
Contributions to nonlinear system modelling and controller synthesis via convex structures



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Author: Ruben Robles Ruiz

Supervisors: D. Antonio Sala Piqueras
D. Miguel Ángel Bernal Reza

February 21, 2018

Abstract

This thesis discusses different modelling methodologies to eke out better performance/stability results than conventional sector-nonlinearity Takagi-Sugeno (also known as quasi-LPV) systems modelling techniques are able to yield.

Indeed, even if LMIs can prove various performance and stability bounds (decay rate, \mathcal{H}_∞ , etc.) for polytopic systems, it is well known that the proven performance depends on the chosen model and, given a nonlinear dynamic systems, the polytopic embeddings available for it are not unique. Thus, explorations on how to obtain the model which is less deleterious for performance are presented.

As a last contribution, extending the polytopic Takagi-Sugeno setup to a gain-scheduled quasi-convex difference inclusion framework allows to improve the results over the polytopic models. Indeed, the non-scheduled convex difference inclusion framework was proposed by a research team in University of Seville (Fiacchini, Alamo, Camacho) as a generalised modelling methodology which included the polytopic one; this thesis poses a further generalised gain-scheduled version of some of these results.

Resumen

Esta tesis discute diferentes metodologías de modelado para extraer mejores prestaciones o resultados de estabilidad que aquéllas que el modelado convencional basado en sector no-lineal de sistemas Takagi-Sugeno (también denominados cuasi-LPV) es capaz de producir.

En efecto, incluso si las LMIs pueden probar distintas cotas de prestaciones o márgenes de estabilidad (tasa de decaimiento, \mathcal{H}_∞ , etc.) para sistemas politópicos, es bien conocido que las prestaciones probadas dependen del modelo elegido y, dado un sistema no-lineal, dicho modelo politópico no es único. Por tanto, se presentan exploraciones hacia cómo obtener el modelo que es menos perjudicial para la medida de prestaciones elegida.

Como una última contribución, mejores resultados son obtenidos mediante la extensión del modelado politópico Takagi-Sugeno a un marco de inclusiones en diferencias cuasi-convexas con planificación de ganancia. En efecto, una versión sin planificación de ganancia fue propuesta por un equipo de investigadores de la Universidad de Sevilla (Fiaccini, Álamo, Camacho) para generalizar el modelado politópico, y esta tesis propone una versión aún más general de algunos de dichos resultados que incorpora planificación de ganancia.

Resum

Aquesta tesi discuteix diferents metodologies de modelatge per extreure millors prestacions o resultats d'estabilitat que aquelles que el modelatge convencional basat en sector no-lineal de sistemes Takagi-Sugeno (també anomenats quasi-LPV) és capaç de produir.

En efecte, fins i tot si les LMIs poden provar diferents cotes de prestacions o marges d'estabilitat (taxa de decaïment, \mathcal{H}_∞ , etc.) per a sistemes politòpics, és ben conegut que les prestacions provades depenen del model triat i, donat un sistema no-lineal, el dit model politòpic no és únic. Per tant, es presenten exploracions cap a com obtenir el model que és menys perjudicial per a la mesura de prestacions triada.

Com una darrera contribució, millors resultats són obtinguts mitjançant l'extensió del modelatge politòpic Takagi-Sugeno a un marc d'inclusions en diferències quasi-convexes amb planificació de guany. En efecte, una versió sense planificació de guany va ser proposada per un equip d'investigadors de la Universitat de Sevilla (Fiaccini, Álamo, Camacho) per a generalitzar el modelatge politòpic, i aquesta tesi proposa una versió més general d'alguns d'aquests resultats que incorpora planificació de guany.

Contents

Abstract	iii
Contents	ix
Summary of the thesis	1
Motivation and objectives	1
Structure of the thesis	2
Contributions	4
I State of the Art	7
1 Linear matrix inequalities in control	9
1.1 Definitions, problems, and properties	9
1.2 LMIs in linear time-invariant systems	14
2 LMI-based nonlinear control	25
2.1 Takagi-Sugeno models	25
2.2 Nonlinear control	31
3 Set-based control theory	51
3.1 Invariant sets	51
3.2 Convex functions and sets	54
3.3 One-step sets	56
3.4 Polytopes and Multi-Parametric programming (MPT)	63
3.5 Convex Difference Inclusions	65
II Contributions	71
4 Choosing a Takagi-Sugeno model for improved performance	73
4.1 Introduction	74

4.2	Preliminaries	75
4.3	Linear Coordinate Transformations	77
4.4	A Hessian-based methodology	78
4.5	Linearization-based methodology	82
4.6	Conclusions	86
5	Subspace-Based Modelling for Improved LMI Performance	89
5.1	Introduction	90
5.2	Preliminaries	91
5.3	Systematic TS modelling	96
5.4	Coordinate transformations	99
5.5	TS models with optimal performance in a subspace	102
5.6	Use of optimal TS models in LMIs	110
5.7	Discussion	113
5.8	Additional Examples	114
5.9	Conclusions	119
6	Optimal-Performance Models via the LMI Null Space	123
6.1	Introduction	124
6.2	Preliminaries	125
6.3	Main result	129
6.4	Decay-Rate analysis example	133
6.5	Conclusion	136
7	On Modelling of Nonlinear Systems as Quasi-LPV Ones	137
7.1	Introduction	138
7.2	Preliminaries	140
7.3	Main Result	144
7.4	Examples	152
7.5	Conclusions	160
8	Gain-Scheduled Control via Polytopic Difference Inclusions	163
8.1	Introduction	164
8.2	Preliminaries	165
8.3	Set-based approach to dynamic system analysis and control design	168
8.4	Main result: computation of gain-scheduled one-step sets	172
8.5	Convexification of product nonlinearities	178
8.6	Numerical Example	179
8.7	Conclusions	182
8.8	Appendix	183

III	Conclusions and Bibliography	187
	Conclusions	189
	Bibliography	193

Summary of the thesis

Motivation and objectives

As is well known in the literature on nonlinear systems modeling, it is possible in a large number of cases to represent these systems as polytopic vertex models (in most cases, locally, over a so-called modelling region), this allows us to make use of linear matrix inequalities (LMI) and other mathematical tools, to guarantee different performance objectives in such systems with a low computation cost. The problem that results in most of the approaches that address this issue, is that their methodologies are not systematic or do not provide any guidance in which, of the infinite options of representations that result from their theories, is more convenient for the nonlinear system that we want to analyze, yielding in a different performance bounds depending on the model choice. To avoid such drawback, earlier literature on the topic used another approaches, which often require gridding on the modelling region and bilinear matrix inequalities (BMI) conditions, in these cases obtain a good performance bound is often equivalent to a huge computational burden.

The central objective of this thesis is to provide some guidelines for the selection and construction of these polytopic vertex models based on some of the most common performance objectives in the control community, considering only the use of linear matrix inequalities. For this purpose we present different modeling methodologies that adopt different approaches in order to satisfy our requirements and overcome the performance of some of the most widely treated approaches in the literature. A last objective of the thesis is departing from linear “vertex

models” to a more general class of models whose bounds have suitable convexity-related properties.

Structure of the thesis

This thesis is divided in three parts:

- Part I summarises the most fundamental results in the literature related to the framework of this thesis. In Chapter 1, an overview of the concept of an LMI and its use for the stability analysis of linear models is conducted, some common matrix properties to transform matrix inequalities into LMIs and some stability concepts are reviewed.

Chapter 2 shows how the direct Lyapunov method is employed altogether with the convex structure of the TS model to express stability and stabilisation conditions in terms of LMIs.

Chapter 3 introduces some concepts related to set theory in the field of control and the analysis of dynamic systems. The basic definitions will be given and some properties of the invariant sets will be described. Properties of convex set and convex functions will be reviewed, these will later be used in later chapters making evident the importance of invariance and methods based on set theory for control.

- Part II contains the contributions of this work. There are two main contribution areas in this part:
 1. Chapters 4 to 7 discuss how to address the problem of achieving minimal performance loss with respect to a given LMI control problem over a linearised model.
 2. Then, Chapter 8 departs from this framework and explores a totally different set-based approach.

The first contribution is presented in Chapter 4, where are introduced two methodologies intended to address the problem of constructing the most adequate Takagi-Sugeno representation of a nonlinear system with respect to a performance goal: maximum decay rate. Both approaches are based on coordinate transformations.

Chapter 5 provides some guidelines on choosing a sector-nonlinearity Takagi-Sugeno model, with provable optimality in the case of quadratic nonlinearities. The approach is based on Hessian and restrictions of a function onto a subspace.

Chapter 6 relaxes the condition of state-dependency of approaches in the previous chapters, thus allowing multiple-LMI setups which might be associated with a wider variety of performance measures. The main idea here is based on a Frobenius-norm bound on the “perturbation” that sector-nonlinearity models produce in the LMI matrices.

Chapter 7 instead of writing LMIs for a given model, it assumes that LMIs for a given performance optimisation problem are already available and sets out for a quest to determine the optimal model that keeps the proven performance as close as possible to the linearised one, which is the best one possible.

The structure of this four chapters is “progressive” in the sense that it mimics the chronological order of submission of different ideas on the problem during the Ph.D. studies. Thus, basically each chapter builds upon the previous one, complementing and refining its approaches; hence, they share common issues in problem statement and initial steps of the developed methodologies.

Chapter 8 presents the generalisation of control design via gain-scheduling from TS models to the so-called parameter-dependent quasiconvex difference inclusions. This will enable to enlarge polyhedral domain of attraction estimates for nonlinear control.

- This thesis ends in part III, drawing some concluding remarks.

Note that most of the content of part II is a verbatim copy of published material (indicated at the beginning of each chapter). Thus, there may be repetitions of preliminary material and notation changes, particularly on the progressive improvements from chapters 4 to 7. This repetition is intentional to keep the chapters basically identical to the published material.

Contributions

Listed below is part of the work done during the period of doctoral studies, presenting only works published/submitted to date. It should be noted that the publications [2,5,6] are outside the main thematic framework of interest of this thesis, done in collaboration with other researchers at the team, but not included in the contribution part of this document.

Journal papers

1. R. Robles, A. Sala, M. Bernal, and T. Gonzalez (2017). “Subspace-Based Takagi-Sugeno Modeling for Improved LMI Performance”. In: *IEEE Transactions on Fuzzy Systems* 25.4, pp. 754–767.
2. T. Gonzalez, A. Sala, M. Bernal, and B. Aguiar (2017c). “Piecewise-Takagi-Sugeno asymptotically exact estimation of the domain of attraction of nonlinear systems”. In: *Journal of the Franklin Institute* 354.3, pp. 1514–1541.

Conference papers

3. R. Robles, A. Sala, M. Bernal, and T. Gonzalez (2015). “Choosing a Takagi-Sugeno model for improved performance”. In: *Proceedings of the 2015 IEEE International Conference on Fuzzy Systems*. Istanbul, Turkey, pp. 1–6.
4. R. Robles, A. Sala, M. Bernal, and T. Gonzalez (2016). “Optimal-Performance Takagi-Sugeno Models via the LMI Null Space”. In: (*Proc. of 4th IFAC Conf. Intell. Control and Autom. Sc.*) *IFAC-PapersOnLine* 49.5, pp. 13–18.
5. T. Gonzalez, A. Sala, M. Bernal, and R. Robles (2015). “Invariant Sets of Nonlinear Models via Piecewise Affine Takagi-Sugeno Model”. In: *Proceedings of the 2015 IEEE International Conference on Fuzzy Systems*. Istanbul, Turkey.
6. T. Gonzalez, A. Sala, M. Bernal, and R. Robles (2017a). “A generalisation of Line-Integral Lyapunov Function for Takagi-Sugeno systems”. In: *Proceedings of the 20th World Congress of the International Federation of Automatic Control*. Toulouse, France.

Contributions under review/in elaboration.

The contents of chapters 7 and 8 are the core of two submissions currently under review. The manuscript data are:

7. R. Robles, A. Sala, M. Bernal, and T. Gonzalez, “On Modelling of Nonlinear Systems as Quasi-LPV Ones”, Under Review.
8. A. Sala, R. Robles, C. Ariño, “Gain-Scheduled Control of Nonlinear Systems via Polytopic Difference Inclusions”, Under Review.

If accepted, the final published version will differ from the one in this thesis, arising from modifications suggested by reviewers and, too, from refinements by the authors after the thesis draft was submitted to the Doctoral School.

Part I

State of the Art

Chapter 1

Linear matrix inequalities in control

The last three decades have witnessed the publication of a lot of results where so-called linear matrix inequalities play an important role for the study of nonlinear control systems. Their usefulness is twofold: (a) many control problems can be cast as linear matrix inequalities, and (b) they belong to the class of convex optimisation problems, for which there exist computationally efficient methods to solve them. This chapter is a short introduction to this subject originally emerged, as many other branches of control theory, in the context of linear time-invariant systems.

1.1 Definitions, problems, and properties

1.1.1 Definitions

As its name indicates, a *linear matrix inequality* (LMI) is an inequality whose sides are linear functions of matrix variables; it is always amenable to the form

$$A_0 + \underbrace{A_1 M_1 B_1 + A_2 M_2 B_2 + \cdots + A_n M_n B_n}_{f(M_1, M_2, \dots, M_n)} > 0, \quad (1.1)$$

where $f(\cdot, \cdot, \dots, \cdot)$ is a symmetric linear function of *decision variables* M_1, M_2, \dots, M_n , which are pre- and post-multiplied by known matrices A_1, A_2, \dots, A_n and

B_1, B_2, \dots, B_n , respectively; A_0 is a known symmetric matrix, i.e., $A_0 = A_0^T$; and (> 0) stands for *positive-definiteness*.

A symmetric matrix M is said to be *positive-definite*, i.e., $M > 0$, if it holds $x^T M x > 0$ for all $x \in \mathbb{R}^n$, $x \neq 0$. A similar definition can be made for a *negative-definite* matrix, denoted as $M < 0$, which with the previous one constitute the class of *strict LMIs*. Non-strict ones arise from the sets of *positive* and *negative semi-definite* matrices denoted as $M \geq 0$ and $M \leq 0$, respectively. Importantly, in all cases $M = M^T$. Positive-definiteness of a symmetric matrix M is equivalent to its eigenvalues having positive real parts; abusing the notation, if $\sigma(M)$ denotes the spectrum of M , we express this fact as $M = M^T > 0 \iff \text{Re}(\sigma(M)) > 0$.

In many references such as (Boyd et al. 1994), an LMI is expressed in terms of a *vector of decision variables* $x \in \mathbb{R}^m$, i.e.:

$$F(x) = F_0 + \sum_{i=1}^n x_i F_i > 0, \quad (1.2)$$

where $F_i = F_i^T \in \mathbb{R}^{n \times n}$, $i = 0, \dots, n$ are given constant symmetric matrices. Again, keep in mind that (> 0) stands for positive-definiteness and that this can be adapted to reflect semi-definiteness if required. Naturally, both definitions 1.1 and 1.2 are equivalent, since the entries of the vector x are the different scalar entries in matrices M_i , $i \in \{1, 2, \dots, n\}$.

Example 1.1. Consider $A = \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix}$ along with symmetric matrix $P = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}$. The LMI $PA + A^T P < 0$, written in the form (1.1), can be put into form (1.2) by doing the following:

$$\begin{aligned} PA + A^T P &= \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix}^T \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \\ &= \begin{bmatrix} -2p_{11} & 2p_{11} - 4p_{12} \\ 2p_{11} - 4p_{12} & 4p_{12} - 6p_{22} \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} -2 & 2 \\ 2 & 0 \end{bmatrix}}_{F_1} \underbrace{p_{11}}_{x_1} + \underbrace{\begin{bmatrix} 0 & -4 \\ -4 & 4 \end{bmatrix}}_{F_2} \underbrace{p_{12}}_{x_2} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & -6 \end{bmatrix}}_{F_3} \underbrace{p_{22}}_{x_3}. \end{aligned}$$

The *feasibility set* of an LMI –using definition (1.2)– is the subset of \mathbb{R}^m such that $S = \{x \in \mathbb{R}^m, F(x) > 0\}$. It turns out that the feasibility set of an LMI is always a convex set (Scherer 2004). A convex set S is such that for every pair $x, y \in S$ it follows that $\lambda x + (1 - \lambda)y \in S$, $0 \leq \lambda \leq 1$.

1.1.2 LMI problems

A *convex optimisation problem* consists in minimising a convex function $f(x)$ on the set $S = \{x \in \mathbb{R}^m, g(x) \leq 0, h(x) = 0\}$ where $g(x) : \mathbb{R}^m \rightarrow \mathbb{R}^k$ is also convex and $h(x) : \mathbb{R}^m \rightarrow \mathbb{R}^l$ is affine. Such sort of problems can be solved in polynomial time and an instance of x holding the constraints be found. Convexity of the sets on which these functions or problems are defined is relevant as it guarantees no local minima will be found and feasibility can always be decided. If no solution exists, the corresponding problem is called *infeasible*. In other words, should a problem be stated as an LMI one, we can rest assured that a solution will be found if it exists or, otherwise stated, that we will know if there is no solution.

The following well-known convex or quasi-convex optimization problems are relevant for analysis and synthesis of control systems (Boyd et al. 1994; Scherer 2004):

1. *Feasibility problem (FP)*: Finding a solution x to the LMI system (1.2) or determine that there is none; it is equivalent to minimising the convex function $f : x \rightarrow \lambda_{\min}(F(x))$ and then deciding whether the solution is positive (strictly feasible solution), zero (feasible solution), or negative (unfeasible case).
2. *Optimization problem*: Minimise a linear combination of the decision variables, i.e., $c_1x_1 + c_2x_2 + \dots + c_mx_m$, where $c = [c_1 \ c_2 \ \dots \ c_m]^T$ is a given vector of coefficients and $x = [x_1 \ x_2 \ \dots \ x_m]^T$ is the decision vector. This problem belongs to semi-definite programming (SDP) as it is a generalisation of linear programming for the cone of positive semi-definite matrices.
3. *Generalised eigenvalue problem (GEVP)*: Minimise eigenvalues of a pair of matrices which depend affinely on a variable, subject to a set of LMI-constraints or determine that the problem is infeasible. More precisely:

$$\min \lambda : \lambda B(x) - A(x) > 0, \quad B(x) > 0, \quad C(x) > 0, \quad (1.3)$$

where $A(x)$, $B(x)$ and $C(x)$ are symmetric and affine with respect to x . The problem can also be rewritten as

$$\min \lambda_{\max}(A(x), B(x)) : B(x) > 0, \quad C(x) > 0, \quad (1.4)$$

where $\lambda_{\max}(X, Y)$ denotes the largest generalised eigenvalue of $\lambda Y - X$ with $Y > 0$. This is a quasi-convex problem which can be solved by bisection on λ and LMIs on x .

1.1.3 Solvers

As mentioned before, stating a problem as an LMI one has a computational advantage; therefore, its solution is no longer analytical (Lur'e 1957) nor graphical (Yakubovich 1965) as intended in the past. Since the early eighties LMIs have been solved *numerically*, first through convex programming (Karmarkar 1984), then through interior-point algorithms (Nesterov and Nemirovsky 1988). The user is no longer faced to the task of programming these algorithms as they are already implemented in commercially available software. Most of them are MATLAB-based, such as the LMI Toolbox (Gahinet et al. 1995), which exploits projective methods and linear algebra; SeDuMi (Sturm 1999) and MOSEK (ApS 2015), which are semi-definite programming solvers whose syntax is that of YALMIP (Löfberg 2004); SDPT3 (Tütüncü, Toh, and Todd 2003), VSDP (Härter, Jansson, and Lange 2012), and LMIRank (Orsi 2005). MAXDET (Boyd et al. 1994) and CSDP do not run under MATLAB-related platforms. The examples in this thesis were obtained using SeDuMi or MOSEK with a Yalmip interface. Here and there pieces of code are given to illustrate certain points, but for deeper information on the solver syntax the reader is referred to the documentation cited in this section.

1.1.4 Properties

Despite the general belief, casting a control problem in terms of LMIs might be a challenging task as the involved conditions may require a number of matrix manipulations before yielding its convex nature (if it is so) or a number of conservative assumptions to obtain a reduced convex formulation. The following are some of the most common matrix properties that can be found in the LMI-related literature:

Property 1.1 (Grouping). *A set of LMIs $F_1(x) > 0, \dots, F_k(x) > 0$ is equivalent to the single LMI $F(x) = \text{diag}[F_1(x) \ \dots \ F_k(x)] > 0$ where $\text{diag}(\dots)$ denotes the block-diagonal matrix with $F_1(x), F_2(x), \dots, F_k(x)$ on its main diagonal.*

Property 1.2 (Schur Complement). *Given a matrix $Q(x) \in \mathbb{R}^{m \times m}$, $Q(x) > 0$, a full rank-by-row matrix $S(x) \in \mathbb{R}^{n \times m}$ and a matrix $R(x) \in \mathbb{R}^{n \times n}$, all of them depending affinely on x , the following inequalities are equivalent:*

$$\begin{bmatrix} R(x) & S(x) \\ S(x)^T & Q(x) \end{bmatrix} > 0 \quad (1.5)$$

$$R(x) - S(x)Q(x)^{-1}S(x)^T > 0. \quad (1.6)$$

Property 1.3 (Slack variables (discrete)). *Given matrices A, G, L, P and Q with appropriate sizes, the following inequalities are equivalent (Oliveira and Skelton*

2001)

$$A^T P A - Q < 0, \quad P > 0 \quad (1.7)$$

$$\begin{bmatrix} -Q & A^T P \\ P A & -P \end{bmatrix} < 0 \quad (1.8)$$

$$\exists G : \begin{bmatrix} -Q & A^T G \\ G^T A & -G - G^T + P \end{bmatrix} < 0, \quad P > 0 \quad (1.9)$$

$$\exists G, L : \begin{bmatrix} -Q + A^T L + L A & -L + A^T G \\ -L^T + G^T A & -G - G^T + P \end{bmatrix} < 0, \quad P > 0 \quad (1.10)$$

Property 1.4 (Slack variables (continuous)). *Given matrices $A, G, L, P = P^T$ and Q with appropriate sizes, the following inequalities are equivalent (Peaucelle et al. 2000):*

$$A^T P + P A + Q < 0 \quad (1.11)$$

$$\exists G, L : \begin{bmatrix} A^T L + L A + Q & P - L + A^T G \\ P - L^T + G^T A & -G - G^T + P \end{bmatrix} < 0 \quad (1.12)$$

Property 1.5 (S-procedure). *Given matrices $F_i = F_i^T, i \in \{0, 1, \dots, p\}$, the following inequalities are equivalent:*

$$x^T F_0 x > 0, \forall x \neq 0 : x^T F_i x \geq 0, \forall i \in \{1, 2, \dots, p\} \quad (1.13)$$

$$\exists \tau_1, \dots, \tau_p \geq 0 : F_0 - \sum_{i=1}^p \tau_i F_i > 0. \quad (1.14)$$

Property 1.6 (Finsler's Lemma). *Given a vector $x \in \mathbb{R}^n$ and matrices $Q = Q^T \in \mathbb{R}^{n \times n}, R \in \mathbb{R}^{m \times n}$ and $S \in \mathbb{R}^{m \times n}$ such that $\text{rank}(R) < n, \text{rank}(S) < n$, the following inequalities are equivalent:*

$$x^T Q x < 0, \quad \forall x \neq 0 : R x = 0, S x = 0 \quad (1.15)$$

$$R_{\perp}^T Q R_{\perp} < 0, \quad S_{\perp}^T Q S_{\perp} < 0 \quad (1.16)$$

$$\exists \sigma \in \mathbb{R} : Q - \sigma R^T R < 0, Q - \sigma S^T S < 0 \quad (1.17)$$

$$\exists X \in \mathbb{R}^{n \times m} : Q + S^T X R + R^T X^T S < 0 \quad (1.18)$$

Getting expression (1.16) from (1.18) is usually referred as the elimination lemma for the obvious reason that expression (1.16) is equivalent to (1.18) without X .

Property 1.7 (Congruence). *Given a matrix $P = P^T$ and a full rank-by-column matrix $Q, P > 0 \Rightarrow Q P Q^T > 0$.*

Property 1.8 (Completion of squares). *Given two matrices X, Y of proper size, $\forall Q = Q^T > 0, X Y^T + Y X^T \leq X Q X^T + Y Q^{-1} Y^T$.*

1.2 LMIs in linear time-invariant systems

Many “classical” problems in standard linear control involving Lyapunov and Riccati equations can be cast as LMI ones. As these LMI problems will be the root of generalisations to polytopic-uncertainty (Takagi-Sugeno) in later developments, they will be briefly reviewed. Of course, LMIs can be applied to nonlinear systems (see next chapter) as well as delayed (Richard 2003; Gonzalez et al. 2013), switching (Sala, Hernández-Mejías, and Ariño 2017), etc. but most of these options root on the basic ideas described below.

1.2.1 Stability

LMIs have not always been recognised as such, for they have appeared in control contexts since the very origins of the Lyapunov theory (Lyapunov 1892) and, in the linear context, through the famous Riccati equation (Willems 1971). Nowadays, we refer to them mainly because of its computational properties which allow solving control problems as convex optimisation ones. To see this in some detail, it is better to begin with linear time-invariant (LTI) systems as stability analysis, controller and observer design, maximisation or minimisation of a performance measure, are all well-known tasks that can be easily cast as LMIs.

Consider an autonomous LTI system of the form

$$\dot{x}(t) = Ax(t), \tag{1.19}$$

with $x \in \mathbb{R}^n$ denoting the state and A the system matrix. If this system has an isolated equilibrium point it can only be globally asymptotically stable (GAS) (which includes stable nodes and foci), stable, but not GAS (center), or unstable (saddle points, unstable nodes and foci) (Chen 1984). GAS stability is equivalent to ensuring that $\text{Re}(\sigma(A)) < 0$, i.e., all the eigenvalues of A must have negative real parts. Such matrices are referred to as Hurwitz matrices. Since A is not necessarily symmetric, this is equivalent to the existence of a matrix P such that

$$PA + A^T P < 0. \tag{1.20}$$

The reader surely identifies the previous inequality with an expression resulting from Lyapunov analysis of LTI systems. We will get later into the Lyapunov theory of stability because it is the basic tool in the nonlinear context; nevertheless, it can also be used in the linear one: an LTI system (1.19) is GAS if and only if there exists a quadratic Lyapunov function of the form $V(x) = x^T P x$ associated with it. In other words, in linear contexts quadratic stability is sufficient and necessary for a GAS equilibrium point. Since $V(x)$ needs to be a Lyapunov function,

it must be positive-definite, i.e., $V(x) > 0 \forall x \neq 0$ while $\dot{V}(x) < 0 \forall x \neq 0$. These conditions straightforwardly translate into LMIs (Boyd et al. 1994):

$$\begin{aligned} V(x) &= x^T P x > 0, \forall x \neq 0 \iff P > 0 \\ \dot{V}(x) &= x^T P \dot{x} + \dot{x}^T P x = x^T (PA + A^T P)x < 0, \forall x \neq 0 \iff PA + A^T P < 0. \end{aligned}$$

Since the Lyapunov function can be scaled without loss of generality, long before the arrival of LMIs, a test equivalent to (1.20) has been established via the Lyapunov equation which is a particular case of the algebraic Riccati one: find $P > 0$ such that for a given $Q > 0$, the resulting system of equations in $PA + A^T P = -Q$ holds.

1.2.2 Stabilisation

For non-autonomous LTI systems of the form

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{1.21}$$

where $u \in \mathbb{R}^m$ is an input vector, state feedback has long ago been proposed as a way of driving the state x to 0, a task which is known as *stabilisation*. A classical linear state feedback is

$$u(t) = Fx(t), \tag{1.22}$$

where $F \in \mathbb{R}^{m \times n}$ is a gain which can be proposed for *a posteriori* checking on stability of the closed-loop system (analysis) or precisely calculated to guarantee that the origin of the closed-loop system is GAS (synthesis).

The closed-loop system seems as

$$\dot{x}(t) = (A + BF)x(t), \tag{1.23}$$

whose origin, according to criteria in (1.20), is GAS if $\exists P = P^T > 0$ such that $P(A + BF) + (A + BF)^T P < 0$. For analysis, the latter is an LMI in P as it assumes F is already given; for synthesis, this expression is not an LMI as it has products of decision variables P and F . This is the occasion of recurring to some of the matrix properties of the previous sections.

Indeed, by congruence (1.7), we have that

$$PA + PBF + A^T P + F^T B^T P < 0 \iff AX + BFX + XA^T + XF^T B^T < 0$$

with $X = P^{-1}$ as a full-rank symmetric matrix that pre- and post-multiplies the original expression. Defining $M = FX$ we have that the latter inequality becomes

$$AX + BM + XA^T + M^T B^T < 0, \tag{1.24}$$

which, if feasible, will produce a solution pair X, M , from which the desired gain $F = MX^{-1}$ is univocally defined. The associated quadratic Lyapunov function will be $V(x) = x^T P x$ with $P = X^{-1}$. This development, though very elementary, shows that properties can be used to reveal the LMI nature of expression that does not seem so at first sight.

Example 1.2. *The linear system*

$$\dot{x}(t) = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} x(t) + \begin{bmatrix} 4 \\ -2 \end{bmatrix} u(t)$$

is unstable. The following MATLAB code in Yalmip calls the SeDuMi solver to search $X > 0$ and M such that (1.24) holds; the interested reader is referred to the software documentation in (Sturm 1999) for details:

```

A=[1 2; 0 3]; B=[4; -2];
X=sdpvar(2); M=sdpvar(1,2,'full'); LMI=[];
LMI=[LMI,X>-(1e-6)*eye(2)];
LMI=[LMI,A*X+B*M+(A*X+B*M)'<-(1e-6)*eye(2)];
5 ops=sdpsettings('solver','sedumi','verbose',0,
'warning',0,'shift',1e-5);
SOL=solvesdp(LMI,[],ops);
if SOL.problem==0
10 X=double(X); P=inv(X); M=double(M); F=M*P;
end

```

The following solution has been found, which means the LMIs above are feasible; correspondingly, the control gain $F = MX^{-1}$ and the Lyapunov matrix $P = X^{-1}$ are also given:

$$X = \begin{bmatrix} 2.3203 & -0.6999 \\ -0.6999 & 0.2207 \end{bmatrix}, P = \begin{bmatrix} 9.9325 & 31.5004 \\ 31.5004 & 104.4338 \end{bmatrix},$$

$$M = \begin{bmatrix} -0.3416 & 0.4728 \end{bmatrix}, F = \begin{bmatrix} 11.5009 & 38.6172 \end{bmatrix}.$$

With these results, the control law (1.22) can be applied to the system above. The trajectories from different initial conditions are shown in Fig. 1.1, which shows that the system was indeed stabilised.

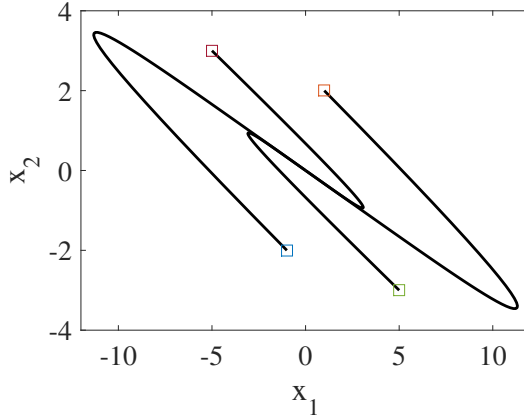


Figure 1.1: Phase portrait.

1.2.3 Observer design

Classical observer design for LTI systems can also be cast as an LMI problem. To see this, consider a system output $y(t) = Cx(t)$, $C \in \mathbb{R}^{p \times n}$, along with the following observer:

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + L(y(t) - \hat{y}(t)), \quad \hat{y}(t) = C\hat{x}(t), \quad (1.25)$$

with $\hat{x} \in \mathbb{R}^n$, $\hat{y} \in \mathbb{R}^p$, $L \in \mathbb{R}^{n \times p}$ being the observer state, the observer output, and the observer gain, respectively. Observer design amounts to calculating L such that $\lim_{t \rightarrow \infty} e(t) = 0$ with $e(t) = x(t) - \hat{x}(t)$; therefore, it follows that

$$\dot{e}(t) = (A - LC)e(t) \quad (1.26)$$

corresponds to the error dynamics. Since this is an LTI system, guaranteeing that $e = 0$ is a GAS equilibrium point is equivalent to the existence of $V(e) = e^T P e$ such that $P = P^T > 0$ and $\dot{V}(e) = e^T P(A - LC)e + e^T (A - LC)^T P e < 0$, $\forall e \neq 0$. The latter inequality is guaranteed if

$$P(A - LC) + (A - LC)^T P = PA - NC + A^T P - C^T N^T < 0, \quad (1.27)$$

where the substitution $N = PL$ has been made as a solution pair P, N univocally defines $L = P^{-1}N$.

1.2.4 Discrete-time systems

Similar results to the ones above can be obtained in the discrete-time case, i.e., systems of the form

$$x(t+1) = Ax(t) + Bu(t). \quad (1.28)$$

Clearly, if $u(t) = 0$ and $V(x) = x^T Px$, $P = P^T > 0$, a Lyapunov function candidate, the resulting autonomous system $x(t+1) = Ax(t)$ is stable if and only if $\Delta V = V(t+1) - V(t) < 0$, a condition that translates into an LMI:

$$\begin{aligned} V(t+1) - V(t) &= x^T(t+1)Px(t+1) - x^T(t)Px(t) \\ &= x^T(t)A^T PAx(t) - x^T(t)Px(t) \\ &= x^T(t) (A^T PA - P) x(t) < 0 \iff A^T PA - P < 0. \end{aligned}$$

The same goes for controller design as linear state feedback (1.22) produces a closed-loop system of the form $x(t+1) = (A + BF)x(t)$ which, when combined with the quadratic Lyapunov function candidate $V(x) = x^T Px$, $P = P^T > 0$ yields the following condition for $\Delta V < 0$:

$$\begin{aligned} (A + BF)^T P(A + BF) - P < 0 &\iff X(A + BF)^T P(A + BF)X - XPX < 0 \\ &\iff X - X(A + BF)^T X^{-1}(A + BF)X > 0 \\ \iff \begin{bmatrix} X & X(A + BF)^T \\ (A + BF)X & X \end{bmatrix} > 0 &\iff \begin{bmatrix} X & XA^T + M^T B^T \\ AX + BM & X \end{bmatrix} > 0, \end{aligned}$$

where $X = P^{-1}$ and $M = BF$. Notice that besides the congruence property 1.7, the Schur complement in property 1.2 has been employed. As before, the substitution of variables univocally defines P and F from the solution pair X, M . Of course, similar developments can be done for observer design (Duan and Yu 2013).

1.2.5 Constraints on input and output

A common requirement in control tasks is to hold a specific bound $\mu > 0$ for the control signal, i.e., $\|u(t)\| < \mu$. Since this condition can be written as

$$\|u(t)\| < \mu \iff \|u(t)\|^2 = u^T u = x^T F^T F x \leq \mu^2 \iff x^T F^T \mu^{-2} F x \leq 1, \quad (1.29)$$

it can be related with the quadratic Lyapunov function by imposing an extra condition on it (without loss of generality):

$$V(x) = x^T Px \leq x^T(0)Px(0) \leq 1. \quad (1.30)$$

Thus, when searching for a control law (1.22), inequalities (1.29) and (1.30) combine as:

$$x^T F^T \mu^{-2} F x \leq x^T P x \leq x^T(0) P x(0) \leq 1 \iff \begin{cases} x^T (P - F^T \mu^{-2} F) x \geq 0 \\ 1 - x^T(0) P x(0) \geq 0 \end{cases} \quad (1.31)$$

By Schur complement –property 1.2– the matrix expression in (1.31) is equivalent to

$$\begin{bmatrix} P & F^T \\ F & \mu^2 I \end{bmatrix} \geq 0, \quad \begin{bmatrix} 1 & x^T(0) \\ x(0) & P^{-1} \end{bmatrix} \geq 0,$$

which, by congruence property 1.7 with $\begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix}$, $X = P^{-1}$ y $M = FX$, becomes

$$\begin{bmatrix} X & M^T \\ M & \mu^2 I \end{bmatrix} \geq 0, \quad \begin{bmatrix} 1 & x^T(0) \\ x(0) & X \end{bmatrix} \geq 0, \quad (1.32)$$

which are already in an LMI form that can be combined with those in (1.24) to guarantee $\|u(t)\| < \mu$.

Similarly, guaranteeing $\|y(t)\| < \lambda$ for an output $y(t) = Cx(t)$ leads to an LMI expression of the form

$$\begin{bmatrix} X & XC^T \\ CX & \lambda^2 I \end{bmatrix} \geq 0, \quad \begin{bmatrix} 1 & x^T(0) \\ x(0) & X \end{bmatrix} \geq 0. \quad (1.33)$$

1.2.6 Performance specification

So far, the control problems just presented let to the LMI solver the task of defining what instance of the solution space (i.e, which specific Lyapunov matrices and gains) is to be proposed. They correspond to the feasibility problem in the previous section, which means they are not concerned with the appropriateness of the solution, but only with its existence. If the LMI solutions above are to be used in real-time-oriented simulations and applications, one must come up with an LMI way to impose restrictions on the control input (actuators energy), the output signals (which might be connected to something else), and a number of performance measures such as decay rate, cost functions, and H_∞ attenuation level.

Pole placement is the basic tool for control performance specification in the linear context as pole location in the complex plane is related with the output settling

time, i.e., the “stabilisation speed” of the controlled signals. Via LMIs, any region in the complex plane of the form

$$\mathcal{R} = \{s \in \mathbb{C} : Q + sS + \bar{s}S^T + \bar{s}Rs\}, \quad (1.34)$$

with \bar{s} being the complex conjugate of s , Q , S , and R matrices of adequate size such that $Q = Q^T$ and $R = R^T$, can be described via LMIs. In the “region template” given by (1.34) we can find strips, circles, ellipses, etc. Since LMIs can be simply stacked together as to hold several restrictions, any intersections of such regions can also be described by LMIs by property 1.1.

As shown in (Scherer 2004), given an autonomous continuous- or discrete-time system of the form (1.19) or (1.28), respectively, the eigenvalues of A are in the region (1.34) if and only if there exists $X > 0$ such that

$$\begin{bmatrix} I \\ A \otimes I \end{bmatrix}^T \begin{bmatrix} X \otimes Q & X \otimes S \\ X \otimes S^T & X \otimes R \end{bmatrix} \begin{bmatrix} I \\ A \otimes I \end{bmatrix} < 0. \quad (1.35)$$

Notice that the stability results given above are particular cases of the LMI above as they intend to guarantee the poles lie in the left half complex plane \mathbb{C}^- for continuous-time systems and within the unitary circle for discrete-time systems.

Decay rate

For controller synthesis, i.e., determining the gain F in (1.22) such that the poles of a continuous- or discrete-time system are within an LMI region (1.34), we begin by referring to the concept of *decay rate*, i.e., the largest Lyapunov exponent α such that $\lim_{t \rightarrow \infty} e^{-\alpha t} |x(t)| = 0$:

1. Continuous-time case: Finding the maximum (or imposing a given) $\alpha > 0$ such that $\dot{V}(x(t)) \leq -2\alpha V(x(t))$.
2. Discrete-time case: Finding the maximum (or imposing a given) $0 \leq \alpha \leq 1$ such that $\Delta V(x(t)) \leq (\alpha^2 - 1)V(x(t))$.

Imposing a decay rate implies placing the poles beyond a certain vertical threshold in the left half complex plane (continuous case) or in a circle smaller than the unitary one (discrete case), which belongs to the class of problems that can be solved via LMIs (1.35). Alternatively, LMIs can be deduced from the conditions listed above; in the continuous-time case (1.23):

$$\dot{V}(x(t)) \leq -2\alpha V(x(t)) \iff AX + BM + XA^T + M^T B^T + 2\alpha X \leq 0$$

or the discrete-time case:

$$\Delta V(x(t)) \leq (\alpha^2 - 1)V(x(t)) \iff \begin{bmatrix} \alpha^2 X & XA^T + M^T B^T \\ AX + BM & X \end{bmatrix} > 0 \quad (1.36)$$

Notice that maximising α subject to the LMIs above is a GEVP problem. In the discrete-case it suffices to substitute α^2 by some new variable $\beta > 0$, then run the GEVP problem over it. GEVP problems have a built-in instruction within the LMI Toolbox; for SeDuMi/MOSEK platforms there is a bisection minimisation function which can be called instead.

Example 1.3. *Obviously, bigger decay rates can be achieved at the cost of bigger size of the control input. To illustrate this, consider the problem of achieving the maximum decay rate for the system*

$$\dot{x}(t) = \begin{bmatrix} 1 & -3 \\ 0 & 5 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 3 \end{bmatrix} u(t), \quad (1.37)$$

with $\|u(t)\| \leq 10$, $x(0) = [-1 \ 2]^T$. LMIs (1.36) should be combined with those in (1.32) as illustrated in the following MATLAB/Yalmip code:

```

A=[1 -3; 0 5]; B=[0; 3]; mu=10; x0=[-1; 2];
sdpvar invalfa;
X=sdpvar(2); M=sdpvar(1,2,'full'); LMI=[];
LMI=[LMI,X>=(1e-6)*eye(2)];
5 LMI=[LMI,invalfa*(A*X+B*M+(A*X+B*M)')+2*X<=-(1e-6)*eye(2)];
LMI=[LMI,[X M'; M mu^2]>=(1e-10)*eye(3)];
LMI=[LMI,[1 x0'; x0 X]>=(1e-10)*eye(3)];
LMI=[LMI invalfa>=0.00001];
Objective = invalfa;
10 SOL=bisection(LMI,Objective,sdpsettings('solver','mosek'))
if SOL.problem==0
    invalfa=double(invalfa); alfa=1/invalfa
    X=double(X); P=inv(X); M=double(M); F=M*P;
15 end

```

It yields a maximum decay rate of $\alpha = 1.7156$. Yet, if the maximum allowable magnitude of the input is $\|u(t)\| \leq 5$, the maximum decay rate falls to $\alpha = 0.0520$; moreover, if $\|u(t)\| \leq 1$, the LMIs become infeasible.

H_∞ disturbance rejection

What is referred as H_∞ disturbance rejection consists in finding a control law (1.22) such that the scaled ℓ_2 gain $\gamma > 0$ (also known as the attenuation level) between the disturbance signal $w(t)$ and output signal $y(t)$ in the LTI system configuration

$$\dot{x}(t) = Ax(t) + Bu(t) + Dw(t), \quad y(t) = Cx(t) + Ew(t), \quad (1.38)$$

subject to a linear state feedback control law $u(t) = Fx(t)$, is minimised (El Ghaoui et al. 1992). In other words, we intend to minimise γ subject to

$$\sup_{\|w(t)\|_2 \neq 0} \frac{\|y(t)\|_2}{\|w(t)\|_2} \leq \gamma. \quad (1.39)$$

Recall that norm ℓ_2 of a signal, let say $y(t)$, is defined as (Doyle, Francis, and Tannenbaum 1992):

$$\left(\int_{-\infty}^{\infty} y^T(t)y(t) \right)^{\frac{1}{2}}. \quad (1.40)$$

Given a Lyapunov function candidate, it can be shown that condition

$$\dot{V} + y^T y - \gamma^2 w^T w \leq 0 \quad (1.41)$$

is sufficient for inequality (1.39) to hold. Therefore,

$$\begin{aligned} \dot{V} + y^T y - \gamma^2 w^T w &= \dot{x}^T P x + x^T P \dot{x} + (Cx + Ew)^T (Cx + Ew) - \gamma^2 w^T w \\ &= x^T P ((A + BF)x + Dw) + ((A + BF)x + Dw)^T P x \\ &\quad + (Cx + Ew)^T (Cx + Ew) - \gamma^2 w^T w \\ &= x^T (PA + PBF + F^T B^T P + A^T P + C^T C) x + x^T (PD + C^T E) w \\ &\quad + w^T (E^T C + D^T P) + w^T (E^T E - \gamma^2 I) w \leq 0 \\ &= \begin{bmatrix} x \\ w \end{bmatrix}^T \begin{bmatrix} PA + PBF + F^T B^T P + A^T P + C^T C & PD + C^T E \\ E^T C + D^T P & E^T E - \gamma^2 I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} \leq 0. \end{aligned}$$

Clearly, the last inequality holds if and only if

$$\begin{bmatrix} PA + PBF + F^T B^T P + A^T P + C^T C & PD + C^T E \\ E^T C + D^T P & E^T E - \gamma^2 I \end{bmatrix} \leq 0,$$

which can be pre- and post-multiplied by $\begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix}$ using the property 1.7, with $X = P^{-1}$ and $M = FX$ to get:

$$\begin{bmatrix} AX + BM + M^T B^T + XA^T + XC^T CX & D + XC^T E \\ D^T + E^T CX & E^T E - \gamma^2 I \end{bmatrix} \leq 0.$$

Applying the Schur complement in property 1.2 to block element (1, 1), we finally arrive to an LMI expression:

$$\begin{bmatrix} AX + BM + M^T B^T + XA^T & XC^T & D + XC^T E \\ & CX & -I \\ D^T + E^T CX & 0 & E^T E - \gamma^2 I \end{bmatrix} \leq 0, \quad (1.42)$$

which, if feasible, sets $F = MX^{-1}$ as the desired gain.

Example 1.4. We intend to apply H_∞ control to minimise the attenuation level between the disturbance signal $w(t)$ and output signal $y(t)$ in the following LTI system

$$\dot{x}(t) = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u(t) + \begin{bmatrix} 0.1 \\ 0.4 \end{bmatrix} w(t), \quad y(t) = x_2.$$

To do so, the following MATLAB code in Yalmip calls the SeDuMi solver to search the minimum γ with $X > 0$ and M such that (1.42) holds:

```

A=[1 2; 2 3]; B=[1; 2]; C=[0 1]; E=0; D=[0.1;0.4];
gamahinfnty=sdpvar(1); X=sdpvar(2); M=sdpvar(1,2,'full'); LMI=[];
LMI=[LMI,X>=(1e-6)*eye(2)]; LMI=[LMI,gamahinfnty>=(1e-10)];
LMI=[LMI,[A*X+B*M+M'*B'+X*A' X*C' D+X*C'*E;
5 C*X -1 0;
D'+E'*C*X 0 E'*E-gamahinfnty]<=-(1e-10)*eye(4)];
ops = sdpsettings('solver','sedumi','verbose',
0,'warning',0,'shift',1e-5);
SOL=optimize(LMI,gamahinfnty); % ,ops);
10 if SOL.problem==0
X=double(X); P=inv(X); M=double(M); F=M*P;
gamahinfnty=sqrt(double(gamahinfnty))
end

```

It was found that $\gamma = 0.2$ is the minimum attenuation level that renders the previous LMIs feasible. Notice that instead of “solvesdp”, the instruction “optimize” has been used.

Chapter 2

LMI-based nonlinear control

For LMIs to be useful in nonlinear contexts, the different constituents –system, controller, observer, etc.– should be amenable to linear-like forms that allow the direct Lyapunov method to be used in a way that mimics the results of the linear context. Linearisation is a first –yet poor– form of achieving such goal; a convex interpolation of models obtained by linearisation at different set points is another one: both of them are approximate. This chapter focuses on the sector nonlinearity approach that allows obtaining an exact convex representation in a compact set of the state space, be in the Takagi-Sugeno or the polynomial form; it then illustrates how analysis and synthesis derived from these forms are valid for the original setup, albeit conditions are only sufficient. The issue of conservativeness is explored to some extent as tackling one of its sources is one of the thesis goals.

2.1 Takagi-Sugeno models

2.1.1 The “fuzzy” origins

Takagi-Sugeno (TS) models originally emerged in the fuzzy literature as a connection between *linguistic fuzzy systems* and ordinary sets of differential or difference equations (Takagi and Sugeno 1985). Up to their appearance, fuzzy systems were considered as sets of “heuristic” linguistic rules intended to describe the behavior of a plant or controller in order to obtain a numerical description of it (Wang

1997). Though powerful for model-free control tasks, fuzzy systems such as the Mamdani one (Mamdani and Assilian 1975) were not fit for mathematical analysis of the sort traditional control systems employed, even if introduction of some “possibilistic uncertainty” was intended to better transcript relevant features of some control problems (Sala and Albertos 1998).

Thus, fuzzy TS models aimed originally to fill the heuristic/formal control gap by proposing a trade-off between linguistic descriptions and traditional state-space representations: instead of describing the actual values of the states, they provided $\dot{x}(t)$ (continuous-time) or $x(t+1)$ (discrete-time) as the consequents of Mamdani fuzzy systems. For illustration, consider the following continuous-time linguistic representation (Tanaka and Wang 2001):

$$\begin{aligned} &\text{IF } z_1(t) \text{ is } M_{i1} \text{ and } \dots \text{ and } z_p(t) \text{ is } M_{ip} \\ &\text{THEN } \dot{x}(t) = A_i x(t) + B_i u(t), \quad i \in \{1, 2, \dots, r\}, \end{aligned}$$

that reduces to the following defuzzified equivalent:

$$\begin{aligned} \dot{x}(t) &= \frac{\sum_{i=1}^r w_i(z(t)) (A_i x(t) + B_i u(t))}{\sum_{i=1}^r w_i(z(t))} \\ &= \sum_{i=1}^r h_i(z(t)) (A_i x(t) + B_i u(t)), \end{aligned}$$

with $M_{ij}(z_j(t))$ being the grade of membership of $z_j(t)$ in fuzzy set M_{ij} , r the number of rules (models), $x(t) \in \mathbb{R}^n$ the state vector, $u(t) \in \mathbb{R}^m$ the input vector, A_i and B_i system matrices of appropriate size (perhaps obtained by linearisation of a nonlinear model at certain operating points, bias neglected), $z_1(t), \dots, z_p(t)$ being known *premise variables* that are functions of the state, exogenous signals, and/or time, $z(t)$ the vector that groups them; $w_i(z(t))$ known as weights and $h_i(z(t))$ known as *membership functions* (MFs), calculated as:

$$w_i(z(t)) = \prod_{j=1}^p M_{ij}(z_j(t)), \quad h_i(z(t)) = \frac{w_i(z(t))}{\sum_{i=1}^r w_i(z(t))}.$$

Normally, fuzzy sets M_{ij} are normalised, which implies that weights $w_i(z(t))$ and MFs $h_i(z(t))$ *always* hold the convex-sum property $\sum_{i=1}^r h_i(z(t)) = 1$, $0 \leq h_i(z(t)) \leq 1$, regardless of the state $x(t)$ or input $u(t)$ values.

Fuzzy TS systems of the sort just described can be subsumed in the broader class of *quasi linear parameter varying* (qLPV) systems as, in addition to possible uncertainties θ , the MFs may depend also on states $x(t)$ (Apkarian and Gahinet

1995). These MFs are viewed then as interpolating functions among linear systems $\dot{x}(t) = A_i x(t) + B_i u(t)$, $i \in \{1, 2, \dots, r\}$. This leads to their interpretation as *polytopic linear differential inclusions* (PLDIs) as the state-, time-, or parameter-nature of the MFs might be ignored by focusing on the MFs convex-sum property. Thus, a TS model can be subsumed in the class of PLDI as they are polytopes of linear systems due to the convexity of the MFs whose dependency is neglected (Boyd et al. 1994).

This introduction was intended to trace back the origins of the TS systems we are about to study in greater detail as well as to clarify common misunderstandings regarding “close-friend” classes of systems, but in the sequel we will not refer to *fuzzy* TS models as our approach is not model-free nor fuzzy. We are interested in TS models as means for analysis and synthesis of nonlinear control systems whose model is *a priori known*. To be useful for this purpose, a TS model must be an *exact convex rewriting* of a nonlinear one, at least in some region. Under mild assumptions, we are about to discover this is possible.

2.1.2 Exact convex modelling

Consider a function $z : \mathbb{R}^n \rightarrow \mathbb{R}$ such that for $x \in \Omega \subset \mathbb{R}^n$, Ω being a compact set, there exist values z^1 and z^0 such that $z(x) \in [z^0, z^1]$, i.e., $z(x)$ is bounded in Ω . In this case, $z(x)$ can be written as a convex sum of its bounds as follows:

$$z(x) = \underbrace{\frac{z^1 - z(x)}{z^1 - z^0}}_{w_0(x)} z^0 + \underbrace{\frac{z(x) - z^0}{z^1 - z^0}}_{w_1(x)} z^1, \quad (2.1)$$

where $w_0(x)$ and $w_1(x)$ hold the convex-sum property $w_0(x) + w_1(x) = 1$, $0 \leq w_0(x), w_1(x) \leq 1$.

Convex sums hold interesting properties when they appear in nested expressions, for instance, given two bounded nonlinearities $z_1(x) \in [z_1^0, z_1^1]$ and $z_2(x) \in [z_2^0, z_2^1]$ modelled as in (2.1) via functions $w_0^1(x)$, $w_1^1(x)$ and $w_0^2(x)$, $w_1^2(x)$, respectively, we have:

$$z_1(x)z_2(x) = \left(\sum_{i=0}^1 w_i^1(x) z_1^i \right) \left(\sum_{j=0}^1 w_j^2(x) z_2^j \right) = \sum_{i=0}^1 \sum_{j=0}^1 w_i^1(x) w_j^2(x) z_1^i z_2^j \quad (2.2)$$

$$z_1(x) + z_2(x) = \sum_{i=0}^1 w_i^1(x) z_1^i + \sum_{i=0}^1 w_i^2(x) z_2^i = \sum_{i=0}^1 \sum_{j=0}^1 w_i^1(x) w_j^2(x) (z_1^i + z_2^j) \quad (2.3)$$

With these tools at hand, consider an affine in-control continuous-time nonlinear system of the form

$$\dot{x}(t) = A(x)x(t) + B(x)u(t), \quad (2.4)$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the input vector, $A(\cdot)$ and $B(\cdot)$ are smooth matrix possibly nonlinear functions of appropriate dimensions. Assume there are p different non-constant terms $z_i(x)$, $i \in \{1, 2, \dots, p\}$, in $A(x)$ and $B(x)$ which are bounded in a compact set $\Omega \subset \mathbb{R}^n$ such that $0 \in \Omega$; they will be the entries of the so-called premise vector $z(x) \in \mathbb{R}^p$.

Let $z_j(x) \in [z_j^0, z_j^1]$, $j \in \{1, 2, \dots, p\}$ be the set of bounded non-constant terms in $A(x)$ and $B(x)$ belonging to Ω . Similarly to (2.1), each of these terms can be written as a convex sum of its bounds, i.e., $z_j(x) = w_0^j(x)z_j^0 + w_1^j(x)z_j^1$ with $w_0^j(x)$, $w_1^j(x)$, $j \in \{1, 2, \dots, p\}$, being weighting functions (WFs) of the form:

$$w_0^j(z_j) = \frac{z_j^1 - z_j(x)}{z_j^1 - z_j^0}, \quad w_1^j(z_j) = 1 - w_0^j(z_j), \quad j \in \{1, 2, \dots, p\}. \quad (2.5)$$

Due to properties (2.2) and (2.3), (2.4) can be exactly rewritten as the following *tensor-product Takagi-Sugeno model*:

$$\dot{x}(t) = \sum_{i_1=0}^1 \sum_{i_2=0}^1 \cdots \sum_{i_p=0}^1 w_{i_1}^1 w_{i_2}^2 \cdots w_{i_p}^p (A_{(i_1, i_2, \dots, i_p)}x(t) + B_{(i_1, i_2, \dots, i_p)}u(t)) \quad (2.6)$$

$$= \sum_{\mathbf{i} \in \mathbb{B}^p} \mathbf{w}_{\mathbf{i}} (A_{\mathbf{i}}x(t) + B_{\mathbf{i}}u(t)), \quad (2.7)$$

where $\mathbf{i} = (i_1, i_2, \dots, i_p)$, $\mathbb{B} \in \{0, 1\}$, $\mathbf{w}_{\mathbf{i}} = w_{i_1}^1 w_{i_2}^2 \cdots w_{i_p}^p$, $A_{\mathbf{i}} = A(x)|_{\mathbf{w}_{\mathbf{i}}=1}$, $B_{\mathbf{i}} = B(x)|_{\mathbf{w}_{\mathbf{i}}=1}$, $\mathbf{1} = \underbrace{(1, 1, \dots, 1)}_{p \text{ ones}}$.

Adding the following definition of *membership functions* (MFs):

$$h_i = h_{1+i_1+i_2 \times 2 + \dots + i_p \times 2^{p-1}} = \prod_{j=1}^p w_{i_j}^j(z_j), \quad (2.8)$$

with $i \in \{1, 2, \dots, r\}$, $r = 2^p$, $i_j \in \{0, 1\}$, which, as WF's, hold the convex-sum property in Ω :

$$\sum_{i=1}^r h_i(\cdot) = 1, \quad h_i(\cdot) \geq 0, \quad i \in \{1, 2, \dots, r\}, \quad (2.9)$$

another exact representation of (2.4) in Ω is given by the following *classical Takagi-Sugeno (TS) model*:

$$\dot{x}(t) = \sum_{i=1}^r h_i(z(x)) (A_i x(t) + B_i u(t)) = A_h x(t) + B_h u(t), \quad (2.10)$$

with $(A_i, B_i) = (A(x), B(x))|_{h_i=1}$, $i \in \{1, 2, \dots, r\}$. Notice the similarities and differences of this model with respect to the fuzzy one in the previous section: the structure is the same, yet dependency of the MFs and the model origin is entirely different. Fuzzy systems arise from linguistic descriptions or linearisation, they are model-free approximations; TS systems as (2.6) and (2.10) are model-based and exact, i.e., they are *not* approximations.

More generally, exact TS models may also be used to rewrite in a convex way any *bounded* expression, whether it depends on the state or not, for instance, parametric uncertainties. If we know that certain parameter θ belongs to the interval $[\theta^0, \theta^1]$ we can always write

$$\theta = \underbrace{\frac{\theta^1 - \theta}{\theta^1 - \theta^0}}_{w_0(\theta)} \theta^0 + \underbrace{\frac{\theta - \theta^0}{\theta^1 - \theta^0}}_{w_1(\theta)} \theta^1,$$

where $w_0(\theta)$ and $w_1(\theta)$ hold the convex-sum property $w_0(\theta) + w_1(\theta) = 1$, $0 \leq w_0(\theta), w_1(\theta) \leq 1$. Similarly, any function of several variables –states or uncertainties– can be written as a convex sum as long as it is bounded.

At first sight, this may seem as an advantageous flexibility: there are many ways to obtain a TS model with a polytopic structure as those above. Nevertheless, we will see that analysis and synthesis of nonlinear control systems based on these convex structures depends critically on the vertex choice of the polytope, i.e., the way nonlinearities are chosen to obtain matrices A_i , B_i in (2.6) or A_i , B_i in (2.10). Moreover, we will see that these structures induce straightforward generalisations of linear state feedback control laws that employ the WFs or the MFs of the controlled model. Clearly, using these functions implies that they are available, which excludes uncertainties (i.e., they cannot be fed back to the system as they are unknown) and may depend on the state availability (if this state is not available, observers may come at hand). Nevertheless, combinations of “known” WFs associated to measured quantities with “unknown” ones (giving rise to unmeasured polytopic uncertainty) can also be handled, see for instance (Kruszewski 2006).

Example 2.1. Consider the following 2nd-order nonlinear system with parametric uncertainty $\theta \in [-2, 2]$ and state $x(t) = [x_1 \ x_2]^T$:

$$\dot{x}_1(t) = x_1^2 x_2 \theta + \sin x_2, \quad (2.11)$$

$$\dot{x}_2(t) = x_1^3 + (x_2 + \theta) u(t), \quad (2.12)$$

with scalar states $x_1(t)$, $x_2(t)$ and input $u(t)$. Suppose we are interested in controlling this system within $\Omega = \{x : -1 \leq x_i \leq 1\}$. Then, defining the premise variables $z_1 = x_1$, $z_2 = x_2$, $z_3 = \theta$, $z_4 = \sin x_2/x_2$, whose extreme values (given or deduced) are listed in table 2.1, the following WFs can be defined:

$$w_0^i(z_i) = \frac{z_i^1 - z_i}{z_i^1 - z_i^0}, w_1^i(z_i) = 1 - w_0^i(z_i), i \in \{1, 2, 3, 4\}. \quad (2.13)$$

Equations (2.11)-(2.12) are thus equivalent to:

$$\begin{aligned} \dot{x}(t) = & \underbrace{\begin{bmatrix} x_1 x_2 \theta & \sin x_2/x_2 \\ x_1^2 & 0 \end{bmatrix}}_{A(x,\theta)} x(t) + \underbrace{\begin{bmatrix} 0 \\ x_2 + \theta \end{bmatrix}}_{B(x,\theta)} u(t) = \sum_{i_1=0}^1 \sum_{i_2=0}^1 \sum_{i_3=0}^1 \sum_{i_4=0}^1 \sum_{i_5=0}^1 w_{i_1}^1(x_1) \\ & \times w_{i_2}^2(x_2) w_{i_3}^3(\theta) w_{i_4}^4(x_2) w_{i_5}^1(x_1) \left(\underbrace{\begin{bmatrix} z_1^{i_1} z_2^{i_2} z_3^{i_3} & z_4^{i_4} \\ z_1^{i_1} z_1^{i_5} & 0 \end{bmatrix}}_{A_i} x(t) + \underbrace{\begin{bmatrix} 0 \\ z_2^{i_2} + z_3^{i_3} \end{bmatrix}}_{B_i} u(t) \right), \end{aligned}$$

with $\mathbf{i} = (i_1, i_2, i_3, i_4, i_5)$, which corresponds to a tensor-product TS model of the form (2.6). Note that a fifth sum is given using the same WF of the first one, $w_j^1(x)$, because x_1^2 is interpreted as the product of two identical convex sums, each of them representing x_1 .

