity for ℓ_{asso}-MPC version 2 is of the same order as for a standard MPC with quadratic cost. In particular, if the given non-linear system can be locally embedded into a linear parameter-varying model, then the ingredients for stability can be computed at best by means of a single linear program. Moreover, if a locally stabilising controller is known then these ingredients can always be computed in finite time as done for quadratic MPC. The online complexity of the new controller is only slightly greater than the one of ℓ_{asso}-MPC version 1 as the new terminal cost can be implemented by introducing a set of additional slack variables. The terminal set is computed independently of the stage cost, thus facilitating online tuning.

For LTI models, ℓ_{asso}-MPC version 2 results in a unique, continuous and piece-wise affine control law as well as ℓ_{asso}-MPC version 1. On the other hand, the guarantee of stability makes sparsity through actuators more difficult to be obtained, albeit possible when the system can be stabilised by means of a subset of inputs. This will be further investigated in the next chapter, where a synthesis method is proposed for systems with a preferred set of actuators.

CHAPTER

SIX

DESIGN OF LASSO MPC FOR PRIORITISED AND AUXILIARY ACTUATORS

6.1 Introduction

This chapter addresses the use of ℓ_{asso}-MPC for control scenarios in which actuators can be clearly divided into two groups: *preferred actuators* and *auxiliary actuators*. A *pre-existing stabilising MPC* is assumed to be given, which considers only the former. These preferred actuators are set to be used for most of the time and are meant to stabilise the system. On the other hand, the auxiliary actuators are meant to play a secondary role. In particular, given a pre-existing MPC the chapter provides a set of numerical tools to construct a new ℓ_{asso}-MPC, which includes the auxiliary actuators among its decision variables.

By means of the proposed tools the control engineer will be able to specify (to a certain extent) a closed region of the state (or the control error) space, $\mathcal{X}_{\mathrm{nom}}$, in which auxiliary actuators are guaranteed not to be used by the resulting ℓ_{asso}-MPC. Outside this region, referred to as the *region of nominal operation*, the use of auxiliary actuators will occur at the discretion of the cost function and the constraints. This remaining region of the feasible space, $\mathcal{X}_{\mathrm{aux}}$, is referred to as the *region of special operation*. The proposed technique allows, for instance, to restrict the use of auxiliary actuators to the times when the main actuators are saturated or when certain constraints become active. At the same time, the original MPC controller is recovered in the nominal operation zone.

As seen in Chapter 2, in the presence of constraints and the finite prediction horizon length impose limitations on the achievable DOA. Similar limitations occur in the achievement of a certain

$\mathcal{X}_{\text{nom}}$. The proposed procedures make use of multi-parametric programming and of explicit MPC techniques.

Throughout the chapter, the notation $\Sigma_{\{\text{i}\}}$ ($\Sigma_{\{\text{ii}\}}$) will refer to the original system Σ when only actuator set $\{\text{i}\}$ (respectively $\{\text{ii}\}$) is used. The set of indices $\{\text{i}\}$, with $m_{\{\text{i}\}} = \text{card}\{\text{i}\}$, represents the *main*, or *preferred*, actuator set, $u^{\{\text{i}\}}$, while $\{\text{ii}\}$, with $m_{\{\text{ii}\}} = \text{card}\{\text{ii}\}$, is for the *auxiliary* set, $u^{\{\text{ii}\}}$.

6.2 Partial regularisation

The proposed strategy uses a ℓ_{asso}-MPC with cost (3.2.2), where

$$S = \begin{bmatrix} 0 & S_{\{\text{ii}\}} \end{bmatrix}, \tag{6.2.1}$$

where the matrix, $S_{\{\text{ii}\}} \in \mathcal{D}_{++}$, has to be computed. Note that, in (6.2.1), the actuators have been grouped as $u = (u^{\{\text{i}\}},\ u^{\{\text{ii}\}})$. We assume that only the first group of actuators is used to stabilise the system, namely, the terminal cost and terminal constraint are designed under the assumption that $u^{\{\text{ii}\}}(k) = 0, \forall k \in \mathbb{I}_{\geq 0}$.

Assumption 8. (A8)

Assume that $R = \text{BlockDiag}(R_1,\ R_2) \succ 0$, and that the MPC constraints are non-degenerate[1].

A unitary permutation, Υ, is used to separate the predictions of $u^{\{\text{i}\}}$ from the ones of $u^{\{\text{ii}\}}$, namely:

$$\Upsilon \underline{\mathrm{u}} = \begin{bmatrix} \underline{\mathrm{u}}_{\{\text{i}\}} \\ \underline{\mathrm{u}}_{\{\text{ii}\}} \end{bmatrix}, \quad \Theta \Upsilon^T = [\Theta_1,\ \Theta_2]. \tag{6.2.2}$$

The permutation can be computed as follows. Define $\mathbb{I}^{\{\text{i}\}}$ and $\mathbb{I}^{\{\text{ii}\}}$ as, respectively, the indices of the columns of B corresponding to the columns of actuator set $\{\text{i}\}$ and $\{\text{ii}\}$. Define $\mathbb{I}^{(\text{I})}$ and $\mathbb{I}^{(\text{II})}$ as

$$\begin{aligned} \mathbb{I}^{(\text{I})} &= \left\{\mathbb{I}^{\{\text{i}\}}, \dots, \mathbb{I}^{\{\text{i}\}} + m(k-1), \dots, \mathbb{I}^{\{\text{i}\}} + m(N-1)\right\}, \\ \mathbb{I}^{(\text{II})} &= \left\{\mathbb{I}^{\{\text{ii}\}}, \dots, \mathbb{I}^{\{\text{ii}\}} + m(k-1), \dots, \mathbb{I}^{\{\text{ii}\}} + m(N-1)\right\}, \end{aligned} \tag{6.2.3}$$

where $k \in [1,\ N]$ is the *block index*, and m is the number of inputs[2]. Suppose that an index $j \in \mathbb{I}^{(\text{I})}$, and define $\bar{j}$ as the corresponding index in the list, that is $j = \mathbb{I}^{(\text{I})}_{\bar{j}}$, with $\bar{j} \in [1,\ m_{\{\text{i}\}}N]$

[1] Degenerate constraints could also be accommodated, for instance using the methods in [Tondel *et al.*, 2003].
[2] In eq. (6.2.3) the "$+m(k-1)$" means that $m(k-1)$ is added to all the indices of the list.

(or respectively, $\bar{j} \in [1, (m - m_{\{\text{i}\}})N]$). Υ is constructed as follows:

$$\Upsilon(i, j) = \begin{cases} 1 & \text{if } j \in \mathbb{I}^{(\text{I})} \text{ and } i = \bar{j}, \\ 1 & \text{if } j \in \mathbb{I}^{(\text{II})} \text{ and } i = m_{\{\text{i}\}}N + \bar{j}, \\ 0 & \text{otherwise.} \end{cases} \tag{6.2.4}$$

6.2.1 Existence of the nominal operation zone

Partition $\Theta\Upsilon^T = [\Theta_1 \ \Theta_2]$ and $\Upsilon\mathbf{R}\Upsilon^T = \text{BlockDiag}(\mathbf{R}_1, \ \mathbf{R}_2)$, $\Omega\Upsilon^T = [\Omega_1, \ \Omega_2]$. The following optimality conditions are obtained for (3.2.1):

$$\Theta_1^T \mathbf{Q}\Psi x(k) + \Theta_1^T \mathbf{Q}(\Theta_1 \underline{\mathbf{u}}_{\{\text{i}\}}^\star + \Theta_2 \underline{\mathbf{u}}_{\{\text{ii}\}}^\star) + \mathbf{R}_1 \underline{\mathbf{u}}_{\{\text{i}\}}^\star + \Omega_1^T v^\star = 0, \tag{6.2.5}$$

$$\Theta_2^T \mathbf{Q}\Psi x(k) + \Theta_2^T \mathbf{Q}(\Theta_1 \underline{\mathbf{u}}_{\{\text{i}\}}^\star + \Theta_2 \underline{\mathbf{u}}_{\{\text{ii}\}}^\star) + \mathbf{R}_2 \underline{\mathbf{u}}_{\{\text{ii}\}}^\star + 1/2 \ \Omega_2^T v^\star \in 1/2 \ W_{\{\text{ii}\}}^T \partial \| W_{\{\text{ii}\}} \underline{\mathbf{u}}_{\{\text{ii}\}}^\star \|_1, \tag{6.2.6}$$

where $W_{\{\text{ii}\}} = I_{(N\times N)} \otimes S_{\{\text{ii}\}}$. Define $\mathcal{O}_\infty^{\{\text{i}\}}$ as the maximal admissible set, for $\Sigma_{\{\text{i}\}}$ in closed loop with the LQR controller, $u^{\{\text{i}\}}(k) = Kx(k)$, with Q, R_1 and consider the following:

Assumption 9. (A9) Assume (A8). Assume also that the matrix P solves the discrete-time LQR Riccati equation (see Chapter 2.6.5.1) associated with $\Sigma_{\{\text{i}\}}$, Q, and R_1. Moreover, $\mathbb{X}_f = \mathcal{O}_\infty^{\{\text{i}\}}$.

Recall that a positive definite diagonal matrix S is said to be of class $\mathcal{D}_{++}$, namely $S \in \mathcal{D}_{++}$. The following result is obtained:

Theorem 6.2.1. Assume (A9), and that $S_{\{\text{ii}\}} \in \mathcal{D}_{++}$. Then, there exists a set $\mathbb{G}$, with $0 \in \text{int}(\mathbb{G})$, such that, if $x(k) \in \mathbb{G}$, the ℓ_{asso}-MPC with S given by (6.2.1) gives $\underline{\mathbf{u}}_{\{\text{ii}\}}^\star = 0$. Moreover, $\underline{\mathbf{u}}_{\{\text{i}\}}^\star$ is equal to the solution of the LQ-MPC designed or system $\Sigma_{\{\text{i}\}}$ with Q, R_1.

Proof. From (6.2.6) the solution $\underline{\mathbf{u}}_{\{\text{ii}\}}^\star = 0$ is obtained if

$$\|\Theta_2^T \mathbf{Q}\Psi x(k) + (\Theta_2^T \mathbf{Q}\Theta_1)\underline{\mathbf{u}}_{\{\text{i}\}}^\star + \frac{1}{2} \ \Omega_2^T v^\star(x))\|_\infty \leq \frac{1}{2} \min_j \left\{ S_{\{\text{ii}\}_{jj}} \right\} \tag{6.2.7}$$

Note that, if constraints are not active, we have $v^\star(x) = 0$. Since $\underline{\mathbf{u}}_{\{\text{i}\}}^\star(x)$ and $v^\star(x)$ are PWA, then (6.2.7) describes a set of states, $\widetilde{\mathbb{G}}$, which clearly contains the origin. Assume now that $x(k) \in \widetilde{\mathbb{G}} \cap \mathcal{O}_\infty^{\{\text{i}\}}$. Then, $\underline{\mathbf{u}}_{\{\text{ii}\}}^\star(x(k)) = 0$ and, from (6.2.5), if follows that

$$\underline{\mathbf{u}}_{\{\text{i}\}}^\star(x(k)) = -(\Theta_1 \mathbf{Q}\Theta_1 + \mathbf{R}_1)^{-1} \Theta_1^T \mathbf{Q}\Psi x(k), \ v^\star(x(k)) = 0, \tag{6.2.8}$$

that is the unconstrained solution of the LQ-MPC in the claim. Hence, we have $\mathbb{G} \supseteq \widetilde{\mathbb{G}} \cap \mathcal{O}_{\infty}^{\{\mathrm{i}\}}$. ■

In Theorem 6.2.1, it is assumed that $S_{\{\mathrm{ii}\}} \in \mathcal{D}_{++}$. For the general case, one could consider the change of coordinates $z = S_{\{\mathrm{ii}\}}\underline{\mathrm{u}}_{\{\mathrm{ii}\}}$.

6.2.2 Formulation as a mpQP

In order for the regularisation matrix $S_{\{\mathrm{ii}\}}$ to be computed, in Section 6.2.4, in this section the considered ℓ_{asso}-MPC problem is formulated as a mpQP. The formulated mpQP problem can be solved offline using standard tools, for instance the open source MPT toolbox V.2.6.3 for Matlab [Kvasnica *et al.* , 2004]. The decision variables, namely the preferred and auxiliary actuators predictions, are for convenience grouped in two separate blocks. The reason for doing this is of practical nature. In particular, this will be useful in Section 6.2.4, where the ℓ_{asso}-MPC is solved under the constraint $\underline{\mathrm{u}}_{\{\mathrm{ii}\}} = 0$, for computing $S_{\{\mathrm{ii}\}}$. Notice that, if the MPC is instead solved online, then more efficient implementations exist, see for instance [Jerez *et al.* , 2011] and reference therein.

6.2.2.1 Using slack variables

A 1-norm cost can be represented by introducing the slack variables $\sigma \in \mathbb{R}^{(m_{\{\mathrm{ii}\}}N)}$ and $2Nm_{\{\mathrm{ii}\}}$ slack constraints. Problem (3.2.1–3.2.2) can now be formulated as a constrained semi-definite QP, for example in the condensed formulation:

$$\begin{aligned} &\chi^{\star} = \arg\min_{\chi} \tfrac{1}{2}\, \chi^T \bar{H} \chi + \chi^T \bar{G}(x) \\ &\text{s.t. } \bar{\Omega}\, \chi \leq b(x),\ x = x(k), \end{aligned} \tag{6.2.9}$$

where

$$\chi = \begin{bmatrix} \underline{\mathrm{u}}_{\{\mathrm{i}\}} \\ \underline{\mathrm{u}}_{\{\mathrm{ii}\}} \\ \sigma \end{bmatrix}, \quad \bar{H} = \begin{bmatrix} 2\mathbf{H} & 0 \\ 0 & 0 \end{bmatrix}, \quad \bar{G}(x) = \begin{bmatrix} 2\Upsilon\Gamma x \\ \underline{1} \end{bmatrix}, \tag{6.2.10}$$

and where $\mathbf{H} = \Upsilon H \Upsilon^T \succ 0$. Constraint matrices for (6.2.9) are

$$\bar{\Omega} = \begin{bmatrix} \Omega\Upsilon^T & 0 \\ \Pi & \end{bmatrix}, \quad b(x) = \begin{bmatrix} \underline{1} - Mx \\ 0_{(2m_{\{\mathrm{ii}\}}N \times 1)} \end{bmatrix}, \quad \Pi = \begin{bmatrix} 0 & W_{\{\mathrm{ii}\}} & -I_{m_{\{\mathrm{ii}\}}N} \\ 0 & -W_{\{\mathrm{ii}\}} & -I_{m_{\{\mathrm{ii}\}}N} \end{bmatrix} \tag{6.2.11}$$

Each time k, the signal applied to the plant is given by $u(k) = K_N(x(k))$, with

$$K_N(x(k)) = [I\ 0\ \dots\ 0]\Upsilon^T \text{BlockDiag}(I,\ [I,\ 0])\chi^\star. \tag{6.2.12}$$

Recall the definition of *primal* and *dual degeneracy*, from Chapter 2.4.4. Since in problem (6.2.9) the matrix $\bar{H}$ is only positive semi-definite, the problem is not strictly convex, meaning that multiple optimal solutions can occur (dual degeneracy) [Bemporad *et al.* , 2002a; Tondel *et al.* , 2003]. Solving an explicit MPC with these features is possible, for instance using the technique of [Alessio & Bemporad, 2009]. However, this technique can be quite computationally demanding. On the other hand, standard tools are available for solving mpQPs with a positive definite Hessian, for instance the open source Matlab MPT toolbox V.2.6.3 [Kvasnica *et al.* , 2004]. The algorithms used by MPT are also less computationally demanding than the ones required for semi-definite mpQPs. In Chapter 4.3.2, a set of arrangements has been proposed so that the considered ℓ_{asso}-MPC problem can be formulated as a strictly convex mpQP. The next section will reformulate the results of Chapter 4.3.2 in order to separate preferred and auxiliary actuators, as done in (6.2.9).

6.2.2.2 Splitting the actuators for strict convexity

As mentioned in the above section, we wish to formulate the considered ℓ_{asso}-MPC problem as a mpQP that can be solved, for instance, by means of the open source Matlab MPT toolbox V.2.6.3 [Kvasnica *et al.* , 2004]. For quadratic problems, the toolbox uses the algorithm of [Bemporad *et al.* , 2002a], that needs the mpQP to be strictly convex. In order to use this algorithm, problem (6.2.9) is modified as follows. First, the regularised variables are split into positive and negative part, namely, $u_+,\ u_-$ with $u = u_+ - u_-$. Then, the additional constraints $u_+ \geq 0\ u_- \geq 0$ are added to the original problem. The case of $R_2 \in \mathcal{D}_{++}$ and $S_{\{\text{ii}\}} \in \mathcal{D}_{++}$ is considered. Finally, problem (6.2.9) can be formulated as the following strictly convex mpQP in condensed formulation:

$$\begin{aligned} &\chi^\star = \arg\min_\chi \tfrac{1}{2}\,\chi^T \bar{H} \chi + \chi^T \bar{G}(x) \\ &\text{s.t. } \bar{\Omega}\,\chi \leq b(x),\ x = x(k), \end{aligned} \tag{6.2.13}$$

where

$$\chi = \begin{bmatrix} \underline{\mathrm{u}}_{\{\text{i}\}} \\ \underline{\mathrm{u}}_{\{\text{ii}\}+} \\ \underline{\mathrm{u}}_{\{\text{ii}\}-} \end{bmatrix}, \quad \bar{G}(x) = 2\bar{\Gamma}x + \begin{bmatrix} 0 \\ W_{\{\text{ii}\}}\underline{1} \\ W_{\{\text{ii}\}}\underline{1} \end{bmatrix}, \tag{6.2.14}$$

and where

$$\bar{H} = 2(\bar{\Theta}^T \mathbf{Q} \bar{\Theta} + \bar{\mathbf{R}}), \quad \bar{\mathbf{R}} = \text{BlockDiag}(I_N \otimes R_1,\ I_N \otimes R_2,\ I_N \otimes R_2), \quad \bar{\Gamma} = \bar{\Theta}^T \mathbf{Q} \Psi, \\ \bar{\Theta} = [\Theta_1,\ \Theta_2,\ -\Theta_2]. \tag{6.2.15}$$

Constraint matrices for (6.2.13) are

$$\bar{\Omega} = \begin{bmatrix} \Omega_1 & \Omega_2 & -\Omega_2 \\ & \Pi & \end{bmatrix}, \quad b(x) = \begin{bmatrix} \underline{1} - Mx \\ 0_{(2m_{\{\text{ii}\}}N \times 1)} \end{bmatrix}, \quad \Pi = - \begin{bmatrix} 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \tag{6.2.16}$$

where $\Omega \Upsilon^T = [\Omega_1,\ \Omega_2]$ and Π has identity blocks of size $m_{\{\text{ii}\}}N$. At each time k, the signal applied to the plant is given by $u(k) = K_N(x(k))$, with

$$K_N(x(k)) = [I\ 0\ \dots\ 0]\Upsilon^T \text{BlockDiag}(I,\ [I,\ -I])\chi^\star. \tag{6.2.17}$$

6.2.3 Limitations due to constraints

Define $\mathcal{Q}^{\{\text{i}\}}$ to be the 1-step operator for $\Sigma_{\{\text{i}\}}$, namely

$$\mathcal{Q}^{\{\text{i}\}}(\mathcal{S}) = \{x \in \mathbb{R}^n\ :\ \exists u \in \mathbb{U} \mid u^{\{\text{ii}\}} = 0,\ Ax + Bu \in \mathcal{S}\} \tag{6.2.18}$$

and recall $\mathcal{K}_j(\mathbb{X}, \mathbb{X}_f)$ to be the j-step controllable set for Σ, from $\mathbb{X}$ to $\mathbb{X}_f$, as in Definition 2.5.20. Similarly, define $\mathcal{K}_N^{\{\text{i}\}}(\mathbb{X}, \mathbb{X}_f)$ as the N-step controllable set for $\Sigma_{\{\text{i}\}}$, from $\mathbb{X}$ to $\mathbb{X}_f$. By definition, $\mathbb{X}_N^{\{\text{i}\}} \equiv \mathcal{K}_N^{\{\text{i}\}}(\mathbb{X}, \mathbb{X}_f)$, where $\mathbb{X}_N^{\{\text{i}\}}$ is the feasible region for the formulated ℓ_{asso}-MPC under the constraints $\underline{\mathbf{u}}_{\{\text{ii}\}} = 0$.

Theorem 6.2.2. The maximal feasible set for which it is possible to have $u^{\{\text{ii}\}} = 0$ is

$$\tilde{\mathbb{X}}_N^{\{\text{i}\}} = \mathcal{Q}^{\{\text{i}\}}(\mathcal{K}_{N-1}(\mathbb{X}, \mathbb{X}_f)) \cap \mathbb{X}. \tag{6.2.19}$$

Proof. The set $\mathcal{K}_{N-1}(\mathbb{X}, \mathbb{X}_f)$ is the set of states in $\mathbb{X}$ for which exists an admissible control sequence $\mathbf{u}_{[N-2]} = \{u(k)\}_0^{N-2} \in \mathbb{U}^{N-1}$, such that $\bar{\phi}(N-1, x, \mathbf{u}_{[N-2]}) \in \mathbb{X}_f$. Therefore, if the state $x(k)$ is steerable to this set in one step by means of only the preferred actuators, then from (6.2.18) we have $x(k) \in \mathbb{X}_N = \mathcal{K}_N(\mathbb{X}, \mathbb{X}_f)$. This means that $\tilde{\mathbb{X}}_N^{\{\text{i}\}} \subseteq \mathbb{X}_N$ and that, at the optimum, $u^{\{\text{ii}\}} \neq 0, \forall x \in \mathbb{X}_N / \tilde{\mathbb{X}}_N^{\{\text{i}\}}$. To prove that the considered set is maximal, recall that $\mathcal{K}_j(\mathbb{X}, \mathbb{X}_f) \supseteq \mathcal{K}_j^{\{\text{i}\}}(\mathbb{X}, \mathbb{X}_f),\ \forall j \in \mathbb{I}_{\geq 0}$ and $\mathcal{K}_j(\mathbb{X}, \mathbb{X}_f) \supseteq \mathcal{K}_j^{\{\text{ii}\}}(\mathbb{X}, \mathbb{X}_f),\ \forall j \in \mathbb{I}_{\geq 0}$, as well as

$\mathcal{X}_1 \subseteq \mathcal{X}_2 \Rightarrow \mathcal{K}_j(\mathbb{X}, \mathcal{X}_1) \subseteq \mathcal{K}_j(\mathbb{X}, \mathcal{X}_2),\ \forall j \in \mathbb{I}_{\geq 0}$. ■

Corollary 6.2.3. The maximal set in which $\underline{\mathrm{u}}^{\star}_{\{\mathrm{ii}\}} = 0$ is attainable is $\mathbb{X}_N^{\{\mathrm{i}\}} = \mathcal{K}_N^{\{\mathrm{i}\}}(\mathbb{X}, \mathbb{X}_f)$.

In this chapter, we focus our attention on $\mathbb{X}_N^{\{\mathrm{i}\}}$. In particular, a method is provided to compute $S_{\{\mathrm{ii}\}}$ so that $\underline{\mathrm{u}}^{\star}_{\{\mathrm{ii}\}} = 0$ for any $\mathcal{X}_{\mathrm{nom}} \subseteq \mathbb{X}_N^{\{\mathrm{i}\}}$. Unfortunately, the method does not guarantee that $u^{\{\mathrm{ii}\}} = 0$ will happen $\forall x \in \tilde{\mathbb{X}}_N^{\{\mathrm{i}\}}$. The reason for this is essentially technical, and due to the fact that our 1-nom penalty is common for all future predictions of $u^{\{\mathrm{ii}\}}$. However, as seen in this Section, a tuning for S could exist, so that $u^{\{\mathrm{ii}\}} = 0$ in $\tilde{\mathbb{X}}_N^{\{\mathrm{i}\}}$. This could be obtained, for instance, by up-scaling the results obtained by the methods in the next sections.

6.2.4 Computation of the regularisation penalty

Recall the brief review of *exact penalty functions* in Chapter 2.4.5, and of *explicit MPC* in Chapter 2.6.8. We wish to use these principles to compute a suitable matrix $S_{\{\mathrm{ii}\}}$ in order for the considered ℓ_{asso}-MPC to provide $\underline{\mathrm{u}}^{\star}_{\{\mathrm{ii}\}} = 0,\ \forall x \in \mathcal{X}_{\mathrm{nom}} \subseteq \mathbb{X}_N^{\{\mathrm{i}\}}$, given the set $\mathcal{X}_{\mathrm{nom}}$. A modified version of (6.2.9) is solved for all $x \in \mathbb{X}_N \cap \mathcal{X}_{\mathrm{nom}}$. To use the theory of exact penalty functions, it is imposed that $\|\underline{\mathrm{u}}_{\{\mathrm{ii}\}}\|_1 = 0$ and then the associated Lagrange multiplier is maximised over x. The Lagrange multipliers are computed using mpQP techniques, similarly to [Kerrigan & Maciejowski, 2000a]. In the following, two possible implementations are discussed.

6.2.4.1 Using slack variables

Consider (6.2.9) and add the constraint $\sigma = 0$, resulting in the mpQP:

$$\begin{aligned}
&\chi^{\star} = \arg\min_{\chi} \tfrac{1}{2}\, \chi^T \bar{H} \chi + \chi^T \bar{G}(x) \\
&\text{s.t. } \bar{\Omega}\, \chi \leq b(x), \\
&\quad [0\ 0\ I_{m_{\{\mathrm{ii}\}}N}]\chi = 0, \\
&\quad x \in \mathbb{X}.
\end{aligned} \tag{6.2.20}$$

In (6.2.20), χ, $\bar{H}$ and $\bar{G}$ are defined by (6.2.10) and the inequality constraints are described by (6.2.11). Recall the definitions in (6.2.10)–(6.2.11). For a given x, the KKT conditions for (6.2.20) are:

$$\begin{aligned}\bar{H}\,\chi + \bar{G}(x) + [\Omega\Upsilon^T,\ 0]^T\nu &= -[0\ 0\ I_{m_{\{\mathrm{ii}\}}N}]^T\gamma \\ -[\Omega_1,\ 0,\ 0]\chi - s^\star &= -\underline{1} + Mx \\ \left[0,\ 0,\ I_{m_{\{\mathrm{ii}\}}N}\right]\chi &= 0 \\ (s^\star)^T\nu &= 0 \\ s^\star &\geq 0 \\ \nu &\geq 0 \\ \gamma &\geq 0.\end{aligned} \tag{6.2.21}$$

Define $\mathbb{X} = \{x \in \mathbb{R}^n \mid Lx \leq \underline{1}\}$. Considering (6.2.20), we have

$$\mathbb{X}_N^{\{\mathrm{i}\}} = \pi_n\left(\left\{(x,\chi)\ \middle|\ \begin{array}{c}\bar{\Omega}\,\chi \leq b(x) \\ [0\ 0\ I_{m_{\{\mathrm{ii}\}}N}]\chi = 0 \\ Lx \leq \underline{1}\end{array}\right\}\right), \tag{6.2.22}$$

where π_n is the polytope projection on the first n-coordinates [Kerrigan & Maciejowski, 2000a], namely, on x. Therefore, constructing this region explicitly can be computationally demanding for large systems or long horizon length. The following is obtained:

Theorem 6.2.4. Consider (6.2.9). The exact penalty weight for $\|\underline{\mathrm{u}}_{\{\mathrm{ii}\}}\|_1 = \sum_{j=1}^{N-1}\|u_j^{\{\mathrm{ii}\}}\|_1$, providing $\underline{\mathrm{u}}_{\{\mathrm{ii}\}} = 0, \forall x \in \mathbb{X}_N^{\{\mathrm{i}\}} \cap \mathcal{X}_{\mathrm{nom}}$, is:

$$\begin{aligned}\gamma^\star = &\max_{\chi,\nu,\gamma,x}\ \|\gamma\|_\infty \\ \text{s.t.}\ &(6.2.21)\text{ holds},\ x \in \mathbb{X}_N^{\{\mathrm{i}\}} \cap \mathcal{X}_{\mathrm{nom}}.\end{aligned} \tag{6.2.23}$$

Proof. Since the ∞-norm is the dual of the 1-norm, then Theorem 2.4.3 can be applied. In particular, computing the maximum ∞-norm of the Lagrange multipliers associated with the constraint $\|\underline{\mathrm{u}}_{\{\mathrm{ii}\}}\|_1 = 0$ provides the exact penalty function weight for the constraint $\sigma = 0$ that is $[0\ 0\ I_{m_{\{\mathrm{ii}\}}N}]\chi = 0$. ■

The above Theorem 6.2.4 leads to the desired result:

Theorem 6.2.5. Assume $S_{\{\mathrm{i}\}} = 0$, $S_{\{\mathrm{ii}\}} = \mathbf{diag}\{s_i\}$. Take $\min_i s_i \geq \gamma^\star$ with $\gamma^\star$ solving (6.2.23). Then, for the ℓ_{asso}-MPC solving (6.2.9) or, equivalently, (6.2.13) it follows that $u^{\{\mathrm{ii}\}} = 0$, $\forall x \in \mathbb{X}_N^{\{\mathrm{i}\}} \cap \mathcal{X}_{\mathrm{nom}}$.

Proof. The theorem is a simple application of the previous results. The ℓ_{asso}-MPC is feasible, under the constraint $\|\underline{\mathrm{u}}_{\{\mathrm{ii}\}}\|_1 = 0$, $\forall x \in \mathbb{X}_N^{\{\mathrm{i}\}} \cap \mathcal{X}_{\mathrm{nom}}$. From Lemma 6.2.4, $\gamma^\star$ is the lower bound for the exact penalty on $\|\underline{\mathrm{u}}_{\{\mathrm{ii}\}}\|_1$, namely, $\min_i s_i \geq \gamma^\star$ provides problem (6.2.9) to have the same solution of (6.2.20), $\forall x \in \mathbb{X}_N^{\{\mathrm{i}\}} \cap \mathcal{X}_{\mathrm{nom}}$. ■

As mentioned in Section 6.2.2.1, the matrix $\bar{H}$ of problem (6.2.9) is only positive semi-definite, meaning that the considered mpQP is not strictly convex. Solving an explicit MPC with this features is possible, for instance using the algorithms in [Alessio & Bemporad, 2009; Tondel *et al.* , 2003], that can be quite computationally demanding. In order to reduce computation, the problem structure will be further exploited in the coming section, similarly to Section 6.2.2.2.

6.2.4.2 Splitting the actuators

As mentioned in Section 6.2.2.2, we wish to compute the solution of a ℓ_{asso}-MPC explicitly, by means of the open source MPT toolbox V.2.6.3 for Matlab [Kvasnica *et al.* , 2004] (which uses the algorithm of [Bemporad *et al.* , 2002a]). This requires the considered mpQP to be strictly convex. Assume that $R_2 \in \mathcal{D}_{++}$, then the ℓ_{asso}-MPC can be formulated as the strictly convex mpQP (6.2.13). The exact penalty weight introduced in Theorem 6.2.4, $\gamma^\star$, is then computed in a similar way to [Kerrigan & Maciejowski, 2000a]. In particular, the constraint $(\underline{\mathrm{u}}_{\{\mathrm{ii}\}+}, \underline{\mathrm{u}}_{\{\mathrm{ii}\}-}) = 0$ is imposed on (6.2.13). The resulting problem, which is smaller than (6.2.13), is then solved explicitly. If the constraints of (6.2.13) are degenerate, one can for instance use YALMIP for Matlab [Lofberg, 2004], which is able to compute the control law for a primal degenerate mpQP.

Recall the definitions (6.2.14)–(6.2.16). For a given x, the KKT conditions for (6.2.13), subject to $(\underline{\mathrm{u}}_{\{\mathrm{ii}\}+}, \underline{\mathrm{u}}_{\{\mathrm{ii}\}+}) = 0$, are:

$$\begin{aligned} \bar{H}\,\chi + \bar{G}(x) + [\Omega_1,\ \Omega_2,\ -\Omega_2]^T\nu &= -\Pi^T\gamma \\ -[\Omega_1,\ 0,\ 0]\chi - s^\star &= -\underline{1} + Mx \\ \Pi\chi &= 0 \\ (s^\star)^T\nu &= 0 \\ s^\star &\geq 0 \\ \nu &\geq 0 \\ \gamma &\geq 0. \end{aligned} \tag{6.2.24}$$

Conditions (6.2.24) (as well as problem (6.2.20)) provide the same solution of the LQ-MPC with only $u^{\{\mathrm{i}\}}$ available, namely, of the given *pre-existing LQ-MPC*. Therefore, we can solve this smaller problem and obtain the necessary active sets and critical regions. Despite the fact that the constraints of (6.2.13) are degenerate, the *pre-existing LQ-MPC*, is non-degenerate and it can be solved, for instance, using the MPT toolbox V.2.6.3 for Matlab [Kvasnica *et al.* , 2004] (this is basically what YALMIP does when trying to solve a primal degenerate problem). Finally, the critical region and active sets are used to compute the solution of (6.2.13), subject to $(\underline{\mathbf{u}}_{\{\text{ii}\}+}, \underline{\mathbf{u}}_{\{\text{ii}\}-}) = 0$ and the required Lagrange multipliers, γ, in the region of interest, $\mathcal{X}_{\text{nom}}$. This provides a set of partitions, $\mathcal{P}_j$, to be used next.

In order to proceed further with the computation, define $H^{\{\text{ii,i}\}} = \Theta_2^T \mathbf{Q} \Theta_1$ and partition $\Upsilon\Gamma$ into $(F^{\{\text{i}\}},\ F^{\{\text{ii}\}})$, with

$$F^{\{\text{i}\}} = \sum_{i=0}^{N-1} B_{\{i\}}^T (A^i)^T Q_{i+1} A^{i+1}, \tag{6.2.25}$$

$$F^{\{\text{ii}\}} = \sum_{i=0}^{N-1} B_{\{ii\}}^T (A^i)^T Q_{i+1} A^{i+1}, \tag{6.2.26}$$

The following LP can be used to compute a candidate penalty, $\bar{\gamma}_j$, for each partition $\mathcal{P}_j \cap \mathcal{X}_{\text{nom}} \neq \emptyset$

$$\begin{aligned} &\bar{\gamma}_j = \max_x \|2F^{\{\text{ii}\}}\, x + 2H^{\{\text{ii,i}\}} \underline{\mathbf{u}}^\star_{\{\text{i}\}}(x) + [\tilde{\Omega}_2^T,\ 0,\ 0] v^\star(x)\|_\infty \\ &\text{s.t. } x \in \mathcal{P}_j \cap \mathcal{X}_{\text{nom}}, \end{aligned} \tag{6.2.27}$$

where $\tilde{\Omega}_2$ (similarly $\tilde{\Omega}_1$) contains the rows of Ω_2, from (6.2.16), corresponding to the active set in $\mathcal{P}_j$. Denote the matrix containing the coefficient for the active inequalities as

$$\Omega_{\text{act}} = \begin{bmatrix} \tilde{\Omega}_1 \ \tilde{\Omega}_2 \ \ -\tilde{\Omega}_2 \\ \Pi \end{bmatrix}, \tag{6.2.28}$$

where Π is defined in (6.2.16). The PWA functions are

$$v^\star(x) = -(\Omega_{\text{act}} \bar{H}^{-1} \Omega_{\text{act}}^T)^{-1} \left(b_{\text{act}} + (2\Omega_{\text{act}} \bar{H}^{-1} \bar{\Gamma} - M_{\text{act}}) x \right), \tag{6.2.29}$$

$$\chi^\star(x) = (\underline{\mathbf{u}}^\star_{\{\text{i}\}}, 0, 0) = -2\bar{H}^{-1}\bar{\Gamma} x - \frac{1}{2} \Omega_{\text{act}}^T v^\star(x), \tag{6.2.30}$$

where $\bar{H}$, $\bar{\Gamma}$ are from (6.2.15). The constant term b_{act} and the gain M_{act} are extrapolated from the rows of $b(x)$, which is also defined in (6.2.16), according to the indices of active set and of Π.

