Deterministic Finite-Automata Abstractions of Time-Variant Sequential Behaviours

Thomas Moor and Stefan Götz

Lehrstuhl für Regelungstechnik,
Friedrich-Alexander Universität Erlangen-Nürnberg, Germany,
e-mail: lrt@fau.de

Abstract: A *sequential behaviour* is a set of infinite length words. Following J.C. Willems *behavioural systems theory*, we interpret the behaviour as the relevant outcome when modelling a phenomenon w.r.t. the progress of discrete time. We propose the notion of an *experiment* as a particular form of inspecting a provided behaviour and ask for a strongest model that can be derived therefrom. The overall construct establishes a behavioural abstraction from the original model which, for finite signal ranges, can be realised as a deterministic finite automaton. The proposed method includes a flexible scheme for abstraction refinement that can be tailored to meet application requirements.

Keywords: Finite automaton abstractions, behavioural systems theory, hybrid systems.

INTRODUCTION

If for a synthesis or verification task at hand the provided system turns out too complex, one option is to resort to a less involved abstraction of the original system. For hybrid systems, there is a particular interest of abstractions with finite automata realisation, as they enable the application of well studied enumeration based methods for both controller synthesis as well as formal verification. Regarding synthesis, a common additional requirement is determinism in the sense that the state of the abstraction is uniquely defined by the past sequence of external symbols. Thus, deterministic abstractions can also be conceived as an alternative to computationally costly exact determinisation procedures for finite automata with large state sets.

A well established approach to obtain an abstraction is to utilise a simulation relation and to base the abstraction on the quotient state set. This is commonly referred to as *quotient based abstraction* and has been extensively studied for various classes of hybrid systems, see e.g. Alur et al. (2000); Pola et al. (2008) and the textbook Tabuada (2009). Provided that the ranges of the external signals are finite and that the quantisation is not considered a degree of freedom in the synthesis problem, an alternative approach are so called *behavioural abstractions* that are defined exclusively in terms of the external signals. Here, l-complete approximations, $l \in \mathbb{N}$, as originally proposed by Moor and Raisch (1999); Raisch and O'Young (1998), are realised on a state set that memorises the l most recent external symbols. A comparative study that elaborates the relationship between behavioural abstractions and quotient based abstractions is given in Schmuck et al. (2015).

In this paper, we further discuss a generalisation of l-complete approximations to memorise a non-uniform number of external symbols which was introduced by Moor et al. (2006). While the given reference is restricted to time-invariant systems and formally requires an exact reachability analysis, we now also address abstractions of time-variant systems that can be obtained by a conservative reachability analysis, and, hence, are applicable to more general classes of hybrid systems. From a practical perspective, we also complement Moor et al. (2006) by establishing an explicit automaton representation. We note that time variant systems are also addressed by *asynchronous l-complete approximations* Schmuck and Raisch (2014). In this regard, the present paper provides a uniform framework for the two independent generalisations of l-complete approximations.

This paper is organised as follows. In Section 1 we develop the notion of an *experiment on a behaviour* from which we obtain *a strongest model*, and we do so for time-variant and for time-invariant systems. This leads to a class of behavioural abstractions that can be refined by adjusting the experiment to the application at hand. In Section 2, we consider the situation where the original system is provided as a state machine. Technically, conducting an experiment then amounts to a recursion of a conservative one-step reachability operator. In Section 3, we establish a deterministic finite automaton realisation of the behavioural abstraction directly in terms of experiments.

NOTATION

Given a *signal space* W and considering discrete time, we denote $W^{\mathbb{N}_0} := \{w \,|\, w : \mathbb{N}_0 \to W\}$ the *universe of signals*. The *left-shift operator* σ^l, $l \in \mathbb{N}_0$, is defined for signals $w \in W^{\mathbb{N}_0}$ by $\sigma^l w \in W^{\mathbb{N}_0}$ with $(\sigma^l w)(k) := w(k + l)$ for all $k \in \mathbb{N}_0$, and we let $\sigma := \sigma^1$. For a signal $w \in W^{\mathbb{N}_0}$, the *restriction* to an integer interval $D \subseteq \mathbb{N}_0$ is denoted $w|_D$ with, e.g., $D = [k_1, k_2) := \{k \in \mathbb{N}_0 \,|\, k_1 \leq k < k_2\}$ and left-open and/or right-closed intervals defined likewise.

