APPROXIMATE BISIMULATION RELATIONS FOR CONSTRAINED LINEAR SYSTEMS

ANTOINE GIRARD AND GEORGE J. PAPPAS

Abstract. In this paper, we define the notion of approximate bisimulation relation between two systems, extending the well established exact bisimulation relations for discrete and continuous systems. Exact bisimulation requires that the observations of two systems are and remain identical, approximate bisimulation allows the observation to be different provided they are and remain arbitrarily close. Approximate bisimulation relations are conveniently defined as level sets of a function called bisimulation function. For the class of linear systems with constrained initial states and constrained inputs, we develop effective characterizations for bisimulation functions that can be interpreted in terms of linear matrix inequalities, set inclusion and games. We derive a computationally effective algorithm to evaluate the precision of the approximate bisimulation between a constrained linear system and its projection. This algorithm has been implemented in a MATLAB toolbox: MATISSE. Two examples of use of the toolbox in the context of safety verification are shown.

1. Introduction

Well established notions of system refinement and equivalence for discrete systems such as language inclusion, simulation and bisimulation relations have been shown useful to reduce the complexity of formal verification [CGP00]. Much more recently, the notions of simulation and bisimulation relations have been extended to continuous and hybrid state spaces resulting in new equivalence notions for nondeterministic continuous [Pap03, TP04, vdS04] and hybrid systems [HTP05, JvdS04, PvdSB04].

These abstraction concepts are all *exact*, requiring external behaviors of two systems to be identical. Approximate relationships which do allow for the possibility of error, will certainly allow for more dramatic system compression while providing more robust relationships between systems. An approach based on approximate versions of simulation and bisimulation relations seems promising and this idea has been explored recently for quantitative [dAFS04], stochastic [DGJP04, vBMOW03] and metric [GP05c] transition systems.

In [GP05c], we developed an approximation framework which applies for both discrete and continuous metric transition systems. We defined an approximate version of bisimulation relations based on a metric on the set of observation. While, exact bisimulation requires that the observations of two systems are and remain identical, approximate bisimulation allows the observation to be different provided they are and remain arbitrarily close (*i.e.* the distance between the observations is and remains bounded by the *precision* of the approximate bisimulation). Approximate bisimulation relations can be characterized as level sets of a function called bisimulation functions. A bisimulation function is a function bounding the distance between the observations of two systems and non-increasing under their parallel evolutions. This Lyapunov-like property allows to design computationally effective methods for the computation of bisimulation functions. Computational approaches have been

This research is partially supported by the Région Rhône-Alpes (Projet CalCel) and the National Science Foundation Presidential Early CAREER (PECASE) Grant 0132716.

developed for metric transition systems derived from constrained linear dynamical systems [GP05a] and nonlinear (but deterministic) dynamical systems [GP05b].

This line of research has been motivated by the algorithmic verification of hybrid systems. The significant progress in the formal verification of discrete systems [BCM$^+$90], has inspired a plethora of sophisticated methods for safety verification of continuous and hybrid systems. The approaches range from discrete and predicate abstraction methods [AHP00, ADI02, TK02], to reachability computations [ABDM00, ADG03, CK99, KV00, MT00, Gir05], to Lyapunov-like barriers [PJ04]. However, progress on continuous (and thus hybrid) systems has been limited to systems of small continuous dimension, motivating research on model reduction [HK04], and projection based methods [AD04] for safety verification. In [GP05a, GP05b], we showed that our approximation framework could be used to reduce significantly the complexity of algorithmic verification of continuous systems allowing to consider dynamics of larger dimension.

In this paper, we improve and extend our work presented in [GP05a] for the computation of bisimulation functions for a class of linear systems with constrained initial states and constrained inputs. We develop a characterization of bisimulation functions based on Lyapunov-like differential inequalities. We show that for a specific class of bisimulation functions based on quadratic forms these inequalities can be interpreted in terms of linear matrix inequalities, set inclusion and optimal values of static games. We derive an efficient algorithm to evaluate the precision of the approximate bisimulation between a constrained linear system and its projection. This algorithm has been implemented in a MATLAB toolbox: MATISSE (Metrics for Approximate TransItion Systems Simulation and Equivalence [GJP05]) available for download.

Compared to other approximation frameworks for linear systems such as traditional model reduction techniques [ASG00, BDG96, HK04], the reduction problem we consider is quite different and much more natural for safety verification for the following reasons. First, the systems we consider have constrained inputs which are internal (and hence they should be thought of as internal disturbances). The fidel reproduction of the input-output mapping is therefore not our main concern. Second, we do not assume that the systems are initially at the equilibrium: contrarily to the model reduction framework, the transient dynamics of the systems are not ignored during the approximation process. From the point of view of verification, the transient phase and the asymptotic phase of a trajectory are of equal importance. In fact, the quality of the approximation may critically depend on initial set of states. Finally, since our research has been motivated by the algorithmic verification of continuous and hybrid systems, the error bounds we compute are based on the L^∞ norm which is the only norm which makes sense for safety verification. In comparison, in [ASG00, BDG96], the error bounds stand for the L^2 norm; in [HK04] the error bound is valid only on a time interval of finite length. We conclude this paper by illustrating this point in the context of safety verification for constrained linear systems.

2. Approximation of transition systems

2.1. Metric transition systems.

The theory has been developed in [GP05c] within the framework of metric transition systems. In this section, we summarize the main results. Metric transition systems can be seen as graphs (possibly with an infinite number of states and transitions) whose set of states and set of observations are metric spaces.

Definition 2.1 (Metric transition system)**.** A (labeled) transition system with observations is a tuple $T = (Q, \Sigma, \rightarrow, Q^0, \Pi, \langle\langle . \rangle\rangle)$ that consists of:

- a (possibly infinite) set Q of states,
- a (possibly infinite) set Σ of labels,
- a transition relation $\rightarrow \subseteq Q \times \Sigma \times Q$,
- a (possibly infinite) set $Q^0 \subseteq Q$ of initial states,
- a (possibly infinite) set Π of observations,
- an observation map $\langle\langle . \rangle\rangle : Q \rightarrow \Pi$.

If (Q, d_Q) and (Π, d_Π) are metric spaces, then T is called a metric transition system. The set of metric transition systems associated to a set of labels Σ and a set of observations Π is denoted $\mathcal{T}_M(\Sigma, \Pi)$.

A transition $(q, \sigma, q') \in \rightarrow$ will be denoted $q \xrightarrow{\sigma} q'$. We assume that the systems we consider are *non-blocking* so that for all $q \in Q$, there exists at least one transition $q \xrightarrow{\sigma} q'$ of T. For all labels $\sigma \in \Sigma$, the σ-successor is defined as the set valued map given by

$$\forall q \in Q,\ \mathrm{Post}^\sigma(q) = \left\{ q' \in Q |\ q \xrightarrow{\sigma} q' \right\}.$$

We denote with $\mathrm{Supp}(\mathrm{Post}^\sigma)$ the support of the σ-successor which is the subset of elements $q \in Q$ such that $\mathrm{Post}^\sigma(q)$ is not empty.

Assumption 2.2. The metric transition systems we consider satisfy the following properties:

(1) the set of initial values Q_0 is a compact subset of Q,
(2) for all $\sigma \in \Sigma$, for all $q \in \mathrm{Supp}(\mathrm{Post}^\sigma)$, $\mathrm{Post}^\sigma(q)$ is a compact subset of Q.

A state trajectory of T is an infinite sequence of transitions,

$$q^0 \xrightarrow{\sigma^0} q^1 \xrightarrow{\sigma^1} q^2 \xrightarrow{\sigma^2} \dots, \text{ where } q^0 \in Q^0.$$

An external trajectory is a sequence of elements of $\Pi \times \Sigma \times \Pi$ of the form $\pi^0 \xrightarrow{\sigma^0} \pi^1 \xrightarrow{\sigma^1} \pi^2 \xrightarrow{\sigma^2} \dots$. An external trajectory is accepted by transition system T if there exists a state trajectory of T, such that for all $i \in \mathbb{N}$, $\pi^i = \langle\langle q^i \rangle\rangle$. The set of external trajectories accepted by transition system T is called the language of T, and is denoted by $L(T)$. The reachable set of T is the subset of Π defined by:

$$\mathrm{Reach}(T) = \left\{ \pi \in \Pi |\ \exists \{\pi^i \xrightarrow{\sigma^i} \pi^{i+1}\}_{i \in \mathbb{N}} \in L(T),\ \exists j \in \mathbb{N},\ \pi^j = \pi \right\}.$$

An important problem for transition systems is the *safety verification problem* which asks whether the intersection of $\mathrm{Reach}(T)$ with an unsafe set $\Pi_U \subseteq \Pi$ is empty or not.

2.2. **Approximate bisimulation relations.** Let $T_1 = (Q_1, \Sigma_1, \rightarrow_1, Q_1^0, \Pi_1, \langle\langle . \rangle\rangle_1)$ and $T_2 = (Q_2, \Sigma_2, \rightarrow_2, Q_2^0, \Pi_2, \langle\langle . \rangle\rangle_2)$ be two metric transition systems with the same set of labels ($\Sigma_1 = \Sigma_2 = \Sigma$) and the same set of observations ($\Pi_1 = \Pi_2 = \Pi$) (*i.e.* T_1 and T_2 are elements of $\mathcal{T}_M(\Sigma, \Pi)$).

The notion of approximate bisimulation relation is obtained from exact bisimulation relation [CGP00] by relaxation of the observational equivalence constraint. Instead of requiring that the observations of the two systems are and remain the same, we require that they are and remain arbitrarily close.

Definition 2.3 (Approximate bisimulation)**.** Let T_1, $T_2 \in \mathcal{T}_M(\Sigma, \Pi)$. A relation $\mathcal{R}_\delta \subseteq Q_1 \times Q_2$ is called a δ-approximate bisimulation relation between T_1 and T_2 if for all $(q_1, q_2) \in \mathcal{R}_\delta$:

(1) $d_\Pi\left(\langle\langle q_1 \rangle\rangle_1, \langle\langle q_2 \rangle\rangle_2\right) \leq \delta$,
(2) for all $q_1 \xrightarrow{\sigma}_1 q_1'$, there exists $q_2 \xrightarrow{\sigma}_2 q_2'$ such that $(q_1', q_2') \in \mathcal{R}_\delta$,

(3) for all $q_2 \xrightarrow{\sigma}_2 q_2'$, there exists $q_1 \xrightarrow{\sigma}_1 q_1'$ such that $(q_1', q_2') \in \mathcal{R}_\delta$.

Since d_Π is a metric, for $\delta = 0$, we recover the established definition of exact bisimulation relations. Parameter δ can thus serve to measure how far T_1 and T_2 are from being exactly bisimilar.

Definition 2.4. Transition systems T_1 and T_2 are approximately bisimilar with precision δ (noted $\mathbb{T}_1 \sim_\delta T_2$), if there exists $\mathcal{R}_\delta$, a δ-approximate bisimulation relation between T_1 and T_2 such that for all $q_1 \in Q_1^0$, there exists $q_2 \in Q_2^0$ such that $(q_1, q_2) \in \mathcal{R}_\delta$ and conversely.

The distance between the languages of two approximately bisimilar transition systems is bounded by the precision of the approximate bisimulation relation:

Theorem 2.5 (adapted from [GP05c])**.** *If $T_1 \sim_\delta T_2$ then for all external trajectory accepted by T_1 (respectively T_2), $\pi_1^0 \xrightarrow{\sigma^0} \pi_1^1 \xrightarrow{\sigma^1} \pi_1^2 \xrightarrow{\sigma^2} \dots$ there exists an external trajectory accepted by T_2 (respectively T_1), with the same sequence of labels, $\pi_2^0 \xrightarrow{\sigma^0} \pi_2^1 \xrightarrow{\sigma^1} \pi_2^2 \xrightarrow{\sigma^2} \dots$ such that for all $i \in \mathbb{N}$, $d_\Pi(\pi_1^i, \pi_2^i) \leq \delta$.*

Proof. If $T_1 \sim_\delta T_2$ then there exists a δ-approximate bisimulation relation $\mathcal{R}_\delta$ as in Definition 2.4. Let $\pi_1^0 \xrightarrow{\sigma^0} \pi_1^1 \xrightarrow{\sigma^1} \pi_1^2 \xrightarrow{\sigma^2} \dots$ be an external trajectory accepted by T_1. Then, there exists a state trajectory of T_1, $q_1^0 \xrightarrow{\sigma^0}_1 q_1^1 \xrightarrow{\sigma^1}_1 q_1^2 \xrightarrow{\sigma^2}_1 \dots$ such that for all $i \in \mathbb{N}$, $\langle\langle q_1^i \rangle\rangle_1 = \pi_1^i$. From Definition 2.4, there exists $q_2^0 \in Q_2^0$ such that $(q_1^0, q_2^0) \in \mathcal{R}_\delta$. By induction, it is easy to see that there exists a state trajectory of T_2 starting from q_2^0 and with the sequence of labels: $q_2^0 \xrightarrow{\sigma^0}_2 q_2^1 \xrightarrow{\sigma^1}_2 q_2^2 \xrightarrow{\sigma^2}_2 \dots$ and such that for all $i \in \mathbb{N}$, $(q_1^i, q_2^i) \in \mathcal{R}_\delta$. Then, the external trajectory, $\pi_2^0 \xrightarrow{\sigma^0} \pi_2^1 \xrightarrow{\sigma^1} \pi_2^2 \xrightarrow{\sigma^2} \dots$, where $\pi_2^i = \langle\langle q_2^i \rangle\rangle_2$ for all $i \in \mathbb{N}$, is accepted by T_2 and, for all $i \in \mathbb{N}$, $d_\Pi(\pi_1^i, \pi_2^i) = d_\Pi\left(\langle\langle q_1^i \rangle\rangle_1, \langle\langle q_2^i \rangle\rangle_2\right) \leq \delta$. □

From Theorem 2.5, it is straightforward that if $T_1 \sim_\delta T_2$ then the distance between the reachable sets of T_1 and T_2 is bounded by the precision δ. In the context of safety verification, this approximation property is of great use since

$$\mathrm{Reach}(T_2) \cap N(\Pi_U, \delta) = \emptyset \implies \mathrm{Reach}(T_1) \cap \Pi_U = \emptyset$$

where $N(\pi, \delta)$ denotes the δ neighborhood of $\pi \in \Pi$.