Obviously, both the decoupling $A(x, \theta)x$ and the way nonlinearities have been chosen afterwards could have been differently done. Consider, for instance, a decou-

Table 2.1: Nonlinearities in (2.11)-(2.12) and their bounds

Nonlinearity	Definition	Lower bounds z_i^0	Upper bounds z_i^1
z_1	x_1	-1	1
z_2	x_2	-1	1
z_3	θ	-2	2
z_4	$\sin x_2/x_2$	0.842	1

pling of the form:

$$A(x, \theta) = \begin{bmatrix} 0 & x_1^2 \theta + \sin x_2 / x_2 \\ x_1^2 & 0 \end{bmatrix},$$

with the choice of nonlinearities $z_1 = x_1^2$, $z_2 = \sin x_2 / x_2$, $z_3 = \theta$, $z_4 = x_2$. Such options will produce 2^4 instead of 2^5 different pairs (A_i, B_i) , something which is highly appreciated in the literature as a lower number of vertex models directly reflects in fewer LMIs. By now the reader may have noticed that the premise variables may be chosen as depending on more than one variable or parameter, for instance, for the same decoupling we can choose $z_1 = x_1^2 \theta + \sin x_2 / x_2$, $z_2 = x_1^2$, and $z_3 = x_2 + \theta$, as to obtain 2^3 different pairs (A_i, B_i) , the bounds will be of course the combined minimum and combined maximum resulting from the minima/maxima of the states and the uncertainty.

The tensor-product TS model above has WFs that depend exclusively on states and another one that depends only on the uncertainty. Such selection is convenient for control purposes as will be discussed in the next section, because only known states can be fed back for control purposes as uncertainties are bounded, but unknown. This distinction will be lost if an ordinary TS model with a single set of MFs (2.8) is constructed, for instance, via the WFs in (2.13), i.e. (arguments omitted for WFs):

$$h_1(x, \theta) = w_0^1 w_0^2 w_0^3 w_0^4 w_0^1, h_2(x, \theta) = w_0^1 w_0^2 w_0^3 w_0^4 w_1^1, \dots, h_{32}(x, \theta) = w_1^1 w_1^2 w_1^3 w_1^4 w_1^1,$$

leading to

$$\dot{x}(t) = \sum_{i=1}^{32} h_i(x, \theta) (A_i x(t) + B_i u(t)),$$

with $A_i = A(x, \theta)|_{h_i=1}$, $B_i = B(x, \theta)|_{h_i=1}$.

Once an exact convex model is at hand, analysis and synthesis via LMIs is at reach since the inequalities involved in Lyapunov-based analysis translate into LMIs over the vertexes of the TS systems. This is the subject of the next section.

2.2 Nonlinear control

2.2.1 Lyapunov stability

Consider a nonlinear system of the form

$$\dot{x}(t) = f(x(t)), x(0) = x_0, \quad (2.14)$$

with $x(t) \in \mathbb{R}^n$ as the system state, $f(\cdot) : \Omega \rightarrow \mathbb{R}^n$ a locally Lipschitz function from a domain $\Omega \subseteq \mathbb{R}^n$ into \mathbb{R}^n , $\psi(t, x_0)$ the solution of (2.14) from initial condition x_0 (which is assumed to be unique). Recall that an *isolated equilibrium point* \bar{x} is a state value such that $\exists r > 0 : \forall x \in \{x : |x| \leq r, x \neq 0\} \Rightarrow f(x) \neq 0$, and $f(\bar{x}) = 0$.

In contrast with linear systems, nonlinear ones may present multiple equilibrium points as well as a variety of phenomena such as limit cycles, finite-time escape, chaos, etc. Without loss of generality, in the sequel we assume that the equilibrium point under analysis is at the origin, i.e., $\bar{x} = 0$. Stability of isolated equilibrium points of nonlinear systems is studied via Lyapunov stability theory, first appeared in “The General Problem of Stability of Motion” (1892) from the Russian mathematician Aleksandr Lyapunov. It consists in two methods: an “indirect” one based on the eigenvalues of the linearised system and a “direct” one based on the existence of a Lyapunov function, i.e., an energy-like function of the states that monotonically decreases over time for stable trajectories.

Some types of stability of equilibrium points are now defined:

Definition 2.1. (Haddad and Chellaboina 2008)

- The equilibrium point \bar{x} is said to be Lyapunov stable, if, for every $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon)$ such that, if $\|x(0) - \bar{x}\| < \delta$, then for every $t \geq 0$ we have $\|x(t) - \bar{x}\| < \varepsilon$.
- The equilibrium point \bar{x} is said to be asymptotically stable if it is Lyapunov stable and there exists $\delta > 0$ such that if $\|x(0) - \bar{x}\| < \delta$, then $\lim_{t \rightarrow \infty} \|x(t) - \bar{x}\| = 0$.
- The equilibrium point \bar{x} is said to be exponentially stable if it is asymptotically stable and there exist $\alpha > 0$, $\beta > 0$, $\delta > 0$ such that if $\|x(0) - \bar{x}\| < \delta$, then $\|x(t) - \bar{x}\| < \alpha \|x(0) - \bar{x}\| e^{-\beta t}$, for $t \geq 0$.
- An equilibrium point \bar{x} is unstable if it is not Lyapunov stable.

The indirect Lyapunov method is now stated:

Theorem 2.1. (Khalil 2002) Consider a nonlinear system (2.14) with Ω being a neighborhood of the isolated equilibrium point $x = 0$. Given the Jacobian matrix of $f(x)$ at $x = 0$

$$A = \left. \frac{\partial f(x)}{\partial x} \right|_{x=0}, \quad \sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_n\},$$

the origin is asymptotically stable if $\operatorname{Re}(\lambda_i) < 0$ for all eigenvalues of A , or unstable if $\operatorname{Re}(\lambda_i) > 0$ for one or more of the eigenvalues of A .

Despite its simplicity, this theorem does not exhaust the different possibilities of location of eigenvalue of A ; it may therefore be useless if a system does not fall in the aforementioned categories. Nevertheless, when it succeeds, it implies the existence of a quadratic Lyapunov function $V(x) = x^T P x$ with $P > 0$ such that $PA + A^T P < 0$, which is also a Lyapunov function of the nonlinear system in some neighborhood of the origin.

Naturally, the direct Lyapunov method is far more general than the previous one. It is not only sufficient and necessary for stability of an equilibrium point $x = 0$, but it helps us determining the *domain of attraction*, which is the region $\mathcal{D} := \{x \in \mathbb{R}^n : \psi(t, x) \in \Omega \forall t \geq 0, \lim_{t \rightarrow \infty} \psi(t, x) = 0\}$. Such estimation provides the designer of a control system with a guaranteed domain (initial conditions) for the desired behaviour.

The direct Lyapunov method is now stated:

Theorem 2.2. (Lyapunov 1992) Consider a nonlinear system (2.14) with Ω being a neighborhood of the isolated equilibrium point $x = 0$. Let $V : \Omega \rightarrow \mathbb{R}$ be a continuously differentiable function in Ω such that the following conditions are fulfilled:

$$V(0) = 0 \tag{2.15}$$

$$V(x) > 0 \quad \forall x \in \Omega, x \neq 0 \tag{2.16}$$

$$\dot{V}(x) = \frac{dV(x)}{dt} < 0 \quad \forall x \in \Omega, x \neq 0 \tag{2.17}$$

then the origin is asymptotically stable in the sense of Lyapunov. If $\Omega \equiv \mathbb{R}^n$ and $V(x)$ being radially unbounded, i.e., $\|x\| \rightarrow \infty \Rightarrow V(x) \rightarrow \infty$, then the origin is globally asymptotically stable.

The existence of a Lyapunov function is a sufficient condition for the stability of an equilibrium point; conversely, for every stable equilibrium point there must exist a Lyapunov function (Hahn 1967). Despite its power and generality, this result has a major drawback: there is no general methodology for searching Lyapunov functions for nonlinear systems. Some forms, such as the quadratic one, have been used for simplicity because they work fine in the linear case.

Theorem 2.2 helps providing an estimate of the DA of the origin: if there exist a Lyapunov function $V(x)$ that satisfies the conditions of asymptotic stability over

a domain Ω , then every trajectory starting in $\mathcal{E}_c := \{x \in \mathbb{R}^n : V(x) \leq c\}$, $\mathcal{E}_c \subset \Omega$, remains in \mathcal{E}_c and approaches the origin as $t \rightarrow \infty$. Therefore, \mathcal{E}_c is an estimate of the DA, i.e. $\mathcal{E}_c \subset \mathcal{D}$.

In the case of linear time-invariant (LTI) systems, the existence of a quadratic Lyapunov function $V(x) = x^T P x$ is a sufficient and necessary condition for global asymptotic stability of $\dot{x} = Ax$ leading to the LMI conditions presented in the previous chapter when linear systems were analysed. In particular, to apply Theorem 2.2 to $\dot{x} = Ax$, consider the Lyapunov function candidate $V(x) = x^T P x$, where $P = P^T > 0$ to satisfy condition (2.16). The time derivative of $V(x)$ is given by:

$$\dot{V}(x) = x^T P \dot{x} + \dot{x}^T P x = x^T (PA + A^T P) x. \quad (2.18)$$

Now, (2.17) is guaranteed if and only if $PA + A^T P < 0$ as it coincides with the definition of a negative-definite matrix. The inequalities $P > 0$ and $PA + A^T P < 0$ are linear matrix expressions; determining whether or not there is an instance of P such that the inequalities hold is an LMI problem.

2.2.2 The quadratic framework

From the previous chapter we already know the remarkable soundness of LMIs within the context of linear control systems. This owns a great deal to the fact that quadratic stability usually implies sufficiency and necessity of LMI conditions. This of course is no longer the case for nonlinear systems; yet, sufficiency of LMI conditions and much of their soundness can be preserved with the aid of TS systems and the convex structure built in them.

Consider the following quadratic Lyapunov function candidate

$$V(x) = x^T(t) P x(t), \quad P = P^T > 0 \quad (2.19)$$

along with a continuous-time nonlinear system (2.4) with $u(t) = 0$, i.e., $\dot{x}(t) = A(x)x(t)$, $x \in \mathbb{R}^n$. By the sector nonlinearity approach presented in section 2.1.2, the following TS model of the system within a compact set of the state space Ω containing the origin, is obtained; it is similar to (2.10) with $u(t) = 0$:

$$\dot{x}(t) = \sum_{i=1}^r h_i(z(x)) A_i x(t) = A_h x(t), \quad (2.20)$$

where $A_i \in \mathbb{R}^{n \times n}$ and h_i , $i \in \{1, 2, \dots, r\}$ are the MFs holding the convex-sum property in Ω . The following result establishes an LMI test for asymptotic stability of the origin $x = 0$ of $\dot{x}(t) = A(x)x(t)$ via the TS model (2.20) within Ω :

Theorem 2.3. (Tanaka and Wang 2001) *The origin $x = 0$ of the autonomous model $\dot{x}(t) = A(x)x(t)$ with TS equivalent (2.20) is asymptotically stable if there exists a matrix $P = P^T > 0$ such that $PA_i + A_i^T P < 0$ for $i \in \{1, 2, \dots, r\}$.*

Proof. According to the direct Lyapunov method in (2.2), the origin of the system under analysis is asymptotically stable if there exists a function $V(x)$ such that $V(x) > 0$, $x \neq 0$ (a condition fulfilled by $V(x) = x^T P x$, $P = P^T > 0$) and $\dot{V}(x) < 0 \forall x \in \Omega$, $x \neq 0$, $0 \in \Omega$, a condition guaranteed by $PA_i + A_i^T P < 0$ for $i \in \{1, 2, \dots, r\}$, since $\sum_{i=1}^r h_i(\cdot) = 1$, $0 \leq h_i \leq 1$ and:

$$\begin{aligned} \dot{V}(t) &= \dot{x}^T(t) P x(t) + x^T(t) P \dot{x}(t) \\ &= \left(\sum_{i=1}^r h_i(z) A_i x(t) \right)^T P x(t) + x^T(t) P x(t) \left(\sum_{i=1}^r h_i(z) A_i x(t) \right) \\ &= \sum_{i=1}^r h_i(z) x^T(t) (P A_i + A_i^T P) x(t), \end{aligned}$$

which concludes the proof. \square

Since the convex sum property holds only in the compact set Ω , any trajectory starting in the outermost Lyapunov level $V(x) = x(t)^T P x(t) = k$, $k \in \mathbb{R}$ within Ω goes to zero. Note that if $\Omega = \mathbb{R}^n$, i.e., if the convex sum property of the MFs hold everywhere, the origin is globally asymptotically stable; this is the case of the TS models in the fuzzy context (Tanaka and Wang 2001). Moreover, should the model (2.6) be used instead of (2.10), with $u(t) = 0$, similar LMI results will follow, namely: $P = P^T > 0$, $PA_i + A_i^T P < 0$, $i \in \mathbb{B}^p$.

Example 2.2. *Consider investigating the stability of the origin of the following continuous-time nonlinear model:*

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 + x_1^2 & -2 - \sin x_2/x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (2.21)$$

Notice that this model has 3 isolated equilibrium points at $(0, 0)$, $(\pm\sqrt{3}, 0)$. Therefore, the origin cannot be globally asymptotically stable.

Consider a TS model of the form (2.10) within the square $\Omega = \{x : |x_i| \leq \bar{x}\}$, with \bar{x} to be defined later; the premise variables $z_1(x) = x_1^2 \in [0, \bar{x}^2]$, $z_2(x) = \sin x_2/x_2 \in [-0.2173, 1]$:

$$\dot{x}(t) = \sum_{i=1}^4 h_i(z(x)) A_i x(t), \quad (2.22)$$

with $A_1 = \begin{bmatrix} 0 & 1 \\ -3 & -1.7827 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ -3 & -3 \end{bmatrix}$, $A_3 = \begin{bmatrix} 0 & 1 \\ -3 + \bar{x}^2 & -1.7827 \end{bmatrix}$, $A_4 = \begin{bmatrix} 0 & 1 \\ -3 + \bar{x}^2 & -3 \end{bmatrix}$, $w_0^1 = (\bar{x}^2 - x_1^2)/\bar{x}^2$, $w_1^1 = 1 - w_0^1$, $w_0^2 = (1 - \sin x_2/x_2)/1.2173$, $w_1^2 = 1 - w_0^2$, $h_1(z(x)) = w_0^1 w_0^2$, $h_2(z(x)) = w_0^1 w_1^2$, $h_3(z(x)) = w_1^1 w_0^2$, and $h_4(z(x)) = w_1^1 w_1^2$.

The following MATLAB code uses *SeDuMi* to search $P > 0$ such that $PA_i + A_i^T P < 0$ for $i \in \{1, 2, 3, 4\}$ when $\bar{x} = 1$:

```

barx=1;
A{1}=[0 1; -3 -1.7827]; A{2}=[0 1; -3 -3];
A{3}=[0 1; -3+barx^2 -1.7827]; A{4}=[0 1; -3+barx^2 -3];
P=sdpvar(2); LMI=[];
5 LMI=[LMI,P>=(1e-6)*eye(2)];
for i=1:4
    LMI=[LMI,P*A{i}+(P*A{i})'<=-(1e-6)*eye(2)];
end
ops=sdpsettings('solver','sedumi','verbose',0,
10 'warning',0,'shift',1e-5);
SOL=solvesdp(LMI,[],ops);
if SOL.problem==0
    P=double(P)
end

```

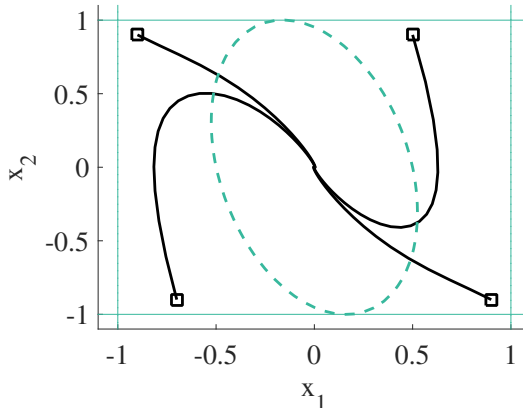


Figure 2.1: Phase portrait.

The LMIs above are feasible with $P = \begin{bmatrix} 1.1805 & 0.1898 \\ 0.1898 & 0.3249 \end{bmatrix}$; this means that every trajectory within the biggest Lyapunov set $V(x) = x^T P x \leq k$ that fits in $\Omega = \{x : |x_i| \leq 1\}$ goes asymptotically to zero. This is confirmed in Fig. 1.1, which shows that the system was indeed stabilised (solid lines from initial conditions marked with squares). The outermost Lyapunov level is the ellipsoid in dashed lines. Obviously, since the nonlinear system has three equilibrium points, \bar{x} cannot go too far before rendering the LMIs above infeasible, which is the case, for instance, if $\bar{x} = 3$.

Deriving LMI conditions for stabilisation of a nonlinear system through an exact TS model (2.10) and quadratic Lyapunov functions is almost as straightforward as in the linear case. Assuming the whole state is available, linear state feedback can be generalised to *parallel distributed compensation* (PDC), first appeared in (Sugeno and Kang 1988), to exploit the information gathered in the system MFs:

$$u(t) = \sum_{j=1}^r h_j(z(x)) F_j x(t) = F_h x(t), \quad (2.23)$$

where $F_j \in \mathbb{R}^{m \times n}$ are gains of adequate size to be determined. Notice this is no longer linear state feedback as nonlinearities enter the control law through the MFs. Similarly, if a tensor-product TS model is employed, the PDC control law has the form $u(t) = \sum_{j \in \mathbb{B}^p} \mathbf{w}_j(z(x)) F_j x(t) = F_{\mathbf{w}} x(t)$, where the set of gains is F_j , $j \in \mathbf{B}^p$.

Substituting the control law (2.23) in the TS model (2.10) gives the following closed-loop system:

$$\begin{aligned} \dot{x}(t) &= \sum_{i=1}^r h_i(z(x)) A_i x(t) + \left(\sum_{i=1}^r h_i(z(x)) B_i \right) \left(\sum_{j=1}^r h_j(z(x)) F_j x(t) \right) \\ &= \sum_{i=1}^r \sum_{j=1}^r h_i(z(x)) h_j(z(x)) (A_i + B_i F_j) x(t) = (A_h + B_h F_h) x(t), \end{aligned} \quad (2.24)$$

where, again, the fact that $\sum_{i=1}^r h_i(\cdot) = 1$ has been taken into account.

Though conditions in the next theorem are not LMIs, they serve as the departure point for introducing the important subject of *sum relaxations* that intend to solve through the less conservative set of LMIs the *co-positivity problem* (Murty and Kabadi 1987) therein stated:

Theorem 2.4. *The origin $x = 0$ of the nonlinear system $\dot{x}(t) = A(x)x(t) + B(x)u(t)$ with TS model (2.10) in Ω , $0 \in \Omega$, under the control law (2.23), is asymptotically stable if $\exists X = X^T > 0$ and M_j , $j \in \{1, 2, \dots, r\}$, such that*

$$\Gamma_{hh} = A_h X + B_h M_h + X A_h^T + M_h^T B_h^T < 0, \quad (2.25)$$

with $M_h = \sum_{j=1}^r h_j(z(x))M_j$. Thus, the controller gains are $F_h = M_h X^{-1}$ (equivalently, $\bar{F}_j = M_j X^{-1}$, $j \in \{1, 2, \dots, r\}$) whereas the Lyapunov function is given by $V(x) = x^T P x$, $P = X^{-1}$. Moreover, any trajectory starting in the outermost Lyapunov level $V(x) \leq k$ inside Ω goes asymptotically to zero.

Proof. Applying the property 1.7 in (2.25) pre- and post-multiplying by $P = X^{-1}$ leads to:

$$P A_h + P B_h M_h P + A_h^T P + P M_h^T B_h^T P < 0,$$

which, defining $M_h = F_h X$, leads to

$$P A_h + P B_h F_h + A_h^T P + F_h^T B_h^T P = P(A_h + B_h F_h) + (A_h + B_h F_h)^T P < 0,$$

which, by theorem 2.3 implies that the closed-loop system (2.24) has its origin $x = 0$ asymptotically stable with $V(x) = x^T P x$, $P = P^T > 0$, as the associated Lyapunov function and the last inequality above the expression of the fact that $\dot{V}(x) < 0 \forall x \in \Omega$, $x \neq 0$. \square

As mentioned before, condition (2.25) is not an LMI, but a double convex sum (dependency of the MFs h_i on $z(x)$ is omitted for brevity):

$$\begin{aligned} \Gamma_{hh} &= A_h X + B_h M_h + X A_h^T + M_h^T B_h^T \\ &= \left(\sum_{i=1}^r h_i A_i \right) X + \left(\sum_{i=1}^r h_i B_i \right) \left(\sum_{j=1}^r h_j M_j \right) + X \left(\sum_{i=1}^r h_i A_i \right)^T + \left(\sum_{j=1}^r h_j M_j \right)^T \left(\sum_{i=1}^r h_i B_i \right)^T \\ &= \sum_{i=1}^r \sum_{j=1}^r h_i h_j \left(\underbrace{A_i X + B_i M_j + X A_i^T + M_j^T B_i^T}_{\Gamma_{ij}} \right). \end{aligned}$$

Clearly, guaranteeing the double convex sum above is negative-definite can be done via LMIs $A_i X + B_i M_j + X A_i^T + M_j^T B_i^T < 0$, $i, j \in \{1, 2, \dots, r\}$, but this is a very clumsy way to do so as, though sufficient, is clearly not necessary. Relaxing this sufficiency is the purpose of the so-called *sum relaxations*. A very obvious one is that in (Tanaka and Sano 1994), which takes into account the fact that $h_i h_j = h_j h_i$:

LEMMA 2.1. *The double convex-sum $\Gamma_{hh} < 0$ is verified if*

$$\begin{aligned} \Gamma_{ii} &< 0, \quad \forall i \in \{1, 2, \dots, r\}, \\ \Gamma_{ij} + \Gamma_{ji} &< 0, \quad \forall (i, j) \in \{1, 2, \dots, r\}^2, \quad i < j. \end{aligned} \quad (2.26)$$

Another common relaxation is that in (Tuan et al. 2001); as the preceding one, it does not use slack matrices (Kim and Lee 2000):

LEMMA 2.2. *The double convex-sum $\Gamma_{hh} < 0$ is verified if*

$$\begin{aligned} \Gamma_{ii} &< 0, \quad \forall i \in \{1, 2, \dots, r\}, \\ \frac{2}{r-1}\Gamma_{ii} + \Gamma_{ij} + \Gamma_{ji} &< 0, \quad \forall (i, j) \in \{1, 2, \dots, r\}^2, \quad i \neq j. \end{aligned} \quad (2.27)$$

The problem of co-positivity has been solved twice with asymptotically sufficient and necessary conditions via the Polya Theorem in (Sala and Ariño 2007) and via triangulation in (Kruszewski et al. 2009); these solutions are not presented here because they will not be used in this thesis as they rapidly exhaust the computational resources of the LMI solvers due to the enormous number of LMIs which emerge close to necessity.

By now, the reader may correctly deduce that the performance measures such as decay rate, cost functions, and H_∞ attenuation level, already defined in the context of linear systems, are easily extrapolated to the nonlinear case with the aid of convexity of TS representations (2.10) or (2.6). Due to the “modularity” of LMI results, such performance specifications can be stacked without further adaptation in order to get a controller holding any subset of them.

Example 2.3. *Consider the Furuta pendulum shown in Fig.2.2. The control task consists in keeping the free beam at the vertical position through the induced*

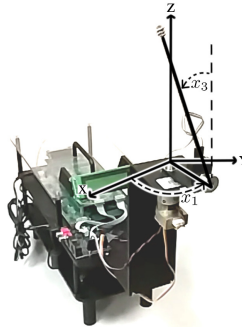


Figure 2.2: Furuta pendulum [source: www.quanser.com]

rotations of the horizontal beam, which is driven at the center by a fixed DC motor. A state-space representation of this plant can be found in (Quanser 2006):

$$\begin{aligned}
 \dot{x}_1 &= x_2 \\
 \dot{x}_2 &= \frac{(\beta + \gamma) (\delta x_4^2 \sin x_3 - 2\beta x_2 x_4 \cos x_3 \sin x_3 + u)}{((\beta + \gamma) \beta + \delta^2) \sin^2 x_3 + (\beta + \gamma) \alpha - \delta^2} \\
 &\quad - \frac{\delta \cos x_3 (\beta x_2^2 \cos x_3 \sin x_3 + \sigma g \sin x_3)}{((\beta + \gamma) \beta + \delta^2) \sin^2 x_3 + (\beta + \gamma) \alpha - \delta^2} \\
 \dot{x}_3 &= x_4 \\
 \dot{x}_4 &= \frac{(\beta \sin^2 x_3 + \alpha) (\beta x_2^2 \cos x_3 \sin x_3 + \sigma g \sin x_3)}{((\beta + \gamma) \beta + \delta^2) \sin^2 x_3 + (\beta + \gamma) \alpha - \delta^2} \\
 &\quad - \frac{\delta \cos x_3 (\delta x_4^2 \sin x_3 - 2\beta x_2 x_4 \cos x_3 \sin x_3 + u)}{((\beta + \gamma) \beta + \delta^2) \sin^2 x_3 + (\beta + \gamma) \alpha - \delta^2}
 \end{aligned} \tag{2.28}$$

with x_1 being the angle of the horizontal beam with respect to a fixed position, x_2 its angular velocity, x_3 the angle between the vertical beam and the vertical upright axis, and x_4 the corresponding angular velocity. We intend to keep the second beam in its upright position, which coincides with driving the system to $x_3 = 0$. The parameters are given by: $\alpha = (J_0 + m_1 L_0^2) / T_c$, $\beta = (m_1 l_1^2) / T_c$, $\gamma = J_1 / T_c$, $\delta = (m_1 L_0 l_1) / T_c$, $\sigma = (m_1 l_1) / T_c$, $L_0 = 0.068m$ and $J_0 = 6.9885e^{-5} kg \cdot m^2$ are length and total moment of inertia of the horizontal link, respectively, $m_1 = 0.02366kg$, $l_1 = 0.08m$, and $J_1 = 1.7590e^{-4} kg \cdot m^2$ are the mass, center of mass, and total moment of inertia of the vertical beam, respectively, $g = 9.81m/s^2$ and $T_c = 0.0049431 N \cdot m/V$ are the gravitational constant and the torque constant, respectively.

Obtaining a TS model of (2.28) requires rewriting every nonlinearity as a convex sum. Consider the choice $z_1 = (\beta + \gamma) \delta x_4^2 - 2\beta (\beta + \gamma) x_2 x_4 \cos x_3 - \beta \delta x_2^2 \cos^2 x_3 - \delta \sigma g \cos x_3$, $z_2 = (\sin x_3) / x_3$, $z_3 = (\beta \sin^2 x_3 + \alpha) (\beta x_2^2 \cos x_3 + \sigma g) - \delta \cos x_3 (\delta x_4^2 - 2\beta x_2 x_4 \cos x_3)$, $z_4 = \cos x_3$, and $z_5 = 1 / ((\beta^2 + \gamma \beta + \delta^2) \sin^2 x_3 + \alpha \beta + \alpha \gamma - \delta^2)$, then, (2.28) is equivalent to:

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & z_1 z_2 z_5 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & z_2 z_3 z_5 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ (\beta + \gamma) z_5 \\ 0 \\ -\delta z_4 z_5 \end{bmatrix} u. \tag{2.29}$$

Consider the operational regime given by the compact $|x_2| \leq 10$ rad/s, $|x_3| \leq 15$ (in degrees), and $|x_4| \leq 3$ rad/s (x_1 is not at the right-hand side of the model).

Thus, the variables in (2.29) can also be bounded as $z_1 \in [-0.2833, -0.0324]$, $z_2 \in [0.9886, 1.0000]$, $z_3 \in [0.1251, 0.2961]$, $z_4 \in [0.9659, 1.0000]$, and $z_5 \in [524.0978, 579.0187]$. Writing each of $z_i \in [z_i^0, z_i^1]$ as a convex sum of its bounds, we have the following identities:

$$z_i = \underbrace{\frac{z_i^1 - z_i}{z_i^1 - z_i^0}}_{w_0^i} (z_i^0) + \underbrace{\frac{z_i - z_i^0}{z_i^1 - z_i^0}}_{w_1^i} (z_i^1), \quad \begin{aligned} w_0^i + w_1^i &= 1, \\ 0 \leq w_j^i &\leq 1, \end{aligned} \quad (2.30)$$

based on which the following TS model in the form (2.10) arises:

$$\dot{x} = \sum_{i_1=0}^1 \sum_{i_2=0}^1 \cdots \sum_{i_5=0}^1 w_{i_1}^1 w_{i_2}^2 \cdots w_{i_5}^5 \left(\underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & z_1^{i_1} z_2^{i_2} z_5^{i_5} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & z_2^{i_2} z_3^{i_3} z_5^{i_5} & 0 \end{bmatrix}}_{A_i} x + \underbrace{\begin{bmatrix} 0 \\ (\beta + \gamma) z_5^{i_5} \\ 0 \\ -\delta z_4^{i_4} z_5^{i_5} \end{bmatrix}}_{B_i} u \right) = \sum_{i=1}^r h_i(x) (A_i x + B_i u), \quad (2.31)$$

where $h_i = w_{i_1}^1 w_{i_2}^2 \cdots w_{i_5}^5$, $[i_1 i_2 \cdots i_5]$ is the 5-digit binary representation of $(i-1)$, $i \in \{1, 2, \dots, r\}$, $r = 2^5 = 32$. Despite its appearance, system (2.31) is not linear nor an approximation of (2.28): its construction proves that it is algebraically equivalent to the original one.

For brevity, only some of the 32 pairs (A_i, B_i) are given below; they correspond to the cases described in table 2.2, where z_i , $i \in \{1, 2, \dots, 5\}$ have taken some of their corresponding extreme values.

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -146.77 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 64.84 & 0 \end{bmatrix}, \quad A_8 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -162.15 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 169.48 & 0 \end{bmatrix},$$

$$A_{16} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -164.02 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 171.43 & 0 \end{bmatrix}, \quad A_{32} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -18.77 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 171.43 & 0 \end{bmatrix}$$

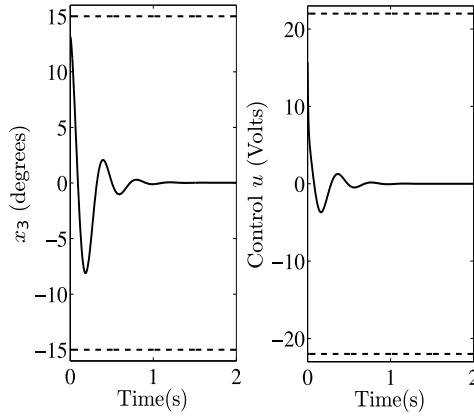


Figure 2.3: Left: angle x_3 (solid-line). Right: control signal $u(t)$

$$B_1 = \begin{bmatrix} 0 \\ 34.695 \\ 0 \\ -13.162 \end{bmatrix}, B_8 = \begin{bmatrix} 0 \\ 38.331 \\ 0 \\ -15.055 \end{bmatrix},$$

$$B_{16} = \begin{bmatrix} 0 \\ 38.331 \\ 0 \\ -15.055 \end{bmatrix}, B_{32} = \begin{bmatrix} 0 \\ 38.331 \\ 0 \\ -15.055 \end{bmatrix}.$$

Using LMIs (2.27) with $\Gamma_{ij} = A_i X + B_i M_j + X A_i^T + M_j^T B_i^T$, $i, j \in \{1, 2, \dots, 32\}$ yields a feasible solution. Due to space limitations, only the Lyapunov matrix and

Table 2.2: Some values for (2.29)

i	z_1	z_2	z_3	z_4	z_5
1	-0.2833	0.9886	0.1251	0.9659	524.0978
8	-0.2833	0.9886	0.2961	1.0000	579.0187
16	-0.2833	1.0000	0.2961	1.0000	579.0187
32	-0.0324	1.0000	0.2961	1.0000	579.0187

some of the controller gains are given:

$$P = \begin{bmatrix} 0.0032 & 0.0078 & 0.1325 & 0.0244 \\ 0.0078 & 0.0272 & 0.4562 & 0.0844 \\ 0.1325 & 0.4562 & 14.3922 & 1.6289 \\ 0.0244 & 0.0844 & 1.6289 & 1.6289 \end{bmatrix}, F_1 = \begin{bmatrix} 0.6351 \\ 2.2343 \\ 78.3388 \\ 8.0989 \end{bmatrix}^T,$$

$$F_8 = \begin{bmatrix} 0.2771 \\ 0.9938 \\ 53.5306 \\ 4.0084 \end{bmatrix}^T, F_{16} = \begin{bmatrix} 0.2776 \\ 0.9954 \\ 53.4616 \\ 4.0117 \end{bmatrix}^T, F_{32} = \begin{bmatrix} 0.1992 \\ 0.7203 \\ 45.3536 \\ 3.0391 \end{bmatrix}^T.$$

On the left, fig. 2.3 shows that the angle x_3 is stabilised from initial conditions $x(0) = [0 \ 0 \ 12.5^\circ \ 0]^T$; on the right, the control signal $u(t)$ is also shown for completeness.

It is worth mentioning that the quadratic framework includes much more than stability analysis and controller design of nonlinear systems: it covers observers (Tanaka, Ikeda, and Wang 1998), output feedback (Yoneyama et al. 2000), descriptors (Taniguchi et al. 1999), delay systems (Cao and Frank 2000), etc., both in the continuous- and discrete-time domain. The interested reader is referred to these works and the references therein for further information (Sala, Guerra, and Babuska 2005; Guerra, Sala, and Tanaka 2015).

Despite its success, the TS-LMI quadratic framework has one caveat when addressed to nonlinear control systems: its conditions are only sufficient. Such issue has a variety of sources and has been the main motivation of the different works appeared in the last two decades. Since this thesis tackles one of this sources, namely, modelling with respect to some performance objective as to optimise it, conservativeness is studied in some detail in the next section.

2.2.3 The issue of conservativeness

Long ago, three sources of conservatism (Sala 2009) have been recognised in the TS-LMI framework:

1. *The treatment of the MFs:* As seen in the previous section, analysis and synthesis of nonlinear control systems via TS models requires embedding them in convex representations in order to focus on the associated vertexes. In other words, the MFs only play a limited role in the derivation of the LMI

conditions as they are dropped off via relaxations or other results intending to involve them more without losing the LMI quality of solutions.

2. *The choice of Lyapunov function:* Converse theorems establish the existence of a Lyapunov function for every isolated asymptotically stable equilibrium point of a nonlinear system, but of course they say nothing about the type of Lyapunov function, let alone limiting ourselves to the class of quadratic Lyapunov functions the results in the previous section are based upon.
3. *The model choice:* It has been seen that TS models can be constructed in a variety of ways, leading to a different set of vertex matrices; feasibility of the LMIs depends on these matrices, which means that different modelling choices may produce radically different results (feasible/unfeasible, for instance) or provide different optimisation values.

The first issue has been already mentioned in the previous section with respect to sum relaxations. The approaches therein mentioned are *shape-independent* in the sense that they do not take into account anything else about the MFs except that they hold the convex-sum property. In contrast, *shape-dependent* approaches intend to exploit the particular form of the MFs in order to achieve less conservative LMI conditions: (Sala and Ariño 2007; Sala and Ariño 2008; Bernal, Guerra, and Kruszewski 2009; Narimani and Lam 2009; Arino and Sala 2015).

Other classes of Lyapunov functions

The second issue has sparked the investigation of different classes of Lyapunov functions that included the quadratic one as a particular case; some of them are briefly described below:

Piecewise Lyapunov functions (PWLF): Originally appeared in the context of fuzzy TS systems, they consist in a set of quadratic forms that get activated according to a pre-defined piecewise partition of the state space (Johansson and Rantzer 1998). This partition was usually made according to the operating or interpolating regimes induced by the MFs. Clearly, if the Lyapunov function candidate is allowed to change according to this partition, it may increase the chances of becoming an actual Lyapunov function not only because it will provide more flexibility (different Lyapunov matrices per partition (Johansson, Rantzer, and Arzen 1999)), but also because there are several ways of including the geometric information of the partition. Discrete-time counterparts can be found in (Feng 2003; Feng et al. 2005).

Basically, PWLFs have the form

$$V(x) = \begin{cases} x^T(t)P_i x(t), & x \in X_i, \quad i \in I_0 \\ \begin{bmatrix} x \\ 1 \end{bmatrix}^T \bar{P}_i \begin{bmatrix} x \\ 1 \end{bmatrix}, & x \in X_i, \quad i \in I_1 \end{cases} \quad (2.32)$$

where the state space is partitioned in a collection of cells X_i indexed by I_0 which will contain the indices of cells that contain the origin and I_1 which will have the indices of cells that do not contain the origin. Continuity of the PWLF across the borders between regimes is guaranteed by parametrisation of matrices $\bar{F}_i = [F_i \ f_i]$, with $F_i \in \mathbb{R}^{n \times n}$, $f_i \in \mathbb{R}^{n \times 1}$, $f_i = 0$ for $i \in I_0$, such that:

$$\bar{F}_i \begin{bmatrix} x \\ 1 \end{bmatrix} = \bar{F}_j \begin{bmatrix} x \\ 1 \end{bmatrix}, \quad x \in \{X_i \cap X_j\}, \quad i, j \in I. \quad (2.33)$$

A systematic procedure for constructing these matrices is given in (Johansson, Rantzer, and Arzen 1999). Then, P_i and \bar{P}_i in (2.32) are parameterised as follows:

$$\begin{aligned} P_i &= F_i^T T F_i, & i \in I_0 \\ \bar{P}_i &= \bar{F}_i^T T \bar{F}_i, & i \in I_1. \end{aligned} \quad (2.34)$$

with T being a symmetric matrix of appropriate dimensions which collects the free parameters of the Lyapunov function. Note that this arrangement is a compromise between continuity and LMI formulation of the results, which is hereby possible.

The partition is usually decided along the TS model structure, which due to locality can be extended to include affine TS models

$$\dot{x}(t) = \sum_{l=1}^r h_l(z(x)) (A_l x(t) + a_l), \quad (2.35)$$

where $A_l \in \mathbb{R}^{n \times n}$ and $a_l \in \mathbb{R}^{n \times 1}$ are the local matrices and affine terms, respectively, and MFs h_l , $l \in \{1, 2, \dots, r\}$ depend on the premise vector $z(\cdot)$, which in turn is assumed to depend linearly on the system state $x(t)$, i.e., $z(x) = Cx$, with $C \in \mathbb{R}^{p \times n}$.

Since matrices P_i or \bar{P}_i are only used to describe the Lyapunov function in cell X_i , it is possible to use the S-procedure in property 1.5 to reduce conservativeness by constructing matrices $\bar{E}_i = [E_i \ e_i]$ with $F_i \in \mathbb{R}^{n \times n}$, $f_i \in \mathbb{R}^{n \times 1}$, $e_i = 0$ for $i \in I_0$, that satisfy

$$\bar{E}_i \begin{bmatrix} x \\ 1 \end{bmatrix} \succeq 0, \quad x \in X_i, \quad i \in I \quad (2.36)$$

where for every matrix W_i with nonnegative entries $W_i \succeq 0$, condition (2.36) implies that

$$\begin{bmatrix} x \\ 1 \end{bmatrix}^T \bar{E}_i^T W_i \bar{E}_i \begin{bmatrix} x \\ 1 \end{bmatrix} > 0, \quad \forall x \in X_i, \quad i \in I.$$

As with matrices \bar{F}_i , these \bar{E}_i can also be systematically constructed (Johansson, Rantzer, and Arzen 1999).

Since this approach was originally addressed to fuzzy TS models, regions where $h_l(x) = 1$ for some l (*operating regimes*) and those in between operating regimes (*interpolation regimes*) were employed to derive LMI conditions; they replace stability analysis LMIs in theorem 2.3 by

$$\begin{aligned} P_i - E_i^T U_i E_i &> 0 \\ A_k^T P_i + P_i A_k + E_i^T W_{ik} E_i &< 0 \end{aligned} \quad (2.37)$$

for $i \in I_0$, $k \in K(i)$, and

$$\begin{aligned} \bar{P}_i - \bar{E}_i^T U_i \bar{E}_i &> 0 \\ \bar{A}_k^T \bar{P}_i + \bar{P}_i \bar{A}_k + \bar{E}_i^T W_{ik} \bar{E}_i &< 0 \end{aligned} \quad (2.38)$$

for $i \in I_1$, $k \in K(i)$, $K(i)$ being the set of indexes of those systems associated with region i . The reader may notice the relaxation is twofold: P is no longer unique and local information has been added through the S-procedure.

Extensions for stability analysis of nonlinear systems modelled as exact affine TS ones appeared in (Gonzalez and Bernal 2016) with interesting geometric refinements in (Gonzalez et al. 2017c). Stabilisation based on PWLFs remains, unfortunately, a BMI problem as shown in (Feng et al. 2005); observer design and output feedback based on PWLFs can be found in (Qiu, Feng, and Gao 2012) and (Qiu, Feng, and Gao 2013), respectively.

Parameter-dependent Lyapunov functions (PDLFs): They incorporate the convex structure provided by the MFs of a TS model to enrich the choice of a common matrix P by replacing it by a convex sum of matrices P_i (Blanco, Perruqueti, and Borne 2001):

$$V(x(t)) = \sum_{i=1}^r h_i(z(x)) x^T(t) P_i x(t) = x^T(t) P_h x(t) \quad (2.39)$$

where $P_i = P_i^T > 0$ and $h_i(z(x))$ are the same MFs of the associated TS model (2.10), for $i \in \{1, 2, \dots, r\}$. PDLFs are also known as fuzzy (Guelton et al. 2010), multiple (Tanaka, Hori, and Wang 2003), non-quadratic (Bernal and Guerra

2010), convex (Fiacchini, Girard, and Jungers 2016), basis-dependent (Su et al. 2014), and poly-quadratic (Pandey and DeOliveira 2017).

Though increasing the flexibility, PDLFs faced a lot of problems in the continuous-time case since the time derivative of the Lyapunov function (2.39) implies deling with the time derivative of P_h which in turn involves the time derivatives of the MFs $h_i(z)$, thus breaking the convexity of expressions and creating algebraic loops. Solutions using direct bounds can be found in (Tanaka, Hori, and Wang 2001; Wu and Dong 2006a; Mozelli et al. 2009); others based on further analysis of the partial derivatives of the MFs \dot{h}_i appeared in (Guerra and Bernal 2009; Bernal and Guerra 2010; Guerra et al. 2012; Pan et al. 2012); some more solved the algebraic loop under mild conditions (Aguiar, Márquez, and Bernal 2016; Gonzalez et al. 2017b).

In contrast, PDLFs have made an impressive progress in the discrete-time domain, as they are not faced with the problem of the time derivative of the MFs: different pairs of control laws and Lyapunov functions were proposed in (Guerra and Vermeiren 2004), the α -samples approach was proposed in (Guerra, Kruszewski, and Bernal 2009), interesting generalisations for observers can be found in (Estrada-Manzo, Lendek, and Guerra 2016).

Closely related with the PDLFs, the proposal in (Rhee and Won 2006) intended to involve the MFs without dealing with their time derivatives by using the following path-independent line-integral Lyapunov function:

$$V(x) = \int_{\Gamma(0,x)} f(\psi) d\psi, \quad (2.40)$$

with $\Gamma(0, x)$ being a path from the origin 0 to the current state x , ψ as a dummy vector for the integral, and $f(x) = \sum_{i=1}^r h_i(z(x)) P_i x = P_h x$. Path-independency conditions translate into $f(x) = [f_1(x), f_2(x), \dots, f_n(x)]^T$ holding the condition $\frac{\partial f_i(x)}{\partial x_j} = \frac{\partial f_j(x)}{\partial x_i}$, $i, j, \in \{1, 2, \dots, n\}$. Extensions to the line-integral proposal appeared in (Marquez et al. 2013; Gonzalez et al. 2017a).

Other models

Recall that the LMI conditions for analysis and synthesis of nonlinear systems in the TS-LMI framework depend almost exclusively on the vertexes of the resulting polytope. Therefore, choosing a set of nonlinearities. Besides the fact that there are infinitely many TS models for a given nonlinear system, far beyond the obvious

choices described in section 2.1.2, there are other convex models that will be briefly described for the sake of completeness.

Exact polynomial TS models: Based on the Taylor series, the work in (Sala and Ario 2009) proposed a generalisation of the well-know TS sector nonlinearity methodology where non-polynomial expressions are rewritten as a convex sum of polynomials (TS models are convex sums of linear terms), leaving vertex polynomial vectors to be used via the sum-of-squares (SOS) tools (Prajna, Papachristodoulou, and Wu 2004).

Systems of the form $\dot{x}(t) = \tilde{f}(\eta(x), x)$, where $\eta(x) = [\eta_1(x) \ \eta_2(x) \ \cdots \ \eta_p(x)]^T$ collects all *non-polynomial* nonlinearities in $f(\cdot)$, are analysed by rewriting the latter as a convex sums of polynomials of arbitrary order. Since for any sufficiently smooth function of one real variable, $\eta(x)$ there exists a Taylor expansion of degree N (Apostol 1967), there must be a point $\psi(x) \in [0, x]$, so that:

$$\eta(x) = \sum_{i=0}^{N-1} \frac{\eta^{[i]}(0)}{i!} x^i + \frac{\eta^{[N]}(\psi(x))}{N!} x^N, \quad (2.41)$$

where $\eta^{[i]}(x)$ denotes the i -th derivative of $\eta(\cdot)$ and $\eta^{[0]}(x)$ is defined, plainly, as $\eta(x)$ (Chesi 2009). Additionally, assume that $\eta^{[N]}(x)$ is continuous in a compact region of interest Ω . Denoting the Taylor approximation of order N of the function $\eta(x)$ by:

$$\eta_N(x) = \sum_{i=0}^{N-1} \frac{\eta^{[i]}(0)}{i!} x^i,$$

and let

$$T_N(x) = \frac{\eta(x) - \eta_N(x)}{x^N}.$$

In the region Ω , $T_N(x)$ is bounded; therefore, the following bounds are well defined:

$$\psi_0 := \sup_{x \in \Omega} T_N(x), \quad \psi_1 := \inf_{x \in \Omega} T_N(x),$$

based on which the following convex rewriting of $T_N(x)$ arises:

$$T_N(x) = w_0(x) \cdot \psi_0 + w_1(x) \cdot \psi_1, \quad (2.42)$$

with

$$w_0(x) = \frac{T_N(x) - \psi_1}{\psi_0 - \psi_1}, \quad w_1(x) = 1 - w_0(x). \quad (2.43)$$

Then, an equivalent convex representation of (2.41) exists in the form:

$$\eta(x) = w_0(x) \cdot p_0(x) + w_1(x) \cdot p_1(x) = \sum_{i=0}^1 w_i(x) \dot{p}_i(x) \forall x \in \Omega, \quad (2.44)$$

where $p_0(x) = \eta_N(x) + \psi_0 x^N$ and $p_1(x) = \eta_N(x) + \psi_1 x^N$ are polynomials of degree N , and $w_0(x)$, $w_1(x)$ are weighting functions which hold the convex sum property in the compact region Ω .

If every $\eta_j(x)$ is rewritten as in (2.44), then $\dot{x}(t) = \tilde{f}(\eta(x), x)$ can be rewritten as the following exact polynomial TS model:

$$\begin{aligned} \dot{x} &= \sum_{i_1=0}^1 \sum_{i_2=0}^1 \cdots \sum_{i_q=0}^1 w_{i_1}^1 w_{i_2}^2 \cdots w_{i_q}^q F_{(i_1, i_2, \dots, i_q)}(x), \\ &= F_w(x(t)), \end{aligned} \quad (2.45)$$

where $F_{(i_1, i_2, \dots, i_q)} = \tilde{f}(\eta(x), x)|_{w_{i_1}^1 = w_{i_2}^2 = \dots = w_{i_q}^q = 1}$. Note that this is a polynomial TS generalisation of the tensor-product structure (2.6) (Ariño and Sala 2007).

Exact descriptor TS models: Descriptor models describe a wider class of dynamical systems than the standard state-space representation. They are constituted by a set of *differential algebraic equations* (Luenberger and Arbel 1977) and have been used to represent implicit or singular systems arising in many physical setups, especially mechanical ones, where they are able to describe parametric perturbations with higher accuracy than standard representations (Guelton, Delprat, and Guerra 2008; Schulte and Guelton 2009). A nonlinear descriptor model has the form:

$$E(x)\dot{x}(t) = A(x)x(t) + B(x)u(t), \quad (2.46)$$

where $E \in \mathbb{R}^{n \times n}$ is a novel element which, depending on its complexity and singularity, may induce interesting nonlinear phenomena such as impasse points and indeterminacies. The sector nonlinearity methodology, even the Taylor-based one in the previous section, can be used to exactly model the previous system as to obtain a TS one, for instance:

$$\sum_{i=1}^{r_e} v_i(z(t)) E_i \dot{x}(t) = \sum_{i=1}^r h_i(z(t)) (A_i x(t) + B_i u(t)) \quad (2.47)$$

where E_i are matrices of proper size, r MFs h_i represent the usual combination instance of extreme values of the nonlinearities in the right-hand side, and r_e MFs v_i have the same meaning for the left-hand side. Of course, due to the convex sum

property, these two sets of models can be combined into a single one if convenient by defining a new set of MFs in terms of the two original sets, i.e., $\bar{h}_k = h_i v_j$, $k \in \{1, 2, \dots, r \times r_e\}$.

Approximate models: A word must be added to give account of the numerous models available in the literature that are close to the spirit of the original TS methodology, most of them approximate. Despite of the lack of exactness, the methodologies involved therein might be of relevance for our contributions.

1. *Linearisation in multiple points:* A TS fuzzy approximation of a given nonlinear model can be obtained by a Taylor series expansion in different representative points (Johansen, Shorten, and Murray-Smith 2000), which may or may not be equilibria. However, this approximation will in general result in affine consequents.
2. *The substitution method:* To avoid affine TS models arising from linearisation while preserving a linear family of models, this method, first proposed in (Kiriakidis 2007), writes the righthand side of the model as $A(x)x + B(x)u(t)$ and, instead of representing each nonlinearity as the convex sum of two linear functions, the values are simply substituted for a chosen grid of points. The resulting points are used instead of vertexes of a polytopic representation for analysis and synthesis.
3. *Least squares fit:* Available data sets from identification processes can serve well for constructing a TS model by conventional least-squares fit (Baruah and Angelov 2011).

This thesis is concerned with contributions on reducing the issue of conservativeness by better choosing the TS model. Yet, the approach is a completely new one, for the model is not chosen a priori but as a result of aiming to achieve the best performance measure once this is given. Before proceeding to these contributions we briefly review the theory of invariant sets, as this is an alternative to the use of Lyapunov theory for stability analysis and controller design in the discrete-time domain.

Chapter 3

Set-based control theory

This chapter will outline the basic ideas underlying set-based stability analysis and control design that will be later on used in the developments in Chapter 8. We will discuss invariant sets, one step sets and polyhedron manipulation.

3.1 Invariant sets

An invariant set for a given dynamic system is a region of the state space such that the trajectory generated by the system remains confined in the set if the initial condition lies within it (Blanchini and Miani 2008).

Definition 3.1. (Fiacchini 2010b) *A set in $\Omega \in \mathbb{R}^n$ is a invariant set for the discrete-time autonomous system (3.1) if $\Omega \subset D$ and $f(x) \in \Omega$ for all $x \in \Omega$.*

Of course, the idea of invariant sets roots at the basics of linear and nonlinear control (Khalil 2000; Blanchini 1999; Blanchini 1991). The importance of invariant sets in control is due to the implicit stability properties of these regions of the state space. Consider the autonomous discrete-time system

$$x_+ = f(x) \tag{3.1}$$

where $x \in \mathbb{R}^n$ is the state, $x_+ \in \mathbb{R}^n$ is the successor state and $f : D \rightarrow \mathbb{R}^n$ is a function defined on the set $D \in \mathbb{R}^n$. Ω is a invariant set if every trajectory generated by (3.1) with initial condition x_0 in Ω , remains confined in the set Ω .

Originally, the term invariant set may denote a set of initial conditions whose trajectory “backward and forward” in time is confined in the set, in this thesis it will refer as a invariant set only the future part of trajectories are required to belong to the set (other authors explicitly define such sets as “positively invariant” (Khalil 2000)).

Analogous definitions can be given for uncertain autonomous systems, that is, in presence of uncertainties, in this case were uncertainties affects the dynamic system, the condition of robust invariance has to be introduced. Consider a discrete-time uncertain autonomous system

$$x_+ = f(x, w) \tag{3.2}$$

where $x \in \mathbb{R}^n$ is the state, $x_+ \in \mathbb{R}^n$ is the successor state, $w \in \mathbb{R}^p$ is the uncertainty, i.e. $w \in W$ with $W \subset \mathbb{R}^p$, and $f : D \times W \rightarrow \mathbb{R}^n$ is a function defined on the set $D \times W \subset \mathbb{R}^{n+p}$.

Definition 3.2. (Fiacchini 2010b) *A set $\Omega \subset \mathbb{R}^n$ is a robust invariant set for the uncertain autonomous system (3.2) if $\Omega \subset D$ and $f(x, w) \in \Omega$ for all $x \in \Omega$ and all $w \in W$.*

That is, robust invariance means that any trajectory of the uncertain system starting inside the set Ω , remains confined in it for every possible realization of the uncertainty $w \in W$, at every time instant. We have that set $\Omega \subset D$ is a robust invariant set for system 3.2 if it is such that $f(\subset, W) \subset \Omega$.

In the same way, definitions can be given for non-autonomous systems, that is, in presence of a manipulable input, this is a control invariant set: a region Ω of the state space such that, for any of its elements $x \in \Omega$, there exists a control input $u(x)$ that maintains the successor state inside Ω .

Given a control invariant set Ω , there exists at least a control law $u(x)$ defined on Ω such that the set is an invariant set for the system in closed-loop with $u(x)$, i.e., consider the non-autonomous system

$$x_+ = f(x, u) \tag{3.3}$$

where $x \in \mathbb{R}^n$ is the state, $x_+ \in \mathbb{R}^n$ is the successor state, $u \in U$ is the control input and $f : D \times U \rightarrow \mathbb{R}^n$ is a function defined on the set $D \times U \subset \mathbb{R}^{n+m}$.

A set $\Omega \subset D$ is a control invariant set if there exists a control law $u = u(x) \in U$, defined for every $x \in \Omega$, such that every trajectory x_k generated by (3.3), in closed-loop with $u(x)$ and with $x_0 \in \Omega$, is such that $x_k \in \Omega$ for all $k \in \mathbb{N}$.

Definition 3.3. A set $\Omega \subset D$ is a control invariant set for the discrete-time system (3.3) if there exists a control law $u = u(x) \in U$ such that $\Omega \subset D$ and $f(x, u(x)) \in \Omega$ for all $x \in \Omega$.

Definition 3.4. (C-set) A C-set is a convex and compact subset of \mathbb{R}^n including the origin as an interior point.

The concept of invariance can be further refined to include geometric contraction.

Definition 3.5. (Contractive set, discrete-time). The C-set S is contractive for the system

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

where $w \in W$, $u \in U$, if and only if there exists a control function $u(x) \in U$ such that, for every $x \in S$, the following condition holds:

$$\Psi_S(f(x, u, w)) < \lambda, \forall w \in W$$

where $\Psi_S(x)$ is the Minkowski function¹ of S , for some $0 < \lambda < 1$. In this case the set S is said to be λ -contractive.

Geometric contraction is closely related to stability of linear systems, as in this case all scalings of a contractive set are contractive. It is also related to decay-rate stability, see Section 3.3.1.

Model predictive control. Due to the relevance of the role that the theory of invariant sets plays in the design of control laws for model predictive control (MPC), this thesis makes mention of it, showing, very briefly, some of its properties and rudimentary notions of how the theory of invariant sets intervenes. Recalling some of the principals characteristics of a standard MPC, it has to be mentioned: *Model based prediction*, where is assume the real system is know, measured and is possible a prediction of the system evolution as a function of the input. *Constrains*, where are gather the state and control input sequences which stated possible trajectories of the system and it is arranged as a optimization problem which is solved on-line. The solution to this optimization problem give a subset of the space of input sequences which represented potential admissible trajectories for the system. *Cost function* have for objective compute the optimization problem providing high performance with low control effort, i.e., obtain admissible control sequences and associated trajectories with low control effort. *Receding horizon*, to avoid discrepancies between the actual behavior of the

¹Minkowski function is defined as $\Psi_S(\psi) := \inf\{\kappa \geq 0 : \psi \in \kappa S\}$.

system and the predicted trajectory, only the first control action of the feasible input sequence obtained is considered, this is possible because a new solution is calculated and obtained at each instant of time.

Convergence for a MPC are often linked to a terminal region and a local control law which guarantees stability. Intuitively, we can consider this terminal region as an invariant set with its corresponding control law ensuring stability once the system reaches such the set, at this point there is no more necessity of the computation of MPC. In addition, it is possible to use assist at the problem of designing control laws for MPC by means of the invariant set methods, for example, assuming that there is an invariant set control for the system of interest, it is enough to replace the constrains of the MPC by a restriction that stipulates that the state belongs to the set at next time step, this guarantees the existence of a controller.

Nevertheless, as multi-step predictive control is out of the scope of this thesis, the reader is referred to, for instance, (Kerrigan 2000) where the set-based methodologies for MPC are developed, the widely-cited book (Camacho and Alba 2013) for an in-depth global coverage of MPC issues, or the work (Ariño, Querol, and Sala 2017) in which these set-based MPC ideas are generalised to a gain-scheduled Takagi-Sugeno (polytopic LDI) framework. In fact, set-based MPC control ideas apply too to switching/Markov jump systems (Sala, Hernández-Mejías, and Ariño 2017) but, again, these classes of models are out of the scope of the present document, so the reader is referred to the cited references for further detail, if interested.

3.2 Convex functions and sets

This section provided some essential notions of convex sets which will be useful in the sequel (Blanchini and Miani 2008).

Definition 3.6. (*Convex set*) A set $\mathcal{S} \in \mathbb{R}^n$ with $x_1 \in \mathcal{S}$ and $x_2 \in \mathcal{S}$, is said to be convex if

$$\alpha x_1 + (1 - \alpha)x_2 \in \mathcal{S} \text{ for all } 0 \leq \alpha \leq 1 \quad (3.4)$$

Note that this definition applies to the empty set \emptyset , therefore the empty set \emptyset is convex.

The above definition make use of the expression of the point $x = \alpha x_1 + (1 - \alpha)x_2$, which is called a convex combination of the pair x_1 and x_2 . This is important

due the set of all such points is the segment connecting the points x_1 and x_2 . An alternative view of a convex set is a set that includes all the segments connecting all the pairs of its points (in case of more of two points).

Definition 3.7. (*Convex hull*). Given a set \mathcal{S} , its convex hull is the smallest convex set containing \mathcal{S} .

The convex hull of a set \mathcal{S} can be alternatively seen as the intersection of all the convex sets $\bar{\mathcal{S}}_i$ containing \mathcal{S} , i.e.,

$$\text{co}(\mathcal{S}) = \bigcap_i \bar{\mathcal{S}}_i$$

Following, is introduces the definition of convexity for functions.

Definition 3.8. (*Convex function*). A real function $\psi \in \mathcal{S}$ with $\mathcal{S} \subset \mathbb{R}^n$, $\psi : \mathcal{S} \rightarrow \mathbb{R}$ is said to be convex if

$$\psi(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha\psi(x_1) + (1 - \alpha)\psi(x_2), \quad 0 \leq \alpha \leq 1 \quad (3.5)$$

holds for all $x_1, x_2 \in \mathcal{S}$.

Any sublevel set $\bar{\phi}[\psi(x), \kappa] := \{x : \psi(x) \leq \kappa, \kappa > 0\}$ of a convex function $\psi(x)$ is a convex set. However, the vice-versa is not true, functions whose sublevel sets are convex are not necessarily convex.