Equivalently, one can directly solve for the multipliers, by solving the LP

$$\begin{aligned} &\bar{\gamma}_j = \max_x \|\Pi\, v^\star(x)\|_\infty \\ &\text{s.t. } x \in \mathcal{P}_j \cap \mathcal{X}_{\text{nom}}, \end{aligned} \tag{6.2.31}$$

which avoids the computation of $\underline{\mathrm{u}}^\star_{\{\mathrm{i}\}}(x)$ and of several components of $v^\star(x)$. The exact penalty is then obtained by taking the largest $\bar{\gamma}_j$ over all feasible partitions. Note that, as shown in [Bemporad *et al.* , 2002a], the partitions of explicit MPC for LTI systems are adjacent (there are no gaps between the partitions). The following result is obtained:

Theorem 6.2.6. Consider the partition $\mathbf{part}\mathcal{P} = \{\mathcal{P}_j\}_{j=1}^{n_p}$ of the feasible region of (6.2.13) subject to $(\underline{\mathrm{u}}_{\{\mathrm{ii}\}+}, \underline{\mathrm{u}}_{\{\mathrm{ii}\}-}) = 0$, obtained from an mpQP solver, namely, with each $\mathcal{P}_j$ corresponding to an optimal active set of the *pre-existing MPC*. Then, the regularisation penalty $\gamma^\star$, satisfying Theorem 6.2.4, is given by:

$$\begin{aligned} \gamma^\star \; &= \max_j \bar{\gamma}_j, \\ &\text{s.t. } \mathcal{P}_j \in \mathbf{part}\mathcal{P}, \\ &\mathcal{P}_j \cap \mathcal{X}_{\text{nom}} \neq \emptyset, \end{aligned} \tag{6.2.32}$$

where $\bar{\gamma}_j$ is computed by either (6.2.31) or (6.2.27), subject to (6.2.29), (6.2.30) and (6.2.28).

Proof. Similar to [Bemporad *et al.* , 2002a], the theorem can be verified by elaborating (6.2.24). Thanks to strict convexity ($\bar{H} \succ 0$), from the first equation of (6.2.24) we can obtain (6.2.30). Then, substituting (6.2.30) into the second equality of (6.2.24), we can obtain (6.2.29), given the active set (since constraints are not degenerate). Then, again from the first line of (6.2.24) it can be easily verified that (6.2.31) or (6.2.27) are equivalent. Finally, in order to obtain the result, the multipliers are to be evaluated for each partition $\mathcal{P}_j$ for which $\mathcal{P}_j \cap \mathcal{X}_{\text{nom}} \neq \emptyset$. ■

6.2.4.3 Algorithms

The following algorithm summarises the required steps to compute $S_{\{\mathrm{ii}\}}$ (see the next page):

Algorithm 3 Computation of $S_{\{\text{ii}\}}$ (partial regularisation)

1. Take any $x \in \mathcal{X}_{\text{nom}} \cap \mathbb{X}_N^{\{\text{i}\}}$
2. Identify the active set $\mathcal{A}_j = \mathcal{A}(x)$ by solving (6.2.13) at x subject to $(\underline{\mathrm{u}}_{\{\text{ii}\}+}, \underline{\mathrm{u}}_{\{\text{ii}\}-}) = 0$
3. Evaluate (6.2.28), (6.2.29), and (6.2.30)
4. Compute $\mathcal{P}_j$ substituting (6.2.30) into $\bar{\Omega}\,\chi \leq b(x)$
5. Compute $\bar{\gamma}_j$ by solving the LP (6.2.31)
6. **IF** $(\mathcal{X}_{\text{nom}} \cap \mathbb{X}_N^{\{\text{i}\}}) \backslash \mathcal{P}_j = \emptyset$ **THEN**
   (a) Compute $\gamma^\star \leftarrow \max_j \bar{\gamma}_j$.
   (b) $S_{\{\text{ii}\}} \leftarrow \gamma^\star I$
7. **ELSE**
   (a) Identify a new active set $\mathcal{A}_{j+1}$, set $j \leftarrow j+1$
   (b) **GOTO** 3.

Algorithm 3 is conceptually similar to the ones in [Bemporad *et al.* , 2002a; Tondel *et al.* , 2003]. In particular, step 2 can be performed by means of an active set QP solver. The search for a new active set, in step 7.a, can be performed for instance with the method of [Bemporad *et al.* , 2002a]. In the degenerate case, the method of [Tondel *et al.* , 2003] can be used for step 7.a, where an appropriate implementation of step 3 is also reviewed based on QP decomposition [Borrelli *et al.* , 2009; Fletcher, 1987]. The complexity of Algorithm 3 is at least the same as for computing the explicit MPC $\forall x \in \mathcal{X}_{\text{nom}}$ with the addition for each region of a few more evaluations in step 3 (more Lagrange multipliers), and of the LP (6.2.31) in step 4. The computation of $\gamma^\star$ can be performed efficiently requiring a comparison per region.

Suppose the explicit solution of (6.2.13), subject to $(\underline{\mathrm{u}}_{\{\mathrm{ii}\}+}, \underline{\mathrm{u}}_{\{\mathrm{ii}\}-}) = 0$, to be available $\forall x \in \mathbb{X}_N^{\{\mathrm{i}\}}$ in terms of $\mathbf{part}\mathcal{P}$ and the active sets $\mathcal{A} = \{\mathcal{A}_j \mid j \in \mathbb{I}_{[1,N_r^{\{\mathrm{i}\}}]}\}$. Then a simple (but inefficient) algorithm for computing the penalty $\gamma^\star$ is the following:

Algorithm 4 Computation of $S_{\{\mathrm{ii}\}}$ given $\mathbf{part}\mathcal{P}$, $\mathcal{A}$ (partial regularisation)

1. $\gamma^\star \leftarrow 0$
2. **FOR** $j \in \mathbb{I}_{[1,N_r^{\{\mathrm{i}\}}]}$
   (a) **IF** $(\mathcal{X}_{\text{nom}} \cap \mathbb{X}_N^{\{\mathrm{i}\}}) \backslash \mathcal{P}_j \neq \emptyset$ **THEN**
      i. Evaluate $\mathcal{A}_j$, and (6.2.28), (6.2.29), and (6.2.30)
      ii. Compute $\bar{\gamma}_j$ by solving the LP (6.2.31)
      iii. **IF** $\bar{\gamma}_j \geq \gamma^\star$ **THEN** $\gamma^\star \leftarrow \bar{\gamma}_j$
3. $S_{\{\mathrm{ii}\}} \leftarrow \gamma^\star I$

Algorithm 4 is based on algorithm 2 of [Borrelli *et al.* , 2008] for evaluation of explicit MPC. The latter has complexity that is combinatorial in the number of partitions ($2nN_C^{\{\mathrm{i}\}}$ flops, with $N_C^{\{\mathrm{i}\}} = \sum_j^{N_r^{\{\mathrm{i}\}}} N_c^j$). The complexity of Algorithm 4 is even higher, since we need to solve the LP (6.2.31) possibly in several partitions. A formal complexity analysis of Algorithm 4 has not been performed, however efficient implementations are possible, for instance based on the partition search algorithm 4 of [Borrelli *et al.* , 2008], which has a complexity proportional to $(2n-1)N_r^{\{\mathrm{i}\}} + N_C^{\{\mathrm{i}\}}$ flops. The partition storage is, however, necessary for Algorithm 4, in contrast with algorithm 4 of [Borrelli *et al.* , 2008]. In any case, the proposed procedures should be performed offline unless the number of partitions for which

$$\mathcal{X}_{\text{nom}} \cap \mathbb{X}_N^{\{\text{i}\}} \backslash \mathcal{P}_j \neq \emptyset$$

is small. In this case one could try to compute $\gamma^\star$ in parallel with the MPC solver.

The reason for taking (6.2.27) into consideration is that, if the redundant actuators are many, one could try to approximate $\gamma^\star$ by plugging into (6.2.27) the solution of the LQ-MPC with preferred actuators only (the *pre-existing LQ-MPC*). This is an approximation only if the dual variables for the two problems are different. In particular, if the active sets of interest don't include state constraints and if input constraints don't involve at the same time elements of $\underline{\mathrm{u}}_{\{\text{i}\}}$ and $\underline{\mathrm{u}}_{\{\text{ii}\}}$. Then solving (6.2.27) with the PWA functions from the *pre-existing LQ-MPC* will return the correct $\gamma^\star$. A significant computational speedup could then be achieved.

Remark 34. In most cases Algorithms 3 and 4 are tractable only for systems with $m_{\{\text{i}\}}, n, N$ less than 5, with possible exceptions up to $n = 10$ [Alessio & Bemporad, 2009].

When using Algorithms 3 and 4 it is possible to determine wether a given $\mathcal{X}_{\text{nom}}$ is contained in $\mathbb{X}_N^{\{\text{i}\}}$ before performing the full search, providing the vertices of $\mathcal{X}_{\text{nom}}$ are used as starting points. In particular, if the MPC problem subject to $(\underline{\mathrm{u}}_{\{\text{ii}\}+}, \underline{\mathrm{u}}_{\{\text{ii}\}+}) = 0$ is feasible at each vertex of $\mathcal{X}_{\text{nom}}$ then $\mathcal{X}_{\text{nom}} \subseteq \mathbb{X}_N^{\{\text{i}\}}$. A reduction in computation could also be obtained by means of different exploration strategies, however, this is not investigated in this thesis.

6.3 Full regularisation

In this Section we consider the case when all actuators are regularised. The proposed strategy uses a ℓ_{asso}-MPC with cost (3.2.2), where

$$S = \begin{bmatrix} S_{\{\text{i}\}} & 0 \\ 0 & S_{\{\text{ii}\}} \end{bmatrix}, \tag{6.3.1}$$

where the matrix $S_{\{\text{i}\}} \in \mathcal{D}_{++}$ is given, and $S_{\{\text{ii}\}} \in \mathcal{D}_{++}$ has to be computed. Note that, in (6.3.1), the actuators have again been grouped as $u = (u^{\{\text{i}\}},\ u^{\{\text{ii}\}})$. We assume that only the first group of actuators is used to stabilise the system, namely, the terminal cost and terminal constraint are designed under the assumption that $u^{\{\text{ii}\}}(k) = 0, \forall k \in \mathbb{I}_{\geq 0}$.

6.3.1 Existence of the nominal operations zone

The existence of a nominal operations zone can be shown in a similar way as done for partial regularisation in Section 6.2.1. Assume, for the moment, to have a mixed ℓ_2^2/ℓ_p terminal cost, $F(x) = x^T P x + \|Zx\|_p$. Partition $\Phi\Upsilon^T = [\Phi_{\{\text{i}\}},\ \Phi_{\{\text{ii}\}}]$, with Φ defined as in (5.5.2). The following optimality conditions are obtained for (3.2.1):

$$\Phi_{\{\text{i}\}}^T \mathbf{r}^\star + \Theta_1^T \mathbf{Q} \Psi x(k) + \Theta_1^T \mathbf{Q}(\Theta_1 \underline{\mathbf{u}}_{\{\text{i}\}}^\star + \Theta_2 \underline{\mathbf{u}}_{\{\text{ii}\}}^\star) + \mathbf{R}_1 \underline{\mathbf{u}}_{\{\text{i}\}}^\star + \frac{1}{2}\Omega_1^T v^\star \in \frac{1}{2} W_{\{\text{i}\}}^T \partial \|W_{\{\text{i}\}} \underline{\mathbf{u}}_{\{\text{i}\}}^\star\|_1, \quad (6.3.2)$$

$$\Phi_{\{\text{ii}\}}^T \mathbf{r}^\star + \Theta_2^T \mathbf{Q} \Psi x(k) + \Theta_2^T \mathbf{Q}(\Theta_1 \underline{\mathbf{u}}_{\{\text{i}\}}^\star + \Theta_2 \underline{\mathbf{u}}_{\{\text{ii}\}}^\star) + \mathbf{R}_2 \underline{\mathbf{u}}_{\{\text{ii}\}}^\star + \frac{1}{2}\Omega_2^T v^\star \in \frac{1}{2} W_{\{\text{ii}\}}^T \partial \|W_{\{\text{ii}\}} \underline{\mathbf{u}}_{\{\text{ii}\}}^\star\|_1, \quad (6.3.3)$$

where $W_{\{\text{i}\}} = I_{(N\times N)} \otimes S_{\{\text{i}\}}$, $\|\mathbf{r}^\star\|_\infty = 1$. The following result provides sufficient conditions for the existence of a nominal operations zone:

Theorem 6.3.1. Assume (A8). Assume also that $S_{\{\text{ii}\}} \in \mathcal{D}_{++}$, with $\min_j \left\{ S_{\{\text{ii}\}_{jj}} \right\} > 2\|\Phi_{\{\text{ii}\}}^T\|_\infty$. Then, there exists a set $\mathbb{G}$, with $0 \in \text{int}(\mathbb{G})$, such that if $x(k) \in \mathbb{G}$ the ℓ_{asso}-MPC with S given by (6.3.1) gives $\underline{\mathbf{u}}_{\{\text{ii}\}}^\star = 0$.

Proof. In particular, from (6.3.3) the solution $\underline{\mathbf{u}}_{\{\text{ii}\}}^\star = 0$ is obtained if

$$\|\Phi_{\{\text{ii}\}}^T\|_\infty + \|\Theta_2^T \mathbf{Q} \Psi x(k) + (\Theta_2^T \mathbf{Q} \Theta_1) \underline{\mathbf{u}}_{\{\text{i}\}}^\star + 1/2\ \Omega_2^T v^\star(x))\|_\infty \le 1/2 \min_j \left\{ S_{\{\text{ii}\}_{jj}} \right\}. \quad (6.3.4)$$

From (A8), it follows that $\underline{\mathbf{u}}_{\{\text{i}\}}^\star(x)$ and $v^\star(x)$ are unique and PWA [Bemporad *et al.* , 2002a]. Then (6.3.4) describes a set of states, $\widetilde{\mathbb{G}}$, which contains the origin, since at the origin the solution is zero. ∎

Theorem 6.3.1 is useful to understand the dependance between $S_{\{\text{ii}\}}$ and $S_{\{\text{i}\}}$ when using the procedure proposed in the next Section. In particular, recalling Lemma 5.2.2, the matrix Z can be expressed as $Z = \|S_{\{\text{i}\}}\|_p \tilde{Z}$, for some $\tilde{Z}$. Therefore, we can expect $S_{\{\text{ii}\}}$ to be dependent on $S_{\{\text{i}\}}$, when using the proposed procedures. This confirms the intuition that $S_{\{\text{ii}\}}$ should be taken to be greater than $S_{\{\text{i}\}}$.

6.3.2 Formulation as a mpQP

Similar to Section 6.2.2, in this section the considered ℓ_{asso}-MPC problem is formulated as a mpQP that can be solved offline using standard explicit MPC tools [Bemporad *et al.* , 2002a]. Once again, the preferred and auxiliary actuators predictions are, for convenience, grouped into two separate blocks. Similarly to Section 6.2.4, the formulation proposed in this section is then used in Section 6.3.3, where $S_{\{\mathrm{ii}\}}$ is computed by solving the considered ℓ_{asso}-MPC subject to $\underline{\mathrm{u}}_{\{\mathrm{ii}\}} = 0$.

6.3.2.1 Using slack variables

Assuming to have a mixed ℓ_2^2/ℓ_1 terminal cost, $F(x) = x^T P x + \|Zx\|_1$, problem (3.2.1–3.2.2) can be formulated as

$$\begin{aligned} &\chi^\star = \arg\min_\chi \tfrac{1}{2}\, \chi^T \bar{H} \chi + \chi^T \bar{G}(x) \\ &\text{s.t. } \bar{\Omega}\, \chi \leq b(x),\ x = x(k), \end{aligned} \tag{6.3.5}$$

where

$$\chi = \begin{bmatrix} \underline{\mathrm{u}}_{\{\mathrm{i}\}} \\ \underline{\mathrm{u}}_{\{\mathrm{ii}\}} \\ \sigma \\ \sigma_t \end{bmatrix}, \quad \bar{H} = \begin{bmatrix} 2\Upsilon H \Upsilon^T & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \bar{G}(x) = \begin{bmatrix} 2\Upsilon\Gamma x \\ \underline{1} \\ \underline{1} \end{bmatrix}, \tag{6.3.6}$$

with matrices

$$\bar{\Omega} = \begin{bmatrix} \Omega\Upsilon^T \quad 0 \quad 0 \\ \Pi \end{bmatrix}, \quad b(x) = \begin{bmatrix} \underline{1} - Mx \\ 0_{(2mN\times 1)} \\ -ZA^N x \\ ZA^N x \end{bmatrix}, \tag{6.3.7}$$

$$\Pi = \begin{bmatrix} W\Upsilon^T & -I_{mN} & 0 \\ -W\Upsilon^T & -I_{mN} & 0 \\ \Phi & 0 & -I_n \\ -\Phi & 0 & -I_n \end{bmatrix}. \tag{6.3.8}$$

At each time k, the signal applied to the plant is given by $u(k) = K_N(x(k))$, with

$$K_N(x(k)) = [I\ 0\ \ldots\ 0]\Upsilon^T [I\ 0\ 0]\chi^\star. \tag{6.3.9}$$

The number of slack variables depends upon the type of terminal cost. For an ℓ_2^2/ℓ_∞ function, σ_t is of size n. On the other hand, for an ℓ_2^2/ℓ_∞ terminal cost the number of slack variables is reduced as σ_t is a scalar.

6.3.2.2 Splitting variables for strict convexity

Assume a Minkowski terminal cost, $F(x) = \alpha(\psi_{\mathbb{X}_f}(x))^2 + \beta\psi_{\mathbb{X}_f}(x)$, with $\alpha, \beta > 0$, and $\mathbb{X}_f = \{x \mid Zx \leq g\underline{1}\}$, for some $g > 0$. The regularised variables are again split into positive and negative part, namely, $u_+,\ u_-$ with $u = u_+ - u_-$, with $u_+ \geq 0\ u_- \geq 0$, $R_1 \in \mathcal{D}_{++}$, $R_2 \in \mathcal{D}_{++}$, $S_{\{\text{i}\}} \in \mathcal{D}_{++}$ and $S_{\{\text{ii}\}} \in \mathcal{D}_{++}$. The following is a condensed formulation of (3.2.1–3.2.2):

$$
\begin{aligned}
&\chi^\star = \arg\min_\chi \tfrac{1}{2}\, \chi^T \bar{H} \chi + \chi^T \bar{G}(x) \\
&\text{s.t. } \bar{\Omega}\, \chi \leq b(x),\ x = x(k),
\end{aligned}
\tag{6.3.10}
$$

where

$$
\chi = \begin{bmatrix} \zeta \\ \underline{\mathrm{u}}_{\{\text{i}\}+} \\ \underline{\mathrm{u}}_{\{\text{i}\}-} \\ \underline{\mathrm{u}}_{\{\text{ii}\}+} \\ \underline{\mathrm{u}}_{\{\text{ii}\}-} \end{bmatrix}, \quad \bar{G}(x) = \begin{bmatrix} 0 \\ 2\bar{\Gamma}x \end{bmatrix} + d,\ d = \left[\begin{array}{c} d_{\{\text{i}\}} \\ \hline d_{\{\text{ii}\}} \end{array}\right] = \left[\begin{array}{c} \beta \\ W_{\{\text{i}\}}\underline{1} \\ W_{\{\text{i}\}}\underline{1} \\ \hline W_{\{\text{ii}\}}\underline{1} \\ W_{\{\text{ii}\}}\underline{1} \end{array}\right], \tag{6.3.11}
$$

and with

$$
\begin{aligned}
&\bar{H} = 2\,\text{BlockDiag}(\alpha,\ (\bar{\Theta}^T \mathbf{Q} \bar{\Theta} + \bar{\mathbf{R}})), \quad \bar{\Gamma} = \bar{\Theta}^T \mathbf{Q} \Psi, \qquad \bar{\Theta} = [\Theta_1,\ -\Theta_1,\ \Theta_2,\ -\Theta_1], \\
&\bar{\mathbf{R}} = \text{BlockDiag}(I_N \otimes R_1,\ I_N \otimes R_1,\ I_N \otimes R_2,\ I_N \otimes R_2).
\end{aligned}
\tag{6.3.12}
$$

Partition $\Phi = [\Phi_{\{\text{i}\}},\ \Phi_{\{\text{ii}\}}]$, with Φ defined as in (5.5.2). The constraint matrices for (6.3.10) are

$$
\bar{\Omega} = \begin{bmatrix} \mho \\ \Pi \end{bmatrix}, \quad b(x) = \begin{bmatrix} \bar{b}(x) \\ 0_{(2mN \times 1)} \end{bmatrix}, \quad \bar{b}(x) = \begin{bmatrix} g\underline{1} - ZA^N x \\ \underline{1} - Mx \end{bmatrix}, \tag{6.3.13}
$$

$$
\Pi = \begin{bmatrix} 0 & -I_{2mN} \end{bmatrix}, \qquad \mho = \begin{bmatrix} -\underline{1} & \Phi_{\{\text{i}\}} & -\Phi_{\{\text{i}\}} & \Phi_{\{\text{ii}\}} & -\Phi_{\{\text{ii}\}} \\ 0 & \Omega_1 & -\Omega_1 & \Omega_2 & -\Omega_2 \end{bmatrix}, \tag{6.3.14}
$$

where $\Omega\Upsilon^T = [\Omega_1,\ \Omega_2]$. At each time k, the signal applied to the plant is given by $u(k) = K_N(x(k))$, with

$$K_N(x(k)) = [I\ 0\ \dots\ 0]\Upsilon^T \text{BlockDiag}([0,\ I,\ -I],\ [I,\ -I])\chi^\star. \tag{6.3.15}$$

6.3.3 Computation of the regularisation penalty

The exact penalty is computed using the actuator splitting method, similarly to Section 6.2.4.2. For a given x, the KKT conditions for (6.3.10) are:

$$\begin{aligned} \bar{H}\,\chi + \bar{G}(x) + \mho^T\nu &= -\Pi\gamma \\ \mho\,\chi &= \bar{b}(x) - s^\star \\ \Pi\chi &= 0 \\ (s^\star)^T\nu &= 0 \\ s^\star &\geq 0 \\ \nu &\geq 0 \\ \gamma &\geq 0. \end{aligned} \tag{6.3.16}$$

Theorem 6.3.2. Assume $S_{\{\text{i}\}} = 0$, $S_{\{\text{ii}\}} = \text{diag}\{s_i\}$. Take $\min_i s_i \geq \gamma^\star$, with $\gamma^\star$ solving (6.2.23). Then, for the ℓ_{asso}-MPC solving (6.2.9) or, equivalently, (6.2.13) it follows that $u^{\{\text{ii}\}} = 0,\ \forall x \in \mathbb{X}_N^{\{\text{i}\}} \cap \mathcal{X}_{\text{nom}}$.

Proof. The theorem is a simple application of the previous results. The ℓ_{asso}-MPC is feasible under the constraint $\|\underline{\mathrm{u}}_{\{\text{ii}\}}\|_1 = 0,\ \forall x \in \mathbb{X}_N^{\{\text{i}\}} \cap \mathcal{X}_{\text{nom}}$. From Lemma 6.2.4, $\gamma^\star$ is the lower bound for the exact penalty on $\|\underline{\mathrm{u}}_{\{\text{ii}\}}\|_1$, namely $\min_i s_i \geq \gamma^\star$, provides problem (6.2.9) the same solution as (6.2.20), $\forall x \in \mathbb{X}_N^{\{\text{i}\}} \cap \mathcal{X}_{\text{nom}}$. ∎

Theorem 6.3.3. Consider (6.3.10). The exact penalty weight for $\|\underline{\mathrm{u}}_{\{\text{ii}\}}\|_1 = \sum_{j=1}^{N-1} \|u_j^{\{\text{ii}\}}\|_1$, providing $\underline{\mathrm{u}}_{\{\text{ii}\}} = 0, \forall x \in \mathbb{X}_N^{\{\text{i}\}} \cap \mathcal{X}_{\text{nom}}$, is:

$$\begin{aligned} \gamma^\star = \max_{\chi,\nu,\gamma,x}\ & \|\gamma\|_\infty \\ \text{s.t.}\ \ & (6.3.16),\ x \in \mathbb{X} \cap \mathcal{X}_{\text{nom}}. \end{aligned} \tag{6.3.17}$$

Proof. Since the ∞-norm is the dual of the 1-norm, the theory of exact penalty functions can be applied, by computing the maximum ∞-norm of the Lagrange multipliers associated with the constraint $\|\underline{\mathrm{u}}_{\{\text{ii}\}}\|_1 = 0$. This constraint corresponds to $\Pi\chi = 0$. ∎

Theorem 6.3.4. Assume $S_{\{\text{ii}\}} = \text{diag}\{s_i\}$. Take $\min_i s_i \geq \gamma^\star$, with $\gamma^\star$ solving (6.3.17). Then, for the ℓ_{asso}-MPC solving (6.3.10), it follows that $u^{\{\text{ii}\}} = 0,\ \forall x \in \mathbb{X}_N^{\{\text{i}\}} \cap \mathcal{X}_{\text{nom}}$.

Proof. Proof is the same as for Theorem 6.2.5, considering Theorem 6.3.3 instead of Theorem 6.2.4. ■

We proceed similarly to Section 6.2.4. Conditions (6.3.16) provide the solution of the ℓ_{asso}-MPC with only $u^{\{\text{i}\}}$ available, namely, of the *pre-existing MPC*. Therefore, we can solve this smaller problem and obtain the necessary active sets and critical regions. This problem can be solved, for instance, using YALMIP [Lofberg, 2004]. Then, the critical region and active sets can be used to compute the solution of (6.3.10), subject to $(\underline{\mathrm{u}}_{\{\text{ii}\}+}, \underline{\mathrm{u}}_{\{\text{ii}\}-}) = 0$ and the required Lagrange multipliers in the region of interest $\mathcal{X}_{\text{nom}}$. This again returns a partition of regions $\mathcal{P}_j$. For each region, denote the matrix of coefficients for the active inequalities as

$$\Omega_{\text{act}} = \begin{bmatrix} \tilde{\mho} \\ \Pi \end{bmatrix}, \tag{6.3.18}$$

where Π is defined in (6.3.13). The PWA functions, now including the extra affine component d, are computed as follows:

$$v^\star(x) = -(\Omega_{\text{act}}\bar{H}^{-1}\Omega_{\text{act}}^T)^{-1}\left(\left(2\Omega_{\text{act}}\bar{H}^{-1}\begin{bmatrix} 0 \\ \bar{\Gamma} \end{bmatrix} - M_{\text{act}}\right)x + b_{\text{act}} + \Omega_{\text{act}}\bar{H}^{-1}\begin{bmatrix} d_{\{\text{i}\}} \\ 0 \end{bmatrix}\right) \tag{6.3.19}$$

$$\chi^\star(x) = (\zeta^\star, \underline{\mathrm{u}}^\star_{\{\text{i}\}+}, \underline{\mathrm{u}}^\star_{\{\text{i}\}-}, 0, 0) = -\bar{H}^{-1}\left(2\begin{bmatrix} 0 \\ \bar{\Gamma} \end{bmatrix}x + \begin{bmatrix} d_{\{\text{i}\}} \\ 0 \end{bmatrix}\right) - \frac{1}{2}\Omega_{\text{act}}^T v^\star(x) \tag{6.3.20}$$

where Π is defined in (6.3.13) and $\bar{H},\ \bar{\Gamma},\ d$ are from (6.3.12). The constant term b_{act} and the gain M_{act} are extrapolated from the rows of $b(x)$ which is also defined in (6.3.13), according to the indices of the active set and of Π. The following result is obtained:

Theorem 6.3.5. Consider the partition $\textbf{part}\mathcal{P} = \{\mathcal{P}_j\}_{j=1}^{n_p}$ of the feasible region of (6.3.10), subject to $(\underline{\mathrm{u}}_{\{\text{ii}\}+}, \underline{\mathrm{u}}_{\{\text{ii}\}-}) = 0$, obtained from an mpQP solver with each $\mathcal{P}_j$ corresponding to an optimal active set of the *pre-existing MPC*. Then, the regularisation penalty $\gamma^\star$, satisfying Theorem 6.3.3, is given by the solution of (6.2.32), where $\bar{\gamma}_j$ is computed by either (6.2.31) or (6.2.27), subject to (6.3.19), (6.3.20), and (6.3.18).

Proof. The theorem can be verified by elaborating (6.3.16), as for Theorem 6.2.6. ■

6.3.3.1 Algorithms

The required steps to compute $S_{\{ii\}}$ are summarised in Algorithm 5 (see the next page). The same considerations as in Section 6.2.4.3 apply here, for what concerns the case of no common constraints between preferred and auxiliary actuators, for which potential complexity reduction can be achieved.

Algorithm 5 Computation of $S_{\{ii\}}$ (full regularisation)

1. Take any $x \in \mathcal{X}_{\text{nom}} \cap \mathbb{X}_N^{\{i\}}$
2. Identify the active set $\mathcal{A}_j = \mathcal{A}(x)$ by solving (6.3.10) at x subject to $(\underline{\mathrm{u}}_{\{ii\}+}, \underline{\mathrm{u}}_{\{ii\}-}) = 0$
3. Evaluate (6.3.18), (6.3.19), and (6.3.20)
4. Compute $\mathcal{P}_j$ substituting (6.3.20) into $\bar{\Omega}\ \chi \leq b(x)$
5. Compute $\bar{\gamma}_j$ by solving the LP (6.2.31)
6. **IF** $(\mathcal{X}_{\text{nom}} \cap \mathbb{X}_N^{\{i\}}) \backslash \mathcal{P}_j = \emptyset$ **THEN**
   (a) Compute $\gamma^\star \max_j \bar{\gamma}_j$.
   (b) $S_{\{ii\}} \leftarrow \gamma^\star I$
7. **ELSE**
   (a) Identify $\mathcal{A}_{j+1}$, set $j \leftarrow j+1$ [Bemporad *et al.* , 2002a; Tondel *et al.* , 2003]
   (b) **GOTO** point 3.

Similar to Section 6.2.4.3, for Algorithm 5 it is possible to determine wether a given $\mathcal{X}_{\text{nom}}$ is contained in $\mathbb{X}_N^{\{i\}}$ by starting the search from the vertices of $\mathcal{X}_{\text{nom}}$.

6.4 Example 1: Computation of the exact penalty

The proposed methodology is demonstrated on the following LTI system:

$$
\begin{aligned}
A &= \begin{bmatrix} 0.2 & 0.1 \\ 0 & 1.1 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \\
\mathbb{I}^{\{\mathrm{i}\}} &= \{2\}, \; Q = I, \; R = I, \; S_{\{\mathrm{i}\}} = 0, \; N = 3, \\
\mathbb{X} &= \{x \mid \|x\|_\infty \leq 20\}, \; \mathbb{U} = \{u \mid \|u\|_\infty \leq 5\}.
\end{aligned} \tag{6.4.1}
$$

The terminal controller, $K^T = [0, \; K_{\{i\}}^T]$, places the closed loop eigenvalues at $0.2, \; 0.9$. We use partial regularisation. The matrix P solves the closed loop Discrete-time Lyapunov equation. All plots are made using the Matlab MPT toolbox [Kvasnica *et al.* , 2004]. Simulations are made using Yalmip for Matlab [Lofberg, 2004]. The control law for the preferred actuator, in Figure 6.1, stabilises the system and resembles the one of a standard LQ MPC. The explicit solution of the pre-existing LQ MPC is computed using the MPT toolbox [Kvasnica *et al.* , 2004], then Algorithm 4 is used to compute the regularisation penalty $S_{\{\mathrm{ii}\}}$.

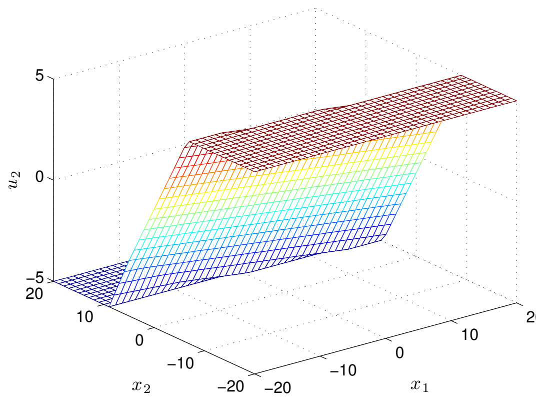

Figure 6.1: Explicit control law. Non-regularised actuator. Constrained case.

Figures 6.2 and 6.3 show the control law for the auxiliary actuator resulting from the procedure, with $\mathcal{X}_{nom} = \{x \mid \pm x \leq \bar{x}\}$ (the dashed box) for, respectively, $\bar{x} = [5 \; 5]^T$ and $\bar{x} = [10 \; 8]^T$. In this case the algorithm provides, respectively, $S_{\{\mathrm{ii}\}} = 7.75$ and $S_{\{\mathrm{ii}\}} = 13.07$. The only critical region explored for $\bar{x} = [5 \; 5]^T$ is the one around the origin, where no constraints are active. The closed-loop trajectory is also shown from four different initial conditions to demonstrate the origin's asymptotic stability. Figure 6.4 shows the results for $\bar{x} = [10\ 15]^T$. In the this case, as well as for $\bar{x} = [10\ 8]^T$, the algorithm explores several critical regions. An approximation is used, in particular the PWA functions are computed by using only the positive identity matrix to represent the redundant actuator constraint. This provides $S_{\{\text{ii}\}} = 14.81$. Despite this approximation, the results are satisfactory, since $u^{\{\text{ii}\}} = 0\ \forall x \in \mathcal{X}_{nom}$.

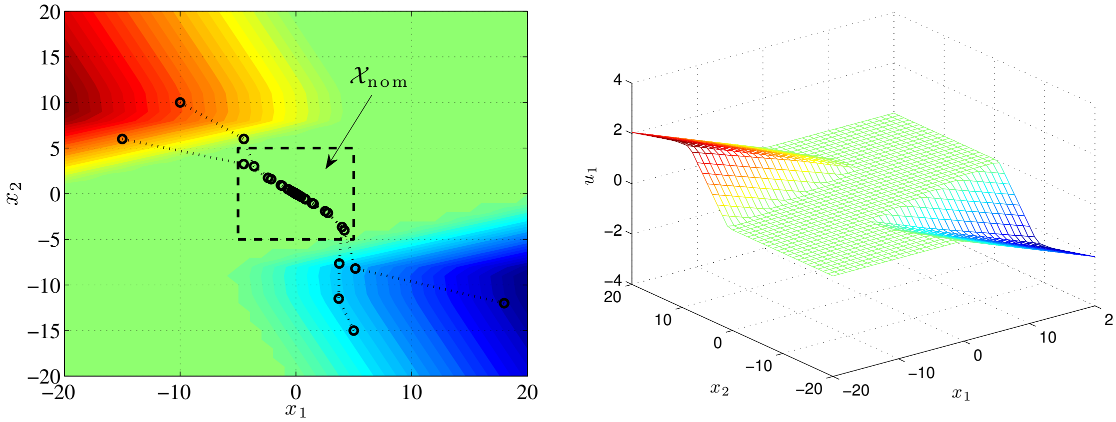

Figure 6.2: Explicit control law. Regularised actuator. Constrained case ($\bar{x} = [5\ 5]^T$, $S_{\{\text{ii}\}} = 7.75$).

