Approximately Bisimilar Abstractions of Incrementally Stable Finite or Infinite Dimensional Systems

Antoine Girard

Abstract—In this paper, we establish a certain number of results for abstraction of a class of incrementally stable dynamical systems, in the framework of approximate bisimulation. Our approach does not rely on a discretization of the state space, it is therefore applicable indifferently to finite dimensional systems such as those modeled by differential equations, or infinite dimensional systems, such as those modeled by time-delay or partial differential equations. Our first result states that the sampled dynamics of an incrementally stable dynamical system is approximately bisimilar to a family of finite dimensional dynamical systems. The second result shows that these finite dimensional systems admit approximately bisimilar symbolic abstractions. In both cases, any precision can be achieved either by increasing the dimension or the number of states of the abstractions.

I. INTRODUCTION

Estimating the behavioral distance of systems of possibly heterogeneous nature can be tackled through the notion of approximate simulation or bisimulation relations, introduced in [1]. These relations have enabled the development of several abstraction based control design methodologies (see [2] and [3] and the references therein) where complex, typically continuous, dynamical systems can be approximated either by simpler, continuous, or even by symbolic (i.e. with finite state-space) dynamical systems. The former are typically obtained by projection of the state-space on a smaller dimensional space (as in [4], [5]), while the latter are usually obtained through state-space quantization (as in [6], [7]). Then, approximate bisimilarity is generally guaranteed under an assumption of incremental stability of the original dynamics. These approaches are typically applied to finite dimensional systems with the notable exception of [8], where symbolic abstractions for time-delay systems are obtained by projecting the infinite dimensional functional state-space on a finite dimensional subspace (e.g. by spline approximation) which is then quantized in order to define symbolic models. One can qualify these approaches as state-based as the states of these abstractions are obtained by aggregation of states of the original system.

An alternative construction of approximately bisimilar symbolic abstractions for incrementally stable switched systems has been proposed in [9]. In that work, the states of the abstractions correspond to input sequences of the original system. In this paper, we generalize this input-based approach to a much wider class of dynamical systems. We propose a generic approach which applies indifferently to finite dimensional systems (e.g. modeled by differential equations) or infinite dimensional systems (e.g. modeled by time-delay or partial differential equations). We propose two main results. Our first result states that incrementally globally asymptotically stable systems admit approximately bisimilar finite dimensional abstractions; this is particularly useful for infinite dimensional dynamical systems. Moreover, any precision can be achieved by increasing the dimension of the abstraction. If the original system is linear, we show that the abstractions are also linear systems. Our second result states that incrementally input-to-state stable systems admit approximately bisimilar symbolic abstractions. Any precision can be achieved by increasing the number of states of the abstraction.

The paper is organized as follows, Section II introduces the necessary background on approximate bisimulation and incremental stability. Then, the constructions of finite dimensional abstractions and symbolic abstractions are presented in III and IV, respectively.

Notations: ||·|| denotes the usual Euclidean vector norm over \( \mathbb{R}^n \) as well as the associated induced matrix norm. Let \( I \subseteq \mathbb{R} \) and \( J \subseteq \mathbb{R}^n \) \((n \geq 1)\), the sets of continuous and piecewise continuous functions from \( I \) to \( J \) are denoted \( C^0(I,J) \) and \( C^0_{pw}(I,J) \) respectively. Given a function \( f \in C^0_{pw}(I,J) \), we denote \( \|f\|_\infty = \sup_{t \in I} \|f(t)\|_\infty \). \( L_2(I,J) \) is the set of functions \( f : I \to J \) such that \( \int_I \|f(t)\|^2 \, dt < +\infty \). \( \mathbb{R}^+_0 \) denotes the set of nonnegative reals. A function \( \gamma \in C^0(\mathbb{R}_0^+, \mathbb{R}^+_0) \) is said to be of class \( K \) if it is strictly increasing and \( \gamma(0) = 0 \), it is of class \( K_\infty \) if in addition \( \gamma(r) \to +\infty \) when \( r \to +\infty \). A function \( \beta : \mathbb{R}^+_0 \times \mathbb{R}^+_0 \to \mathbb{R}^+_0 \) is of class \( KL \) if it is continuous; for all \( t \in \mathbb{R}^+_0 \), \( \beta(.,t) \) is a \( K_\infty \) function; for all \( r \in \mathbb{R}^+_0 \), \( \beta(r,.) \) is strictly decreasing and \( \beta(r,t) \to 0 \) when \( t \to +\infty \).

II. PRELIMINARIES

A. Approximate bisimulation

We consider transition systems which will serve as a common modeling framework for symbolic, finite or infinite-dimensional dynamical systems.

Definition 1: A transition system is a tuple \( T = (X,U,S,X^0,Y,O) \) consisting of a set of states \( X \); a set of inputs \( U \); a transition map \( S : X \times U \to 2^X \); a set of initial states \( X^0 \subseteq X \); a set of outputs \( Y \); and an output map \( O : X \to Y \).