2.3. **Bisimulation functions.** The problem of system approximation can be handled more practically by a dual approach to the one based on approximate bisimulation relations. It is based on a class of functions called bisimulation functions. A bisimulation function between T_1 and T_2 is a positive function defined on $Q_1 \times Q_2$, bounding the distance between the observations associated to the couple (q_1, q_2) and non increasing under the dynamics of the systems.

Definition 2.6 (Bisimulation function)**.** A function $V : Q_1 \times Q_2 \to \mathbb{R}^+ \cup \{+\infty\}$ is called a bisimulation function between T_1 and T_2 if its level sets are closed, and for all $(q_1, q_2) \in Q_1 \times Q_2$:

$$V(q_1, q_2) \geq \max\left(d_\Pi\left(\langle\langle q_1 \rangle\rangle_1, \langle\langle q_2 \rangle\rangle_2\right), \sup_{q_1 \xrightarrow{\sigma}_1 q_1'} \inf_{q_2 \xrightarrow{\sigma}_2 q_2'} V(q_1', q_2'), \sup_{q_2 \xrightarrow{\sigma}_2 q_2'} \inf_{q_1 \xrightarrow{\sigma}_1 q_1'} V(q_1', q_2') \right).$$

Before we are able to give our main results, we need to prove the following Lemma:

Lemma 2.7. *Let $V : Q_1 \times Q_2 \to \mathbb{R}^+ \cup \{+\infty\}$ be a function with closed level sets. Let C_2 be a compact subset of Q_2, then*

$$\forall q_1 \in Q_1,\ \exists q_2 \in C_2,\ \text{such that } V(q_1, q_2) = \inf_{q_2 \in C_2} V(q_1, q_2).$$

Proof. Let $q_1 \in Q_1$, and $\delta = \inf_{q_2 \in C_2} V(q_1, q_2)$. Let $\{\varepsilon_i\}_{i \in \mathbb{N}}$ be a decreasing sequence of real numbers converging to 0. Then, for all $i \in \mathbb{N}$, there exists $q_2^i \in C_2$ such that $V(q_1, q_2^i) \leq \delta + \varepsilon_i$. Since C_2 is compact, there exists a subsequence of $\{q_2^i\}_{i \in \mathbb{N}}$ which we will also note $\{q_2^i\}_{i \in \mathbb{N}}$ and which converges to a limit $q_2 \in C_2$. Let $\varepsilon > 0$, there exists $n \in \mathbb{N}$, such that for all $i \geq n$, $\varepsilon_i \leq \varepsilon$. Then, for all $i \geq n$, $V(q_1, q_2^i) \leq \delta + \varepsilon$. Since the levels sets of V are closed subsets of $Q_1 \times Q_2$, it follows that $V(q_1, q_2) \leq \delta + \varepsilon$. Finally, since this holds for all $\varepsilon > 0$, $V(q_1, q_2) \leq \delta$. □

The duality of the approach using approximate bisimulation relations and the approach using bisimulation functions is captured by the following result:

Theorem 2.8. [GP05c] *Let V be a bisimulation function between T_1 and T_2. Then for all $\delta \geq 0$,*

$$\mathcal{R}_\delta = \{(q_1, q_2) \in Q_1 \times Q_2 |\ V(q_1, q_2) \leq \delta\}$$

is a δ-approximate bisimulation relation between T_1 and T_2.

Proof. It is clear that $\mathcal{R}_\delta$ satisfies the first property of Definition 2.3. Let $(q_1, q_2) \in \mathcal{R}_\delta$, let $q_1 \xrightarrow{\sigma}_1 q_1'$, from the regularity assumptions we made on T_2, the set $\mathrm{Post}_2^\sigma(q_2)$ is compact and therefore from Lemma 2.7, there exists $q_2 \xrightarrow{\sigma}_2 q_2'$ such that

$$V(q_1', q_2') = \inf_{q_2 \xrightarrow{\sigma}_2 q_2'} V(q_1', q_2') \leq \sup_{q_1 \xrightarrow{\sigma}_1 q_1'} \inf_{q_2 \xrightarrow{\sigma}_2 q_2'} V(q_1', q_2') \leq V(q_1, q_2) \leq \delta.$$

Then, (q_1', q_2') is an element of $\mathcal{R}_\delta$ and the second property as well as the third property (using symmetrical arguments) of Definition 2.3 hold. $\mathcal{R}_\delta$ is consequently a δ-approximate bisimulation relation between T_1 and T_2. □

Let us remark that particularly the zero set of a bisimulation function is an exact bisimulation relation. From Theorem 2.5, it is clear that for good approximation results it is necessary to have a tight evaluation of the precision δ for which $T_1 \sim_\delta T_2$. Given a bisimulation function between T_1 and T_2, this can be done by solving a game.

Theorem 2.9 (adapted from [GP05c])**.** *Let V be a bisimulation function between T_1 and T_2 and*

$$\delta = \max \left(\sup_{q_1 \in Q_1^0} \inf_{q_2 \in Q_2^0} V(q_1, q_2),\ \sup_{q_2 \in Q_2^0} \inf_{q_1 \in Q_1^0} V(q_1, q_2) \right). \tag{2.1}$$

If the value of δ is finite, then $T_1 \sim_\delta T_2$.

Proof. Let $q_1 \in Q_1^0$, from the regularity assumptions we made on T_2, the set Q_2^0 is compact and therefore from Lemma 2.7, there exists $q_2 \in Q_2^0$ such that

$$V(q_1, q_2) = \inf_{q_2 \in Q_2^0} V(q_1, q_2) \leq \sup_{q_1 \in Q_1^0} \inf_{q_2 \in Q_2^0} V(q_1, q_2) = \delta.$$

Hence, for all $q_1 \in Q_1^0$, there exists $q_2 \in Q_2^0$ such that (q_1, q_2) is in $\mathcal{R}_\delta$, the δ-approximate bisimulation relation of T_1 by T_2 defined in Theorem 2.8. Similarly, we can show that for all $q_2 \in Q_2^0$, there exists $q_1 \in Q_1^0$ such that (q_1, q_2) is in $\mathcal{R}_\delta$. Then, $T_1 \sim_\delta T_2$. □

Let us remark that for any $\delta' \geq \delta$, we also have $T_1 \sim_{\delta'} T_2$. It appears that one of the challenge of this theory of system approximation is the computation of bisimulation functions. The purpose of this paper is to address this problem for the class of continuous-time constrained linear systems.

3. Bisimulation functions for constrained linear systems

Let us consider the following class of linear systems with constrained inputs and constrained initial states:

$$\Delta_i : \left\{ \begin{array}{lcl} \dot{x}_i(t) & = & A_i x_i(t) + B_i u_i(t),\ x_i(t) \in \mathbb{R}^{n_i},\ u_i(t) \in U_i,\ x_i(0) \in I_i \\ y_i(t) & = & C_i x_i(t),\ y_i(t) \in \mathbb{R}^p \end{array} \right. ,\ i = 1, 2 \tag{3.1}$$

where I_i is a compact subset of $\mathbb{R}^{n_i}$ and U_i is a convex compact subset of $\mathbb{R}^{m_i}$. One may want to think of I_i and U_i as bounded polytopes. The constrained inputs of systems Δ_1 and Δ_2 are to be understood as disturbances rather than control variables. In the spirit of [Pap03], the dynamical system Δ_i can be written as a non-deterministic metric transition system $T_{\Delta_i} = (Q_i, \Sigma_i, \to_i, Q_i^0, \Pi_i, \langle\langle . \rangle\rangle_i)$ where:

- The set of states is $Q_i = \mathbb{R}^{n_i}$,
- The set of labels is $\Sigma_i = \mathbb{R}^+$,
- The transition relation $\to_i \subseteq Q_i \times \Sigma_i \times Q_i$ is given by $x \xrightarrow{t}_i x'$ if and only if there exists a locally measurable function $u(.)$ such that

$$\text{for all } s \in [0, t],\ u(s) \in U_i \text{ and } x' = e^{A_i t} x + \int_0^t e^{A_i (t-s)} B_i u(s) ds,$$

- The set of initial states is $Q_i^0 = I_i$,
- The set of observations is $\Pi_i = \mathbb{R}^p$,
- The observation map is given by $\langle\langle x \rangle\rangle_i = C_i x$.

The sets of states and observations are equipped with the traditional Euclidean distance. Note that T_{Δ_1} and T_{Δ_2} are elements of the set of metric transition systems $\mathcal{T}_M(\mathbb{R}^+, \mathbb{R}^p)$. We can check that they satisfy Assumption 2.2 (see for instance [Aub01]). Let us introduce the following notations:

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \overline{B}_1 = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \overline{B}_2 = \begin{bmatrix} 0 \\ B_2 \end{bmatrix}, C = \begin{bmatrix} C_1 & -C_2 \end{bmatrix} \text{ and } x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

For such metric transition systems, Definition 2.6 is not a very convenient way to characterize bisimulation functions. In this section, we derive a characterization of bisimulation functions based on Lyapunov-like differential inequalities. The different characterizations of bisimulation functions given in this paper are derived from the following result:

Proposition 3.1. *Let $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^+$ be a C^1 function and $\mathcal{H}$ a subspace of $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ such that for all $u_1 \in U_1$ there exists $u_2 \in U_2$ satisfying $(u_1, u_2) \in \mathcal{H}$. Let us assume that there exists $\eta \geq 0$, such that for all (x_1, x_2) satisfying $f(x_1, x_2) \geq \eta$,*

$$\sup_{u_1 \in U_1} \left(\inf_{u_2 \in U_2, (u_1, u_2) \in \mathcal{H}} \nabla f(x)(Ax + \overline{B}_1 u_1 + \overline{B}_2 u_2) \right) \leq 0. \tag{3.2}$$

Then, for all $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, for all $x_1 \xrightarrow{t}_1 x_1'$, there exists $x_2 \xrightarrow{t}_2 x_2'$ such that

$$f(x_1', x_2') \leq \max(f(x_1, x_2), \eta).$$

Moreover, there exist inputs $u_i(.)$ $(i = 1, 2)$ leading Δ_i from x_i to x_i' at time t and such that for all $s \in [0, t]$, $(u_1(s), u_2(s)) \in \mathcal{H}$.

Let us remark that a symmetrical result holds when the maximization is done over U_2 and the minimization over U_1. The proof of this proposition is quite technical and is therefore stated in appendix. In the following, based on Proposition 3.1, we will derive several characterizations for bisimulation functions. Particularly, for smooth bisimulation functions with finite values on $\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$, we have:

Theorem 3.2. *Let $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^+$ be a C^1 function. If for all $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$,*

$$f(x) \geq x^T C^T C x, \tag{3.3}$$

$$\sup_{u_1\in U_1} \inf_{u_2\in U_2} \nabla f(x)(Ax + \overline{B}_1 u_1 + \overline{B}_2 u_2) \leq 0, \tag{3.4}$$

$$\sup_{u_2\in U_2} \inf_{u_1\in U_1} \nabla f(x)(Ax + \overline{B}_1 u_1 + \overline{B}_2 u_2) \leq 0. \tag{3.5}$$

Then, $V(x_1, x_2) = \sqrt{f(x_1, x_2)}$ is a bisimulation function between T_{Δ_1} and T_{Δ_2}.

Proof. Let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $x_1 \xrightarrow{t}_1 x_1'$. It follows from equation (3.4) and Proposition 3.1 (with $\mathcal{H} = \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ and $\eta = 0$) that there exists $x_2 \xrightarrow{t}_2 x_2'$ such that $f(x_1', x_2') \leq f(x_1, x_2)$. Symmetrically, we can show from equation (3.5) that for all $x_2 \xrightarrow{t}_2 x_2'$, there exists $x_1 \xrightarrow{t}_1 x_1'$ such that $f(x_1', x_2') \leq f(x_1, x_2)$. Together with equation (3.3), this allows to conclude that $V(x_1, x_2) = \sqrt{f(x_1, x_2)}$ is a bisimulation function between T_{Δ_1} and T_{Δ_2}. □

Remark 3.3. There are similarities between the notions of bisimulation function and robust control Lyapunov function [FK96, LSW02] as well as some significant conceptual differences. Indeed, let us consider the input u_1 as a disturbance and the input u_2 as a control variable in equation (3.4). Then, the interpretation of this inequality is that for all disturbances there exists a control such that the bisimulation function does not increase during the evolution of the system. This means that the choice of u_2 can be made with the knowledge of u_1. In comparison, a robust control Lyapunov function requires that there exists a control u_2 such that for all disturbances u_1, the function decreases during the evolution of the system. Thus, it appears that robust control Lyapunov functions require stronger conditions than bisimulation functions.