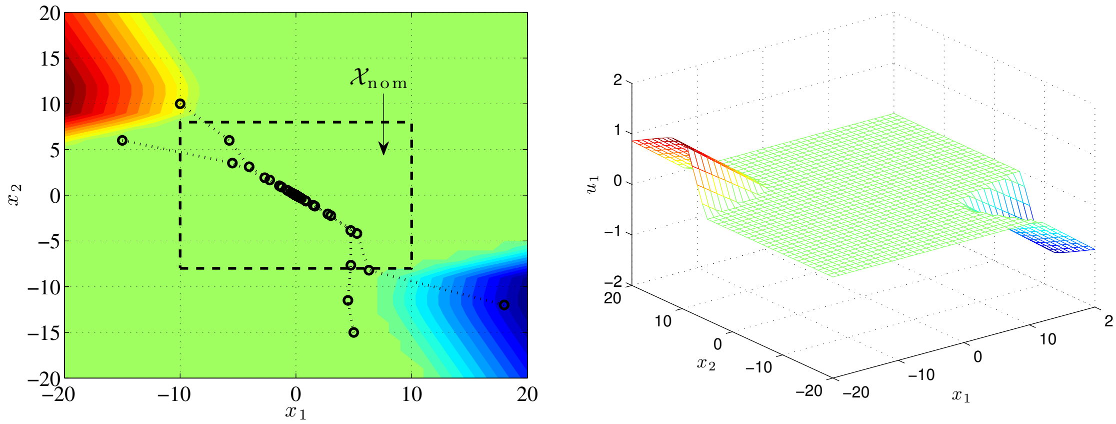

Figure 6.3: Explicit control law. Regularised actuator. Constrained case ($\bar{x} = [10\ 8]^T$, $S_{\{\text{ii}\}} = 13.07$).

Figure 6.5 shows the result for $\bar{x} = [10\ 5]^T$ in the considered case (left), as well as the control law for $u^{\{\text{ii}\}}$ in the unconstrained case with the same $S_{\{\text{ii}\}} = 9.43$ (right). In particular, it can be noticed that in the latter case we have $u^{\{\text{ii}\}} = 0\ \forall x \in \mathcal{X}_{unc}$, where $\mathcal{X}_{unc} \supseteq \mathcal{X}_{nom}$. This can be expected since $S_{\{\mathrm{ii}\}}$ is computed for the constrained case. If on the other hand, the matrix $S_{\{\mathrm{ii}\}}$ is computed for the unconstrained case, in the constrained case we might not have $u^{\{\mathrm{ii}\}} = 0 \; \forall x \in \mathcal{X}_{nom}$ if some constraints are active in $\mathcal{X}_{nom}$. In this case $S_{\{\mathrm{ii}\}}$ could be smaller than the lower bound $\gamma^\star$ required by Theorem 6.2.4.

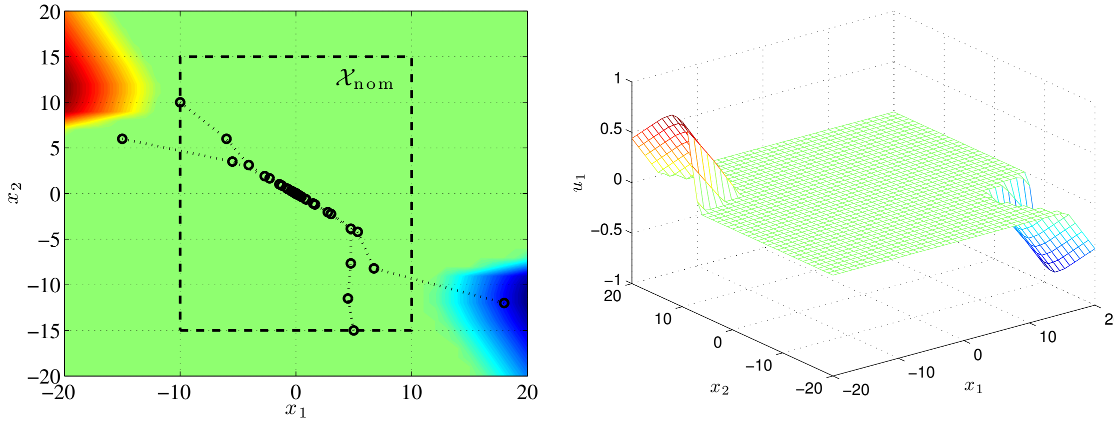

Figure 6.4: Explicit control law. Regularised actuator. Constrained case ($\bar{x} = [10\ 15]^T$, $S_{\{\mathrm{ii}\}} = 14.81$).

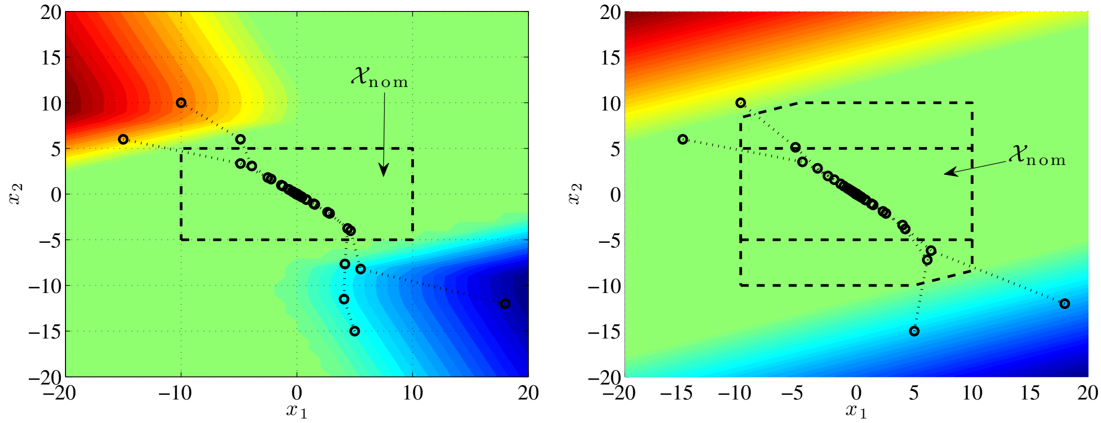

Figure 6.5: Explicit control law. Regularised actuator ($\bar{x} = [10\ 5]^T$, $S_{\{\mathrm{ii}\}} = 9.43$). Constrained (left), unconstrained (right).

Figure 6.6 shows the closed loop trajectories from the constrained case from different initial conditions with $\bar{x} = [10\ 5]^T$, $S_{\{\mathrm{ii}\}} = 9.43$, confirming that the system origin is asymptotically stable with $u^{\{\mathrm{ii}\}} = 0 \; \forall x \in \mathcal{X}_{nom}$.

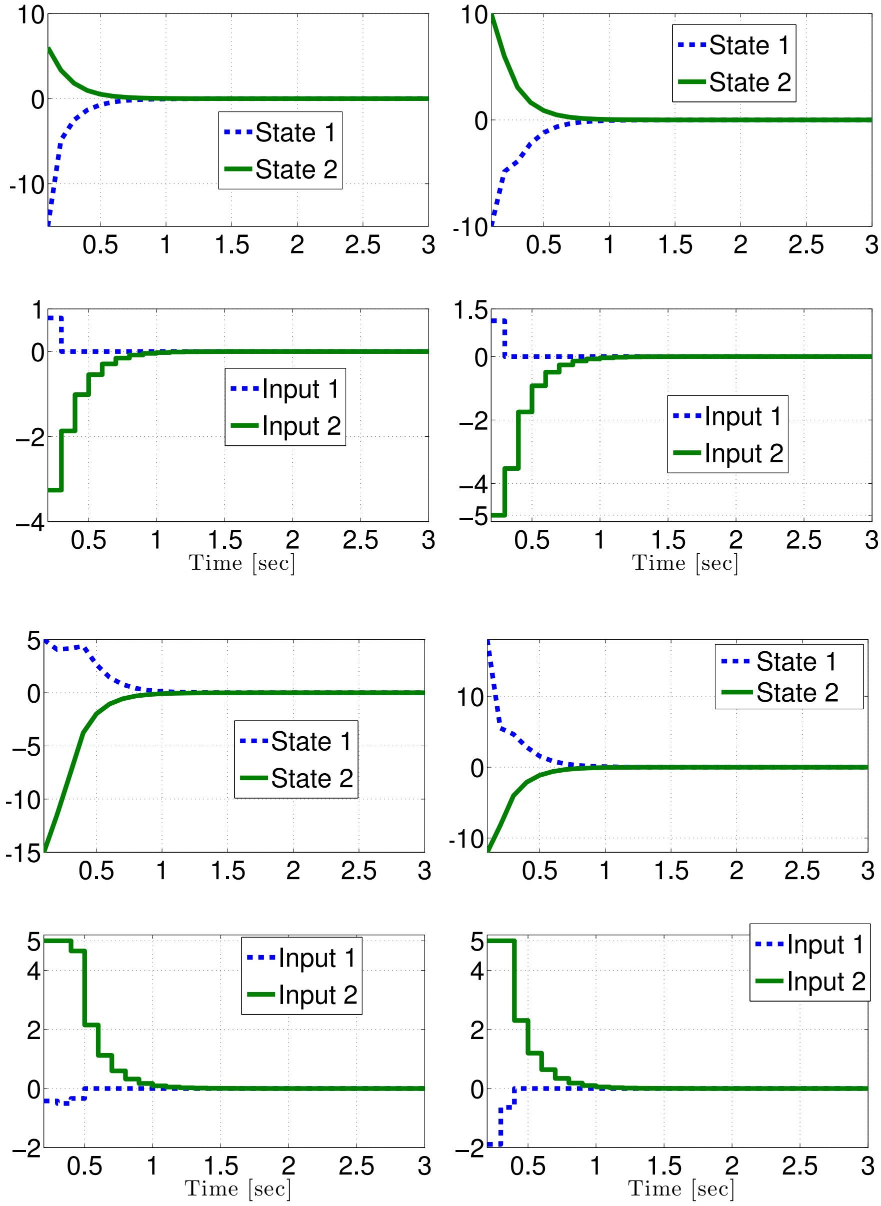

Figure 6.6: I/O trajectories for different initial conditions ($\bar{x} = [10\ 5]^T$, $S_{\{\mathrm{ii}\}} = 9.43$).

Results are summarised in Table 6.1, where T_{MPT} is the time (seconds) required by the MPT to obtain the explicit solution for the pre-existing MPC with 47 regions ($N_r^{\{\mathrm{i}\}}$) N_s is the number of regions where Algorithm 4 computes the Lagrange multipliers, and T_{search} is the execution time (seconds) for Algorithm 4, averaged over 20 simulations. Notice that Algorithm 4 is very inefficient

Table 6.1: Computation complexity using Algorithm 4 and the explicit solution from the MPT toolbox [Kvasnica *et al.* , 2004]

$\bar{x}$	$S_{\{\text{ii}\}}$	$N_s/N_r^{\{\text{i}\}}$	T_{MPT} $[sec]$	T_{search} $[sec]$
$[5\ 5]^T$	7.75	1/47	0.84	1.02
$[10\ 5]^T$	9.43	1/47	0.86	1.06
$[10\ 8]^T$	13.07	3/47	0.80	1.87
$[10\ 15]^T$	14.81	5/47	0.82	2.75

as it requires more time than the computation of the entire explicit solution of the (smaller) pre-existing MPC performed by the MPT toolbox. As discussed in Section 6.2.4.3, Algorithm 4 could be substantially improved, for instance considering Algorithm 4 of [Borrelli *et al.* , 2008]. Further improvements could be obtained by directly implementing Algorithm 3, for which existing routines from the MPT toolbox could be used.

6.5 Example 2: Limits of the technique

Consider the LTI system with the following matrices:

$$\begin{aligned} A &= \begin{bmatrix} 0.8 & 0 \\ -0.6 & 1.1 \end{bmatrix},\ B = \begin{bmatrix} 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \end{bmatrix}, \\ \mathbb{I}^{\{\text{i}\}} &= \{1,2,3\},\ Q = I,\ R = I,\ S_{\{\text{i}\}} = 0,\ N = 3, \\ \mathbb{X} &= \{x \mid \|x\|_\infty \leq 100\},\ \mathbb{U} = \{u \mid \|u\|_\infty \leq 1\}, \end{aligned} \tag{6.5.1}$$

for which the proposed algorithm gives $\gamma^\star = 44 \cdot I$. A scalar, $\eta \geq 1$, is used to refine the tuning, namely, $S_{\{\text{ii}\}} = \eta\gamma^\star$. Figure 6.7 shows the phase plot of the system and the sets of interest, as discussed in Section 6.2.3. Figure 6.8 shows the I/0 behaviour for $\eta = 1$, and $\eta = 10$. In particular, in the first case input 4 is used when in $\tilde{\mathbb{X}}_N^{\{\text{i}\}}$ (first move), while for $\eta = 10$ input 4 is never used. As mentioned in Section 6.2.3, this cannot be guaranteed a priori when $x \notin \tilde{\mathbb{X}}_N^{\{\text{i}\}}$, even though it is achieved for this simple example. The advantage of using ℓ_{asso}-MPC over LQ-MPC is evident, since the actuators behaviour in Figure 6.8 cannot be achieved using only quadratic penalties.

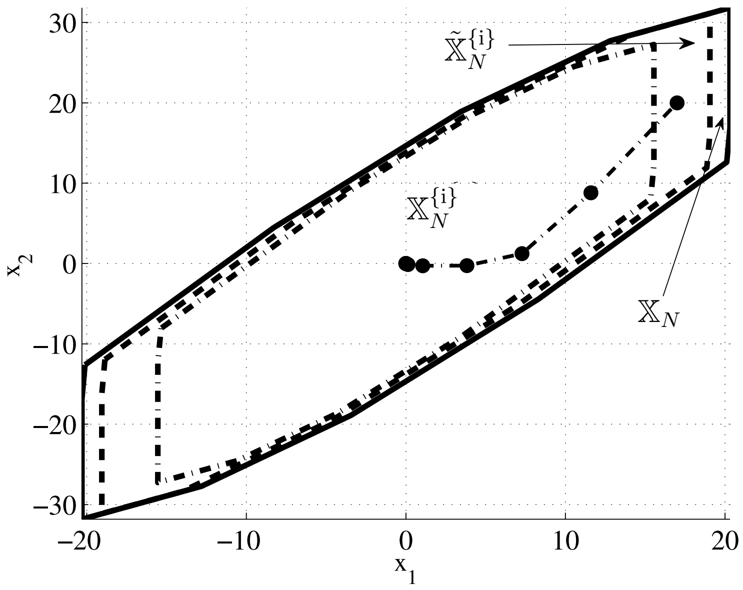

Figure 6.7: Example 2: Sets, $\mathbb{X}_N$ (solid), $\tilde{\mathbb{X}}_N^{\{i\}}$ (dashed), $\mathbb{X}_N^{\{i\}}$ (dash-dot), and trajectory for $\eta = 1$.

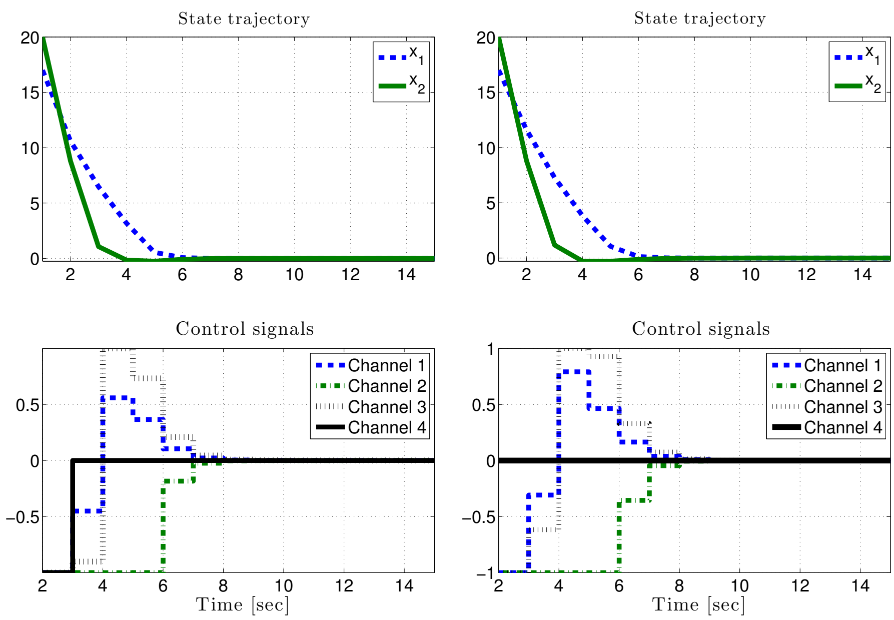

Figure 6.8: Example 2: I/O trajectories (left: $\eta = 1$, right: $\eta = 10$).

6.6 Example 3: Multi-source distribution

A distribution-like problem is studied to demonstrate the proposed methodology. Consider a network consisting of 2 servers and 4 clients. Each server is assigned to a pair of clients, but it can also serve the other 2 if necessary. This is effectively a multi-objective control problem, where the main control objective is to provide data to the assigned clients. The second objective is to assist the other server when its part of the line is experiencing particularly "high" demand. The ℓ_{asso}-MPC is used to regulate the tradeoff between the two control objectives, and the regularisation parameter will be used to quantify how high the demand has to be for one server to help the other one. The network structure is shown in Figure 6.9, where "S" denotes the servers and "c" the clients.

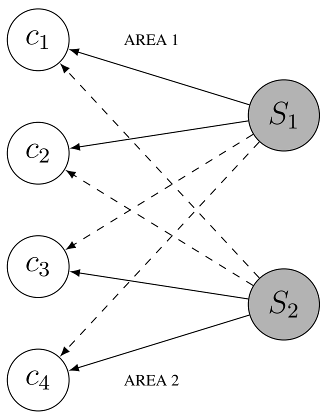

Figure 6.9: Network example. Main links (solid), auxiliary links (dashed).

The dynamics of the network are assumed to be the discrete-time equivalent of the composition of the links dynamics. The single link is assumed to be described by

$$G(s) = \frac{s+1}{s(s^2+1.2s+1)}. \tag{6.6.1}$$

The network links are grouped as

$$\mathcal{G}(z) = \left[\begin{array}{c|c} \mathrm{L}_1^1(z) & \mathrm{L}_2^1(z) \\ \mathrm{L}_2^2(z) & \mathrm{L}_1^2(z) \end{array}\right] = \left[I_p \otimes G(z) \mid I_p \otimes G(z)\right], \tag{6.6.2}$$

where p is the number of clients, and L_1^2 represents the two links from the server 1 to the second pair of clients. The blocks L_1^1 and L_2^2 are chosen to have higher priority. For $G(z)$, the following state-space realisation is used ($T_s = 0.5$ sec.):

$$A = \begin{bmatrix} 2.3647 & -1.9135 & 0.5488 \\ 1.0000 & 0 & 0 \\ 0 & 1.0000 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0.1191 & 0.0426 & -0.0696 \end{bmatrix}, \; D = 0. \tag{6.6.3}$$

The full state-space model of the system has 12 states, 4 outputs, and 8 positive-valued inputs. The inputs corresponding to links starting from the same server must satisfy a joint ℓ_1 constraint representing the server throughput (50 units/sample).

The ℓ_{asso}-MPC with partial regularisation is used. Since each link dynamics has integral action, the steady-state input reference is zero. Given a reference output series, a target calculator is used online to compute the current reference state vector, x_T. No state constraints are considered, and the state x and $\mathcal{O}_\infty$ can be simply shifted of $-x_T$, since the plant has internal integrators ($u_T = 0$). Full state measurements are assumed. The blocks of the matrix Q contain $C^T C$ (output regulation), $N = 10$, and $R = 100 \cdot I$, for which the associated LQR controller has provided and acceptably large $\mathcal{O}_\infty^{\{i\}}$. A scaled box-shaped set $\mathcal{X}_{\text{nom}} = \{x \mid \|x\|_\infty \leq c\}$, is used for computing $S_{\{ii\}}$ that will provide the primary links to be preferred when $x \in \mathcal{X}_{\text{nom}}$. For a more specific design, the set $\mathcal{X}_{\text{nom}}$ could also be taken as the intersection of the constraint $\|Cx\|_\infty \leq c$ with a box. Since the system is large, computing the explicit MPC solution can be computationally demanding. To reduce the computational burden we approximate $\gamma^\star$ by using only the unconstrained MPC solution.

The simulation scenario is the following. Initially, all clients demand 300 service-units. At time 60, the demand of client 2 goes up to 1000 units. At time 70, client 3 also increases its demand to 1000. Figure 6.10–6.11 shows the behaviour of links 1 and 2 for $c = 65000$ $S_{\{ii\}} = 15152 \cdot I$. In this case, server 2 does not transmit to clients 1 and 2, according to its priority.

The behaviour of links 1,2, for $c = 55000$, $S_{\{ii\}} = 12928 \cdot I$, is shown in Figure 6.12–6.13. When idle, server 2 helps server 1 (only for one sample at time 60), when the demand increases. On the other hand, at time 70, server 1 doesn't help server 2.

Figure 6.14 shows the cumulative sum of the input signals from the two servers for $c = 65000$ (left) and $c = 55000$ (right).

The priorities given to the main links are directly proportional to $S_{\{ii\}}$. Simulations have shown that increasing $S_{\{ii\}}$ can reduce the size of the impulse given by server 2 at time 60. On the other hand, decreasing $S_{\{ii\}}$ further can cause server 2 to be used for longer than a single sample.

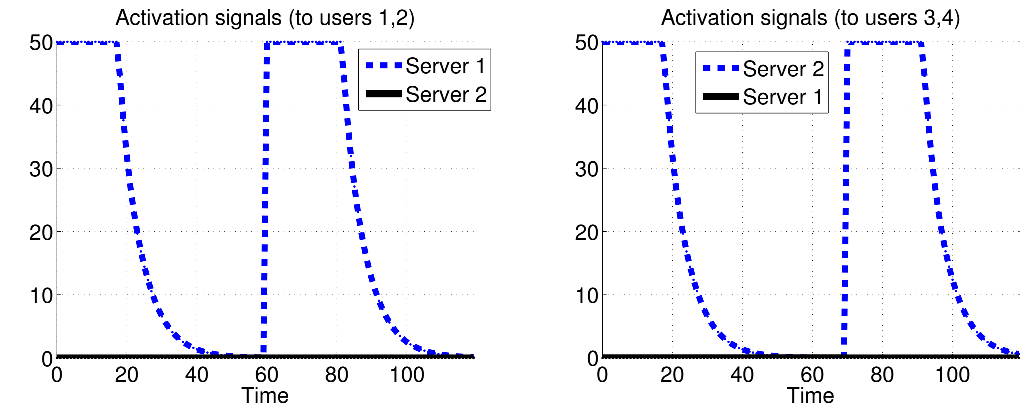

Figure 6.10: Distribution problem: Input trajectories, $S_{\{ii\}} = 15152 \cdot I$.

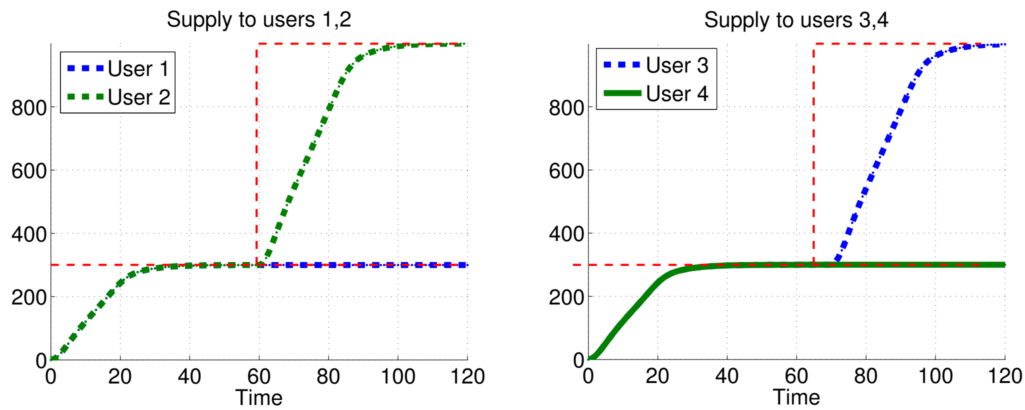

Figure 6.11: Distribution problem: Output trajectories, $S_{\{ii\}} = 15152 \cdot I$.

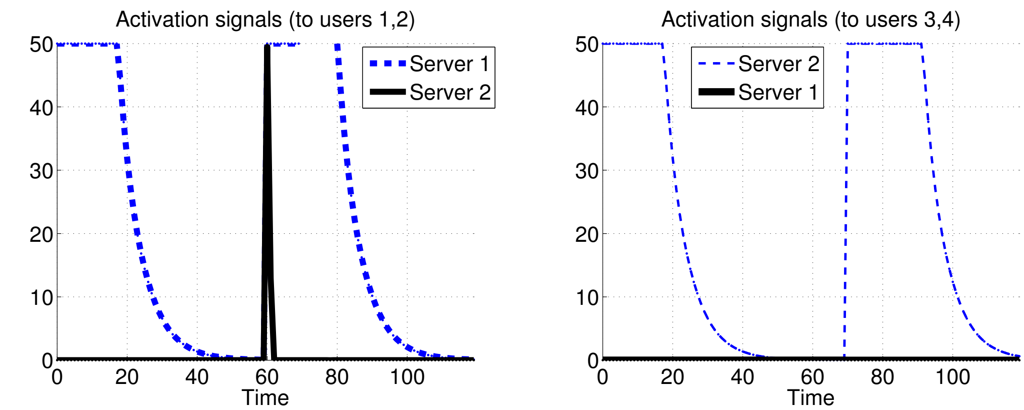

Figure 6.12: Distribution problem: Input trajectories, $S_{\{ii\}} = 12928 \cdot I$.

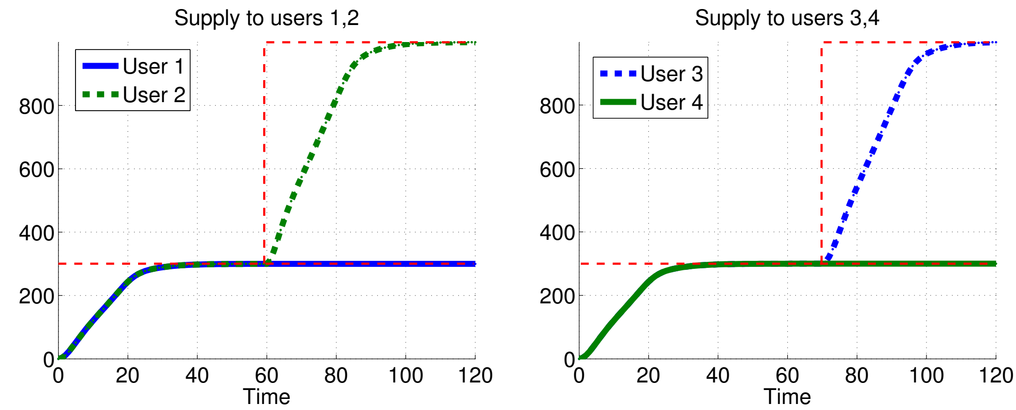

Figure 6.13: Distribution problem: Output trajectories, $S_{\{\mathrm{ii}\}} = 12928 \cdot I$ (80% of original penalty).

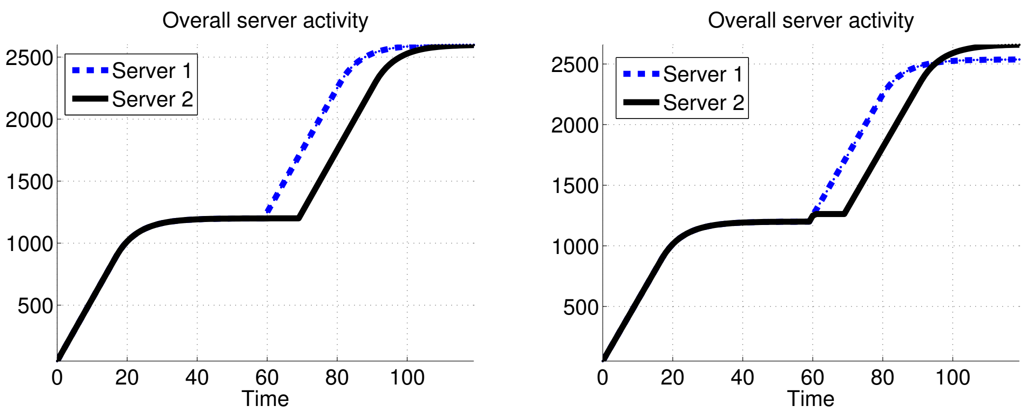

Figure 6.14: Distribution problem: Overall server activity, $S_{\{\mathrm{ii}\}} = 15152 \cdot I$ (left) $S_{\{\mathrm{ii}\}} = 12928 \cdot I$ (right).

Similarly, server 1 can be forced to help server 2 at time 70. Several behaviours are possible, depending on the demands and on the size of c (or $S_{\{\mathrm{ii}\}}$).

6.7 Example 4: Aircraft roll control

This example demonstrates the use of ℓ_{asso}-MPC for the control of the lateral dynamics of a linearised aircraft model at different operating points. A 6-DOF model of a Boeing 747 [Edwards *et al.* , 2010], in Figure 6.15, is linearised around different altitudes (z $[m]$) and values of true airspeed (V_{TAS} $[m/s]$). The objective of the control design is to track a roll rate command from the pilot, p_{ref}, while keeping the magnitude of yaw rate and sideslip angle moderate. A ℓ_{asso}-MPC controller is designed using partial regularisation. The preferred actuators to perform the task are the ailerons (4 in total) and the upper and lower rudders. A nominal quadratic MPC is designed for these actuators taking into account their input magnitude and rate constraints. Then, the spoilers (12 in total) are introduced in the design as auxiliary actuators, with the aim of helping out the ailerons when they are close to saturation or stall, at the expense of additional drag. The main variables

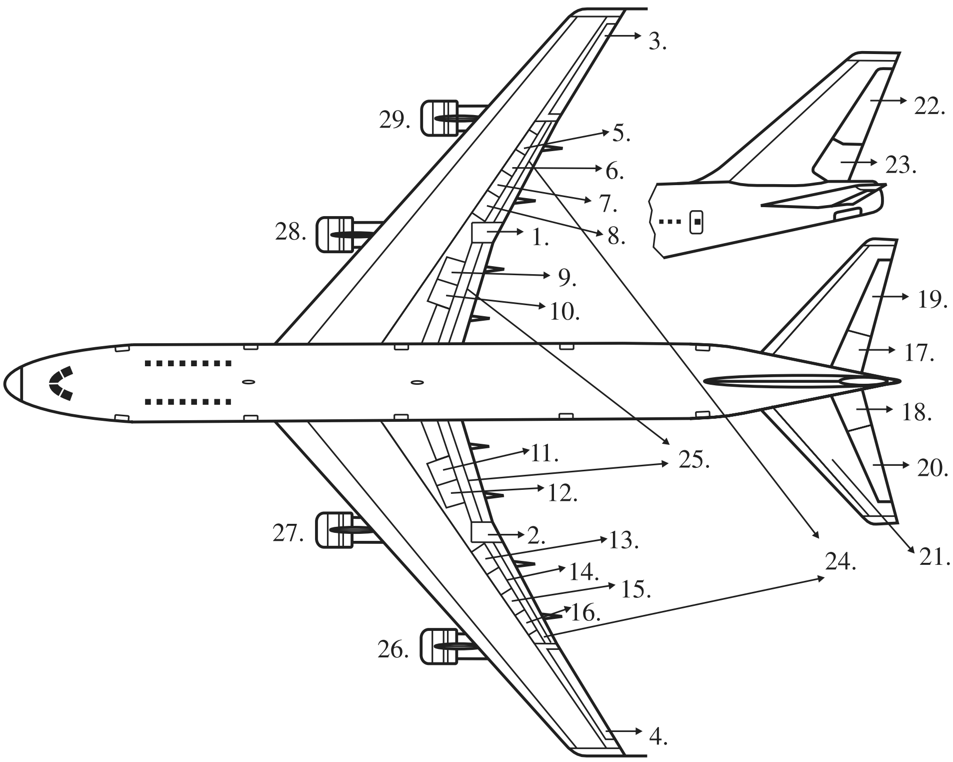

Figure 6.15: Control inputs available on a Boeing 747. Based on [Edwards *et al.* , 2010].

are summarised in Table 6.2, together with the cost function parameters, labelled ℓ_2^2 and ℓ_1 penalty, which have been tuned by trial and improvement for the linearisation at $z = 5000,\ V_{TAS} = 180$.

For this operating point, the ℓ_1 penalty has been tuned in order to have the spoilers at rest when the roll rate tracking error is less than about 6 $[Deg/s]$. The system is sampled each $T_s = 0.25$ seconds.

Variable	Description	Constraint	Rate constraint	ℓ_2^2 penalty	ℓ_1 penalty
p	Roll rate $[Deg/s]$	-	-	100	0
r	Yaw rate $[Deg/s]$	-	-	100	0
ζ	Sideslip angle $[Deg]$ -	-	-	100	0
ϕ	Roll angle $[Deg]$	-	-	200	0
$A1$–$A2$	Rx/Lx inb. aileron $[Deg]$	[-20, 20]	[-45, 40]	0.001	0
$A3$–$A4$	Rx/Lx outb. aileron $[Deg]$	[-25, 15]	[-55, 45]	0.001	0
UR–LR	Rudders $[Deg]$	[-25, 25]	[-50, 50]	0.001	0
$S1$–$S4$	Spoilers 1–4 $[Deg]$	[0, 45]	[-75, 75]	0.001	0.01
$S5$–$S8$	Spoilers 5–8 $[Deg]$	[0, 20]	[-25, 25]	0.001	0.01
$S9$–$S12$	Spoilers 9–12 $[Deg]$	[0, 45]	[-75, 75]	0.001	0.01
$\lvert UR - LR\rvert$	Rudders differential $[Deg]$	-	-	100	0

Table 6.2: Considered aircraft variables and MPC setup

6.7.1 Nominal altitude and airspeed

The control is first simulated on the linearisation at the operating point ($z = 5000$, $V_{TAS} = 180$). Figure 6.16 shows the results for a moderate command (dashed line). In particular, p and ϕ are tracked with a good accuracy by means of the ailerons and the rudders. At the same time, ζ and r are always less than 1 degree. Ailerons on opposite wings have opposite directions, and the rudders operate together. Noticeably, the spoilers are at zero for the whole time. A larger command is then applied in Figure 6.17, which causes some spoilers to be brought into operation. Note that the spoilers command is small at the beginning, and slightly larger when the command changes in sign, consistent with the control law being a continuous function. At the same time, when the control error becomes less than 5.5, the spoiler command is zero, and they literally disappear from the loop. The same behaviour can be noticed in Figure 6.18, with an even larger reference from the pilot (double than in Figure 6.17). The spoilers commands are increased proportionally to the reference signal (roughly speaking, they are also doubled), and a third peak appears at 8.5 seconds. Overall, the spoilers never extend for more than 15 degrees, and are not used for longer than 1.5 seconds over a 10 seconds simulation.

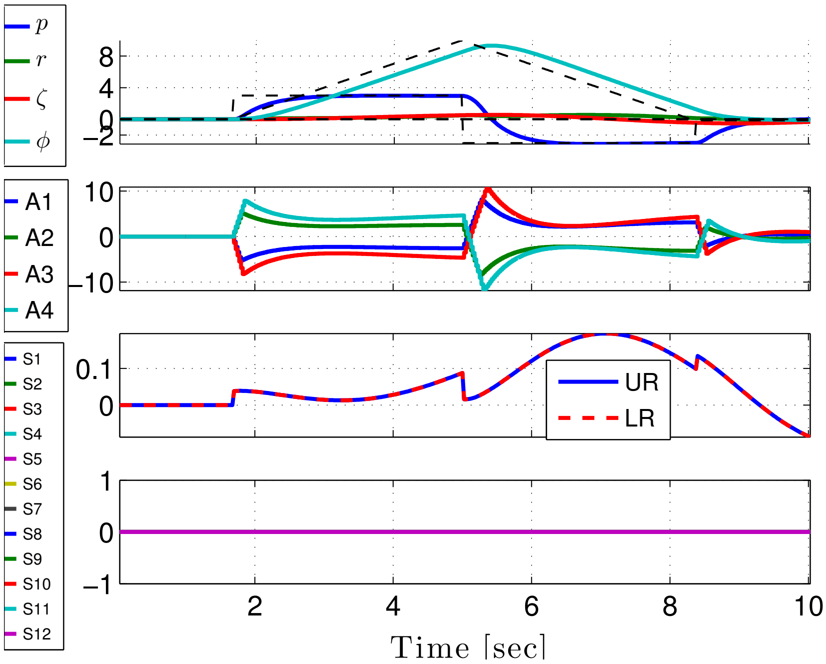

Figure 6.16: Aircraft roll control. Nominal case. Moderate roll rate command (± 3 $[deg/s]$).

