August 2007

Approximate Bisimulation Relations for Constrained Linear Systems

Antoine Girard
Université Joseph Fourier

George J. Pappas
University of Pennsylvania, pappasg@seas.upenn.edu

Postprint version. Published in *Automatica*, Volume 43, Issue 8, August 2007, 1307-1317.
Publisher URL: http://dx.doi.org/10.1016/j.automatica.2007.01.019

Approximate Bisimulation Relations for Constrained Linear Systems

Abstract

In this paper, we define the notion of approximate bisimulation relation between two continuous systems. While exact bisimulation requires that the observations of two systems are and remain identical, approximate bisimulation allows the observations to be different provided the distance between them remains bounded by some parameter called precision. Approximate bisimulation relations are conveniently defined as level sets of a so-called bisimulation function which can be characterized using Lyapunov-like differential inequalities. For a class of constrained linear systems, we develop computationally effective characterizations of bisimulation functions that can be interpreted in terms of linear matrix inequalities and optimal values of static games. We derive a method to evaluate the precision of the approximate bisimulation relation between a constrained linear system and its projection. This method has been implemented in a Matlab toolbox: MATISSE. An example of use of the toolbox in the context of safety verification is shown.

Keywords

abstractions, approximation, bisimulation, lyapunov techniques, safety

Comments

Postprint version. Published in *Automatica*, Volume 43, Issue 8, August 2007, 1307-1317.
Publisher URL: http://dx.doi.org/10.1016/j.automatica.2007.01.019

Approximate Bisimulation Relations for Constrained Linear Systems ⋆

Antoine Girard [a], George J. Pappas [b],

[a] *Université Joseph Fourier, Laboratoire de Modélisation et Calcul, B.P. 53, 38041 Grenoble Cedex 9, France*

[b] *Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA*

Abstract

In this paper, we define the notion of approximate bisimulation relation between two continuous systems. While exact bisimulation requires that the observations of two systems are and remain identical, approximate bisimulation allows the observations to be different provided the distance between them remains bounded by some parameter called precision. Approximate bisimulation relations are conveniently defined as level sets of a so-called bisimulation function which can be characterized using Lyapunov-like differential inequalities. For a class of constrained linear systems, we develop computationally effective characterizations of bisimulation functions that can be interpreted in terms of linear matrix inequalities and optimal values of static games. We derive a method to evaluate the precision of the approximate bisimulation relation between a constrained linear system and its projection. This method has been implemented in a Matlab toolbox: MATISSE. An example of use of the toolbox in the context of safety verification is shown.

Key words: Abstractions, Approximation, Bisimulation, Lyapunov techniques, Safety.

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 [6]. 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 [21,26] and hybrid systems [15,22]. These concepts are all exact, requiring observed behaviors of two systems to be identical. For systems observed over a metric space, approximate concepts which give the possibility of an error, certainly allow for more dramatic system compression while providing more robust relationships between systems. Several approaches based on approximate versions of simulation and bisimulation relations have been explored recently for quantitative [7], stochastic [8] and metric [14] transition systems.

In [14], we developed an approximation framework which applies for both discrete and continuous transition systems. We defined an approximate version of bisimulation relations based on a metric on the set of observations by relaxing the observational equivalence required by exact bisimulation relations. Approximate bisimulation relations can be characterized as level sets of a so-called bisimulation function. 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 the design of computationally effective methods for the computation of bisimulation functions. Computational approaches have been developed for constrained linear dynamical systems [12] and nonlinear (but deterministic) dynamical systems [13].

In this paper, we improve and extend our work presented in [12] on approximate bisimulation relations for a class of linear systems with constrained initial states

⋆ This paper was not presented at any IFAC meeting. This research is partially supported by the Région Rhône-Alpes (Projet CalCel) and the National Science Foundation Presidential Early CAREER (PECASE) Grant 0132716. Corresponding author A. Girard. Tel. +33-476514342. Fax +33-476631263.

Email addresses: Antoine.Girard@imag.fr (Antoine Girard), pappasg@seas.upenn.edu (George J. Pappas).

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 and optimal values of static games. We derive a method which evaluates the precision of the approximate bisimulation relation between a constrained linear system and its projection. This method has been implemented in a Matlab toolbox: MATISSE [11] available for download. We conclude this paper by applying our framework in the context of safety verification of constrained linear systems.

2 Approximate Bisimulation Relations

The notion of approximate bisimulation relations allows one to quantify how far two systems are from being bisimilar. The theory has been developed in [14] within the framework of metric transition systems which makes it possible to consider in a unified setting discrete, continuous and hybrid systems. In this paper, we focus on continuous systems of the following form:

$$\Delta_i : \begin{cases} \dot{x}_i(t) = f_i(x_i(t), u_i(t)), \\ y_i(t) = g_i(x_i(t)), \end{cases} , \; i = 1, 2 \tag{1}$$

with $x_i(t) \in \mathbb{R}^{n_i}$, $y_i(t) \in \mathbb{R}^p$ and $x_i(0) \in I_i$ where I_i is a compact subset of $\mathbb{R}^{n_i}$. The inputs $u_i(.)$ are measurable functions with values in U_i, a compact subset of $\mathbb{R}^{m_i}$. We assume that the functions f_i are Lipschitz-continuous and that for all $x_i \in \mathbb{R}^{n_i}$, $f_i(x_i, U_i)$ is a convex set. The functions g_i are assumed to be continuous. Note that both systems are observed on the same space (*i.e.* $\mathbb{R}^p$).

The notion of approximate bisimulation relation is obtained from the exact notion 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 the distance between these observations is and remains bounded by a given parameter δ.

Definition 1 *A relation $\mathcal{R}_\delta \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ is called a δ-approximate bisimulation relation between Δ_1 and Δ_2 if for all $(x_1, x_2) \in \mathcal{R}_\delta$:*

(1) $\|g_1(x_1) - g_2(x_2)\| \leq \delta$,
(2) for all $T > 0$, for all inputs $u_1(.)$ of Δ_1 there exists an input $u_2(.)$ of Δ_2, such that the solutions of $\dot{x}_i(t) = f_i(x_i(t), u_i(t))$, $x_i(0) = x_i$ satisfy for all $t \in [0, T]$, $(x_1(t), x_2(t)) \in \mathcal{R}_\delta$.
(3) for all $T > 0$, for all inputs $u_2(.)$ of Δ_2 there exists an input $u_1(.)$ of Δ_1, such that the solutions of $\dot{x}_i(t) = f_i(x_i(t), u_i(t))$, $x_i(0) = x_i$ satisfy for all $t \in [0, T]$, $(x_1(t), x_2(t)) \in \mathcal{R}_\delta$.

For $\delta = 0$, we recover the definition of exact bisimulation relation. Parameter δ can thus serve to measure how far Δ_1 and Δ_2 are from being exactly bisimilar.