When taking finite restrictions we drop absolute time, i.e., we associate $w|_{[k_1,k_2)}$ with the *finite sequence* $\langle w(k_1), \ldots, w(k_2-1)\rangle \in W^l$ of length $l := k_2 - k_1 > 0$ for $k_1 < k_2$. We denote the *empty sequence* $\varepsilon \notin W$ to let $w|_\emptyset := \varepsilon$ and $W^0 := \{\varepsilon\}$. The set of all finite sequences is defined as $W^* := \cup\{W^l \,|\, l \in \mathbb{N}_0\}$. For a sequence $s \in W^l \subseteq W^*$, let $|s|$ denote its length l. For two finite sequences $s = \langle \omega_1, \ldots, \omega_l\rangle \in W^l$ and $r = \langle \varrho_1, \ldots, \varrho_n\rangle \in W^n$, the *concatenation* is defined by $\langle s, r\rangle := \langle \omega_1, \ldots, \omega_l, \varrho_1, \ldots, \varrho_n\rangle \in W^{l+n}$. For the empty sequence, let $\langle s, \varepsilon\rangle := s =: \langle \varepsilon, s\rangle$. The concatena-

tion of the finite sequence $s = \langle \omega_1, \ldots, \omega_l \rangle \in W^l$ with a signal $w \in W^{\mathbb{N}_0}$ is denoted $v = \langle s, w \rangle$, with $v(k) = \omega_{k+1}$ for $k < l$ and $v(k) = w(k - l)$ for $k \geq l$. Again, for the empty sequence let $\langle \varepsilon, w \rangle := w$.

A sequence $r \in W^*$ is a *prefix* of $s \in W^*$ if there exists $u \in W^*$ such that $\langle r, u \rangle = s$; we then write $r \leq s$. If, in addition, $r \neq s$. we say that r is a *strict prefix* of s and write $r < s$. Likewise, $r \in W^*$ is a prefix of a signal $w \in W^{\mathbb{N}_0}$, if there exits $v \in W^{\mathbb{N}_0}$ such that $\langle r, v \rangle = w$; we then write $r < w$. The set of all prefixes of a given sequence $s \in W^*$ or a given signal $w \in W^{\mathbb{N}_0}$ is denoted $\operatorname{pre} s \subseteq W^*$ or $\operatorname{pre} w \subseteq W^*$, respectively. A sequence $u \in W^*$ is a *suffix* of $s \in W^*$ if there exists $r \in W^*$ such that $\langle r, u \rangle = s$.

Taking point-wise images, all operators and set-valued maps in this paper are identified with their respective extension to set-valued arguments; e.g., $\langle S, W^{\mathbb{N}_0} \rangle = \{ \langle s, w \rangle \mid s \in S,\ w \in W^{\mathbb{N}_0} \}$ for $S \subseteq W^*$, and $\operatorname{pre} \mathfrak{B} = \cup \{ \operatorname{pre} w \mid w \in \mathfrak{B} \}$ for $\mathfrak{B} \subseteq W^{\mathbb{N}_0}$.

1. BEHAVIOURAL ABSTRACTIONS BY EXPERIMENTS

Following the terminology from Willems (1991), a *dynamical system* is a mathematical model of a phenomenon with a particular focus on variables that change their respective value over time. Formally, a dynamical system amounts to a mathematical expression, such as a conjunction of equations, that encodes on which trajectories the variables may evolve. We associate a dynamical system with the set of all trajectories that satisfy the expression and refer to this set as the *behaviour*. J.C. Willems' *behavioural systems theory* proposes to discuss and to categorise dynamical systems in terms of their behaviours. For the scope of this paper, we restrict considerations to the time axis $\mathbb{N}_0$ and use the following definition of a *sequential behaviour*.

