A Set-Based Cosimulation Method to Overapproximate the Reachable Set of an Interconnection of Dynamical Systems

Paul Rousse

Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Institut Supérieur de l'Aéronautique et de l'Espace

---

Présentée et soutenue par

Paul ROUSSE

Le 5 novembre 2020

Cosimulation ensembliste d'une interconnexion de systèmes.

Ecole doctorale : **AA - Aéronautique, Astronautique**

Spécialité : **Automatique**

Unité de recherche :
ISAE-ONERA MOIS MOdélisation et Ingénierie des Systèmes

Thèse dirigée par
Pierre-Loïc GAROCHE et Didier HENRION

Jury

M. Anders RANTZER, Rapporteur
Mme Thao DANG, Rapporteure
Mme Sylvie PUTOT, Examinatrice
M. Eric FERON, Examinateur
M. Alexandre CHAPOUTOT, Examinateur
M. Pierre-Loïc GAROCHE, Directeur de thèse
M. Didier HENRION, Co-directeur de thèse
Mme Sonia CAFIERI, Présidente

ONERA
THE FRENCH AEROSPACE LAB

A Set-Based Cosimulation Method to Overapproximate the Reachable Set of an Interconnection of Dynamical Systems

Paul Rousse

Remerciements

Tout d'abord, je souhaite remercier mes encadrants Pierre-Loïc Garoche et Didier Henrion. Leur expertise académique, leur soutien constant et leur optimisme sans faille durant mes travaux de thèse m'ont été précieux. Vous m'avez donné beaucoup de liberté dans le choix des sujets que j'ai abordés et je vous en suis infiniment reconnaissant. Enfin, vous m'avez permis de rencontrer différentes communautés de chercheurs (lors de collaborations et lors de conférences), c'était pour moi un sacré coup de pouce.

Cette thèse a été possible grâce au soutien financier de la Direction Générale de l'Armement (DGA) et celui de l'Office National d'Études et de Recherches Aérospatiales (ONERA). Je remercie également l'École Doctorale Aéronautique Astronautique (EDAA) et l'École Supérieure d'Aéronautique(ISAE-Supaéro).

I would like to thanks Thao Dang and Anders Rantzer for their review of this manuscript and their interesting feedback on my work. I also thank each member of the jury for attending my thesis defense which was a wonderful moment for me despite the complex sanitary condition.

Je remercie grandement Alexandre Chapoutot et Julien Alexandre Dit Sandretto de m'avoir accueilli dans leur équipe à l'ENSTA ParisTech. Travailler à vos côtés est toujours un réel plaisir, et nos discussions sont souvent fructueuses.

I am grateful to John Hauser for hosting me at Boulder. Besides his expert knowledge about control and aerospace systems, John is a wonderful person who introduced me to *sous-vide* cooking, outdoor rock-climbing and gave me valuable English and skying lessons.

J'ai eu le plaisir de travailler avec Pierre-Loïc Garoche, Xavier Thirioux et Loïc Dubois à l'occasion du stage de L2 de ce dernier. Malgré la durée epsilonesque du stage, Loïc a travaillé avec ardeur et nos échanges ont été passionnants.

Durant cette thèse, j'ai rencontré plusieurs communautés de chercheurs au travers de différents projets: à Feanicses, à CAFEIN, dans COVNI au DTIS et dans l'équipe MAC au LAAS. Ces projets ont été pour moi autant d'ouvertures à des domaines qui

m'étaient souvent inconnus et ils m'ont permis de faire des rencontres formidables.

Je remercie également les doctorants du DTIS, Guillaume, Hugo, David, Kosta, Nour, Jeanne, César, Nathanaël, Pauline, Sébastien, Valentin, Sovanna, Oktay, Thomas, Quentin, Félix, Arthur, Margot, Mathias, Clément, Matthieu et Emmanuelle; en votre compagnie, la Sodexo avait bon goût. Et bien entendu, j'ai une pensée affectueuse pour Lucien et Nadir, pour nos ballades dans les Pyrénnées, celles dans les rues de Toulouse, celles sur l'eau et celles sur les murs d'escalade.

Je remercie mes amis. Quentin, Matthieu, Quentin, Thomas, Nicolas, Clément, nos pérégrinations nous rassemblent jamais suffisamment, rendez-vous à nos 30ans. Mes amis français de Suède avec qui j'explore le monde, Guillaume, Alice, Jérémy, Anaïs, Éliane, Rémy, Nicolas, Juliette, Félix. Marie et Claire, avec toutes ces ballades, les Pyrénnées n'ont désormais plus de secrets pour vous. Michel, pour nos discussions d'électronique et de saumon fumé. Antoine, pour nos randos en France et ailleurs. Louis, un cycliste hors pair, sans aucun doute un futur maillot jaune. Et enfin, Thierry, Magali et Tess, des voisins d'une qualité rare.

Durant cette thèse j'ai eu la chance de vivre en colocation. D'abord avec Mathieu, un infatigable optimiste et chercheur talentueux (rendez-vous au prochain Hash-code). Puis avec Anne-Catherine, ma *covid buddy*, un véritable rayon de soleil en période de pandémie mondiale (Paul Bocuse te remercie par ailleurs). Et enfin, Astrid, débordante d'énergie, une adversaire coriace au Catan et capable de se mesurer à Anne-Catherine (je ne sais laquelle des deux est la meilleure)!

Pour terminer, je remercie ma famille. Mes parents, ma sœur et mon frère, pour leur amour inconditionnel, leur support dans mes projets (même insensés) et pour leur présence. Mon père, de m'avoir donné le goût de l'ingénierie et de la technique. Ma mère, pour m'avoir appris à travailler. Ma sœur et mon frère, pour leur joie de vivre, leur humour, et les moments que l'on partage.

Contents

Abstract

The reachable set of a cyber-physical system is of great interest when it comes to verification of safety properties. Such systems usually models dynamical systems (a physical system) controlled by an embedded system (computer program). Its reachable set is known to be a complex geometrical object (sometimes non-convex and/or disconnected even for simple cases) and computing it is challenging. This thesis proposes three methods to overapproximate the reachable set of cyber-physical systems. The first method, the "ellipsoidal method", studies linear systems subject to a disturbance bounded in 2-norm or ∞-norm. The second method, the "interval method", studies non-linear integral quadratic constraint (IQC) systems by means of interval sets. In the last method, the "cosimulation method" studies a broader family of systems (an interconnection of dynamical systems) with an abstract interpretation approach.

The "ellipsoidal method" describes the computation of reachable sets for linear time-invariant systems with an unknown input bounded by IQC, which can models delays, rate limiters, energy bounds, or sector inequalities. The reachable set is over-approximated with a family of time-varying conics. The parameters of the conic are solutions to a Differential Riccati Equation (DRE). Our approach unifies ellipsoidal methods (for bounded disturbances) and storage function methods (classically used for IQC systems).

The "interval method" describes the use of a Runge-Kutta validated integration scheme to overapproximate the reachable set of nonlinear IQC systems. The reachable tube is overapproximated as a union of intervals in the time and state space. The IQC is used to define a contractor over each interval overapproximating the reachable tube. This contractor and a propagation step are successively applied on an a priori given overapproximation of the reachable tube until a fixed point is reached. We evaluated our algorithm with DynIbex library to simulate a delayed system, i.e., an infinite-dimensional system that can be modeled as a linear time-invariant system subject to an IQC. Our approach is shown to be tractable and it enables the use of

interval arithmetic and validated integration for a richer set of dynamical systems.

Finally, the "cosimulation method" merges the two previous methods into a unique one that can study a broader family of systems: an interconnection of systems. Each system in the interconnection is considered as an operator on a signal space (continuous-/discrete-time) and the proposed method is formalized within the abstract interpretation framework. The system reachable tube is expressed as the solution of a fixed point equation. This reachable set is overapproximated in the abstract domain with time-varying sets (e.g. time-varying intervals or ellipsoidal sets) and computed by solving a greatest fixed point equation. We apply the cosimulation method to overapproximate the reachable tube of several non-linear systems subject to IQC-bounded disturbances.

Notations

$\mathbb{N}$	natural numbers,
$\mathbb{R}$	reals,
$\mathbb{R}_+$	positive reals,
$\mathbb{R}_+^*$	strictly positive reals,
$\mathbb{R}^n$	real-valued vectors of dimension n,
$\mathbb{R}^{n\times m}$	real-valued matrix of dimension $n \times m$,
$\mathbb{S}^{n\times n}$	real-valued symmetric matrix of dimension $n \times n$,
$\wp(A)$	power set of A,
$\mathbb{P}^n$	conic sets in $\mathbb{R}^n$,
$\mathbb{IR}$	interval sets in $\mathbb{R}$,
$\underline{[\![\mathbf{x}]\!]}$	norm of $\mathbb{IR}$, $\underline{[\![\mathbf{x}]\!]} = \max_{x\in[\mathbf{x}]} \|x\|$, for $[\mathbf{x}] \in \mathbb{IR}^n$,
$\overline{[\mathbf{x}]}$	for $[\mathbf{x}] \in \mathbb{IR}$, let $\overline{[\mathbf{x}]} = \sup_{x\in[\mathbf{x}]} x$,
$\|x\|$	Euclidean norm of $x \in \mathbb{R}^n$,
$A \succeq 0$	semidefinite positive matrix $A \in \mathbb{S}^{n\times n}$, $A \succeq 0$ iff $\forall x \in \mathbb{R}^n, x^\top A x \succeq 0$,
$A \succeq B$	semidefinite positive partial order $A, B \in \mathbb{S}^{n\times n}$, $A \succeq B$ iff $A - B \succeq 0$,
Tr	trace,
det	determinant,
$A^\top$	transpose,
$\|A\|$	Frobenius norm, $\|A\| = \sqrt{tr(A^\top A)}$,
$L(I;X)$	continuous-time functions from $I = [0,T] \subseteq \mathbb{R}_+$, $T \in \mathbb{R}_+ \cup \{\infty\}$ to X,
$l(I;X)$	discrete-time functions from $I = \{1,\dots,T\} \subseteq \mathbb{N}$, $T \in \mathbb{N} \cup \{\infty\}$ to X,
$\mathscr{C}^1(I;X)$	continuously differentiable functions from I to X,
$\|f\|_\infty$	the ∞-norm, $\|f\|_\infty = \max_{t\in\mathbb{R}_+} \|f(t)\|$,
$\mathscr{L}_\infty(I;X)$	functions from I to X, bounded by the $\|\cdot\|_\infty$ norm,
$\|f\|$	the 2-norm, $\|f\| = \left(Tr\left(\int_0^\infty f^\top(t)f(t)dt\right)\right)^{1/2}$,
$\mathscr{L}_2(I;X)$	functions from I to X, square integrable over I,
$\mathscr{L}_{2,\mathrm{loc}}(I;X)$	functions from I to X, locally square integrable over I,
$\langle f,g\rangle$	scalar product $\langle f,g\rangle = Tr\left(\int_0^\infty f^\top(\tau)g(\tau)d\tau\right)$,
$\langle f,g\rangle_t$	truncated scalar product, $\langle f,g\rangle_t = Tr\left(\int_0^t f^\top(\tau)g(\tau)d\tau\right)$,

Introduction

In a project life cycle, there is a high interest to find errors at the early stages of the development. Take the example of software development. When the error is detected during the test phase, every new fix in the code requires reallocating human resources. Subparts of the faulty element might as well require new development to correct the error. After the error is fixed, the project should go back through the entire manufacturing, integration, and testing phases cycle. Thus, the *cost-to-fix* is consequent. The cost-to-fix during the project has been analyzed within the aerospace industry in [Haskins et al., 2004]. Table 1 reports the average cost-to-fix of an error during the project phase in software development for aerospace companies. This "exponential growth" of the cost justifies investing in solutions that can identify these errors at an early stage in the development. In the case of software development for the automotive industry, these solutions should analyze a complex class of systems: *embedded systems*.

An embedded system generally represents a mechanical system controlled by a computer program, e.g. an automatic car drove by an embedded computer. At a given time, such a system can be represented by its state (e.g. the position of a car, its velocity, and the state of the on-board computer). The evolution of the state

Project Phase	**Cost-to-fix Factor**
Requirements	1×
Design	4×
Build	16×
Test	61×
Operations	157×

Table 1: *Average cost-to-fix of errors during the project phase in software development (Table 8 in [Haskins et al., 2004]).*

through time is described by the system dynamic (e.g. a part of the dynamic of the car might be represented with the equation of motion described with the theory of rigid body, the other part might be represented as a state machine modeling the computer program). Any trajectory of the system is then a time-dependent function that associates to a time instant, the system state evolving according to the system dynamic.

Embedded systems are usually subject to rigorous safety requirements. Often, these requirements can be expressed as a property over the state space of the system. The system is *safe* if its trajectories avoid the set of states, the *unsafe* set, where this property no longer holds. If there are only a few system trajectories, it is sufficient to compute them by simulating the system in order to prove or disprove that the system is safe. When such a solution is not possible (because it is too computationally demanding, or because there is an infinite number of trajectories), verifying the safety requirement in every possible scenario becomes more challenging.

A possible approach is to verify only a subset of the trajectories. This approach is formalized using Monte-Carlo methods in [Rubinstein and Kroese, 2016]. In this framework, the probability distribution of the system trajectories is estimated by randomly choosing a set of trajectories with respect to probability distribution of uncertain system parameters or inputs. Then the probability of a safety property can be numerically estimated. The Monte-Carlo method is versatile and can be applied to a large variety of systems to verify complex safety properties (e.g. temporal properties for hybrid systems in [Sankaranarayanan and Fainekos, 2012]). These simulation-based methods require a sample of input sets. Usually, these input sets are large. For example, they can be uncountable sets (e.g. the initial position of a car) and/or they can have an infinite number (e.g. a continuous-time real signal might model the longitudinal effect of the wind over the car). It is therefore not possible to cover all the behaviors of the system.

Another method is to manipulate infinite sets of trajectories instead of reasoning over a finite set of trajectories. In such an approach, trajectories are represented as a subset of the time- and state-space: the *reachable tube*. Then the system satisfies the safety properties if the reachable tube does not intersect with the unsafe set (i.e. the set of states violating the safety property).

Computing this reachable tube is usually difficult. For most systems, there exist no "out-of-the-box" methods to compute the reachable set. The first step is to abstract the dynamical systems with a class of uncertain systems that encapsulate the set of behaviors. However, even for simple systems, the reachable set is a complex geometrical object that does not admit a simple geometrical representation (e.g. the reachable set of a linear time-invariant system is not a semi-algebraic). Most of the time, the reachable tube is therefore only overapproximated, and if the safety properties are satisfied by the overapproximation, then the system satisfies these safety

properties.

In practice, to compute these overapproximations, we use computer-representable geometrical sets (e.g. intersection of hyperplanes, ellipsoidal sets, superlevel set of polynomial functions). The most precise is the overapproximation, the more properties can be proved over the reachable set. Therefore, there is a high interest to overapproximate the reachable set with a geometrical template that correctly fits the reachable set. Finding the "good" geometrical template is complex. Since these templates do not exactly represent the set of reachable states, they introduce some non-existing trajectories into the overapproximation. If these overapproximations are too pessimistic, it might be impossible to prove that the system satisfies some safety property. At the same time, these geometrical templates should not be too costly to compute. The "good" template is highly dependent on the system (its dynamic, external noises sets, the initial set of states).

This thesis proposes methods to overapproximate the reachable tube for specific classes of embedded systems.

1 Related works

In the case of dynamical systems where trajectories are solutions to an ordinary differential equation with bounded unknown input disturbances, the set of reachable states can be computed by solving an optimal control problem [Lee and Markus, 1969, Gusev and Zykov, 2018]. For a given state and a given cost function that associates to each initial state a positive cost (and a negative cost if the state is outside the set of initial states), if the maximal cost leading to a given state is positive, then this state is reachable. When such optimal control problem is solved (using Hamilton-Jacobi-Bellman -HJB- viscosity subsolutions, see [Soravia, 2000]), the set of states associated with a positive cost corresponds to the reachable set of the system. However, HJB solutions are difficult to compute. They rely on numerical integration of (partial) differential equations and these solutions are usually not regular.

HJB based methods propagate a cost function along with the flow of the dynamical system. Occupation measures and barrier certificates methods aim at finding constraints over the reachable tube of a dynamical system: [Wang et al., 2016] uses Integral Quadratic Constraints (IQC) for verification purposes using barrier certificates where the positivity of the energetic state is ensured by using a nonnegative constant multiplier: [Henrion and Korda, 2014, Korda, 2016] use an occupation measure approach where the IQC can potentially be incorporated as a constraint over the moment of the trajectories (note however that these references do not deal explicitly with IQCs). A hierarchy of semi-definite conditions is derived for polynomial dynamics. Then, off-the-shelf Semi-Definite Programming (SDP) solvers are used to solve

the feasibility problem. Optimization-based methods do not usually take advantage of the model structure as they consider a large class of systems (convex, Lipschitz, or polynomial dynamics for example). Similarly than for HJB methods, moment-based methods can be used for a large class of systems, but they do not scale well, i.e. they are limited to systems with a small number of states.

When the dynamical system has a convex reachable set, the set can be described by an intersection of hyperplanes. These hyperplanes are obtained by finding trajectories maximizing a linear cost. This optimization problem can be solved in practice using the Pontryagin Maximum Principle -PMP- (see [Lee and Markus, 1969, Graettinger and Krogh, 1991, Varaiya, 2000, Gusev and Zykov, 2018]). By solving this problem for different cost direction, it is possible to describe the reachable set as an intersection of hyperplanes.

The case of Linear Time-Varying (LTV) systems with ellipsoidal bounded inputs is studied in [Chernous'ko, 1999, Kurzhanski and Varaiya, 2002, Kurzhanskiy and Varaiya, 2007]. Such systems can model infinity norm bounded input-output Linear Time-Invariant (LTI) systems. The reachable set (which is convex and bounded; see [Kurzhanski and Varaiya, 2002]) can be overapproximated with time-varying ellipsoidal sets. Each ellipsoid is described by its parameters (center and radius) that are solutions to an Initial Value Problem (IVP). These parameters produce tight ellipsoids (i.e., ellipsoids touching the reachable set) which are external approximations of the reachable set. When multiple ellipsoids with different touching trajectories are considered, their intersection is a strictly smaller overapproximation of the reachable set. The accuracy of the overapproximation can be made arbitrarily small by adding more well-chosen ellipsoids. The exact representation of the reachable set is possible by using an uncountable set of ellipsoids.

The study of LTI systems with IQC constraint is closely related to the Linear Quadratic Regulator (LQR) problem. In the LQR problem, a quadratic integral is minimized at the terminal time. Optimal trajectories belong to a time-varying parabolic surface, whose quadratic coefficients are a solution to a Differential Riccati Equation (DRE). [Savkin and Petersen, 1996b, Guseinov and Nazlipinar, 2011, Gusev and Zykov, 2018] describe the reachable set of LTI systems with terminal IQC. [Jönsson, 2002] formalizes the problem with a game theory approach. [Seiler et al., 2019] solves the differential Riccati inequality over a finite horizon using a basis of polynomial functions, then an SDP solver (such as Sedumi in [Sturm, 1999]) searches for a solution that minimizes the final volume of the overapproximation. This algorithm has been implemented in available tools (see LTVTools toolbox, [Seiler et al., 2017]). In all these works, the overapproximation of the reachable set is conditioned by the existence of a solution to the DRE over the interval of integration. In the case of unstable systems, there exists no stable solution to the continuous algebraic Riccati equation. Any reachable set overapproximation is then defined only over a

finite interval of time.

A second approach to reachability analysis is set-based methods. Here, we try to find a set invariant to the trajectories of the system. Usually, these methods are expressed as a greatest fixed point equation and are iteratively solved by removing infeasible trajectories from the set. In these frameworks, the sets are not necessarily defined as level sets and operations are focused over set operations (e.g. intersection, union, Minkowski sum, and inclusion). Contrary to previous level set methods, the family of systems that can be analyzed is usually larger since no system structure is assumed. [Moore et al., 2009] introduced the validated numerical integration framework where numerical integration schemes (such as the Runge-Kutta one) are redefined over the set of real intervals. The system trajectories are overapproximated as a union of intervals in the state- and time-space. Other shapes (e.g. zonotopes -the projection of a hypercube-, polytopes) are used for their ability to compute the Minkowski (see [Girard, 2005]).

These fix-point algorithms are also used in compositional methods where the system is decomposed into a closed-loop composition of systems as in [Chen and Sankaranarayanan, 2016]. In such cases, a prior overapproximation of each internal signal is iteratively refined through computation. In such an approach, interacting dynamics between subsystems are neglected, leading to more conservative overapproximations, however, systems of higher dimensions can be treated. It has been used for stability analysis of continuous-time systems [Platzer and Clarke, 2009, Eqtami and Girard, 2019] in level set-based approaches for reachability analysis.

2 Outline and Contribution

The thesis is organized in three parts (see Table 2), Part I and Part II derive methods to overapproximate the reachable tube for linear systems subject to bounded disturbances (bounded by a quadratic inequality over the signal space). Part I overapproximates by mean of quadratic superlevel sets, Part II overapproximates with interval sets. Part III describes a general framework to reason about the interconnection of systems.

In Part I, we study linear systems subject to bounded disturbances. The disturbance set is defined with a quadratic inequality between the disturbance signal, the state signal, and the input signal. Chapter 1 derives the general framework of overapproximation of the reachable set using the level set method with quadratic forms. Chapter 2 and Chapter 3 are applying previously derived results for two specific disturbance sets. Chapter 2 details the case where the set of disturbances is bounded at any time. Since there an infinite number of overapproximations, we show how the minimal volume overapproximation can be computed using a contin-

Chapter	System	Framework	Sets
Chapter 1	CT LTV + bounded disturbance	levelset method	conic sets
Chapter 2	CT LTV + QC	levelset method	conic sets
Chapter 3	CT LTV + IQC	levelset method	conic sets
Chapter 4	CT NL + NLIC	interval arithmetic	intervals
Chapter 5	interconnection of systems	abstract interpretation	various

Table 2: *Summary of the systems studied in this thesis. Abbreviations: CT (Continuous Time), LTV (Linear Time-Varying), NL (Non Linear), QC (Quadratic Constraint), IQC (Integral Quadratic Constraint), and NLIC (Non Linear Integral Constraint).*

uation method. Chapter 3 studies the case where the disturbance is constrained by an Integral Quadratic Constraint (IQC). We show that the reachable set is exactly characterized as an intersection of previously defined overapproximations. The work presented in Chapter 3 has been published in the paper "Rousse, P., Garoche, P.-L., and Henrion, D. (2019). Parabolic set simulation for reachability analysis of linear time invariant systems with integral quadratic constraint. In *2019 18th European Control Conference, ECC 2019*", and in the paper "Rousse, P., Garoche, P.-L., and Henrion, D. (2020b). Parabolic Set Simulation for Reachability Analysis of Linear Time-Invariant Systems with Integral Quadratic Constraint. *European Journal of Control*".

In Part II, we study how the validated numerical integration method builds upon interval arithmetic can be used to overapproximate the reachable set of an IQC system. Chapter 4 presents the classical interval arithmetic framework and the guaranteed integration framework, and extend it to study a nonlinear IQC system. This work has been published in the paper "Rousse, P., Alexandre dit Sandretto, J., Chapoutot, A., and Garoche, P.-L. (2020a). Guaranteed Simulation of Dynamical Systems with Integral Constraints and Application on Delayed Dynamical Systems. In *Lecture Notes in Computer Science*, volume 11971 LNCS".

In Part III, we more specifically study the interconnection of systems In Chapter 5, we present the classical abstract interpretation framework and introduce the concrete semantic of the interconnection of systems, then, we describe abstract domains for signal spaces that will be used to represent the trajectories of the interconnection of the system. Finally, we treat several examples to compute the reachable set of the interconnection of systems.

Part I

Ellipsoidal Methods

Chapter 1

Ellipsoidal Methods

Contents

This chapter proposes a common framework for ellipsoidal methods applied for the reachability analysis of linear time-varying systems subject to bounded disturbances. Ellipsoidal methods have been studied in the '80s mainly for dynamical systems subject to ∞-norm bounded disturbances (at any time the disturbance belongs to a bounded ellipsoidal set). Since then, it has been extended disturbances subjects to state-input-disturbance inequality, namely to QC disturbances (Quadratic Constraint disturbances) and IQC disturbances (Integral Quadratic Constraint disturbances). These models have been widely used in the robust control community for stability analysis of nonlinear systems (among other applications). We show that such sets of disturbances can be described by a set of quadratic constraints in the signal space. It results in an elegant approach to present ellipsoidal methods for a wide family of models.

Section 1.1 defines the system of interest, the disturbance set, and the set of reachable states. Section 1.2 introduces the level set approach in order to overapproximate

the set of reachable states. Section 1.3 introduces conic templates and there corresponding quadratic value functions. Section 1.4 applies the results of Section 1.2 to define overapproximating time-varying conic sets. The time-varying coefficients of their corresponding value function are the solution of a DRE (Differential Riccati Equation) parametrized by positive multipliers. This section provides two results about overapproximating the reachable set (Theorem 1.1 and Theorem 1.2). The domain of definition of the DRE is then analyzed in Section 1.5. This chapter ends with Section 1.6 with a discussion about the conservatism of the approach (the sources of pessimism) and a comparison between the proposed approach with state of the art found in the literature.

Theorem 1.1 and Theorem 1.2 are then applied within Chapter 2 for QC systems (i.e. with ∞-norm constrained disturbances) and in Chapter 3 for IQC systems (i.e. with 2-norm constrained disturbances).

1.1 Dynamical systems

In this part, the system of interest is a linear time-varying system perturbed by an unknown disturbance that satisfies a set of constraints with the state and input trajectories. The constraints are expressed as a quadratic form in the signal space for the

Definition 1.1. Dynamical system

For a given input signal $u \in \mathscr{L}_\infty(\mathbb{R}_+;\mathbb{R}^p)$, given time-varying matrices $A \in \mathscr{L}_\infty(\mathbb{R}_+;\mathbb{R}^{n\times n})$, $B \in \mathscr{L}_\infty(\mathbb{R}_+;\mathbb{R}^{n\times m})$, $C \in \mathscr{L}_\infty(\mathbb{R}_+;\mathbb{R}^{n\times p})$, let the system $\mathcal{S}$ be defined as

$$\mathcal{S}: \begin{cases} \dot{x}(t) &= A(t)x(t) + B(t)w(t) + C(t)u(t) \\ y &= (x,u,w) \in \mathcal{D} \end{cases} \tag{1.1}$$

where $w \in \mathscr{L}_2(\mathbb{R}_+;\mathbb{R}^m)$ is an unknown disturbance described by $\mathcal{D} \subseteq \mathscr{L}_2(\mathbb{R}_+;\mathbb{R}^{n+m+p})$, the set of disturbances

$$\mathcal{D} = \{y \in \mathscr{L}_2(\mathbb{R}_+;\mathbb{R}^{n+m+p}) | \forall \mu \in \mathcal{D}^*, \forall t > 0, \langle y, y\rangle_{M,\mu|_t} \geq 0\}$$

where

$$\langle y, y\rangle_{M,\mu|_t} = \int_0^t y(s)^\top M(s) y(s) \mu(s) ds$$

where M is a quadratic form negative in the disturbance dimension and $\mathcal{D}^* \subseteq \mathscr{L}_{2,\mathrm{loc}}(\mathbb{R}_+;\mathbb{R}_+^*)$ is a subset of positive functions from $\mathbb{R}_+ \to \mathbb{R}_+^*$ locally integrable over $\mathbb{R}_+$.

The system $\mathcal{S}$ is a linear system since for any $(x,u,w) \in \mathcal{S}$ and $\alpha > 0$, $\alpha(x,u,w) \in \mathcal{S}$. The set of disturbance is chosen *causal* since the constraint satisfied by any signal y in $\mathcal{D}$, i.e. $\langle y,y\rangle_{M,\mu|_t} \geq 0$, for all $t \geq 0$, only depends on past values of y. One should notice that any signal y of $\mathcal{D}$ satisfies

$$\forall t > 0, \langle y,y\rangle_{M,\eta|_t} \geq 0$$

where $\eta = a\mu + b\nu$ with $\mu,\nu \in \mathcal{D}^*$ and $a,b \geq 0$. Without loss of generality, we can therefore assume that $\mathcal{D}^*$ is a convex cone.

Let a block decomposition of M be

$$M(t) = \begin{bmatrix} M_x(t) & M_{x,u}(t) & M_{x,w}(t) \\ M_{x,u}^\top(t) & M_u(t) & M_{u,w}(t) \\ M_{x,w}^\top(t) & M_{w,u}^\top(t) & M_w(t) \end{bmatrix}. \tag{1.2}$$

Let the decomposition of M in the basis $[x,1,w]$ be such that

$$\begin{aligned} M_{x1}(t) &= \pi_{x1}^\top M(t)\pi_{x1}(t), \\ M_{w,x1}(t) &= \pi_w^\top M(t)\pi_{x1}(t), \\ M_w(t) &= \pi_w^\top M(t)\pi_w^\top, \\ M_{x1w}(t) &= \pi_{x1w}^\top(t)M(t)\pi_{x1w}(t) \end{aligned} \tag{1.3}$$

with the projections

$$\pi_w = \begin{bmatrix} 0 \\ 0 \\ I_w \end{bmatrix}, \pi_{x1}(t) = \begin{bmatrix} I_x & 0 \\ 0 & u(t) \\ 0 & 0 \end{bmatrix} \text{ and } \pi_{x1w}(t) = \begin{bmatrix} \pi_{x1}(t) & \pi_w \end{bmatrix}$$

(the input $u(\cdot)$ is contained in the definition of $M_{x1}(\cdot)$, $M_{w,x1}(\cdot)$, and $M_{x1w}(\cdot)$).

The set of reachable states is defined as the time-varying set that associates to a time-instant $t \geq 0$ the set x of states that are reachable starting from a given set of initial states X_0 subset of $\mathbb{R}^n$:

Definition 1.2. Reachable set

The set of reachable states is

$$\mathcal{R}(t;X_0) = \{x(t) \,|\, (x,u,w) \in \mathcal{S},\, x(0) \in X_0\}$$

where $X_0 \subset \mathbb{R}^n$ is the set of initial states.

This chapter and the two following one propose a method to overapproximate the set of reachable states $\mathcal{R}(t;X_0)$.

Problem 1.1. Reachability problem

Find a $\mathcal{P}$ time-dependent set $t \in \mathbb{R}_+ \mapsto \mathcal{P}(t) \subseteq \mathbb{R}^n$ that overapproximates $\mathcal{R}(t;X_0)$ at any time $t \geq 0$, i.e. $\mathcal{R}(t;X_0) \subseteq \mathcal{P}(t)$.

1.2 Set-based simulation

In this part, the reachable set $\mathcal{R}(t; X_0)$, for $t \in \mathbb{R}_+$, is overapproximated by the superlevel set $\mathcal{P}(t) = \{x \in \mathbb{R}^n \mid p(t,x) \geq 0\}$ of a value function $p : \mathbb{R}_+ \times \mathbb{R}^n \mapsto \mathbb{R}$ expressed in the time and state space domain. $\mathcal{P}(t)$ overapproximates $\mathcal{R}(t; X_0)$ for any $t \geq 0$ if for any system trajectory x of $\mathcal{S}$ and any time instant $t \geq 0$, it holds $p(t, x(t)) \geq 0$, i.e.

$$x \in \mathcal{S} \Rightarrow \forall t \in \mathbb{R}_+, p(t, x(t)) \geq 0. \tag{1.4}$$

When the set of disturbances is reduced to a singleton $w = 0$, the conditions

1. $p(0, x_0)$ is positive for any initial state $x_0 \in X_0$;
2. $t \to p(t, x(t))$ is an increasing function over $\mathbb{R}_+$ for any system trajectory $x \in \mathcal{S}$

are sufficient to enforce (1.4), and therefore to have $\mathcal{R}(t; X_0) \subseteq \mathcal{P}(t)$ (i.e. $\mathcal{P}(t)$ is an overapproximation of the set of reachable states of $\mathcal{S}$, see Figure 1.1). Condition 2)

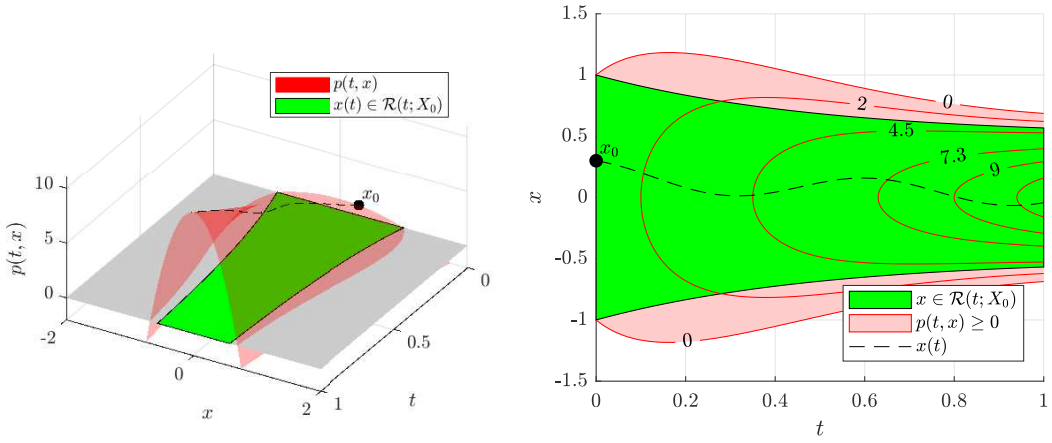


Figure 1.1: *The value function p is positive (increasing and starting from zero at the initial time) along each trajectory x of the system $\mathcal{S}$. The superlevel set $\mathcal{P}(t)$ of $x \mapsto p(t,x)$ overapproximates the reachable set $\mathcal{R}(t; X_0)$ of the system $\mathcal{S}$ at any given time $t > 0$.*

is expressed over the system trajectories. To avoid enforcing condition 2) over the set of trajectories, it can be replaced with a stronger one: for any $t \in \mathbb{R}_+$ and any $x \in \mathbb{R}^n$, the value function of p increases along the flow of $\mathcal{S}$. When p is continuous and differentiable in time and space, this condition is written

$$\frac{\partial p}{\partial t} + \frac{\partial p}{\partial x}^\top f(t,x) \geq 0, \text{ for all } x \in \mathbb{R}^n \text{ and all } t \in \mathbb{R}_+ \tag{1.5}$$

where $f(t,x) = A(t)x + C(t)u(t)$ is the vector flow of $\mathcal{S}$ (for the case where $w = 0$).

When the set $\mathcal{D}$ is not trivial, it is still possible to use the previous approach. A sufficient condition such that (1.4) holds can be derived using a S-procedure approach (see [Boyd et al., 1994], Section 2.6.3). If there is a $\mu \in \mathcal{D}^*$ such that for any $t \in \mathbb{R}_+$

$$p(t, x(t)) \geq \lambda(t)\langle y, y\rangle_{M,\mu|_t} \tag{1.6}$$

where $y = (x, u, w)$ is a system trajectory and $\lambda : \mathbb{R}_+ \mapsto \mathbb{R}_+^*$ is a strictly positive time-dependent function, then (1.4) holds and $\mathcal{R}(t; X_0) \subseteq \mathcal{P}(t)$. Similarly than in the trivial case, the following conditions are sufficient for (1.6) to hold:

1. $p(0, x_0)$ is positive for any initial state $x_0 \in X_0$
2. $t \to p(t, x(t)) - \lambda(t)\langle y, y\rangle_{M,\mu|_t}$ is an increasing function over $\mathbb{R}_+$ for any system trajectory $x \in \mathcal{S}$.

When p is continuous and differentiable in time and space, and when λ is continuous and differentiable, a sufficient condition for 2) to hold is

$$\frac{\partial p}{\partial t} + \frac{\partial p}{\partial x}^\top f(t, x_t, w_t) \geq \dot{\lambda}(t)\langle y, y\rangle_{M,\mu|_t} + \lambda(t)\mu(t)y(t)^\top M(t)y(t) \tag{1.7}$$

for all $x_t \in \mathbb{R}^n$ and all $t \in \mathbb{R}_+$, where $f(t, x_t, w_t) = A(t)x_t + B(t)w_t + C(t)u(t)$ is the vector flow of $\mathcal{S}$. Moreover, when $\dot{\lambda}$ is positive over $\mathbb{R}_+$ and when $p(t, x(t)) \geq \lambda(t)\langle y, y\rangle_{M,\mu|_t}$

$$\frac{\partial p}{\partial t} + \frac{\partial p}{\partial x}^\top f(t, x_t, w_t) \geq \dot{\lambda}(t)\lambda(t)^{-1}p(t, x_t) + \lambda(t)\mu(t)y(t)^\top M(t)y(t), \tag{1.8}$$

for any $w_t \in \mathbb{R}^m$ and any $x_t \in \mathbb{R}^n$, it implies that (1.7) holds. Contrary to (1.7), (1.8) is expressed over the state space and the disturbance space. Finally, the problem of overapproximation is reduced to proving the positivity of some function over a space of finite dimension.

In the general case, proving the positivity of some function is difficult. In the next part, this problem will be solved for a specific family of p, namely the time-varying quadratic forms for which proving the positivity of a function can be equivalently solved in its dual form by proving that the minimum is positive.

Remark 1.1. Relationship with Liouville equations

Equation (1.7) *can be interpreted in many ways. It corresponds to the Koopman equation (dual of the Liouville equation). It is as well the value function of the minimization problem* $\min_{x_0, y=(x,u,w)} p(t, x)$ *such that* y *is a system trajectory. We do not investigate further these connections in this manuscript.*

We summarize the above discussion in Proposition 1.1.

Proposition 1.1. Level-set overapproximation
For a given $\mu \in \mathcal{D}^*$, and a given $\lambda : \mathbb{R}_+ \mapsto \mathbb{R}^{+*}$ continuous, differentiable, and increasing over $\mathbb{R}_+$, let $p : \mathbb{R}_+ \times \mathbb{R}^n \mapsto \mathbb{R}$ be a continuous and differentiable function over $\mathbb{R}_+ \times \mathbb{R}^n$ satisfying

$$\frac{\partial p}{\partial t} + \frac{\partial p}{\partial x}^\top f(t,x,w) \geq \dot{\lambda}(t)\lambda(t)^{-1}p(t,x(t)) + \lambda(t)\mu(t)y(t)^\top M(t)y(t)$$

for all $t \in \mathbb{R}_+$, $x \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$ where $f(t,x,w) = A(t)x + B(t)w + C(t)u(t)$, $y(t) = \begin{bmatrix} x^\top(t) & u^\top(t) & w^\top(t) \end{bmatrix}^\top$,

Then, the 0-superlevel set $\mathcal{P}(t)$ of $x \to p(t,x)$, for a given $t \geq 0$, defined by

$$\mathcal{P}(t) = \{x \in \mathbb{R}^n \mid p(t,x) \geq 0\},$$

overapproximates the set of reachable states of $\mathcal{S}$, i.e.

$$\mathcal{R}(t; X_0) \subseteq \mathcal{P}(t), \text{ for any } t \geq 0.$$

Among the trajectories of the system, some might belongs to the boundary $\partial\mathcal{P}(t)$ of an overapproximation $\mathcal{P}$ at any time $t \geq 0$. Such trajectories are called *touching trajectories*.

Definition 1.3. Touching trajectory
A trajectory x of the system $\mathcal{S}$ is a touching trajectory of $\mathcal{P}$ if $x(t)$ belongs to the surface of $\mathcal{P}(t)$ at every time $t \in I$, i.e. $x(t) \in \partial\mathcal{P}(t)$.

Since a touching trajectory is a trajectory of the system, it belongs as well to the set of reachable states. Therefore, the overapproximation $\mathcal{P}$ locally touches the set of reachable states.

1.3 Conic sets

The next sections apply Proposition 1.1 to time-varying conic sets.

Definition 1.4. Conic set
Let $P = P^\top \in \mathbb{R}^{(n+1)\times(n+1)}$ be the coefficient of the quadratic form over $\mathbb{R}^n$

$$\begin{aligned} p : \mathbb{R}^n &\mapsto \mathbb{R} \\ x &\to \begin{bmatrix} x \\ 1 \end{bmatrix}^\top P \begin{bmatrix} x \\ 1 \end{bmatrix}. \end{aligned}$$

We define the conic set as

$$\mathcal{P} = \{x \in \mathbb{R}^n \mid p(x) \geq 0\}.$$

Let $\mathbb{P}$ be the set of conic sets of $\mathbb{R}^n$ and let `Conic` $: \mathbb{S}^{(n+1)\times(n+1)} \mapsto \mathbb{P}$ *be the function that associates to any quadratic coefficient $P \in \mathbb{S}^{(n+1)\times(n+1)}$ the conic set $\mathcal{P} =$* `Conic`(P).

Let a block decomposition of P be

$$P = \begin{bmatrix} E & f \\ f^\top & g \end{bmatrix}. \tag{1.9}$$

$\mathcal{P}$ is a conic subset of $\mathbb{R}^n$ centered around $x_c = -E^{-1}f$. E corresponds to the (signed) curvature of $\mathcal{P}$. When E is not negative definite, i.e. $E \not\prec 0$, $\mathcal{P}$ is unbounded, otherwise $\mathcal{P}$ is bounded (see Figure 1.2). When $E \prec 0$, $\mathcal{P}$ is a (bounded) ellipsoidal subset of $\mathbb{R}^n$. In this case, $\mathcal{P} \neq \emptyset$ iff $g - f^\top E^{-1} f \geq 0$ (i.e. x_c belongs to $\mathcal{P}$). In such a case, the ellipsoidal set can be equivalently described by the relation

$$x \in \mathcal{P} \Leftrightarrow (x - x_c)^\top Q^{-1}(x - x_c) \leq 1$$

where $Q = (g - f^\top E f)(-E)^{-1}$.

The volume of $\mathcal{P}$ is then equal to

$$\mathrm{Vol}(P) = \frac{\pi^{n/2}\sqrt{(g - f^\top E f)det(-E)^{-1}}}{\Gamma(n/2+1)}$$

where Γ is the gamma function.

In this report, the volume is only used as a minimizing criterion within sets of ellipsoids of fixed dimensions. Most of the time the constant in n is being neglected and since we only want to minimize the volume, we use the *pseudo volume* that we define by

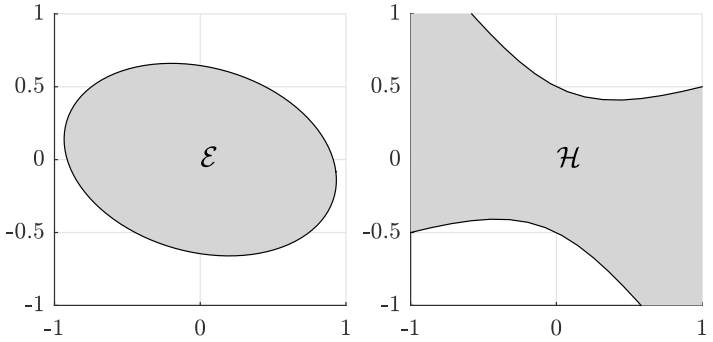
$$\widetilde{\mathrm{Vol}}(P) = (g - f^\top E f)det(-E)^{-1}. \tag{1.10}$$

Let $\mathrm{TrSq} : \mathbb{S}^{(n+1)\times(n+1)} \to \mathbb{R}$ be the map that associates to an ellipsoid $\mathcal{P}$ parametrized by $P \in \mathbb{S}^{(n+1)\times(n+1)}$ the squared sum of its semi-axes:

$$\mathrm{TrSq}(P) = (g - f^\top E f)trace(-E^{-1}). \tag{1.11}$$

Remark 1.2. Representation of the conic set

Conic sets could be represented differently, i.e. their center and the curvature arguments. However, we observed that alternative representations are less convenient. The time-varying conic sets are not bounded in the general case and their center is not a continuous function of time. Associated ordinary equations are most of the time difficult to manipulate and analyze.



(a) *Ellipse* ***(b)*** *Hyperboloid*

Figure 1.2: *When E (see the block decomposition of P in (1.9)) is negative definite (subfigure 1.2a) and $\mathcal{P} \neq \emptyset$, $\mathcal{P}$ is an ellipsoidal set, to this respect $\mathcal{P}$ is a bounded and convex set. When E is not sign-definite (subfigure 1.2b), $\mathcal{P}$ and $\mathcal{P} \neq \emptyset$, the $\mathcal{P}$ is unbounded.*

1.4 Overapproximation with time-varying conics

This section applies Proposition 1.1 to overapproximate $\mathcal{R}(t; X_0)$ with time-varying conic sets (defined in Section 1.3). The coefficients of the time-varying conics are expressed as the solution of a DRE (Differential Riccati Equation) that is parametrized by the positive multipliers λ and μ introduced in Section 1.2. When generating all the time-varying conics for all the possible parameters (λ, μ), one obtains a family of time-dependent overapproximations of $\mathcal{R}(t; X_0)$, we define their intersection that is a tighter overapproximation. The next section will further analyze the DRE to determine the domain of definitions of these overapproximations.

Let the time-dependent conic $\mathcal{P} : \mathbb{R}_+ \to \mathbb{P}$ associated with its value function $p : \mathbb{R}_+ \times \mathbb{R}^n \mapsto \mathbb{R}$ defined by

$$p(t, x) = \begin{bmatrix} x \\ 1 \end{bmatrix}^\top P(t) \begin{bmatrix} x \\ 1 \end{bmatrix}$$

where $P : \mathbb{R}_+ \mapsto \mathbb{S}^{n \times n}$ the time-varying coefficient associated with the quadratic form p.

We will use Proposition 1.1 to find sufficient conditions over $\dot{P}$ such that $\mathcal{P}$ overapproximates the set of reachable states $\mathcal{R}(t; \mathcal{P}_0)$. Let q be the overapproximation of the variation of $p(t, x(t)) - \lambda(t)\langle y, y\rangle_{M,\mu_t}$ along a system trajectory $y = (x, u, w) \in \mathcal{S}$ as defined in Property 1.1

$$q(t, x, w) = \frac{\partial p}{\partial t} + \frac{\partial p}{\partial x}^\top f(t, x, w) - \dot{\lambda}(t)\lambda(t)^{-1} p(t, x(t)) - \lambda(t)\mu(t) y^\top M(t) y. \quad (1.12)$$

Contrary to $p(t,x(t)) - \lambda(t)\langle y,y\rangle_{M,\mu_t}$ which depends on trajectories, q only depends on states at the given time t. Moreover, q is a quadratic function of $\begin{bmatrix} x \\ 1 \\ w \end{bmatrix}$

$$\begin{aligned} q(t,x,w) = &\begin{bmatrix} x \\ 1 \end{bmatrix}^\top \dot{P}(t) \begin{bmatrix} x \\ 1 \end{bmatrix} + 2 \begin{bmatrix} x \\ 1 \end{bmatrix}^\top P(t) \begin{bmatrix} A(t)x + B(t)w + C(t)u(t) \\ 0 \end{bmatrix} \\ &- \dot{\lambda}(t)\lambda(t)^{-1} \begin{bmatrix} x \\ 1 \end{bmatrix}^\top P(t) \begin{bmatrix} x \\ 1 \end{bmatrix} - \lambda(t)\mu(t) \begin{bmatrix} x \\ u(t) \\ w \end{bmatrix}^\top M(t) \begin{bmatrix} x \\ u(t) \\ w \end{bmatrix}. \end{aligned} \tag{1.13}$$

Let us fix a pair $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^n$. To prove the positivity of $h : w \mapsto q(t,x,w)$, we prove that the minimum of h is positive. For any $t \geq 0$, $M_w(t) \leq 0$, therefore the minimum of the quadratic function $w \mapsto q(t,x,w)$ exists. The minimum is reached at w^*,

$$\begin{aligned} w^*(t,x) &= \arg\min_{w \in \mathbb{R}^p} q(t,x,w) \\ &= \lambda(t)^{-1}\mu(t)^{-1} M_w(t)^{-1} \left(B_1^\top(t) P(t) - \lambda(t)\mu(t) M_{w,x1}(t) \right) \begin{bmatrix} x \\ 1 \end{bmatrix}. \end{aligned} \tag{1.14}$$

Let $q^*(t,x) = q(t,x,w^*(t,x))$, we can rewrite q^* as

$$q^*(t,x) = \begin{bmatrix} x \\ 1 \end{bmatrix}^\top Q(t) \begin{bmatrix} x \\ 1 \end{bmatrix} \tag{1.15}$$

where

$$\begin{aligned} Q(t) = &\dot{P}(t) + P(t)A_1(t) + A_1^\top(t)P(t) - \dot{\lambda}(t)\,\lambda(t)^{-1}\,P(t) - \lambda(t)\,\mu(t)\,M_{x1}(t) \\ &+ \lambda(t)^{-1}\,\mu(t)^{-1} \left(B_1^\top(t)P(t) - \lambda(t)\mu(t)M_{w,x1}(t) \right)^\top M_w^{-1}(t) \left(B_1^\top(t)P(t) - \lambda(t)\mu(t)M_{w,x1}(t) \right) \end{aligned}$$

with the following matrices

$$A_1(t) = \begin{bmatrix} A(t) & C(t)u(t) \\ 0 & 0 \end{bmatrix} \text{ and } B_1(t) = \begin{bmatrix} B(t) \\ 0 \end{bmatrix} \tag{1.16}$$

and the projection and block decomposition of M into the basis $\begin{bmatrix} x^\top & 1 & w^\top \end{bmatrix}^\top$ is defined in (1.3).

When $Q(t) = 0$ for all $t \geq 0$, Proposition 1.1 holds.

Definition 1.5. Time-varying conic

For a given $\mu \in \mathcal{D}^*$, λ a continuous, differentiable function strictly positive and

increasing over $\mathbb{R}_+$*, a given initial condition* $P_0 \in \mathbb{S}^{(n+1)\times(n+1)}$*, let* $P : [0,T] \mapsto \mathbb{S}^{(n+1)\times(n+1)}$ *be the solution of the initial value problem (time dependences are omitted for readability)*

$$\begin{aligned} 0 =& \dot{P} + PA_1 + A_1^\top P - \dot{\lambda}\,\lambda^{-1}\,P - \nu\,M_{x1} \\ &+ \nu^{-1}\,\left(B_1^\top P - \nu M_{w,x1}\right)^\top M_w^{-1}\left(B_1^\top P - \nu M_{w,x1}\right) \end{aligned} \tag{1.17}$$

where $\nu(\cdot) = \lambda(\cdot)\,\mu(\cdot)$ *and* $\gamma(\cdot) = \dot{\lambda}(\cdot)\,\lambda^{-1}(\cdot)$*, with initial condition*

$$P(0) = P_0.$$

The ordinary differential equation (1.17) is a Differential Riccati Equation (DRE). Since all the parameters of (1.17) are measurable, the solution P to (1.17) exists locally and is unique. The convergence properties and continuity of the solution to the DRE (1.17) are studied in [Kučera, 1973]. Depending on the initial condition P_0 and on the parameters of the system, the solution P might diverge in finite-time. In such a case, the corresponding overapproximation $\mathcal{P}(\cdot)$ is only defined over a finite time-horizon as well. In Section 1.5, we will show that an appropriate choice of λ and ν allows to have a solution P to DRE (1.17) defined over any time-horizon.

We can then define the time-varying conic set as

$$\mathcal{P}(t) = \texttt{Conic}(P(t)). \tag{1.18}$$

By Proposition 1.1, since along any trajectory $x \in \mathcal{S}$, $t \to p(t,x)$ is an increasing function of time, the following property holds

Theorem 1.1. Overapproximation with a time-varying conic
The set of reachable states $\mathcal{R}(t;\mathcal{P}_0)$ *of* $\mathcal{S}$ *is overapproximated at any time instant* $t \in \mathbb{R}_+$ *and for any conic set of initial of states* $\mathcal{P}_0 \in \mathbb{P}$

$$\mathcal{R}(t;\mathcal{P}_0) \subseteq \mathcal{P}(t)$$

where $\mathcal{P}$ *is the time-varying paraboloid defined in* (1.18) *with the time-varying coefficient* P *solution of the DRE* (1.17) *of Definition 1.5,* $A_1(\cdot)$ *and* $B_1(\cdot)$ *are defined in* (1.16)*.*

The ODE (1.17) depends on the multipliers $\lambda \in \mathscr{C}^1(\mathbb{R}_+;\mathbb{R}_+^*)$ and $\mu \in \mathcal{D}^*$. For each multiplier, Theorem 1.1 ensures that the corresponding $\mathcal{P}(t;\lambda,\mu)$ is a valid overapproximation of the reachable set $\mathcal{R}(t;\mathcal{P}_0)$. We define the time-varying set Π that associates to any time-instant $t > 0$ a subset $\Pi(t) \subseteq \mathbb{R}^n$

$$\Pi(t) = \bigcap_{\substack{\lambda \in \mathscr{C}^1(\mathbb{R}_+;\mathbb{R}_+^*) \\ \mu \in \mathcal{D}^*}} \mathcal{P}_{\lambda,\mu}(t) \tag{1.19}$$

as well as the set time-varying paraboloids:

$$\Pi^* = \{\mathcal{P}_{\lambda,\mu}(t) | \lambda \in \mathscr{C}^1(\mathbb{R}_+; \mathbb{R}_+^*), \mu \in \mathcal{D}^*\}. \tag{1.20}$$

The following theorem is a direct consequence of Theorem 1.1.

Theorem 1.2. Intersection of overapproximations
The set of reachable states $\mathcal{R}(t; \mathcal{P}_0)$ of $\mathcal{S}$ is overapproximated by the intersection of time-varying paraboloids $\Pi(t)$ generated by the sets of multipliers $(\lambda, \mu) \in \mathscr{C}^1(\mathbb{R}_+; \mathbb{R}_+^*) \times \mathcal{D}^*$, i.e.

$$\mathcal{R}(t; \mathcal{P}_0) \subseteq \Pi(t),$$

for every $t \geq 0$.

Proof. Direct consequence of Theorem 1.1 and the definition (1.19) of Π. ♢

1.5 Domain of definition of the overapproximations

In the case of reachability analysis, the existence and boundedness of the overapproximation over a given time-horizon is of great interest. This part first expresses the block decomposition of $P(t)$. We explicit the ODEs satisfied by each block. This formulation shows the structure of the DRE (1.17) and is then used to study the domain of definition of the overapproximations $\mathcal{P}$ and therefore of their point-wise intersection Π.

1.5.1 Coefficient expansion of the DRE

The expression of (1.17) highlights the "Riccati formulation" of the ODE satisfied by P, it conveniently formulates the ODE in a one-line equation. In this part, we give an alternative useful form of (1.17) that will help to characterize the domain of definition of the overapproximations $\mathcal{P}$.

Let $P(\cdot)$ be a solution of (1.17) for a given initial condition $P_0 \in \mathbb{P}$. The associated value function $p(t, x) = \begin{bmatrix} x \\ 1 \end{bmatrix}^\top P(t) \begin{bmatrix} x \\ 1 \end{bmatrix}$ is a quadratic function of $\mathbb{R}^n$. Let $E : \mathbb{R}_+ \mapsto \mathbb{S}^{n \times n}$ be the time-varying quadratic coefficient of p, let $f : \mathbb{R}_+ \mapsto \mathbb{R}^n$ be the affine coefficient and let $g : \mathbb{R}_+ \mapsto \mathbb{R}$ be the constant coefficient. Namely,

$$P(t) = \begin{bmatrix} E(t) & f(t) \\ f^\top(t) & g(t) \end{bmatrix}.$$

In this section, we expand (1.17) to express the ODE satisfied by E, f, and g. Expressions (1.21), (1.22), and (1.23) are more verbose than (1.17) but exhibit a more meaningful structure.

Using (1.17), E satisfies DRE

$$\begin{aligned} 0 =& \dot{E}(t) + E(t)A(t) + A^\top(t)E(t) - \dot{\lambda}(t)\lambda(t)^{-1}E(t) - \lambda(t)\mu(t)M_x(t) \\ &+ \lambda(t)^{-1}\mu(t)^{-1}(B^\top(t)E(t) + M_{x,w}(t))^\top M_w(t)^{-1}(B^\top(t)E(t) + M_{x,w}(t)). \end{aligned} \tag{1.21}$$

E is independent of f and g, therefore the quadratic coefficient of p is independent of the center and the offset of $\mathcal{P}_0$. f satisfies the linear time-varying differential equation

$$\begin{aligned} 0 =& \dot{f}(t) + A^\top(t)f(t) - \dot{\lambda}(t)\lambda(t)^{-1}f(t) - (M_{x,u}(t) + E(t)C(t))u(t) \\ &+ (E(t)B(t) + M_{x,w})M_w(t)^{-1}(B^\top(t)f(t) - M_{u,w}^\top u(t)). \end{aligned} \tag{1.22}$$

When E is measurable over $\mathbb{R}_+$, there exists a solution to (1.22) over $\mathbb{R}_+$ which is bounded over any interval $[0, T]$, $T > 0$. g satisfies the linear time-varying differential equation

$$0 = \dot{g}(t) - \dot{\lambda}(t)\lambda(t)^{-1}g(t) + \begin{bmatrix} f(t) \\ u(t) \end{bmatrix}^\top G(t) \begin{bmatrix} f(t) \\ u(t) \end{bmatrix} \tag{1.23}$$

where

$$G(t) = \begin{bmatrix} B(t)M_w(t)^{-1}B^\top(t) & C(t) - B(t)M_w(t)^{-1}M_{u,w}(t) \\ (C(t) - B(t)M_w^{-1}M_{u,w}^\top(t))^\top & -M_u(t) + M_{u,w}(t)M_w^{-1}(t)M_{u,w}^\top(t) \end{bmatrix}.$$

Similarly than for f, when E is measurable over $\mathbb{R}_+$, there exists a solution f to (1.23) over $\mathbb{R}_+$, this solution is bounded over any interval $[0, T]$, $T > 0$.