Definition 2 *Δ_1 and Δ_2 are approximately bisimilar with precision δ (noted $\Delta_1 \sim_\delta \Delta_2$), if there exists $\mathcal{R}_\delta$, a δ-approximate bisimulation relation between Δ_1 and Δ_2 such that for all $x_1 \in I_1$, there exists $x_2 \in I_2$ such that $(x_1, x_2) \in \mathcal{R}_\delta$ and conversely.*

Therefore, if $\Delta_1 \sim_\delta \Delta_2$, then for all outputs $y_1(.)$ observed from Δ_1, there exists an output $y_2(.)$ observed from Δ_2, such that their distance remains bounded by the precision δ. Thus, the problem of computing a tight evaluation of the precision of the approximate bisimilarity of two systems is important and can be handled more practically by a dual approach based on functions rather than on relations.

2.1 Bisimulation functions

A bisimulation function is a function whose level sets define approximate bisimulation relations.

Definition 3 *A function $V : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^+ \cup \{+\infty\}$ is a bisimulation function between Δ_1 and Δ_2 if for all $\delta \geq 0$:*

$$\mathcal{R}_\delta = \{(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} | \; V(x_1, x_2) \leq \delta\}$$

is a closed set and is a δ-approximate bisimulation relation between Δ_1 and Δ_2.

Let us remark that the zero set of a bisimulation function is an exact bisimulation relation. Given a bisimulation function, a tight upper-bound of the smallest δ such that $\Delta_1 \sim_\delta \Delta_2$ can be evaluated by solving two static games:

Theorem 1 *[14] Let V be a bisimulation function between Δ_1 and Δ_2 and*

$$\delta \geq \max\left(\sup_{x_1 \in I_1} \inf_{x_2 \in I_2} V(x_1, x_2), \sup_{x_2 \in I_2} \inf_{x_1 \in I_1} V(x_1, x_2) \right). \tag{2}$$

If the value of δ is finite, then $\Delta_1 \sim_\delta \Delta_2$.

Before giving an effective characterization of bisimulation functions, we define the following notations:

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \; f(x, u_1, u_2) = \begin{bmatrix} f_1(x_1, u_1) \\ f_2(x_2, u_2) \end{bmatrix},$$
$$g(x) = g_1(x_1) - g_2(x_2).$$

Intuitively, a bisimulation function is a function which bounds the distance between the observations of Δ_1 and

Δ_2 and which does not increase during the parallel evolution of the systems. More formally, for smooth functions with finite values on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ we can show the following result:

Theorem 2 *Let $q : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^+$ be a continuously differentiable function and $\alpha \geq 0$. If for all $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$,*

$$q(x) \geq \|g(x)\|^2, \tag{3}$$

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

$$\sup_{u_1 \in U_1} \inf_{u_2 \in U_2} \nabla q(x) \cdot f(x, u_1, u_2) \leq 0, \tag{4}$$

$$\sup_{u_2 \in U_2} \inf_{u_1 \in U_1} \nabla q(x) \cdot f(x, u_1, u_2) \leq 0. \tag{5}$$

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

The proof of this result is stated in appendix. An interpretation of the form of the bisimulation function can be given as follows: the term $q(x)$ stands for the error of approximation of the transient dynamics whereas α stands for the error of approximation of the asymptotic dynamics and is thus independent of the initial states of the systems.

2.2 Related notions

Compared to other approximation frameworks for continuous systems such as model reduction techniques [2], the problem we consider is quite different and much more natural for some applications such as safety verification which motivated this work. First, approximation of the input-output mapping is not our main concern. In general, the systems we compare even have different sets of inputs. Second, contrarily to the model reduction framework, the transient dynamics of the systems are not ignored during the approximation process. In fact, the quality of the approximation may critically depend on the set of initial states. Finally, the error bounds we compute are based on the L^∞ norm whereas standard model reduction techniques [2] deal with L^2 or H^∞ norms. In philosophy, our approach is closer to the regulator problem [27] and more generally to the model matching problem [23]. The main difference is that we do not want the systems behaviors to match exactly nor asymptotically but that they remain within given error bounds for the L^∞ norm.

Also, Theorem 2 allows us to relate bisimulation functions and approximate bisimilarity to some other notions in control theory such as robust control Lyapunov functions, input to output stability and incremental stability. We give a short informal discussion of these relations as a rigorous analysis of the connections between these notions is out of the scope of this paper.

There are similarities between the notions of bisimulation function and of robust control Lyapunov function [9] for output stabilization of the composite system given by vector field f and observation function g. Let us consider the input $u_1(.)$ as a disturbance and the input $u_2(.)$ as a control variable in equation (4). Then, the interpretation of this inequality is that for all disturbances there exists a control such that the bisimulation function decreases when the output is far from 0. 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 when the output is far from 0. Thus, it appears that robust control Lyapunov functions require stronger conditions than bisimulation functions. If Δ_1 and Δ_2 are input to output stable [24], then the composite system is also input to output stable, thus there exists a function which decreases for all inputs $u_1(.)$ and $u_2(.)$ when the system output is far from 0. In spirit, it is clear that this function is also a bisimulation function. This should imply that two input to output stable systems are approximately bisimilar. Further evidence of this will be given in the following section for the class of linear systems. Finally, let us remark that if $\Delta_1 = \Delta_2$ is an incremental globally asymptotically stable system [1], then there exists a function which decreases for all inputs $u_1(.)$ and $u_2(.)$ such that $u_1(.) = u_2(.)$. This function can thus be viewed as bisimulation function between Δ_1 and itself.

3 Bisimulation Functions for Linear Systems

In this section, we show that for the class of constrained linear systems, computationally effective characterizations of bisimulation functions can be given. Let us now consider systems of the form:

$$\Delta_i : \begin{cases} \dot{x}_i(t) = A_i x_i(t) + B_i u_i(t), \\ y_i(t) = C_i x_i(t), \end{cases}, \; i = 1, 2 \tag{6}$$

with $x_i(t) \in \mathbb{R}^{n_i}$, $y_i(t) \in \mathbb{R}^p$ and $x_i(0) \in I_i$ where I_i is a compact subset of $\mathbb{R}^{n_i}$. The inputs $u_i(.)$ are measurable functions with values in U_i, a compact convex subset of $\mathbb{R}^{m_i}$. A_i, B_i and C_i are constant matrices of appropriate dimension. We define the following notations

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

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 later in the paper.

3.1 Truncated quadratic bisimulation functions

Regarding Lyapunov-like differential inequalities (4) and (5) in Theorem 2 it is natural, for constrained linear systems to cast the bisimulation functions in the class of truncated quadratic functions:

$$V(x_1, x_2) = \max(\sqrt{x^T M x}, \alpha) \tag{7}$$

where M is a positive semidefinite matrix. Then, a characterization of V under the form of linear matrix inequalities and optimization problems is given by the following result:

Theorem 3 *If there exists $\lambda > 0$, such that*

$$M \geq C^T C \tag{8}$$

$$A^T M + MA + 2\lambda M \leq 0 \tag{9}$$

$$\alpha \geq \frac{1}{\lambda} \sup_{x^T M x = 1} \left(\sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \right) \tag{10}$$

$$\alpha \geq \frac{1}{\lambda} \sup_{x^T M x = 1} \left(\sup_{u_2 \in U_2} \inf_{u_1 \in U_1} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \right) \tag{11}$$

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

Proof: Let $q(x) = x^T M x$, equation (8) is equivalent to equation (3). Let $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that $x^T M x \geq \alpha^2$. Then, equation (10) implies that

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

Therefore,

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

Then, from equation (9)

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

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) holds. Similarly, from equations (9) and (11), we can show that equation (5) holds. Then, from Theorem 2, V is a bisimulation function between Δ_1 and Δ_2. ■

An important consequence of Theorem 3 is that the class of truncated quadratic bisimulation functions are universal for the class of stable constrained linear systems.