Example 3.4. Let us consider the following three dimensional dynamical system Δ_1 whose dynamics is given by:

$$\begin{cases} \dot{z}_1(t) = -8z_1(t) + 7z_2(t) - 7z_3(t) + u_1(t) \\ \dot{z}_2(t) = 3z_1(t) + z_2(t) + 4z_3(t) + u_2(t) \\ \dot{z}_3(t) = 2z_1(t) + 3z_2(t) + 2z_3(t) - u_1(t) + u_2(t) \end{cases}$$

The system is observed through the output variable $y_1(t) = z_1(t)$. The value of the inputs are constrained in the following way: $u_1(t) \in [-1, 1]$, $u_2(t) \in [0, 2]$. The set of initial states is the polytope I_1 defined by

$$I_1 = \{6 \leq z_1(0) \leq 8,\ -2 \leq z_2(0) \leq -3,\ -1 \leq z_1(0) - z_2(0) + z_3(0) \leq 1\}.$$

As stated previously, we can derive from Δ_1 a metric transition system $T_{\Delta_1} \in \mathcal{T}_M(\mathbb{R}^+, \mathbb{R})$. We want to show that T_{Δ_1} can be approximated by the metric transition system $T_{\Delta_2} \in \mathcal{T}_M(\mathbb{R}^+, \mathbb{R})$ derived from the one dimensional dynamical system Δ_2 whose dynamics is given by:

$$\dot{x}(t) = -x(t) + v(t).$$

Δ_2 is observed through the variable $y_2(t) = x(t)$. The value of the input $v(t)$ is constrained in the interval $[-1, 1]$. The set of initial states is the interval $[5.5, 8.5]$. Let us show that the function

$$V(z_1, z_2, z_3, x) = \sqrt{(z_1 - x)^2 + (z_2 - z_3 - x)^2}$$

is a bisimulation function between T_{Δ_1} and T_{Δ_2}. Let $f(z_1, z_2, z_3, x) = (z_1 - x)^2 + (z_2 - z_3 - x)^2$, we first remark that

$$f(z_1, z_2, z_3, x) \geq (z_1 - x)^2.$$

Hence, equation (3.3) holds. Moreover, we can check that

$$\frac{\partial f}{\partial z_1}\dot{z}_1 + \frac{\partial f}{\partial z_2}\dot{z}_2 + \frac{\partial f}{\partial z_3}\dot{z}_3 + \frac{\partial f}{\partial x}\dot{x} = -4(2z_1 - z_2 + z_3 - x)^2 + 2(z_1 + z_2 - z_3 - 2x)(u_1 - v).$$

Hence, equations (3.4) and (3.5) also hold. Therefore V is a bisimulation function between T_{Δ_1} and T_{Δ_2}. Let us use this bisimulation function to evaluate the precision of the approximate bisimulation between T_{Δ_1} and T_{Δ_2}. We can check that

$$\begin{aligned}\sup_{(z_1,z_2,z_3)\in I_1} \inf_{x\in I_2} V(z_1, z_2, z_3, x) &= \sup_{(z_1,z_2,z_3)\in I_1} V\left(z_1, z_2, z_3, \frac{z_1 + z_2 - z_3}{2}\right)\\ &= \sup_{(z_1,z_2,z_3)\in I_1} \frac{|z_1 - z_2 + z_3|}{\sqrt{2}} = \frac{1}{\sqrt{2}}.\end{aligned}$$

On the other hand, we have

$$\sup_{x\in I_2} \inf_{(z_1,z_2,z_3)\in I_1} V(z_1, z_2, z_3, x) = \sup_{x\in I_2} \inf_{z_1\in[6,8]} \sqrt{(z_1 - x)^2} = \frac{1}{2}.$$

From Theorem 2.9, T_{Δ_1} and T_{Δ_2} are approximately bisimilar with the precision $\delta = 1/2$. As explained in the previous section, we can use this result to compute an approximation of the reachable set of T_{Δ_1}. We have that $\text{Reach}(T_{\Delta_1}) \subseteq N(\text{Reach}(T_{\Delta_2}), 1/2)$ and $\text{Reach}(T_{\Delta_2}) \subseteq N(\text{Reach}(T_{\Delta_1}), 1/2)$. The reachable set of T_{Δ_2} is easily computable and is equal to $(-1, 8.5]$. Therefore, we get that

$$(-0.5, 8] \subseteq \text{Reach}(T_{\Delta_1}) \subseteq (-1.5, 9].$$

4. Quadratic and truncated quadratic bisimulation functions

In this section, we show how Proposition 3.1 can be used to derive effective characterizations of bisimulation functions for constrained linear systems. Let us assume that both systems Δ_1 and Δ_2 are asymptotically stable (*i.e.* all the eigenvalues of A_1 and A_2 have strictly negative real parts). The non-stable case will be considered further in the paper.

4.1. **Quadratic bisimulation functions.** Let us remark that equations (3.4) (3.5) of Theorem 3.2 are Lyapunov-like differential inequalities. For autonomous linear systems, it is well known that quadratic forms provide universal and computationally effective Lyapunov functions. Hence, it seems reasonable to search a bisimulation function of the form:

$$\forall (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2},\ V(x_1, x_2) = \sqrt{x^T M x}. \tag{4.1}$$

Then, Theorem 3.2 becomes

Proposition 4.1. *If for all* $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$*,*

$$x^T M x \geq x^T C^T C x, \tag{4.2}$$

$$x^T M A x + x^T A^T M x + 2 \sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \leq 0, \tag{4.3}$$

$$x^T M A x + x^T A^T M x + 2 \sup_{u_2 \in U_2} \inf_{u_1 \in U_1} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \leq 0. \tag{4.4}$$

Then, $V(x_1, x_2) = \sqrt{x^T M x}$ *is a bisimulation function between* T_{Δ_1} *and* T_{Δ_2}*.*

We can find equivalent conditions in terms of linear matrix inequalities and set equalities:

Theorem 4.2. *Equations (4.2) (4.3) and (4.4) are equivalent to*

$$M \geq C^T C, \tag{4.5}$$

$$A^T M + M A \leq 0, \tag{4.6}$$

$$\ker(M) + \overline{B}_1 U_1 = \ker(M) - \overline{B}_2 U_2. \tag{4.7}$$

If equations (4.5), (4.6) and (4.7) hold then $V(x_1, x_2) = \sqrt{x^T M x}$ *is a bisimulation function between* T_{Δ_1} *and* T_{Δ_2}*.*

Proof. First, equation (4.5) is equivalent to equation (4.2). Let us assume that equations (4.6) and (4.7) hold. Then, for all $u_1 \in U_1$, there exists $v \in \ker(M)$ and $u_2 \in U_2$ such that $\overline{B}_1 u_1 = v - \overline{B}_2 u_2$. Then,

$$\forall x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, \ \sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \leq 0.$$

Similarly, for all $u_2 \in U_2$, there exists $v \in \ker(M)$ and $u_1 \in U_1$ such that $-\overline{B}_2 u_2 = v + \overline{B}_1 u_1$. Then,

$$\forall x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, \ \sup_{u_2 \in U_2} \inf_{u_1 \in U_1} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \leq 0.$$

From the linear matrix inequality (4.6), it follows that equations (4.3) and (4.4) hold. Conversely, let us assume that equation (4.3) holds. Let $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, then

$$\forall \lambda > 0, \ \lambda^2 (x^T A^T M x + x^T M A x) + 2\lambda \sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \leq 0. \tag{4.8}$$

If $x^T A^T M x + x^T M A x > 0$, then for λ sufficiently large, inequality (4.8) cannot hold. Necessarily, we have the linear matrix inequality (4.6). If

$$\sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) > 0$$

then for λ sufficiently small equation (4.8) cannot hold. Hence,

$$\forall x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, \ \sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \leq 0. \tag{4.9}$$

Since U_1 and U_2 are compact and then bounded, we have

$$\begin{aligned} \sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) &= \sup_{u_1 \in U_1} x^T M \overline{B}_1 u_1 + \inf_{u_2 \in U_2} x^T M \overline{B}_2 u_2 \\ &= \sup_{v \in M\overline{B}_1 U_1} x^T v - \sup_{v \in -M\overline{B}_2 U_2} x^T v \\ &= \mathcal{S}_{M\overline{B}_1 U_1}(x) - \mathcal{S}_{-M\overline{B}_2 U_2}(x) \end{aligned}$$

where $\mathcal{S}_{M\overline{B}_1U_1}$ and $\mathcal{S}_{-M\overline{B}_2U_2}$ denote the support functions of the sets $M\overline{B}_1U_1$ and $-M\overline{B}_2U_2$. Then, inequality (4.9) becomes

$$\forall x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2},\ \mathcal{S}_{M\overline{B}_1U_1}(x) \leq \mathcal{S}_{-M\overline{B}_2U_2}(x)$$

which is equivalent to say, since $M\overline{B}_1U_1$ and $-M\overline{B}_2U_2$ are compact convex sets, that $M\overline{B}_1U_1$ is a subset of $-M\overline{B}_2U_2$. Then, for all $u_1 \in U_1$, there exists $u_2 \in U_2$ such that $M\overline{B}_1u_1 = -M\overline{B}_2u_2$ which means that $\overline{B}_1u_1 + \overline{B}_2u_2 \in \ker(M)$ which implies that $\overline{B}_1U_1 + \subseteq \ker(M) + \overline{B}_2U_2$. Similarly, we can show from equation (4.4) that $-\overline{B}_2U_2 \subseteq \ker(M) + \overline{B}_1U_1$. Hence, equation (4.7) also holds. □

Quadratic bisimulation functions are particularly useful for autonomous systems (*i.e.* $B_1 = 0$, $B_2 = 0$). Indeed, in that case, equation (4.7) is always satisfied. Then, the characterization of a bisimulation function between T_{Δ_1} and T_{Δ_2} reduces to a set of linear matrix inequalities. Moreover, for stable autonomous linear systems, we can show that a quadratic bisimulation function always exists.

Proposition 4.3. *Let Δ_1 and Δ_2 be asymptotically stable autonomous linear systems, then there exists a bisimulation function of the form (4.1) between T_{Δ_1} and T_{Δ_2}.*

Proof. Linear matrix inequality (4.5) is equivalent to say that $M = C^TC + N$ where N is a positive semi-definite symmetric matrix. Then, linear matrix inequality (4.6) becomes

$$A^TN + NA \leq -(A^TC^TC + C^TCA). \tag{4.10}$$

Let us remark that $A^TC^TC + C^TCA$ is a symmetric matrix and then can be written as the difference between two positive semi-definite symmetric matrices Q^+ and Q^-: $A^TC^TC + C^TCA = Q^+ - Q^-$. Let us consider the Lyapunov equation

$$A^TN + NA = -Q^+.$$

Since Δ_1 and Δ_2 are asymptotically stable, there exists a unique solution N to this Lyapunov equation. This solution is positive semi-definite symmetric and clearly satisfies inequality (4.10). Therefore, for $M = C^TC + N$, $V(x_1, x_2) = \sqrt{x^TMx}$ is a bisimulation function between T_{Δ_1} and T_{Δ_2}. □

Corollary 4.4. *Let Δ_1 and Δ_2 be asymptotically stable autonomous linear systems, then T_{Δ_1} and T_{Δ_2} are approximately bisimilar.*

Proof. The proof is straightforward from the fact that the game given by equation (2.1) has obviously a finite value since I_1 and I_2 are compact sets. □

Considering quadratic bisimulation functions for linear systems with inputs is actually quite restrictive. Indeed, the value of quadratic functions at $x = 0$ is always 0. Particularly, this means that if Δ_1 and Δ_2 start from 0, the outputs of both systems will be identical. Equivalently, this means that Δ_1 and Δ_2 have identical asymptotic behaviors and that only their transient behaviors can differ. Therefore, we need to consider more general classes of functions so that bisimulation functions exist even if Δ_1 and Δ_2 do not have identical asymptotic behaviors.

4.2. **Truncated quadratic bisimulation functions.** A natural extension of quadratic bisimulation functions is the class of truncated quadratic bisimulation functions:

$$V(x_1,x_2)=\max(\sqrt{x^TMx},\alpha). \tag{4.11}$$

The choice of such a form is motivated by the following remark. Trajectories of stable constrained linear systems can be decomposed into two phases: the transient phase and the asymptotic phase. The initial states affect only the transient phase. Here, the term $\sqrt{x^TMx}$ in $V(x_1,x_2)$ can be interpreted as the error of approximation due to the transient phase. Then, the term α accounts for the error of approximation due to the asymptotic phase and is thus independent of the initial states of the systems.

Proposition 4.5. *If for all $x\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$,*

$$x^TMx\geq x^TC^TCx \tag{4.12}$$

and for all $x\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$, such that $x^TMx\geq\alpha^2$,

$$x^TMAx+x^TA^TMx+2\sup_{u_1\in U_1}\inf_{u_2\in U_2}x^TM(\overline{B}_1u_1+\overline{B}_2u_2)\leq 0, \tag{4.13}$$

and

$$x^TMAx+x^TA^TMx+2\sup_{u_2\in U_2}\inf_{u_1\in U_1}x^TM(\overline{B}_1u_1+\overline{B}_2u_2)\leq 0. \tag{4.14}$$

Then, $V(x_1,x_2)=\max(\sqrt{x^TMx},\alpha)$ is a bisimulation function between T_{Δ_1} and T_{Δ_2}.