Definition 3.9. (*Quasi-convex function*). A function $\psi(x)$ with $x \in \mathbb{R}^n$ and $\phi(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ is quasi-convex if the following condition holds

$$\psi(\alpha x_1 + (1 - \alpha)x_2) \leq \max\{\psi(x_1), \psi(x_2)\}, \quad \forall x_1, x_2, 0 \leq \alpha \leq 1. \quad (3.6)$$

Note: Convexity implies quasi-convexity, while the opposite is not true. A function is quasi-convex if and only if its sublevel sets are convex.

Definition 3.10. (*Support function*). Given a convex set \mathcal{S} , a support function is defined as

$$\phi_{\mathcal{S}}(z) := \sup_{x \in \mathcal{S}} z^T x. \quad (3.7)$$

It is possible represented a convex and closed set in terms of its support function, i.e., if the set \mathcal{S} is convex and closed,

$$\mathcal{S} = \{x : z^T x \leq \phi_{\mathcal{S}}(z) \quad \forall z \in \mathbb{R}^n\} \quad (3.8)$$

3.2.1 Set operations

(Blanchini and Miani 2008) Some basic operations between sets are presented. Consider \mathcal{A} and \mathcal{B} as subsets of \mathbb{R}^n , i.e., $\mathcal{A}, \mathcal{B} \subset \mathbb{R}^n$

Definition 3.11. (*Operations on convex sets*)

1. *Sum of sets \mathcal{A} and \mathcal{B} : the set*

$$\mathcal{A} \oplus \mathcal{B} := \{x = a + b, a \in \mathcal{A}, \text{ and } b \in \mathcal{B}\}$$

where λ is a real number.

2. *Scaled set of \mathcal{A} : the set*

$$\lambda \mathcal{A} := \{x = \lambda a, a \in \mathcal{A}\}, \lambda \geq 0$$

3. *Image of a set under an affine map M : the set*

$$M(\mathcal{A}) := \{y = M(x), x \in \mathcal{A}\}$$

with $M : \mathbb{R}^n \rightarrow \mathbb{R}^m$

4. *Projection of \mathcal{A} on a subspace \mathcal{X} : the set*

$$\text{Proj}_{\mathcal{X}}(\mathcal{A}) := \{b \in \mathcal{X} : \exists a \in \mathcal{A} : a = b + c, \text{ with } c \in \mathcal{X}^\perp\}$$

where \mathcal{X}^\perp denotes the subspace orthogonal to \mathcal{X} .

Theorem 3.1. *All sets in the above definition are convex if \mathcal{A} and \mathcal{B} are convex.*

3.3 One-step sets

Consider a dynamic system

$$x_+ = f(x, u) \tag{3.9}$$

defined for $x \in \mathbb{X}$ and for $u \in \mathbb{U}$ being \mathbb{X} and \mathbb{U} known sets.

Given a set $\Omega \subset \mathbb{X}$ in the state space and a dynamic system, the one-step set $Q(\Omega)$ relates Ω to the set of points in X whose evolution through the dynamic function is contained in Ω . That is, given $\Omega \subset \mathbb{X}$, a point x belongs to the one-step set $Q(\Omega)$ if $x \in \mathbb{X}$ and $f(x) \in \Omega$. Hence, $\mathbb{X}_1 = Q(\mathbb{X})$ is the set of points in X which remain in \mathbb{X} at least at the first instant. It is clear that iterative application

generates a sequence of sets $\mathbb{X}_{k+1} = Q(\mathbb{X}_k) \cap \mathbb{X}_k$, such that a point x belongs to \mathbb{X}_k if and only if the trajectory generated with initial condition $x_0 = x$ remains in \mathbb{X} at least during the first k steps, for all $k \in N$. It should be also evident that the **maximal invariant set** can be obtained iterating the procedure for an infinite number of steps.

The result would be not very useful unless the maximal invariant set can be obtained after a finite number of iterations. In this case the invariant set is said to be finitely determined and the number of iterations is denoted as determination index. Important contributions have been provided in literature, mainly for linear systems, which permit to establish conditions under which the maximal invariant set is finitely determined. Basically, if LMIs can be proven to achieve some (quadratic decay rate) $V(x_+) \leq \lambda_*^2 V(x)$, then the maximal invariant (or contractive) set can be proven to converge in a finite number of iterations for any geometric decay $\lambda > \lambda_*$.

The basic ideas above can be formalised in the definitions below:

Definition 3.12. *Given an arbitrary target set Ω , the one-step set $\mathcal{Q}(\Omega)$ is the set of states x in \mathbb{X} from which the next state of system (3.9) can be driven to Ω with an admissible $u \in \mathbb{U}$, i.e.,*

$$\mathcal{Q}(\Omega) := \{x \in \mathbb{X} \mid \exists u \in \mathbb{U} : f(x, u) \in \Omega\}$$

Note that $x \in \mathcal{Q}(\Omega)$ iff $U_\Omega(x) \neq \emptyset$. Also, Ω is control λ -contractive iff $\Omega \subset Q(\lambda\Omega)$.

Definition 3.13 ((Gilbert and Tan 1991)). *The so-called i -step set $\mathcal{C}_i^\lambda(\Omega)$ is recursively defined, starting with $\mathcal{C}_0^\lambda(\Omega) := \Omega$ as $\mathcal{C}_{i+1}^\lambda(\Omega) := \mathcal{Q}(\lambda\mathcal{C}_i^\lambda(\Omega)) \cap \Omega$, for $i \geq 0$.*

If there exists a finite i such that $\mathcal{C}_{i+1}^\lambda(\Omega) = \mathcal{C}_i^\lambda(\Omega)$, it can be proved (Kerrigan 2000) that $\mathcal{C}_i^\lambda(\Omega)$ is the maximal control λ -contractive set. Such set will be denoted as $\mathcal{C}_\infty^\lambda(\Omega)$. Also, in case such finite i does not exist, but there exists $\mathcal{C}_\infty^\lambda$, for any $1 \geq \lambda^* > \lambda$, there exist a finite i^* such that \mathcal{C}_i^λ is control λ^* -contractive for all $i \geq i^*$, albeit possibly non-maximal (Blanchini 1991, Theorem 3.2).

Efficient computational characterisation of the one-step set \mathcal{Q} in Definition 3.12 can only be easily carried out for special cases of f ; for instance, the linear case (Kerrigan 2000).

Example 3.1. *In a linear system $x_+ = Ax + Bu$, for $\mathbb{X} = \mathbb{R}^n$, $\Omega := \{R_x x \leq s_x\}$, $\mathbb{U} := \{R_u u \leq s_u\}$ the one-step set is the projection onto the x subspace of the set*

$$\mathcal{Q}^* := \{(x, u) : R_x(Ax + Bu) \leq s_x, R_u u \leq s_u\}$$

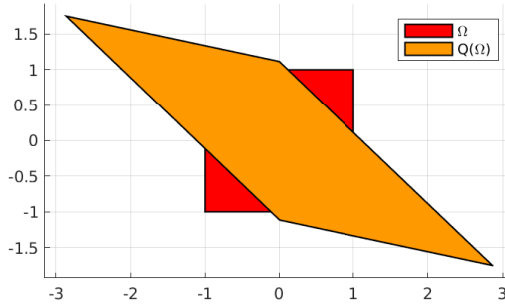


Figure 3.1: One-step set computation for a linear-polyhedral case.

Using the *Multiparametric Toolbox MPT 3.0* (Herceg et al. 2013), the following code computes such one-step set for a particular example:

```
A=[0.9 0.9;.2 0.9];B=[0.2;-0.3];

Rx=[eye(2);-eye(2)];
sx=[1 1 1 1]';
Ru=[1;-1]; su=[.01;.01];

omega=Polyhedron(Rx,sx);
unpasomegaxu=Polyhedron([ Rx*[A B]; [zeros(2,2) Ru]] ,[sx;su]);
unpasoproj=projection(unpasomegaxu,[1 2])

plot([omega unpasoproj])
legend('\Omega','Q(\Omega)')
```

resulting in the plot in Figure 3.1 where the target set Ω , a square, appears in red, and its one-step set appears in orange.

Another important property of the one-step operator is the fact that applying the operator to a set which is invariant, generates another invariant set which contains the previous one. Thus, the iterative application of the one-step operator, with a given invariant set as initial element, produces a growing sequences of invariant sets. Notice that the same iterative process with $X_0 = X$ as initial element, generates a sequence of sets not necessarily invariant, which entails that invariance of the current set is not guaranteed until the determination index is reached (if finite).

In some cases, it can be proved that iterations initialized with an invariant set converge to the domain of attraction of an equilibrium point, that is, to the set of points which converge to the equilibrium. Clearly, assumptions on stability of the system are required in this case.

3.3.1 Convex Lyapunov functions and contractive sets

Consider a discrete-time system $x_+ = f(x, u)$.

Definition 3.14. (Ariño et al. 2017) *A convex function $V(x)$ such that $V(0) = 0$ is a (local) control Lyapunov function (CLF) ensuring geometric contraction rate λ for system (3.9) if there exists a convex set $\Omega \subset \mathbb{X}$ including the origin in which, for all $x \in \Omega \sim \{0\}$, $V(x) > 0$ and there exists $u \in \mathbb{U}$ such that $V(\lambda^{-1}f(x, u)) \leq V(x)$.*

The above definition is an adaptation to the discrete-time and contraction-rate setting of well-known concepts defined in, for instance, (Sontag 1999) for continuous-time stabilization. The motivation of the definition is the fact that V is a Lyapunov function as, by convexity, $V(f(x, u)) \leq \lambda V(\lambda^{-1}f(x, u)) \leq \lambda V(x)$ and the level sets are control λ -contractive, as a level set $\Omega := \{V(x) \leq \gamma\}$ scaled would be $\lambda\Omega = \{V(\lambda^{-1}x) \leq \gamma\}$, so if $x \in \Omega$, next state $f(x, u)$ will lie in $\lambda\Omega$ for some u . If $f(x, u)$ were linear, condition in the definition could be stated as $V(f(x, u)) \leq V(\lambda x)$, because $V(f(x, u)) = V(\lambda^{-1}f(\lambda x, \lambda u)) \leq V(\lambda x)$.

In many common cases, $V(x)$ is a homogeneous degree- q polynomial in x , then $V(\lambda^{-1}x) = \lambda^{-q}V(x)$; standard discrete decay-rate formulas $V(\lambda^{-2}x_{k+1}) \leq V(x_k)$ arise with, for instance, $q = 2$. In the homogeneous case, we have

$$V(\lambda^{-k}x_k) = \lambda^{-(k-1)q}V(\lambda^{-1}x_k) \leq \lambda^{-(k-1)q}V(x_{k-1}) = V(\lambda^{-(k-1)}x_{k-1})$$

so, by induction, we can easily prove $V(\lambda^{-k}x_k) \leq V(x_0)$ or, equivalently, by multiplication by λ^{kq} we get $V(x_k) \leq V(\lambda^k x_0)$.

Corollary 3.1.1. *Any level set of a local CLF in \mathbb{X} ensuring contraction rate λ is a subset of the maximal control λ -contractive set in \mathbb{X} .*

Proof. Evident, because of the above-mentioned fact that the referred level sets are control λ -contractive and all such sets are subsets of the maximal one. \square

3.3.2 One-step sets for Takagi-Sugeno systems

Consider a Takagi-Sugeno model:

$$x_+ = \sum_{i=1}^r \mu_i(x)(A_i x + B_i u) \quad (3.10)$$

For this TS system (which is a “rewriting” of a nonlinear dynamics), the true one-step set is:

$$\mathcal{Q}(\Omega) = \{x \in \mathbb{X} \mid \exists u \in \mathbb{U} : \sum_{i=1}^r \mu_i(x)(A_i x + B_i u) \in \Omega\} \quad (3.11)$$

The shape of $\mathcal{Q}(\Omega)$ may be very hard to compute, due to the nonlinearities in the membership functions $\mu_i(x)$.

A reasonable approach, in order to deal with this drawback, is disregarding the information about the actual value of the membership functions, dealing with the Takagi-Sugeno model for *any* possible value of μ_i –assumed known to the controller, as done in most TS literature (i.e., a shape-independent analysis (Sala 2009))–. Hence, the one-step set in Definition 3.12 should be *replaced* by the one below (Ariño et al. 2017):

Definition 3.15. *The shape-independent one-step set of a TS system (3.10) is*

$$\mathcal{Q}_{si}(\Omega) := \{x \in \mathbb{X} \mid \forall \mu \in \Delta \exists u \in \mathbb{U} : \sum_{i=1}^r \mu_i(A_i x + B_i u) \in \Omega\} \quad (3.12)$$

A shape-independent definition of λ -contractiveness also follows: Ω is shape-independent λ -contractive if $\Omega \subseteq \mathcal{Q}_{si}(\Omega)$.

The definition ensures that for each $(x, \mu) \in \mathcal{Q}_{si}(\Omega) \times \Delta$ there exists a non-empty set of membership-dependent control actions defined as:

$$U_\Omega(x, \mu) := \{u \in \mathbb{U} \mid \tilde{f}(\mu, x, u) \in \Omega\} \quad (3.13)$$

If Ω is polyhedral, the set $U_\Omega(x, \mu)$ is itself a polytope, for fixed x and μ . Unfortunately, exact computation of \mathcal{Q}_{si} is still cumbersome, due to the nonlinearities involving products of μ_i with x and u .

Let us show that $\mathcal{Q}_{si}(\Omega) \subset \mathcal{Q}(\Omega)$. Indeed,

$$\mathcal{Q}(\Omega) = \{x \in \mathbb{X} \mid \text{for } \mu \equiv \mu(x) \exists u \in \mathbb{U} : \sum_{i=1}^r \mu_i(A_i x + B_i u) \in \Omega\} \supset \mathcal{Q}_{si}(\Omega) \quad (3.14)$$

as the conditions in the left-hand side of (3.14) involve only the single point $\mu(x)$, instead of the whole simplex in (3.12).

The following facts hold for the above shape-independent sets, see (Ariño et al. 2017) for details and proofs.

Proposition 3.1. *If Ω is shape-independent control λ -contractive for the TS system (3.10), so it is its convex hull $Co(\Omega)$. Thus, the maximal shape-independent control λ -contractive set is convex.*

Proposition 3.2. *If Ω is shape-independent control λ -contractive for the TS system (3.10), then any linear scaling $\gamma\Omega$, with $0 < \gamma \leq 1$, is shape-independent control λ -contractive, too.*

Hence, as shape-independent control λ -contractive sets are control λ -contractive, the following well-known result and Proposition 3.2 can be joined to induce a control Lyapunov function, if a shape-independent control λ -contractive set is found:

Proposition 3.3. *(Blanchini 1999) Consider $\Omega = \{x \in \mathbb{R}^n \mid \max_{1 \leq i \leq n_h} H_i x \leq 1\}$. If $\gamma\Omega$ is control λ -contractive, for the TS system (3.10), for all $0 \leq \gamma$ such that $\gamma\Omega \in \mathbb{X}$ then*

$$V(x) := \max_{1 \leq i \leq n_h} (H_i x) \quad (3.15)$$

is a control Lyapunov function ensuring contraction rate λ .

3.3.3 Inner approximation of shape-independent control λ -contractive sets for TS systems

The above shape-independent sets need choosing a particular controller parametrisation $u(x, \mu)$ in order to be computable with available computational geometry software such as Multi-Parametric Toolbox (MPT) (Herceg et al. 2013). This is the topic of this section.

The simplest approximation is choosing u not depending on memberships. Indeed, let us consider:

$$\mathcal{Q}_{si}^0(\Omega) := \{x \in \mathbb{X} \mid \exists u \in \mathbb{U} : A_i x + B_i u \in \Omega \quad \forall i = 1 \dots r\} \quad (3.16)$$

The above expression comes from plugging a membership-independent $u(x, \mu) := u(x)$ into (3.12) and considering that $\sum_{i=1}^r \mu_i (A_i x + B_i u) \in \Omega$ if and only if $A_i x + B_i u \in \Omega$ for all i . Obviously, $\mathcal{Q}_{si}^0(\Omega) \subset \mathcal{Q}_{si}(\Omega)$.

In fact, the set $\mathcal{Q}_{si}^0(\Omega)$ is the *robust* one-step set in uncertain polytopic systems literature (Pluymers et al. 2005): its main drawback is its conservativeness coming from the fact that, for a given state, the control action should be the same for any value of the membership functions, i.e., the resulting controllers will be in the form $u(x)$ and not $u(\mu, x)$.

Given that $\mu_i(x_k)$ are actually known, a clear improvement is using gain-scheduling in μ , i.e., defining a so-called parallel distributed control parametrisation in the form:

$$u(x) = \sum_{j=1}^r \mu_j(x) u_j(x) \quad (3.17)$$

which defines a different “vertex controller” $u_j(x)$ for each model. This well-known formula is, of course, the key idea behind “fuzzy” controllers since the 1990s, building a “complicated” nonlinear controller from “simpler” components $u_j(x)$; in particular, there exist design methods for vertex controllers based on convex optimisation.

The closed-loop system with the parametrisation (3.17) can be written as

$$x_{k+1} = \sum_{i=1}^r \sum_{j=1}^r \mu_i(x_k) \mu_j(x_k) (A_i x_k + B_i u_j(x_k)) \quad (3.18)$$

Let us introduce the augmented notation

$$\bar{u}(x) = \begin{pmatrix} u_1(x) \\ \vdots \\ u_r(x) \end{pmatrix}, \quad E_j = [0_{m \times m} \ 0_{m \times m} \ \dots \ I_{m \times m} \ \dots \ 0_{m \times m}] \quad (3.19)$$

being E_j an $m \times (mr)$ matrix with an identity matrix in the j -th block position, for $1 \leq j \leq r$. In this way, we have $u_j(x_k) = E_j \bar{u}(x_k)$, so the closed-loop system can be written as the augmented-input one:

$$x_{k+1} = \sum_{i=1}^r \sum_{j=1}^r \mu_i(x_k) \mu_j(x_k) (A_i x_k + B_i E_j \bar{u}(x_k)) \quad (3.20)$$

where the new input is a vector of length $r \times m$. In this case, disregarding again the fact that memberships depend on state, the shape-independent one-step set of system (3.20), to be denoted as $\mathcal{Q}_{si}^1(\Omega)$, is readily expressed as:

$$\mathcal{Q}_{si}^1(\Omega) := \left\{ x \in \mathbb{X} \mid \exists \bar{u} \in \mathbb{U}^r, \sum_{i=1}^r \sum_{j=1}^r \mu_i \mu_j (A_i x + B_i E_j \bar{u}) \in \Omega \quad \forall \mu \in \Delta \right\} \quad (3.21)$$

where \bar{u} is understood as a length- r vector whose elements belong to \mathbb{U} ; such elements must be the same for all membership values but might be different for different states; in an analogous way to (3.13), a set $\bar{U}_\Omega(x)$ could be suitably defined, and a continuous $\bar{u}(x) : \mathcal{Q}_{si}^1(\Omega) \mapsto \mathbb{U}^r$ can be proven to exist, thus justifying the chosen parametrisation (3.17), obtained from $\bar{u}(x)$ by reverting back the vertical stacking to a fuzzy summation.

Now, $\mathcal{Q}_{si}^0(\Omega) \subseteq \mathcal{Q}_{si}^1(\Omega) \subset \mathcal{Q}_{si}(\Omega)$ because forcing all u_j to be equal converts (3.21) into the particular case (3.16) and, on the other hand, the parametrisation of the underlying $u(x, \mu)$ in (3.12) is generic, not restricted to being linear as (3.17) postulates. Notation \mathcal{Q}_{si}^1 is used to emphasise that the candidate controller is a polynomial of degree 1 in the memberships.

In order to obtain a computable approximation of $\mathcal{Q}_{si}^1(\Omega)$, a Polya relaxation of (3.21) must be carried out. For instance, denoting $\Xi_{ij}(x, \bar{u}) := A_i x + B_i E_j \bar{u}$

$$\tilde{\mathcal{Q}}_{si}^1(\Omega) := \left\{ x \in \mathbb{X} \mid \exists \bar{u} \in \mathbb{U}^r, \frac{1}{2}(\Xi_{ij}(x, \bar{u}) + \Xi_{ji}(x, \bar{u})) \in \Omega \right\} \quad (3.22)$$

is a Polyhedron that can be guaranteed to be a subset of $\mathcal{Q}_{si}^1(\Omega)$, but more relaxed options do exist, see (Ariño et al. 2017). The iterations on the above approximations to one-step sets, when converge, provide polytopic inner approximations to the maximal contractive sets in which there exist membership-dependent control actions stabilising the TS system. As these polytope manipulations are carried out with the Multi-Parametric Toolbox (MPT), it is shortly reviewed next.

3.4 Polytopes and Multi-Parametric programming (MPT)

The Multi-parametric toolbox (MPT) is a tool whose objective is to provide efficient computational means for a wide variety of complex computational problems, between these, constrained optimal control problems with great results belonging to the lines of interest of this thesis.

3.4.1 Polytopes

Some of the definitions and basics operations with polytopes that involved the use of MPT are presented (Kvasnica et al. 2004).

Definition 3.16. (*Polyhedron*): A polyhedron $Q \subseteq \mathbb{R}^n$ is a convex set given as the intersection of a finite number of hyper-planes and half-spaces (or a convex

combination of a finite number of vertices)

$$S = \{x \in \mathbb{R}^n | Ax \leq b\} \quad (3.23)$$

Definition 3.17. (*Polytope*): A polytope $\mathcal{P} \subset \mathbb{R}^n$ is a bounded (convex) polyhedron

$$\mathcal{P} = \{x \in \mathbb{R}^n | Ax \leq b\} \quad (3.24)$$

From the above definitions, every polytope represents a convex and compact set. A polytope can also be described by its vertices

$$\mathcal{P} = \{x \in \mathbb{R}^n | x = \sum_{i=1}^{v_P} \lambda_i V_P^{(i)}, 0 \leq \lambda_i \leq 1, \sum_{i=1}^{v_P} \lambda_i = 1\} \quad (3.25)$$

where $V_P^{(i)}$ denotes the i -th vertex of \mathcal{P} , and v_P is the total number of vertices of \mathcal{P} .

Definition 3.18. (*face*) Linear inequality $a'x \leq b$ is called valid for a polyhedron \mathcal{P} if $a'x \leq b$ holds for all $x \in \mathcal{P}$. A subset of a polyhedron is called a face of \mathcal{P} if it is represented as

$$\mathcal{F} = \mathcal{P} \cap \{x \in \mathbb{R}^n | a'x = b\},$$

for some valid inequality $a'x \leq b$. The faces of polyhedron \mathcal{P} of dimension 0, 1, $(n - 2)$ and $(n - 1)$ are called vertices, edges, ridges and facets, respectively.

We say that a polytope $P \subset \mathbb{R}^n$, $\mathcal{P} = \{x \in \mathbb{R}^n | Ax \leq b\}$ is in a minimal representation if a removal of any of the rows in $Ax \leq b$ would change it (i.e., there are no redundant halfspaces).

3.4.2 Basic Polytope Manipulation

The Set-difference of two polytopes \mathcal{P}_1 and \mathcal{P}_2 is a union of polytopes $\mathcal{R} = \bigcup_i \mathcal{R}_i$

$$\mathcal{R} = \mathcal{P}_1 \setminus \mathcal{P}_2 := \{x \in \mathbb{R}^n | x \in \mathcal{P}_1, x \notin \mathcal{P}_2\} \quad (3.26)$$

The Minkowski-Addition of two polytopes \mathcal{P} and \mathcal{W} is a polytope

$$\mathcal{P} \oplus \mathcal{W} := \{x + w \in \mathbb{R}^n | x \in \mathcal{P}, w \in \mathcal{W}\} \quad (3.27)$$

The convex hull of a union of polytopes $\mathcal{P}_i \subset \mathbb{R}^n$, $i = \{1, \dots, p\}$, is a polytope

$$\text{hull} \left(\bigcup_{i=1}^p \mathcal{P}_i \right) := \{x \in \mathbb{R}^n | x = \sum_{i=1}^p \lambda_i x_i, x_i \in \mathcal{P}_i, 0 \leq \lambda_i \leq 1, \sum_{i=1}^p \lambda_i = 1\} \quad (3.28)$$

The envelope of two \mathcal{H} -polyhedra $\mathcal{P} = \{x \in \mathbb{R}^n | A^p x \leq b^p\}$ and $\mathcal{Q} = \{x \in \mathbb{R}^n | A^q x \leq b^q\}$ is an \mathcal{H} -polyhedron

$$\text{env}(\mathcal{P}, \mathcal{Q}) = \{x \in \mathbb{R}^n | \bar{A}^p x \leq \bar{n}^p, \bar{A}^q x \leq \bar{b}^q\} \quad (3.29)$$

where $\bar{A}^p x \leq \bar{b}^p$ is the subsystem of $A^p x \leq b^p$ obtained by removing all the inequalities not valid for the polyhedron \mathcal{Q} , and $A^q x \leq b^q$ are defined in the similar way with respect to $\bar{A}^q x \leq \bar{b}^q$ and \mathcal{P} .

3.5 Convex Difference Inclusions

In (Fiacchini 2010b; Alamo et al. 2009; Fiacchini, Alamo, and Camacho 2012; Fiacchini 2010a) an extended modelling framework was presented so that invariant and contractive sets could be computed for a larger class of systems than the polytopic ones. In Chapter 8, a gain-scheduled extension of these ideas will be laid out. Here, we will review the main ideas of the “convex difference inclusion” approach the authors of the above-cited works introduced.

Invariant and contractive sets can be easily determined for linear systems and uncertain linear systems characterised as polytopic Linear difference inclusions (LDI) (Kouramas et al. 2005), either from a set-based approach as discussed in earlier sections of this chapter, or in a linear matrix inequality framework (LMI), (Boyd and Vandenberghe 2004; Guerra and Vermeiren 2004; Sala 2009; Wu and Dong 2006b) discussed in earlier chapters. As many nonlinear systems can be embedded in a polytopic LDI via a systematic modelling procedure, described in the previous chapter of this Thesis, the above results apply to them.

The works (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2010; Fiacchini, Alamo, and Camacho 2012) show that set-based LDI ideas can be generalised to the so-called convex differential inclusions (CDI), which are more general models capable of representing nonlinear and uncertain systems with lower conservatism (overbounding) than LDI, being these a particular case. Furthermore, they present iterative scaling/shooting algorithms to compute contractive sets, both in analysis and *robust* control design settings. In the polytopic case, such steps can be carried out with polytope manipulation software such as the multi-parametric toolbox (MPT, (Herceg et al. 2013)).

Consider a discrete-time dynamic system where the successor of a current state $x \in \mathbb{R}^n$, denoted by x_+ , is given by:

$$x_+ = f(x, d, u) \quad (3.30)$$

being u is the so-called input vector, with values $u \in \mathbb{U} \subset \mathbb{R}^m$, and $d \in \mathbb{D} \subset \mathbb{R}^v$ a set of time varying parameters or disturbances. In many cases an uncertain “*embedding model*” of the dynamics (3.30) can be built so that an inclusion condition can be stated as follows:

$$f(x, d, u) \in \mathcal{F}(x, u) \quad \forall d \in \mathbb{D} \quad \forall x \in \mathbb{X} \quad \forall u \in \mathbb{U} \quad (3.31)$$

where $\mathcal{F}(x, u) \subset \mathbb{R}^n$ is the image of (x, u) under a known set-valued map, $\mathcal{F} : \mathbb{X} \times \mathbb{U} \mapsto \mathcal{P}(\mathbb{R}^n)$ being $\mathcal{P}(\cdot)$ the power set. The set \mathbb{X} will be denoted as the *modelling region*. Under this embedding $\mathcal{F}(x, u)$ is a *set* of possible successor states. The motivation towards using these uncertain embeddings is that they will exchange conservatism for computational advantages, if \mathcal{F} fulfills some convexity-related properties, to be later discussed.

Uncertain dynamic systems in the form:

$$x_+ \in \mathcal{F}(x, u) \quad (3.32)$$

are denoted as difference inclusions (DI) (Fiacchini, Alamo, and Camacho 2010). If there is no input vector $x_+ \in \mathcal{F}(x)$, the DI is called autonomous. Of course, if $\mathcal{F}(x, u)$ were a single point, the above DI would reduce to a deterministic dynamic system.

We will assume that the origin ($x = 0, u = 0$) is an equilibrium point of the undisturbed system, i.e. $f(0, 0) = 0$ and, additionally, that $0 \in \mathbb{X}$, $0 \in \mathbb{U}$ and $0 \in \mathbb{D}$. Also, \mathbb{X} , \mathbb{U} and \mathbb{D} will be assumed to be compact polyhedra.

Let us consider a set-valued map $\mathcal{F} : \mathcal{D} \mapsto \mathcal{P}(\mathbb{R}^n)$; the domain \mathcal{D} will be assumed to be a convex set in the sequel. So, the image under such a map of $x \in \mathcal{D}$ is a set $\mathcal{F}(x) \subset \mathbb{R}^n$. The image of a set $\Omega \subseteq \mathcal{D}$ under a set-valued map \mathcal{F} will be understood as $\mathcal{F}(\Omega) := \cup_{\xi \in \Omega} \mathcal{F}(\xi)$.

Definition 3.19. A set-valued map \mathcal{F} is **convex-valued** if $\mathcal{F}(\xi)$ is convex for every $\xi \in \mathcal{D}$. Additionally, it will be denoted as **polytopic** if $\mathcal{F}(\xi)$ is a compact polytope for all ξ in \mathcal{D} .

Definition 3.20. A set-valued map is **linear** if $\mathcal{F}(\lambda\xi) = \lambda\mathcal{F}(\xi)$ and $\mathcal{F}(\lambda_1\xi_1 + \lambda_2\xi_2) = \lambda_1\mathcal{F}(\xi_1) \oplus \lambda_2\mathcal{F}(\xi_2)$, for all $\lambda, \lambda_1, \lambda_2$ in \mathbb{R} .

Definition 3.21. A set-valued map is **S-convex** if $\mathcal{F}(\lambda\xi_1 + (1-\lambda)\xi_2) \subseteq \lambda\mathcal{F}(\xi_1) \oplus (1-\lambda)\mathcal{F}(\xi_2)$ for all ξ_1, ξ_2 in \mathcal{D} and for all $0 \leq \lambda \leq 1$.

Proposition 3.4 (cf. Prop. 1 in (Fiacchini, Alamo, and Camacho 2010)). A set-valued map \mathcal{F} with $\mathcal{F}(x)$ being convex and compact, is **S-convex** iff $\hat{f}(\eta, x) := \sup_{x \in \mathcal{F}(x)} \eta^T x$ is a convex function of x when η is fixed, for all $\eta \in \mathbb{R}^n$.

Theorem 3.2. *Given a polytope P , and an S -convex set-valued map \mathcal{F} , we have*

$$\mathcal{F}(\xi) \subseteq \Xi := \text{co} \left(\bigcup_{\zeta \in \text{vert}(P)} \mathcal{F}(\zeta) \right), \quad \forall \xi \in P$$

Proof. See the original references by Fiacchini and coworkers; also a more general quasi-convex version will be proven in Chapter 8. \square

Classification of DI. The properties of the set-valued map $\mathcal{F}(x, u)$ associated to a DI $x_+ \in \mathcal{F}(x, u)$, defined in $\mathcal{D} := \mathbb{X} \times \mathbb{U}$, will define different types of DI. Based on the nomenclature in (Kerrigan 2001; Fiacchini, Alamo, and Camacho 2010), a difference inclusion (3.32) will be denoted as linear (LDI), uncertain-convex (UCDI), if \mathcal{F} is linear or S -convex, respectively. Additionally, a convex DI (CDI), as defined in (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2012) will be understood as having an associated S -convex map where $\mathcal{F}(0) = \{0\}$.

3.5.1 Modelling: embedding nonlinear functions in set-valued maps

In literature, embedding a nonlinear system onto a convex combination of linear “vertex” models is a widely-used approach to prove properties of nonlinear systems based on the properties of the linear vertices, see the quasi-LPV or Takagi-Sugeno modelling approaches in (Tanaka and Wang 2004; Rugh and Shamma 2000a; Kwiatkowski and Werner 2008a), for instance. In particular, given a model in the form:

$$x_+ = [A(x) \quad G(x) \quad E(x)] \begin{pmatrix} x \\ u \\ d \end{pmatrix} := \Psi(x) \begin{pmatrix} x \\ u \\ d \end{pmatrix} \quad (3.33)$$

the above references show that a polyhedral bound for Ψ in a compact modelling region \mathbb{X} can be found under mild continuity assumptions. Let us denote such polyhedral bound as $\hat{\Psi} := \text{Co}(\{(A_i \ G_i \ E_i) \mid i = 1, \dots, N_\Psi\})$, being N_Ψ the number of vertices, such that $\Psi(x) \in \hat{\Psi}$ for all $x \in \mathbb{X}$. Elementary convexity argumentations prove that the above nonlinear system can be rewritten as a parameter-dependent expression:

$$x_+ = \sum_{i=1}^{N_\Psi} h_i(x) (A_i x + G_i u + E_i d) \quad (3.34)$$

being h a parameter vector ranging in the standard simplex $\Delta := \{h \in \mathbb{R}^{N_\Psi} : \sum_{i=1}^{N_\Psi} h_i = 1, h_i \geq 0\}$. Abusing the notation, in the sequel Δ will denote the

standard simplex in any vector space, not just \mathbb{R}^{N_Ψ} . Disregarding the actual “shape” of $h(x)$, we can obtain an embedding in DI form (3.31):

$$x_+ \in Co(\{A_i x + G_i u\} \oplus E_i \mathbb{D}, i = 1, \dots, N_\Psi) \quad (3.35)$$

It can be proved that the above DI is linear (LDI) if $\mathbb{D} := \{0\}$ (undisturbed case) and UC DI in the case D is a compact convex set.

There are other more general techniques to embed a nonlinear system onto a CDI, UC DI, QC DI, \dots , see (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2010).

LDI can be built for multiple-argument functions in several ways. The topic is well studied, so the reader is referred to Chapter 5 (Robles et al. 2017) for contributions on “optimal” modelling, the previous chapter of this Thesis and references therein for ample detail.

This section will outline methodologies for embedding a (non-uncertain) nonlinearity $f(x)$ inside a set-valued map $\mathcal{F}(x)$ with a CDI structure. Parametric/additive uncertainty elements can be added on top of that, with minor modifications, see sections 2.1.1.2 and 2.1.1.3 in (Fiacchini 2010a), as well as some other classes of uncertainty such as ellipsoidal (section 2.2 in the cited work).

Let us consider the case where a function $\rho(x) : \Omega \mapsto \mathbb{R}$, with $\Omega \subset \mathbb{R}^n$, can be bounded by known f_l, f_u such that $f_l(x) \leq \rho(x) \leq f_u(x)$. If the bounds fulfill some convexity-related properties, the associated set-valued maps will accordingly do.

Proposition 3.5. (Fiacchini 2010a, Property 3.3) *The set-valued map defined by $\mathcal{F}(x) = [f_l(x), f_u(x)]$ where $f_l(x) \leq f_u(x)$ and f_l is concave and f_u is convex is S -convex.*

For instance, the 1st-order dynamic system $x_+ = x^3$ can be embedded in $x_+ \in [f_l(x), f_u(x)]$ with $f_l(x) = \min(x^3, 0)$ and $f_u(x) = \max(x^3, 0)$. As f_l is concave and f_u is convex, and $f_l(0) = f_u(0) = 0$ the embedding is a CDI.

Proposition 3.6. (Fiacchini 2010a, Property 3.18) *Assume that $\rho(x)$ can be expressed as $\rho(x) = g(x) - h(x)$ being g and h convex functions (the so-called difference-of-convex functions (Fiacchini 2010a)). Let $\xi_g(x)$ be any linear function such that $\xi_g(x) \leq g(x)$ and let $\xi_h(x)$ be any linear function such that $\xi_h(x) \leq h(x)$. Then, under the above assumption, the set-valued map $\mathcal{F}(x) = [\xi_g(x) - h(x), g(x) - \xi_h(x)]$ is S -convex.*

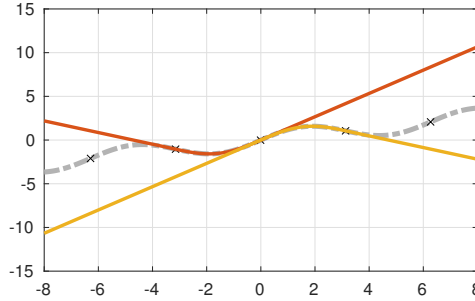


Figure 3.2: A convex(red)-concave(yellow) bound inducing a S-convex set-valued map embedding $\sin(x) + x/3$.

Proof. Indeed, note that ξ_g and ξ_h exist by supporting hyperplane results. Then, $\rho = g - h \geq \xi_g - h$ (which is a concave function) and $\rho = g - h \leq g - \xi_h$ (which is convex). \square

There are other ways to obtain CDI/UCDI bounds for nonlinear systems, the reader is referred to Fiacchini 2010a and other related works of this author for details.

Note that, as in the well-developed linear LDI case, the result of an S-convex embedding is not unique, and some resulting maps will give different performance and contractive set (domain of attraction) estimates. This is, indeed, a well-known fact (Sala 2009).

Stability

In the case of CDIs (no uncertainty at the origin in the model), stability can be proved if a λ -contractive set is found.

Theorem 3.3. *If there exists a control λ -contractive set Ξ for a CDI $x_+ \in \mathcal{F}(x, u)$, then there exists a Lyapunov function $V(x)$, $V(x) > 0$ for $x \neq 0$, which is homogeneous, i.e., $V(\kappa x) = \kappa V(x)$, and a control law $\tilde{u}(x)$ such that $V(x_+) \leq \lambda V(x)$ for all $x \in \Xi$, for all $x_+ \in \mathcal{F}(x, \tilde{u}(x))$.*

Proof. See original references. A more general version will be proved in Chapter 8. \square

3.5.2 Contractive set computation by shooting algorithms

Theorem 3.4. *Assume that $x_+ \in \mathcal{F}(x, u)$ is a UCIDI. Given a known “target” polytope Ω , assume that a known polytope Ξ has been proved to fulfill $\Xi \subseteq \mathcal{Q}(\Omega)$. Given a point $\hat{x} \notin \Xi$, if there exists $\hat{u} \in \mathbb{U}$ such that $\mathcal{F}(\hat{x}, \hat{u}) \subseteq \Omega$, then $Co(\Xi \cup \{\hat{x}\}) \subseteq \mathcal{Q}(\Omega)$.*

Proof. See original references. A more general version will be proved in Chapter 8. \square

The above results can be used to enlarge an initial estimate Ξ of $\mathcal{Q}(\Omega)$ by drawing a collection of new \hat{x} in random directions, via shooting algorithms (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2012).

In order to implement algorithms from Theorem 3.4, a polynomial-time sufficient condition to check, given a target set Ω and \hat{x} , if there exists \hat{u} such that $\mathcal{F}(\hat{x}, \hat{u}) \in \Omega$ is needed. In the case (3.35), such problem amounts to solving a set of linear programming feasibility conditions. The same case would hold if $\mathcal{F}(x, u)$ were affine in u , i.e., in the form $\mathcal{F}(x, u) = \mathcal{F}_1(x) \oplus \mathcal{F}_2(x)u$ for some set-valued maps \mathcal{F}_1 and \mathcal{F}_2 . In a general case, if the support function of $\mathcal{F}(x, u)$ is convex in u , computationally viable conditions can be cast.

As a conclusion of Theorem 3.2, the shooting algorithm in (Fiacchini, Alamo, and Camacho 2010) can obtain progressively larger λ -contractive sets if one of them is available, with the following algorithm:

[contractive set expansion]

1. Obtain an initial λ -contractive polyhedral “seed” set Ω_0 overbounding with an LDI and using set-based LDI approaches ((Fiacchini, Alamo, and Camacho 2010, §4.1)) . Set $k = 0$.
2. Choose a random direction vector ρ .
3. Determine, by bisection, the largest scaling $\gamma > 0$ such that there exists \hat{u} such that $\mathcal{F}(\gamma\rho, \hat{u}) \in \lambda\Omega_k$.
4. Set $\Omega_{k+1} = Co(\Omega_k \cup \{\gamma\rho\})$, and $k = k + 1$. If $k < k_{max}$, go to step 2.
5. End. The set Ω_k is control λ -contractive.

Part II

Contributions

Chapter 4

Choosing a Takagi-Sugeno model for improved performance

This chapter presents two methodologies intended to address the problem of constructing the most adequate Takagi-Sugeno representation of a nonlinear system with respect to a performance goal: maximum decay rate. Both approaches are based on coordinate transformations: the first one uses the Hessian of the system in an attempt to represent nonlinearities as sum of squares of independent variables; the second one takes into account the direction provided by a “failing” eigenvector with respect to the desired performance acting on the linearization of the system. It is shown that, in some respects, both methodologies offer better solutions than ordinary “blind” TS modelling.

The contents of this chapter appeared in the conference article:

- R. Robles, A. Sala, M. Bernal, and T. Gonzalez (2015). “Choosing a Takagi-Sugeno model for improved performance”. In: *Proceedings of the 2015 IEEE International Conference on Fuzzy Systems*. Istanbul, Turkey, pp. 1–6.

4.1 Introduction

Originally intended for fuzzy logic heuristic applications, Takagi-Sugeno (TS) models (Takagi and Sugeno 1985) have evolved during the last three decades until they became a well-established nonlinear control methodology (Tanaka and Wang 2001; Sala, Guerra, and Babuska 2005; Lendek et al. 2010). For model-based methodologies, the sector nonlinearity approach allows TS models to exactly represent nonlinear systems in a compact subset of the state space (Taniguchi, Tanaka, and Wang 2001) as a convex combination of some vertex models, being the combination coefficients functions of the state. If the number of nonlinearities is high, a sensible fuzzy model simplification strategy might be advisable. A number of methods is available to reduce complexity in a TS model: empirical (Setnes et al. 1998), more or less systematic though depending on the number of inputs (Gegov 2007), and systematic based on the functional principal component analysis (Matía, Marichal, and Jiménez 2014). These issues are, however, out of the scope of this contribution, even if they are advisable as an *a posteriori* step to some of our results.

Once a TS model is available, stability analysis and design conditions for TS systems can often be expressed as linear matrix inequalities (LMIs), see (Wang, Tanaka, and Griffin 1996), which can be efficiently solved via convex optimization techniques (Boyd et al. 1994). Most results pose LMIs in terms of the TS model vertices; such results are called membership-shape independent (Sala 2009) and actually prove stability for a family of linear time-varying systems in which the plant is embedded. This step introduces quite a lot of conservatism. Recent shape-dependent options (Bernal, Guerra, and Kruszewski 2009; Kruszewski et al. 2009) try to partially avoid such issues.

It is well known that TS representations, even if exact in a compact set of interest $\Omega : \Omega \subset \mathbb{R}^n$, may not be unique (Feng 2006). Thus, a natural question arises: is there a way to decide whether a TS model will be the most appropriate for a particular objective? As basically all shape-independent LMI results for TS systems depend on the vertex models, different choices of supposedly “equivalent” TS representations of a particular nonlinear system may result in different performance levels being proved (because of the shape-independent conservatism).

The objective of this chapter is indagation, in a systematic way, about the causes of the non-uniqueness of the TS models. Given a nonlinear system, a family of TS models will be crafted using as design parameter a coordinate transformation. Based on that idea, some guidelines for choosing the “best” coordinate transformation in order to optimise a decay-rate performance objective will be presented.

This chapter is organized as follows: next section introduces notation, preliminaries and motivates the problem statement; the intermediate section 4.3 explains how a systematic linear coordinate transformation can be made to “decouple” nonlinearities in order to search for the best under some performance requirements; Section 4.4 introduces the first approach towards the work’s objective: it is based on the Hessian of the system; Section 4.5 presents the linearization-based methodology which transforms the original nonlinear system into a narrower representation close to the origin which proves to be better than other approaches to a certain extent; examples are offered all along the chapter to illustrate these techniques.

4.2 Preliminaries

Consider a given nonlinear model

$$\dot{x}(t) = f(x(t)) \quad (4.1)$$

with $x(t) \in \mathbb{R}^n$ being the state and $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ being a continuously differentiable nonlinear vector-valued function. Without loss of generality, the origin $x = 0$ will be assumed to be the equilibrium point and Ω a region of interest containing the origin.

Usually, the sector nonlinearity methodology (Taniguchi, Tanaka, and Wang 2001) to obtain a TS model starts by grouping all the nonlinearities of $f(x)$ (assume there are s of them) in a vector

$$\rho(x) = [\rho_1(x) \quad \rho_2(x) \quad \cdots \quad \rho_s(x)] \in \mathbb{R}^s, \quad (4.2)$$

so that (4.1) can be expressed as:

$$\dot{x}(t) = \hat{f}(x, \rho(x)) \quad (4.3)$$

being $\hat{f}(\cdot, \cdot)$ a linear function.

Then, considering the modelling region Ω each nonlinearity is bounded in two sectors, by expressing:

$$\rho_j(x) = \frac{\rho_j(x)}{\eta_j(x)} \eta_j(x) := \tilde{\rho}_j(x) \eta_j(x), \quad j = 1, \dots, s \quad (4.4)$$

being $\eta_j(x)$ a *linear* function of the state. The choice of η_j is not arbitrary: as the origin must be an equilibrium point and linearity of η_j forces $\eta_j(0) = 0$, the

limit of $\tilde{\rho}_j$ must exist. Then, the bounding:

$$\left(\min_{q \in \Omega} \tilde{\rho}_j(q) \right) \eta_j(x) \leq \rho_j(x) \leq \left(\max_{q \in \Omega} \tilde{\rho}_j(q) \right) \eta_j(x) \quad (4.5)$$

entails that ρ_j can be expressed as an interpolation between the minimum and maximum value in the above expression, i.e., $\rho_j = w_j * \overline{\tilde{\rho}_j} + (1 - w_j) * \underline{\tilde{\rho}_j}$, where:

$$\overline{\tilde{\rho}_j} = \max_{q \in \Omega} \tilde{\rho}_j(q), \quad \underline{\tilde{\rho}_j} = \min_{q \in \Omega} \tilde{\rho}_j(q) \quad (4.6)$$

Subsequently, a TS model with 2^s vertices is built. A total of $r = 2^s$ membership functions (MFs) $h_i(x)$, $i = 1, \dots, r$, lying on the standard simplex $\Gamma = \{h_i : 0 \leq h_i \leq 1, \sum_{i=1}^r h_i = 1\}$, are constructed and a TS model

$$\dot{x}(t) = \sum_{i=1}^r h_i(x) A_i x(t), \quad (4.7)$$

is crafted so that it is *equivalent* to the nonlinear one in the region Ω . For brevity, membership functions will be grouped in a vector $h(x)$, $h \in \Gamma$.

The basic idea motivating this manuscript is the fact that there might be more than one option to build (4.2) (for instance $\sin(x_1) - \cos(x_1)$ may be considered a single nonlinearity or two of them) and even with the same choice of $\rho(x)$, there might be more than one option when choosing $\eta_j(x)$ to build the “sector” bound (for instance $x_1 x_2$ can be modeled as $[\mu x_{1,min} + (1 - \mu)x_{1,max}]x_2$ or $[\mu x_{2,min} + (1 - \mu)x_{2,max}]x_1$), i.e., there might be many equivalent TS models of the same nonlinear system. Another problem in TS modelling is the fact that the number of vertices grows rapidly with the number of nonlinearities, thus leading to higher complexity in modelling (Matía, Marichal, and Jiménez 2014).

From a philosophical point of view, being (4.7) and (4.1) equivalent expressions, simply rewritten in a different way, for whatever choice of ρ , a non-conservative analysis of (4.7) would yield exactly the same results with any TS model fed to it. However, most stability analysis and control design methodologies in literature are *membership-shape independent*, in the sense that, instead of studying stability of (4.7), stability of the following system is actually studied:

$$\dot{x}(t) = \left(\sum_{i=1}^r h_i(t) \cdot A_i \right) x(t), \quad (h_1(t), \dots, h_r(t)) \in \Gamma \quad (4.8)$$

and stability conditions hold *for all* $h(t) \in \Gamma$, i.e., disregarding the fact that h_i are actually functions of x and, simply, considering (4.8) to be a linear-time-varying

system whose interpolating coefficients may jump arbitrarily in time from one point of the simplex to other (though refinements considering bounded variation rates do exist).

In the most basic interpretation, shape-independent conditions are usually cast as LMI conditions involving only the vertex models A_1, \dots, A_r . So, the properties of the vertex models A_i are key in determining the properties of (4.1) (conservatively, because of shape-independence). However, the parameters of the models A_i heavily depend on the way the vector (4.2) and the linear functions η have been chosen. However, systematic criteria for such choices are not present in prior fuzzy literature to the author's knowledge. This motivates the following problem statement:

In concrete terms, shape-independent conditions consider all the convex hull of vertices A_i as “uncertainty”. Different models will, hence, have different shape and orientation of such “uncertainty” polytope. The overall objective of this chapter is suggesting a “good” TS model, with respect to some criteria regarding how the uncertainty is allocated in different directions.

4.3 Linear Coordinate Transformations

As TS modelling depends on the choice of some linear transformations η_j in (4.4), in this work, linear coordinate transformations $\eta = Tx$ will be considered, with invertible T , and η_i will be considered to be

$$\eta_j = T_j x$$

being T_j the j -th row of matrix T . Dependence on x of the functions η_j has been removed for notational simplicity.

In order to be systematic, given a scalar function $q : \mathbb{R}^n \mapsto \mathbb{R}$, $q(x)$ can be expressed in the η coordinates as $g(\eta) = q(T^{-1}\eta)$. Under these coordinate transformation, the functions ρ in (4.2) will be chosen as:

$$g(\eta) = \sum_{j=1}^n \rho_j(\eta) \tag{4.9}$$

where the choice of ρ is, specifically:

$$\begin{aligned}\rho_1(\eta) &= g(\eta) - g(0, \eta_2, \dots, \eta_n) \\ \rho_2(\eta) &= g(0, \eta_2, \dots, \eta_n) - g(0, 0, \eta_3, \dots, \eta_n) \\ \rho_3(\eta) &= g(0, 0, \eta_3, \dots, \eta_n) - g(0, 0, 0, \eta_4, \dots, \eta_n) \\ &\vdots \\ \rho_n(\eta) &= g(0, \dots, 0, \eta_n).\end{aligned}$$

With such choice of ρ , if g has continuous partial derivatives at the origin, then it is straightforward to see that:

$$\lim_{\eta_j \rightarrow 0} \tilde{\rho}_j(\eta) = \lim_{\eta_j \rightarrow 0} \frac{\rho_j(\eta)}{\eta_j} = \left. \frac{\partial \rho_j}{\partial \eta_j} \right|_{\eta_j=0} \quad (4.10)$$

so the required limit in (4.4) exists, and in fact $\rho_j(\eta)$ and $\tilde{\rho}_j(\eta)$ are functions of only (η_j, \dots, η_n) . So, using (4.6), the associated sector-nonlinearity 2-rule model of ρ_j is:

$$\rho_j(\eta) = \left(w_j \cdot \overline{\tilde{\rho}_j} + (1 - w_j) \cdot \underline{\tilde{\rho}_j} \right) \eta_j \quad (4.11)$$

where w_j is only dependent of (η_j, \dots, η_n) .

In this way, a family of TS models of $q(x)$ is obtained, parameterised by the coordinate transformation T . The above proposal is systematic in the sense that apart from T no other modelling decision is involved.

Now, searching for the best T under some performance requirements can be proposed.

4.4 A Hessian-based methodology

This section is motivated by the idea that near to the origin all smooth nonlinear functions have a quadratic behavior given by the Taylor series up to degree 2. Thereby, it is possible to approximate these functions in a spherical ball of arbitrary radius containing the origin by the Hessian matrix, which describes the local characteristics of a multi-variable function from quadratic elements.

As the Hessian is a symmetric matrix, there exists an orthonormal basis which diagonalises it: hence, all nonlinearities close to the origin can be represented as a sum of squares of single independent variables. So, the fuzzy model of $q(x) = \lambda x^2$, expressed as $\lambda(\mu_1 \underline{x} + (1 - \mu_1) \bar{x}) x$ being \underline{x} and \bar{x} the minimum and

maximum values of x in a modelling region can be generalised to the Hessian coordinates, as follows.

Consider again a nonlinearity $q(x)$, being an smooth and differentiable scalar nonlinear function in Ω , dependent of the state $x \in \mathbb{R}^n$.

By this way, the respective Hessian matrix $H \in \mathbb{R}^{n \times n}$ at the origin is obtained, i.e.

$$H = \frac{1}{2} \left. \frac{\partial^2 q(x)}{\partial x_i \partial x_j} \right|_{x=0}, \quad i, j \in \{1, \dots, n\}. \quad (4.12)$$

which is symmetric.

Consider the set of normalized eigenvectors of H , grouped in so-called eigenvector matrix, $V = [v_1 \ v_2 \ \dots \ v_n] \in \mathbb{R}^{n \times n}$, to define the following coordinate transformation:

$$\eta = Vx \Leftrightarrow x = V^{-1}\eta \quad (4.13)$$

The motivation of the above transformation is the fact that, if $q(x)$ is quadratic, then

$$q(x) = g(\eta) = \lambda_1 \eta_1^2 + \dots + \lambda_n \eta_n^2 \quad (4.14)$$

where $\lambda_1, \dots, \lambda_n$ are the Hessian eigenvalues.

In this way, the functions ρ_j in (4.9) verify:

$$\rho_j(\eta) = \lambda_j \eta_j^2 \quad (4.15)$$

and its TS model would involve using $\tilde{\rho}_j = \lambda_j \eta_j$ in (4.4), and the required vertices would simply be obtained from the maximum and minimum of η_j in the modelling region in consideration.

Example 4.1. Consider the nonlinear function $q(x) = x_1 x_2$ to be modeled for a general purpose; it is assumed that the space of interest $\Omega : \Omega \subset \mathbb{R}^2$ is a circle of radius 1. In order to implement the methodology of the approach, it is necessary to obtain its Hessian transformation (4.12):

$$H = \begin{bmatrix} 0 & 0.5 \\ 0.5 & 0 \end{bmatrix}.$$

Once H is calculated, the associated normalized eigenvector matrix V is obtained such that

$$H = V\Lambda V^{-1},$$

yielding

$$V = \begin{bmatrix} -0.7071 & 0.7071 \\ 0.7071 & 0.7071 \end{bmatrix} \quad \text{and} \quad \Lambda = \begin{bmatrix} -0.5 & 0 \\ 0 & 0.5 \end{bmatrix}.$$

Thus, the system can be rewritten according to (4.13) as

$$\begin{aligned} x_1 &= -0.7071\eta_1 + 0.7071\eta_2 \\ x_2 &= 0.7071\eta_1 + 0.7071\eta_2 \end{aligned}$$

where $q(x)$ can be rewritten in terms of η as

$$g(\eta) = \frac{\eta_2^2}{2} - \frac{\eta_1^2}{2}. \quad (4.16)$$

Note that the resulting function is expressed in quadratic terms and consequently the election of these terms, to perform the nonlinear sector part, is given in a natural way. Following the methodology, (4.16) is treated as shown in (4.9), this is

$$g(\eta) = \tilde{\rho}_1(\eta)\eta_1 + \tilde{\rho}_2(\eta)\eta_2,$$

where $|\tilde{\rho}_i(\eta)| \leq 0.5$, $i = \{1, 2\}$: a corresponding TS representation with suitable membership functions can be thus performed, yielding:

$$q(x) = \sum_{i=1}^2 \sum_{j=1}^2 w_i(\eta_1)w_j(\eta_2)A_{ij}x \quad (4.17)$$

with

$$\begin{aligned} A_{11} &= [0 \quad 0.7071], \quad A_{12} = [-0.7071 \quad 0] \\ A_{21} &= [0 \quad -0.7071], \quad A_{22} = [0.7071 \quad 0] \end{aligned}$$

In order to compare the Hessian based methodology with alternative linear transformations, we considered the set of transformations:

$$T(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad (4.18)$$

where $\alpha \in [-\pi/2, \pi/2]$ and actually, the Hessian corresponds to $\alpha = \pi/4$.

The measure of ‘‘uncertainty’’ around the actual value of $q(x) = x_1x_2$ induced by the shape-independent TS model can be assessed by replacing the actual $w_1(\eta_1)$

and $w_2(\eta_2)$ in (4.17) by arbitrary scalars in the unit simplex, and evaluating:

$$\delta(x) = \max_{w_1, w_2 \in \Gamma} \left| \sum_{i=1}^2 \sum_{j=1}^2 w_i w_j A_{ij} x - q(x) \right| \quad (4.19)$$

as, indeed, $\delta(x)$ only depends on the vertices A_{ij} .

Figure 4.1 depicts the value of uncertainty $\delta(x)$ when x lies in the frontier of the modelling region (unit circle), i.e., $x = x(\beta) = (\cos(\beta), \sin(\beta))$, for $\beta \in [0, 2\pi]$ for the different choices of $T(\alpha)$ in (4.18).

In the referred figure, β appears in the abscissa axis, and $\delta(x(\beta))$ in the ordinate one. The modelling error of the proposed Hessian approach appears in thick solid line, whereas the modelling error of the rest of transformations is presented with thin dotted lines. Clearly, the Hessian has the lowest maximum modelling error: other coordinate transformations have regions below the Hessian and regions above it.

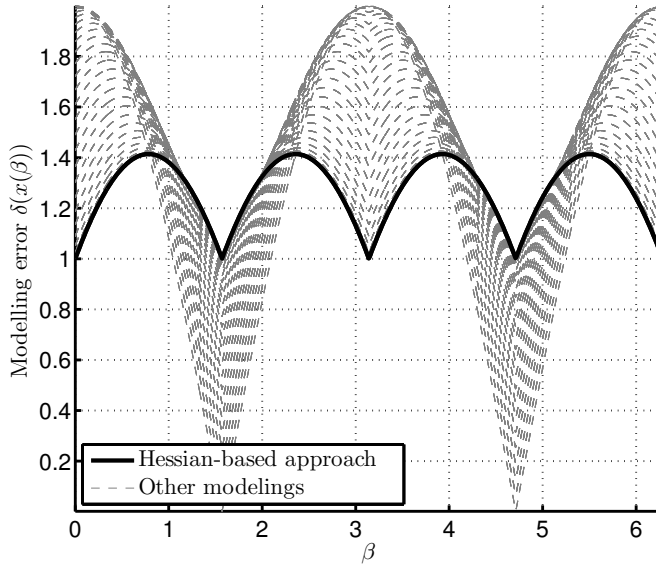


Figure 4.1: Comparison of modelling error

With generic nonlinear functions, close enough to the origin the TS model will be similar to the above case, so there will be an approximate “decoupling” with the proposed Hessian-based transformation, as well as minimising the worst-case uncertainty figure.

In the case we have a vector of several nonlinearities (indeed, in a system's equations (4.1), we have n nonlinear functions to model as $f : \mathbb{R}^n \mapsto \mathbb{R}^n$) we might consider modelling each of it in terms of its respective Hessian transformation in order to provide almost-optimal modelling when close to the origin (in terms of minimum worst-case “error”).

However, the idea has drawbacks due to the need of multiple transformations and the fact that the dynamic performance might depend strongly on the “uncertainty” in some particular directions and not in other ones. This motivates the next section.

4.5 Linearization-based methodology

In this section, alternative modelling approaches will be compared in terms of a goal to meet, such as a performance criterion. For simplicity, decay-rate performance will be chosen for further development.

In particular, we seek a modelling methodology able to keep the linearised system performance when the radius of the modelling region expands (such region will be assumed to be spherical). Indeed, nonlinearity “opens” the TS vertex models as distance from the origin increases and the linearised system belongs to the convex hull of the models. Obviously, the performance provable with the TS model will decrease with the size of the modelling region, and it will be always worse or equal to the linearised one.

The basic idea of this section is considering the directional information from the LMIs arising from the linearised model as coordinate transformation for the TS modelling.

Consider the system in (4.1), and obtain its linearization as:

$$A = \left. \frac{\partial f(x)}{\partial x} \right|_{x=0} \quad (4.20)$$

The maximum decay rate, which we have chosen as performance objective, is obtained by the GEVP below (Boyd et al. 1994):

$$\begin{aligned} \max \quad & \gamma > 0 \\ \text{subject to} \quad & P_0 > 0, \\ & (A^T P_0 + P_0 A) < -2\gamma P_0 \end{aligned} \quad (4.21)$$

with $P_0 = P_0^T$. Then, the state can be bounded by $\|x(t)\| \leq ce^{-\gamma t}$ for some $c > 0$.

Once the LMIs are solved for the optimal γ , compute the Cholesky factor Q_0 such that $P_0 = Q_0^T Q_0$.