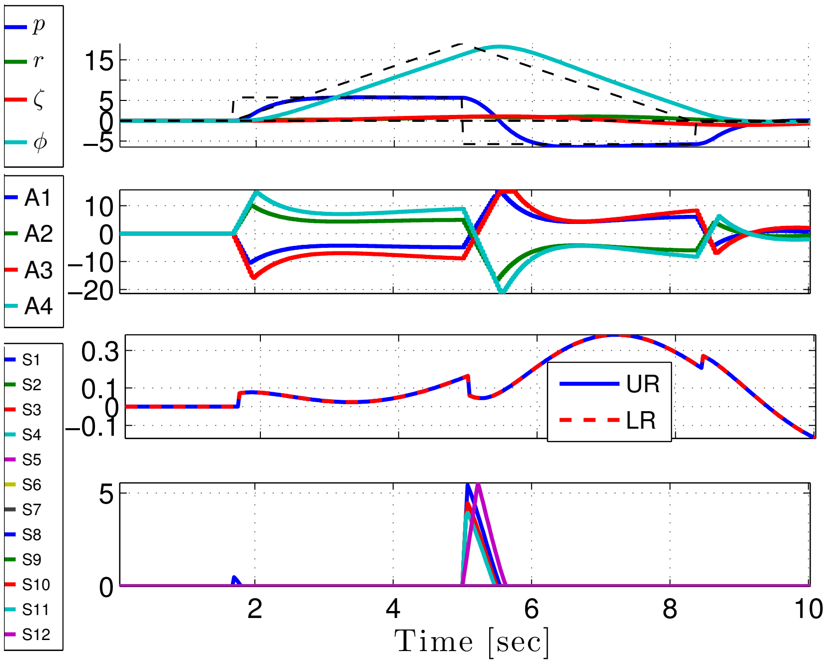

Figure 6.17: Aircraft roll control. Nominal case. Large roll rate command (± 6 $[deg/s]$).

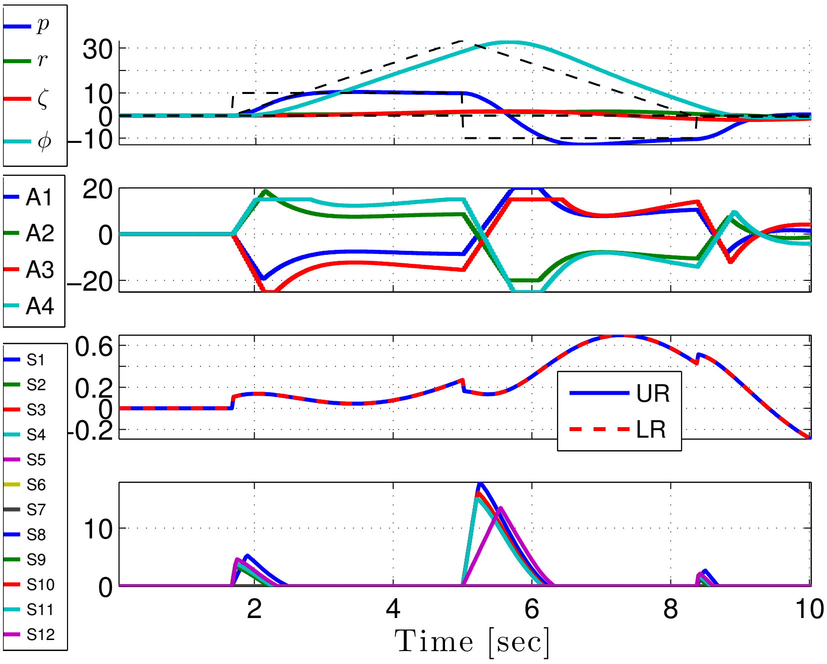

Figure 6.18: Aircraft roll control. Nominal case. Very large roll rate command (± 10 $[deg/s]$).

6.7.2 Different altitude and airspeed

In this Section, we demonstrate the effect of using, at different operating points, the tuning obtained at a particular one. Results are summarised in Table 6.3 for the nominal case, and for 12 different operating points, which are divided into 3 blocks. Two indices are considered. The first one, labelled $\text{RMS}(p(\cdot) - p_{\text{ref}}(\cdot))/T_{\text{SIM}}$, is used to monitor the control performance and it consists of the RMS of the closed loop roll rate error trajectory (about 4000 samples) and it is divided by the simulation time in seconds $T_{\text{SIM}} = 10$. From Table 6.3, it can be seen that all considered cases have comparable performance in term of this index. Simulations have also confirmed that ζ and r are still of small magnitude. The second index, labelled $\Sigma_k \|Su(k)\|_1 / T_{\text{SIM}}$, consists of the sum of the 1-norm of the spoilers command trajectory again divided by T_{SIM}, and it is used to monitor the total use of spoilers.

In the first block (cases 1–3) of Table 6.3, the airspeed is kept at its nominal value, and the altitude is decreased. It can be noticed that, as desired, the usage of spoilers increases with the decrease of altitude. The system trajectory is in Figure 6.19 and 6.20 for cases 1 and 3, where the increase in spoilers activity can be noticed. In the second block of Table 6.3 (cases 4–5), the altitude is kept to 10000 $[m]$ and the airspeed is increased, with an increase of the spoilers activity.

Case	Altitude	V_{TAS}	$\mathrm{RMS}(p(\cdot) - p_{\mathrm{ref}}(\cdot))/T_{\mathrm{SIM}}$	$\Sigma_k \|u^{\{\mathrm{ii}\}}(k)\|_1/T_{\mathrm{SIM}}$
Nominal	5000	180	2.49	0
1	2500	180	2.45	0.049
2	750	180	2.46	0.176
3	250	180	2.47	0.235
4	10000	180	2.67	0
5	10000	220	2.52	0.1
6	12000	220	2.61	0.019
7	1000	200	2.5	0.741
8	1000	160	2.46	0
9	1000	240	2.62	7.04
10	1000	280	2.16	0.411
11	5000	240	2.47	1.63
12	5000	280	2.57	15.86

Table 6.3: Aircraft control for different operating points

If the altitude is increased, as in case 6, the spoilers are again used less, as expected. An increase from the nominal case is certainly expected, as the airspeed is rather increased.

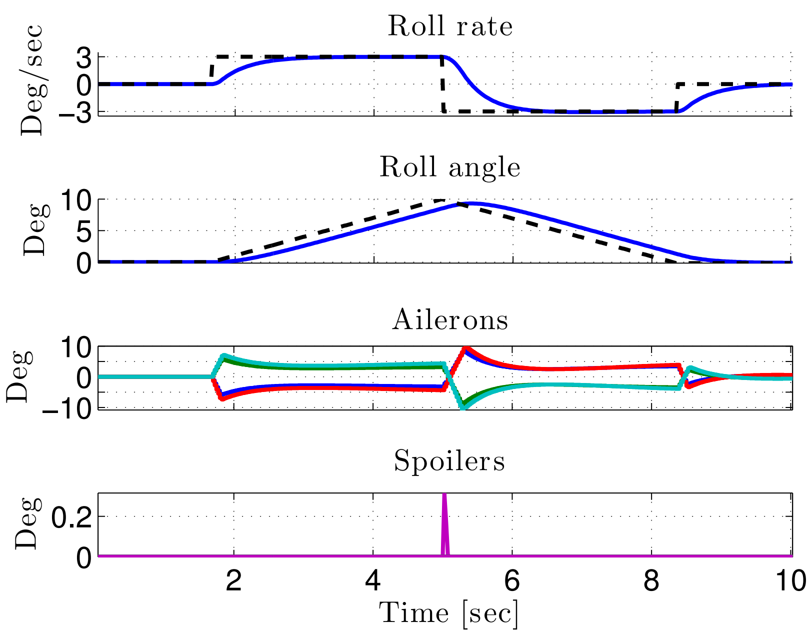

Figure 6.19: Aircraft roll control. Case 1. Moderate roll rate command (± 3 $[deg/s]$).

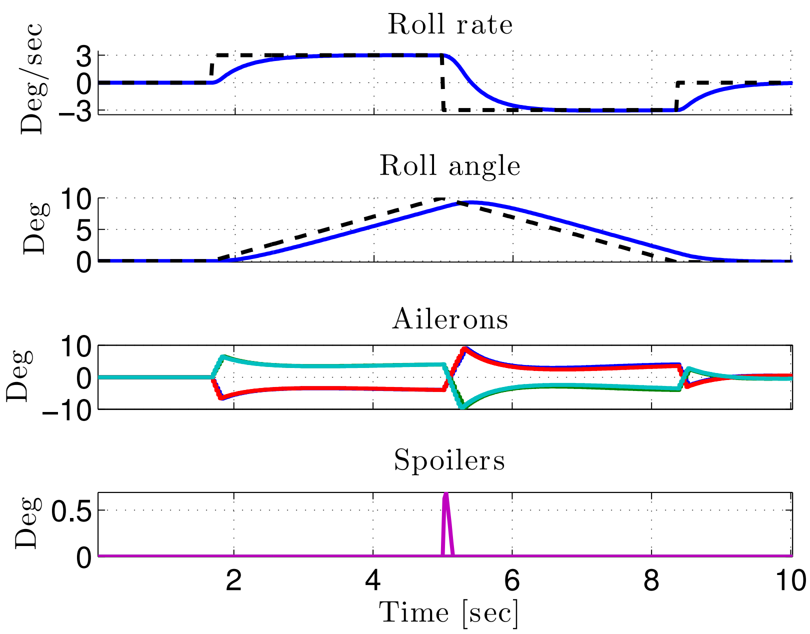

Figure 6.20: Aircraft roll control. Case 3. Moderate roll rate command (± 3 $[deg/s]$).

The third block of Table 6.3 presents a similar trend to the previous ones, with the exception of case 9 and case 12. Case 9 has a value of airspeed that is in between case 8 and 10, with same altitude, however, the spoiler indices has a much higher value than both of latter cases. From a closer look at the system behaviour of case 8 and case 9, in Figure 6.21 and 6.22 respectively, it is possible to see that the outboard ailerons are used normally in the former, and almost not used in the latter case, differently from the spoilers. This behaviour can be explained by inspecting the magnitude of the B matrix elements mapping the outboard ailerons to the roll rate for both linearisations. By doing so, it has been noticed that for case 9 the inboard ailerons have only 30% of the values of case 8, thus causing the other ailerons and the spoilers to be preferred. This can be a symptom of being close to stall for these surfaces. It can be noticed that, since the ailerons have only a quadratic penalty, the MPC is not setting them exactly to zero, differently from the spoilers. This suggests that a further level of priority could be inserted in the design, at the discretion of the aerospace engineer. Finally, in case 12, the highest use of spoilers is made, among the investigated ones. However, the use of spoilers is much higher than in any other case and, for instance, the gap with case 11 is considerable. A comparison with case 11 has been be made by inspecting the B matrix again. In this case, the considered elements have, for case 9, only 7% of the magnitude for case 8, meaning the aircraft is close to stall condition. On the other hand, the gain of spoilers 5 and 8 has increased almost three times from the nominal case. This results in the behaviour in Figure 6.23, which shows again a reduced use of the outboard ailerons, and an increased use of spoilers with a maximum of 5 degrees. While using the spoilers at lower altitude is desirable, it is not desirable to use them at higher speed (even though the magnitude is less than 5 degrees). This could be handled in two ways. The first is to tune the ℓ_1-penalty considering a linearisation at higher airspeed, as table 6.3 showed that a decrease of speed would not cause them to operate, while one in altitude does. The second could be to schedule the ℓ_1 penalty if a speed estimate is available. As this example is for illustrative purpose, these final considerations are left to the more experienced designer.

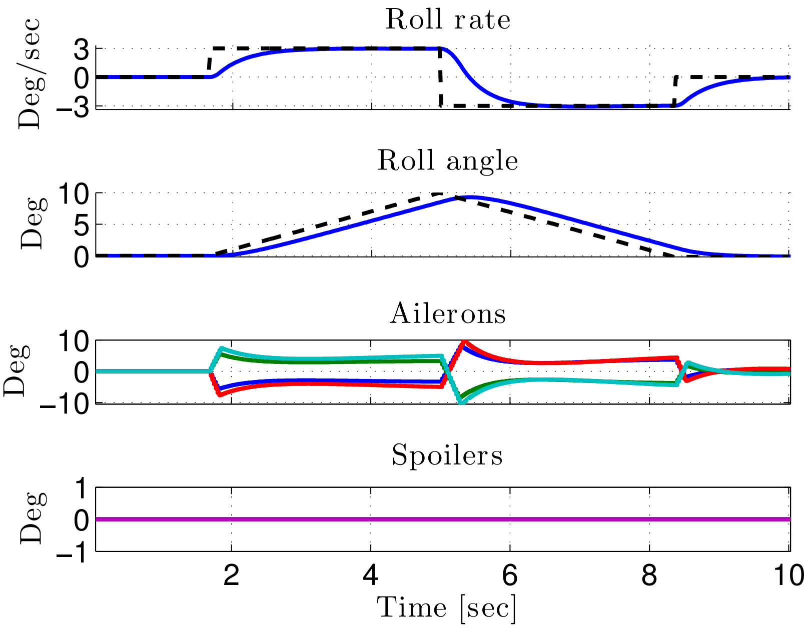

Figure 6.21: Aircraft roll control. Case 8. Moderate roll rate command (± 3 $[deg/s]$)).

6.7.3 Considerations

The preceeding example suggests the possibility of using the same tuning for different operating conditions (the systems model being available). In particular, ℓ_{asso}-MPC seems to respond sensibly to variations in the systems dynamics, and to reserve the use of auxiliary actuators to situations that

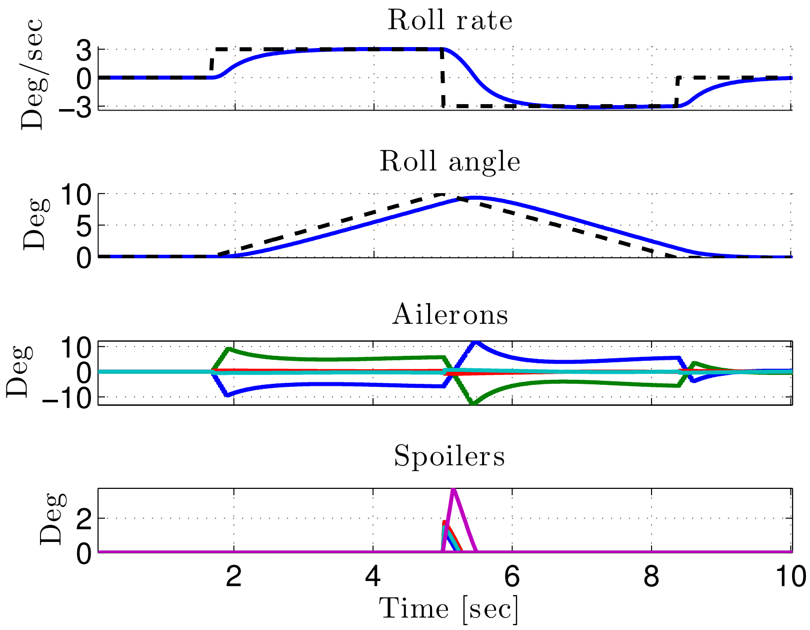

Figure 6.22: Aircraft roll control. Case 9. Moderate roll rate command (± 3 $[deg/s]$).

are sensibly compatible to the ones considered for its design. It is interesting to see how spoilers can play an important role to counteract a loss of gain (or loss of lift) in some of the ailerons, while they disappear from the picture in nominal operations. This is thanks to ℓ_{asso}-MPC, that allows one to consider all possible actuators, and to restrict their use to when the system is outside regular operations. Finally, it is worth noticing that the investigated behaviour, in particular the compensation of a loss of gain in the main actuators through the use of the auxiliary ones, appears to indicate a certain potential of ℓ_{asso}-MPC for handling faults in redundantly-actuated systems if a reconfiguration scheme for the model is available.

6.8 Conclusions

This chapter has investigated the control of constrained linear systems with preferred and auxiliary actuators by means of ℓ_{asso}-MPC. The considered scenario is the control of systems for which is desirable to operate only a subset of the actuators for most of the time and to use the remaining ones for support only when the control error is large or when the constraints are active. This could be, for instance, the case of a communication network with priorities links or of an over-actuated

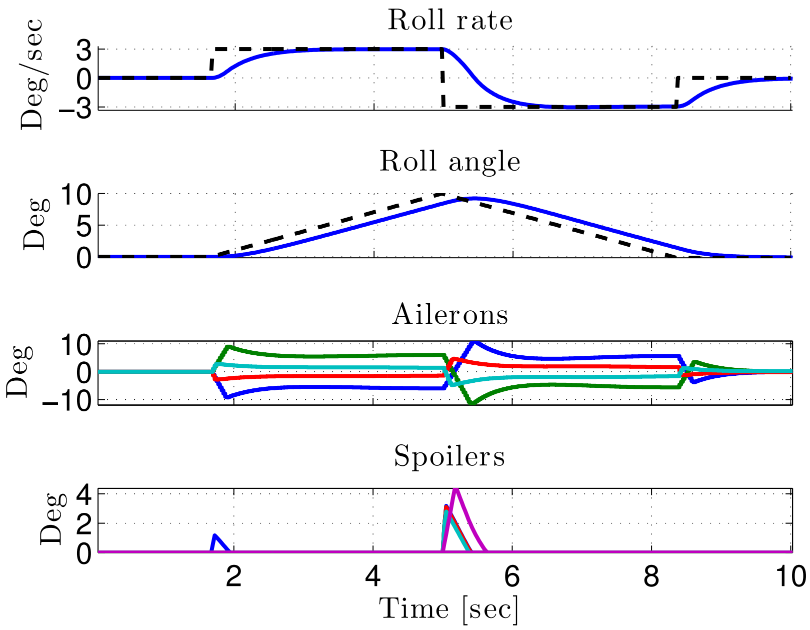

Figure 6.23: Aircraft roll control. Case 12. Moderate roll rate command (± 3 $[deg/s]$)).

vehicle.

The main result of the chapter is a procedure for computing the regularisation matrix for the auxiliary actuators given a pre-existing MPC for the preferred ones. By means of this regularisation penalty the auxiliary actuators are not used when the control error is in a certain *nominal region*. This set is a closed and convex polytopic neighbourhood of the origin specified by the designer. The complexity of the presented algorithms depends on the size of the nominal region and is comparable to the one of computing an explicit MPC controller for the preferred actuators.

Two synthesis procedures have been presented for, respectively, ℓ_{asso}-MPC version 1 and version 2. These make use of multi-parametric quadratic programming and of the theory of exact penalty functions. The preexisting MPC controller is assumed to stabilise the plant by only means of the preferred inputs. This controller can be either a LQ-MPC or a ℓ_{asso}-MPC version 2. In particular, while the former can be suitable for most LQR scenarios the latter can be used to have an additional level of priority among the main actuators.

A set-theoretic control framework has been used to characterise the limitations on the attainable nominal region, which are inevitably imposed by state and input constraints and by the finite prediction horizon. A further but possibly minor limitation comes from the assumption of the regularisation penalty being constant through the prediction horizon. In particular, the attainable nominal region is in most cases larger than the one guaranteed by the proposed procedures. If this is the case, then increasing the obtained regularisation penalty can help one achieving the maximal region. For the same reason, increasing the horizon length could also be considered.

The procedures have been demonstrated for the case of partial regularisation by means of two illustrative examples. The first one is the control of an abstracted client-server distribution network with preferred and auxiliary links. In particular, each server has full knowledge of the demand and is fully linked (with a finite transfer rate) to all clients. Only two servers and four clients have been considered, the control input being the rate of transfer though each link. The main control objective for each server is to satisfy the demand of the assigned set of clients. The second goal is to help the other server when the demand is too high. In this case it has been shown that ℓ_{asso}-MPC allows the designer to regulate the level of cooperation achieved by the servers, by implicitly determining how high the demand has to be for cooperation to occur.

The second example investigated is the control of the roll motion and lateral dynamics of a linearised Boeing 747 model, where the ailerons and rudders are preferred upon the spoilers. The use of spoilers can in fact be beneficial in anomalous situations, for instance when a very large roll rate is commanded, at the expense of additional drag. The control objective is to track a roll rate command from the pilot while keeping the yaw rate and sideslip angle small. Given the high number of inputs the proposed procedures are computationally demanding for this example. Therefore, a tuning has been performed by trial and error so that the spoilers are not used when the control error is less than 3 radians per second. Simulation results have shown that the spoilers come smoothly into play when the pilot gives a higher command. The control signals are piece-wise continuous and the system is free from chattering. Simulations have been performed also for different linearisation points using the same cost functions. In this case the results suggest that a different tuning might be needed. The results of the case studies are overall encouraging, and the proposed procedures seem to be suitable for the considered control scenarios.

CHAPTER

SEVEN

ROBUST TRACKING WITH SOFT-CONSTRAINTS

7.1 Introduction

It is becoming standard among researchers to use a terminal constraint in MPC. This constraint generally depends on the particular set-point we want to regulate the plant to. For changing set-points this means changing the terminal constraint, which could make the problem infeasible. The MPC for tracking, introduced by [Ferramosca *et al.* , 2008; Limon *et al.* , 2008] for LTI systems and extended to nonlinear systems in [Ferramosca *et al.* , 2009a], overcomes the terminal constraint limitation by means of a common invariant set for a range of feasible targets. A virtual set-point and a penalty function are also included in the MPC cost. These modifications can also significantly enlarge the feasible region of the MPC controller with respect to the one of standard MPC for regulation. The recent development of [Zeilinger *et al.* , 2010] combined the approach of [Ferramosca *et al.* , 2008] with the concept of soft-constraints [Kerrigan & Maciejowski, 2000b; Maciejowski, 2002]. The approach of [Zeilinger *et al.* , 2010] results in an even larger feasible region than [Ferramosca *et al.* , 2008], and it has also been demonstrated to provide a level of intrinsic robustness to bounded uncertainties in the form of Input-to-State Stablility (ISS) [Grimm *et al.* , 2004; Jiang & Wang, 2001; Lazar, 2006; Limon *et al.* , 2009].

In this chapter, a modified version of the LTI approach of [Zeilinger *et al.* , 2010] is formulated. We assume, similarly to [Ferramosca *et al.* , 2008] but in contrast to [Zeilinger *et al.* , 2010], the terminal state constraint to be inside the set of admissible states. The proposed approach, informally dubbed ℓ_{asso}-MPC for LTI tracking, features a novel terminal cost which improves the results of [Gallieri & Maciejowski, 2013b]. The formulation allows the terminal set to be computed independently from the cost function. This allows one to tune the MPC cost for performance without modifying the constraints, thus providing a consistent feasible region and a drastic reduction of required computation for online tuning. The control action can be computed by solving at each sampling time a positive definite quadratic program (QP). The resulting closed-loop system is ISS. The main contribution of the present chapter is a set of conditions for robust feasibility under bounded additive uncertainties. These conditions require minor modification of the terminal constraint of [Ferramosca *et al.* , 2008] and, if satisfied, they provide an uncertainty bound for which the problem is always feasible and the entire feasible region is Robustly Positively Invariant (RPI). This RPI region is allowed to be greater than the state constraints and, in most cases, is greater than the feasible region of standard MPC under similar assumptions A local ISS gain, which depends on the system and the cost parameters, is then computed. This (conservative) ISS gain is based on worst-case open-loop disturbance propagation through the MPC predictions. The approach is compared to the robust MPC of [Ferramosca *et al.* , 2012], upon which it presents computational advantages. A small example is considered, for which the proposed strategy provides a larger terminal set and a larger set of attainable steady-states than the one of [Ferramosca *et al.* , 2012].

7.2 ℓ_{asso}-MPC for LTI tracking

This chapter concerns the control of uncertain Linear Time-Invariant (LTI) systems of the form 2.5.2), subject to convex constraints of the form (2.5.3). The aim is the asymptotic regulation of $y(k) = Cx(k)$, to a constant, $y_t \in \mathbb{Y} \subset \mathbb{R}^{n_y}$. We consider a particular steady target $z_t = (x_t,\ u_t)$, satisfying which $y_t = [C\ 0]z_t$. This choice is made in order to accommodate the cases where different values of z_t could give the same y_t (for instance if $m \geq n$). For example, u_t can be taken as a sparse vector. As discussed in the previous chapters, ℓ_{asso}-MPC can be used to have preferred actuators. For the same purpose, let us assume the target z_t to be obtained from a "LASSO target calculator":

$$
\begin{aligned}
z_t = \arg\min\nolimits_{x_t, u_t}\ & u_t^T \tilde{R} u_t + \|\tilde{S} u_t\|_1 \\
\text{s.t.}\ & \begin{bmatrix} I - A & -B \\ C & 0 \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} = \begin{bmatrix} 0 \\ y_t \end{bmatrix}.
\end{aligned}
\tag{7.2.1}
$$

The 1-norm term in the target calculator cost encourages u_t to be sparse for some y_t, if for instance we take $\tilde{S} \succ 0$ and diagonal, with large entries for some actuators. It is assumed that:

Assumption 10. **(A10)** $\tilde{R} \succ 0$, $\tilde{S} \in \mathcal{D}_{++}$, $\mathbb{Y}$ is a C-set.

Theorem 7.2.1. Assume **(A10)**, (A, B) is stabilisable, (C, A) detectable, and the matrix

$$\bar{G} = \begin{bmatrix} I - A & -B \\ C & 0 \end{bmatrix}$$

has full row rank $(n + n_y)$. Then (7.2.1) has a unique minimiser z_t, continuous and PWA in y_t.

Proof. (Theorem 7.2.1) Assume for simplicity $\tilde{R} \in \mathcal{D}_{++}$ (for the general case note that (7.2.1) can be expressed in terms of $\tilde{u}_t = \tilde{R}^{1/2} u_t$, since $\tilde{R}$ is invertible). Then, define the variables $(u_t^+,\ u_t^-)$ such that $u_t = u_t^+ - u_t^-$. Since $\tilde{S} \in \mathcal{D}_{++}$, then (7.2.1) is equivalent to a positive definite QP, in $\chi = \mathrm{col}(x_t, u_t^+, u_t^-)$, subject to $\mathrm{col}(u_t^+, u_t^-) \geq 0$ and $G\chi = \mathrm{col}(0, y_t)$, with

$$G = \begin{bmatrix} I - A & -B & B \\ C & 0 & 0 \end{bmatrix}.$$

From the assumption on $\bar{G}$, the matrix G has full row rank and the problem is solvable, as in [Muske, 1997; Pannocchia & Bemporad, 2007; Pannocchia & Rawlings, 2003]. The solution characteristic follows from the arguments in Theorem 1 of [Alessio & Bemporad, 2009]. ■

Remark 35. A more general $\tilde{S}$ can also be used in (7.2.1), by means of similar arguments to the ones of Theorem 4.4.2. This could be useful, for instance, to limit differential actuation in steady-state.

Throughout the chapter, we will assume z_t to be known. Denote:

$$\delta x = x - x_t. \tag{7.2.2}$$

Recall the definition of M_θ from Lemma 2.6.11:

$$M_\theta = \mathbf{null}([A - I,\ B]).$$

Consider also the following:

Assumption 11. **(A11)** The sets $\mathbb{X}, \mathbb{U}$ are polytopes, with $\mathbb{X} = \{x \in \mathbb{R}^n : Lx \leq \underline{1}\}$, $L \in \mathbb{R}^{n_L \times n}$, for some $n_L > 0$.

The following MPC formulation is used throughout the chapter:

Definition 7.2.1. (ℓ_{asso}–MPC for robust LTI tracking)
Consider the following finite-horizon constrained optimal control problem (FHCOCP)

$$\begin{aligned}
V_N^o(\delta x) = \min_{\underline{\mathbf{u}},\theta,\zeta,\underline{\mathbf{s}}} &\{V_N(x,\underline{\mathrm{u}},\theta,\zeta,\underline{\mathrm{s}})\} \\
s.t. \quad \hat{x}_{j+1} = A\hat{x}_j + B\hat{u}_j,\ &\delta\hat{x}_j = \hat{x}_j - M_x\theta, \\
\hat{u}_j \in \mathbb{U},\ L\hat{x}_j \leq \underline{1} + s_j,\ &j = 1,\dots,N-1, \\
\delta\hat{u}_j = \hat{u}_j - M_u\theta,\ &Z\hat{x}_N = X\zeta, \\
(\delta\hat{x}_N,\theta) \in \mathbb{X}_f^t,\ &\hat{x}_0 = x,
\end{aligned} \tag{7.2.3}$$

$$\begin{aligned}
V_N(x,\underline{\mathrm{u}},\theta,\underline{\mathrm{s}}) &= F(\zeta,\delta\hat{x}_N) + \sum_{j=0}^{N-1}\ell(\delta\hat{x}_j,\delta\hat{u}_j) \\
&\quad + V_s(\underline{\mathrm{s}}) + V_O(M_\theta\theta - z_t), \\
V_s(\underline{\mathrm{s}}) = \sum_{j=0}^{N-1}\ell_s(s_j),\ &\ell_s(s_j) = s_j^TQ_cs_j + \gamma_c\|s_j\|_1, \\
V_O(z) &= z^TQ_tz + \gamma_t\|\bar{S}z\|_1, && (7.2.4) \\
\ell(x,u) &= x^TQx + u^TRu + \|Su\|_1, && (7.2.5) \\
F(\zeta,x) &= \alpha F_1(\zeta,x) + \beta F_2(\zeta,x),\ \alpha,\beta \geq 0, && (7.2.6) \\
Q_c \succ 0,\ Q_t \succ 0,\ &\bar{S} \in \mathcal{D}_{++},\ \gamma_t > 0,\ \gamma_c > 0, && (7.2.7)
\end{aligned}$$

with $\underline{\mathrm{s}}^T = \begin{bmatrix} s_0^T, & \cdots, & s_{N-1}^T \end{bmatrix}$. At each iteration k, the ℓ_{asso}-MPC for tracking applies to the plant the first move of the optimal policy, $u(k) = \hat{u}_0^\star$, obtained by online solution of (7.2.3-7.2.4) at the current state $x = x(k)$. The generated implicit control law is referred to as $K_N^t(x, z_t) \equiv \hat{u}_0^\star$. The closed-loop evolution of (2.5.2) is then

$$x(k+1) = Ax(k) + BK_N^t(x(k), z_t) + B_w w(k), \tag{7.2.8}$$

and its one-step evolution, with $w(k) = 0$ is denoted as x^+.

Note that, in problem (7.2.3), the set $\mathbb{X}_f^t$ (the *terminal set for tracking*), the function F and the matrices Z, X are to be determined to obtain closed-loop stability and robustness. The variable ζ is used to implement the Minkowski function terminal cost discussed in Chapter 5. A novel formulation is proposed to assess a level of intrinsic robustness. In particular, the terminal set is the only ingredient of the approach upon which robustness depends, and it is computed offline prior to cost function definition, similarly to Chapter 5. The approach extends the ideas of [Limon *et al.* , 2002] to the MPC for tracking. Moreover, in this chapter we make use of *polytopic Lyapunov functions* [Blanchini & Miani, 2008]. Soft-constraints are used, as in [Zeilinger *et al.* , 2010], to allow for momentary constraint violation (for example in the MPC predictions). The resulting ℓ_{asso}-MPC for tracking will be shown to be robustly feasible and ISS, for a certain disturbance bound μ, returned by the proposed method.

The slack variable, $\underline{s}$, allows for feasibility feasibility to be maintained when $w(k)$ drives the system state (or its future MPC predictions) out of $\mathbb{X}$. In fact, thanks to the use of $\underline{s}$, the feasible region can be larger than $\mathbb{X}$. Soft-constraints are common in MPC literature and applications, see for instance [Kerrigan & Maciejowski, 2000b; Zeilinger *et al.* , 2010] and reference therein. Differently from [Zeilinger *et al.* , 2010], a quadratic slack constraint is not included here. This, besides dispensing the solver from additional complexity, expresses our intention to keep the system's trajectory within constraints under nominal conditions. This choice is taken at the expense of a possibly smaller feasible region than in [Zeilinger *et al.* , 2010].

In order to proceed further, consider the following:

Assumption 12. (A12) For system (2.5.2)

(H0) (A, B) is stabilisable,

(H1) $Q \succ 0$, $R \succ 0$, $S \in \mathbb{R}^{n_s \times n_u}$,

(H2) $\mathbb{X}$, $\mathbb{U}$, $\mathbb{X}_f^t$ are C-sets (convex, compact sets),

(H3) $\mathbb{W}$ is a C-set, with $0 \in \text{int}(\mathbb{W})$.

Recall that, from Definition 2.5.1, $\bar{f}^i_{\underline{\mathbf{u}}}(x)$, is the i-step evolution of (2.5.2), under a policy $\underline{\mathbf{u}}$, with $w(k) = 0, \forall k \geq 0$. For a given $\theta_t = M_\theta^T z_t$, we can define:

$$\mathbb{X}_f(\theta_t) = \{\delta x : (\delta x, \theta_t) \in \mathbb{X}_f^t\}. \tag{7.2.9}$$

The feasible region of (7.2.3) is a polytope, given by

$$\mathbb{X}_N^t = \{x : \exists \underline{\mathbf{u}} \in \mathbb{U}^N, \theta \in \mathbb{R}^{n_\theta} \mid f^N_{\underline{\mathbf{u}}}(x) - M_x\theta \in \mathbb{X}_f(\theta)\}. \tag{7.2.10}$$

For a fixed θ_t, the feasible region becomes

$$\mathbb{X}_N(\theta_t) = \{x : \exists \underline{\mathbf{u}} \in \mathbb{U}^N \mid f_{\underline{\mathbf{u}}}^N(x) - M_x\theta_t \in \mathbb{X}_f(\theta_t)\}, \tag{7.2.11}$$

which is the typical feasible region for an MPC for regulation with softened state constraints.

7.2.1 Nominal asymptotic stability

In order to analyse the ℓ_{asso}-MPC, a set of assumptions are made on the terminal set $\mathbb{X}_f^t$, the computation of which will be discussed in Section 7.3. Define the notation:

$$f_K^i(\delta\hat{x}) = (A + BK)^i \delta\hat{x}, \tag{7.2.12}$$
$$\delta\hat{x} = x - M_x\theta, \tag{7.2.13}$$

as well as the auxiliary system

$$\delta\hat{x}(k+1) = (A + BK)\delta\hat{x}(k). \tag{7.2.14}$$

It is assumed that:

Assumption 13. (A13) $|\lambda_{\max}(A + BK)| < 1$.

Denote

$$\bar{F}(\delta\hat{x}) = \min_{\{\zeta, Z\delta\hat{x} = X\zeta\}} F(\zeta, \delta\hat{x}), \tag{7.2.15}$$

and consider the following candidate terminal controller[1]:

$$\kappa(x, \theta) = K\delta\hat{x} + M_u\theta. \tag{7.2.16}$$

Assumption 14. (A14) Assume $Zf_K(\delta\hat{x}) = X\zeta^+(\delta\hat{x})$, for some $\zeta^+(\delta\hat{x})$ and that $\forall(\delta\hat{x}, \theta) \in \mathbb{X}_f^t$:

$$\tilde{\alpha}_1(\|\delta\hat{x}\|_2) \leq \bar{F}(\delta\hat{x}) \leq \tilde{\alpha}_2(\|\delta\hat{x}\|_2), \tag{7.2.17}$$
$$F(\zeta^+(\delta\hat{x}), f_K^1(\delta\hat{x})) - \bar{F}(\delta\hat{x}) \leq -\ell(\delta\hat{x}, K\delta\hat{x}), \tag{7.2.18}$$
$$\kappa(x, \theta) \in \mathbb{U}, \tag{7.2.19}$$

[1]This controller is used for proving the proposed results and is not applied to the plant during operations.

where $\tilde{\alpha}_1(r)$ and $\tilde{\alpha}_2(r)$ are $\mathcal{K}_\infty$-functions.