T is metric if the set of outputs \( Y \) is equipped with a metric. If the set of states \( X \) is a finite set (respectively a finite or infinite dimensional vector space), then \( T \) is...
called symbolic (respectively finite or infinite dimensional). The transition map captures the dynamics of the system: $x' \in S(x,u)$ means that the state of the system can evolve from $x$ to $x'$ under the action of input $u$. Input $u \in U$ belongs to the set of enabled inputs at state $x \in X$, denoted $\text{Enab}(x)$, if $S(x,u) \neq \emptyset$. $T$ is said to be deterministic if for all $x \in X$, for all $u \in \text{Enab}(x)$, $S(x,u)$ consists of a unique element. State $x \in X$ is said to be blocking if $\text{Enab}(x) = \emptyset$, otherwise it is said to be non-blocking. $T$ is said to be non-blocking if all states are non-blocking.

A state trajectory is a finite or infinite sequence of states and inputs of the form $(x^0, u^0), (x^1, u^1), \ldots$ where $x^0 \in X^0$ and $x^{i+1} \in S(x^i, u^i)$ for $i \geq 0$; the associated output trajectory is a sequence of outputs and inputs $(y^0, u^0), (y^1, u^1), \ldots$ where $y^i = O(x^i)$ for $i \geq 0$.

We want to quantify the distance between the output trajectories of two metric transition systems, possibly of different type (symbolic, finite or infinite dimensional). Let us consider two metric transition systems $T_j = (X_j, U, S_j, X_0^j, Y, O_j)$, $j \in \{1, 2\}$, with common sets of inputs and outputs equipped with metric $d$. The comparison of transition systems can be handled through the notion of approximate bisimulation [11].

**Definition 2:** Let $\varepsilon \geq 0$, a relation $R \subseteq X_1 \times X_2$ is called an $\varepsilon$-approximate bisimulation relation between $T_1$ and $T_2$, if for all $(x_1, x_2) \in R$:

1) $d(O_1(x_1), O_2(x_2)) \leq \varepsilon$,

2) $\forall u \in U, \forall x'_1 \in S_1(x_1, u), \exists x'_2 \in S_2(x_2, u)$ such that $(x'_1, x'_2) \in R$,

3) $\forall u \in U, \forall x'_2 \in S_2(x_2, u), \exists x'_1 \in S_1(x_1, u)$ such that $(x'_1, x'_2) \in R$.

$T_1$ and $T_2$ are approximately bisimilar with precision $\varepsilon$ (denoted $T_1 \sim_\varepsilon T_2$), if there exists $R$, an $\varepsilon$-approximate bisimulation relation between $T_1$ and $T_2$, such that for all $x_1 \in X_1^0$, there exists $x_2 \in X_2^0$ such that $(x_1, x_2) \in R$, and conversely.

In particular, approximate bisimulation guarantees that for each output trajectory of $T_1$ (respectively $T_2$), there exists an output trajectory of $T_2$ (respectively $T_1$) with the same sequence of inputs and with sequences of outputs whose distance is bounded by $\varepsilon$ (see [11]).

**B. Incrementally stable systems**

In this paper, we consider the following general notion of dynamical system:

**Definition 3:** A dynamical system is a tuple $\Sigma = (X, U, \Phi, X^0)$ where the state-space $X$ is a vector space equipped with a norm $\| \cdot \|_X$. $U \subseteq \mathbb{R}^p$ is the set of inputs, $X^0 \subseteq X$ is a set of initial states, and the flow $\Phi: \mathbb{R}_0^+ \times X \times C^0_{pu}(\mathbb{R}_0^+, U) \rightarrow X$ satisfying for all $x \in X$, $u \in C^0_{pu}(\mathbb{R}_0^+, U)$,

$$\Phi(0, x, u) = x; \quad (1)$$

$$\forall t, \tau \in \mathbb{R}_0^+, \Phi(t + \tau, x, u) = \Phi(t, \Phi(\tau, x, u), u_\tau) \quad (2)$$

where $u_\tau(t) = u(\tau + t)$, for all $t \in \mathbb{R}_0^+$.

$\Sigma$ is called finite dimensional, (respectively infinite dimensional), if the dimension of $X$ is finite (respectively infinite). The functions $\Phi(\cdot, x, u)$ are called the trajectories of $\Sigma$. For all $x \in X$, $u \in U$, we denote by $\Phi(\cdot, x, u)$ the trajectory $\Phi(t, x, u)$ where $u$ is the constant input $u$. A pair $(x, u) \in X \times U$, is an equilibrium of $\Sigma$ if for all $t \in \mathbb{R}_0^+$, $\Phi(t, x, u) = x$. The reachable set from $x \in X$ at time $t \in \mathbb{R}_0^+$ is the set

$$\text{Reach}(t, x) = \{ x' \in X | \exists u \in C^0_{pu}(\mathbb{R}_0^+, U), x' = \Phi(t, x, u) \}.$$ 