Taking Q_o as a matrix transformation such that $\tilde{x} = Q_o x$, the second condition in (4.21) can be rewritten¹ as

$$\Xi = \tilde{A}^T + \tilde{A} + 2\gamma I \leq 0 \quad (4.22)$$

where $\tilde{A} = Q_o A Q_o^{-1}$. If γ is optimal, then Ξ has at least one eigenvalue equal to zero. Consider now the eigenvalue-eigenvector decomposition of Ξ , i.e:

$$\Xi = Q D Q^T. \quad (4.23)$$

where D is diagonal, positive semi-definite, with diagonal elements in ascending order.

Then, generate the matrix transformation

$$T = Q^T Q_o \quad (4.24)$$

This transformation will “position” the nonlinear system (4.1) in the direction of interest based on the analysis of the linearization. The intuitive justification lies in the fact that the first elements of the transformed state will be the ones associated to the smallest (zero if optimal γ) eigenvalues of Ξ .

First, from (4.1) obtain, the “nonlinear” part

$$f_x(x) = f(x) - Ax$$

so the linearisation of $f_x(x)$ is zero.

Then, transform the model into $\dot{x} = Ax + g(\eta)$, with $g(\eta) : \mathbb{R}^n \mapsto \mathbb{R}^n$ defined as $g(\eta) := f_x(x) = f_x(T^{-1}\eta)$, using the transformation matrix T in (4.24). Let us denote each of the elements of $g(\eta)$ as

$$g(\eta) = [g_1(\eta) \quad g_2(\eta) \quad \cdots \quad g_n(\eta)]^T \quad (4.25)$$

In order to apply sector-nonlinearity TS modelling and ensure the existence of the required limits, each of these entries is written as (4.9), i.e.,

$$g_k(\eta) = \sum_{i=1}^n \tilde{\rho}_{ki}(\eta) \cdot \eta_i, \quad (4.26)$$

¹Expression (4.22) for the Lyapunov equation could have been obtained directly via diagonalisation of A . However, generalising the proposal to TS systems with several vertex models far from the origin (currently under research) cannot resort to such diagonalisation, as it would be different for each vertex; hence, the Lyapunov inequality has been presented in the form (4.21) as the basis of development, as it might be better suited for future generalisations.

where ρ_{ki} (and, subsequently, $\tilde{\rho}_{ki}$) are chosen as proposed in (4.9), but individually for each component g_k so the double sub-index appears. In this way we can construct the following equivalent matricial form

$$f_x(x) = g(\eta) = \mathcal{G}(\eta) \cdot \eta = \begin{bmatrix} \tilde{\rho}_{11}(\eta) & \cdots & \tilde{\rho}_{1n}(\eta) \\ \vdots & \ddots & \vdots \\ \tilde{\rho}_{n1}(\eta) & \cdots & \tilde{\rho}_{nn}(\eta) \end{bmatrix} \cdot \eta \quad (4.27)$$

Obtaining the maximum and minimum of each element of the matrix $\mathcal{G}(\eta)$ in a chosen modelling region (details omitted for brevity, as they are well known) the final TS model ensues:

$$\dot{x} = Ax + \mathcal{G}(\eta)\eta, \quad (4.28)$$

where, letting $q = n^2$,

$$\mathcal{G}(\eta) = \sum_{i_1=1}^2 \cdots \sum_{i_q=1}^2 w_{i_1} \dots w_{i_n} G_{i_1 \dots i_q} \quad (4.29)$$

so, transforming back from η to x , a TS model in tensor-product form can be written as:

$$\dot{x} = \sum_{i_1=1}^2 \cdots \sum_{i_q=1}^2 w_{i_1} \dots w_{i_n} (A + G_{i_1 \dots i_q} T) x \quad (4.30)$$

Now, if the original decay-rate optimisation problem for the nonlinear system is formulated with the above proposed TS model –LMIs are well known, replacing A by the above vertices in (4.21)–, the proposed selection of vertices can attain better performance than other options, as the example below discusses.

Example 4.2. Consider the following nonlinear system

$$\dot{x} = \begin{bmatrix} -2x_1 - \sin(0.5x_1x_2) - x_1x_2^2 - x_2 \\ -0.7x_2 - 0.5x_1 \sin(x_1^2) - 0.5x_2 \sin(x_2^2) \end{bmatrix} \quad (4.31)$$

whose linearization is given by:

$$A = \begin{bmatrix} -2 & -1 \\ 0 & -0.7 \end{bmatrix},$$

whose eigenvalue of 0.7 will, of course, give rise to a decay rate γ of 0.7 in (4.21). Indeed, after solving the LMIs in (4.21), the optimal Lyapunov matrix and the associated decay rate are achieved for the linearised system:

$$P_0 = \begin{bmatrix} 1.609 & 1.237 \\ 1.237 & 3.5136 \end{bmatrix}, \quad \gamma = 0.7.$$

The resulting Cholesky factorisation of P_0 is

$$Q_o = \begin{bmatrix} 1.2686 & 0.9756 \\ 0 & 1.6 \end{bmatrix}$$

from which diagonalisation (4.23) is:

$$Q = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, D = \begin{bmatrix} -2.6 & 0 \\ 0 & 0 \end{bmatrix}.$$

Now we build the matrix transformation as $T = Q^T Q_o$, i.e.,

$$T = \begin{bmatrix} -1.2686 & -0.9757 \\ 0 & -1.6 \end{bmatrix}. \quad (4.32)$$

Following the methodology, once we have T , we proceed to extract the nonlinear part of the system in (4.31) and apply $\eta = Tx$ to obtain the TS model (4.30), with 16 rules.

Comparative analysis with alternative TS models The proposed model is compared with other more “classical” ways of obtaining the sector-nonlinearity TS model, such as the four equivalent expressions of (4.31) below:

$$\dot{x} = \begin{bmatrix} -2 - \frac{\sin(0.5x_1x_2) - x_1x_2^2}{0.5 \sin(x_1^2)} & -1 \\ 0.5 \sin(x_1^2) & -0.7 - 0.5 \sin(x_2^2) \end{bmatrix} x \quad (4.33)$$

$$\dot{x} = \begin{bmatrix} -2 & -\frac{\sin(0.5x_1x_2)}{x_2} - x_1x_2 - 1 \\ 0.5 \sin(x_1^2) & -0.7 - 0.5 \sin(x_2^2) \end{bmatrix} x \quad (4.34)$$

$$\dot{x} = \begin{bmatrix} -2 - \frac{\sin(0.5x_1x_2)}{0.5 \sin(x_1^2)} & -x_1x_2 - 1 \\ 0.5 \sin(x_1^2) & -0.7 - 0.5 \sin(x_2^2) \end{bmatrix} x \quad (4.35)$$

$$\dot{x} = \begin{bmatrix} -2 - x_2^2 & -\frac{\sin(0.5x_1x_2)}{x_2} - 1 \\ 0.5 \sin(x_1^2) & -0.7 - 0.5 \sin(x_2^2) \end{bmatrix} x \quad (4.36)$$

and obtaining the maximum and minimum of each of the elements of the above matrices to get the corresponding TS representation, with 8 rules in (4.33) and (4.34), and 16 rules in the last two cases.

Figure 4.2 depicts the maximum decay-rate performance achieved by the proposed Cholesky-based approach and the other TS modelling options as the radius of a proposed circular modelling region increases: the idea of emphasising accuracy in the key direction provided by the linearised LMIs proves fruitful and performance

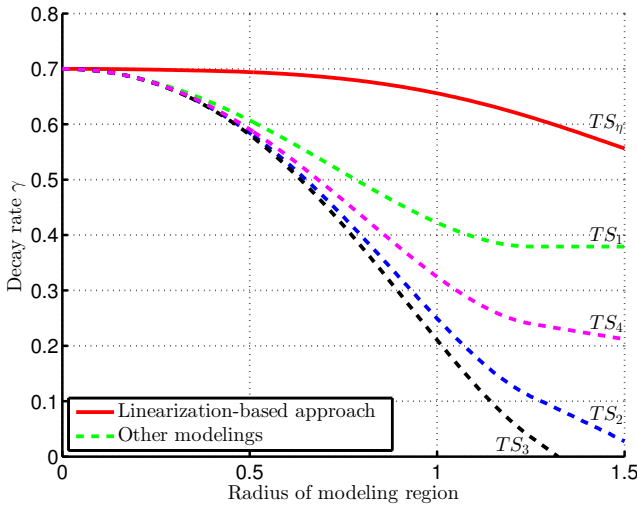


Figure 4.2: Performance comparison

keeps better even at a significant distance from the origin with the proposed coordinates.

In Figure 4.2 the modelling with the proposed Cholesky-based approach is presented in a thick solid line referred as TS_η , whereas the models in (4.33,4.34,4.35,4.36) referred as TS_1 to TS_4 , respectively, are presented with thin dotted lines.

4.6 Conclusions

This chapter is a first approach towards bringing into systematic procedures the idea that TS models might be selected so that the generated “uncertainty” is suitably managed.

A Hessian-based approach has been presented, which is performed in a systematic way, and it “decouples” multivariable nonlinearities (up to 2nd degree approximation) providing minimum maximal error.

Hessian might be a good option when modelling a nonlinearity without knowledge on how its error might influence later results; however, its complexity lies in the fact that multiple Hessians might be needed in a systems with multiple nonlinearities thereby increasing the complexity and number of linear models of the resulting Takagi-Sugeno model.

In the case where it is required to keep a performance measure as close as possible to the linearised case, the proposed linearisation-based approach in section 4.5 is an advantageous alternative over “blind” methodologies since, in a systematic way, the system’s “orientation”, through a linear transformation, is sought so that it is the most beneficial to the performance measure. Obviously, increased performance improvements due to *nonlinear* transformations might be achievable; they are left as a matter of future research.

Chapter 5

Subspace-Based Modelling for Improved LMI Performance

Given a nonlinear system, the sector-nonlinearity methodology provides a systematic way of transforming it in an equivalent Takagi-Sugeno model. However, such transformation is not unique: conservatism of shape-independent performance conditions in the form of linear matrix inequalities results in some models yielding better results than others. This chapter provides some guidelines on choosing a sector-nonlinearity Takagi-Sugeno model, with provable optimality (in a particular sense) in the case of quadratic nonlinearities. The approach is based on Hessian and restrictions of a function onto a subspace.

This chapter extends and improves the results of conference work presented in the Chapter 4. The contents of this chapter appeared in the journal article:

- R. Robles, A. Sala, M. Bernal, and T. Gonzalez (2017). “Subspace-Based Takagi-Sugeno Modeling for Improved LMI Performance”. In: *IEEE Transactions on Fuzzy Systems* 25.4, pp. 754–767.

5.1 Introduction

Analysis and design of nonlinear control systems via Takagi-Sugeno (TS) models is well developed, evolving from model-free heuristics (Takagi and Sugeno 1985; Wang 1997) to model-based exact representations, combined with the direct Lyapunov method in order to obtain linear matrix inequalities (LMIs) (Tanaka and Wang 2001; Sala, Guerra, and Babuska 2005; Lendek et al. 2010). The latter case is based on the *sector nonlinearity* approach, obtaining an exact TS model via maximum and minimum bounds of a nonlinearity in a compact modelling region. Approximate TS models can, too, be obtained, via linearisation at several points (Tanaka and Wang 2001), or based on approximate fitting via \mathcal{H}_2 or SVD argumentations (linear or polynomial in the scheduling parameters (Pettersson and Löfberg 2009), tensor-product summation (Nagy et al. 2009)); SVD-based techniques for rule reduction of complex TS systems appear in (Escaño and Bordons 2014). However, these “approximate” TS models are intentionally left out of the scope of this chapter, concentrating on presenting improvements to the exact sector-nonlinearity technique. Polynomial-fuzzy models (Sala and Ario 2009) will also not be considered in the present work.

Although the models are exact, a first drawback comes from the conservatism of considering only the vertices and not the combination coefficients in the stability conditions. Thus, stability is actually proven for a family of linear time-varying (LTV) systems in which the plant is embedded; these results are therefore called *shape-independent* (Sala 2009). Shape independency is the easiest way to get conditions in the form of LMIs, which in turn are advantageous because they belong to the class of convex optimization problems, which are efficiently solved¹ (Boyd et al. 1994). A few shape-dependent options are available (Bernal, Guerra, and Kruszewski 2009), not considered here.

Another drawback of the TS/LMI methodology is the well-known fact that TS representations may not be unique (Feng 2006). Hence, different performance levels can be proven with shape-independent LMIs for the same nonlinear system, depending on the chosen TS model. To handle this issue, apart from naive trial-and-error, no systematic procedure of choosing a “good” TS model from the many options (infinitely many, actually) is available in literature, to the authors’ knowledge.

¹In some cases, performance optimisations are in generalised-eigenvalue problem form (GEVP); bisection plus LMI is a well-known option: this chapter will understand “LMI” as the convex or quasi-convex problems involving matrix inequalities, including GEVP, discussed in (Boyd et al. 1994).

From the above discussion, the objective of this chapter is choosing an appropriate TS model in order to maximise a performance objective in regions close to the origin. The approach is based in first and second partial derivatives (Jacobian and Hessian). As shape-independent conditions consider all the convex hull of vertices A_i as polytopic uncertainty, different TS models will, hence, have different shape and orientation of such uncertainty polytope. The key idea to be presented is making the intersection of such polytope with some vector subspaces (appearing in performance-related LMIs) as small as possible. A preliminary approach appears in Chapter 4 (Robles et al. 2015).

This chapter is organized as follows: section 5.2 introduces preliminaries and motivates the problem; section 5.3 presents shape-independent uncertainty measures; section 5.4 shows how a Hessian based transformation can optimise such measures; restrictions onto a subspace are discussed in section 5.5; consequences in an LMI context appear in section 5.6. Discussion and examples, are sections 5.7 and 5.8, respectively.

5.2 Preliminaries

Consider a nonlinear dynamic system in the form

$$\dot{x}(t) = f(x(t)) \quad (5.1)$$

with $f : \mathbb{R}^n \mapsto \mathbb{R}^n$, having continuous second derivatives and $f(0) = 0$. Consider the linearised model of (5.1) to be:

$$\dot{x} = Ax, \quad A := \left. \frac{\partial f(x)}{\partial x} \right|_{x=0} \quad (5.2)$$

Takagi-Sugeno modelling

The well-known sector nonlinearity methodology (Taniguchi, Tanaka, and Wang 2001) allows algebraically rewriting (5.1) as an *equivalent* convex sum of linear models

$$\dot{x} = \sum_{i=1}^r h_i(x) A_i x, \quad (5.3)$$

where the membership functions (MFs), grouped in a vector $h \in \mathbb{R}^r$, belong to the $r - 1$ -dimensional standard simplex:

$$\Delta := \{h \in \mathbb{R}^r : \sum_{i=1}^r h_i = 1, h_i \geq 0 \forall i\} \quad (5.4)$$

Basically, each $A_i \in \mathbb{R}^{n \times n}$ is the matrix corresponding to a particular combination of maxima/minima of previously defined nonlinearities in a compact set of the state space Ω . Although the methodology is well known, a brief outline, introducing some notation needed later on, appears in Appendix, justifying that the number of rules in (5.3) is a power of two, see (5.72). For later developments, let us denote as \mathcal{A} the ordered list of consequents matrices $\mathcal{A} := \{A_1, \dots, A_r\}$. The sector-nonlinearity technique may not produce a unique TS model, resulting in possible conservatism (Sala 2009). It is also well known that the linearised matrix A can be cast as a convex combination of the vertices \mathcal{A} of any TS model of (5.1).

From the developments in the Appendix, the actual structure of h_i coming from sector-nonlinearity TS model of $f : \mathbb{R}^n \mapsto \mathbb{R}$ is either in the form (5.71), as the sum of s two-rule models:

$$f(x) = \sum_{i=1}^s \sum_{j=0}^1 \mu_{ij}(x) A_{ij} x \quad (5.5)$$

where the MFs belong to:

$$\Delta_s = \{\mu_{ij} : \mu_{i0} = 1 - \mu_{i1}, \mu_{ij} \geq 0, i=1, \dots, s, j=0, 1\}$$

or as a tensor-product (5.72), where h_i in (5.3) has the structure:

$$h_i(x) = \prod_{l=1}^s \mu_{l \text{ bit}_l(i)}(x) \quad (5.6)$$

The reader is referred to the Appendix for detailed definition and obtention of the above expressions.

Performance measures

Once a TS model is obtained, analysis and design can be done taking advantage of its convex structure combining them with Lyapunov functions such as $V = x^T P x$, $P = P^T > 0$, which naturally leads to conditions in the form of LMIs in P . A generic assumption on the problem structure will be made:

Assumption 5.1. *The pursued control objective is the minimisation of a performance measure γ subject to some model-independent matrix-definiteness constraints*

$$\Psi(D) \succ 0 \quad (5.7)$$

and model-dependent constraints:

$$x^T \Upsilon(A_i, D, \gamma) x \geq 0 \quad \forall x \neq 0, \forall i = 1, \dots, r \quad (5.8)$$

where D denotes the decision variables (Lyapunov function, controller gains, etc.) and all $A_i \in \mathcal{A}$ are given by the TS model under consideration. Matrix expression $\Upsilon(\cdot, \cdot, \cdot)$ will be assumed symmetric, convex in its first argument, and linear in the third argument. We will assume, too, that Ψ and Υ can be transformed to tractable problems such as, for instance, LMI, so that suitable convex optimisation software will find the optimal γ and D .

Many contributions, referred to in the introduction, set up problems which can be expressed as the above assumption (for instance, decay-rate or \mathcal{H}_∞ norm computations for continuous- and discrete-time TS systems, see example section). Note that problem (5.8) is *shape-independent* as memberships h do *not* appear there: in (conservative) *shape-independent* analysis non-uniqueness of sector-nonlinearity models ends up in different performance levels being proven for different (supposedly equivalent) TS models (Sala 2009).

Under the above assumption, linearity in the third argument will force that the optimal solution of (5.8), to be denoted as $\gamma^{opt,TS}$, hits the boundary of the constraint set: there will exist i^* such that $\Upsilon(A_{i^*}, D, \gamma^{opt,TS})$ will be positive *semi-definite* (non-empty nullspace) for all feasible D .

Relation with performance of linearised model

When problem (5.8) is solved with single matrix A , the optimal performance of the linearised model is obtained.

Proposition 5.1. *The optimal performance measure for (5.2), say γ^{opt} , is obtained when there exists D^{opt} such that the above conditions (5.8), particularised to a single matrix A , are*

$$x^T \Upsilon(A, D^{opt}, \gamma^{opt}) x = 0 \quad \forall x \neq 0, x \in \mathcal{C} \quad (5.9)$$

$$x^T \Upsilon(A, D^{opt}, \gamma^{opt}) x > 0 \quad \forall x \neq 0, x \in \mathcal{C}^\perp \quad (5.10)$$

for some vector subspace $\mathcal{C} \subset \mathbb{R}^n$, being \mathcal{C}^\perp its orthogonal complement.

Proof. As $\Upsilon(\cdot, \cdot, \cdot)$ is a symmetric matrix, it has an orthonormal basis of eigenvectors, so when conditions cease to be strictly feasible they will be equal to zero in a subspace (associated to the null eigenvalues of $\Upsilon(A, D^{opt}, \gamma^{opt})$, denoted as \mathcal{C}) and strictly positive in vectors associated to the non-zero eigenvalues of $\Upsilon(A, D^{opt}, \gamma^{opt})$ (which will belong to the orthogonal complement of \mathcal{C} , i.e., \mathcal{C}^\perp). \square

With a suitable change of variable in the original linearised dynamics, there is no loss of generality in assuming that the linearised dynamics yields some constraints which fail when x lies in the canonical q -dimensional subspace²

$$\mathcal{C} = \{x \in \mathbb{R}^n : x = \begin{pmatrix} 0 \\ I \end{pmatrix} \eta, \quad \eta \in \mathbb{R}^q\} \quad (5.11)$$

In these coordinates, the constraints for the linearised system (5.9) and (5.10) can be equivalently written as a single one in the form below, for some matrix Υ_{11} :

$$x^T \Upsilon(A, D^{opt}, \gamma^{opt}) x = x_c^T \begin{pmatrix} \Upsilon_{11} & 0 \\ 0 & 0 \end{pmatrix} x_c \quad (5.12)$$

Proposition 5.2. *The (shape-independent) optimal performance for (5.3) proven with (5.8), $\gamma^{opt,TS}$, is equal or worse than γ^{opt} for the linearised system (5.2) proven by (5.9)–(5.10): $\gamma^{opt,TS} \geq \gamma^{opt}$.*

Proof. Note that the linearised A is in the convex hull of the matrices in the TS consequents \mathcal{A} in exact sector-nonlinearity models. Convexity in the first argument of $\Upsilon(\cdot, \cdot, \cdot)$ entails that for any D such that $\Upsilon(A_i, D, \gamma^{opt,TS}) \geq 0$ we would have $\Upsilon(A, D, \gamma^{opt,TS}) \geq 0$. Evidently, then, the best performance provable with (5.8) will be larger or equal than γ^{opt} from the linearised model (5.9)–(5.10). \square

5.2.1 Other preliminary results

Through this section, some other results/notation will be used.

Proposition 5.3. *For any two vectors ψ, x in \mathbb{R}^n ,*

$$\max_{\|x\| \leq 1} \psi^T x = \|\psi\|, \quad \min_{\|x\| \leq 1} \psi^T x = -\|\psi\|.$$

Proof is trivial from scalar product properties.

Given $M \in \mathbb{R}^{m \times n}$, the *Frobenius norm* of M is defined as $\|M\|_F := \sqrt{\sum_{i=1}^m \sum_{j=1}^n m_{ij}^2}$, where m_{ij} denotes the element of M at row i and column j ; it verifies $\|M\|_F = \sqrt{\sum_{i=1}^{\min\{m,n\}} \lambda_i^2}$, where λ_i are the *singular* values of M (Meyer 2000). If M is square and symmetric, λ_i are, actually, its eigenvalues.

²Indeed, if Υ fail in a subspace \mathcal{C} (in original coordinates), canonical expression (5.11) is obtained by conforming a transformation matrix $x = \bar{T}x_c$, with $x_c = (\xi^T, \eta^T)^T$ as $\bar{T} = (T_{\mathcal{C}^\perp} \quad T_{\mathcal{C}})$, where columns of $T_{\mathcal{C}}$ are a basis of \mathcal{C} , and those of $T_{\mathcal{C}^\perp}$ are a basis of \mathcal{C}^\perp .

Classical interval arithmetic: An interval $\eta = [a, b]$ is a convex subset of the real line, with minimum a and maximum b , $a \leq b$. The sum $[a, b] + [c, d]$ will be defined as $[a+c, b+d]$. The product $x \times [a, b]$ will be defined as $[ax, bx]$ if $x \geq 0$, and as $[bx, ax]$ otherwise. The width of an interval $[a, b]$, with $b \geq a$, will be denoted as $w([a, b]) := b - a$. The absolute value will be defined as $|[a, b]| = \max(|a|, |b|)$. Let us denote $\text{sym}([a, b]) = [-|[a, b]|, |[a, b]|]$. Obviously, $w([a, b]) \leq 2|[a, b]|$, and $[a, b] \subset \text{sym}([a, b])$.

Proposition 5.4. *Let $\eta = \sum_{i=1}^n x_i \times [a_i, b_i]$ be an interval. Then,*

$$\max_{\|x\| \leq 1} w(\eta) \leq 2\sqrt{\sum_{i=1}^n |[a_i, b_i]|^2}$$

Proof. Obviously, $\eta \subset \sum_{i=1}^n x_i \times \text{sym}([a_i, b_i])$. Elementary manipulations and Proposition 5.3 yield the required result. \square

Second-order approximation. Smooth functions $f(x)$ around the origin based on Taylor series can be seen as:

$$f(x) \approx Jx + \frac{1}{2}x^T Hx,$$

where J stands for the Jacobian while H stands for the Hessian matrix, evaluated at the origin:

$$H := \left. \frac{\partial^2 \rho(x)}{\partial x_i \partial x_j} \right|_{x=0}, \quad i, j \in \{1, \dots, n\}.$$

Since the Hessian is a symmetric matrix, there exists an orthonormal basis which diagonalises it; this implies that each nonlinearity close to the origin can be represented as a sum of squares of single independent variables. If $H = V^T \Lambda V$, the transformation $\eta = Vx$ can express

$$x^T Hx = g(\eta) = \lambda_1 \eta_1^2 + \dots + \lambda_n \eta_n^2, \quad (5.13)$$

where $\lambda_1, \dots, \lambda_n$ are the Hessian eigenvalues.

5.2.2 Motivation

The basic idea arising from the above propositions is that, as TS vertex matrices A_i drift away from the linearisation A , then, the “closer” the matrices A_i could be made to the such linearisation, the better the obtained proven performance $\gamma^{\text{opt}, TS}$ might be. As the worst-case directions are those in subspace \mathcal{C} , the goal of

the TS modelling will be fitting “as closely as possible” the model in the subspace \mathcal{C} . This idea motivates this manuscript.

Given the non-uniqueness of the TS modelling, the first objective of this chapter is providing a systematic methodology to build an infinite family of possible TS models based on some coordinate transformations, from a set of functions $\eta_i(x)$, see (5.69) in Appendix.

Then, the next objective will be defining what the above-mentioned fit means in formal terms, proposing a choice of the aforementioned $\eta_i(x)$ derived from the equations which define subspace \mathcal{C} , proving optimality for quadratic nonlinearities. Given that all smooth functions are locally quadratic by the Taylor series up to degree 2 when close enough to the origin, the proposal in this chapter allow to find the optimal TS model in the above settings in small enough modelling regions.

Examples will show that the proposed TS models preserve performance of the linearised model (the best one, from Proposition 5.2) in a better way than other arbitrary choices as the modelling region increases.

5.3 Systematic TS modelling

Let us consider a TS model in box form (5.5) with a suitably ordered consequent list \mathcal{A} .

Definition 5.1. *The shape-independent TS model $f^{si}(x, \mathcal{A})$ is defined as a set-valued map $f^{si} : \mathbb{R}^n \mapsto \mathbb{C}(\mathbb{R}^m)$, where $\mathbb{C}(\mathbb{R}^m)$ denotes the convex subsets of \mathbb{R}^m , given by:*

$$f^{si}(x, \mathcal{A}) := \{y \in \mathbb{R} : \exists \mu_{ij} \in \Delta_s \text{ s.t. } y = \sum_{i=1}^s \sum_{j=1}^2 \mu_{ij} A_{ij} x\} \quad (5.14)$$

where \mathcal{A} in the left-hand side has been implicitly used to denote the whole list of consequent models.

With the above definition, the following is evident:

Proposition 5.5. *For any of the possible TS models of a given $f(x)$, evidently, $f(x) \in f^{si}(x, \mathcal{A})$.*

The core step in the sector-nonlinearity methodology deals with single-output nonlinear functions $\rho_i : \mathbb{R}^n \mapsto \mathbb{R}$, conforming a vector ρ such that $f(x) = Ax + M\rho(x)$, see (5.68). Each ρ_i is a nonlinearity with *one* output, to be bounded

between *two* linear functions, see (5.70); therefore, this chapter focuses first on analysing these mappings $\mathbb{R}^n \mapsto \mathbb{R}$ in order to tackle the problem of choosing a TS model for improved performance. Later on, the case of multiple nonlinearities is discussed.

Let us, then, consider a function $f : \mathbb{R}^n \mapsto \mathbb{R}$, with $f(0) = 0$. Due to convexity, $f^{si}(x, \mathcal{A})$ is an interval for the chosen class of one-output functions. The width of such interval will be related to the conservatism of shape-independent developments with the TS model.

As TS consequents are linear, considering properties of the TS model over $\{\|x\| = 1\}$ will be informative enough. This motivates the definition below:

Definition 5.2. *The normalised worst-case width (WCW) of a one-output TS model with consequent list \mathcal{A} is:*

$$\bar{\sigma}(\mathcal{A}) := \max_{x \in \Omega, x \neq 0} \frac{w(f^{si}(x, \mathcal{A}))}{\|x\|} \quad (5.15)$$

Definition 5.3. *A TS model with consequent list \mathcal{A} is WCW-optimal if there is no other choice of consequent matrices with better $\bar{\sigma}(\mathcal{A})$.*

Actually, this chapter will prove that a Hessian-based methodology obtains such optimal TS model if $f(x)$ is quadratic.

Later, the optimality criteria will be recast to finding the model having the lowest uncertainty width in the intersection of modelling region Ω with a particular subspace³ \mathcal{C} .

Definition 5.4. *Given a vector subspace \mathcal{C} , the subspace-constrained WCW of a one-output TS model is defined as:*

$$\bar{\sigma}_{\mathcal{C}}(\mathcal{A}) := \max_{x \in \Omega \cap \mathcal{C}, x \neq 0} \frac{w(f^{si}(x, \mathcal{A}))}{\|x\|}$$

So, the optimal TS model will be redefined to be the one minimising $\bar{\sigma}_{\mathcal{C}}$. Again, in a quadratic case, the solution to the minimum $\bar{\sigma}_{\mathcal{C}}$ will be provided in this chapter.

³Such subspace will actually come from constraints (5.9).

5.3.1 Optimal shape-independent TS model for SISO nonlinearity

Consider a single-variable nonlinearity with $f : \mathbb{R} \mapsto \mathbb{R}$, $f(0) = 0$ and its classical sector-nonlinearity TS model, (5.67) in Appendix, here repeated for convenience:

$$f(x) = h(x)\tilde{f}_0x + (1 - h(x))\tilde{f}_1x \quad (5.16)$$

being $\tilde{f}_0 = \max_{x \in \Omega} \tilde{f}(x)$ and $\tilde{f}_1 = \min_{x \in \Omega} \tilde{f}(x)$, with $\tilde{f}(x) = f(x)/x$. Consider now any other possible consequent models \hat{q}_0 and \hat{q}_1 such that there exists $\hat{h}(x)$ allowing writing

$$f(x) = \hat{h}(x)\hat{q}_0x + (1 - \hat{h}(x))\hat{q}_1x \quad (5.17)$$

Using (5.14), with $s = m = 1$, so f^{si} is an interval, we have:

$$f(x) \in f^{si}(x, \{\hat{q}_0, \hat{q}_1\}) = \begin{cases} [\hat{q}_0x, \hat{q}_1x], & x \leq 0 \\ [\hat{q}_1x, \hat{q}_0x], & x \geq 0. \end{cases}$$

In order for the above \hat{h} to exist, the consequents must verify (proof is straightforward, omitted for brevity):

$$\hat{q}_0 \geq \tilde{f}_0, \quad \hat{q}_1 \leq \tilde{f}_1 \quad (5.18)$$

Then (5.18) translates to:

Proposition 5.6. *The shape-independent TS model (5.16) fulfills*

$$f^{si}(x, \{\tilde{f}_0, \tilde{f}_1\}) \subset f^{si}(x, \{\hat{q}_0, \hat{q}_1\})$$

for any \hat{q}_0, \hat{q}_1 such that $f(x) \in f^{si}(x, \{\hat{q}_0, \hat{q}_1\})$ for all $x \in \Omega$.

Note: on the sequel, $f^{si}(x, \mathcal{A})$ will be shorthanded to $f^{si}(x)$ when no confusion on the consequent parameters arises. In functions of one variable, the only reasonable choice of consequents is that in (5.67), because of the above proposition. Such model $f^{si}(x, \{\tilde{f}_0, \tilde{f}_1\})$ will be, also, shorthanded to $f^{si,opt}(x)$.

The objective of this chapter is generalising the easily provable Proposition 5.6 above to functions of several variables. In order to do that, a reformulation of the sector nonlinearity methodology, altogether with a coordinate transformation will be presented in next section.

5.4 Coordinate transformations

In ordinary TS modelling, as discussed in the previous section and the Appendix, the selection of a particular element of $\rho(x)$ in model (5.68), say $\rho_i(x)$, assumes the existence of a specific linear function of the state $\eta_i(x)$; these selections are usually chosen by “manual inspection”, such that, defining

$$\tilde{\rho}_i(x) := \frac{\rho_i(x)}{\eta_i(x)}, \quad \rho_i(x) = \tilde{\rho}_i(x)\eta_i(x)$$

suitable limits of $\tilde{\rho}_i(x)$ exist, so a 2-rule model of ρ_i can be crafted (see Appendix).

Example 5.1 (Ad-hoc modelling). *Consider $f(x) := -4x_1 + 4x_1x_2$, expressed as $f(x) = -4x_1 + 4\rho(x)$, $\rho(x) := x_1x_2$, to be modelled in the unit circle. Either $\{\tilde{\rho}(x) := x_1, \eta(x) := x_2\}$ or $\{\tilde{\rho}(x) := x_2, \eta(x) := x_1\}$ could be reasonable choices to craft a TS model. These two possible choices for η can be visually found in the aforementioned inspection, leading to:*

- $f(x) = -4x + h_1(x_1)x_2 + h_2(x_1) \cdot (-x_2)$, $h_1 = 0.5(x_1 + 1)$, $h_2 = 1 - h_1$, with the associated shape-independent model $f^{si}(x) = [-x_2, x_2]$, or
- $f(x) = -4x + h_1(x_2)x_1 + h_2(x_2) \cdot (-x_1)$, $h_1 = 0.5(x_2 + 1)$, $h_2 = 1 - h_1$, being $f^{si}(x) = [-x_1, x_1]$.

Introducing generic coordinate transformations, the above TS models can be expressed as a particular case of an infinite family of choices, as the discussed below.

Example 5.2 (i.e., Example 1, continued). *The motivation of this section is that, actually, there are infinitely many other choices for the TS models of $f(x)$ in example 5.1. If we express*

$$f(x) = -4x_1 + \frac{1}{\alpha\beta}\rho_1(x) - \frac{1}{\alpha\beta}\rho_2(x)$$

with $\rho_1(x) = (\alpha x_1 + \beta x_2)^2$ and $\rho_2(x) = (\alpha x_1 - \beta x_2)^2$, we could also think of $\eta_1 = (\alpha x_1 + \beta x_2)$ for the TS model of $\rho_1(x) = \eta_1^2(x)$, and, on the other hand, choose $\eta_2 = (\alpha x_1 - \beta x_2)$ for the TS model of $\rho_2(x) = \eta_2^2(x)$. For notational convenience, let us define $v_1 := (\alpha \ \beta)$, $v_2 := (-\alpha \ \beta)$, so we have $\eta_1 = v_1x$, $\eta_2 = v_2x$.

The resulting TS model, in box form (5.5), would be a four vertex representation:

$$f(x) = ((\mu_{10}A_{10} + \mu_{11}A_{11}) + (\mu_{20}A_{20} + \mu_{21}A_{21}))x$$

where $\mu_{10} + \mu_{11} = 1$, $\mu_{20} + \mu_{21} = 1$, and

$$A_{i0} = (-2 \ 0) + \frac{1}{\alpha\beta} \xi_{i0} v_i, \quad A_{i1} = (-2 \ 0) + \frac{1}{\alpha\beta} \xi_{i1} v_i$$

$\xi_{i0} = \max_{x \in \Omega} \eta_i(x)$, $\xi_{i1} = \min_{x \in \Omega} \eta_i(x)$, details omitted for brevity. Note that the two prior “manually” obtained TS models in Example 5.1 correspond to $(\alpha = 1, \beta = 0)$ or $(\alpha = 0, \beta = 1)$, respectively.

As the number of possible models is infinite, the question of which is the “best” one arises. Using the WCW-optimality criteria in Definition 5.3, in the quadratic case, such best model can be found via eigenvalue decomposition, leading to the main result in this section below.

Theorem 5.1. *Consider a quadratic nonlinearity $f : \mathbb{R}^n \mapsto \mathbb{R}$, $f(x) := x^T M x$, with M symmetric, with an eigenvalue-eigenvector decomposition $M = V \Lambda V^T$ with Λ diagonal and V orthonormal matrices. Consider, too, a modelling region $\Omega := \{x : \|x\| \leq 1\}$. Then, the WCW-optimal TS model is given by expressing*

$$f(x) = \sum_{i=1}^n \lambda_i \rho_i(x) \tag{5.19}$$

being λ_i the eigenvalues of M and $\rho_i(x) = \eta_i^2(x)$, for $\eta_i(x) = V_i^T x$, i.e. η_i being the projection of x over the unit-norm eigenvector V_i^T associated to λ_i . Then, the optimal TS model has the form:

$$f(x) = \sum_{i=1}^n \sum_{j=0}^1 \mu_{ij}(x) \psi_{ij} V_i^T x \tag{5.20}$$

where $\psi_{i0} := \max_{x \in \Omega} \lambda_i \eta_i(x)$, $\psi_{i1} := \min_{x \in \Omega} \lambda_i \eta_i(x)$ and membership functions are:

$$\mu_{i0}(x) := \frac{V_i^T x - \psi_{i1}}{\psi_{i0} - \psi_{i1}}, \quad \mu_{i1}(x) := 1 - \mu_{i0}(x), \quad i = \{1, \dots, n\}.$$

Proof. For symmetric M , $f(x)$ can be expressed as:

$$\begin{aligned} f(x) &= \sum_{i=1}^n \left(M_{ii} x_i^2 + \sum_{j>i}^n 2M_{ij} x_i x_j \right) \\ &= \sum_{i=1}^n \left(M_{ii} x_i + \sum_{j>i}^n 2M_{ij} x_j \right) x_i. \end{aligned} \tag{5.21}$$

The last expression can be equivalently written as

$$f(x) = \sum_{i=1}^n \tilde{\rho}_i(x) x_i, \tag{5.22}$$

with $\tilde{\rho}_i$ defined as:

$$\tilde{\rho}_i(x) := M_{ii} x_i + \sum_{j>i}^n 2M_{ij} x_j. \tag{5.23}$$

Note that each $\tilde{\rho}_i(x)$ is linear in x ; its maximum and minimum over the unit ball, defined as $\underline{\rho}_i, \bar{\rho}_i$ respectively, so $\underline{\rho}_i \leq \tilde{\rho}_i(x) \leq \bar{\rho}_i$, are from Proposition 5.3:

$$\bar{\rho}_i := \sqrt{M_{ii}^2 + 4 \sum_{j>i}^n M_{ij}^2}, \quad \underline{\rho}_i := -\bar{\rho}_i, \quad (5.24)$$

hence the TS model arising from (5.22) will be:

$$f(x) = \sum_{i=1}^n \left(\mu_{i0}(x) \bar{\rho}_i + \mu_{i1}(x) \underline{\rho}_i \right) x_i$$

and the map in (5.14) will be the interval:

$$f^{si}(x, \mathcal{A}) = \left[-\sum_{i=1}^n \bar{\rho}_i |x_i|, \sum_{i=1}^n \bar{\rho}_i |x_i| \right] \quad (5.25)$$

So, from Definition 5.2 we have

$$\bar{\sigma} = \max_{\|x\|=1} \sum_{i=1}^n 2\bar{\rho}_i \cdot |x_i|$$

and such maximum on the unit sphere (Proposition 5.3 again) is given by:

$$\bar{\sigma} = \sqrt{\sum_{i=1}^n 4\bar{\rho}_i^2}$$

which, substituting (5.24), results in:

$$\bar{\sigma}^2 = \sum_{i=1}^n 4 \left(M_{ii}^2 + 4 \sum_{j>i}^n M_{ij}^2 \right) = 4 \sum_{i=1}^n \left(M_{ii}^2 + 2 \sum_{j \neq i}^n M_{ij}^2 \right)$$

which can be bounded as

$$\bar{\sigma} \geq 2 \sqrt{\sum_{i=1}^n \left(M_{ii}^2 + \sum_{j \neq i}^n M_{ij}^2 \right)} \quad (5.26)$$

where the term at the right-hand side of the inequality is, actually, twice the Frobenius-norm, i.e., $\bar{\sigma} \geq 2\|M\|_F$.

The key idea for the theorem is the fact that the above bound is tight if M is diagonal, i.e.,

$$\bar{\sigma} = 2\|M\|_F = 2\sqrt{\sum_{i=1}^n \lambda_i^2}. \quad (5.27)$$

Hence, if the representation of $f(x) = x^T M x$ had been chosen in diagonalised coordinates $x^T V \Lambda V^T x$, $V^T x = \eta$, $f(\eta) = \eta^T \Lambda \eta$, the resulting $\bar{\sigma}$ would have been the lowest possible one. Note that, as M is symmetric, there exists an orthonormal basis of eigenvectors so the transformation $\eta = V^T x$ preserves the norm and, hence, exploration over $\|x\| = 1$ is the same as exploring over $\|\eta\| = 1$. This proves that the diagonalised representation is WCW-optimal. \square

Remark 5.1. *If the modelling region is not the unit ball, a scaling/change of variable should be carried out before obtaining the optimal TS model for a quadratic nonlinearity so that the new modelling region coincides with the one required in the above theorem.*

Remark 5.2. *If the function to be modelled is non-quadratic, the diagonalisation-based approach no longer applies. However, close to $x = 0$, the function may be approximated to:*

$$f(x) \approx Jx + \frac{1}{2}x^T Hx$$

where J is the Jacobian and H is the Hessian at $x = 0$. Hence, the coordinate changes arising from diagonalisation of the Hessian would obtain a TS model guaranteed to be optimal in a small enough sphere around the origin (so that higher-order terms can be neglected).

In later developments, the optimal shape-independent TS model from (5.20) will be denoted as $f^{si,opt}(x)$.

5.5 TS models with optimal performance in a subspace

As discussed earlier, an accurate fit of the shape-independent TS model in the performance-critical subspace \mathcal{C} is often needed. This issue will be now addressed.

5.5.1 Restrictions

Definition 5.5 (Restriction of a function or set-valued map). *The restriction of a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ to a linear q -dimensional ($q < n$) vector subspace \mathcal{C} , will be denoted as $f|_{\mathcal{C}} : \mathcal{C} \mapsto \mathbb{R}$, trivially defined as $f|_{\mathcal{C}}(x) := f(x) \forall x \in \mathcal{C}$. An analogous definition will be assumed for the restriction of a set valued map, such as f^{si} , i.e., $f^{si}|_{\mathcal{C}}(x) := f^{si}(x) \forall x \in \mathcal{C}$.*

Consider now the subspace \mathcal{C} being defined as:

$$\mathcal{C} = \{x \in \mathbb{R}^n : \exists \eta \in \mathbb{R}^q \text{ s.t. } x = H\eta\} \quad (5.28)$$

thus, being H a $n \times q$ matrix mapping from canonical coordinates in \mathbb{R}^q to \mathcal{C} .

Then, abusing the notation, the restriction can be also defined in terms of q -dimensional subspace coordinates η , as

$$\begin{aligned} f|_{\mathcal{C}} : \mathbb{R}^q &\mapsto \mathbb{R} \\ f|_{\mathcal{C}}(\eta) &:= f(H\eta), \quad \eta \in \mathbb{R}^q \end{aligned} \quad (5.29)$$

Given a function and its restriction, a complementary function can be defined fulfilling the following proposition.

Proposition 5.7 (Complementary function). *For any subspace \mathcal{C} , a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ can be decomposed as:*

$$f(x) = f_{-\mathcal{C}}(x) + f|_{\mathcal{C}}(x)$$

where $f_{-\mathcal{C}}(x)$, denoted as complementary function, fulfills $f_{-\mathcal{C}}(x) = 0$ for $x \in \mathcal{C}$.

Proof. It is evident, setting $f_{-\mathcal{C}}(x) := (f(x) - f|_{\mathcal{C}}(x))$. □

Consider an invertible $n \times n$ matrix $\bar{T} := \begin{pmatrix} H_{\xi} & H \end{pmatrix}$, formed by completing H with suitable $n - q$ linearly independent columns, so the following linear transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ is set up:

$$x = T(\xi, \eta) := \bar{T} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (5.30)$$

with $\eta \in \mathbb{R}^q$ the subspace coordinates, and $\xi \in \mathbb{R}^{n-q}$ being dummy complementary coordinates. Then, using the above transformation T between x and (ξ, η) and notation (5.29), we can express:

$$f(x) = f(T(\xi, \eta)) = f^{[1]}(\xi, \eta) + f^{[2]}(\eta) \quad (5.31)$$

being $f^{[1]}$ and $f^{[2]}$ defined as:

$$\begin{aligned} f^{[1]}(\xi, \eta) &:= f(T(\xi, \eta)) - f(T(0, \eta)) = f_{-\mathcal{C}}(x) \\ f^{[2]}(\eta) &:= f(T(0, \eta)) = f|_{\mathcal{C}}(\eta) \end{aligned}$$

If the original function f verifies $f(0) = 0$ then these functions verify $f^{[1]}(0, \eta) = 0$, $f^{[1]}(0, 0) = 0$, $f^{[2]}(0) = 0$.

Definition 5.6 (Restriction of a TS model). *Consider a box-TS model of $f(x)$ given by (5.5). The restriction of such model to a linear subspace \mathcal{C} given in (5.28) results in:*

$$f|_{\mathcal{C}}(\eta) = \sum_{i=1}^s \sum_{j=0}^1 \mu_{ij}(H\eta) A_{ij} H\eta \quad (5.32)$$

The shape-independent TS model $f^{si}(x, \mathcal{A})$ from (5.14) can be also restricted to \mathcal{C} , allowing to prove:

Proposition 5.8. *The restriction of f to \mathcal{C} is contained in the restriction of the shape-independent TS model, i.e.,*

$$f|_{\mathcal{C}}(\eta) \in f^{si}|_{\mathcal{C}}(H\eta, \mathcal{A})$$

Proof is obvious from Definition 5.6 and Proposition 5.5,

Proposition 5.9. *If the subspace \mathcal{C} is one-dimensional, denoting $g(\eta) := f|_{\mathcal{C}}(\eta)$, the optimal TS model of the univariate function $g(\eta) : \mathbb{R} \mapsto \mathbb{R}$ fulfills*

$$g^{si,opt}(\eta) \subseteq f^{si}|_{\mathcal{C}}(H\eta, \mathcal{A}) \quad (5.33)$$

for any choice of consequents \mathcal{A} in the original model.

Proof. Representation (5.32) is a TS model of $f|_{\mathcal{C}}(\eta)$, i.e., of $g(\eta)$. Then, Proposition 5.6 yields the required result. \square

In subspaces whose dimension might not be one, we can assert the following result for quadratic functions, which extends Theorem 5.1 in order to consider restrictions:

Theorem 5.2. *Consider $f(x)$ being a quadratic function. Denote as $g(\eta) := f|_{\mathcal{C}}(\eta)$ the restriction of f to \mathcal{C} , being $g^{si,opt}(\eta)$ the Hessian-based optimal shape-independent TS model of g . Then:*

$$\bar{\sigma}(g^{si,opt}) \leq \bar{\sigma}(f^{si}|_{\mathcal{C}}(\eta, \mathcal{A})) \quad (5.34)$$

Proof. Proof is analogous to Proposition 5.9: as $g(\eta) \subset f^{si}|_{\mathcal{C}}(\eta, \mathcal{A})$, the Hessian-based representation of $g(\eta)$ has the lowest maximum uncertainty width, by Theorem 5.1. \square

Basically, as intuitively expected, Proposition 5.9 and Theorem 5.2 say that it is better (or at least equal) to get a direct TS model on the restriction of a function –left-hand side of (5.33) and (5.34)– than, first, modelling on a larger space and, later, restricting the resulting TS model –right-hand side of the referred inequalities–. This motivates using the decomposition in Proposition 5.7 –equivalently, (5.31)– to obtain WCW-optimal TS models of the restrictions, as detailed in next section.

5.5.2 Global models with optimal performance in \mathcal{C}

Up to now, results in previous section have discussed optimality of certain TS models defined *only* on a subspace \mathcal{C} (indeed, restrictions from Definition 5.5 are meaningless outside \mathcal{C}). However, applications usually require TS modelling in all of \mathbb{R}^n and not only in \mathcal{C} . Expression (5.31) comes handy now.

Consider a quadratic function $f(x) = x^T M x$ and the q -dimensional subspace defined in (5.28). Consider, too, any of the possible linear transformations T and its associated matrix \bar{T} in (5.30), and express f in the new coordinates as:

$$f(x) = (\xi^T \quad \eta^T) F \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (5.35)$$

for $F = \bar{T}^T M \bar{T}$. Using the new coordinates (ξ, η) , the subspace \mathcal{C} becomes the canonical subspace $(0, \eta)$, $\eta \in \mathbb{R}^q$.

Express, then $f(x)$ decomposed in the form (5.31). Given the fact that $f^{[1]}(\eta, 0) = 0$, as $f^{[1]}$ is quadratic we can express, with trivial manipulations of matrix F above:

$$f^{[1]}(\xi, \eta) = (\xi^T \quad \eta^T) \begin{pmatrix} F_{11} & 0 \\ F_{21} & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (5.36)$$

$$f^{[2]}(\eta) = \eta^T F_{22} \eta \quad (5.37)$$

With suitable choices of H and H_ξ when conforming \bar{T} , are assumed F_{11} and F_{22} diagonal, without loss of generality⁴.

⁴Indeed, if it were not the case, consider the diagonalisations $F_{22} = V_\eta^T \Lambda_\eta V_\eta$, $F_{11} = V_\xi^T \Lambda_\xi V_\xi$. Then, the change of variable $\eta^* = V_\eta \eta$, $\xi^* = V_\xi \xi$ would render a representation:

$$f^{[1]}(\xi^*, \eta^*) = (\xi^{*T} \quad \eta^{*T}) \begin{pmatrix} \Lambda_\xi & 0 \\ V_\eta F_{21} V_\xi^T & 0 \end{pmatrix} \begin{pmatrix} \xi^* \\ \eta^* \end{pmatrix}$$

$$f^{[2]}(\eta^*) = \eta^{*T} \Lambda_\eta \eta^*$$

So, replacing a supposed (initial guess) \hat{H}_η by a corrected $H_\eta = V_\eta \hat{H}_\eta$ and \hat{H}_ξ by $H_\xi = V_\xi \hat{H}_\xi$, the requested diagonal form would be obtained.

Denote as $\lambda_i^{[2]}$, $i = 1, \dots, q$ the diagonal elements (i.e., eigenvalues) of F_{22} ; denote as $\lambda_i^{[1]}$, $i = 1, \dots, n-q$ the diagonal elements of F_{11} . Then, in the η, ξ coordinates, we can write

$$f^{[1]}(\xi, \eta) = (\xi^T \quad \eta^T) \begin{pmatrix} F_{11} \\ F_{21} \end{pmatrix} \xi \quad (5.38)$$

Let us denote $f^{[1]}(\xi, \eta) = \mathcal{F}(\xi, \eta)\xi$, being \mathcal{F} the linear function $\mathcal{F} : \mathbb{R}^n \mapsto \mathbb{R}^{n-q}$ multiplying ξ in (5.38), i.e.,

$$\mathcal{F}(\xi, \eta) := (\xi^T \quad \eta^T) \begin{pmatrix} F_{11} \\ F_{21} \end{pmatrix}$$

actually expressed as a $1 \times (n-q)$ row vector, and denote as \mathcal{F}_i its elements, for $1 \leq i \leq n-q$. Then, (5.38) can be expressed as:

$$f^{[1]}(\xi, \eta) = \sum_{i=1}^{n-q} \mathcal{F}_i(\xi, \eta)\xi_i$$

Denoting $\bar{\mathcal{F}}_{i0} := \max_{x \in \Omega} \mathcal{F}_i(T^{-1}x)$, and $\bar{\mathcal{F}}_{i1} := \min_{x \in \Omega} \mathcal{F}_i(T^{-1}x)$, we can express $f^{[1]}$ as the TS model:

$$f^{[1]}(\xi, \eta) = \sum_{i=1}^{n-q} \sum_{j=0}^1 \mu_{ij}(x) \bar{\mathcal{F}}_{ij} \xi_i \quad (5.39)$$

Manipulating $f^{[2]}(\eta) = \sum_{i=1}^q \lambda_i^{[2]} \eta_i^2$, we can express it as:

$$f^{[2]}(\eta) = \sum_{i=1}^q \sum_{j=0}^1 \mu_{ij}(x) \bar{\mathcal{G}}_{ij} \eta_i \quad (5.40)$$

with $\bar{\mathcal{G}}_{i0} := \max_{x \in \Omega} \lambda_i^{[2]} \eta_i$ and $\bar{\mathcal{G}}_{i1} := \min_{x \in \Omega} \lambda_i^{[2]} \eta_i$.

From Theorem 5.2, because of the diagonal representation of $f^{[2]}(\eta)$, the above TS model (5.40) is the WCW-optimal one, in the sense of Definition 5.3, for the restriction $f|_{\mathcal{C}}$ in η coordinates.

Combination of the optimal $f^{[2]}$ with its complementary function $f^{[1]}$ results in the main theorem of this section:

Theorem 5.3. *Given a quadratic function $f(x) = x^T M x$, subspace \mathcal{C} defined in (5.28), and the change of variable (5.30), the TS model*

$$f(x) = f^{[1]}(\xi, \eta) + f^{[2]}(\eta) = \sum_{i=1}^{n-q} \sum_{j=0}^1 \mu_{ij}(\xi, \eta) \bar{\mathcal{F}}_{ij} \xi_i + \sum_{i=1}^q \sum_{j=0}^1 \mu_{ij}(\eta) \bar{\mathcal{G}}_{ij} \eta_i \quad (5.41)$$

is WCW-optimal in subspace \mathcal{C} in the sense of Definition 5.4.

Proof. As the restriction of $f^{[1]}$ onto \mathcal{C} is zero, and so it is the restriction to \mathcal{C} of the shape-independent TS model arising from (5.39), i.e., $f^{[1],si}(0, \eta) = \{0\}$, it is easily seen that:

$$f^{si}|_{\mathcal{C}}(\xi, \eta) = f^{[2],si}|_{\mathcal{C}}(\eta)$$

so optimality in \mathcal{C} is not lost when adding (5.39) and (5.40) as the complementary function (and its TS model) is zero on \mathcal{C} . \square

Note that the diagonal form in F_{11} is, actually, not needed in the proof. However, there is no loss of generality in assuming it: it has been stated as such in the above discussion because, being both F_{11} and F_{22} diagonal, the same matrices F_{11} , F_{21} and F_{22} in (5.36) and (5.37) would be usable for obtaining the optimal model in either the subspace $\xi = 0$ or $\eta = 0$.

Note also that the change of variable should be reverted in ξ_i , η_i , in order to have the TS model (5.41) depending on the original x coordinates instead of the transformed ones (details omitted for brevity).

So, basically the procedure to obtain the optimal model for a subspace \mathcal{C} of a quadratic function would be:

1. Get a basis H of \mathcal{C} , complete it and obtain \bar{T} , F , and subsequently, $f^{[1]}$ and $f^{[2]}$.
2. Obtain the Hessian eigenvectors of $f^{[1]}$ (yielding diagonal F_{11}) and $f^{[2]}$ (yielding diagonal F_{22}).
3. Combine both steps in a single change of variable.
4. [optional] Express the box-TS model (5.41) in original coordinates.

The outline of the procedure is illustrated in example below.

Example 5.3. Let us model the function $f : \mathbb{R}^5 \mapsto \mathbb{R}$ given by $f(x) = x_1^2$ in the 2-dimensional subspace⁵ spanned by $\mathcal{C} = \{x \in \mathbb{R}^5 : x = H\eta, \eta \in \mathbb{R}^2\}$ with

$$H = \begin{pmatrix} 1 & 1 & -1 & 0.5 & 1 \\ 2 & 0 & 2 & -4 & -2 \end{pmatrix}^T.$$

⁵The actual subspace is generated via performance optimisation (5.9) of a 5th-order linearised model. The actual model matrices and LMIs are omitted because they are not relevant for the time being. Full examples will appear on Section 5.8.

In order to avoid scalings so that the unit circle in original coordinates keeps being the unit circle in transformed ones, let us obtain an orthonormal basis of \mathcal{C} (overwriting H with Matlab command $H=\text{orth}(H)$, for instance), as well as an orthonormal basis of its complementary space (Matlab $H_{\perp}=\text{null}(H')$). With these two basis, we can form the change of variable \bar{T} in (5.30), with \bar{T} actually being an orthonormal matrix. So, variables x_1 to x_5 will be mapped to $(\xi_1, \xi_2, \xi_3, \eta_1, \eta_2)$ and the above subspace will be, in the new coordinates:

$$\mathcal{C} = \{(\xi_1, \xi_2, \xi_3, \eta_1, \eta_2) \in \mathbb{R}^5 : \xi_1 = \xi_2 = \xi_3 = 0\}.$$

Let us check the accuracy on \mathcal{C} of several possible TS models:

- a) **Inspection-based coordinates:** The first TS model to be considered is the “inspection” one given by:

$$f(x) = h_1(x_1)A_0x + h_2(x_1)A_1x \quad (5.42)$$

with $h_1(x_1) = 0.5(x_1+1)$, $h_2(x_1) = 1-h_1(x_1)$, $A_1 = -A_0$, and $A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}$. The model yields

$$f^{si}(x) = [-x_1, x_1]. \quad (5.43)$$

Of course, this would end the classical way of TS modelling in prior literature. The issue under discussion is, however, how accurate is such model in subspace \mathcal{C} .

In order to assess the accuracy of (5.42) on \mathcal{C} , let us trivially carry out the change of coordinates T by rewriting f^{si} in (5.43) replacing x_1 by its expression on the new coordinates (η, ξ) arising from matrix \bar{T}

$$\begin{aligned} x_1 = & -0.1379\xi_1 + 0.6081\xi_2 + 0.1379\xi_3 \\ & + 0.3385\eta_1 + 0.6911\eta_2 \end{aligned}$$

As η are subspace coordinates, the restriction $f^{si}|_{\mathcal{C}}$ is

$$f^{si}|_{\mathcal{C}} = [-(0.3385\eta_1 + 0.6911\eta_2), 0.3385\eta_1 + 0.6911\eta_2]$$

for η ranging in the unit ball (\bar{T} is orthonormal), the worst-case width (Definition 5.4) on the subspace \mathcal{C} is given by twice the norm of $\kappa := (0.3385, 0.6911)^T$, by Proposition 5.3 and the fact that all intervals are symmetric. The result for the TS model (5.42) is:

$$\bar{\sigma}_{\mathcal{C}} = 2\|\kappa\| = 2 \times 0.7695 = 1.5391 \quad (5.44)$$

- a) **Non-optimised change of variable:** Now, let us test the ideas in Section 5.5.1, i.e., that carrying out TS modelling after the change of variable will lead to a better model than the one above, where TS modelling was carried before the change $x = T(\eta, \xi)$.

Writing now $f(x) = x^T M x$ being M the matrix with all its entries equal to zero except the (1,1) term (equal to 1), we can express it in the new coordinates as $F = \bar{T}^T M \bar{T}$ (not displayed, to save space). The restriction to \mathcal{C} , in such η coordinates is $f|_{\mathcal{C}}(x) = \eta^T F_{22} \eta$, with:

$$F_{22} = \begin{pmatrix} 0.1146 & 0.2339 \\ 0.2339 & 0.4777 \end{pmatrix} \quad (5.45)$$

where, as expected, $F_{22} = \kappa \cdot \kappa^T$. So, we can express $g(\eta) = f|_{\mathcal{C}}(H\eta) = (0.1146\eta_1 + 0.4679\eta_2)\eta_1 + (0.4777\eta_2)\eta_2$. Computing the norms of $(0.1146, 0.4679)^T$ and $(0, 0.4777)^T$, by Proposition 5.3, the resulting TS model yields on the unit circle a shape-independent interval:

$$g^{si} = [-0.4817, 0.4817] \times \eta_1 + [-0.4777, 0.4777] \times \eta_2 \quad (5.46)$$

It can be shown that, as η range on the unit ball, the worst-case width in this case will be:

$$\bar{\sigma}_{\mathcal{C}} = 2\sqrt{0.4817^2 + 0.4777^2} = 1.357 \quad (5.47)$$

which is lower than that from (5.44), as expected.

- a) **Optimised coordinates:** Obtaining three orthonormal eigenvectors of the top-left 3×3 block of F , arranged in a 3×3 matrix V_1 , as well as two orthonormal eigenvectors of the bottom-right 2×2 block of said F –the F_{22} matrix in (5.45) above–, in a 2×2 matrix V_2 , the matrix $V = \text{blockdiag}(V_1, V_2)$ is the Hessian-based coordinate transform of each of the subspaces so that the resulting TS model is optimal in the sense of Theorem 5.1. The overall coordinate change is $x = \bar{T} \cdot V \cdot (\xi^T \ \eta^T)^T$, yielding the transformed model of x_1^2 as:

$$\begin{aligned} f(x) &= 0.4078 \xi_3^2 + 0.9828 \xi_3 \eta_2 + 0.592 \eta_2^2 \\ &= (0.4078 \xi_3 + 0.9828 \eta_2) \xi_3 + 0.592 \eta_2^2 \end{aligned}$$

which, as $\sqrt{0.4078^2 + 0.9828^2} = 1.064$ would lead to a 4-rule TS model on the unit circle given by:

$$\begin{aligned} f^{si}(x) &= f^{si} \left(\bar{T} \cdot V \cdot (\xi^T \ \eta^T)^T \right) \\ &= [-1.064, 1.064] \times \xi_3 + [-0.592, 0.592] \times \eta_2 \end{aligned} \quad (5.48)$$

The worst-case width on the subspace $\xi_1 = \xi_2 = \xi_3 = 0$ results, in this case:

$$\bar{\sigma}_{\mathcal{C}} = 2 \times 0.592 = 1.184 \quad (5.49)$$

which is, as expected, lower than that from (5.44) and (5.47). In fact, as $f(x)$ was quadratic, Theorem 5.3 states that there is no other linear coordinate change which gives a better figure for $\bar{\sigma}_{\mathcal{C}}$ than that in (5.49).