The domain of definition of P is therefore only dependent over the domain of definition of (1.21). The differential equation (1.21) has been well studied in control. In particular, it is known that solutions of (1.21) might diverge in finite-time (independently of the regularity of the coefficient). The next section shows that under some hypothesis over $\mathcal{D}^*$, we can show that Π is bounded at any time $t \in \mathbb{R}_+$.

1.5.2 Domain of definition of the time-varying ellipsoids

Since the domain of definition of P is the domain of definition of its quadratic parameter E solution of (1.21), we study the domain of definition of E. Let the dynamical function of E in (1.21) be $\text{Ricc} : \mathbb{S}^{n\times n} \times \mathbb{R}_+ \times \mathbb{R}_+$ such that (1.21) is equivalently formulated by

$$\dot{E} = \text{Ricc}(E(t), \lambda(t), \mu(t)).$$

A sufficient condition for E to exist over $\mathbb{R}_+$ is that $E_t \mapsto \text{Ricc}(E_t, \lambda(t), \mu(t))$ is Lipschitz over $\mathbb{S}^{n\times n}$ for every $t \geq 0$ (see the Cauchy-Lipschitz Theorem in Proposition 1 of

[Zeidler, 1995a]). This is not the case since Ricc is a quadratic function in E_t and thus there exists no constant that can bound the slope of this operator. However, if the solution E to (1.21) is bounded over $\mathbb{R}_+$, then $E_t \mapsto \text{Ricc}(E_t, \lambda(t), \mu(t))$ is Lipschitz over the bounded set of possible solutions for every $t \geq 0$.

Assumption 1.1. Well-posed disturbance set
There is a $K > 0$ such that for any $\kappa \geq K$, $\mu_\kappa \in \mathcal{D}^*$ where $\mu_\kappa(t) = exp(-\kappa t)$.

Such an assumption about $\mathcal{D}^*$ will be satisfied in the cases presented in Chapter 2 and Chapter 3. When Assumption 1.1 holds, for a $\kappa \geq K$ such that for the pair $(\lambda_\kappa, \mu_\kappa) \in \mathscr{C}^1(\mathbb{R}_+; \mathbb{R}_+^*) \times \mathcal{D}^*$, with $\lambda_\kappa(t) = exp(\kappa t)$, $\mu_\kappa(t) = exp(-\kappa t)$, it holds $\dot{\lambda}_\kappa(t)\, \lambda_\kappa(t)^{-1} = \kappa$ and $\lambda_\kappa(t)\mu_\kappa(t) = 1$ for any $t \geq 0$.

In this section, we show that, when Assumption 1.1 holds and when $E_0 \succ 0$, then there is a solution E to (1.21) that is upper-bounded $E(t) \preceq \overline{E}(t)$ (see Proposition 1.2) and lower-bounded $\underline{E}(t) \preceq E(t)$ (see Proposition 1.3) at every $t \geq 0$. These two bounds are sufficient to prove that E is bounded by the matrix norm $\|\cdot\|$ (see Proposition 1.4). We conclude that the solution E exists over $\mathbb{R}_+$ (Proposition 1.5).

To prove that E is upper-bounded, we use (1.21) and the fact that $M_w \prec 0$. By integration of (1.21) over the interval $[0, t]$, $t > 0$, it holds

$$\begin{aligned}\psi(t,0)^\top E(t)\psi(t,0) - E_0 = \\ \int_0^t \psi(t,\tau)^\top \left(-M_x + (B^\top E(\tau) + M_{xw}^\top) M_w^{-1} (B^\top E(\tau) + M_{xw}^\top)\right) \psi(t,\tau) d\tau\end{aligned} \tag{1.24}$$

where ψ is the transition matrix of $t \to A(t) - \frac{\kappa}{2} I$. Since $M_w \prec 0$,

$$\psi(t,0)^\top E(t)\psi(t,0) \preceq E_0 - \int_0^t \psi(t,\tau) M_x \psi(t,\tau) d\tau.$$

The transition matrix ψ is invertible over $[0, T]$, therefore the following property holds

Proposition 1.2. ***E*'s upper-bound**
For any $t \geq 0$, $E(t) \preceq \overline{E}(t)$ where

$$\overline{E}(t) = (\psi(t,0)^{-1})^\top \left(E_0 - \int_0^t \psi(t,\tau) M_x \psi(t,\tau) d\tau\right) \psi(t,0)^{-1}.$$

To prove that E is lower-bounded, we show that for some $\kappa \geq K$ large enough, E_0 is a lower-bound to E.

For any $E_0 \succ 0$, since A and B are bounded at any time (see Definition 1.1), there is a $\kappa \geq K$ large enough s.t.

$$-E_0 A(t) - A(t)^\top E_0 - M_x + (B(t)^\top E_0 + M_{xw}^\top) M_w^{-1} (B(t)^\top E_0 + M_{xw}^\top) + \kappa E_0 \succeq 0 \tag{1.25}$$

for every $t \geq 0$ Let $\Delta(t) = E(t) - E_0$. By definition of E, $\Delta(0) = 0$. Using 1.25, Δ satisfies the following ordinary differential inequality

$$-E_0A(t)-A(t)^\top E_0-M_x+(B(t)^\top E_0+M_{xw}^\top)^\top M_w^{-1}(B(t)^\top E_0+M_{xw}^\top)+\kappa E_0 \succeq \dot{E}(t)-\dot{\Delta}(t).$$

Using the definition of E and the one of Δ, it holds

$$\begin{aligned}\dot{\Delta}(t) \succeq &- \Delta(t)A(t) - A(t)^\top\Delta(t) + (B(t)^\top E(t) + M_{xw}^\top)^\top M_w^{-1}(B(t)^\top E(t) + M_{xw}^\top) \\ &- (B(t)^\top E_0 + M_{xw}^\top)^\top M_w^{-1}(B(t)^\top E_0 + M_{xw}^\top)\end{aligned} \tag{1.26}$$

We use the following expansion in (1.26)

$$Y^\top RY - X^\top RX = -\Gamma^\top RY - Y^\top R\Gamma - \Gamma^\top R\Gamma$$

with $\Gamma = X - Y$, $X = B^\top E_0 + M_{xw}^\top$, $Y = B^\top E + M_{xw}^\top$, and $R = M_w^{-1}$. Then, (1.26) gives

$$\dot{\Delta}(t) \succeq -\Delta(t)\tilde{A}(t) - \tilde{A}(t)^\top\Delta(t) - \Delta(t)^\top B(t)M_w^{-1}B(t)^\top\Delta(t)$$

where $\tilde{A}(t) = A(t) - (B(t)^\top E(t) + M_{xw}^\top)M_w^{-1}$ for every $t \geq 0$. Since $M_w \prec 0$, it holds

$$\dot{\Delta}(t) + \Delta(t)\tilde{A}(t) + \tilde{A}(t)^\top\Delta(t) \succeq 0$$

By integration over $[0, t]$, it holds

$$\tilde{\psi}(t,0)^\top\Delta(t)\tilde{\psi}(t,0) \succeq 0$$

where $\tilde{\psi}(t,0)$ is the transition matrix of $\tilde{A}(\cdot)$ from 0 to t. Since $\tilde{\psi}(t,0)$ is invertible, $\Delta(t) \succeq 0$, i.e. $E(t) \succeq E_0$.

Proposition 1.3. ***E*'s lower-bound**
There exists a $\kappa \geq K$ such that the solution E to (1.21) with $\lambda(t) = exp(\kappa t)$ and $\mu(t) = exp(-\kappa t)$ is lower-bounded by E_0, i.e. for every $t \geq 0$, $E(t) \succeq \underline{E}(t)$ where $\underline{E}(t) = E_0$.

Properties 1.2 and 1.3 provide an upper and lower bound for the positive semidefinite matrix order. Proposition 1.4 shows that it is a sufficient condition for E to be bounded by the $\|\cdot\|$ norm.

Proposition 1.4. Boundedness of an upper and lower bounded matrix
Let $A, B, C \in \mathbb{S}^{n\times n}$, if $A \preceq B \preceq C$ then $\|B - \tilde{B}\| \leq \|A - C\|$ where $\tilde{B} = \frac{A+C}{2}$.

Proof. $B - \tilde{B} = \frac{B-A}{2} + \frac{B-C}{2}$. The inequality satisfied by A, B, and C gives $0 \preceq B - A \preceq C - A$ and $0 \preceq C - B \preceq C - A$. Since for any $X, Y \in \mathbb{S}^{n\times n}$, $0 \preceq X \preceq Y$ implies that $\|X\| \leq \|Y\|$, we have

$$\begin{aligned}\|B - \tilde{B}\| &= \left\| \frac{B-A}{2} + \frac{B-C}{2} \right\| \\ &\leq \frac{\|B-A\|}{2} + \frac{\|B-C\|}{2} \\ &\leq \frac{\|C-A\|}{2} + \frac{\|A-C\|}{2} \\ &\leq \|C-A\|\end{aligned}$$

♢

Using Properties 1.2, 1.3, and 1.4, when condition (1.25) holds, E is bounded over its domain of definition $[0, T]$. It implies that $E_t \mapsto \mathrm{Ricc}(E_t, \lambda_\kappa(t), \mu_\kappa(t))$ is Lipschitz over the set $\mathscr{E} = \bigcup_{t\in[0,T]} \left\{ E_t \in \mathbb{S}^{n\times n} \,|\, \|E_t - \tilde{E}(t)\| \leq \|\overline{E}(t) - \underline{E}(t)\| \right\}$, where $\tilde{E}(t) = \frac{\underline{E}(t)+\overline{E}(t)}{2}$ for time instants in $[0, T]$. By using a contradiction argument, if E diverges at a time instant $t_d \in \mathbb{R}_+$, then E is not continuous at $t_d \in [0, T]$. Since Ricc is Lipschitz over $\mathscr{E}$, this contradicts the Cauchy-Lipschitz Theorem (see Proposition 1 in [Zeidler, 1995a]) and thus E is defined over $\mathbb{R}_+$.

Proposition 1.5. Domain of definition of the solution to the DRE
When $E_0 \succ 0$ and when Assumption 1.1 holds, there is a pair of multipliers (λ, μ), $\lambda \geq 0$ and $\mu \in \mathcal{D}^*$ s.t. E is defined over $\mathbb{R}_+$.

1.6 Discussion

1.6.1 Related works

Linear systems subject to disturbances bounded by quadratic constraints have been studied in the verification of dynamical systems (as in [Chaudenson, 2013]), in guaranteed state estimation (as in [Bertsekas and Rhodes, 1971, Savkin and Petersen, 1995]) and in stability analysis (as in [Jönsson, 1996]). Reachability analysis of such systems has been derived within three different approaches: a set-based approach, an optimal control approach, and a level-set approach. Each method leads to fundamentally the same result which is a time-varying ellipsoid overapproximating the reachable tube and whose parameter is the solution to a Differential Riccati Equation.

In the *set-based* approach, the reachable set is expressed as operations (Minkowski sum and affine transformation) over the set of initial states and the disturbance set.

[Chernous'ko, 1999] studies the case of linear time-varying system subjects to a disturbance bounded by an ∞-norm constraint (i.e. $w^\top(t)R(t)w(t) \leq 1$ with $R(t) \succeq 0$). The set of trajectories is overapproximated by a time-varying ellipsoid. Its center and radius satisfy an ordinary differential equation similar to (1.17). These differential equations are obtained using a set-based reasoning. The reachable set is described with set operations involving the initial set and the set of disturbance. More precisely the reachable set is expressed as a Minkowski sum between the flow of the initial set and the flow of the disturbance set. The first set corresponds to the image of the initial state through the autonomous dynamic and is simply an affine transformation of the set of initial states. The second set describes the influence of the exogenous disturbance w. For small time-step increase, the reachable set can be soundly approximated with an ellipsoidal set whose center and radius evolve according to an ordinary differential equation. This differential equation is parametrized by a free positive time-varying signal.

The original ellipsoidal method was firstly derived in [Schweppe, 1973]. Since then, this set-based approach has been extended to different geometrical shapes (such as zonotopes [Girard, 2005]) where the set operations could be overapproximated. The differential equation satisfied by the radius of the time-varying ellipsoid is a Differential Riccati Equation (DRE). This DRE and its associated Continuous Algebraic Riccati Equation (CARE, the equilibrium solutions of the DRE) are crucial for the control community and have been extensively studied in many works. The reader can refer to [Bittanti et al., 1991] for an exhaustive survey.

Such an approach is difficult to reproduce when it comes to more complex disturbances as the one studied in this thesis.

The optimal control approach uses the following observation: let a real-valued function V defined over the state space be such that $V(x_0) \geq 0$ for any x_0 in the initial state, for any reachable state $x_t \in \mathbb{R}^n$, there exists a system trajectory $x(\cdot)$ such that $x(t) = x_t$ and $V(x(0)) \geq 0$. Then, (t, x_t) belongs to the reachable set iff

$$\max V(x(0)) \text{ for } x(\cdot) \text{ a system trajectory s.t. } x(t) = x_t$$

When the system is subject to bounded disturbances, the optimization problem is constrained.

Such a constrained optimization problem can be studied using the Hamilton-Jacobi-Bellman (HJB) equation. The HJB equation is a nonlinear partial differential equation (PDE). This PDE models the propagation of the cost function along the flow of the system. Solutions to the PDE are difficult to approximate in practice (its solution is often not smooth).

When the cost is quadratic, the constrained optimization problem is known as the constrained Linear Quadratic Regulator problem (LQR). [Matveev and Yakubovich,

1997] studied it using a Lagrangian relaxation. When the Lagrangian multipliers are independent of the state, i.e. are time-dependent functions, the relaxed optimization problem is the well-known LQR problem that has been extensively studied in the literature (see [Lee and Markus, 1969], Chapter 5). [Jönsson, 2002] applied this approach to study Integral Quadratic Constraint systems (that fall into the class of system of interest of this thesis). Lagrangian multipliers are here constant weights.

The *level-based* approach represents an overapproximation of the reachable tube as the superlevel-set of a real-valued function defined over the time and space state. The overapproximation relationship is obtained by choosing a level-set function positive over the initial set and of increasing value along each system trajectory (see Section 1.2). Contrary to the optimal control approach, the level-set function does not solve any optimization problem.

In [Yin et al., 2020], the authors overapproximate the reachable set of a nonlinear system subject to IQC constraints using polynomial level sets.

[Seiler et al., 2019] shows that the solution to the DRE is overapproximated by solutions to a Differential Riccati Inequality (DRI). This DRI can be equivalently expressed as Differential Linear Matrix Inequality (DLMI) by using the Schur complement. Then, the authors find a solution that minimizes the input-output gain. Since the solution to the DLMI is a time-dependent matrix, the optimization problem has decision variables and constraints that belong to a signal space (of infinite dimension). To solve this optimization problem in practice, the authors express the signals in a finite-dimensional signal basis (with splines) and defined a time-sampled version of the DLMI constraint. The resulting optimization problem can be solved with semidefinite programming solvers (such as [Sturm, 1999]). The solution to the DRE can then be chosen by successively finding an optimal multiplier value with the DLMI and the DRE to minimize the input-output gain.

In previously cited works, two representation of the ellipsoidal sets are used:

- the *centered representation* where the ellipsoidal set $\mathcal{E}$ is characterized by their radius (or their curvature, i.e. the inverse of the radius), i.e. $\mathcal{E} = \{x \in \mathbb{R}^n \mid (x - c)^\top R^{-1}(x - c) \leq 1\}$;

- the *homogeneous coordinates representation* where the ellipsoidal set is expressed as the superlevel set of a quadratic form $y \to y^\top P y$, $P \in \mathbb{S}^{n+1}$ in homogeneous coordinates $y = \begin{bmatrix} x \\ 1 \end{bmatrix}$, $\mathcal{E} = \{x \in \mathbb{R}^n \mid \begin{bmatrix} x \\ 1 \end{bmatrix}^\top P \begin{bmatrix} x \\ 1 \end{bmatrix} \geq 0$ in coordinates $y = \begin{bmatrix} x \\ 1 \end{bmatrix}$.

The center-radius representation is frequently used in a set-based approach (as in [Chernous'ko, 1999, Kurzhanski and Varaiya, 2014]) whereas the center-curvature is more frequent in optimal control and level set approaches (as in [Seiler et al., 2019]). The homogeneous coordinate representation is less used in the literature (see [Savkin and Petersen, 1995, Savkin and Petersen, 1996a]).

The ODE satisfied by the coefficient of the time-varying ellipsoid depends on the chosen representation. When the disturbance set is centered, the center and radius of the ellipsoid are independent from each other (i.e. the center x_c satisfies an ODE $\dot{x}_c = f_x(t, x_c)$ independent of the radius, and the radius Q satisfies another ODE $\dot{Q} = f_Q(t, Q)$ independent from the center x_c). Such representation is therefore convenient for centered disturbances. However, in the general case (where disturbances are not necessarily centered), the center and radius ODEs are coupled. In our work, we found the homogeneous coordinate representation to lead to well-defined ODEs compared to the centered representation. Also, in this representation, the ODE satisfied by the coefficient of time-varying ellipsoid has an elegant form which is a DRE.

1.6.2 Conservatism

In this chapter, the conservatism of the ellipsoidal method is not addressed. It is studied in the two following chapters. Here, we give few hints about the different sources of conservatism and their impact over the overapproximation Π.

Incomplete dual description Few sources of uncertainties due to an incomplete dual space characterization can be identified. These incomplete dual only have a consequence of more conservatism. Since we are only interested in overapproximations, it does not compromise our analysis.

The disturbance constraint is taken into account in (1.6) by using a positive multiplier. A more general approach would choose the multiplier in the dual set of functions. And therefore, λ would be a strictly positive function of (t, x, w). Instead, (1.6) uses a multiplier λ chosen in a set of time-dependent functions. This approach allows us to derive the equations in the case of conic overapproximation case but induces some conservatism. Typically, the impact of disturbances correlated to the state is neglected.

The characterization of $\mathcal{D}^*$ is provided to the model. In practice, disturbances are rarely described by their dual. Different disturbances will be addressed in the case of ∞-norm constraints (Chapter 2) and 2-norm constraints (Chapter 3). In practice, an underapproximation $\tilde{\mathcal{D}}^*$ of $\mathcal{D}^*$ produces an overapproximation $\tilde{\mathcal{D}}$ of $\tilde{\mathcal{D}}$. Therefore, the reachable set $\mathcal{R}(t; X_0)$ of $\mathcal{S}$ with the set of disturbances $\mathcal{D}$ is overapproximated by the reachable set $\tilde{\mathcal{R}}(t; X_0)$ of $\tilde{\mathcal{S}}$ with the set of disturbances $\tilde{\mathcal{D}}$. Overapproximation

$\tilde{\mathcal{P}}(t)$ computed with Theorem 1.1 or $\tilde{\Pi}(t)$ of $\tilde{\mathcal{R}}(t; X_0)$ are valid overapproximation of $\mathcal{R}(t; X_0)$.

Conic overapproximation For some specific $\mathcal{D}^*$, the reachable set $\mathcal{R}(t; \mathcal{P}_0)$ is not a pure conic set when $\mathcal{P}_0 \in \mathbb{P}$ (see Chapter 2). Several facts motivate the choice of conic sets to describe the set of reachable states:

- when the set of disturbances is trivial (singleton set, as at the beginning of Section 1.2), the set of reachable states is exactly described by a time-varying ellipsoidal set;
- any set can be described as a (possibly non-finite, uncountable) intersection of conic sets;
- computing the minimum (and therefore proving the positivity) of a quadratic function can be done analytically;

Contrary to many works in reachability analysis, we use the general set of conics (instead of the subset of ellipsoidal sets). It allows to describe non-convex reachable sets and hopefully reduce the conservatism of the global approach.

1.6.3 Conclusion

This chapter proposed to overapproximate the reachable set of a linear time-varying system with time-varying conic sets. The overapproximation relationship is obtained with a "Lyapunov" approach: the value function associated with the time-varying conic is chosen to be increasing along every trajectory and positive over the set of initial states. The coefficient (in homogeneous coordinates) of the time-varying conic satisfies a parametrized Differential Riccati Equation (DRE). When some assumption about the system holds, these parameters can be chosen such that the DRE has a solution (i.e. the overapproximation exists) over the entire time-domain. Such a result is obtained by finding a lower and an upper bound to the DRE.

The next two chapters apply Theorem 1.1 and Theorem 1.2 to two subclasses of the system defined in (1.1). Chapter 2 studies linear systems subject to Quadratic Constraint (QC systems), i.e. systems where the state-input-disturbance signal $y(t) = (x(t), u(t), w(t))$ satisfies the quadratic constraint $y^\top(t)M(t)y(t) \geq 0$ at any time $t \geq 0$. Chapter 3 studies linear systems subject to Integral Quadratic Constraint (IQC), i.e. $\int_0^t y^\top(s)M(s)y(s)ds \geq 0$ at each $t \geq 0$.

Future works The DRE of interest satisfies many interesting properties one of them being that when $t \mapsto E(t)$ solves the DRE (1.21), $P' : t \mapsto k(t)P(t)$ solves the same DRE (1.21) provided some conditions over k and the parameters of the system hold, in other words, the DRE is separable. Such a result provides a lot of information on how the reachable tube of such systems behaves depending on the initial set of states. On another note, the DRE 1.21 is closely related to its associated Continuous Algebraic Riccati Equation (CARE; the CARE corresponds to constant solutions to DRE 1.21 with a free initial condition). The CARE appears in the Kalman-Yakubovitch-Popov Lemma when studying the robust stability of uncertain systems. In this case, the quantity $\dot{\lambda}\lambda^{-1}$ is related to the input-to-output $\mathscr{L}_2$ gain. A clear understanding of this connection would help to choose an appropriate value for λ.

Chapter 2

Quadratic Constraints

Contents

In Chapter 1, we formalized the general framework of reachability analysis by means of the ellipsoidal method. This chapter applies this framework to a specific subclass of disturbances: point-wise constrained disturbances. Such a model frequently appears for verification purposes since local linearization and overapproximation of the system dynamic provides such an abstraction.

This chapter is organized as follows. The QC system is presented in Section 2.1, Section 2.2 provides motivating examples and computes their reachable set. Section 2.3 applies Theorem 1.1, Chapter 1, to the QC system. Section 2.4 identifies touching trajectories and their associated support ellipsoid. Section 2.5 studies the problem of finding optimal time-varying ellipsoids (e.g. that minimize the end volume

of the overapproximation), Section 2.5.2 provides a continuation method to find this optimal overapproximation. Section 2.6 presents several study cases. This chapter ends with a discussion in Section 2.7 and a conclusion in Section 2.8.

2.1 Definition of the system

In this chapter, we study an LTV system subject to a disturbance that satisfies a point-wise quadratic constraint with the state and the input.

Definition 2.1. Quadratic constraint system

Let the system $\mathcal{S}$ be defined by

$$\mathcal{S}: \begin{cases} \dot{x} = A(t)x + B(t)w + C(t)u(t) \\ \begin{bmatrix} x(t) \\ u(t) \\ w(t) \end{bmatrix}^\top M(t) \begin{bmatrix} x(t) \\ u(t) \\ w(t) \end{bmatrix} \geq 0 \text{ for any } t \geq 0 \end{cases} \tag{2.1}$$

where M is a quadratic form that is negative in the disturbance dimension and s.t.

$$M_{x1\backslash w}(t) = M_{x1}(t) - M_{x1,w}(t)M_w(t)^{-1}M_{x1,w}^\top(t) \succ 0 \tag{2.2}$$

for any $t \geq 0$ (with the representation of M and u defined in (1.3), $M_{x1\backslash w}$ is the Schur complement of the block decomposition of M).

The assumption (2.2) ensures that the disturbance set is not empty at any time. It guarantees that

$$\max_{w_t \in \mathbb{R}^m} \begin{bmatrix} x_t \\ u(t) \\ w_t \end{bmatrix}^\top M(t) \begin{bmatrix} x_t \\ u(t) \\ w_t \end{bmatrix} \geq 0$$

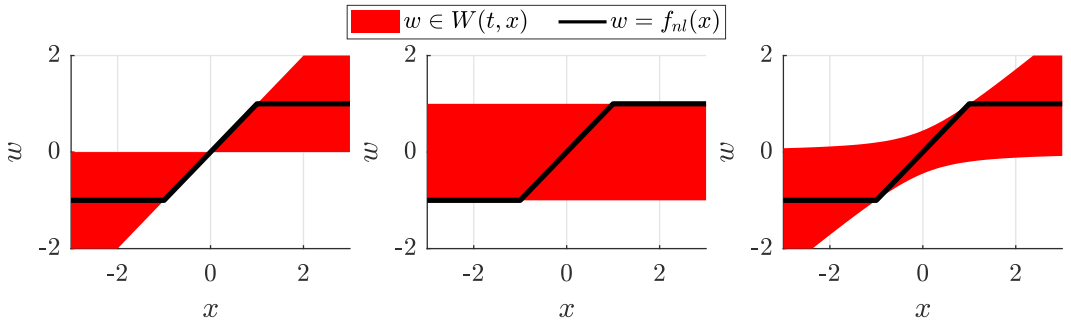
for any state $x_t \in \mathbb{R}^n$ at any time-instant $t \geq 0$.

Remark 2.1. Link with IQC systems

In the field of robust control, such inequalities have been analyzed with the framework of Integral Quadratic Constraints. Such inequality is known as static IQC, sector inequalities.

2.2 Motivation of the QC model

To motivate the use of the quadratic constrained model of Definition 2.1, Section 2.2.1 presents some use of QC systems as an abstraction for nonlinear systems. Section 2.2.2 computes the reachable set of a 1-dimensional linear time-invariant system for different QC constraints (i.e. for different M matrices in Definition 2.1).



(a) Pure conic set *(b) Cylinder set* *(c) Hyperbola set*

Figure 2.1: *Different QC abstractions for some nonlinearities in dynamical systems.*

2.2.1 Abstraction

At a given time, admissible states and disturbances belong to a conic set $\mathcal{D}_1(t) \subset \mathbb{R}^{n\times m}$ described by

$$(x_t, w_t) \in \mathcal{D}_1(t) \Leftrightarrow \begin{bmatrix} x_t \\ 1 \\ w_t \end{bmatrix}^\top M_{x1w}(t) \begin{bmatrix} x_t \\ 1 \\ w_t \end{bmatrix} \geq 0.$$

For a given x_t, $t \geq 0$, let $W(t, x_t) \subset \mathbb{R}^m$ be the set of admissible disturbances w_t at time instant t and state x_t. Using the QC constraint in Definition 2.1, $W(t, x_t)$ is equivalently described by

$$w_t \in W(t, x_t) \Leftrightarrow (w_t - r_t)^\top R_t (w_t - r_t) \leq 1$$

where $R_t = -\left(\begin{bmatrix} x_t \\ 1 \end{bmatrix}^\top M_{x1\backslash w}(t) \begin{bmatrix} x_t \\ 1 \end{bmatrix}\right)^{-1} M_w$ and $r_t = M_w^{-1} M_{x1,w}^\top \begin{bmatrix} x_t \\ 1 \end{bmatrix}$. Since $M_w \prec 0$ and $M_{x1\backslash w} \succ 0$, it holds $R_t \succ 0$. To this respect, $W(t, x_t)$ is an ellipsoidal set, bounded (since $R_t \succ 0$) and not empty (since $R_t \prec \infty$).

The following paragraphs show how the QC systems can be used as an abstraction to nonlinear systems.

Let the following system

$$\dot{x} = Ax + Bsat(c^\top x)$$

where *sat* is the saturation operator defined by

$$sat : \begin{cases} sat(y) = y & \text{when } -1 \leq y \leq 1 \\ sat(y) = sign(y) & \text{otherwise} \end{cases}$$

Let $w = sat(y)$ and $y = c^\top x$, for any $x \in \mathbb{R}^n$, it holds

$$w \geq 0 \Rightarrow w \leq y$$
$$w \leq 0 \Rightarrow w \geq y$$

It implies $(y - w)w \geq 0$. Which corresponds to

$$M_p = \begin{bmatrix} 0 & 0 & \frac{c}{2} \\ 0 & 0 & 0 \\ \frac{c^\top}{2} & 0 & -1 \end{bmatrix}.$$

The associated set of disturbance $\mathcal{D}_{1,p}(t)$ is a pure conic set (represented in Figure 2.1a).

For any $y \in \mathbb{R}$, it also holds that $w \leq 1$ and $w \geq -1$. Therefore, it holds $1 - w^2 \geq 0$ and

$$M_c = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

is a valid QC constraint for the system G_{sat}. The associated set of disturbance $\mathcal{D}_{1,c}(t)$ is a cylinder (represented in Figure 2.1b).

Any convex combinations of the QC constraint is a valid constraint as well, and therefore $M_\lambda = (1-\lambda)M_p + \lambda M_c$ for any $\lambda \in [0, 1]$ is a valid QC constraint for G_{sat} as well. For any given $\lambda \in [0, 1]$, the associated set $\mathcal{D}_{1,\lambda}$ is a hyperbolic set (represented in Figure 2.1c).

2.2.2 A simple example

In the case of a 1-dimensional system, the reachable set of $\mathcal{S}$ of Definition 2.1 can be easily characterized. The reachable set corresponds to a time-varying interval whose bounds are trajectories of $\mathcal{S}$. These extremal trajectories $\underline{x}$ and $\overline{x}$ are generated by disturbances that steer the state away from a center trajectory. Example 2.1 and Example 2.2 compute the reachable set for two different disturbance sets (i.e. two different QC constraints). Example 2.1 compute the reachable set for a sector inequality, Example 2.2 for an exogenous disturbance.

Example 2.1.

Let the following 1-dimensional LTI system parametrized by a $k > 0$

$$\mathcal{S}_k^1 : \begin{cases} \dot{x} = -x + w \\ x(0) \in [-1, 1] \\ w(t)^2 \leq kx(t)^2, \ \textit{for all } t \geq 0 \end{cases} \tag{2.3}$$

Extremal trajectories of $\mathcal{S}_k^1$ *are respectively associated with the disturbance*

$$\overline{w}(t) = \max_{|x_t| \leq x(t)} \sqrt{k} x_t, \text{ when } x \geq 0$$

and

$$\underline{w}(t) = \min_{|x_t| \leq x(t)} \sqrt{k} x_t, \text{ when } x \leq 0$$

Therefore, extremal trajectories $\overline{x}$ *and* $\underline{x}$ *satisfies*

$$\begin{aligned} \dot{\overline{x}} &= (\sqrt{k} - 1)\overline{x} \qquad \overline{x}(0) = 1, \\ \dot{\underline{x}} &= (\sqrt{k} - 1)\underline{x} \qquad \underline{x}(0) = -1. \end{aligned}$$

Every other disturbance $w \in [\underline{w}, \overline{w}]$ *steers the state in-between these two trajectories* $\underline{x}$ *and* $\overline{x}$*. The reachable set can be exactly derived*

$$\mathcal{R}_k^1(t) = [\underline{x}, \overline{x}].$$

We can then identify the following cases:

- *when* $\sqrt{k} - 1 > 0$*, trajectories* $\underline{x}$ *and* $\overline{x}$ *are exponentially unstable, the reachable set is unbounded when* $t \to \infty$ *and the system* $\mathcal{S}_k^1$ *is said unstable;*
- *when* $\sqrt{k} - 1 > 0$*, trajectories* $\underline{x}$ *and* $\overline{x}$ *are exponentially unstable, the reachable set is unbounded when* $t \to \infty$ *and the system* $\mathcal{S}_k^1$ *is said stable;*
- *when* $\sqrt{k} - 1 = 0$*, extremal trajectories satisfies* $\underline{x} = -1$ *and* $\overline{x} = 1$*, the reachable set is a constant tube* $\mathcal{R}(t) = [-1, 1]$*.*

Example 2.2.

Let the following 1-dimensional LTI system, parametrized by a $k > 0$*, be defined by*

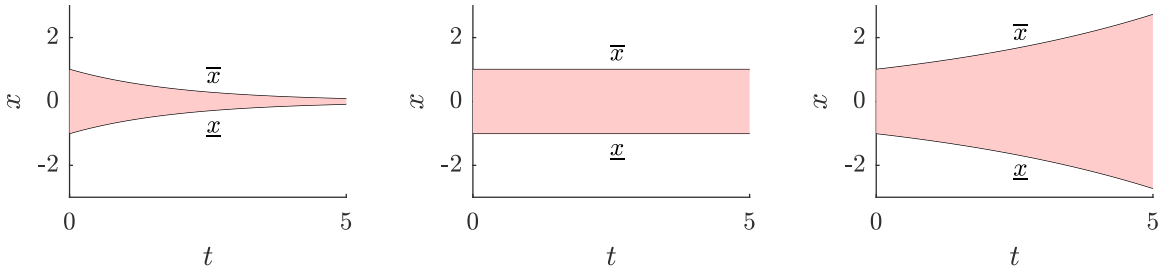
$$\mathcal{S}_k^2 : \begin{cases} \dot{x} = -x + w \\ x(0) \in [-1, 1] \\ w(t)^2 \leq k, \text{ for all } t \geq 0 \end{cases} \tag{2.4}$$

As in Example 2.1, extremal trajectories of $\mathcal{S}_k^2$ *corresponds to disturbances that steer the state away of* 0*. I.e. the disturbance is equal to*

$$\overline{w}(t) = \sqrt{k}, \text{ when } x \geq 0$$

and

$$\underline{w}(t) = -\sqrt{k}, \text{ when } x \leq 0$$



(a) $\mathcal{S}^1_{0.25}$ ***(b)*** $\mathcal{S}^1_1$ ***(c)*** $\mathcal{S}^1_{1.44}$

Figure 2.2: *Reachable set of dynamical system* $\mathcal{S}^1_k$*, defined in* (2.4)*, for different values of* k*. For* $k = 0.5$*,* $\mathcal{S}^1_{0.5}$ *is a stable system, the reachable set* $\mathcal{R}(t)$ *is a tube converging to* $\{0\}$ *when* $t \to \infty$*. For* $k = 1$*,* $\mathcal{S}^1_1$ *the reachable set* $\mathcal{R}(t)$ *is a constant tube equals to* $[-1, 1]$ *for all* $t \geq 0$*. For* $k = 1.2$*,* $\mathcal{S}^1_{0.5}$ *is an unstable system, the reachable set* $\mathcal{R}(t)$ *is a tube diverging when* $t \to \infty$*.*

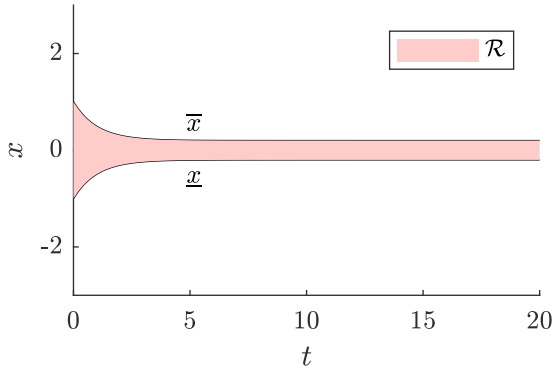


Figure 2.3: *Reachable set of dynamical system* $\mathcal{S}^2_{0.04}$ *defined in* (2.4)*. The reachable tube is not contracting when* $t \to \infty$*.*

Therefore, extremal trajectories $\overline{x}$ *and* $\underline{x}$ *satisfies*

$$\begin{aligned} \dot{\overline{x}} &= -\overline{x} + \sqrt{k}, \qquad \overline{x}(0) = 1, \\ \dot{\underline{x}} &= -\underline{x} - \sqrt{k}, \qquad \underline{x}(0) = -1. \end{aligned}$$

Every other disturbance $w \in [\underline{w}, \overline{w}]$ *steers the state in-between these two trajectories* $\underline{x}$ *and* $\overline{x}$*. The reachable set can be exactly derived*

$$\mathcal{R}^2_k(t) = [\underline{x}, \overline{x}].$$

2.3 Ellipsoidal method

In order to apply Theorem 1.1 of Chapter 1 for the QC system of Definition 2.1, we first identify the dual space $\mathcal{D}^*$ of the set of signals.

The set positive functions can be fully characterized as the subset of square-integrable signals $f \in \mathscr{L}_2(\mathbb{R}_+;\mathbb{R})$ such that the scalar product with any positive measurable square-integrable signal $\mu \in \mathscr{L}_2(\mathbb{R}_+;\mathbb{R})$ is positive.

Proposition 2.1. Positive signal duality
Any square-integrable measurable $f \in \mathscr{L}_2(\mathbb{R}_+;\mathbb{R})$ signal is positive over $\mathbb{R}_+$ iff

$$\langle f, \mu \rangle \geq 0$$

for any square-integrable measurable μ positive over $\mathbb{R}_+$.

Proof. ($\Rightarrow$) if $f(\cdot), \mu(\cdot) \geq 0$ over $\mathbb{R}_+$, then $\langle f, \mu \rangle \geq 0$ ($\Leftarrow$) if there is a $t \geq 0$ s.t. $f(t) \leq 0$, then since f is measurable, there is an interval $I \subset \mathbb{R}_+$ centered over t where $\langle f, 1 \rangle_I \leq 0$, therefore the property holds for $\mu = \delta_I$. $\diamondsuit$

Therefore, the system $\mathcal{S}$ of Definition 2.1 is strictly equivalent to the system of Definition 1.1 where $\mathcal{D}^*$ is the set of square-integrable functions, measurable and positive over $\mathbb{R}_+$. Therefore, the following Theorem 1.1 holds.

Corollary 2.1. Overapproximation of the reachable set, application of Theorem 1.1
The set of reachable states $\mathcal{R}(t;\mathcal{P}_0)$ of $\mathcal{S}$ of Definition 2.1 is overapproximated at any time instant $t \in \mathbb{R}_+$ and for any conic set of initial of states $\mathcal{P}_0 \in \mathbb{P}$ with the coefficient $P_0 \in \mathbb{S}^{(n+1)\times(n+1)}$

$$\mathcal{R}(t;\mathcal{P}_0) \subseteq \mathcal{P}(t)$$

where $\mathcal{P}$ is the time-varying paraboloid defined by its time-varying coefficient P solution of the DRE (time dependence is omitted for readability)

$$\begin{aligned} 0 =& \dot{P} + PA_1 + A_1^\top P - \mu\, M_{x1} \\ &+ \mu^{-1} \left(B_1^\top P - \mu M_{w,x1}\right)^\top M_w^{-1} \left(B_1^\top P - \mu M_{w,x1}\right) \end{aligned} \tag{2.5}$$

with initial condition $P(0) = P_0$ where μ is any measurable function strictly positive over $\mathbb{R}_+$, $A_1(\cdot)$, and $B_1(\cdot)$ are defined in (1.16).

Corollary 2.1 is a weaker form of Theorem 1.1 since the multiplier λ is taken as the constant function $\lambda(t) = 1$, $t \geq 0$. The following parts establish that this subset

of overapproximations is satisfactory enough to overapproximate the reachable set $\mathcal{R}(t;\mathcal{P}_0)$.

Next sections will manipulate the dynamic function of P in (2.5). We name it below. For given $t \geq 0$, $P_t \in \mathbb{S}^{(n+1)\times(n+1)}$, $\mu_t > 0$, let the operator $\text{Ricc} : \mathbb{R}_+ \times \mathbb{S}^{(n+1)\times(n+1)} \times \mathbb{R}_+^* \mapsto \mathbb{S}^{(n+1)\times(n+1)}$ be defined by

$$\begin{aligned}\text{Ricc}(t, P_t, \mu_t) =& P_t A_1(t) + A_1^\top(t) P_t - \mu_t\, M_{x1}(t) \\ &+ \mu_t^{-1} \left(B_1^\top(t) P_t - \mu_t M_{w,x1}(t)\right)^\top M_w^{-1}(t) \left(B_1^\top(t) P_t - \mu_t M_{w,x1}(t)\right).\end{aligned} \tag{2.6}$$

Then, for a given $\mu(t) \in \mathcal{D}^*$, solutions of (2.5) satisfies $\dot{P}(t) + \text{Ricc}(t, P(t), \mu(t)) = 0$ at any $t \geq 0$.

Remark 2.2. Comparison with the ellipsoidal method

In the case where $M_x = 0$, $M_{x,u} = 0$, $M_u = diag([0,\ldots,0,1])$, $M_{w,u} = M_w^{-1} w_c$. The system $\mathcal{S}$ of Definition 2.1 falls into the scope of ellipsoidal methods developed in [Chernous'ko, 1999]. In these works, the reachable sets are overapproximated with time-varying ellipsoids defined by their time-dependent center $x_c(\cdot)$ and time-dependent radius $Q(\cdot)$. (x_c, Q) and (E, f, g) are linked by the following equations

$$\begin{aligned} Q &= (g - f^\top E f)^{-1} E^{-1} \\ x_c &= -E^{-1} f \end{aligned}$$

By deriving (x_c, Q) according to time and using (2.5) (with the block decomposition (1.9)), we obtain similar differential equations than the one presented in [Chernous'ko, 1999]. Such a remark gives further insight into (2.5). The μ corresponds to a weight that either drives the system toward the disturbance direction or toward the direction of the system dynamic.

2.4 Support ellipsoids

This section defines *touching trajectories* and *support conics*. Touching trajectories are trajectories of $\mathcal{S}$ staying in contact with the boundary of the reachable set $\mathcal{R}(t;\mathcal{P}_0)$. Such trajectories are associated with a time-dependent conic overapproximation $\mathcal{P}$ (defined in Corollary 2.1) defined such that the state $x(t)$ of the touching trajectory stays on the boundary of $\mathcal{P}(t)$ at any time $t \geq 0$. Since the reachable set $\mathcal{R}(t;\mathcal{P}_0)$ and its overapproximation $\mathcal{P}(t)$ touches at $x(t)$, their normals coincide at $x(t)$ and the conic $\mathcal{P}(t)$ is said to be supported by the reachable set $\mathcal{R}(t;\mathcal{P}_0)$ (see Figure 2.4).

The optimal disturbance defined in (1.14) when $\lambda = 1$ is

$$w(t) = \mu(t)^{-1} M_w(t)^{-1} \left(B_1^\top(t) P(t) - \mu(t) M_{w,x1}(t)\right) \begin{bmatrix} x \\ 1 \end{bmatrix}. \tag{2.7}$$

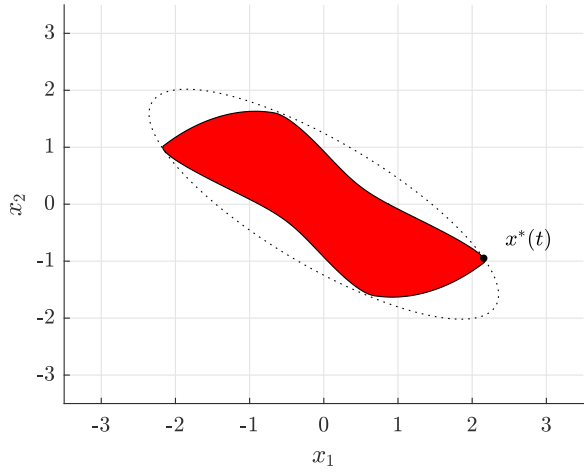


Figure 2.4: *The touching trajectory x touches the boundary of the reachable set $\mathcal{R}(t;\mathcal{P}_0)$ and the boundary of the overapproximation $\mathcal{P}(t)$ at any $t \geq 0$.*

Let x be the trajectory generated by the feedback w defined in (2.7). The integration of (1.15) gives

$$p(t, x(t)) = \langle y, y \rangle_{M,\mu_t} + p(0, x(0)). \tag{2.8}$$

Any point $x_t \in \mathbb{R}^n$ such that $p(t, x_t) = 0$ belongs to the boundary of the conic set $\mathcal{P}(t)$. Therefore, using (2.8), sufficient conditions for x to be a touching trajectory are

a) $y^\top(t)M(t)y(t) = 0$ for all $t \geq 0$,

b) $p(0, x(0)) = 0$, i.e. $x(0)$ belongs to the boundary of $\mathcal{P}(0)$.

Since w is chosen such that (2.7), the condition a) corresponds to an equation that depends over $\mu(t)^{-1}$, with $\mu(t) > 0$. The expansion of a) using (2.7) gives

$$\begin{aligned} y^\top(t)M(t)y(t) = z_t^\top M_{x1}(t) z_t + 2 z_t^\top M_{x1,w}(t) M_w(t)^{-1} \left(\mu(t)^{-1} B_1^\top(t) P(t) - M_{w,x1}(t)\right) z_t \\ + z_t^\top \left(\mu(t)^{-1} B_1^\top(t) P(t) - M_{w,x1}(t)\right) M_w(t)^{-1} \left(\mu(t)^{-1} B_1^\top(t) P(t) - M_{w,x1}(t)\right) z_t \end{aligned} \tag{2.9}$$

where $z_t = \begin{bmatrix} x(t) \\ 1 \end{bmatrix}$. If, at any time $t \geq 0$, there is a strictly positive root $\mu^*(t)$ to the quadratic equation (2.9), then a) is satisfied. We now show that such a $\mu^*(t) \in (0, \infty)$ exists. Let $h : \mathbb{R} \mapsto \mathbb{R}$ be the quadratic function defined by $h(\mu(t)^{-1}) = y^\top(t)M(t)y(t)$. When $\mu(t)^{-1} = 0$, expression (2.9) becomes

$$h(0) = z_t^\top \left(M_{x1}(t) - M_{x1,w}(t) M_w(t)^{-1} M_{w,x1}(t)\right) z_t$$

Since $M_{x1}(t) - M_{x1,w}(t)M_w(t)^{-1}M_{w,x1}^{\top}(t) \succ 0$, it holds $h(0) > 0$. Since $M_w \prec 0$, when $\mu(t) \to 0$, $h(\mu(t)^{-1}) \to -\infty$. h is a continuous function, $h(0) > 0$ and $h(\infty) = -\infty$, by the intermediate value theorem, there is always a strictly positive solution $\mu^*(t)$ to the equation $y^{\top}(t)M(t)y(t) = 0$ and therefore condition a) can be satisfied.

The above discussion is summarized in the following property.

Proposition 2.2. Touching trajectories and support conic
When w, μ are chosen such that conditions a) and b) holds, then x is a touching trajectory and $\mathcal{P}$ is a support conic.

Proposition 2.2 defines some touching trajectory of the reachable set. Touching trajectories are defined as long as their associated support conic is defined.

2.5 Optimal ellipsoids

In this section, we address the problem of finding overapproximations minimizing some given criteria. Necessary optimality conditions are derived out of the Pontryagin's Maximum Principle (PMP) in Section 2.5.1. These conditions lead to locally optimal solutions of the initial optimization problem. Section 2.5.2 provides a numerical method to compute sub-optimal overapproximations.

In this section, we are interested in solving the following optimization problem

$$\begin{array}{ll} \text{Minimize} & J(P(\cdot)) \\ \text{such that} & P(\cdot) \text{ is a solution of (2.5)} \\ & \text{with } \mu(t) > 0 \text{ over } [0,T] \end{array} \tag{2.10}$$

where $J : \mathscr{L}_2([0,T]; \mathbb{S}^{(n+1)\times(n+1)}) \to \mathbb{R}$ associates to a time-dependent conic $\mathcal{P}(\cdot)$ a cost in $\mathbb{R}$. The cost is composed of an integral cost and a final cost as follow

$$J(P(\cdot)) = \Psi(P(T)) + \int_0^T L(P(t))dt$$

The following specializations of (2.10) will be detailed within this section:

a) the minimal pseudo-volume (value proportional to the squared volume, defined in (1.10)) of the terminal conic

$$\begin{cases} \Psi(P_T) = \widetilde{\text{Vol}}(P_T) \\ L(P_t) = 0 \end{cases} \tag{2.11}$$

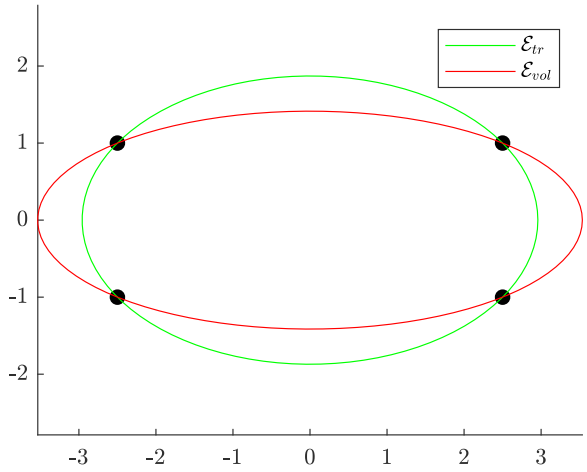


Figure 2.5: *Ellipsoids overapproximating a set of points: minimum volume ellipsoid (in red, corresponds to the objective a)), minimum sum of square semi-axis (in green, corresponds to the objective c)).*

b) the minimal pseudo-volume of the tube $t \to \mathcal{P}(t)$

$$\begin{cases} \Psi(P_T) = 0 \\ L(P_t) = \widetilde{\mathrm{Vol}}(P_t) \end{cases} \tag{2.12}$$

c) the sum of squared semi-axis (defined in (1.11))

$$\begin{cases} \Psi(P_T) = \mathrm{TrSq}(P_T) \\ L(P_t) = 0 \end{cases} \tag{2.13}$$

Figure 2.5 illustrates the difference between the different objectives of cases a) and c).

2.5.1 Application of the Pontryagin's Minimum Principle

Necessary conditions at the optimal solution of Problem (2.10) can be derived using the Minimum Pontryagin's Principle for matrices as presented in [Athans, 1967].

The PMP is an approach to solve optimal control problems, i.e. to find a control signal for system trajectory such that some cost function is minimized. To find the necessary conditions of optimality, the Hamiltonian of the dynamical system together with a co-state variable are introduced. The Hamiltonian is a storage function that preserves the final cost of a trajectory. The co-state variable measures the sensitivity

of the final cost to variation in the control signal. The co-state is the solution of a linear time-varying system with final state conditions. Necessary conditions of optimality imply constraints between the Hamiltonian, the co-state, and the control signal.

In the present case, the dynamical system is the time-dependent parameter of the conic and is described by differential equations (2.5), the control input is the μ parameter. Contrary to the classical application of the PMP for dynamical systems where the state is represented with vectored variables, P is a symmetric matrix real-valued variable. The co-state can be expressed as a solution to a matrix ordinary differential equation. The scalar product corresponds to the inner product of the space, i.e. $trace(A^\top B)$ where $A, B \in \mathbb{S}^{n\times n}$. Note that the matrix representation is equivalent to a vectored representation of each variable. However, the formulation is more convenient in the matrix form.

In what follows, we define the Hamiltonian H, the co-state Q, the optimal control μ^*. This section ends with the necessary conditions satisfied by an optimal solution P^* to (2.10). Those necessary conditions are formulated as a Boundary Value Problem (BVP) in Theorem 2.1. The next subsection provides numerical methods to compute sub-optimal solutions to (2.10) using the BVP.

Let the Hamiltonian be defined by

$$H(P(t), Q(t), \mu(t), t) = trace(Q(t)\mathrm{Ricc}(t, P(t), \mu(t))) + L(P(t)) \tag{2.14}$$

where $Ricc$ is defined in (2.6). Let Q be the co-state solution of

$$\dot{Q} - A_Q(t)Q - QA_Q^\top(t) - U_P(t) = 0 \tag{2.15}$$

where

$$A_Q(t) = A_1(t) + \mu(t)^{-1}\left(B_1^\top(t)P(t) - \mu(t)M_{w,x1}(t)\right)^\top M_w^{-1}(t)B_1^\top(t)$$

and

$$U_P(t) = L_P(P(t))$$

with final conditions

$$Q(T) = \Psi_P(P(T)).$$

Functions L_P and Ψ_P can be easily computed in cases a), b) and c) ($\widetilde{\mathrm{Vol}}_P$ and T_P are respectively defined in (2.22) and (2.23), their differential are defined in Appendix 2.A):

a) $\begin{cases} L_P(P_t) = 0 \\ \Psi_P(P_T) = \widetilde{\mathrm{Vol}}_P(P_T) \end{cases}$

b) $\begin{cases} L_P(P_t) = \widetilde{\mathrm{Vol}}_P(P_t) \\ \Psi_P(P_T) = 0 \end{cases}$

c) $\begin{cases} L_P(P_t) = 0 \\ \Psi_P(P_T) = \mathrm{TrSq}_P(P_T) \end{cases}$

The optimal control μ^* is a solution of the following minimization problem

$$\mu^*(t) = \underset{\mu(t)>0}{\arg\min}\, H(P(t), Q(t), \mu(t), t) \tag{2.16}$$

Let the function $h_t : \mu_t \to H(P(t), Q(t), \mu_t, t)$, it holds

$$h_t = h_t^0 - \mu_t trace\left(Q(t)M_{x1\backslash w}(t)\right) + \mu_t^{-1} trace\left(Q(t)P(t)B_1(t)M_w^{-1}(t)B_1(t)P(t)\right)$$

where h_t^0 is a term independent of μ_t. Deriving h_t gives the following minimizer $\mu^*(t)$ for (2.16)

$$\mu^*(t) = \left[\frac{-trace\left(Q(t)P(t)B_1(t)M_w^{-1}(t)B_1^\top(t)P(t)\right)}{trace\left(Q(t)M_{x1\backslash w}(t)\right)}\right]^{1/2} \tag{2.17}$$

Since, $M_w(t) \prec 0$ and $M_{x1\backslash w}(t) \succ 0$, when $Q(t) \succ 0$, $\mu^*(t)$ is correctly defined (i.e, the square root exists) and the minimization problem (2.16) does have a solution.

Theorem 2.1. Necessary conditions for optimality
Let P^* be an optimal trajectory associated with the optimal control μ^* as defined in (2.17) *of* (2.10).

Necessary conditions for P^* are expressed as the existence of a solution to a boundary value problem. The co-state is a symmetric matrix function.

2.5.2 A continuation method to solve the PMP

The BVP defined in Theorem 2.1 is not easy to solve. Solutions of (2.5) might have a finite escape time. Even when the solution $E(\cdot)$ is correctly defined for a given $\mu(\cdot)$, there are no guarantees about the sign of $E(\cdot)$. Considering the different optimality criterion presented in Section 2.5.1, when $E(\cdot)$ is not invertible along the trajectory, the integration of the co-state variable is compromised.

In this part, we investigate the use of a continuation method to solve the BVP in Theorem 2.1.

The BVP can be equivalently defined as the following root-finding problem parametrized by the final time of integration $T > 0$

$$F(Q_0, T) = Q(T) - \Psi_P(P(T)). \tag{2.18}$$

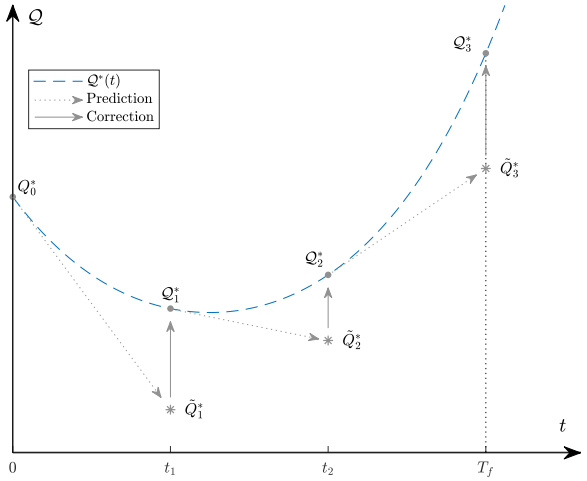


Figure 2.6: *In the continuation method, we follow the curve $\mathcal{Q}$ over $[0, T_f]$ by using a prediction-correction algorithm.*

where P is a solution of (2.5) for a given initial state Q_0, Q is a solution of (2.15) with the initial state $Q(0) = Q_0$.

The continuation method proposes to find the solution Q_{0,T_f} of (2.18) where $T = T_f$, $T_f > 0$ given, by following the curve of solutions

$$\begin{array}{rccc} \mathcal{Q}: & [0, T_f] & \mapsto & \mathbb{S}^{(n+1)\times(n+1)} \\ & T & \to & Q_{0,T} \end{array} \tag{2.19}$$

from a trivial point (that will be at $T = 0$ in our case) to the point of interest $T = T_f$. When the curve $\mathcal{Q}$ is regular enough, a prediction correction algorithm can be used to follow the curve over $[0, T_f]$ (see Figure 2.6).

When $T = 0$, the root Q_0 of the equation $F(Q_0, 0) = 0$ is indeed trivial (using (2.18)).

Proposition 2.3. Initial point the co-state
The solution of (2.18) *for* $T = 0$ *is* $Q(0) = \Psi_P(P_0)$.