$\Sigma$ is said to have bounded reachable sets if for any $t \in \mathbb{R}_0^+$, $x \in X$, $\text{Reach}(t, x)$ is a bounded subset of $X$. Finally, $\Sigma$ is called linear if there exists $\overline{\Phi}: \mathbb{R}_0^+ \times X \times C^0_{pu}(\mathbb{R}_0^+, U) \rightarrow X$ such that $\Phi$ is the restriction of $\overline{\Phi}$ to $\mathbb{R}_0^+ \times X \times C^0_{pu}(\mathbb{R}_0^+, U)$ and for any $t \in \mathbb{R}_0^+$, $x \in X$, $u \in C^0_{pu}(\mathbb{R}_0^+, U)$, the applications $\overline{\Phi}(t, x, u)$ and $\overline{\Phi}(t, x, \cdot)$ are linear maps and

$$\overline{\Phi}(t, x, u) = \overline{\Phi}(t, x, 0) + d(t, x, u).$$

$\Phi$ is called linear extension of flow $\overline{\Phi}$.

We associate to a dynamical system $\Sigma$ a family of transition systems $T_\tau(\Sigma)$ parameterized by $\tau > 0$. Intuitively, $T_\tau(\Sigma)$ describes the transitions of $\Sigma$ where the input value can only change at multiples of time step $\tau$. In other words, $T_\tau(\Sigma)$ describes the sampled dynamics of $\Sigma$. Formally, $T_\tau(\Sigma) = (X, U, S, X^0, X, id_{\mathbb{R}})$ where the set of states is the state-space $X$, the set of inputs is $U$, the successor map is given by

$$x' \in S(x, u) \iff x' = \Phi(\tau, x, u),$$

the set of initial states is $X^0$, the set of outputs is the state-space $X$ and the observation map is the identity map over $X$. It is easy to see that $T_\tau(\Sigma)$ is deterministic and non-blocking, it is finite dimensional (respectively infinite dimensional) if $\Sigma$ is finite dimensional (respectively infinite dimensional). It is metric when $X$ is equipped with the metric associated to the norm $\| \cdot \|_X$.

In the following, we will deal with dynamical systems satisfying an incremental stability property [10]. Intuitively, a dynamical system is incrementally stable if two trajectories associated with the same input converge to each other independently of their initial states:

**Definition 4:** $\Sigma$ is said to be incrementally globally asymptotically stable (\(\delta\)-GAS) if there exists a $KL$ function $\beta$ such that for any $t \in \mathbb{R}_0^+$, $x_1, x_2 \in X$ and $u \in C^0_{pu}(\mathbb{R}_0^+, U)$ the following condition holds:

$$\| \Phi(t, x_1, u) - \Phi(t, x_2, u) \|_X \leq \beta(\| x_1 - x_2 \|, x, t).$$ \quad (3)$$

$\Sigma$ is said to be incrementally input-to-state stable (\(\delta\)-ISS) if there exists a $KL$ function $\beta$ and a $K_\infty$ function $\gamma$ such that for any $t \in \mathbb{R}_0^+$, $x_1, x_2 \in X$ and $u_1, u_2 \in C^0_{pu}(\mathbb{R}_0^+, U)$ the following condition holds:

$$\| \Phi(t, x_1, u_1) - \Phi(t, x_2, u_2) \|_X \leq \beta(\| x_1 - x_2 \|, x, t) + \gamma(\| u_1 - u_2 \|_\infty).$$ \quad (4)$$

It is clear that a $\delta$-ISS system is $\delta$-GAS. Moreover, the following property holds:
Proposition 1: Let us assume that $\Sigma$ is $\delta$-ISS, has an equilibrium, and that the set of inputs $U$ is bounded. Then $\Sigma$ has bounded reachable sets.

Proof: Let $(x_0, u_0)$ denote an equilibrium of $\Sigma$. Let $r_U$ be given by $r_U = \sup_{u \in U} \|u - u_0\|$. Since $U$ is bounded, $r_U$ is finite. Then, for all $t \in \mathbb{R}^+_0$, $x \in X$, $u \in C_{pw}(\mathbb{R}^+_0, U)$, it holds by (4)

$$\|\Phi(t, x, u) - x_0\|_X \leq \beta(\|x - x_0\|_X, t) + \gamma(r_U)$$

which implies that $\Sigma$ has bounded reachable sets. ■

C. Examples

In this section, we present two examples of infinite dimensional dynamical systems which can be considered within our framework. Stability results are adapted from [11] and [12]. Proofs are omitted due to space limitation.

1) Linear time-delay system: Let us consider the time-delay system $\Sigma$ given by:

$$\dot{x}(t) = A_0x(t) + A_1x(t - h) + Bu(t), \quad t \in \mathbb{R}^+_0$$

(5)

where $x(t) \in \mathbb{R}^n$, $u(t) \in U \subseteq \mathbb{R}^p$, the time delay $h > 0$, and $A_0$, $A_1$ and $B$ are matrices of compatible dimensions. The state of $\Sigma$ at time $t \in \mathbb{R}^+_0$ can be described by the function $x_t \in X = C^0([t - h, 0], \mathbb{R}^n)$, such that $x_t(s) = x(t + s)$, for all $s \in [-h, 0]$. Hence, $\Sigma$ is an infinite dimensional linear dynamical system. Moreover, we can give the following sufficient conditions for $\Sigma$ to be $\delta$-ISS.