Proposition 1 *Let Δ_1 and Δ_2 be asymptotically stable constrained linear systems, then there exists a bisimulation function of the form (7) between Δ_1 and Δ_2.*

Proof: First, let us remark that (9) is equivalent to

$$A_\lambda^T M + MA_\lambda \leq 0 \tag{12}$$

where $A_\lambda = A + \lambda I$. Since all the real parts of the eigenvalues of A_1 and A_2 are strictly negative, it follows that there exists λ small enough such that, the real parts of the eigenvalues of A_λ are all strictly negative. Linear matrix inequality (8) is equivalent to say that $M = C^T C + N$ where N is a positive semidefinite matrix. Then, linear matrix inequality (12) becomes

$$A_\lambda^T N + NA_\lambda \leq -(A_\lambda^T C^T C + C^T C A_\lambda). \tag{13}$$

Let us remark that $A_\lambda^T C^T C + C^T C A_\lambda$ is a symmetric matrix and then can be written as the difference between two positive semidefinite matrices Q^+ and Q^-: $A_\lambda^T C^T C + C^T C A_\lambda = Q^+ - Q^-$. Let us consider the Lyapunov equation $A_\lambda^T N + NA_\lambda = -Q^+$. Since the real parts of the eigenvalues of A_λ are all strictly negative, there exists a unique solution N to this Lyapunov equation. This solution is positive semidefinite and clearly satisfies (13). Thus M satisfies both linear matrix inequalities (8) and (9). Moreover,

$$\begin{aligned}\sup_{x^T M x = 1} \left(\sup_{u_1 \in U_1} \inf_{u_2 \in U_2} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \right) &\leq \\ \sup_{u_1 \in U_1} \sup_{u_2 \in U_2} \left(\sup_{x^T M x = 1} x^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2) \right) &\leq \\ \sup_{u_1 \in U_1} \sup_{u_2 \in U_2} \sqrt{(\overline{B}_1 u_1 + \overline{B}_2 u_2)^T M(\overline{B}_1 u_1 + \overline{B}_2 u_2)}. &\end{aligned}$$

Since, U_1 and U_2 are compact sets, it is easy to see that there exists $\alpha \geq 0$ such that (10) holds. By a symmetric reasoning, there exists $\alpha \geq 0$ such that (11) also holds. ■

Corollary 1 *Let Δ_1 and Δ_2 be asymptotically stable constrained linear systems, then Δ_1 and Δ_2 are approximately bisimilar.*

Proof: The proof is straightforward from the fact that the games given by equation (2) have obviously finite values since I_1 and I_2 are compact sets. ■

The previous result states that any two asymptotically stable constrained linear systems Δ_1 and Δ_2 can be seen as approximations of each other. However, the precision δ with which $\Delta_1 \sim_\delta \Delta_2$ can be very large. An evaluation of this precision is thus necessary in order to get useful information on how well Δ_2 approximates Δ_1 and conversely.

3.2 Handling instability

When Δ_1 and Δ_2 are not asymptotically stable, the results of the previous sections cannot be used. Theorem 2 gives a characterization for a bisimulation function between Δ_1 and Δ_2 with finite values on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Particularly, 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. When dealing with unstable dynamics, this is generally not the case and therefore, bisimulation functions with finite values on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ cannot exist. In the following, we search for bisimulation 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 with the 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

$$A_i = \begin{bmatrix} A_{u,i} & 0 \\ 0 & A_{s,i} \end{bmatrix}, B_i = \begin{bmatrix} B_{u,i} \\ B_{s,i} \end{bmatrix}, C_i = [C_{u,i} \; C_{s,i}], \tag{14}$$

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_i

$$\Delta_{u,i} : \begin{cases} \dot{x}_{u,i}(t) = A_{u,i} x_{u,i}(t) + B_{u,i} u_i(t), \\ y_{u,i}(t) = C_{u,i} x_{u,i}(t) \end{cases}$$

where $x_{u,i}(t) \in E_{u,i}$, $y_{u,i}(t) \in \mathbb{R}^p$, $x_{u,i}(0) \in P_{u,i} I_i$ and the inputs $u_i(.)$ are measurable functions with values in U_i. For $j \in \{u, s\}$, we define the matrices

$$A_j = \begin{bmatrix} A_{j,1} & 0 \\ 0 & A_{j,2} \end{bmatrix}, \; C_j = \begin{bmatrix} C_{j,1} \middle| -C_{j,2} \end{bmatrix} \tag{15}$$

$$\overline{B}_{j,1} = \begin{bmatrix} B_{j,1} \\ 0 \end{bmatrix}, \; \overline{B}_{j,2} = \begin{bmatrix} 0 \\ B_{j,2} \end{bmatrix}.$$

and the projection defined by

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

The following theorem generalizes the result of Theorem 2 to the class of constrained linear systems with unstable modes.

Theorem 4 *Let $\mathcal{R}_u \subseteq E_{u,1} \times E_{u,2}$ be a subspace satisfying:*

$$\mathcal{R}_u \subseteq \ker(C_u), \tag{16}$$
$$A_u \mathcal{R}_u \subseteq \mathcal{R}_u, \tag{17}$$
$$\mathcal{R}_u + \overline{B}_{u,1} U_1 = \mathcal{R}_u - \overline{B}_{u,2} U_2. \tag{18}$$

Let $q_s : E_{s,1} \times E_{s,2} \to \mathbb{R}^+$ be a continuously differentiable, and $\alpha_s \geq 0$. If for all $x_s \in E_{s,1} \times E_{s,2}$,

$$q_s(x_s) \geq x_s^T C_s^T C_s x_s \tag{19}$$

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

$$\sup_{u_1 \in U_1} \inf_{\substack{u_2 \in U_2 \\ \overline{B}_{u,1}u_1 + \overline{B}_{u,2}u_2 \in \mathcal{R}_u}} \nabla q_s^T(x_s) \left(A_s x_s + \overline{B}_{s,1} u_1 + \overline{B}_{s,2} u_2 \right) \leq 0 \tag{20}$$

$$\sup_{u_2 \in U_2} \inf_{\substack{u_1 \in U_1 \\ \overline{B}_{u,1}u_1 + \overline{B}_{u,2}u_2 \in \mathcal{R}_u}} \nabla q_s^T(x_s) \left(A_s x_s + \overline{B}_{s,1} u_1 + \overline{B}_{s,2} u_2 \right) \leq 0 \tag{21}$$

Then, the function $V : \mathbb{R}^{n_1+n_2} \to \mathbb{R}^+ \cup \{+\infty\}$ defined by $V(x_1, x_2) = \max(\sqrt{q_s(P_s x)}, \alpha_s)$ if $P_u x \in \mathcal{R}_u$ and $V(x_1, x_2) = +\infty$ otherwise, is a bisimulation function between Δ_1 and Δ_2.