For illustration, reverting the change of variable by suitable inversion of the transformation matrices, we can write:

$$\begin{aligned} \xi_3 &= (0.6386 \quad -0.5473 \quad 0.1672 \quad 0.4865 \quad -0.1672) x \\ \eta_2 &= (0.7696 \quad 0.4542 \quad -0.1388 \quad -0.4037 \quad 0.1388) x \end{aligned}$$

so, replacing these expressions in (5.48) and multiplying by 1.064 and 0.592, respectively, we can define consequent matrices:

$$\begin{aligned} A_{10} &= (-0.8591 \quad 0.2088 \quad 0.1779 \quad -0.5449 \quad -0.1477) \\ A_{20} &= (-0.1104 \quad 0.4794 \quad -0.0990 \quad 0.3032 \quad 0.0822) \\ A_{11} &= -A_{10}, \quad A_{21} = -A_{20} \end{aligned}$$

such that a final (WCW-optimal in the requested subspace) box model (5.5) with $s = 2$ can be written in the original x coordinates, as an alternative to the naive (5.42).

The above example has shown how a rewriting of x_1^2 has reduced the uncertainty due to shape-independence from 1.54 to 1.18 in a particular subspace \mathcal{C} . These manipulations will be able to improve associated LMI results in fuzzy control, as discussed in next section.

5.6 Use of optimal TS models in LMIs

Let us consider a nonlinear system (5.1), its linearisation A in (5.2) and the nonlinearities $\rho(x)$ in (5.68). The objective of this section is using the previous developments to suggest a TS model which preserves performance of the linearised system proven with some LMIs (arising by suitable transformations of (5.7) and (5.8), if needed) by avoiding larger than necessary uncertainty of f^{si} in key subspaces, given by Proposition 5.1.

5.6.1 Effect of the nonlinearity in the Lyapunov equations

When the optimal solution of a LMI for the linearised system $\dot{x} = Ax$ has been obtained, we are in the situation in (5.9)–(5.10). However, the actual performance proved for a nonlinear system would require replacing the linearised state derivatives $\dot{x} = Ax$ by the nonlinear ones $\dot{x} = Ax + M\rho(x)$. Of course, that would destroy the LMI form as ρ is nonlinear, so the objective is generating a sector-nonlinearity TS model of $M\rho$ with low conservatism which still allows proving good enough performance with LMIs.

As the performance limit in (5.9) and (5.10) is hit for $x \in \mathcal{C}$, that means that the restriction of $\rho(x)$ onto subspace \mathcal{C} must be modelled with precision in order to lose the least possible performance (ideally). So, at first glance, applying Theorem 5.3 to each element of ρ might seem a viable solution and so it is, indeed. However, further improvements may be crafted by considering the structure of matrix M and the obtained Lyapunov function. Such ideas will be detailed next.

First, note that, actually, it is not each component of $\rho(x)$ the magnitude to be precisely modelled. Indeed, let us assume there exists a Lyapunov function whose time-derivative will require, in turn, the use of the state derivatives. Say, such Lyapunov function having the form $V(x) := x^T Px$ will give rise to $\dot{V} = 2x^T P\dot{x}$, so the difference between (a) the linearised system's behaviour $\dot{V}_L := 2x^T PAx$ and (b) that from the nonlinear system $\dot{V}_{NL} := 2x^T P(Ax + M\rho(x))$ is

$$\Xi(x) := \dot{V}_{NL} - \dot{V}_L = 2x^T PM\rho(x) \quad (5.50)$$

So, the actual term whose uncertainty must be small, when x lies in subspace \mathcal{C} , is $\Xi(x)$.

Of course, qualitatively speaking, if each element (nonlinearity) in vector $\rho(x)$ is modelled with a “precise enough” TS system, the overall Ξ will be precise. However, the different intervals of uncertainty in ρ will result in a cumulative uncertainty in $PM\rho$ given by the rules of classical interval arithmetic (Moore, Bierbaum, and Schwietz 1979). Such uncertainty is larger than that arising from the joint evaluation of each element of the vector $\tilde{g}(x) := PM\rho$, because interval arithmetic assumes all intervals may vary independently, which is not usually the case. Suitable canonical structure choices for \tilde{g} will be discussed below to try to avoid such a source of conservatism.

Remark 5.3. *In a discrete-time case, Lyapunov equations would have considered $\Delta V = (Ax + M\rho)^T P(Ax + M\rho)$ so the difference between linear and non-linear would be: $2x^T A^T PM\rho + \rho^T M^T PM\rho$. As ρ is $O(x^2)$, then the first term is $O(x^3)$ and the second one is $O(x^4)$. So, in order to minimise the discrepancy close the*

origin, concentrating on the terms of the order of x^3 will suggest setting, in this case,

$$\Xi(x) = 2x^T A^T P M \rho(x) \quad (5.51)$$

5.6.2 Diagonalisation of $\Xi(x)$

With $PM=I$ (or $A^T PM=I$ in the discrete case), we have:

$$\Xi(x) = 2x^T \rho(x) \quad (5.52)$$

Consider, too, that a suitable TS model for each element of $\rho(x)$, yielding a shape-independent interval of uncertainty $\rho_i^{si}(x)$, is available. Denote the interval of uncertainty of Ξ as Ξ^{si} , given by:

$$\Xi^{si}(x) = 2 \sum_{i=1}^n x_i \rho_i^{si}(x) \quad (5.53)$$

obtained with standard interval arithmetic from intervals ρ_i^{si} .

Proposition 5.10. *In the above case, the interval of uncertainty Ξ^{si} fulfills:*

$$w(\Xi^{si}(x)) \leq 2\|x\| \cdot \sqrt{\sum_{i=1}^n |\rho_i^{si}(x)|^2} \quad (5.54)$$

Proof. Proof is a consequence of Proposition 5.4 and linearity in x of Ξ . □

Otherwise, with $PM \neq I$, such bound would need to include terms regarding the norm (singular values) of PM , and not all elements of ρ would have the same relevance in Ξ (depending on alignment with the worst-case singular vectors). As handling $PM \neq I$ results quite cumbersome, the objective of the next developments is showing that some changes of variable can lead to a TS model in which $\Xi(x)$ has the expression (5.52) above (in the relevant q -dimensional subspace given by the failing LMIs).

In that way, all directions would have the same influence (so measuring uncertainty on the unit circle is meaningful) and sums in (5.54) will be carried only for i ranging from 1 to q .

Restrictions

As discussed in Section 5.2, constraints (5.8) actually fail in a q -dimensional subspace \mathcal{C} . Considering the change of variable \bar{T} leading to canonical form

(5.12), the expression of $\Xi(x)$ in (5.50) can be written in the new coordinates $x_c = (\xi^T, \eta^T)^T$ being η the subspace coordinates as:

$$\begin{aligned}\Xi(x) &= 2(\xi^T \ \eta^T)\bar{T}^T PM\rho(\bar{T}x_c) \\ &= \xi^T \Gamma_1 \rho(\bar{T}x_c) + \eta^T \Gamma_2 \rho(\bar{T}x_c) := x_c^T \Gamma \rho(\bar{T}x_c)\end{aligned}\quad (5.55)$$

where matrix Γ_1 is formed by selecting the first $n - q$ rows of $\Gamma := 2\bar{T}^T PM$ and Γ_2 is built with the last q rows of Γ . Abusing the notation, we will define $\rho(\xi, \eta) := \rho(\bar{T}x_c)$.

As subspace \mathcal{C} is the one with $\xi = 0$, the proposed goal of the optimal TS modelling is producing a WCW-optimal model of $\tilde{\rho}(\eta) := \Gamma_2 \rho(\xi, \eta)$ in \mathcal{C} in order to avoid losing performance with respect to (5.12), because the restriction of Ξ to \mathcal{C} is

$$\Xi|_{\mathcal{C}}(\eta) = \eta^T \tilde{\rho}(0, \eta) \quad (5.56)$$

Such optimal model of $\tilde{\rho}$ can be obtained by the techniques in previous sections.

Note that (5.56) is a reduced-dimensionality version of (5.52). Hence, with straightforward modifications Proposition 5.10 applies to bound uncertainty in $\Xi|_{\mathcal{C}}(\eta)$, simply changing $\rho(x)$ to $\tilde{\rho}(0, \eta)$ and summing over q dimensions (instead of n) in (5.54).

Last, regarding the remaining nonlinearities in $\Gamma_1 \rho(\xi, \eta)$, a TS model of them can be crafted using any available technique and choice of coordinates as, from (5.55), it will not influence modelling accuracy in the requested subspace.

5.7 Discussion

To conclude the theoretical part of the chapter, let us discuss potential limitations and future lines of research.

The first issue is the resulting number of rules. For instance, the inspection-based model of x_1^2 -case (a) in Example 5.3- has two rules, whereas the ones arising from our recommended changes of variables have four rules. In a quadratic case, it can be shown that, for a nonlinearity $x^T M x$, with M being of rank m , the overall number of rules associated to the q -dimensional subspace \mathcal{C} where it should be optimal will be equal to 2^ϑ , being $\vartheta = \min(q, m)$. Now, regarding the model on the orthogonal subspace \mathcal{C}^\perp , which must forcedly be carried out to build a global model, the nonlinearities should be modelled in the simplest possible way, but, of course, it would require at least two more rules, so the total number of rules will be $2^{\vartheta+1}$ or higher, for each of the involved nonlinearities. Hence, our approach

may result in a larger number of rules than inspection-based ones looking for “simple” TS representations. In complex cases, combining our approach with the approximate complexity reduction techniques mentioned in the introduction might be needed.

Now, the discussion on what is more conservative, few *inexact* rules (classical approach) versus our new proposal of more rules with *optimal fit in a subspace* should be addressed. If the modelling region is small enough, it has been proved that our model will fit inside the projections onto \mathcal{C} of the vertex models of any other TS model. Of course, this fit might not be true on \mathcal{C}^\perp but such possible worse accuracy in the complementary subspace is irrelevant due to the excess performance margin implicit in the *strict* inequality (5.10). So, for small-enough modelling regions our approach will yield performance equal to or better than alternative options.

In larger modelling regions, for a heavily nonlinear system, the geometry of the state trajectories might change substantially from that around the origin; hence, our proposal cannot claim optimality in such a situation, because the directions critical for performance cease to be those in subspace \mathcal{C} from the linearised LMIs (5.9). Further research is needed in order to generalise the idea in Proposition 5.1 to, for instance, LMIs arising from preexisting TS models, incorporating a set of A_i instead of the linearisation A ; however, some technical difficulties arise. Detailed analysis of the issues arising with non-quadratic nonlinearities far from the origin is also needed in further research.

A last observation is the fact that, if LMIs were shape-dependent (such as (Bernal, Guerra, and Kruszewski 2009)), in an ideal case, as “all” possible TS models are equivalent rewritings of the nonlinearities, a “perfect” shape-dependent approach should give the same results whichever the TS model used; however such perfect algorithm has not yet been discovered.

5.8 Additional Examples

Example 5.4. *This first example illustrates the advantages of the proposed approach in order to obtain an optimal model in the sense of finding the maximum decay rate of the following continuous nonlinear system:*

$$\dot{x} = \begin{pmatrix} -3x_1 - 2x_2 - x_3 \\ -4x_2 - x_3 \\ -8(x_1^2 + x_1) + x_2 - 2x_3 \end{pmatrix} \quad (5.57)$$

where $x \in \Omega$, $\Omega \subset \mathbb{R}^3$ being a spherical modelling region (several values for its radius being tested later on). Taking into account the single nonlinearity x_1^2 in (5.57), a conventional inspection-based 2-rule TS model can be obtained:

$$\dot{x} = \sum_{i=1}^2 \mu_i(x) \begin{pmatrix} -3 & -2 & -1 \\ 0 & -4 & -1 \\ -8(\alpha_i + 1) & 1 & -2 \end{pmatrix} x \quad (5.58)$$

with $\mu_1(x) = \frac{x_1 - \alpha_2}{\alpha_1 - \alpha_2}$, $\mu_2(x) = 1 - \mu_1(x)$, $\alpha_1 = \max_{x \in \Omega} x_1$, $\alpha_2 = \min_{x \in \Omega} x_1$. Maximum decay rate γ^{opt} for $V = x^T P x$ under TS model (5.58) is obtained maximising $\gamma > 0$ subject to:

$$P > 0, \quad A_i^T P + P A_i \leq -2\gamma P, \quad i = \{1, 2\} \quad (5.59)$$

These conditions fulfill Assumption 5.1 and are in GEVP form.

The above model's performance will be compared with the maximum decay rate obtained from the proposed modelling technique. To that end, (5.57) is rewritten as (5.68), i.e.,

$$\dot{x} = A x + M \rho(x) \quad (5.60)$$

where A is the linearization of (5.57), $M = (0 \ 0 \ -8)^T$ and, $\rho(x) = x_1^2$. Once $\rho(x)$ is defined, the GEVP test of decay rate (4.21) was applied for the linearised system, i.e. (5.60) with $\rho(x) = 0$, and achieved a $\gamma^{opt} = 1$. No TS model will be able, of course, of getting a faster decay (Proposition 5.2).

The one-dimensional subspace $\mathcal{C} = \{x \in \mathbb{R}^3 : x = (-0.1764 \ 0.3014 \ 0.9371)^T \eta, \eta \in \mathbb{R}\}$ is the one which prevents the linearised decay-rate problem to progress any further: its basis correspond to the eigenvector of the matrix $A^T P + P A + 2\gamma^{opt} P$ associated to the minimal eigenvalue ($\lambda_{min} = 0$ due to γ^{opt}).

Following identical modelling procedures to those in Example 5.3 with the above subspace, the resulting TS model (4 rules), when used in decay-rate optimisation, gives the results in Figure 5.1 (solid red line), which clearly improve over the first proposed TS model (5.58) (dashed-blue line). For instance, the improved modelling can prove marginal stability ($\gamma = 0$) up to a radius of the spherical modelling region of 3, whereas the conventional non-optimised model only proves stability up to radius 0.7. For any of the radius values in the figure, the proposed TS model gets a faster decay (larger γ).

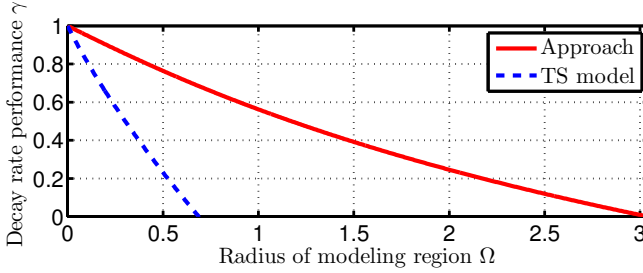


Figure 5.1: Comparison of performance of the proposed approach in a solid line against the TS model (5.58) in dotted lines.

Example 5.5. *This example illustrates the advantages of the proposed approach for \mathcal{H}_∞ control synthesis. Consider the following discrete-time nonlinear system*

$$\begin{aligned} x[k+1] &= \begin{pmatrix} 0.5x_1 + 0.8x_2 + x_1 \sin x_2 + x_2^2 + w_1 - 0.2u \\ 0.25x_1 - 0.45x_2 + 1.5x_1 \sin x_2 + w_2 + 0.1u \end{pmatrix} \\ y[k] &= (0.5x_1 \quad 0.5u)^T \end{aligned} \quad (5.61)$$

where $x \in \Omega$, $\Omega \subset \mathbb{R}^2$ is the state, $w \in \mathbb{R}^2$ is the perturbation, $u \in \mathbb{R}$ is the control input. For comparison a set of 4-rule TS models are proposed, based on extracting the state as a common factor in three different ways:

$$\begin{aligned} x[k+1] &= A_{[i]}^{TS} x + Bu + Ew, \quad i = \{1, 2, 3\} \\ y[k] &= Cx + Du \end{aligned} \quad (5.62)$$

where $A_{[i]}^{TS}$ denotes the arrangement of nonlinearities of the i -th TS model as follows (see Remark 5.4 in Appendix):

$$\begin{aligned} A_{[1]}^{TS} &= \begin{pmatrix} 0.5 & 0.8 + x_1 \operatorname{sinc} x_2 + x_2 \\ 0.25 + 1.5 \sin x_2 & -0.45 \end{pmatrix} \\ A_{[2]}^{TS} &= \begin{pmatrix} 0.5 + \sin x_2 & 0.8 + x_2 \\ 0.25 + 1.5 \sin x_2 & -0.45 \end{pmatrix} \\ A_{[3]}^{TS} &= \begin{pmatrix} 0.5 & 0.8 + x_1 \operatorname{sinc} x_2 + x_2 \\ 0.25 & -0.45 + 1.5x_1 \operatorname{sinc} x_2 \end{pmatrix} \end{aligned}$$

being $\operatorname{sinc}(\alpha) := \frac{\sin(\alpha)}{\alpha}$, and the respective constant matrices:

$$B = \begin{pmatrix} -0.2 \\ 0.1 \end{pmatrix}, \quad C = \begin{pmatrix} 0.5 & 0 \\ 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 \\ 0.5 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Of course, all representations are equivalent. Subsequently, corresponding $TS_{[i]}$ models are obtained using the maximum and minimum in Ω of each element of its respective 2×2 matrix $A_{[i]}^{TS}$ (details omitted).

Now, in order to apply the methodology in this chapter, let us first linearise and then rewrite (5.61) in the form (5.68), i.e.,

$$\begin{aligned} x[k+1] &= Ax + Bu + Ew + M\rho(x) \\ y[k] &= Cx + Du \end{aligned} \quad (5.63)$$

with

$$A = \begin{pmatrix} 0.5 & 0.8 \\ 0.25 & -0.45 \end{pmatrix}, \quad M = \begin{pmatrix} 1 & 1 \\ 0 & 1.5 \end{pmatrix}, \quad \rho(x) = \begin{pmatrix} x_2^2 \\ x_1 \sin x_2 \end{pmatrix}$$

where $\rho(x)$ is the vector of nonlinearities present in the system.

Then, sufficient conditions for the well-know \mathcal{H}_∞ Lyapunov inequality $\Delta V + y^T y - \gamma_{opt}^2 w^T w \leq 0$ are posed minimising γ subject to:

$$\begin{pmatrix} -X & (*) & (*) \\ A_i X + B F_i & -X + E \gamma^{-2} E^T & (*) \\ C X + D F_i & 0 & -I \end{pmatrix} \leq 0 \quad (5.64)$$

where X and F_i are decision variables given by the Lyapunov function $V = x^T X^{-1} x$ and the control gains, respectively (Tanaka and Wang 2001). $(*)$ refers to completion to obtain a symmetric matrix.

Let us now consider the linearised case, i.e., (5.63) with $\rho(x) = 0$, in order to obtain the relevant subspaces where modelling must be precise. In that case, from straightforward Schur complement manipulations, it can be proved that the above problem is feasible if and only if it is so for the worst-case disturbance $w^* = (\gamma^2 I - E^T P E)^{-1} E^T P (A + B K) x$. Replacing such disturbance in the \mathcal{H}_∞ inequality, it can be proved that if (5.64) holds, equivalently, $P > 0$ and

$$\begin{aligned} x^T & \left((A + B K)^T (P^{-1} - E \gamma^{-2} E^T)^{-1} (A + B K) - P \right. \\ & \left. + (C + D K)^T (C + D K) \right) x \leq 0 \quad \forall x. \end{aligned} \quad (5.65)$$

also hold, with $P := X^{-1}$ and $K := F X^{-1}$. These conditions, too, fulfill Assumption 5.1 even if (5.65) is not directly in LMI form; note that convexity in A_i is ensured by the presence of $-X + E \gamma^{-2} E^T$ in the constraints in (5.64).

As (5.64) is an LMI problem in (X, F_i, γ^{-2}) , solving it for the linearized system results in $\gamma^{opt} = 1.4644$. Replacing the obtained optimal decision variables in (5.65), the one-dimensional subspace where (5.65) marginally holds, see expression

(5.9), is $\mathcal{C} = \{x \in \mathbb{R}^2 : x = (0.9518 \quad 0.3068)^T \eta, \eta \in \mathbb{R}\}$, this allows to obtain a transformation matrix $x = \bar{T}x_c$ that rewrites (5.65) in the form (5.12), i.e.,

$$\bar{T} = \begin{pmatrix} -0.3068 & 0.9518 \\ 0.9518 & 0.3068 \end{pmatrix}.$$

Now, in order to generate the optimal TS model, following analogous reasoning to (5.51) and (5.55) in Section 5.6, nonlinearity enters the Lyapunov equation (5.65), disregarding quadratic terms in ρ , as:

$$\Xi = 2x_c^T \bar{T} (A + BK)^T (P^{-1} - E\gamma^{-2}E^T)^{-1} M \rho(\bar{T}x_c)$$

yielding

$$\Xi = x_c^T \Gamma \rho(\bar{T}x_c) = x_c^T \begin{pmatrix} 0.7051 & -0.4719 \\ 1.3140 & 2.4020 \end{pmatrix} \rho(\bar{T}x_c) \quad (5.66)$$

Defining now $\hat{\rho}(\xi, \eta)$ as $\hat{\rho}(\xi, \eta) := \Gamma \rho(\bar{T}x_c)$, we can express $\Xi = x_c^T \hat{\rho}(\xi, \eta) = (\xi^T \quad \eta^T) \hat{\rho}(\xi, \eta)$. In this way, (5.66) has the overall form (5.52) and the restriction can be written as (5.56), after the proposed coordinate changes. By decomposing $\hat{\rho}$ as in (5.31), standard TS modelling in the coordinates ξ and η of $\hat{\rho}^{[1]}$ and $\hat{\rho}^{[2]}$, respectively, is used to find an expression analogous to (5.41), which concludes the modelling step (details omitted).

This allows finding a 16-rule TS model fulfilling the bound in Proposition 5.10 for subspace \mathcal{C} , with the sum ranging in a single dimension. Such model has been used for locally searching for controllers guaranteeing an \mathcal{H}_∞ norm for the nonlinear system in circular regions.

For illustrative purposes, we present Fig. 5.2, which describes the performance bound γ^{opt} using the TS models in LMIs (5.64), as the radius of modelling region Ω increases. Results corresponding to the proposed modelling technique are presented in a solid line, whereas those three TS models in (5.62) are presented with dotted lines. As expected, all TS models match the performance of the linearization at the origin, but as they move away from it, the proposed subspace-based approach outperforms the alternative ones (which disregard the linearised geometry), yielding a lower disturbance-rejection bound (theoretical optimality only for small modelling regions). For completeness, a time simulation presenting $\|y[k]\| = \|Cx + Du\|$ (whose squared sum should be minimised, according to the requested performance criterion) appears in figure 5.3. The original nonlinear dynamics is simulated with four different controllers arising from each of the TS modelling alternatives for a step disturbance. Results show that the achieved figures for $\|y[k]\|$ with our proposed methodology are the lowest ones, and time response is better damped. Of course, we cannot prove that the chosen disturbance is the worst-case one for the

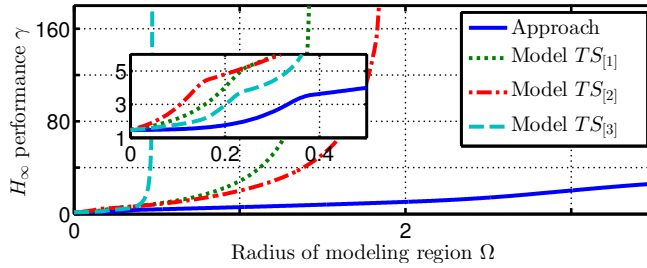


Figure 5.2: Comparison of proven LMI performance between the proposed approach in solid line and the TS models in (5.62) in dotted lines labelled as $TS_{[i]}$. A close-up zoom of the radius range $[0, 0.5]$ is also provided.

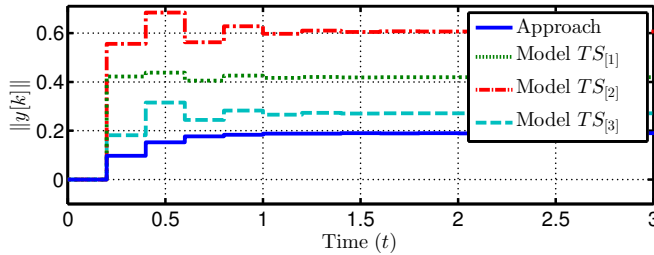


Figure 5.3: Time response of $\|y[k]\|$ for a step $w = (-0.17, -0.09)^T$ of the nonlinear system, with the four controllers arising from each of the considered TS modelling options, using a modelling region radius of 0.3.

nonlinear system (it is an unsolved problem) but, at least with the chosen $w[k]$, the observed performance of our proposal is better than the classical ones.

Note that optimality in the above examples is only claimed “close enough” to the origin: in fact, changing parameters in system matrices, alternative setups can be built in which the proven performance of the proposed approach is only optimal up to a certain radius.

5.9 Conclusions

This chapter has presented a Hessian and subspace based methodology to minimise the conservatism of TS models to be later used in shape-independent LMI conditions. The presented procedure is optimal (in minimax worst-case uncertainty width) for TS models of quadratic functions, hence approximately optimal for any smooth nonlinearity close enough to the origin. Different examples prove

that, indeed, the uncertainty width measures, as well as decay-rate and \mathcal{H}_∞ performance figures are better with the proposed TS modelling technique than those obtained with frequently used “inspection” and “extraction of factors” ideas. The LMIs discuss only the preservation of linearised performance. Other LMI setups and reduction/simplification of the number of rules, as well as optimality for large modelling regions, are matter of future research.

Appendix

5.9.1 sector nonlinearity approach

Consider a single-input nonlinear function $f : \mathbb{R} \mapsto \mathbb{R}$. If $f(0) = 0$ and its derivative is continuous, then the function:

$$\tilde{f}(x) = \begin{cases} f(x)x^{-1} & x \neq 0 \\ \lim_{x \rightarrow 0} f(x)x^{-1} & x = 0 \end{cases}$$

can be defined, because the required limit exists; furthermore, $\tilde{f}(x)$ is continuous. As $f(x) = \tilde{f}(x)x$, we can trivially express, in any compact region $\Omega \subset \mathbb{R}$:

$$f(x) = h(x)\tilde{f}_0x + (1 - h(x))\tilde{f}_1x \quad (5.67)$$

being $\tilde{f}_0 = \max_{x \in \Omega} \tilde{f}(x)$ and $\tilde{f}_1 = \min_{x \in \Omega} \tilde{f}(x)$.

In a multi-input case, a nonlinear model (5.1), as f is linearisable at the origin, can be expressed as:

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

being $\hat{f}(\cdot, \cdot)$ a linear function and ρ a vector of nonlinearities, i.e., there exists $\rho(x) : \mathbb{R}^n \mapsto \mathbb{R}^s$ such that (5.1) is equivalent to

$$\dot{x} = Ax + M\rho(x) \quad (5.68)$$

being A the Jacobian of f at $x = 0$, from (5.2). Note that representation (5.68) may be not unique (there might be several options in choosing M and ρ). Each element of ρ , denoted as $\rho_i(x)$ will be a function $\rho_i(x) : \mathbb{R}^n \mapsto \mathbb{R}$; subindex will be omitted in notation if ρ has a single element.

Consider a modelling region Ω . The objective of sector-nonlinearity TS modelling is bounding in Ω each element $\rho_i(x)$, by two linear functions: actually, finding linear $\eta_i(x)$ such that $\alpha\eta_i(x) \leq \rho_i(x) \leq \beta\eta_i(x)$ for some α, β in \mathbb{R} .

Considering, then, a particular $\rho_i(x)$. Assume there exists a function of the state $\eta_i(x) : \mathbb{R}^n \mapsto \mathbb{R}$, such that $\rho_i(x) = 0$ for all x in $B = \{x \in \Omega : \eta_i(x) = 0\}$. Defining $\tilde{\rho}_i(x)$ as:

$$\tilde{\rho}_i(x) = \frac{\rho_i(x)}{\eta_i(x)}, \quad \text{for } x \notin B \quad (5.69)$$

Then, if the following limit exists for all $x \in B$:

$$\gamma(x) = \lim_{\xi \rightarrow x} \tilde{\rho}_i(\xi)$$

then, the definition of $\tilde{\rho}_i(x)$ can be extended, i.e., defined everywhere in Ω (including the set B), by defining $\tilde{\rho}_i(x) = \gamma(x)$ for $x \in B$, and (5.69) elsewhere. It is well known that the limit $\gamma(x)$ exists, and the resulting extended $\tilde{\rho}_i(x)$ is continuous in Ω if $\rho_i(x)$ has continuous first derivatives (which it does, by assumption). Hence, the relationship

$$\rho_i(x) = \tilde{\rho}_i(x)\eta_i(x)$$

holds in all Ω . By compactness of Ω , the bounding:

$$\left(\min_{y \in \Omega} \tilde{\rho}_i(y) \right) \eta_i(x) \leq \rho_i(x) \leq \left(\max_{y \in \Omega} \tilde{\rho}_i(y) \right) \eta_i(x) \quad (5.70)$$

entails that each $\rho_i(x)$ can be expressed as an interpolation $\rho_i(x) = w_i(x) \times \tilde{\rho}_{i0} + (1 - w_i(x)) \times \tilde{\rho}_{i1}$ where:

$$\tilde{\rho}_{i0} = \max_{y \in \Omega} \tilde{\rho}_i(y), \quad \tilde{\rho}_{i1} = \min_{y \in \Omega} \tilde{\rho}_i(y), \quad w_i(x) = \frac{\tilde{\rho}_i(x) - \tilde{\rho}_{i1}}{\tilde{\rho}_{i0} - \tilde{\rho}_{i1}}$$

When bounding each ρ_i , $i = \{1, \dots, s\}$ as above discussed, $\eta_i(x)$ are linear, say $\eta_i(x) = N_i^T x$, denoting by $M_{[i]}$ the i -th column of M in (5.68), and $\mu_{i0}(x) = w_i(x)$, $\mu_{i1}(x) = 1 - w_i(x)$, the result will be an expression in the form:

$$\dot{x} = \sum_{i=1}^s \sum_{j=0}^1 \mu_{ij} (A + M_{[i]} \tilde{\rho}_{ij} N_i^T) x := \sum_{i=1}^s \sum_{j=0}^1 \mu_{ij} A_{ij} x \quad (5.71)$$

which is a box-like parameter uncertainty description (Gahinet, Apkarian, and Chilali 1996, Eq. (2)). Converting box representations to tensor-product ones (Tanaka and Wang 2001; Ariño and Sala 2007; Nagy et al. 2009) can be done in a straightforward way. Indeed, as $\sum_{j=0}^1 \mu_{ij}(x) = 1$ we can express (5.71) as:

$$\begin{aligned} \dot{x} &= \sum_{i=1}^s \sum_{j=0}^1 \left(\mu_{ij} \left(\prod_{k \neq i} \sum_{j=1}^2 \mu_{kj} \right) A_{ij} \right) x \\ &= \sum_{k=1}^{2^s} \left(\prod_{i=1}^s \mu_i \text{bit}_i(k) A_{\text{bit}_i(k)i} \right) x = \sum_{k=1}^{2^s} h_k A_k x \end{aligned} \quad (5.72)$$

where $\text{bit}_i(k) \in \{0, 1\}$ is the i -th bit ($1 \leq i \leq s$) of the binary representation of integer $k - 1$, and h_k is thus formed as the product of a specific combination of w_i (or $1 - w_i$, depending on the corresponding binary digit).

In summary, once M , $\rho(x)$ and the linear functions $\eta_i(x)$ are chosen, the above well-known steps end up in a TS system with power-of-two vertex models (box or tensor-product form).

Ideally, the TS model of a scalar expression $\rho(x) = \lambda x^2$ is $\lambda(\mu_1 \underline{x} + (1 - \mu_1) \bar{x}) x$, being \underline{x} and \bar{x} the minimum and maximum values of x in a modelling region. For quadratic multivariable functions, the idea is generalised to the Hessian coordinates applying the above to each of the squares in (5.13).

Remark 5.4. *In some references, TS models are generated from a representation*

$$\dot{x} = A(x)x \tag{5.73}$$

instead of (5.68), such that $\lim_{x \rightarrow 0} A(x)$ exists; then, maximum and minimum in the modelling region of each element of matrix $A(x)$ are sought. Evidently, this is a particular case of the above procedure, considering $\eta_i = x_i$.

Chapter 6

Optimal-Performance Models via the LMI Null Space

The problem of achieving optimal performance for nonlinear systems by constructing the most adequate exact Takagi-Sugeno model of the plant and considering its relationship with the linear matrix inequalities it gives rise to, is considered in this report. In contrast with recent approaches on the subject, the performance goal can be chosen from a wide variety of definitions while constraints are no longer required to be state-dependent. As before, this approach is based on coordinate transformations that isolate the effects of the system nonlinearities and allow optimisation with respect to some performance level, by keeping some norm close enough to the linearised performance. It is shown that the proposed methodology outperforms both ordinary “blind” TS modelling as well as former similar approaches.

The contents of this chapter appeared in the conference publication:

- R. Robles, A. Sala, M. Bernal, and T. Gonzalez (2016). “Optimal-Performance Takagi-Sugeno Models via the LMI Null Space”. In: *(Proc. of 4th IFAC Conf. Intell. Control and Autom. Sc.) IFAC-PapersOnLine* 49.5, pp. 13–18

6.1 Introduction

Takagi-Sugeno (TS) models appeared as a way to incorporate mathematical knowledge about nonlinear plants in the construction of a rule-based fuzzy approximation. Quadratic Lyapunov-based methodologies for analysis and design of TS models were developed (Wang, Tanaka, and Griffin 1996): their focus was put on exploiting the convex structure of TS models as to obtain conditions in the form of linear matrix inequalities (LMIs), which could be efficiently solved by convex optimization techniques (Boyd et al. 1994; Tanaka and Wang 2001). No attention was paid to the actual modelling since it was assumed a TS model was readily available: whether this came from a nonlinear system or a parameter-dependent structure, was irrelevant: functions of the nonlinearities/parameters were assumed to lie in a simplex where the convex sum property held.

In the seminal work of (Taniguchi, Tanaka, and Wang 2001), a modelling technique called *sector nonlinearity* was presented: in contrast with former approaches, this methodology allows obtaining exact convex representations of nonlinear systems within a compact set of the state space; results thus obtained were directly valid for the original nonlinear setup without further adjustment. Therefore, during the next years, mainstream research abandoned earlier model-free ideas, to concentrate on model-based methodologies, where exact TS models and LMI conditions were seen as vehicles to analyse and control general nonlinear systems (Sala, Guerra, and Babuska 2005; Guerra, Sala, and Tanaka 2015).

Nevertheless, sector nonlinearity was not without disadvantages. Modelling was identified as one of the main reasons for conservativeness in the TS/LMI framework since a family of systems can be represented by the same TS model and different TS representations can be found for a given nonlinear expression; conservatism issues in fuzzy control are discussed in (Sala 2009). This led to the study of wider classes of convex structures such as descriptors (Guerra, Estrada-Manzo, and Lendek 2015) and polynomial models (Sala and Ario 2009). Moreover, since the complexity of the TS model critically depends on the number of nonlinearities in the nonlinear system, a number of methods were developed to reduce the number of rules: empirical (Setnes et al. 1998), depending on the number of inputs (Gegov 2007), minimising the H_2 norm between a possibly complex expression of the gain-scheduling parameters and a linear-fractional one (Pettersson and Löfberg 2009), based on higher-order singular value decomposition (SVD) in order to approximate the system with another one in tensor-product form (Nagy et al. 2009), and based on the functional principal component analysis (Esaño and Bordons 2014), among others. The former are shape-independent approaches; when the shape of the convex functions capturing the system nonlinearities is taken into

account, some relaxations can also be achieved (Bernal, Guerra, and Kruszewski 2009; Kruszewski et al. 2009). These issues are, however, out of the scope of this contribution, since some of them are *a posteriori* methodologies (once a first model has been obtained) and others deal with *approximations* at a finite set of points rather than *exact* models.

This work inverts the usual approach on the TS/LMI framework: instead of writing LMIs for a given TS model, it assumes that LMIs for a given (linearisation-based) optimisation problem are already available and sets out for a quest to determine the optimal TS model that keeps the proved performance as close as possible to the linearised one (which, as proven later, is the ideal). A first answer to this problem has been offered in former works by the authors, based on subspaces of the state space in which performance is more sensible to modelling errors, see Chapter 4 and Chapter 5 (Robles et al. 2015; Robles et al. 2017).

This work relaxes the condition of state-dependency of former approaches, thus allowing multiple-LMI setups which might be associated with a wider variety of performance measures. The main idea here is based on a Frobenius-norm bound on the “perturbation” that sector-nonlinearity models produce in the LMI matrices.

This chapter is organized as follows: section 6.2 defines the class of nonlinear systems under consideration and the problem of linearised performance optimisation, establishing their relationship with TS modelling and performance optimisation based on such models; section 6.3 develops the main results in this report, this is to say, a more direct approach where several constraints involving the model can be exploited towards an optimisation objective to keep results as close as possible to the linearised case; the methodology is illustrated in section 6.4 via suitable examples.

6.2 Preliminaries

Nonlinear affine-in-control dynamic systems will be considered in this report, i.e.

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t) \quad (6.1)$$

with $x(t) \in \mathbb{R}^n$ being the state, $u(t) \in \mathbb{R}^m$ being a control input, $f : \mathbb{R}^n \mapsto \mathbb{R}^n$, having continuous second-order derivatives and $f(0) = 0$. Also, consider the linearised model of (6.1):

$$\dot{x} = Ax + Bu, \quad A := \left. \frac{\partial f(x)}{\partial x} \right|_{x=0}, \quad B := \left. \frac{\partial g(x)}{\partial x} \right|_{x=0} \quad (6.2)$$

6.2.1 Linearised performance optimisation

Assumption 6.1. *The pursued control objective is the optimisation of a performance measure γ subject to LMI constraints:*

$$\begin{aligned} & \text{minimise } \gamma, \\ & \text{subject to } \gamma > 0, \zeta^T LMI(L, D, \gamma) \zeta \geq 0, \forall \zeta \neq 0 \end{aligned} \quad (6.3)$$

where $\zeta \in \mathbb{R}^q$, D denotes the decision variables (usually matrix variables associated to the Lyapunov function, controller gains, etc.) and L stands for constant model matrices associated to the system under consideration; for instance, L could be defined as $L := \{A, B\}$ given by the linearised system in (6.2)¹. Expression $LMI(\cdot, \cdot, \cdot)$ will be assumed to be (separately) linear in its arguments. Suitable convex optimisation² software will be employed to find the optimal γ and D .

Proposition 6.1. *The optimal performance measure γ^{opt} for the linearised model (6.2) is obtained when there exists D^{opt} such that conditions (6.3), are*

$$\zeta^T LMI(L, D^{opt}, \gamma^{opt}) \zeta = 0 \quad \forall \zeta \neq 0, \zeta \in \mathcal{C} \quad (6.4)$$

$$\zeta^T LMI(L, D^{opt}, \gamma^{opt}) \zeta > 0 \quad \forall \zeta \neq 0, \zeta \in \mathcal{C}^\perp \quad (6.5)$$

for some vector subspace $\mathcal{C} \subset \mathbb{R}^q$, being \mathcal{C}^\perp its orthogonal complement.

Proof is trivial, as symmetric matrices have an orthonormal basis of eigenvectors.

6.2.2 Takagi-Sugeno modelling

The well-known sector nonlinearity methodology rewrites the nonlinear expression on the right-hand side of (6.1) as an algebraically *equivalent* convex sum of linear models

$$\dot{x}(t) = \sum_{i=1}^r h_i(x) (A_i x + B_i u), \quad (6.6)$$

where the membership functions (MFs) h_i , grouped in a vector $h \in \mathbb{R}^r$, belong to the $r - 1$ -dimensional standard simplex

$$\Delta := \{h \in \mathbb{R}^r : \sum_{i=1}^r h_i = 1, h_i \geq 0 \forall i\},$$

¹Other problems would include additional model matrices in L , for example, matrices relating the system with output-feedback settings, disturbance rejection, exact convex representations of (6.1), etc. Details omitted for brevity.

²Although the second constraint in problem (6.3) is named LMI, any tractable matrix inequality constraint, such as GEVP problems, can be considered in the referred expression.

provided the nonlinearities belong to a compact set of the state space Ω , including the origin.

Basically, sector nonlinearity methodology begins by taking the nonlinearities of (6.1), each of these, say $\rho_j(x)$, as in Chapter 4 (Robles et al. 2015), is decomposed as follows:

$$\rho_j(x) = \frac{\rho_j(x)}{\eta_j(x)} \eta_j(x) := \tilde{\rho}_j(x) \eta_j(x), \quad j = 1, \dots, s \quad (6.7)$$

where $\eta_j(x)$ is any *linear* function of the state such that $\eta_j(0) = 0$, thus enforcing that the limit of $\tilde{\rho}_j$ exists when $x \rightarrow 0$. The number of such nonlinearities has been denoted with s . Then, bounding $\rho_j(x)$ by two sectors

$$\left(\min_{x \in \Omega} \tilde{\rho}_j(x) \right) \eta_j(x) \leq \rho_j(x) \leq \left(\max_{x \in \Omega} \tilde{\rho}_j(x) \right) \eta_j(x) \quad (6.8)$$

means that the nonlinearity ρ_j can be expressed as an interpolation between the minimum and maximum value in the above expression, i.e., $\rho_j = w_j \overline{\tilde{\rho}_j} + (1-w_j) \underline{\tilde{\rho}_j}$, where:

$$\overline{\tilde{\rho}_j} = \max_{x \in \Omega} \tilde{\rho}_j(x), \quad \underline{\tilde{\rho}_j} = \min_{x \in \Omega} \tilde{\rho}_j(x). \quad (6.9)$$

Consider a vector $\rho(x) \in \mathbb{R}^s$ whose entries are the terms ρ_j defined above; then, every combination of maxima/minima of the s nonlinear terms defined as (6.8), when substituted in $f(x(t))$ and $g(x(t))$, produces A_i and B_i , respectively, $i \in \{1, 2, \dots, r\}$, $r=2^s$. Each of them corresponds to a MF h_i which is defined as the product of w_j and/or $1-w_j$, $j \in \{1, 2, \dots, s\}$, according to the corresponding combination. For later developments, let us denote as \mathcal{L} the ordered list of TS consequents $\mathcal{L} := \{L_1, L_2, \dots, L_r\}$, where $L_i = \{A_i, B_i\}$.

Remark 6.1. *Note that the freedom in choosing η_j for the nonlinearities in $f(x)$ is not present for those in $g(x)$. Indeed, considering nonlinearities of $g(x(t))u(t)$ in (6.1), when following the structure of (6.7), we notice that u can be considered already “extracted” as a right factor, taking the role of $u \equiv \eta_j$, so $\tilde{\rho}$ are also forcedly set as the nonlinear elements in $g(x(t))$. As there is no freedom in g , for simplicity, in the sequel this work will consider only systems in the form:*

$$\dot{x} = f(x) + Bu \quad (6.10)$$

trying to produce a wise choice of the η_i in $f(x)$.

6.2.3 Performance optimisation in TS models

Assumption 6.2. *The pursued control objective for a TS model is the optimisation of a performance measure γ subject to LMI constraints:*

$$\begin{aligned} & \text{minimise } \gamma > 0, \\ & \text{subject to } \zeta^T \sum_{i=1}^r \sum_{j=1}^r h_i h_j \text{LMI}(L_i, D_j, \gamma) \zeta \geq 0, \\ & \forall \zeta \neq 0 \in \mathbb{R}^q, \forall i, j \in \{1, 2, \dots, r\}. \end{aligned} \quad (6.11)$$

where $L_i \in \mathcal{L}$ denote model vertices and $\sum_{j=1}^r h_j D_j$ denotes the decision variables, which have been made membership-dependent, contrarily to the single set of variables in (6.3). Simple sufficient conditions for positivity of the double convex summation in (6.11) are given in (Tanaka and Wang 2001):

$$\begin{aligned} & \text{minimise } \gamma > 0, \text{ subject to} \\ & \zeta^T (\text{LMI}(L_i, D_j, \gamma) + \text{LMI}(L_j, D_i, \gamma)) \zeta \geq 0 \end{aligned}$$

More powerful relaxed conditions approaches for (6.11) are presented in (Ariño and Sala 2007; Tuan et al. 2001).

Note that setting up (6.11) with a single L (from the linear model) but multiple D_j will *not* prove better performance than the single D in (6.3). This is a well-known result.

As discussed in the introduction TS modelling is not unique, and some choices render better proven performance with latter LMIs.

This work will assume that LMIs (6.3) are available. Then, the goal will be determining the optimal TS model of (6.10), wishing to keep the proved performance as close as possible to the linearised one when LMIs (6.11) are later posed.

This problem has been addressed in prior chapters: in Chapter 4 and Chapter 5 (Robles et al. 2015; Robles et al. 2017), a procedure for decay-rate optimisation (analysis only) was presented based on some Cholesky factors of the Lyapunov function; relevant performance-critical subspaces of the state space were thus identified and suitably exploited; an \mathcal{H}_∞ example was also proposed in the second work.

The basic drawback of such approaches is that constraints should be state-dependent, i.e., (6.3) required $\zeta \equiv x$ in the cited works. That cannot be achieved if model matrices appear at different places in the LMIs, such as, for

instance, Schur complements in discrete-time systems; In some cases, transformations can be made so that the cited procedures can be applied, but this chapter presents a more direct approach applicable to a generic ζ , with the possibility of several constraints involving the model (dimension of ζ needs not be the same as the state one).

6.3 Main result

Consider the optimal solution of the linearised LMIs (6.3). We can express the model (6.10) as $\dot{x} = Ax + Bu + (f(x) - Ax)$ and, suitably extracting the nonlinearities, we can set up a representation

$$\dot{x} = Ax + Bu + M\rho(x) \quad (6.12)$$

With this, note that (6.12) expresses the linear and nonlinear components of (6.1) as two separated terms.

Example 6.1. Consider the following first-order nonlinear system

$$\dot{x} = \begin{pmatrix} -x_1 - x_2 + x_2^2 \\ x_1 - x_2^2 + x_2 \sin(x_1) + u \end{pmatrix}, \quad (6.13)$$

with

$$f(x) = \begin{pmatrix} -x_1 - x_2 + x_2^2 \\ x_1 - x_2^2 + x_2 \sin(x_1) \end{pmatrix}, \text{ and } B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

A representation as (6.12) of this model can be obtained from its linearisation (6.2), which results in $A = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$; then, by rewriting (6.13) as (6.12), it is obtain:

$$\dot{x} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u + \begin{pmatrix} x_2^2 \\ -x_2^2 + x_2 \sin(x_1) \end{pmatrix}$$

where it is easy to see, following (6.12), that $M = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ and $\rho(x) = \begin{pmatrix} x_2^2 \\ x_2 \sin(x_1) \end{pmatrix}^T$.

Now, once the expression for $\rho(x)$ has been selected, with the aim of obtaining a TS representation, express it as $\rho(x) = Q(x)x$ with

$$Q(x) = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{s1} & q_{s2} & \cdots & q_{sn} \end{pmatrix} \quad (6.14)$$

For notational brevity, dependence on x of each element q_{ij} has been omitted. Notice that $\lim_{x \rightarrow 0} Q(x)$ does exist because the linearisation of ρ is zero. This means that (6.12) can be expressed as

$$\dot{x} = Ax + Bu + M\rho(x) = (A + MQ(x))x + Bu. \quad (6.15)$$

Note that, as expression (6.7) in the previous section suggests, there might be infinitely many options to express the original nonlinearities as (6.15). The goal is to obtain the “optimal” decomposition exploring in a set of coordinate transformations, in the line of Chapter 4 (Robles et al. 2015).

In order to obtain an optimal TS model it is suggested to pay special attention to the directions that are critical for performance measure in Proposition 6.1. Considering this proposition, there exists a matrix $H \in \mathbb{R}^{q \times p}$ such that $\mathcal{C} = \{\zeta : \zeta = H\eta, \eta \in \mathbb{R}^p\}$, i.e., whose columns form an orthonormal basis of the null space of $LMI(L, D^{opt}, \gamma^{opt})$. Hence,

$$W = H^T LMI(L, D^{opt}, \gamma^{opt})H \equiv \mathbf{0} \in \mathbb{R}^{p \times p}.$$

Now, in order to find out how the nonlinear part of the original system ($M\rho(x)$) affects the performance of its linearisation, we can use (6.15) to define:

$$W_{NL}(x) = H^T LMI(\{A + MQ(x), B\}, D^{opt}, \gamma^{opt})H \quad (6.16)$$

The linearity of LMI in its first argument, altogether with $W=0$, explains the fact that the contribution of the nonlinearities to the performance “loss” in the LMIs is given by:

$$W_{NL}(x) = H^T LMI(MQ(x), D^{opt}, \gamma^{opt})H \quad (6.17)$$

Then, informally stated, the objective of the “optimal” TS modelling will be obtaining a TS model (i.e., an expression for $Q(x)$) such that the norm of matrix $W_{NL}(x)$ will be “small”. For geometrical reasons, to be discussed later, we will choose the *Frobenius* matrix norm.

Note that in an idealistic “optimal” case, for a TS representation, the norm of matrix W_{NL} is zero for $Q(x) = 0$, which coincides with the linearisation. Thus, to maintain the linearisation performance, the goal is keeping the Frobenius norm of W_{NL} as low as possible for $MQ(x) \neq 0$ by a reasonable choice of the elements of $Q(x)$. Also, consider that each entry of W_{NL} can be expressed as:

$$W_{NL}^{ij} = N^{ij} \cdot \text{vec}(Q(x))$$

where $\text{vec}(Q(x)) \in \mathbb{R}^{(n \cdot s) \times 1}$ is the vectorization of $Q(x)$ (column form), this is

$$\text{vec}(Q(x)) = (q_{11} \quad q_{12} \quad \cdots \quad q_{1n} \quad \cdots \quad q_{s1} \quad q_{s2} \quad \cdots \quad q_{sn})^T$$

and row vector $N^{ij} \in \mathbb{R}^{1 \times n \cdot s}$ is the constant Jacobian vector of the entry W_{NL}^{ij} with respect to $\text{vec}(Q(x))$. Note that N^{ij} contains only known numerical constants. Generalising, the following structure is obtained:

$$\text{vec}(W_{NL}) = N \cdot \text{vec}(Q(x)) \quad (6.18)$$

with $\text{vec}(W_{NL})$ being the vectorization of W_{NL} , and $N \in \mathbb{R}^{p^2 \times (n \cdot s)}$ being the resulting constant matrix whose rows are the row vectors N^{ij} , in the same order as the corresponding elements of $\text{vec}(W_{NL})$.

Through the singular value decomposition (SVD) of N , the representation of $\text{vec}(W_{NL})$ in (6.18) will provide in a weighted way the directions of interest (principal components) in which the modelling of $\rho(x)$ has a greater effect on the standard 2-norm of $\text{vec}(W_{NL})$ which is, actually, the Frobenius norm of matrix W_{NL} . Let us discuss the details below.

Given that the same nonlinearity (through the corresponding q_{ij} elements) might affect several elements of W_{NL} , the decomposition of N in principal components separates the effect of the nonlinearity in ‘‘orthogonal’’ elements (in the vectorised space), as follows

$$N = USV^T \quad (6.19)$$

where $S \in \mathbb{R}^{p^2 \times (n \cdot s)}$ is a matrix containing the singular values of N in the diagonal, and $V \in \mathbb{R}^{(n \cdot s) \times (n \cdot s)}$ is a square matrix whose columns correspond to the critical directions for $LMI(\cdot, \cdot, \cdot)$ in (6.16). Then, consider matrix $Z := SV^T$, $Z \in \mathbb{R}^{p^2 \times (n \cdot s)}$,

$$Z = \begin{pmatrix} z_{1,1} & \cdots & z_{1,n} & \cdots & z_{1,(s-1) \cdot n+1} & \cdots & z_{1,s \cdot n} \\ z_{2,1} & \cdots & z_{2,n} & \cdots & z_{2,(s-1) \cdot n+1} & \cdots & z_{2,s \cdot n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ z_{p^2,1} & \cdots & z_{p^2,n} & \cdots & z_{p^2,(s-1) \cdot n+1} & \cdots & z_{p^2,s \cdot n} \end{pmatrix} \quad (6.20)$$

which represents the vectors of interest weighted by importance. Notice that the norm of Z is equal to the norm of N , due to the fact that $U \in \mathbb{R}^{p^2 \times p^2}$ is an orthogonal matrix; hence, we will concentrate on the effect of the nonlinearities on the Frobenius norm of Z .

Now, for each row i , $i \in \{1, \dots, p^2\}$ of Z in (6.20), to be denoted as Z_i , a matrix $\Xi_i \in \mathbb{R}^{s \times n}$ is built as follows:

$$\Xi_i = \begin{pmatrix} z_{i,1} & \cdots & z_{i,n} \\ z_{i,n+1} & \cdots & z_{i,2 \cdot n} \\ \vdots & \ddots & \vdots \\ z_{i,(s-1) \cdot n+1} & \cdots & z_{i,s \cdot n} \end{pmatrix}. \quad (6.21)$$

The above is motivated by the following relationship:

$$Z_i \text{vec}(Q(x)) = \text{tr}(\Xi_i Q(x)^T),$$

where $\text{tr}(\cdot)$ stands for the trace. Thus, modelling the nonlinearities so that the projection over Ξ_1 is as small as possible is, hence, suggested.

Example 6.2. Assume there is an interest on applying the \mathcal{H}_∞ controller problem and obtain the lower disturbance-rejection bound of a TS representation for the Example 6.1 with the presented approach. Considering the adequate set of restrictions for the mentioned problem, there are obtained the following matrices:

$$H = \begin{pmatrix} 0 & 0 & 0.0243 & 0.7067 & 0.0243 & -0.7067 \\ 0 & 0 & -0.7067 & 0.0243 & -0.7067 & -0.0243 \end{pmatrix}^T,$$

which is an orthonormal basis of the null space of the assumed set of restrictions for the \mathcal{H}_∞ problem,

$$W_{NL}(x) = \begin{pmatrix} W_{NL}^{11} & W_{NL}^{12} \\ W_{NL}^{21} & W_{NL}^{22} \end{pmatrix},$$

described in (6.17) with

$$\begin{aligned} W_{NL}^{11} &= 0.034q_{21} - 0.964q_{12} - 0.033q_{11} + 0.998q_{22}, \\ W_{NL}^{12} &= 0.464q_{11} - 0.533q_{12} - 0.498q_{21} + 0.034q_{22}, \\ W_{NL}^{21} &= 0.464q_{11} - 0.533q_{12} - 0.498q_{21} + 0.034q_{22}, \\ W_{NL}^{22} &= 1.033q_{11} - 0.035q_{12} - 0.034q_{21} + 0.001q_{22}. \end{aligned}$$

Then, $W_{NL} = N \cdot \text{vec}(Q(x))$ where

$$N = \begin{pmatrix} -0.033 & -0.964 & 0.034 & 0.998 \\ 0.464 & -0.533 & -0.498 & 0.034 \\ 0.464 & -0.533 & -0.498 & 0.034 \\ 1.033 & -0.035 & -0.034 & 0.001 \end{pmatrix}, \quad (6.22)$$

Then, following the idea, after the first SVD of (6.22), Z is obtained as in (6.20), resulting in:

$$Z = \begin{pmatrix} 0.723 & -1.170 & -0.447 & 0.723 \\ 0.948 & 0.316 & -0.316 & -0.632 \\ 0.276 & 0.170 & 0.447 & 0.276 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and consequently, the following matrices Ξ_i , as in (6.21) are obtained:

$$\begin{aligned} \Xi_1 &= \begin{pmatrix} 0.723 & -1.170 \\ -0.447 & 0.723 \end{pmatrix}, \Xi_2 = \begin{pmatrix} 0.948 & 0.316 \\ -0.316 & -0.632 \end{pmatrix}, \\ \Xi_3 &= \begin{pmatrix} 0.276 & 0.170 \\ 0.447 & 0.276 \end{pmatrix}, \Xi_4 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Hence, we recommend modelling the first nonlinearity x_2^2 so that its projection on the first row of Ξ_1 is minimised, and modelling the second one $-x_2^2 + x_2 \sin(x_1)$ so that its projection on the second row of Ξ_1 is minimised.

6.4 Decay-Rate analysis example

Example 6.3. This example illustrates the advantages of the proposed approach in order to obtain an optimal model in the sense of finding the maximum decay rate of the following continuous nonlinear system:

$$\dot{x} = \begin{pmatrix} -4x_1 - x_2 + 4x_3 \\ -2x_1 - 4x_2 + 3x_3 \\ 2x_1 + 3x_2 + 4x_3^2 - 4x_3 \end{pmatrix} \quad (6.23)$$

where $x \in \Omega$, $\Omega \subset \mathbb{R}^3$ being a spherical modelling region (several values for its radius being tested later on). Taking into account the single nonlinearity x_3^2 in (6.23), a conventional 2-rule TS model can be obtained:

$$\dot{x} = \sum_{i=1}^2 \mu_i(x) \begin{pmatrix} -4 & -1 & 4 \\ -2 & -4 & 3 \\ 2 & 3 & 4\alpha_i - 4 \end{pmatrix} x \quad (6.24)$$

with $\alpha_1 = \max_{x \in \Omega} x_3$, $\alpha_2 = \min_{x \in \Omega} x_3$, $\mu_1(x) = \frac{x_3 - \alpha_2}{\alpha_1 - \alpha_2}$, and $\mu_2(x) = 1 - \mu_1(x)$. The maximum decay rate $\gamma^{opt} > 0$ for $V = x^T P x$ under the TS model (6.24) via sufficient conditions

$$\begin{aligned} & \text{maximise } \gamma > 0 \\ & \text{subject to } \begin{cases} P > I, \\ A_i^T P + P A_i \leq -2\gamma P, \quad i = \{1, 2\} \end{cases} \end{aligned} \quad (6.25)$$

is obtained.

The above model's performance will be compared with the maximum decay rate obtained from the proposed modelling technique. To that end, (6.23) is rewritten as (6.15), i.e.,

$$\dot{x} = Ax + M\rho(x) \quad (6.26)$$

where A is the linearisation of (6.23), $M = (0 \ 0 \ 4)^T$ and $\rho(x) = x_3^2$. Once $\rho(x)$ is defined, the LMI test of decay rate (6.25) was applied for the linearised system, i.e. (6.26) with $\rho(x) = 0$, which for this particular case reduces to:

$$\begin{aligned} & \text{LMI}(A, P^{opt}, \gamma^{opt}) \\ & = \begin{pmatrix} P^{opt} - I & 0 \\ 0 & -A^T P^{opt} - P^{opt} A - 2\gamma^{opt} P^{opt} \end{pmatrix} \end{aligned}$$

and achieved a $\gamma^{opt} = 1$ (numerical values for P^{opt} are omitted for brevity). No TS model will be able, of course, of proving a faster decay as the linearisation should be in the convex hull of the vertices.