Proposition 2: Let us assume that there exist $n \times n$ positive definite symmetric matrices $P$ and $Q$ and $\mu > 0$ such that the following inequality holds:

$$\begin{pmatrix} A_0^TP + PA_0 + Q + \mu P & PA_1 \\ A_1^TP & -Qe^{-\mu h} \end{pmatrix} \leq 0.$$ 

(6)

Then, $\Sigma$ is $\delta$-ISS when $X$ is equipped with the norm $\|x\|_X = \sqrt{V(x)}$ where

$$V(x_t) = x_t(0)^TPx_t(0) + \int_0^h x_t(s)^TQx_t(s)e^{\mu s}ds.$$ 

2) Linear hyperbolic distributed parameter system: Let us consider the hyperbolic distributed parameter system given by:

$$\partial_t x(t, s) + \Lambda \partial_s x(t, s) = Fx(t, s), \quad s \in [0, 1], \quad t \in \mathbb{R}^+_0$$

(7)

with the boundary conditions

$$x(t, 0) = Gx(t, 1) + Hu(t), \quad t \in \mathbb{R}^+_0$$

(8)

where $x(t) \in \mathbb{R}^n$, $u(t) \in U \subseteq \mathbb{R}^p$, $\Lambda$ is an $n \times n$ positive definite diagonal matrix and $F$, $G$ and $H$ are matrices of compatible dimensions. For $u \in C_{pw}(\mathbb{R}^+_0, U)$, there may not exist traditional solutions to (7) and (8); however one can define solutions in a weak sense (see [12]). Then, the state of $\Sigma$ at time $t \in \mathbb{R}^+_0$ can be described by the function $x(t, \cdot) \in X = L^2([0, 1], \mathbb{R}^n)$. Hence, $\Sigma$ is an infinite dimensional linear dynamical system. Moreover, we can give the following sufficient conditions for $\Sigma$ to be $\delta$-ISS.

Proposition 3: Let us assume that there exist a $n \times n$ positive definite diagonal matrix $D$ and $\mu \in \mathbb{R}$, $\nu > 0$ such that the following inequality holds:

$$F^TD + DF - \mu DA < 0,$$ 

(9)

$$G^TDA < e^{-\nu D}.$$ 

(10)

Then, $\Sigma$ is $\delta$-ISS when $X$ is equipped with the norm $\|x(t, \cdot)\|_X = \sqrt{V(x(t, \cdot))}$ where

$$V(x(t, \cdot)) = \int_0^1 x(t, s)^TDx(t, s)e^{-\nu s}ds.$$ 

III. FINITE DIMENSIONAL ABSTRACTIONS

In this section, we state our first result on the existence of finite dimensional abstractions for $T_\tau(\Sigma)$ when $\Sigma$ is $\delta$-GAS. We first try to give an intuition of the proposed construction. Intuitively, we can see from (3), that for long times, the state of a $\delta$-GAS system $\Sigma$ does not depend on its initial state but only on the input that has been applied to $\Sigma$. Similarly, it can be shown that only the most recent values of the input affect significantly the state of $\Sigma$. Hence, the main idea of our construction is to identify the state of $\Sigma$ at some time $t$ with the input signal that was applied on the time interval $[t - N\tau, t]$, where $N \in \mathbb{N}$ is some time horizon. Since we deal with sampled inputs, this signal can be represented by an element of $U^N \subseteq \mathbb{R}^{Np}$ which has finite dimension. A similar approach has been proposed for computing symbolic abstractions of incrementally stable switched systems in [9] where states are identified with mode sequences.

A. Construction of the abstraction

Let $N \in \mathbb{N}$, $N \geq 1$, let $x_{\cdot} \in X$, and let us define the following transition system $T_\tau(\Sigma) = (X_f, U, S_f, X_0, O_f)$ where the set of states is $X_f = U^N$, the set of inputs is $U$, the successor map is given for $v = (v_1, \ldots, v_N), v' = (v'_1, \ldots, v'_N) \in X_f$ and $u \in U$ by

$$v' \in S_f(v, u) \iff \begin{cases} v'_i = v_{i+1}, & i = 1, \ldots, N-1 \\
 v'_N = u, 
\end{cases}$$

(11)

the set of initial states is $X_0 \subseteq X_f$, the set of outputs is $X$ and the output map is given for all $v = (v_1, \ldots, v_N) \in X_f$ by

$$O_f(v) = \Phi(N\tau, x_S, u_v)$$

(12)

where $u_v$ is the input of $\Sigma$ given by

$$u_v(t) = v_i, \quad \forall t \in \left((i - 1)\tau, i\tau\right), \quad i = 1, \ldots, N.$$ 

(13)

The parameters of the abstraction $N$ and $x_{\cdot}$ are called time horizon and source state of the abstraction respectively. It is easy to see that $T_\tau(\Sigma)$ is deterministic, non-blocking and finite dimensional. It is metric when $X$ is equipped with the metric associated to the norm $\|\cdot\|_X$.