Proof. Let $f(x)=x^TMx$, let $(x_1,x_2)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$. From equation (4.13) and Proposition 3.1 (with $\mathcal{H}=\mathbb{R}^{m_1}\times\mathbb{R}^{m_2}$ and $\eta=\alpha^2$), we have that for all $x_1\xrightarrow{t}_1 x_1'$, there exists $x_2\xrightarrow{t}_2 x_2'$ such that $f(x_1',x_2')\leq\max(f(x_1,x_2),\alpha^2)$ which implies that $V(x_1',x_2')\leq V(x_1,x_2)$. Similarly, from equation (4.14), we can show that for all $x_2\xrightarrow{t}_2 x_2'$, there exists $x_1\xrightarrow{t}_1 x_1'$ such that $V(x_1',x_2')\leq V(x_1,x_2)$. Equation (4.12) allows to conclude that $V(x_1,x_2)$ is a bisimulation function between T_{Δ_1} and T_{Δ_2}. □

A more effective characterization result can be given for truncated quadratic bisimulation functions.

Theorem 4.6. *If there exists $\lambda>0$, such that*

$$M\geq C^TC, \tag{4.15}$$

$$A^TM+MA+2\lambda M\leq 0, \tag{4.16}$$

$$\alpha\geq\frac{1}{\lambda}\sup_{x^TMx=1}\left(\sup_{u_1\in U_1}\inf_{u_2\in U_2}x^TM(\overline{B}_1u_1+\overline{B}_2u_2)\right), \tag{4.17}$$

$$\alpha\geq\frac{1}{\lambda}\sup_{x^TMx=1}\left(\sup_{u_2\in U_2}\inf_{u_1\in U_1}x^TM(\overline{B}_1u_1+\overline{B}_2u_2)\right). \tag{4.18}$$

Then, the function $V(x_1,x_2)=\max(\sqrt{x^TMx},\alpha)$ is a bisimulation function between T_{Δ_1} and T_{Δ_2}.

Proof. Equation (4.15) is equivalent to equation (4.12). Let $(x_1,x_2)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ such that $x^TMx\geq\alpha^2$. Then, equation (4.17) implies that

$$\sup_{u_1\in U_1}\inf_{u_2\in U_2}x^TM(\overline{B}_1u_1+\overline{B}_2u_2)\leq\lambda\alpha\sqrt{x^TMx}.$$

Therefore,

$$x^T A^T M x + x^T M A x + 2 \sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \leq x^T A^T M x + x^T M A x + 2\lambda\alpha\sqrt{x^T M x}.$$

Then, from equation (4.16)

$$\begin{aligned} x^T A^T M x + x^T M A x + 2 \sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) &\leq -2\lambda x^T M x + 2\lambda\alpha\sqrt{x^T M x} \\ &\leq -2\lambda\sqrt{x^T M x}(\sqrt{x^T M x} - \alpha) \leq 0. \end{aligned}$$

Thus, for all $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that $x^T M x \geq \alpha^2$, equation (4.13) holds. Similarly, from equations (4.16) and (4.18), we can show that equation (4.14) holds. Then, from Proposition 4.5, $V(x_1, x_2) = \max(\sqrt{x^T M x}, \alpha)$ is a bisimulation function between T_{Δ_1} and T_{Δ_2}. □

Example 4.7. Let us consider that following systems:

$$\Delta_1 : \begin{cases} \dot{x}_1(t) = -x_1(t) + u_1(t), & u_1(t) \in [-1, 1] \\ y_1(t) = x_1(t) \end{cases}$$

$$\Delta_2 : \begin{cases} \dot{x}_2(t) = -x_2(t) \\ y_2(t) = x_2(t) \end{cases}$$

Let us define

$$M = C^T C = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Equation (4.15) holds. We can check that $A^T M + MA = -2M$. Hence, equation (4.16) holds for $\lambda = 1$. Equation (4.17) becomes

$$\alpha \geq \sup_{(x_1 - x_2)^2 = 1} \left(\sup_{u_1 \in [-1,1]} (x_1 - x_2) u_1 \right) = 1.$$

Equation (4.18) becomes

$$\alpha \geq \sup_{(x_1 - x_2)^2 = 1} \left(\inf_{u_1 \in [-1,1]} (x_1 - x_2) u_1 \right) = -1.$$

From Theorem 4.6, $V(x_1, x_2) = \max(|x_1 - x_2|, 1)$ is a bisimulation function between T_{Δ_1} and T_{Δ_2}. Let us remark that this example illustrates an important difference between approximate bisimulation and model reduction techniques. Indeed, our approach allows to *abstract* (*i.e.* to ignore) the input of a system which does not make sense in the model reduction framework.

Note that our approach is consistent with our previous results. Indeed, if equation (4.7) holds, then we can choose $\alpha = 0$ and have a bisimulation function of the form $V(x_1, x_2) = \sqrt{x^T M x}$. The advantage of considering truncated quadratic simulation functions over purely quadratic simulation functions is that they are universal for the class of stable constrained linear systems.

Proposition 4.8. *Let Δ_1 and Δ_2 be asymptotically stable constrained linear systems, then there exists a bisimulation function of the form (4.11) between T_{Δ_1} and T_{Δ_2}.*

Proof. First, let us remark that equation (4.16) is equivalent to

$$(A + \lambda I)^T M + M(A + \lambda I) \leq 0.$$

Then, since all the real parts of the eigenvalues of A_1 and A_2 are strictly negative, it follows that for λ small enough, the real parts of the eigenvalues of $A + \lambda I$ are all strictly negative. Hence, similar to the proof of Proposition 4.3, we can show that there exists a matrix M such that equations (4.15) and (4.16) hold. Moreover,

$$\begin{aligned}
\sup_{x^TMx=1}\left(\sup_{u_1\in U_1}\inf_{u_2\in U_2} x^TM(\overline{B}_1u_1+\overline{B}_2u_2)\right) &\leq \sup_{x^TMx=1}\left(\sup_{u_1\in U_1}\sup_{u_2\in U_2} x^TM(\overline{B}_1U_1+\overline{B}_2U_2)\right)\\
&\leq \sup_{u_1\in U_1}\sup_{u_2\in U_2}\left(\sup_{x^TMx=1} x^TM(\overline{B}_1u_1+\overline{B}_2u_2)\right)\\
&\leq \sup_{u_1\in U_1}\sup_{u_2\in U_2}\sqrt{(\overline{B}_1u_1+\overline{B}_2u_2)^TM(\overline{B}_1u_1+\overline{B}_2u_2)}.
\end{aligned}$$

Since, U_1 and U_2 are compact sets, it is easy to see that there exists $\alpha \geq 0$ such that (4.17) holds. By a symmetric reasoning, we can show that there exists $\alpha \geq 0$ such that (4.18) also holds. □

Corollary 4.9. *Let Δ_1 and Δ_2 be asymptotically stable constrained linear systems, then T_{Δ_1} and T_{Δ_2} are approximately bisimilar.*

Proof. The proof is straightforward from the fact that the game given by equation (2.1) has obviously a finite value since I_1 and I_2 are compact sets. □

4.3. **Handling instability.** If Δ_1 and Δ_2 are not asymptotically stable, the results of the previous section cannot be used directly. Indeed, it is implicitly assumed that there exists a bisimulation function between T_{Δ_1} and T_{Δ_2} with finite values on $\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$. From Theorem 2.5, this implies that for any $(x_1,x_2)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$, for any trajectory of Δ_1 starting in x_1, there exists a trajectory of Δ_2 starting in x_2 and such that the distance between the observations of these trajectories remains bounded (and conversely). When dealing with unstable dynamics, it is not hard to see that this is generally not the case and that bisimulation functions with finite values on $\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ cannot exist. In the following, we search for simulation functions whose values are finite on a subspace of $\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$.

Let $E_{u,i}$ (respectively $E_{s,i}$) be the subspace of $\mathbb{R}^{n_i}$ spanned by the generalized eigenvectors of A_i associated to eigenvalues whose real part is positive (respectively strictly negative). Note that we have $E_{u,i}\oplus E_{s,i}=\mathbb{R}^{n_i}$. Let $P_{u,i}$ and $P_{s,i}$ denote the associated projections. $E_{u,i}$ and $E_{s,i}$ are invariant under A_i and are called the unstable and the stable subspaces of the system Δ_i. Using a change of coordinates, the matrices of system Δ_i can be transformed into the following form

$$A_i=\begin{bmatrix} A_{u,i} & 0\\ 0 & A_{s,i}\end{bmatrix}, B_i=\begin{bmatrix} B_{u,i}\\ B_{s,i}\end{bmatrix}, C_i=[C_{u,i}\ C_{s,i}]\,, \tag{4.19}$$

where all the eigenvalues of $A_{u,i}$ have a positive real part and all the eigenvalues of $A_{s,i}$ have a strictly negative real part. Let us define the unstable subsystems of Δ_1 and Δ_2

$$\Delta_{u,i}:\left\{\begin{array}{rcl}\dot{x}_{u,i}(t) &=& A_{u,i}x_{u,i}(t)+B_{u,i}u_i(t),\ x_{u,i}(t)\in E_{u,i},\ u_i(t)\in U_i,\ x_{u,i}(0)\in P_{u,i}I_i\\ y_{u,i}(t) &=& C_{u,i}x_{u,i}(t),\ y_{u,i}(t)\in\mathbb{R}^p\end{array}\right.$$

Let $T_{\Delta_{u,1}}$ and $T_{\Delta_{u,2}}$ be the associated metric transition systems. For $j\in\{u,s\}$, we define the matrices

$$A_j=\begin{bmatrix} A_{j,1} & 0\\ 0 & A_{j,2}\end{bmatrix},\ \overline{B}_{j,1}=\begin{bmatrix} B_{j,1}\\ 0\end{bmatrix},\ \overline{B}_{j,2}=\begin{bmatrix} 0\\ B_{j,2}\end{bmatrix},\ C_j=\begin{bmatrix} C_{j,1} & -C_{j,2}\end{bmatrix}. \tag{4.20}$$

and the projections defined by

$$P_j x = \begin{bmatrix} P_{j,1}x_1 \\ P_{j,2}x_2 \end{bmatrix}.$$

In the following, we show that if there exists subspace $\mathcal{R}_u \subseteq E_{u,1} \times E_{u,2}$ which is an exact bisimulation relation between $T_{\Delta_{u,1}}$ and $T_{\Delta_{u,2}}$, then we are able to compute a bisimulation function between T_{Δ_1} and T_{Δ_2}.

Lemma 4.10. *Let $\mathcal{R}_u \subseteq E_{u,1} \times E_{u,2}$ be a subspace satisfying:*

$$\mathcal{R}_u \subseteq \ker(C_u), \tag{4.21}$$

$$A_u\mathcal{R}_u \subseteq \mathcal{R}_u, \tag{4.22}$$

$$\mathcal{R}_u + \overline{B}_{u,1}U_1 = \mathcal{R}_u - \overline{B}_{u,2}U_2. \tag{4.23}$$

Then, $\mathcal{R}_u$ is an exact bisimulation relation between $T_{\Delta_{u,1}}$ and $T_{\Delta_{u,2}}$.

Proof. Let $x_u = (x_{u,1}, x_{u,2}) \in \mathcal{R}_u$, equation (4.21) implies that $C_{u,1}x_{u,1} = C_{u,2}x_{u,2}$. Let $x_{u,1} \xrightarrow{t}_{u,1} x'_{u,1}$, we note $u_1(.)$ the input that leads $\Delta_{u,1}$ from $x_{u,1}$ to $x'_{u,1}$. From equation (4.23), there exists $u_2(.)$ an input of $\Delta_{u,2}$ such that for all $s \in [0,t]$, $\overline{B}_{u,1}u_1(s) + \overline{B}_{u,2}u_2(s) \in \mathcal{R}_u$. The input $u_2(.)$ leads $\Delta_{u,2}$ from $x_{u,2}$ to $x'_{u,2}$. Let us remark that $x'_u = (x'_{u,1}, x'_{u,2})$ satisfies,

$$x'_u = e^{A_u t}x_u + \int_0^t e^{A_u(t-s)}(\overline{B}_{u,1}u_1(s) + \overline{B}_{u,2}u_2(s))\, ds.$$

From equation (4.22), it is then clear that $(x'_{u,1}, x'_{u,2}) \in \mathcal{R}_u$. Using symmetrical arguments, we can show that for all $x_{u,2} \xrightarrow{t}_{u,2} x'_{u,2}$, there exist $x_{u,1} \xrightarrow{t}_{u,1} x'_{u,1}$ such that $(x'_{u,1}, x'_{u,2}) \in \mathcal{R}_u$. Therefore, $\mathcal{R}_u$ is an exact bisimulation relation between $T_{\Delta_{u,1}}$ and $T_{\Delta_{u,2}}$. □

Remark 4.11. The characterization of an exact bisimulation relation given by Lemma 4.10 slightly differs from those that can be found in the literature [Pap03, vdS04]. This is due to the fact that the systems considered in these papers do not have constraints on the inputs.