Now in order to obtain an optimal TS model through the described idea, decompose the single element $\rho(x)$ as indicated (6.14), i.e.,

$$\rho(x) = (\tilde{\rho}_{11} \quad \tilde{\rho}_{12} \quad \tilde{\rho}_{13}) x.$$

Then, rewrite as (6.15), this is

$$\begin{aligned} \dot{x} &= (A + M\tilde{\rho}(x)) x \\ &= \begin{pmatrix} -4 & -1 & 4 \\ -2 & -4 & 3 \\ 4\tilde{\rho}_{11} + 2 & 4\tilde{\rho}_{12} + 3 & 4\tilde{\rho}_{13} - 4 \end{pmatrix} x \end{aligned}$$

Matrix H is built from the nullspace of $LMI(A, P^{opt}, \gamma^{opt})$, which resulted to be one-dimensional with the LMI solver, yielding:

$$H = (0 \quad 0 \quad 0 \quad 0.7864 \quad 0.0873 \quad 0.6116)^T$$

It is easily verifiable that $W = H^T LMI(A, P, \gamma^{opt}) H \equiv 0$. Then, proceed to introduce the nonlinear terms into $LMI(A, P, \gamma^{opt})$ to obtain $W(M\tilde{\rho})$ as follows

$$\begin{aligned} W(M\tilde{\rho}) &= H^T LMI(A + M\tilde{\rho}(x), P, \gamma^{opt}) H \\ &= 87.5\tilde{\rho}_{11} + 9.71\tilde{\rho}_{12} - 68\tilde{\rho}_{13} \end{aligned}$$

Now, we rewrite $W(M\tilde{\rho})$ as (6.18), i.e.:

$$W(M\tilde{\rho}) = (87.5 \quad 9.71 \quad -68) \begin{pmatrix} \tilde{\rho}_{11} \\ \tilde{\rho}_{12} \\ \tilde{\rho}_{13} \end{pmatrix}$$

in this case N in (6.19) is $N = (87.5 \quad 9.71 \quad -68)$, and the direction of interest for the TS model is given by vector associated to the largest singular values of N , which correspond to the first row of V in (6.19); for this case it is just one vector:

$$V = (-0.7864 \quad -0.0873 \quad -0.6116)^T$$

Now, complete the transformation matrix T with two more columns, to make it full rank orthonormal; in this example the transformation matrix has been chosen as:

$$T = \begin{pmatrix} -0.6116 & -0.0873 & -0.7864 \\ -0.0299 & 0.9957 & -0.0873 \\ 0.7906 & -0.0299 & -0.6116 \end{pmatrix} \quad (6.27)$$

Then, at this point the presented approach is reduced to follow the methodology described in Chapter 4 (Robles et al. 2015) from (24) therein and onwards, as a transformation matrix has been crafted. This is that, the nonlinearity in (6.23) is expressed and decomposed in new coordinates $x = T\eta$ with T given by (6.27). Thus, we can express

$$x_3 = 0.7906\eta_1 - 0.0299\eta_2 - 0.6116\eta_3$$

and, hence, the nonlinearity will be expressed as:

$$\begin{aligned} x_3^2 = & (0.625\eta_1 - 0.0473\eta_2 - 0.967\eta_3)\eta_1 \\ & + (0.0008942\eta_2 + 0.03658\eta_3)\eta_2 + (0.374\eta_3)\eta_3. \end{aligned}$$

Notice that this arrangement in terms of η correspond to what it is described in Section 6.2.2. This leads us to 3 functions to be modelling through sector nonlinearity, these are $\tilde{\rho}_1(\eta) = (0.625\eta_1 - 0.0473\eta_2 - 0.967\eta_3)$, $\tilde{\rho}_2(\eta) = (0.0008942\eta_2 + 0.03658\eta_3)$, and $\tilde{\rho}_3(\eta) = (0.374\eta_3)$.

With these three $\tilde{\rho}_1$, $\tilde{\rho}_2$, $\tilde{\rho}_3$, maximum and minimum on the requested modelling region are computed and, as a result, an 8-rule TS model is obtained.

The resulting TS model, when used in LMI decay-rate optimisation, gives the results in Fig. 6.1 (solid-blue line), which clearly improve over the first proposed “inspection-based” 2-rule TS model (6.24) (dotted-red line). The figure depicts the proven decay rate bound as the radius of the modelling region (chosen to be a circle) increases. The obtained results are almost identical to those in Chapter 4

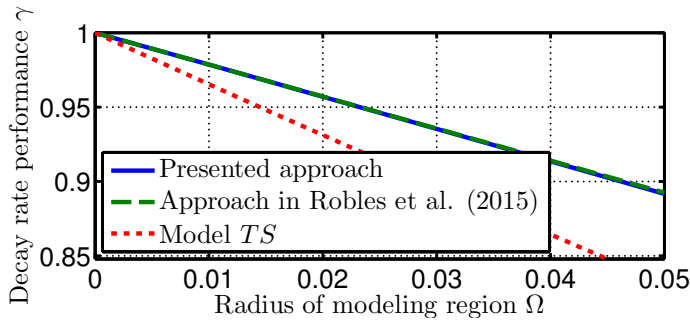


Figure 6.1

(Robles et al. 2015). This was intentional, presenting a comparative example in order to assert that no loss is incurred with the here discussed procedure, when applied to the prior considered cases. Note, however, that the new methodology allows a more general LMI setup for more complex cases; such extensive examples and methodological refinements are left for further research.

6.5 Conclusion

Generating an “optimal” TS model from nonlinear equations is an overlooked aspect in literature.

This chapter has presented a null-space based methodology to craft a coordinate transformation from which a TS model is generated. There are prior works by the team which discuss related issues, but such works can only be applied in very restrictive setups (no LMI transformations). The presented procedure allows for arbitrary LMI setups, with the model appearing as many times as wished in Schur/multicriteria problems. Numerical examples seem to show that performance is comparable to the prior solutions, with a more general applicability.

Chapter 7

On Modelling of Nonlinear Systems as Quasi-LPV Ones

Polytopic quasi-LPV models of nonlinear processes allow using linear matrix inequalities (LMI) to guarantee some performance goal on them (in most cases, locally, over a so-called modelling region). In order to get a finite number of LMIs, nonlinearities are embedded on the convex hull of a finite set of linear models. However, for a given system, these quasi-LPV representations are not unique, yielding different performance bounds depending on the model choice. To avoid such drawback, earlier literature on the topic used annihilator-based approaches, which require gridding on the modelling region, and non-convex BMI conditions for controller synthesis: optimal performance bounds are obtained, but with a huge computational burden. This chapter proposes building a model by minimising the projection of the nonlinearities onto directions which are deleterious for performance. For a small modelling region, these directions are obtained from LMIs with the linearised model. Additionally, these directions will guide the selection of the polytopic embedding's vertices. The procedure allows gridding-free LMI controller synthesis, as in standard LPV setups, with a very reduced performance loss with respect to the above BMI+gridding approaches, at a fraction of the computational cost.

The contents of this chapter improve over those of the previous one. Indeed, the Frobenius norm of some uncertainty-related matrices in an LMI may be only

indirectly related to performance. Thus, this chapter presents an alternative approach. Its contents appear in a draft journal article, currently under review.

7.1 Introduction

Many nonlinear systems can be embedded into a linear parameter-varying (LPV) dynamics, resulting in the so-called quasi-LPV (Rugh and Shamma 2000b; Huang and Jadbabaie 1999; Lendek et al. 2010) or Takagi-Sugeno (TS) (Tanaka and Wang 2001; Guerra and Vermeiren 2004) systems. Efficient linear matrix inequality (LMI) techniques for LPV systems have been developed in order to analyse their properties or synthesise gain-scheduled controllers for them, see the above-cited works, (Apkarian and Gahinet 1995; Scherer 2001; Guerra, Estrada-Manzo, and Lendek 2015; Mohammadpour and Scherer 2012) and references therein. Nonlinearities can also be embedded into the convex hull of polynomials (Sala 2009), amenable to sum-of-squares optimisation (Prajna, Papachristodoulou, and Wu 2004); these approaches are intentionally out of the scope of this chapter. Function-approximation setups and identification-based LPV models (Tóth, Heuberger, and Hof 2012) will also be left out of the present discussion, focusing on obtaining quasi-LPV models from first-principle nonlinear ODEs.

Rewriting a nonlinear model as a quasi-LPV one consists in decomposing a \mathcal{C}^1 nonlinearity $\rho(z)$, present in a first-principle model, as $\rho(z) = Q(z)z$, see (Rugh and Shamma 2000b; Tanaka and Wang 2001; Robles et al. 2017) and references therein. The knowledge about the relationship between the factors $Q(z)$ and z is usually disregarded, so $Q(z(t))$ is considered as an uncertain linear time-varying element $\Delta(t)$ (assumed measurable in gain-scheduled designs (Apkarian, Gahinet, and Becker 1995; Rugh and Shamma 2000b)). Subsequently, some LMIs involving $Q(z)$ must hold for all z in a modelling region Ω_z . If feasible decision variables are found, system properties will be proven inside some bounded invariant set (usually a Lyapunov level set in the undisturbed case, or an inescapable set if disturbances are present and amplitude or integral bounds for them are known (Sala and Pitarch 2016a)).

As the above problem would require solving an infinite set of LMIs, in order to get computable results, either evaluating $Q(z)$ in a “dense enough” grid (Wu et al. 1996; Huang and Jadbabaie 1999; Sala 2005) or obtaining a polytopic bound¹

¹In some LPV developments, stability conditions are relaxed via additional *a priori* bounds on the time-derivatives of the arguments of Q , see (Gahinet, Apkarian, and Chilali 1996; Huang and Jadbabaie 1999; Wu and Dong 2006a). In the quasi-LPV setup, this might be justified in stability analysis if bounding \dot{z} is possible; however, in controller synthesis, as \dot{x} depends on the to-be-designed control u , such derivative bound assumption may imply *a posteriori* checks of their fulfillment (for

with a finite number of vertices (Tanaka and Wang 2001; Kwiatkowski and Werner 2008b) are routinely used in applications.

The main motivation of this chapter is the fact that the decomposition $\rho(z) = Q(z)z$ is, in general, not unique. For instance, in $\rho(z) = \sin(z_1 z_2)$ both $(\sin(z_1 z_2)/z_1 \ 0)$ and $(0 \ \sin(z_1 z_2)/z_2)$ can be equally used as two alternative options for $Q(z)$. However, different performance results for the ensuing analysis or control design steps might be proven for each choice. This is one of the several causes of the inherent conservatism of the quasi-LPV/TS approach with respect to an ideal nonlinear controller (Sala 2009). Actually, there are infinitely many of options for $Q(z)$: the use of annihilators $N(z)z = 0$, such that $\rho(z) = (Q(z) + N(z))z$, has been proposed in (Huang and Jadbabaie 1999; Bruzelius, Pettersson, and Breitholtz 2004; Coutinho, Fu, and Trofino 2004; Trofino and Dezuo 2014), where decision variables in $N(z)$ are used to diminish conservatism; however, most results require gridding and are BMI in control synthesis problems, so these approaches will be not pursued here. An alternative option is a subspace-based selection of the optimal factorisation in order to diminish the “spread” of $Q(z)$ on some directions deleterious for performance; these idea is exploited in (Bruzelius, Pettersson, and Breitholtz 2004), and in the previous chapters, Chapter 4 and Chapter 5 (Robles et al. 2015; Robles et al. 2017), which present preliminary approaches to the developments here.

The objective of this chapter is reversing the usual approach on the quasi-LPV framework: instead of writing LMIs for a given model, it assumes that LMIs for a given performance optimisation problem are already available and sets out for a quest to determine the optimal model that keeps the proven performance as close as possible to the linearised one (indeed, as the linearisation is included in the LPV embedding, the proven performance cannot be better than the linearised one). A general subspace-based method is presented, determining a worst-case matrix perturbation to the linearised LMIs.

The structure of this chapter is as follows: section 7.2 introduces preliminaries and motivates the problem; section 7.3 presents the main result of this chapter and the assumptions involved. Examples in sections 7.4. For readability, some ideas and notation are presented in an appendix.

instance, by simulation (Huang 1999, Chap. 4)). In some cases, the time-derivative components may be cancelled (Gonzalez et al. 2016). Last, conditions on the gradient on the polytopic interpolation coefficients (Lee and Kim 2014; Bernal and Guerra 2010) seem more natural in the quasi-LPV context than plain time-derivative bounds. Nevertheless, these issues will not be considered in the scope of the present work.

7.2 Preliminaries

Let us consider a class of nonlinear systems expressed as:

$$\dot{x} = Ax + M_x v + Bu + Ew \quad (7.1)$$

$$z = C_z x + D_z u + F_z w \quad (7.2)$$

$$y = C_y x + M_y v + D_y u + F_y w \quad (7.3)$$

$$v = \rho(z) \quad (7.4)$$

where $x \in \mathbb{R}^n$, $v \in \mathbb{R}^m$, $u \in \mathbb{R}^q$, $w \in \mathbb{R}^p$, $z \in \mathbb{R}^s$, $y \in \mathbb{R}^r$, and $\rho(z)$ fulfills the following assumption:

Assumption 7.1. *The function $\rho : \mathbb{R}^s \mapsto \mathbb{R}^m$ is a vector of continuously differentiable nonlinearities $\rho \in \mathcal{C}^1$, with $\rho(0) = 0$. Also, without loss of generality, we will assume that ρ has null Jacobian².*

Under the above assumption, $x = 0$ is an equilibrium point of the above system for $u = w = 0$ and the classical linearisation of the above system is obtained replacing (7.4) by $v = 0$.

From Taylor's theorem (Peano form), the following result is well known (Huang and Jadbabaie 1999, Lemma 3.1):

Proposition 7.1. *Under Assumption 7.1, there exists a factorisation*

$$\rho(z) = Q(z)z \quad (7.5)$$

with $Q(z)$ being a continuous matrix-valued function, with $\lim_{z \rightarrow 0} Q(z) = 0$.

Finding $Q(z)$ is trivial in the case $s = 1$, as the only option for $Q(z)$ is $Q(z) = \rho(z)/z$, taking suitable limits when $z \rightarrow 0$, guaranteed to exist by the assumptions³.

Actually, a systematic construction of a particular $Q(z)$ from an explicit $\rho(z)$ appears in the Appendix of this chapter, where an invertible change of variable $z = T\eta$ has been assumed (to be later discussed).

Nevertheless, as discussed in the introduction, the above decomposition is not unique if the dimension of z is equal or larger than two, as infinitely many $\aleph(z)$

²Indeed, if an original model had nonzero linearisation of ρ , such linearisation could be embedded into the linear part of the dynamics (7.1)–(7.3), via straightforward manipulations.

³On the sequel, we will assume an expression $\rho(z)/z$ to be understood, with a slight abuse of notation, for $z = 0$, as $\lim_{z \rightarrow 0} \rho(z)/z$.

such that $\aleph(z)z = 0$ can be found (Bruzelius, Pettersson, and Breitholtz 2004; Huang and Jadbabaie 1999) so that given $\rho(z) = Q_0(z)z$, the decomposition below is still valid:

$$\rho(z) = (Q_0(z) + \aleph(z))z \quad (7.6)$$

Once a decomposition (7.5) is chosen, the representation (7.1)–(7.4) replacing the last equation by (7.5) is named in the literature as quasi-LPV system (Bianchi, Mantz, and Christiansen 2005a), because the original nonlinear dynamics can be embedded into the LPV system conformed by equations (7.1)–(7.3) plus

$$v = \Delta(t) \cdot z \quad (7.7)$$

being $\Delta(t) := Q(z(t))$. Now, if the linear expression (7.7) is substituted on (7.1)–(7.3), a representation:

$$\dot{x} = \tilde{A}(\Delta)x + \tilde{B}(\Delta)u + \tilde{E}(\Delta)w \quad (7.8)$$

$$y = \tilde{C}(\Delta)x + \tilde{D}_y(\Delta)u + \tilde{F}_y(\Delta)w \quad (7.9)$$

can be obtained where system matrices are *affine*⁴ in Δ (i.e., affine in $Q(z)$ in the quasi-LPV setup).

Note that, in most cases, the LMIs arising from posing some control problems to (7.8)–(7.9) disregard the relationship between Δ and z . This, of course, introduces conservatism with respect to an “ideal” nonlinear control *explicitly* using $\rho(z)$. This is the reason on why different choices of $Q(z)$ influence the final LPV obtained performance.

As discussed in the introduction, in order to get a finite number of LMI conditions, if z is assumed to lie in a so-called *modelling region* Ω_z either a “dense enough grid” of values of $Q(z)$ (Leite and Peres 2003) or a polytopic embedding (Angeli, Casavola, and Mosca 2000) are pursued in many cases. The most straightforward polytopic bounding would be obtaining the $2^{m \times s}$ vertices from the infimum and supremum of each element of matrix $Q(z)$ when z ranges in Ω_z (Tanaka and Wang 2001); for instance, these vertices are guaranteed to exist if Ω_z is compact. However, there might be better options with a lower penalty on the achieved performance (see later discussion on the topic). These polytopic quasi-LPV systems are also denoted as Takagi-Sugeno ones in literature (Tanaka and Wang 2001; Guerra and Vermeiren 2004; Sala 2009).

⁴The linear-fractional representation (Packard 1994) is a more general, widely studied, LPV representation. See footnote 7 for a discussion on the applicability of the proposals in this work to such setting.

Annihilator approaches in prior literature. Options for selecting $\aleph(z)$ in (7.6) have been proposed in (Huang and Jadbabaie 1999; Bruzelius, Pettersson, and Breitholtz 2004; Coutinho, Fu, and Trofino 2004; Trofino and Dezuo 2014), where $\aleph(z)$ is denoted as “annihilator”. Note that, if $\aleph(z)z = 0$, we have $R(z)\aleph(z)z = 0$ for any matrix function $R(z)$, so elements in $R(z)$ can be decision variables. For instance, the above approach, in a state-feedback control synthesis problem, would translate into:

$$(A + Q_0(z) + R(z)\aleph(z))X + BF + (*) \preceq 0$$

where $(*)$ denotes the transpose needed to conform a symmetric matrix.

However, the above expression hints the main drawbacks of the annihilator proposal:

- Non-convex synthesis problems, due to the products between decision variables in $R(z)$ and those in X .
- Need of setting a grid on Ω_z , yielding approximate results, as there is the risk of not hitting close enough to the “worst-case” point⁵.

Due to the above computational concerns, the annihilator approach will not be pursued in this work; only some comparisons with it will be made in the example section.

7.2.1 Performance optimisation problem

Under the conditions discussed above, considering that $\Delta(t) := Q(z(t))$ in the usual LPV analysis of nonlinear systems, some matrix inequalities (actually, in many cases, LMIs) can be asserted on the model matrices $L(Q) := \{\tilde{A}(Q), \tilde{B}(Q), \tilde{C}(Q), \tilde{D}(Q), \tilde{E}(Q), \tilde{F}(Q)\}$ arising from (7.8)–(7.9). The structure of problem to be solved over the original nonlinear system is in the form given by the assumption below:

Assumption 7.2. *The pursued objective over the model (7.1)–(7.4), whose LPV embedding is (7.8)–(7.9) is the optimisation of a performance measure γ subject*

⁵ Gridding can be avoided only with polytopic Ω_z on some classes of polynomial systems, setting constant R and affine \aleph (Coutinho, Fu, and Trofino 2004; Trofino and Dezuo 2014). That makes the approach related to the sum-of-squares techniques in (Prajna, Papachristodoulou, and Wu 2004), as pinpointed in (Trofino and Dezuo 2014). Anyway, the relationship with polynomial approaches is intentionally left out of the scope of this chapter, restricted to LPV/TS settings.

to some constraints in the form of a matrix inequality, i.e., obtaining γ^{LPV} below:

$$\begin{aligned} \gamma^{LPV} := \inf_{D_1, D_2(\cdot)} \quad & \gamma, \\ \text{subject to} \quad & \mathcal{MI}(L(Q(z)), \{D_1, D_2(Q(z))\}, \gamma) \succeq 0 \quad \forall z \in \Omega_z \end{aligned} \quad (7.10)$$

where $\{D_1, D_2(Q(z))\}$ are the decision variables, usually arranged in matrices with elements ranging in \mathbb{R} , with D_1 denoting a set of decision variables which cannot depend on z (usually matrix variables associated to the Lyapunov function), and $D_2(Q(z))$ are decision variables which can depend on $Q(z)$ (for instance, those related to controller gains in gain-scheduled setups). $\mathcal{MI}(\cdot, \cdot, \cdot) \succeq 0$ will be understood as \mathcal{MI} being positive semi-definite, with $\mathcal{MI}(\cdot, \cdot, \cdot)$ assumed to be a continuous matrix-valued expression, monotonic in γ , i.e., for fixed L , D_1 and D_2 , condition $\gamma_1 \geq \gamma_2$ implies $\mathcal{MI}(L, \{D_1, D_2\}, \gamma_1) \succeq \mathcal{MI}(L, \{D_1, D_2\}, \gamma_2)$. The modelling region Ω_z will be assumed to be the unit ball around the origin⁶.

Of course, problems stated as multiple matrix inequalities will be assumed to be equivalently cast as a single block-diagonal one.

Robust versus gain-scheduled designs. In stability analysis or robust linear control (u independent of $\Delta(t) \equiv Q(z(t))$) there is no need of assuming the value of $Q(z(t))$ being explicitly known (Scherer 2004). On the other hand, with a gain-scheduled controller, say $u = \bar{u}(Q(z(x, u, w)), x)$, factorisation (7.5) should be set up so that unmeasurable components of w are not present in $Q(z)$. Also, if Q depends on u then an algebraic loop would need to be solved at each instant of time in order to compute u (and well-posedness conditions would need to be added). If $D_z = 0$ such algebraic loop is avoided. Alternatively, if $Q(z)$ is forced to be in the form $Q(x)$, i.e., a nonlinearity $v = (1 + x^2)u$ is forcedly factored with $Q = (1 + x^2)$, then such algebraic loops would also be avoided. As there is no “freedom” in choosing such “pinned” components of Q , the discussion in this work will be centered on giving good options for the “unrestricted” elements of Q , so pinned components of Q will not be considered any further.

Due to the non-uniqueness of the decomposition (7.5), different performance results can be obtained for different choices of Q . In any case, as the modelling region includes the origin, $Q(0) = 0$ will be included in the inequalities (7.10), so proven performance (in the setting in Assumption 7.2) will always be worse than or equal to that of the linearised model.

⁶If the modelling region were expressed in terms of ellipsoids containing a modelling region of variables x, u, w , there is no loss of generality on assuming Ω_z to be the unit ball, after linear changes of variable.

The main goal of this chapter is providing guidelines on how to select a suitable $Q(z)$ so that “performance loss” with respect to the linear case is reduced, at least when the modelling region Ω_z is small. Also, as a by-product, some worst-performance directions are obtained (projection of $Q(z)$ over them must be minimised); these directions can, too, contribute to minimising the performance loss when using polytopic embeddings.

7.3 Main Result

Consider the linearisation of (7.1)–(7.4) at the origin, i.e.:

$$\dot{x} = Ax + Bu + Ew \quad (7.11)$$

$$z = C_z x + D_z u + F_z w \quad (7.12)$$

$$y = C_y x + D_y u + F_y w \quad (7.13)$$

The restriction to $Q(0)=0$ of the problem stated in Assumption 7.2 gives rise to the following definition:

Definition 7.1 (linearised performance). *The pursued objective over the above linearised model is the optimisation of a performance measure γ subject to some constraints in the form of matrix inequalities in decision variables D_1 and D_2 , obtaining γ^{lin} below:*

$$\begin{aligned} \gamma^{lin} := & \inf_{D_1, D_2} \gamma, \\ & \text{subject to } \mathcal{MI}(L(0), \{D_1, D_2\}, \gamma) \succeq 0 \end{aligned} \quad (7.14)$$

Assumption 7.3. *We will assume that the above linearised performance problem is feasible and that the objective function is bounded from below ($\gamma^{lin} > -\infty$). The decision variable values achieving γ^{lin} will be denoted by $\{D_1^{lin}, D_2^{lin}\}$ (or any arbitrary selection of them if non-unique). By assumption, too, conditions \mathcal{MI} will be such that the linearised problem has a finite solution $\{D_1^{lin}, D_2^{lin}\}$ achieving optimality.*

Let us denote the null space of the optimal solution as:

$$\mathcal{C} := \text{null}(\mathcal{MI}(L(0), \{D_1^{lin}, D_2^{lin}\}, \gamma^{lin})) \quad (7.15)$$

which, by continuity of \mathcal{MI} , boundedness of γ^{lin} and linearity of the objective function will be non-empty (linear functions attain a minimum in the frontier of the feasible set).

Now, let us consider the nonlinear model (7.1)–(7.4) and consider expressing $\rho(z) = Q(z)z$ with

$$Q(z) = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ q_{m1} & q_{m2} & \cdots & q_{ms} \end{pmatrix} \quad (7.16)$$

For notational brevity, dependence on z of each element q_{ij} has been omitted. Considering the control problem in Assumption 7.2, given that the dependence on z of the model matrices will be only through $Q(z)$, the problem in the referred assumption will be equivalently recast as:

$$\begin{aligned} \gamma^{\text{LPV}} &= \inf_{D_1, D_2(\cdot)} \gamma, \\ \text{subject to } & \mathcal{MI}(L(Q), \{D_1, D_2(Q)\}, \gamma) \succeq 0 \quad \forall Q \in Q(\Omega_z) \end{aligned} \quad (7.17)$$

where, with a slight abuse of notation, we define $Q(\Omega_z)$ as

$$Q(\Omega_z) := \{\xi \in \mathbb{R}^{m \times s} \mid \exists z \in \Omega_z \text{ s.t. } \xi = Q(z)\}. \quad (7.18)$$

As earlier discussed, evidently $\gamma^{\text{LPV}} \geq \gamma^{\text{lin}}$.

Now, let us assert additional assumptions for the conditions (7.17), i.e., (7.10), which will allow us to bound $\gamma^{\text{LPV}} - \gamma^{\text{lin}}$ for small Ω_z from geometric considerations over (7.10) and (7.15). On the sequel, $\|\cdot\|_F$ will denote the Frobenius norm.

Now, keeping the linearised solution for D_1 , i.e., D_1^{lin} , introducing an “incremental” notation in D_2 and in performance, a conservative version of problem (7.17) may be written as:

$$\Delta\gamma^* := \sup_{Q \in Q(\Omega_z)} \Delta\bar{\gamma}(Q) \quad (7.19)$$

$$\Delta\bar{\gamma}(Q) := \inf_{\Delta D_2} \Delta\gamma \quad (7.20)$$

$$\text{subject to } \mathcal{MI}(L(Q), \{D_1^{\text{lin}}, D_2^{\text{lin}} + \Delta D_2\}, \gamma^{\text{lin}} + \Delta\gamma) \succeq 0 \quad (7.21)$$

where $\Delta\bar{\gamma}(Q)$ is the performance loss (or gain, if negative) if system matrices were modified to the currently tested Q (keeping D_1^{lin}). Obviously a different ΔD_2 for each Q can be used in (7.21), analogously to (7.17). Conservatism arises from fixing D_1 to the linearised solution, instead of searching for a common D_1 in (7.17), not necessarily equal to D_1^{lin} ; this implies $\gamma^{\text{LPV}} \leq \gamma^{\text{lin}} + \Delta\gamma^*$. In exchange for such conservatism, conditions (7.21) can be conceptually solved *independently* for each Q , motivating the definition of $\Delta\bar{\gamma}(Q)$ in (7.20); this was not possible for (7.17).

Assumption 7.4. We will consider \mathcal{MI} to be:

1. (Jointly) affine ⁷ in arguments (Q, D_2, γ) ,
2. The linearised solution is robust, in the sense that for any $\varepsilon > 0$, there exists $\delta > 0$ such that for any Q such that $\|Q\|_F \leq \delta$, condition (7.21) is feasible and the optimal $\Delta\bar{\gamma}(Q)$ and its associated ΔD_2 in (7.20) fulfill $\|\Delta D_2\|_F \leq \varepsilon$, $|\Delta\bar{\gamma}(Q)| \leq \varepsilon$.

Let H be a matrix whose vectors form a basis of the subspace \mathcal{C} defined in (7.15), and let H_\perp a basis of the orthogonal subspace to \mathcal{C} . Hence, considering the optimal conditions for the linearised model, introducing notation in incremental variables and a congruence matrix, we have an equivalent condition to (7.21):

$$\begin{aligned} \begin{pmatrix} H^T \\ H_\perp^T \end{pmatrix} \mathcal{MI}(L(Q), \{D_1^{\text{lin}}, D_2^{\text{lin}} + \Delta D_2\}, \gamma^{\text{lin}} + \Delta\gamma) \begin{pmatrix} H_\perp & H \end{pmatrix} \\ = \begin{pmatrix} \Xi_{11}(Q, \Delta D_2, \Delta\gamma) & \Xi_{12}(Q, \Delta D_2, \Delta\gamma) \\ \Xi_{12}(Q, \Delta D_2, \Delta\gamma)^T & \mathcal{W}(Q, \Delta D_2, \Delta\gamma) \end{pmatrix} \succeq 0 \end{aligned} \quad (7.22)$$

And, by definition, we have $\Xi_{11}(0, 0, 0) \succ 0$, $\Xi_{12}(0, 0, 0) = 0$, $\mathcal{W}(0, 0, 0) = 0$.

The sorted eigenvalues of a matrix are Lipschitz continuous with respect to its elements (Bronshtein 1979); therefore, there exists ε such that for all $\|Q\|_F \leq \varepsilon$, $\|\Delta D_2\|_F \leq \varepsilon$, $|\Delta\gamma| \leq \varepsilon$ the inequality $\Xi_{11}(Q, \Delta D_2, \Delta\gamma) \succ 0$ holds. Now, from Assumption 7.4.(2), for all $\|Q\|_F \leq \min(\varepsilon, \delta)$, the LMIs (7.22), equivalent to (7.21), are feasible, and the optimal solution $\Delta\bar{\gamma}(Q)$ in (7.20) is attained with $\|\Delta D_2\|_F \leq \varepsilon$, $\Delta\bar{\gamma}(Q) \leq \varepsilon$. Then, as $\Xi_{11}(\cdot) \succ 0$, by Schur complement, the optimisation problem of minimising $\Delta\gamma$ subject to:

$$\Xi_{11}(Q, \Delta D_2, \Delta\gamma) \succ 0 \quad (7.23)$$

$$\mathcal{W}(Q, \Delta D_2, \Delta\gamma) - \Xi_{12}(\cdot)\Xi_{11}(\cdot)^{-1}\Xi_{12}(\cdot)^T \succeq 0 \quad (7.24)$$

⁷ In order to approximately fulfill Assumption 7.4.1 in gain-scheduled cases where $B(Q)u(Q, x)$ appears, for small Q we can linearise $B(Q)u(Q, x) \approx B(0)(u(Q, x) - u(0, x)) + B(Q)u(0, x)$. Similarly, if the model equation (7.2) were changed to $z = C_z x + M_z v + D_z u + F_z w$, then the model would be a more general linear-fractional representation (Packard 1994), which would result in model matrices in (7.8) and (7.9) being rational in $\Delta \equiv Q$. For instance, $A(Q)$ would be $A(Q) = A + M_x Q(1 - M_z Q)^{-1} C_z$. As this model matrix is rational in Q , Assumption 7.4.(1) does not apply. Again, the 1st-order approximation to $A(Q)$ is $A(Q) \approx A + M_x Q C_z$ so the results from our proposal to the model with $M_z = 0$ can, approximately, be applied to LFT models with $M_z \neq 0$ for small modelling regions.

will be feasible, too, for all $\|Q\|_F \leq \min(\varepsilon, \delta)$, and the optimal solution $\Delta\bar{\gamma}(Q)$ will be limited only by active constraints in (7.24). The above reasoning, then, proves the following proposition:

Proposition 7.2. *Under assumptions 7.1 to 7.4, there exists $\varepsilon > 0$ such that, for all Q verifying $\|Q\|_F \leq \varepsilon$, the solution of (7.20) subject to (7.21) is the same as (7.20) subject to (7.24).*

Note, however, that (7.24) is a nonlinear matrix inequality. Expressing the affine $\Xi_{11}(Q, \Delta D_2, \Delta\gamma)$ as $\Xi_{11}(Q, \Delta D_2, \Delta\gamma) = M + \mathcal{N}(Q, \Delta D_2, \Delta\gamma)$, being $M := \Xi_{11}(0, 0, 0)$, and \mathcal{N} a linear function of its arguments, the nonlinear term in (7.24) can be written as:

$$\begin{aligned}\widehat{\Xi} &:= -\Xi_{12}(\cdot)\Xi_{11}(\cdot)^{-1}\Xi_{12}(\cdot)^T \\ &= -\Xi_{12}(\cdot)M^{-1}\Xi_{12}(\cdot)^T + \Xi_{12}(\cdot)M^{-1}\mathcal{N}(\cdot)M^{-1}\Xi_{12}(\cdot)^T + \dots\end{aligned}$$

which shows that it contains quadratic, cubic, etc. terms in Q and the decision variables, which will be negligible for small enough ε : as the eigenvalues of Ξ_{12} are $O(\varepsilon)$, the eigenvalues of $\widehat{\Xi}$ are $O(\varepsilon^2)$. Thus, the solution to (7.19)–(7.21), for small enough Ω_z will be approximately equal to the netx one:

$$\Delta\gamma_L^* := \sup_{Q \in \mathcal{Q}(\Omega_z)} \Delta\bar{\gamma}_L(Q) \quad (7.25)$$

$$\Delta\bar{\gamma}_L(Q) := \inf_{\Delta D_2} \Delta\gamma \quad (7.26)$$

$$\text{subject to } \mathcal{W}(Q, \Delta D_2, \Delta\gamma) \succeq 0 \quad (7.27)$$

because the difference between (7.27) and (7.24) shrinks as $O(\varepsilon^2)$.

Note, however, that the above \mathcal{W} is linear in all its arguments, so if we have a solution of (7.26) subject to (7.27) for some Q , the solution for a “scaled” $Q' := \lambda Q$ would be $\Delta\bar{\gamma}_L(Q') = \lambda\Delta\bar{\gamma}_L(Q)$. This scaling and convexity arguments will allow us to give a bound on $\Delta\gamma_L^*$ considering the solution of the above problem over a canonical basis of Q , as follows:

Let us now define as $E_{jk} \in \mathbb{R}^{s \times n}$ the matrix with the same size as Q whose element (j, k) , is one, being the rest of elements zero (j denotes row, k denotes column number).

LEMMA 7.1. *Let us denote by $\Delta\gamma^{[jk]}$ the solution to the LMI problem*

$$\Delta\gamma^{[jk]} := \inf_{\Delta D_2} \Delta\gamma \quad (7.28)$$

$$\text{subject to } \mathcal{W}(E_{jk}, \Delta D_2, \Delta\gamma) \succeq 0 \quad (7.29)$$

Defining $\Delta\Gamma \in \mathbb{R}^{m \times s}$ as the matrix whose element (j, k) is $\Delta\gamma^{[jk]}$, and the set of matrices

$$\mathcal{J} := \{Q \in \mathbb{R}^{s \times n} : \text{tr}(\Delta\Gamma Q^T) \leq \delta\},$$

it follows that, if $Q(\Omega_z) \in \mathcal{J}$, then $\Delta\gamma_L^* \leq \delta$.

Proof. By linearity of \mathcal{W} in its arguments, conditions (7.27) are (convex) LMIs. Hence, for an arbitrary Q with structure (7.16), which can be trivially expressed as $Q = \sum_{j=1}^m \sum_{k=1}^s q_{jk} E_{jk}$, we can state that the optimal $\Delta\bar{\gamma}_L(Q)$ in (7.26) will fulfill, by convexity:

$$\Delta\bar{\gamma}_L(Q) \leq \sum_{j=1}^m \sum_{k=1}^s q_{jk} \Delta\gamma^{[jk]} = \text{tr}(\Delta\Gamma Q^T) \quad (7.30)$$

Thus, for any $Q \in \mathcal{J}$ we have $\Delta\bar{\gamma}_L(Q) \leq \delta$, which implies the assertion in the lemma. \square

7.3.1 Quasi-LPV modelling methodology

The above argumentations sum up asserting that for a small enough Ω_z , the solution $\Delta\gamma_L^*$ from (7.25) will be approximately equal to $\Delta\gamma^*$ from (7.19), so $\gamma^{\text{LPV}} \leq \gamma^{\text{lin}} + \Delta\gamma^* \approx \gamma^{\text{lin}} + \Delta\gamma_L^*$.

From the fact that $\text{tr}(\Delta\Gamma Q^T)$ is the scalar product on the vector space of matrices, and $\|Q\|_F = \text{tr}(QQ^T)^{1/2}$, we can assert that, for a given Q such that $\|Q\|_F \leq \varepsilon$, $\text{tr}(\Delta\Gamma Q^T) \leq \|\Delta\Gamma\|_F \|Q\|_F$ (Schwartz inequality) and that equality holds if $Q = \alpha \Delta\Gamma$, being α a positive scalar. So the largest performance loss would be incurred if the components of Q “collinear” with $\Delta\Gamma$ are large.

Thus, the problem addressed next is finding a factorisation $\rho(z) = Q(z)z$, being $Q(z)$ a row vector, such that $Q(\Omega_z)$ has the smallest possible orthogonal projection over a 1-dimensional subspace.

Consider a real-valued nonlinearity $v = \rho(z)$, $\rho : \mathbb{R}^s \mapsto \mathbb{R}$, fulfilling assumptions in Section 7.2 (i.e., being \mathcal{C}^1 and $\rho(0) = 0$) and a modelling “direction” given by column vector $\psi \in \mathbb{R}^s$, normalised so that $\psi^T \psi = 1$ without loss of generality⁸.

Let us denote the 1-dimensional vector subspace generated by ψ as $\Xi := \{\eta\psi, \eta \in \mathbb{R}\}$. Considering the projection matrix onto Ξ , $\Pi_\Xi := \psi\psi^T$, the orthogonal pro-

⁸From the above-discused geometric considerations, the actual model direction to be evaluated comes from transposing a $s \times 1$ row of $\Delta\Gamma$ in Lemma 7.1.

jection of a set S of row vectors over Ξ will be understood as:

$$\text{proj}(S, \Xi) := \{y\Pi_{\Xi} : y \in S\} \quad (7.31)$$

Defining the line segment $\mathbb{L} := \Omega_z \cap \Xi$, evidently $\text{proj}(Q(\mathbb{L}), \Xi) \subseteq \text{proj}(Q(\Omega_z), \Xi)$ because $\mathbb{L} \subseteq \Omega_z$.

Definition 7.2. A factorisation $\rho(z) = Q(z)z$, $Q(z) = [q_1(z) \dots q_s(z)]$ is tight on subspace Ξ if

$$\text{proj}(Q(\Omega_z), \Xi) = \text{proj}(Q(\mathbb{L}), \Xi).$$

Proposition 7.3. Under the assumptions in Section 7.2, given any arbitrary vector ψ , generating the 1-dimensional subspace Ξ , there exists a factorisation $\rho(z) = Q(z)z$ which is tight on Ξ .

Proof. Considering Π_{Ξ} , and $\Pi_{\perp} := I - \Pi_{\Xi}$, z can be decomposed on two orthogonal components $z = z_{\Xi} + z_{\perp}$, being $z_{\Xi} = \Pi_{\Xi}z$ and $z_{\perp} = \Pi_{\perp}z$. Also, if $z \in \Omega_z$, $z_{\Xi} \in \mathbb{L}$. Consider now the decomposition $\rho(z) = \rho(z_{\Xi}) + \delta(z)$, being $\delta(z) = \rho(z) - \rho(\Pi_{\Xi}z)$. Obviously, for $z \in \mathbb{L}$, we have $z = z_{\Xi}$ so $\delta(z) = 0$.

If, for fixed z_{Ξ} , we define $\bar{\delta}(z_{\perp}) := \delta(z_{\perp} + z_{\Xi})$ it fulfills Assumption 7.1 so, from Proposition 7.1, there exists a factorisation $\bar{\delta}(z_{\perp}) = \bar{Q}(z_{\perp}, z_{\Xi})z_{\perp} = \delta(z)$. Thus, we can decompose:

$$\rho(z) = \rho(\Pi_{\Xi}z) \frac{\psi^T z}{\psi^T z} + \bar{Q}(z_{\perp}, z_{\Xi})\Pi_{\perp}z$$

Now, denoting:

$$Q(z) = \left(\frac{\rho(\Pi_{\Xi}z)}{\psi^T z} \psi^T + \bar{Q}(z_{\perp}, z_{\Xi})\Pi_{\perp} \right) \quad (7.32)$$

we can, indeed, express $\rho(z) = Q(z)z$. Now,

$$\text{proj}(Q(z_{\Xi} + z_{\perp}), \Xi) = \frac{\rho(z_{\Xi})}{\psi^T z_{\Xi}} \psi^T = \text{proj}(Q(z_{\Xi}), \Xi)$$

so the projection of $Q(z)$ only depends on the component $z_{\Xi} \in \mathbb{L}$. Thus, $\text{proj}(Q(\Omega_z), \Xi) \subseteq \text{proj}(Q(\mathbb{L}), \Xi)$ and, as inclusion in the other sense comes from $\Omega_z \supseteq \mathbb{L}$, the proof is complete. \square

The above proof constructs the component of $Q(z)$ collinear with ψ^T , but does not consider how to build \bar{Q} . A recursive application of the idea is outlined in the Appendix to completely specify Q , after making a change of variable $z = T\eta$

where the last row of T is ψ^T and the remaining rows are built to conform an orthogonal matrix.

Once the methodology to obtain $Q(z)$ has been discussed, Algorithm 1 is proposed to minimise the performance bound in Lemma 7.1. In this way, the spread of $Q(z)$ in the directions more deleterious for performance (at least for small enough modelling regions) will be minimised. Examples in Section 7.4 will show the performance improvement achieved with the resulting model from the referred algorithm.

Algorithm 1 Factorisation $\rho(z) = Q(z)z$

- 1: Solve the linearised performance problem, obtaining $\{D_1^{\text{lin}}, D_2^{\text{lin}}\}$ and γ^{lin} .
 - 2: Using D_1^{lin} , obtain matrix $\Delta\Gamma$ defined on Lemma 7.1.
 - 3: Denoting as $\Delta\Gamma^{[i]}$ the i -th row of $\Delta\Gamma$ and, likewise, $Q^{[i]}$ the i -th row of Q , model each nonlinearity as $\rho_i(z) = Q^{[i]}(z)z$, in such a way so that $Q^{[i]}$ is tight in the direction $\Delta\Gamma^{[i]}$, in the sense of Definition 7.2.
-

Note that only the geometry of the null space \mathcal{C} is used in the factorisation proposal (Algorithm 1), and not any “shape” information on $\rho(z)$. Thus, final performance loss will depend on (a) the actual range of $Q(z)$, (b) the improvement a new D_1 in problem (7.10) might be able to achieve and, (c) the influence of other directions not in the null space \mathcal{C} as modelling region size increases (so $Q(z)$ significantly departs from zero). Hence, the ideas inspiring Algorithm 1 only apply for small enough modelling regions.

7.3.2 Polytopic bounds for $Q(z)$

Once $Q(z)$ has been chosen, a so-called polytopic embedding (a.k.a. TS model) of $Q(z)$ is a set of n_v vertex points \hat{Q}_i such that $Q(z) \in \text{Co}\{\hat{Q}_i, i=1, \dots, n_v\}$ for all $z \in \Omega_z$. Of course, the best embedding would be $\text{Co}\{Q(\eta) | T\eta \in \Omega_z\}$ but that set might not have a finite number of vertices⁹.

Figure 7.1 illustrates the idea for a case of a single nonlinearity $\rho(z_1, z_2) = Q(z)z = (q_{11}(z) \ q_{12}(z))z$, where $Q(\Omega_z)$ is depicted as a blue region, and its convex hull is shown with a dotted-black boundary.

⁹In practice, the convex hull of a set of (dense enough) grid points might be the best reasonable approximation but it might, nevertheless, have a large number of vertices. Such grid appears as a collection of white dots on Figure 7.1.

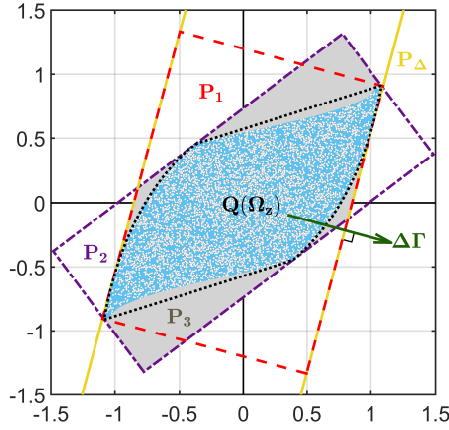


Figure 7.1: Illustration of options for polytopic bounding of $Q(\Omega_z)$ [blue] for $Q \in \mathbb{R}^{2 \times 1}$. Best bounding is the convex hull [dotted-black line]. Given direction $\Delta\Gamma$ [green arrow], the rectangle with minimal projection on $\Delta\Gamma$, equation (7.53), is P_1 . The principal-component box from (Kwiatkowski and Werner 2008b) is P_2 [dash-dotted violet rectangle]. The intersection between P_2 and P_Δ from (7.34) is the shaded gray region P_3 .

However, some simple alternatives exist to avoid gridding, by bounding $Q(\Omega_z)$ in a “box” (hyperrectangle), obtaining infimum and supremum over some orthogonal directions, as follows:

- (a) *Worst-performance directions:* expression (7.53) in the Appendix provides a polytopic hyper-rectangle bounding, given some change of variable. For the case $\Delta\Gamma = (0.966 \ -0.259)$, which appears as a green arrow in the figure, the resulting bound would be the rectangle delimited by dashed-red lines, labelled as P_1 in the figure.
- (a) *Principal-component (PCA) directions:* the work (Kwiatkowski and Werner 2008b) proposes vectorising $Q(z)$ and obtaining the principal directions (eigenvectors of the covariance matrix), to reduce possible overbounding arising from choosing boxes not “aligned” with the principal components of $Q(\Omega_z)$; indeed, this does improve the obtained performance over other box orientations in many cases, see the cited work for details and examples. The result of their algorithm would be the rectangle whose edges are outlined in dash-dotted violet line on Figure 7.1, labelled P_2 .

Combined PCA plus worst-performance direction embedding. The above approaches have drawbacks arising from the fact that principal directions of the dataset $Q(z)$ might not be aligned with the worst-performance directions arising from the rows $\Delta\Gamma^{[i]}$. Thus, choosing just one of the two options may lead to either overbounding, in case (a), or nonminimal projection over $\Delta\Gamma^{[i]}$, in case (b) losing the “tightness” pursued in the previous subsection.

Thus, to avoid any overbounding in the most influential direction, our proposal is to obtain both the principal directions and the supporting hyperplanes given by

$$l_1 := \inf_{z \in \Omega_z} \text{tr}(\Delta\Gamma Q(z)^T), \quad l_2 := \sup_{z \in \Omega_z} \text{tr}(\Delta\Gamma Q(z)^T) \quad (7.33)$$

so that the principal-component polytope P_2 , is intersected with the set

$$P_\Delta := \{l_1 \leq \text{tr}(\Delta\Gamma Q^T) \leq l_2\} \quad (7.34)$$

being P_Δ represented as a yellow band on the figure.

So we can state that $Q(\Omega_z) \subseteq P_3$, $P_3 := P_2 \cap P_\Delta$. The set P_3 is the shaded gray region on the figure¹⁰.

Let us summarise, in the form of an algorithm, the proposed polytopic bounding, see Algorithm 2. Note: in the algorithm, we denote as $\text{vec} : \mathbb{R}^{m \times s} \mapsto \mathbb{R}^{1 \times (ms)}$ the vectorisation operation so that a matrix is transformed to a row vector by concatenating its rows. The inverse operation (building the matrix again), will be denoted as vec^{-1} .

7.4 Examples

Consider the following nonlinear system, assuming $z = x$ in (7.2):

$$\begin{aligned} \dot{x} &= \begin{pmatrix} -2.6x_1 + 0.7x_2 - 0.5u + 1.2v_1 - 1.7v_2 + w_1 \\ -3.4x_1 - 3.4x_2 + 0.9u - 0.1v_1 + 0.5v_2 + w_2 \end{pmatrix} \\ Y_\infty &= \begin{pmatrix} 0.9x_2 - 0.5x_1 \\ u \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 0.9x_2 + 0.5x_1 \\ u \end{pmatrix}, \\ \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= \begin{pmatrix} \rho_1(x) \\ \rho_2(x) \end{pmatrix} = \begin{pmatrix} x_1x_2 \\ \sin(0.4x_1x_2) \end{pmatrix} \end{aligned} \quad (7.35)$$

¹⁰Obviously, the polyhedral bound could be further trimmed by actually intersecting P_1 and P_2 or obtaining supporting hyperplanes in more directions; nevertheless, that would increase the number of vertices but their effect on the performance (for small Ω_z) would likely not be significant, according to the ideas in Lemma 7.1. Thus, the recommended option is the set P_3 , and it will be used that way in the examples.

Algorithm 2 Performance-aware polytopic embedding

1: Obtain $\Delta\Gamma$ from the linearised problem using Algorithm 1, as well as the explicit expression of $Q(z)$.

2: Denote

$$\Theta := \{vec(Q(z)) : z \in \Omega_z\} \subset \mathbb{R}^{ms}$$

Obtain the covariance matrix of Θ , defined as $\Sigma = \int_{\Theta} \phi^T \phi d\phi$, see (Kwiatkowski and Werner 2008b). The dimensions of Σ are $ms \times ms$.

3: Obtain the eigenvalue decomposition $\Sigma = V\Lambda V^T$.

4: Denoting the i -th column of V as $V^{[i]}$, compute the bounds:

$$\lambda_{i1} = \inf_{\phi \in \Theta} \phi V^{[i]} \quad \lambda_{i2} = \sup_{\phi \in \Theta} \phi V^{[i]}$$

5: Using the above bounds and (7.33), form the polytopic subset of \mathbb{R}^{ms} given by:

$$P_3 := \{l_1 \leq vec(\Delta\Gamma)\phi \leq l_2, \quad l_{i1} \leq (V^{[i]})^T \phi \leq l_{i2}, i = 1, \dots, ms\}$$

6: Obtain the vertices of P_3 , so $P_3 = Co(\{v^{[1]}, \dots, v^{[j]}\})$ for some finite j .

7: Obtain the polytopic bounding of $Q(\Omega_z)$ inverting the vectorisation, so we can state that:

$$Q(z) \in Co(\{vec^{-1}(v^{[1]}), \dots, vec^{-1}(v^{[j]})\}) \quad \forall z \in \Omega_z$$

Once a quasi-LPV representation $v = Q(x)x$ is available, in this example, the performance objective will be minimising γ subject to conditions (7.37)–(7.41) on top of next page, being

$$A(Q) = \begin{pmatrix} -2.6 & 0.7 \\ -3.4 & -3.4 \end{pmatrix} + \begin{pmatrix} 1.2 & -1.7 \\ -0.1 & 0.5 \end{pmatrix} Q \quad (7.36)$$

where $A(Q)$ comes from trivially rewriting the state equation in (7.35) in matrix form, as in (7.1).

$$X \succeq 0 \quad (7.37)$$

$$\gamma - 3\gamma_{\infty} - \gamma_2 \geq 0 \quad (7.38)$$

$$\begin{pmatrix} A(Q)X + BF(Q) + (A(Q)X + BF(Q))^T & (C_2X + D_2F(Q))^T \\ C_2X + D_2F(Q) & -\gamma_2 I \end{pmatrix} \preceq 0 \quad (7.39)$$

$$(7.40)$$

$$\begin{pmatrix} A(Q)X + BF(Q) + (A(Q)X + BF(Q))^T & (C_\infty X + D_\infty F(Q))^T & E^T \\ C_\infty X + D_\infty F(Q) & -\gamma_\infty I & 0 \\ E & 0 & -\gamma_\infty I \end{pmatrix} \preceq 0 \quad (7.41)$$

This problem verifies the conditions in Assumption 7.2, with $D_1 \equiv X$, $D_2 \equiv F(Q)$, indicating that a gain-scheduled state feedback controller $u(Q, x) := F(Q)X^{-1}x$ is sought, being $V(x) := x^T X^{-1}x$ a common quadratic Lyapunov function, independent of Q .

If the above-referred conditions are feasible, they guarantee that $\|Y_\infty\|_2 \leq \gamma_\infty \|w\|_2$ under zero initial conditions and, too, that with $w = 0$ but nonzero initial state $x(0)$, $\|Y_2\|_2 \leq \gamma_2 x^T(0)x(0)$, where 2-norm of signals are understood in the integral sense $\|y\|_2^2 = \int_0^\infty y(t)^T y(t) dt$. In the proposed problem, the importance weight of the γ_∞ performance bound has been arbitrarily set to be three times that of γ_2 , in order to write (7.38).

Solving the above problem for the linearized system ($Q = 0$), the obtained optimal linearised performance is $\gamma^{\text{lin}} = 1.4121$, jointly with Lyapunov function $D_1^{\text{lin}} \equiv X^{\text{lin}}$ and linear controller decision variable $D_2^{\text{lin}} \equiv F^{\text{lin}}$, omitted for brevity; so, Assumption 7.3 is fulfilled. The used LMI software was SeDuMi 1.3 in Matlab R2014a. With these variables, obtaining the suitable basis H of the relevant null space of the optimal LMIs $\Omega(Q_z)$, and checking four vertex values for Q :

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

on the incremental version of the LMIs, sensitivity matrix $\Delta\Gamma$ in Lemma 7.1 results in

$$\Delta\Gamma = \begin{bmatrix} 0.6435 & -0.7654 \\ -0.6333 & 0.7739 \end{bmatrix} \quad (7.42)$$

Hence, the proposal in this work suggests that the nonlinearities ρ_1 and ρ_2 are modelled for minimal projection onto the subspace spanned for the first and second row of $\Delta\Gamma$, respectively, following Algorithm 1. As a result, the finally proposed factorisation is (7.43), understanding the fractions at $\xi_3 = 0$ or $\xi_4 = 0$ in a limit sense, see the Appendix.

$$\begin{aligned}
\rho_1(x) &= \begin{pmatrix} 0.493\xi_1 - 0.172\xi_2 & -0.493\xi_2 \\ \xi_2 \end{pmatrix}, \\
\rho_2(x) &= \begin{pmatrix} \frac{\sin(0.196\xi_4^2) + \sin(\lrcorner)}{\xi_3} & -\frac{\sin(0.196\xi_4^2)}{\xi_4} \end{pmatrix} \begin{pmatrix} \xi_3 \\ \xi_4 \end{pmatrix}, \\
\xi_1 &:= 0.7654x_1 + 0.6435x_2, \quad \xi_2 := \Delta\Gamma^{[1]}x = 0.6435x_1 - 0.7654x_2, \\
\xi_3 &:= 0.7739x_1 + 0.6333x_2, \quad \xi_4 := \Delta\Gamma^{[2]}x = -0.6333x_1 + 0.7739x_2, \\
\lrcorner &:= 0.196\xi_3^2 + 0.0792\xi_3\xi_4 - 0.196\xi_4^2
\end{aligned} \tag{7.43}$$

The proposal in Algorithm 1 to generate $Q(x)$ will be compared with other ones in literature.

Gain-scheduled controller parametrisation. In gridding-based approach, one controller gain will be sought for each grid point. In later polytopic-based embeddings, one controller gain F_i will be obtained for each vertex: given the structure of LMIs (7.39) and (7.41), it is well-known (De Caigny et al. 2010) that these vertex gains can be used to synthesise a gain-scheduled controller with the same interpolation coefficients as the model matrices.

7.4.1 Comparison with alternative options for $Q(z)$

In order to compare with alternative approaches, a grid of 481 points $\{x^{[k]}, k = 1, \dots, 481\}$ over a modelling region Ω_x being a circle of radius r in the state space \mathbb{R}^2 has been chosen: the origin plus a grid of 480 points in 10 radius values $\{0.1, 0.2, \dots, 1\} \times r$ and 48 equally-spaced angle steps. In gridding-based approaches, the proposed $Q(z)$ would be evaluated at these 481 points, and compared with other possibilities for $Q(z)$; of course, later on, gridding will be avoided by embedding $Q(z)$ in a polytope with a small number of vertices. Obviously, a naive scaling transforms Ω_z to the unit ball required in the assumptions.

Annihilator approach. Let us first compare with BMI+gridding annihilator approaches in earlier literature (Coutinho et al. 2008). A factorisation $v = (Q_0(x) + \aleph(x))x$ has been used, for any annihilator $\aleph(x)$ such that $\aleph(x)x = 0$. We chose $Q_0(x)$ to be the one suggested by our proposed subspace approach. The choice for $\aleph(x)$, at each grid point, was the linear annihilator:

$$\aleph(x^{[k]}) := \begin{pmatrix} -l_1^{[k]}x_2^{[k]} & l_1^{[k]}x_1^{[k]} \\ -l_2^{[k]}x_2^{[k]} & l_2^{[k]}x_1^{[k]} \end{pmatrix} \quad (7.44)$$

being $l_1^{[k]}$ and $l_2^{[k]}$ decision variables for each grid point. The total number of BMI decision variables is 1928: 481×2 for $l_i^{[k]}$, 481×2 for $F^{[k]} \in \mathbb{R}^{1 \times 2}$, three for the symmetric x , plus γ , γ_∞ and γ_2 , and the total number of LMIs arise from (7.37), (7.38) and repeating 481 times, for each $Q(x^{[k]})$ the LMIs (7.39) and (7.41); this can be expressed as a big block-diagonal BMI of size 4813×4813 . For the sake of comparison, in our approach to search for the best $Q(z)$, each element of $\Delta\Gamma$ was obtained with a 13×13 LMI, with 8 decision variables.

Note that, starting with $l_1^{[k]} = l_2^{[k]} = 0$, the resulting BMI search, if properly converged, should always yield a better performance bound than our proposal: this is intentional, as the issue is determining whether the achieved improvement is significant or worthwhile, given the increase of computational cost arising from the iterative BMI steps. Results will be later discussed after other alternatives options are presented.

Inspection-based factorisation. For the sake of comparison, four other easy “common sense factorisations” will be evaluated, either extracting x_1 or x_2 as a factor, as follows:

$$v = \begin{pmatrix} \frac{\rho_1}{x_1} & 0 \\ \frac{\rho_2}{x_1} & 0 \end{pmatrix} x = \begin{pmatrix} x_2 & 0 \\ \frac{\sin(0.4x_1x_2)}{x_1} & 0 \end{pmatrix} x \quad (7.45)$$

$$v = \begin{pmatrix} 0 & \frac{\rho_1}{x_2} \\ \frac{\rho_2}{x_1} & 0 \end{pmatrix} x = \begin{pmatrix} 0 & x_1 \\ \frac{\sin(0.4x_1x_2)}{x_1} & 0 \end{pmatrix} x \quad (7.46)$$

$$v = \begin{pmatrix} \frac{\rho_1}{x_1} & 0 \\ 0 & \frac{\rho_2}{x_2} \end{pmatrix} x = \begin{pmatrix} x_2 & 0 \\ 0 & \frac{\sin(0.4x_1x_2)}{x_2} \end{pmatrix} x \quad (7.47)$$

$$v = \begin{pmatrix} 0 & \frac{\rho_1}{x_2} \\ 0 & \frac{\rho_2}{x_2} \end{pmatrix} x = \begin{pmatrix} 0 & x_1 \\ 0 & \frac{\sin(0.4x_1x_2)}{x_2} \end{pmatrix} x \quad (7.48)$$

Trivially, these options can give a simple 4-vertices polytopic embedding representation, later evaluated.

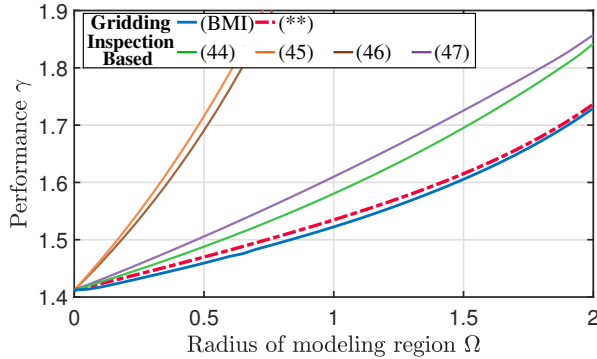


Figure 7.2: Performance comparison of choices for $Q(z)$, gridding approach. Our proposal from Algorithm 1, i.e., $Q(z)$ in (7.43), labelled as (**), achieves a very good performance with significantly less computational cost than the (BMI) one.

Comparison of results (gridding approach). Comparing all options for $Q(z)$ in the same grid of 481 points results in Figure 7.2. It can be clearly seen that our proposal for $Q(z)$, labelled as (**), provides the best results except the marginal improvements of the costly annihilator-BMIs, as expected (at least for small modelling region radius). The common-sense factorisations above incur in a clear performance penalty. Note that, as intuitively expected, performance is close to the linearised one (1.4121) for small modelling region radius, but it worsens as such radius increases (indeed, the set $Q(\Omega_z)$ grows larger as the radius of Ω_z expands).

Computation time (gridding approach). On a Core I5-4690 processor, computation of $\Delta\Gamma$ with our LMI proposals takes around 0.088 seconds. Once $Q(z)$ is crafted from the resulting $\Delta\Gamma$, solving the LMIs with SeDuMi 1.3 over the 481-point grid to obtain the new gain-scheduled controller takes 8.76 seconds. On the other hand, using PenBMI2.1 32 bit with default configuration, a 481-point grid takes around 420 seconds in average for each tested radius.

Note that the computational cost would exacerbate for higher-order systems because of the larger dimension of the matrix inequalities and the likely need of an exponentially larger number of grid points; thus, polytopic bounding to avoid gridding is important, and it will be discussed below in the context of this example.