Denote:

$$\mathcal{X}_f = \pi_{\delta x}(\mathbb{X}_f^t), \tag{7.2.20}$$

where $\pi_{\delta x}$ is the polytopic projection on the $\delta\hat{x}$ coordinate (see Definition 2.2.17). Define also the set of *admissible steady states* as

$$\mathbb{Z}_s = \{z_s = M_\theta \theta : \theta \in \pi_\theta(\mathbb{X}_f^t)\}, \tag{7.2.21}$$

where $\pi_\theta(\mathbb{X}_f^t)$ is the projection of $\mathbb{X}_f^t$ on the θ coordinate. The proposed assumptions on the terminal set and terminal controller provide the following:

Theorem 7.2.2. Assume **(A14)**. Then the following results are obtained

1. For all θ such that $(x - M_x\theta, \theta) \in \mathbb{X}_f^t$, there exists an admissible set, $\mathbb{X}_f(\theta) + M_x\theta$, which is positively invariant (PI) for system (2.5.7) under the candidate terminal controller (7.2.16).

2. $\delta\hat{x} \in \mathbb{X}_f(\theta) \Rightarrow f_K^i(\delta\hat{x}) \in \mathbb{X}_f(\theta) \subseteq \mathcal{X}_f, \ \forall i \in \mathbb{I}_{\geq 0}$.

3. The set $\mathcal{X}_f$ is PI for system (7.2.14).

Proof. (Theorem 7.2.2)

1. Notice that $\bar{F}(x - M_x\theta)$ is by **(A14)** a CLF in[1] $\mathbb{X}_f(\theta) + M_x\theta$, for any θ such that $(x - M_x\theta, \theta) \in \mathbb{X}_f^t$, with the associated control law (7.2.16). Then, from Theorem 2.5.1 the set $\mathbb{X}_f(\theta) + M_x\theta$ is a domain of attraction and, by Definition 2.5.9, is positively invariant for the considered system.

2. Point 1 implies directly that $x(i) - M_x\theta \in \mathbb{X}_f(\theta), \ \forall i \in \mathbb{I}_{\geq 0}$. In order to obtain $f_K^i(\delta\hat{x}) \in \mathbb{X}_f(\theta), \ \forall i \in \mathbb{I}_{\geq 0}$, consider that for $i = 1$ we have

$$\begin{aligned} f_K^1(\delta\hat{x}(k)) = x(k+1) - M_x\theta &= Ax(k) + B(Kx(k) - M_x\theta) + BM_u\theta - M_x\theta && (7.2.22)\\ &= A(x(k) - M_x\theta) + BK(x(k) - M_x\theta), && (7.2.23) \end{aligned}$$

where we have used the fact that

$$BM_u\theta - M_x\theta = -AM_x\theta, \tag{7.2.24}$$

[1]The notation $\mathbb{X}_f(\theta) + M_x\theta$ stands for the set $\mathbb{X}_f(\theta)$ with its centre shifted to $M_x\theta$.

from Lemma 2.6.11. This holds for all i, since the system is linear. The fact that $\mathbb{X}_f(\theta) \subseteq \mathcal{X}_f$ follows by Definition 2.2.17.

3. From point 2 the trajectory of (7.2.14) satisfies the positively invariance condition in $\mathbb{X}_f(\theta)$ for any admissible θ. By definition 2.2.17 we have $\mathcal{X}_f = \cup_{(\theta \in \pi_\theta(\mathbb{X}_f^t))} \mathbb{X}_f(\theta)$, which is a C-set. The invariance of $\mathcal{X}_f$ for system (7.2.14) follows from the fact that the system is linear and the constraints convex, therefore any convex combination of points $ax_1 + (1-a)x_2 \in \mathcal{X}_f,\ a \in [0,1]$ is inside an invariant set $\mathbb{X}_f(a\theta_1 + (1-a)\theta_2)$ with $M_\theta(a\theta_1 + (1-a)\theta_2) \in \mathbb{Z}_s$. Conversely, by Definition 2.2.17 we have $x \in \mathcal{X}_f \Rightarrow \exists \theta : M_\theta \theta \in \mathbb{Z}_s,\ x \in \mathbb{X}_f(\theta)$.

■

In order to proceed with the stability result concerning ℓ_{asso}-MPC for tracking, define the following quantities:

$$\bar{r} = \min_{\theta} \max_{r} \{ r \mid \{\delta\hat{x} : \|\delta\hat{x}\|_2 \leq r\} \subseteq \mathbb{X}_f(\theta) \}, \tag{7.2.25}$$
$$\bar{V}_N = \max_{(x \in \mathbb{X}_N^t,\ z_t \in \mathbb{R}^{n+m})} \{ V_N^o(\delta x) - V_O(z_s^\star - z_t) - V_s(\underline{s}^\star) \}, \tag{7.2.26}$$

where $z_s^\star$ and $\underline{s}^\star$ are evaluated at the optimum of (7.2.3) and $\delta x = x - x_t$ as in (7.2.2). Define also

$$\bar{z}_t = \arg \min_{z_s \in \mathbb{Z}_s} V_O(z_s - z_t). \tag{7.2.27}$$

Denote $\delta z_s^\star = z_s^\star - z_t$ and $\delta x^\star = x - M_\theta \theta^\star$, where $\theta^\star$ is optimal for (7.2.3). The following is obtained:

Lemma 7.2.3. Assume **(A12)–(A14)**. Then, it follows that:

$$\alpha_1(\|\delta x^\star\|_2) \leq V_N^o(\delta x) - V_O(\delta z_s^\star) - V_s(\underline{s}^\star) \leq \alpha_2(\|\delta x^\star\|_2),\ \forall x \in \mathbb{X}_N^t, \tag{7.2.28}$$

where, $\alpha_1(r) = \lambda_{\min}(Q) r^2$, $\alpha_2(r) = \epsilon \tilde{\alpha}_2(r)$, $\epsilon = \max\left(1, \frac{\bar{V}_N}{\tilde{\alpha}_2(\bar{r})}\right)$ and $\epsilon > 0$ is finite.

Proof. (Lemma 7.2.3) Trivially, $\alpha_1(\|x\|_2) = \lambda_{\min}(Q)\|x\|_2^2$ follows from $\ell(x,u) \geq \lambda_{\min}(Q)\|x\|_2^2, \forall u$, and $V_N^o(x) - V_O(\delta z_s^\star) \geq \ell(x, K_N(x)), \forall x$. In order to compute α_2, we proceed similarly to Lemma 4.5.1. First, consider the case when $(x - M_x\theta^\star, \theta^\star) \in \mathbb{X}_f^t$, where x is the system's state. In this case, from **(A14)** and Theorem 7.2.2, the sequence $\delta\tilde{x} = \{\delta\hat{x}_j = f_K^j(x - M_x\theta^\star),\ j = 0, \ldots, N\}$ stays in $\mathbb{X}_f(\theta^\star)$ and, together with $\delta\tilde{u} = \{\delta\hat{u}_j = K f_K^j(x - M_x\theta^\star),\ j = 0, \ldots, N-1\}$ and $\tilde{s} = 0$, provides a feasible solution for problem (7.2.3). Denote the MPC cost for these sequences as $\widetilde{V}_N(x)$. This is an upper bound for the optimal MPC cost, therefore, by optimality and by recursive application of (7.2.18) we have, similarly to Lemma 4.5.1, that $\forall(x - M_x\theta^\star, \theta^\star) \in \mathbb{X}_f^t$:

$$\begin{aligned} V_N^o(\delta x) - V_O(\delta z_s^\star) - V_s(\underline{\mathrm{s}}^\star) &\leq \widetilde{V}_N(x) - V_O(\delta z_s^\star) \\ \leq \bar{F}(x - M_x\theta^\star) &\leq \tilde{\alpha}_2(\|x - M_x\theta^\star\|_2) = \tilde{\alpha}_2(\|\delta x^\star\|_2). \end{aligned} \tag{7.2.29}$$

To complete the proof, we need to show the existence of the upper bounding $\mathcal{K}_\infty$-function for the case when $(x - M_x\theta^\star, \theta^\star) \notin \mathbb{X}_f^t$. The proof proceeds in a similar way to the one of Lemma 1 of [Limón *et al.* , 2006]. Recall that $\mathbb{X}$, $\mathbb{X}_f^t$ and $\mathbb{U}$ are C-sets. This fact and **(A14)** guarantee that there exist a finite upper bound, $\bar{V}_N$, given by (7.2.26) for which $V_N^o(\delta x) - V_O(\delta z_s^\star) - V_s(\underline{\mathrm{s}}^\star) \leq \bar{V}_N$, $\forall x \in \mathbb{X}_N^t$, $\forall z_t \in \mathbb{R}^{n+m}$. Notice that in the last step the possibly unbounded offset function $V_O(\delta z_s^\star)$ has been subtracted from the optimal cost, thus having a bounded function. Since $(\delta x^\star, \theta^\star) \notin \mathbb{X}_f^t$ then $\|\delta x^\star\|_2 > \bar{r}$, and $\alpha_2(\|\delta x^\star\|_2) > \alpha_2(\bar{r})$. This implies the following:

$$V_N^o(\delta x) - V_O(\delta z_s^\star) - V_s(\underline{\mathrm{s}}^\star) \leq \bar{V}_N \leq \bar{V}_N \frac{\alpha_2(\|\delta x^\star\|_2)}{\alpha_2(\bar{r})} \leq \epsilon\alpha_2(\|\delta x^\star\|_2). \tag{7.2.30}$$

Notice that the above inequality holds $\forall x \in \mathbb{X}_N^t, \forall\theta^\star : M_\theta\theta^\star \in \mathbb{Z}_s$. ■

Building upon [Ferramosca *et al.* , 2008, 2009b; Limon *et al.* , 2008; Zeilinger *et al.* , 2010], the following can be stated for a constant z_t:

Theorem 7.2.4. Assume **(A11)**–**(A14)**, $w(k) = 0, \forall k \geq 0$. Then:

1. If $z_t \in \mathbb{Z}_s$, then z_t is the sole equilibrium for system (7.2.8) with $x(k) \to x_t$ for $k \to \infty$, $\forall x(0) \in \mathbb{X}_N^t$. The set $\mathbb{X}_N^t$ is PI for (7.2.8).

2. If $z_t \notin \mathbb{Z}_s$ then $\bar{z}_t = (\bar{x}_t, \bar{u}_t)$ the unique solution of (7.2.27), is the sole equilibrium for (7.2.8), and $x(k) \to \bar{x}_t$ for $k \to \infty$, $\forall x(0) \in \mathbb{X}_N^t$.

3. $\forall x(0) \in \mathbb{X}_N^t$, $x(k) \to \mathbb{X}$, as $k \to \infty$.

Proof. (Theorem 7.2.4)

1. Convergence to an equilibrium, $\bar{z}_s \in \mathbb{Z}_s$, follows from a direct method Lyapunov argument [Ferramosca *et al.* , 2008; Limon *et al.* , 2008]. In particular, at time $k+1$ the control move $\tilde{u}_{N-1|k+1} = K\delta x_{N|k} + M_u\theta_k^\star$ can be added to the tail of the solution at time k to obtain a

new feasible solution (similar to Theorem 2.6.4). This argument provides nominal recursive feasibility for the MPC problem. Take $\bar{z}_s = M_\theta \theta_k^\star$. Then $\forall x(k) \in \mathbb{X}_N^t$ we have:

$$\begin{aligned} V_N^o(\delta x^+(k)) - V_N^o(\delta x(k)) &\leq -\alpha_1(\|\delta x^\star(k)\|_2) - \ell_s(\tilde{s}_1(k)) \\ &\leq -\alpha_1(\|\delta x^\star(k)\|_2), \end{aligned} \tag{7.2.31}$$

where $\delta x^+(k)$ is the 1-step system evolution under the MPC controller, from $\delta x(k) = x(k) - \theta_k^\star$, where $\delta x^\star(k)$, $\tilde{s}_1(k)$ are taken from the optimal solution at time k. Inequalities (7.2.31), (7.2.28) imply that $\delta x^\star(k) \to 0$ and $V_N^o(\delta x(k)) \Rightarrow V_O(\delta z_s^\star(k))$. Moreover, condition (7.2.31) implies that $\mathbb{X}_N^t$ is positively invariant (notice that $M_x\theta_k^\star \in \mathbb{X}_N^t$ by definition of $\mathbb{X}_f^t$).

Convergence to the specified z_t follows from the contradiction argument of [Ferramosca *et al.* , 2009b], since all assumptions are satisfied.

2. Convergence to $\bar{z}_t$ is discussed in [Ferramosca *et al.* , 2009b]. The cost function is strongly convex (since $R \succ 0$ is assumed), and the optimal solution is unique, hence the equilibrium point for (7.2.8) is unique.

3. The claim follows by definition of $\mathbb{Z}_s$ and by part 1.

■

For a time-varying z_t, the following result is obtained:

Corollary 7.2.5. Assume a non-constant $z_t(k)$ which tends asymptotically to a constant z_∞. If the assumptions for Theorem 7.2.4 are satisfied, then the ℓ_{asso} MPC for tracking is recursively feasible in $\mathbb{X}_N^t$, which is positively invariant for system (7.2.8). The trajectory of (7.2.8) converges asymptotically to the unique solution of (7.2.27) for $z_t = z_\infty$.

7.3 Computation of the terminal constraint

Two approaches are discussed, with similar arguments to those of Chapter 5.

7.3.1 Using the maximal admissible set

Building upon [Limon *et al.* , 2008; Zeilinger *et al.* , 2010], define the extended state space $(\delta x, \theta)$. Our contribution, for the purpose of robustness, is based on the introduction of a $\lambda \in (0, 1)$ and of:

$$\bar{A}_{\lambda} = \begin{bmatrix} \frac{1}{\lambda}(A+BK) & 0 \\ 0 & I \end{bmatrix}. \tag{7.3.1}$$

Define now, similar to [Ferramosca *et al.* , 2008; Limon *et al.* , 2008]:

$$\mathbb{V}_{\alpha_z} = \{(\delta x, \theta) : (\delta x, K\delta x) + M_\theta \theta \in \mathbb{Z},\ M_\theta \theta \in \alpha_z \mathbb{Z} \subset \mathbb{Z}\}. \tag{7.3.2}$$

Assumption 15. **(A15)** $\lambda \in (|\lambda_{\max}(A+BK)|, 1)$, $\alpha_z \in (0, 1)$.

Then, the following is a λ*-contractive set for tracking*:

$$\mathcal{O}^t_{\infty,\lambda,\alpha_z} = \{(\delta x, \theta) : \bar{A}^j_\lambda(\delta x, \theta) \in \mathbb{V}_{\alpha_z},\ \forall j \in \mathbb{I}_{\geq 0}\}. \tag{7.3.3}$$

Assumption 16. **(A16)** $\mathbb{X}^t_f = \{(\delta x, \theta) : (\frac{1}{\lambda}\delta x, \theta) \in \mathcal{O}^t_{\infty,\lambda,\alpha_z}\}$.

The following Lemmas are key for proving the main result:

Lemma 7.3.1. (Contractivity of $\mathbb{X}_f(\theta)$) Assume **(A11)–(A16)**. Then, $\mathcal{O}^t_{\infty,\lambda,\alpha_z}$ is a finitely determined polytope and $\delta x \in \mathbb{X}_f(\theta) \Rightarrow f^j_K(\delta x) \in \lambda^j \mathbb{X}_f(\theta),\ \forall j \in \mathbb{I}_{\geq 0}$.

Proof. (Lemma 7.3.1) From (A15), we have $\lambda_{\max}(\frac{1}{\lambda}(A+BK)) < 1$. Since $\bar{A}_\lambda$ has also unitary eigenvalues, as noticed in [Limon *et al.* , 2008], the set $\mathcal{O}^t_{\infty,\lambda,\alpha_z}$ can be computed for any $\alpha_z \in (0,1)$ as the intersection of a finite number of inequalities [Gilbert & Tan, 1991].

Given $(\gamma\delta x, \theta) \in \mathcal{O}^t_{\infty,\lambda,\alpha_z}$ with $\gamma = 1/\lambda$, from (7.3.3), it follows that $(\frac{1}{\lambda^j}(A+BK)^j\gamma\delta x, \theta) \in \mathcal{O}^t_{\infty,\lambda,\alpha_z}$, $\forall j \in \mathbb{I}_{\geq 0}$. By (A16), we also have $(\delta x, \theta) \in \mathbb{X}^t_f$ and $\delta x \in \mathbb{X}_f(\theta)$. Then, it also follows that $\frac{1}{\lambda^j}(A+BK)^j\delta x \in \mathbb{X}_f(\theta)$, from which the claim is obtained. ■

Denote:

$$\mathbb{X}^t_f = \left\{(\delta x, \theta) \in \mathbb{R}^{n+n_\theta} : \frac{1}{\lambda}G_x\delta x + G_\theta\theta \leq \underline{1}\right\}. \tag{7.3.4}$$

Clearly, from (7.2.20) we have $\mathcal{X}_f = \lambda\pi_{\delta x}(\mathcal{O}^t_{\infty,\lambda,\alpha_z})$, from which the following is obtained:

Lemma 7.3.2. (Contractivity of $\mathcal{X}_f$) Assume **(A11)–(A16)**. Then,

1. $\pi_{\delta x}(\lambda\mathbb{X}^t_f) = \lambda\pi_{\delta x}(\mathbb{X}^t_f)$,

2. $(\delta x, \lambda\theta) \in \lambda\mathbb{X}^t_f \Rightarrow \delta x \in \lambda\pi_{\delta x}(\mathbb{X}^t_f)$,

3. $\delta x \in \pi_{\delta x}(\mathbb{X}^t_f) \Rightarrow \exists\theta :\ (\frac{1}{\lambda}f^1_K(\delta x), \theta) \in \mathbb{X}^t_f$ and $f^j_K(\delta x) \in \lambda^j\pi_{\delta x}(\mathbb{X}^t_f),\ \forall j \in \mathbb{I}_{\geq 0}$.

Proof. (Lemma 7.3.2)

1. It is assumed that $\exists \theta : \frac{1}{\lambda} G_x \delta x + G_\theta \theta \leq \lambda \underline{1}$. Therefore, $\exists \bar{\theta} = \theta/\lambda : (\frac{1}{\lambda}\delta x, \bar{\theta}) \in \mathbb{X}_f^t$. The converse is also easily verified. Then, from Definition 2.2.17, we have $\frac{1}{\lambda}\delta x \in \mathcal{X}_f$, from which the claim is obtained.

2. This is a direct consequence of point 1.

3. From Definition 2.2.17 we have $\delta x \in \pi_{\delta x}(\mathbb{X}_f^t) \Rightarrow \exists \theta : (\delta x, \theta) \in \mathbb{X}_f^t$. From Lemma 7.3.1, this implies that $\frac{1}{\lambda} G_x (A + BK)^j \delta x \leq \lambda^j (\underline{1} - G_\theta \theta)$, namely, $((A+BK)^j \delta x, \theta \lambda^j) \in \lambda^j \mathbb{X}_f^t$. This, from point 2, implies that $f_K^j(\delta x) \in \lambda^j \pi_{\delta x}(\mathbb{X}_f^t)$.

■

7.3.2 Using the control at the vertices

An alternative way to compute a contractive set for tracking is to consider the matrices V, Θ the columns of which are vertices in the δx and in the θ space, together with the scalars λ, γ, so that $\mathbb{X}_f^t = \mathbf{conv}(\lambda\gamma V) \times \mathbf{conv}(\Theta)$. Then Algorithm 1 can be modified to obtain Algorithm 6:

Algorithm 6 Terminal set for tracking (alternative)

1. Choose V, $\lambda \in (0, 1)$, and $\epsilon \geq 0$.

2. Solve the LP
$$\begin{array}{ll} \max_\chi & \|\Theta\|_1 + \gamma - \epsilon \|\bar{K}\|_1 \\ \text{s.t.:} & A\gamma V + B\bar{K}V = V\bar{Y}, \\ & \underline{1}^T \bar{Y} \leq \gamma\lambda \underline{1}^T, \\ & E(\bar{K}V + M_u\Theta) \leq [\underline{1}, \dots, \underline{1}], \\ & L(\gamma V + M_x \Theta) \leq [\underline{1}, \dots, \underline{1}], \\ & 0 < \gamma \leq \gamma_{\max}, \ \bar{Y} \geq 0, \ \|\Theta\|_1 > 0. \end{array} \tag{7.3.5}$$
where $\chi = (\Theta,\ \gamma,\ \bar{K},\ \bar{Y})$.

3. **IF** (7.3.5) is feasible **THEN** Take $K = \frac{1}{\gamma^\star}\bar{K}^\star$ **ELSE** goto 1.

4. Take $\mathbb{X}_f^t = \mathbf{conv}(\lambda\gamma^\star V, \Theta^\star)$.

If in Algorithm 6 problem (7.3.5) has a full dimensional solution then it can be shown, by application of Proposition 7.34, p. 254 of [Blanchini & Miani, 2008], that Lemmas 7.3.1, 7.3.2

hold. As in Chapter 5, the main difficulty when using Algorithm 6 is to start from a suitable V. This could be overcome, for instance, by using similar techniques to [Cannon *et al.* , 2003; Fiacchini *et al.* , 2012; Lazar & Jokic, 2010]. On the other hand, the objectives in (7.3.5) are competing. In particular, increasing the set of θ is certainly desirable in steady state, however, it will be soon seen that robustness depends upon the size of $\mathcal{X}_f$. Definition of an optimal compromise is left for future research.

7.4 Computation of the terminal penalty

The terminal cost is computed in a similar way to the previous chapters. In particular, both the cases of partial and total regularisation are considered. Assume (for the moment) that $\mathbb{X}$ and $\mathbb{U}$ are symmetric. In this case, $\mathcal{X}_f$ is also symmetric and it admits the irreducible representations:

$$\mathcal{X}_f = \{\delta x : \|G\delta x\|_\infty \leq \lambda\} \equiv \{\delta x = \bar{X}\zeta : \|\zeta\|_1 \leq \lambda\}, \tag{7.4.1}$$

where $\bar{X}$ is a matrix containing the vertices of $\frac{1}{\lambda}\mathcal{X}_f$. This set is λ-contractive by Lemma 7.3.2, and its Minkowski function, $\Psi_{\mathcal{X}_f}(\delta x) = \|G\delta x\|_\infty \equiv \min_{\{\delta x = X\zeta\}} \|\zeta\|_1$, is a control Lyapunov function in this set, with $\Psi_{\mathcal{X}_f}(f_K(\delta x)) \leq \lambda\Psi_{\mathcal{X}_f}(\delta x)$. Assume, for brevity, that $\beta F_2(x,\zeta) = x^T P x$, with (P, K) satisfying:

Assumption 17. **(A17)** Given $\bar{n}_u$ columns of B, $\bar{B} \in \mathbb{R}^{n\times\bar{n}_u}$, with $1 \leq \bar{n}_u \leq m$, $(A, \bar{B})$ stabilisable, $B = [\bar{B},\ \bullet]$, $K^T = [\bar{K}^T\ 0]$, assume:

$$(A + \bar{B}\bar{K})^T P (A + \bar{B}\bar{K}) - P \leq -(\tau I + Q + K^T R K),\ \tau > 0. \tag{7.4.2}$$

Define the scalar:

$$\bar{\alpha} = \frac{n_s \|SKG^L\|_\infty}{1-\lambda}, \text{with } G^L = (G^T G)^{-1} G^T, \tag{7.4.3}$$

and assume the following:

Assumption 18. **(A18)** Assume that in Definition 7.2.1, $\beta F_2(x,\zeta) = x^T P x$ and that one of the following holds:

1. $S = [0_{n_s\times\bar{n}_u},\ \bullet]$, $K^T = [\bullet,\ 0_{n_x\times(m-\bar{n}_u)}]$, $\alpha = Z = X = 0$, $(F_1(x,\zeta) = 0, \zeta = 0\ \forall x)$,

2. $X = I$, $Z = G$, $F_1(x, \zeta) = \|\zeta\|_\infty$, $\alpha \geq \bar{\alpha}$, $(\zeta = G\delta x_N \; \forall x)$,

3. $Z = I$, $X = \bar{X}$, $F_1(x, \zeta) = \|\zeta\|_1$, $\alpha \geq \bar{\alpha}$.

The following is obtained:

Theorem 7.4.1. Assume **(A11)**, **(A12)**, $\mathbb{X}$, $\mathbb{U}$ symmetric, **(A17)**, **(A18)**, **(A15)**, **(A16)**. Then, **(A13)**, **(A14)** and Theorem 7.2.4 hold.

Proof. (Theorem 7.4.1) Consider the three cases in **(A18)**:

1. In the terminal set it is feasible to use just $\bar{n}_u$ actuators, which are non-regularised, and to set the remaining ones to their steady states and $\zeta = 0$. The results follow by application of standard quadratic MPC arguments ([Rawlings & Mayne, 2010] Theorem 2.24, p. 123) thanks to **(A17)**.

2. We have $F(\zeta, \delta x_N) = \bar{F}(\delta x_N) = \alpha \|G\delta x_N\|_\infty + x_N^T P x_N$. The quadratic part is accommodated by **(A17)**. To satisfy **(A14)**, we need $\alpha\|G(A+BK)x\|_\infty - \alpha\|Gx\|_\infty + \|SKx\|_1 \leq 0, \; \forall x \in \mathcal{X}_f$. Since $\|Gx\|_\infty$ is the Minkowski function of (the λ-contractive) $\mathcal{X}_f$, and G has full column rank, from α satsfing (7.4.3) we have $0 \geq \alpha(\lambda - 1) + n_s\|SK(G^TG)^{-1}G^T\|_\infty$. Multiplying for $\|Gx\|_\infty$ provides $0 \geq \alpha(\lambda-1)\|Gx\|_\infty + n_s\|SK(G^TG)^{-1}G^T\|_\infty\|Gx\|_\infty \geq \alpha(\|G(A+BK)x\|_\infty - \|Gx\|_\infty) + n_s\|SKx\|_\infty \geq \alpha(\|G(A+BK)x\|_\infty - \|Gx\|_\infty) + \|SKx\|_1$, that is the desired result.

3. We have $F(\zeta, \delta x_N) = \alpha\|\zeta\|_1 + x_N^T P x_N$, $Gx = \bar{X}\zeta$. The quadratic part is accommodated by **(A17)**, and $\min_{\{x = \bar{X}\zeta\}} \|\zeta\|_1 \equiv \|Gx\|_\infty$. From contractivity, the ζ^+ required by **(A14)** exists, and we can proceed as in 2).

■

Remark 36. Non-symmetric constraints can be accommodated with minor modifications of **(A18)**, for instance by considering Lemma 5.3.2. It is also possible to take $F_2(x, \zeta) = (F_1(x, \zeta))^2$ and to take β as in Lemma 5.3.2 (similar to [Grammatico & Pannocchia, 2013]) provided that $n\tau\|G^L\|_2^2$ is added the numerator of β. Details are omitted for brevity.

7.5 Enlarged domain of attraction and local optimality

The approach provides an enlargement of the domain of attraction, with respect to an MPC for regulation. Namely:

Lemma 7.5.1. $\mathbb{X}_N^t \cap \mathbb{X} \supseteq \mathbb{X}_N(\theta) \cap \mathbb{X}$, for all $\theta : M_\theta \theta \in \mathbb{Z}_s$.

This has been widely discussed in, among others, [Ferramosca *et al.* , 2008; Zeilinger *et al.* , 2010], to which the reader is referred for the sake of brevity. Local optimality of admissible references can also be characterised, as follows:

Assumption 19. **(A19)** [Ferramosca *et al.* , 2008; Zeilinger *et al.* , 2010]. The parameters γ_t, γ_c are taken to be greater then the supremum over x of the ∞-norm of the respective Lagrange multiplier, for (7.2.3)–(7.2.4) subject to $\|\underline{s}\|_1 = 0$ and $\|z_t - z_s\|_1 = 0$, for $z_t \in \mathbb{Z}^o \subseteq \mathbb{Z}_s$.

Theorem 7.5.2. Assume **(A19)**. Then, for $z_t \in \mathbb{Z}^o$, at the optimum of (7.2.3)–(7.2.4) we have, for $z_t = M_\theta \theta$, that $z_s^\star = z_t, \forall x(0) \in \mathbb{X}_N(\theta)$.

Notice that, for $z_t \in \mathbb{Z}_s \backslash \mathbb{Z}^o$, $\mathbb{X}$ can also be expected to be invariant, and $z_s^\star = z_t$ holds locally [Ferramosca *et al.* , 2008]. The required (exact) penalties can be computed, as discussed in [Ferramosca *et al.* , 2008; Zeilinger *et al.* , 2010] as well as in the previous chapter, by means of the algorithms in [Alessio & Bemporad, 2009]. A proof of Theorem 7.5.2 requires only minor modification of the results in [Ferramosca *et al.* , 2008; Zeilinger *et al.* , 2010] and is therefore omitted.

7.6 Robustness bound and local ISS gain

As shown in [Zeilinger *et al.* , 2010], system (7.2.8) is intrinsicly ISS. This section provides sufficient conditions for robust feasibility and computes a local worst-case ISS gain.

The proposed robustness bound follows:

Assumption 20. **(A20)** $\mathbb{W} \subseteq \{w \in \mathbb{R}^q : \|w\|_\infty \leq \mu\}$, with

$$\mu \leq \frac{1-\lambda}{\|GA^{N-1}B_w\|_\infty}. \tag{7.6.1}$$

Denote the optimal values for $\delta\hat{x}_j,\ \hat{u}_j$ as $\delta\hat{x}_j^\star, \delta\hat{u}_j^\star$. Define also $\nabla V_N^o(\delta x) = V_N^o(\delta x^+) - V_N^o(\delta x)$. Define $\delta\theta = \tilde{\theta} - \bar{\theta}$, $\bar{A} = (A + BK)$, $E = (-KM_x + M_u)$. The main result follows:

Theorem 7.6.1. Assume **(A11)**–**(A16)**, **(A20)**. Then, $\forall w \in \mathbb{W}$,

1. $\mathbb{X}_N^t$ is RPI, and system (7.2.8) is ISS, $\forall x(0) \in \mathbb{X}_N^t$,

2. If **(A19)** holds and $z_t = M_\theta \bar{\theta} \in \mathbb{Z}^o$, then $\forall x(0) \in \mathbb{X}_N(\bar{\theta})$ and for any $\tau_Q \in (0,\ \lambda_{\min}(Q))$ the value function satisfies:

$$\nabla V_N^o(\delta x(k)) \leq -\tau_Q \|\delta x(k)\|_2^2 + \sigma(\|w(k)\|_2),$$

with local ISS gain

$$\begin{aligned}
\sigma(r) = \sigma_Q(r) + \sigma_S(r) + \sigma_R(r) + \sigma_P(r) + \sigma_G(r) + \sigma_s(r) + \sigma_O(r),\\
\sigma_Q(r) = 2\epsilon_Q \left(\sum_{j=1}^{N} \|Q^{1/2} A^{j-1} B_w\|_2^2 + \gamma_Q \right) r^2,\\
\sigma_S(r) = (\|SKA^{N-1}B_w\| + \sqrt{q} + \gamma_S r) r,\\
\sigma_R(r) = \epsilon_Q \left(2\|R^{1/2} K A^{N-1} B_w\|_2^2 + \gamma_R \right) r^2\\
\sigma_P(r) = 2\left(1 + \frac{1}{\epsilon_P}\right) (\|P^{1/2} \bar{A} A^{N-1} B_w\|_2^2 + \gamma_P) r^2,\\
\sigma_G(r) = (\|G\bar{A}A^{N-1}B_w\|_\infty + \gamma_G) r,\\
\sigma_s(r) = \gamma_s r,\ \sigma_O(r) = \gamma_O r.
\end{aligned}$$

The (worst case) gains $\gamma_\bullet$ are:

$$\begin{aligned}
\gamma_Q = 2N\|Q\|_2 \max \|M_x(\tilde{\theta} - \bar{\theta})\|_2^2 / \|w\|_2^2,\\
\gamma_S = (N-1)\|S\|_1 (\max \|M_u(\tilde{\theta} - \bar{\theta})\|_1 / \|w\|_2 + \max \|(E - M_u)(\tilde{\theta} - \bar{\theta})\|_1 / \|w\|_2),\\
\gamma_R = \|R\|_2 ((N-1) \max \|M_u(\tilde{\theta} - \bar{\theta})\|_2^2 / \|w\|_2^2 + 2 \max \|(E - M_u)(\tilde{\theta} - \bar{\theta})\|_1 / \|w\|_2),\\
\gamma_P = \|P\|_2 \max \|(BE - M_x)(\tilde{\theta} - \bar{\theta})\|_2^2 / \|w\|_2^2,\\
\gamma_G = \|G\|_\infty \max \|(BE - M_x)(\tilde{\theta} - \bar{\theta})\|_\infty / \|w\|_2,
\end{aligned}$$

where the (finite) $\max$ is taken over $(x_N,\ w)$, where $\tilde{\theta}$ solves the following:

$$\max_{\tilde{\theta}} V_O(M_\theta(\tilde{\theta} - \bar{\theta})), \tag{7.6.2}$$

$$\text{s.t. } \frac{1}{\lambda} G_x(x_N - M_x\bar{\theta}) + G_\theta\, \bar{\theta} \leq \underline{1}, \tag{7.6.3}$$

$$G_x(x_N - M_x\tilde{\theta} + A^{N-1}B_w w) + G_\theta \tilde{\theta} \leq \underline{1},\ w \in \mathbb{W}. \tag{7.6.4}$$

The remaining constants are $\epsilon_P = \tau/\|P\|_2$, and

$$\gamma_O = \max V_O(M_\theta(\tilde{\theta} - \bar{\theta}))/\|w\|_2,$$
$$\gamma_s = \max |V_s(\tilde{s}) - V_s(\bar{s})|/\|w\|_2,$$
$$\epsilon_Q = \frac{N(\kappa_x^2\|Q\|_2 + \kappa_u^2\|R\|_2)}{\lambda_{\min}(Q) - \tau_Q} + 1,$$

with $\kappa_x, \kappa_u > 0$ from Lemma 7.6.2. For γ_s, the (finite) max is taken over $(\underline{x}, w)$, $\underline{x} = (\hat{x}_0, \ldots, \hat{x}_{N-1})$, under $(x_j, w) \in \mathbb{X}_N^t \times \mathbb{W}$ for $j \in \mathbb{I}_{N-1}$ and where, given a pair $(\underline{x}, w)$, the vectors $\bar{s}$, $\tilde{s}$ minimise V_s subject to, respectively,

$$Lx_j - \underline{1} \leq \bar{s}_j,\ L(x_j + A^j B_w w) - \underline{1} \leq \tilde{s}_j.$$

In order to prove Theorem 7.6.1, a couple of lemmas are needed.

Lemma 7.6.2. $\forall x \in \mathbb{X}_N^t, \forall z_t \in \mathbb{R}^{n+m},\ \exists \kappa_x, \kappa_u > 0 : \|\delta\hat{x}_j^\star\|_2 \leq \kappa_x\|\delta x\|_2,\ \forall j \in \mathbb{I}_{[0,\ N]}$ and $\|\delta u_j^\star\|_2 \leq \kappa_u\|\delta x\|_2,\ \forall j \in \mathbb{I}_{[0,\ N-1]}$.