The regularity of the curve $\mathcal{Q}$ is difficult to assess globally since the curve is an implicit function; it would require to study the regularity of the inverse function of $Q_0 \to F(Q_0, T)$ for $T \in [0, T_f]$. However, for a given point $(T, Q_{0,T})$ on the curve $\mathcal{Q}$, if F is linearizable at $(T, \mathcal{Q}(T))$ and if $Q_0 \to \widetilde{F}(Q_0, T)$ can be inverted, then the curve $\mathcal{Q}$ exists in the neighborhood of T. This is stated in the implicit function theorem as described in [Zeidler, 1995b, Chapter 4.8].

The linearization $\widetilde{F}$ of F can be derived using variational calculus (see Fréchet derivatives [Zeidler, 1995b, Chapter 2.1]). Let $R : [0, T] \to \mathbb{S}^{(n+1)\times(n+1)}$ and S :

$[0,T] \to \mathbb{S}^{(n+1)\times(n+1)}$ (resp.) be variations of P and Q (resp.) for a given variation $S_0 \in \mathbb{S}^{(n+1)\times(n+1)}$ to the initial co-state Q_0. R and S are solutions of an LTV system, details of how (R,S) are derived are given in Annexe 2.B. The linearization $\widetilde{F}$ along (S_0, dt) is

$$\widetilde{F} = \Delta F_{Q_0} \cdot S_0 + \Delta F_T \, dt$$

where

$$\Delta F_{Q_0} \cdot S_0 = R(T) - \Delta\Psi_P(P(T)) \cdot S(T)$$
$$\Delta F_T = \dot{Q}(T) - \Delta\Psi_P(P(T)) \cdot \dot{P}(T)$$

where the Fréchet derivatives $\Delta\Psi_P(P) \cdot S$ are defined in Appendix 2.A.

Proposition 2.4. Tangent of the co-state

When ΔF_{Q_0} is invertible, the tangent S_0 of $\mathcal{Q}$ satisfies

$$0 = \Delta F_{Q_0} \cdot S_0 + \Delta F_T. \tag{2.20}$$

The BVP is solved for an increasing sequence of time-horizons $\{T_k\}$ until the time-horizon T_f is reached. For each time -horizon T_k, we compute the solution of the BVP problem, i.e. the initial state of the co-state Q_{0,T_k}. We use the continuity of the curve (guaranteed by using the inverse function theorem) to predict the next $Q_{0,T_{k+1}}$.

Remark 2.3. Corrector step

The corrector step numerically solving (2.18) *uses a BVP solver such as* bvp5c *[Kierzenka and Shampine, 2008].*

Remark 2.4. Complexity of the BVP algorithm

P, Q, R and S are symmetric matrices. Therefore, the differential equations have a $2n(n+1)$ dimensional state.

2.6 Example

We study a linear time-invariant system of two dimensions defined by (2.1) with the following parameters

$$\begin{aligned} A &= \begin{bmatrix} -2 & 1 \\ -3 & -3 \end{bmatrix} \\ B &= \begin{bmatrix} 1 & 1 \end{bmatrix}^\top \\ C &= \begin{bmatrix} 1 & 1 \end{bmatrix}^\top \\ u &= 0 \\ M &= diag\Big(\begin{bmatrix} 1 & 2 & 0 & -1 \end{bmatrix} \Big) \end{aligned} \tag{2.21}$$

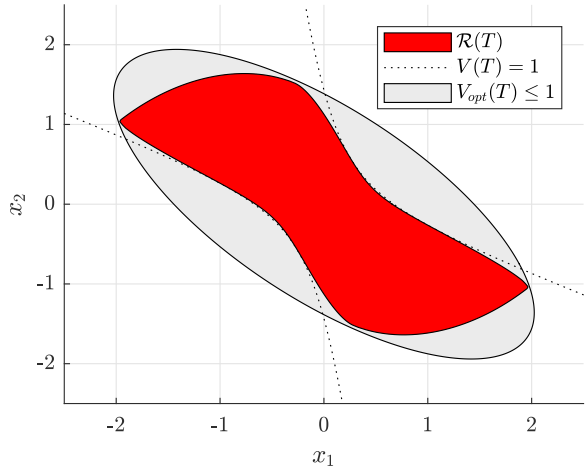


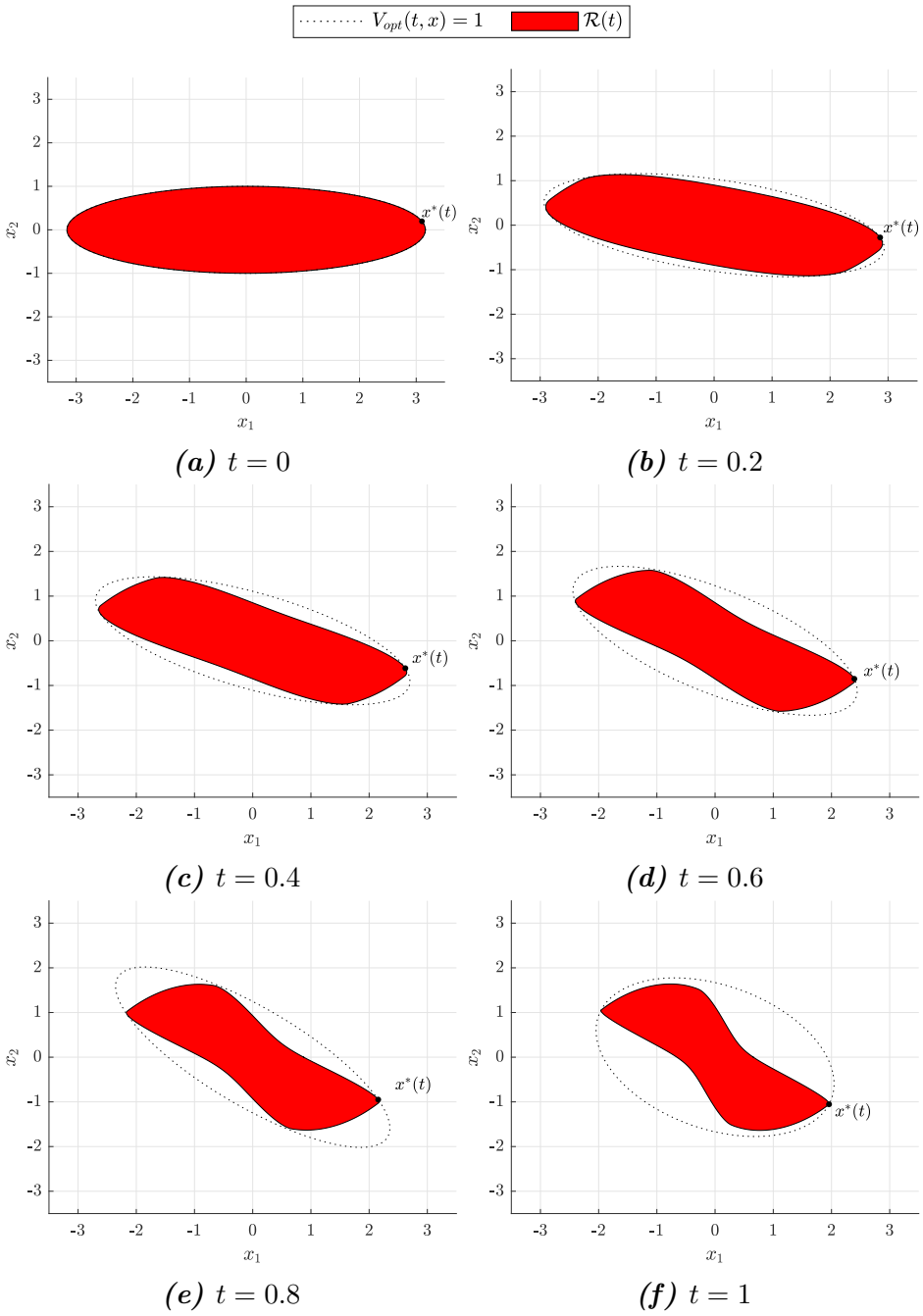
Figure 2.7: *Reachable set of the system of Definition 2.1 with parameters defined in* (2.21)*. The gray area corresponds to the minimum volume ellipsoid. Since the Pontryagin's Maximum Principle only gives necessary conditions for optimality, and since the ellipsoidal method only provides overapproximations, the $\mathcal{E}_{vol}$ is not the actual minimum volume ellipsoid overapproximating the reachable tube $\mathcal{R}(T; \mathcal{E}_0)$. The dotted line represents an overapproximating hyperbola touching the reachable set.*

The reachable set at $T = 1$ and the minimal volume overapproximating ellipsoid are drawn in Figure 2.7. The minimal volume overapproximating ellipsoid is drawn at different times in Figure 2.8.

2.7 Discussion

The QC constraint is a ∞-norm constraint over the state-input-disturbance signal. Such QC system (or subfamilies of QC systems) has been studied many times in the literature. First studies goes back to the 60's with the *unknown-but-bounded* disturbance model (in [Schweppe, 1973], Section 7.5), in this model the disturbance belongs to an ellipsoidal set at any time, i.e. $(w(t) - w_c(t))^\top R(t)(w(t) - w_c(t)) \leq 1$. In other works, the inequality $\alpha v^2 \leq vw \leq \beta v^2$ which is equivalent to the QC constraint $\left\|w - \frac{\alpha+\beta}{2}\right\| \leq \frac{\beta-\alpha}{2}\|v\|$, is referred as the *sector inequality* (as in [Megretski and Rantzer, 1997]). [Boyd et al., 1994] (in Section 4.2.3) studies the stability of the *norm-bound* linear time-invariant system, a system subject to disturbance $w = \Delta x$ where $\|\Delta(t)\| \leq 1$. Such constraint is equivalent to the QC constraint $\|w(t)\| \leq \|x(t)\|$.

Any QC constraint $y(\cdot)^\top M(\cdot)y(\cdot) \geq 0$ is equivalent to the set of IQC constraints $\forall \mu(\cdot) \geq 0, \int_0^T y(t)^\top M(t)y(t)\mu dt \geq 0$ for any $T \in \mathbb{R}_+ \cup \{+\infty\}$. For this reason,



(a) $t = 0$ ***(b)*** $t = 0.2$

(c) $t = 0.4$ ***(d)*** $t = 0.6$

(e) $t = 0.8$ ***(f)*** $t = 1$

Figure 2.8: *Reachable set $\mathcal{R}$ (in red) of the dynamical system $\mathcal{S}$, defined in Section 2.6, at different time instant t. V_{opt} is the level-set associated with the minimal volume overapproximating ellipsoids and x^* is its associated touching trajectory.*

the study of QC system is closely related to the study of IQC systems. Moreover, the integration of the QC constraint is a necessary step that happens during the Lagrangian relaxation of the LQR problem. [Bertsekas and Rhodes, 1971] identified this link in the case of guaranteed state estimation where the system is subject to hard bound constraint or to energy constraint. In the robust control community, a lot of effort in the study of QC constraints has been dedicated to the search of a class of multiplier μ that can be expressed as state-space model of finite dimension (see [Veenman et al., 2016] Section 6).

The most common way to handle QC constraints is the use of constant multipliers, i.e. $\mu(\cdot) = \mu_0 > 0$ as in [Jönsson, 1996]. Such an approach is commonly referred in the literature as the S-procedure (see [Boyd et al., 1994], Section 2.6.3).

More recently, [Fetzer et al., 2018, Veenman et al., 2016] used so-called *dynamic multipliers* that allowed to reduce the conservatism of robust stability analysis and reachability analysis of IQC systems. The inequality $w(t)^2 \leq v(t)^2$ is equivalently replaced by $\langle w, Hw \rangle \leq \langle v, Hv \rangle$ where H is a positive operator over the set of signal (i.e. a positive transfer function).

For QC systems, multipliers belong to the set of positive functions over the time interval. Each multiplier is associated with a valid overapproximation of the reachable set. Since the set of multipliers is infinite, it is interesting to find one "good" overapproximation.

In [Chernous'ko, 1999], the author tries to search for minimal volume overapproximation (for the case of exogenous disturbances, i.e. $w^\top(t)R(t)w(t) \leq 1$ with $R(t) \succeq 0$). This optimization problem is an optimal control problem for which necessary conditions can be derived using the Pontryagin's Maximum Principle (PMP, as in Section 2.5). In our work, we consider a wider family of systems, namely the LTV systems with a disturbance that satisfies a quadratic constraint at any time. The ellipsoidal method is then a sub-case of our approach.

In the context of discrete-time IQC models, [Fry et al., 2017] chooses the overapproximation that maximizes the input-to-output robust $\mathscr{L}_2$ gain (where the $\mathscr{L}_2$ norm of the signal is computed over a finite time interval). A similar approach method is applied in [Seiler et al., 2019] for the case of continuous-time system with IQC systems. In both of these works, since the Riccati equation (DRE in the continuous-time case, difference Riccati equation for the discrete-time case) is expressed as an LMI constraint, SDP solvers can then be used to find a solution minimizing a given criterion by using an SDP solver.

2.8 Conclusion

This chapter applied the framework introduced in Chapter 1 for the specific case of disturbances with a point-wise quadratic constraint. Since the set of time-varying conic overapproximations is infinite, we proposed two methods to find one "tight" overapproximation of the reachable set. In the first method, we compute a touching time-varying conic overapproximation, this overapproximation touches the reachable set along system trajectories (so-called touching trajectories). In the second method, we find a time-varying paraboloid that minimizes a given criterion (e.g. its volume at a given time). We derive necessary conditions by using the Pontryagin's Maximum Principle (PMP) and proposed a continuation method to solve its associated boundary value problem.

Future Works In this chapter, we only computed overapproximations minimizing their volume at a given time. An interesting extension of this chapter would be to compute be the time-varying conic minimizing its volume at each time over the interval of integration. To address this problem, it is possible to parametrize optimal overapproximations $\mathcal{P}(t, t^*)$ with two time indexes t and t^*. t^* corresponds to the time where the time-varying conic $t \mapsto \mathcal{P}(t, t^*)$ minimizes its volume, t corresponds to the regular time index. In such a situation, the overapproximation $t^* \mapsto \mathcal{P}(t^*, t^*)$ is a time-varying conic with a minimal volume at each time instant t^*. Another problem of interest would be to adapt the work of [Seiler et al., 2019], where optimal multipliers are searched by solving a Differential Linear Matrix Inequality (DLMI), to our optimization problem in order to find the optimal positive multiplier μ.

Appendices of Chapter 2

2.A Derivatives of the cost functions

We hereby detail the computations of the first derivative $\widetilde{\text{Vol}}_P$ and TrSq_P (resp.) of $\widetilde{\text{Vol}}$, defined in (1.10), and TrSq, defined in (1.11) (resp.). Then, the Fréchet derivatives of $\widetilde{\text{Vol}}_P$ and TrSq_P at $P \in \mathbb{S}^{(n+1)\times(n+1)}$ in a given direction $S \in \mathbb{S}^{(n+1)\times(n+1)}$.

Derivative of $\widetilde{\text{Vol}}$ The derivative of $\widetilde{\text{Vol}}$ at P in the direction $\delta P \in \mathbb{S}^{(n+1)\times(n+1)}$ is $\widetilde{\text{Vol}}_P(P)\cdot\delta_P = det(-E)^{-1}(\delta_g - 2\delta_f^\top Ef - f^\top \delta_E f) - det(-E)^{-1}(g - f^\top Ef)trace(-E^{-1}\delta_E)$.

$\widetilde{\text{Vol}}_P(P)\cdot\delta_P$ can be expressed with the inner product

$$\widetilde{\text{Vol}}_P(P)\cdot\delta_P = trace\left(det(-E)^{-1}\begin{bmatrix} -ff^\top + (g - f^\top Ef)E^{-1} & -Ef \\ -(Ef)^\top & 1 \end{bmatrix}\delta_P\right)$$

Therefore, we can express the differential $\widetilde{\text{Vol}}_P(P)$ of $\widetilde{\text{Vol}}$ at P in its matrix form (we hereby use an abuse of notation)

$$\widetilde{\text{Vol}}_P(P) = det(-E)^{-1}\begin{bmatrix} -ff^\top + (g - f^\top Ef)E^{-1} & -Ef \\ -(Ef)^\top & 1 \end{bmatrix} \tag{2.22}$$

We compute the Fréchet derivative $\Delta_P\widetilde{\text{Vol}}_P \cdot S$ of $\widetilde{\text{Vol}}_P$ in the direction $S \in \mathbb{R}^{(n+1)\times(n+1)}$. Let the block decomposition of S

$$S = \begin{bmatrix} T & u \\ u^\top & v \end{bmatrix}$$

To compute $\Delta_P\widetilde{\text{Vol}}_P \cdot S$, we first the following functions

$$\begin{aligned} d(P) &= det(-E)^{-1} \\ r(P) &= g - f^\top Ef \\ V_P(P) &= \begin{bmatrix} -ff^\top + r(P)E^{-1} & -Ef \\ -(Ef)^\top & 1 \end{bmatrix} \end{aligned}$$

then $\widetilde{\text{Vol}}_P = d(P)V_P(P)$. Fréchet derivatives of d, r and V_P in the direction S are

$$\begin{aligned} \Delta_P d(P)\cdot S &= -det(-E)^{-1}trace(E^{-1}T) \\ \Delta_P r(P)\cdot S &= v - u^\top Ef - f^\top Tf - f^\top Eu \\ \Delta_P V_P(P)\cdot S &= \begin{bmatrix} -uf^\top - fu^\top + \Delta r(P)\cdot S\, E^{-1} - r(P)E^{-1}TE^{-1} & -Tf - Eu \\ -(Tf + Eu)^\top & 0 \end{bmatrix} \end{aligned}$$

Therefore,

$$\Delta_P \widetilde{\mathrm{Vol}}_P \cdot S = \Delta_P d(P) \cdot S\, V_P(P) + d(P)\, \Delta V_P(P) \cdot S.$$

Derivative of TrSq similarly than for $\widetilde{V}$, we can derive the differential TrSq_P of TrSq at P (in its matrix representation)

$$\mathrm{TrSq}_P(P) = trace(E^{-1}) \begin{bmatrix} ff^\top & Ef \\ (Ef)^\top & -1 \end{bmatrix} + (g - f^\top Ef) \begin{bmatrix} E^{-2} & 0 \\ 0 & 0 \end{bmatrix} \tag{2.23}$$

Similarly, we can compute the Fréchet derivative of TrSq_P at P in the direction S.

$$\begin{aligned} \mathrm{TrSq} =& trace(-E^{-1}TE^{-1}) \begin{bmatrix} ff^\top & Ef \\ (Ef)^\top & -1 \end{bmatrix} + trace(E^{-1}) \begin{bmatrix} uf^\top + fu^\top, Tf + Eu & \\ (Tf + Eu)^\top & 0 \end{bmatrix} \\ & - \Delta r(P) \cdot S \begin{bmatrix} E^{-2} & 0 \\ 0 & 0 \end{bmatrix} - (g - f^\top Ef) \begin{bmatrix} -E^{-1}TE^{-2} - E^{-2}TE^{-1} & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

2.B Variations of the state and co-state

Hereby, we explicit the computation of the variation (R, S) of the state and co-state (P, Q) with respect to the initial state $(0, S_0)$. R and S are solutions of

$$\begin{cases} \dot{R} + A_Q^\top(t)R + RA_Q(t) + V_R(t) + trace(ST_S + RT_R)H_R = 0 \\ \dot{S} - A_Q(t)S - SA_Q^\top(t) + V_S(t) + trace(ST_S + RT_R)H_S = 0 \end{cases} \tag{2.24}$$

with initial conditions

$$\begin{cases} R(0) = 0 \\ S(0) = S_0 \end{cases}$$

For the following computations, all signals are time-dependent. We omit its notation to simplify the following computations. Let $k = \frac{a}{b}$, $\mu = \sqrt{k}$, then $d\mu = \frac{da}{a}\mu - \frac{db}{b}\mu$, where

$$\begin{aligned} a &= -trace(QM_{x1\backslash w}) \\ b &= trace(QPB_1M_w^{-1}B_1^\top P) \end{aligned}$$

$$\begin{aligned} da &= \Delta a \cdot (R, S) &&= -trace(SM_{x1\backslash w}) \\ db &= \Delta b \cdot (R, S) &&= trace(SPB_1M_w^{-1}B_1^\top P + R\, B_1M_w^{-1}B_1^\top PQ + QPB_1M_w^{-1}B_1^\top\, R) \end{aligned}$$

$$\begin{aligned} T_S &= \frac{\mu\, M_{x1\backslash w}}{trace(QM_{x1\backslash w})} - \frac{\mu\, PB_1M_w^{-1}B_1^\top P}{trace(QPB_1M_w^{-1}B_1^\top P)} \\ T_R &= -\frac{\mu\,(B_1M_w^{-1}B_1^\top PQ + QPB_1M_w^{-1}B_1^\top)}{trace(QPB_1M_w^{-1}B_1^\top P)} \end{aligned}$$

Finally,

$$d\mu = \Delta b \cdot (R, S) = trace(S T_S + R T_R).$$

$$H_R = -M_{x1\backslash w} - \mu^{-2} P B_1 M_w^{-1} B_1^\top P$$
$$H_S = -\mu^{-2}(P B_1 M_w^{-1} B_1^\top Q + Q B_1 M_w^{-1} B_1^\top P)$$

$$V_S = \Delta_P L_P \cdot R + \mu^{-1} R B_1 M_{w^{-1}} B_1^\top Q + \mu^{-1} Q B_1 M_{w^{-1}} B_1^\top R$$

2.C Contracting property in the centered case

We show that in the specific case of centered stable LTI system (see the definition on the paragraph below), we can find an overapproximation that is contracting as t goes to ∞, i.e. $\mathcal{P}(t) \to \{0\}$. The following chapter shows that this result is not available in the case of IQC systems.

Let $\mathcal{S}$ a *centered stable LTI QC system* be a QC system as in Definition 2.1 with:

1. time-invariant A, B, C and M matrices (the linear time-invariant system part);
2. with a centered set of initial states (i.e. with $f_0 = 0$ in the block decomposition of P_0);
3. with a centered set of disturbances (i.e. $M_{x1,w}(\cdot) = 0$);
4. with a null input signal, i.e. $u(\cdot) = 0$;
5. such that $\mathcal{S}$ is exponentially stable.

These assumptions are restrictive but will simplify the following proofs. When this specific case, DRE (1.21) is

$$0 = \dot{E} + EA + A^\top E - \mu(t) M_x + \mu(t)^{-1} E B M_w^{-1} B^\top E. \tag{2.25}$$

Assumptions 1 and 5 let us express a negative definite solution $\bar{E}(\cdot)$ to the DRE (2.25) with initial condition $\bar{E}(0) = \bar{E}_0 \prec 0$ such that $\bar{E}^{-1}$ converges to 0 when $t \to \infty$. Assumptions 2, 3 and 4 are chosen such that any overapproximation $\mathcal{P}$ is centered at any time, when they holds, $f(\cdot) = 0$ and $g(\cdot) = g_0$. Since $\bar{E}(\cdot) \leq 0$, $\mathcal{P}(\cdot)$ is an ellipsoidal set at any time. The radius of $\mathcal{P}$ is $\sqrt{\bar{E}^{-1}}$, and if $\bar{E}(t)^{-1} \to 0$ when $t \to \infty$, then $\mathcal{P} \to \{0\}$. The following paragraph will demonstrate this result. First, we express such a $\bar{E}$ and show that $\bar{E}(t)^{-1} \to 0$ when $t \to \infty$. Then we show that $\bar{E}$ is an upperbound for solutions of (2.25) with any initial conditions $E_0 \prec 0$.

Let $\bar{E}(t) = \exp \lambda t \bar{E}_0$ and $\mu(t) = \bar{\mu}(t) = \kappa \exp \lambda t$ with $\lambda, \kappa > 0$. $\bar{E}$ satisfies (2.25) if and only if $\bar{E}_0$ is solution to the following Continuous Algebraic Riccati Equation

$$0 = \lambda \bar{E}_0 + \bar{E}_0 A + A^\top E - \kappa M_x + \kappa^{-1} \bar{E}_0 B M_w^{-1} B^\top \bar{E}_0. \tag{2.26}$$

Since the system is exponentially stable, there is a $\bar{E}_0 \prec 0$ and $\kappa > 0$ such that (2.26) is solved for $\lambda = 0$. By continuity, there is a $\bar{E}_0 \prec 0$, a $\kappa > 0$ and a $\lambda > 0$ such that (2.26) holds.

It will appear that $\bar{E}(t)$ is an upperbound for any solution to (2.25). Let $\Delta(t) = \bar{E}(t) - E(t)$, we aim at proving that $\Delta(t) \succ 0$. Using (2.25) and the definition of $\bar{E}$, it holds

$$\dot{\Delta} = -A^\top \Delta - \Delta A - \bar{\mu}(t)^{-1} \bar{E}(t) B M_w^{-1} B^\top \bar{E}(t) + \bar{\mu}(t)^{-1} E(t) B M_w^{-1} B^\top E(t).$$

Using the following identity

$$\bar{E} R \bar{E} - E R E = \Delta R \Delta + E R \Delta + \Delta R E$$

where $R = -\bar{\mu}(t)^{-1} B M_w^{-1} B^\top$, $R(t) \succ 0$, we have

$$\dot{\Delta} + A_R(t)^\top \Delta + \Delta A_R(t) = \Delta(t) R(t) \Delta(t)$$

By integration, it holds

$$\Phi_R^\top(T) \Delta(T) \Phi_R(T) = \Delta(0) + \int_0^T \Phi^\top(t) \Delta(t) R(t) \Delta(t) \Phi^\top(t) dt$$

where Φ_R is the transition matrix associated with A_R. Therefore, when $\Delta(0) \succ 0$, since $R \succ 0$, it holds $\Delta \succ 0$.

By definition, $\bar{E}^{-1} \to 0$ when $t \to \infty$. Since $\bar{E}(t) \succ E(t)$, it holds

$$0 \prec E(t)^{-1} \prec \bar{E}(t)^{-1}$$

therefore, $E(t)^{-1} \to 0$. By consequence, $\mathcal{P}(t) \to \{0\}$ when $t \to \infty$.

Chapter 3

Integral Quadratic Constraints

Contents

Chapter 1 introduced the general framework of a set-based simulation method for linear systems subject to bounded disturbances. This chapter applies it for a specific

subclass of systems: LTV system subjects to a disturbance that satisfies an Integral Quadratic Constraint (IQC) on the state and input signal. IQC models are a classical tool of robust control theory (see e.g. [Megretski, 2010, Megretski and Rantzer, 1997]). They can model infinite-dimensional states, nonlinear dynamics, delays, rate limiters, uncertain systems (see [Helmersson, 1999], [Rantzer and Megretski, 1998], [Peaucelle et al., 2014] and [Ariba et al., 2018]).

In Section 3.1, we first define IQC systems. In Section 3.2, we motivate the use of IQC systems as a modeling tool. In Section 3.3, we apply Theorem 1.1 of Chapter 1 to define overapproximations of the reachable set of an IQC system. Contrary to overapproximations defined in Chapter 2 that are supported by the reachable set at any time, it is not possible to satisfy such property for an IQC system. The next sections are then dedicated to the investigation of another approach in order to find tight overapproximations to the reachable set. In Section 3.4, we define an extended system of the original IQC system. In Section 3.5, we show that the reachable set of this extended system can be overapproximated with time-varying paraboloids and we identify touching trajectories and their associated overapproximations supported by the reachable set. In Section 3.6, we define the intersection of all the supporting time-varying paraboloids, Theorem 3.1 states that the reachable set of the extended system is exactly characterized by this intersection of overapproximations. A projection of this overapproximation is the exact reachable set of the initial IQC system. In Section 3.7, we detail the practical implementation of the overapproximation of the reachable set of the IQC system. In Section 3.8, we present some examples. In Section 3.9 and Section 3.10, we discuss and conclude this chapter.

3.1 Definition of the system

In this chapter, the system of interest is an LTV system subjects to a disturbance that satisfies an Integral Quadratic Constraint (IQC) with the state and input signal. This IQC constraint can be seen as an energetic relationship on the disturbance and the state/input signals (see Remark 3.1). This system is a subclass of the system described in Definition 1.1.

Definition 3.1. IQC system
Let the system $\mathcal{S}$ be defined by

$$\mathcal{S}: \begin{cases} \dot{x}(t) &= A(t)x(t) + B(t)w(t) + C(t)u(t) \\ \displaystyle\int_0^t \begin{bmatrix} x(\tau) \\ u(\tau) \\ w(\tau) \end{bmatrix}^\top M(\tau) \begin{bmatrix} x(\tau) \\ u(\tau) \\ w(\tau) \end{bmatrix} d\tau \geq 0 \text{ for every } t \geq 0 \end{cases} \tag{3.1}$$

where M is a quadratic form negative in the disturbance dimension, i.e. such

that $M_w(\cdot) \prec 0$ *with the block decomposition of* M

$$M(t) = \begin{bmatrix} M_x(t) & M_{x,u}(t) & M_{x,w}(t) \\ M_{x,u}^\top(t) & M_u(t) & M_{u,w}(t) \\ M_{x,w}^\top(t) & M_{w,u}^\top(t) & M_w(t) \end{bmatrix} \in \mathbb{S}^{(n+m+p)\times(n+m+p)}.$$

Since for two system trajectories (x_1, w_1, u_1) and (x_2, w_2, u_2) of $\mathcal{S}$, the trajectory $(x_1, w_1, u_1) + (x_2, w_2, u_2)$ does not necessarily satisfy the IQC constraint, i.e. the system $\mathcal{S}$ is not linear. However, $(\alpha\, x_1, \alpha\, w_1, \alpha\, u_1)$, for $\alpha \in \mathbb{R}$, is a system trajectory, and therefore the set of trajectories of $\mathcal{S}$ is conic.

Remark 3.1. Energetic constraint
The IQC constraint in (3.1) *can be reformulated as follows*

$$\int_0^t (w(\tau) - w_c(\tau))^\top R\,(w(\tau) - w_c(\tau))\,d\tau \leq \int_0^t \left[\begin{smallmatrix} x(\tau) \\ u(\tau) \end{smallmatrix}\right]^\top Q \left[\begin{smallmatrix} x(\tau) \\ u(\tau) \end{smallmatrix}\right] d\tau \qquad (3.2)$$

where $R = -M_w \succ 0$,

$$Q = \begin{bmatrix} M_x & M_{xu} \\ M_{xu}^\top & M_u \end{bmatrix} - \begin{bmatrix} M_{xw} \\ M_{uw} \end{bmatrix}^\top M_w^{-1} \begin{bmatrix} M_{xw} \\ M_{uw} \end{bmatrix}$$

and

$$w_c = M_w^{-1} \begin{bmatrix} M_{xw} \\ M_{uw} \end{bmatrix}.$$

Then, (3.2) *can be reformulated in terms of norm 2 constraints*

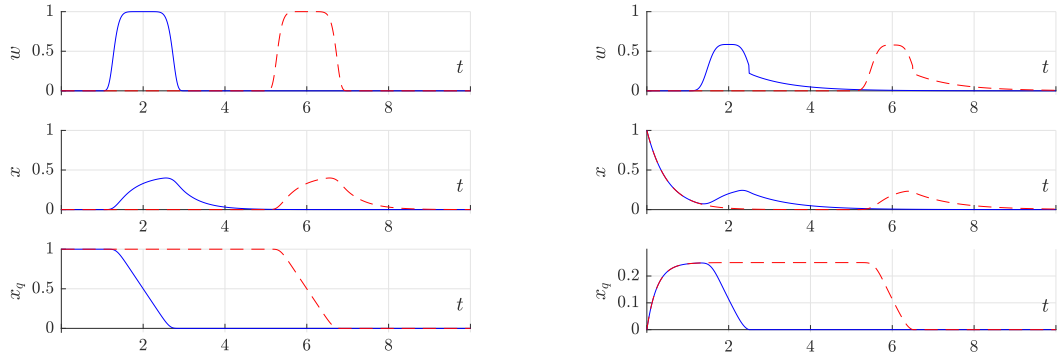
$$\langle w - w_c, R(w - w_c)\rangle_t \;\leq x_{q0} + \langle \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right], Q \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\rangle_t$$

where

$$\langle y, y\rangle_t = \int_0^t y(\tau)^\top y(\tau) d\tau.$$

Let $v = \sqrt{R}(w - w_c)$, *the term* $\langle v, v\rangle_t^{1/2}$ *corresponds to the energy of* v *over* $[0, t]$. *The IQC constraint* (3.1) *can therefore be understood as the energetic constraint*

$$\langle v, v\rangle_t^{1/2} \leq \left(x_{q0} + \langle \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right], Q \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\rangle_t\right)^{1/2}.$$



(a) Unit energy exogenous input. *(b) Unit H_∞ norm feedback.*

Figure 3.1: *Two trajectories of the unit energy response system and the unit H_∞ feedback loop. The second trajectory (dashed red line) is delayed, with a delay of $\tau = 4$, compared to the first trajectory (plain blue line). This is true for any delay τ. Therefore, the state of such a system can be steered away from its equilibrium position at any time in the future.*

3.2 Motivation and examples of IQC systems

IQC models are widely used in robust control theory to assert the stability of dynamical systems. Usually, they model nonlinearities or uncertainty within the model. They have strong connections with fundamental system theory and physical concepts (energy and passivity, see Remark 3.1). IQC models can be used to model physical quantities such as the error of an observation (as in [Savkin and Petersen, 1996b]). They also arise when only the H_∞ gain of a feedback loop is available (see [Megretski, 2010]).

Contrary to examples of QC systems in Section 2.2.2, we could not find a simple expression of the reachable set. Hereby, we represent a few trajectories for two systems of one state dimension. The first system corresponds to a linear stable LTI system disturbed by a unit energy disturbance. The second system corresponds to the same stable LTI system with a feedback loop of unit H_∞ gain.

Unit energy exogenous input: the system described by (3.3) corresponds to a linear stable system disturbed by a unit energy noise.

$$\begin{cases} \dot{x} = -2x + w \\ \displaystyle\int_0^t w(\tau)^2 d\tau \leq 1 \\ x(0) = 0 \end{cases} \tag{3.3}$$

The reachable set of such a system is known to be bounded (see [Boyd et al., 1994, Chapter 6.1.1]). For a trajectory (x, w) of the system (3.3), let (x_τ, w_τ) be the translated signal, i.e. $(x_\tau(t+\tau), w_\tau(t+\tau)) = (x(t), w(t))$, for any $t > 0$, $(x_\tau(t), w_\tau(t)) = (0, 0)$ elsewhere (see Figure 3.1a). Since the system is time-invariant, the translated signal (x_τ, w_τ) is as well a trajectory of (3.3). The state can be steered away from its stable equilibrium at every time $\tau > 0$. Therefore, we expect the reachable set of such an IQC system to be not contractive when $t \to \infty$.

Unit H_∞ gain feedback loop: the system described by (3.4) corresponds to a linear stable system with a unit H_∞ feedback loop.

$$\begin{cases} \dot{x} = -2x + w \\ \displaystyle\int_0^t w(\tau)^2 \leq \int_0^t x(\tau)^2 \\ x(0) = 1 \end{cases} \tag{3.4}$$

Contrary to the system described by (3.3) any translated signal of a system trajectory (x, w) is not necessarily a trajectory of (3.4). However, the same observation applies to the reachable set. For any trajectory, since $x(0) \neq 0$, for a null w over a given interval, the system accumulates energy along time. This energy can then be used at any time in the future (see Figure 3.1b). As for the Unit energy case, we expect the reachable set of such an IQC system to be not contractive when $t \to \infty$.

3.3 Overapproximation of the reachable set with time-varying conics

Theorem 1.1 of Chapter 1 derives a time-varying conic overapproximation to the reachable set of a system subject to a bounded disturbance. The coefficient of the time-varying conic is the solution of a differential equation parametrized by an element of the dual space $\mathcal{D}^*$ of the disturbance set $\mathcal{D}$. In this section, we apply Theorem 1.1 for the IQC system of Definition 3.1. To do so, we first identify the dual space $\mathcal{D}^*$.

The dual space $\mathcal{D}^*$ is characterized by

$$\mu \in \mathcal{D}^*$$

iff

$$\int_0^\infty y(\tau)^\top M(\tau) y(\tau) \mu(\tau) d\tau \geq 0$$

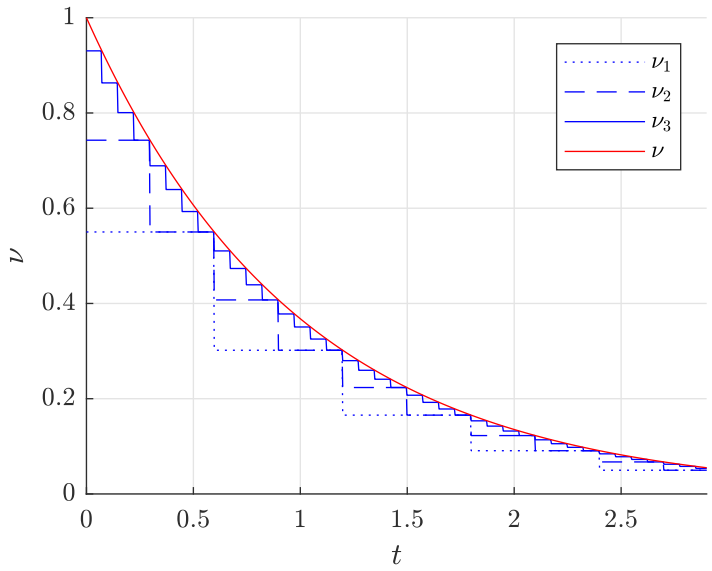


Figure 3.2: *An infinite sum of functions corresponds to the set of positive and decreasing functions $\mathcal{D}^*$.*

for all $y \in \mathcal{D}$. Hereby, a given signal $y = (x, u, w)$ belongs to the disturbance set $\mathcal{D}$ iff $F(t) = \int_0^t f(\tau)d\tau$ is positive for every $t \geq 0$, where $f(\tau) = y(\tau)^\top M(\tau)y(\tau)$. This condition is strictly equivalent to $\langle f, \mu_t \rangle \geq 0$ for every $t \geq 0$ where

$$\mu_t(\tau) = \begin{cases} 1 \text{ if } \tau \leq t \\ 0 \text{ otherwise.} \end{cases}$$

Therefore, $\mu_t \in \mathcal{D}^*$ for any $t \geq 0$. Since the weighted sum of μ_t and $\mu_{t'}$ belongs to $\mathcal{D}^*$, every function ν defined by $\nu(t) = \int_0^\infty \kappa(\tau)\mu_\tau(t)d\tau$, where κ is a positive function over $\mathbb{R}_+$, belongs to $\mathcal{D}^*$ as well (see Figure 3.2). To this respect, we can choose $\mathcal{D}^*$ as the set of functions from $\mathbb{R}_+$ to $\mathbb{R}$ that are positive and decreasing over $\mathbb{R}_+$.

Proposition 3.1. Positive integral duality

Any square-integrable measurable signal $f \in \mathscr{L}_2(\mathbb{R}_+; \mathbb{R})$ is of positive integral over any interval $[0, t]$, $t > 0$, iff

$$\langle f, \mu \rangle \geq 0$$

for any square-integrable measurable μ positive and decreasing over $\mathbb{R}_+$.

Proof. ($\Rightarrow$) if $f(\cdot), \mu(\cdot) \geq 0$ over $\mathbb{R}_+$, then $\langle f, \mu \rangle \geq 0$. ($\Leftarrow$) if there is a $t \geq 0$ s.t. $f(t) \leq 0$, then since f is measurable, there is an interval $I \subset \mathbb{R}_+$ centered over t where $\langle f, 1 \rangle_I \leq 0$, therefore the property holds for $\mu = \delta_I$. ♢

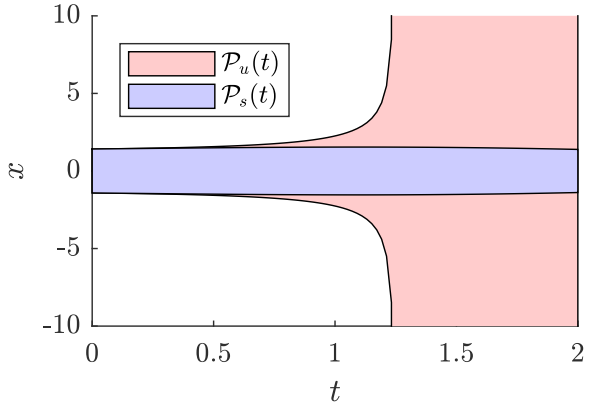


Figure 3.3: *Two overapproximations of the reachable set for the system of Example 3.1. Depending on the values of λ and μ, the overapproximation is bounded over the interval of integration ($\mathcal{P}_s$ in blue) or is diverges in finite-time ($\mathcal{P}_u$ in red).*

The system $\mathcal{S}$ of Definition 3.1 is strictly equivalent to the system of Definition 1.1 where $\mathcal{D}^*$ is the set of square-integrable functions, measurable and positive over $\mathbb{R}_+$. Therefore, the following Theorem 1.1 holds.

Corollary 3.1. Application of Theorem 1.1

The set of reachable states $\mathcal{R}(t;\mathcal{P}_0)$ of $\mathcal{S}$ of Definition 3.1 is overapproximated at any time instant $t \in \mathbb{R}_+$ and for any conic set of initial of states $\mathcal{P}_0 \in \mathbb{P}$ with coefficient $P_0 \in \mathbb{S}^{(n+1)\times(n+1)}$, i.e.

$$\mathcal{R}(t;\mathcal{P}_0) \subseteq \mathcal{P}(t)$$

Example 3.1.

Let $A = -1$, $B = 1$, $M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$, $C = 0$ and $u : [0, \infty[\mapsto 0$. Two overapproximations, as defined in Definition 1.5, are represented in Figure 3.3. The overapproximation $\mathcal{P}_s$ corresponds to the case where $\lambda(t) = exp(0.3t)$ and $\mu(t) = 1$, for all $t \geq 0$. The overapproximation $\mathcal{P}_u$ corresponds to the case where $\lambda(t) = 0$ and $\mu(t) = 1$, for all $t \geq 0$. $\mathcal{P}_s$ is defined over $\mathbb{R}_+$ and $\mathcal{P}_u$ diverges in finite-time.

To assert the quality of overapproximations defined in Corollary 3.1, as for QC systems in Chapter 1, we would like to identify touching trajectories and their associated support conic. To find such trajectories, we should find a valid couple (λ, μ) and an optimal disturbance w such that the level set function associated with the support conic is equal to zero. According to (1.12), it implies that λ is a positive increasing function of $\mathbb{R}_+$ and where μ is a positive and decreasing function over $\mathbb{R}_+$

(i.e. $\mu \in \mathcal{D}^*$), such that $y(t)^\top M(t) y(t) = 0$. It results that $\mu = \mu^*$ where μ^* solves (2.9). However, since μ^* might be increasing, μ^* does not necessarily belong to $\mathcal{D}^*$. Therefore, we are only able to identify a subset of touching trajectories. It is not possible to find all conics supported by the reachable set in a similar way than for the QC system in Chapter 2. The next sections show that it is possible to define a time-varying paraboloid overapproximating and supported by the reachable set $\widetilde{\mathcal{R}}$ of an extended system $\widetilde{\mathcal{S}}$. We show that the intersection of these overapproximations leads to the exact characterization of the reachable set of $\mathcal{S}$.

3.4 Extended system

In the following sections, we study the system $\widetilde{\mathcal{S}}$ which extends $\mathcal{S}$, composed of the state signal x solution of

$$\begin{cases} \dot{x}(\tau) = Ax(\tau) + Bw(\tau) + Cu(\tau) \quad \text{with } \tau \in [0,t] \\ x(0) = x_0 \end{cases} \tag{3.5}$$

with a signal $x_q \in \mathscr{L}_2([0,t];\mathbb{R})$ (corresponding to the IQC constraint) defined for $\tau \in [0,t]$ by

$$x_q(\tau) = x_{q0} + \int_0^\tau \begin{bmatrix} x(s) \\ u(s) \\ w(s) \end{bmatrix}^\top M \begin{bmatrix} x(s) \\ u(s) \\ w(s) \end{bmatrix} ds, \tag{3.6}$$

and that satisfies the state constraint

$$x_q(\tau) \geq 0 \text{ for all } \tau \in [0,t]. \tag{3.7}$$

The constrained dynamical system $\widetilde{\mathcal{S}}(\mathcal{Z}_0, t)$ is then defined for a given set of initial states $\mathcal{Z}_0 \subset \mathbb{R}^n \times \mathbb{R}$ and a terminal time $t > 0$

$$z = (x, x_q) \in \widetilde{\mathcal{S}}(\mathcal{Z}_0, t) \Leftrightarrow \begin{cases} x \text{ solves (3.5)} \\ \text{and } x_q \text{ solves (3.6)} \\ \text{with } (x_0, x_{q0}) \in \mathcal{Z}_0, \text{ and} \\ x_q \text{ satisfies (3.7)} \end{cases} \tag{3.8}$$

Let the reachable set of $\widetilde{\mathcal{S}}(\mathcal{Z}_0, t)$ be

$$\widetilde{\mathcal{R}}(\mathcal{Z}_0, t) = \left\{ z(t) \,\middle|\, z \in \widetilde{\mathcal{S}}(\mathcal{Z}_0, t) \right\}. \tag{3.9}$$

Then, $\widetilde{\mathcal{R}}(\mathcal{Z}_0, t) \subseteq \mathcal{Z}_+$ where $\mathcal{Z}_+ = \mathbb{R}^n \times \mathbb{R}_+$.

3.4.1 Paraboloids

We overapproximate the reachable set $\widetilde{\mathcal{R}}(\mathcal{Z}_0, t)$ of $\widetilde{\mathcal{S}}(\mathcal{Z}_0, t)$ with *paraboloids*.

Definition 3.2. Paraboloid
Given $(E, f, g) \in \mathbb{S}^{n\times n} \times \mathbb{R}^n \times \mathbb{R}$, we define the value function $\widetilde{p}$ by

$$\begin{array}{rccl} \widetilde{p}: & \mathbb{R}^n \times \mathbb{R} & \to & \mathbb{R} \\ & \widetilde{x} = (x, x_q) & \mapsto & p(x) - x_q, \end{array}$$

and the paraboloid:

$$\widetilde{P} = \left\{ \widetilde{x} = (x, x_q) \in \mathbb{R}^{n+1} \,\middle|\, \widetilde{p}(\widetilde{x}) \geq 0 \right\}.$$

Definition 3.3. Scaled paraboloid
For $\mathcal{P} = \texttt{Parab}(E, f, g) \in \widetilde{\mathbb{P}}$ and a scaling factor $\lambda > 0$, let $\lambda\mathcal{P} \in \widetilde{\mathbb{P}}$ be the scaled paraboloid defined by $\lambda\mathcal{P} = \texttt{Parab}(\lambda E, \lambda f, \lambda g)$.

Scaled paraboloids satisfy the following:

Proposition 3.2. Overapproximation relationship of a scaled paraboloid
Given $\mathcal{P} \in \widetilde{\mathbb{P}}$ and $\lambda \geq 1$, it holds $\mathcal{P} \cap \mathcal{Z}_+ \subseteq \lambda\mathcal{P} \cap \mathcal{Z}_+$.

Proof. Let h and h' (resp.) be the value functions of $\mathcal{P} = \texttt{Parab}(E, f, g)$ and $\lambda\mathcal{P}$ (resp.) evaluated at $(x, x_q) \in \mathcal{P}$. Since $(x, x_q) \in \mathcal{P}$, $h \leq 0$, i.e. $x^\top Ex - 2f^\top x + g \leq -x_q$. Then, $h' = \lambda(x^\top Ex - 2f^\top x + g) + x_q \leq -(\lambda - 1)x_q$. Since $(x, x_q) \in \mathcal{Z}_+$ and since $\lambda - 1 \geq 0$, we have $(\lambda - 1)x_q \geq 0$ i.e. $h' \leq 0$ meaning that $(x, x_q) \in \lambda\mathcal{P} \cap \mathcal{Z}_+$. $\diamondsuit$

Remark 3.2. On the paraboloid sets
In what follow, a paraboloid $\mathcal{P}_t = \texttt{Parab}(E_t, f_t, g_t)$ (see Definition 3.2) is used to overapproximate the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ (for a given $\widetilde{\mathcal{P}}_0 \in \widetilde{\mathbb{P}}$) at a given $t > 0$. $\mathcal{P}_t$ is a paraboloid centered around the ray $x_q \mapsto (x_t, x_q)$ with $x_t = E_t^{-1} f_t$ with a summit at (x_t, x_{qs}) where $x_{qs} = f_t^\top E_t^{-1} f_t - g_t$. Since $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t) \subset \mathcal{Z}_+$, depending on the parameters (E_t, f_t, g_t) of $\mathcal{P}_t = \texttt{Parab}(E_t, f_t, g_t)$, the overapproximation $\mathcal{P}_t \cap \mathcal{Z}_+$ might describe an empty, unbounded, convex or not convex set. Studying the subset $\mathcal{P}_t \cap \mathcal{Z}_*$ of $\mathbb{R}^n$ where $\mathcal{Z}_* = \mathbb{R}^n \times \{0\}$ (i.e. the section of the paraboloid in the cone $\mathbb{R}^*_+ = \mathbb{R}^n \times \{0\}$) gives more insight into the shapes of these overapproximations. When $x_{qs} > 0$ and $E_t \succ 0$, $\mathcal{P}_t \cap \mathcal{Z}_*$ is an ellipsoid of quadratic coefficient $(f_t^\top E_t^{-1} f_t - g_t)^{-1} E_t$. When E_t is not sign-definite, $\mathcal{P}_t \cap \mathcal{Z}_*$ is unbounded and not convex. When $E_t \prec 0$ and $x_{qs} \geq 0$, $\mathcal{P}_t \cap \mathcal{Z}_* = \mathbb{R}^n$.

Contrary to ellipsoidal sets used in e.g. [Savkin and Petersen, 1996b, Savkin and Petersen, 1996a, Scherer and Veenman, 2018], overapproximations used

in this work are not necessarily bounded and convex (e.g. when E_t is sign-undefined). Since the reachable set is not always convex (see Example 3.3 and Section 3.8.2), it allows us to define tight overapproximations that are in contact with the reachable set even where the surface of the reachable set is not locally convex.

3.5 Overapproximation with time-varying paraboloids

In this section, we define time-varying paraboloids that overapproximate the reachable set $\widetilde{\mathcal{S}}$. This overapproximation relationship arises from Theorem 1.1 of Chapter 1. These overapproximations are then used in Section 3.6 to define tighter overapproximations of the reachable set. We prove that the overapproximations $\mathcal{P}$ are *tight* since there are touching trajectories of $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ that both belong to the surface of $\mathcal{P}(t)$ and to the surface of $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ for $t \in I$. Finally, the method is presented for a simple example.

To overapproximate the reachable set $\widetilde{\mathcal{R}}$, we choose λ and μ such that $\lambda(t)\,\mu(t) = 1$, for every $t \geq 0$. This constraint is compatible with the necessity for λ to be positive and increasing and μ to be strictly positive and decreasing. Let $\gamma = \dot{\lambda}\,\lambda^{-1}$, by hypothesis over λ, γ can be any positive function from $\mathbb{R}_+$ to $\mathbb{R}_+$.

Definition 3.4. Time-varying paraboloid
For an initial paraboloid $\widetilde{\mathcal{P}}_0 \in \widetilde{\mathbb{P}}$, let the time-varying paraboloid $\mathcal{P}$ be defined as

$$\begin{aligned} \mathcal{P}\colon I &\to \widetilde{\mathbb{P}} \\ t &\mapsto \texttt{Parab}(P(t)) \end{aligned}$$

where the time-varying coefficient P solves

$$0 = \dot{P} + PA_1 + A_1^\top P - \gamma\, P - \, M_{x1} + \left(B_1^\top P - M_{w,x1}\right)^\top M_w^{-1} \left(B_1^\top P - M_{w,x1}\right) \quad (3.10)$$

with the initial condition $P(0) = P_0$, where $\texttt{Parab}(P_0) = \lambda_0 \widetilde{\mathcal{P}}_0$. Let $\mathcal{T}$ be the function that associates to the initial paraboloid $\widetilde{\mathcal{P}}_0 \in \widetilde{\mathbb{P}}$ the time-varying paraboloid $\mathcal{P}$. Let $T_P(\mathcal{P}) = T_E(E_0)$ and $\mathcal{I}(\mathcal{P}) = [0, T_P(\mathcal{P})[$ be the interval of definition of $\mathcal{P}$.

Then, by integration of (1.7) along a system trajectory, for the case described in Definition 3.4, it holds

$$d\widetilde{p}(t, \widetilde{x}) \;\geq \gamma\,(\widetilde{p}(t, \widetilde{x}) + x_q)$$

where $\widetilde{x} = (x, x_q)$. Since $\gamma(\cdot) \geq 0$, it holds that $\widetilde{p}(0, \widetilde{x}(0)) \geq 0$ implies that for any $t \geq 0$, then $\widetilde{p}(t, \widetilde{x}(t)) \geq 0$.

For $\gamma \in \mathscr{L}_{2,\text{loc}}(\mathbb{R}_+;\mathbb{R}_+)$, $\lambda_0 \geq 1$, let $\mathcal{P} = \mathcal{T}(\widetilde{\mathcal{P}}_0, \lambda_0, \gamma)$ be the time-varying paraboloid with time-varying parameters defined by (3.10) for initial conditions defined by $\widetilde{\mathcal{P}}_0$. Equation (3.28) in Appendix 3.A proves that when the quadratic coefficient E diverges in finite-time, at $t^* > 0$, the interval of definition of the time-varying paraboloids, which is $[0, t^*[$ can be as well prolongated to the closed interval $[0, t^*]$.

The worst disturbance, given by (1.14), is associated with the maximal variation of the value function.

Proposition 3.3. Variation of the value function
For $\gamma \in \mathscr{L}_{2,\text{loc}}(\mathbb{R}_+;\mathbb{R}_+)$, $\lambda_0 \geq 1$, and $\widetilde{\mathcal{P}}_0 \in \widetilde{\mathbb{P}}$, let $\mathcal{P} = \mathcal{T}(\widetilde{\mathcal{P}}_0, \lambda_0, \gamma)$. For an optimal trajectory z^* generated by the disturbance w defined in (1.14) *s.t. $z^*(0) \in \partial\widetilde{\mathcal{P}}_0$, for any $t \geq 0$ it holds*

$$\dot{\widetilde{p}}_{z^*}(t) = \gamma(t)(\widetilde{p}_{z^*}(t) - x_q^*(t)). \tag{3.11}$$

Proof. Direct derivation from (3.10). ♢

For any $t \in \mathcal{I}(\mathcal{P})$, the solution to the ODE (3.11) is

$$\widetilde{p}_{z^*}(t) = (1-\lambda_0)(x_q^*(0) - \widetilde{p}^0_{z^*})e^{\int_0^t \gamma(r)dr} - \int_0^t \gamma(s)x_q^*(s)e^{\int_s^t \gamma(r)dr}ds \tag{3.12}$$

where $\widetilde{p}^0_{z^*}$ is the evaluation at $z^*(0)$ of the value function of $\widetilde{\mathcal{P}}_0$.

Equations (3.11) and (3.12) provide a convenient way to a) prove the overapproximation relationship; b) identify touching trajectories; c) reject invalid trajectories.

a) $\underline{\widetilde{p}_z(t) \leq 0}$: for a valid system trajectory z, since $z(0) \in \widetilde{\mathcal{P}}_0$ (i.e. $\widetilde{p}^0_z \leq 0$) and $\underline{\text{(3.7) holds}}$, (3.12) ensures that $z(t) \in \mathcal{P}(t)$ for any t (stated in Corollary 3.2);

b) $\underline{\widetilde{p}_z(t) = 0}$: optimal trajectories z^* as defined in Proposition 3.3 are touching trajectories when $z(0)$ belongs to the boundary of $\widetilde{\mathcal{P}}_0$ and $\gamma(t)x_q^*(t) = 0$, for all $t \geq 0$, and $(1-\lambda_0)x_q^*(t)$ (stated in Proposition 3.4);

c) $\underline{\widetilde{p}_z(t) > 0}$: when z violates the constraint (3.7), there is a $t \geq 0$ s.t. $x_q^*(t) < 0$, one can choose a $\gamma(t) > 0$ such that $\dot{\widetilde{p}}_z(t) \geq 0$ eventually leading to $\widetilde{p}_z(t') > 0$, $t' > t$, proving that $z(t') \notin \widetilde{\mathcal{R}}(t; \widetilde{\mathcal{P}}_0)$ (this will be used in Proposition 3.7).

Intuitively, time-varying paraboloid $\mathcal{P}$ is contracting for valid trajectories (i.e. when $z(0) \in \widetilde{\mathcal{P}}_0$ and (3.7) holds) and expanding for invalid trajectories (i.e. when either $z(0) \in \widetilde{\mathcal{P}}_0$ or either (3.7) is violated).

Since $\int_0^t \gamma(s)x_q^*(s)e^{\int_s^t \gamma(r)dr}$ might not be equal to 0, trajectories generated by the worst-case disturbance w^* do not necessarily stay in contact with the time-varying paraboloid and therefore are not touching trajectories (see Definition 1.3). For this reason, we call *optimal trajectories* the trajectories generated by w^* given by (1.14). When some conditions hold (see below), an optimal trajectory might be a touching trajectory.