Proposition 4.12. *Let $\mathcal{R}_u \subseteq E_{u,1} \times E_{u,2}$ be a subspace satisfying equations (4.21), (4.22) and (4.23). Let M_s be a positive semi-definite symmetric matrix and α_s a positive number such that for all for all $x_s \in E_{s,1} \times E_{s,2}$,*

$$x_s^T M_s x_s \geq x_s^T C_s^T C_s x_s \tag{4.24}$$

and for all $x_s \in E_{s,1} \times E_{s,2}$, such that $x_s^T M_s x_s \geq \alpha_s^2$,

$$x_s^T M_s A_s x_s + x_s^T A_s^T M_s x_s + 2 \sup_{u_1 \in U_1} \left(\inf_{u_2 \in U_2,\ \overline{B}_{u,1}u_1 + \overline{B}_{u,2}u_2 \in \mathcal{R}_u} x_s^T M_s(\overline{B}_{s,1}u_1 + \overline{B}_{s,2}u_2) \right) \leq 0, \tag{4.25}$$

and

$$x_s^T M_s A_s x_s + x_s^T A_s^T M_s x_s + 2 \sup_{u_2 \in U_2} \left(\inf_{u_1 \in U_1,\ \overline{B}_{u,1}u_1 + \overline{B}_{u,2}u_2 \in \mathcal{R}_u} x_s^T M_s(\overline{B}_{s,1}u_1 + \overline{B}_{s,2}u_2) \right) \leq 0. \tag{4.26}$$

Then, the function $V : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^+ \cup \{+\infty\}$ given by

$$V(x) = \begin{cases} +\infty, & \text{if } P_u x \notin \mathcal{R}_u \\ \max(\sqrt{x^T P_s^T M_s P_s x}, \alpha_s), & \text{if } P_u x \in \mathcal{R}_u \end{cases} \tag{4.27}$$

is a bisimulation function between T_{Δ_1} and T_{Δ_2}.

Proof. Let $(x_1,x_2)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$, if $P_u x\notin\mathcal{R}_u$ it is clear that $V(x)\geq\|Cx\|$. If $P_u x\in\mathcal{R}_u$, from equation (4.21) we have that $C_uP_ux=0$. Then, $\|Cx\|=\|C_uP_ux+C_sP_sx\|=\|C_sP_sx\|$. From equation (4.24), we have that $\|Cx\|\leq\sqrt{x^TP_s^TM_sP_sx}\leq V(x)$. Let $x_1\xrightarrow{t}_1 x_1'$, if $P_ux\notin\mathcal{R}_u$ then for any $x_2\xrightarrow{t}_2 x_2'$, $V(x_1',x_2')\leq+\infty=V(x_1,x_2)$. If $P_ux\in\mathcal{R}_u$, then let $f(x)=x^TP_s^TM_sP_sx$. From equation (4.25) and Proposition 3.1 (with $\mathcal{H}=\{(u_1,u_2)|\ \overline{B}_{u,1}u_1+\overline{B}_{u,2}u_2\in\mathcal{R}_u\}$, $\eta=\alpha_s^2$) we have that there exists $x_2\xrightarrow{t}_2 x_2'$ such that $f(x_1',x_2')\leq\max(f(x_1,x_2),\alpha_s^2)$. Moreover, there exist inputs $u_i(.)$ $(i=1,2)$ leading Δ_i from x_i to x_i' at time t and such that for all $s\in[0,t]$, $\overline{B}_{u,1}u_1(s)+\overline{B}_{u,2}u_2(s)\in\mathcal{R}_u$. Now let us remark that

$$P_ux'=e^{A_ut}P_ux+\int_0^t e^{A_u(t-s)}(\overline{B}_{u,1}u_1(s)+\overline{B}_{u,2}u_2(s))\,ds.$$

From equation (4.22), it follows that $P_ux'\in\mathcal{R}_u$. Hence $V(x_1',x_2')=\max(\sqrt{f(x_1',x_2')},\alpha_s)\leq\max(\sqrt{f(x_1,x_2)},\alpha_s)=V(x_1,x_2)$. By symmetry, we also have that for all $x_2\xrightarrow{t}_2 x_2'$, there exists $x_1\xrightarrow{t}_1 x_1'$ such that $V(x_1',x_2')\leq V(x_1,x_2)$. Therefore, V is a bisimulation function between T_{Δ_1} and T_{Δ_2}. □

The following Theorem gives another characterization of such bisimulation functions. The proof is similar to the one of Theorem 4.6 and is therefore not stated here.

Theorem 4.13. *Let $\mathcal{R}_u\subseteq E_{u,1}\times E_{u,2}$ be a subspace satisfying equations (4.21), (4.22) and (4.23). If there exists $\lambda_s>0$, such that*

$$M_s\geq C_s^TC_s, \tag{4.28}$$

$$A_s^TM_s+M_sA_s+2\lambda_sM_s\leq 0, \tag{4.29}$$

$$\alpha_s\geq\frac{1}{\lambda_s}\sup_{x_s^TM_sx_s=1}\left(\sup_{u_1\in U_1}\left(\inf_{u_2\in U_2,\ \overline{B}_{u,1}u_1+\overline{B}_{u,2}u_2\in\mathcal{S}_u} x_s^TM_s(\overline{B}_{s,1}u_1+\overline{B}_{s,2}u_2)\right)\right), \tag{4.30}$$

$$\alpha_s\geq\frac{1}{\lambda_s}\sup_{x_s^TM_sx_s=1}\left(\sup_{u_2\in U_2}\left(\inf_{u_1\in U_1,\ \overline{B}_{u,1}u_1+\overline{B}_{u,2}u_2\in\mathcal{S}_u} x_s^TM_s(\overline{B}_{s,1}u_1+\overline{B}_{s,2}u_2)\right)\right). \tag{4.31}$$

Then, the function $V(x_1,x_2)$ given by equation (4.27) is a bisimulation function between T_{Δ_1} and T_{Δ_2}.

We also have, similar to Proposition 4.8:

Proposition 4.14. *If there exists a subspace $\mathcal{R}_u$ satisfying equations (4.21), (4.22) and (4.23), then there exists a bisimulation function of the form (4.27) between T_{Δ_1} and T_{Δ_2}.*

Then, it follows that two systems with exactly bisimilar unstable subsystems are approximately bisimilar.

Corollary 4.15. *If there exists a subspace $\mathcal{R}_u$ satisfying equations (4.21), (4.22) and (4.23), and such that for all $x_{u,1}\in P_{u,1}I_1$ there exists $x_{u,2}\in P_{u,2}I_2$ satisfying $(x_{u,1},x_{u,2})\in\mathcal{R}_u$ and conversely (i.e. $T_{\Delta_{u,1}}$ and $T_{\Delta_{u,2}}$ are exactly bisimilar), then T_{Δ_1} and T_{Δ_2} are approximately bisimilar.*

Proof. Let $V(x_1,x_2)$ be a bisimulation function of the form (4.27) between T_{Δ_1} and T_{Δ_2}. For all $x_1 \in I_1$, there exists $x_2 \in I_2$ such that $P_u x \in \mathcal{R}_u$ then,

$$\sup_{x_1\in I_1}\inf_{x_2\in I_2} V(x_1,x_2) = \sup_{x_1\in I_1}\left(\inf_{x_2\in I_2,\, P_u x\in\mathcal{R}_u} \max(\sqrt{x^T P_s^T M_s P_s x}, \alpha_s)\right).$$

Since I_1 and I_2 are compact sets, this game has a finite value. Symmetrically, we also have that the value of

$$\sup_{x_2\in I_2}\inf_{x_1\in I_1} V(x_1,x_2) = \sup_{x_2\in I_2}\left(\inf_{x_1\in I_1,\, P_u x\in\mathcal{R}_u} \max(\sqrt{x^T P_s^T M_s P_s x}, \alpha_s)\right)$$

is finite and thus from Theorem 2.9, T_{Δ_1} and T_{Δ_2} are approximately bisimilar. □

5. Approximation of linear systems using approximate bisimulation

In this section, we use the previous results to compute the precision of the approximate bisimulation between a linear system with constrained inputs Δ_1 of the form (3.1) and a projection Δ_2. Let us assume, without loss of generality that the system Δ_1 has been decomposed into a stable and unstable subsystems and that the matrices A_1, B_1, C_1 are of the form given by equation (4.19). Given a surjective map $x_2 = Hx_1$, we define the projection of Δ_1 as the linear system with constrained inputs Δ_2 given by:

$$A_2 = HA_1H^+,\ B_2 = HB_1,\ C_2 = C_1H^+,\ U_2 = U_1 \text{ and } I_2 = HI_1 \tag{5.1}$$

where H^+ denotes the Moore-Penrose pseudoinverse of H. For simplicity, we will assume that the map H is of the form:

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

Then,

$$A_2 = \begin{bmatrix} H_uA_{u,1}H_u^+ & 0 \\ 0 & H_sA_{s,1}H_s^+ \end{bmatrix},\ B_2 = \begin{bmatrix} H_uB_{u,1} \\ H_sB_{s,1} \end{bmatrix} \text{ and } C_2 = \begin{bmatrix} C_{u,1}H_u^+ & C_{s,1}H_s^+ \end{bmatrix}.$$

Hence, the matrices A_2, B_2, C_2 are also of the form given by equation (4.19).

Lemma 5.1. *The subspace $\mathcal{R}_u \subseteq E_{u,1}\times E_{u,2}$ given by*

$$\mathcal{R}_u = \{(x_{u,1},x_{u,2})|\ x_{u,2} = H_u x_{u,1}\}$$

satisfies equations (4.21), (4.22) and (4.23) if and only if

$$C_{u,1} = C_{u,1}H_u^+H_u, \tag{5.2}$$

$$H_uA_{u,1} = H_uA_{u,1}H_u^+H_u. \tag{5.3}$$

In that case, $T_{\Delta_{u,1}}$ and $T_{\Delta_{u,2}}$ are exactly bisimilar.

Proof. First, let us remark that

$$C_{u,1} = C_{u,1}H_u^+H_u \iff C_{u,1} - C_{u,2}H_u = 0 \iff \mathcal{R}_u \subseteq \ker(C_u).$$

Secondly,

$$H_uA_{u,1} = H_uA_{u,1}H_u^+H_u \iff H_uA_{u,1} = A_{u,2}H_u \iff A_u\mathcal{R}_u \subseteq \mathcal{R}_u.$$

Finally, for all $u \in U_1$, $H_uB_{u,1}u = B_{u,2}u$. Since $U_1 = U_2$, equation (4.23) holds. From Lemma 4.10, $\mathcal{R}_u$ is an exact bisimulation relation between $T_{\Delta_{u,1}}$ and $T_{\Delta_{u,2}}$. From the specific form of H, we have for all $x_1 \in \mathbb{R}^{n_1}$, $H_uP_{u,1}x_1 = P_{u,2}Hx_1$. Then, for all $x_{u,1} \in P_{u,1}I_1$, $x_{u,1} = P_{u,1}x_1$ with $x_1 \in I_1$. Let $x_{u,2} = H_ux_{u,1} = H_uP_{u,1}x_1 = P_{u,2}Hx_1$, hence $x_{u,2} \in P_{u,2}I_2$ and $(x_{u,1},x_{u,2}) \in \mathcal{R}_u$.

Similarly, for all $x_{u,2} \in P_{u,2}I_2$, $x_{u,2} = P_{u,2}Hx_1$ with $x_1 \in I_1$. Let $x_{u,1} = P_{u,1}x_1$, then $x_{u,1} \in P_{u,1}I_1$ and $H_u x_{u,1} = H_u P_{u,1} x_1 = P_{u,2} H x_1 = x_{u,2}$ and hence $(x_{u,1}, x_{u,2}) \in \mathcal{R}_u$. Thus, $T_{\Delta_{u,1}}$ and $T_{\Delta_{u,2}}$ are exactly bisimilar. □

Let us assume that the map H_u is chosen such that equations (5.2) and (5.3) hold and that the map H_s is such that the eigenvalues of the matrix $H_s A_{s,1} H_s^+$ have all a strictly negative real part. Then, from Proposition 4.14, we know that there exists a bisimulation function between T_{Δ_1} and T_{Δ_2} of the form (4.27). Let A_s, $\overline{B}_{s,1}$, $\overline{B}_{s,2}$ and C_s be defined as in equation (4.20). There exist a matrix M_s and a real number $\lambda_s > 0$ satisfying equations (4.28) and (4.29). Let us define the matrix

$$Q_s = \left[\begin{array}{cc} I & H_s^T \end{array}\right] M_s \left[\begin{array}{c} I \\ H_s \end{array}\right].$$

Theorem 5.2. *Let α_s be defined by*

$$\alpha_s = \frac{1}{\lambda_s} \sup_{u_1 \in U_1} \sqrt{u_1^T B_{s,1}^T Q_s B_{s,1} u_1}. \tag{5.4}$$

Then, the function V defined by

$$V(x) = \begin{cases} +\infty, & \text{if } P_u x \notin \ker([H_u \quad -I]) \\ \max\left(\sqrt{x^T P_s^T M_s P_s x}, \alpha_s\right), & \text{if } P_u x \in \ker([H_u \quad -I]) \end{cases}$$

is a bisimulation function between T_{Δ_1} and T_{Δ_2}.