Theorem 1: Let us assume that $\Sigma$ is $\delta$-GAS and has bounded reachable sets. Then, let time step $\tau > 0$, time horizon $N \geq 1$ source state $x_S \in X$ and $\varepsilon > 0$ be such that

$$\beta(\varepsilon, \tau) + \beta(r(x_S), N\tau) \leq \varepsilon$$

(14)
Hence, \( r(x) = \sup_{u \in U} \| \Phi(\tau, x, u) - x \|_X \). \hspace{1cm} (15)

Then, the relation \( R^f \) given by

\[
R^f = \{ (x, v) \in X \times X_f \mid \| x - O_f(v) \|_X \leq \varepsilon \}
\]

is a \( \varepsilon \)-approximate bisimulation relation between \( T_f(\Sigma) \) and \( T_f^f(\Sigma) \).

**Proof:** Let us start by remarking that the fact that \( \Sigma \) has bounded reachable sets implies that \( r(x) \) is finite for any \( x \in X \). Let \( (x, v) \in R^f \), then \( \| O(x) - O_f(v) \|_X = \| x - O_f(v) \|_X \leq \varepsilon \). Hence, the first condition of Definition 2 holds. Let us remark that since \( T_f(\Sigma) \) and \( T_f^f(\Sigma) \) are deterministic, the second and third conditions of Definition 2 are equivalent. Let \( v = (v_1, \ldots, v_N) \), let \( u \in U \), and \( x' \in S_f(v, u) \). Then by definition of the transition and output maps, we have \( O(x') = \Phi(\tau, x, u) \) and \( O_f(v') = \Phi(N\tau, x, u, v) \). Let \( u_{v, u} \) denote the input given by \( u_{v, u} = v_t \) for \( t \in [(i - 1)\tau, i\tau], i = 1, \ldots, N \), and \( u_{v, u} = u \) for \( t \in [N\tau, (N + 1)\tau] \). Let us remark that \( u_{v, u}(t) = u_v(t) \) for \( t \in [0, N\tau) \) and \( u_{v, u}(t) = u_v(t - \tau) \) for \( t \in [N\tau, (N + 1)\tau) \). Then,

\[
\| O(x') - O_f(v') \|_X = \Phi(\tau, x, u) - \Phi(N\tau, x, u, v) \leq \| \Phi(N + 1)\tau, x, u, v \|_X + \| \Phi(\tau, x, u) - \Phi(N\tau, x, u, v) \|_X
\]

\[
\leq \| \Phi(\tau, x, u) - \Phi(\tau, x, u, v) \|_X + \beta(\| x - \Phi(\tau, x, u) \|_X, \tau)
\]

\[
\leq \beta(\| x - \Phi(\tau, x, u) \|_X, \tau) + \beta(\| x - \Phi(\tau, x, u) \|_X, \tau)
\]

\[
\leq \beta(v, \tau) + \beta(x, \tau) \leq \varepsilon.
\]

Hence, \( (x', v') \in R^f \). Hence, the second and third conditions of Definition 2 are satisfied and \( R^f \) is an \( \varepsilon \)-approximate bisimulation relation between \( T_f(\Sigma) \) and \( T_f^f(\Sigma) \). \( \Box \)

The previous result gives a relation between the parameters \( \tau, N \) and \( x_f \) and the precision of the approximate bisimulation between \( T_f(\Sigma) \) and \( T_f^f(\Sigma) \). Let \( \varepsilon > 0 \) be some desired precision. It appears from (14) that we should have \( \beta(\varepsilon, \tau) < \varepsilon \). This is always true for \( \tau \) sufficiently large; this may impose a lower bound on the time step \( \tau \). Also, there are cases (such as the examples presented in the previous section) where we have for all \( \tau > 0 \), \( \beta(\varepsilon, \tau) < \varepsilon \); in that case any positive value can be chosen for the time step \( \tau \). Once \( \varepsilon \) and \( \tau \) are chosen it remains to choose the time horizon \( N \geq 1 \) and the source state \( x_s \in X \). It appears from (14) that \( x_s \) should be chosen in order to minimize \( r(x) \) given by (15). If \( \beta(\varepsilon, \tau) < \varepsilon \), then one can always choose \( N \) such that (14) holds. Hence, we can see that any precision \( \varepsilon \) can be achieved for \( N \) sufficiently large. More precise abstractions hence have sets of higher dimension.

The following corollary provides conditions for choosing a suitable set of initial states \( X^0 \) for the abstraction, given the set of initial states \( X^0 \) of \( \Sigma \). It is a direct consequence of Theorem 1 and Definition 2.

**Corollary 1:** Under the assumptions of Theorem 1, let us assume that \( X^0 \subseteq O_f(X_f) \). Let \( X^0_f = O_f^{-1}(X^0) \), then \( T_f(\Sigma) \approx_{\varepsilon} T_f^f(\Sigma) \).