Proof. (Lemma 7.6.2) This follows from the fact that $\mathbb{X}$, $\mathbb{U}$ are bounded and, from Theorem 2.6.10, the explicit solution of (7.2.3) is PWA in x and z_t. ■

The following norm identity is key for computing the proposed ISS gain:

Lemma 7.6.3. Given $a, b \in \mathbb{R}^n$, then

$$\|a + b\|_2^2 \leq (1 + \epsilon^2)\|a\|_2^2 + (1 + 1/\epsilon^2)\|b\|_2^2,\ \forall \epsilon \in \mathbb{R}_{>0}. \tag{7.6.5}$$

Proof. (Lemma 7.6.3) Recall that $\|a + b\|_2^2 = \|a\|_2^2 + \|b\|_2^2 + 2a^Tb$. Suppose first the result to hold for $\epsilon = 1$. Then, we have $2a^Tb \leq a^Ta + b^Tb, \forall a, b \in \mathbb{R}^n$ and we need to show that this holds also for $\hat{a} = \epsilon a$ and $\hat{b} = 1/\epsilon b$, if $\epsilon \neq 0$. In particular, differentiating $\hat{a}^T\hat{a} + \hat{b}^T\hat{b}$ with respect to ϵ provides $2\epsilon a^2 - 1/2\frac{1}{\epsilon^3}b^2$ which is equal to zero for $\epsilon^2 = \frac{b}{2a}$, thus providing (the minimum) $\hat{a}^T\hat{a} + \hat{b}^T\hat{b} = 5/2a^Tb \geq 2a^Tb = 2\hat{a}^T\hat{b}$. To show that this is a minimum, we differentiate further in ϵ to have $2a^2 - \frac{1}{6\epsilon^4}b^2$, which is positive for $\epsilon^2 = \frac{b}{2a}$ (and equal to $5/3a^2$). The choice of $\epsilon > 0$ is made without loss of generality.

Finally, it has to be proven that $2a^Tb = a^Tb + b^Ta \leq a^Ta + b^Tb, \forall a, b \in \mathbb{R}^n$. If $a = b$, this is trivial. Let us first assume that $a = b + c$, with $a_i - b_i = c_i \geq 0, \forall i \in \mathbb{I}_{[1,n]}$. Then we have

$(b+c)^T b + (a-c)^T a = b^T b + a^T a + c^T(b-a) \leq b^T b + a^T a$. A similar result (with strict inequality) can be obtained by taking $b = a + c$ with $c_i > 0, \forall i \in \mathbb{I}_{[1,n]}$. ■

We are now ready to prove the main result:

Proof. (Theorem 7.6.1)

1. Suppose $w(k) \neq 0$, for some $k \geq 0$. Define a candidate $N-1$ steps-ahed MPC prediction, from time $k+1$, as $\tilde{x}_{N-1|k+1} = \hat{x}^\star_{N|k} + A^{N-1}B_w w(k)$, where $\hat{x}^\star_{N|k}$ denotes $\hat{x}^\star_N$ at time k. In order to show that $\mathbb{X}^t_N$ is RPI it is sufficient to show that, given $\|w(k)\|_\infty \leq \mu$, there exists a feasible control move $u_{\tilde{\theta}}$ which steers $\tilde{x}_{N-1|k+1}$ to a new point $\tilde{x}_{N|k+1} \in \mathbb{X}_f(\tilde{\theta}) + M_x\tilde{\theta}$ for some $\tilde{\theta}$. Define the Minkowski function of $\mathcal{X}_f$ as $\psi(x) = \max_i G_i x$. Then $\psi(\delta\tilde{x}_{N-1|k+1}) = \psi(\delta\hat{x}^\star_{N|k} + A^{N-1}B_w w) \leq \psi(\delta\hat{x}^\star_{N|k}) + \psi(A^{N-1}B_w w) \leq \lambda + \|GA^{N-1}B_w w\|_\infty \leq \lambda + \|GA^{N-1}B_w\|_\infty \|w\|_\infty \leq \lambda + \|GA^{N-1}B_w\|_\infty \mu$. Recall (7.4.1). Taking μ from (7.6.1) provides $\psi(\delta\tilde{x}_{N-1|k+1}) \leq 1$, and $\delta\tilde{x}_{N-1|k+1} \in \frac{1}{\lambda}\mathcal{X}_f$. From Definition 2.2.17 it follows that $\exists\tilde{\theta} : \tilde{x}_{N-1|k+1} - M_x\tilde{\theta} \in \frac{1}{\lambda}\mathbb{X}_f(\tilde{\theta})$. Recall (7.2.16). By Lemma 7.3.1 the control move $u_{\tilde{\theta}} = \kappa(\tilde{x}_{N-1|k+1}, \tilde{\theta})$ is feasible and steers $\tilde{x}_{N-1|k+1}$ to $\tilde{x}_{N|k+1}$, which satisfies $\tilde{x}_{N|k+1} - M_x\tilde{\theta} \in \mathbb{X}_f(\tilde{\theta})$. Therefore the MPC problem (7.2.3) is robustly feasible and $\mathbb{X}^t_N$ is RPI for the closed-loop system (7.2.8) which, from Theorem 2.6.9, is Input-to-State Stable.

2. Take $\bar{u} = \{\hat{u}^\star_{1|k}, \ldots, \hat{u}^\star_{N-1|k}, \kappa(\hat{x}^\star_{N|k}, \theta^\star_k)\}$ and $\bar{s} = \{s^\star_{1|k}, \ldots, s^\star_{N-1|k}, 0\}$ as the tails of the previous optimal control and slacks sequences, plus the terminal control law and zero. Recall that $z_t = M_\theta\bar{\theta}$. From Theorem 7.5.2 we have $\theta^\star = \bar{\theta}, \forall x \in \mathbb{X}_N(\bar{\theta})$ and, $\forall x \in X_f(\bar{\theta})$. Denote the perturbed predictions under $w(k)$ as $\tilde{x}_{i-1|k+1} = \hat{x}^\star_{i|k} + A^{i-1}B_w w(k)$ for $i = 1, \ldots N$. From part 1, $\exists\tilde{\theta}$ providing feasibility at time $k+1$ of the control $\tilde{u}_{N|k} = K(\tilde{x}_{N-1|k+1} - M_x\tilde{\theta}) + M_u\tilde{\theta}$, and of the new prediction $\tilde{x}_{N|k+1} = A\tilde{x}_{N-1|k+1} + B\tilde{u}_{N|k}$. Denote $\tilde{u}$ as $\bar{u}$ with terminal move replaced by the feasible $\tilde{u}_{N|k}$. Denote $\tilde{s}$ as the new vector of slack variables, for the perturbed predictions $\tilde{x}_{\bullet|k+1}$. Define $\delta\tilde{x}_{\bullet|k+1} = \tilde{x}_{\bullet|k+1} - M_x\bar{\theta}$ and $\delta\tilde{u}_{j|k+1} = \tilde{u}_{j|k+1} - M_u\bar{\theta}$, where $\tilde{u}_{j|k+1}$ is the $j+1$ element of $\tilde{u}$. Similarly define $\delta\bar{u}_{i|k} = \bar{u}_{i|k} - M_u\bar{\theta}$ where $\bar{u}_{i|k}$ is the i-th element of $\bar{u}$. Denote $\bar{x}_{i|k}$ as the i-th element of the list $\bar{x} = \{\hat{x}^\star_{1|k}, \ldots, \hat{x}^\star_{N|k}, \bar{x}_{N+1|k}\}$ with $\bar{x}_{N+1|k} = A\hat{x}^\star_{N|k} + \kappa(\hat{x}^\star_{N|k}, \bar{\theta})$, and $\delta\bar{x}_{\bullet|k} = \bar{x}_{\bullet|k} - M_x\theta$. Define $\phi_N(w) = A^{N-1}B_w w$.

Notice that, by the above definitions and by means of (7.2.24), we have:

$$\tilde{u}_{N-1|k+1} = \kappa(\bar{x}_{N|k}, \bar{\theta}) + E\delta\theta + K\phi_N(w) \tag{7.6.6}$$

$$\delta\tilde{u}_{N-1|k+1} = \delta\bar{u}_{N|k} + E\delta\theta + K\phi_N(w), \tag{7.6.7}$$

$$\tilde{x}_{N|k+1} = \bar{A}\bar{x}_{N+1|k} + BE\tilde{\theta} + \bar{A}\phi_N(w), \tag{7.6.8}$$

$$\delta\tilde{x}_{N|k+1} = \bar{A}\delta\bar{x}_{N+1|k} + BE\delta\theta + \bar{A}\phi_N(w), \tag{7.6.9}$$

Then by optimality:

$$\begin{aligned}
\nabla V_N^o(\delta x) \leq V_N(\tilde{x}_{0|k+1}, \tilde{u}, \tilde{\theta}, \tilde{s}) - V_N^o(\delta x) = \\
V_N(\tilde{x}_{0|k+1}, \tilde{u}, \tilde{\theta}, \tilde{s}) - V_N(\hat{x}^\star_{1|k}, \bar{u}, \bar{\theta}, \bar{s}) \\
+V_N(\hat{x}^\star_{1|k}, \bar{u}, \bar{\theta}, \bar{s}) - V_N^o(\delta x) \leq \\
\sum_{i=1}^{N} \|Q^{1/2}(\delta\tilde{x}_{i-1|k+1} - M_x\delta\theta)\|_2^2 - \|Q^{1/2}\delta\hat{x}^\star_{i|k}\|_2^2 \qquad (7.6.10)\\
+\sum_{i=1}^{N-1} \|R^{1/2}(\delta\tilde{u}_{i-1|k+1} - M_u\delta\theta)\|_2^2 - \|R^{1/2}\delta\bar{u}_{i|k}\|_2^2 \qquad (7.6.11)\\
+\|R^{1/2}(\delta\tilde{u}_{N-1|k+1} - M_u\delta\theta)\|_2^2 - \|R^{1/2}\delta\bar{u}_{N|k}\|_2^2 \qquad (7.6.12)\\
+\sum_{i=1}^{N-1} \|S(\delta\tilde{u}_{i-1|k+1} - M_u\delta\theta)\|_1 - \|S\delta\bar{u}_{i|k}\|_1 \qquad (7.6.13)\\
+\|S(\delta\tilde{u}_{N-1|k+1} - M_u\delta\theta)\|_1 - \|S\delta\bar{u}_{N|k}\|_1 \qquad (7.6.14)\\
+\|G(\delta\tilde{x}_{N|k+1} - M_x\delta\theta)\|_\infty - \|G\delta\bar{x}_{N+1|k}\|_\infty \qquad (7.6.15)\\
+\|P^{1/2}(\delta\tilde{x}_{N-1|k+1} - M_x\delta\theta)\|_2^2 - \|P^{1/2}\delta x_{N+1|k}\|_2^2 \qquad (7.6.16)\\
-\tau\|\delta\bar{x}_{N|k}\|_2^2 + V_O(-M_\theta\delta\theta) + V_s(\tilde{s}) - V_s(\bar{s}) \qquad (7.6.17)\\
-(\lambda_{\min}(Q) - \tau_Q + \tau_Q)\|\delta x\|_2^2 - \ell(s^\star_{1|k}), \qquad (7.6.18)
\end{aligned}$$

where (7.6.18) provides the same condition as for the nominal case. We want to upper-bound the functions above (7.6.18) to construct the ISS gain $\sigma(\|w\|_2)$. Recall that $\delta\tilde{x}_{i-1|k+1} = \delta\hat{x}^\star_{i|k}$ for $i = 1, \ldots, N$ and $\delta\tilde{u}_{i-1|k+1} = \delta\bar{u}_{i|k}$ for $i = 1, \ldots, N-1$. Using the triangle inequality, (7.6.13) is upper-bounded by

$$(N-1)\|SM_u\delta\theta\|_1. \tag{7.6.19}$$

Substituting (7.6.7) into (7.6.14) and using again the triangle inequality we have the upper-bound

$$\|S(E-M_u)\delta\theta\|_1 + \|SKA^{N-1}B_w\|_1\sqrt{q}\|w(k)\|_2. \tag{7.6.20}$$

Substituting (7.6.9) into (7.6.15), the latter is upper bounded by

$$\|G(BE-M_x)\delta\theta\|_\infty + \|G\bar{A}A^{N-1}B_w\|_\infty\|w(k)\|_2. \tag{7.6.21}$$

For any $\epsilon \neq 0$ Lemma 7.6.3 provides the following upper-bound, respectively for (7.6.10) and (7.6.11):

$$\begin{aligned}\epsilon^2\sum_{i=1}^{N}\|Q\|_2\|\delta\hat{x}^\star_{i|k}\|_2^2 + 2\left(1+\frac{1}{\epsilon^2}\right)\sum_{i=1}^{N}\|Q^{1/2}A^{i-1}B_w\|_2^2\|w(k)\|_2^2&\\ +2N\left(1+\frac{1}{\epsilon^2}\right)\|Q^{1/2}M_x\delta\theta\|_2^2,& \qquad (7.6.22)\\ \epsilon^2\sum_{i=1}^{N-1}\|R\|_2\|\delta\hat{u}^\star_{i|k}\|_2^2 + (N-1)\left(1+\frac{1}{\epsilon^2}\right)\|R^{1/2}M_u\delta\theta\|_2^2.& \qquad (7.6.23)\end{aligned}$$

For (7.6.12) and (7.6.16) we have, respectively:

$$\begin{aligned}\epsilon^2\|R\|_2\|\delta\hat{u}^\star_{N|k}\|_2^2 + 2\left(1+\frac{1}{\epsilon^2}\right)\|R^{1/2}(E-M_u)\delta\theta\|_2^2&\\ +2\left(1+\frac{1}{\epsilon^2}\right)\|R^{1/2}KA^{N-1}B_w\|_2^2\|w(k)\|_2^2,& \qquad (7.6.24)\\ \epsilon^2\|P\|_2\|\delta\bar{x}_{N+1|k}\|_2^2 + 2\left(1+\frac{1}{\epsilon^2}\right)\|P^{1/2}(BE-M_x)\delta\theta\|_2^2&\\ +2\left(1+\frac{1}{\epsilon^2}\right)\|P^{1/2}\bar{A}A^{N-1}B_w\|_2^2\|w(k)\|_2^2,& \qquad (7.6.25)\end{aligned}$$

The next step consists of eliminating the terms depending on the nominal predictions from the above inequalities. First of all recall that from Lemma 7.6.2 there exist $\kappa_u, \kappa_x : \|\delta u^\star_{i|k}\|_2^2 \leq \kappa_u^2\|\delta x(k)\|_2^2$ and $\|\delta x^\star_{i|k}\|_2^2 \leq \kappa_x^2\|\delta x(k)\|_2^2, \forall i$. Substitute this identify for the predictions in (7.6.22), (7.6.23), (7.6.24). In these equations, we want all the constants multiplying $\|\delta x(k)\|_2^2$ to be less or equal to $\lambda_{\min}(Q) - \tau_Q$, for some appropriate τ_Q, so that the positive terms can be cancelled out, leaving just $-\tau_Q\|\delta x\|_2^2$. It can be easily verified that this happens for $\left(1+\frac{1}{\epsilon^2}\right) \geq \epsilon_Q$. For (7.6.25) we can proceed in a similar way to verify that

$\epsilon^2 \leq \tau/\|P\|_2 = \epsilon_P$ (where $\tau > 0$ is from **(A17)**) is sufficient to cancel out the term depending on $\bar{x}_{N+1|k}$.

The above arguments allow one to extract from (7.6.19-7.6.25) the elements of the gains $\sigma_{\bullet}(\|w\|_2)$, with the exception of the $\gamma_{\bullet}$. In order to conclude the proof, we will show that the result of the maximisation in $\gamma_{\bullet}$ is finite, so that the $\gamma_{\bullet}$ can upper-bound the terms of (7.6.19-7.6.25) depending on $\delta\theta$ (this can then be verified by inspection). As discussed in the proof of Lemma 7.6.2, the sets $\mathbb{X}$, $\mathbb{U}$, $\mathbb{X}_f^t$, $\mathbb{X}_N^t$, $\mathbb{W}$ are closed and bounded and, from Theorem 2.6.10, the explicit solution of (7.2.3) is continuous and PWA in x and z_t. Recall that $z_t = M_\theta\bar{\theta}$ is fixed by assumption. Then $\theta_k^\star$ and $\underline{s}_k^\star$ are continuous in $x \in \mathbb{X}_N^t$ and, if $w(k) = 0$, then $\theta_k^\star = \bar{\theta}$ by **(A19)** and $\mathbb{X}_N(\bar{\theta})$ is positively invariant. Therefore $w(k) = 0 \Rightarrow \theta_{k+1}^\star = \theta_k^\star \Rightarrow \delta\theta = 0$. This means that $\max\|\delta\theta\|/\|w(k)\|$ with $\tilde{\theta}$ solving (7.6.2–7.6.4) is finite. The same holds for γ_s, notice that if $w(k) = 0$ then $\tilde{s} = \bar{s}$ is also feasible. This means that the $\max$ for the $\gamma_{\bullet}$ are finite, since each of them is taken from a continuous function over a bounded domain. ■

Condition (7.6.1) is only sufficient for robust feasibility. In general, decreasing λ is beneficial to μ, unless the terminal set becomes too small. Larger values of μ could also be allowed by taking a smaller α_z. The proposed approach can provide a larger RPI set than the one obtained with the regulation approach, in other words, we have

$$\mathbb{X}_N^t \supseteq \mathbb{X}_N(\theta), \; \forall\theta \in \mathbb{R}^{n_\theta}.$$

Note that the above results holds also when $\mathbb{X}_N(\bar{\theta}) \supset \mathbb{X}$, which is possible using soft-constraints. On the other hand, it should be noticed that the proposed local ISS gain, $\sigma(\cdot)$, can be quite conservative.

Remark 37. If $F_2(x,\zeta) = (F_1(x,\zeta))^2$ is used, as in Remark 36, then the results of Theorem 7.6.1 still hold, with the substitution of $\beta\bar{b}_2$ for $\|P\|_2$.

The ISS gain allows one to speculate about the closed-loop behaviour of the system, under the effect of vanishing or persistent disturbances. In particular, from Theorem 7.6.1 it can be noticed that increasing $\lambda_{\min}(Q)$, as well as τ, can reduce the gain and provide a stronger disturbance rejection and a smaller offset for constant disturbances.

7.6.1 Achievable steady states and constraint violation

The proposed terminal set, $\mathbb{X}_f$, is based on the contraction of its projection on δx. In (A16), robustness is obtained by only restricting the space of δx (scaling by $\lambda < 1$). On the other hand, the set of achievable terminal states is the same as the ones obtained by using the maximum contractive set, $\mathcal{O}^t_{\infty,\lambda,\alpha_z}$. By definition, these steady states are in the interior of $\mathbb{X}$, and could even consist of the whole set without its boundary. However, with the proposed approach no guarantees are provided for state constraint satisfaction, under a disturbance satisfying (A20). The only guarantees are robust feasibility under soft constraints and ISS. Increasing the Q matrix (and in a sense the MPC "gain") can provide stronger disturbance rejection, however, if we are operating at the limit of constraints, the repeated occurrence of disturbances could cause further constraint violation. If one wishes to prevent or limit constraint violation, it is possible for instance to restrict the set of admissible steady states by taking the following choice as an alternative to (A16):

$$\mathbb{X}^t_f = \{(\delta x, \theta) : (\delta x, \theta) \in \lambda \mathcal{O}^t_{\infty,\lambda,\alpha_z}\}. \tag{7.6.26}$$

Thanks to part 1. of Lemma 7.3.2, if (7.6.26) hold it follows that $\mathcal{X}_f \equiv \lambda \pi_{\delta x}(\mathcal{O}^t_{\infty,\lambda,\alpha_z})$ satisfies Theorem 7.6.1. Conditions for recursive feasibility and ISS remain unchanged, however, we have sacrificed some of the achievable steady states for having higher chances to stay inside $\mathbb{X}$ when a disturbance hits the system.

7.6.1.1 Comparison with constraint restriction

In the next example, it will be shown that the proposed approach, even when (7.6.26) is used instead of (A16), can be less conservative than another existing strategy. The considered approach, proposed in [Ferramosca *et al.* , 2012], makes use of constraints restriction, namely, constraints are time varying and defined by $\tilde{\mathbb{X}}_{j+1} = \mathbb{X}_j \sim (A + B\tilde{K})B_w\mathbb{W}$, $\tilde{\mathbb{U}}_{j+1} = \mathbb{U}_j \sim K(A + B\tilde{K})B_w\mathbb{W}$ for some stabilising $\tilde{K}$, with $\tilde{\mathbb{X}}_0 = \mathbb{X} \sim B_w\mathbb{W}$, $\tilde{\mathbb{U}}_0 = \mathbb{U} \sim KB_w\mathbb{W}$. Several versions of this have been considered in MPC for regulation, for instance [Blanchini & Miani, 2008; Chisci *et al.* , 2001; Goulart *et al.* , 2006; Kuwata *et al.* , 2007; Richards & How, 2006; Shekhar & Maciejowski, 2012]. In particular, in [Shekhar & Maciejowski, 2012] an auxiliary control policy is constructed to obtain less conservative predictions and a larger DOA than most standard approaches. In [Ferramosca *et al.* , 2012], the terminal set is taken to be a a particular RPI set that considers the new constraints, then further restricted (N times) by $\mathbb{W}$. This approach provides robust constraints satisfaction, as shown in [Ferramosca *et al.* , 2012]. However, the obtained region of attraction and attainable steady states set can be quite small. This comes from the fact that $(\mathbb{X} \sim \mathbb{W}) \oplus \mathbb{W} \subseteq \mathbb{X}$ (see Lemma 2.2.2).

The use of constraint restriction is quite standard in robust MPC theory. However, the required offline computation is certainly increased from the nominal case in order to obtain the RPI set. In many cases, the RPI set could be empty, or the number of iterations required for its computation could be excessively large, meaning a different set $\mathbb{W}$ needs to be found. On the other hand, our proposed approach makes use of a λ, the range of which is known to return a sufficient disturbance bound, μ, for which robust feasibility is guaranteed. The terminal set is always guaranteed to exist, and its computation requires similar complexity as for the nominal case. In the next Section, it will be demonstrated that the proposed approach can provide a set of achievable steady states significantly larger than with the benchmark approach, while the same constraints satisfaction can be achieved (with no formal guarantees).

7.7 Illustrative example

The approach is demonstrated for a double integrator, $n_x = 2,\ n_y = n_u = n_\theta = 1$, $B_w = [1\ 1]^T$, $y = [1\ 0]x$, sampled at 1 Hz. The parameters are $Q = 2I,\ R = 2$, $\tau = 1$, $S = 1$, $\bar{\alpha} = 6.7$, $N = 5$. Constraints are $\|x\|_\infty \leq 10$, $|u| \leq 5$. The candidate terminal controller is an LQR placing the eigenvalues at about 0.4. All plots are made using the the Matlab MPT toolbox [Kvasnica *et al.* , 2004]. Considered scenarios are:

- `Case 1` - Robust soft-constrained LTI tracking using (A16).
- `Case 2` - Robust soft-constrained LTI tracking using (7.6.26).
- `Case 3` - Approach of [Ferramosca *et al.* , 2012] with the addition of soft-constraints.

Figure 7.1 shows the projections on δx of, in order of size from the largest, the nominal invariant set for tracking, the contractive set computed with the proposed approach (same for both `Case 1` and `Case 2`) with $\lambda = 0.7, \mu = 0.4$, and the RPI set computed with the constraint restriction approach of [Ferramosca *et al.* , 2012] (`Case 3`) for the same μ and the LQR as disturbance affine feedback ($\bar{K} = K$). The full sets $\mathbb{X}_f^t$ are shown in Figure 7.2 and the attainable steady states are plotted against δx_1 in Figure 7.3.

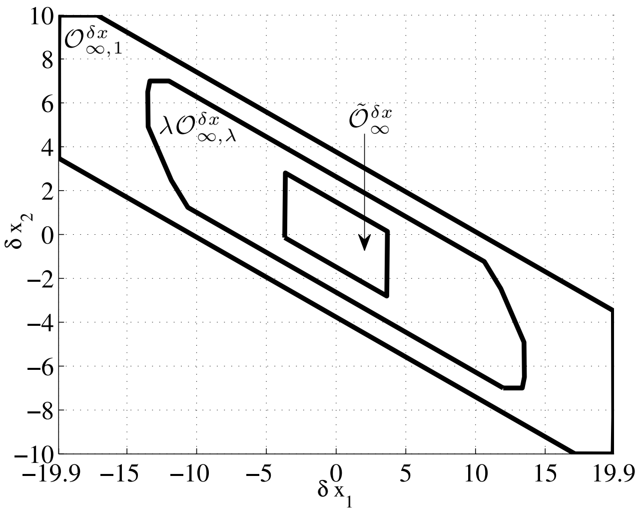

Figure 7.1: Terminal sets (on δx): nominal, proposed (same for both `Case 1` and `Case 2`), based on [Ferramosca *et al.* , 2012] (`Case 3`). $\lambda = 0.7,\ \mu = 0.4$.

Figure 7.4 shows the closed loop trajectory of the system for `Case 1`, under admissible disturbances (w_1, w_2) in the vanishing and constant case. Notice that constraints are initially violated for two samples because of the high initial velocity. The target is initially set to the edge of the constraints and then to an infeasible position, for which all algorithms achieve the minimum offset steady state. For vanishing disturbances, in `Case 1` constraints are violated at each occurrence for only one sample. For constant disturbances, constraints are violated for most of the time and in steady state, with an offset of 1.5. The feasible region is, however, robustly invariant as expected, and the input and velocity (x_2) are non-zero at steady, having $x_2 = -\mu$. For `Case 2`, in Figure 7.5, even tracking the edge of $\mathbb{Z}_s$, the constraints are not violated in steady state. This is because we have sacrificed some of the achievable steady states. The system is robustly stable with a slightly increased offset and $x_2 = -\mu$. Constraints are violated only for the first 2 samples, when a position disturbance is experienced (in practice the system tolerates greater disturbances). For `Case 3`, in Figure 7.6, constraints are violated initially, as this is inevitable. The steady state offset is relatively large. For `Case 1`, any feasible y_t is achievable. For `Case 2`, the target must satisfy $|y_t| \leq 6.9$, while `Case 3` can achieve only $|y_t| \leq 1.1$. For `Case 1` and `Case 3`, the attainable steady-states do not seem to change with λ and μ, while they reduce for `Case 2`. Therefore, `Case 2` can become less inconvenient for very small values of λ. For $\lambda = 0.65,\ \mu = 0.4677$, the terminal sets are all reduced on δx, as shown in Figure 7.7, the set becoming very small for `Case 3`. For

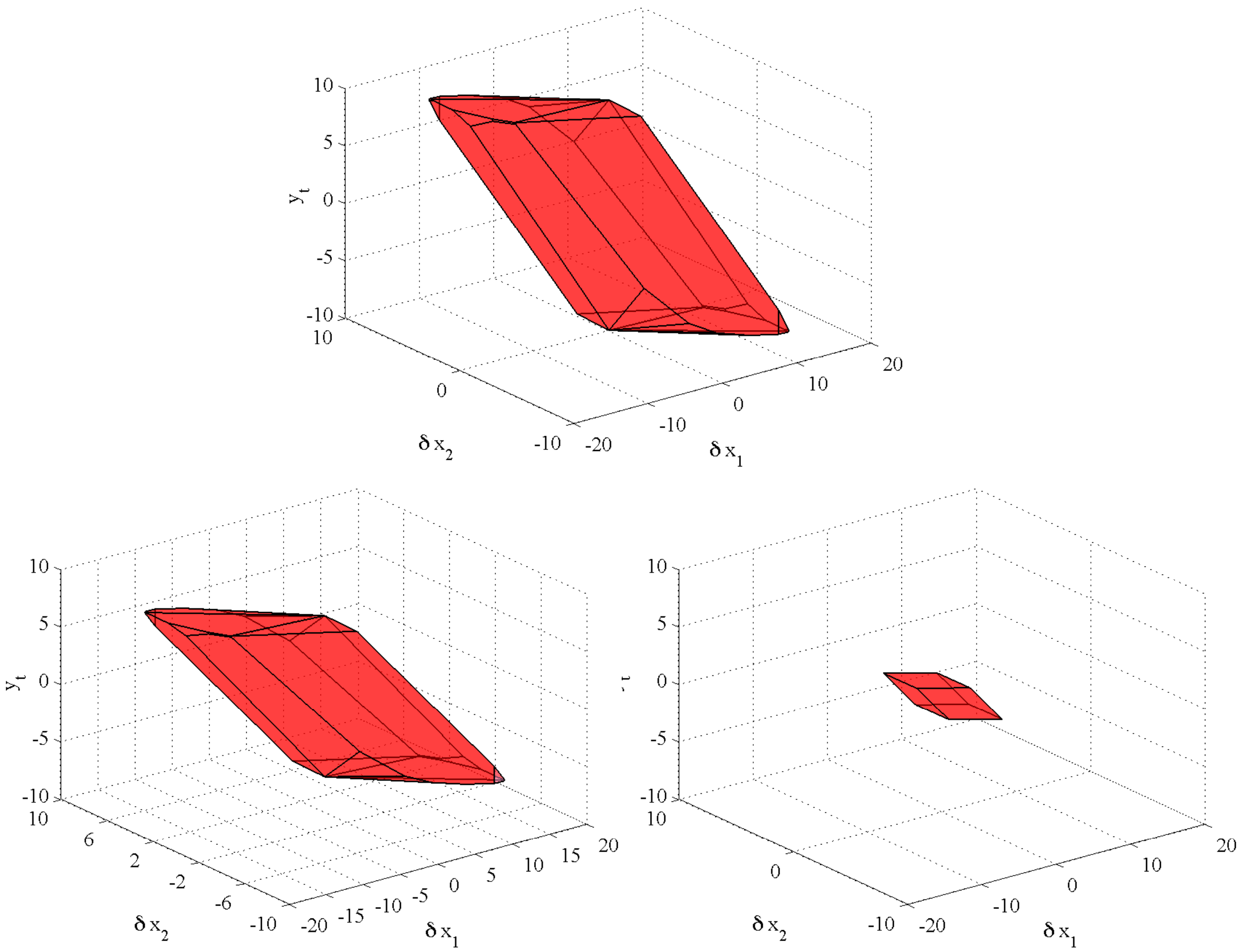

Figure 7.2: Terminal set, $\mathbb{X}_f^t$, for `Case 1` (top), `Case 2` (left) and `Case 3` (right). $\lambda = 0.7,\ \mu = 0.4$.

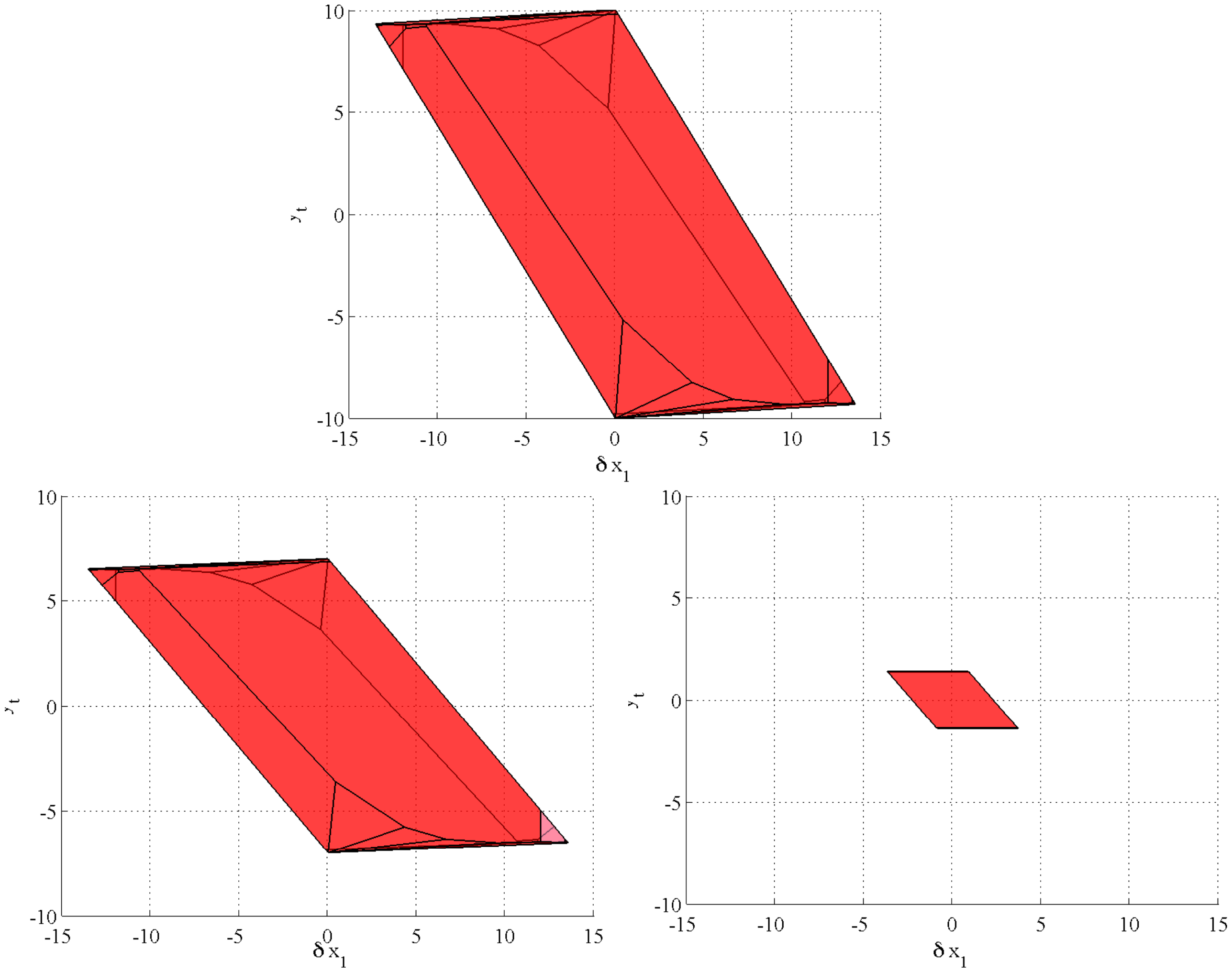

Figure 7.3: Attainable steady states, for `Case 1` (top), `Case 2` (left) and `Case 3` (right). $\lambda = 0.7,\ \mu = 0.4$.

$\lambda = 0.6,\ \mu = 0.5$ the RPI set for `Case 3` is empty and the approach cannot be used. Notice that the results for `Case 3` depend on $\bar{K}$, as well as on K.

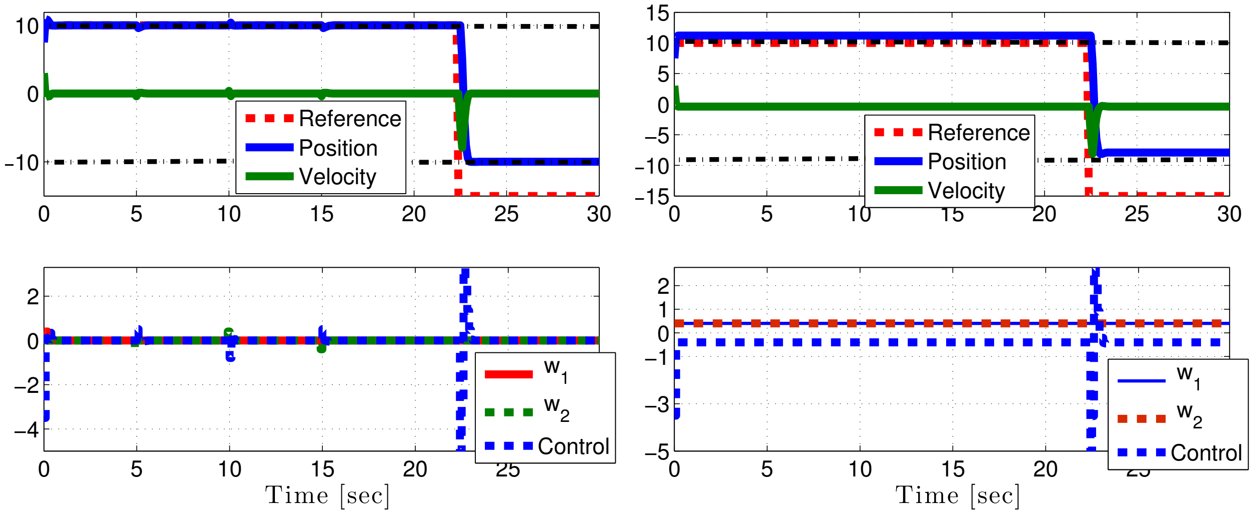

Figure 7.4: Closed-loop trajectory with a moving reference and disturbance (vanishing - left, constant - right). `Case 1`, $\lambda = 0.7,\ \mu = 0.4,\ S = 1$.