The proof of this result is stated in appendix. It can be shown [21,26] that the subspace $\mathcal{R}_u$ is actually an exact bisimulation relation between the unstable subsystems $\Delta_{u,1}$ and $\Delta_{u,2}$.

Similar to the case of stable systems, we can cast the function q_s in the class of quadratic forms. The proof of the following result is similar to that of Theorem 3 and is not stated here.

Theorem 5 *Let $\mathcal{R}_u \subseteq E_{u,1} \times E_{u,2}$ be a subspace satisfying equations (16), (17) and (18). If there exists $\lambda_s > 0$, such that*

$$M_s \geq C_s^T C_s \tag{22}$$
$$A_s^T M_s + M_s A_s + 2\lambda_s M_s \leq 0 \tag{23}$$

$$\alpha_s \geq \frac{1}{\lambda_s} \sup_{x_s^T M_s x_s = 1} \left(\sup_{u_1 \in U_1} \inf_{\substack{u_2 \in U_2 \\ \overline{B}_{u,1}u_1 + \overline{B}_{u,2}u_2 \in \mathcal{R}_u}} x_s^T M_s (\overline{B}_{s,1} u_1 + \overline{B}_{s,2} u_2) \right) \tag{24}$$

$$\alpha_s \geq \frac{1}{\lambda_s} \sup_{x_s^T M_s x_s = 1} \left(\sup_{u_2 \in U_2} \inf_{\substack{u_1 \in U_1 \\ \overline{B}_{u,1}u_1 + \overline{B}_{u,2}u_2 \in \mathcal{R}_u}} x_s^T M_s (\overline{B}_{s,1} u_1 + \overline{B}_{s,2} u_2) \right) \tag{25}$$

Then, the function $V : \mathbb{R}^{n_1+n_2} \to \mathbb{R}^+ \cup \{+\infty\}$ defined by $V(x_1, x_2) = \max(\sqrt{x^T P_s^T M_s P_s x}, \alpha_s)$ if $P_u x \in \mathcal{R}_u$ and $V(x_1, x_2) = +\infty$ otherwise, is a bisimulation function between Δ_1 and Δ_2.

If there is a subspace $\mathcal{R}_u \subseteq E_{u,1} \times E_{u,2}$ satisfying equations (16), (17) and (18) then, similar to Proposition 1, we can show that there always exists a bisimulation function as in Theorem 5 between Δ_1 and Δ_2. As a consequence, we have:

Corollary 2 *If there exists a subspace $\mathcal{R}_u$ satisfying equations (16), (17) and (18), 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. the unstable subsystems $\Delta_{u,1}$ and $\Delta_{u,2}$ are exactly bisimilar), then Δ_1 and Δ_2 are approximately bisimilar.*

Proof: Let us consider the games given by equation (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. ■

4 Linear Systems Approximation

Projections are often used for linear systems approximation, in classical model reduction techniques [2] but also in approaches based on exact simulation and bisimulation relations [21,26]. In this section, we use the previous results to compute the precision of the approximate bisimulation relation between a linear system with constrained inputs Δ_1 of the form (6) and a projection Δ_2. Let us assume that the system Δ_1 has been decomposed into stable and unstable subsystems and that the matrices A_1, B_1, C_1 are of the form given by equation (14). 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 the matrices $A_2 = HA_1H^+$, $B_2 = HB_1$, $C_2 = C_1H^+$, and the sets of initial states and inputs $I_2 = HI_1$ and $U_2 = U_1$, 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,

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

Hence, the matrices A_2, B_2, C_2 are also of the form given by equation (14).

Lemma 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 (16), (17) and (18) if and only if

$$C_{u,1} = C_{u,1} H_u^+ H_u, \tag{26}$$
$$H_u A_{u,1} = H_u A_{u,1} H_u^+ H_u. \tag{27}$$

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

Proof: Firstly, let us remark that equation (26) means that $C_{u,1} - C_{u,2}H_u = 0$ which is equivalent to $\mathcal{R}_u \subseteq \ker(C_u)$. Secondly,equation (27) means that $H_u A_{u,1} = A_{u,2} H_u$ which is equivalent to $A_u \mathcal{R}_u \subseteq \mathcal{R}_u$. Finally, for all $u \in U_1$, $H_u B_{u,1} u = B_{u,2} u$. Since $U_1 = U_2$, equation (18) holds. Therefore, $\mathcal{R}_u$ 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 $x_1 \in \mathbb{R}^{n_1}$, $H_u P_{u,1} x_1 = P_{u,2} H x_1$. 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 $x_{u,2} = H_u x_{u,1} = H_u P_{u,1} x_1 = P_{u,2} H x_1$, hence $x_{u,2} \in P_{u,2} I_2$ and $(x_{u,1}, x_{u,2}) \in \mathcal{R}_u$. Similarly, for all $x_{u,2} \in P_{u,2} I_2$, $x_{u,2} = P_{u,2} H x_1$ 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, $\Delta_{u,1}$ and $\Delta_{u,2}$ are exactly bisimilar. ■

Let us assume that the map H_u is chosen such that equations (26) and (27) 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 previous sections, we know that there exists a bisimulation function between Δ_1 and Δ_2 as in Theorem 5. Let A_s, $\overline{B}_{s,1}$, $\overline{B}_{s,2}$ and C_s be defined as in equation (15). There exist a matrix M_s and a real number $\lambda_s > 0$ satisfying equations (22) and (23). Let us define the matrix

$$Q_s = \begin{bmatrix} I & H_s^T \end{bmatrix} M_s \begin{bmatrix} I \\ H_s \end{bmatrix}.$$

Theorem 6 *Let α_s be defined by*

$$\alpha_s = \frac{1}{\lambda_s} \sup_{u_1 \in U_1} \sqrt{u_1^T B_{s,1}^T Q_s B_{s,1} u_1}. \tag{28}$$

Then, the function $V : \mathbb{R}^{n_1+n_2} \to \mathbb{R}^+ \cup \{+\infty\}$ defined by $V(x_1, x_2) = \max(\sqrt{x^T P_s^T M_s P_s x}, \alpha_s)$ if $P_u x \in \mathcal{R}_u$ and $V(x_1, x_2) = +\infty$ otherwise is a bisimulation function between Δ_1 and Δ_2.

Proof: We assumed that H_u is such that $\mathcal{R}_u$ satisfies equations (16), (17), (18). Furthermore, M_s and λ_s satisfy equations (22) and (23). Now, let us remark that

$$\begin{aligned}\alpha_s &= \frac{1}{\lambda_s}\sup_{u_1\in U_1}\sqrt{u_1^T(\overline{B}_{s,1}+\overline{B}_{s,2})^T M_s(\overline{B}_{s,1}+\overline{B}_{s,2})u_1} \\ &= \frac{1}{\lambda_s}\sup_{x_s^T M_s x_s}\left(\sup_{u_1\in U_1} x_s^T M_s(\overline{B}_{s,1}+\overline{B}_{s,2})u_1\right).\end{aligned}$$

Since $U_1 = U_2$, this equation implies that equations (24) and (25) hold. Then, from Theorem 5, V is a bisimulation function between Δ_1 and Δ_2. ∎

From Theorem 1, the precision of the approximate bisimulation relation between Δ_1 and Δ_2 can then be evaluated by solving the games given by equation (2).