Note that the condition \( X^0 \subseteq O_f(X_f) \) is necessary for defining \( X^0_f = O_f^{-1}(X^0) \). This condition might not hold and we may not be able to choose \( X^0_f \) such that \( T_f(\Sigma) \approx_{\varepsilon} T_f^f(\Sigma) \). However, in that case, we can still quantify the distance between the trajectories of \( T_f(\Sigma) \) and \( T_f^f(\Sigma) \):

**Corollary 2:** Under the assumptions of Theorem 1, for all state trajectories of \( T_f(\Sigma), (x^0, u^0), (x^1, u^1), \ldots, \) and \( T_f^f(\Sigma), (v^0, w^0), (v^1, w^1), \ldots, \) with identical sequences of inputs, it holds for all \( i \geq 0 \):

\[
\| x^i - O_f(v^i) \|_X \leq \beta(\| x^0 - O_f(v^0) \|_X, i\tau) + \varepsilon.
\]

**Proof:** Let \( \bar{x}^0 = O_f(v^0) \) and consider the state trajectory of \( T_f(\Sigma), (\bar{x}^0, u^0), (\bar{x}^1, u^1), \ldots \). Then we have \( \| x^0 - O_f(v^0) \|_X \leq \varepsilon \) and \( (\bar{x}^i, v^i) \in R^f \). From Theorem 1, it follows that \( (\bar{x}^i, v^i) \in R^f \), for all \( i \geq 0 \). Then,

\[
\| x^i - O_f(v^i) \|_X \leq \| x^i - \bar{x}^i \|_X + \| \bar{x}^i - O_f(v^i) \|_X \leq \beta(\| x^0 - \bar{x}^0 \|_X, i\tau) + \varepsilon
\]

The corollary follows from \( \bar{x}^0 = O_f(v^0) \). \( \Box \)

**B. Case of linear systems**

In this section, we examine the specific case of linear systems. In particular, we show that if \( \Sigma \) is linear then the abstraction \( T_f(\Sigma) \) is actually a finite-dimensional linear system. First, let us remark that the transition relation (11) can be always (even if \( \Sigma \) is nonlinear) described by a linear system: for all \( v \in X_f, u \in U, v' \in S_f(v, u) \):

\[
v' = Av + Bu
\]

where the \( A \) and \( B \) are block matrices:

\[
A = \begin{bmatrix}
0 & I_p & \cdots & 0_p \\
\vdots & \ddots & \ddots & \vdots \\
0_p & \cdots & I_p & 0_p \\
\end{bmatrix},
B = \begin{bmatrix}
0_p \\
\vdots \\
0_p \\
\end{bmatrix}
\]

where \( 0_p \) and \( I_p \) denote the \( p \times p \) zero and identity matrices, respectively.

Assuming now that \( \Sigma \) is linear, let us consider the output map given by (12). Let \( v \in X_f, v = (v_1, \ldots, v_N) \), for \( i = 1, \ldots, N, j = 1, \ldots, p \) let \( v_{i,j} \) denote the \( j \)-th coordinate of \( v_i \in U \subseteq \mathbb{R}^p \). We can write

\[
u_v = \sum_{i=1}^N \sum_{j=1}^p v_{i,j} u_{i,j}
\]

where \( u_{i,j} : \mathbb{R}_0^+ \to \mathbb{R}^p \) is defined by

\[
u_{i,j,k}(t) = \begin{cases} 
1 & \text{if } k = j \text{ and } t \in [(i - 1)\tau, i\tau) \\
0 & \text{otherwise}
\end{cases}
\]
where $u_{i,j,k}$ is the $k$th coordinate of $u_{i,j}$. Let $\Phi$ be the linear extension of the flow $\Phi$, then, using the linearity of $\Phi$ the output map $O_f$ is given by
\[
O_f(v) = \Phi(N\tau, x_S, 0) + \sum_{i=1}^{N} \sum_{j=1}^{p} v_{i,j} \Phi(N\tau, 0, u_{i,j}). \tag{17}
\]
Hence, $O_f$ is an affine map with respect to the state $v$. In particular, we can see that it is completely determined by $p \times N + 1$ values of $\Phi$. As a consequence, the abstraction $T^f_\tau(\Sigma)$ can be computed by simulating $p \times N + 1$ trajectories of $\Sigma$.

Let us remark that if we choose the source state $x_S = 0$, then by linearity we have $\Phi(N\tau, x_S, 0) = 0$ and the abstraction $T^f_\tau(\Sigma)$ is in a finite-dimensional discrete-time linear system.

IV. SYMBOLIC ABSTRACTIONS

In this section, we state our second result on the existence of symbolic abstractions for $T_\tau(\Sigma)$. Let us remark that if $U$ is a finite subset of $\mathbb{R}^p$ (e.g. when $\Sigma$ is a quantized or switched system), then the abstractions $T^f_\tau(\Sigma)$ introduced in the previous section are already symbolic. If $U$ is infinite, then we can still prove the existence of symbolic abstractions under the assumption that $\Sigma$ is $\delta$-ISS. We first explain the main idea of the construction.