As a conclusion, our subspace approach has two worthwhile advantages (at least in this example): first, it clearly outperforms *ad-hoc* inspection-based factorisa-

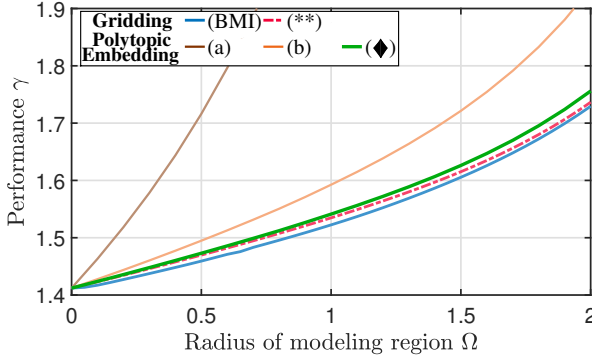


Figure 7.3: Performance degradation due to different options for polytopic embedding, for the fixed $Q(z)$ arising from Algorithm 1 in (7.43). See enumeration in the text for label interpretation. Our proposal (c) achieves comparable performance with a 10-fold improvement in computational time over the gridding results (**) and 700-fold over (BMI), both copied from Figure 7.2.

tions; second, BMI+gridding options in other literature achieve only marginal improvements and the computational cost is much higher.

7.4.2 Effect of polytopic embedding

Once $Q(z)$ is chosen from our proposal, in order to avoid gridding, performance of polytopic embeddings will be tested. On Figure 7.3, the three different approaches discussed on Section 7.3.2 are compared with the gridding-based “best-possible” performance¹¹. First, the simpler options:

- (a) the 16-vertices box arising from an orthogonal transformation induced by worst-performance directions in $\Delta\Gamma$, using (7.53); the methodology is analogous to the one yielding P_2 in the simplified example on Figure 7.1.
- (a) the PCA approach from (Kwiatkowski and Werner 2008b) (with, too, 16 vertices), symbolically depicted as P_1 on Figure 7.1.

and, finally, our proposed combined approach, labelled as (♦) in the Figure, intersecting the principal-component box with the supporting hyperplanes (7.33),

¹¹If the grid were “dense enough”; indeed, otherwise, such an estimate would be optimistic if leaving out the worst-case points from the grid. In this example, denser grids were tested with no appreciable performance improvement.

Table 7.1: Performance versus computation time (radius 2)

Method	Num. of vertices	Computation time (s)	Performance (penalty%)
Earlier literature:			
BMI-Gridding (Coutinho et al. 2008)	Grid 481	420	1.729
LMI-Gridding (7.45)	Grid 481	8.73	1.842 (6.5%)
PCA bound (7.45)	4	0.14	1.929 (11.5%)
inf/sup bound of (7.45)	4	0.11	11.60 (570%)
LMI-Gridding (7.46)	Grid 481	7.14	3.832 (121%)
LMI-Gridding (7.47)	Grid 481	7.19	4.819 (178%)
LMI-Gridding (7.48)	Grid 481	6.68	1.857 (7.4%)
Proposals in this work:			
Alg. 1 $Q(z)$ +LMI	Grid 481	8.76	1.742 (0.7%)
Alg. 1+Alg. 2. LMI	24	0.56	1.756 (1.5%)

which results in a polytope with 24 vertices (symbollically depitec as P_3 on Figure 7.1).

In order to better assess the performance loss due to the poytopic embedding, the original BMI+grid approach and the LMI+grid on our proposed $Q(z)$ have been copied over from Figure 7.2 to Figure 7.3, and additional lines added.

Computation time comparison. As intuitively expected, polytopic embedding achieves more conservative performance bounds than gridding options, but at a much lower computational cost.

Table 7.1 depicts an overall comparison between the different approaches showing computation time and performance (and the number of considered vertices/grid points). Let us discuss the different table entries below:

The first rows describe proposals in prior literature. Row 1 lists data for the already discussed BMI-gridding, whose performance is, obviously, the best one. Rows 2 to 4 use the inspection-based factorisation (7.45): directly on a grid (row 2), bounding it with the PCA-based approach in (Kwiatkowski and Werner 2008b) (row 3), or using the most straightforward bounding by computing the minimum

and maximum over Ω_z of the matrix elements ¹² (row 4). Rows 5 to 7 are the results of alternative inspection-based options.

The last two rows present the data from our proposals: first, a gridding approach (row 8) to point out the effect of the choice of $Q(z)$ with respect to the other gridding options (rows 1, 2, 5, 6, 7). Apart from the reference row 1 (BMI), the performance from our proposal beats by a large margin the rest of the referred rows. Finally, row 9 presents the combination of our choice of $Q(z)$ plus the polytopic embedding from Algorithm 2, with a minor performance loss with respect to rows 1 and 8, and a significantly faster execution.

In conclusion, our proposals yield 0.7% performance penalty with a 46 times lower computation time (Algorithm 1 with LMI gridding), and 1.5% performance penalty with more than 700 times lower computation time (Algorithm 2, polytopic embedding) with respect to the ideal BMI-gridding results. These computational advantages would, likely, accentuate for higher dimensions.

On the other hand, for sake of comparison, the inspection-based polytopic embeddings were unable to approach the optimal performance points. For instance, (7.45) was the best factorisation when evaluated on the 481-point grid, yielding a performance of 1.84 (6.4% penalty); however, the 4-vertices PCA bound of (7.45) incurred in a 11.5% penalty, and a much larger penalty figure appeared for the straightforward inf/sup bound of each matrix component over the circle Ω_z . Similar results were achieved with the other options (7.46)–(7.48), omitted for brevity.

7.5 Conclusions

This chapter has addressed the problem of choosing a good quasi-LPV model for a nonlinear system. Our proposals are able to mitigate the performance loss from a reference linearised design, compared to other “inspection-based” modelling choices, while avoiding BMIs needed in prior literature for controller synthesis problems. The main idea is based on obtaining the more harmful perturbations of the model matrices in the linearised design, proposing then a coordinate transformation so that the projection of the uncertainty over these model perturbations is minimised. This direction plays, too, a fundamental role in later polytopic embedding options which are built to avoid the need of gridding. As the root of the

¹²This approach is routinely proposed in many applications (Shamma and Cloutier 1992; Kwiatkowski and Werner 2008b); in this particular case, in the ball of radius 2 we can bound: $-2 \leq x_2 \leq 2$, and $-0.8 \leq \frac{\sin(0.4x_1x_2)}{x_1} \leq 0.8$ yielding a 4-vertices polytopic LPV model, details left for the reader.

procedure is the linearised design, the results are only valid, formally, for small modelling regions around the origin.

Appendix: Systematic factorisation of a nonlinearity

Let us consider a nonlinearity $\tilde{\rho} : \mathbb{R}^s \rightarrow \mathbb{R}$:

$$v = \tilde{\rho}(\eta) \quad (7.49)$$

and present a factorisation option inspired in Chapter 4 (Robles et al. 2015) and (Gao et al. 2012) and the proof of Proposition 7.3.

Let us define $\xi_1 := \tilde{\rho}(\eta_1, \dots, \eta_s)$ and $\xi_i(\eta_i, \dots, \eta_s) := \tilde{\rho}(0, \dots, 0, \eta_i, \dots, \eta_s)$ for $i = \{2, \dots, s\}$, $\xi_{s+1} := 0$ and let us denote, for $i = \{1, \dots, s\}$:

$$\zeta_i(\eta_i, \dots, \eta_s) := \xi_i(\eta_i, \dots, \eta_s) - \xi_{i+1}(\eta_{i+1}, \dots, \eta_s) \quad (7.50)$$

so, as $\zeta_i(0, \eta_{i+1}, \dots, \eta_s) = 0$ and $\zeta_i(\cdot) \in \mathcal{C}^1$, by Proposition 7.1, there exists a factorisation

$$\zeta_i(\eta_i, \dots, \eta_s) = \frac{\zeta_i(\eta_i, \dots, \eta_s)}{\eta_i} \cdot \eta_i$$

taking suitable limits when $\eta_i \rightarrow 0$. Thus, as trivially:

$$\xi_i(\eta_i, \dots, \eta_s) = \zeta_i(\eta_i, \dots, \eta_s) + \xi_{i+1}(\eta_{i+1}, \dots, \eta_s)$$

with these definitions, we can express:

$$\tilde{\rho}(\eta) = \sum_{i=1}^s \zeta_i(\eta_i, \dots, \eta_s) = \sum_{i=1}^s \frac{\zeta_i(\cdot)}{\eta_i} \cdot \eta_i \quad (7.51)$$

Considering the expression in the form (7.32) below:

$$Q(\eta) = \frac{\zeta_s}{\eta_s} \begin{pmatrix} 0 & \dots & 0 & 1 \end{pmatrix} + \begin{pmatrix} \frac{\zeta_1}{\eta_1} & \dots & \frac{\zeta_{s-1}}{\eta_{s-1}} & 0 \end{pmatrix} \quad (7.52)$$

it can be easily proved that the resulting $Q(\eta)$ is tight on the subspace generated by $(0 \ \dots \ 0 \ 1)^T$, details left to the reader.

Polytopic embedding via inf/sup bounds. A straightforward embedding with $n_v=2^s$ vertices can be built by obtaining the infimum and supremum of $\zeta_i(\eta_i, \dots, \eta_s)/\eta_i$ for $z \in \Omega_z$, denoted as $\underline{\zeta}_i^*$ and $\overline{\zeta}_i^*$, respectively (assuming they exist). Indeed, straightforward interpolation allows us to express:

$$\tilde{\rho}(\eta) = \sum_{i=1}^s \left(\mu_{i1}(\eta) \underline{\zeta}_i^* + \mu_{i2}(\eta) \overline{\zeta}_i^* \right) \eta_i \quad (7.53)$$

with $\mu_{i1} + \mu_{i2} = 1$, $\mu_{i1} \geq 0$, $\mu_{i2} \geq 0$. Apart from this “box” embedding, other possibilities exist, as discussed in Section 7.3.2 on the main text.

Gain-Scheduled Control via Polytopic Difference Inclusions

Many works in literature are based on the idea that a nonlinear system with sector-bounded nonlinearities can be expressed as a quasi-LPV system (convex combination of linear models, with state-dependent interpolation coefficients). The convex difference inclusion (CDI) modelling framework proposed by M. Fiacchini and coworkers in several of their works generalises the modelling procedure and proposes robust controllers for them. This work generalises the approach allowing for gain-scheduled controllers applicable to a wider class of models than the quasi-LPV ones. As most set-based approaches, the proposal is tractable in low-dimensional cases.

This chapter proposes a radical departure from the modelling approaches in the previous contributions. Its objective is to address a larger class of models than the TS one discussed until now. Instead of searching for the optimal model given some LMIs, a set-based approach will be pursued to show that larger domain of attraction estimates are possible with the new modelling framework.

Its contents form the basis of a journal article, currently under review.

8.1 Introduction

Determining invariant and contractive sets for dynamic systems is a control problem which has been addressed in a variety of settings (Blanchini 1999). Such sets can be easily determined for linear systems and uncertain linear systems characterised as polytopic Linear difference inclusions (LDI) (Kouramas et al. 2005), either from a set-based approach (Blanchini 1999; Mayne 2001; Kerrigan 2001; Rakovic et al. 2004; Herceg et al. 2013) or in a linear matrix inequality framework (LMI), (Boyd and Vandenberghe 2004; Guerra and Vermeiren 2004; Sala 2009; Wu and Dong 2006b). As some nonlinear systems can be embedded in a polytopic LDI via a systematic modelling procedure (quasi-LPV embedding (Rugh and Shamma 2000a; Bianchi, Mantz, and Christiansen 2005b; Kwiatkowski and Werner 2008a), Takagi-Sugeno models (Tanaka and Wang 2004; Sala 2009)), the above results apply to some robust nonlinear control problems and gain-scheduled extensions, with, of course, a dose of conservatism.

This work will pursue an approach based in polyhedron manipulation, as LMIs do not apply to an extended class of models to be discussed below. The works (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2010; Fiacchini, Alamo, and Camacho 2012) show that set-based LDI ideas can be generalised to the so-called convex differential inclusions (CDI), which are more general models capable of representing nonlinear and uncertain systems with lower conservatism (overbounding) than LDI, being these a particular case. Furthermore, they present iterative scaling/shooting algorithms to compute contractive sets, both in analysis and *robust* control design settings. In the polytopic case, such steps can be carried out with polytope manipulation software such as the multiparametric toolbox (MPT, (Herceg et al. 2013)). The preliminary conference paper (Sala 2017) outlines a generalisation of some concepts to a quasi-convex setup so that the above-referred algorithms still work.

The goal of this chapter is generalising the ideas in (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2010; Fiacchini, Alamo, and Camacho 2012; Sala 2017), in order to incorporate gain-scheduling options. Definitions of shape-independent gain-scheduled one-step sets will be given to generalise the standard one-step set definitions in the cited works, and they will be used to obtain invariant/contractive estimates achievable with gain-scheduling controllers. In this way, set-based gain-scheduling developments for LDI (Ariño et al. 2017) will be generalised to the wider class of convex and quasi-convex DI models.

8.2 Preliminaries

This chapter will consider a discrete-time dynamic system where the successor of a current state $x \in \mathbb{R}^n$, denoted by x_+ , is given by:

$$x_+ = f(x, d, u) \quad (8.1)$$

being u is the so-called input vector, with values $u \in \mathbb{U} \subset \mathbb{R}^m$, and $d \in \mathbb{D} \subset \mathbb{R}^v$ a set of time varying parameters or disturbances. In many cases an uncertain “*embedding model*” of the dynamics (8.1) can be built so that an inclusion condition can be stated as follows:

$$f(x, d, u) \in \mathcal{F}(x, u) \quad \forall d \in \mathbb{D} \forall x \in \mathbb{X} \forall u \in \mathbb{U} \quad (8.2)$$

where $\mathcal{F}(x, u) \subset \mathbb{R}^n$ is the image of (x, u) under a known set-valued map, $\mathcal{F} : \mathbb{X} \times \mathbb{U} \mapsto \mathcal{P}(\mathbb{R}^n)$ being $\mathcal{P}(\cdot)$ the power set. The set \mathbb{X} will be denoted as the *modelling region*. Under this embedding $\mathcal{F}(x, u)$ is a *set* of possible successor states. The motivation towards using these uncertain embeddings is that they will exchange conservatism for computational advantages, if \mathcal{F} fulfills some convexity-related properties, to be later discussed.

Uncertain dynamic systems in the form:

$$x_+ \in \mathcal{F}(x, u) \quad (8.3)$$

are denoted as difference inclusions (DI) (Fiacchini, Alamo, and Camacho 2010). If there is no input vector $x_+ \in \mathcal{F}(x)$, the DI is called autonomous. Of course, if $\mathcal{F}(x, u)$ were a single point, the above DI would reduce to a deterministic dynamic system.

Assumption 8.1. *We will assume that the origin ($x = 0, u = 0$) is an equilibrium point of the undisturbed system, i.e. $f(0, 0) = 0$ and, additionally, that $0 \in \mathbb{X}$, $0 \in \mathbb{U}$ and $0 \in \mathbb{D}$. Also, \mathbb{X} , \mathbb{U} and \mathbb{D} will be assumed to be compact polyhedra.*

Convexity-related properties of set-valued maps. Let us first introduce some notation: Given a set $\Sigma \subset \mathbb{R}^n$ and a matrix M with compatible dimensions (or a real scalar), $M\Sigma := \{z : \exists s \in \Sigma \text{ s.t. } z = Ms\}$ will denote the linear scaling of it. Minkowsky sum of two sets will use notation: $A \oplus B := \{z \in \mathbb{R}^n : \exists a \in A, b \in B \text{ s.t. } z = a + b\}$. Sets of matrices (or vectors) will be also considered. Given sets of matrices \mathbf{C} and \mathbf{D} with compatible dimensions, its product $\mathbf{C} \otimes \mathbf{D}$ will be understood analogously to the Minkowski sum, replacing addition by matrix product.

Let us consider a set-valued map $\mathcal{F} : \mathcal{D} \mapsto \mathcal{P}(\mathbb{R}^n)$; the domain \mathcal{D} will be assumed to be a convex set in the sequel. The image of a set $\Omega \subseteq \mathcal{D}$ under a set-valued map \mathcal{F} will be understood as $\mathcal{F}(\Omega) := \cup_{\xi \in \Omega} \mathcal{F}(\xi)$.

Definition 8.1. A set-valued map \mathcal{F} is **convex-valued** if $\mathcal{F}(\xi)$ is convex for every $\xi \in \mathcal{D}$. Additionally, it will be denoted as **polytopic** if $\mathcal{F}(\xi)$ is a compact polytope for all ξ in \mathcal{D} .

Definition 8.2. A set-valued map is **linear** if $\mathcal{F}(\lambda\xi) = \lambda\mathcal{F}(\xi)$ and $\mathcal{F}(\lambda_1\xi_1 + \lambda_2\xi_2) = \lambda_1\mathcal{F}(\xi_1) \oplus \lambda_2\mathcal{F}(\xi_2)$, for all $\lambda, \lambda_1, \lambda_2$ in \mathbb{R} .

Definition 8.3. A set-valued map is **S-convex** if $\mathcal{F}(\lambda\xi_1 + (1 - \lambda)\xi_2) \subseteq \lambda\mathcal{F}(\xi_1) \oplus (1 - \lambda)\mathcal{F}(\xi_2)$ for all ξ_1, ξ_2 in \mathcal{D} and for all $0 \leq \lambda \leq 1$.

Definition 8.4. A set-valued map is **S-quasiconvex** if $\mathcal{F}(\lambda\xi_1 + (1 - \lambda)\xi_2) \subseteq \text{Co}(\mathcal{F}(\xi_1) \cup \mathcal{F}(\xi_2))$ for all ξ_1, ξ_2 in \mathcal{D} and for all $0 \leq \lambda \leq 1$.

Definition 8.5. A set-valued map is **sub-homogeneous** if $\mathcal{F}(\lambda\xi) \subseteq \lambda\mathcal{F}(\xi)$, for all $\xi \in \mathcal{D}$ and for all $\lambda \in [0, 1]$.

Evidently from the definitions, all linear set-valued maps are S-convex, and all S-convex ones are S-quasiconvex¹. Also, linear maps are sub-homogeneous and so they are S-convex maps if, additionally, $\mathcal{F}(0) = \{0\}$ holds.

Theorem 8.1. Given a polytope P , and an S-quasiconvex set-valued map \mathcal{F} , we have

$$\mathcal{F}(\xi) \subseteq \Xi := \text{co} \left(\bigcup_{\zeta \in \text{vert}(P)} \mathcal{F}(\zeta) \right), \quad \forall \xi \in P$$

Proof. Denote the vertices of P as $\{\zeta_1, \dots, \zeta_r, \dots, \zeta_{N_P}\}$. Obviously, the image of the vertices fulfils the inclusion stated in the theorem, and so for the image of the convex combinations of two vertices directly from Definition 8.4. Now, let us assume that the convex combination of $r - 1$ vertices fulfills the assertion. Any point $\xi \in P$ which is not a vertex and is a convex combination of r vertices can be expressed, for some $\eta_i(x) > 0, i = 1, \dots, r$ as

$$\xi = \sum_i^r \eta_i(\xi) \zeta_i = \eta_r \zeta_r + (1 - \eta_r) \hat{\zeta}_{r-1} \quad (8.4)$$

¹Notions of convexity and quasi-convexity of set-valued maps are discussed in, for instance, (Kuroiwa 1996) so that they extend the concepts for real- and vector-valued mappings in (Jeyakumar, Oettli, and Natividad 1993). These notions are such that the definitions, when $\mathcal{F}(\xi)$ is a single point, reduce to the standard convex and quasiconvex function cases. However, in order to apply later results in this manuscript, stronger conditions than these in the mentioned works must be posed, in definitions 8.3 and 8.4.

being

$$\hat{z}_{r-1} := \left(\sum_{i=1}^{r-1} \frac{\eta_i(x)}{1 - \eta_r(\xi)} \zeta_i \right) \in P \quad (8.5)$$

i.e., as the convex combination of a vertex of P and a point \hat{z}_{r-1} which is a convex combination of $r - 1$ vertices. By the induction assumption, $\mathcal{F}(\hat{z}_{r-1}) \subseteq \Xi$; thus, as $\mathcal{F}(\zeta_r) \subseteq \Xi$, Definition 8.4 entails $\mathcal{F}(\xi) \subseteq \Xi$. As it is true for $r = 2$ and $r = 3$, so it is for any integer r . As every point of P can be expressed as a convex combination of its vertices, the proof is complete. \square

The theorem enables to bound the image of a polyhedron $\mathcal{F}(P)$ by just computing the image of its vertices.

Classification of DI. The properties of the set-valued map $\mathcal{F}(x, u)$ associated to a DI $x_+ \in \mathcal{F}(x, u)$, defined in $\mathcal{D} := \mathbb{X} \times \mathbb{U}$, will define different types of DI. Based on the nomenclature in (Kerrigan 2001; Fiacchini, Alamo, and Camacho 2010; Sala 2017), a difference inclusion (8.3) will be denoted as linear (LDI), uncertain-convex (UCDI), or uncertain quasi-convex (UQCDI) if \mathcal{F} is linear, S-convex, or S-quasiconvex, respectively². Additionally, a convex DI (CDI), as defined in (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2012) will be understood as having an associated S-convex and sub-homogeneous $\mathcal{F}(x, u)$.

Embedding a nonlinear model onto a DI. In literature, embedding a nonlinear system onto a convex combination of linear “vertex” models is a widely-used approach to prove properties of nonlinear systems based on the properties of the linear vertices, see the quasi-LPV or Takagi-Sugeno modelling approaches in (Tanaka and Wang 2004; Rugh and Shamma 2000a; Kwiatkowski and Werner 2008a), for instance. In particular, given a model in the form:

$$x_+ = [A(x) \quad G(x) \quad E(x)] \begin{pmatrix} x \\ u \\ d \end{pmatrix} := \Psi(x) \begin{pmatrix} x \\ u \\ d \end{pmatrix} \quad (8.6)$$

the above references show that a polyhedral bound for Ψ in a compact modelling region \mathbb{X} can be found under mild continuity assumptions. Let us denote such

²The approach to set-valued maps in (Fiacchini, Alamo, and Camacho 2010; Fiacchini, Alamo, and Camacho 2012) is strongly based on the so-called support function of a non-empty, compact set (Fiacchini, Alamo, and Camacho 2010, Def. 1). The equivalence between that approach and the above definitions is discussed in Appendix 8.8.2.

polyhedral bound as $\widehat{\Psi} := \text{Co}(\{(A_i \ G_i \ E_i) \ i = 1, \dots, N_\Psi\})$, being N_Ψ the number of vertices, such that $\Psi(x) \in \widehat{\Psi}$ for all $x \in \mathbb{X}$. Elementary convexity argumentations prove that the above nonlinear system can be rewritten as a parameter-dependent expression:

$$x_+ = \sum_{i=1}^{N_\Psi} h_i(x)(A_i x + G_i u + E_i d) \quad (8.7)$$

being h a parameter vector ranging in the standard simplex $\Delta := \{h \in \mathbb{R}^{N_\Psi} : \sum_{i=1}^{N_\Psi} h_i = 1, h_i \geq 0\}$. Abusing the notation, in the sequel Δ will denote the standard simplex in any vector space, not just \mathbb{R}^{N_Ψ} . Disregarding the actual “shape” of $h(x)$, we can obtain an embedding in DI form (8.2):

$$x_+ \in \text{Co}(\{A_i x + G_i u\} \oplus E_i \mathbb{D}, \ i = 1, \dots, N_\Psi) \quad (8.8)$$

It can be proved that the above DI is linear (LDI) if $\mathbb{D} := \{0\}$ (undisturbed case) and UCIDI in the case D is a compact convex set.

There are other more general techniques to embed a nonlinear system onto a CDI, UCIDI, QCDI, \dots , see (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2010; Sala 2017). For completeness, Appendix 8.8.3 in this manuscript recalls some of these, but the reader is referred to the original sources for further discussion in modelling issues, important but out of the scope of this thesis for brevity. For instance, the 1st-order dynamic system $x_+ = x^3$ can be embedded in $x_+ \in [f_l(x), f_u(x)]$ with $f_l(x) = \min(x^3, 0)$ and $f_u(x) = \max(x^3, 0)$. As f_l is concave and f_u is convex, and $f_l(0) = f_u(0) = 0$ the embedding is a CDI (Proposition 8.2). Actually, as x^3 is quasilinear, the embedding $x_+ \in \{x^3\}$ is a UQCDI in which no conservatism in the modelling step has arisen.

8.3 Set-based approach to dynamic system analysis and control design

A plethora of results exist for DIs in the form (8.8) proving stability (LDI case), finding inescapable sets, state-feedback controller design, etc. both using set-manipulation software (Kerrigan 2001; Herceg et al. 2013) and convex linear matrix inequalities (LMI), see for instance (Boyd and Vandenberghe 2004; Guerra and Vermeiren 2004; Sala 2009; Wu and Dong 2006b) and references therein. However, LMIs do not easily apply to more general³ non-LDI models (CDI, UCIDI,

³Another way of obtaining more general models is using polynomial bounds and using Sum-of-squares convex programming techniques for them (Prajna, Papachristodoulou, and Wu 2004; Chesi 2009; Sala and Pitarch 2016b). The sum-of-squares generalisation of LMIs, however, will not be considered in the scope of this thesis, either.

UQCIDI) so, intentionally, the scope of this chapter will be directed to the set-based approach.

The advantages of the set-based approach are its applicability to more general systems than (8.8) –discussed later on–, as well as the ability to naturally consider saturation and possible non-symmetric constraints in state and input. The disadvantage comes from the need of vertex enumeration steps, which make it impractical for high-dimensional systems in a general case. The basic definitions in the set-based approach are recalled below.

Definition 8.6 (one-step set). *Given a target set Ω , the disturbance-rejection one-step set of Ω under model (8.1) is defined as:*

$$\mathcal{Q}(\Omega) := \{x \in \mathbb{R}^n : \exists u \in \mathbb{U} \text{ s.t. } f(x, d, u) \in \Omega \forall d \in \mathbb{D}\} \quad (8.9)$$

The above definition implies that there exists a nonlinear state-feedback law $u = \nu(x)$ such that for all $x \in \mathcal{Q}(\Omega)$ and for all $d \in \mathbb{D}$, $f(x, d, \nu(x)) \in \Omega$.

Definition 8.7 (robust one-step set). *Given a target set Ω , the (robust) one-step set of Ω under model (8.3) is defined as:*

$$\mathcal{Q}_{\mathcal{F}}(\Omega) := \{x \in \mathbb{R}^n : \exists u \in \mathbb{U} \text{ s.t. } \mathcal{F}(x, u) \subseteq \Omega\} \quad (8.10)$$

Obviously, if \mathcal{F} is an embedding of (8.1), i.e., (8.2) holds, then $\mathcal{Q}_{\mathcal{F}}(\Omega) \subseteq \mathcal{Q}(\Omega)$.

8.3.1 Geometric invariance and Stability

The concepts of stability (undisturbed systems) and invariance/inescapability (disturbed case) are closely related to the geometric contractiveness definition below.

Definition 8.8. *(Kerrigan 2001; Fiacchini, Alamo, and Camacho 2012) A set $\Omega \subseteq \mathbb{X}$ is control λ -contractive for (8.1) if $\Omega \subseteq \mathcal{Q}(\lambda\Omega)$. If $\lambda = 1$ the sets are named control invariant (or, plainly, invariant, in the autonomous case). Analogous definitions for (8.3) arise if $\Omega \subseteq \mathcal{Q}_{\mathcal{F}}(\lambda\Omega)$.*

Trivially, due to the embedding, if a set is control λ -contractive set for (8.3), and (8.2) is fulfilled, so it is for (8.1).

If a control λ -contractive set Ω can be found, there exists a state-feedback control law that keeps the state into Ω at all future time, if initial conditions lie in Ω . The following theorem generalises the CDI stability with a given geometric decay rate in (Fiacchini, Alamo, and Camacho 2010, Prop. 6) or (Fiacchini, Alamo, and Camacho 2012, Cor. 4.17) to sub-homogeneous (maybe non S-convex) maps.

Theorem 8.2. *If there exists a control λ -contractive set Ξ for $x_+ \in \mathcal{F}(x, u)$ with sub-homogeneous \mathcal{F} , then there exists a Lyapunov function $V(x)$, $V(x) > 0$ for $x \neq 0$, which is homogeneous, i.e., $V(\kappa x) = \kappa V(x)$, and a control law $\tilde{u}(x)$ such that $V(x_+) \leq \lambda V(x)$ for all $x \in \Xi$, for all $x_+ \in \mathcal{F}(x, \tilde{u}(x))$.*

Proof. Let us prove that the function:

$$V(x) := \min \{ \gamma : \gamma \geq 0, x \in \gamma \Xi \} \quad (8.11)$$

is a control Lyapunov function the required properties in the theorem statement. Under the definition of $V(x)$, $\kappa \Xi = \{x : V(x) \leq \kappa\}$ for any $\kappa \geq 0$. As Ξ is λ -contractive, there exists a control law $u(x)$ such that, for any $x \in \Xi$, $V(x_+) \leq \lambda$ for all $x_+ \in \mathcal{F}(x, u(x))$. Also, for any $x \in \Xi$, we have $\frac{1}{V(x)}x \in \Xi$. Let us define:

$$\tilde{u}(x) := V(x) \cdot u\left(\frac{1}{V(x)}x\right) \quad (8.12)$$

Now, consider that:

$$\mathcal{F}(x, \tilde{u}(x)) = \mathcal{F}\left(V(x) \cdot \frac{1}{V(x)}x, V(x) \cdot u\left(\frac{1}{V(x)}x\right)\right) \quad (8.13)$$

Thus, as \mathcal{F} is sub-homogeneous, and $0 \leq V(x) \leq 1$ we can assert that:

$$\mathcal{F}(x, \tilde{u}(x)) \subseteq V(x) \cdot \mathcal{F}\left(\frac{1}{V(x)}x, u\left(\frac{1}{V(x)}x\right)\right) \quad (8.14)$$

Hence, for any $x_+ \in \mathcal{F}(x, \tilde{u}(x))$, there exists $\tilde{x} \in \mathcal{F}\left(\frac{1}{V(x)}x, u\left(\frac{1}{V(x)}x\right)\right)$ such that $x_+ = V(x)\tilde{x}$. Thus, as V is homogeneous, $V(x_+) = V(x)V(\tilde{x}) \leq V(x)\lambda$. \square

Computation of one-step sets for DI. From the above results, computation of one-step sets is the key step in the set-based approach to control. For instance, for DIs in the form (8.8) and polyhedral $\Omega := \{x : Hx \leq 1\}$ and \mathbb{D} , the one-step set is the polyhedron

$$Q(\Omega) = Proj_x (\{(x, u) : H(A_i x + B_i u + E_i \epsilon) \leq 1, \forall i = 1, \dots, N_\phi, \forall \epsilon \in vert(\mathbb{D})\}) \quad (8.15)$$

where $Proj_x$ denotes the component x of a pair (x, u) . Software based on this idea (Herceg et al. 2013) has been developed to design state-feedback and predictive

controllers for (8.8), as well as computing the so-called maximal λ -contractive sets, fulfilling $\Omega = \mathcal{Q}(\lambda\Omega)$, see (Kerrigan 2001).

The works (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2010; Fiacchini, Alamo, and Camacho 2012) propose a generalisation of the one-step set computation to CDI and UCDI. In these works, the fact that an *explicit* expression analogous to (8.15) cannot be easily given is suitably justified; as an alternative, they provide means to compute a polytopic approximation, based on the results below (*cf.* (Sala 2017, Prop. 6), (Fiacchini, Alamo, and Camacho 2010, Cor. 1)):

Theorem 8.3. *Assume that $x_+ \in \mathcal{F}(x, u)$ is a UQCDI. Given a known “target” polytope Ω , assume that a known polytope Ξ has been proved to fulfill $\Xi \subseteq \mathcal{Q}_{\mathcal{F}}(\Omega)$. Given a point $\hat{x} \notin \Xi$, if there exists $\hat{u} \in \mathbb{U}$ such that $\mathcal{F}(\hat{x}, \hat{u}) \subseteq \Omega$, then $Co(\Xi \cup \{\hat{x}\}) \subseteq \mathcal{Q}_{\mathcal{F}}(\Omega)$.*

Proof. For every $\xi \in \text{vert}(\Xi)$ we have u_{ξ} such that $\mathcal{F}(\xi, u_{\xi}) \in \Omega$. Thus, Theorem 8.1 ensures that every point in the augmented polyhedron

$$\hat{\Xi} := Co \left((\hat{x}, \hat{u}) \cup \bigcup_{\xi \in \text{vert}(\Xi)} (\xi, u_{\xi}) \right) \subseteq \mathbb{X} \times \mathbb{U}$$

maps to Ω . Now, $Co(\Xi \cup \{\hat{x}\}) = \text{Proj}_x(\hat{\Xi})$. □

Corollary 8.3.1. *If Ξ is control λ -contractive for a UQCDI (8.3) and, given a point $\hat{x} \notin \Xi$, there exists $\hat{u} \in \mathbb{U}$ such that $\mathcal{F}(\hat{x}, \hat{u}) \in \lambda\Xi$, then $Co(\Xi \cup \{\hat{x}\})$ is, too, control λ -contractive for (8.3).*

Proof. Evident from Theorem 8.3, replacing Ω for $\lambda\Xi$. □

The above results can be used to enlarge an initial estimate Ξ of $\mathcal{Q}(\Omega)$ by drawing a collection of new \hat{x} in random directions, via shooting algorithms (Fiacchini 2010a; Fiacchini, Alamo, and Camacho 2012). For completeness, Appendix 8.8.1 outlines the main ideas in them.

8.3.2 Problem statement

In some cases, although the full dynamics (8.1) might not be known (unmeasurable d), partial knowledge of it can be available, in the form of some scheduling parameters comprising a vector $h(x, d)$. This setup encompasses, for instance, the partial knowledge of some elements of d , or the knowledge of some nonlinearities

$h(x)$ involved in system dynamics⁴; thus, control action can depend on h , giving rise to the so-called gain-scheduled controllers.

Indeed, gain-scheduling is well-developed with LMI approaches for the LDIs in literature (Rugh and Shamma 2000a; Tanaka and Wang 2004; Mohammadpour and Scherer 2012); a set-based approach for LDIs appears in (Ariño et al. 2017). However, the convex/quasi-convex DI framework developed in the works by Fiacchini and coworkers does not consider the possibility of such scheduling, as in Theorem 8.3 a *single* control action u must map all the uncertain DI output onto Ω , whereas in a gain-scheduled setup such u can depend on “scheduling parameters”.

The objective of this work is extending to a gain-scheduled setup the robust set-based methods from the convex and quasi-convex difference inclusion (Fiacchini, Alamo, and Camacho 2010; Sala 2017) cases, and including prior set-based LDI/quasi-LPV (Ariño et al. 2017) results in them. The basic idea will be relaxing Theorem 8.3 with further flexibility in the control action depending on scheduling parameters. In this way, one-step and λ -contractive sets which are larger than the robust ones in prior literature can be obtained.

8.4 Main result: computation of gain-scheduled one-step sets

A parameter-dependent DI (PDDI) is defined as:

$$x_+ \in \tilde{\mathcal{F}}(x, u, h) \tag{8.16}$$

being $\tilde{\mathcal{F}}$ a set-valued map. By assumption, there is a known *polytopic* set-valued map $\mathcal{H}(x)$ such that $h \in \mathcal{H}(x)$ and h will be assumed computable in operation from available measurements.

As $x_+ \in \tilde{\mathcal{F}}(x, u, h) \subseteq \tilde{\mathcal{F}}(x, u, \mathcal{H}(x))$, gain-scheduling $u(x, h)$ provides an intermediate approach between the robust approach (designing controllers $u(x)$ for $\mathcal{F}(x, u) := \tilde{\mathcal{F}}(x, u, \mathcal{H}(x))$) and an (idealistic) pure-nonlinear controller design for (8.1).

The focus of this work will be *shape-independent* gain scheduling, which amounts to assuming that an explicit expression for h is not known at design time but its numerical value will be available for control computation, and that these numerical values take values on the known set⁵ $\mathcal{H}(x)$. This is a conservative assumption

⁴The case (8.7) is the most well-known example of this situation (LPV/quasi-LPV systems).

⁵For instance, in shape-independent quasi-LPV gain-scheduled control of (8.7), only the fact that $h \in \Delta$ is used, but that allows well-known gain-scheduling LMIs (Scherer 2006; Sala 2009).

which, in exchange, can ease subsequent control design steps under suitable convexity structures of \mathcal{H} and $\tilde{\mathcal{F}}$.

Definition 8.9. *Given a target set Ω , the **gain-scheduled shape-independent one-step set** of Ω under model (8.16) is defined as:*

$$\mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega) := \{x \in \mathbb{R}^n : \exists v_x : \mathcal{H}(x) \mapsto \mathbb{U} \text{ such that} \\ \tilde{\mathcal{F}}(x, v_x(h), h) \subseteq \Omega \forall h \in \mathcal{H}(x)\} \quad (8.17)$$

The gain-scheduled one-step set requires, for each x , the existence of a *function* of h , $u = v_x(h)$, defined over all $\mathcal{H}(x)$, instead of a single value of u , to be used as control action in Definitions 8.6 and 8.7. As v_x will be, in general, different for different x , the above definition implies that there exists a gain-scheduled controller $u = \nu(h, x)$ such that for all $x \in \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$, and for all $h \in \mathcal{H}(x)$ we have $x_+ \in \tilde{\mathcal{F}}(x, \nu(h, x), h) \subseteq \Omega$. Actually the robust one-step set (Definition 8.7) is the particularisation of Definition 8.9 for constant v_x (i.e., control action depending on the state but not on the scheduling parameters h).

Definition 8.10. *A set $\Omega \subseteq \mathbb{X}$ is said to be (shape-independent) gain-scheduled control λ -contractive for (8.16) if $\Omega \subseteq \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$.*

Theorem 8.4. *Assume that the model (8.16) is such that $\tilde{\mathcal{F}}(x, u, h)$ is S -quasiconvex for fixed x , considering (u, h) as arguments. Denote the elements of $\text{vert}(\mathcal{H}(x))$ as ζ_i , $i = 1, \dots, N_{\mathcal{H}}$, being $N_{\mathcal{H}}$ the number of such vertices (dependence on x of such vertices omitted for notational clarity).*

If, for each ζ_i there exists $v_i \in \mathbb{U}$ such that $\tilde{\mathcal{F}}(x, v_i, \zeta_i) \subseteq \Omega$ then $x \in \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$. Furthermore, for any $h \in \mathcal{H}(x)$ denoting as $\eta(h) \in \Delta$ any arbitrary choice of convex coordinates such that $h = \sum_{i=1}^{N_{\mathcal{H}}} \eta_i(h) \zeta_i$ the control function:

$$v_x(h) := \sum_{i=1}^{n_{\eta}} \eta_i(h) v_i \quad (8.18)$$

achieves $\tilde{\mathcal{F}}(x, v_x(h), h) \subseteq \Omega$ for all $h \in \mathcal{H}(x)$.

Proof. Indeed, Theorem 8.1 ensures that for any $h \in \mathcal{H}(x)$, the function $v_x(h) := \sum_{i=1}^{n_{\eta}} \eta_i(h) v_i$ will ensure $\tilde{\mathcal{F}}(x, v_x(h), h) = \tilde{\mathcal{F}}(x, \sum_{i=1}^{n_{\eta}} \eta_i(h) v_i, \sum_{i=1}^{n_{\eta}} \eta_i(h) \zeta_i) \subseteq \Omega$. \square

Now, the result can be used to assert gain-scheduled versions of Theorem 8.3.

Notation: The following results will discuss conditions for a polytope Ξ to be a subset of the gain-scheduled one-step set of another set Ω . The vertices of Ξ will be enumerated as ξ_j , $j = 1, \dots, N_\Xi$. Subsequently, the vertices of $\mathcal{H}(\xi_j)$ will be enumerated as ζ_{ji} , for $i = 1, \dots, N_{\xi_j}$.

Theorem 8.5. *Assume that $x_+ \in \tilde{\mathcal{F}}(x, u, h)$ is such that $\tilde{\mathcal{F}}$ is an S-quasiconvex set-valued map, and $h \in \mathcal{H}(x)$, being \mathcal{H} a polytopic S-convex set-valued map. Consider a known “target” polytope Ω , and a candidate polytope Ξ . If for each ζ_{ji} there exists $v_{ji} \in \mathbb{U}$ su that $\tilde{\mathcal{F}}(\xi_j, v_{ji}, \zeta_{ji}) \subseteq \Omega$, then $\Xi \subseteq \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$.*

Proof. First, Theorem 8.4 ensures that vertices $\xi_j \in \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$ for all j , as $\tilde{\mathcal{F}}$ is S-quasiconvex. We will now prove that a scheduling control law can be obtained for any point in Ξ from the vertex ones. By S-convexity of \mathcal{H} , as $\Xi = \text{Co}(\xi_j)$, denoting by $\theta(x) \in \Delta$ the convex coordinates such that:

$$x = \sum_{j=1}^{N_\Xi} \theta_j(x) \xi_j \quad (8.19)$$

we can assert that:

$$\mathcal{H}(x) \subseteq \Phi(x) := \bigoplus_{j=1}^{n_\Xi} \theta_j(x) \mathcal{H}(\xi_j) \quad \forall x \in \Xi$$

Now, the direct-sum expression of $\Phi(x)$ above entails that, for any $h \in \mathcal{H}(x)$, there exist points $\hat{\zeta}_j(x, h) \in \mathcal{H}(\xi_j)$, such that, we can express $h = \sum_{j=1}^{N_\Xi} \theta_j(x) \hat{\zeta}_j(x, h)$. Given that each $\hat{\zeta}_j$ is in the convex hull of $\{\zeta_{ji} | i = 1, \dots, N_{\xi_j}\}$, there exist convex coordinates $\eta_j \in \Delta$ such that

$$\hat{\zeta}_j(x, h) = \sum_{i=1}^{N_{\xi_j}} \eta_{ji}(x, h) \zeta_{ji} \quad (8.20)$$

In summary, there exist θ and η such that the expression below holds for any $h \in \mathcal{H}(x)$:

$$h = \sum_{j=1}^{N_\Xi} \sum_{i=1}^{N_{\xi_j}} \theta_j(x) \eta_{ji}(x, h) \zeta_{ji} \quad (8.21)$$

From (8.19), we also have, because the η coordinates add one, that:

$$x = \sum_{j=1}^{N_\Xi} \sum_{i=1}^{N_{\xi_j}} \theta_j(x) \eta_{ji}(x, h) \xi_j \quad (8.22)$$

Now, S-quasiconvexity of $\tilde{\mathcal{F}}$ entails that the scheduling function:

$$v_x(h) := \sum_{j=1}^{N_{\Xi}} \sum_{i=1}^{N_{\xi_j}} \theta_j(x) \eta_{ji}(x, h) v_{ji} \quad (8.23)$$

drives the successor state to Ω for every $h \in \mathcal{H}(x)$, as required by (8.17), because S-quasiconvexity of $\tilde{\mathcal{F}}$ entails that:

$$\tilde{\mathcal{F}}(x, v_x(h), h) = \tilde{\mathcal{F}} \left(\sum_{j=1}^{N_{\Xi}} \sum_{i=1}^{N_{\xi_j}} \theta_j(x) \eta_{ji}(x, h) \xi_j, \sum_{j=1}^{N_{\Xi}} \sum_{i=1}^{N_{\xi_j}} \theta_j(x) \eta_{ji}(x, h) v_{ji}, \sum_{j=1}^{N_{\Xi}} \sum_{i=1}^{N_{\xi_j}} \theta_j(x) \eta_{ji}(x, h) \zeta_{ji} \right) \in \Omega \quad (8.24)$$

which fulfills the conditions in (8.17) for $x \in \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$. \square

If $\tilde{\mathcal{F}}$ does not depend on the state, conditions on \mathcal{H} can be relaxed, as follows.

Theorem 8.6. *Assume that $x_+ \in \tilde{\mathcal{F}}(u, h)$ is such that $\tilde{\mathcal{F}}$ is an S-quasiconvex set-valued map, and $h \in \mathcal{H}(x)$, being \mathcal{H} a polytopic S-quasiconvex set-valued map. Consider a known “target” polytope Ω , and a candidate polytope Ξ . If for each ζ_{ji} there exists $v_{ji} \in \mathbb{U}$ such that $\tilde{\mathcal{F}}(v_{ji}, \xi_{ji}) \subseteq \Omega$, then $\Xi \subseteq \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$.*

Proof. Theorem 8.4 ensures that each $\xi_j \in \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$, i.e., the vertices of Ξ can have a gain-scheduled law mapping them onto Ω .

Now, S-quasiconvexity of \mathcal{H} ensures that for $x \in \Xi$, we will have $\mathcal{H}(x) \subseteq \text{Co}\{\zeta_{ji} \mid j = 1, \dots, N_{\Xi}, i = 1, \dots, N_{\xi_j}\}$, thus there exist coordinates $\eta_{ji}(h) \in \Delta$ such that, for any $h \in \mathcal{H}(x)$:

$$h = \sum_{j=1}^{N_{\Xi}} \sum_{i=1}^{N_{\xi_j}} \eta_{ji}(h) \zeta_{ji} \quad (8.25)$$

Hence S-quasiconvexity of $\tilde{\mathcal{F}}$ ensures that

$$v_x(h) := \sum_{j=1}^{N_{\Xi}} \sum_{i=1}^{N_{\xi_j}} \eta_{ji}(h) v_{ji} \quad (8.26)$$

fulfills:

$$\tilde{\mathcal{F}} \left(\sum_{j=1}^{N_{\Xi}} \sum_{i=1}^{N_{\xi_j}} \eta_{ji}(h) v_{ji}, \sum_{j=1}^{N_{\Xi}} \sum_{i=1}^{N_{\xi_j}} \eta_{ji}(h) \zeta_{ji} \right) \subseteq \Omega \quad (8.27)$$

hence the condition for $x \in \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$ required in (8.17) is fulfilled for any $x \in \Xi$. \square

Remark: in Theorem 8.6, although the dependence of $\tilde{\mathcal{F}}$ on the state is not explicit, the gain-scheduled control does depend on the state due to the role of $\mathcal{H}(x)$.

Theorem 8.7 (gain-scheduled stabilization). *If a gain-scheduled control λ -contractive set Ξ for (8.16) is available with both $\tilde{\mathcal{F}}$ and \mathcal{H} being sub-homogeneous maps, then there exists a Lyapunov function $V(x)$, $V(x) > 0$ for $x \neq 0$, which is homogeneous, i.e., $V(\kappa x) = \kappa V(x)$, and a control law $\tilde{u}(x, h)$ such that $V(x_+) \leq \lambda V(x)$ for all $x \in \Xi$, for all $h \in \mathcal{H}(x)$, for all $x_+ \in \mathcal{F}(x, \tilde{u}(x, h), h)$.*

Proof. Let us prove that the function:

$$V(x) := \min \{ \gamma : \gamma \geq 0, x \in \gamma \Xi \} \quad (8.28)$$

is a control Lyapunov function with the required properties in the theorem statement. Under the definition of $V(x)$, $\kappa \Xi = \{x : V(x) \leq \kappa\}$ for any $\kappa \geq 0$. Also, $V(x)$ is homogeneous, in the sense that $V(\kappa x) = \kappa V(x)$.

As Ξ is λ -contractive, there exists a control law $u(x, h)$ such that, for any $x \in \Xi$, for any $h \in \mathcal{H}(x)$, we have $V(x_+) \leq \lambda$ for all $x_+ \in \mathcal{F}(x, u(x, h), h)$. Also, for any $x \in \Xi$, we have $\frac{1}{V(x)}x \in \Xi$ and, as \mathcal{H} is sub-homogeneous, also $\mathcal{H}(x) \subseteq V(x)\mathcal{H}(\frac{1}{V(x)}x)$. Let us define:

$$\tilde{u}(x, h) := V(x) \cdot u \left(\frac{1}{V(x)}x, \frac{1}{V(x)}h \right) \quad (8.29)$$

Now, denoting $\hat{x} := V(x)^{-1}x$, $\hat{h} := V(x)^{-1}h$, we can write:

$$\begin{aligned} \tilde{\mathcal{F}}(x, \tilde{u}(x, h), h) \\ = \tilde{\mathcal{F}} \left(V(x) \cdot \hat{x}, V(x) \cdot u(\hat{x}, \hat{h}), V(x) \cdot \hat{h} \right) \end{aligned} \quad (8.30)$$

Thus, as $\tilde{\mathcal{F}}$ is sub-homogeneous, and $0 \leq V(x) \leq 1$ we can assert that:

$$\tilde{\mathcal{F}}(x, \tilde{u}(x, h), h) \subseteq V(x) \cdot \tilde{\mathcal{F}} \left(\hat{x}, u(\hat{x}, \hat{h}), \hat{h} \right) \quad (8.31)$$

Hence, for any $h \in \mathcal{H}(x)$ and any $x_+ \in \tilde{\mathcal{F}}(x, \tilde{u}(x, h), h)$, there exists $\hat{h} \in \mathcal{H}(\hat{x})$ and $\tilde{x} \in \tilde{\mathcal{F}}(\hat{x}, u(\hat{x}, \hat{h}), \hat{h})$ such that $x_+ = V(x)\tilde{x}$. Thus, as V is homogeneous, $V(x_+) = V(x)V(\tilde{x}) \leq V(x)\lambda$. \square

Algorithm modifications. Given that conditions in theorems 8.6 and 8.5 require independent “control decision variables” ν_{ji} for each vertex of Ξ , algorithms in Appendix 8.8.1 can be applied, changing $Q_{\mathcal{F}}$ by $Q_{\tilde{\mathcal{F}}}$, so that if conditions in the statement of Theorem 8.4 are fulfilled for a point in a random direction, such a point can be added to a one-step set (or to enlarge an initial λ -contractive seed set). If all involved maps are polytopic and linear in u , conditions end up being plain linear-programming feasibility problems.

Example 8.1. *In order to show a sensible setup where the above gain-scheduling theorems apply, consider the model $x_+ = f(x) + Gu + e(x, d)$ such that $f(x)$ is a vector of known functions, and $e(x, d)$ contains uncertainty/disturbances⁶.*

Consider that $f(x)$ and $e(x, d)$ are bounded by known polytopic set-valued maps $f(x) \in \mathcal{F}(x)$ and $e(x, d) \in \mathcal{E}(x)$, respectively.

If the state is measurable, we can set up a gain-scheduled law with $h \equiv f(x)$. In order to apply the above theorems, we can express the model as the PDDI:

$$x_+ \in \tilde{\mathcal{F}}(x, u, h) := \{h + Gu\} \oplus \mathcal{E}(x) \quad (8.32)$$

with $\mathcal{H}(x) := \mathcal{F}(x)$.

Let us consider some particular cases:

Case 1: *undisturbed system $\mathcal{E}(x) := \{0\}$, and S -quasiconvex \mathcal{F} . In that case, Theorem 8.6 applies as $\tilde{\mathcal{F}}(x, u, h) = \{h + Gu\}$ is a deterministic linear system (hence, S -quasiconvex).*

Case 2: *uncertain system with both \mathcal{E} and \mathcal{F} being S -convex. In that case, $\tilde{\mathcal{F}}(x, u, h)$ is an S -convex map and conditions of Theorem 8.5 apply.*

For instance, a quasi-LPV model $\mathcal{F}(x) := Co(\{A_i x, i = 1, \dots, r\})$, $\mathcal{E}(x) := M\Delta N \cdot x \oplus \mathbb{D}$ being Δ a polytope of matrices⁷ would readily fit the above case 2.

Note, too, that Theorem 8.7 would apply if both \mathcal{E} and \mathcal{F} were sub-homogeneous, thus proving stabilisability if a λ -contractive set is found.

⁶Nonlinearities $g(x)u$ will be considered in Section 8.5.

⁷Of course, Δ can be an ellipsoid if the set-based approach were to be pursued with LMI software, details omitted for brevity.

8.5 Convexification of product nonlinearities

In general, when expressions such as $f(x, u)g(x, u)$ appear, if separate maps such that $f(x, u) \in \mathcal{F}(x, u)$ and $g(x, u) \in \mathcal{G}(x, u)$ are built, then the product $\mathcal{F} \otimes \mathcal{G}$ loses the convexity properties that might be present in \mathcal{F} or \mathcal{G} . Thus, a product $g(x)u$ may lose whichever convexity properties g and u have.

Scheduling in input-channel nonlinearities. To apply prior results, we should embed the nonlinearities in input onto a constant polytope of input uncertainty (the associated columns of $\bar{\Sigma}$) in order to preserve convexity structures.

In order to be able to design a control u which depends on the specific shape of $G(h, x, d)$ and not only on a polyhedral bound of it, we can now consider a controller with the form:

$$u = \Upsilon(x, h)\hat{u}, \quad (8.33)$$

with $\Upsilon \in \mathbb{R}^s \mapsto \mathbb{R}^m$ being a matrix of user-defined “basis functions” and \hat{u} being “artificial” input variables; note that the size s of \hat{u} may not be coincident with that of u , if so wished⁸. The idea can remove conservatism if $\Upsilon(x, h)$ is suitably selected to carry out, for instance, cancellation/linearisation of some elements of $G(h, x, d)$.

Once (8.33) we can replace input nonlinearity $G(h, x, d)$ by $G(h, x, d)\Upsilon(x, h)$ and try to bound it with a polyhedron. Note that input constraints $u \in \mathbb{U}$ should be enforced now by $\Upsilon(x, h)\hat{u} \in \mathbb{U}$, so another polytopic bound for $\Upsilon(\mathbb{X}, \mathcal{H}(\mathbb{X}))$ would need to be computed in order to get sufficient conditions expressed as linear inequalities.

As a simple example, consider a system $d_+ = f(x) + G(x)u$, with $G(\cdot)$ being a known function depending only on the state, so we can express $G(x)$ as a convex combination of some vertex matrices, i.e.:

$$G(x) = \sum_{i=1}^r \mu_i(x)G_i \quad (8.34)$$

with $\mu \in \Delta$. Considering a vector ϖ of monomials in μ , and setting $\Upsilon = \varpi \otimes I_{m \times m}$ where \otimes denotes Kronecker product, an homogeneous polynomial representation in μ can be built and Polya relaxations (Scherer 2006; Sala and Arino 2007; Ariño et al. 2017) can be applied. Full detail on a generic case is omitted for brevity

⁸Evidently, results will be equal or better than the non-transformed approach if $\Upsilon(x, h)$ contains the identity as a submatrix.

(see the cited references): for illustration and details on a particular example, see Example 8.2 below.

8.6 Numerical Example

Example 8.2. Consider the 2nd-order nonlinear system:

$$\begin{pmatrix} x_{1+} \\ x_{2+} \end{pmatrix} = f(x) + \begin{pmatrix} \frac{0.1}{1+x_2^2} \\ 0.05 \end{pmatrix} u \quad (8.35)$$

where

$$f(x) := \begin{pmatrix} x_1 - 0.05x_2 \\ 0.05x_1 + 0.95x_2 - 0.05x_2x_1^2 \end{pmatrix} \quad (8.36)$$

with a control input u constrained to $\mathbb{U} := [-0.5, 0.5]$. We will consider a modeling region $\mathbb{X} = \{(x_1, x_2) : x_1 \in [-a, a], x_2 \in [-b, b]\}$, for different (positive) values of a and b .

Let us first consider modeling the input-channel nonlinearity $\frac{0.1}{1+x_2^2}u$, fulfilling:

$$G(x) := \begin{pmatrix} \frac{0.1}{1+x_2^2} \\ 0.05 \end{pmatrix} \in \tilde{\mathbf{G}}_{[1]} := Co(\{G_1, G_2\}) \quad (8.37)$$

with $G_1 := (0.1 \ 0.05)^T$ and $G_2 := (\frac{0.1}{1+b^2} \ 0.05)^T$. Thus, we can express the system as the difference inclusion

$$x_+ \in \mathcal{F}(x) \oplus \tilde{\mathbf{G}}_{[1]} \cdot u \quad (8.38)$$

for any set-valued map $\mathcal{F}(x)$ such that $f(x) \in \mathcal{F}(x)$. Hence, using an S -convex bound \mathcal{F} would obtain a non-scheduled (robust) controller for (8.38).

The proposal in this chapter will generalise such results, using that fact that we can find $\mu_1(x_2)$ and $\mu_2(x_2)$ such that $(\mu_1, \mu_2) \in \Delta$ and $G(x) = \mu_1(x_2)G_1 + \mu_2(x_2)G_2$. Indeed, the fact that $\mu_2(x_2) := \frac{x_2^2(1+b^2)}{b^2(1+x_2^2)}$ and $\mu_1(x_2) = 1 - \mu_2(x_2)$ can be easily proven.

Thus, the functions μ can be considered to be part of the scheduling vector h in state-feedback control, and, according to the ideas in Section 8.5, we can propose a control law depending on μ : $u(\mu, x) := \mu_1(x_2)\hat{u}_1(x) + \mu_2(x_2)\hat{u}_2(x)$ which would correspond to using $\Upsilon(x, h) := (\mu_1 \ \mu_2)$.

Hence, the transformed dynamics can be expressed as:

$$x_+ = f(x) + \mu_1^2 G_1 \hat{u}_1 + \mu_1 \mu_2 (G_2 \hat{u}_1 + G_1 \hat{u}_2) + \mu_2^2 G_2 \hat{u}_2 \quad (8.39)$$

Now, as $\mu_1 + \mu_2 = 1$, we can write that $(\mu_1 + \mu_2)^2 = \mu_1^2 + 2\mu_1\mu_2 + \mu_2^2 = 1$, too. Hence, by defining the matrix polyhedron $\bar{\mathbf{G}}_{[2]} := \text{Co}([G_1 \ 0], [0 \ G_2], 0.5[G_2 \ G_1])$, the dynamics can be embedded on the set-valued map:

$$x_+ \in \mathcal{F}(x) \oplus \bar{\mathbf{G}}_{[2]} \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} \quad (8.40)$$

which is a generalisation of (8.38), obtained from (8.40) if we enforce $\hat{u}_1 = \hat{u}_2$ (details omitted for brevity).

Different options for $\mathcal{F}(x)$ will be now considered, arising from different boundings of the nonlinearity $\varsigma(x) := -x_2x_1^2$.

LDI case. In the previously-defined modelling region \mathbb{X} , we can bound⁹ $-x_2x_1^2 \in [-ba, ba] \cdot x_1$, thus we can ensure that:

$$f(x) \in \mathcal{F}_{LDI}(x) := \begin{pmatrix} x_1 - 0.05x_2 \\ 0.05x_1 + 0.95x_2 + 0.05[-ba, ba] \cdot x_1 \end{pmatrix} \quad (8.41)$$

CDI case. If the nonlinearity $\varsigma(x)$ is bounded as $x_2x_1^2 \in [-b, b] \cdot x_1^2$ we can assert:

$$f(x) \in \mathcal{F}_{CDI}(x) = \begin{pmatrix} x_1 - 0.05x_2 \\ 0.05x_1 + 0.95x_2 + 0.05[-bx_1^2, bx_1^2] \end{pmatrix}$$

where $\mathcal{F}_{CDI}(x)$ is an S -convex set-valued map (because $-bx_1^2$ is concave and bx_1^2 is convex, from Proposition 8.2), and $\mathcal{F}_{CDI}(0) = \{0\}$.

The CDI bound is tighter than the LDI one, because for any $x \in \mathbb{X}$, it is straightforward to realise that $\mathcal{F}_{CDI}(x) \subseteq \mathcal{F}_{LDI}(x)$.

Note, additionally, that only the bound on x_2 has been used to craft the above CDI bounded. Thus, the above mapping can apply to a modelling region $\hat{\mathbb{X}} := \{(x_1, x_2) : -b \leq x_2 \leq b\}$ with x_1 being unconstrained.

⁹Of course, the LDI modelling is not unique; we made the choice arbitrarily for this example. For instance, another option might have been the bounding $-x_2x_1^2 \in [0, a^2] \cdot x_2$.

Computation of λ -contractive sets. *In order to compute shape-independent gain-scheduled contractive sets, Theorem 8.6 was used¹⁰ with $h \equiv f$, $\mathcal{H}(x) = \mathcal{F}_{CDI}(x)$, and $\tilde{\mathcal{F}}(u, h) = h + \mathbf{G}_{[2]}\hat{u}$. Note that Poly-a-based bounding has eliminated the input-channel scheduling variables μ , although, of course, they are implicitly present: they are needed to reconstruct the actual u from \hat{u} .*

Note that $\tilde{\mathcal{F}}$ is linear (hence sub-homogeneous), and \mathcal{F}_{CDI} is, too, sub-homogeneous as $[-b(\lambda x)^2, b(\lambda x)^2] \subseteq \lambda[-bx^2, bx^2]$ for $\lambda \in [0, 1]$. Thus, λ -contractive sets can be considered to be level-sets of homogeneous control Lyapunov functions (Theorem 8.7).