Theorem 7 *Let α_s be defined as in equation (28), let β_s be defined as*

$$\beta_s = \sup_{x_1\in I_1}\sqrt{x_1^T P_{s,1}^T Q_s P_{s,1} x_1}. \qquad (29)$$

Let $\delta = \max(\alpha_s, \beta_s)$. Then, $\Delta_1 \sim_\delta \Delta_2$.

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,

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

Similarly, we also have,

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

Hence, the values of the games in equation (2) are bounded by $\max(\alpha_s, \beta_s)$ which implies, from Theorem 1, that the systems Δ_1 and Δ_2 are approximately bisimilar with the precision δ. ∎

We presented a method to evaluate the precision of the approximate bisimulation relation between a constrained linear system and its projection. From the computational point of view, it requires to solve the linear matrix inequalities (22) (23). Then, if we assume that I_1 and U_1 are polytopes, the precision of the approximate bisimulation relation between a constrained linear system and its projection can be computed by solving two linear quadratic programs given by equations (28) and (29). Solving the linear matrix inequalities can be done using semi-definite programming [25]. 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 relation between Δ_1 and 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. An important parameter in this method 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 relation between Δ_1 and Δ_2.

An open question is how do we choose the surjective map H so that the precision of the approximate bisimulation relation between Δ_1 and its projection Δ_2 of desired dimension is minimal. 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 [21,26], 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 [2]. 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 (28) and (29). 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 map H_s.

Our method has been implemented in a Matlab toolbox available for download: MATISSE (Metrics for Approximate TransItion Systems Simulation and Equivalence [11]). It uses several toolboxes such as the Multi-Parametric Toolbox [18] for polytopes manipulation, the interface YALMIP [19] to translate linear matrix inequalities into semi-definite programs which are solved by the toolbox SEDUMI [25]. MATISSE allows the reduction of a constrained linear system Δ_1 to a system Δ_2 of given dimension, and the computation of the precision of the approximate bisimulation relation between Δ_1 and Δ_2.

5 Application to Safety Verification

In this section, we show an example of application of the toolbox MATISSE [1] . Let us consider Δ_1, a constrained linear system as in equation (6) where the matrices

$$A_1 = \begin{bmatrix} 1 & 0 & -0.4 & 2 & 0.24 & 1.6 & -0.6 & 0 & 0.54 & 0 \\ 0 & 0.8 & -2 & -0.3 & 4 & -0.5 & 0 & 0.3 & 0 & -0.18 \\ 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.2 & -8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & -0.2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

B_1 is the 10×10 identity matrix and C_1 is the projection matrix over the first two components of $\mathbb{R}^{10}$. The set of inputs is $U_1 = [-0.1, 0.1]^{10}$ and the set of initial states is

$$I_1 = [2.9, 3.1] \times [-0.1, 0] \times [1.9, 2]^5 \times [2.4, 2.6] \times [1.9, 2.1]^2$$

Let $T > 0$, we define the reachable set of Δ_1 on $[0, T]$ as $\text{Reach}_{[0,T]}(\Delta_1)$, the subset of $\mathbb{R}^p$ consisting of the points y_1 such that there exists an input $u_1(.)$ of Δ_1, an initial state $x_1 \in I_1$ and a time $\tau \in [0, T]$ such that the solution of $\dot{x}_1(t) = A_1 x_1(t) + B_1 u_1(t)$, $x_1(0) = x_1$ satisfies $y_1 = C_1 x_1(\tau)$. We would like to verify that the system satisfy the following safety property:

$$\text{Reach}_{[0,T]}(\Delta_1) \cap \text{Unsafe} = \emptyset$$

where Unsafe is a set of observations associated with unsafe states of the system. Here, the inputs $u_1(.)$ have to be seen as internal disturbances introducing non-determinism in the behavior of Δ_1 rather than control inputs. Safety verification can be handled by reachability analysis for which several computational techniques have been developed [3,5,17,20]. Though recent progress has been made in the reachability analysis of high dimensional systems [10,16,28], it remains one of the most challenging issues of the verification of continuous and hybrid systems, motivating the use of simple approximate models for the verification of complex systems. Let Δ_2 be a constrained linear system such that $\Delta_1 \sim_\delta \Delta_2$, then it is easy to show that

$$\text{Reach}_{[0,T]}(\Delta_1) \subseteq \mathcal{N}(\text{Reach}_{[0,T]}(\Delta_2), \delta)$$

where $\mathcal{N}(., \delta)$ denotes the δ neighborhood of a set. Consequently, to prove that Δ_1 is safe, it is sufficient to verify that

$$\text{Reach}_{[0,T]}(\Delta_2) \cap \mathcal{N}(\text{Unsafe}, \delta) = \emptyset.$$

[1] This example is available as a demo file in MATISSE.

Δ_1 has a four dimensional unstable subsystem $\Delta_{u,1}$. From Corollary 2, Δ_1 and $\Delta_{u,1}$ are approximately bisimilar. Following the method described in the previous section we evaluate the precision of the approximate bisimulation relation between these two systems. The computations give $\delta = 1.9027$. We computed the reachable sets (for $T = 2$) of both systems using zonotope techniques for reachability analysis of constrained linear systems [10] implemented in MATISSE. In Figure 1, we represented the reachable sets of the ten dimensional system and of its four dimensional approximation. We can see that the approximation does not allow us to conclude though Δ_1 is actually safe.

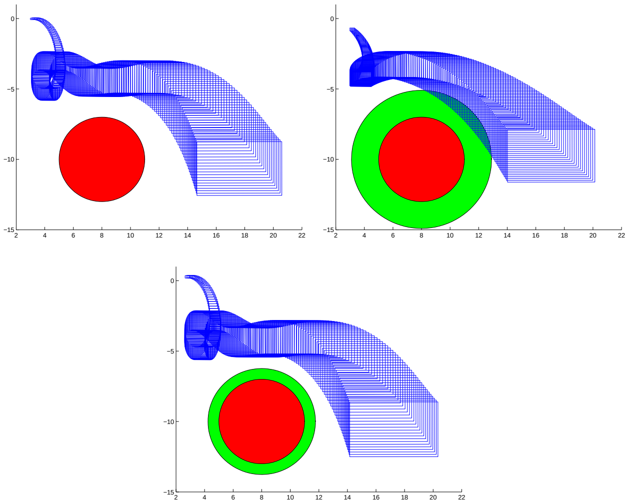

Fig. 1. Reachable sets of the original ten dimensional system (top left) and of its four dimensional and six dimensional approximations (top right and bottom). The disk on the left figure and the inner disks on the right and bottom figure represent the set Unsafe. The outer disks on the right and bottom figures consist of the set of points whose distance to Unsafe is smaller than the precision of the approximate bisimulation relation between Δ_1 and its approximation.