Proposition 3.4. Touching optimal trajectories
Let z^* be an optimal trajectory of $\mathcal{P}$ s.t. $z^*(0) \in \partial\widetilde{\mathcal{P}}_0$, if $(1-\lambda_0)x_q^*(0) = 0$, and $\gamma(t)x_q^*(t) = 0$ for any $t \geq 0$, then z^* is a touching trajectory of $\mathcal{P}$.

Proof. Since, $z^*(0) \in \partial\widetilde{\mathcal{P}}_0$, it holds $\widetilde{p}^0_{z^*} = 0$. Using (3.12), we get that $\widetilde{p}_{z^*}(t) = 0$ for all $t \geq 0$. ♢

Theorem 1.1 can be rewritten in the specific IQC case for the extended system $\widetilde{\mathcal{S}}$

Corollary 3.2. Overapproximation of the set of reachable states
For a set of initial states $\widetilde{\mathcal{P}}_0$, a time-varying multiplier $\gamma \in \mathscr{L}_{2,\mathrm{loc}}(\mathbb{R}_+;\mathbb{R}_+)$ and an initial multiplier $\lambda_0 \geq 1$, let $\mathcal{P} = \mathcal{T}(\widetilde{\mathcal{P}}_0, \lambda_0, \gamma)$. The reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ of $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0, t)$, $t > 0$, is overapproximated by $\mathcal{P}(t)$, i.e.

$$\forall t \in \mathcal{I}(\mathcal{P}), \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t) \subseteq \mathcal{P}(t) \cap \mathcal{Z}_+.$$

Proof. For any trajectory z of $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0, t)$, $z(0) \in \widetilde{\mathcal{P}}_0$ implies that $\widetilde{p}_{z^*}(0) \leq 0$, since (3.7) holds, using (3.12), $\widetilde{p}_{z^*}(t) \leq 0$, i.e. $z(t) \in \mathcal{P}(t)$. Any trajectory z satisfies the state constraint (3.7) over $\mathbb{R}_+$, so $z(t) \in \mathcal{Z}_+$, for any $t \in \mathcal{I}(\mathcal{P})$ and therefore $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t) \subseteq \mathcal{Z}_+$. ♢

Example 3.2.
Let $A = -1$, $B = 1$, $M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$, $C = 0$ and $u : [0, \infty[\mapsto 0$. Solutions to IVP (1.21) (that is $\dot{E} = -\frac{1}{2}E^2 + 2E - 1$) diverge when $E_0 \prec E^-$ and converge to E^+ when $E_0 \succ E^-$ (see Figure 3.4) where $E^- \prec E^+$ are the roots of the equation $-\frac{1}{2}E^2 + 2E - 1 = 0$ for $E \in \mathbb{R}$, $E^- = 2 - \sqrt{2}$ and $E^+ = 2 + \sqrt{2}$. Since time-varying paraboloids of Definition 3.4 are defined over the domain of definition of E, for a time-varying paraboloid $\mathcal{P}$ with the initial value (E_0, f_0, g_0), depending on whether E_0 is in the stable region (i.e. $E_0 \succ E^-$) or the unstable region (i.e. $E_0 \prec E^-$), the time-varying paraboloid might be defined over $\mathbb{R}_+$ or a finite horizon only. Figure 3.5 shows the trajectory of the time-varying paraboloid $\mathcal{P}$ for E_0 in the stable region. Figure 3.6 shows the trajectory of the paraboloid for E_0 in the unstable region.

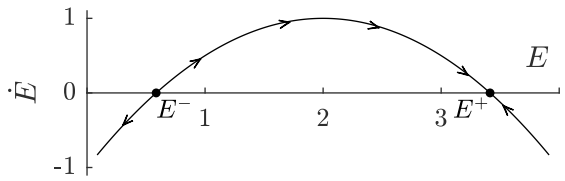


Figure 3.4: *Convergence analysis of the DRE for Example 3.2*

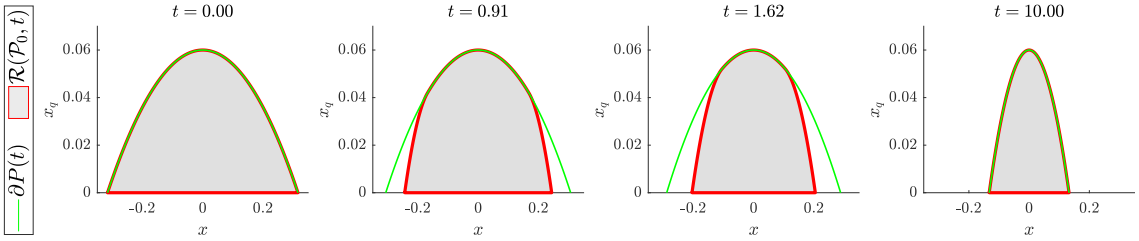


Figure 3.5: *Time-varying paraboloid (its boundary is the green line) overapproximating the reachable set (the gray shaded area with the red boundary) at different time instants t in $\{0.00, 0.91, 1.62, 10.00\}$ for an initial maximum energetic level of $x_{q,0} = 0.06$. The solution to (1.21) converges to a constant value when $t \to +\infty$. The shaded regions are the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ (with $\widetilde{\mathcal{P}}_0 = (E_0, f_0, g_0)$, $E_0 = 0.6$, $f = 0$ and $g_0 = -0.06$), the thin lines are the boundary of the overapproximation $\mathcal{P}(t)$ of Corollary 3.2.*

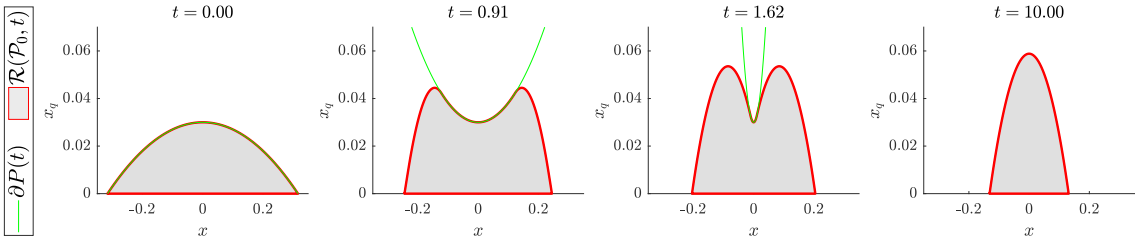


Figure 3.6: *Time-varying paraboloid (its boundary is the green line) overapproximating the reachable set (the grey shaded area with the red boundary) at different time instants t in $\{0.00, 0.91, 1.62, 10.00\}$ for an initial maximum energetic level of $x_{q,0} = 0.03$. The solution to (1.21) has a finite escape time and diverges at $t = 1.68$. The shaded regions are the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ (with $\widetilde{\mathcal{P}}_0 = (E_0, f_0, g_0)$, $E_0 = 0.3$, $f = 0$ and $g_0 = -0.03$), the thin lines are the boundary of the overapproximation $\mathcal{P}(t)$ of Corollary 3.2.*

Remark 3.3. Domain of definition

Previous works in reachability analysis of IQC systems (such as [Savkin and Petersen, 1995, Savkin and Petersen, 1996b, Jönsson, 2002, Seiler et al., 2019]) derive overapproximations as in Section 3.5 but with a scaling function $\gamma(t) = 0$, for all $t \geq 0$. In such a case, the DRE (3.10) *might have a finite escape time depending on the initial set (more precisely, depending on the initial condition E_0 of E) and on the system's parameters. When the initial set belongs to the unstable region, no time-varying paraboloid is defined over $\mathbb{R}_+$ (even if the reachable set is bounded at any time).*

In Example 3.2, the overapproximations are derived using a null scaling function $\gamma(t) = 0$, their domain of definition that depends on the initial condition E_0 of the associated coefficient of E (see the convergence analysis in Figure 3.4). When E_0 is greater than the unstable equilibrium of DRE (1.21)*, E is defined over $\mathbb{R}_+$, otherwise, E has a finite escape time. Such an unstable equilibrium of DRE* (1.21) *exists for stable systems, but it does not exist for unstable systems: the DRE* (1.21) *does not have any equilibrium at all. In this case, solutions E diverge in finite-time for every $E_0 \in \mathbb{R}$. Example 3.3 shows that by using a non-zero scaling function γ in the DRE* (1.21)*, it is possible to find a γ large enough such that the overapproximation is defined over the entire time-interval $\mathbb{R}_+$.*

Thus, our method to overapproximate the reachable set of an IQC system extends the scope of systems that can be studied to unstable systems.

Example 3.3.

Let the system $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0)$ defined by parameters:

$$A = -1,\ B = 1,\ B_u = 0,\ M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -0.9 \end{bmatrix}$$

and a zero input signal u, and let the set of initial states be the paraboloid $\widetilde{\mathcal{P}}_0 =$ $\texttt{Parab}(E_0, f_0, g_0)$ with parameters:

$$E_0 = 1,\ f_0 = 0 \text{ and } g_0 = -0.015.$$

Let the disturbance signal $w = \frac{x}{\sqrt{0.9}}$, w satisfy the IQC (3.7) *since $\int_0^t x^2(\tau) - 0.9w^2(\tau)d\tau = 0$ for every $t \geq 0$. Trajectories associated with such w satisfy $\dot{x} = \alpha x$ with $\alpha = \frac{1}{\sqrt{0.9}} - 1$. Since $\alpha > 0$, every trajectory starting from a non zero initial condition $x(0) = x_0 \neq 0$ diverges when $t \to \infty$, and the system is said unstable.*

Figure 3.7 shows plots of the reachable set of this system and several overapproximations at different time instants. Overapproximations are derived using Definition 3.4 for different scaling functions and initial scaling factors. These

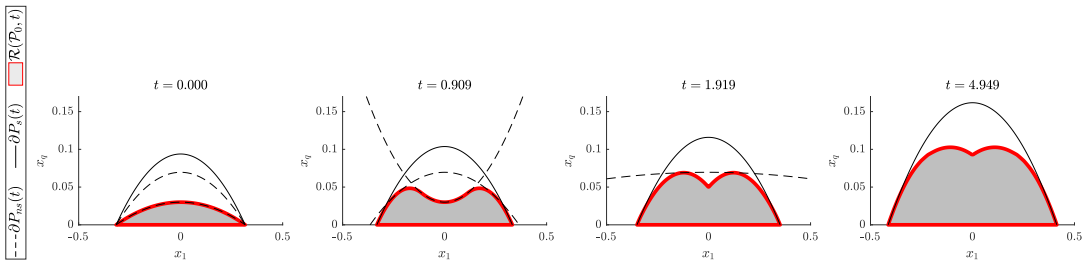


Figure 3.7: *Stabilized constraint $\mathcal{P}_s(t)$ versus finite escape time constraints $\mathcal{P}_{ns}(t)$. The time-varying paraboloid $\mathcal{P}_{ns}$ is defined over $[0, 3]$ whereas $\mathcal{P}_s$ is defined over $\mathbb{R}_+$.*

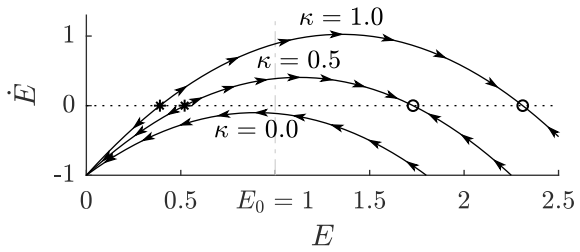


Figure 3.8: *Convergence analysis of the DRE (1.21) for the unstable system defined in Example 3.3 for different scaling functions $\gamma(\cdot) = \kappa$, with $\kappa \in \{0, 0.5, 1\}$. Stable equilibriums are marked with circles, unstable equilibrium are marked with stars. For $\kappa = 0$, the DRE does not have any stable equilibrium, and solutions have a finite escape time. When $\kappa \in \{0.5, 1\}$, the DRE has a stable equilibrium, the solution E with initial condition E_0 converges to this equilibrium and it is defined over $\mathbb{R}_+$.*

overapproximations have different domains of definitions depending on these scalings. The solution to DRE (3.10) for $\gamma(t) = 0$, for all $t \geq 0$, and $\lambda_0 = 0$ has a finite escape time and diverges at $T_P(\mathcal{P}_{ns}) = 1.7$. The solution to DRE (3.10) for $\gamma(t) = 1$, for all $t \geq 0$, and $\lambda_0 = 0$ is defined over $\mathbb{R}_+$.

Their domains of definitions can be studied by conducting a stability analysis of DRE (1.21). Figure 3.9 plots the phase portrait of DRE (1.21) for different constant scaling functions. Figure 3.8 plots the domain of convergence of DRE (1.21) depending on the scaling function γ and the initial condition E_0. For $\gamma = 0$, solutions to the DRE escape in finite-time for every initial value E_0. When $\gamma(t) = \kappa$, for all $t \geq 0$, with $\kappa \geq 0$, for every E_0, there is a value of κ such that the associated solution E converges to a stable equilibrium of the DRE.

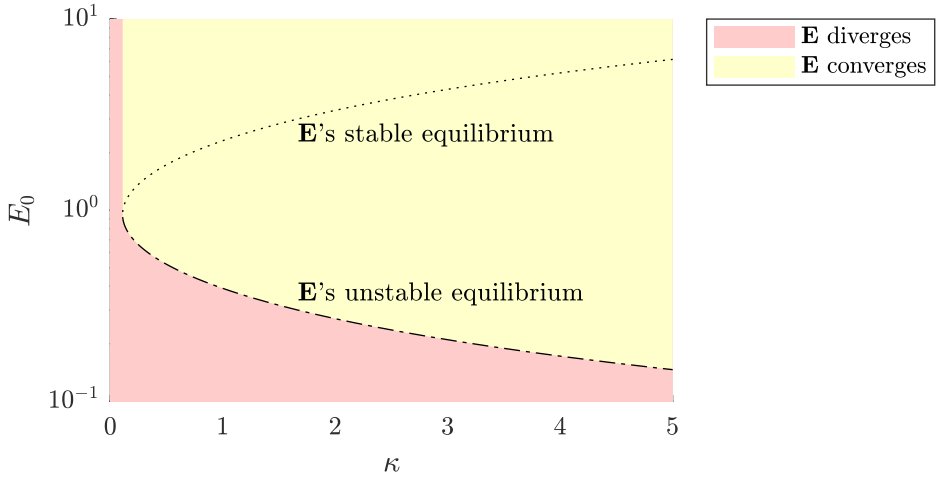


Figure 3.9: *Domain of convergence of the solutions to the DRE* (1.21) *for the unstable system defined in Example 3.3 for different scaling functions $\gamma(\cdot) = \kappa$, and different initial conditions E_0. The red area corresponds to solutions with finite escape time, the yellow area corresponds to solutions converging to a stable equilibrium of the DRE.*

3.6 Exact reachable set

The previous section introduced time-varying paraboloids that overapproximate the reachable set $\widetilde{\mathcal{R}}$ of $\widetilde{\mathcal{S}}$. In this section, for the set of time-varying paraboloid defined in Definition 3.4, at each time instant, we show that the intersection of these paraboloids is an overapproximation of the reachable set (in Section 3.6.2). Then, we prove that when some topological assumption holds about the reachable set, our overapproximation is equal to the reachable set (in Sections 3.6.3, 3.6.4, and 3.6.5).

3.6.1 Overapproximation with an intersection of time-varying paraboloids

In this section, a set of time-varying paraboloids is defined. At a given time, the intersection of the paraboloids gives better overapproximations of the reachable set. With additional assumptions about the topology of the reachable set, the reachable set is exactly characterized. This approach relies on the use of Corollary 3.2 and preliminary results showing that for any state of the overapproximation, there exists a trajectory in $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0, t)$, $t > 0$, leading to this state.

Let $\widetilde{\Pi}^*$ be defined as follows

$$\widetilde{\Pi}^* = \{\mathcal{T}(\widetilde{\mathcal{P}}_0, \lambda_0, \gamma) \,|\, \gamma \in \mathscr{L}_{2,\mathrm{loc}}(\mathbb{R}_+; \mathbb{R}_+), \gamma \geq 0, \lambda_0 \in \mathbb{R}, \lambda_0 \geq 1\}. \tag{3.13}$$

$\widetilde{\Pi}^*$ corresponds to the set of all time-varying paraboloids with initial conditions $\widetilde{\mathcal{P}}_0$, generated by the set of positive time-varying multipliers $\gamma \in \mathscr{L}_{2,\text{loc}}(\mathbb{R}_+;\mathbb{R}_+)$, and the set of initial multipliers $\lambda_0 \geq 1$. Let

$$\widetilde{\Pi}^*(t) = \left\{ \mathcal{P} \in \widetilde{\Pi}^* \,\middle|\, t \in \mathcal{I}(\mathcal{P}) \right\} \tag{3.14}$$

be the set of all the defined time-varying paraboloids at time $t \geq 0$.

For $t \geq 0$, let

$$\widetilde{\Pi}(t) = \bigcap_{\mathcal{P} \in \widetilde{\Pi}^*(t)} \mathcal{P}(t) \tag{3.15}$$

the intersection of all the defined time-varying paraboloids $\mathcal{P}$ of $\widetilde{\Pi}^*$ at time t (see Figure 3.10). Since $\widetilde{\Pi}^*(\cdot)$ is defined over $\mathbb{R}_+$, $\widetilde{\Pi}(\cdot)$ is defined over $\mathbb{R}_+$.

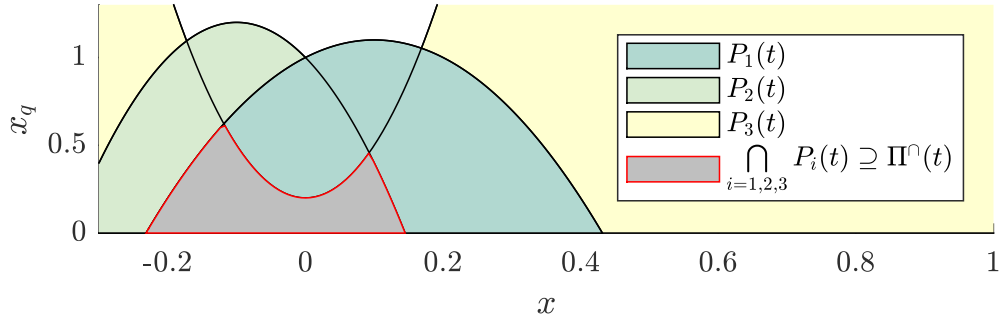


Figure 3.10: *For a given $t \geq 0$, consider three time-varying paraboloids $\mathcal{P}_i \in \widetilde{\Pi}^*$, $i = 1,2,3$. Light color shaded areas are their corresponding parabolic sets $\mathcal{P}_i(t)$ at t, $i = 1,2,3$. The grey color shaded area is their intersection. By (3.15), $\widetilde{\Pi}(t)$ is a subset of $\mathcal{P}_1(t) \cap \mathcal{P}_2(t) \cap \mathcal{P}_3(t)$.*

Since Assumption 1.1 is satisfied in the case of IQC systems (indeed, $t \to exp(-kt)$ is a decreasing and positive function), therefore Proposition 1.5 holds, and the following holds:

Corollary 3.3. Domain of definition of the intersection of time-varying paraboloid

When $E_0 \succ 0$ (i.e. the set of initial states is bounded), $\widetilde{\Pi}$ is defined over $\mathbb{R}_+$.

Proof. By Proposition 1.5. ♢

We now prove that, when some assumptions about the topology of $\widetilde{\Pi}$ hold (Assumption 3.1 and 3.2), we have $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t) = \widetilde{\Pi}(t) \cap \mathcal{Z}_+$, for any $t \geq 0$ (Theorem 3.1, Section 3.6.5). To achieve that:

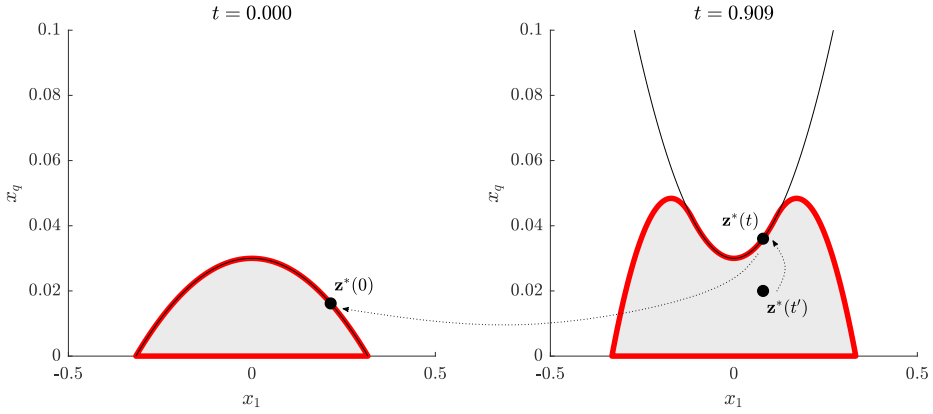


Figure 3.11: *Past trajectory z^* constructed to prove that any state within $\widetilde{\Pi}(t) \cap \mathcal{Z}_+$ belongs to $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ (in grey color shaded area with red boundary). The trajectory z^* connects the endpoint $z(t')$ to trajectory $z(t)$ which is a touching trajectory of some time-varying paraboloid $\mathcal{P} \in \widetilde{\Pi}^*$ (its boundary is represented with the plain black line).*

- we prove the overapproximation relationship $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t) \subseteq \widetilde{\Pi}(t)$ (Section 3.6.2);
- we prove that any state $(x, x_q) \in \widetilde{\Pi}(t)$ is reachable from a state $(x, x_q') \in \partial\widetilde{\Pi}(t)$ with $x_q \leq x_q'$ (Section 3.6.3);
- for a state $z_t \in \partial\widetilde{\Pi}(t)$, we find a touching trajectory $z^* = (x^*, x_q^*)$ of $\widetilde{\Pi}$ such that $z^*(t) = z_t$. This touching trajectory (x^*, x_q^*) of $\widetilde{\Pi}$ satisfies the state constraint $x_q(\cdot) \geq 0$ over $[0, t]$ (Section 3.6.4);
- finally, we conclude that any $z_t \in \widetilde{\Pi}(t)$ is reachable from $\widetilde{\mathcal{P}}_0$, thus $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t) = \widetilde{\Pi}(t) \cap \mathcal{Z}_+$ (Section 3.6.5).

The three last bullet-points characterize a system trajectory z that associates a given state $(x, x_q) \in \widetilde{\Pi}(t)$ to an initial state that belongs to the set of initial states. We illustrate this in Figure 3.11.

3.6.2 Overapproximation relationship

Corollary 3.2 states that each time-varying paraboloid of Definition 3.4 is an overapproximation of the reachable set. An intersection of many time-varying paraboloids is as well an overapproximation of the reachable set.

Proposition 3.5. Overapproximation relationship
$\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t) \subseteq \widetilde{\Pi}(t) \cap \mathcal{Z}_+$ *for any* $t \geq 0$.

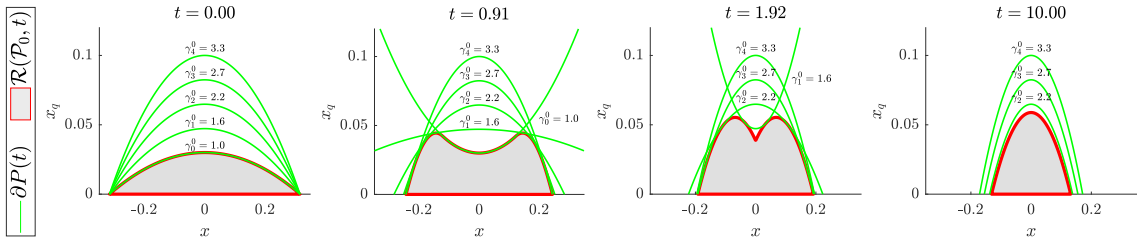


Figure 3.12: *Time-varying paraboloids (its boundary is the green line) overapproximating the reachable set (the grey shaded area with the red boundary) at different time instants t in $\{0.00, 0.91, 1.62, 10.00\}$ for different initial multipliers. Time-varying multipliers (see Definition 3.4) are equal to the null signal, $\gamma_i = 0$, and initial multiplier $\lambda_i^0 \geq 1$ are respectively equal to 1.0, 1.6, 2.2, 2.7 and 3.3 for $i = 0, \ldots, 4$. The shaded regions are the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ (with $\widetilde{\mathcal{P}}_0 = (E_0, f_0, g_0)$, $E_0 = 0.3$, $f = 0$ and $g_0 = -0.03$), the thin lines are the boundary of the overapproximation $\mathcal{P}(t)$ of Theorem 1.1.*

Proof. This is a direct consequence of Corollary 3.2 and (3.15). ♢

Example 3.4.

Example continued from Example 3.2. In the case where the solution to (1.21) *does not converge (i.e. $E_0 < E^-$), Figure 3.12 shows several paraboloid trajectories with different initial multipliers. Time-varying multipliers are equal to the 0 function and initial multiplier λ_i are greater than 1, $\widetilde{\mathcal{P}}_0 \cap \mathcal{Z}_+ \subset \lambda_i \widetilde{\mathcal{P}}_0 \cap \mathcal{Z}_+$. Therefore, each time-varying paraboloid is a valid constraint that bounds $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$, $t \in \mathcal{I}(\widetilde{\Pi}^*)$ (Theorem 1.1). Therefore, $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t) \subseteq \mathcal{P}^*(t) = \mathcal{P}_0(t) \cap \mathcal{P}_1(t) \cap \cdots \cap \mathcal{P}_4(t)$ where $\mathcal{P}_i = \mathcal{T}(\widetilde{\mathcal{P}}_0, \lambda_i, 0)$, and λ_i are resp. equal to 1, 1.6, 2.2, 2.7 and 3.3 for $i = 0, \ldots, 4$. In this case, the overapproximation $\mathcal{P}^*(t)$ is strictly included in $\mathcal{P}_0(t)$.*

Observations in Example 3.4 motivate the use of multiple time-varying paraboloids to get better overapproximations of the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$, $t > 0$.

3.6.3 Past trajectory for states in the interior of the overapproximation

Proposition 3.6 shows that the state $(x, \alpha x_q)$ is reachable from the given state (x, x_q) for any given $\alpha \in [0, 1]$.

Proposition 3.6. Reachability of states in the interior of the overapproximation

For $t \geq 0$, if $(x, x_q) \in \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ then $(x, \alpha x_q) \in \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ for all $\alpha \in [0, 1]$.

Proof. Let $f : t, x \mapsto Ax + Bw(t) + Cu(t)$. Since for any $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^n$, $f(., x)$ is locally measurable over $\mathbb{R}_+$, $f(t, .)$ is Lipschitz over $\mathbb{R}^n$, (3.5) has a unique solution x (see [Schuricht and von der Mosel, 2000], Theorem 1.1) that is time-continuous. Therefore, for a trajectory $(x, x_q) \in \widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0, T)$, $T > 0$, x is time-continuous.

For $\epsilon > 0$, let $w \in \mathscr{L}_2([0, t+\epsilon]; \mathbb{R}^m)$, s.t. $w^\top(s) M_w w(s) = -(1-\alpha) x_q(t) \frac{1}{\epsilon}$ when $s \in [t, t+\epsilon]$. Then

$$\int_t^{t+\epsilon} w^\top(s) M_w w(s) ds \to -(1-\alpha) x_q(t)$$

when $\epsilon \to 0$. Using the Cauchy-Schwartz inequality

$$\left| \int_t^{t+\epsilon} (-M_w)^{\frac{1}{2}} w(s) ds \right| \leq \sqrt{\epsilon} \sqrt{\int_t^{t+\epsilon} -w^T(s) M_w w(s) ds}$$

and the time-continuity of x, the quantity

$$\int_t^{t+\epsilon} \begin{bmatrix} x(s) \\ u(s) \\ 0 \end{bmatrix}^\top M \begin{bmatrix} x(s) \\ u(s) \\ w(s) \end{bmatrix} ds \to 0$$

when $\epsilon \to 0$. By integration, $x_q(t+\epsilon) \to \alpha x_q(t)$ when $\epsilon \to 0$. Since x is time-continuous, $x(t+\epsilon) \to x(t)$ when $\epsilon \to 0$. Since u is bounded at any time ($u \in \mathscr{L}_\infty(\mathbb{R}_+; \mathbb{R}^p)$) and since x is continuous, w is bounded over $[t, t+\epsilon]$. Therefore, u, x and w are bounded over $[t, t+\epsilon]$ x_q is continuous over $[t, t+\epsilon]$. Then, there exists a $t' \in [t, t+\epsilon]$ such that $x_q(\tau) \geq \alpha x_q(t) \geq 0$ for all $\tau \in [t, t']$ and $x_q(t') \to \alpha x_q(t)$ when $\epsilon \to 0$. Therefore, the constraint $x_q(\cdot) \geq 0$ is satisfied over $[t, t']$ and the trajectory (x, x_q) is a valid trajectory of $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0, t')$ for all $t \leq t'$. ♢

3.6.4 Past trajectory for states on the boundary of the over-approximation

In this section, touching trajectories of $\widetilde{\Pi}$ are identified. We show that all these touching trajectories satisfy the state constraint (3.7).

The value function $\tilde{h}$ of a time-varying paraboloid $\tilde{\mathcal{P}} \in \widetilde{\Pi}^*$ can be approximated at the first-order along a touching trajectory z^* of another time-varying paraboloid $\mathcal{P} \in \widetilde{\Pi}^*$ when their time-varying multiplier $\tilde{\gamma}$ and γ and initial multiplier $\tilde{\lambda}_0$ and λ_0 are close to each other. In this part, we compute this first-order approximation when $\tilde{\gamma} = \gamma + \delta$ and $\tilde{\lambda}_0 = \lambda_0 + \delta_0$ for small variations $\delta \in \mathscr{L}_2(\mathbb{R}_+; \mathbb{R})$ and $\delta_0 \in \mathbb{R}$ (i.e. when $\|\delta\| + |\delta_0|$ tends to 0).

To show that the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$, $t > 0$, is exactly described by $\widetilde{\Pi}(t)$, we use a proof by contradiction. For any optimal trajectory z^* of $\mathcal{P} \in \widetilde{\Pi}$ s.t. z^* is violating the constraint (3.7), there is a $\tilde{\mathcal{P}} \in \widetilde{\Pi}^*$ such that the endpoint $z^*(t)$ does not belong to $\tilde{\mathcal{P}}(t)$ and therefore to $\widetilde{\Pi}(t)$. To do so, we will study the value function of a time-varying paraboloid $\tilde{\mathcal{P}}$ for touching trajectories of $\mathcal{P}$.

Proposition 3.7. First-order expansion of the value function
For $\gamma \in \mathscr{L}_{2,\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}_+)$ and $\lambda_0 \geq 1$, let the corresponding time-varying paraboloid $\mathcal{P} = \mathcal{T}(\widetilde{\mathcal{P}}_0, \lambda_0,\ \gamma)$. For any t in the open set of $\mathcal{I}(\mathcal{P})$, it exists $\epsilon > 0$ and $H > 0$ s.t. for any $\delta \in \mathscr{L}_2(\mathbb{R}_+; \mathbb{R})$, $\|\delta\| \leq \epsilon$, for any $\delta_0 \in \mathbb{R}$, $|\delta_0| \leq \epsilon$, $\tilde{\mathcal{P}} = \mathcal{T}(\widetilde{\mathcal{P}}_0, \tilde{\lambda}_0, \tilde{\gamma})$ where $\tilde{\gamma} = \gamma + \delta$ and $\tilde{\lambda}_0 = \lambda_0 + \delta_0$ s.t. t belongs to the open set of $\mathcal{I}(\tilde{\mathcal{P}})$, let $\tilde{h}_t$ be the value function of $\tilde{\mathcal{P}}(t)$ and $z^* = (x^*, x_q^*)$ an optimal trajectory of $\mathcal{P}$, it holds

$$\left|\tilde{h}_t(z^*(t)) - \beta(t)\right| \leq H\epsilon^2$$

where

$$\beta(t) = \frac{\delta_0}{\lambda_0} x_q^*(0) + \int_0^t \delta(s)\psi(s) e^{\int_s^t (\gamma(r)+\delta(r))dr} ds$$

and

$$\psi(s) = x_q^*(s) - \int_s^t \gamma(\tau) x_q^*(\tau) e^{\int_\tau^s \gamma(r)dr} d\tau.$$

Proof. Let (E, f, g) and $(\tilde{E}, \tilde{f}, \tilde{g})$ (resp.) be parameters of $\mathcal{P}$ and $\tilde{\mathcal{P}}$(resp.), and

$$\nu = (E - \tilde{E})x^* - (f - \tilde{f}).$$

Using (3.5), (1.14), and (3.10), ν satisfies the linear time-varying differential equation

$$\dot{\nu}(\tau) = A_\nu(\tau)\nu(\tau) - \delta(\tau)n(\tau). \tag{3.16}$$

with $n = Ex^* - f$ and $A_\nu(\tau) = -A^\top + M_{xw}M_w^{-1}B^\top + \tilde{E}(\tau)BM_w^{-1}B^\top + \gamma(\tau)I$. By Definition 3.4, initial values of $\mathcal{P}$ and $\tilde{\mathcal{P}}$ satisfy

$$\frac{1}{\lambda_0}\mathcal{P}(0) = \frac{1}{\tilde{\lambda}_0}\tilde{\mathcal{P}}(0) = \widetilde{\mathcal{P}}_0,$$

therefore $\nu(0) = \delta_0(E_0x^*(0) - f_0)$ where $(E_0, f_0, g_0) = \widetilde{\mathcal{P}}_0$. Since t belongs to the open set of $\mathcal{I}(\tilde{\mathcal{P}})$, $\tilde{E}(\cdot)$ is bounded over $[0, t]$ (the discontinuity of $\tilde{E}$ can only occur at the final integration time). By time-continuity of $\tilde{E}(\cdot)$ over $[0, t]$, there is a scalar $K > 0$

that bounds $\left\|\tilde{E}(\cdot)\right\|$ over $[0,t]$. Then, since γ is measurable, there exists a measurable function $L \in \mathscr{L}_2([0,t];\mathbb{R}_+)$ such that

$$\|A_\nu(\tau)\| \leq L(\tau) \tag{3.17}$$

over $\tau \in [0,t]$. We can integrate (3.17)

$$\|\nu(t) - \nu(0)\| \leq \int_0^t L(\tau)\,\|\nu(\tau)\|\,d\tau + \int_0^t |\delta(\tau)|\,\|n(\tau)\|\,d\tau.$$

Since $l \mapsto \int_0^l |\delta(\tau)|\,\|n(\tau)\|\,d\tau$ is a non-decreasing function over $[0,t]$, by applying the Grönwall inequality, we get

$$\|\nu(t)\| \leq \left(\|\nu(0)\| + \int_0^t |\delta(\tau)|\,\|n(\tau)\|\,d\tau\right) e^{\int_0^t L(\tau)d\tau}. \tag{3.18}$$

Let

$$q(t) = h_t(z^*(t)) - \tilde{h}_t(z^*(t)). \tag{3.19}$$

z^* is an optimal trajectory of $\mathcal{P}$ s.t. $z^*(t) \in \partial\mathcal{P}(t)$, therefore, $h_t(z^*(t)) = 0$, therefore, using (3.11)

$$h_\tau(z^*(\tau)) = \int_\tau^t \gamma(s) x_q^*(s) e^{\int_\tau^s \gamma(r)dr} ds. \tag{3.20}$$

Using (3.5, 1.14, 3.10, 3.20), q satisfies

$$\begin{aligned}\dot{q}(\tau) = &-\nu^\top(\tau) B M_w^{-1} B^\top \nu(\tau) + \gamma(\tau) h_\tau(z^*(\tau)) \\ &- \tilde{\gamma}(\tau)\tilde{h}_\tau(z^*(\tau)) - \gamma(\tau) x_q^*(\tau) + \tilde{\gamma}(\tau) x_q^*(\tau).\end{aligned}$$

Using (3.19) and (3.20):

$$\begin{aligned}\dot{q}(\tau) = &-\nu^\top(\tau) B M_w^{-1} B^\top \nu(\tau) \\ &+ \delta(\tau)\psi(\tau) + (\gamma(\tau) + \delta(\tau)) q(\tau)\end{aligned} \tag{3.21}$$

where

$$\psi(\tau) = x_q^*(\tau) - \int_\tau^t \gamma(s) x_q^*(s) e^{\int_\tau^s \gamma(r)dr} ds$$

with the initial condition $q(0) = h_0(z^*(0)) - \tilde{h}_0(z^*(0))$. Since z^* is a touching trajectory of $\mathcal{P}$, it holds $h_0(z^*(0)) = 0$, therefore

$$x_q^*(0) = -\lambda_0(x^*(0)^\top E_0 x^*(0) - 2f_0^\top x^*(0) + g_0).$$

Therefore, $\tilde{h}_0(z^*(0))$ satisfies

$$\tilde{h}_0(z^*(0)) = -\frac{\delta_0}{\lambda_0} x_q^*(0)$$

and $q(0) = \frac{\delta_0}{\lambda_0} x_q^*(0)$.

Since t belongs to the open set of $\mathcal{I}(\tilde{\mathcal{P}})$, the optimal trajectory z^* and ν are defined and continuous over $[0, t]$. Moreover, since γ and δ are measurable over $[0, t]$, the solution to the linear time-varying equation (3.21) exists over $[0, t]$ and is

$$q(\tau) = -\tilde{h}_0(z^*(0)) + \int_0^\tau \Big(-\nu(s)^\top B M_w^{-1} B^\top \nu(s) + \delta(s)\psi(s) \Big) e^{\int_s^t (\gamma(r)+\delta(r))dr} ds.$$

Then, using (3.18)

$$\left| q(t) - \frac{\delta_0}{\lambda_0} x_q^*(0) - \int_0^t \delta(s)\psi(s) e^{\int_s^t (\gamma(r)+\delta(r))dr} ds \right| \leq H\epsilon^2$$

with

$$H = tRK^2 N(\|n(0)\|^2 + \|n\|^2) \tag{3.22}$$

a finite constant where $R = \left\| B M_w^{-1} B^\top \right\|$, $K = \exp \int_0^t L(\tau)d\tau$ and $N = \int_0^t e^{\int_s^t (\gamma(r)+\delta(r))dr} ds$. This ends the proof. ♢

Proposition 3.8 gives conditions where the sign of $\tilde{h}_t(z^*(t))$ is only determined by its first-order approximation defined in Proposition 3.7.

Proposition 3.8. Sign of the value function of perturbed trajectories
Let z^* be a touching trajectory of $\mathcal{P} = \mathcal{T}(\widetilde{\mathcal{P}}_0, \lambda_0, \gamma)$ for $\gamma \in \mathscr{L}_{2,\mathrm{loc}}(\mathbb{R}_+; \mathbb{R}_+)$, $\lambda_0 \geq 1$ given and $t \in \mathcal{I}(\mathcal{P})$ given. If there is a $\delta \in \mathscr{L}_{2,\mathrm{loc}}(\mathbb{R}_+; \mathbb{R})$ and a $\delta_0 \in \mathbb{R}$, s.t. $\|\delta\| \leq \epsilon$ and $|\delta_0| \leq \epsilon$ and $t \in \mathcal{I}(\tilde{\mathcal{P}})$ (where $\tilde{\mathcal{P}} = \mathcal{T}(\widetilde{\mathcal{P}}_0, \lambda_0 + \delta_0, \gamma + \delta)$) and

$$H\epsilon^2 \leq \left| \frac{\delta_0}{\lambda_0} x_q^*(0) + \int_0^t \left[\delta(s)\psi(s) e^{\int_s^t (\gamma(r)+\delta(r))dr} \right] ds \right|, \tag{3.23}$$

then the sign of

$$-\frac{\delta_0}{\lambda_0}x_q^*(0) - \int_0^t \left[\delta(s)\psi(s)e^{\int_s^t(\gamma(r)+\delta(r))dr}\right] ds$$

is equal to the sign of $\tilde{h}_t(z^*(t))$ *where* $\tilde{h}_t$ *is the value function of* $\tilde{\mathcal{P}}(t)$ *and* $H > 0$ *defined in* (3.22) *and*

$$\psi(s) = x_q^*(s) - \int_s^t \gamma(\tau)x_q^*(\tau)e^{\int_\tau^s \gamma(r)dr}d\tau.$$

Proof. This is a direct consequence of Proposition 3.7 and of the property: $(|a-b| \leq c) \wedge (c < |b|) \Rightarrow \text{sign}(a) = \text{sign}(b)$ for $a, b, c \in \mathbb{R}$. ♢

Provided the existence of a $(\delta, \delta_0) \in \mathscr{L}_2(\mathbb{R}_+;\mathbb{R}) \times \mathbb{R}$ such that $\gamma + \delta \geq 0$ and $\lambda_0 + \delta_0 \geq 1$, the first-order approximation of the value function of $\tilde{\mathcal{P}} = \mathcal{T}(\widetilde{\mathcal{P}}_0, \lambda_0+\delta_0, \gamma+\delta)$ gives a way to identify time-varying paraboloids $\tilde{\mathcal{P}}$ that belongs to $\widetilde{\Pi}^*$ such that an invalid trajectory with an end state $z_t \in \partial\mathcal{P}(t)$ (meaning with an initial state outside of the initial set $\widetilde{\mathcal{P}}_0$ or a trajectory violating the constraint) does not belong to $\tilde{\mathcal{P}}(t)$ and therefore, does not belong to $\widetilde{\Pi}(t)$.

Proposition 3.9 states that the touching trajectories of $\widetilde{\Pi}$ satisfy the state constraint (3.7). Proposition 3.9 is proven by choosing a valid trajectory candidate. If this trajectory violates the state constraint (3.7), then Proposition 3.8 provides a proof that this trajectory does not belong to the overapproximation $\widetilde{\Pi}$.

Proposition 3.9. Valid touching trajectory
For $\mathcal{P} \in \widetilde{\Pi}$*, if* $z_t \in \partial\widetilde{\Pi}(t)$ *and* $z_t \in \partial\mathcal{P}(t)$ *for* t *in the open set of* $\mathcal{I}(\mathcal{P})$*, then the optimal trajectory* z^* *of* $\mathcal{P}$ *such that* $z^*(t) = z_t$ *is a valid touching trajectory of* $\mathcal{P}$ *and* $\widetilde{\Pi}$*.*

Proof. Let $\psi : \mathbb{R}_+ \mapsto \mathbb{R}$ be defined for $s \geq 0$ by

$$\psi(s) = x_q^*(s) - \int_s^t \gamma(\tau)x_q^*(\tau)e^{\int_\tau^s \gamma(r)dr}d\tau.$$

Let $\tau \in [0, t]$ and $I = [\tau, t]$.

- *Case 1,* $x_q^*(0) < 0$: with $\delta_0 > 0$, using Proposition 3.8, $z^*(t) \notin \tilde{\mathcal{P}}(t)$ where $\tilde{\mathcal{P}} \in \widetilde{\Pi}^*$ since $\delta_0 + \lambda_0 \geq 1$, so $z^*(t) \notin \widetilde{\Pi}(t)$.

- *Case 2, $\psi(\cdot) < 0$ over I*: any $\delta(\cdot) \geq 0$ over I and $\delta(\cdot) = 0$ elsewhere such that $\int_0^t \delta(s)\psi(s)ds \neq 0$ and for $\delta_0 = 0$, using Proposition 3.8, $z^*(t) \notin \tilde{\mathcal{P}}(t)$ where $\tilde{\mathcal{P}} \in \widetilde{\Pi}^*$ since $\gamma + \delta \geq 0$, so $z^*(t) \notin \widetilde{\Pi}(t)$.
- *Case 3, $\psi(\cdot) > 0$ over the open of I and there is a $l \in I$, s.t. $\int_{s\in[\tau,l]} \gamma(s)x_q^*(s)ds \neq 0$*: since $\gamma \geq 0$, there exists a $\delta \leq 0$ such that $\gamma + \delta \geq 0$ and for $\delta_0 = 0$, using Proposition 3.8, $z^*(t) \notin \widetilde{\Pi}(t)$.
- *Case 4, $\psi(\cdot) = 0$ over I*: since x_q^* is continuous over $\mathbb{R}_+$, and since γ is locally square-integrable, $\psi(t) = 0 \Rightarrow x_q^*(t) = 0$, therefore $x_q^*(\cdot) = 0$ over I. Consequently $h_\tau(z^*(\tau)) = 0$ for $\tau \in I$.

Cases 1 to 4 show that for $z^*(t) \in \partial\widetilde{\Pi}(t)$:

- either $\forall l \in I, \int_\tau^l \gamma(\tau)x_q^*(\tau)d\tau = 0$ and $\psi(\cdot) = x_q^*(s) > 0$;
- nor $x_q^*(l) = 0$ for $l \in I$.

Let a partition $[0,t] = \bigcup_{i\in\mathbb{N}} I_i$ be such that over each open interval I_i, $\psi(\cdot) \bowtie_i 0$ with $\bowtie_i \in \{<, >, =\}$. We deduce that for every $s \in [0,t]$, $x_q^*(s) \geq 0$ and $\int_I \gamma(\tau)x_q^*(\tau)d\tau = 0$. z^* is a valid trajectory, i.e. the constraint (3.7) is satisfied. By (3.11), z^* is a touching trajectory of $\mathcal{P}$. Moreover, since $z^*(0) \in \widetilde{\mathcal{P}}_0$, z^* is as well a touching trajectory of $\widetilde{\Pi}$. ♢

Since $\widetilde{\Pi}(t)$ is an intersection of closed sets, $\widetilde{\Pi}(t)$ is closed as well. In the general case, for an infinite intersection $\mathcal{Y}^* = \bigcap_{i\in\mathbb{N}} Y_i$ of closed sets Y_i, $i \in \mathbb{N}$, any boundary point $y \in \partial\mathcal{Y}^*$ does not necessarily belong to the boundary of any Y_i, $i \in \mathbb{N}$ (e.g. $\bigcap_{\epsilon\in]1,2]}[-\epsilon,\epsilon] = [-1,1]$, but there is no $\epsilon \in]1,2]$ such that $1 \in \partial[-\epsilon,\epsilon]$). The following assumption states that for every state on the boundary of the overapproximation $\widetilde{\Pi}(t)$, $t > 0$, there exists a time-varying paraboloid $\mathcal{P}$ such that this state belongs as well to the boundary of the $\mathcal{P}(t)$.

Assumption 3.1. Closedness of $\widetilde{\Pi}(t)$.
For every $z_t \in \partial\widetilde{\Pi}(t)$, there is a $\mathcal{P} \in \widetilde{\Pi}^*$ such that $z_t \in \partial\mathcal{P}(t)$.

This assumption is not a strong one and it is satisfied for simpler cases (see [Rousse et al., 2019, Property 11]).

In Proposition 3.9, the existence of $\tilde{\gamma}$ and λ_0 is conditioned by t belonging to the open domain $\mathcal{I}(\tilde{\mathcal{P}})$; to ensure this, $\|E(\cdot)\|$ is assumed to be bounded over $[0,T]$ (by considering the case where t is in the open set of $\mathcal{I}(\mathcal{P})$). In the general case, the boundedness of $\|E(\cdot)\|$ is not granted (see the unstable case in Example 3.2 and Figure 3.6). Assumption 3.2 states that for any state on the boundary of the overapproximation $\widetilde{\Pi}(t)$, $t > 0$, there is a neighbor state on the boundary of $\tilde{\mathcal{P}}(t)$ where $\tilde{\mathcal{P}}$ is a time-varying paraboloid of $\widetilde{\Pi}^*$ not diverging at t (i.e. t belongs to the interior of $T_P(\mathcal{P})$).

Assumption 3.2. Unbounded time-varying paraboloids
For $t > 0$, for all $\epsilon > 0$, for all $z_t \in \partial\widetilde{\Pi}(t)$ such that $z_t \in \partial\mathcal{P}(t)$, $\mathcal{P} \in \widetilde{\Pi}^*$ with $\mathcal{P}(t)$ unbounded, there is a $\tilde{z}_t$ that belongs to the boundary of $\tilde{\mathcal{P}}(t)$, $\tilde{z}_t \in \partial\tilde{\mathcal{P}}(t)$, such that $\|z_t - \tilde{z}_t\| < \epsilon$.

Lemma 1 shows that all state $z_t \in \partial\widetilde{\Pi}(t)$ (with $t \in \mathcal{I}(\widetilde{\Pi}^*)$ given) is the terminal state of a touching trajectory z^* of $\widetilde{\Pi}$ with an initial state $z^*(0) \in \partial\widetilde{\Pi}(0)$.

Lemma 1. *If Assumption 3.1 and Assumption 3.2 hold, every state $z_t \in \partial\widetilde{\Pi}(t)$ has a past touching trajectory z^* of $\widetilde{\Pi}$ s.t. $z^*(t) = z_t$.*

Proof. Let $z_t \notin \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ such that for any $\epsilon > 0$, there is a touching trajectory $\tilde{z}$ of $\tilde{\mathcal{P}}$, $\tilde{\mathcal{P}}$ finite, with $\tilde{z}(t) \in \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ and $\|\tilde{z}(t) - z_t\| < \epsilon$. For $x_q(t) > 0$, we can define the optimal trajectory z^* with $z^*(t) = z_t$. For all $\tau \in [0, t]$, $\mathcal{P}(\tau)$ is finite. Then, Proposition 3.9 can be used over $[0, \tau]$. Therefore, if $z_t \in \partial\mathcal{P}$ such that $\mathcal{P}$ diverges at t, it holds

$$z_t \in \partial\widetilde{\Pi}(t) \Leftrightarrow z_t \in \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$$

For states not belonging to a diverging time-varying paraboloid, the property is a direct consequence of Assumption 3.1, Proposition 3.9. ♢

Lemma 1 shows that any point on the boundary belongs to the reachable set since, for any given terminal state, we found a past trajectory (the touching trajectory) that satisfies the constraint (3.7) and with an initial condition in the set of initial states.

3.6.5 Exact reachable set

We now state the main result of the chapter.

Theorem 3.1. Exact reachability
When Assumption 3.1 and Assumption 3.2 hold, the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ of system $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0, t)$ (defined in Section 3.4) is equal to the set $\widetilde{\Pi}$ defined in (3.15)*, namely*

$$\widetilde{\Pi}(t) = \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$$

for all $t \geq 0$.

Proof. Corollary 3.2 states that $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t) \subseteq \widetilde{\Pi}(t)$. By Proposition 3.6, for $z_t \in \widetilde{\Pi}(t)$, we can construct a trajectory z such that $z(t) = z_t$, $z(t^-) = z_t^* \in \partial\widetilde{\Pi}(t)$ (Proposition 3.6). Since $z_t^* \in \partial\widetilde{\Pi}(t)$, using Lemma 1, there exists a trajectory z such that $z(t^-) = z_t^*$ and z is a touching trajectory of $\widetilde{\Pi}$ on $[0, t[$. Since z is a touching trajectory of $\widetilde{\Pi}$, $z(0) \in \partial\widetilde{\Pi}(0)$ with $\widetilde{\Pi}(0) = \widetilde{\mathcal{P}}_0 = \widetilde{\mathcal{R}}(0)$. By Proposition 3.8, the trajectory z is valid (i.e. satisfies the energy constraint (3.7)) $z_t \in \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$. ♢

3.7 Implementation

In this part, we discuss the practical implementation of the reachable sets overapproximation using Theorem 3.1. To do so, we compute a subset $\overline{\Pi}^*$ of $\widetilde{\Pi}^*$,

$$\overline{\Pi}^* \subseteq \widetilde{\Pi}^*. \tag{3.24}$$

$\overline{\Pi}^*$ corresponds to the time-varying paraboloid set generated by a finite subset of time-varying multiplier and initial multiplier. Then, the intersection of each time-varying paraboloid evaluated at a given $t > 0$ is an overapproximation of the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$. Finally, the DRE numerical integration is detailed for the case of non-negative solutions to the DRE. We propose an algorithm (Algorithm 3.1) that computes $\overline{\Pi}^*$. Its implementation in Matlab is available online [Rousse, 2019].

Subset of time-varying multipliers and initial multipliers In this work, the time-horizon is partitioned into intervals of width $T_c > 0$ and the construction of the over-approximation is restarted for each interval with an appropriate initial multiplier while the time-varying multiplier is set to the zero function $\gamma(\cdot) = 0$. A multiplier is then described by a sequence of initial multipliers $\{\lambda_k\}_{k\in\mathbb{N}}$, $\lambda_k \geq 1$, $k \in \mathbb{N}$.

In the ideal case, the multipliers (time-varying and initial) would be chosen such that the following property is satisfied

$$\exists \epsilon > 0, \forall \tau \in [t, t+\epsilon], x_q^*(\tau) \geq 0 \tag{3.25}$$

where (x^*, x_q^*) corresponds to the touching trajectory associated with the time-varying multiplier γ and initial multiplier λ_0 and such that $(x^*(t), x_q^*(t)) = (x, x_q)$. In practice, since there might be an infinite number of states (x, x_q) satisfying $\dot{x}_q \geq 0$, only a finite number of states are checked. These states are chosen as projections of a given point in given directions over $\mathcal{Z}_* \cap \partial\overline{\Pi}$ where $\mathcal{Z}_* = \mathbb{R}^n \times \{0\}$. These points are then used to evaluate a range of initial multipliers λ_0 to enforce $\dot{x}_q^*(kT_c) \geq 0$. γ is not used.

Paraboloid numerical integration Consider two paraboloids $\mathcal{P} = \mathcal{T}(\widetilde{\mathcal{P}}_0, 1, \gamma)$, $\tilde{\mathcal{P}} = \mathcal{T}(\widetilde{\mathcal{P}}_0, 1, \tilde{\gamma})$. If $\gamma(.) = \tilde{\gamma}(.)$ over an interval $[0, t_i]$, $t_i > 0$, then $\mathcal{P}(.) = \tilde{\mathcal{P}}(.)$ over $[0, t_i]$. Let $t_i \geq 0$ correspond to the maximal time instant where there is $\tilde{\mathcal{P}} \in \widetilde{\Pi}^*$ such that $\tilde{\mathcal{P}}|_{[0,t_i]} = \mathcal{P}|_{[0,t_i]}$ (i.e. such that the restriction of $\tilde{\mathcal{P}}$ on $[0, t_i]$ is equal to the one of $\mathcal{P}$ on the same interval). And let $t_f \geq 0$ correspond either to the integration horizon $T > 0$, or to the maximal of the interval of definition of $\mathcal{P}$. For implementation purposes, each time-varying paraboloid is defined over the interval $[t_i, t_f] \subseteq [0, T]$.

The solution to $\mathcal{P} = \mathcal{T}(\widetilde{\mathcal{P}}_0, 1, \gamma)$ is then described by parameters (E, f, r) with

$$(E(t), f(t), g(t)) = (E_k(t), f_k(t), g_k(t))$$

for each $t \in [kT_c, (k+1)T_c]$, $k \in \mathbb{N}$.

Cardinal limitation of $\overline{\Pi}^*$ In order to have a tractable integration of the reachable set computation, we limit the cardinality of $\overline{\Pi}^*$ in the following way

- at each time step kT_c, we consider only the N_{new} scaled paraboloids of the largest initial multiplier;
- $\overline{\Pi}^*$ below N_P, oldest time-varying paraboloids are dismissed in benefit of more recent ones;

N_{new} and N_P are user-defined parameters. Choosing the paraboloids with this heuristic showed good results in practice. These rules try to consider only elements of $\widetilde{\Pi}^*$ that are more stable. Since for two solutions E and $\tilde{E}$ of (1.21) respectively defined over $[0,T]$ and $[0,\tilde{T}]$ where $T,\tilde{T} \in \mathbb{R} \cup \{\infty\}$, if $E(0) \preceq \tilde{E}(0)$, then $E(t) \preceq \tilde{E}(t)$ for t in the interval of definition of E and $\tilde{E}$, we have $T \geq \tilde{T}$ (this property follows directly by writing the corresponding value function of the basic LQR optimization problem). Therefore, for a time-varying paraboloid that is positive definite at $t > 0$, its scaled time-varying paraboloid at t will be defined for a longer time horizon.

DRE numerical integration DRE integration is subject to numerical instability. Numerical integration of the DRE (1.21) does not produce good results in practice (see [Kenney and Leipnik, 1985]). Experiments presented in this works make use of the Chandrasekhar method [Lainiotis, 1976]. This method integrates the Ordinary Differential Equation (ODE) (1.21) E using an intermediate ODE over the time-dependent matrix L in $\mathscr{L}_2(\mathbb{R}_+, \mathbb{R}^{n\times n})$ as follow

$$\begin{aligned}\dot{E}(t) &= L(t)L(t)^\top \\ \dot{L}(t) &= (E(t)BM_w^{-1}B^\top - A^\top - M_{xw}M_w^{-1}B^\top)L(t)\end{aligned}$$

with

$$\begin{aligned}E(0) &= E_0 \\ L(0)L(0)^\top &= \dot{E}_0\end{aligned}$$

where $\dot{E}_0 = \dot{E}(0)$ given by (1.21). Then E is a solution to (1.21).