In general, for decreasing λ we can, to a certain extent, expect μ to increase, until $\mathcal{X}_f$ becomes very small (as λ approaches $|\lambda_{\max}(A + BK)|$, with possible computational issues). The proposed approaches allow for a wider range of steady states than the benchmark, which, on the other hand, has the advantage of robust constraint satisfaction. In practice, the two strategies behave quite similarly when constraints are (in most cases inevitably) softened, the proposed one benefiting from lower computation requirements (comparable to the nominal case).

Simulations performed for different values of S and asymptotically vanishing disturbances are shown, for `case` 1 and `case` 2, in Figure 7.8, 7.9. In particular, the reference is set to be on the boundary of the achievable steady states. Noticeably, for $S = 1$, the MPC counteracts disturbances by means of a few control moves, the first being the largest, the following ones being smaller and of opposite sign. For $S = 10$, the smallest moves are eliminated, and the time to reject disturbances increases (0.5 to 1 sec). In particular, the controller waits for one more step before performing the second move. In summary, we can say that the ℓ_{asso} cost allows for a tradeoff between "input consumption" and speed of convergence. Moreover, the resulting control signals are sparse in time. From Figure 7.9, it can be noticed that the control action is far more aggressive when constraint

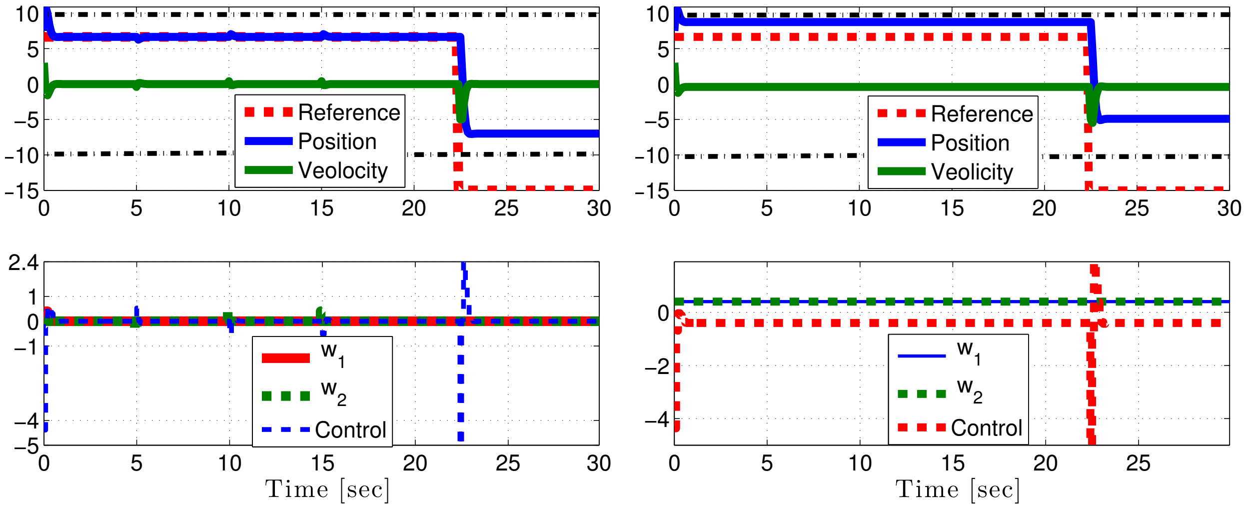

Figure 7.5: Closed-loop trajectory with a moving reference and disturbance (vanishing - left, constant - right). `Case 2`. $\lambda = 0.7,\ \mu = 0.4$, $S = 1$.

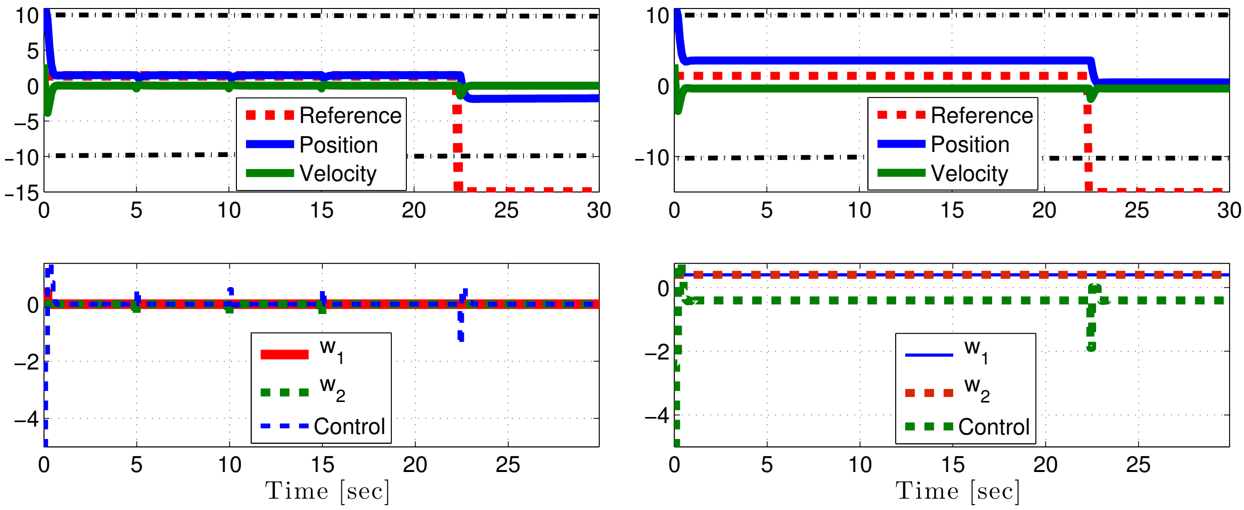

Figure 7.6: Closed-loop trajectory with a moving reference and disturbance (vanishing - left, constant - right). Considered benchmark (`Case 3`), $\lambda = 0.7,\ \mu = 0.4$, $S = 1$.

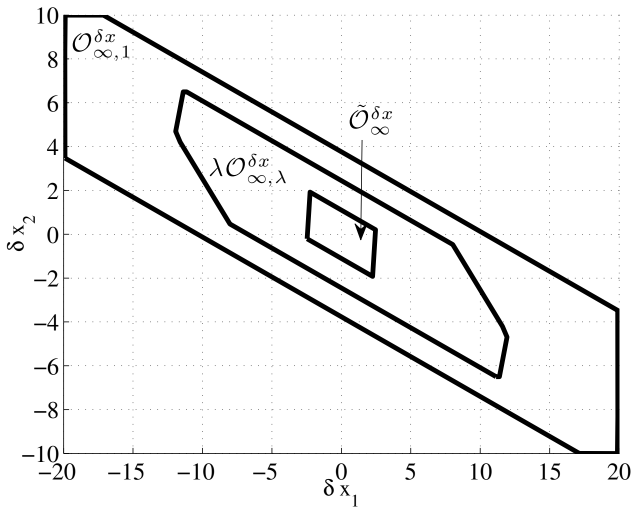

Figure 7.7: Terminal sets (on δx): nominal, proposed, based on [Ferramosca *et al.* , 2012]. $\lambda = 0.65,\ \mu = 0.47$.

are violated because of a disturbance. In this case in fact the state x_1 is rapidly steered back within constraints, even at the cost of an undershoot when $S = 1$.

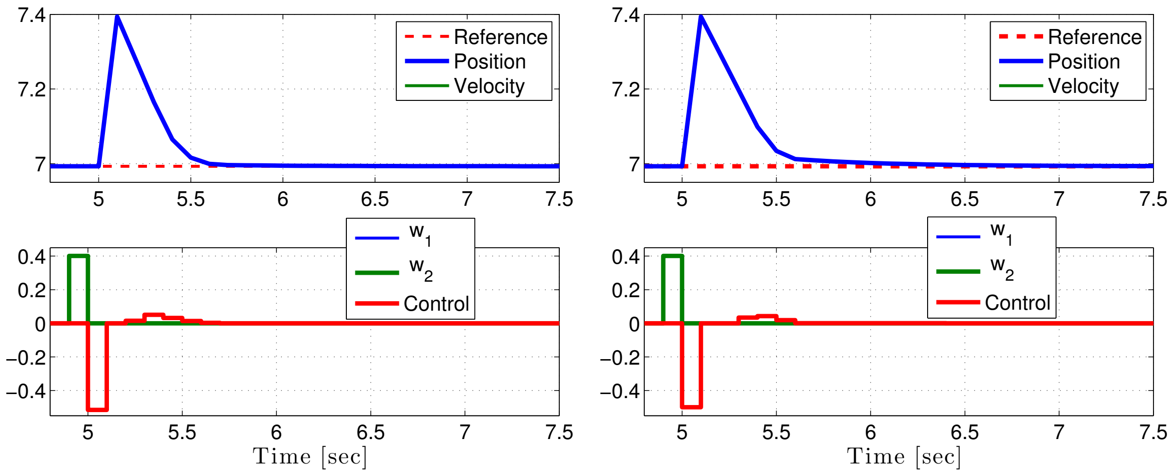

Figure 7.8: Closed-loop trajectory with a moving reference and vanishing disturbance. `case` 1: $S = 1 \Rightarrow$ faster response (left). $S = 100 \Rightarrow$ cheaper control (right). $\|w\|_\infty = \mu$.

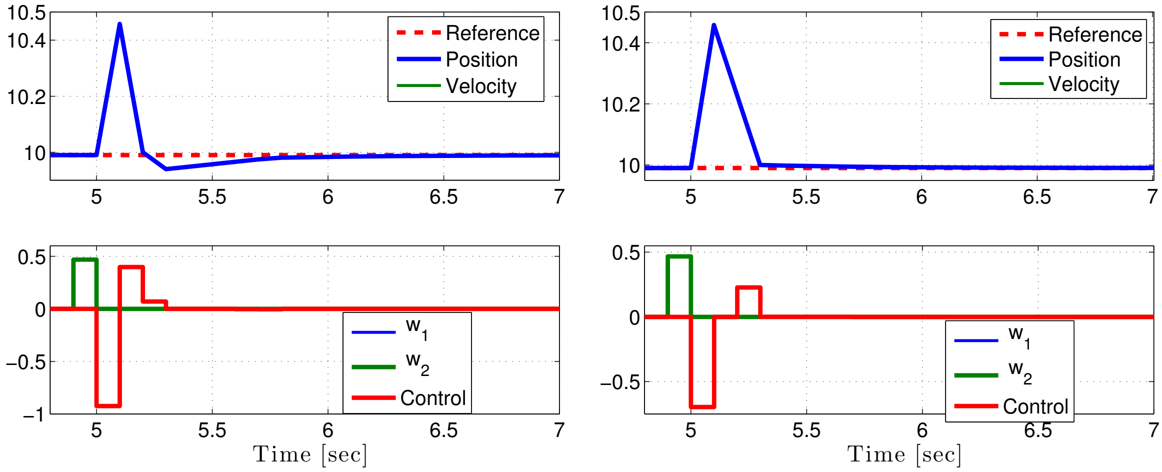

Figure 7.9: Closed-loop trajectory with a moving reference and vanishing disturbance. `case` 2: $S = 1 \Rightarrow$ faster response (left). $S = 100 \Rightarrow$ cheaper control (right). $\|w\|_\infty = \mu$.

7.8 Conclusions

This chapter presented a soft-constrained ℓ_{asso}-MPC strategy for tracking an asymptotically constant reference for linear time-invariant uncertain systems. Nominal asymptotically convergence of the closed-loop system to a constant admissible steady-state has been proven as well as the robust invariance of the feasible region for bounded additive uncertainties. Moreover, the closed-loop system is locally Input-to-State Stability (ISS). The strategy is based upon the recent literature on MPC for tracking and its extension to incorporate soft-constraints. The former uses a virtual set-point as an additional variable and a particular type of terminal set, which is known as the invariant set for tracking. The relaxation of the state constraints (not the terminal constraint) is performed by means of a set of slack variables, the magnitude of which is penalised in the MPC cost. This is inevitable for most real-world applications of MPC. Soft-constraints allow the MPC to remain feasible if a disturbance drives the state out of the constraints and, provided that the system can tolerate this, to recover from constraint violation. The total cost of the considered strategy features an offset penalty for exact convergence to the closest admissible reference.

The main result of the chapter is a bound for the infinity norm of the tolerable uncertainty, which is based on worst case open-loop disturbance propagation through the MPC predictions. Robustness has been obtained by means of a novel invariant set for tracking. In particular, this set is positively invariant for a suitable artificial system, the states of which being the state error from a virtual set-point (δx) and the virtual set-point itself (θ). As a result of the proposed procedures the terminal set is a subset of the new invariant set for tracking and it is λ-contractive in the δx coordinate. This fact is then used to show the existence of a feasible control sequence that steers the δx variable back into the terminal set starting from any point of the invariant set for tracking. Finally, a robustness bound is obtained by perturbing the tail of the MPC predictions at the previous time-step. In particular, if the open-loop propagation of the disturbance through the predictions causes the end of the perturbed sequence to be inside the invariant set for tracking, then the MPC is robustly feasible and the feasible region is a robustly positively-invariant set.

The second result proposed is a local ISS gain for the system, which relies on the assumption of local optimality of the external reference (assumed constant) with respect to the offset cost. This assumptions can be fulfilled if the offset penalty weight is computed using the theory of exact penalty functions, as discussed in the previous chapter. This ISS gain allows one to speculate on the closed-loop trajectory of the system, although in a conservative way. Conservativeness may arise from the fact that, differently from standard MPC theory, recursive feasibility is obtained for the new approach by assuming the parameter θ to be perturbed from the previous solution. Similar to the previous chapters, the solution of the proposed MPC is a continuous piece-wise affine function of the state and this time also of the external reference. This argument is used to upper-bound the required perturbation of θ for the worst-case disturbance. The quantification of conservatism and the investigation of potentially less conservative approximations are left for future research.

Soft-constrained MPC for tracking provides a potentially larger region of attraction (DOA) than standard regulation and tracking MPC. In this chapter, a new terminal cost and terminal set for tracking have been computed by means of contractive polytopes. As a result of this the DOA is non-empty by construction. This is in contrast to standard robust MPC methods, where the designer specifies a disturbance set. In the latter case in fact the terminal set could be empty, and this situation cannot be detected a priory. The proposed method for achieving robustness requires less computation than other robust MPC methods, for instance constraint restriction. The latter strategy and the proposed one have been compared for a simple system. Simulations have shown that the proposed algorithm can be less conservative than constraint restriction in terms of the size of the terminal set and of the achievable steady states. On the other hand, the proposed method does not guarantee robust constraints satisfaction. For this purpose, a combination of the two strategies is possible, although not investigated here. An alternative method has also been proposed, which restricts the allowable steady-states with no additional computation and the same robustness properties. A formal result is left for future research, however, using the modified approach the trajectory of the simulated example stays inside constraints despite disturbances once the terminal set is reached. The use of the ℓ_{asso}-cost has been shown to be beneficial to the regulation of the tradeoff between the speed of convergence and the magnitude of the control action, especially when constraints are violated.

CHAPTER

EIGHT

SHIP ROLL REDUCTION WITH RUDDER AND FINS

8.1 Introduction

Roll motion in ships is a major cause of problems to passengers and human operators. The reduction of vessel roll, or roll damping, has been widely addressed in the control literature as surveyed in [Perez, 2005]. The considered setup is shown in Figure 8.1. The system, controlled by rudder and fins, is *fully-actuated*. The necessity of handling constraints motivated the recent development of MPC-based fin stabilisers [Perez & Goodwin, 2008]. In particular, MPC allows one to avoid undesirable lift hysteresis phenomena due to excessive angle of attack of the fins [Perez, 2005].

The control task consists of the reduction of the high frequency wave induced roll motion as well as the regulation of the low frequency yaw motion. Quadratic MPC causes all actuators to be in use for all of the time, whereas it would be preferable for only fins to act to reduce the roll and hence most of the time, with the rudder applying additional torque only when the fins have reached the limit of their authority. Less roll-induced rudder action is preferable in order to have less yaw interference and to reduce drag. Moreover, rudder-based roll reduction suffers from performance limitations due to non-minimum phase dynamics (as well as to constraints) [Perez, 2005]. At the same time, the rudder should also act unimpeded in its primary role of steering the ship. The above remarks motivate our consideration of the problem of roll damping with rudder and fins as a 'demonstrator' for ℓ_{asso}-MPC.

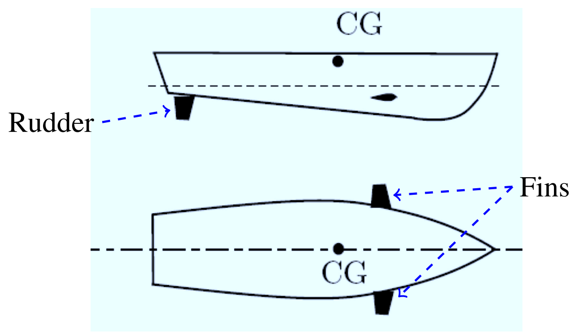

Figure 8.1: Rudder/fins setup for roll reduction. Based on [Perez & Goodwin, 2008].

8.2 Ship dynamics

The dynamics of marine vessels can be formulated using two different approaches: *motion superposition*, and *force superposition* [Fossen, 2002; Perez, 2005]. For control design, the first approach is generally used, according to marine craft *maneuvring theory*. The vessel equations of motion include rigid body dynamics and some hydrodynamic effects, which are given by the so-called *hydrodynamic derivatives* – a set of parameters obtained via system identification. The wave motion is then treated as an output disturbance, thus the name "motion superposition". The vessel motion has 6 degrees of freedom (DOF), and it is generally described by the standard SNAME coordinates in a reference frame set described in [Fossen, 2002]. In particular, the motion will be expressed in an Earth-fixed inertial frame, the n-frame, while the equations of motion are formulated in a body-fixed b-frame. The SNAME coordinates for positions and velocities are summarised in Table 8.1. The next section will briefly summarise the theory behind the manoeuvring equations of motion, while the rest of the chapter will highlight the application of the ℓ_{asso}-MPC.

8.2.1 Equations of motion

A manoeuvring model is used to describe the relationship between the control inputs (in terms of torques) and the motion response, in *relatively calm water*. The equations of motion are defined in

Variable name	Description
$n,\ e,\ d$	North, east, down positions, n-frame
$\phi,\ \theta,\ \psi$	Roll, pitch, yaw (Euler) angles, $n \to b$
$u,\ v,\ w$	Surge, sway, heave velocities, b-frame
$p,\ q,\ r$	Roll, pitch, yaw rate, b-frame
ν	Vector of gen. velocities and rates, b-frame
η	Vector of positions and attitude, n-frame

Table 8.1: SNAME coordinates

the b-frame and n-frame, as given by [Perez, 2005]

$$\begin{aligned} \left[M^b_{RB} + M^b_A\right]\dot{\nu} + C^b(\nu)\nu + D^b(\nu)\nu + g^b(\eta) = \tau^b \\ \dot{\eta} = J^n_b(\Theta_{nb})\nu \end{aligned} , \qquad (8.2.1)$$

where M^b_{RB} and M^b_A are respectively the rigid body and hydrodynamic added mass, $C^b(\nu)$, $D^b(\nu)$ are respectively the (velocity dependant) Centripetal and Coriolis force matrix and the hydrodynamic damping matrix, $g^b(\eta)$ provides the (position and attitude dependant) restoring forces, all expressed in the body-fixed frame. In (8.2.1) the matrix $J^n_b(\Theta_{nb})$ is a rotation matrix from the b-frame to the n-frame, with $\Theta_{nb} = [\phi,\ \theta,\ \psi]^T$. A full description of these matrices can be found in Chapter 4 of [Perez, 2005]. The forces acting on the body are represented by the linear combination:

$$\tau^b = \tau^b_{hyd} + \tau^b_{hs} + \tau^b_c + \tau^b_p\ ,$$

where τ^b_{hyd} are the generalized hydrodynamic forces, τ^b_{hs} are the generalized hydrostatics forces, τ^b_c are the vessel motion control forces, and τ^b_p are the propeller forces. According to [Perez, 2005], the control designer must be aware of the following:

Assumption 21. The τ^b_p compensates for the hydrodynamic resistance of the hull.

Assumption 22. The dynamics of the surge motion are much slower than the ones corresponding to other DOFs.

The latter allows one to decouple the surge component and take the variable u to be equal to the vessel service speed, which corresponds to the seakeeping average speed U, such that $u \approx U$. On the other hand, the model (8.2.1) will not be suitable for applications featuring frequent changes of forward speed.

It should be also noted that:

Remark. The hydrodynamic forces, τ^b_{hyd}, are generally non-linear functions of attitude, forward speed and wave frequency. The model will provide a rough approximation of some of them, for instance *radiation* and *viscous* forces, valid for low vessel-relative wave frequency (encounter frequency). The memory effects due to radiation forces are not taken into account for this application. However, for high speed application, like high speed manoeuvring, these effects cannot be neglected. See [Perez, 2005] for further details. The vector τ^b_{hyd} will also include the restoring force, which, for roll stabilisation, will only include the roll stabilising moment.

The parameters for the manoeuvring model (8.2.1) are generally obtained by system identification over a scale model. The resulting system is non-linear. Assuming a *slender ship*, one can use the linear Taylor series approximation, in the b-frame. The parameters corresponding to such linear model, are called *hydrodynamics derivatives*. The linear part of the model includes the couplings between roll, sway and yaw. The coefficients of the non-linear terms of (8.2.1) are simply obtained by curve fitting, and can be neglected for control design, not for testing.

The response of (8.2.1), in which the coefficients are obtained only by system identification and numerical methods, depends on the particular input signal used during the system identification, and it might differ from the vessel response, in different sea conditions.

8.2.2 Control surfaces

A hydrofoil in an irrotational inviscid flow is subject to the following lift and drag forces, [Perez, 2005]

$$L = \tfrac{1}{2}\rho V_f^2 A_f C_L(\alpha_e), \quad D = \tfrac{1}{2}\rho V_f^2 A_f \left(C_{D0} + \tfrac{C_L(\alpha_e)^2}{0.9\pi a}\right) \tag{8.2.2}$$

where ρ is the water density and V_f is the flow velocity, generally assumed to be equal to the vessel forward speed U. Therefore, it will be assumed that $V_f = U$. In (8.2.2) A_f is the foil area, C_L is the lift coefficient, C_{D0} is a drag constant, a is the foil aspect ratio, and α_e is the effective angle of attack, given by

$$\alpha_e = -\alpha_m - \alpha_f \tag{8.2.3}$$

where α_m is the actuator mechanical angle (also referred as α), while α_f is the flow angle. For small angles of attack, the lift coefficient can be approximated by the steady flow equation:

$$C_L = \left.\frac{\partial C_L}{\partial \alpha_e}\right|_{\alpha_e=0} \alpha_e + \frac{C_{DC}}{a}\left(\frac{\alpha_e}{57.3}\right)^2, \tag{8.2.4}$$

where the second term, describing the cross flow drag, is smaller than the linear component, and generally neglected. For large angles of attack, the lift coefficient has a nonlinear characteristic, mainly due to flow separation, with a consequent hysteresis. This phenomenon is generally neglected for a rudder, without much loss of accuracy. This is not so for the fins, when used for roll stabilisation, where the lift hysteresis induces significant performance limitations. This condition is also known as dynamic stall, and occurs for angle of attack larger than the stall angle, α^f_{stall}, [Perez & Goodwin, 2008].

The rudder forces acting on the vessel, expressed in body fixed coordinates, are given by

$$\tau_1^R \approx -D, \quad \tau_2^R \approx L, \quad \tau_4^R \approx -r_R L, \quad \tau_6^R \approx -L_{CG} L \tag{8.2.5}$$

where r_R is the rudder roll arm, and L_{CG} is the longitudinal distance from the vessel centre of gravity. The flow angle is assumed to be zero, therefore, for the rudder, $\alpha_e^R = -\alpha^R$ (where superscript R refers to the rudder).

For the fins, define the vector of normal and tangential forces as:

$$N = L\cos(\alpha_e) + D\sin(\alpha_e), \quad T = D\cos(\alpha_e) - L\sin(\alpha_e). \tag{8.2.6}$$

For the fins, the flow angle cannot be assumed to be null, but can be approximated by

$$\alpha_f = \arctan\left(\frac{r_F p}{U}\right), \tag{8.2.7}$$

where p is the roll rate of the vessel, r_F is the fins roll centre arm length, and U is the vessel forward speed. For control design, the latter is generally linearised. The fin induced forces acting on the hull can be approximated by:

$$\begin{aligned} &\tau_1^F \approx -\tfrac{1}{2}(T^{SF} + T^{PF}), \quad \tau_2^F \approx -\tfrac{1}{2}(N^{SF} + N^{PF})\sin(\beta), \\ &\tau_4^F \approx r_F(N^{SF} + N^{PF}), \quad \tau_6^F \approx F_{CG}\tfrac{1}{2}(N^{SF} + N^{PF})\sin(\beta), \end{aligned} \tag{8.2.8}$$

where the superscript SF and PF stand for, respectively, starboard fin and port fin, while β is the fin tilt angle, and F_{CG} is the longitudinal distance between the fin centre of pressure and the vessel centre of gravity.

8.2.3 Linearised manoeuvring equations of motion in 3 DOF

The equations of motion are formulated for the sway, roll and yaw DOFs, assuming the propeller action maitains the vessel travelling at a constant forward speed, U. Choosing the following coordinates:

$$\nu = \begin{bmatrix} v & p & r \end{bmatrix}^T, \quad \eta = \begin{bmatrix} \phi & \psi \end{bmatrix}^T, \tag{8.2.9}$$

according to the motion superposition formulation, the vessel state can be partitioned in order to describe control induced and wave induced motion as:

$$x = \begin{bmatrix} (x^c)^T & (x^w)^T \end{bmatrix}^T \quad x^c = \begin{bmatrix} \nu^T & \eta^T \end{bmatrix}^T, \tag{8.2.10}$$

where x^w represents the wave induced disturbances on the roll rate and angle. A course keeping problem consists into a setpoint tracking of the low frequency vessel yaw, $\bar{\psi}$, where:

$$\begin{aligned} \psi &= \bar{\psi} + \psi^w \\ \bar{\psi} &= \psi^c + \psi^B \end{aligned} \tag{8.2.11}$$

where ψ^w, ψ^c, and ψ^B represent, respectively, the first order wave induced (high frequency), the control induced yaw motion and a bias, due to slowly varying disturbances, like wind, current and second order wave drift. On the other hand, the roll stabilisation, or roll damping problem aims to control both the high and low frequency roll motion, in terms of angle, rate and/or acceleration. Roll motion influences the performance of the vessel, and the comfort of the crew. For a particular manoeuvre, where the yaw setpoint changes in time, the actual roll angle might not be strictly required to be zero for all the time, while its velocity and acceleration have to be kept small.

To formulate the system dynamics, the fin normal force is first linearised for small angles of attack as follows:

$$N = \underbrace{(K_L + K_{D1})}_{K_N} \alpha_e^F, \tag{8.2.12}$$

where

$$K_L = \frac{1}{2}\rho U^2 A_f \frac{\partial C_L}{\partial \alpha_e}\bigg|_{\alpha_e=0}, \quad K_{D1} = \frac{1}{2}\rho V_f^2 A_f C_{D0}, \quad \alpha_e^F = \frac{r_{FP}}{U}. \tag{8.2.13}$$

The vessel dynamics is linearised around $u = U$, $\phi = 0$, $p = 0$, $r = 0$. This approximation will be valid for slow manoeuvring (slow turning). Elaborating the formulation in [Perez, 2005] to add the fins dynamics, the linearised control induced dynamics can be represented by:

$$\dot{x}^c = M_{inv}F(U)x^c + M_{inv}H(U)u, \tag{8.2.14}$$

where

$$M_{inv} = \begin{bmatrix} M^{-1} & 0_{2\times 2} \\ 0_{2\times 2} & I_{2\times 2} \end{bmatrix}, \tag{8.2.15}$$

and

$$M = \begin{bmatrix} m - Y_{\dot{v}} & -(mZ_G + Y_{\dot{p}}) & mX_G - Y_{\dot{r}} \\ -(mZ_G + K_{\dot{v}}) & -(I_{xx} - K_{\dot{p}}) & -K_{\dot{r}} \\ -(mX_G + N_{\dot{v}}) & -N_{\dot{p}} & I_{zz} - N_{\dot{r}} \end{bmatrix} \tag{8.2.16}$$

represents the rigid body mass and inertia, plus the first order hydrodynamic derivatives (corresponding to the added mass). In equation (8.2.14), $F(U) = F_1(U) + F_2(U)$, where:

$$F_1(U) = \begin{bmatrix} Y_{|u|v}U & 0 & (Y_{ur} - m)U & Y_{\phi uu}U^2 & 0 \\ K_{|u|v}U & K_p + K_{|u|p}|U| & (K_{ur} + mZ_G)U & K_{\phi uu}U^2 - \rho g \Delta GM_t & 0 \\ N_{|u|v}U & N_p + N_{|u|p}|U| & N_{|u|r}U - mX_GU & N_{\phi u|u|}U|U| & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{8.2.17}$$

represents the centripetal, Coriolis and linearised second order hydrodynamic forces, while

$$F_2(U) = \begin{bmatrix} 0 & K_N \frac{r_F}{U} \sin(\beta) & 0 & 0 & 0 \\ 0 & -2K_N \frac{r_F^2}{U} & 0 & 0 & 0 \\ 0 & -F_{GC} K_N \frac{r_F}{U} \sin(\beta) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{8.2.18}$$

provides the fin effect due to flow directionality. Defining the input vector as

$$u = \begin{bmatrix} \alpha^R \\ \alpha^{SF} \\ \alpha^{PF} \end{bmatrix} \tag{8.2.19}$$

results in

$$H(U) = \begin{bmatrix} K_L^R & \frac{1}{2} K_N \sin(\beta) & \frac{1}{2} K_N \sin(\beta) \\ -r_R K_L^R & -r_F K_N & -r_F K_N \\ -L_{CG} K_L^R & -F_{GC} K_N \sin(\beta) & -F_{GC} K_N \sin(\beta) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{8.2.20}$$

A discrete time equivalent of (8.2.14) can be obtained, having

$$x_{k+1}^c = A x_k^c + B u_k. \tag{8.2.21}$$

8.3 Simulation Results

Simulations are performed on a realistic vessel model in ocean waves, from the 2009 version of the open source Marine System Simulator (MSS) toolbox for Matlab [Fossen & Perez, 2004; Perez *et al.* , 2006]. As explained in the previous section, the vessel dynamics are formulated according to the marine craft maneuvring theory [Fossen, 2002]. The waves are treated as a Gaussian input disturbance whose distribution is based on a power spectrum, the choice of which depends on several factors [Fossen, 2002]. A JONSWAP spectrum [Fossen, 2002] is chosen as an example. Full state measurement is assumed. The model parameters are in the Appendix of [Perez, 2005]. The controller is required to maintain a desired average yaw angle, while reducing the roll variance despite the action of sea waves. The discretised linearisation for small angles is used as shown in the previous section. A wave model is not used for the MPC predictions, in contrast with [Perez, 2005]. A terminal constraint, the computation of which is discussed in [Gallieri & Maciejowski, 2012], is not imposed here as the presence of uncertainty can often make this constraint infeasible. The ℓ_{asso}-MPC is compared to the more common LQ, ℓ_1 and ℓ_∞-MPC. The horizon length is $N = 15$, and the sampling time is $T_s = 0.1$ sec. Constraints involve both inputs and states [Perez, 2005] and are described in Table 8.2. In Table 8.2 a constraint is imposed on the fins angle of

Variable	Description	Constraint	Rate constraint
α_R	Rudder angle $[Deg]$	45	20
α_F	Fins angle $[Deg]$	35	35
α_e	Fins angle of attack $[Deg]$	23	-

Table 8.2: Vessel constraints

attack, which is defined in (8.2.3), (8.2.7). If this limit is exceeded then the fins are subject to a large loss of lift as well as to lift hysteresis [Perez & Goodwin, 2008]. In order to avoid these phenomena (8.2.7) is linearised around $p = 0$ and then included into the MPC (for each fin) as [Perez & Goodwin, 2008]:

$$\left| -\frac{r_F}{U} p - \alpha_F \right| \leq 23. \tag{8.3.1}$$

The weights for the cost function are summarised in the upper half of Table 8.3 for the selected MPC strategies. Performance is evaluated by computing the mean and standard deviation (STD) of the roll angle, and of the low-frequency yaw, obtained by a low-pass filter with a bandwidth of $1.5\ \omega_0$, where ω_0 is the mean wave frequency. The mean rudder and starboard fin angle are also computed. Simulations are performed for an irregular sea, namely, a Sea State 5 (SS5) [Perez, 2005]. Simulation results, averaged over 40 simulations, are shown the lower half of Table 8.3 as well as in Figure 8.2.

Table 8.3: Cost function parameters and simulation results per MPC strategy

weight/MPC	ℓ_1	ℓ_∞	LQ	ℓ_{asso}
q_ϕ (roll angle)	10	20	100	100
q_r (yaw rate)	1	1	10	10
q_ψ (yaw angle)	10	10	100	100
q_R (rudder angle)	200	20	20	2
q_F (fins angle)	100	10	10	1
S (all inputs)	-	-	-	$2.2 \cdot I$
Results/MPC	ℓ_1	ℓ_∞	LQ	ℓ_{asso}
Roll angle std $[Deg]$	0.21	0.4	0.24	0.20
Average roll angle $[Deg]$	-0.01	0.003	-0.03	-0.001
Roll angle reduction $\%$	76	57	73	78
Yaw angle std $[Deg]$	0.58	0.6	0.78	0.7
Average yaw angle $[Deg]$	0.14	0.12	0.23	0.28
Average rudder angle magnitude $[Deg]$	26	17.5	7.2	5.6
Average fin angle magnitude $[Deg]$	22	28.1	5.5	7.8

For the ℓ_{asso}-MPC, the tradeoff between control error statistics and input amplitude can be regulated by setting $S = \eta \cdot I$. The best results have been obtained for $\eta = 2.2$. The dead-zone approach from Section 4.5 is used. To challenge the new approach, the R matrix for quadratic MPC is chosen to be 10 times the one used in ℓ_{asso}-MPC. Despite this, the results obtained with ℓ_{asso}-MPC are better than the ones achieved with quadratic MPC, in terms of the tradeoff between error standard deviation and input magnitude. In particular, ℓ_{asso}-MPC achieves the maximum roll reduction and prefers the use of the fins over the rudder, which is left at rest for most of the time. Note that a reduction of the ℓ_1-norm of the input signals, expected in the nominal LTI case, is still achieved for the nonlinear stochastic system. On the other hand, ℓ_∞-MPC has very poor performances in terms of low roll reduction and very high input 1-norm, despite being reasonably good for the linearised model. This is mainly due to the model mismatch and the controller being much more aggressive then to the ones with quadratic state penalty. The ℓ_1-MPC offers a good roll reduction, but this comes at the expense of the highest usage of the rudder and the fins (the latter is second only to the ℓ_∞-MPC). The above considerations make the considered ℓ_1 and ℓ_∞-MPC

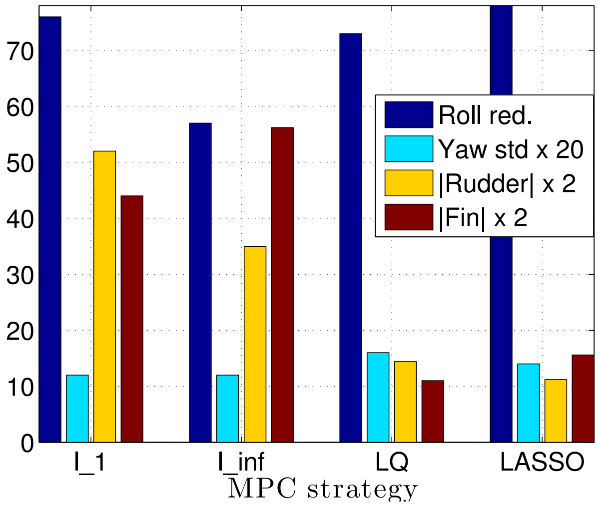

Figure 8.2: Obtained performance for the considered MPC strategies

unsuitable for this application. On the other hand, the ℓ_{asso}-MPC seems to be a good candidate solution. Figure 8.3 shows the control signals obtained with the selected MPC strategies, in a short time interval. As expected, in contrast to quadratic MPC, the ℓ_{asso}-MPC is capable of holding the rudder angle exactly at zero for most of the time. Similarly for the fins — namely, the solution is 'sparse in time' as well as 'sparse in actuators'. On the other hand, quadratic MPC seems to generate rather unrealistic commands, with magnitudes lower than 1 degree. These are not likely to be achievable by ether the rudder or the fin machinery. The LP-based MPC controllers are, on the other hand, far more aggressive. The ℓ_{asso}-MPC seems to appeal to this application.