Therefore, we need to refine the approximation. We consider a six dimensional approximation Δ_2 which is a combination of the unstable subsystem $\Delta_{u,1}$ with a stable subsystem. Then, from Corollary 2, we know that Δ_1 and Δ_2 are approximately bisimilar. The better the stable subsystem of Δ_2 approximates the stable subsystem of Δ_1, the better the system Δ_2 approximates system Δ_1. For our example, we chose the stable subsystem of Δ_2 as the projection of the stable subsystem of Δ_1 on the two dimensional space spanned by the eigenvectors associated to the two largest eigenvalues of the matrix $A_{s,1}$. The precision of the approximate bisimulation relation between Δ_1 and Δ_2 evaluated by the method presented in the previous section is $\delta = 0.76329$. We can see on Figure 1 that the approximation of Δ_1 by the six dimensional system Δ_2 allows us to check the safety of Δ_1.

This example also illustrates the important point that robustness simplifies verification. Indeed, if the distance between $\mathrm{Reach}_{[0,T]}(\Delta_1)$ and Unsafe would have been larger then the approximation of Δ_1 by its unstable subsystem might have been sufficient to check the safety of Δ_1. 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.

6 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 method to compute the precision of the approximate bisimulation relation between a system and a projection. This method only requires to solve a set of linear matrix inequalities and two linear quadratic programs and is therefore computationally effective.

This method has been implemented within a Matlab toolbox, MATISSE [11]. MATISSE allows the reduction of a constrained linear system to a system of given dimension and the computation of the precision of the approximate bisimulation relation between the original system and its approximation. An example of application was shown. 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

[1] D. Angeli. A Lyapunov approach to incremental stability properties. *IEEE Trans. Auto. Control*, 47:410–422, 2002.

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

[3] E. Asarin, O. Bournez, T. Dang, and O. Maler. Approximate reachability analysis of piecewise linear dynamical systems. In *HSCC*, volume 1790 of *LNCS*, pages 21–31. Springer, 2000.

[4] J. P. Aubin. *Viability Theory*. Birkhauser, 1991.

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

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

[7] 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.

[8] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled markov processes. *Theo. Comp. Sc.*, 318(3):323–354, June 2004.

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

[10] A. Girard. Reachability of uncertain linear systems using zonotopes. In *HSCC*, volume 3414 of *LNCS*, pages 291–305. Springer, 2005.

[11] A. Girard, A. A. Julius, and G. J. Pappas. MATISSE, 2005. `http://www.seas.upenn.edu/~agirard/Software/MATISSE/`.

[12] A. Girard and G. J. Pappas. Approximate bisimulations for constrained linear systems. In *Proc. CDC and ECC*, pages 4700–4705, 2005.

[13] A. Girard and G. J. Pappas. Approximate bisimulations for nonlinear dynamical systems. In *Proc. CDC and ECC*, pages 684–689, 2005.

[14] A. Girard and G. J. Pappas. Approximation metrics for discrete and continuous systems. *IEEE Trans. on Auto. Control*, 2007. to appear.

[15] 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.

[16] Z. Han and B.H. Krogh. Reachability analysis of large-scale affine systems using low-dimensional polytopes. In *HSCC*, volume 3927 of *LNCS*, pages 287–301. Springer, 2006.

[17] A. Kurzhanski and P. Varaiya. Ellipsoidal tehcniques for reachability analysis. In *HSCC*, volume 1790 of *LNCS*. Springer, 2000.

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

[19] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proc. CACSD*, 2004. `http://control.ee.ethz.ch/~joloef/yalmip.php`.

[20] I. Mitchell and C. Tomlin. Level set methods for computation in hybrid systems. In *HSCC*, volume 1790 of *LNCS*. Springer, 2000.

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

[22] G. Pola, A.J. van der Schaft, and M. D. Di Benedetto. Bisimulation theory for switching linear systems. In *Proc. CDC*, pages 555–569, 2004.

[23] S. Sastry. *Nonlinear Systems: Analysis, Stability and Control*. Springer, 1999.

[24] E. Sontag and Y. Wang. Lyapunov characterizations of input to output stability. *SIAM J. Control and Opt.*, 39:226–249, 2001.

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

[26] A. van der Schaft. Equivalence of dynamical systems by bisimulation. *IEEE Trans. on Auto. Control*, 49(12):2160–2172, 2004.

[27] W.M. Wonham. *Linear Multivariable Control: a Geometric Approach*. Springer, 1979.

[28] H. Yazarel and G. J. Pappas. Geometric programming relaxations for linear system reachability. In *Proc. ACC*, 2004.

A Proof of Theorem 2

The proof of Theorem 2 requires several preliminary results.

Lemma 2 *Let $i = 1, 2$, $x_i \in \mathbb{R}^{n_i}$ and $T > 0$, then for all inputs $u_i(.)$ of Δ_i, the solution of $\dot{x}_i(t) = f_i(x_i(t), u_i(t))$, $x_i(0) = x_i$ satisfies for all $t, t' \in [0, T]$, with $t \le t'$*

$$\|x_i(t') - x_i(t)\| \le \sup_{u_i \in U_i} \|f_i(x_i(t), u_i)\| \frac{e^{\lambda_i(t'-t)} - 1}{\lambda_i}$$

where λ_i is the Lipschitz constant of f_i.

The proof of this result is not stated here but is a straightforward consequence of Filippov's Theorem (see [4], p.170).

Lemma 3 *Let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ and $T > 0$, 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, the solutions of $\dot{x}_i(t) = f_i(x_i(t), u_i(t))$, $x_i(0) = x_i$ satisfy for all $u_1 \in U_1$, $u_2 \in U_2$, $t, t' \in [0, T]$, with $t \le t' \le t + h$*

$$|\nabla q(x(t)) \cdot f(x(t), u_1, u_2) - \nabla q(x(t')) \cdot f(x(t'), u_1, u_2)| \le \frac{\varepsilon}{T}$$

where $x(t) = (x_1(t), x_2(t))$.