Since $L(t)L(t)^\top \succeq 0$, this method is only applicable to strictly increasing solutions of the DRE. As seen in Example 3.2, the solutions to ODE (1.21) are not strictly increasing over the time horizon, even for a positive definite initial condition. Therefore, the Chandrasekhar method cannot be used directly. We instead use the following approach, let $L, K \in \mathscr{L}_2(\mathbb{R}_+; \mathbb{R}^{n\times n})$ be such as:

$$\begin{aligned}\dot{E}(t) &= L(t)L(t)^\top - K(t)K(t)^\top \\ \dot{L}(t) &= (E(t)BM_w^{-1}B^\top - A^\top - M_{xw}M_w^{-1}B^\top)L(t) \\ \dot{K}(t) &= (E(t)BM_w^{-1}B^\top - A^\top - M_{xw}M_w^{-1}B^\top)K(t)\end{aligned}$$

with

$$L(0)L(0)^\top = \dot{E}_0^+$$
$$K(0)K(0)^\top = -\dot{E}_0^-$$

where $\dot{E}_0 = \dot{E}(0) = \dot{E}_0^+ + \dot{E}_0^-$ given by (1.21), with $\dot{E}_0^+ \succeq 0$ and $\dot{E}_0^- \preceq 0$ The increasing and decreasing parts of E are respectively represented by the terms L and K. Our Chandrasekhar inspired method performs better since the square root term L and K are much smaller than E, and they produce less numerical errors.

For f and g, integration of the ODE as given in (1.22) and (1.23) is used.

Algorithm 3.1 summarizes the computation of $\overline{\Pi}^*$. An implementation on Matlab is available online [Rousse, 2019].

Algorithm 3.1: Computation of $\overline{\Pi}^*$ defined by (3.24), in Section 3.7, as the subset of $\widetilde{\Pi}^*$ defined by (3.14), in Section 3.6.

Inputs :

-A paraboloid $\widetilde{\mathcal{P}}_0 \in \widetilde{\mathbb{P}}$ of initial of states

-A time-horizon of simulation $T > 0$

-Sample time $T_c > 0$ of constraint addition

-The maximum $\mathtt{N_P}$ cardinal of $\overline{\Pi}^*$

Output :

-A set of overapproximating time-varying paraboloids $\overline{\Pi}^*$

Algorithm:

$\overline{\Pi}^* = \{\mathcal{T}(\widetilde{\mathcal{P}}_0, 1, 0)\}$
$t = 0$
$\mathtt{Sim\_Parab} = \{(\widetilde{\mathcal{P}}_0, 0)\}$
while $t < T$ **do**
  `/* Find the new time-varying paraboloids to consider and add them to Sim_Parab (see Algorithm 3.2) */`
  $\mathtt{Sim\_Parab} = Update\_Sim\_Parab(\mathtt{Sim\_Parab}, \overline{\Pi})$
  `/* Simulate the paraboloid for` T_c `*/`
  for $(P_\tau, \tau) \in \mathtt{Sim\_Parab}$ **do**
    Simulate $\mathcal{P}(\cdot)$ over $[t, t + T_c]$ with $\mathcal{P}(\tau) = P_\tau$
    Add $\mathcal{P}(\cdot)$ to $\overline{\Pi}^*$
    if $\mathcal{P}(\cdot)$ *diverges* **then**
      Remove $(\mathcal{P}, t)$ from $\mathtt{Sim\_Parab}$
  $t = t + T_c$

Algorithm 3.2: Computation of the new paraboloids `New_Parab` and update the set of paraboloids `Sim_Parab`.

Parameters:
-Searching directions `Search_Dir` $\subset \mathbb{R}^n$ to add constraints
-The maximum $\mathtt{N_{new}}$ paraboloids to add at each step
Inputs :
-The current overapproximation $\overline{\Pi}_t^*$
-The updated set of paraboloids `Sim_Parab`
Output :
-The updated set of paraboloids to simulate`Sim_Parab`
Algorithm :

```
New_Parab = {}
for n ∈ Search_Dir do
    project x_c on ∂Π̄_t in the direction n
    let x* be this projection and P* ∈ Π̄* its corresponding touching paraboloid
    compute λ̄ given (x*,P*)
    for λ = 1 + dλ, 1 + 2dλ, ..., λ̄ do
        add (P*, λ) to New_Parab
Sort New_Parab according to the values of λ
Keep N_new elements of New_Parab with the largest value of λ
for (P*, λ) ∈ New_Parab do
    add (λP*(t), t) to Sim_Parab
```

3.8 Examples

Algorithm 3.1 deduced from Corollary 3.2 and Theorem 3.1 is used to compute the overapproximation $\widetilde{\Pi}$ defined in (3.24) (subset of $\widetilde{\Pi}^*$ defined in (3.14)) of the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ of the system $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0, t)$ (described in Section 3.4), $t \geq 0$. Several examples are treated. With these examples, we provide some performance evaluations of our approach.

3.8.1 Examples from COMPleib

To evaluate the performance of our approach, we compute an overapproximation of the reachable set for several real-life systems from the COMPleib library [Leibfritz, 2006]. For each system, a stabilizing controller is generated for the generalized plant (by using the `h2syn` function of Matlab), then the system is reduced using a balanced

truncation method to a given state space size (by using the `balred` function of Matlab). The set of initial states is chosen such that the quadratic term belongs to the set of stable solutions to the associated Continuous Algebraic Riccati Equation. The simulation are ran for an input $u(t) = \left[\begin{smallmatrix} 1 & \dots & 1 \end{smallmatrix}\right]^\top \exp(-t)$ for $t \in [0, 2]$. Each ODE is numerically integrated using the `ode113` solver in Matlab. Finally, we run the simulation with one time-varying paraboloid and then multiple time-varying paraboloids. CPU time performances for a computer with an Intel i5 2.5GHz are presented in Table 3.1 and Table 3.2.

In Figure 3.13, we show several runs for the examples. Each paraboloid is over-approximated with a box, we show the intersection of these intervals.

Performance is mainly dependent on the number of paraboloids that we consider, and our ability to efficiently solve the DRE.

System size	Helicopter (HE7)	Aircraft (AC10)	Coupled Spring (CSE1)
5	4.32	4.64	3.65
10	5.12	5.96	3.86
19	7.42	10.62	7.92
30	*n.a.*	28.85	*n.a.*
40	*n.a.*	50.66	*n.a.*
49	*n.a.*	88.00	*n.a.*

Table 3.1: *Computation times (in seconds) of the overapproximation for different systems sizes, using a unique time-varying paraboloid. (When the size of the original system is smaller than the required reduced system size, then the model reduction is not applicable -n.a.-.)*

3.8.2 System verification

We study the stable IQC system $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0, t)$, defined in (3.8), at a given time t in $[0, 1]$, for a parabolic set of initial states $\widetilde{\mathcal{P}}_0 = \texttt{Parab}(E_0, f_0, g_0)$, with $E_0 = \left[\begin{smallmatrix} a+b & a \\ a & a+b \end{smallmatrix}\right]$, $f_0 = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$, $g_0 = -0.015$, $a = 10^{-2}$ and $b = 10^{-6}$, and for the following parameters

$$A = -I,\ B = I,\ C = 0,\ M = \begin{bmatrix} I & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2I \end{bmatrix}$$

where $I = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, and with a zero input signal u.

The reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ of $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0, t)$, defined in (3.9), is computed using (3.13) and Theorem 3.1, for $t \in [0, 1]$. Figures 3.14a and 3.14b show the reachable set

System size	Helicopter (HE7)	Aircraft (AC10)	Coupled Spring (CSE1)
5	83.63 (66)	36.64 (13)	213.88 (232)
10	89.55 (57)	25.77 (9)	261.32 (197)
19	167.53 (52)	27.67 (4)	21.97 (4)
30	*n.a.*	113.72 (7)	*n.a.*
40	*n.a.*	117.60 (4)	*n.a.*

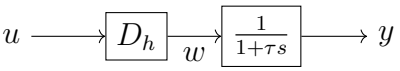
Table 3.2: *Computation times (in seconds) and number of paraboloids (in parenthesis) of the overapproximation for different systems sizes. (When the size of the original system is smaller than the required reduced system size, then the model reduction is not applicable -n.a.-.)*

$\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ set at time $t = 0.794$ and its projection $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)|_x$ over the LTI state space (i.e. projection over (x_1, x_2) states). In Figure 3.14b, the constraints boundaries $\partial\mathcal{P}(t)$ (for $\mathcal{P} \in \widetilde{\Pi}^*$, $\widetilde{\Pi}^*$ defined in Section 3.6) are touching the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$. The non-convexity of $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ arises from the non-positive solutions to the DRE (1.21). Figure 3.14c represents the projection of the reachable tube $t \mapsto \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0, t)$ projected over the LTI dimension (x_1, x_2).

3.8.3 Delayed system

IQC relationships can be derived for delayed systems (see [Seuret and Gouaisbaut, 2015]). The delay operator is a linear time-invariant system that has a state of infinite dimension. The states dynamic can be described by a wave partial differential equation. Projections of the state over a base of Legendre polynomials (of maximal degree $r \in \mathbb{N}$) have a linear dynamic that only depends on smaller degree projection. Moreover, the error between the true state and the projections satisfies energetic constraints (that is derived from Jensen inequality). By increasing the maximal degree r of the considered polynomials, for similar input, the set of reached output is strictly reduced. For each degree r, the IQC falls into the context of this work since $M_w^r \prec 0$.

In the sequel, the reachable set of this overapproximating model is computed and plots of the reachable outputs are given. Consider the following delayed system



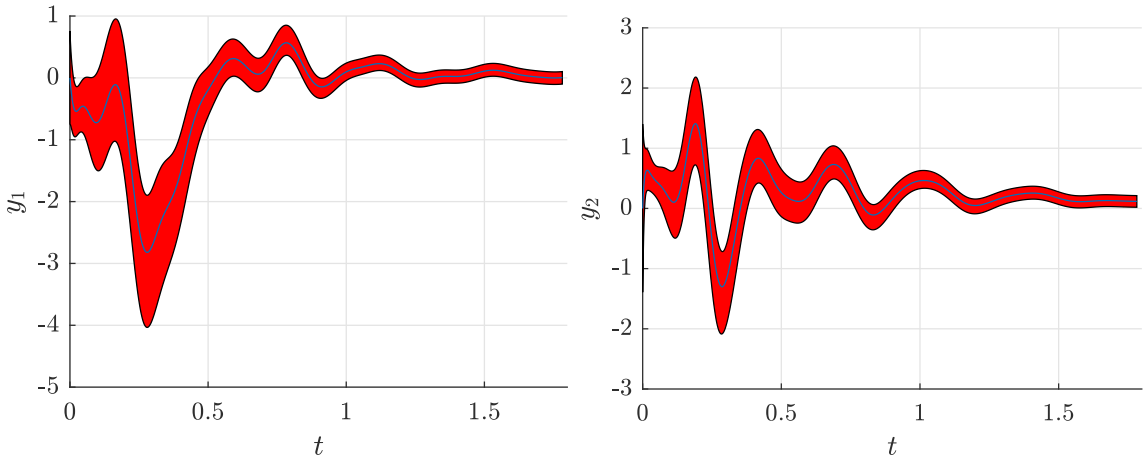
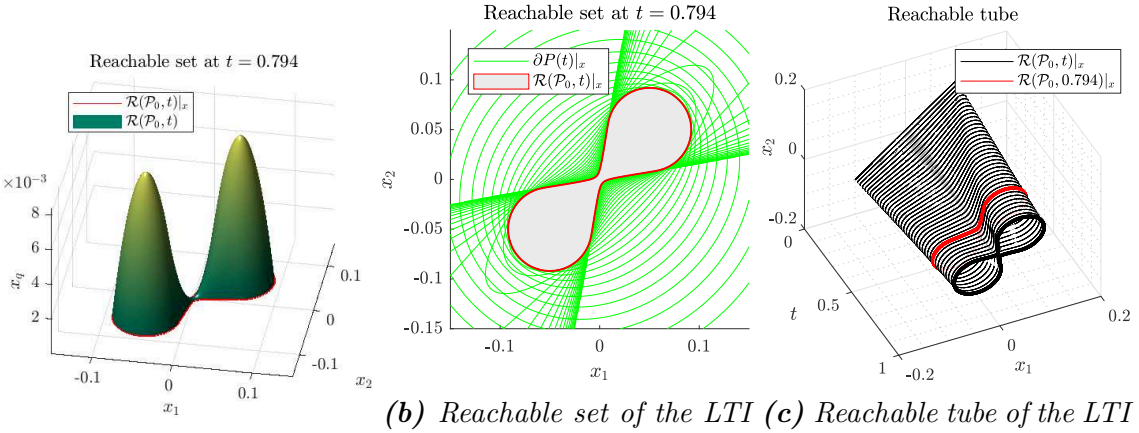


Figure 3.13: *Overapproximation of the output reachable set (projection of the reachable set $\mathbb{R}(t)$ through the observation map; the red area) of the AC10 example from the COMPleib library. The plain black line corresponds to the unperturbed trajectory of the system.*



(a) *Reachable set* ***(b)*** *Reachable set of the LTI system* ***(c)*** *Reachable tube of the LTI system*

Figure 3.14: *The green surface in (a) is the reachable set $\widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0,t)$ at $t = 0.794$ of $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0,t)$ computed using Theorem 3.1. Its projection over the LTI state space (x_1,x_2) (in solid red line) is shown in (b), each green line corresponds to one constraint $\mathcal{P} \in \widetilde{\Pi}^*$ computed with Theorem 1.1. (c) is the reachable tube $t \to \widetilde{\mathcal{R}}(\widetilde{\mathcal{P}}_0,t)$ of $\widetilde{\mathcal{S}}(\widetilde{\mathcal{P}}_0,t)$ projected over the LTI state space (x_1,x_2) for $t \in [0,1]$. The red section corresponds to the time $t = 0.794$.*

where $s \in \mathbb{C}$ is the Laplace variable, $\tau > 0$ and a delay $h > 0$ and the input signal is

defined as follows

$$u(t) = \begin{cases} 0 & \text{if } t < 0 \text{ or } t > 5 \\ 0.3(1 - \frac{t}{5})t + 0.1sin(2\pi t) & \text{otherwise.} \end{cases}$$

Let $\widetilde{\mathcal{S}}^r(\widetilde{\mathcal{P}}_0, T)$ for $g \in \mathbb{N}$ be the relationship of maximal polynomial degree r, and let $\widetilde{\mathcal{R}}^r(\mathcal{Z}_0, t)$ be the associated reachable set at time $t > 0$ for a given set of initial states. The reachable set $\widetilde{\mathcal{R}}^r$ for a set of initial states $\widetilde{\mathcal{P}}_0$ is computed using (3.13) and Theorem 3.1 for different orders r of the hierarchy. Figure 3.15 is a plot of the projections of the reachable set over the output map (where $r = 1, 2, 3$) together with the trajectory of the delayed system. We have the following relationship: $\widetilde{\mathcal{R}}^3(\widetilde{\mathcal{P}}_0, t) \subset \widetilde{\mathcal{R}}^2(\widetilde{\mathcal{P}}_0, t) \subset \widetilde{\mathcal{R}}^1(\widetilde{\mathcal{P}}_0, t)$.

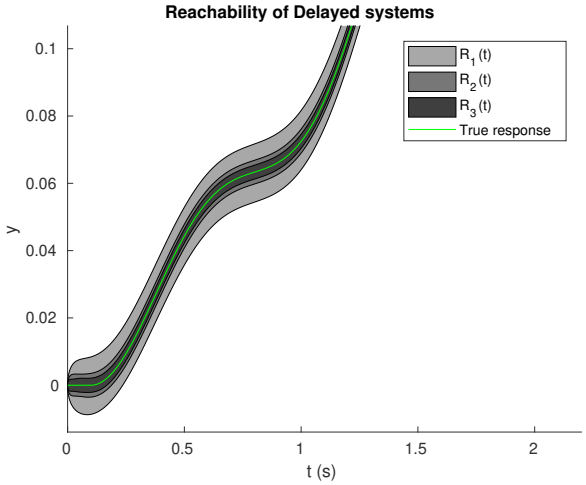


Figure 3.15: *Reachable set computation of the delayed system*

3.9 Discussion

IQC theory was introduced in [Rantzer and Megretski, 1998] as a method to assert robust stability and performance for uncertain LTI systems. It merges the results from [Yakubovich, 1967] and [Zames, 1966b, Zames, 1966a]. The uncertain system is represented as an interconnection between an LTI system with an unknown block that is described by a set of to input-output energy relationships. The Kalman-Yakubowich-Popov lemma gives equivalence between the stability of the interconnection and a Linear Matrix Inequality.

These IQCs are represented in a frequency domain. However, some of them can be equivalently expressed over a finite-horizon in the temporal domain. Such IQCs

are called *hard IQCs* (or *complete IQCs*) and have been studied in [Megretski, 2010]. These IQCs allows to derive bounds over the state at any given time and are therefore of high interest in the verification literature. More recently, [Scherer and Veenman, 2018], showed that any IQCs can always be expressed as a finite-horizon IQC plus a state-dependent quantity. Such an inequality allows to derive bounds over the reachable tube of the uncertain system.

When computing bounds over the reachable set of an IQC system, the IQC is integrated into a storage function. The IQC is scaled with multiplier and then integrated into the storage function. The class of multipliers used to handle IQC constraint is closely related to the method used to solve the DRE. [Jönsson, 2002, Savkin and Petersen, 1995, Savkin and Petersen, 1996a, Savkin and Petersen, 1996b] use constant multipliers, i.e. the IQC constraint is integrated using the S-procedure (see [Boyd et al., 1994], Section 2.6.3). In the study of LTI IQC systems, multipliers can be chosen as positive operators over the signal space (i.e. positive LTI systems, see [Fetzer et al., 2018, Scherer and Veenman, 2018, Yin et al., 2020]). Recursive methods based over integration of the DRE allow to study LTV systems and thus time-varying multipliers.

The existence of an overapproximation over a given time-horizon is dependent on the existence of a solution to a DRE or a DLMI.

The existence of a solution to the DRE has been extensively studied in several works (see a review of the Riccati equation in [Kučera, 1973] and in [Bittanti et al., 1991]). The DRE equation is known to diverge in finite-time depending on its initial condition and on the parameters. [Savkin and Petersen, 1995] provides the existence of a solution to the DRE over any time-horizon for a subclass of IQC. [Jönsson, 2002] gives equivalent conditions between the existence of a solution to the DRE and a full-rank condition over the Hamiltonian.

[Seiler et al., 2019] shows (in Theorem 1) that the DRE is equivalent to a DRI (Differential Riccati Inequality). Contrary to the DRE, the DRI can be expressed as a DLMI by using the Schur complement, thus the existence of a solution to the DRE is then expressed as a feasibility problem of an infinite-dimensional linear problem. The problem can then be solved over a finite basis of functions. An SDP solver can then be used to find the optimal weights and solution of the DRE.

In our work, we provide (for the LTI case) a theorem stating that overapproximations of the reachable always exist over any time-horizon. Such a result is obtained by using time-varying multipliers for the IQC constraint.

3.10 Conclusion

This chapter applied the framework described in Chaper 1 to overapproximate the reachable sets of a linear time-varying systems with an unknown input bounded by integral quadratic constraints, modeling e.g. delay, rate limiter, or energy bounds. We define a family of paraboloidal overapproximations. These paraboloids are supported by the reachable tube on touching trajectories. Parameters of each paraboloid are expressed as a solution to an initial value problem.

Our integration scheme is not guaranteed and the paraboloids we compute are subject to the error of the differential equation numerical integration. The next chapter, Chapter 4, investigates the use of interval arithmetic and validated integration scheme to derive a guaranteed overapproximation to the reachable set of a system subject to an IQC disturbance.

Future work We showed that the reachable set can be described as an intersection of uncountably many paraboloids. In our implementation, a subset of these time-varying paraboloids is computed to overapproximate the reachable set. Then, we compute a minimal volume paraboloid that contains the intersection of all the paraboloids. The computation time of our method is directly dependent on the number of time-varying paraboloids. Finding only one time-varying paraboloid which minimizes its end volume would avoid integrating multiple time-varying paraboloids. Solutions exist for this optimization problem.

The differential Riccati equation can be weakly solved using a basis of polynomial solutions (as in [Seiler et al., 2019]). Then Sum-Of-Square relaxation provides a suboptimal overapproximating paraboloid. Previous works implementing this approach use conservative overapproximations that do not fully incorporate the state constraint. In future works, we will develop such an approach with the results presented in this chapter. A locally optimal solution of the optimization problem can be derived using the Pontryagin's Maximum Principle. We provided such a solution for in the previous chapter, Chapter 2. This could be adapted as well for the IQC case.

In our implementation, the time-varying multipliers and initial multiplier are chosen such that some touching trajectory validate the constraint in the future. Other criteria could be derived such as studying the average behaviors of the trajectories. Since the computational complexity is linear in the number of time-varying paraboloids that need to be simulated, an efficient choice of the multipliers can lead to algorithms that demand less computational resources.

Appendices of Chapter 3

3.A Continuous extension of the domain of definition of the time-varying paraboloids

This part introduces an intermediate result that is used for the proof of Proposition 3.7 in Section 3.6.

By Definition 3.4, the domain of a time-varying paraboloid $\mathcal{P}$ is the domain of its quadratic time-varying coefficient E. Since the solution of the DRE (1.21) might diverge in a finite-time $T_E(E_0) < \infty$ (where E_0 is the initial condition of (1.21)), $\mathcal{P}$ is defined only in the right-open interval $[0, T_E(E_0)[$. In this part, we show that since the touching trajectories of $\mathcal{P}$ are defined over the closed interval $[0, T_E(E_0)]$, i.e. the definition of $\mathcal{P}$ can be prolonged to the same closed interval.

(3.10) can be derived solving the following optimal control problem (for $t > 0$)

$$\begin{aligned} \max_{w \in \mathscr{L}_2([0,t];\mathbb{R}^m)} \quad & \int_0^t \begin{bmatrix} x(\tau) \\ u(\tau) \\ w(\tau) \end{bmatrix} M \begin{bmatrix} x(\tau) \\ u(\tau) \\ w(\tau) \end{bmatrix} d\tau - x_{q,t} \\ \text{s.t.} \quad & \dot{x} = Ax + Bw + Cu \\ & x(t) = x_t \end{aligned}$$

for given $(x_t, x_{q,t}) \in \mathcal{Z}_+$. This is a special instance of the LQR problem (see e.g. [Savkin and Petersen, 1996b]). For $x \in \mathscr{L}_2([0,T];\mathbb{R}^n)$ a touching trajectory, let

$$n = Ex - f. \tag{3.26}$$

Using (3.10), n satisfies the following differential equation

$$\begin{bmatrix} \dot{x} \\ \dot{n} \end{bmatrix} = L \begin{bmatrix} x \\ n \end{bmatrix} + Nu$$

where

$$L = \begin{bmatrix} A - BM_w^{-1}M_{xw}^\top & -BM_w^{-1}B^\top \\ -(M_x - M_{xw}M_w^{-1}M_{xw}^\top) & -A^\top + M_{xw}M_w^{-1}B^\top \end{bmatrix}$$

and

$$N = \begin{bmatrix} C - BM_w^{-1}M_{uw}^\top \\ -(M_{xu} - M_{xw}M_w^{-1}M_{uw}^\top) \end{bmatrix}.$$

The value function evaluated along the touching trajectory x is then obtained by introducing the parameter

$$r = g - f^\top x \tag{3.27}$$

which satisfies

$$\dot{r} = u \begin{pmatrix} H & R \end{pmatrix} \begin{bmatrix} x \\ n \\ u \end{bmatrix}$$

with

$$H = \begin{pmatrix} M_{uw} M_w^{-1} M_{xw} - M_{xu}^\top & -(C - M_{uw} M_w^{-1} B^\top) \end{pmatrix}$$

and

$$R = M_u - M_{uw} M_w^{-1} M_{uw}^\top.$$

Using (3.26) and (3.27), the value function satisfies:

$$\widetilde{p}(t, z(t)) = x(t)^\top n(t) + r(t) + x_q(t).$$

Let the time-varying paraboloid $\mathcal{P} = \mathcal{T}(\widetilde{\mathcal{P}}_0)$ such that $\mathcal{P}$ diverges in finite-time, i.e. $T_P(\widetilde{\mathcal{P}}_0) < \infty$. Since all the touching trajectories are continuous in time, each touching trajectory is defined over $[0, T_P(\widetilde{\mathcal{P}}_0)]$. Their corresponding value function h evaluated along the touching trajectory is as well continuous over $[0, T_P(\widetilde{\mathcal{P}}_0)]$. Therefore, one can extend the definition of $\mathcal{P}$ until $T_P(\widetilde{\mathcal{P}}_0)$ using the continuity of the value function

$$\mathcal{P}(T) = \{z \in \mathbb{R}^{n+1} \mid \lim_{\substack{t \to T \\ t < T}} \widetilde{p}(t, z) \leq 0\}. \tag{3.28}$$

where $T = T_P(\widetilde{\mathcal{P}}_0)$. We state this result in the following property

Proposition 3.10. Continuous extension
For any $\mathcal{P} = \mathcal{T}(\widetilde{\mathcal{P}}_0)$, if the quadratic coefficient of the time-varying paraboloid set $\mathcal{P}$ diverges in finite-time, then the extension to the right of $\mathcal{P}$ exists and is defined by (3.28).

Part II

Interval Analysis Methods

In Part I, we applied the levelset method to overapproximate the reachable tube of a linear time-varying system subject to bounded disturbances. In this part, we propose another approach based on the framework of Interval Arithmetic and validated integration to overapproximate the reachable tube of a nonlinear system subject to a disturbance bounded by an Integral Quadratic Constraint (IQC, see Chapter 3).

The Interval Arithmetic has been introduced in the '50s and '60s as a numerical method to evaluate mathematical expressions while embedding rounding errors. These rounding errors were represented as intervals of values and propagated by structural induction within the mathematical expression. Thus, the exact evaluation of this expression is guaranteed to belongs to this interval despite the use of approximated numerical real arithmetic (e.g. the floating-point arithmetic). The initial Interval Arithmetic was designed to evaluate expressions with additions, subtractions, and multiplications, and was later on extended to divisions, Taylor series and integrals. More complicated problems have been studied, and in particular, root equations and fixed-point equations can as well be addressed within this framework. For such equations, the exact solution is often not computer-representable (it can be an irrational number) and Interval Arithmetic provides a practical tool to find bounds for which the exact solution is guaranteed to lie in.

In dynamical systems analysis, continuous trajectories are solutions to an Initial Value Problem (IVP)

$$\begin{cases} \dot{x}(t) = f(t, x) \\ x(0) = x_0 \end{cases}$$

where $x \in \mathscr{L}_{2,\mathrm{loc}}(I; \mathbb{R}^n)$, with I is the domain definition of the solution x. Such an IVP can be highly sensitive on the numerical error and providing bounds containing the exact solutions is often necessary. This IVP can, in fact, be conveniently and equivalently expressed by the following fixed-point equation:

$$x(t) = x_0 + \int_0^t f(\tau, x(\tau)) d\tau.$$

In this form, it is possible to use the Interval Arithmetic framework to compute a guaranteed solution to the IVP. [Moore et al., 2009] described a method where the trajectory x is represented as a union of interval in the time and state space. Since then, more problems in the field of analysis of dynamical systems have been addressed such as Differential Algebraic Equations (DAE), guaranteed estimations (where the measurement noise is chosen in an interval of value), and, in particular, for system verifications. The interval evaluation allows to model unknown inputs such as unknown initial value (when x_0 belongs to a set of initial states) and input disturbances (such as a disturbance $w(\cdot)$ where $w(t) \in [-1, 1]^m$ at every $t \in I$).

In such a case, the same validated numerical integration method can be applied to compute an overapproximation of the reachable tube of a dynamical system.

In this part, we overapproximate the reachable set of a dynamical system subject to a disturbance bounded by an integral constraint. Contrary to Chapter 3, the system of interest has a non-linear dynamic, and the set of disturbances is described by a non-linear integral constraint.

Non-linear systems can already be studied with IQC models. To do so, the image of the non-linear block is modeled as an unknown but bounded disturbance. Then, the methods presented in Part I can be used for the class of systems studied in this chapter. However, there is a practical difficulty. Getting this IQC model cannot be done automatically out of the dynamical function. Most of the time, the IQC model is obtained manually by identifying the non-linearities in the model and by overapproximating them with known IQCs.

To automatize the analysis of such systems, a possible approach is to study the syntax of the dynamical system models. Usually, the non-linear system is given as a formula involving only a few elementary operations. When these operations can be overapproximated, the resulting expression can be as well overapproximated. The interval arithmetic and its associated validated numerical integration framework use such an approach to overapproximate the reachable tube of a non-linear system. The dynamical function (whether it is linear or not) can be automatically overapproximated by syntactic decomposition over its expression.

In the classical validated numerical integration framework introduced in [Moore et al., 2009], models of interest are usually dynamical systems with an unknown disturbance bounded by an ∞-norm constraint. These frameworks do not take into account disturbances defined by an integral constraint (as presented in Chapter 3). However such models are interesting as they can model complex systems (such as systems with internal delays).

In Chapter 4, we present the classical interval arithmetic framework and the validated numerical integration framework, then, we extend the last to overapproximate the reachable set of non-linear system subject to a disturbance described by a non-linear integral constraint.

Chapter 4

Validated Integration of Nonlinear Systems subject to Integral Constraint

Contents

In this chapter, we present a method to compute the reachable set of a nonlinear dynamical system subject to an unknown disturbance described by an integral constraint between the disturbance and the state trajectory. The Interval Arithmetic and validated numerical integration frameworks are used. With additional assumptions about the dynamic of the disturbance, the integral constraint gives bounds over the set of disturbances. A contractor over the set of reachable states is defined out of these bounds. This contractor is then used in a fixed point algorithm with a propagation step (as described in [Alexandre dit Sandretto and Chapoutot, 2016]). Our algorithm is implemented using the DynIbex library [Dit Sandretto and Chapoutot, 2016] and applied to overapproximate the reachable tube of a dynamical system with

an internal delay.

In Section 4.1, we describe the basics of *interval analysis*. In Section 4.2, we give a short introduction of *validated numerical integration*. In Section 4.3, we present the nonlinear system with disturbances subject to an integral constraint. In Section 4.4, we extend the validated numerical integration presented in Section 4.2. In Section 4.5, we present an example and compare it to the results obtained in Chapter 3. In Section 5.10 we conclude this chapter.

4.1 Interval arithmetic

A simple and common way to represent and manipulate sets of values is *interval arithmetic* (see [Moore et al., 2009]). An interval $[\mathbf{x_i}] = [\underline{x_i}, \overline{x_i}]$ defines the set of reals x_i such that $\underline{x_i} \leq x_i \leq \overline{x_i}$. $\mathbb{IR}$ denotes the set of all intervals over reals. The size (or width) of $[\mathbf{x_i}]$ is denoted by $w([\mathbf{x_i}]) = \overline{x_i} - \underline{x_i}$.

Interval arithmetic extends to $\mathbb{IR}$ elementary functions over $\mathbb{R}$. For instance, the interval sum, i.e., $[\mathbf{x_1}]+[\mathbf{x_2}] = [\underline{x_1}+\underline{x_2}, \overline{x_1}+\overline{x_2}]$, encloses the image of the sum function over its arguments.

An interval vector or a *box* $[\mathbf{x}] \in \mathbb{IR}^n$, is a Cartesian product of n intervals. The enclosing property defines what is called an *interval extension* or an *inclusion function*.

Definition 4.1. Inclusion function

Consider a function $f : \mathbb{R}^n \to \mathbb{R}^m$, then $[f] : \mathbb{IR}^n \to \mathbb{IR}^m$ is said to be an inclusion function of f to intervals if

$$\forall [\mathbf{x}] \in \mathbb{IR}^n, \quad [f]([\mathbf{x}]) \supseteq \{f(\mathbf{x}), \mathbf{x} \in [\mathbf{x}]\} \ .$$

It is possible to define inclusion functions for all elementary functions such as $\times$, $\div$, sin, cos, exp, etc. The *natural* inclusion function is the simplest to obtain: all occurrences of the real variables are replaced by their interval counterpart and all arithmetic operations are evaluated using interval arithmetic. More sophisticated inclusion functions such as the centered form, or the Taylor inclusion function may also be used (see [Jaulin et al., 2001] for more details).

Example 4.1.

A few examples of arithmetic operations between interval values are given

$$\begin{aligned}
[-2,5]+[-8,12] &= [-10,17]\\
[-10,17]-[-8,12]=[-10,17]+[-12,8] &= [-22,25]\\
[-10,17]-[-2,5] &= [-15,19]\\
\frac{[-2,5]}{[-8,12]} &= [-\infty,\infty]\\
\frac{[3,5]}{[8,12]} &= \left[\frac{3}{12},\frac{5}{8}\right]\\
\left[\frac{3}{12},\frac{5}{8}\right]\times[8,12] &= \left[2,\frac{15}{2}\right].
\end{aligned}$$

In the first example of division, the result is the interval containing all the real numbers because the denominator contains 0.

As an example of inclusion function, we consider a function p defined by

$$p(x,y)=xy+x\ .$$

The associated natural inclusion function is

$$[p]([\mathbf{x}],[\mathbf{y}])=[\mathbf{x}][\mathbf{y}]+[\mathbf{x}],$$

in which variables, constants and arithmetic operations have been replaced by its interval counterpart. And so
$p([0,1],[0,1])=[0,2]\subseteq\{p(x,y)\mid x,y\in[0,1]\}=[0,2]$. ■

4.1.1 System of equations

The Interval Arithmetic framework can be used to find bounds containing the solutions of a set of equations. In Section 4.1.1, we present a method to solve a fixed-point equation. In Section 4.1.1, we present a method to solve a system of equations that is not given in the form of a fixed point equation.

Fixed-point equation To solve the fixed-point equation

$$x=f(x) \tag{4.1}$$

where $f:\mathbb{R}^n\mapsto\mathbb{R}^n$, it is possible to compute the fixed-point iterations

$$x_{k+1}=f(x_k)$$

for $x_0 \in \mathbb{R}^n$. If the self-map f is contractive, i.e. if

$$\|f(x) - f(y)\| \leq k\|x - y\|$$

where $0 \leq k < 1$, then

- the equation (4.1) has a unique fixed point x^*, and
- the sequence $\{x_k\}_{k\in\mathbb{N}}$ converges to this fixed point x^*.

This result is known as the Banach Theorem (Section 1.6 in [Zeidler, 1995a]). It gives a practical algorithm to compute an approximate solution to (4.1).

This iteration method has been adapted to the framework of interval arithmetic (e.g. in Chapter 6 of [Moore et al., 2009]). Let $[f]$ be the interval evaluation of f. If $[f]$ is contractive (with respect to the norm $dist$[1]), i.e. for every $[\mathbf{x}] \in \mathbb{IR}^n$, it holds $dist\,([f]([\mathbf{x}]), [f]([\mathbf{y}])) \subseteq k\,dist\,([\mathbf{x}], [\mathbf{y}])$, where $0 \leq k < 1$, then the sequence of $[\mathbf{x_{k+1}}]$ defined by

$$\begin{cases} [\mathbf{x_{k+1}}] = [f]([\mathbf{x_k}]) \\ [\mathbf{x_0}] \in \mathbb{IR}^n \end{cases} \tag{4.2}$$

converges to the singleton $\{x^*\}$, i.e.

$$[\mathbf{x_k}] \to \{x^*\}. \tag{4.3}$$

The convergence property (4.3) is not satisfying since the sequence of $[\mathbf{x_k}]$ are not guaranteed to contain the fixed point solution x^*. Such a result can be obtained by observing that

$$x^* \in [\mathbf{x}] \Rightarrow x^* \in [f]([\mathbf{x}])$$

for every $[\mathbf{x}] \in \mathbb{IR}^n$. Then, the interval version of the iteration method can be applied for the case where

- $[f]$ is contractive (with respect to the set inclusion), i.e. $[f]([\mathbf{x}]) \subseteq [\mathbf{x}]$, and
- $[\mathbf{x_0}]$ contains the solution x^* to the fixed-point equation (4.1).

in such a case, the sequence of $[\mathbf{x_k}]$ defined by

$$[\mathbf{x_{k+1}}] = [f]([\mathbf{x_k}]) \tag{4.4}$$

satisfies

$$x^* \in [\mathbf{x_{k+1}}] \subseteq [\mathbf{x_k}] \subseteq [\mathbf{x_{k-1}}] \subseteq \cdots \subseteq [\mathbf{x_0}].$$

[1] where $dist([\mathbf{x}], [\mathbf{y}]) = \sup_{x\in[\mathbf{x}], y\in[\mathbf{y}]} \|x - y\|$.

Thus, each iterate $[\mathbf{x_k}]$ is a sound approximation of x^*.

When the contractive property is not available, the operator $[f]$ can be enforced to be contractive by using the operator $\overline{[f]}$ defined by $\overline{[f]}([\mathbf{x}]) = [f]([\mathbf{x}]) \cap [\mathbf{x}]$.

This iterative approach to solve the fixed point equation (4.1) has been improved by using other existing fixed-point algorithm. Chapter 8 of [Moore et al., 2009] uses a Newton gradient descent to improve the contractive properties of $[f]$.

Example 4.2.

Let f be a univariate function $f : \mathbb{R} \mapsto \mathbb{R}$ defined by

$$f(x) = 0.5(2 - x). \tag{4.5}$$

The solution x^* to the fixed point equation (4.1) is $x^* = \frac{2}{3}$. We compute the fixed-point iterates $[\mathbf{x_k}]$ defined by (4.4) starting from $[\mathbf{x_0}] = [0.5, 1]$. The first iterates are (see Figure 4.1)

$$\begin{aligned}
[\mathbf{x_0}] &= [0.500, 1.000] = 0.750 \pm 5.00e-01 \\
[\mathbf{x_1}] &= [0.500, 0.750] = 0.625 \pm 2.50e-01 \\
[\mathbf{x_2}] &= [0.625, 0.750] = 0.688 \pm 1.25e-01 \\
[\mathbf{x_3}] &= [0.625, 0.688] = 0.656 \pm 6.25e-02 \\
[\mathbf{x_4}] &= [0.656, 0.688] = 0.672 \pm 3.12e-02
\end{aligned}$$

The iteration sequence converges to $[\mathbf{x}_\infty] = 0.666666 \pm 4e-6$. The error is due to numerical imprecision introduced by the floating-point arithmetic.

Contractors When the equation to solve can be represented as a fixed point equation, the previous section proposes iterative methods that can refine an a priori bound containing a solution of this equation. When the equation is not a fixed point equation, it is possible to use similar concepts: the contractor. A contractor is an operator that associates to every given set, a subset that contains all the points where the constraint is verified (see [Chabert and Jaulin, 2009]).

Definition 4.2. Contractor

For a constraint f that maps $\mathbb{R}^n$ to a truth value, a contractor Ctc of f associates to a subset of $\mathbb{R}^n$ to a subset of $\mathbb{R}^n$. For any $[\mathbf{b}], [\mathbf{b}'] \in \mathbb{IR}^n$, Ctc must verify the following properties:

- *the contraction: $Ctc([\mathbf{b}]) \subseteq [\mathbf{b}]$,*
- *the conservativeness: $\forall x \in [\mathbf{b}] \setminus Ctc([\mathbf{b}]), f(x)$ is not satisfied,*
- *the monotonicity: $[\mathbf{b}'] \subseteq [\mathbf{b}] \Rightarrow Ctc([\mathbf{b}']) \subseteq Ctc([\mathbf{b}])$*

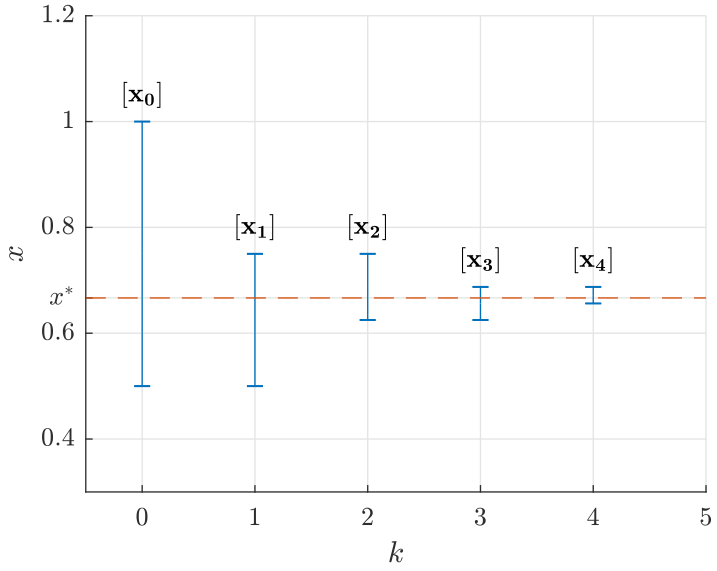


Figure 4.1: *Sequence of* $[\mathbf{x_k}]$ *defined by* (4.4) *overapproximating the fixed point solution* x^* *of* (4.1) *with* f *defined in* (4.5) *for* $[\mathbf{x_0}] = [0.5, 1]$.

As for the fixed point equation and thanks to the contraction and monotonicity property, if a solution x^* belongs to the truth value of f, then every sequence of $[\mathbf{x_k}]$ defined by

$$[\mathbf{x_{k+1}}] = Ctc\left([\mathbf{x_{k+1}}]\right),$$

such that $x^* \in [\mathbf{x_0}]$, satisfies

$$x^* \in [\mathbf{x_{k+1}}] \subseteq [\mathbf{x_k}] \subseteq [\mathbf{x_{k-1}}] \subseteq \cdots \subseteq [\mathbf{x_0}].$$

The precision and the speed of convergence of the sequence of $[\mathbf{x_k}]$ depends on the contraction of Ctc.

Example 4.3.

To find the set of x *such that* $f(x) \geq 0$ *where*

$$f(x) = 2 - x^4 \tag{4.6}$$

it is possible to define a contractor that exploit the concavity of f *(see Figure 4.2). The resulting contractor produces a decreasing (in the set inclusion sens) sequence of iterates* $[\mathbf{x_k}]$ *for an initial* $[\mathbf{x_0}] = [-2, 2]$. *The sequence converges to* $[\mathbf{x}_\infty] = [-1.189207, 1.189207]$ *(see Figure 4.3).*

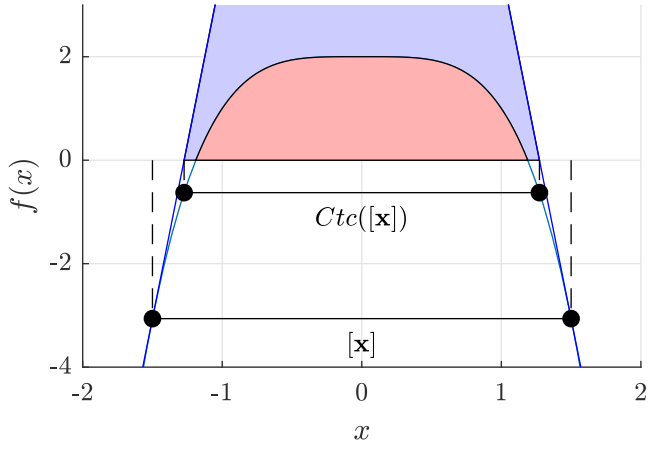


Figure 4.2: *The red area corresponds to the set of points $(x, f(x))$ where $f(x) \geq 0$. The blue area corresponds to the set of points under the linear approximation of the concave function f at the boundaries of $[\mathbf{x}]$.*

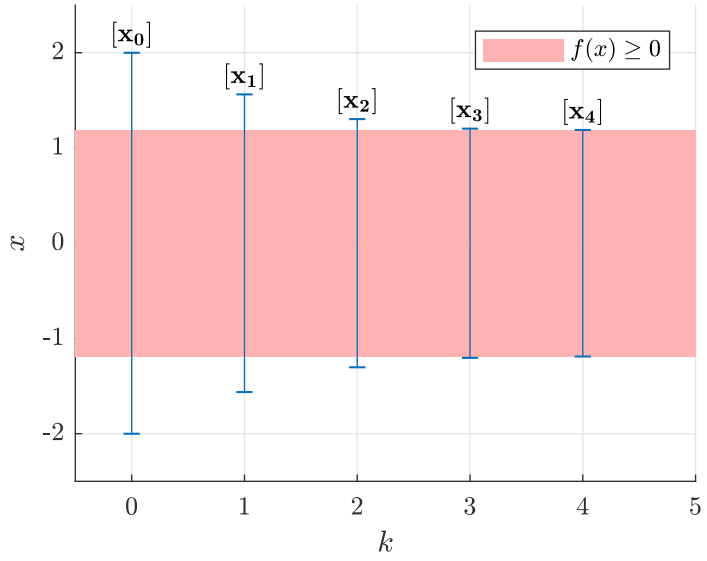


Figure 4.3: *The red area corresponds to the set of points x where $f(x) \geq 0$..*

4.2 Validated numerical integration methods

The Initial Value Problem (IVP)

$$\begin{cases} \dot{\mathbf{y}}(t) = F(t, \mathbf{y}(t)) \\ \mathbf{y}(0) = y_0 \end{cases} \tag{4.7}$$

can be equivalently expressed with the fixed point equation

$$\mathbf{y} = \Phi(\mathbf{y}) \tag{4.8}$$

where

$$\Phi(\mathbf{y})(t) = y_0 + \int_0^t F(s, \mathbf{y}(s))ds$$

with $F : \mathbb{R}_+ \times \mathbb{R}^n \mapsto \mathbb{R}^n$, $y_0 \in \mathbb{R}^n$, and with the unknown $\mathbf{y} : \mathbb{R}_+ \mapsto \mathbb{R}^n$. The function $\mathbf{y} : \mathbb{R}_+ \mapsto \mathbb{R}^n$ can be represented with a union of intervals in the time and state space, i.e.

$$\{(t, \mathbf{y}(t)) \,|\, t > 0\} \in \bigcup_k [\widetilde{\mathbf{y}}_k]$$

where $[\widetilde{\mathbf{y}}_k] \in \mathbb{IR}^{n+1}$ (see Figure 4.4). With such a representation, one can solve the fixed point equation (4.8) using the fixed point iteration proposed in Section 4.1.1. This section details this approach for a more general IVP than (4.7).

Mathematically, differential equations have no explicit solutions, except for a few particular cases. Nevertheless, the solution can be numerically approximated with the help of integration schemes such as Taylor series [Nedialkov et al., 1999] or Runge-Kutta methods [Sandretto and Chapoutot, 2016, Dit Sandretto and Chapoutot, 2016].

In the following, we consider a *generic* parametric differential equation as an *interval initial value problem* (IIVP) defined by

$$\begin{cases} \dot{\mathbf{y}} = F(t, \mathbf{y}, \mathbf{x}, \mathbf{p}, \mathbf{u}) \\ 0 = G(t, \mathbf{y}, \mathbf{x}, \mathbf{p}, \mathbf{u}) \\ \mathbf{y}(0) \in \mathcal{Y}_0, \mathbf{x}(0) \in \mathcal{X}_0, \mathbf{p} \in \mathcal{P}, \mathbf{u} \in \mathcal{U}, t \in [0, t_{\text{end}}] \ , \end{cases} \tag{4.9}$$

with $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^r \times \mathbb{R}^s \mapsto \mathbb{R}^n$ and $G : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^r \times \mathbb{R}^s \mapsto \mathbb{R}^m$. The vector variable $\mathbf{y}$ of dimension n is the differential variable while the variable $\mathbf{x}$ is an algebraic variable of dimension m with an initial condition $\mathbf{y}(0) \in \mathcal{Y}_0 \subseteq \mathbb{R}^n$ and $\mathbf{x}(0) \in \mathcal{X}_0 \subseteq \mathbb{R}^m$. In other words, differential-algebraic equations (DAE) are considered, and in the case of $m = 0$, this differential equation simplifies to an ordinary differential equation (ODE). Note that usually, the initial values of algebraic variable

$\mathbf{x}$ are computed by numerical algorithms used to solve DAE but we consider it fixed here for simplicity. Variable $\mathbf{p} \in \mathcal{P} \subseteq \mathbb{R}^r$ stands for parameters of dimension r and variable $\mathbf{u} \in \mathcal{U} \subseteq \mathbb{R}^s$ stands for a control vector of dimension s. We assume standard hypotheses on F and G to guarantee the existence and uniqueness of the solution to such a problem.

A validated simulation of a differential equation consists of a discretization of time, such that $t_0 \leqslant \cdots \leqslant t_{\text{end}}$, and a computation of enclosures of the set of states of the system $\mathbf{y}_0, \ldots, \mathbf{y}_{\text{end}}$, by the help of a validated integration scheme. In details, a validated integration scheme is made of

- an integration method $\Phi(F, G, \mathbf{y}_j, t_j, h)$, starting from an initial value $\mathbf{y}_j$ at time t_j and a finite time horizon h (the step-size), producing an approximation $\mathbf{y}_{j+1}$ at time $t_{j+1} = t_j + h$, of the exact solution $\mathbf{y}(t_{j+1}; \mathbf{y}_j)$, i.e., $\mathbf{y}(t_{j+1}; \mathbf{y}_j) \approx \Phi(F, G, \mathbf{y}_j, t_j, h)$;

- a truncation error function $\text{lte}_\Phi(F, G, \mathbf{y}_j, t_j, h)$, such that

$$\mathbf{y}(t_{j+1}; \mathbf{y}_j) = \Phi(F, G, \mathbf{y}_j, t_j, h) + \text{lte}_\Phi(F, G, \mathbf{y}_j, t_j, h).$$

Basically, a validated numerical integration method is based on a numerical integration scheme such as Taylor series [Nedialkov et al., 1999] or Runge-Kutta methods [Sandretto and Chapoutot, 2016, Dit Sandretto and Chapoutot, 2016] which is extended with interval analysis tools to bound the *local truncation error*, i.e., the distance between the exact and the numerical solutions. Such methods work in two stages at each integration step, starting from an enclosure $[\mathbf{y}_j] \ni \mathbf{y}(t_j; \mathbf{y}_0)$ at time t_j of the exact solution, we proceed by:

i. a computation of an *a priori* enclosure $[\widetilde{\mathbf{y}}_{j+1}]$ of the solution $\mathbf{y}(t; \mathbf{y}_0)$ for all t in the time interval $[t_j, t_{j+1}]$. This stage allows one to prove the existence and the uniqueness of the solution.

ii. a computation of a tightening of state variable $[\mathbf{y}_{j+1}] \ni \mathbf{y}(t_{j+1}; \mathbf{y}_0)$ at time t_{j+1} using $[\widetilde{\mathbf{y}}_{j+1}]$ to bound the local truncation error term $\text{lte}_\Phi(F, G, \mathbf{y}_j, t_j, h)$.

A validated simulation starts with the interval enclosures $[\mathbf{y}(0)]$, $[\mathbf{x}(0)]$, $[\mathbf{p}]$ and $[\mathbf{u}]$ of respectively, $\mathcal{Y}_0$, $\mathcal{X}_0$, $\mathcal{P}$, and $\mathcal{U}$. It produces two lists of boxes:

- the list of discretization time steps: $\{t_0, \ldots, t_{\text{end}}\}$;

- the list of state enclosures at the discretization time steps: $\{[\mathbf{y}_0], \ldots, [\mathbf{y}_{\text{end}}]\}$;

- the list of *a priori* enclosures: $\{[\widetilde{\mathbf{y}}_0], \ldots, [\widetilde{\mathbf{y}}_{\text{end}}]\}$.

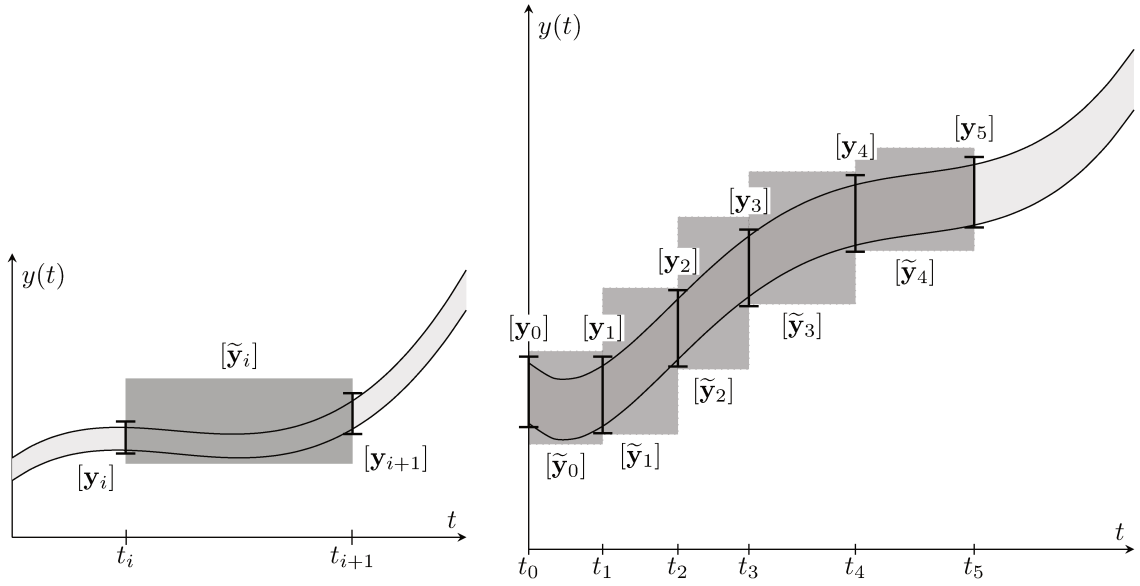


Figure 4.4: *The trajectories (in light gray) are overapproximated by* $[\mathbf{y}_i]$ *(thick line segment) at time step* t_i*. The a priori enclosure* $[\widetilde{\mathbf{y}}_i]$ *(in gray) contains the trajectories over the time interval* $[t_i, t_{i+1}]$*.*

Figure 4.4 represents the enclosures $[\widetilde{\mathbf{y}}_i]$ and $[\mathbf{y}_i]$ and their membership properties with the trajectories of the dynamical system.

Example 4.4.

We consider an Initial Value Problem

$$\begin{cases} \dot{\mathbf{y}} = -6\mathbf{y} + \mathbf{w} \\ \mathbf{y}(0) \in [\mathbf{y_0}] \\ \mathbf{w}(t) \ \in [\mathbf{w}_t] \end{cases} \tag{4.10}$$

A plot of the result of the validated simulation is given in Figure 4.5 for

$$\begin{cases} [\mathbf{y_0}] = [-2, 2] \\ [\mathbf{w_t}] = [-1, 1] \end{cases} \tag{4.11}$$

and in Figure 4.6 for

$$\begin{cases} [\mathbf{y_0}] = [2, 2] \\ [\mathbf{w_t}] = [0] \end{cases} \tag{4.12}$$

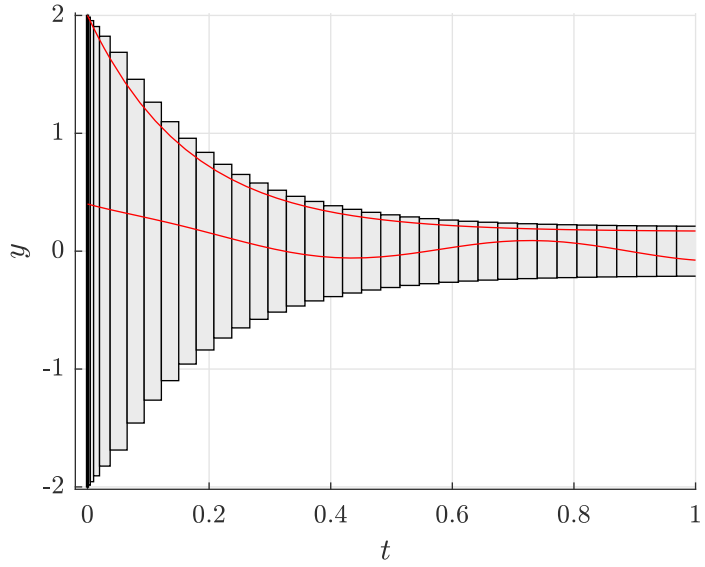


Figure 4.5: *Results of the validated simulation for the system described in* (4.10) *with parameters* (4.11)*. The system trajectories are overapproximated by the union of intervals* $[\widetilde{\mathbf{y}}_k]$ *(in grey). Two trajectories are represented (in red): one with* $y_0 = 0.4$ *and* $w(t) = cos(t)$*; and another one with* $y_0 = 2$ *and* $w(t) = 1$*.*

4.3 System with integral constraint over the state

We presented the classical framework of interval arithmetic and the framework of validated numerical integration. Such a framework handle systems that are subject to disturbances bounded by the ∞-norm (i.e. nonlinear systems subject to a disturbance w such that $\|w(t)\| \leq 1$ for example). However, the case where the disturbance is subject to an integral constraint has never been addressed until now.

The next sections propose a method to overapproximate the set of reachable states of a nonlinear system subject to a disturbance bounded by an integral constraint. We assume the disturbance and its time derivative to be bounded in ∞-norm (i.e. to be bounded at any time). These bounds are used to get a first overapproximation of the reachable tube. This coarse overapproximation might contain a set of trajectories not satisfying the integral constraint. We use a contractor operator (as introduced in Section 4.1.1) in order to reduce the initial prior overapproximation of the disturbance set. This new disturbance is then reused to get a new (smaller) overapproximation of the reachable tube. These two operations are iterated until a fixed point is reached.