Proof. We assumed that H_u is such that $\ker([H_u \quad -I])$ satisfies equations (4.21), (4.22), (4.23). Furthermore, M_s and λ_s satisfy equations (4.28) and (4.29). Now, let us remark that

$$\begin{aligned}
\alpha_s &= \frac{1}{\lambda_s} \sup_{u_1 \in U_1} \sqrt{u_1^T (\overline{B}_{s,1} + \overline{B}_{s,2})^T M_s (\overline{B}_{s,1} + \overline{B}_{s,2}) u_1} \\
&= \frac{1}{\lambda_s} \sup_{x_s^T M_s x_s} \left(\sup_{u_1 \in U_1} x_s^T M_s (\overline{B}_{s,1} + \overline{B}_{s,2}) u_1 \right) \\
&\geq \frac{1}{\lambda_s} \sup_{x_s^T M_s x_s} \left(\sup_{u_1 \in U_1} \left(\inf_{u_2 \in U_2,\, \overline{B}_{s,1} u_1 + \overline{B}_{s,2} u_2 \in \ker([H_u \quad -I])} x_s^T M_s (\overline{B}_{s,1} u_1 + \overline{B}_{s,2} u_2) \right) \right).
\end{aligned}$$

Then, equation (4.30) holds. Since $U_1 = U_2$, equation (4.31) holds as well. Then, from Theorem 4.13, V is a bisimulation function between T_{Δ_1} and T_{Δ_2}. □

From Theorem 2.9, the precision of the approximate bisimulation between T_{Δ_1} and T_{Δ_2} can be evaluated by solving the game (2.1).

Theorem 5.3. *Let α_s be defined as in equation (5.4), let β_s be defined as*

$$\beta_s = \sup_{x_1 \in I_1} \sqrt{x_1^T P_{s,1}^T Q_s P_{s,1} x_1}. \tag{5.5}$$

Let $\delta = \max(\alpha_s, \beta_s)$. Then, T_{Δ_1} and T_{Δ_2} are approximately bisimilar with the precision δ.

Proof. Let us remark that

$$\beta_s = \sup_{x_1 \in I_1} \sqrt{\left[\begin{array}{cc} x_1^T P_{s,1}^T & x_1^T P_{s,1}^T H_s^T \end{array}\right] M_s \left[\begin{array}{c} P_{s,1} x_1 \\ H_s P_{s,1} x_1 \end{array}\right]}.$$

From the block diagonal structure of H we have that $P_{s,2}H = H_s P_{s,1}$. Hence,

$$\begin{aligned}\beta_s &= \sup_{x_1 \in I_1} \sqrt{\left[\begin{array}{cc} x_1^T & x_1^T H^T \end{array}\right] P_s^T M_s P_s \left[\begin{array}{c} x_1 \\ Hx_1 \end{array}\right]} \\ &= \sup_{x_1 \in I_1} \left(\inf_{x_2 \in I_2,\, x_2 = H_1} \sqrt{x^T P_s^T M_s P_s x} \right) \\ &\geq \sup_{x_1 \in I_1} \left(\inf_{x_2 \in I_2,\, P_u x \in \ker([H_u \quad -I])} \sqrt{x^T P_s^T M_s P_s x} \right).\end{aligned}$$

Similarly, we also have,

$$\begin{aligned}\beta_s &= \sup_{x_2 \in H I_1} \left(\inf_{x_1 \in I_1,\, x_2 = H_1} \sqrt{x^T P_s^T M_s P_s x} \right) \\ &\geq \sup_{x_2 \in I_2} \left(\inf_{x_1 \in I_1,\, P_u x \in \ker([H_u \quad -I])} \sqrt{x^T P_s^T M_s P_s x} \right).\end{aligned}$$

Hence, the value of the game (2.1) is bounded by $\max(\alpha_s, \beta_s)$ which implies, from Theorem 2.9, that T_{Δ_1} and T_{Δ_2} are approximately bisimilar with the precision δ. □

We presented a method to evaluate the precision of the approximate bisimulation between a constrained linear system and its projection. From the computational point of view, it requires to solve a set of linear matrix inequalities which can be done using semi-definite programming [Stu99]. Then, if we assume that I_1 and U_1 are polytopes, the precision of the approximate bisimulation between a constrained linear system and its projection can be computed by solving two linear quadratic programs given by equations (5.4) and (5.5). The method is summarized in the following algorithm:

Algorithm 5.4. *Let Δ_1 be a constrained linear system and Δ_2 its projection given by equation (5.1).*

(1) *Check that $C_{u,1} = C_{u,1} H_u^+ H_u$ and $H_u A_{u,1} = H_u A_{u,1} H_u^+ H_u$.*
(2) *Choose $\lambda_s > 0$ such that the eigenvalues of $A_s + \lambda_s I$ have a strictly negative real part. Then, solve the linear matrix inequalities:*

$$M_s \geq C_s^T C_s,$$
$$A_s^T M_s + M_s A_s + 2\lambda_s M_s \leq 0,$$

and set

$$Q_s = \left[\begin{array}{cc} I & H_s^T \end{array}\right] M_s \left[\begin{array}{c} I \\ H_s \end{array}\right].$$

(3) *Solve the linear quadratic program*

$$\gamma_1 = \max_{u_1 \in U_1} u_1^T B_{s,1}^T Q_s B_{s,1} u_1$$

and set $\alpha_s = \sqrt{\gamma_1}/\lambda_s$.
(4) *Solve the linear quadratic program*

$$\gamma_2 = \max_{x_1 \in I_1} x_1^T P_{s,1}^T Q_s P_{s,1} x_1$$

and set $\beta_s = \sqrt{\gamma_2}$.
(5) *Let $\delta = \max(\alpha_s, \beta_s)$.*

Then, T_{Δ_1} and T_{Δ_2} are approximately bisimilar with the precision δ.

An important parameter in this algorithm is the strictly positive scalar λ_s. On one hand, λ_s must be chosen small enough so that the eigenvalues of $A_s + \lambda_s I$ have a strictly negative real part. On the other hand, it appears experimentally that the larger λ_s, the better the evaluation of the precision of the approximate bisimulation between T_{Δ_1} and T_{Δ_2}.

The resolution of the linear matrix inequalities can be done using semi-definite programming [Stu99]. It should be noted that the smaller the matrix Q_s the smaller the precision δ. Hence, to get a tight evaluation of the precision of the approximate bisimulation between T_{Δ_1} and T_{Δ_2}, it is useful to add to the semi-definite program a linear objective function which can be, for instance, the trace of Q_s. For very large systems, the resolution of the semi-definite program can be costly. In such cases, the linear matrix inequalities can be solved by replacing them by a Lyapunov equation as in the proof of Proposition 4.3. The evaluation of the precision δ is not as tight, but the computations are much faster.

The remaining question is how do we choose the surjective map H so that the precision of the approximate bisimulation between T_{Δ_1} and its projection T_{Δ_2} of desired dimension is as small as possible. First, it is to be noted that the choice of H_u is quite restricted. Any bijective map is obviously an admissible choice for H_u. Using exact bisimulation reduction techniques [Pap03, TP03, vdS04], admissible surjective but non-bijective maps H_u can be chosen.

The choice of H_s is much less constrained and thus much more difficult. For instance, it can be chosen according to traditional model reduction techniques such as balanced truncation [ASG00]. It appears that in the context of approximate bisimulation these techniques have quite poor results. This is due to the fact that traditional model reduction techniques aim to approximate the input-output mapping associated to a linear system: the transient behavior is completely ignored (the initial state is assumed to be 0). We have seen that in the context of approximate bisimulation, the transient phase is as important as the asymptotic phase. Therefore, it is not surprising that model reduction techniques are not of great help for the choice of the map H_s. Then, H_s can be chosen using the following heuristic. Define H_s as the projection on the subspace of $E_{s,1}$ of desired dimension, invariant under $A_{s,1}$ and which is the most likely to minimize the optimal value of the optimization problems (5.4) and (5.5). Experimentally, it appears that, most of the time, this heuristic gives better result than model reduction techniques. However, it is clearly not optimal. Further research is definitely needed to design better methods to find a good surjective map H_s.

Our method has been implemented in a MATLAB toolbox available for download: MATISSE (Metrics for Approximate TransItion Systems Simulation and Equivalence [GJP05]). It uses several toolboxes such as the Multi-Parametric Toolbox [KGB04] for polytopes manipulation, the interface YALMIP [Löf04] to translate linear matrix inequalities into semi-definite programs which are solved by the toolbox SEDUMI [Stu99]. MATISSE allows to reduce a constrained linear system Δ_1 to a system Δ_2 of given dimension, and to compute the precision of the approximate bisimulation between T_{Δ_1} and T_{Δ_2}.

6. Examples

In this section, we show two examples of application of the toolbox MATISSE[1]. The first one deals with a middle-scale system (dimension ten). It is shown how MATISSE can be used in the context of safety verification to reduce the complexity of the problem. The second one deals with a large-scale system (about a hundred continuous variables).

[1]This examples are available as demo files in MATISSE.

6.1. **Middle-scale system.** Let us consider Δ_1, the ten dimensional system with a one dimensional input given by the following matrices:

$$A_1 = \begin{bmatrix} -0.1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -0.1 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.1 & -8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & -0.1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.1 & -0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.6 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -0.6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.1 \end{bmatrix}, \; B_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \; C_1^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The input is constrained in the interval $[-0.05,\ 0.05]$ and the initial value is constrained in the rectangle

$$I_1 = [9,\ 10] \times [0,\ 1] \times [-0.1,\ 0.1]^2 \times [-2,\ 1] \times [-0.1,\ 0.1]^5.$$

Δ_1 is asymptotically stable, thus it is already of the form (4.19). We compute approximations Δ_2 of dimension five and Δ_3 of dimension seven using the heuristic described in the previous section and implemented in ASILIS. Then, using Algorithm 5.4, we evaluate the precision of the approximate bisimulation between T_{Δ_1} and its approximations. If the linear matrix inequalities are solved using semi-definite programming, the evaluation of the precision of the approximate bisimulation between T_{Δ_1} and T_{Δ_2} is 2.016. The computation requires 5.48 seconds. If the linear matrix inequalities are solved using Lyapunov equations, the evaluation of the precision of the approximate bisimulation between T_{Δ_1} and T_{Δ_2} is 4.636 and the computation requires 0.12 seconds. Similarly, the evaluation of the precision of approximate bisimulation between T_{Δ_1} and T_{Δ_3} using semi-definite programming is 0.359 (computation time: 6.11 seconds). Using Lyapunov equations, it is 4.072 (computation time: 0.12 seconds). Thus, we can see that the method using semi-definite programming gives much better evaluations of the precision of the approximate bisimulation between T_{Δ_1} and its approximations than the method using Lyapunov equations. However, the latter requires much less computation and can therefore handle larger systems.

Reachability routines based on zonotope computation [Gir05] have been implemented in MATISSE. On figure 1, we represented the reachable set of the original ten dimensional system (left) and of its five dimensional and seven dimensional approximations (center and right). We also represented the unsafe set Π_U. For the approximations, this set is bloated by the precision of the approximate bisimulation between T_{Δ_1} and its approximations (evaluated using semi-definite programming). It follows that if an approximation is safe then the original system is safe. We can see that the approximation of dimension 7 allows to conclude that the original system is safe, whereas the approximation of dimension 5 does not.

The example also illustrates the important point that robustness simplifies verification. Indeed, if the distance between the reachable set of the original system and the set of unsafe states would have been larger, then the approximation of the original system by its five dimensional approximation T_{Δ_2} might have been sufficient to check the safety. Further, we might have been able to conclude that the system is safe using the precision of the approximate bisimulation between T_{Δ_1} and T_{Δ_2} evaluated using Lyapunov equations. Generally, the more robustly safe a system is, the larger the distance from the unsafe safe, resulting in larger model compression and easier safety verification.

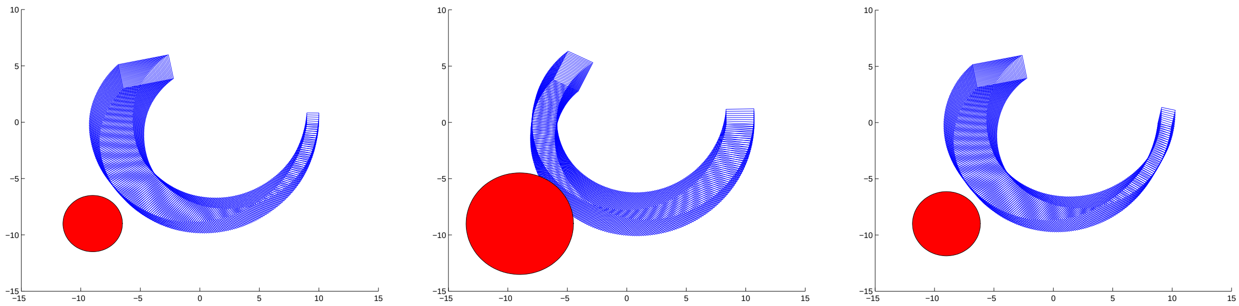

Figure 1. Reachable sets of the original ten dimensional system (left) and of its five dimensional and seven dimensional approximations (center and right). The disk on the left figure represent the unsafe set Π_U. The disk on the center and right figure consists of the set of points whose distance to Π_U is smaller than the precision of the approximate bisimulation between T_{Δ_1} and its approximation.