Definition 1. Given some set W, referred to as the *signal space*, a *sequential behaviour* over W is a set $\mathfrak{B} \subseteq W^{\mathbb{N}_0}$. □

Prototypical representations of sequential behaviours are sampled data systems, finite automata or other variants of transition systems. Here, we think of the state variable as "internal" and use W as the range of transition labels. To this end, we only assume that the representation provided allows us to test whether or not $s \in \operatorname{pre} \mathfrak{B}$ for any specific finite sequence $s \in W^*$. We follow a method proposed by Moor et al. (2006) and examine the behaviour by suitably chosen tests to derive an alternative representation. The outcome of the examination is then referred to as an *experiment conducted on the behaviour*.

Definition 2. A *conservative experiment* on a behaviour $\mathfrak{B} \subseteq W^{\mathbb{N}_0}$ is a set $S \subseteq W^*$ of sequences such that $\mathfrak{B} \subseteq \langle S, W^{\mathbb{N}_0} \rangle$. [1] □

The technical condition $\mathfrak{B} \subseteq \langle S, W^{\mathbb{N}_0} \rangle$ requires the tests to be "exhaustive" in the sense that we identify at least some prefix $s < w$ of every possible trajectory $w \in \mathfrak{B}$; see Fig. 1 for an illustration.

1.1 Naive Abstractions

We consider the situation where we are provided an experiment S but not the original model $\mathfrak{B}$, and we ask what S reveals about $\mathfrak{B}$. While we can not expect to recover $\mathfrak{B}$ exactly, the condition $\mathfrak{B} \subseteq \langle S, W^{\mathbb{N}_0} \rangle$ enables us to obtain a *behavioural abstraction*, i.e., a model $\mathfrak{B}_S \subseteq W^{\mathbb{N}_0}$ that satisfies the inclusion $\mathfrak{B} \subseteq \mathfrak{B}_S$.

[1] In the original definition, as proposed by Moor et al. (2006), there is an additional requirement, namely that $S \subseteq \operatorname{pre} B$. Practically, this implies that the test on an individual sequence $s \in W^*$ for containment in $\operatorname{pre} \mathfrak{B}$ must not report "false positives". We drop the additional requirement for the present paper.

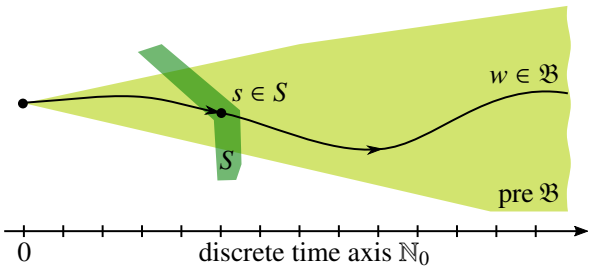

Fig. 1. Experiment $S \subseteq W^*$ on $\mathfrak{B} \subseteq W^{\mathbb{N}_0}$

Definition 3. Given an experiment $S \subseteq W^*$ on a behaviour over W, we say that $\mathfrak{B}_S \subseteq W^{\mathbb{N}_0}$ is a *model obtained from* S if

$$\forall\, \mathfrak{B}' \subseteq W^{\mathbb{N}_0} : \quad S \text{ is experiment on } \mathfrak{B}' \;\Rightarrow\; \mathfrak{B}' \subseteq \mathfrak{B}_S \,. \tag{1}$$

If, in addition, we have $\mathfrak{B}_S \subseteq \mathfrak{B}'_S$ for all models $\mathfrak{B}'_S$ obtained from S, then $\mathfrak{B}_S$ is a *strongest model obtained from* S. □

Given an experiment S on a behaviour $\mathfrak{B}$, the following proposition establishes the unique existence of the strongest model $\mathfrak{B}_S$ obtained from S. Thus, $\mathfrak{B}_S$ is the tightest behavioural abstraction of $\mathfrak{B}$ that can be stated exclusively in terms of S.