Proof: From Lemma 2, we have for all $t \in [0, T]$,

$$\|x_i(t)\| \le \|x_i\| + \sup_{u_i \in U_i} \|f_i(x_i, u_i)\| \frac{e^{\lambda_i T} - 1}{\lambda_i} = m_i.$$

Note that $\mathcal{C}_i = \{z_i \in \mathbb{R}^{n_i}, \|z_i\| \le m_i\}$ is a compact set. Then, since $\nabla q(z) \cdot f(z, u_1, u_2)$ is continuous on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ it is uniformly 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 $u_1 \in U_1$, $u_2 \in U_2$, $z_1, z_1' \in \mathcal{C}_1$, $\|z_1 - z_1'\| \le \xi$ and $z_2, z_2' \in \mathcal{C}_2$, $\|z_2 - z_2'\| \le \xi$,

$$|\nabla q(z) \cdot f(z, u_1, u_2) - \nabla q(z') \cdot f(z', u_1, u_2)| \le \frac{\varepsilon}{T}. \quad \text{(A.1)}$$

From Lemma 2, we have for all $t, t' \in [0, T]$, with $t \le t'$,

$$\|x_i(t') - x_i(t)\| \le \sup_{x_i \in \mathcal{C}_i, u_i \in U_i} \|f_i(x_i, u_i)\| \frac{e^{\lambda_i(t'-t)} - 1}{\lambda_i}$$

Therefore, there exists $h > 0$, such that for all $t, t' \in [0, T]$, with $t \le t' \le t + h$,

$$\|x_1(t') - x_1(t)\| \le \xi \text{ and } \|x_2(t') - x_2(t)\| \le \xi. \quad \text{(A.2)}$$

Then, equations (A.1) and (A.2) allow us to conclude. ∎

Lemma 4 *Let q be a function as in Theorem 2. Let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ satisfying $q(x) \ge \alpha^2$, and $T > 0$, then for all inputs $u_1(.)$ of Δ_1, for all $\varepsilon > 0$, there exists an input $u_2(.)$ of Δ_2, such that the solutions of $\dot{x}_i(t) = f_i(x_i(t), u_i(t))$, $x_i(0) = x_i$ satisfy*

$$\forall t \in [0, T],\ q(x(t)) \le q(x) + \varepsilon.$$

Proof: Let $h > 0$ be given as in Lemma 3 (we assume without loss of generality that $T/h = N \in \mathbb{N}$). From equation (4), there exists an input $u_2(.)$ of Δ_2 such that for all $t \in [0, h]$, $\nabla q(x) \cdot f(x, u_1(t), u_2(t)) \le 0$. Let us remark that for all $t \in [0, h]$,

$$q(x(t)) - q(x) = \int_0^t \nabla q(x(s)) \cdot f(x(s), u_1(s), u_2(s)) ds.$$

Then, from Lemma 3, for all $t \in [0, h]$,

$$q(x(t)) - q(x) \le \int_0^t \nabla q(x(0)) \cdot f(x(0), u_1(s), u_2(s)) + \varepsilon/T\, ds \le \frac{h\varepsilon}{T}.$$

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

$$\forall t \in [0, ih],\ q(x(t)) - q(x) \le \frac{ih\varepsilon}{T}. \quad \text{(A.3)}$$

We showed that this is true for $i = 1$. If $q(x(ih)) \ge \alpha^2$, then, according to equation (4), we can choose $u_2(.)$ of Δ_2 such that

$$\forall t \in [ih, (i+1)h],\ \nabla q(x(ih)) \cdot f(x(ih), u_1(t), u_2(t)) \le 0.$$

Then, from Lemma 3, for all $t \in [ih, (i+1)h]$,

$$q(x(t)) - q(x(ih)) \le \int_{ih}^t \nabla q(x(ih)) \cdot f(x(ih), u_1(s), u_2(s)) + \varepsilon/T\, ds \le \frac{h\varepsilon}{T}.$$

Together with equation (A.3), we have

$$\forall t \in [ih, (i+1)h],\ q(x(t)) - q(x) \le \frac{(i+1)h\varepsilon}{T}.$$

If $q(x(ih)) < \alpha^2$. Let $v_2(.)$ be an input of Δ_2, let $z_2(.)$ be the solution of $\dot{z}_2(t) = f_2(z_2(t), v_2(t))$, $z_2(ih) = x_2(ih)$. If for all $t \in [ih, (i+1)h]$, $q(x_1(t), z_2(t)) \le \alpha^2$, then we choose for all $t \in [ih, (i+1)h]$, $u_2(t) = v_2(t)$ and therefore for all $t \in [ih, (i+1)h]$,

$$q(x(t)) - q(x) \le \alpha^2 - q(x) \le 0 \le \frac{(i+1)h\varepsilon}{T}.$$

Otherwise, let $t^* \in (ih, (i+1)h)$ be the first time when $q(x_1(t^*), z_2(t^*)) = \alpha^2$. Let $x^* = (x_1(t^*), z_2(t^*))$. Then, according to equation (4), we can choose $u_2(.)$ of Δ_2 such that for all $t \in [ih, t^*)$, $u_2(t) = v_2(t)$ and for all $t \in [t^*, (i+1)h]$, $\nabla q(x^*) \cdot f(x^*, u_1(t), u_2(t)) \leq 0$. Then, from Lemma 3, for all $t \in [t^*, (i+1)h]$,

$$q(x(t)) - q(x(t^*)) \leq \int_{t^*}^{t} \nabla q(x^*) \cdot f(x^*, u_1(s), u_2(s)) + \varepsilon/T \, ds \leq \frac{h\varepsilon}{T}.$$

Hence, for all $t \in [ih, (i+1)h]$,

$$q(x(t)) - q(x) \leq q(x(t)) - q(x(t^*)) \leq \frac{h\varepsilon}{T} \leq \frac{(i+1)h\varepsilon}{T}.$$

Then equation (A.3) holds for all $i \in \{1, \dots N\}$ and particularly (for $i = N$) there exists an input $u_2(.)$ of Δ_2 such that for all $t \in [0, T]$, $q(x(t)) - q(x) \leq \varepsilon$. ■

Lemma 5 *Let q be a function as in Theorem 2. Let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ satisfying $q(x) \geq \alpha^2$, and $T > 0$, then for all inputs $u_1(.)$ of Δ_1, there exists an input $u_2(.)$ of Δ_2, such that the solutions of $\dot{x}_i(t) = f_i(x_i(t), u_i(t))$, $x_i(0) = x_i$ satisfy*

$$\forall t \in [0, T], \; q(x(t)) \leq q(x).$$

Proof: Let $\{\varepsilon_n\}_{n\in\mathbb{N}}$ be a decreasing sequence converging to 0. From Lemma 4, for all $n \in \mathbb{N}$, there exists an input $u_2^n(.)$ of Δ_2 such that the solution of $\dot{x}_2^n(t) = f_2(x_2^n(t), u_2^n(t))$, $x_2^n(0) = x_2$ satisfy for all $t \in [0, T]$, $q(x_1(t), x_2^n(t)) \leq q(x) + \varepsilon_n$. We can prove (see [4], p.101) that the set $\mathcal{S}_2(x_2)$ consisting of the functions $z_2(.)$ such that $\dot{z}_2(t) = f_2(x_2(t), u_2(t))$ with $z_2(0) = x_2$ for some input $v_2(.)$ of Δ_2 is a compact subset of the space of continuous functions supplied with the topology of uniform convergence on compact intervals. Therefore, there exists a subsequence $\{x_2^{n_k}(.)\}_{k\in\mathbb{N}}$ which converges uniformly on $[0, T]$ to some $x_2(.)$ in $\mathcal{S}_2(x_2)$. The end of the proof is straightforward. ■