8.4 Conclusions

This chapter has addressed the problem of vessel course keeping and roll reduction by means of rudder and fins, the commands of which are generated by ℓ_{asso}-MPC. High fidelity simulations have been performed on a realistic vessel model in ocean waves, the latter being modelled as a multi-directional Gaussian noise. The considered scenario is a benchmark from the marine systems control literature. The proposed ℓ_{asso}-MPC has also been compared to the more common LQ (used as a benchmark), ℓ_1 and ℓ_∞ MPC, the latter two providing particularly poor performance. Controller tuning has been performed by trial and improvement. The ℓ_{asso}-MPC provides an addi-

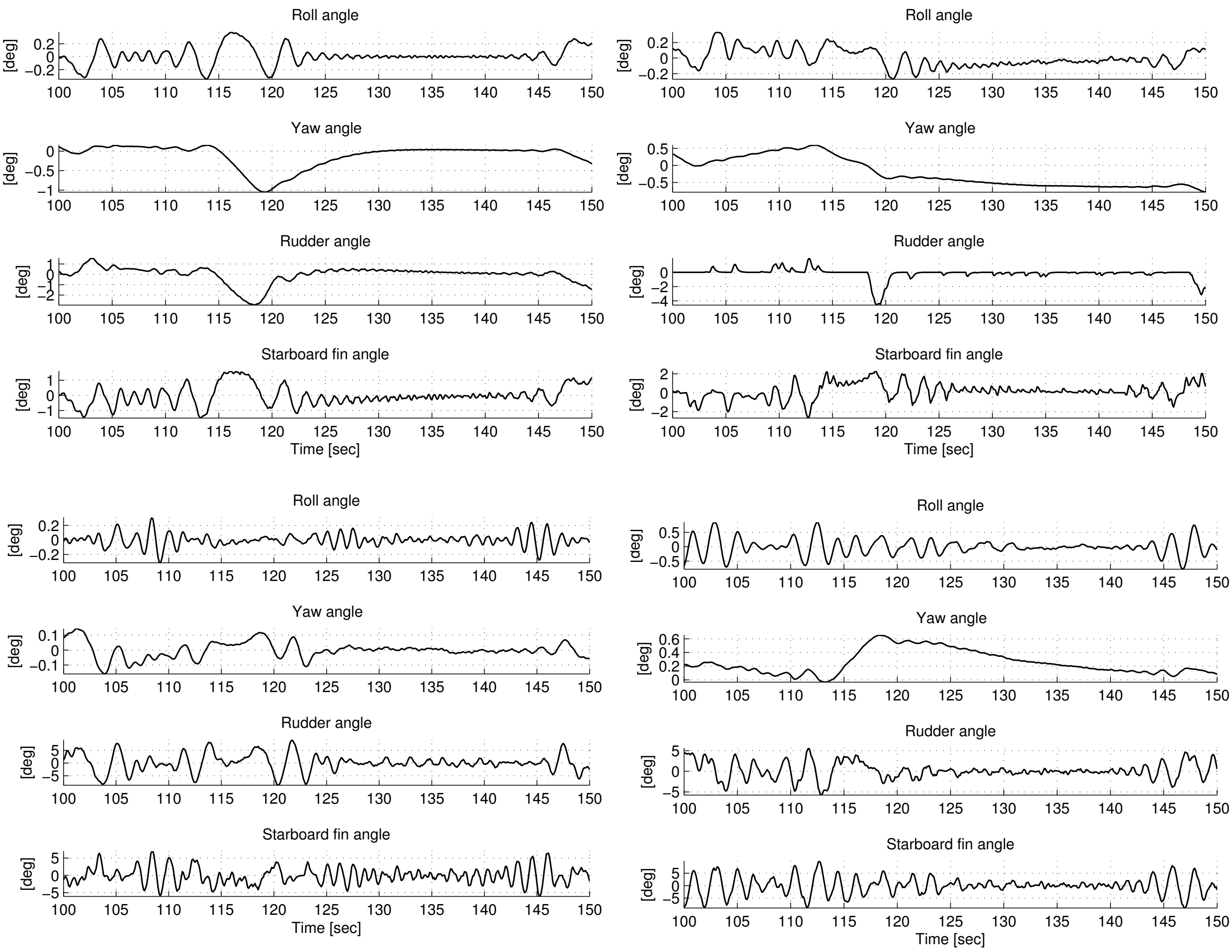

Figure 8.3: LQ-MPC (top left), ℓ_{asso}-MPC (top right), ℓ_1-MPC (bottom left) and ℓ_∞-MPC (bottom right) results in SS5

tional 6% reduction of the roll variance with respect to LQ-MPC as well as 28% reduction of the rudder activity, the fins being preferred instead. The implementation of ℓ_{asso}-MPC requires to add to the LQ-MPC as many variables and constraints as the number of actuators times the prediction horizon length (in this case 15). A reduction of the computational requirement could be possible, if for instance a longer sampling time (and a shorter horizon) is used and/or if the fins are not regularised. This case is left for future investigation. Simulation results are particularly encouraging for a possible future implementation of ℓ_{asso}-MPC on the physical system.

CHAPTER

NINE

CONCLUDING REMARKS

9.1 Summary of contribution

This thesis presented a theoretical framework for the use of ℓ_1 regularised least squares in Model Predictive Control (MPC), informally referred to as ℓ_{asso}-MPC. This is suitable, for instance, for control applications where the control signals are required to be sparse in their magnitudes or in their time-increments. This is the case where the use of some actuators is particularly expensive, or where it is important to minimise actuator wear and tear. In particular, the thesis has focused on sparsity through input channels, which can be desirable for the control of systems such as distribution networks and redundantly-actuated marine or air vehicles. The main contributions of the thesis follow:

- The geometry of the LASSO and elastic-net regression cost functions have been studied, showing that ℓ_{asso}-MPC has a smooth and unique solution (as for quadratic MPC) if a positive definite quadratic input penalty is included in the cost. At the same time, the solution of ℓ_{asso}-MPC can be sparse as for MPC based on ℓ_p-norm costs. This is achieved by ℓ_{asso}-MPC without the limitations of ℓ_p-MPC, such as actuator chattering.
- Regularising an existing LQ-MPC (ℓ_{asso}-MPC version 1) has been shown to introduce a soft-thresholding effect in the control law, that is unique, continuous and piece-wise affine (PWA), as for LQ-MPC. In contrast to the latter, the unconstrained solution of ℓ_{asso}-MPC version 1 is not a linear controller (like LQR), but is still a PWA function. In particular, the

commands of the regularised actuator are set to zero in a neighbourhood of the origin (for the case of a regulator).

- Stability of ℓ_{asso}-MPC version 1 has been investigated for non-linear systems with differentiable vector field. First, it has been shown that ℓ_{asso}-MPC version 1 can achieve ultimate boundedness of the solution, when all actuators are regularised. Asymptotic stability is then proven for a particular case, referred to as *partial regularisation*. This assumes the system to be stabilised by a strict subset of actuators, which are not regularised.

- The ℓ_{asso}-MPC version 2 has been proposed to guarantee asymptotic stability of the closed-loop system for any regularisation penalties. This approach uses a different terminal cost from the original version, which is computed by means of set-theoretic control methods. Two candidate terminal costs have been investigated for non-linear systems with differentiable vector field. The final choice has fallen upon a scaled version of the Minkowski function of the terminal set, which is a λ-contractive convex polytope. For LTI systems ℓ_{asso}-MPC version 2 can be formulated, as well as version 1, in terms of a strictly convex mpQP. This guarantees a unique, continuous and PWA solution. Conditions have been given under which ℓ_{asso}-MPC version 2 can exhibit sparsity through actuator channels. This requires the system to be locally stabilisable through a strict subset of the available actuators.

- Tractable procedures have been reviewed, so that the ingredients for stability of ℓ_{asso}-MPC version 2 can be computed offline and prior to the cost function definition. This facilitates online tuning, when compared to most common MPC assumptions.

- It has been shown how ℓ_{asso}-MPC versions 1 and 2 can be used for systems with prioritised actuators, namely systems where a set of actuators is to be operated only when the control error is large. In particular, a procedure for LTI systems has been proposed for computing the regularisation penalty so that the mentioned actuators are not used when the error is in a given neighbourhood of the origin. The procedure is based upon standard results on multi-parametric programming and exact penalty functions. The size of the attainable neighbourhood is shown to be influenced by the horizon length as well as the constraints.

- The ℓ_{asso}-MPC for prioritised actuators has been demonstrated to be effective for the minimisation of the use of spoilers for the roll control of the linearised lateral dynamics of a Boeing 747 at different flight points as well as for the control of an abstracted distribution network with prioritised links.

- A robust ℓ_{asso}-MPC for tracking has been formulated which makes use of soft-constraints on the state predictions with a hard terminal constraint. Sufficient conditions have been given to obtain a bound on the infinity norm of the allowable additive disturbances, for which the MPC is recursively feasible. The feasible region is a robustly positively invariant set and the closed-loop system is locally Input-to-State Stable. The proposed conditions are based on soft-constraints, which can be inevitable in the presence of uncertainty, and on the introduction of a novel terminal set for tracking. This new set is formulated so that its projection on the state error space is a λ-contractive set. Conditions are also given for the finite time (offline) computation of the required ingredients. For a simple example, the proposed controller has been shown to provide a much larger region of attraction as well as to track a much larger set of references than a standard robust MPC based on constraint restriction.

- High fidelity simulations have been performed for the problem of roll reduction of an ocean vessel by means of rudder and fins. The control objective consists of the reduction of the high frequency wave induced roll motion, preferably by means of the fins, as well as the regulation of the low frequency yaw motion by means of the rudder. In particular, ℓ_{asso}-MPC version 1 has been shown to outperform the more common LQ, ℓ_1 and ℓ_∞-MPC in terms of the reduction of roll variance and low frequency yaw as well as for the minimisation of the rudder activity.

9.2 Directions for future research

The following directions are proposed for future research:

- The size of the achievable region where the secondary actuators are not used could be further increased. This, if possible, would require a different method to compute the regularisation penalty or the use of a different penalty at different prediction steps.

- Sparsity in time could be investigated. The use of a time-varying stage cost might be a possible solution. As an alternative, information on the past control moves could be incorporated in the optimisation, with the aim of reducing the gap between open-loop and closed-loop solution. Stability of the new solution should then be investigated.

- ℓ_{asso}-MPC could be investigated as a convex surrogate of hybrid MPC, for the control of systems with inputs subject to binary constraints.
- The potential for using ℓ_{asso}-MPC for the reduction of the effect of disturbances and measurement noise could be addressed through a stochastic framework.
- Robustness to multiplicative uncertainties could be investigated, within the proposed framework. The ISS gain of soft-constrained ℓ_{asso}-MPC could be investigated further, to aim towards a less conservative result.
- Further case studies could be addressed, for instance, applications where using the actuators frequently is particularly expensive or where actuators' wear and tear must be minimised.
- Further improvement may be possible for the computation of the required contractive sets.

REFERENCES

ALESSIO, A., & BEMPORAD, A. 2009. A Survey on Explicit Model Predictive Control. *Nonlinear Model Predictive Control: Lecture Notes in Control and Information Sciences*, **38**, 345–369. 42, 43, 44, 52, 74, 76, 77, 97, 125, 129, 134, 163, 175

ASTROM, K. J. 1970. *Introduction to Stochastic Control Theory*. New York, NY : Academic Press. 37

BEMPORAD, A., MORARI, M., DUA, V., & PISTIKOPOULOS, E.N. 2002a. The explicit Linear quadratic regulator for constrained systems. *Automatica*, **38**, 3–20. 22, 32, 37, 42, 43, 44, 52, 76, 77, 78, 125, 129, 131, 133, 135, 136, 140

BEMPORAD, A., BORRELLI, F., & MORARI, M. 2002b. Model predictive control based on linear programming - the explicit solution. *IEEE Transaction on Automatic Control*, **47**(12), 1974 –1985. 74, 76

BLANCHINI, F. 1994. Ultimate boundedness control for uncertain discrete-time systems via set-induced Lyapunov functions. *IEEE Transaction on Automatic Control*, **39**, 428 – 433. 102, 119

BLANCHINI, F. 1999. Set invariance in control. *Automatica*, **35**, 1747–1767. 12, 27, 29, 40, 89

BLANCHINI, F., & MIANI, S. 2003. Stabilisation of LPV systems: state feedback, state estimation and duality. *In: Proceedings of the Conference on Decision and Control (CDC)*. 119

BLANCHINI, F., & MIANI, S. 2008. *Set-theoretic methods in control*. Birkhäuser. 29, 30, 37, 38, 40, 45, 86, 89, 96, 97, 98, 101, 102, 112, 114, 117, 118, 119, 165, 172, 182

BORRELLI, F., BAOTIC, M., BEMPORAD, A., & MORARI, M. 2008. Efficient on-line computation of explicit model predictive control. *SIAM Journal of Control and Optimization*, **47**, 2470–2489. 42, 45, 133, 145

BORRELLI, F., BAOTIĆ, M., PEKAR, J., & STEWART, G. 2009. On the complexity of explicit MPC laws. *Pages 2408–2413 of: Proceedings of the European Control Conference (ECC)*. 42, 43, 45, 133

BOYD, S., & VANDENBERGHE, L. *Notes for EE364b, Stanford University, Winter 2006-2007, Subgradients*. 16

BOYD, S., & VANDENBERGHE, L. 2004. *Convex Optimization*. Cambridge University Press. 11, 14, 16, 17, 18, 19, 20, 22, 23

BOYD, S., EL GHAOUI, L., FERON, E., & BALAKRISHNAN, V. 1994. *Linear Matrix Inequalities in System and Control Theory*. SIAM. 38, 39, 112, 113

BRAVO, J.M., LIMON, D., ALAMO, T., & CAMACHO, E.F. 2005. On the computation of invariant sets for constrained nonlinear systems: An interval arithmetic approach. *Automatica*, **41**, 1583–1589. 40

CANNON, M., DESHMUKH, & KOUVARITAKIS, B. 2003. Nonlinear model predictive control with polytopic invariant sets. *Automatica*, **39**, 1487 – 1494. 101, 118, 173

CHEN, H., & ALLGÖWER, F. 1998. A quasi-infinite horizon Nonlinear Model Predictive Control Scheme with guaranteed stability. *Automatica*, **34**(10), 1205 – 1217. 32, 36, 37, 38

CHISCI, L., ROSSITER, J.A., & ZAPPA, G. 2001. Systems with persistent disturbances: predictive control with restricted constraints. *Automatica*, **37**, 1019 – 1028. 45, 182

CHMIELEWSKI, D., & MANOUSIOUTHAKIS, V. 1996. On constrained infinite-time linear quadratic optimal control. *System & Control Letters*, 121–129. 32, 37

DI CAIRANO, S., LIANG, W., KOLMANOVSKY, I. V., KUANG, M. L., & PHILLIPS, A. M. 2013. Power Smoothing Energy Management and Its Application to a Series Hybrid Powertrain. *IEEE Transaction on Control System Technology*, **21**(6), 2091–2103. 42

EDWARDS, C., LOMBAERT, T., & HAFID, S. 2010. *Fault Tolerant Flight Control: A Benchmark Challenge*. Springer-Verlag Berlin Heidelberg. xii, xiv, 4, 5, 151

EFRON, BRADLEY, HASTIE, TREVOR, JOHNSTONE, IAIN, & TIBSHIRANI, ROBERT. 2004. Least angle regression. *The Annals of Statistics*, **32**(2), 407 – 499. 51, 52

FERRAMOSCA, A., LIMON, D., ALVARADO, I., ALAMO, T., & CAMACHO, E.F. 2008. MPC for tracking with optimal closed-loop performance. *Pages 4055 – 4060 of: Proceedings of the Conference on Decision and Control (CDC)*. 46, 47, 161, 162, 169, 171, 175

FERRAMOSCA, A., LIMON, D., ALVARADO, I., ALAMO, T., & CAMACHO, E.F. 2009a. MPC for tracking of constrained nonlinear systems. *Pages 7978–7983 of: Proceedings of the Conference on Decision and Control (CDC)*. 46, 161

FERRAMOSCA, A., LIMON, D., ALVARADO, I., ALAMO, T., & CAMACHO, E. F. 2009b. MPC for tracking with optimal closed-loop performance. *Automatica*, **45**, 1975 – 1978. 46, 47, 169, 170

FERRAMOSCA, A., LIMON, D., GONZÁLEZ, I., ALVARADO, I., & CAMACHO, E. F. 2012. Robust MPC for tracking zone regions based on nominal predicitons. *Journal of Process Control*, **22**, 1966–1974. xiv, 162, 182, 183, 184, 189

FIACCHINI, M., ALAMO, T., & CAMACHO, E.F. 2012. Invariant sets computation for convex difference inclusions systems. *Systems & control letters*, **8**, 819 – 826. 40, 118, 173

FLETCHER, R. 1987. *Practical Methods of Optimization, 2nd edition*. Wiley. 23, 45, 133

FOSSEN, T. I. 2002. *Marine Control Systems: Guidance, Navigation and Control of Ships, Rigs and Underwater Vehicles*. Trondheim: Marine Cybernetics AS. 194, 201

FOSSEN, T. I., & PEREZ, T. 2004. *Marine Systems Simulator (MSS)*. www.marinecontrol.org. 201

GALLIERI, M., & MACIEJOWSKI, J. 2013a. Soft-Constrained Lasso-MPC for Robust LTI Tracking: Enlarged Feasible Region and an ISS Gain Estimate. *Pages 7101 – 7106 of: Proceedings of the Conference on Decision and Control (CDC)*. 8, 40, 48, 92, 103

GALLIERI, M., & MACIEJOWSKI, J.M. 2012. ℓ_{asso} MPC: Smart Regulation of Overactuated Systems. *Pages 1217 – 1222 of: Proceedings of the American Control Conference (ACC)*. 8, 53, 92, 201

GALLIERI, M., & MACIEJOWSKI, J.M. 2013b. Stabilising Terminal Cost and Terminal Controller for ℓ_{asso}-MPC: Enhanced Optimality and Region of Attraction. *Pages 524 – 529 of: Proceedings of the European Control Conference (ECC)*. 8, 40, 92, 101, 162

GILBERT, E.G., & TAN, K.T. 1991. Linear systems with state and control constraints: The theory and application of maximal output admissible sets. *IEEE Transaction on Automatic Control*, **36**, 1008–1020. 171

GOODWIN, G., SERON, M. M., & DE DONÁ, J. A. 2005. *Constrained Control and Estimation. An Optimisation Approach*. Springer. 32, 54

GOULART, P. J., KERRIGAN, E. C., & MACIEJOWSKI, J. M. 2006. Optimization Over State Feedback Policies for Robust Control with Constraints. *Automatica*, **42**(4), 523 –533. 45, 182

GRAMMATICO, S., & PANNOCCHIA, G. 2013. Achieving a large domain of attraction with short-horizon linear MPC via polyhedral Lyapunov functions. *Pages 1059 – 1064 of: Proceedings of the European Control Conference (ECC)*. 40, 41, 96, 102, 103, 119, 174

GRIMM, G., MESSINA, M. J., TUNA, S. E., & TEEL, A. R. 2004. Examples when nonlinear model predictive control is nonrobust. *Automatica*, **40**(10), 1729 – 1738. 41, 47, 161

GRÜNE, L. 2011. *Economic receding horizon control without terminal constraints*. Research Report. University of Bayreuth. 46

GRÜNE, L., & PANNEK, J. 2011. *Nonlinear Model Predictive Control: Theory and Algorithms*. First edn. Springer-Verlag, London. 46

HARTLEY, E.N., GALLIERI, M., & MACIEJOWSKI, J.M. 2013. Terminal Spacecraft Rendezvous and Capture with LASSO Model Predictive Control. *International Journal of Control*, **86**(11), 2104 – 2113. 9, 75

HASTIE, T., TIBSHIRAN, R., & FRIEDMAN, J. 2011. *The Elements of Statistical Learning*. Springer. xii, 49, 50, 51, 52

HINDI, H. 2004. A tutorial on convex optimization. *Pages 3252 – 3265 of: Proceedings of the American Control Conference (ACC)*. 11, 14, 16, 17, 18, 19

HINDI, H. 2006. A tutorial on convex optimization II: Duality and interior point methods. *Pages 686 – 696 of: Proceedings of the American Control Conference (ACC)*. 18, 20, 21

HORN, R. A., & JOHNSON, C. R. 2010. *Matrix Analysis*. Cambridge university press. 18, 96, 97

IBM. 2013. *ILOG CPLEX v.12.2.0.2*. 22

JEREZ, J. L., KERRIGAN, E. C., & A., CONSTANTINIDES G. 2011. A Condensed and Sparse QP Formulation for Predictive Control. *In: Proceedings of the Conference on Decision and Control (CDC)*. 5217-5222. 124

JIANG, ZHONG-PING, & WANG, YUAN. 2001. Input-to-state stability for discrete-time nonlinear systems. *Automatica*, **37**(6), 857 – 869. 47, 161

JOHANSEN, T. A.AND GRANCHAROVA A, & GRANCHAROVA, R. 2003. Approximate Explicit Constrained Linear Model Predictive Control Via Orthogonal Search Tree. *IEEE Transaction on Automatic Control*, **48**, 810–815. 45

KAKADE, S. M., SHALEV-SHWARTZ, S., & TEWARI, A. 2012. Regularization techniques for Learning with Matrices. *Journal of Machine Learning Research*, **13**(1), 1865–1890. 16

KARMARKAR, N. 1984. A new polynomial-time algorithm for linear programming. *Combinatorica*, **4**(4), 373 – 395. 18, 22

KERRIGAN, E. 2000. *Robust constraint satisfaction : Invariant sets and predictive control*. Ph.D. thesis, University of Cambridge - St. John's college. 23, 30, 31, 37, 40, 102, 118, 119

KERRIGAN, E., & MACIEJOWSKI, J. 2000a. Invariant sets for constrained nonlinear discrete-time systems with application to feasibility in model predictive control. *Pages 4951 – 4956 of: Proceedings of the Conference on Decision and Control (CDC)*. 30, 31, 32, 34, 35, 37, 41, 45, 127, 128, 129

KERRIGAN, E. C., & MACIEJOWSKI, J. M. 2000b. Soft constraints and exact penalty functions in Model Predictive Control. *In: Proceedings of the UKACC International Conference (Control 2000)*. 44, 45, 47, 161, 165

KIM, S., KOH, K., LUSTING, M., & BOYD, S. 2007a. An efficient method for compressed sensing. *Pages II:117 – III:120 of: IEEE International Conference on Image Processing*. 49, 51

KIM, S.J., LUSTING, M., BOYD, STEPHEN, & GORINEVSKY, D. 2007b. An interior-point method for large-scale ℓ_1-regularised least squares. *Journnal of Selected Topics in Signal Processing*, **1**(4), 606 – 617. 49, 50, 71, 106

KOTHARE, MAYURESH V., BALAKRISHNAN, VENKATARAMANAN, & MORARI, MANFRED. 1996. Robust constrained model predictive control using linear matrix inequalities. *Automatica*, **32**(10), 1361 – 1379. 39, 113

KUWATA, Y., RICHARDS, A., & HOW, J. 2007. Robust Receding Horizon Control using Generalized Constraint Tightening. *Pages 4482 – 4487 of: Proceedings of the American Control Conference (ACC)*. 45, 182

KVASNICA, M., GRIEDER, P., & BAOTIĆ, M. 2004. *Multi-Parametric Toolbox (MPT) Version 2.6.3*. 78, 124, 125, 129, 130, 141, 145, 183

LAZAR, M. 2006. *Model Predictive Control of Hybrid Systems: Stability and Robustness*. Ph.D. thesis, Technische Universiteit Eindhoven. 27, 47, 161

LAZAR, M. 2010. On Infinity Norms as Lyapunov Functions: Alternative Necessary and Sufficient Conditions. *Pages 5936–5942 of: Proceedings of the Conference on Decision and Control (CDC)*. 101

LAZAR, M., & HEEMELS, W.P.M.H. 2009. Predictive control of hybrid systems: Input-to-state stability results for sub-optimal solutions. *Automatica*, **45**(1), 180 – 185. 34

LAZAR, M., & JOKIC, A. 2010. On infinity norms as Lyapunov functions for piecewise affine systems. *Pages 131 – 141 of: Proceedings of the Hybrid Systems Computation and Control Conference*. 118, 173

LAZAR, M., HEEMELS, W.P.M.H, WEILAND, S., & BEMPORAD, A. 2006. Stabilizing Model Predictive Control of Hybrid Systems. *IEEE Transaction on Automatic Control*, **51**(11), 1813–1818. 34, 92, 113

LAZAR, M., HEEMELS, W.P.M.H., MUÑOZ DE LA PEÑA, D., & ALAMO, T. 2009. Further Results on Robust MPC Using Linear Matrix Inequalities. *Pages 89–98 of:* MAGNI, LALO, RAIMONDO, DAVIDEMARTINO, & ALLGWER, FRANK (eds), *Nonlinear Model Predictive Control*. Lecture Notes in Control and Information Sciences, vol. 384. Springer Berlin Heidelberg. 38

LERNER, D. 2010. *Lecture notes of Math 500 – Introduction to mathematical analysis*. University of Kansas. 14

LIMON, D., ALAMO, T., & CAMACHO, E.F. 2002. Input-to-state stable MPC for constrained discrete-time nonlinear systems with bounded additive uncertainties. *Pages 4619 – 4624 of: Proceedings of the Conference on Decision and Control (CDC)*. 165

LIMON, D., ALAMO, T., & CAMACHO, EDUARDO F. 2005. Enlarging the domain of attraction of MPC controllers. *Automatica*, **41**(4), 629–635. 106

LIMÓN, D., ALAMO, T., SALAS, F., & CAMACHO, E. F. 2006. Input to state stability of min-max MPC controllers for nonlinear systems with bounded uncertainties. *Automatica*, **42**(5), 797–803. 26, 28, 83, 169

LIMON, D., ALAMO, T., SALAS, F., & CAMACHO, E.F. 2006. On the stability of constrained MPC without terminal constraints. *IEEE Transaction on Automatic Control*, **51**(6), 832 – 836. 46

LIMON, D., ALVARADO, I., ALAMO, T., & CAMACHO, E.F. 2008. MPC for tracking piecewise constant references for constrained linear systems. *Automatica*, **44**(9), 2382–2387. 46, 161, 169, 170, 171

LIMON, D., ALAMO, T., RAIMONDO, D. M., MUÑOZ DE LA PEÑA, D., BRAVO, J. M., FERRAMOSCA, A., & CAMACHO, E. F. 2009. Input-to-State Stability: A Unifying Framework for Robust Model Predictive Control. *Nonlinear Model Predictive Control: Lecture Notes in Control and Information Sciences*, **384**, 1–26. 41, 42, 47, 161

LIN, F., FARDAD, M., & JOVANOVIC, M.R. 2012. Sparse feedback synthesis via the alternating direction method of multipliers. *Pages 4765–4770 of: Proceedings of the American Control Conference (ACC)*. 52, 53

LJUNG, L., HJALMARSSON, H., & OHLSSON, H. 2011. Four Encounters with System Identification. *European Journal of Control*, **17**(5-6), 449–471. 50, 51

LOFBERG, J. 2004. YALMIP: a toolbox for modeling and optimization in MATLAB. *Pages 284–289 of: IEEE International Symposium on Computer Aided Control Systems Design*. 60, 129, 139, 141

MACIEJOWSKI, J.M. 2002. *Predictive Control with Constraints*. Pearson Education. 22, 31, 32, 33, 46, 47, 54, 161

MARIÉTHOZ, S., DOMAHIDI, A., & MORARI, M. 2012. High-Bandwidth Explicit Model Predictive Control of Electrical Drives. *IEEE Transaction on Industry Applications*, **48**(6), 1980–1992. 42

MAYNE, D. Q., RAWLINGS, J., RAO, C., & SCOKAERT, P. 2000. Constrained model predictive control: Stability and optimality. *Automatica*, **36**(6), 789 – 814. 32, 85, 88

MUSKE, K. R. 1997. Steady-State Target Optimization in Linear Model Predictive Control. *Pages 3597–3601 of: Procedings of the American Control Conference (ACC)*. 163

NAGAHARA, M., & QUEVEDO, D. E. 2011. Sparse Representations for Packetized Predictive Networked Control. *Pages 84 – 89 of: IFAC World Congress*. 52, 53, 70, 86, 87

NESTEROV, Y. 2009. Primal-dual subgradient methods for convex problems. *Mathematical Programming Series B*, **120**(1), 221 – 259. 22

NOCEDAL, J., & WRIGHT, S. J. 1999. *Numerical Optimisation*. Operations Research Series. Springer. 22

OHLSSON, H. 2010. *Regularization for Sparseness and Smoothness*. Ph.D. thesis, Department of Electrical Engineering, Linkping University, Sweden. 49

OHLSSON, H., LJUNG, L., & BOYD, S. 2010a. Segmentation of ARX-Models Using Sum-of-Norms Regularization. *Automatica*, **46**(6), 1107–1111. 49, 50

OHLSSON, H., GUSTAFSSON, F., LJUNG, L., & BOYD, S. 2010b. Trajectory Generation Using Sum-of-Norms Regularization. *Pages 540 – 545 of: Proceedings of the Conference on Decision and Control (CDC)*. 49, 50, 52

OHLSSON, H., GUSTAFSSON, F., LJUNG, L., & BOYD, S. 2012. Smoothed state estimates under aboupt changes using sum-of-norms. *Automatica*, **48**(4), 595–0605. 50, 52

OSBORNE, M.R., & PRESNELL, B. 2000. A new approach to variable selection in least squares problems. *IMA Journal of Numerical Analysis*, **20**(3), 389–403. 49, 50, 51, 52, 71, 73

OSBORNE, M.R., PRESNELL, B., & TURLACH, A. 2000. On the LASSO and its dual. *Journal of Computational and Graphical Statistics*, **9**(2), 319–337. 16

PANNOCCHIA, G., & BEMPORAD, A. 2007. Combined Design of Disturbance Model and Observer for Offset-Free Model Predictive Control. *IEEE Transaction on Automatic Control*, **52**(6), 1048–1053. 163

PANNOCCHIA, G., & RAWLINGS, J.B. 2003. Disturbance models for offset-free model predictive control. *AIChE Journal*, **49**(2), 426 – 437. 163

PANNOCCHIA, G., RAWLINGS, J.B., & WRIGHT, S. J. 2011. Conditions under which suboptimal nonlinear MPC is inherently robust. *Systems & Control Letters*, **60**(9), 747–755. 34

PEREZ, T. 2005. *Ship Motion Control. Course Keeping and Roll Stabilization Using Rudder and Fins*. Springer. 193, 194, 195, 196, 199, 201, 202

PEREZ, T., & GOODWIN, G.C. 2008. Constrained predictive control of ship fin stabilisers to prevent dynamic stall. *Control Engineering Practice*, **16**(4), 482–494. xii, xv, 3, 193, 194, 197, 201

PEREZ, T., SMOGELI, O.N., FOSSEN, T.I., & SORENSEN, A. J. 2006. An Overview Of The Marine System Simulator MSS: A Simulink Toolbox For Marine Control Systems. *Journal of Modelling, Identification and Control*, **27**(4), 259–275. 201

RAKOVIĆ, S., & LAZAR, M. 2012. Minkowski terminal cost functions for MPC. *Automatica*, **48**(10), 2721 – 2725. 92, 96

RAO, C.V., & RAWLINGS, J.B. 2000. Linear programming and model predictive control. *Journal of Process Control*, **10**(2-3), 283–289. 33

RAWLINGS, J. B., & MAYNE, D. Q. 2010. *Model Predictive Control Theory and Design*. Nob Hill Pub, Llc. 26, 27, 32, 33, 34, 35, 36, 41, 83, 85, 88, 119, 174

RICHARDS, A., & HOW, J. 2006. Robust stable model predictive control with constraint tightening. *Pages 1558 – 1562 of: Proceedings of the American Control Conference (ACC)*. 45, 182

RICHTER, S., JONES, C. N., & MORARI, M. 2013. Certification aspects of the fast gradient method for solving the dual of parametric convex programs. *Mathematical Methods of Operations Research*, **77**(3), 305–321. 18, 22

ROBBINS, J. W. *Lecture notes of Math 521: Advanced calculus*. University of Wisconsin. 14

SCHMIDT, M. 2010. *Graphical Model Structure Learning with L1-Regularization, PhD Thesis*. Ph.D. thesis, Univ. of British Columbia. 49

SCHUET, S. 2010. Wiring Diagnostics via ℓ_1-regularized Least Sqaures. *IEEE Sensor Journal*, **10**(7), 1218 – 1225. 49, 52

SCOKAERT, P.O.M., MAYNE, D.Q., & RAWLINGS, J.B. 1999. Suboptimal model predictive control (feasibility implies stability). *IEEE Transaction on Automatic Control*, **44**(3), 648 –654. 34

SHEKHAR, ROHAN C., & MACIEJOWSKI, JAN M. 2012. Optimal constraint tightening policies for robust variable horizon model predictive control. *Pages 5170 –5175 of: Proceedings of the Conference on Decision and Control (CDC)*. 45, 182

SPJTVOLD, J., TNDEL, P., & JOHANSEN, T. A. 2007. Continuous Selection and Unique Polyhedral Representation of Solutions to Convex Parametric Quadratic Programs. *Journal of Optimization Theory and Applications*, **134**(2), 177–189. 97

SUMMERS, S., JONES, COLIN, LYGEROS, J., & MORARI, M. 2011. A multiresolution approximation method for fast explicit model predictive control. *IEEE Transactions on Automatic Control*, **56**(11), 2530–2541. 45

TIBSHIRANI, R. 1996. Regression Shrinkage and Selection via the LASSO. *Journal of the Royal Statistical Society*, **58**(1), 267–288. 49, 50, 52, 74, 75

TONDEL, P., JOHANSEN, T.A., & BEMPORAD, A. 2003. Further results on multiparametric quadratic programming. *Pages 3173 – 3178 of: Proceedings of the Conference on Decision and Control (CDC)*. 42, 43, 77, 122, 125, 129, 133, 140

TRENCH, W. F. 2013. *Introduction to Real Analysis*. Pearson Education. 39

VIDYASAGAR, M. 1993. *Nonlinear Systems Analysis*. 2nd edn. Classics in Applied Mathematics. New Jersey: SIAM: Society for Industrial and Applied Mathematics. 26

VINNICOMBE, G. 2000. *Uncertainty And Feedback, H_∞ Loop Shaping And The ν Gap Metric*. Imperial College press. 45, 58

VINTER, R. 2000. *Optimal Control*. Birkhäuser. 16

WALKER, M. *Lecture notes of Economics 519*. University of Arizona. 14

WANG, Y., & BOYD, S. 2010. Fast Model Predictive Control Using Online Optimization. *IEEE Transaction on Control Systems Technology*, **18**(2), 267–278. 18, 22

WILLEMS, J.C. 2004. Deterministic least squares filtering. *Journal of Econometrics*, **118**(1-2), 341–373. 49

ZEILINGER, M.N., JONES, C.N., & MORARI, M. 2010. Robust stability properties of soft constrained MPC. *Pages 5276 – 5282 of: Proceedings of the Conference on Decision and Control (CDC)*. 46, 47, 161, 165, 169, 170, 175

ZOU, HUI, & HASTIE, TREVOR. 2005. Regularization and variable selection via the Elastic Net. *Journal of the Royal Statistical Society, Series B*, **67**(2), 301–320. 49, 51, 52