A. Construction of the abstraction

Let $N \in \mathbb{N}$, $N \geq 1$ and $x_S \in X$ be, in the previous section, the time horizon and source state of the abstraction. In addition, let $\eta > 0$ be the quantization resolution of the abstraction. Let $[U]_\eta \subseteq U$ be a finite set such that
\[
\forall u \in U, \exists u' \in [U]_\eta, \|u - u'\| \leq \eta.
\]
Let us remark that such a set always exists if $U$ is bounded.

Then, let us consider a quantizer $[.]_\eta : U \rightarrow [U]_\eta$ such that for all $u \in U$,
\[
\|u - [u]_\eta\| = \min_{u' \in [U]_\eta} \|u - u'\| \leq \eta.
\]

We now define the following transition system $T^*_\tau(\Sigma) = (X_S, U, S, X^0_S, X, O_S)$ where the set of states is $X_S = [U]_\eta$, the set of inputs is $U$, the successor map is given for $w = (w_1, \ldots, w_N)$, $w' = (w'_1, \ldots, w'_N) \in X_S$ and $u \in U$ by
\[
\begin{align*}
    w' \in S_u(v,u) & \iff \begin{cases} 
        w'_i = w_{i+1}, i = 1, \ldots, N-1 \\
        w'_N = [u]_\eta,
    \end{cases}
\end{align*} \tag{18}
\]
the set of initial states is $X^0_S \subseteq X$, the set of outputs is $X$ and the output map is given for all $w = (w_1, \ldots, w_N) \in X_S$ by
\[
O_s(w) = \Phi(N\tau, x_S, u_w) \tag{19}
\]
where the input $u_w$ is defined similarly to (13).

It is easy to see that $T^*_\tau(\Sigma)$ is deterministic, non-blocking and symbolic. It is metric when $X$ is equipped with the metric associated to the norm $\|\cdot\|_X$. The following Lemma shows the existence of an approximate bisimulation between the transitions systems $T^f_\tau(\Sigma)$, defined in the previous section, and $T^*_\tau(\Sigma)$.

**Lemma 1:** Let us assume that $\Sigma$ is $\delta$-ISS and that $U$ is bounded. Then, for all time step $\tau > 0$, time horizon $N \geq 1$, source state $x_S \in X$ and quantization resolution $\eta$, the relation $R^{fs}$ given by
\[
R^{fs} = \{ (v,w) \in X_f \times X_s \mid w_i = [v_i]_\eta, i = 1, \ldots, N \}
\]
is a $\gamma(\eta)$-approximate bisimulation relation between $T^f_\tau(\Sigma)$ and $T^*_\tau(\Sigma)$.

**Proof:** Let $(v,w) \in R^{fs}$, then $\|u_v - u_w\|_\infty \leq \eta$. By the $\delta$-ISS property, it holds:
\[
\begin{align*}
    \|O_f(v) - O_s(w)\|_X &= \|\Phi(N\tau, x_S, u_v) - \Phi(N\tau, x_S, u_w)\|_X \\
    &\leq \gamma(\|u_v - u_w\|_\infty) \leq \gamma(\eta).
\end{align*}
\]
Hence, the first condition of Definition 2 holds. Let us remark that since $T^f_\tau(\Sigma)$ and $T^*_\tau(\Sigma)$ are deterministic, the second and third conditions of Definition 2 are equivalent.

Let $u \in U$, and $w' \in S_u(v,u)$, $w' \in S_u(v,u)$, by definition of the transition relations $S_f$ and $S_s$, it is obvious that $(v',w') \in R^{fs}$. Hence, the second and third conditions of Definition 2 are satisfied and $R^{fs}$ is a $\gamma(\eta)$-approximate bisimulation relation between $T^f_\tau(\Sigma)$ and $T^*_\tau(\Sigma)$.

We can now state the main result of the section, which is an immediate consequence of Theorem 1, Lemma 1 and of transitivity of approximate bisimulation (see Proposition 4 in [1]):

**Theorem 2:** Let us assume that $\Sigma$ is $\delta$-ISS, has bounded reachable sets and that $U$ is bounded. Then, let time step $\tau > 0$, time horizon $N \geq 1$ source state $x_S \in X$ and $\varepsilon > 0$ be such that (14) holds. Then, for all quantization resolution $\eta$, the relation $R^{fs}$ given by
\[
\begin{align*}
    &R^{fs} = \{ (x,w) \in X \times X_s \mid \exists v \in X_f, (x,v) \in R^f \land (v,w) \in R^{fs} \}
\end{align*}
\]
is a $\gamma(\varepsilon + \gamma(\eta))$-approximate bisimulation relation between $T_\tau(\Sigma)$ and $T^*_\tau(\Sigma)$.

Referring to the discussion after Theorem 1, it appears that any precision can be achieved by choosing $N$ sufficiently large and $\eta$ sufficiently small. More precise symbolic abstractions hence have sets of states of higher cardinality.