6.2. **Large-scale system.** Let us now consider the following problem [CD02]. We consider the heat diffusion equation on a rod:

$$\begin{cases} \frac{\partial}{\partial t} T(x,t) = \alpha \frac{\partial^2}{\partial x^2} T(x,t) + u(x,t), & x \in (0,1),\ t > 0, \\ T(0,t) = 0 = T(1,t), & t > 0, \\ T(x,0) = 0, & x \in (0,1) \end{cases}$$

where $T(x,t)$ represents the temperature field on the rod. We assume that the heat source is of the form $u(x,t) = \delta_{1/3}(x)u(t)$ where $u(t) \in [1, 1.1]$. The system is observed through the temperature at the point 2/3: $y(t) = T(2/3, t)$. The partial differential equation is discretized in space (101 nodes). This 101 dimensional linear system with a one-dimensional input is our original system Δ_1. We compute approximations Δ_2 of dimension ten and Δ_3 of dimension twenty. The evaluation of the precision of the approximate bisimulation between T_{Δ_1} and its approximations is done using Lyapunov equations. It requires respectively 1.81 and 1.92 seconds. T_{Δ_1} and T_{Δ_2} are approximately bisimilar with the precision 1.27 whereas T_{Δ_1} and T_{Δ_3} are approximately bisimilar with the precision 0.32. On figure 2, we represented the evolution of the reachable sets of T_{Δ_1} and T_{Δ_2} against time. It is clear that the distance between the reachable sets is actually much smaller than the precision of the approximate simulation between T_{Δ_1} and T_{Δ_2} given by Algorithm 5.4. This is due to the use of Lyapunov equations to solve the linear matrix inequalities which gives a large evaluation of the precision. However, in the context of safety verification, if T_{Δ_1} is robustly safe then this evaluation of the precision might well be sufficient to conclude that T_{Δ_1} is safe by performing the reachability analysis on T_{Δ_2}.

7. Conclusion

In this paper, we applied the framework of system approximation based on approximate versions of bisimulation relations to a class of constrained linear systems. We presented a class of functions which provide universal bisimulation functions for such systems. An important consequence, is that any two systems with exactly bisimilar unstable subsystems are approximately bisimilar. We gave effective characterizations for this class of bisimulation functions allowing us to develop an efficient algorithm to compute the precision of the approximate bisimulation between a system and its projection. This

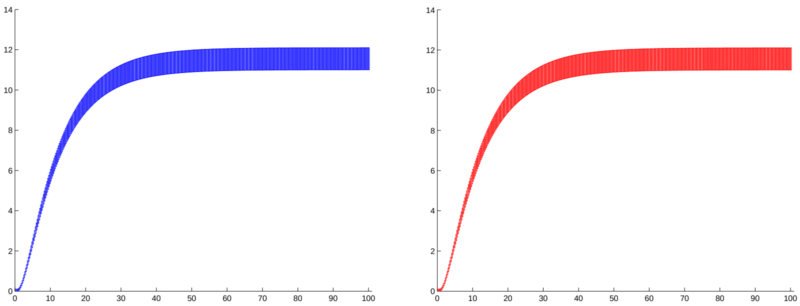

FIGURE 2. Evolution of the reachable sets of the original 101-dimensional system (left) and of its 10-dimensional approximation (right).

algorithm only requires the resolution of a set of linear matrix inequalities and of two linear quadratic programs and is therefore computationally effective.

This algorithm has been implemented within a MATLAB toolbox, MATISSE [GJP05]. MATISSE allows to reduce a constrained linear system to a system of given dimension and to compute the precision of the approximate bisimulation between the original system and its approximation. Two examples a application of MATISSE were showed. Particularly, we saw that, coupled to reachable set computation methods, it can be used to solve more efficiently the safety verification problem of linear systems.

Future research includes extending the results for linear systems to stochastic linear systems. We also aim to develop such computational techniques for nonlinear and hybrid systems.

References

[ABDM00] E. Asarin, O. Bournez, T. Dang, and O. Maler. Approximate reachability analysis of piecewise linear dynamical systems. In *Hybrid Systems: Computation and Control*, volume 1790 of *LNCS*, pages 21–31. Springer, 2000.

[AD04] E. Asarin and T. Dang. Abstraction by projection and application to multi-affine systems. In *Hybrid Systems: Computation and Control*, volume 2993 of *LNCS*, pages 32–47. Springer, 2004.

[ADG03] E. Asarin, T. Dang, and A. Girard. Reachability of non-linear systems using conservative approximations. In *Hybrid Systems: Computation and Control*, volume 2623 of *LNCS*, pages 22–35. Springer, 2003.

[ADI02] R. Alur, T. Dang, and F. Ivancic. Reachability analysis of hybrid systems via predicate abstraction. In *Hybrid Systems: Computation and Control*, volume 2289 of *LNCS*, pages 35–48. Springer, 2002.

[AHP00] R. Alur, T.A. Henzinger, and G. Lafferriere G.J. Pappas. Discrete abstractions of hybrid systems. *Proceedings of the IEEE*, 88(7):971–984, July 2000.

[ASG00] A. C. Antoulas, D. C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. *Contemporary Mathematics*, 280:193–219, 2000.

[Aub01] J. P. Aubin. Viability kernels and capture basins of sets under differential inclusion. *SIAM J. Control and Optimization*, 40(3):853–881, 2001.

[BCM+90] J.R. Burch, E.M. Clarke, K.L. McMillan, D.L. Dill, and L.J. Hwang. Symbolic Model Checking: 10^{20} States and Beyond. In *Proceedings of the Fifth Annual IEEE Symposium on Logic in Computer Science*, pages 1–33, Washington, D.C., 1990.

[BDG96] C. L. Beck, J. Doyle, and K. Glover. Model reduction of multi-dimensional and uncertain systems. *IEEE Transactions on Automatic Control*, 41(10):1466–1477, October 1996.

[CD02] Y. Chahlaoui and . Van Dooren. A collection of benchmark examples for model reduction of linear time invariant dynamical systems, 2002. SLICOT Working Note 2002-2.

[CGP00] E. M. Clarke, O. Grumberg, and D. A. Peled. *Model Checking*. MIT Press, 2000.

[CK99] A. Chutinan and B.H. Krogh. Verification of polyhedral invariant hybrid automata using polygonal flow pipe approximations. In *Hybrid Systems: Computation and Control*, volume 1569 of *LNCS*, pages 76–90. Springer, 1999.

[dAFS04] L. de Alfaro, M. Faella, and M. Stoelinga. Linear and branching metrics for quantitative transition systems. In *ICALP'04*, volume 3142 of *LNCS*, pages 1150–1162. Springer, 2004.

[DGJP04] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled markov processes. *Theoretical Computer Science*, 318(3):323–354, June 2004.

[FK96] R. A. Freeman and P. V. Kokotovic. Inverse optimality in robust stabilization. *SIAM J. Control and Optimization*, 34(4):1365–1391, July 1996.

[Gir05] A. Girard. Reachability of uncertain linear systems using zonotopes. In *Hybrid Systems: Computation and Control*, volume 3414 of *Lecture Notes in Computer Science*, pages 291–305. Springer, 2005.

[GJP05] A. Girard, A. A. Julius, and G. J. Pappas. MATISSE: Metrics for Approximate TransItion Systems Simulation and Equivalence, 2005. Available from `http://www.seas.upenn.edu/~agirard/Software/MATISSE/` .

[GP05a] A. Girard and G. J. Pappas. Approximate bisimulations for constrained linear systems. In *Proc. IEEE Conference on Decision and Control and European Control Conference*, Seville, Spain, December 2005. To appear.

[GP05b] A. Girard and G. J. Pappas. Approximate bisimulations for nonlinear dynamical systems. In *Proc. IEEE Conference on Decision and Control and European Control Conference*, Seville, Spain, December 2005. To appear.

[GP05c] A. Girard and G. J. Pappas. Approximation metrics for discrete and continuous systems. Technical Report MS-CIS-05-10, Dept. of CIS, University of Pennsylvania, May 2005.

[HK04] Z. Han and B. H. Krogh. Reachability of hybrid control systems using reduced-order models. In *Proceedings of the American Control Conference*, Boston, MA, July 2004.

[HTP05] E. Haghverdi, P. Tabuada, and G. J. Pappas. Bisimulation relations for dynamical, control, and hybrid systems. *Theoretical Computer Science*, 342(2-3):229–261, 2005.

[JvdS04] A.A. Julius and A.J. van der Schaft. A behavioral framework for compositionality: linear systems, discrete event systems and hybrid systems. In *Proceedings of the Sixteenth International Symposium on Mathematical Theory of Networks and Systems*, Leuven, Belgium, July 2004.

[KGB04] M. Kvasnica, P. Grieder, and M. Baotić. Multi-Parametric Toolbox (MPT), 2004. Available from `http://control.ee.ethz.ch/~mpt/`.

[KV00] A. Kurzhanski and P. Varaiya. Ellipsoidal tehcniques for reachability analysis. In *Hybrid Systems: Computation and Control*, volume 1790 of *LNCS*. Springer, 2000.

[Löf04] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. Available from `http://control.ee.ethz.ch/~joloef/yalmip.php`.

[LSW02] D. Liberzon, E. D. Sontag, and Y. Wang. Universal construction of feedback laws achieving ISS and integral-ISS disturbance attenuation. *Systems and Control Letters*, 46:111–127, 2002.

[MT00] I. Mitchell and C. Tomlin. Level set methods for computation in hybrid systems. In *Hybrid Systems: Computation and Control*, volume 1790 of *LNCS*. Springer, 2000.

[Pap03] G. J. Pappas. Bisimilar linear systems. *Automatica*, 39(12):2035–2047, December 2003.

[PJ04] S. Prajna and A. Jadbabaie. Safety verification of hybrid systems using barrier certificates. In *Hybrid Systems: Computation and Control*, volume 2993 of *Lecture Notes in Computer Science*, pages 477 – 492. Springer, 2004.

[PvdSB04] G. Pola, A.J. van der Schaft, and M. D. Di Benedetto. Bisimulation theory for switching linear systems. In *Proceedings of the 43rd IEEE Conference on Decision and Control*, pages 555–569, Nassau, Bahamas, December 2004.

[Stu99] J. F. Sturm. Using SEDUMI 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Softwares*, 11-12:625–653, 1999.

[TK02] A. Tiwari and G. Khanna. Series of abstractions for hybrid automata. In *Hybrid Systems: Computation and Control*, volume 2289 of *LNCS*, pages 465–478. Springer, 2002.

[TP03] H. Tanner and G. J. Pappas. Abstractions of constrained linear systems. In *Proc. American Control Conference*, Denver, CO, June 2003.

[TP04] P. Tabuada and G. J. Pappas. Bisimilar control affine systems. *Systems and Control Letters*, 52:49–58, 2004.

[vBMOW03] F. van Breugel, M. Mislove, J. Ouaknine, and J. Worrell. An intrinsic characterization of approximate probabilistic bisimilarity. In *Foundations of Software Science and Computation Structures*, volume 2620 of *LNCS*, pages 200–215. Springer, 2003.

[vdS04] A. van der Schaft. Equivalence of dynamical systems by bisimulation. *IEEE Transactions on Automatic Control*, 49(12):2160–2172, December 2004.

Appendix

Proof of Proposition 3.1. The proof of Proposition 3.1 requires several preliminary results.

Lemma 7.1. *Let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, let $t > 0$, then for all $\varepsilon > 0$, there exists $h > 0$ such that for all inputs $u_1(.)$ and $u_2(.)$ of Δ_1 and Δ_2, and associated trajectories*

$$\forall s \in [0,t],\ z_i(s) = e^{A_i s}x_i + \int_0^s e^{A_i(s-\tau)}B_i u_i(\tau)d\tau,\ i = 1,2$$

we have for all $u_1 \in U_1$, $u_2 \in U_2$, $s, s' \in [0,t]$,

$$s \le s' \le s+h \implies |\nabla f(z(s))(Az(s) + \overline{B}_1u_1 + \overline{B}_2u_2) - \nabla f(z(s'))(Az(s') + \overline{B}_1u_1 + \overline{B}_2u_2)| \le \varepsilon/t$$

where $z(s) = (z_1(s), z_2(s))$.

Proof. First let us remark that for all inputs $u_1(.)$ and $u_2(.)$ of Δ_1 and Δ_2, the associated trajectories are bounded on $[0,t]$:

$$\forall s \in [0,t],\ \|z_i(s)\| \le e^{\|A_i\|t}\|x_i\| + \int_0^t e^{\|A_i\|(t-\tau)}\|B_i\|d\tau \sup_{u_i \in U_i} \|u_i\| = m_i,\ i = 1,2. \tag{7.1}$$

Note that $\mathcal{C}_1 = \{z_1 \in \mathbb{R}^{n_1} | \|z_1\| \le m_1\}$ and $\mathcal{C}_2 = \{z_2 \in \mathbb{R}^{n_2} | \|z_2\| \le m_2\}$ are compact sets. Then, since $\nabla f(z)(Az + \overline{B}_1u_1 + \overline{B}_2u_2)$ is continuous on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ it is absolutely continuous on $\mathcal{C}_1 \times \mathcal{C}_2 \times U_1 \times U_2$. Particularly, for all $\varepsilon > 0$, there exists $\xi > 0$ such that for all $u_1 \in U_1$, $u_2 \in U_2$, $z_1, z_1' \in \mathcal{C}_1$, $z_2, z_2' \in \mathcal{C}_2$:

$$\|z_1 - z_1'\| \le \xi \text{ and } \|z_2 - z_2'\| \le \xi \implies |\nabla f(z)(Az + \overline{B}_1u_1 + \overline{B}_2u_2) - \nabla f(z')(Az' + \overline{B}_1u_1 + \overline{B}_2u_2)| \le \varepsilon/t \tag{7.2}$$

where $z = (z_1, z_2)$, $z' = (z_1', z_2')$.