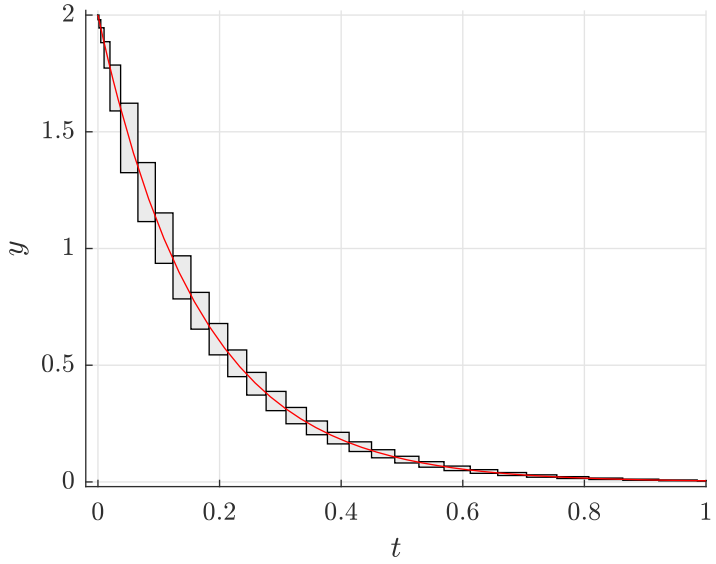


Figure 4.6: *Results of the validated simulation for the system described in* (4.10) *with parameters* (4.11)*. The system has a unique trajectory (in red) that is overapproximated by the union of intervals* $[\widetilde{\mathbf{y}}_k]$ *(in grey).*

Consider the system:

$$\begin{cases} \dot{x} = f(t, x, w) \\ x(0) \in \mathbf{x_0} \end{cases} \tag{4.13}$$

where w is an unknown disturbance in $\mathscr{L}_{2,\text{loc}}(\mathbb{R}_+; \mathbb{R}^m)$ that satisfies the integral constraint, for any $\tau \geq 0$:

$$\int_0^\tau \|w(s)\|^2 \, ds \leq \int_0^\tau g(s, x(s)) ds \tag{4.14}$$

where $g : \mathbb{R}_+ \times \mathbb{R}^n$ is a given function.

Remark 4.1. IQC bounds

The integral constraint does not give any bound on the disturbance as it can be easily understood from the unit energy disturbed system

$$\begin{cases} \dot{x} &= -x + w \\ x(0) &= 0 \\ 1 &\geq \displaystyle\int_0^1 w^2(\tau) d\tau \end{cases}. \tag{4.15}$$

Let w be defined for any $\epsilon > 0$ by

$$\begin{cases} w(\tau) = \dfrac{1}{\epsilon} & \text{when } \tau \in [0, \epsilon] \\ w(\tau) = 0 & \text{otherwise.} \end{cases}$$

Since $\int_0^1 w^2(\tau)d\tau = 1$, the inequality in (4.15) is satisfied for every $\epsilon > 0$, however, no bounds can be determined for w since $w(0) \to \infty$ when $\epsilon \to 0$. Please note that the system defined in (4.15) has a bounded reachable set even if the disturbance cannot be bounded at any given time (see [Boyd et al., 1994, Chapter 8.1.2]).

4.4 Validated numerical integration for dynamical systems subject to integral constraints

This section presents the main contribution of this chapter. For a system described by (4.13) and subjects to the integral constraint defined by (4.14), we compute an overapproximation of its reachable tube over the time domain $[0, T]$, where the time horizon $T > 0$ is given. A first overapproximation of the reachable tube is computed using pessimistic bounds over the disturbances. The integral constraint in (4.14) is used to derive a contractor. This contractor and a propagation step are applied in a fixed point algorithm until a contraction factor is reached. We run the algorithm over a simple example.

4.4.1 Extended system

We extend the state of the system with the integral value corresponding to the integral constraint in (4.14):

$$\begin{cases} \dot{z}(t) = g(t, x(t)) - \|w(t)\|^2 \\ z(0) = 0 \end{cases} \tag{4.16}$$

Then, (4.14) can be equivalently expressed for z:

$$\forall t \in \mathbb{R}_+, z(t) \geq 0. \tag{4.17}$$

As mentioned in Remark 4.1, no L_∞ bounds can be derived for L_2 bounded signals. To study such systems, we make further assumptions about the disturbance:

Assumption 4.1. Continuous disturbance signal
w is continuous, differentiable, and of continuous derivative over $\mathbb{R}_+$.

This assumption seems reasonable in the case of real systems modeling since disturbances modeled by integral constraints correspond to physical quantities. Since the continuity of a function over a closed interval implies its boundedness, Assumption 4.1 implies that the signal w is bounded and of bounded variation over $[0,T]$. Therefore, there exists $[\mathbf{w}] \in \mathbb{IR}^m$ and $[\mathbf{w}'] \in \mathbb{IR}^m$ such that for all $t \in [0,T]$:

$$\begin{cases} w(t) \in [\mathbf{w}] \\ \dot{w}(t) \in [\mathbf{w}'] \end{cases} \tag{4.18}$$

Using Assumption 4.1 and (4.16), the following system will be studied:

$$\mathcal{S} : \begin{cases} \dot{x}(t) = f(t, x(t), w(t)) \\ \dot{z}(t) = g(t, x(t), w(t)) - \|w(t)\|^2 \\ \dot{w}(t) \in [\mathbf{w}'] \\ x(0) \in [\mathbf{x_0}] \\ z(0) = 0 \\ \quad 0 \leq z(t) \\ w(t) \in [\mathbf{w}] \end{cases} \tag{4.19}$$

where $[\mathbf{x_0}] \in \mathbb{IR}^n$ is the set of initial states. We use the following notation $(x,z,w) \in \mathcal{S}$ iff $(x,w) \in \mathscr{L}_{2,\text{loc}}([0,T];\mathbb{R}^n) \times \mathscr{L}_{2,\text{loc}}([0,T];\mathbb{R}) \times \mathscr{L}_{2,\text{loc}}([0,T];\mathbb{R}^m)$ is a trajectory of $\mathcal{S}$.

(4.18) gives prior bounds on the disturbance w. They can be used to propagate the trajectories using standard validated numerical integration frameworks. Thanks to this, we get a first a priori overapproximation of the reachable set. In the next section, we use this first overapproximation and a contractor (defined out of the integral inequality) in a fixed point algorithm in order to get a tighter overapproximation of the reachable set.

4.4.2 Bounds over the disturbance

In this section, (4.18) and the integral constraint in (4.17) are used to derive bounds over the disturbance w. These bounds are then used to define a contractor over the *a priori* enclosure of the trajectories.

We present a preliminary result before deriving bounds over the disturbance w:

Proposition 4.1. Overapproximate intersection
For $[\mathbf{v}] \in \mathbb{IR}^p$, $p \in \mathbb{N}$ *and* $r > 0$, *if* $[\![\mathbf{v}]\!] \leq r$ *then* $[\mathbf{v}] \subset [-r,r]^p$.

Proof. In an Euclidean space, the norm 1 and norm 2 satisfy $\sqrt{v_1^2 + \cdots + v_p^2} \leq |v_1| + \cdots + |v_p|$ for any $(v_1, \ldots, v_p) \in \mathbb{R}^p$. ♢

When w satisfies (4.18) and a given integral constraint, hard bounds (i.e. in ∞-norm) can be derived over w:

Proposition 4.2. Disturbance bounds

For a $w \in \mathscr{L}_{2,\text{loc}}([0,h];\mathbb{R}^m)$ defined over an interval of length $h > 0$. If w satisfies (4.18) (with given bounds $[\mathbf{w}], [\mathbf{w}'] \in \mathbb{IR}^m$), then for any $r > 0$:

$$\int_0^h \|w(\tau)\|^2 d\tau \leq r \Rightarrow \forall \tau \in [0,h], w(\tau) \in [\mathbf{W_r}],$$

where $[\mathbf{W_r}] = [-k,k]^n$ with $k = \sqrt{\frac{r}{h}} + \frac{h}{2}[\![\mathbf{w}']\!]$ (where $[\![\mathbf{w}']\!]$ is the maximum Euclidean norm over the elements of $[\mathbf{w}']$).

Proof. By applying the Cauchy-Schwartz inequality between the signal w and $t \mapsto 1$ for the inner product of square-integrable function, we have:

$$\left\| \int_0^h w(\tau) d\tau \right\|^2 \leq h \int_0^h \|w(\tau)\|^2 d\tau \leq hr.$$

By (4.18), $w(\tau) = w_0 + \int_0^\tau w_1(\kappa) d\kappa$ with $w_0 \in [\mathbf{w}]$ and $w_1(\cdot) \in [\mathbf{w}']$. Using the reverse triangular inequality, we have:

$$\left\| \int_0^h w_0 d\tau \right\| \leq \sqrt{rh} + \left\| \int_0^h \int_0^\tau w_1(\kappa) d\kappa \right\|.$$

Then, we get:

$$\|hw_0\| \leq \sqrt{hr} + \frac{h^2}{2}[\![\mathbf{w}']\!]. \tag{4.20}$$

This relationship is derived over $[0,h]$ but is also valid for any time interval $[t,t+h]$ of width h, $t > 0$. Therefore, by using Proposition 4.1 and (4.20), we have: $\forall \tau \in [0,h], w(\tau) \in [\mathbf{W_r}]$. ♢

We then use Proposition 4.2 to derive bounds in the specific case of (4.16). Consider a system trajectory $(x,z,w) \in \mathcal{S}$, such that at a given $t \in [0,T]$ and $h > 0$ s.t. $t + h \in [0,T]$, and for all $\tau \in [t,t+h]$:

$$\begin{cases} (x(t),z(t),w(t)) \in [\mathbf{y_t}] \\ (x(\tau),z(\tau),w(\tau)) \in [\widetilde{\mathbf{y}}_t] \end{cases} \quad \text{where} \quad \begin{cases} [\mathbf{y_t}] = [\mathbf{x_t}] \times [\mathbf{z_t}] \times [\mathbf{w_t}] \\ [\widetilde{\mathbf{y}}_t] = [\widetilde{\mathbf{x}}_t] \times [\widetilde{\mathbf{z}}_t] \times [\widetilde{\mathbf{w}}_t] \end{cases}. \tag{4.21}$$

The trajectories belong to $[\mathbf{y_t}]$ at t and are in $[\widetilde{\mathbf{y}}_t]$ between $[t,t+h]$. At $t+h$, for a given $t \geq 0$ and a given $h \geq 0$, (4.19) implies that z satisfies:

$$z(t+h) = z(t) + \int_t^{t+h} g(t,x(t)) d\tau - \int_t^{t+h} \|w(\tau)\|^2 d\tau.$$

By applying (4.17) at $t+h$ implies that $z(t+h) \geq 0$, we have the following relationship:

$$\int_t^{t+h} \|w(\tau)\|^2 \, d\tau \leq z(t) + \int_t^{t+h} g(\tau, x(\tau)) d\tau. \tag{4.22}$$

Let a function

$$q(z, x) = z + \int_t^{t+h} g(\tau, x(\tau)) d\tau. \tag{4.23}$$

By using an interval evaluation $[q]$ of q, the upperbound of $q(z, x)$ can be evaluated for $z \in [\mathbf{z_t}]$ and $x \in [\widetilde{\mathbf{x}}_t]$. We denote by $\overline{[q]([\mathbf{z_t}], [\widetilde{\mathbf{x}}_t])}$ this upperbound. For any $w \in \mathbf{L}_{2,loc}([t, t+h], [\widetilde{\mathbf{w}}_t])$, 4.22 implies:

$$\int_t^{t+h} \|w(\tau)\|^2 \, d\tau \leq \overline{[q]([\mathbf{z_t}], [\widetilde{\mathbf{x}}_t])}.$$

Then, Proposition 4.2 can be used to derive bounds over the disturbance w:

Proposition 4.3. Disturbance overapproximation
For a $w \in \mathscr{L}_{2,\mathrm{loc}}([t, t+h]; \mathbb{R}^m)$ defined over an interval of length $h > 0$, $t > 0$. If w satisfies (4.18) (with given bounds $[\mathbf{w}], [\mathbf{w}'] \in \mathbb{IR}^m$), then for any $\tau \in [t, t+h]$:

$$w(\tau) \in [\mathbf{W_q}], \tag{4.24}$$

where $[\mathbf{W_q}]([\widetilde{\mathbf{x}}_t], [\mathbf{z_t}]) = [-r, r]^m$ with $r = \sqrt{\frac{\overline{[q]([\mathbf{z_t}], [\widetilde{\mathbf{x}}_t])}}{h}}$ and q defined in (4.23).

Proof. This is a direct application of Proposition 4.2. ♢

We then define the operator over $[\mathbf{y_t}]$ and $[\widetilde{\mathbf{y}}_t]$

$$\mathcal{C}([\mathbf{y_t}], [\widetilde{\mathbf{y}}_t]) = ([\mathbf{y_t}] \cap [\mathbf{Y_g}]([\widetilde{\mathbf{x}}_t], [\mathbf{z_t}]), [\widetilde{\mathbf{y}}_t] \cap [\mathbf{Y_g}]([\widetilde{\mathbf{x}}_t], [\mathbf{z_t}])) \tag{4.25}$$

where $[\mathbf{y_t}]$ and $[\widetilde{\mathbf{y}}_t]$ are defined in (4.21),

$$[\mathbf{Y_g}] = [-\infty, \infty]^n \times [0, \infty] \times [\mathbf{W_q}],$$

with $[\mathbf{W_q}]$ defined in Proposition 4.3.

Proposition 1. *$\mathcal{C}$ defined in (4.25) is a contractor.*

Proof. By Proposition 4.3, we have, for $\tau \in [t, t+h]$,

$$w(\tau) \in [\mathbf{W_q}],$$

i.e., all the disturbance signals of $\mathcal{S}$ belong to $[\mathbf{W_q}]$, so the contractor is conservative. Since the contractor is defined as an intersection with $[\mathbf{y}_t]$ and $[\widetilde{\mathbf{y}}_t]$ respectively, we have

$$([\mathbf{y}_t],[\widetilde{\mathbf{y}}_t]) \subseteq \mathcal{C}([\mathbf{y}_t],[\widetilde{\mathbf{y}}_t]),$$

$\mathcal{C}$ is contractive. For any $([\mathbf{y}'_t],[\widetilde{\mathbf{y}}'_t])$ such that $[\mathbf{y}'_t] \subseteq [\mathbf{y}_t]$ and $[\widetilde{\mathbf{y}}'_t] \subseteq [\widetilde{\mathbf{y}}_t]$,

$$\mathcal{C}([\mathbf{y}'_t],[\widetilde{\mathbf{y}}'_t]) \subseteq \mathcal{C}([\mathbf{y}_t],[\widetilde{\mathbf{y}}_t]),$$

i.e. $\mathcal{C}$ is monotone. ♢

4.4.3 Integral constraint propagation

The contractor defined by (4.25) is used in a fixed point algorithm as in [Alexandre dit Sandretto and Chapoutot, 2016]. A priori enclosure of the trajectory is computed using bounds (4.18) over w. The integration algorithm gives

- the discretization time steps: $\{t_0,\ldots,t_{end}\}$;
- the state enclosure at the discretization time steps: $\mathcal{Y}^0 = \{[\mathbf{y}_0^0],\ldots,[\mathbf{y}_{end}^0]\}$;
- the *a priori* enclosures: $\widetilde{\mathcal{Y}}^0 = \{[\widetilde{\mathbf{y}}_0^0],\ldots,[\widetilde{\mathbf{y}}_{end}^0]\}$.

We then apply the contractor over each couple of discretized time-step boxes $[\mathbf{y_i^0}] \in \mathcal{Y}^0$ and their associated *a priori* enclosures $[\widetilde{\mathbf{y}}_i^0] \in \widetilde{\mathcal{Y}}^0$. These two steps are repeated in a fixed point algorithm until the contraction factor is lower than a given value. In this approach, time steps are computed at the first iteration of the algorithm and are not updated.

Example 4.5.
We study the following linear time-invariant system disturbed by an unknown signal w constrained by a 2-norm inequality:

$$\begin{cases} \dot{x}(t) = -x(t) + w(t) \\ \displaystyle\int_0^t w(\tau)^2 d\tau \leq \int_0^t 0.01 x(\tau)^2 d\tau \\ x(0) \in [-1,1] \end{cases} \tag{4.26}$$

with $[\mathbf{w}] = [-1,1]$ and $[\mathbf{w}'] = [-1,1]$ in (4.18) for $t \in [0,2.5]$. Figure 4.7 shows the reachable set of this dynamical system computed with the method described in this section.

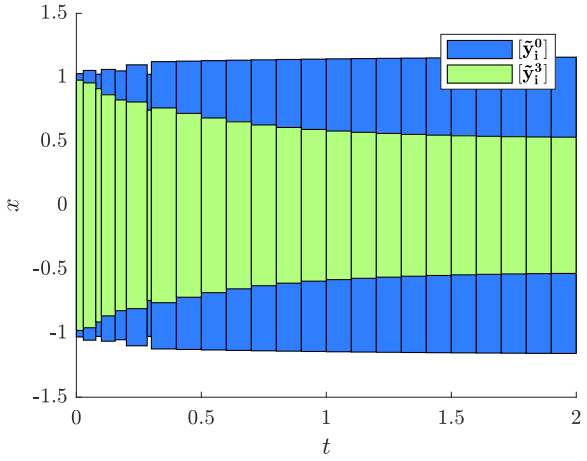


Figure 4.7: *Computation of the overapproximation of the reachable set of Example 4.5 using the algorithm presented in Section 4.4. Blue boxes correspond to the a priori enclosures at the first iteration of the algorithm $\widetilde{\mathcal{Y}}^0$, green boxes are the a priori enclosure at the 3^{rd} iteration $\widetilde{\mathcal{Y}}^3$ of the algorithm.*

4.5 Example

In this part, we present an application of the method described in Section 4.4 to a system with a time delay.

For $u, v \in \mathbf{L}_{2,loc}(\mathbb{R}_+;\mathbb{R})$, the delay operator D_h over an input signal u is defined by the following relationship:

$$v = D_h(u) \Leftrightarrow \begin{cases} v(t) = u(t-h) & \text{for all } t \geq h \\ v(t) = 0 & \text{otherwise.} \end{cases} \tag{4.27}$$

Validated numerical integration of differential equations with delays is challenging. Since they act as a memory of the past input signal over an interval of width h, the state of the delay belongs to $\mathbf{L}_{2,loc}([0,h],\mathbb{R})$. The dimension of the system state space is therefore non-finite.

The stability of linear time-invariant (LTI) systems with internal delays is studied in [Seuret and Gouaisbaut, 2015]. The state of the delay is projected over a finite Legendre polynomial basis. These projections are time-dependent values since the state of the delay is also time-varying. The time derivative of these projections only depends on the input of the delay operator. Then the norm of the state is overapproximated using a Bessel inequality. By integrating this inequality, we get an Integral Quadratic Constraint (IQC) between the output of the delay operator, its

input, the derivative of its inputs, the projections over the truncated basis of Legendre polynomial and an error signal. The IQC models the energy of the remaining of the Legendre expansion (i.e. the error signal). In [Seuret and Gouaisbaut, 2015], the stability of the delayed LTI system is assessed for all possible error signals satisfying the derived IQC. We use this IQC to overapproximate the reachable set of such systems.

In what follows, we use the first order of the IQC relationship described in [Seuret and Gouaisbaut, 2015, Theorem 5]. The state ξ corresponds to the average value of the state of the delay. The remaining energy of the state is bounded by an integral quadratic constraint.

$$\begin{cases} \dot{\xi}(t) = -15\xi(t) + 1.5v(t) - w(t) \quad \text{with} \quad \xi(0) = 0 \\ \text{under the IQC } \int_0^t w(s)^2 ds \leq \int_0^t \left[0.0025\dot{v}(s)^2 - 0.75\left(v\left(s\right) - \xi\left(s\right)\right)^2\right] ds \end{cases} \tag{4.28}$$

The IQC system (4.28) is used to overapproximate the delay in the following system:

$$\begin{cases} \dot{x} &= -x - k_c D_h(x) + u \\ x(0) &= 0 \end{cases} \tag{4.29}$$

where $k_c = 4$, $h = 0.01$ and $u(t) = 1 - t$. (4.27,4.28,4.29) are then combined in a unique linear time-invariant system with an integral quadratic constraint.

$$\begin{cases} \dot{X}(t) &= AX + B_w w(t) + B_u u(t) \\ X(0) &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ & \int_0^t w(\tau)^2 d\tau \leq \int_0^t \begin{bmatrix} X(\tau) \\ u(\tau) \end{bmatrix}^\top M \begin{bmatrix} X(\tau) \\ u(\tau) \end{bmatrix} \end{cases} \tag{4.30}$$

where the matrices are defined by

$$A = \begin{bmatrix} 1.0417 & 15.6250 \\ -6.0417 & -15.6250 \end{bmatrix}, B_w = \begin{bmatrix} 1.0000 \\ -1.0000 \end{bmatrix}, B_u = \begin{bmatrix} 1.0417 \\ -0.0417 \end{bmatrix}$$

and

$$M = \begin{bmatrix} -12.4566 & -30.5990 & 0.0434 \\ -30.5990 & -68.3594 & 0.6510 \\ 0.0434 & 0.6510 & 0.0434 \end{bmatrix}.$$

The bounds in Eq.(4.24) are $[\mathbf{w}] = [-10, 10]$ and $[\mathbf{w}'] = [-1, 1]$. The initial disturbance set is defined such that $[\mathbf{w}_0] = [\mathbf{w}]$.

Figure 4.8 corresponds to the reachable tube of the delayed system modeled with the integral quadratic constraint. Y_{IQC} is the reachable tube of the corresponding system.

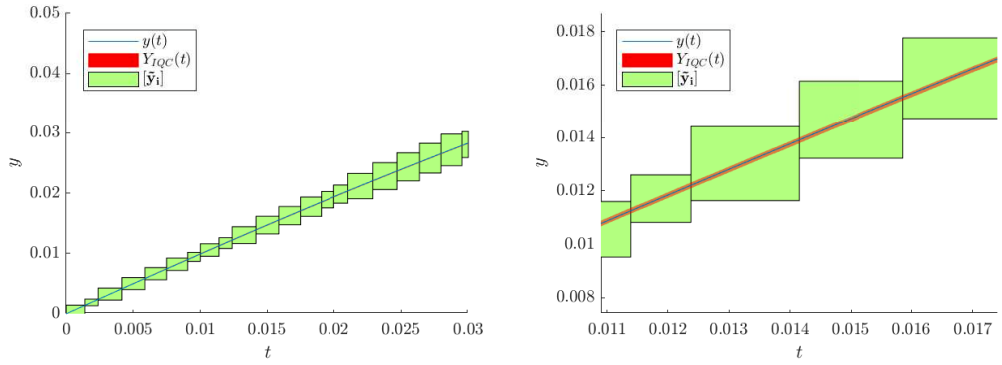


Figure 4.8: *Computation of the reachable tube of the system* (4.30) *using the validated numerical integration framework described in the previous section and the contractor $\mathcal{C}$ introduced in Section 4.4. y (the blue line) corresponds to the response of the delayed system. Y_{IQC} is the exact reachable tube of system computed using the paraboloid method presented in Chapter 3.*

Discussion One motivation of this work is to use Integral quadratic constraint (IQC) models in a validated numerical integration framework. IQC models are widely used in the robust control community for stability analysis of dynamical systems. When the IQC system is stable, there exists an invariant over the set of states (x, z), and the maximal reachable z value (i.e. the maximal integral value reachable) is bounded for any trajectory.

In our approach, such an invariant does not exist. The overapproximation of the maximal reachable z is constantly increasing in size. Consequently, bounds provided by the fixed point algorithm are also strictly increasing in size. When these bounds reach the prior bounds given by (4.18) over the disturbance, the reachable set tends to the reachable set computed without the integral constraint. Figure 4.9 corresponds to the reachable set of Example 4.5 for a larger horizon of integration. The integral constraints provide bounds over w. However, when the energy level is too high, these bounds are strictly included in the bounds given by (4.18). At $t = 15s$, the reachable set converges to the reachable set of the system with no integral constraint between the disturbance and the state.

The bounds of the input disturbance depend on the result of the used guaranteed set integration method. Therefore, if the later is too pessimistic, the proposed contraction method will only rely on the bounds $[\mathbf{w}]$ and $[\mathbf{w}']$ of Eq.(4.18).

In our approach, a larger class of systems is considered compared to the linear case treated in Chapter 3. Contrary to IQC models, only the dependence in the disturbance needs to be quadratic for the integral constraint.

In terms of scalability, our approach needs the state of the original dynamical

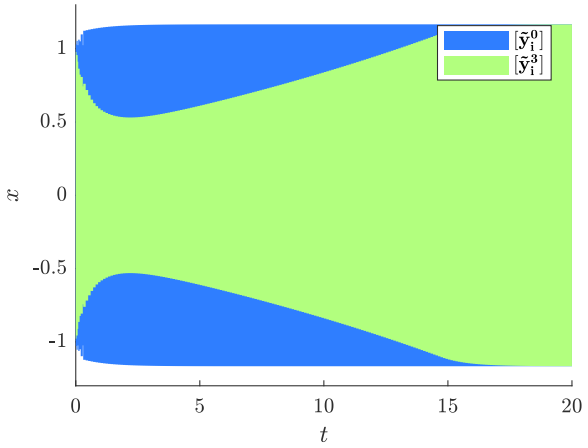


Figure 4.9: *Computation of the reachable tube of the system* (4.15) *in Example 4.5 over* $[0, 20]$ *using the validated numerical integration framework described in Chapter 4 and the contractor* $\mathcal{C}$ *introduced in Section 4.4. In blue, the reachable set when only* (4.18) *is used (i.e., when the integral constraint is not used). In green, the reachable set of the system when the integral constraint is taken into account.*

system to be extended from n variables to $p = n + m + 1$ variables (m states for w, 1 state for z). Since the disturbance signal spans in a subspace of R^n, m is always smaller than n. Since m is often close to 1 (the delay modeled as an integral quadratic constraint introduces a 1-dimensional disturbance signal), p is close to n (or $2n$ in the worst case). However, only the integration part can suffer from the dimension of the system. Based on the advantage of our approach, a less expansive integration method can be used for large systems for a similar result.

4.6 Conclusion

We presented a method to compute an overapproximation of the reachable tube for dynamical systems with integral constraints over the input set. To overapproximate the reachable set, we use a Runge-Kutta validated numerical integration scheme with pessimistic bounds over the input. It provides a first conservative bound over the reachable tube. Then, the integral constraint is used to define a contractor over the reachable tube. This contractor and a propagation step are successively applied to the overapproximation until a fixed point is reached. We evaluated our algorithm with DynIbex library to simulate a delayed system, i.e., an infinite-dimensional system that can be modeled as a linear time-invariant system subject to an integral quadratic constraint.

Future works The method developed in this chapter is guaranteed (we compute an overapproximation of the reachable tube). However, our overapproximations tend to constantly grow in size, even when the reachable set is known to be bounded (see Section 4.5). Such an issue originates from the use of intervals to overapproximate the reachable set. Intervals are too conservative and the use of affine forms to overapproximate the reachable set should provide better-conditioned overapproximations. One purpose of this work was to investigate how models introduced by the robust control community can be used in the field of validated numerical integration. We proposed the use of an IQC model that bounds the input to output $\mathscr{L}_2$ gain of a system with an internal constant-delay. Many complex systems can be modeled in a similar approach. More specifically, the error of approximation in a reduced system can be expressed with an input to output $\mathscr{L}_2$ gain constraint. Simplification of models is very appealing for validated numerical integration since the computational time is mainly dependent on the system dimension. Being able to reduce the order of the system and to bound the error with a 2-norm gain would lead to a more efficient algorithm.

Part III

Set-based Cosimulation: Abstract Interpretation for Reachability Analysis

Part I presents a set-based simulation method that uses time-varying ellipsoids to compute overapproximations of the reachable set of a system. The systems of interest are continuous-time linear time-varying systems subject to disturbances bounded by quadratic constraints in a Hilbert space. Two subclasses of models have been more specifically developed, the (point-wise) quadratic constraint model and the Integral Quadratic Constraint (IQC) case. These classes of systems are of particular interest since they are widely studied in the field of robust control theory. In particular, these models can be used as abstractions to complex models that do not fall into the initial scope of considered systems (e.g. nonlinear systems, systems with internal delays).

Many other reachability analysis frameworks are available. They differ by the family of systems they can analyze, the geometrical sets used to overapproximate the reachable set, and the available tools to compute these overapproximations. Complex systems might involve different family of dynamical systems (discrete-time and continuous-time, nonlinear and linear behaviors for example). In this case, one might need to use several reachability analysis frameworks to compute the overapproximation of the reachable set.

In this part, we propose to analyze an interconnection of systems. Each subsystem in the interconnection corresponds to a system that can be analyzed with its associated reachability analysis framework (such as the ones developed in Part I and Part II). This interconnection of systems can be described with two basic operations, a composition of systems and a feedback operator. One theoretical and practical difficulty in the analysis of such an interconnection is to "close the loop", i.e. to study the following system

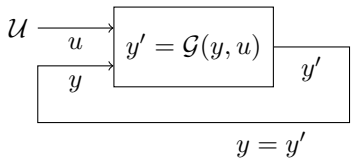


Figure 4.10: *Block diagram of a closed-loop system*

For such systems, one should overapproximate the set of solution $\mathcal{Y}$ of the fixed point equation

$$y = \mathcal{G}(y, u) \tag{4.31}$$

for every input u in a set of inputs $\mathcal{U}$.

Chapter 5 describes the classical framework of abstract interpretation and its application to study an interconnection of systems.

Chapter 5

Set-based Cosimulation

Contents

This chapter presents the abstract interpretation framework applied to the analysis of an interconnection of systems. Such interconnection frequently appears in system modeling and usually involves systems of heterogeneous nature. For example, a continuous plant controlled by a discrete-time controller is an interconnection of two systems. Hereby, the behavior of the continuous plant might be described as solutions to an ordinary differential equation and the behavior controller by a computer program. Even if there exists tools to individually analyze each system, analyzing their interconnection is still challenging. In this chapter, we show that the abstract

interpretation framework is suited to address such issues.

Abstract interpretation is a common tool in the verification of computer programs. To prove that a program satisfies a set of properties, abstract interpretation starts from these two following observations: a computer program can be exactly described by its *concrete semantic*, i.e. the set of its executions; and the specification to verify is, as well, a subset in the domain of the concrete semantic. If the concrete semantic is a subset of this property, then the program is valid. However, this concrete semantic is usually *not computable*. Abstract interpretation aims at computing an *abstract semantic*, a superset of this concrete semantic. The abstract semantic is also called sound approximation (or over-approximation) of the concrete semantic. Contrary to the concrete semantic, the abstract semantic is designed to be a simpler and hopefully a computable mathematical object. If the abstract semantic is a subset of the property, then the concrete semantic, i.e. the program, verifies the property.

Abstract interpretation has been mainly used in program analysis where the concrete semantic domain corresponds to the set of finite or infinite sequences of symbols. However, this notion of semantic is versatile enough and can be used in other fields. In the case of dynamical systems, the semantic domain corresponds to a signal space (such as the set of finite or infinite horizon, discrete or continuous-time, real-valued vectored signals).

This chapter is organized as follows. In Section 5.1, we define a language for an interconnection of systems. In Section 5.2, we define the concrete semantic of an interconnection of systems. Section 5.2, Section 5.3 and Section 5.4 present the usual framework and classical results of Abstract Interpretation. Most of their content is available in [Cousot and Cousot, 1979] and is restated hereby for completeness. In Section 5.2, we define the concrete semantic that is equivalent to the semantic of the interconnection of systems. In Section 5.3, we define abstract domains. In Section 5.4, we define the abstract semantic. In Section 5.5, Section 5.6, and Section 5.7, we present abstracts domains that we use to describe signals. In Section 5.8, we apply the framework to find a sound approximation of the reachable tube of several interconnections of systems. In Section 5.9 and Section 5.10, we present the works related to the presented framework and conclude.

Notations are not classical and are chosen to be consistent with the rest of the document. Instead of the classical $\cdot^{\#}$ for abstract elements and $\cdot^{\flat}$ for concrete elements, we use a starred notation for abstract elements (x^*, X^*, ...) and no superscript notation (x, X, ...) for the concrete elements.

5.1 Syntax and semantic of an interconnection of systems

We study an interconnection of systems that can be expressed with two constructions between subsystems: a serial connection and a feedback connection. Each connection is associated with a signal and an equation satisfied by this signal. The problem of interest is then to identify the set of signals of the interconnection that satisfies this set of equations.

In this part, we present the *syntax*, or *language*, used to describe our interconnection of systems. This syntax is introduced for two reasons. First, it identifies the class of systems we can analyze. Second, it gives a structure to the set of equations described by the interconnection of systems. This structure is, later on, reused to analyze the interconnection of systems. Once the syntax is defined, the actual mathematical meaning of the syntax is expressed by its *semantic*. We associate to each construct in our syntax an equation that should be satisfied by the signals of the interconnection.

An interconnection of systems is described with an $\texttt{ISys}$ expressed within the syntax detailed in Table 5.1.

$\texttt{ISys}$	:=	$\texttt{Src} \xrightarrow{\texttt{u}} \texttt{Snk}$	$\texttt{u} \in \mathbb{V}$	*(Interconnection of systems)*
$\texttt{Src}$	:=	$\mathcal{U}$	$\mathcal{U} \subseteq \texttt{Dom}_\texttt{u}$ where $\texttt{u} \in \mathbb{V}$	*(Source)*
	|	$\texttt{Src} \xrightarrow{\texttt{u}} \texttt{Sys}$	$\texttt{u} \in \mathbb{V}$	*(Serial connection)*
$\texttt{Snk}$	:=	$\circ$	a *sink* symbol	*(Sink)*
	|	$\texttt{Sys} \xrightarrow{\texttt{u}} \texttt{Snk}$	$\texttt{u} \in \mathbb{V}$	*(Serial connection)*
$\texttt{Sys}$	:=	$\mathcal{S}$	$\mathcal{S} : \texttt{Dom}_\texttt{u} \to \texttt{Dom}_\texttt{y}$ with $\texttt{u}, \texttt{y} \in \mathbb{V}$	*(System)*
	|	$\texttt{Sys} \xrightarrow{\texttt{v}} \texttt{Sys}$		*(Serial connection)*
	|	$\mu_\texttt{x}\{\texttt{Sys}\}$		*(Feedback connection)*
	|	$(\texttt{Src}, \texttt{Sys})$		*(Concatenation of signals)*
$\mathbb{V}$	a set of labels.			*(Labels)*
$\texttt{Dom}_\texttt{x}$	set of signals associated with $\texttt{x} \in \mathbb{V}$ and equal to the Cartesian product $X_1 \times X_2 \times \ldots$ where X_i are a finite real-valued vectorial space $\mathbb{R}^n$, a discrete-time signal space $l(\{0, \ldots, T\}; \mathbb{R}^n)$, or a continuous-time signal space $L([0, t]; \mathbb{R}^n)$, $n \in \mathbb{N}$, $T \in \mathbb{N}$, $t \in \mathbb{R}$.			*(Domain)*

Table 5.1: *Syntax of an interconnection of systems* $\texttt{ISys}$.

We assume that any `ISys` expressed in the syntax described by Table 5.1 identify each connection (a serial connection or a feedback connection) with a unique label in the set of labels $\mathbb{V}$ (in other words, a label does not appear two times in the formula). An interconnection of systems `ISys` is then described with a set of systems, a set of sources, and connections in-between each of them.

Signal The syntax describing an interconnection of systems is used to define the connections in-between each system. A connection is labeled with a $\mathtt{u} \in \mathbb{V}$ and associated with a signal u that is an element of a domain $\mathtt{Dom}_\mathtt{u}$. These signals are chosen as a concatenation of constant (e.g. a parameter, an initial state), continuous-time (e.g. the output or input of a continuous-time dynamical system), and/or discrete-time signals (e.g. a signal generated by a computer program). Signals might be defined over a finite or infinite time-horizon. Then, $\mathtt{Dom}_\mathtt{u}$ corresponds to the Cartesian product of each domain of each type that composes u. All the signals of the interconnection of systems are then described with a single variable ρ, called an *environment*, $\rho = (u, x, y, \dots)$, $\mathbb{V} = \{\mathtt{u}, \mathtt{x}, \mathtt{y}, \dots\}$, that belongs to the Cartesian product of all the domains

$$\mathtt{Dom} = \prod_{\mathtt{v} \in \mathbb{V}} \mathtt{Dom}_\mathtt{v}.$$

For an environment $\rho \in \mathtt{Dom}$, a set of labels $\mathbb{V}$, and a label $\mathtt{u} \in \mathbb{V}$, let $\rho_\mathtt{u}$ be the projection of ρ over $\mathtt{Dom}_\mathtt{u}$.

Source A source $\mathcal{U} \subseteq \mathtt{Dom}_\mathtt{u}$ is a subset associated with a signal $u \in \mathtt{Dom}_\mathtt{u}$, and a label $\mathtt{u} \in \mathbb{V}$. A source describes an exogenous input such as an initial state of a dynamical system (or of a program), or an unknown bounded input signal (e.g. $\mathcal{U}$ might be the set of signals from $\mathbb{R}_+$ to $\mathbb{R}$ bounded at any time by $[-1, 1]$).

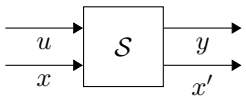
Sink The sink $\circ$ is a terminal point used such that every signal in the interconnection is connected to another element.

System The systems are defined as operators between an input domain and an output domain. For example, the system $\mathcal{S}$ in " $\dots \xrightarrow{\mathtt{u}} \mathcal{S} \xrightarrow{\mathtt{y}} \dots$ ", where $\mathtt{u}, \mathtt{y} \in \mathbb{V}$, defines a mapping

$$\mathcal{S} : \mathtt{Dom}_\mathtt{u} \to \mathtt{Dom}_\mathtt{y}$$

from $\mathtt{Dom}_\mathtt{u}$ (the domain of the signal u associated with the label $\mathtt{u} \in \mathbb{V}$) and $\mathtt{Dom}_\mathtt{y}$ (the domain of the signal y associated with the label $\mathtt{y} \in \mathbb{V}$). An interconnection of systems is then described with serial connections $\mathcal{S}_1 \xrightarrow{\mathtt{v}} \mathcal{S}_2$, feedback connections $\mu_\mathtt{x}\{\mathcal{S}\}$ and exogenous inputs in $\mathcal{U}$. The connections between systems are assumed to be correctly defined such that every types are compatible. A serial connection

" $\ldots \xrightarrow{\mathtt{u}} \mathcal{S}_1 \xrightarrow{\mathtt{v}} \mathcal{S}_2 \xrightarrow{\mathtt{w}} \ldots$ " involves two systems $\mathcal{S}_1 : \mathtt{Dom}_\mathtt{u} \to \mathtt{Dom}_\mathtt{v}$ and $\mathcal{S}_2 : \mathtt{Dom}_\mathtt{v} \to \mathtt{Dom}_\mathtt{w}$ (i.e. the output set $\mathtt{Dom}_\mathtt{v}$ of $\mathcal{S}_1$ corresponds to the input set of $\mathcal{S}_2$). A feedback connection " $\ldots \xrightarrow{\mathtt{u}} \mu_\mathtt{x}\{\mathcal{S}\} \xrightarrow{\mathtt{y}} \ldots$ " involves an *open-loop* system $\mathcal{S} : \mathtt{Dom}_\mathtt{u} \times \mathtt{Dom}_\mathtt{x} \to \mathtt{Dom}_\mathtt{y} \times \mathtt{Dom}_\mathtt{x}$ where the input and output signals both contain the same *state signal* x associated with the label $\mathtt{x} \in \mathbb{V}$.



The resulting closed-loop system " $\ldots \xrightarrow{\mathtt{u}} \mu_\mathtt{x}\{\mathcal{S}\} \xrightarrow{\mathtt{y}} \ldots$ " defines a system mapping $\mathtt{Dom}_\mathtt{u}$ to $\mathtt{Dom}_\mathtt{y}$,

Interconnection of systems An interconnection of systems $\mathscr{S}$ of type `ISys` is then described as a source `Src` connected to a sink `Snk`.

Semantic of the interconnection of systems For an interconnection of systems `ISys` denoted by $\mathscr{S}$ and expressed in the syntax defined in Table 5.1, the semantic of $\mathscr{S}$ corresponds to the set of environments $\overline{[\mathscr{S}]} \subseteq \mathtt{Dom}$ defined by

$$\overline{[\mathscr{S}]} = \{\rho \in \mathtt{Dom} \,|\, [\mathscr{S}](\rho; \varnothing, \varnothing) \text{ is } \mathtt{True}\}$$

where $\rho \mapsto [\mathscr{S}](\rho; \varnothing, \varnothing)$ is defined by Table 5.2. For each expression $\mathcal{S}$ of the syntax, Table 5.2 defines a function

$$[\mathcal{S}] : \mathtt{Dom} \times \overline{\mathbb{V}} \times \overline{\mathbb{V}} \to \{\mathtt{True}, \mathtt{False}\}.$$

by structural induction over the syntax of $\mathcal{S}$. The set $\overline{\mathbb{V}}$ is defined by

$$\overline{\mathbb{V}} = \{\varnothing\} \cup \mathbb{V} \cup \mathbb{V}^2, \tag{5.1}$$

$\overline{\mathbb{V}}$ contains a symbol $\varnothing$, the labels in $\mathbb{V}$, and the pairs of labels in $\mathbb{V}^2$. For a $\rho \in \mathtt{Dom}$, each operator $[\mathcal{S}](\rho; \mathtt{u}, \mathtt{y})$ defined in Table 5.2 evaluates the truth value that the signal $\rho_\mathtt{y}$ is an output of the system $\mathcal{S}$ for the input $\rho_\mathtt{u}$. The symbol $\varnothing$ is associated with expression in the syntax which ends with a sink symbol or begin with a source. A `Src` denoted by $\mathcal{U}$ has only an output connection and is therefore associated with the function $[\mathcal{U}](\rho; \varnothing, \mathtt{u})$. A `Snk` denoted by $\mathcal{Y}$ has only an input connection and is therefore associated with the function $[\mathcal{Y}](\rho; \mathtt{y}, \varnothing)$. A system `Sys` denoted by $\mathcal{S}$ has an input and an output connection and is therefore associated with the function $[\mathcal{S}](\rho; \mathtt{u}, \mathtt{y})$. The pair of labels are used in the concatenation "$(\mathcal{U}, \mathcal{S})$" of a source $\mathcal{U}$ and a system $\mathcal{S}$.

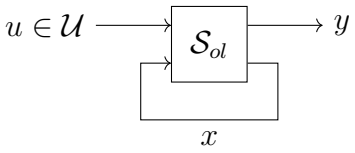
$$
\begin{array}{rl}
[\circ](\rho;\mathtt{y},\varnothing) := & \mathtt{True} \\
[\mathcal{U}_b](\rho;\varnothing,\mathtt{u}) := & \rho_{\mathtt{u}} \in \mathcal{U}_b \\
[\mathcal{S}_b](\rho;\mathtt{v},\mathtt{w}) := & \rho_{\mathtt{w}} = \mathcal{S}_b(\rho_{\mathtt{v}}) \\
[\mathcal{S}_1 \xrightarrow{\mathtt{v}} \mathcal{S}_2](\rho;\mathtt{u},\mathtt{w}) := & [\mathcal{S}_1](\rho;\mathtt{u},\mathtt{v}) \wedge [\mathcal{S}_2](\rho;\mathtt{v},\mathtt{w}) \\
[\mu_{\mathtt{x}}\{\mathcal{S}\}](\rho;\mathtt{u},\mathtt{y}) := & [\mathcal{S}](\rho;(\mathtt{u},\mathtt{x}),(\mathtt{y},\mathtt{x})) \\
[(\mathcal{U},\mathcal{S})](\rho;\mathtt{u},(\mathtt{v},\mathtt{y})) := & [\mathcal{S}](\rho;\mathtt{v},\mathtt{y}) \wedge [\mathcal{U}](\rho;\varnothing,\mathtt{u})
\end{array}
$$

Table 5.2: *An environment ρ belongs to the semantic $\overline{[\mathscr{S}]}$ of an interconnection of systems $\mathscr{S}$ whenever $[\mathscr{S}](\rho;\varnothing,\varnothing)$ is true. $[\mathscr{S}](\rho;\varnothing,\varnothing)$ is computed by structural induction over the syntax of $\mathscr{S}$. $\mathcal{U}_b \subseteq \mathtt{Dom}_u$, $\mathcal{S}_b : \mathtt{Dom}_{\mathtt{v}} \to \mathtt{Dom}_{\mathtt{w}}$, $\mathcal{S}_1$ and $\mathcal{S}_2$ are either a* `Src` *and a* `Sys`*, two* `Sys`*, or a* `Sys` *and a* `Snk`*, $\mathcal{S}$ is a* `Sys`*, and* `u`*,* `v`*,* `w`*,* `x`*, and* `y` *are labels in $\mathbb{V} \cup \{\varnothing\}$.*

In practice, the semantic on the interconnection of systems defines a set of equations over the signals. These equations are of three kinds: the inclusion (defined by the sources "$\mathcal{U}_b \xrightarrow{\mathtt{u}} \ldots$", i.e. $u \in \mathcal{U}_b$), the mapping (defined by the system operator, "$\ldots \xrightarrow{\mathtt{u}} \mathcal{S}_b \xrightarrow{\mathtt{y}} \ldots$", i.e. $y = \mathcal{S}_{ol}(u)$), and a fixed-point equation (defined by the feedback connection, "$\ldots \xrightarrow{\mathtt{u}} \mu_{\mathtt{x}}\{S\} \xrightarrow{\mathtt{y}} \ldots$" i.e. $(y,x) = \mathcal{S}(u,x)$). For any given expression, it is possible to compose the systems, and concatenate the signals such that the semantic is defined by "$\mathcal{U} \xrightarrow{\mathtt{u}} \mu_{\mathtt{x}}\{\mathcal{S}_{ol}\} \xrightarrow{\mathtt{y}} \circ$" (see Example 5.1), which is equivalent to the following set of equations

$$
\begin{cases}
u \in \mathcal{U} & \text{(5.2a)} \\
x = \mathcal{S}_{ol,x}(u,x) & \text{(5.2b)} \\
y = \mathcal{S}_{ol,y}(u,x) & \text{(5.2c)}
\end{cases}
$$

and the following block diagram



where $\mathcal{S}_{ol,x}$ and $\mathcal{S}_{ol,y}$ are projections of the system $\mathcal{S}_{ol}$ over x and y (resp.). In such a case, the semantic of the system is defined as the set of environments $\rho = (u,x,y) \in \mathtt{Dom}$ such that (5.2a,5.2b,5.2c) are satisfied. The goal of this chapter is therefore to find the set of solutions to the fixed-point equation (5.2b) for all the inputs u in $\mathcal{U}$. However, the fixed-point equation (5.2b) is difficult to solve in practice. This is especially true when $\mathtt{Dom}_{\mathtt{x}}$ is a combination of signals of heterogeneous types (e.g.

when x is a combination of discrete-time and continuous-time signals). The following section expresses the system semantic with operators lifted to sets. This semantic results as well in a fixed-point equation in a particular structured set: a partially ordered set. In such a structure, there exist practical ways to solve this fixed-point equation.

Example 5.1.

The formula

$$\text{“}\mathcal{U} \xrightarrow{\mathtt{u}} \mu_{\mathtt{x}} \{\mathcal{A} \xrightarrow{\mathtt{e}} \mu_{\mathtt{x}} \{\mathcal{B}\} \xrightarrow{\mathtt{v}} \mathcal{C}\} \xrightarrow{\mathtt{z}} \circ \text{”} \tag{5.3}$$

describes the interconnection of systems represented in Figure 5.1. The set of

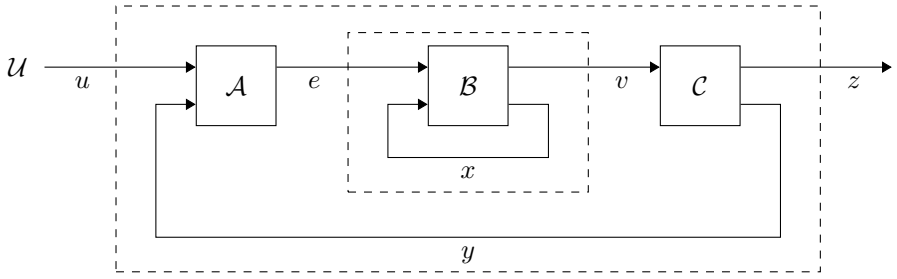


Figure 5.1: *Block diagram of the system described by the formula* (5.3).

signals of the interconnection can be described by the solutions $(u, e, x, y, v, z) \in$ `Dom` *of*

$$\begin{cases} u \in \mathcal{U} \\ e = \mathcal{A}(u, y) \\ (v, x) = \mathcal{B}(e, x) \\ (z, y) = \mathcal{C}(z) \end{cases}$$

The system can be equivalently described by $\text{“}\mathcal{U} \xrightarrow{\mathtt{u}} \mu_{\mathtt{r}} \{\mathcal{S}_{ol}\} \xrightarrow{\mathtt{z}} \circ \text{”}$ *where the block diagram of the open-loop system* $\mathcal{S}_{ol}$ *is represented in Figure 5.2.*

Remark 5.1. Construction of the semantic

For interconnections of systems as well as for computer programs, the formalization of the semantic of a syntax is not unique. [Bouissou and Martel, 2008] considers an interconnection of a computer program and a dynamical system. The semantic of the computer program part is constructed by extending each trace with the next possible reachable state (by using the transition function). In between two transitions of the program, the dynamical system is an autonomous system and can just run toward the future until a new event happens. Then, the

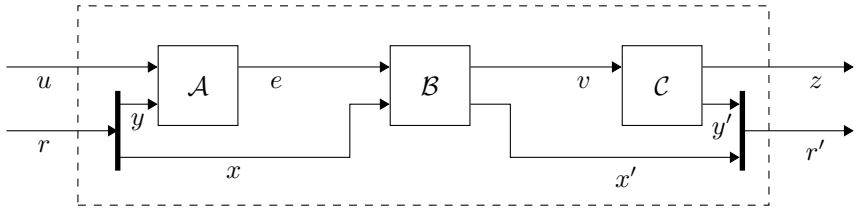


Figure 5.2: *Open-loop system $\mathcal{S}_{ol}$ of* (5.3).

dynamical system can just be modeled as a discrete-time system, and the semantic of the interconnection is expressed in a similar way than for the computer program.

Such an approach is possible because the continuous-time system (the dynamical system) and the discrete-time system (the computer program) defines equations that can be "decoupled" in time. In our case, we do not only consider interconnections of one discrete-time system with one continuous-time system but also interconnections of multiple instances of each type. Therefore, signals can interact with each other, and it is not possible to define a semantic that uses a forward propagation scheme for each systems. We define the semantic of the interconnection of systems as the set of signals that satisfies a set of equations.

Remark 5.2. Initial value problem as fixed-point equation
For $\mathbb{V} = \{\mathtt{x}, \mathtt{u}, \mathtt{y}\}$, $\mathtt{Dom}_\mathtt{x} = \mathscr{L}_2(I; \mathbb{R}^{n_x})$, $\mathtt{Dom}_\mathtt{u} = \mathscr{L}_2(I; \mathbb{R}^{n_u})$, $\mathtt{Dom}_\mathtt{y} = \mathscr{L}_2(I; \mathbb{R}^{n_y})$, and $I = [0, T]$ with $T > 0$. Let $\mathcal{S}_{ol}$ be the system that associates to a time-varying input $(u, x) \in \mathtt{Dom}_\mathtt{u} \times \mathtt{Dom}_\mathtt{y}$, the output $(y, x) \in \mathtt{Dom}_\mathtt{y} \times \mathtt{Dom}_\mathtt{x}$ where

$$\begin{cases} y(t) = h(x(t), u(t)) \\ x(t) = x_0 + \displaystyle\int_0^t f(x(s), u(s))ds \quad \textit{for all } t \geq 0 \end{cases}$$

for a given initial state $x_0 \in \mathbb{R}^{n_x}$, a given observation function $h : \mathbb{R}^{n_u} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}$, and a given dynamical function $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$. Then, the signal x of the feedback connection of "$\ldots \xrightarrow{\mathtt{u}} \mu_\mathtt{x}\{\mathcal{S}_{ol}\} \xrightarrow{\mathtt{y}} \ldots$" satisfies the fixed-point equation

$$x(t) = x_0 + \int_0^t f(x(s), u(s))ds$$

for any $t \in I$. When the solution x is differentiable, this fixed-point equation is equivalent to the initial value problem

$$\begin{cases} \dot{x}(t) \ = f(x(t), u(t)), \ \textit{for all } t \ \in I \\ x(0) = x_0. \end{cases}$$

5.2 Concrete domain and semantic

Now that we have defined the syntax for an interconnection of systems and its associated semantic, we introduce the *concrete domain* and *concrete semantic*. The concrete semantic is defined to be equivalent to the system semantic except that its construction involves operators over sets of signals. The next section defines sufficient conditions to derive a sound approximation of this concrete semantic: the *abstract semantic*.

Extended interconnection of systems In order to solve the fixed-point equation with a set-based approach, we study an extended version of the interconnection of systems (see Remark 5.3). To do so, for each signal u in $\mathtt{Dom}_{\mathtt{u}}$ associated with the label $\mathtt{u} \in \mathbb{V}$, we introduce a new signal u' in $\mathtt{Dom}_{\mathtt{u}'} = \mathtt{Dom}_{\mathtt{u}}$ associated with the new label $\mathtt{u}' \in \mathbb{V}' = \{\mathtt{v}' \,|\, \mathtt{v} \in \mathbb{V}\}$. This extended version of the interconnection of systems redirects any output y associated with $\mathtt{y} \in \mathbb{V}$ of each system to the new signal y' associated with $\mathtt{y}' \in \mathbb{V}$. Then, the solutions to the fixed-point equation are searched in the set of environments ρ satisfying $\rho_{\mathtt{y}} = \rho_{\mathtt{y}'}$ for any label $\mathtt{y} \in \mathbb{V}$. This extended system has a set of labels $\widetilde{\mathbb{V}} = \mathbb{V} \cup \mathbb{V}'$, and a domain $\widetilde{\mathtt{Dom}} = \mathtt{Dom} \times \mathtt{Dom}$.

In the rest of this document, the interconnection of systems implicitly refers to this extended interconnection of systems. We keep the notation $\mathbb{V}$ and $\mathtt{Dom}$ to respectively refer to the set of labels (i.e. to $\widetilde{\mathbb{V}}$) and to the domain (i.e. to $\widetilde{\mathtt{Dom}}$).

Remark 5.3. Lift to sets of a fixed-point equation
Consider the fixed-point equation

$$x = f(x) \tag{5.4}$$

and its set of solutions $\widetilde{X} \subseteq \mathtt{Dom}$, where $f : \mathtt{Dom} \to \mathtt{Dom}$. Such a fixed-point equation can be studied in a set-based approach. Let the lift to sets F of f be defined by

$$F(X) = \{f(x) \,|\, x \in X\}. \tag{5.5}$$

The solutions to the fixed-point equation

$$X = F(X) \tag{5.6}$$

are closely related to the fixed-points of f. Every subset $X \subseteq \widetilde{X}$ of fixed-points to (5.4) *is a fixed-point of F. For every pair of fixed-points X, Y of F, $X \cup Y$ is as well a fixed-point of F. And therefore, if Z is a fixed-point of F, then $Y = Z \cup \widetilde{X}$ is as well a fixed-point of F, i.e. there is a fixed-point Y to F larger (or equal) than $\widetilde{X}$. Thus,* (5.6) *might introduce elements of* $\mathtt{Dom}$ *that are not solutions to*

the fixed-point equation (see Example 5.2). To avoid this, we reformulate (5.4) *with*

$$\lambda = g(\lambda) \tag{5.7}$$

where $\lambda = (x, x') \in \mathtt{Dom}^2$ *and* g *is an extended version of* f *defined by*

$$g(\lambda) = (x', f(x)).$$

In fact, $\lambda = (x, x')$ *is a fixed-point of* g *iff* x *is a fixed-point of* f *(the proof is direct from* (5.7) *as* $x = x' = f(x)$*). Let* $\widetilde{\Lambda}$ *be the set of fixed-points of* g*. Similarly than for* f*, the lift to sets* G *of* g *can be used to characterize* $\widetilde{\Lambda}$*. Let* $\Sigma_x = \{(x, x') \in \mathtt{Dom}^2 | x = x'\}$ *and* Λ *be a fixed-point of* G *such that* $\Lambda \subseteq \Sigma_x$*. For* $\lambda_1 = (x_1, x_1') \in \Lambda$*, since* Λ *is a fixed-point of* G*, there is a* $\lambda_2 \in \Lambda$ *such that* $\lambda_1 = g(\lambda_2)$*, i.e.*

$$\begin{cases} x_1 = x_2' \\ x_1' = f(x_2) \end{cases}$$

Since λ_1 *and* λ_2 *belong to* $\Lambda \subseteq \Sigma_x$*, it holds* $x_1 = x_1'$ *and* $x_2 = x_2'$ *and therefore,*

$$x_1' = f(x_1).$$

To summarize, when $G(\Lambda) = \Lambda$ *and* $\Lambda \subseteq \Sigma_x$*, every* x*, such that* $\lambda = (x, x') \in \Lambda$*, is a fixed-point of* f*. The projection* $\Lambda|_x$ *of* Λ *is a subset of* $\widetilde{X}$*, i.e.* $\Lambda|_x \subseteq \widetilde{X}$*. For this reason,* $\widetilde{\Lambda}$ *(and* $\widetilde{\Lambda}$*) can be characterized by searching for the largest fixed-point of* G *lower than* Σ_x*.*

Example 5.2.
Let f *be the polynomial defined by* $f(x) = x^2 - 1$*. The fixed-points of* f *are the roots of the second degree equation* $f(x) - x = 0$*. Thus,* f *has two fixed-points* $\widetilde{X} = \{x_1, x_2\}$*, where* $x_1 = \frac{-1+\sqrt{5}}{2}$ *and* $x_2 = \frac{-1-\sqrt{5}}{2}$*. Similarly, the fixed-points of* $f \circ f$ *are the roots of the fourth degree equation* $f \circ f(x) - x = 0$*.* $f \circ f$ *has four fixed-points which are* $X = \{-1, 0, x_1, x_2\}$ *(see Figure 5.3). Let the lift to sets of* f *be* F *(as defined in Remark 5.3).* X *is a fixed-point of* F *since* $F(X) = \{0, -1, x_1, x_2\} = X$*. In a more general way, all the sets defined by* $X = \{x_1, x_2, \ldots, x_m\}$ *that are composed of elements of a periodic sequence* $\{x_n\}_{n\in\mathbb{N}}$ *(of period* $m \in \mathbb{N}$*) defined by* $x_{n+1} = f(x_n)$ *are fixed-points of* F *(indeed* $F(X) = \{f(x_1), f(x_2), \ldots, f(x_m)\} = \{x_2, x_3, \ldots, x_1\} = X$*).*

5.2.1 Concrete domain

The concrete domain $\mathcal{A}$ is defined as the powerset of environments, i.e. $\mathcal{A} = \wp(\mathtt{Dom})$. We equip this concrete domain with a complete lattice structure.