The following corollary provides conditions for choosing a suitable set of initial states $X^0_S$ for the abstraction, given
the set of initial states $X^0$ of $\Sigma$. It is a direct consequence of Theorem 2 and Definition 2.

**Corollary 3:** Under the assumptions of Theorem 2, let us assume that $X^0 \subseteq O_f(X_f)$. Let $X_s = [O_f^{-1}(X^0)]_\eta$, where

$$[O_f^{-1}(X^0)]_\eta = \left\{ w \in X_s \mid \forall v \in O_f^{-1}(X^0), \ w_i = [v_i]_\eta, \ i = 1, \ldots, N \right\}.$$ 

Then, $T_\eta(\Sigma) \sim_{\varepsilon + \gamma(\eta)} T_\eta^{a}(\Sigma)$.

Similar to Corollary 1, the condition $X^0 \subseteq O_f(X_f)$ is necessary in order to be able to define a suitable set of initial states $X^0$. In other cases, similar to Corollary 2, we can still quantify the distance between the trajectories of $T_\eta(\Sigma)$ and $T_\eta^{a}(\Sigma)$:

**Corollary 4:** Under the assumptions of Theorem 2, for all state trajectories of $T_\eta(\Sigma)$, $(x^0, u^0, (x^1, u^1), \ldots$, and of $T_\eta^{a}(\Sigma)$, $(w^0, u^0), (w^1, u^1), \ldots$, with identical sequences of inputs, it holds for all $i \geq 0$:

$$\|x^i - O_s(w^i)\|_X \leq \beta(\|x^0 - O_s(w^0)\|_X, i\tau) + \varepsilon + \gamma(\eta).$$

**Proof:** Let $v^0 = w^0$ and let us consider the trajectory of $T_\eta^{a}(\Sigma)$, $(v^0, u^0), (v^1, u^1), \ldots$. By Lemma 1, it follows from $(v^0, u^0) \in R^{fs}$ that for all $i \geq 0$ $(v^i, u^i) \in R^{fs}$. Then, by Corollary 2, we have

$$\|x^i - O_s(w^i)\|_X \leq \|x^i - O_f(v^i)\|_X + \|O_f(v^i) - O_s(w^i)\|_X \leq \beta(\|x^0 - O_f(v^0)\|_X, i\tau) + \varepsilon + \gamma(\eta).$$

**B. Case of linear systems**

For the specific case of linear systems, we can obtain a tighter estimate of the precision using the expression of the output map given by (17). In the following let us consider the vectors $\psi_i \in \mathbb{R}^p$, $i = 1, \ldots, N$, given by

$$\psi_i = (\|\Phi(N\tau, 0, u_{i,1})\|_X, \ldots, \|\Phi(N\tau, 0, u_{i,p})\|_X).$$

Then, we can show the following result in place of Lemma 1:

**Lemma 2:** Let us assume that $\Sigma$ is linear and $U$ is bounded. Then, for all time step $\tau > 0$, time horizon $N \geq 1$, source state $x_S \in X$ and quantization resolution $\eta$, the relation $R^{fs}$ given by

$$R^{fs} = \{(v, w) \in X_f \times X_s \mid w_i = [v_i]_\eta, \ i = 1, \ldots, N\}$$

is an $\alpha \eta$-approximate bisimulation relation between $T_\eta^{a}(\Sigma)$ and $T_\eta^{a}(\Sigma)$ where

$$\alpha = \sum_{i=1}^{N} \|\psi_i\|.$$

**Proof:** Let $(v, w) \in R^{fs}$, then $\|v_i - w_i\| \leq \eta$, for $i = 1, \ldots, N$. By (17), we have

$$\|O_f(v) - O_s(w)\|_X \leq \sum_{i=1}^{N} \sum_{j=1}^{p} |v_{i,j} - w_{i,j}| \|\Phi(N\tau, 0, u_{i,j})\|_X \leq \sum_{i=1}^{N} \eta \|\psi_i\|.$$ 

Hence, the first condition of Definition 2 holds. The end of the proof follows that of Lemma 1.

Let us remark that for linear systems, we do not need to assume that $\Sigma$ is $\delta$-ISS. Similar to the general case, we can derive from the previous lemma, results similar to Theorem 2, Corollaries 3 and 4 for the specific case of linear systems.

V. CONCLUSIONS

In this paper, we have presented an approach for constructing approximately bisimilar abstractions of incrementally stable dynamical systems. Our approach, based on the input space, is quite generic and thus can be applied to a large variety of dynamical systems with finite or infinite dimensional state-spaces. We have considered two types of abstractions. Firstly, we have shown that $\delta$-GAS systems admit finite-dimensional abstractions; secondly, we have shown that $\delta$-ISS systems admit symbolic abstractions. The specific case of linear system has also been considered.

In future work, we will extend our approach to systems subject to disturbances and apply our approach to leverage finite dimensional and symbolic control techniques in order to control infinite dimensional systems.

REFERENCES