Now, let us remark that for all inputs $u_1(.)$ and $u_2(.)$ of Δ_1 and Δ_2, the associated trajectories satisfy for all $s, s' \in [0,t]$, with $s \le s'$,

$$\begin{aligned} \|z_i(s') - z_i(s)\| &\le e^{\|A_i\|s}(e^{\|A_i\|(s'-s)} - 1)\|x_i\| + \int_s^{s'} e^{\|A_i\|(s'-\tau)}\|B_i\|d\tau \sup_{u_i \in U_i}\|u_i\| \\ &\le (e^{\|A_i\|(s'-s)} - 1)\left(e^{\|A_i\|t}\|x_i\| + \frac{\|B_i\|}{\|A_i\|}\sup_{u_i \in U_i}\|u_i\|\right),\ i = 1,2. \end{aligned} \tag{7.3}$$

Therefore, there exists $h > 0$, such that for all inputs $u_1(.)$ and $u_2(.)$ of Δ_1 and Δ_2, the associated trajectories $z_1(.)$ and $z_2(.)$ satisfy for all $s, s' \in [0,t]$

$$s \le s' \le s+h \implies \|z_1(s) - z_1(s')\| \le \xi \text{ and } \|z_2(s) - z_2'(s)\| \le \xi. \tag{7.4}$$

Moreover from equation (7.1), for all $s, s' \in [0,t]$, we have $z_1(s), z_1(s') \in \mathcal{C}_1$, $z_2(s), z_2(s') \in \mathcal{C}_2$. Then, equations (7.2) and (7.4) allow to conclude. □

Lemma 7.2. *Let f be a function and $\mathcal{H}$ a subspace satisfying assumptions of Proposition 3.1. Then, for all (x_1, x_2) satisfying $f(x_1, x_2) \geq \eta$, for all $x_1 \xrightarrow{t}_1 x_1'$, for all $\varepsilon > 0$, there exists $x_2 \xrightarrow{t}_2 x_2'$ such that*

$$f(x_1', x_2') \leq f(x_1, x_2) + \varepsilon. \tag{7.5}$$

Moreover, there exist inputs $u_i(.)$ $(i = 1, 2)$ leading Δ_i from x_i to x_i' at time t and such that for all $s \in [0, t]$, $(u_1(s), u_2(s)) \in \mathcal{H}$.

Proof. Let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that $f(x_1, x_2) \geq \eta$, let $x_1 \xrightarrow{t}_1 x_1'$, let $u_1(.)$ be an input which leads Δ_1 from x_1 to x_1' and $z_1(.)$ the associated trajectory of Δ_1. Let $\varepsilon > 0$, let $h > 0$ be given as in Lemma 7.1 (we assume without loss of generality that $t/h = N \in \mathbb{N}$). From equation (3.2), there exists an input $u_2(.)$ for Δ_2 such that

$$\forall s \in [0, h],\ (u_1(s), u_2(s)) \in \mathcal{H} \text{ and } \nabla f(x)(Ax + \overline{B}_1 u_1(s) + \overline{B}_2 u_2(s)) \leq 0.$$

Let $z_2(.)$ be the associated trajectory of Δ_2, then

$$f(z(h)) - f(x) = \int_0^h \nabla f(z(s))(Az(s) + \overline{B}_1 u_1(s) + \overline{B}_2 u_2(s))ds.$$

Then, from Lemma 7.1,

$$f(z(h)) - f(x) \leq \int_0^h \nabla f(z(0))(Az(0) + \overline{B}_1 u_1(s) + \overline{B}_2 u_2(s)) + \varepsilon/t \, ds \leq \frac{h\varepsilon}{t}.$$

Now let us assume that for some $i \in \{1, \ldots, N-1\}$ there exists an input $u_2(.)$ for Δ_2 such that

$$\forall s \in [0, ih],\ (u_1(s), u_2(s)) \in \mathcal{H} \text{ and } f(z(ih)) - f(x) \leq \frac{ih\varepsilon}{t}. \tag{7.6}$$

We showed that this is true for $i = 1$.

If $f(z(ih)) \geq \eta$, then, according to equation (3.2), we can choose $u_2(.)$ on $[ih, (i+1)h]$ such that

$$\forall s \in [ih, (i+1)h],\ (u_1(s), u_2(s)) \in \mathcal{H} \text{ and } \nabla f(z(ih))(Az(ih) + \overline{B}_1 u_1(s) + \overline{B}_2 u_2(s)) \leq 0.$$

Then from Lemma 7.1,

$$f(z((i+1)h)) - f(z(ih)) \leq \int_{ih}^{(i+1)h} \nabla f(z(ih))(Az(ih) + \overline{B}_1 u_1(s) + \overline{B}_2 u_2(s)) + \varepsilon/t \, ds \leq \frac{h\varepsilon}{t}.$$

Hence,

$$\forall s \in [0, (i+1)h],\ (u_1(s), u_2(s)) \in \mathcal{H} \text{ and } f(z((i+1)h)) - f(x) \leq \frac{(i+1)h\varepsilon}{t}.$$

Let us assume that $f(z(ih)) < \eta$. Let $v_2(.)$ be an input of Δ_2 such that

$$\forall s \in [ih, (i+1)h],\ (u_1(s), v_2(s)) \in \mathcal{H}.$$

Let $w_2(.)$ be the solution of the differential equation

$$\forall s \in [ih, (i+1)h],\ \dot{w}_2(s) = A_2 w_2(s) + B_2 v_2(s),\ w_2(ih) = z_2(ih).$$

If $f(z_1((i+1)h), w_2((i+1)h)) \leq \eta$, then we choose for all $s \in [ih, (i+1)h]$, $u_2(s) = v_2(s)$ and therefore

$$f(z((i+1)h)) - f(x) \leq \eta - f(x) \leq 0 \leq \frac{(i+1)h\varepsilon}{t}.$$

If $f(z_1((i+1)h), w_2((i+1)h)) > \eta$, there exists $s^* \in (ih, (i+1)h)$, such that $f(z_1(s^*), w_2(s^*)) = \eta$. Let $z^* = (z_1(s^*), w_2(s^*))$. Then, according to equation (3.2), we can choose $u_2(.)$ such that

$$\begin{array}{ll} \forall s \in [ih, s^*), & u_2(s) = v_2(s), \\ \forall s \in [s^*, (i+1)h], & (u_1(s), u_2(s)) \in \mathcal{H} \text{ and } \nabla f(z^*)(Az^* + \overline{B}_1 u_1(s) + \overline{B}_2 u_2(s)) \leq 0. \end{array}$$

Then, from Lemma 7.1,

$$f(z((i+1)h)) - f(z(s^*)) \leq \int_{s^*}^{(i+1)h} \nabla f(z^*)(Az^* + \overline{B}_1 u_1(s) + \overline{B}_2 u_2(s)) + \varepsilon/t \, ds \leq \frac{h\varepsilon}{t}.$$

Hence, for all $s \in [0, (i+1)h]$, $(u_1(s), u_2(s)) \in \mathcal{H}$ and

$$f(z((i+1)h)) - f(x) \leq f(z((i+1)h)) - \eta \leq \frac{h\varepsilon}{t} \leq \frac{(i+1)h\varepsilon}{t}.$$

Then equation (7.6) holds for all $i \in \{1, \dots N\}$ and particularly (for $i = N$) there exists an input $u_2(.)$ for Δ_2 such that

$$\forall s \in [0, t],\ (u_1(s), u_2(s)) \in \mathcal{H} \text{ and } f(z_1(t), z_2(t)) - f(x_1, x_2) \leq \varepsilon.$$

□

Lemma 7.3. *Let $x_1 \xrightarrow{t}_1 x_1'$, we define*

$$Post_2^{\mathcal{H}}(x_2, x_1 \xrightarrow{t}_1 x_1') = \Big\{ x_2' |\ x_2 \xrightarrow{t}_2 x_2' \text{ and for all } s \in [0, t],\ (u_1(s), u_2(s)) \in \mathcal{H} \Big\}$$

where $u_i(.)$ is an input which leads Δ_i from x_i to x_i' at time t (i=1,2). Then, $Post_2^{\mathcal{H}}(x_2, x_1 \xrightarrow{t}_1 x_1')$ is a compact set.

Proof. Let us define the set

$$\mathrm{Post}^{t,\mathcal{H}}(x_1, x_2) = \Big\{ (x_1', x_2') |\ x_1 \xrightarrow{t}_1 x_1',\ x_2 \xrightarrow{t}_2 x_2' \text{ and for all } s \in [0, t],\ (u_1(s), u_2(s)) \in \mathcal{H} \Big\}$$

Let us remark that $\mathrm{Post}^{t,\mathcal{H}}(x_1, x_2)$ is the set of reachable points at time t of a linear system whose input $u(.) = (u_1(.), u_2(.))$ is constrained in the compact convex set $\mathcal{H} \cap (U_1 \times U_2)$. Hence, it can be shown (see *e.g.* [Aub01]) that $\mathrm{Post}^{t,\mathcal{H}}(x_1, x_2)$ is a compact set. Let $x_1 \xrightarrow{t}_1 x_1'$, then we have

$$\mathrm{Post}_2^{\mathcal{H}}(x_2, x_1 \xrightarrow{t}_1 x_1') = \mathrm{Post}^{t,\mathcal{H}}(x_1, x_2) \cap (\{x_1'\} \times \mathbb{R}^{n_2}).$$

Hence, it is clear that $\mathrm{Post}_2^{\mathcal{H}}(x_2, x_1 \xrightarrow{t}_1 x_1')$ is a compact set. □

We can now prove Proposition 3.1. Let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, let $x_1 \xrightarrow{t}_1 x_1'$. If $f(x_1, x_2) \geq \eta$, then from Lemma 7.2,

$$\text{for all } \varepsilon > 0, \text{ there exists } x_2' \in \mathrm{Post}_2^{\mathcal{H}}(x_2, x_1 \xrightarrow{t}_1 x_1') \text{ such that } f(x_1', x_2') \leq f(x_1, x_2) + \varepsilon.$$

From Lemma 7.3, $\mathrm{Post}_2^{\mathcal{H}}(x_2, x_1 \xrightarrow{t}_1 x_1')$ is a compact subset, moreover f is a positive C^1 function and therefore it has closed level sets. Then, from Lemma 2.7, it is clear that

$$\text{there exists } x_2' \in \mathrm{Post}_2^{\mathcal{H}}(x_2, x_1 \xrightarrow{t}_1 x_1') \text{ such that } f(x_1', x_2') \leq f(x_1, x_2).$$

Hence, there exists $x_2 \xrightarrow{t}_2 x_2'$ such that $f(x_1', x_2') \leq f(x_1, x_2)$ and there exist inputs $u_i(.)$ $(i = 1, 2)$ leading Δ_i from x_i to x_i' at time t and such that for all $s \in [0, t]$, $(u_1(s), u_2(s)) \in \mathcal{H}$.

If $f(x_1,x_2)<\eta$, let $v_1(.)$ be an input which leads Δ_1 from x_1 to x_1' at time t, let $z_1(.)$ the associated trajectory of Δ_1. Let $v_2(.)$ be an input of Δ_2 such that for all $s\in[0,t]$, $(v_1(s),v_2(s))\in\mathcal{H}$ and $z_2(.)$ the associated trajectory of Δ_2 starting from x_2.

If $f(x_1',z_2(t))\leq\eta$, then we can choose $x_2\stackrel{t}{\longrightarrow}_2 x_2'$ with $x_2'=z_2(t)$. If $f(x_1',z_2(t))>\eta$, then there exists s^* in $(0,t)$ such that $f(z_1(s^*),z_2(s^*))=\eta$. Note that $z_1(s^*)\stackrel{t-s^*}{\longrightarrow}_1 x_1'$. Since $f(z_1(s^*),z_2(s^*))=\eta$, we know that there exists $z_2(s^*)\stackrel{t-s^*}{\longrightarrow}_2 x_2'$ such that $f(x_1',x_2')\leq f(z_1(s^*),z_2(s^*))$. Moreover, there exist inputs $v_i'(.)$ leading Δ_i from $z_i(s^*)$ to x_i' $(i=1,2)$ and such that for all $s\in[s^*,t]$, $(v_1'(s),v_2'(s))\in\mathcal{H}$. Then, for $i=1,2$, let the input $u_i(.)$ be defined by

$$\forall s\in[0,s^*],\ u_i(s)=v_i(s) \text{ and } \forall s\in[s^*,t],\ u_i(s)=v_i'(s).$$

Then, $u_i(.)$ leads system Δ_i from x_i to x_i' at time t and for all $s\in[0,t]$, $(u_1(s),u_2(s))\in\mathcal{H}$ and $f(x_1',x_2')\leq\eta$.

Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104

E-mail address: agirard@seas.upenn.edu

Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104

E-mail address: pappasg@ee.upenn.edu