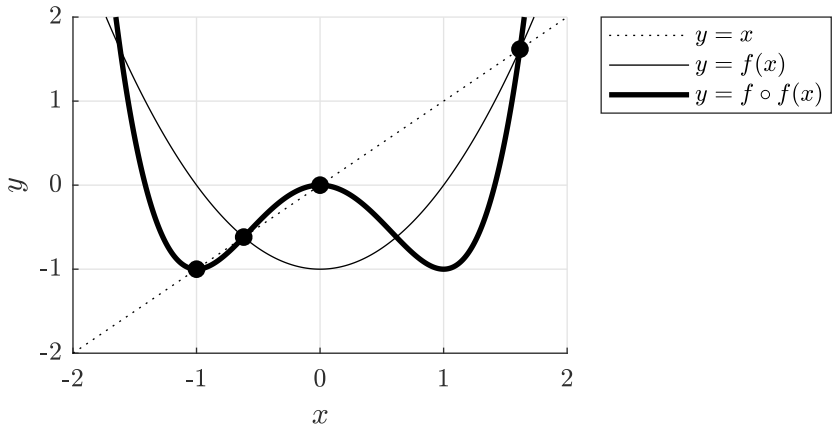


Figure 5.3: *Fixed-points of f and $f \circ f$ in Example 5.2.*

Definition 5.1. Partially ordered set

A partially ordered set (poset) is a tuple $(\mathcal{A}, \sqsubseteq)$ where $\mathcal{A}$ is a set and $\sqsubseteq$ is a partial order relationship, i.e.

$$\begin{array}{ll} \forall x \in \mathcal{A}, x \sqsubseteq x & \textit{(reflectivity)} \\ \forall x, y \in \mathcal{A}, x \sqsubseteq y \wedge y \sqsubseteq x \Rightarrow x = y & \textit{(antisymmetry)} \\ \forall x, y, z \in \mathcal{A}, x \sqsubseteq y \wedge y \sqsubseteq z \Rightarrow x \sqsubseteq z & \textit{(transitivity)} \end{array}$$

Definition 5.2. Lattice

A lattice is a tuple $(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap)$ with

$(\mathcal{A}, \sqsubseteq)$ a poset,

$\sqcup$ a binary upperbound, i.e.

$\forall x, y \in \mathcal{A}, x \sqsubseteq x \sqcup y \wedge y \sqsubseteq x \sqcup y$

$\forall x, y, z \in \mathcal{A}, x \sqsubseteq z \wedge y \sqsubseteq z \Rightarrow x \sqcup y \sqsubseteq z$

$\sqcap$ a binary lowerbound, i.e.

$\forall x, y \in \mathcal{A}, x \sqcap y \sqsubseteq x \wedge x \sqcap y \sqsubseteq y$

$\forall x, y, z \in \mathcal{A}, z \sqsubseteq x \wedge z \sqsubseteq y \Rightarrow z \sqsubseteq x \sqcup y$

Definition 5.3. Complete lattice

A complete lattice is a tuple $(\mathcal{A}, \sqsubseteq, \sqcup)$ where

$\mathcal{A}$ a set,

$\sqsubseteq$ a partial order relationship (reflective, antisymmetric and transitive),

$\sqcup$ a lowerbound, i.e., for all subset S of $\mathcal{A}$,

$\forall a \in S, a \subseteq \sqcup S$

$\forall b \in S, (\forall a \in S, a \sqsubseteq b) \Rightarrow \sqcup S \sqsubseteq b$

The concrete domain $(\mathcal{P}(\texttt{Dom}), \subseteq, \cup, \cap, \texttt{Dom}, \emptyset)$ is a complete lattice.

5.2.2 Concrete semantic

Each construction in the syntax (defined in Section 5.1) can be associated with an operator over the concrete domain. The concrete semantic is then defined by structural induction over the syntax. We will see that this concrete semantic can be expressed as the greatest fixed-point of a monotonic operator in the concrete domain. In the first paragraph, we present known results for the computation of such a greatest fixed-point. In the second paragraph, we define the concrete semantic of the interconnection of systems.

Operators and fixed-points in a complete lattice Let a complete lattice $(\mathcal{A}, \sqsubseteq, \sqcap, \top, \bot)$ and an operator $F : \mathcal{A} \to \mathcal{A}$. F is *monotonic* when F preserves the ordering relationship, i.e. $X \sqsubseteq Y \Rightarrow F(X) \sqsubseteq F(Y)$ for every $X, Y \in \mathcal{A}$. A fixpoint Y of F is an element that satisfies $Y = F(Y)$. Let $\mathrm{gfp}_X \{F\}$ be the greatest fixpoint of F that is lower than X. Y is a *pre-fixpoint* of F if $Y \sqsubseteq F(Y)$, and Y is a *post-fixpoint* of F if $F(Y) \sqsubseteq Y$. The following theorem provides the existence of such a fixed-point.

Theorem 5.1. Knaster-Tarski
The set of fixed-points of a monotonic operator F in a complete lattice is a complete lattice. Moreover, the greatest fixed-point of F that is lower than X is

$$\mathrm{gfp}_X \{F\} = \bigsqcup \{Y \mid Y \sqsubseteq X \text{ and } Y \sqsubseteq F(Y)\}.$$

Proof. See [Tarski, 1955]. ♢

By Theorem 5.1, the set of fixed-points of F lower than $X \in \mathcal{A}$ is a complete lattice, there is a lowest and a greatest fixed-point. The monotonicity of the operator F can be used to define a decreasing sequence of Y_k in $\mathcal{A}$ starting from a post-fixpoint of F. When the initial element Y_0 is greater than X, then each element is greater than $\mathrm{gfp}_X \{F\}$.

Proposition 5.1. Descending chains
If F is a monotonic operator and X is a post-fixpoint of F, then

$$\mathrm{gfp}_X \{F\} \sqsubseteq F^{k+1}(X) \sqsubseteq F^k(X) \sqsubseteq \cdots \sqsubseteq F(X) \sqsubseteq X$$

.

Proof. It is a direct consequence of the monotonicity of F applied to the post-fixpoint X of F. ♢

Concrete semantic The concrete semantic is defined in Definition 5.4 as the greatest fixed-point of the concrete operator $X \mapsto [\![\mathscr{S}]\!](X;\varnothing,\varnothing)$ defined by Table 5.3 by structural induction over the syntax of an interconnection of systems $\mathscr{S}$ (as introduced in Table 5.1). Similarly than for the semantic of the interconnection of systems, for every expression $\mathcal{S}$ of the syntax of Table 5.1, we define an operator $[\![\mathcal{S}]\!](R;\mathtt{u},\mathtt{y})$ in the concrete semantic that associates to a set of environments $R \subseteq \mathtt{Dom}$, an input label $\mathtt{u}$, and an output label $\mathtt{y}$, another set of environments in $\mathtt{Dom}$:

$$[\![\mathcal{S}]\!] : \wp(\mathtt{Dom}) \times \overline{\mathbb{V}} \times \overline{\mathbb{V}} \to \wp(\mathtt{Dom})$$

where $\overline{\mathbb{V}}$ is defined by $\overline{\mathbb{V}} = \{\varnothing\} \cup \mathbb{V} \cup \mathbb{V}^2$.

$$\begin{array}{rcl}
[\![\circ]\!](R;\mathtt{y},\varnothing) := & R \\
[\![\mathcal{U}_b]\!](R;\varnothing,\mathtt{u}) := & \{\rho \in R \,|\, \rho_{\mathtt{u}'} \in \mathcal{U}\} \\
[\![\mathcal{S}_b]\!](R;\mathtt{v},\mathtt{w}) := & \{\rho[\mathtt{w}' \leftarrow \mathcal{S}_b(\rho_{\mathtt{v}}), \mathtt{v} \leftarrow \rho_{\mathtt{v}'}] \mid \rho \in R\} \\
[\![\mathcal{S}_1 \xrightarrow{\mathtt{v}} \mathcal{S}_2]\!](R;\mathtt{u},\mathtt{w}) := & [\![\mathcal{S}_2]\!]\Big([\![\mathcal{S}_1]\!](R;\mathtt{u},\mathtt{v});\mathtt{v},\mathtt{w}\Big) \\
[\![\mu_{\mathtt{x}}\{\mathcal{S}\}]\!](R;\mathtt{u},\mathtt{y}) := & [\![\mathcal{S}]\!](R;(\mathtt{u},\mathtt{x}),(\mathtt{y},\mathtt{x})) \\
[\![(\mathcal{U},\mathcal{S})]\!](R;\mathtt{u},(\mathtt{v},\mathtt{y})) := & [\![\mathcal{S}]\!](R;\mathtt{v},\mathtt{y}) \cap [\![\mathcal{U}]\!](R;\varnothing,\mathtt{u})
\end{array}$$

Table 5.3: *Concrete operator for an interconnection of systems.* $\mathcal{U}_b \subseteq \mathtt{Dom}_{\mathtt{u}}$, $\mathcal{S}_b : \mathtt{Dom}_v \to \mathtt{Dom}_w$, $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *are either a* `Src` *and a* `Sys`*, two* `Sys`*, or a* `Sys` *and a* `Snk`*,* $\mathcal{S}$ *is a* `Sys`*, and* $\mathcal{U}$ *is a* `Src`*.* `u`*,* `v`*,* `w`*,* `x`*, and* `y` *are labels in* $\mathbb{V} \cup \{\varnothing\}$*.*

We now study the fixed-points lower that $\Sigma \subseteq \mathtt{Dom}$ of the concrete operator $X \mapsto [\![\mathscr{S}]\!](X)$ where Σ is defined as the set of environments where each signal is equal to its prime signal (see Remark 5.3), i.e.

$$\Sigma = \{\rho \in \mathtt{Dom} \,|\, \forall \mathtt{v} \in \mathbb{V}, \rho_{\mathtt{v}} = \rho_{\mathtt{v}'}\}\,. \tag{5.8}$$

Each environment ρ of the semantic is a fixed-point of any concrete operator since, by construction, $\rho \in \Sigma$, and since each proposition satisfied in Table 5.1 implies that $\{\rho\}$ is a fixed-point of its corresponding operator in Table 5.2. More generally, any subset of the semantic of the interconnection is a fixed-point of the concrete semantic.

Proposition 5.2. Sets of trajectories are fixed-points of the semantic
Every $R \subseteq \overline{[\mathcal{S}]}$ *is a fixed-point of* $X \mapsto [\![\mathscr{S}]\!](X)$*.*

Proposition 5.3. Fixed-point of the concrete semantic
Every fixed-point R *of* $X \mapsto [\![\mathscr{S}]\!](X)$ *is a subset of* $\overline{[\mathscr{S}]}$*.*

Proof. Let $\rho \in R = [\![\mathscr{S}]\!](R)$, we want to show that ρ is a system trajectory, i.e. $[\mathscr{S}](\rho; \varnothing, \varnothing)$ is true. Since R is a fixed-point of $X \mapsto [\![\mathscr{S}]\!](X)$ and each operator only modifies a projection of R, and R is as well a fixed-point for each operator. Moreover, for every label $\mathtt{u} \in \mathbb{V}$, $R \subseteq \Sigma$, i.e. $\rho_{\mathtt{u}} = \rho_{\mathtt{u}'}$. We prove that ρ is a trajectory by induction. For each operator, we prove that $R = [\![\mathcal{S}]\!](R; \mathtt{u}, \mathtt{y})$ implies that $[\mathcal{S}](\rho; \mathtt{u}, \mathtt{y})$ is true:

- $R = [\![\mathcal{U}_b]\!](R; \mathtt{u})$ implies that $\rho_{\mathtt{u}'} \in \mathcal{U}_b$, since $\rho \in \Sigma$, it holds $\rho_{\mathtt{u}} = \rho_{\mathtt{u}'} \in \mathcal{U}_b$, therefore, $[\mathcal{U}_b](\rho; \mathtt{u})$ is true;
- $R = [\![\mathcal{S}_b]\!](R; \mathtt{v}, \mathtt{w})$ implies that there is a ρ' such thats $\rho_{\mathtt{w}'} = \mathcal{S}_b(\rho'_{\mathtt{v}})$ and $\rho_{\mathtt{v}} = \rho'_{\mathtt{v}'}$, since $\rho, \rho' \in \Sigma$, $\rho_{\mathtt{v}'} = \rho_{\mathtt{v}}$ and $\rho'_{\mathtt{v}'} = \rho'_{\mathtt{v}}$, therefore, $\rho_{\mathtt{w}} = \mathcal{S}_b(\rho_{\mathtt{v}})$, i.e. $[\mathcal{S}_b](\rho; \mathtt{v}, \mathtt{w})$ is true and moreover $\rho'_{\mathtt{l}} = \rho_{\mathtt{l}}$ for $\mathtt{l} \in \{\mathtt{v}, \mathtt{v}', \mathtt{w}, \mathtt{w}'\}$;
- $R = [\![\mathcal{S}_1 \xrightarrow{\mathtt{v}} \mathcal{S}_2]\!](R; \mathtt{u}, \mathtt{w})$, since R is also a fixed a point of $[\![\mathcal{S}_1]\!](\cdot; \mathtt{u}, \mathtt{v})$, it holds $R = [\![\mathcal{S}_1]\!](R; \mathtt{u}, \mathtt{v})$ and $R = [\![\mathcal{S}_2]\!](R; \mathtt{v}, \mathtt{w})$, for $\rho \in R$, by induction, if $[\mathcal{S}_1](\rho; \mathtt{u}, \mathtt{v})$ and $[\mathcal{S}_2](\rho; \mathtt{v}, \mathtt{w})$ are true, then $[\mathcal{S}_1 \xrightarrow{\mathtt{v}} \mathcal{S}_2](\rho; \mathtt{u}, \mathtt{w})$ is true;
- the last operators have a similar proof than the one for the serial connection. ♢

Proposition 5.2 and Proposition 5.3 imply that $\overline{[\mathscr{S}]}$ is a fixed-point of $X \mapsto [\![\mathscr{S}]\!](X)$ and that every fixed-point $R \subseteq \Sigma$ of $X \mapsto [\![\mathscr{S}]\!](X)$ is a subset of $\overline{[\mathscr{S}]}$, i.e. $R \subseteq \overline{[\mathscr{S}]}$. Therefore, the semantic of the interconnection of systems is equal to the greatest fixed-point of $X \mapsto [\![\mathscr{S}]\!](X)$.

Definition 5.4. Concrete semantic as the greatest fixed-point of a monotonic operator

Let $\Sigma \subseteq \mathtt{Dom}$ defined as in (5.8)*. The concrete semantic $\overline{[\![\mathscr{S}]\!]}$ is defined by*

$$\overline{[\![\mathscr{S}]\!]} = \mathrm{gfp}_{\Sigma} \{X \mapsto [\![\mathscr{S}]\!](X; \varnothing, \varnothing)\}$$

where the monotonic operator $X \mapsto [\![\mathscr{S}]\!](X; \varnothing, \varnothing)$ is defined in Table 5.3 by structural induction over the syntax of the interconnection of systems $\mathscr{S}$.

Proposition 5.4. Concrete semantic

The semantic of the interconnection of systems $\mathscr{S}$ is equal to the concrete semantic, i.e.

$$\overline{[\mathscr{S}]} = \overline{[\![\mathscr{S}]\!]}.$$

Remark 5.4. Concrete semantic as the greatest fixed-point of a monotonic operator

The approach to express the system trajectories (or the computer program traces) as a fixed-point of a monotonic operator in a complete lattice structure is classical. However, the fixed-point of interest is usually the lowest fixed-point of a monotonic operator.

Expressing the semantic as the fixed-point has been initially introduced to study computer program [Cousot and Cousot, 1979]. In these works, the concrete domain is represented with the sets of finite and/or infinite traces, the fixed-point is iteratively constructed by extending each trace through time. Then, the fixed-point of interest is the lowest fixed-point that contains these trajectories. [Bouissou and Martel, 2008] has extended this approach to the case of an interconnection of a controller (modeled as a computer program) and dynamical systems. The trajectories of the dynamical system are simulated over a small time-horizon.

As we pointed out in Remark 5.1, such an approach is possible because systems are be decoupled in time. In our case, the system trajectories are expressed as the solution to a fixed-point equation that depends over an input. For each input, there might be one (or more) solution to this fixed-point equation. The semantic of the system is then defined as the union of all these fixed-point solutions for every possible input. The system semantic can be equivalently defined with operators over sets. For these monotonic operators, the union of fixed-points corresponds to the greatest fixed-point (by the Theorem 5.1).

5.3 Abstract domains

The previous section defines the concrete semantic as the greatest fixed-point of an operator in the concrete domain. It is then possible to use an iterative method to overapproximate this fixed-point (see Proposition 5.1). However, elements of the concrete domain are too complex to be computer-represented and therefore, the concrete semantic cannot be calculated. In this section, we soundly approximate it with elements chosen in a subset of the concrete domain. This sound approximation of the concrete domain is called the *abstract domain*. The abstract domain is supplied with a complete lattice structure where elements and operators (resp.) can be represented and calculated (resp.) on a computer program. The *abstract semantic* is then deduced from the concrete semantic. Each operator of the abstract semantic mimics its associated operator in the concrete semantic.

The concrete domain and abstract domain are linked with a so-called *Galois connection*. This Galois connection enforces the soundness property. Each element of the concrete domain is associated with an abstract element that is a sound approximation (i.e. an overapproximation). The abstract semantic is derived such that the Galois connection is preserved through each evaluation in the concrete semantic.

To simplify notations, the complete lattice structure associated with the concrete domain $(\wp(\mathtt{Dom}), \subseteq, \cup, \cap)$ is denoted with $(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap)$. Each element in $\mathcal{A}$ corre-

sponds to a set of signals of the interconnection. Abstracting the concrete semantic corresponds to describing this set of signals with elements in a $\overline{\mathcal{A}} \subseteq \mathcal{A}$. In practice, elements of $\overline{\mathcal{A}}$ are chosen to be computer representable (such as time-varying intervals or time-varying ellipsoids for example). By using an isomorphism from $\bar{\mathcal{A}}$ to a set $\mathcal{A}^*$, each element $\bar{P} \in \bar{\mathcal{A}}$ is associated with an abstract element $\bar{P}^*$ of an *abstract domain* $\mathcal{A}^*$.

Definition 5.5. Galois connection
Let $(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap)$ *and* $(\mathcal{A}^*, \sqsubseteq^*, \sqcup^*, \sqcap^*)$ *be two complete lattices. A pair of functions* (α, γ)*, with* $\alpha : \mathcal{A} \to \mathcal{A}^*$ *and* $\gamma : \mathcal{A}^* \to \mathcal{A}$*, is a Galois connection if it holds*

$$\forall x \in \mathcal{A}, \forall x^* \in \mathcal{A}^*, \alpha(x) \sqsubseteq^* x^* \Leftrightarrow x \sqsubseteq \gamma(x^*)$$

We can enumerate a few properties of the Galois connection.

Theorem 5.2. Galois connection
Let $(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap)$ *and* $(\mathcal{A}^*, \sqsubseteq^*, \sqcup^*, \sqcap^*)$ *be two complete lattices and* (α, γ) *be a Galois connection between* $\mathcal{A}$ *and* $\mathcal{A}^*$

$$(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} (\mathcal{A}^*, \sqsubseteq^*, \sqcup^*, \sqcap^*).$$

Then, the following properties are satisfied
α *is a monotonic function,*
γ *is a monotonic function,*
$\forall x^* \in \mathcal{A}^*, \alpha \circ \gamma(x^*) \sqsubseteq^* x^*$*, and*
$\forall x \in \mathcal{A}, x \sqsubseteq \gamma \circ \alpha(x)$.

The *abstraction function* α associates to any concrete element $x \in \mathcal{A}$ an abstract element $x^* = \alpha(x)$. The *concretisation function* γ associates to any abstract element $x^* \in \mathcal{A}^*$ a concrete element $x = \gamma(x^*)$.

Sometimes the existence of a Galois connection between two complete lattices is too strong of a requirement as the abstraction function α might not exist. For such cases, [Cousot and Cousot, 1992] proposes to relax the Galois connection framework to work only with the concretisation function γ.

Definition 5.6. Concretisation function
Let $(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap)$ *and* $(\mathcal{A}^*, \sqsubseteq^*, \sqcup^*, \sqcap^*)$ *be two complete lattices. A concretisation function is monotonic function* $\gamma : \mathcal{A}^* \to \mathcal{A}$*.* X^* *is an abstraction of* X *when* $X \subseteq \gamma(X^*)$.

5.4 Abstract semantic

The previous section showed that the concrete domain and abstract domain can be both provided with a mathematical structure of complete lattices. In what follows, we assume to have a Galois connection between the concrete domain $(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap)$ and an abstract domain $(\mathcal{A}^*, \sqsubseteq^*, \sqcup^*, \sqcap^*)$

$$(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} (\mathcal{A}^*, \sqsubseteq^*, \sqcup^*, \sqcap^*).$$

Now that abstract domains are defined, we should define the abstract semantic. We will see that the abstract semantic mimics the concrete semantic: each function evaluation in the concrete domain gives rise to a function evaluation in the abstract domain, each fixed-point computation in the concrete domain gives rise to a fixed-point computation in the abstract domain. For a given concrete semantic $\overline{[\![\mathscr{S}]\!]}$, we can compute an abstract semantic $\overline{[\![\mathscr{S}]\!]}^*$. This abstract semantic is built by structural induction over the syntax of the system interconnection. This semantic is *sound* when $\alpha(\overline{[\![\mathscr{S}]\!]}) \sqsubseteq^* \overline{[\![\mathscr{S}]\!]}^*$, or equivalently when $\overline{[\![\mathscr{S}]\!]} \sqsubseteq \gamma(\overline{[\![\mathscr{S}]\!]}^*)$. The soundness property is obtained by using sound operator evaluation and fixed-point computation over abstract elements.

In Section 5.4.1, we detail the abstract counterpart of the concrete functions. In Section 5.4.2, we detail the abstract fixed-point computation.

5.4.1 Abstract evaluation

For a function $F : \mathcal{A} \to \mathcal{A}$ in the concrete domain, it is sometimes possible to define $F^* : \mathcal{A}^* \to \mathcal{A}^*$ an "abstract evaluation" of F. To ensure the soundness of the abstract semantic, the soundness of F^* with respect to the Galois connection and F should be ensured.

Definition 5.7. Sound approximation of a function
For a function $F : \mathcal{A} \to \mathcal{A}$, $F^* : \mathcal{A}^* \to \mathcal{A}^*$ is a sound approximation of F whenever

$$\forall X^* \in \mathcal{A}^*, F \circ \gamma(X^*) \sqsubseteq \gamma \circ F^*(X^*)$$

5.4.2 Abstract fixed-points

The previous section details how the abstract counterpart of concrete functions can be defined. We now detail the computation of the greatest fixed-point that appears in the concrete semantic. Since the greatest fixed-point in the concrete domain exists but is not necessarily computable, we aim at computing a sound approximation of the greatest fixed-point in the abstract domain by using a fixed-point transfer theorem:

Theorem 5.3. Fixed-point transfer
Given a Galois connection

$$(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} (\mathcal{A}^*, \sqsubseteq^*, \sqcup^*, \sqcap^*),$$

for $F^* : \mathcal{A}^* \to \mathcal{A}^*$ a monotonic function and $F : \mathcal{A} \to \mathcal{A}$ a monotonic function. When F^* is a sound approximation of F then

$$\mathrm{gfp}_{\gamma(X^*)} \{F\} \sqsubseteq \gamma\big(\mathrm{gfp}_{X^*} \{F^*\}\big)$$

for every $X^* \in \mathcal{A}^*$.

Proof. See Theorem 2 in [Cousot, 2002] ♢

When each operator is sound, the greatest fixed-point in the concrete domain is soundly approximated by the abstract greatest fixed-point in the abstract domain.

Practical computation of fixed-point

Theorem 5.3 gives a way to soundly approximate the fixed-point of the concrete domain by computing a greatest fixed-point in the abstract domain. Contrary to the concrete domain, the abstract domain is computer-representable and abstract evaluations are computable. This greatest fixed-point can be computed using an iterative approach as in Property 5.1: let assume that for a $D^* \in \mathcal{A}^*$, $\gamma(D^*)$ overapproximates $\mathrm{gfp}_X \{F\}$, i.e.

$$\mathrm{gfp}_X \{F\} \sqsubseteq \gamma(D^*),$$

that D^* is a post-fixpoint of F^*, i.e.

$$F^*(D^*) \sqsubseteq^* D^*,$$

and that F^* is a monotonic operator, then the sequence $\{Y_k^*\}$ of iterates, defined by

$$\begin{cases} Y_{k+1}^* = F^*(Y_k^*) \\ Y_0^* = D^* \end{cases} \tag{5.9}$$

is monotonically decreasing, i.e. $Y_{k+1}^* \sqsubseteq^* Y_k^*$ for every $k \geq 0$, and each iterate is a sound approximation of $\mathrm{gfp}_X \{F\}$, i.e.

$$\mathrm{gfp}_X \{F\} \sqsubseteq \cdots \sqsubseteq \gamma(Y_k^*) \sqsubseteq \gamma(Y_{k-1}^*) \sqsubseteq \cdots \sqsubseteq \gamma(D^*),$$

they correspond to a refinement of $\gamma(D^*)$.

Partial narrowing operator In practice, when Y_0^* is an abstraction of $\mathrm{gfp}_X\,\{F\}$, it is always possible to refine Y_0^* with a sound approximation of $Y \mapsto F(Y) \sqcap Y$. The produced sequence decreases more rapidly than the sequence derived by (5.9) and is still a sound approximation of $\mathrm{gfp}_X\,\{F\}$. An abstract operator $\triangle^*$ is a *narrowing operator* if it is a sound approximation of the operator $\sqcap^*$ and such that for any $X^*, Y^* \in \mathcal{A}^*$ it holds $X^* \triangle^* Y^* \sqsubseteq^* Y^*$. Classical definitions of the narrowing operator usually require that the sequence $\{Y_k^*\}$ generated by

$$\begin{cases} Y_{k+1}^* = F^*(Y_k^*) \triangle^* Y_k^* \\ Y_0^* = D^* \end{cases}$$

is ultimately stationary in finite-time (see Definition 2.2.4 in [Miné, 2004]), such property ensure the termination of iterates. In this work, we use a relaxed form of the narrowing operator where the ultimately stationary property is not required and where the property $X^* \triangle^* Y^* \sqsubseteq^* Y^*$ for every X^* and Y^* in $\mathcal{A}^*$ holds for a different partial order $\sqsubseteq$. Then, iterates Y_k^* decreases with respect to the partial ordering $\sqsubseteq$.

Definition 5.8. Partial narrowing operator

An abstract binary operator $\triangle^*$ is a narrowing with respect to the partial order $\sqsubseteq^*$ if and only if, for all $X^*, Y^* \in \mathcal{A}^*$, it holds

$(Y^* \sqcap^* X^*) \sqsubseteq^* (Y^* \triangle^* X^*)$, *and*

$(Y^* \triangle^* X^*) \sqsubseteq^* X^*$.

5.4.3 Abstract semantic

In Section 5.4.1 and Section 5.4.2, we defined the abstract counterparts of each operator of the concrete semantic (as defined in Section 5.2). The abstract semantic is then defined by structural induction overt the syntax (defined in Section 5.1). When there is a Galois connection between the concrete domain and when each operator used within the abstract semantic is sound, then the abstract semantic is sound with respect to the concrete semantic.

Proposition 5.5. Soundness of the abstract semantic

Provided the abstract operator are sound, the abstract semantic is sound with respect to the concrete one:

$$\overline{[\![\mathscr{S}]\!]} \sqsubseteq \gamma\big(\overline{[\![\mathscr{S}]\!]}^*\big).$$

5.5 Abstractions for vectorial space of finite dimension

The previous sections detailed the classical framework of abstract interpretation and the concrete semantic of an interconnection of dynamical systems. The following two sections describe abstract domains used to abstract elements of the concrete domains. This section details abstractions for constant signals such as system parameters or an initial state of a system. The next section uses a point-wise lift of these abstract domains to represents time-varying signals.

5.5.1 Ellipsoidal domain

In this section, we define the domain of convex sets, this domain can be supplied with a complete lattice structure. Since every convex set can be described as an intersection of ellipsoids, we then define the abstract domain of ellipsoidal sets. Finally, we define a narrowing operator that allows to describe a sound approximation of convex sets with fewer ellipsoidal sets.

Let $\texttt{Conv}^n$ be the set of convex subsets of $\mathbb{R}^n$. $(\texttt{Conv}^n, \subseteq)$ is a partially ordered set. Since an intersection of convex sets is convex, $c_1, c_2 \in \texttt{Conv}^n$ implies that $c_1 \cap c_2 \in \texttt{Conv}^n$. A union $c_1 \cup c_2$ of convex sets c_1 and c_2 is not convex in the general case. However, its convex hull $\texttt{hull}(c_1 \cup c_2)$ (where $\texttt{hull}(X)$ is defined as the intersection of all the convex sets greater than X, $\texttt{hull}(X) = \bigcap_{X \subseteq C} C$) is convex and it soundly approximates the union $\cup$. Let $\cup_c$ be the convex hull-union operator, i.e. $c_1 \cup_c c_2 = \texttt{hull}(c_1 \cup c_2)$. To this respect

Proposition 5.6. Convex set domain
$(\texttt{Conv}^n, \subseteq, \cup_c, \cap, \emptyset, \mathbb{R}^n)$ *is a complete lattice.*

Proposition 5.7. Galois connection of the convex set domain
The pair of functions $\alpha_{\texttt{C}} = \texttt{hull}$ *and* $\gamma_{\texttt{C}} = id$ *is a Galois connection between the two complete lattices* $(\wp(\mathbb{R}^n), \subseteq, \cup, \cap, \emptyset, \mathbb{R}^n)$ *and* $(\texttt{Conv}^n, \subseteq, \cup_c, \cap, \emptyset, \mathbb{R}^n)$.

Proof. Since $\texttt{hull}(X)$ is defined as the intersection of all the convex sets greater than X, the equivalence $\texttt{hull}(X) \subseteq C \Leftrightarrow X \subseteq C$ holds for all $X \subseteq \mathbb{R}^n$ and $C \in \texttt{Conv}^n$. $\diamondsuit$

Any convex set $C \in \texttt{Conv}^n$ can be described as an intersection of ellipsoids, i.e. $C = \bigcap_{E \in \mathcal{E}} E$ for $\mathcal{E} \subseteq \texttt{Elli}^n$. But contrary to $\texttt{Conv}^n$, ellipsoids in $\texttt{Elli}^n$ can be represented with a symmetric matrix in $\mathbb{S}^{(n+1)\times(n+1)}$. Since matrices can be represented and manipulated in a computer program, the manipulation of convex sets as an intersection of ellipsoidal sets is easier.

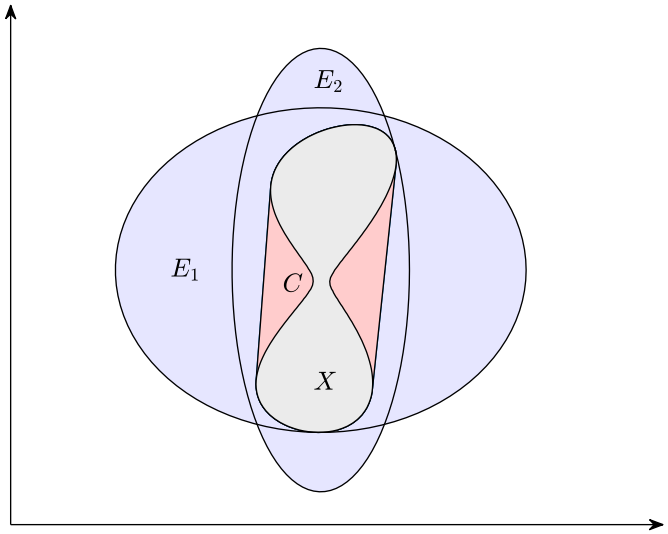


Figure 5.4: *The set $X \subseteq \mathbb{R}^2$ is soundly approximated by its convex envelop $C = \alpha_{\mathtt{C}}(X) = \mathtt{hull}(X)$. The set C is equivalently described with $C = \gamma_{\mathtt{E}}(\mathcal{E})$ the intersection of all ellipsoids overapproximating C (such as $E_1, E_2 \in \mathcal{E}$).*

The power set of ellipsoids $\wp(\mathtt{Elli}^n)$ can be provided with a complete lattice structure. For $\mathcal{E}_1, \mathcal{E}_\in \in \wp(\mathtt{Elli}^n)$, we define the following relationships $\mathcal{E}_1 \sqsubseteq_{\mathtt{E}} \mathcal{E}_2$ iff $\mathcal{E}_2 \subseteq \mathcal{E}_1$, $\mathcal{E}_1 \sqcup_{\mathtt{E}} \mathcal{E}_2 = \mathcal{E}_1 \cap \mathcal{E}_2$, and $\mathcal{E}_1 \sqcap_{\mathtt{E}} \mathcal{E}_2 = \mathcal{E}_1 \cup \mathcal{E}_2$.

Proposition 5.8. Ellipsoidal domain
$(\wp(\mathtt{Elli}^n), \sqsubseteq_{\mathtt{E}}, \sqcup_{\mathtt{E}}, \sqcap_{\mathtt{E}}, \emptyset, \mathbb{R}^n)$ is a complete lattice.

Proposition 5.9. Galois connection of ellipsoidal domain
The pair of functions $\alpha_{\mathtt{E}}(C) = \{E \in \mathtt{Elli}^n \,|\, C \subseteq E\}$ and $\gamma_{\mathtt{E}}(\mathcal{E}) = \bigcap_{E \in \mathcal{E}} E$ is a Galois connection between the two complete lattices $(\mathtt{Conv}^n, \subseteq, \cup_c, \cap, \emptyset, \mathbb{R}^n)$ and $(\wp(\mathtt{Elli}^n), \sqsubseteq_{\mathtt{E}}, \sqcup_{\mathtt{E}}, \sqcap_{\mathtt{E}}, \emptyset, \mathbb{R}^n)$.

The Galois connections can be composed, and therefore, there is a Galois connection between the complete lattice $(\wp(\mathbb{R}^n), \subseteq, \cup, \cap, \emptyset, \mathbb{R}^n)$ and $(\wp(\mathtt{Elli}^n), \sqsubseteq_{\mathtt{E}}, \sqcup_{\mathtt{E}}, \sqcap_{\mathtt{E}}, \emptyset, \mathbb{R}^n)$ (see Figure 5.4).

Narrowing operator The intersection $\sqcap_{\mathtt{E}}$ introduces many new terms. Each time, we compute the intersection between two elements $\mathcal{E}_1$ and $\mathcal{E}_2$ of $\wp(\mathtt{Elli}^n)$, $\mathcal{E}_1 \cap \mathcal{E}_2$ requires as many ellipsoids as the one describing $\mathcal{E}_1$ plus the ones that describe $\mathcal{E}_2$. To avoid having abstract elements that accumulate terms because of the intersection $\sqcap_{\mathtt{E}}$, we introduce a narrowing operator that soundly approximates the intersection $\cap$. This narrowing operator associates an intersection of ellipsoids with a sound approximating

ellipsoid. Since there is an infinite number of ellipsoids that overapproximates an intersection of ellipsoids, and since the set of ellipsoids greater than this intersection (with respect to $\subseteq$) is not a pointed ordered set (i.e. there is no least element), we choose the least element with respect to another partial order (such as the volume of the ellipsoid for example).

We will talk more specifically about the case where ellipsoids are ordered by a cost function $J : \texttt{Elli}^n \to \mathbb{R}$. For example, J might be the volume of the ellipsoids. Then, we define the narrowing operator $\triangle_J^*$ by

$$Y^* \triangle_J^* X^* = \begin{array}{ll} \arg\inf & J(Z^*) \\ \text{s.t.} & Y^* \cap X^* \subseteq Z^* \end{array}$$

for $Y^*, X^*, Z^* \in \wp(\texttt{Elli}^n)$. By construction, $Y^* \cap X^* \subseteq Y^* \triangle_J^* X^*$, moreover, since X^* belongs to the feasible set of this optimization problem, $J(Y^* \triangle_J^* X^*) \leq J(X^*)$. $\triangle_J^*$ is a narrowing operator with respect to the order $\sqsubseteq_{\approx J}^*$ defined by $Z^* \sqsubseteq_{\approx J}^* X^*$ iff $J(Z^*) \leq J(X^*)$.

5.6 Abstractions for time-varying signals

In this section, we detail an abstract domain used to describe continuous-time or discrete-time signals.

Let $\texttt{Dom}_\texttt{x}$ be the concrete domain of a time-varying signal x (continuous- or discrete-time) of the interconnection of systems. The signal x is a function that associates to a time-domain $\mathcal{T}$ (that is a subset of $\mathbb{R}_+$) a value in $\mathbb{R}^n$. We can abstract such signals with the domain of *time-varying sets* which corresponds to a point-wise lift of the complete lattice $(\wp(\mathbb{R}^n), \subseteq, \cup, \cap, \emptyset, \mathbb{R}^n)$ over the time-domain $\mathcal{T}$.

Definition 5.9. Point-wise lifting

If $(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap, \bot, \top)$ *is a complete lattice and* $\mathcal{T}$ *is a set, then* $(\mathcal{T} \to \mathcal{A}, \dot{\sqsubseteq}, \dot{\sqcup}, \dot{\sqcap}, \dot{\bot}, \dot{\top})$ *is also a complete lattice if we define the dotted operator by* point-wise lifting*:*

$X \dot{\sqsubseteq} Y$ *is defined by* $\forall t \in \mathcal{T}, X(t) \sqsubseteq Y(t)$,

$(\dot{\bigsqcup}\mathcal{X})(t)$ *is defined by* $\bigsqcup \{X(t) \,|\, X \in \mathcal{X}\}$, $\qquad \dot{\top}(t) = \top$,

$(\dot{\bigsqcap}\mathcal{X})(t)$ *is defined by* $\bigsqcap \{X(t) \,|\, X \in \mathcal{X}\}$, $\qquad \dot{\bot}(t) = \bot$.

Then, we define the *Time-varying set domain* as the complete lattice $(\mathcal{T} \to \mathcal{A}, \dot{\sqsubseteq}, \dot{\sqcup}, \dot{\sqcap}, \dot{\bot}, \dot{\top})$. that corresponds to the point-wise lift of the complete lattice $(\mathcal{A}, \sqsubseteq, \sqcup, \sqcap, \bot, \top)$. Time-varying set domains are an abstraction of time-varying signals domains.

Proposition 5.10. Galois connection of the domain of time-varying sets
If the pair (α, γ) defines a Galois connection between the two complete lattices $(\wp(\mathbb{R}^n), \subseteq, \cup, \cap, \bot, \top)$ and $(\mathcal{A}^*, \sqsubseteq^*, \sqcup^*, \sqcap^*, \bot^*, \top^*)$, then $(\alpha_{\mathtt{TV}}, \gamma_{\mathtt{TV}})$, defined by

$$\begin{array}{lll} X^* & = \alpha_{\mathtt{TV}}(X), & \textit{where} \quad X^* : t \mapsto \alpha(X(t)) \\ X & = \gamma_{\mathtt{TV}}(X^*), & \textit{where} \quad X = \{x \in \mathtt{Dom_x} \,|\, x(t) \in \gamma(X^*(t))\} \end{array}$$

defines a Galois connection between the domain of time-varying signals $(\wp(\mathcal{T} \to \mathbb{R}^n), \subseteq, \cup, \cap, \bot, \top)$ and $(\mathcal{T} \to \mathcal{A}^*, \dot{\subseteq}^*, \dot{\cup}^*, \dot{\cap}^*, \dot{\bot}^*, \dot{\top}^*)$.

Proof. ♢

Time-varying signals can be approximated with time-varying ellipsoids which is a point-wise lift of the ellipsoidal domain introduced in Section 5.5.1.

Example 5.3.
The set of signals $X \subseteq (\mathbb{R}_+ \mapsto \mathbb{R})$

$$X = \{t \mapsto sin(t + \psi)e^{\frac{t}{10}} \,|\, \psi \in [0, 2\pi]\} \tag{5.10}$$

can be soundly approximated with the time-varying ellipsoid $X^* = \alpha_{\mathtt{TV}}(X)$ (see Figure 5.5) defined by

$$X^*(t) = E(t)$$

where E is the time-varying ellipsoid defined by its time-varying center $c(t) = 0$ and its time-varying radius $r(t) = e^{\frac{t}{10}}$, for ever $t \geq 0$. The set $\gamma_{\mathtt{TV}}(X^*)$ contains all the signals $x : \mathbb{R}_+ \to \mathbb{R}$ belonging to $(x(t) - c(t))^2 \leq r(t)^2$ at every $t \in \mathbb{R}_+$.

5.7 Piecewise linear abstraction

Let us assume that the interconnection of systems involves a source signal "$\mathcal{P} \xrightarrow{\mathtt{p}}$" where p is of finite dimension (e.g. p is a real-valued vector) and where $\mathcal{P}$ is bounded (e.g. p could be an unknown initial state or an uncertain parameter of the system). The concrete domain $\mathtt{Dom}$ corresponds to the Cartesian product of $\mathtt{Dom_p}$ and of the domain $\mathtt{Dom_x}$ of other signals x labeled by $\mathtt{x}$ in the interconnection. Each value $p \in \mathcal{P}$ is associated with a subset of the semantic, i.e. a subset of $\mathtt{Dom_x}$. In this section, we propose to define a *piecewise linear abstraction* that expresses this relationship between p (over a partition of $\mathcal{P}$) and x in the concrete semantic. Section 5.8.2 uses such abstraction to derive tight approximation.

We first define the partition of $\mathcal{P}$, then we introduce the piecewise linear abstraction, and finally, we detail the Galois connection of this abstraction with the concrete domain.

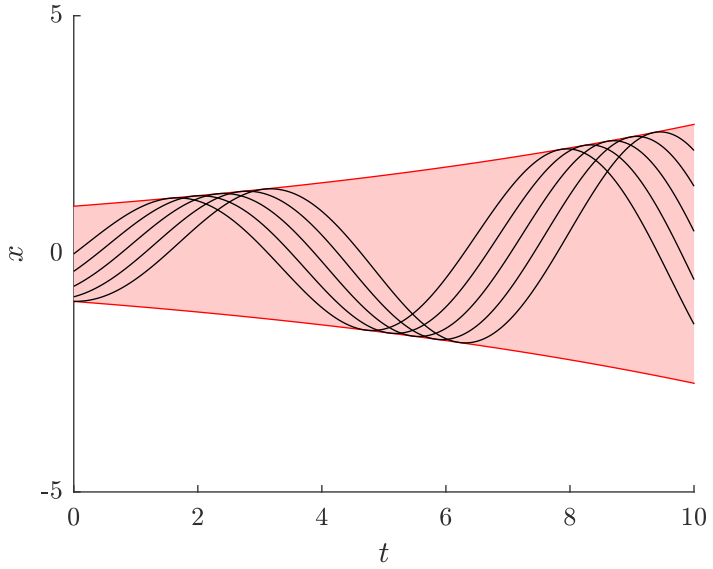


Figure 5.5: *The set of signals X defined by* (5.10) *(few signals are drawn in black) with a time-varying ellipsoidal set (in red).*

Partition Let a *partition* of $\mathcal{P}$ be described by a finite set of cells $P_i \subseteq \mathcal{P}$, $i \in I$, with I a set of labels associated with each cell, such that:

- cells are convex polytopes, to this respect, each cell P_i can be described as the convex hull of a finite set of points;
- cells are disjoint, i.e. $i \neq j \Rightarrow P_i^o \cap P_j^o = \emptyset$ (where P^o is the open set associated with the set P);
- cells cover the entire set $\mathcal{P}$, i.e. $\mathcal{P} = \bigcup_{i \in I} P_i$.

Let $\{p_k\}_{k \in K}$ be the vertices of the partition with K a set of labels identifying the vertices. For a cell P_i, let $K_i \subseteq K$ the labels of the vertices of P_i. For each vertex p_k, let the cell neighborhood of p_k, denoted by $\tilde{P}_k$, be the union of cells touching p_k (see Figure 5.6).

Piecewise linear abstraction Let the three complete lattices $(\mathcal{A}_p, \subseteq, \cup, \cap, \bot, \top)$, $(\mathcal{A}_x, \subseteq, \cup, \cap, \bot, \top)$, and $(\mathcal{A}, \subseteq, \cup, \cap, \bot, \top)$ where $\mathcal{A}_p = \wp(\mathtt{Dom_p})$, $\mathcal{A}_x = \wp(\mathtt{Dom_x})$, and $\mathcal{A} = \wp(\mathtt{Dom_p} \times \mathtt{Dom_x})$ such that $\mathtt{Dom_p}$ and $\mathtt{Dom_x}$ are vector spaces. Let the piecewise linear (PWL) domain be defined as the point-wise lift (introduced in Definition 5.9) of $\mathcal{A}_x$ over the set of vertices K, i.e.

$$(\mathcal{A}_{\mathtt{PWL}}, \dot{\subseteq}, \dot{\cup}, \dot{\cap}, \dot{\bot}, \dot{\top}).$$

where $\mathcal{A}_{\mathtt{PWL}}$ is a subset of $K \to \mathcal{A}$, the set of functions that associates to each K an element of $\mathcal{A}$. Elements $Z_{\mathtt{PWL}} \in \mathcal{A}_{\mathtt{PWL}}$ associates to a $k \in K$ an element $\{p_k\} \times X_k$

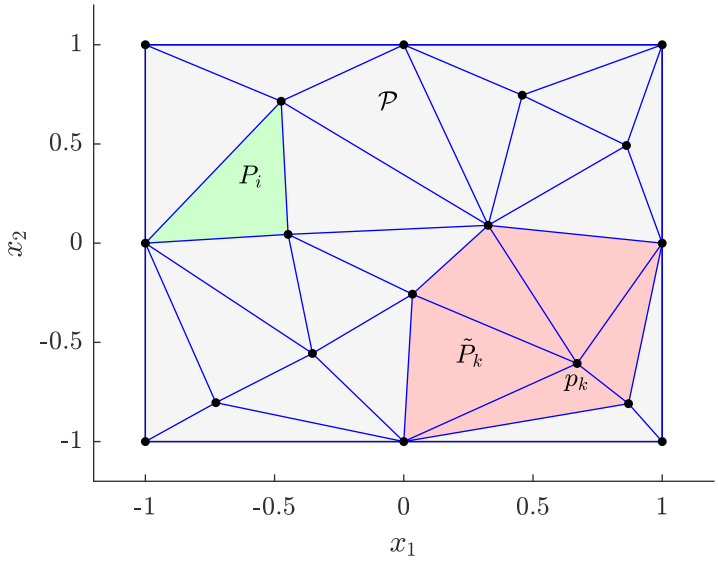


Figure 5.6: *A partition of $\mathcal{P} = [-1, 1]^2$. P_i is a cell, $\tilde{P}_j$ is the neighbor cells of the vertex p_k.*

where $X_k \in \mathcal{A}_x$. By Definition 5.9, the PWL domain inherits from the complete lattice structure of $(\mathcal{A}, \subseteq, \cup, \cap, \bot, \top)$. We define a concretization function $\gamma_{\texttt{PWL}}$. This concretization function associates to $Z_{\texttt{PWL}} \in \mathcal{A}_{\texttt{PWL}}$, an element of $\mathcal{A}$, and is defined

$$\gamma_{\texttt{PWL}}(Z_{\texttt{PWL}}) = \bigcup_{j \in I} \gamma_{\texttt{PWL}}(Z_{\texttt{PWL}}, j) \tag{5.11}$$

where

$$\gamma_{\texttt{PWL}}(Z_{\texttt{PWL}}, j) = \bigcup \left\{ \sum_{k \in K_j} \lambda_{j,k} \cdot Z_{\texttt{PWL}}(k) \;\middle|\; \forall k \in K_j, \lambda_{j,k} \geq 0 \text{ and } \sum_{k \in K_j} \lambda_{j,k} = 1 \right\}$$

where $\sum_{k \in K_j} \lambda_{j,k} \cdot Z_{\texttt{PWL}}(k)$ is a weighted Minkowski sum[1]. In other words, let $(p, x) \in \texttt{Dom}$ such that p belongs to the cell P_j, $j \in I$. The point (p, x) belongs to $\gamma_{\texttt{PWL}}(Z_{\texttt{PWL}})$ whenever (p, x) can be expressed as a convex combination of the points (p_k, x_k), with $k \in K_j$, where p_k is a vertex of P_j and where $(p_k, x_k) \in Z_{\texttt{PWL}}(k)$ (see Example 5.4).

Example 5.4.

Let the set Z be

$$Z = \{(p, x) \in \mathbb{R}^2 \mid x = sin(1.5p)^2\}, \tag{5.12}$$

[1]For two subsets A and B of a vectorial space, $x \in \lambda \cdot A + \mu \cdot B$ whenever $x = \lambda a + \mu b$ for some $a \in A$ and $b \in B$, $\lambda, \mu \geq 0$ and $\lambda + \mu = 1$.

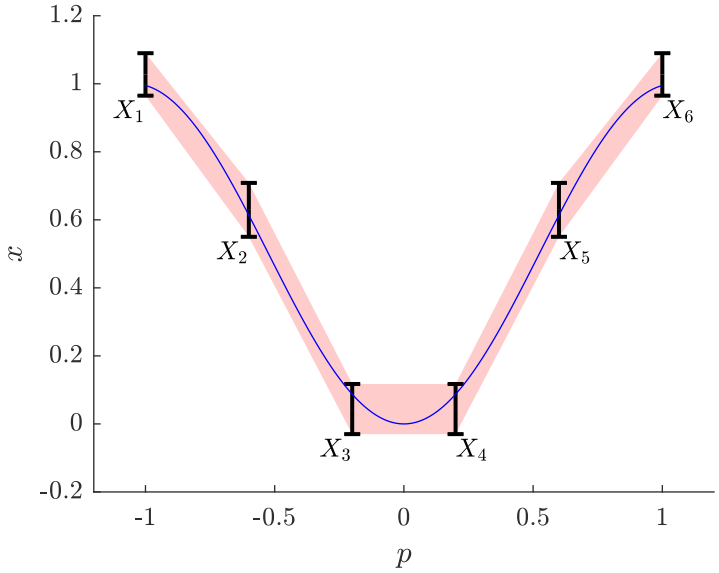


Figure 5.7: *The PWL abstraction* $Z_{\texttt{PWL}}$ *in* (5.13) *defines a sound approximation* $\gamma_{\texttt{PWL}}(Z_{\texttt{PWL}})$ *of the set* Z *in* (5.12).

with $p \in \mathcal{P} = [-1, 1]$. $\mathcal{P}$ *is partitioned with a regular mesh* $\{p_k\}_{k \in K} = \{-1.0, -0.6, -0.2, 0.2, 0.6, 1.0\}$, *where* $K = 1, \ldots, 6$. *The set* Z *can be soundly approximated by* $\gamma_{\texttt{PWL}}(Z_{\texttt{PWL}})$ *where* $Z_{\texttt{PWL}}$ *is defined by*

$$Z_{\texttt{PWL}}(k) \mapsto \{p_k\} \times X_k \tag{5.13}$$

with

$$\begin{aligned} X_1 = X_6 &= [1.09, 0.96] \\ X_2 = X_5 &= [0.71, 0.55] \\ X_3 = X_4 &= [0.12, -0.03] \end{aligned}$$

where $X_k \subseteq \mathbb{R}$ *(see Figure 5.7).*

Operator evaluation Let a function $F : \mathcal{A} \to \mathcal{A}$. When F satisfies

$$F\left(\sum_k \lambda_k Y_k\right) \subseteq \sum_k \lambda_k F(Y_k) \tag{5.14}$$

then the function $F_{\texttt{PWL}}$ defined by

$$W_{\texttt{PWL}} = F_{\texttt{PWL}}(Z_{\texttt{PWL}})$$

with

$$W_{\texttt{PWL}}(k) = F \circ Z_{\texttt{PWL}}(k)$$

is a sound approximation of F. The condition (5.14) is held for a particular class of functions. Let $f : \texttt{Dom} \mapsto \texttt{Dom}$ be an affine function over the vectorial space $\texttt{Dom} = \texttt{Dom}_p \times \texttt{Dom}_x$. Let F the lift to sets of f, i.e. $F(A) = \{f(a) \,|\, a \in A\}$. Since f is affine, $f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)$. For $\lambda, \mu \geq 0$ s.t. $\lambda + \mu = 1$, the Minkowski sum $\lambda A + \mu B$ is well defined. For $x \in F(\lambda A + \mu B)$, there is a $a \in A$ and $b \in B$ such that $x = f(\lambda a + \mu b) = \lambda f(a) + \mu f(b) \in \lambda F(A) + \mu F(B)$. Therefore, $F(\lambda A + \mu B) = \lambda F(A) + \mu F(B)$.

5.8 Examples

In this section, we present two examples that explains the concepts introduced in the previous sections. In Section 5.8.1, we present a simple closed-loop LTI discrete-time system analyzed through Interval Arithmetic (presented in Chapter 4), In Section 5.8.2, we present an application of the piecewise linear abstract domain.

5.8.1 Closed-loop of a discrete-time system

We study the interconnection of systems $\mathscr{S}$ (see its block diagram in Figure 5.8) over the discrete-time interval $\mathcal{T}_d = \{1, \ldots, T\} \subseteq \mathbb{N}$ for $T > 0$

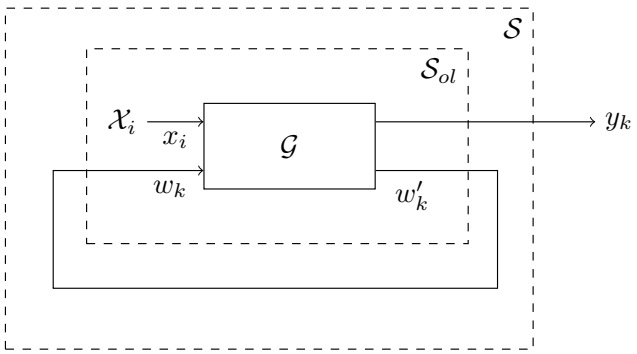


Figure 5.8: *Block diagram of the closed-loop discrete-time LTI system.*

$$\mathscr{S} : \text{“}\, \mathcal{X}_i \xrightarrow{\texttt{x}_\texttt{i}} \mu_{\texttt{w}} \{\mathcal{G}\} \xrightarrow{\texttt{y}} \circ \,\text{”} \tag{5.15}$$

where $\mathcal{G} : \mathbb{R} \times l(\mathcal{T}_d; \mathbb{R}) \to \left(l(\mathcal{T}_d; \mathbb{R})\right)^2$ is the discrete-time system that maps an initial state $x_i \in \mathbb{R}$, and a discrete-time input signal $w \in l(\mathcal{T}_d; \mathbb{R})$, to the two output signals

$y \in l(\mathcal{T}_d;\mathbb{R})$ and $w' \in l(\mathcal{T}_d;\mathbb{R})$. $\mathcal{G}$ is defined by

$$\mathcal{G} : \begin{cases} x_{t+1} = ax_t + bw_t \\ x_1 = x_i \\ y_t = x_t \\ w'_t = x_t \end{cases} \tag{5.16}$$

with $a = 0.9$ and $b = 0.02$. The set of initial states belongs to $\mathcal{X}_1 = [-1, 1]$.