A contraction rate of 0.999 was sought. Using the LDI representation Using $\mathcal{F}(x) \equiv \mathcal{F}_{LDI}(x)$ from (8.41) in the overall model (8.40), the maximal shape-independent gain-scheduled λ -contractive set for the modelling region \mathbb{X} with $a = b = 0.9$ can be obtained as standard gain-scheduled LDI algorithms (Ariño et al. 2017) converge in a finite number of steps. The maximal set corresponds to the inner red region in Figure 8.1. Actually, trying with $a = b = 1$, these LDI algorithms did not succeed.

As $\mathcal{F}_{CDI}(x) \subseteq \mathcal{F}_{LDI}(x)$, the above maximal gain-scheduled contractive set for the LDI is, too, contractive for the more precise CDI model. Thus, it can be used as a seed set in the shooting algorithms (Appendix 8.8.1).

When the algorithms are actually implemented, given a point x , checking that $x \in \mathcal{Q}_{\tilde{\mathcal{F}}}(\Omega)$, for a polyhedral $\Omega := \{x : Rx \leq s\}$ is straightforward, as it reduces to computing the two vertices of $\mathcal{F}(x)$, denoted as ζ_1, ζ_2 , and checking the feasibility of some linear inequalities in two decision variables.

After 1500 iterations in random directions, an enlarged λ -contractive set for modelling region $\hat{\mathbb{X}}$ is obtained; such a set is the union of the brown and red regions in Figure 8.1.

Increasing the size of the modelling region. *Noting that the evolution of the growth of the sets in Figure 8.1 seems to be limited by the bound in x_2 , we propose to enlarge the modeling region in such a way that $|x_2| \leq 2$. We can remodel with a larger b and use the model $\mathcal{F}_{CDI}(x)$ to try to enlarge the proven domain of attraction in the vertical direction. However, in such a case, if we use the brown region as a seed set, we have no guarantee that such sets are contractive with the new model as there is more uncertainty. Hence, only computation of one-step sets*

¹⁰There are no disturbances or state-dependent uncertainty so there is no need of using Theorem 8.5.

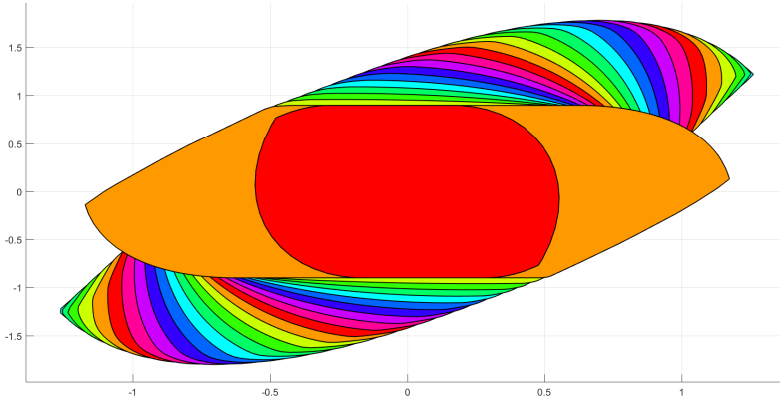


Figure 8.1: Shooting algorithm results with LDI, CDI and a new CDI valid in a larger region.

is a viable solution. With search directions equally spaced 5 degrees, computing 24 one-step set approximations results in the multi-colored regions in the referred figure which expand the proven domain of attraction of the origin. Of course, controller implementation would try to steer a state in the N -th one-step set to the $N - 1$ th one, guaranteed feasible, until the brown region is reached. Then, a control with geometric decay rate 0.999 and a scaled version of the brown set as control Lyapunov function is, too, guaranteed feasible.

8.7 Conclusions

This chapter has presented a gain-scheduled generalisation of set/shooting algorithms to compute control λ -contractive sets in prior literature. The gain-scheduled solutions incorporate the quasi-LPV solutions and its Polya relaxations, as well as the robust controllers for convex difference inclusions, resulting in larger contractive set estimates given the additional information available to decide the control action. Modelling issues have been left out of the scope of the chapter; difference inclusion boundings of nonlinear systems are not unique and, hence, different results can be obtained with different models, as shown in the numerical example. Actual computations of the proposed algorithms for affine-in-control systems are based on polyhedron manipulation software.

8.8 Appendix

8.8.1 Shooting algorithms

In order to implement algorithms from Theorem 8.3 and Corollary 8.3.1, a polynomial-time sufficient condition to check, given a target set Ω and \hat{x} , if there exists \hat{u} such that $\mathcal{F}(\hat{x}, \hat{u}) \in \Omega$ is needed. In the case (8.8), such problem amounts to solving a set of linear programming feasibility conditions. The same case would hold if $\mathcal{F}(x, u)$ were affine in u , i.e., in the form $\mathcal{F}(x, u) = \mathcal{F}_1(x) \oplus \mathcal{F}_2(x)u$ for some set-valued maps \mathcal{F}_1 and \mathcal{F}_2 . In a general case, if the support function of $\mathcal{F}(x, u)$ is convex or quasi-convex in u , computationally viable conditions may be cast (Eppstein 2004).

As a conclusion of Theorem 8.1, in order to determine an inner approximation of the one-step set of a set Ω for a DI $x_+ \in \mathcal{F}(x, u)$, we can use the following algorithm:

Algorithm 1 [one-step set approximation]

1. Generate a set of K vectors ρ_1, \dots, ρ_K either randomly or in a suitable grid of the unit hypersphere in \mathbb{R}^n .
2. For each $\rho_k, k = 1, \dots, K$, determine by bisection the largest scaling γ_k such that $\xi_k := \gamma_k \rho_k \in \mathcal{Q}_{\mathcal{F}}(\Omega)$.
3. Form $\Xi := \mathcal{Co}(\{\xi_k, k = 1, \dots, K\})$.
4. End. Theorem 8.1 ensures that $\Xi \subseteq \mathcal{Q}_{\mathcal{F}}(\Omega)$.

As a variation, the shooting algorithm in (Fiacchini, Alamo, and Camacho 2010) can obtain progressively larger λ -contractive sets if one of them is available:

Algorithm 2 [contractive set expansion]

1. Obtain an initial λ -contractive polyhedral “seed” set Ω_0 overbounding with an LDI and using set-based LDI approaches ((Fiacchini, Alamo, and Camacho 2010, §4.1)). Set $k = 0$.
2. Choose a random direction vector ρ .
3. Determine, by bisection, the largest scaling $\gamma > 0$ such that there exists \hat{u} such that $\mathcal{F}(\gamma\rho, \hat{u}) \in \lambda\Omega_k$.

4. Set $\Omega_{k+1} = Co(\Omega_k \cup \{\gamma\rho\})$, and $k = k + 1$. If $k < k_{max}$, go to step 2.
5. End. Corollary 8.3.1 ensures that Ω_k is control λ -contractive.

8.8.2 Support function approach

Proposition 8.1 (cf. Prop. 1 in (Fiacchini, Alamo, and Camacho 2010)). *A set-valued map \mathcal{F} with $\mathcal{F}(x)$ being convex and compact, is **S-convex** (**S-quasiconvex**) iff $\check{f}(\eta, x) := \sup_{x \in \mathcal{F}(x)} \eta^T x$ is a convex (quasiconvex) function of x when η is fixed, for all $\eta \in \mathbb{R}^n$.*

Proof. From support function properties, a convex and compact set Ω can be equivalently defined as $\Omega = \{x : \eta^T x \leq \Phi_\Omega(\eta) \forall \eta \in \mathbb{R}^n\}$. If Σ and Γ are compact and convex sets, support function has the properties :

$$\Sigma \subseteq \Gamma \Leftrightarrow \Phi_\Sigma(\eta) \leq \Phi_\Gamma(\eta) \quad \forall \eta \quad (8.42)$$

$$\Phi_{\lambda\Sigma}(\eta) = |\lambda| \Phi_\Sigma(\text{sign}(\lambda)\eta) \quad (8.43)$$

$$\Phi_{\Omega \oplus \Gamma}(\eta) = \Phi_\Omega(\eta) + \Phi_\Gamma(\eta) \quad (8.44)$$

$$\Phi_{Co(\Omega \cup \Gamma)}(\eta) = \max(\Phi_\Omega(\eta), \Phi_\Gamma(\eta)) \quad (8.45)$$

Thus, the condition in Definition 8.3 can be stated as

$$\Phi_{\mathcal{F}(\lambda x_1 + (1-\lambda)x_2)}(\eta) \leq \lambda \Phi_{\mathcal{F}(x_1)}(\eta) + (1-\lambda) \Phi_{\mathcal{F}(x_2)}(\eta)$$

and that in Definition 8.4 as:

$$\Phi_{\mathcal{F}(\lambda x_1 + (1-\lambda)x_2)}(\eta) \leq \max(\Phi_{\mathcal{F}(x_1)}(\eta), \Phi_{\mathcal{F}(x_2)}(\eta))$$

so, with suitable notation changes to \check{f} the convexity and quasi-convexity in the theorem statement follow, respectively. \square

8.8.3 Modelling: embedding nonlinear functions in set-valued maps

LDI can be built for multiple-argument functions in several ways. The topic is well studied, so the reader is referred to Chapter 5 (Robles et al. 2017) and the works (Tanaka and Wang 2004; Rugh and Shamma 2000a; Sala 2009; Kwiatkowski and Werner 2008a) for ample detail.

This section will outline methodologies for embedding a (non-uncertain) nonlinearity $f(x)$ inside a set-valued map $\mathcal{F}(x)$ with an S-convex/S-quasiconvex structure. Of course, parametric/additive uncertainty elements can be added on top

of that, with minor modifications, see sections 2.1.1.2 and 2.1.1.3 in (Fiacchini 2010a), as well as some other classes of uncertainty such as ellipsoidal (section 2.2 in the cited work).

Let us consider the case where a function $\rho(x) : \Omega \mapsto \mathbb{R}$, with $\Omega \subset \mathbb{R}^n$, can be bounded by known f_l, f_u such that $f_l(x) \leq \rho(x) \leq f_u(x)$.

Proposition 8.2. (Fiacchini 2010a, Property 3.3) *The set-valued map defined by $\mathcal{F}(x) = [f_l(x), f_u(x)]$ where $f_l(x) \leq f_u(x)$ and f_l is concave and f_u is convex is S-convex.*

Proposition 8.3. (Sala 2017) *The set-valued map defined by $\mathcal{F}(x) = [f_l(x), f_u(x)]$ where $f_l(x) \leq f_u(x)$ and f_l is quasiconcave and f_u is quasiconvex is S-quasiconvex.*

Proposition 8.4. (Fiacchini 2010a, Property 3.18) *Assume that $\rho(x)$ can be expressed as $\rho(x) = g(x) - h(x)$ being g and h convex functions (the so-called difference-of-convex functions (Fiacchini 2010a)). Let $\xi_g(x)$ be any linear function such that $\xi_g(x) \leq g(x)$ and let $\xi_h(x)$ be any linear function such that $\xi_h(x) \leq h(x)$. Then, under the above assumption, the set-valued map $\mathcal{F}(x) = [\xi_g(x) - h(x), g(x) - \xi_h(x)]$ is S-convex.*

Proposition 8.5 (stacking). *If $\mathcal{F}_1(x) \subset \mathbb{R}^m$ and $\mathcal{F}_2(x) \subset \mathbb{R}^q$ are S-convex set-valued maps, so it is the map $(\mathcal{F}_1(x), \mathcal{F}_2(x)) \subset \mathbb{R}^{m+q}$.*

Proof. Straightforward from definition. However, the S-quasiconvex version of the proposition is *not* true. \square

Proposition 8.6. *Consider an expression $\rho(x) = A(x)f(x)$ and assume that vector $f(x)$ can be embedded onto an S-quasiconvex (S-convex) map, i.e., $f(x) \in \mathcal{F}(x)$. Then, obtaining a polytopic bound \mathbf{A} such that $A(\Omega) \subseteq \mathbf{A}$, with $\text{vert}(\mathbf{A}) = \{A_i, i = 1, \dots, r\}$, the map $\mathcal{G}(x) := \mathbf{A} \otimes \mathcal{F}(x) = \text{Co}(\cup_{i=1}^r A_i \cdot \mathcal{F}(x))$ is S-quasiconvex (S-convex).*

Proof. Note that for any set $\mathbb{S} \subseteq \mathbb{R}^n$ and a linear transformation given by matrix A , we have that $\text{Co}(A\mathbb{S}) = A\text{Co}(\mathbb{S})$. Thus, consider $\tilde{x} = \gamma x_1 + (1 - \gamma)x_2$, with $0 \leq \gamma \leq 1$. As $\mathcal{F}(\tilde{x}) \subseteq \text{Co}(\mathcal{F}(x_1) \cup \mathcal{F}(x_2))$, we can assert that

$$\begin{aligned} \mathcal{G}(\tilde{x}) &\subseteq \text{Co}(\cup_{i=1}^r \cup_{j=1}^2 A_i \cdot \mathcal{F}(x_j)) \\ &= \text{Co}(\text{Co}(\cup_{i=1}^r A_i \cdot \mathcal{F}(x_1)) \cup \text{Co}(\cup_{i=1}^r A_i \cdot \mathcal{F}(x_2))) \\ &= \text{Co}(\mathcal{G}(x_1) \cup \mathcal{G}(x_2)) \quad (8.46) \end{aligned}$$

The proof of the S-convex case roots on the fact that a linear transformation of an S-convex map is S-convex and so it is the union of them, but details are left to the reader, for brevity. \square

Note that, as in the well-developed linear LDI case, the result of an S-convex or S-quasiconvex embedding is not unique, and some resulting maps will give different performance and contractive set (domain of attraction) estimates. This is, indeed, a well-known fact (Sala 2009). Note also that S-quasiconvex modelling of multi-dimensional outputs is a difficult task, as quasiconvexity is neither preserved after linear combinations nor after mere stacking. Thus, even if some of the results in this chapter apply to a quasiconvex case, systematic quasiconvex modelling is an open issue. As DI modelling is not the main topic of this chapter, the reader is referred to (Fiacchini 2010a; Sala 2017) for further information.

Part III

Conclusions and Bibliography

Conclusions

This thesis is a first approach towards bringing into systematic procedures the idea that TS models might be selected so that the generated “uncertainty” is suitably managed.

A Hessian-based approach has been presented, which is performed in a systematic way, and it “decouples” multivariable nonlinearities (up to 2nd degree approximation) providing “minimum maximal error”.

Hessian might be a good option when modelling a nonlinearity without knowledge on how its error might influence later results; however, its complexity lies in the fact that multiple Hessians might be needed in a systems with multiple nonlinearities thereby increasing the complexity and number of linear models of the resulting Takagi-Sugeno model.

In the case where it is required to keep a performance measure as close as possible to the linearised case, the proposed linearisation-based approach is an advantageous alternative over “blind” methodologies since, in a systematic way, the system’s “orientation”, through a linear transformation, is sought so that it is the most beneficial to the performance measure. Obviously, increased performance improvements due to *nonlinear* transformations might be achievable; they are left as a matter of future research.

This thesis has also presented a subspace based methodology to minimise the conservatism of TS models to be later used in shape-independent LMI conditions. The presented procedure is optimal (in minimax worst-case uncertainty width) for TS models of quadratic functions, hence approximately optimal for any smooth

nonlinearity close enough to the origin. Different examples prove that, indeed, the uncertainty width measures, as well as decay-rate and \mathcal{H}_∞ performance figures are better with the proposed TS modelling technique than those obtained with frequently used “inspection” and “extraction of factors” ideas. The LMIs discuss only the preservation of linearised performance. Other LMI setups and reduction/simplification of the number of rules, as well as optimality for large modelling regions, are matter of future research.

Generating an “optimal” TS model from nonlinear equations is an overlooked aspect in literature. This paper has presented a null-space based methodology to craft a coordinate transformation from which a TS model is generated. There are prior works by the team which discuss related issues, but such works can only be applied in very restrictive setups (no LMI transformations). The presented procedure allows for arbitrary LMI setups, with the model appearing as many times as wished in Schur/multicriteria problems. Numerical examples seem to show that performance is comparable to the prior solutions, with a more general applicability.

This thesis also addressed the problem of choosing a good quasi-LPV model for a nonlinear system. Our proposals are able to mitigate the performance loss from a reference linearised design, compared to other “inspection-based” modelling choices, while avoiding BMIs needed in prior literature for controller synthesis problems. The main idea is based on obtaining the more harmful perturbations of the model matrices in the linearised design, proposing then a coordinate transformation so that the projection of the uncertainty over these model perturbations is minimised. This direction plays, too, a fundamental role in later polytopic embedding options which are built to avoid the need of gridding. The root of all the procedure is the linearised design, the results are only valid, formally, for small modelling regions around the origin.

The last contribution of this thesis presents a radical departure from the TS setup, using bounds for nonlinearities that are no longer required to be linear, but convex or quasi-convex. In this framework, the linear (quasi-LPV, TS) case is a particularisation of the general theory. In the same way that ellipsoidal domain of attraction estimates from LMIs can be “beaten” by set-based methodologies approaching the “maximal invariant set” in polytopic models (both robust and gain-scheduled versions), the proposals here enable to further enlarge such estimates with gain-scheduled controller over this expanded class of models.

Open issues, discussion

This thesis approached the problem of modelling in an “optimal” way a nonlinear system to be controlled with convex optimisation or convex set manipulation.

However, the concept of “optimality” is only proved in a restrictive sense and, in some cases, theorems only apply to quadratic nonlinearities; all functions are quadratic close to the origin but, formally, we have not truly departed from “infinitesimal” modelling regions, and actually the target optimal performance is that of a LTI linearisation with quadratic Lyapunov function.

Also, in practice, when comparing an “optimal” model with high numerical complexity and a “less optimal” model with reasonable complexity, the designers may end up choosing the latter one, regardless of the involved mathematical developments behind the optimal model. Thus, a practical compromise between model complexity, controller computation requirements, closed loop performance, etc. may be needed.

Additionally, given that some control problems are amenable to polynomial approaches via Sum-of-squares (convex) optimisation, improvements over the linearised model may be achievable by considering its higher-degree Taylor series in a polynomial-only approach. However, Taylor series models abruptly diverge as distance to the origin increases, so, in some cases, the actual performance in “large” modelling regions is worse than the TS one. Combining/extending the proposals in this paper to the polynomial case might be an interesting future work.

Regarding the issues on the set-based gain-scheduled difference inclusions, the basic drawback is that some convex-hull and vertex enumeration steps are computationally demanding problem so the results, even if elegant, may apply in real applications only to low-order systems. Also, although the results apply to “quasi-convex” setups generalising earlier “convex” ones, given that the mere “stacking” of two quasiconvex functions is not quasiconvex, a systematic methodology for quasiconvex modelling in several dimensions may be a difficult endeavour.

In summary, this thesis tried to approach a mostly open issue in “nonlinear control via convex structures” regarding the different modelling options available, yielding different performance results. Even if we first tried to settle some ideas for a TS/quasi-LPV framework, later on, the breadth of possible modelling options was opened up again with the difference-inclusion approach. Each framework has its own advantages and drawbacks and, actually, the “best convex structure” for a generic nonlinear control problem was, and still is, a challenging issue.

Bibliography

- Aguiar, Braulio, Raymundo Márquez, and Miguel Bernal (2016). “An LMI-based global non-quadratic observer design via Takagi-Sugeno models and Levant’s robust differentiators”. In: *American Control Conference (ACC), 2016*. IEEE, pp. 119–124 (cit. on p. 47).
- Alamo, T, A Cepeda, Mirko Fiacchini, and Eduardo F Camacho (2009). “Convex invariant sets for discrete-time Lurâe systems”. In: *Automatica* 45.4, pp. 1066–1071 (cit. on p. 65).
- Angeli, David, Alessandro Casavola, and Edoardo Mosca (2000). “Constrained predictive control of nonlinear plants via polytopic linear system embedding”. In: *International Journal of Robust and Nonlinear Control* 10.13, pp. 1091–1103 (cit. on p. 141).
- Apkarian, P. and P. Gahinet (1995). “A convex characterization of gain-scheduled H_∞ controllers”. In: *IEEE Transactions on Automatic Control* 40.5, pp. 853–864 (cit. on pp. 26, 138).
- Apkarian, Pierre, Pascal Gahinet, and Greg Becker (1995). “Self-scheduled H_8 control of linear parameter-varying systems: a design example”. In: *Automatica* 31.9, pp. 1251–1261 (cit. on p. 138).
- Apostol, T. (1967). *Calculus, Vol. 1: One-Variable Calculus with an Introduction to Linear Algebra*. 2nd ed. Wiley (cit. on p. 48).

- ApS, MOSEK (2015). *The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28)* (cit. on p. 12).
- Ariño, C. and A. Sala (2007). “Relaxed LMI conditions for closed-loop fuzzy systems with tensor-product structure”. In: *Engineering Applications of Artificial Intelligence* 20.8, pp. 1036–1046 (cit. on pp. 49, 121, 128).
- Ariño, Carlos, Andrés Querol, and Antonio Sala (2017). “Shape-independent model predictive control for Takagi–Sugeno fuzzy systems”. In: *Engineering Applications of Artificial Intelligence* 65, pp. 493–505 (cit. on p. 54).
- Ariño, Carlos, Antonio Sala, Emilio Pérez, Fernando Bedate, and Andrés Querol (2017). “Asymptotically exact stabilization for constrained discrete Takagi–Sugeno systems via set-invariance”. In: *Fuzzy Sets and Systems* 316, pp. 117–138 (cit. on pp. 59–61, 63, 164, 172, 178, 181).
- Arino, C.V. and A. Sala (2015). “Shape-dependent maximal controllable sets for constrained discrete-time TS systems”. In: *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*. IEEE, pp. 1–7 (cit. on p. 44).
- Baruah, Rashmi Dutta and Plamen Angelov (2011). “Evolving fuzzy systems for data streams: a survey”. In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1.6, pp. 461–476 (cit. on p. 50).
- Bernal, M. and T. M. Guerra (2010). “Generalized non-quadratic stability of continuous-time Takagi-Sugeno models”. In: *IEEE Transactions on Fuzzy Systems* 18.4, pp. 815–822 (cit. on pp. 46, 47, 139).
- Bernal, Miguel, Thierry Marie Guerra, and Alexandre Kruszewski (2009). “A membership-function-dependent approach for stability analysis and controller synthesis of Takagi-Sugeno models”. In: *Fuzzy sets and systems* 160.19, pp. 2776–2795 (cit. on pp. 44, 74, 90, 114, 125).
- Bianchi, F.D., R.J. Mantz, and C.F. Christiansen (2005a). “Gain scheduling control of variable-speed wind energy conversion systems using quasi-LPV models”. In: *Control Engineering Practice* 13.2, pp. 247–255 (cit. on p. 141).
- Bianchi, FD, RJ Mantz, and CF Christiansen (2005b). “Gain scheduling control of variable-speed wind energy conversion systems using quasi-LPV models”. In: *Control Engineering Practice* 13.2, pp. 247–255 (cit. on p. 164).

-
- Blanchini, F. and S. Miani (2008). *Set-Theoretic Methods in Control*. Boston, USA: Birkhäuser Boston (cit. on pp. 51, 54, 56).
- Blanchini, Franco (1991). “Ultimate boundedness control for uncertain discrete-time systems via set-induced Lyapunov functions”. In: *Decision and Control, 1991., Proceedings of the 30th IEEE Conference on*. IEEE, pp. 1755–1760 (cit. on pp. 51, 57).
- (1999). “Set invariance in control”. In: *Automatica* 35.11, pp. 1747–1767 (cit. on pp. 51, 61, 164).
- Blanco, Y., W. Perruquetti, and P. Borne (2001). “Stability and stabilization of nonlinear systems and Takagi-Sugeno fuzzy models”. In: *Mathematical Problems in Engineering* 7.3, pp. 221–240 (cit. on p. 46).
- Boyd, S., L.E. Ghaoui, E. Feron, and V. Belakrishnan (1994). *Linear Matrix Inequalities in System and Control Theory*. Vol. 15. Philadelphia, USA: SIAM: Studies In Applied Mathematics (cit. on pp. 10–12, 15, 27, 74, 82, 90, 124).
- Boyd, Stephen and Lieven Vandenberghe (2004). *Convex optimization*. Cambridge University Press (cit. on pp. 65, 164, 168).
- Bronshtein, MD (1979). “Smoothness of roots of polynomials depending on parameters”. In: *Siberian Mathematical Journal* 20.3, pp. 347–352 (cit. on p. 146).
- Bruzelius, Fredrik, Stefan Pettersson, and Claes Breitholtz (2004). “Linear parameter-varying descriptions of nonlinear systems”. In: *American Control Conference, 2004. Proceedings of the 2004*. Vol. 2. IEEE, pp. 1374–1379 (cit. on pp. 139, 141, 142).
- Camacho, Eduardo F and Carlos Bordons Alba (2013). *Model predictive control*. Springer Science & Business Media (cit. on p. 54).
- Cao, Y.Y. and P.M. Frank (2000). “Analysis and Synthesis of Nonlinear Time-Delay Systems via Fuzzy Control Approach”. In: *IEEE Transactions on Fuzzy Systems* 8.2, pp. 200–211 (cit. on p. 43).
- Chen, C. T. (1984). *Linear System Theory and Design*. New York, USA: Saunders College Publishing (cit. on p. 14).

Chesi, G. (2009). “Estimating the domain of attraction for non-polynomial systems via LMI optimizations”. In: *Automatica* 45.6, pp. 1536–1541 (cit. on p. 48).

Chesi, Graziano (2009). “Estimating the domain of attraction for non-polynomial systems via LMI optimizations”. In: *Automatica* 45.6, pp. 1536–1541. ISSN: 0005-1098. DOI: <https://doi.org/10.1016/j.automatica.2009.02.011> (cit. on p. 168).

Coutinho, D. F., M. Fu, A. Trofino, and P. Danes (2008). “ \mathcal{L}_2 -Gain analysis and control of uncertain nonlinear systems with bounded disturbance inputs”. In: *International Journal of Robust and Nonlinear Control* 18.1, pp. 88–110 (cit. on pp. 156, 159).

Coutinho, Daniel Ferreira, Minyue Fu, and Alexandre Trofino (2004). “Robust analysis and control for a class of uncertain nonlinear discrete-time systems”. In: *Systems & control letters* 53.5, pp. 377–393 (cit. on pp. 139, 142).

De Caigny, Jan, Juan F Camino, Ricardo CLF Oliveira, Pedro Luis Dias Peres, and Jan Swevers (2010). “Gain-scheduled \mathcal{H}_2 and \mathcal{H}_∞ control of discrete-time polytopic time-varying systems”. In: *IET control theory & applications* 4.3, pp. 362–380 (cit. on p. 155).

Doyle, J. C., B. A. Francis, and A. Tannenbaum (1992). *Feedback control theory*. Vol. 1. Macmillan Publishing Company New York (cit. on p. 22).

Duan, Guang-Ren and Hai-Hua Yu (2013). *LMIs in control systems: analysis, design and applications*. CRC press (cit. on p. 18).

El Ghaoui, L, V Balakrishnan, E Feron, and S Boyd (1992). “On maximizing a robustness measure for structured nonlinear perturbations”. In: *American Control Conference, 1992*. IEEE, pp. 2923–2924 (cit. on p. 22).

Eppstein, David (2004). “Quasiconvex Programming”. In: *arXiv preprint* <http://arxiv.org/abs/cs/0408001> (cit. on p. 183).

Escaño, Juan Manuel and Carlos Bordons (2014). “Complexity Reduction in Fuzzy Systems Using Functional Principal Component Analysis”. In: *Fuzzy Modeling and Control: Theory and Applications*. Springer, pp. 49–65 (cit. on pp. 90, 124).

-
- Estrada-Manzo, V., Zs. Lendek, and T. M. Guerra (2016). “Generalized LMI observer design for discrete-time nonlinear descriptor models”. In: *Neuro-computing* 182, pp. 210–220 (cit. on p. 47).
- Feng, G. (2003). “Controller synthesis of fuzzy dynamic systems based on piecewise Lyapunov functions”. In: *IEEE Transactions on Fuzzy Systems* 11.5, pp. 605–612 (cit. on p. 44).
- (2006). “A Survey on Analysis and Design of Model-Based Fuzzy Control Systems”. In: *IEEE Transactions on Fuzzy Systems* 14.5, pp. 676–697 (cit. on pp. 74, 90).
- Feng, G., C.L. Chen, D. Sun, and Y. Zhu (2005). “ H_∞ controller synthesis of fuzzy dynamic systems based on piecewise Lyapunov functions and bilinear matrix inequalities”. In: *IEEE Transactions on Fuzzy Systems* 13.1, pp. 94–103 (cit. on pp. 44, 46).
- Fiacchini, M (2010a). “Convex difference inclusions for systems analysis and control”. PhD thesis. Universidad de Sevilla, Spain (cit. on pp. 65, 67–70, 164, 167, 168, 171, 185, 186).
- Fiacchini, Mirko (2010b). *Convex difference inclusions for systems analysis and design* (cit. on pp. 51, 52, 65).
- Fiacchini, Mirko, Teodoro Alamo, and Eduardo F Camacho (2010). “On the computation of convex robust control invariant sets for nonlinear systems”. In: *Automatica* 46.8, pp. 1334–1338 (cit. on pp. 65–68, 70, 164, 165, 167–169, 171, 172, 183, 184).
- (2012). “Invariant sets computation for convex difference inclusions systems”. In: *Systems & Control Letters* 61.8, pp. 819–826 (cit. on pp. 65, 67, 70, 164, 167, 169, 171).
- Fiacchini, Mirko, Antoine Girard, and Marc Jungers (2016). “On the stabilizability of discrete-time switched linear systems: Novel conditions and comparisons”. In: *IEEE Transactions on Automatic Control* 61.5, pp. 1181–1193 (cit. on p. 47).
- Gahinet, P., A. Nemirovski, A. J. Laub, and M. Chilali (1995). *LMI Control Toolbox*. Natick, USA: MathWorks (cit. on p. 12).

- Gahinet, Pascal, Pierre Apkarian, and Mahmoud Chilali (1996). “Affine parameter-dependent Lyapunov functions for real parametric uncertainty”. In: *Automatic Control, IEEE Transactions on* 41.3, pp. 436–442 (cit. on pp. 121, 138).
- Gao, Q., X.J Zeng, G. Feng, Y. Wang, and J. Qiu (2012). “T-S-Fuzzy-Model-Based Approximation and Controller Design for General Nonlinear Systems”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 42.4, pp. 1143–1154 (cit. on p. 161).
- Gegov, Alexander (2007). *Complexity management in fuzzy systems*. 211. Springer (cit. on pp. 74, 124).
- Gilbert, Elmer G and K Tin Tan (1991). “Linear systems with state and control constraints: The theory and application of maximal output admissible sets”. In: *IEEE Transactions on Automatic control* 36.9, pp. 1008–1020 (cit. on p. 57).
- Gonzalez, Antonio, Antonio Sala, Pedro Garcia, and Pedro Albertos (2013). “Robustness analysis of discrete predictor-based controllers for input-delay systems”. In: *International Journal of Systems Science* 44.2, pp. 232–239 (cit. on p. 14).
- Gonzalez, T. and M. Bernal (2016). “Progressively better estimates of the domain of attraction for nonlinear systems via piecewise Takagi-Sugeno models: Stability and stabilization issues”. In: *Fuzzy Sets and Systems* 297, pp. 73–95 (cit. on p. 46).
- Gonzalez, T., A. Sala, M. Bernal, and R. Robles (2015). “Invariant Sets of Nonlinear Models via Piecewise Affine Takagi-Sugeno Model”. In: *Proceedings of the 2015 IEEE International Conference on Fuzzy Systems*. Istanbul, Turkey (cit. on p. 4).
- Gonzalez, T., R. Marquez, M. Bernal, and T.M. Guerra (2016). “Nonquadratic Controller and Observer Design for Continuous TS Models: A Discrete-Inspired Solution”. In: *International Journal of Fuzzy Systems* 18.1, pp. 1–14 (cit. on p. 139).
- Gonzalez, T., A. Sala, M. Bernal, and R. Robles (2017a). “A generalisation of Line-Integral Lyapunov Function for Takagi-Sugeno systems”. In: *Proceedings*

-
- of the 20th World Congress of the International Federation of Automatic Control. Toulouse, France (cit. on pp. 4, 47).
- Gonzalez, T., M. Bernal, A. Sala, and B. Aguiar (2017b). “Cancellation-Based Nonquadratic Controller Design for Nonlinear Systems via Takagi-Sugeno Models”. In: *IEEE Transactions on Cybernetics* 47.9, pp. 2628–2638 (cit. on p. 47).
- Gonzalez, T., A. Sala, M. Bernal, and B. Aguiar (2017c). “Piecewise-Takagi-Sugeno asymptotically exact estimation of the domain of attraction of nonlinear systems”. In: *Journal of the Franklin Institute* 354.3, pp. 1514–1541 (cit. on pp. 4, 46).
- Guelton, K., S. Delprat, and T.M. Guerra (2008). “An alternative to inverse dynamics joint torques estimation in human stance based on a Takagi-Sugeno unknown-inputs observer in the descriptor form”. In: *Control Engineering Practice* 16.12, pp. 1414–1426 (cit. on p. 49).
- Guelton, K., T.M. Guerra, M. Bernal, Bouarar T., and N. Manamanni (2010). “Comments on Fuzzy Control Systems Design via Fuzzy Lyapunov Functions”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 40.3, pp. 970–972 (cit. on p. 46).
- Guerra, Thierry Marie, Victor Estrada-Manzo, and Zsófia Lendek (2015). “Observer design for Takagi–Sugeno descriptor models: An LMI approach”. In: *Automatica* 52, pp. 154–159 (cit. on pp. 124, 138).
- Guerra, Thierry Marie and Laurent Vermeiren (2004). “LMI-based relaxed nonquadratic stabilization conditions for nonlinear systems in the Takagi-Sugeno’s form”. In: *Automatica* 40.5, pp. 823–829. ISSN: 0005-1098. DOI: <https://doi.org/10.1016/j.automatica.2003.12.014> (cit. on pp. 65, 164, 168).
- Guerra, T.M. and M. Bernal (2009). “A way to escape from the quadratic framework”. In: *Proceedings of the IEEE International Conference on Fuzzy Systems*. Jeju, Korea, pp. 784–789 (cit. on p. 47).
- Guerra, T.M., A. Kruszewski, and M. Bernal (2009). “Control Law Proposition for the Stabilization of Discrete Takagi-Sugeno Models”. In: *IEEE Transactions on Fuzzy Systems* 17.3, pp. 724–731 (cit. on p. 47).

- Guerra, T.M., A. Sala, and K. Tanaka (2015). “Fuzzy control turns 50: 10 years later”. In: *Fuzzy Sets and Systems* 281.3, pp. 168–182 (cit. on pp. 43, 124).
- Guerra, T.M. and L. Vermeiren (2004). “LMI-based relaxed non-quadratic stabilization conditions for nonlinear systems in Takagi-Sugeno’s form”. In: *Automatica* 40.5, pp. 823–829 (cit. on pp. 47, 138, 141).
- Guerra, T.M., M. Bernal, K. Guelton, and S. Labiod (2012). “Nonquadratic local stabilization for continuous-time Takagi-Sugeno models”. In: *Fuzzy Sets and Systems* 201, pp. 40–54 (cit. on p. 47).
- Haddad, W. M. and V. Chellaboina (2008). *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*. Princeton, NJ: Princeton University Press (cit. on p. 32).
- Hahn, Wolfgang (1967). *Stability of motion*. Vol. 138. Springer (cit. on p. 33).
- Härter, V., C. Jansson, and M. Lange (2012). *VSDP: A Matlab toolbox for verified semidefinite-quadratic-linear programming*. Tech. rep. Institute for Reliable Computing, Hamburg University of Technology (cit. on p. 12).
- Herceg, M., M. Kvasnica, C.N. Jones, and M. Morari (2013). “Multi-Parametric Toolbox 3.0”. In: *Proc. of the European Control Conference*. Zürich, Switzerland, pp. 502–510 (cit. on pp. 58, 61, 65, 164, 168, 170).
- Huang, Yun (1999). “Nonlinear optimal control: an enhanced quasi-LPV approach”. PhD thesis. California Institute of Technology (cit. on p. 139).
- Huang, Yun and Ali Jadbabaie (1999). “Nonlinear H8 control: An enhanced quasi-LPV approach”. In: *Proceedings of the IFAC World Congress*. Cite-seer, pp. 85–90 (cit. on pp. 138–142).
- Jeyakumar, V, W Oettli, and M Natividad (1993). “A solvability theorem for a class of quasiconvex mappings with applications to optimization”. In: *Journal of Mathematical Analysis and Applications* 179.2, pp. 537–546 (cit. on p. 166).
- Johansen, T.A., R. Shorten, and R. Murray-Smith (2000). “On the interpretation and identification of dynamic Takagi-Sugeno fuzzy models”. In: *IEEE Transactions on Fuzzy Systems* 8.3, pp. 297–313 (cit. on p. 50).

- Johansson, M. and A. Rantzer (1998). “Computation of piecewise quadratic Lyapunov functions for hybrid systems”. In: *IEEE Transactions on Automatic Control* 43.4, pp. 555–559 (cit. on p. 44).
- Johansson, M., A. Rantzer, and K. Arzen (1999). “Piecewise quadratic stability of fuzzy systems”. In: *IEEE Transactions on Fuzzy Systems* 7.6, pp. 713–722 (cit. on pp. 44–46).
- Karmarkar, Narendra (1984). “A new polynomial-time algorithm for linear programming”. In: *Proceedings of the sixteenth annual ACM symposium on Theory of computing*. ACM, pp. 302–311 (cit. on p. 12).
- Kerrigan, E. (2000). “Robust Constraint Satisfaction: Invariant Sets and Predictive Control”. PhD thesis. Cambridge, UK: Department of Engineering, University of Cambridge (cit. on pp. 54, 57).
- Kerrigan, Eric Colin (2001). “Robust constraint satisfaction: Invariant sets and predictive control”. PhD thesis. University of Cambridge (cit. on pp. 67, 164, 167–169, 171).
- Khalil, H. (2000). *Nonlinear Systems*. New Jersey, USA: Pearson Education (cit. on pp. 51, 52).
- (2002). *Nonlinear Systems*. 3rd ed. New Jersey, USA: Prentice Hall (cit. on p. 32).
- Kim, E. and H. Lee (2000). “New approaches to relaxed quadratic stability condition of fuzzy control systems”. In: *IEEE Transactions on Fuzzy Systems* 8.5, pp. 523–533 (cit. on p. 39).
- Kiriakidis, Kiriakos (2007). “Nonlinear modeling by interpolation between linear dynamics and its application in control”. In: *Journal of Dynamic Systems, Measurement, and Control* 129.6, pp. 813–824 (cit. on p. 50).
- Kouramas, Koustas I, Sasa V Rakovic, Eric C Kerrigan, JC Allwright, and David Q Mayne (2005). “On the minimal robust positively invariant set for linear difference inclusions”. In: *44th IEEE Conf. on Decision and Control, European Control Conf.* IEEE, pp. 2296–2301 (cit. on pp. 65, 164).

- Kruszewski, A., A. Sala, T.M. Guerra, and C. Arino (2009). “A Triangulation Approach to Asymptotically Exact Conditions for Fuzzy Summations”. In: *IEEE Transactions on Fuzzy Systems* 17.5, pp. 985–994 (cit. on pp. 39, 74, 125).
- Kruszewski, Alexandre (2006). “Lois de commande pour une classe de modèles non linéaires sous la forme Takagi-Sugeno: Mise sous forme LMI”. PhD thesis. Université de Valenciennes et du Hainaut-Cambresis (cit. on p. 29).
- Kuroiwa, D (1996). “Convexity for set-valued maps”. In: *Applied Mathematics Letters* 9.2, pp. 97–101 (cit. on p. 166).
- Kvasnica, Michal, Pascal Grieder, Mato Baotić, and Manfred Morari (2004). “Multi-parametric toolbox (MPT)”. In: *International Workshop on Hybrid Systems: Computation and Control*. Springer, pp. 448–462 (cit. on p. 63).
- Kwiatkowski, Andreas and Herbert Werner (2008a). “PCA-based parameter set mappings for LPV models with fewer parameters and less overbounding”. In: *IEEE Transactions on Control Systems Technology* 16.4, pp. 781–788 (cit. on pp. 67, 164, 167, 184).
- (2008b). “PCA-based parameter set mappings for LPV models with fewer parameters and less overbounding”. In: *IEEE Transactions on Control Systems Technology* 16.4, pp. 781–788 (cit. on pp. 139, 151, 153, 158–160).
- Lee, D. H. and D. W. Kim (2014). “Relaxed LMI conditions for local stability and local stabilization of continuous-time Takagi-Sugeno fuzzy systems”. In: *IEEE Transactions on Cybernetics* 44.3, pp. 394–405 (cit. on p. 139).
- Leite, Valter JS and Pedro LD Peres (2003). “An improved LMI condition for robust D-stability of uncertain polytopic systems”. In: *IEEE Transactions on Automatic Control* 48.3, pp. 500–504 (cit. on p. 141).
- Lendek, Z., T.M Guerra, R. Babuska, and B. De-Schutter (2010). *Stability Analysis and Nonlinear Observer Design Using Takagi-Sugeno Fuzzy Models*. Netherlands: Springer-Verlag (cit. on pp. 74, 90, 138).
- Löfberg, J. (2004). “YALMIP : A Toolbox for Modeling and Optimization in MATLAB”. In: *Proceedings of the 2004 IEEE International Symposium on*

-
- Computer Aided Control Systems Design*. Taipei, Taiwan, pp. 284–289 (cit. on p. 12).
- Luenberger, David G and Ami Arbel (1977). “Singular dynamic Leontief systems”. In: *Econometrica: Journal of the Econometric Society*, pp. 991–995 (cit. on p. 49).
- Lur’e, Anatolii Isakievich (1957). *Some non-linear problems in the theory of automatic control*. Her Majesty’s stationery office (cit. on p. 12).
- Lyapunov, A. (1992). *The general problem of the stability of motion*. London; Washington, DC: Taylor & Francis (cit. on p. 33).
- Lyapunov, A.M. (1892). “The general problem of the stability of motion”. in Russian. PhD thesis. Kharkov Mathematical Society (cit. on p. 14).
- Mamdani, E. and S. Assilian (1975). “An Experiment in Linguistic synthesis with a fuzzy logic controller”. In: *International Journal of Man-Machine Studies* 7.1, pp. 1–13 (cit. on p. 26).
- Marquez, R., T.M. Guerra, A. Kruszewski, and M. Bernal (2013). “Improvements on Non-quadratic Stabilization of Takagi-Sugeno Models via Line-Integral Lyapunov Functions”. In: *Proceedings of the 2013 IFAC International Conference on Intelligent Control and Automation Science*. Chengdu, China, pp. 473–478 (cit. on p. 47).
- Matía, Fernando, G Nicolás Marichal, and Emilio Jiménez (2014). *Fuzzy Modeling and Control: Theory and Applications*. Springer (cit. on pp. 74, 76).
- Mayne, David Q (2001). “Control of constrained dynamic systems”. In: *European Journal of Control* 7.2, pp. 87–99 (cit. on p. 164).
- Meyer, Carl D (2000). *Matrix analysis and applied linear algebra*. Siam (cit. on p. 94).
- Mohammadpour, J. and C.W. Scherer (2012). *Control of Linear Parameter Varying Systems with Applications*. New York, USA: Springer (cit. on pp. 138, 210).

- Mohammadpour, Javad and Carsten W Scherer, eds. (2012). *Control of linear parameter varying systems with applications*. Springer (cit. on p. 172).
- Moore, Ramon E, Fritz Bierbaum, and Klaus-Peter Schwiertz (1979). *Methods and applications of interval analysis*. Vol. 2. SIAM (cit. on p. 111).
- Mozelli, L.A., R.M. Palhares, F.O. Souza, and E.M. Mendes (2009). “Reducing conservativeness in recent stability conditions of TS fuzzy systems”. In: *Automatica* 45.6, pp. 1580–1583 (cit. on p. 47).
- Murty, J.D. and S.N. Kabadi (1987). “Some NP-complete problems in quadratic and nonlinear programming”. In: *Mathematical Programming* 39.2, pp. 117–129 (cit. on p. 37).
- Nagy, Szabolcs, Zoltán Petres, Péter Baranyi, and Hideki Hashimoto (2009). “Computational relaxed TP model transformation: restricting the computation to subspaces of the dynamic model”. In: *Asian Journal of Control* 11.5, pp. 461–475 (cit. on pp. 90, 121, 124).
- Narimani, M. and H.K. Lam (2009). “Relaxed LMI-Based Stability Conditions for Takagi-Sugeno Fuzzy Control Systems Using Regional-Membership-Function-Shape-Dependent Analysis Approach”. In: *IEEE Transactions on Fuzzy Systems* 17.5, pp. 1221–1228 (cit. on p. 44).
- Nesterov, Yurii and A Nemirovsky (1988). *A general approach to polynomial-time algorithms design for convex programming*. Tech. rep. (cit. on p. 12).
- Oliveira, M.C. and R.C. Skelton (2001). “Stability tests for constrained linear systems”. In: *Perspectives in robust control*. Vol. 268. Lecture Notes in Control and Information Sciences. Berlin: Springer-Verlag, pp. 241–257 (cit. on p. 12).
- Orsi, R. (2005). *LMIRank: Software for rank constrained LMI problems*. <http://users.cecs.anu.edu.au/~robert/lmirank/> (cit. on p. 12).
- Packard, A. (1994). “Gain scheduling via linear fractional transformations”. In: *Systems & Control Letters* 22.2, pp. 79–92 (cit. on pp. 141, 146).

-
- Pan, J.T, T.M. Guerra, S.M. Fei, and A. Jaadari (2012). “Nonquadratic Stabilization of Continuous T-S Fuzzy Models: LMI Solution for a Local Approach”. In: *IEEE Transactions on Fuzzy Systems* 20.3, pp. 594–602 (cit. on p. 47).
- Pandey, A. and M.C. DeOliveira (2017). “Quadratic and Poly-Quadratic Discrete-Time Stabilizability of Linear Parameter-Varying Systems”. In: *IFAC-PapersOnLine* 50.1, pp. 8624–8629 (cit. on p. 47).
- Peaucelle, D., D. Arzelier, O. Bachelier, and J. Bernussou (2000). “A new robust D-stability condition for real convex polytopic uncertainty”. In: *Systems and Control Letters* 40.1, pp. 21–30 (cit. on p. 13).
- Petersson, Daniel and Johan Löfberg (2009). “Optimization based LPV-approximation of multi-model systems”. In: *Control Conference (ECC), 2009 European*. IEEE, pp. 3172–3177 (cit. on pp. 90, 124).
- Pluymers, B, JA Rossiter, JAK Suykens, and Bart De Moor (2005). “The efficient computation of polyhedral invariant sets for linear systems with polytopic uncertainty”. In: *American Control Conference, 2005. Proceedings of the 2005*. IEEE, pp. 804–809 (cit. on p. 62).
- Prajna, S., A. Papachristodoulou, and F. Wu (2004). “Nonlinear control synthesis by sum of squares optimization: A Lyapunov-based approach”. In: *Proceedings of the 5th Asian Control Conference*. Vol. 1. Melbourne, Australia, pp. 157–165 (cit. on pp. 48, 138, 142).
- Prajna, Stephen, Antonis Papachristodoulou, and Fen Wu (2004). “Nonlinear control synthesis by sum of squares optimization: A Lyapunov-based approach”. In: *Control conference, 2004. 5th asian*. Vol. 1. IEEE, pp. 157–165 (cit. on p. 168).
- Qiu, J., G. Feng, and H. Gao (2012). “Observer-Based Piecewise Affine Output Feedback Controller Synthesis of Continuous-Time T-S Fuzzy Affine Dynamic Systems Using Quantized Measurements”. In: *IEEE Transactions on Fuzzy Systems* 20.6, pp. 1046–1062 (cit. on p. 46).
- (2013). “Static-Output-Feedback H_∞ Control of Continuous-Time T-S Fuzzy Affine Systems Via Piecewise Lyapunov Functions”. In: *IEEE Transactions on Fuzzy Systems* 21.2, pp. 245–261 (cit. on p. 46).

- Quanser, Inc (2006). *Mechatronics Control Kit User's Manual (Instructor)*. Natick, MA: Mathworks, inc (cit. on p. 40).
- Rakovic, SV, P Grieder, M Kvasnica, DQ Mayne, and M Morari (2004). "Computation of invariant sets for piecewise affine discrete time systems subject to bounded disturbances". In: *43rd IEEE Conf. on Decision and Control*. Vol. 2. IEEE, pp. 1418–1423 (cit. on p. 164).
- Rhee, B.J. and S. Won (2006). "A new fuzzy Lyapunov function approach for a Takagi-Sugeno fuzzy control system design". In: *Fuzzy Sets and Systems* 157.9, pp. 1211–1228 (cit. on p. 47).
- Richard, Jean-Pierre (2003). "Time-delay systems: an overview of some recent advances and open problems". In: *automatica* 39.10, pp. 1667–1694 (cit. on p. 14).
- Robles, R., A. Sala, M. Bernal, and T. Gonzalez (2015). "Choosing a Takagi-Sugeno model for improved performance". In: *Proceedings of the 2015 IEEE International Conference on Fuzzy Systems*. Istanbul, Turkey, pp. 1–6 (cit. on pp. 4, 73, 91, 125, 127, 128, 130, 135, 139, 161).
- (2016). "Optimal-Performance Takagi-Sugeno Models via the LMI Null Space". In: (*Proc. of 4th IFAC Conf. Intell. Control and Autom. Sc.*) *IFAC-PapersOnLine* 49.5, pp. 13–18 (cit. on pp. 4, 123).
- (2017). "Subspace-Based Takagi-Sugeno Modeling for Improved LMI Performance". In: *IEEE Transactions on Fuzzy Systems* 25.4, pp. 754–767 (cit. on pp. 4, 68, 89, 125, 128, 138, 139, 184).
- Rugh, Wilson J and Jeff S Shamma (2000a). "Research on gain scheduling". In: *Automatica* 36.10, pp. 1401–1425 (cit. on pp. 67, 164, 167, 172, 184).
- (2000b). "Research on gain scheduling". In: *Automatica* 36.10, pp. 1401–1425 (cit. on p. 138).
- Sala, A. (2009). "On the conservativeness of fuzzy and fuzzy-polynomial control of nonlinear systems". In: *Annual Reviews in Control* 33.1, pp. 48–58 (cit. on pp. 43, 74, 90, 92, 93, 124, 138, 139, 141).

- Sala, A. and C. Ariño (2007). “Relaxed Stability and Performance Conditions for Takagi-Sugeno Fuzzy Systems With Knowledge on Membership Function Overlap”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 37.3, pp. 727–732 (cit. on p. 44).
- (2008). “Relaxed Stability and Performance LMI Conditions for Takagi-Sugeno Fuzzy Systems With Polynomial Constraints on Membership Function Shapes”. In: *IEEE Transactions on Fuzzy Systems* 16.5, pp. 1328–1336 (cit. on p. 44).
- Sala, A. and C. Ario (2009). “Polynomial fuzzy models for nonlinear control: a Taylor series approach”. In: *IEEE Transactions on Fuzzy Systems* 17.6, pp. 1284–1295 (cit. on pp. 48, 90, 124).
- Sala, A. and C. Ariño (2007). “Asymptotically necessary and sufficient conditions for stability and performance in fuzzy control: Applications of Polyá’s theorem”. In: *Fuzzy Sets and Systems* 158.24, pp. 2671–2686 (cit. on p. 39).
- Sala, A., T.M. Guerra, and R. Babuska (2005). “Perspectives of fuzzy systems and control”. In: *Fuzzy Sets and Systems* 156.3, pp. 432–444 (cit. on pp. 43, 74, 90, 124).
- Sala, Antonio (2005). “Computer control under time-varying sampling period: An LMI gridding approach”. In: *Automatica* 41.12, pp. 2077–2082 (cit. on p. 138).
- (2009). “On the conservativeness of fuzzy and fuzzy-polynomial control of nonlinear systems”. In: *Annual Reviews in Control* 33.1, pp. 48–58 (cit. on pp. 60, 65, 69, 164, 168, 172, 184, 186).
- (2017). “Generalising quasi-LPV and CDI models to Quasi-Convex Difference Inclusions”. In: *IFAC-PapersOnLine* 50.1. 20th IFAC World Congress, pp. 7560–7565. ISSN: 2405-8963 (cit. on pp. 164, 167, 168, 171, 172, 185, 186).
- Sala, Antonio and Pedro Albertos (1998). “Fuzzy systems evaluation: The inference error approach”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 28.2, pp. 268–275 (cit. on p. 26).

- Sala, Antonio and Carlos Arino (2007). “Asymptotically necessary and sufficient conditions for stability and performance in fuzzy control: Applications of Polyá’s theorem”. In: *Fuzzy Sets and Systems* 158.24, pp. 2671–2686 (cit. on p. 178).
- Sala, Antonio, Manuel Hernández-Mejías, and Carlos Ariño (2017). “Stable receding-horizon scenario predictive control for Markov-jump linear systems”. In: *Automatica* 86, pp. 121–128 (cit. on pp. 14, 54).
- Sala, Antonio and José Luis Pitarch (2016a). “Optimisation of transient and ultimate inescapable sets with polynomial boundaries for nonlinear systems”. In: *Automatica* 73, pp. 82–87 (cit. on p. 138).
- (2016b). “Optimisation of transient and ultimate inescapable sets with polynomial boundaries for nonlinear systems”. In: *Automatica* 73, pp. 82–87 (cit. on p. 168).
- Scherer, C. (2004). *Linear Matrix Inequalities in Control Theory*. Delf, The Netherlands: Delf University (cit. on pp. 10, 11, 20, 143).
- Scherer, Carsten (2001). “LPV control and full block multipliers”. In: *Automatica* 37.3, pp. 361–375 (cit. on p. 138).
- Scherer, Carsten W (2006). “LMI relaxations in robust control”. In: *European Journal of Control* 12.1, pp. 3–29 (cit. on pp. 172, 178).
- Schulte, Horst and Kevin Guelton (2009). “Descriptor modelling towards control of a two link pneumatic robot manipulator: AT–S multimodel approach”. In: *Nonlinear Analysis: Hybrid Systems* 3.2, pp. 124–132 (cit. on p. 49).
- Setnes, Magne, Robert Babuška, Uzay Kaymak, and Hans R van Nauta Lemke (1998). “Similarity measures in fuzzy rule base simplification”. In: *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 28.3, pp. 376–386 (cit. on pp. 74, 124).
- Shamma, J. S. and J. R. Cloutier (1992). “A linear parameter varying approach to gain scheduled missile autopilot design”. In: *American Control Conference, 1992*. IEEE, pp. 1317–1321 (cit. on p. 160).

-
- Sontag, Eduardo D (1999). “Control-lyapunov functions”. In: *Open problems in mathematical systems and control theory*. Springer, pp. 211–216 (cit. on p. 59).
- Sturm, J.F. (1999). “Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones”. In: *Optimization Methods and Software* 11.1-4, pp. 625–653 (cit. on pp. 12, 16).
- Su, Xiaojie, Ligang Wu, Peng Shi, and Yong-Duan Song (2014). “A novel approach to output feedback control of fuzzy stochastic systems”. In: *Automatica* 50.12, pp. 3268–3275 (cit. on p. 47).
- Sugeno, M. and G.T. Kang (1988). “Structure Identification of Fuzzy Model”. In: *Fuzzy Sets and Systems* 28.1, pp. 15–33 (cit. on p. 37).
- Takagi, T. and M. Sugeno (1985). “Fuzzy identification of systems and its applications to modeling and control”. In: *IEEE Transactions on Systems, Man and Cybernetics* 15.1, pp. 116–132 (cit. on pp. 25, 74, 90).
- Tanaka, K., T. Hori, and H.O. Wang (2001). “A fuzzy Lyapunov approach to fuzzy control system design”. In: *Proceedings of the 2001 American Control Conference*. Vol. 6. Arlington, USA, pp. 4790–4795 (cit. on p. 47).
- (2003). “A multiple Lyapunov function approach to stabilization of fuzzy control systems”. In: *IEEE Transactions on Fuzzy Systems* 11.4, pp. 582–589 (cit. on p. 46).
- Tanaka, K., T. Ikeda, and H.O. Wang (1998). “Fuzzy regulators and fuzzy observers: Relaxed stability conditions and LMI-based designs”. In: *IEEE Transactions on Fuzzy Systems* 6.2, pp. 250–264 (cit. on p. 43).
- Tanaka, K. and M. Sano (1994). “A robust stabilization problem of fuzzy control systems and its application to backing up control of a truck-trailer”. In: *IEEE Transactions on Fuzzy Systems* 2.2, pp. 119–134 (cit. on p. 38).
- Tanaka, K. and H.O. Wang (2001). *Fuzzy Control Systems Design and Analysis. A linear matrix inequality approach*. New York: John Wiley and Sons (cit. on pp. 26, 35, 74, 90, 117, 121, 124, 128, 138, 139, 141).

- Tanaka, Kazuo and Hua O Wang (2004). *Fuzzy control systems design and analysis: a linear matrix inequality approach*. John Wiley & Sons (cit. on pp. 67, 164, 167, 172, 184).
- Taniguchi, T., K. Tanaka, and H.O. Wang (2001). “Model construction, rule reduction and robust compensation for generalized form of Takagi-Sugeno fuzzy systems”. In: *IEEE Transactions on Fuzzy Systems* 9.2, pp. 525–537 (cit. on pp. 74, 75, 91, 124).
- Taniguchi, T., K. Tanaka, K. Yamafuji, and H.O. Wang (1999). “Fuzzy descriptor systems: stability analysis and design via LMIs”. In: *Proceedings of the 1999 American Control Conference*. Vol. 3. San Diego, USA, pp. 1827–1831 (cit. on p. 43).
- Tóth, Roland, Peter SC Heuberger, and Paul MJ Van den Hof (2012). “Prediction-error identification of LPV systems: Present and beyond”. In: *In Mohammadpour and Scherer 2012*. Springer, pp. 27–58 (cit. on p. 138).
- Trofino, A and TJM Dezuo (2014). “LMI stability conditions for uncertain rational nonlinear systems”. In: *International Journal of Robust and Nonlinear Control* 24.18, pp. 3124–3169 (cit. on pp. 139, 142).
- Tuan, H.D., P. Apkarian, T. Narikiyo, and Y. Yamamoto (2001). “Parameterized linear matrix inequality techniques in fuzzy control system design”. In: *IEEE Transactions on Fuzzy Systems* 9.2, pp. 324–332 (cit. on pp. 39, 128).
- Tütüncü, R. H., K. C. Toh, and M. J. Todd (2003). “Solving semidefinite-quadratic-linear programs using SDPT3”. In: *Mathematical programming* 95.2, pp. 189–217 (cit. on p. 12).
- Wang, H.O., K. Tanaka, and M. Griffin (1996). “An approach to fuzzy control of nonlinear systems: Stability and design issues”. In: *IEEE Transactions on Fuzzy Systems* 4, pp. 14–23 (cit. on pp. 74, 124).
- Wang, L.X. (1997). *A Course in Fuzzy Systems and Control*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc. (cit. on pp. 25, 90).
- Willems, Jan (1971). “Least squares stationary optimal control and the algebraic Riccati equation”. In: *IEEE Transactions on Automatic Control* 16.6, pp. 621–634 (cit. on p. 14).

- Wu, Fen and Ke Dong (2006a). “Gain-scheduling control of LFT systems using parameter-dependent Lyapunov functions”. In: *Automatica* 42.1, pp. 39–50 (cit. on pp. 47, 138).
- (2006b). “Gain-scheduling control of LFT systems using parameter-dependent Lyapunov functions”. In: *Automatica* 42.1, pp. 39–50 (cit. on pp. 65, 164, 168).
- Wu, Fen, Xin Hua Yang, Andy Packard, and Greg Becker (1996). “Induced \mathcal{L}_2 -norm control for LPV systems with bounded parameter variation rates”. In: *International Journal of Robust and Nonlinear Control* 6.9-10, pp. 983–998 (cit. on p. 138).
- Yakubovich, V.A. (1965). “The method of matrix inequalities in the stability theory of nonlinear control systems. ii- absolute stability in a class of nonlinearities with a condition on the derivative(Matrix inequalities to obtain frequency condition for improving stability of nonlinear control systems)”. In: *Automation and Remote control* 26, pp. 577–592 (cit. on p. 12).
- Yoneyama, J., M. Nishikawa, H. Katayama, and A. Ichikawa (2000). “Output stabilization of Takagi-Sugeno fuzzy systems”. In: *Fuzzy Sets and Systems* 111.2, pp. 253–266 (cit. on p. 43).