We can now prove Theorem 2. Let $\delta \geq 0$, let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, such that $V(x_1, x_2) = \max(\sqrt{q(x)}, \alpha) \leq \delta$. First, let us remark that from equation (3), we have $\|g_1(x_1) - g_2(x_2)\| \leq \sqrt{q(x)} \leq V(x_1, x_2) \leq \delta$. Thus, the first property of Definition 1 is satisfied. Let $T > 0$ and $u_1(.)$ an input of Δ_1, if $q(x) \geq \alpha^2$ then from Lemma 5, there exists an input $u_2(.)$ of Δ_2, such that the solutions of $\dot{x}_i(t) = f_i(x_i(t), u_i(t))$, $x_i(0) = x_i$ satisfy for all $t \in [0, T]$, $V(x_1(t), x_2(t)) \leq \sqrt{q(x)} \leq \delta$. If $q(x) < \alpha^2$, let $v_2(.)$ be an input of Δ_2, let $z_2(.)$ be the solution of $\dot{z}_2(t) = f_2(z_2(t), v_2(t))$, $z_2(0) = x_2$. If for all $t \in [0, T]$, $q(x_1(t), z_2(t)) \leq \alpha^2$, then we choose for all $t \in [0, T]$, $u_2(t) = v_2(t)$ and therefore for all $t \in [0, T]$, $V(x_1(t), x_2(t) \leq \alpha^2 \leq \delta$. Otherwise, let $t^* \in (0, T)$ be the first time when $q(x_1(t^*), z_2(t^*)) = \alpha^2$. Let $x^* = (x_1(t^*), z_2(t^*))$. Then, from Lemma 5, we can choose an input $u_2(.)$ of Δ_2 such that for all $t \in [0, t^*)$, $u_2(t) = v_2(t)$, and for all $t \in [t^*, T]$, the solution of $\dot{x}_2(t) = f_2(x_2(t), u_2(t))$, $x_2(t^*) = z_2(t^*)$ satisfies for all $t \in [t^*, T]$, $V(x_1(t), x_2(t)) \leq \sqrt{q(x(t^*))} = \alpha^2$. Then, for all $t \in [0, T]$, $V(x_1(t), x_2(t)) \leq \alpha^2 \leq \delta$. Then, the second property of Definition 1 holds. Similarly, we can show that the third property of Definition 1 holds as well which leads to the conclusion of Theorem 2.

B Proof of Theorem 4

The technical details of the proof are similar to that of Theorem 2: using the same kind of arguments than the ones leading to Lemma 5, we can show the following result.

Lemma 6 *Let $\mathcal{R}_u \subseteq E_{u,1} \times E_{u,2}$ be a subspace satisfying equations (16), (17) and (18), let q_s be a function as in Theorem 4. Let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ satisfying $P_u x \in \mathcal{R}_u$ and $q_s(P_s x) \geq \alpha_s^2$, let $T > 0$, then for all inputs $u_1(.)$ of Δ_1, there exists an input $u_2(.)$ of Δ_2, such that*

$$\forall t \in [0, T], \; \overline{B}_{u,1} u_1(t) + \overline{B}_{u,2} u_2(t) \in \mathcal{R}_u$$

and the solutions of $\dot{x}_i(t) = A_i x_i(t) + B_i u_i(t)$, $x_i(0) = x_i$ satisfy

$$\forall t \in [0, T], \; q_s(P_s x(t)) \leq q_s(P_s x).$$

Let us prove Theorem 4. Let $\delta \geq 0$, let $(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, such that $V(x_1, x_2) \leq \delta$. Then, we must have $P_u x \in \mathcal{R}_u$ and therefore $V(x_1, x_2) = \max(\sqrt{q_s(P_s x)}, \alpha)$. First, let us remark that from equation (16),

$$\|C_1 x_1 - C_2 x_2\| = \|C_s P_s x + C_u P_u x\| = \|C_s P_s x\|.$$

Then, from equation (19), we have $\|C_1 x_1 - C_2 x_2\| \leq \sqrt{q_s(P_s x)} \leq V(x_1, x_2) \leq \delta$. Thus, the first property of Definition 1 is satisfied. Let $T > 0$ and $u_1(.)$ an input of Δ_1, if $q_s(P_s x) \geq \alpha_s^2$ then from Lemma 6, there exists an input $u_2(.)$ of Δ_2, such that the solutions of $\dot{x}_i(t) = A_i x_i(t) + B u_i(t)$, $x_i(0) = x_i$ satisfy for all $t \in [0, T]$, $q_s(P_s x(t)) \leq q_s(P_s x) \leq \delta$ and $\overline{B}_{u,1} u_1(t) + \overline{B}_{u,2} u_2(t) \in \mathcal{R}_u$. Moreover since $E_{u,1}$ and $E_{u,2}$ are invariant under A_1 and A_2, we have that

$$P_u x(t) = e^{A_u t} P_u x + \int_0^t e^{A_u (t-s)} \left(\overline{B}_{u,1} u_1(s) + \overline{B}_{u,2} u_2(s)\right) ds.$$

Thus, for all $t \in [0, T]$, it is clear that $P_u x(t) \in \mathcal{R}_u$ and therefore for all $t \in [0, T]$, $V(x_1(t), x_2(t)) = \max(\sqrt{q_s(P_s x(t))}, \alpha_s) \leq \delta$. If $q_s(P_s x) < \alpha_s^2$, let

$v_2(.)$ be an input of Δ_2 such that for all $t \in [0,T]$ $\overline{B}_{u,1}u_1(t) + \overline{B}_{u,2}v_2(t) \in \mathcal{R}_u$. Let $z_2(.)$ be the solution of $\dot{z}_2(t) = A_2 z_2(t) + B_2 v_2(t)$, $z_2(0) = x_2$. Clearly, for all $t \in [0,T]$, $(P_{u,1}x_1(t), P_{u,2}z_2(t)) \in \mathcal{R}_u$. If for all $t \in [0,T]$, $q_s(P_{s,1}x_1(t), P_{s,2}z_2(t)) \leq \alpha_s^2$, then we choose for all $t \in [0,T]$, $u_2(t) = v_2(t)$ and therefore for all $t \in [0,T]$, $V(x_1(t), x_2(t) \leq \alpha_s^2 \leq \delta$. Otherwise, let $t^* \in (0,T)$ be the first time when $q_s(P_{s,1}x_1(t^*), P_{s,2}z_2(t^*)) = \alpha^2$. Let $x^* = (x_1(t^*), z_2(t^*))$. Then, from Lemma 5, we can choose an input $u_2(.)$ of Δ_2 such that for all $t \in [0,t^*)$, $u_2(t) = v_2(t)$, and for all $t \in [t^*,T]$, the solution of $\dot{x}_2(t) = A_2 x_2(t) + B_2 u_2(t)$, $x_2(t^*) = z_2(t^*)$ satisfies for all $t \in [t^*,T]$, $q_s(P_s x(t)) \leq q_s(P_s x(t^*)) = \alpha^2$ and $\overline{B}_{u,1}u_1(t) + \overline{B}_{u,2}u_2(t) \in \mathcal{R}_u$. Then, for all $t \in [0,T]$, $V(x_1(t), x_2(t)) \leq \alpha^2 \leq \delta$. Then, the second property of Definition 1 holds. Similarly, we can show that the third property of Definition 1 holds as well which leads to the conclusion of Theorem 4.