Thanks to its simple form ($\mathcal{G}$ is 1 state dimension discrete-time linear system), its associated lifted operator $[\![\mathcal{G}]\!]$ can be exactly computed using the interval arithmetic abstract domain as introduced in Chapter 4. Its exact computation is given in Remark 5.5.

Remark 5.5. Exact fixed-point expression
For this specific system, the $\mathscr{S}$ represents a discrete-time system

$$\begin{cases} x_{t+1} = (a+b)x_t \\ x_1 = x_i \in \mathcal{X}_i \\ y_t = x_t \end{cases} \tag{5.17}$$

$\mathtt{Dom}$ *can be described as a* $1 + 2T$ *vectorial space where each element can be represented by a vector* $(x_i, w_1, w_2, \ldots, w_T, y_1, y_2, \ldots, y_T)$. *Each trajectory* $\{x_t\}$ *of* $\mathscr{S}$ *is a sequence that linearly depends on the initial value* $x_1 = x_i \in \mathcal{X}_i$. *For every* $k = 1, \ldots, T$, *it holds* $x_t = x_i c^{k-1}$, *where* $c = a + b$, *and thus, the semantic of the interconnection of systems* $\overline{[\mathscr{S}]} \subseteq \mathtt{Dom}$ *is then exactly equal to*

$$\overline{[\mathscr{S}]} = \{x_i \cdot (1, 1, c, \ldots, c^{T-1}, 1, c, \ldots, c^{T-1}) \,|\, x_i \in \mathcal{X}_i\}.$$

In what follows, we detail the computation of the concrete semantic $\overline{[\![\mathscr{S}]\!]}$ of the interconnection of systems $\mathscr{S}$. The set of labels of $\mathscr{S}$ is $\mathbb{V} = \{\mathtt{x_i}, \mathtt{w}, \mathtt{y}, \mathtt{x'_i}, \mathtt{w'}, \mathtt{y'}\}$, each label is associated with its respective domain $\mathtt{Dom}_{\mathtt{x_i}} = \mathtt{Dom}_{\mathtt{x'_i}} = \mathbb{R}$, $\mathtt{Dom}_{\mathtt{l}} = l(\mathcal{T}_d;\mathbb{R})$ for $\mathtt{l} \in \{\mathtt{w}, \mathtt{y}, \mathtt{w'}, \mathtt{y'}\}$. The concrete domain $\mathcal{A} = \wp(\mathtt{Dom})$ corresponds to the power set of the domain of all internal signals. The concrete semantic $\overline{[\![\mathscr{S}]\!]} \subseteq \mathtt{Dom}$ is then deduced by syntactic decomposition of the expression (5.15). It corresponds to the greatest

fixed-point of a monotonic operator $X \mapsto [\![\mathscr{S}]\!](X;\varnothing,\varnothing)$. Its detailed computation is

$$\begin{aligned}
\overline{[\![\mathscr{S}]\!]} &= \mathrm{gfp}_{\Sigma}\,\{X \mapsto [\![\mathscr{S}]\!](X;\varnothing,\varnothing)\}\\
[\![\mathscr{S}]\!](X;\varnothing,\varnothing) &= [\![\circ]\!](*;\mathtt{y},\varnothing)\circ[\![\mu_{\mathtt{w}}\,\{\mathcal{G}\}]\!](*;\mathtt{x_i},\mathtt{y})\circ[\![\mathcal{X}_i]\!](X;\varnothing,\mathtt{x_i})\\
[\![\mathcal{X}_i]\!](X;\varnothing,\mathtt{x_i}) &= \{\rho\in X\,|\,\rho_{\mathtt{x}'_\mathtt{i}}\in\mathcal{X}_i\}\\
[\![\mathcal{G}]\!](X;(\mathtt{x_i},\mathtt{w}),(\mathtt{y},\mathtt{w})) &= \{\rho[(\mathtt{x}'_\mathtt{i},\mathtt{w}')\leftarrow\mathcal{G}(\rho_{\mathtt{y}},\rho_{\mathtt{w}}),(\mathtt{x_i},\mathtt{w})\leftarrow(\rho_{\mathtt{x}'_\mathtt{i}},\rho_{\mathtt{w}'})]\,|\,\rho\in X\}\\
[\![\circ]\!](X;\mathtt{y},\varnothing) &= X\\
[\![\mu_{\mathtt{w}}\,\{\mathcal{G}\}]\!](X;\mathtt{x_i},\mathtt{y}) &= [\![\mathcal{G}_{ol}]\!](X;(\mathtt{x_i},\mathtt{w}),(\mathtt{y},\mathtt{w}))
\end{aligned}$$

where $\Sigma=\{\rho\,|\,\sigma_{\mathtt{l}}=\sigma_{\mathtt{l}'},\mathtt{l}\in\{\mathtt{x_i},\mathtt{w},\mathtt{y}\}\}$.

The concrete domain $\mathcal{A}$ is abstracted with the time-varying abstraction (as introduced in Section 5.6) where sets are intervals and the time domain is $\mathcal{T}_d$

$$\mathcal{A}^* = (\mathcal{T}_d \to \mathtt{Int}).$$

Since the interval domain is a complete lattice structure, this instance of the time-varying set abstraction has a complete lattice structure as well. The two domains $\mathcal{A}$ and $\mathcal{A}^*$ are linked by the concretisation function

$$\gamma(X^*) = (\gamma_{\mathtt{Int}}(X^*(1)) \times (\gamma_{(\mathcal{T}_d\to\mathtt{Int})}(X^*))^2)^2$$

As the definition of the concretisation function γ suggests, for $\rho=(x_i,w,y,x'_i,w',y')\in\mathtt{Dom}$, the signals x_i and x'_i are represented with the interval $\gamma_{\mathtt{Int}}(X^*(1))\subseteq\mathbb{R}$, and the signals w,w', y, and y' are represented with the interval set $\gamma_{(\mathcal{T}_d\to\mathtt{Int})}(X^*)$.

Let $\mathcal{S}$ be an expression appearing in the syntactic decomposition of $\mathscr{S}$ ($\mathcal{S}$ is expressed in the syntax given in Table 5.1). Since we are searching for a fixed-point lower than Σ, each operator $X\to[\![\mathcal{S}]\!](X;\varnothing,\varnothing)$ in the concrete semantic can be soundly approximated by

$$[\![\mathcal{S}]\!]_{\Sigma}(X;\mathtt{u},\mathtt{y}) = [\![\mathcal{S}]\!](X;\mathtt{u},\mathtt{y})\cap\Sigma. \tag{5.18}$$

where $\mathtt{u},\mathtt{y}\ \in\overline{\mathbb{V}}$ (where $\overline{\mathbb{V}}$ is defined by (5.1)). A sound approximation $[\![\mathcal{S}]\!]^*_{\Sigma}$ of $[\![\mathcal{S}]\!]_{\Sigma}$ can be evaluated within the abstract domain of interval arithmetic as defined in Section 4.1 of Chapter 4. The greatest fixed-point is computed using descending chains as defined in Proposition 5.1. Starting with a post-fixpoint $X^*_0(k)=[-10,10]$ for all $t\in\mathcal{T}_d$, we compute the sequence X^*_k defined by

$$X^*_{k+1} = [\![\mathscr{S}]\!]^*_{\Sigma}(X^*_k;\varnothing,\varnothing)$$

for few iterations. Figure 5.9 shows plots of the abstract iterates projected over the domain of the signals $w=w'=y=y'$ (these signals coincide in this specific case

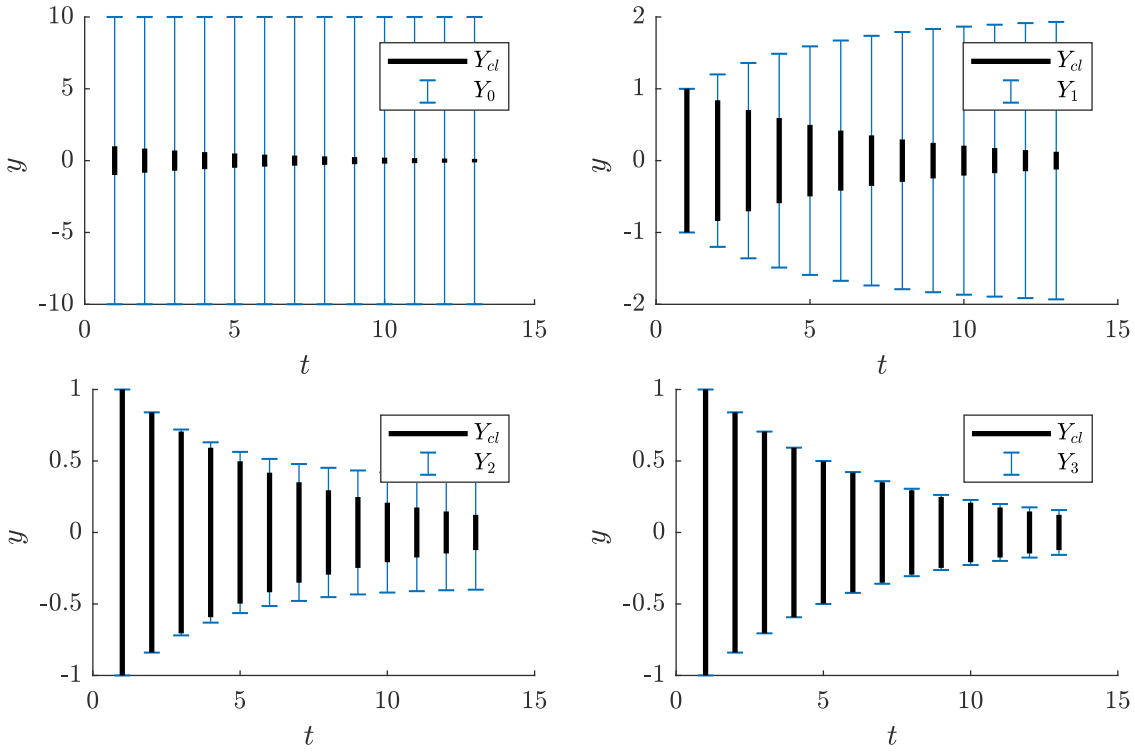


Figure 5.9: *Decreasing sequence* $\{Y_k\}$ *of overapproximations of the reachable tube* Y_{cl} *of the signal* y *of the interconnection of systems* S_{cl}*.*

due to the form of $\mathcal{G}$), these iterates are $\mathcal{Y}_k = \gamma_{\mathcal{T}_d \to \mathtt{Int}}(X_k^*)$, for $k = 0, 1, 2, \ldots$ and we compare them to $\mathcal{Y}_{cl}$ that corresponds to the exact semantic of the interconnection of systems (projected over $w = w' = y = y'$). Since every function evaluation is sound, by Proposition 5.5, and if X_0^* is a post-fixpoint of $\mathrm{gfp}_{\Sigma^*} \{X \mapsto [\![\mathscr{S}]\!]_\Sigma^*(X; \varnothing, \varnothing)\}$, then $\gamma_{\mathcal{T}_d \to \mathtt{Int}}(X_k^*)$ are a sound approximation of $\overline{[\mathscr{S}]}$. Also, the time-varying interval $\mathcal{Y}_k$ is a sound approximation of $y = y'$ and $w = w'$ where $\rho = (x_i, w, y, x_i, w, y) \in \overline{[\mathscr{S}]}$. Figure 5.10 compares the volume of the overapproximation with the volume of the exact reachable set of $\mathscr{S}$ (see Remark 5.5) for the 6 first iterates. After three iterations, the volume error is less than 1%. The error exponentially decreases with iterations and converges to a value close to the numerical precision of the floating-point arithmetic.

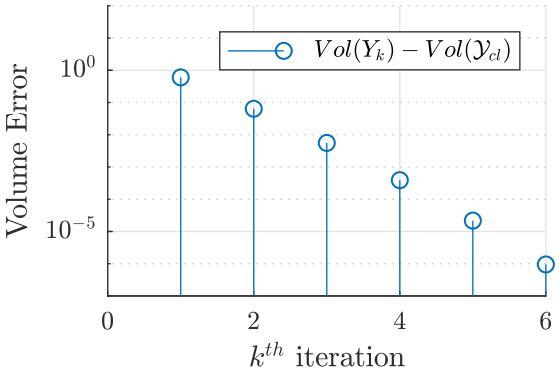


Figure 5.10: *The average volume error at each iteration exponentially decreases and converges to* 0.

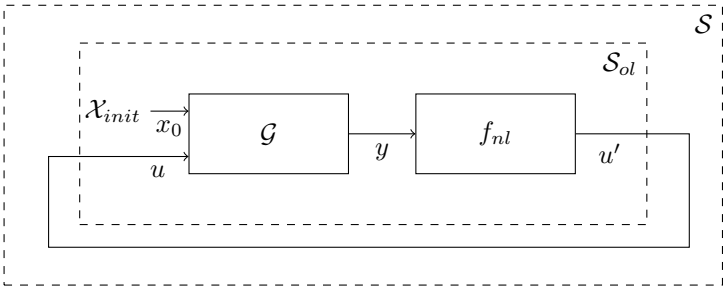


Figure 5.11: *Block diagram representation of the interconnection of systems $\mathscr{S}$ described by* (5.19)*. $\mathcal{S}_{ol}$ is the open-loop system described by* (5.22)*.*

5.8.2 Piecewise linear models

In this section, we study a nonlinear system described as the interconnection of a linear time-invariant system with a nonlinear feedback. The response of the linear system is overapproximated with results of Part I, and the one of the nonlinear feedback is overapproximated with the interval arithmetic framework (as presented in Section 4.1 of Chapter 4). The concrete domain of this system is abstracted with a PWL abstraction as presented in Section 5.7.

Description of the interconnection of systems We study the interconnection of systems $\mathscr{S}$ described by the expression (5.19) (see Figure 5.11) over the time-interval $\mathcal{T} = [0, T]$, $T > 0$.

$$\mathscr{S} : \text{“ } \mathcal{X}_{init} \xrightarrow{\mathtt{x_0}} \mu_{\mathtt{u}} \left\{ \mathcal{G} \xrightarrow{\mathtt{y}} f_{nl} \right\} \xrightarrow{\mathtt{z}} \circ \text{ ”} \tag{5.19}$$

where $\mathbb{V} = \{\mathtt{x_0}, \mathtt{u}, \mathtt{y}\}$ are labels of $\mathscr{S}$. The system $\mathcal{G} : \mathbb{R}^2 \times L(\mathcal{T}; \mathbb{R}) \mapsto L(\mathcal{T}; \mathbb{R})$ is a linear time-invariant system defined by $y = \mathcal{G}(x_0, u)$ where

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ x(0) = x_0 \\ y(t) = Cx(t) \end{cases}$$

for an initial state $x_0 \in \mathbb{R}^2$, an input $u \in L(\mathcal{T}; \mathbb{R})$, and an output $y \in L(\mathcal{T}; \mathbb{R})$. Parameters A, B, and C are defined by

$$A = \begin{bmatrix} -2 & 3 \\ 2 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \text{ and } C = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

The system $f_{nl} : L(\mathcal{T}; \mathbb{R}) \mapsto L(\mathcal{T}; \mathbb{R})^2$ is the nonlinear function defined by $(z, u) = f_{nl}(y)$ where

$$z(t) = u(t) = sin(3y(t))^3. \tag{5.20}$$

The set $\mathcal{X}_{init}$ of initial states is defined by

$$\mathcal{X}_{init} = \{(x_{1,init}, 0) \,|\, x_{1,init} \in [-1, 1]\}. \tag{5.21}$$

Let the open-loop system $\mathcal{S}_{ol}$ be

$$\mathcal{S}_{ol} : \text{“}(\mathcal{X}_{init}, id_{\mathtt{u}}) \xrightarrow{(\mathtt{x_0},\mathtt{u})} \mathcal{G} \xrightarrow{\mathtt{y}} f_{nl} \text{”}. \tag{5.22}$$

The system $\mathscr{S}$ can equivalently be expressed with “$\mu_{\mathtt{u}}\,\{\mathcal{S}_{ol}\}$”. Its domain is $\mathtt{Dom} = \mathtt{Dom}_{\mathtt{x_0}} \times \mathtt{Dom}_{\mathtt{u}} \times \mathtt{Dom}_{\mathtt{y}}$, where $\mathtt{Dom}_{\mathtt{x_0}} = \mathbb{R}^2$, and $\mathtt{Dom}_{\mathtt{u}} = \mathtt{Dom}_{\mathtt{y}} = L(\mathcal{T}; \mathbb{R})$. An environment $\rho = (x_0, u, y)$ belongs to the semantic of the system $\overline{[\mathscr{S}]}$ (i.e. $[\mathscr{S}](\rho; \varnothing, \varnothing)$ is true) whenever the following relations are satisfied

$$\begin{cases} x_0 \in \mathcal{X}_{init} \\ y = \mathcal{G}(x_0, u) \\ u = f_{nl}(y) \end{cases}$$

Concrete semantic The concrete domain $\mathcal{A}$ is the power set of $\mathtt{Dom}$, i.e. $\mathcal{A} = \wp(\mathtt{Dom})$. Each element $\rho \in X$ in a concrete element $X \in \mathcal{A}$ is an environment $\rho = (x_0, u, y, x_0', u', y')$ (i.e. an environment of the system extended with the “primes variables”). The concrete semantic is then defined by induction over the syntax.

Abstract semantic We study two possible abstractions of the interconnection of systems $\mathscr{S}$.

The first abstraction $(\mathcal{Y}_0, \mathcal{U}_0)$ makes use of the time-varying ellipsoidal domain. Since y and u are 1-dimensional signals, the time-varying ellipsoidal domain coincides with the interval domain. The operator $\mathcal{G}$ is a linear time-invariant system and can be abstracted using the ellipsoidal method presented in Chapter 1. The operator f_{nl} is a nonlinear system, we abstract it with $\mathcal{U}_0$ as the set of signals with values belonging to $[-1, 1]$.

The second abstraction $(\mathcal{Y}_i, \mathcal{U}_i)$ (for i in $\{5, 10, 20\}$) makes use of the piecewise linear template allowing to describe the relationship between the initial condition x_0 and the signals (u, y). The set of initial condition is partitioned using a uniform grid $\{x_0^k\}_{k \in K_i}$ where I have i elements. To each initial condition, we associate a trajectory (y^k, u^k). Then, the signals u and y are abstracted using the ellipsoidal domain, and, as above, the operator $\mathcal{G}$ is abstracted using the ellipsoidal method as presented in Chapter 1. The operator f_{nl} is abstracted using a linear piecewise model between each centered trajectories (y_k, u_k).

Figure 5.13 and Figure 5.14 show the overapproximation of signals y and u (resp.) with the ellipsoidal domain (resp. $\mathcal{Y}_0$ and $\mathcal{U}_0$) and with the piecewise linear abstraction (resp. $\mathcal{Y}_i$ and $\mathcal{U}_i$ for i in $\{5, 10, 20\}$). Figure 5.12 compares the volume of the overapproximations with the volume of the exact reachable set of $\mathcal{Y}$.

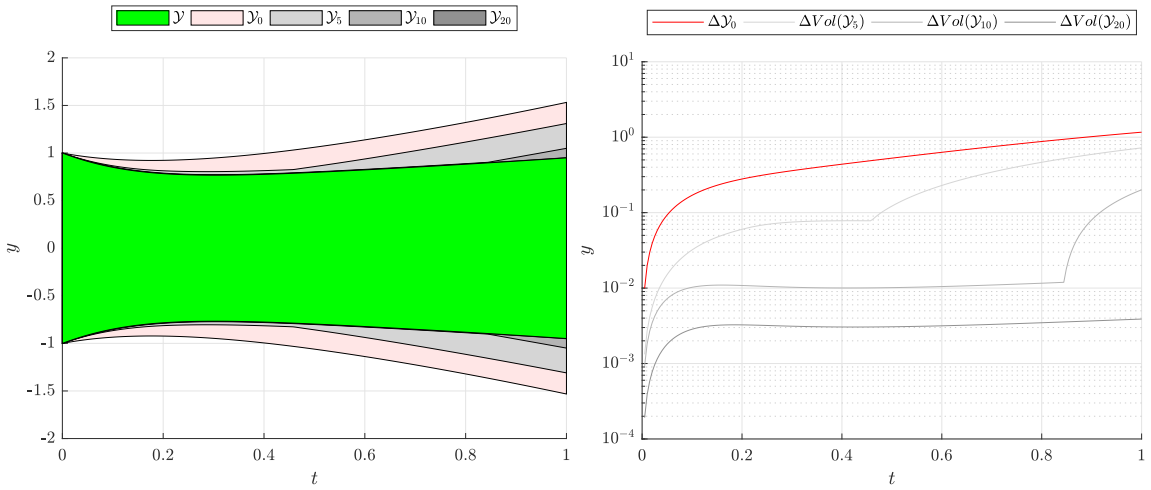
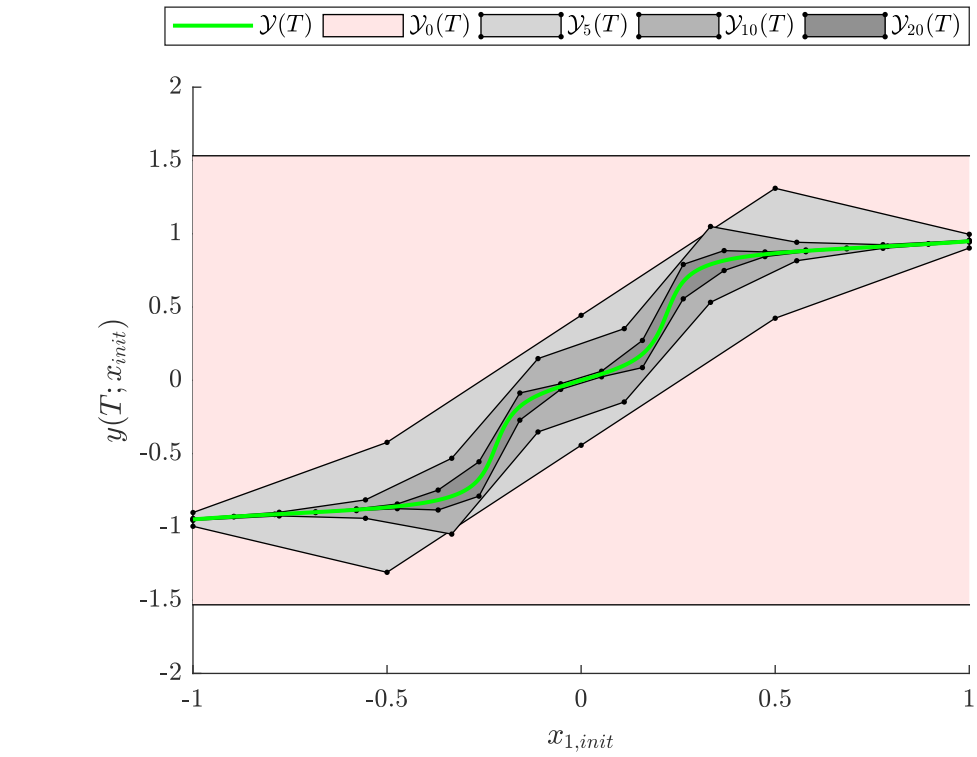


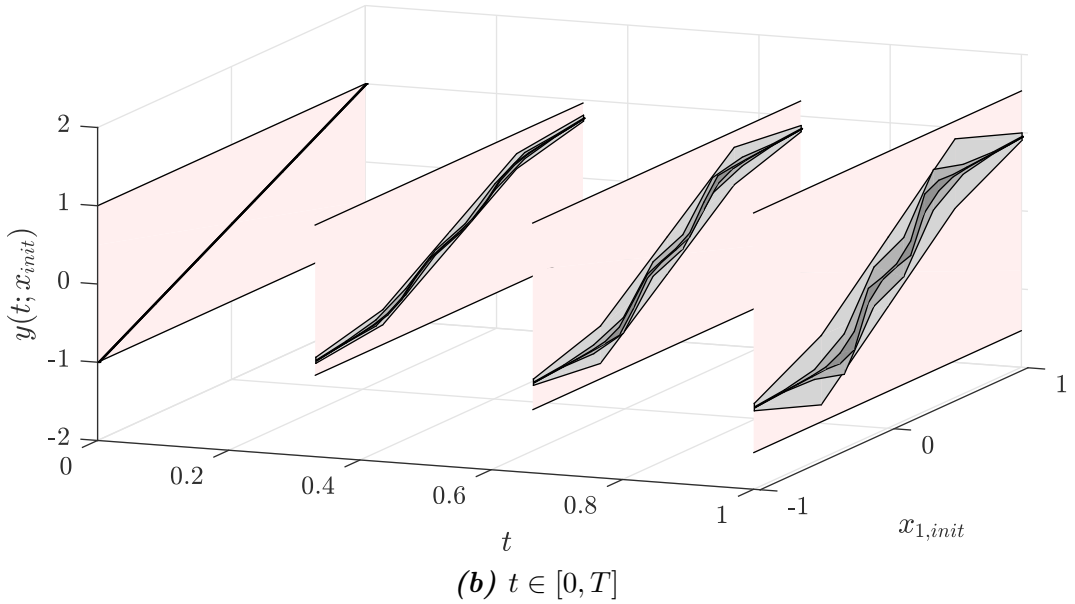
Figure 5.12: *Comparison of the precision when using a PWL abstraction. $\mathcal{Y}_i$ for i in $\{5, 10, 20\}$ corresponds to overapproximations of $\mathcal{Y}$ using the PWL abstraction (i corresponds to the number of points in the mesh of the initial state), $\mathcal{Y}_0$ corresponds to the ellipsoidal domain.*

Thanks to the Galois connection between the concrete and abstract domain, we know that the abstract semantic is a sound approximation of the concrete semantic. Since $\mathcal{Y}$ spans over $[-0.7, 0.7]$ (see Figure 5.12), using (5.20), $\mathcal{U}$ spans $[-1, 1]$. Therefore, the abstraction $\mathcal{U}_0$ of $\mathcal{U}$ with the ellipsoidal domain cannot be smaller than $\mathcal{U}$. And to this respect, the smallest fixed-point using an ellipsoidal domain is often strictly larger than the actual set of trajectories. The ellipsoidal fail to capture the dependencies between the signals u and y.

When using the PWL, abstractions $\mathcal{Y}_i$ and $\mathcal{U}_i$ better represent the set of trajectories of the system $\mathscr{S}$ (see Figure 5.13, Figure 5.14 and Figure 5.12).

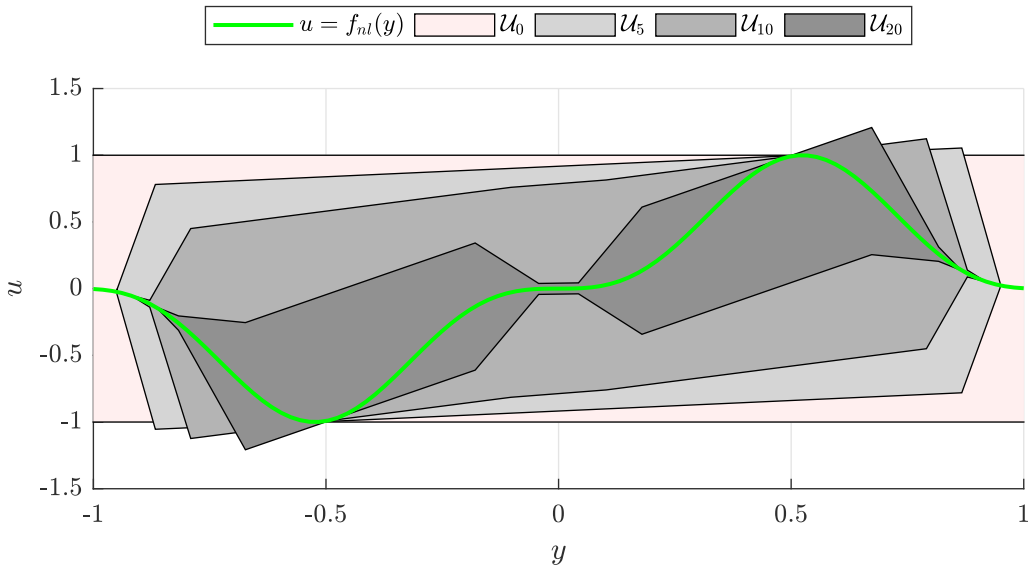


(a) $t = T$

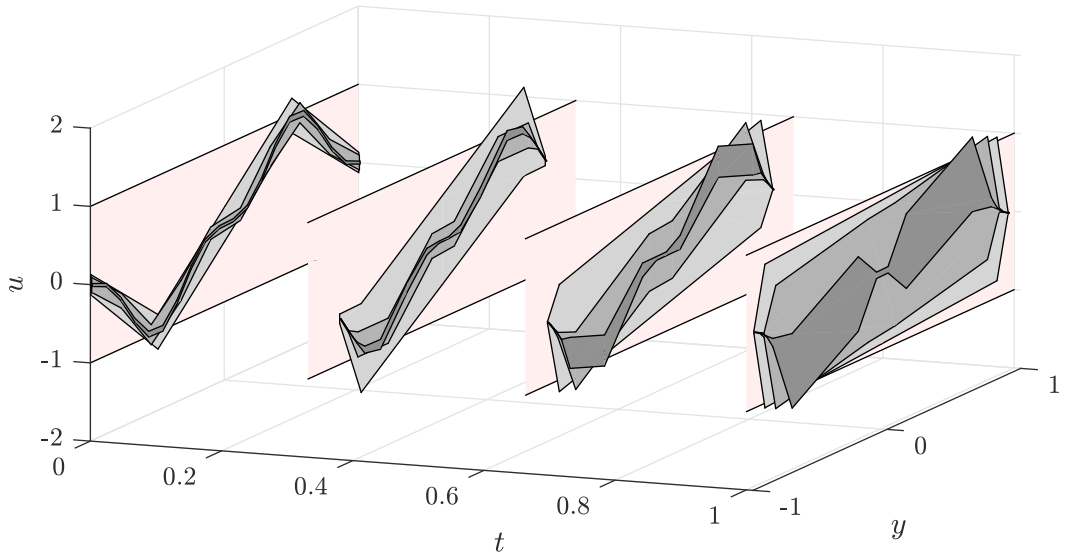


(b) $t \in [0, T]$

Figure 5.13: *Sound approximation of the relationship* $y(t) \mapsto u(t)$.



(a) $t = T$



(b) $t \in [0, T]$

Figure 5.14: *Sound approximation of the relationship* $y(t) \mapsto u(t)$ *at final time.*

5.9 Related works

[Oulamara and Venet, 2015] introduces the gauge domain to overapproximate the reachable tube of discrete-time systems (that are computer programs in this case). The reachable tube is abstracted with ellipsoids and ellipsoidal cones oriented in the time axis direction. In this work, we consider a larger class of abstractions: time-varying subsets of the state space, examples treated make use of time-varying ellipsoids as well as intervals.

An efficient reachability analysis framework is often specialized for a very narrow family of models. However, modern automotive systems are a mixture of many systems that are modeled in different frameworks. Trying to find one common model is often cumbersome. In this work, we propose an approach combining different analysis frameworks to overapproximate the reachable set of an interconnection of complex systems. These frameworks are articulated around the abstract interpretation framework.

In our work, the interconnection of systems is expressed with two connections: the serial connection of systems and the feedback connection. Such constructions have been studied in the field of reachability analysis under the name of compositional methods. [Eqtami and Girard, 2019] computes invariants for an interconnection of two systems. This invariant is expressed as a level set in the state space. Then each system of the interconnection is used to refine a prior overapproximation (i.e. post-fixpoint) until a fixed-point is found. [Chen and Sankaranarayanan, 2016] applies the same reasoning to reachability analysis in the case of multi-robot applications. The reachable set of each robot is overapproximated using a prior overapproximation of the reachable set of other robots. The reachable tubes are computed using intervals propagated with Taylor models (using the tool Flow* described in [Chen et al., 2013]). Previously cited works use compositional methods in order to analyze higher dimensional systems. Our motivation is different, we use compositional methods to couple different reachability analysis frameworks.

[Bogomolov et al., 2019] expresses the reachable set of an hybrid system as a fixed-point equation over sets of system trajectories. The hybrid system is expressed as an interconnection of a continuous system and a discrete system. Then, trajectories are abstracted as a union of convex sets and the reachable tube is expressed as the least fixed-point of a discrete-time operator and the continuous-time operator.

In our approach, we produce a decreasing sequence of sets that overapproximates the reachable tube of the interconnection of systems. It is possible to think of these iterates as a set-based version of the Picard iterates that arise in ordinary differential equation integration (see Remark 5.2). In the field of validated numerical integration, [Moore et al., 2009] has introduced a set-based Picard operator evaluated by using

the interval arithmetic framework (as presented in Chapter 4).

5.10 Conclusion

In this chapter, we study an interconnection of systems formulated within an abstract interpretation approach. The interconnection of systems is described with a syntax that defines a set of systems (a mapping from an input signal to an output signal), a set of sources (i.e. a set of signal), and a set of connections between a pair of these two (a serial connection and a feedback connection). Each connection is associated with: a signal, that is a constant, or a time-varying, continuous- or discrete-time variable; and a relationship that this signal satisfies with other signals in the interconnection of systems. We show that the set of signals of the interconnection of systems is the greatest fixed-point of a monotonic operator in a complete lattice. Such a problem is suited to the abstract interpretation framework. We introduce the point-wise lift abstract domains in order to represent time-varying signals. The method is applied to two toy examples.

Future works This work can be extended in several ways.

We have abstracted signals with time-varying sets. When studying an interconnection of systems, this approach neglects the correlation between two systems. The piecewise linear (PWL) abstraction allows to represent this correlation between two variables. However, we only applied the PWL abstraction between a constant signal and a time-varying signal. This use is easier since the constant variable belongs to a space of low dimensions for which it was straightforward to define a partition. The use of the PWL abstraction between two time-varying signals is more complex since time-varying signals belong to spaces of higher dimensions (of infinite dimension continuous-time signals and of large dimension for discrete-time signals). Future works could focus on an extension of the PWL abstraction for time-varying signals.

The precision of our method has been evaluated for specific systems where the actual reachable tube was computable. For more complex systems, the precision can be evaluated between each system by computing the distance between some existing output signals and the actual overapproximation of the trajectory transformer. Future work would focus on a more systematic approach to evaluate the precision of produced overapproximations.

Perspectives

Conclusion

During this thesis, we developed three methods to overapproximate the reachable set of an embedded system.

Our first contribution concerns results around overapproximations with time-varying conics of the reachable set of a Linear Time-Varying (LTV) system subject to bounded disturbances. The time-varying coefficient of this conic is the solution to an initial value problem. Contrary to previous works for such a class of systems, we do not restrict the overapproximations to time-varying ellipsoids (i.e. overapproximations might be unbounded), and we work with homogeneous coordinates. In this coordinate system, the initial value problem has an elegant expression which is a Differential Riccati Equation (DRE). Such expression is interesting as the DRE has been heavily studied. We more specifically studied two subclasses of systems: LTV system subjects to Quadratic Constrained (QC) disturbances, and LTV systems subjects to Integral Quadratic Constrained (IQC) disturbances. For these subclasses, we proved that it exists an overapproximation over any time-horizon. Such a guaranty of existence has been rarely addressed. Usually, the existence of an overapproximation is either conditioned by the non-emptiness of a feasible set of a linear optimization problem or is either conditioned by the existence of a non-diverging solution to a DRE.

For QC systems, we proposed the use of positive time-varying multipliers. Each multiplier is associated with a time-varying conic overapproximation. Since the set of multipliers is infinite, the set of time-varying conics is as well infinite. We provide two methods to choose a multiplier that produce a tight time-varying conic overapproximation. In the first method, the multiplier is chosen such that the boundary of the overapproximation touches the reachable set along a so-called touching trajectory. In the second method, the multiplier is chosen such that the overapproximation has a minimal volume at a given time. Suboptimal solutions of this minimization problem are expressed with an application of the Pontryagin's Maximum Principle and solved

with a continuation algorithm. These two methods were already developed for the case of systems with disturbances bounded by an exogenous signal, we extended it for the more general case of QC disturbances.

For the IQC systems, we introduced the extended system. This extended system embeds the integral quadratic into a new state of the system and the IQC is then expressed as a constraint over this new state. We then compute overapproximations of the reachable set with time-varying paraboloids (which corresponds to an extension of the previously introduced time-varying conics). For this extended system, it is possible to define touching time-varying paraboloid overapproximations and their associated touching trajectories. As for the QC system, the boundary of the touching time-varying paraboloids stays in contact with the reachable set of the extended system on the touching trajectory. We prove that the intersection of all the time-varying paraboloid overapproximations (generated by all the multipliers) exactly describes the reachable set of the extended system. This result is interesting as it shows that the set of multipliers is *sufficient* to exactly describe this reachable set. The existing literature in IQC systems never addressed the precision of overapproximations in such a way. Usually, since the IQC framework originated from stability analysis, multipliers where chosen in frequency domains, and touching trajectories and touching time-varying paraboloids could not be defined. The precision of the overapproximations was only appreciated on a practical example by comparing the different obtained overapproximations. By choosing multipliers in a temporal domain, we can define these touching trajectories and touching time-varying paraboloids. We developed and implemented an algorithm to compute a tight overapproximation of this reachable set. The set of multipliers is chosen according to the behavior of the touching trajectories that violate the state constraint in the future.

Possibles extensions of these works include the use of semidefinite solvers to derive overapproximations (optimal in volume for example). Such methods have been already developed for multipliers chosen in a frequency basis and can be adapted to the temporal domain multipliers used in this thesis. Optimal time-varying conics can as well be improved. We only considered optimal time-varying conic that minimizes the volume at a given time. A natural extension of our method would be to compute time-varying overapproximation minimizing the volume at any time. For such a case, the coefficient of the overapproximation satisfies a partial differential equation that involves two time-indexes, the regular time-index and the time-index at which the time-varying conic is of minimal volume.

Our second contribution is an extension of the validated numerical integration framework based upon interval arithmetic to the analysis of nonlinear systems subject to a disturbance bounded by an integral constraint. Validated numerical integration frameworks provide methods to compute a valid tube that contains the solution to an

initial value problem. These methods can be used as well to compute overapproximations of the reachable set of a dynamical system subject to an unknown but bounded input. Previous works in this field studied the case of input with a point-wise bound (i.e. bounded at any time), we studied the case of an input bounded by an integral constraint. Such integral models are interesting as they can model complex systems (such as systems with inner delay). To do so, we use the integral constraint to define a contractor over the set of system trajectories. The contractor associates with a set of trajectory, a subset of these trajectories that satisfies the integral constraint. We then use this contractor in a fixed point algorithm to refine a prior overapproximation of the reachable tube.

A natural extension would explore the use of other geometrical shapes to overapproximate the reachable set. In particular, affine forms can be appropriate to express the relationship between the integral value and the state of the system (this relationship is not handled by the arithmetic interval framework and introduces some conservatism). Furthermore, in our work, we assume that the unknown disturbance evolves according to a model, such a hypothesis allows us to define the contractor out of the integral constraint. Future works should consider to remove such an assumption.

Our last contribution is a framework to study interconnections of systems. This framework is formulated within an abstract interpretation approach.

The interconnection of systems is described by a set of systems (which are mappings from a signal space to another signal space), a set of sources (a set of input signals), and a set of connections between these two. A connection is associated with a signal that corresponds to a constant, or time-varying (discrete-time or continuous-time) variable. Each connection defines a relationship between signals of the interconnection. The problem of interest is then to overapproximate the set of signals satisfying these equations. We show that these equations can be expressed as a fixed point equation over the signal space and that the set of signals of the interconnection of systems is the greatest fixed point of the lift to sets of this fixed point equation. Contrary to other approaches in this field, this set of signals is not defined as a sequence that evolves according to a transition function (as it is usually done in analysis of computer programs), a dynamic function (as for dynamical systems), or a combination of these two (as for hybrid systems). The semantic cannot be iteratively constructed, and furthermore, the concrete semantic is not defined as the least fixed point of a monotonic operator.

To overapproximate the reachable set of signals, we compute the greatest fixed point in an abstract domain. We propose to abstract time-varying signals with a point-wise lifting of an interval domain, or of an ellipsoidal domain, and to abstract each system with an overapproximating method (such as the two first methods pro-

posed in this thesis). Then, we overapproximate the greatest fixed point in the abstract domain with a refinement approach, i.e. starting from a prior overapproximation of the greatest fixed point and by refining it until a fixed point is reached. To enhance the precision of our overapproximations, we introduce a so-called piecewise linear (PWL) abstract domain which corresponds to abstractions defined by a weighted sum (in the sense of the Minkowski sum) at chosen interpolating points. This PWL abstraction improved the analysis of systems where signals were correlated.

Extensions of this work could investigate new abstract domains for time-varying signals. Since abstract interpretation has been introduced to study computer programs, there are not so many options to abstract continuous-time signals. We used a point-wise lift of the subsets of the state space. One could think to provide a frequency-based abstraction where signals are characterized by there frequency spectrum. Also, we used the PWL abstraction to represents the relationship between a scalar and a time-varying signal. A similar approach to model the relationship between two time-varying signals is necessary to improve the conservatism of overapproximations.

Bibliography

[Alexandre dit Sandretto and Chapoutot, 2016] Alexandre dit Sandretto, J. and Chapoutot, A. (2016). Contraction, propagation and bisection on a validated simulation of ODE. In *Summer Workshop on Interval Methods*, Lyon, France.

[Ariba et al., 2018] Ariba, Y., Gouaisbaut, F., Seuret, A., and Peaucelle, D. (2018). Stability analysis of time-delay systems via Bessel inequality: A quadratic separation approach. *International Journal of Robust and Nonlinear Control*, 28(5):1507–1527.

[Athans, 1967] Athans, M. (1967). The matrix minimum principle. *Information and Control*, 11(5-6):592–606.

[Bertsekas and Rhodes, 1971] Bertsekas, D. P. and Rhodes, I. B. (1971). Recursive State Estimation for a Set-Membership Description of Uncertainty. *IEEE Transactions on Automatic Control*, 16(2):117–128.

[Bittanti et al., 1991] Bittanti, S., Laub, A. J., and Willems, J. C. (1991). *The Riccati Equation*, volume 32. Springer, Berlin, Heidelberg.

[Bogomolov et al., 2019] Bogomolov, S., Forets, M., Frehse, G., Potomkin, K., and Schilling, C. (2019). Reachability analysis of linear hybrid systems via block decomposition. *arXiv:1905.02458*.

[Bouissou and Martel, 2008] Bouissou, O. and Martel, M. (2008). A hybrid denotational semantics for hybrid systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4960 LNCS:63–77.

[Boyd et al., 1994] Boyd, S., El Ghaoui, L., Feron, E., and Balakrishnan, V. (1994). *Linear Matrix Inequalities in System and Control Theory*, volume 15. Society for Industrial and Applied Mathematics.

[Chabert and Jaulin, 2009] Chabert, G. and Jaulin, L. (2009). Contractor programming. *Artificial Intelligence*, 173(11):1079–1100.

[Chaudenson, 2013] Chaudenson, J. (2013). *Robustness analysis with Integral Quadratic Constraints, application to space launchers.* Theses, Supélec, France.

[Chen et al., 2013] Chen, X., Ábrahám, E., and Sankaranarayanan, S. (2013). Flow*: An analyzer for non-linear hybrid systems. *Lecture Notes in Computer Science*, 8044 LNCS:258–263.

[Chen and Sankaranarayanan, 2016] Chen, X. and Sankaranarayanan, S. (2016). Decomposed Reachability Analysis for Nonlinear Systems. In *2016 IEEE Real-Time Systems Symposium (RTSS)*, pages 13–24. IEEE, IEEE.

[Chernous'ko, 1999] Chernous'ko, F. L. (1999). What is ellipsoidal modelling and how to use it for control and state estimation? In Elishakoff, I., editor, *Whys and Hows in Uncertainty Modelling*, pages 127–188, Vienna. Springer.

[Cousot, 2002] Cousot, P. (2002). Constructive design of a hierarchy of semantics of a transition system by abstract interpretation. *Theoretical Computer Science*, 277(1-2):47–103.

[Cousot and Cousot, 1979] Cousot, P. and Cousot, R. (1979). Systematic design of program analysis frameworks. In *Conference Record of the Annual ACM Symposium on Principles of Programming Languages*, pages 269–282, San Antonio, Texas. ACM Press, New York, NY.

[Cousot and Cousot, 1992] Cousot, P. and Cousot, R. (1992). Abstract interpretation frameworks. *Journal of Logic and Computation*, 2(4):511–547.

[Dit Sandretto and Chapoutot, 2016] Dit Sandretto, J. A. and Chapoutot, A. (2016). Validated explicit and implicit Runge-Kutta methods. *Reliable Computing*, 22:78–103.

[Eqtami and Girard, 2019] Eqtami, A. and Girard, A. (2019). A quantitative approach on assume-guarantee contracts for safety of interconnected systems. In *2019 18th European Control Conference, ECC 2019*, pages 536–541.

[Fetzer et al., 2018] Fetzer, M., Scherer, C. W., and Veenman, J. (2018). Invariance with dynamic multipliers. *IEEE Transactions on Automatic Control*, 63(7):1929–1942.

[Fry et al., 2017] Fry, J. M., Farhood, M., and Seiler, P. (2017). IQC-based robustness analysis of discrete-time linear time-varying systems. *International Journal of Robust and Nonlinear Control*, 27(16):3135–3157.

[Girard, 2005] Girard, A. (2005). Reachability of uncertain linear systems using zonotopes. In *Lecture Notes in Computer Science*, volume 3414, pages 291–305. Springer.

[Graettinger and Krogh, 1991] Graettinger, T. J. and Krogh, B. H. (1991). Hyperplane method for reachable state estimation for linear time-invariant systems. *Journal of Optimization Theory and Applications*, 69(3):555–588.

[Guseinov and Nazlipinar, 2011] Guseinov, K. G. and Nazlipinar, A. S. (2011). An algorithm for approximate calculation of the attainable sets of the nonlinear control systems with integral constraint on controls. *Computers and Mathematics with Applications*, 62(4):1887–1895.

[Gusev and Zykov, 2018] Gusev, M. I. and Zykov, I. V. (2018). On Extremal Properties of the Boundary Points of Reachable Sets for Control Systems with Integral Constraints. In *Proceedings of the Steklov Institute of Mathematics*, volume 300, pages 114–125. Elsevier.

[Haskins et al., 2004] Haskins, B., Corp, N. G., Moroney, G., and Dabney, J. (2004). Error Cost Through the Project Life Cycle. Technical report, NASA.

[Helmersson, 1999] Helmersson, A. (1999). An IQC-based stability criterion for systems with slowly varying parameters. In *IFAC Proceedings Volumes*, volume 32, pages 3183–3188. Elsevier.

[Henrion and Korda, 2014] Henrion, D. and Korda, M. (2014). Convex computation of the region of attraction of polynomial control systems. *IEEE Transactions on Automatic Control*, 59(2):297–312.

[Jaulin et al., 2001] Jaulin, L., Kieffer, M., Didrit, O., and Walter, É. (2001). *Applied Interval Analysis*. Springer.

[Jönsson, 1996] Jönsson, U. (1996). *Robustness Analysis of Uncertain and Nonlinear Systems*. PhD thesis, Lunds Universitet, Sweden.

[Jönsson, 2002] Jönsson, U. (2002). Robustness of trajectories with finite time extent. *Automatica*, 38(9):1485–1497.

[Kenney and Leipnik, 1985] Kenney, C. S. and Leipnik, R. B. (1985). Numerical Integration of the Differential Matrix Riccati Equation. *IEEE Transactions on Automatic Control*, 30(10):962–970.

[Kierzenka and Shampine, 2008] Kierzenka, J. and Shampine, L. F. (2008). A BVP solver that controls residual and error. *Journal of Numerical Analysis, Industrial and Applied Mathematics*, 3(1-2):27–41.

[Korda, 2016] Korda, M. (2016). *Moment-sum-of-squares hierarchies for set approximation and optimal control*. PhD thesis, EPFL, Switzerland.

[Kučera, 1973] Kučera, V. (1973). A Review of the Matrix Riccati Equation. *Kybernetika*, 9(1):42–61.

[Kurzhanski and Varaiya, 2002] Kurzhanski, A. B. and Varaiya, P. (2002). On ellipsoidal techniques for reachability analysis. Part II: Internal approximations box-valued constraints. *Optimization Methods and Software*, 17(2):207–237.

[Kurzhanski and Varaiya, 2014] Kurzhanski, A. B. and Varaiya, P. (2014). *Dynamics and Control of Trajectory Tubes*, volume 85 of *Systems & Control: Foundations & Applications*. Springer International Publishing, Cham.

[Kurzhanskiy and Varaiya, 2007] Kurzhanskiy, A. A. and Varaiya, P. (2007). Ellipsoidal techniques for reachability analysis of discrete-time linear systems. *IEEE Transactions on Automatic Control*, 52(1):26–38.

[Lainiotis, 1976] Lainiotis, D. G. (1976). Generalized Chandrasekhar Algorithms: Time-Varying Models. *IEEE Transactions on Automatic Control*, 21(5):728–732.

[Lee and Markus, 1969] Lee, E. B. and Markus, L. (1969). *Foundations of Optimal Control Theory.*, volume 132. John Wiley & Sons, New York.

[Leibfritz, 2006] Leibfritz, F. (2006). COMPleib: COnstrained Matrix optimization Problem library. Technical report, University of Trier, Department of Mathematics, Trier, Germany.

[Matveev and Yakubovich, 1997] Matveev, A. S. and Yakubovich, V. (1997). Nonconvex Problems of Global Optimization: Linear-Quadratic Control Problems with Quadratic Constraints. *Dynamics and Control*, 7(2):99–134.

[Megretski, 2010] Megretski, A. (2010). KYP Lemma for Non-Strict Inequalities and the associated Minimax Theorem. *arXiv preprint arXiv:1008.2552*, pages 1–24.

[Megretski and Rantzer, 1997] Megretski, A. and Rantzer, A. (1997). System analysis via integral quadratic constraints. *IEEE Transactions on Automatic Control*, 42(6):819–830.

[Miné, 2004] Miné, A. (2004). Weakly relational numerical abstract domains. page 322.

[Moore et al., 2009] Moore, R. E., Kearfott, R. B., and Cloud, M. J. (2009). *Introduction to Interval Analysis*, volume 15. Society for Industrial and Applied Mathematics.

[Nedialkov et al., 1999] Nedialkov, N. S., Jackson, K. R., and Corliss, G. F. (1999). Validated solutions of initial value problems for ordinary differential equations. *Applied Mathematics and Computation*, 105(1):21–68.

[Oulamara and Venet, 2015] Oulamara, M. and Venet, A. J. (2015). Abstract interpretation with higher-dimensional ellipsoids and conic extrapolation. In *Lecture Notes in Computer Science*, volume 9206, pages 415–430.

[Peaucelle et al., 2014] Peaucelle, D., Baudouin, L., and Gouaisbaut, F. (2014). Integral Quadratic Separators for performance analysis. In *2009 European Control Conference, ECC 2009*, pages 788–793, Budapest.

[Platzer and Clarke, 2009] Platzer, A. and Clarke, E. M. (2009). Computing differential invariants of hybrid systems as fixedpoints. In *Formal Methods in System Design*, volume 35, pages 98–120. Springer.

[Rantzer and Megretski, 1998] Rantzer, A. and Megretski, A. (1998). Analysis of Rate Limiters Using Integral Quadratic Constraints. *IFAC Proceedings Volumes*, 31(17):669–673.

[Rousse, 2019] Rousse, P. (2019). IQCARUS: IQC for Analysis of Reachability for Uncertain Systems. https://github.com/roussePaul/IQCARUS.

[Rousse et al., 2020a] Rousse, P., Alexandre dit Sandretto, J., Chapoutot, A., and Garoche, P.-L. (2020a). Guaranteed Simulation of Dynamical Systems with Integral Constraints and Application on Delayed Dynamical Systems. In *Lecture Notes in Computer Science*, volume 11971 LNCS.

[Rousse et al., 2019] Rousse, P., Garoche, P.-L., and Henrion, D. (2019). Parabolic set simulation for reachability analysis of linear time invariant systems with integral quadratic constraint. In *2019 18th European Control Conference, ECC 2019*.

[Rousse et al., 2020b] Rousse, P., Garoche, P.-L., and Henrion, D. (2020b). Parabolic Set Simulation for Reachability Analysis of Linear Time-Invariant Systems with Integral Quadratic Constraint. *European Journal of Control*.

[Rubinstein and Kroese, 2016] Rubinstein, R. Y. and Kroese, D. P. (2016). *Simulation and the Monte Carlo Method*. Wiley Series in Probability and Statistics. John Wiley & Sons, Hoboken, NJ, USA.

[Sandretto and Chapoutot, 2016] Sandretto, J. A. D. and Chapoutot, A. (2016). Validated simulation of differential algebraic equations with runge-kutta methods. *Reliable Computing*, 22:56–77.

[Sankaranarayanan and Fainekos, 2012] Sankaranarayanan, S. and Fainekos, G. (2012). Falsification of temporal properties of hybrid systems using the cross-entropy method. *Hybrid Systems: Computation and Control (HSCC)*, pages 125–134.

[Savkin and Petersen, 1995] Savkin, A. V. and Petersen, I. R. (1995). Recursive State Estimation for Uncertain Systems with an Integral Quadratic Constraint. *IEEE Transactions on Automatic Control*, 40(6):1080–1083.

[Savkin and Petersen, 1996a] Savkin, A. V. and Petersen, I. R. (1996a). Model validation for robust control of uncertain systems with an integral quadratic constraint. *Automatica*, 32(4):603–606.

[Savkin and Petersen, 1996b] Savkin, A. V. and Petersen, I. R. (1996b). Robust state estimation for uncertain systems with averaged integral quadratic constraints. *International Journal of Control*, 64(5):923–939.

[Scherer and Veenman, 2018] Scherer, C. W. and Veenman, J. (2018). Stability analysis by dynamic dissipation inequalities: On merging frequency-domain techniques with time-domain conditions. *Systems and Control Letters*, 121:7–15.

[Schuricht and von der Mosel, 2000] Schuricht, F. and von der Mosel, H. (2000). Ordinary differential equations with measurable right-hand side and parameters in metric spaces. In *Sonderforschungsbereich 256*. Universität Bonn.

[Schweppe, 1973] Schweppe, F. C. (1973). *Uncertain dynamic systems*.

[Seiler et al., 2017] Seiler, P., Buch, J., Moore, R. M., Meissen, C., Arcak, M., and Packard, A. (2017). LTVTools (Beta), A MATLAB Toolbox for Linear Time-Varying System. https://github.com/buchjyot/LTVTools.

[Seiler et al., 2019] Seiler, P., Moore, R. M., Meissen, C., Arcak, M., and Packard, A. (2019). Finite horizon robustness analysis of LTV systems using integral quadratic constraints. *Automatica*, 100:135–143.

[Seuret and Gouaisbaut, 2015] Seuret, A. and Gouaisbaut, F. (2015). Hierarchy of LMI conditions for the stability analysis of time-delay systems. *Systems and Control Letters*, 81:1–7.

[Soravia, 2000] Soravia, P. (2000). Viscosity solutions and optimal control problems with integral constraints. *Systems and Control Letters*, 40(5):325–335.

[Sturm, 1999] Sturm, J. F. (1999). Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11(1):625–653.

[Tarski, 1955] Tarski, A. (1955). A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5(2):285–309.

[Varaiya, 2000] Varaiya, P. (2000). Reach Set Computation Using Optimal Control. In *Verification of Digital and Hybrid Systems*, pages 323–331. Springer.

[Veenman et al., 2016] Veenman, J., Scherer, C. W., and Köroğlu, H. (2016). Robust stability and performance analysis based on integral quadratic constraints. *European Journal of Control*, 31:1–32.

[Wang et al., 2016] Wang, G., Liu, J., Sun, H., Liu, J., Ding, Z., and Zhang, M. (2016). Safety verification of state/time-driven hybrid systems using barrier certificates. In *2016 35th Chinese Control Conference (CCC)*, pages 2483–2489. Springer, IEEE.

[Yakubovich, 1967] Yakubovich, V. (1967). Frequency conditions for the absolute stability of control systems with several nonlinear or linear nonstationary blocks. *Avtomat. i Telemekh.*, (6):5–30.

[Yin et al., 2020] Yin, H., Packard, A., Arcak, M., and Seiler, P. (2020). Reachability analysis using dissipation inequalities for uncertain nonlinear systems. *Systems and Control Letters*, 142:104736.

[Zames, 1966a] Zames, G. (1966a). On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems Part I: Conditions Derived Using Concepts of Loop Gain, Conicity, and Positivity. *IEEE Transactions on Automatic Control*, 11(2):228–238.

[Zames, 1966b] Zames, G. (1966b). On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems-Part II: Conditions Involving Circles in the Frequency Plane and Sector Nonlinearities. *IEEE Transactions on Automatic Control*, 11(3):465–476.

[Zeidler, 1995a] Zeidler, E. (1995a). *Applied Functional Analysis*, volume 108 of *Applied Mathematical Sciences*. Springer New York, New York, NY.

[Zeidler, 1995b] Zeidler, E. (1995b). *Applied Functional Analysis*, volume 109 of *Applied Mathematical Sciences*. Springer New York, New York, NY.

List of Figures

List of Tables