Proposition 4. Given an experiment $S \subseteq W^*$ on some behaviour over W, then

$$\mathfrak{B}_S := \{ w \in W^{\mathbb{N}_0} \mid \exists\, s \in S \,:\, s < w \} \tag{2}$$

is the unique strongest model obtained from S.

Proof. As a preliminary observation, note that strongest models obtained from S include each other. Hence, if a strongest model exists, it must be unique. Moreover, it is immediate from Eq. (2) that S is an experiment on $\mathfrak{B}_S$. Thus, we have $\mathfrak{B}_S \subseteq \mathfrak{B}'_S$ for any model $\mathfrak{B}'_S$ obtained from S. If $\mathfrak{B}_S$ is a model obtained from S, it must be the strongest such model. Now consider any behaviour $\mathfrak{B}'$ such that S is an experiment on $\mathfrak{B}'$ and pick an arbitrary $w \in \mathfrak{B}'$. By Definition 2, we can choose $s \in S$ such that $s < w$. This implies $w \in \mathfrak{B}_S$ and, hence, $\mathfrak{B}' \subseteq \mathfrak{B}_S$. Therefore, $\mathfrak{B}_S$ is indeed a model obtained from S. □

An intuitive interpretation of Eq. (2) is that $\mathfrak{B}_S$ tracks a sequence within $\operatorname{pre} S$ until it first reaches S and from then on allows for any arbitrary behaviour. For its simplicity, we also refer to $\mathfrak{B}_S$ as the *naive abstraction* of $\mathfrak{B}$ obtained from S.

We point out some technical consequences of Eq. (2). Provided an experiment $S \subseteq W^*$ consider $S' \subseteq S$ by removing all sequences from S that have a strict prefix within S:

$$S' := \{ s \in S \mid \forall\, t \in W^* \,:\, \langle s, t \rangle \in S \;\Rightarrow\; t = \varepsilon \} \,. \tag{3}$$

Then the naive abstraction $\mathfrak{B}_{S'}$ obtained from S' matches $\mathfrak{B}_S$ and, as long as the only concern is the naive abstraction, we may without loss of generality restrict considerations to *prefix-free* experiments, i.e experiments $S \in W^*$ that satisfy

$$\forall s, r \in S \,:\, s \leq r \;\Rightarrow\; s = r \,. \tag{4}$$

Another consequence observed from Eq. (2) is that $\varepsilon \in S$ implies $\mathfrak{B}_S = W^{\mathbb{N}_0}$, i.e., from a *trivial experiment* S, $\varepsilon \in S$, we obtain the *trivial abstraction* $\mathfrak{B}_S = W^{\mathbb{N}_0}$ as the strongest model. Hence, trivial experiments are of little practical value.

For a prefix-free experiment $S \subseteq W^*$ on $\mathfrak{B} \subseteq W^{\mathbb{N}_0}$ we can construct a *refinement* $S' \subseteq W^*$ by replacing an arbitrary finite

sequence $s \in S$ by all one-symbol extensions that pass the test for containment in pre $\mathfrak{B}$, i.e.,

$$S' := (S \setminus \{s\}) \cup \{\langle s, \omega\rangle \,|\, \omega \in W \text{ and } \langle s, \omega\rangle \in \operatorname{pre}\mathfrak{B}\}, \quad (5)$$

where, by Eq. (2), we obtain $\mathfrak{B}_{S'} \subseteq \mathfrak{B}_S$ for the respective naive abstractions. The construction of an experiment can be organised as successive refinements of the trivial experiment $S := \{\varepsilon\}$, where the refinements can either be uniform or guided by application specific requirements.

1.2 Abstractions under the assumption of time-invariance

The situation becomes more interesting when we assume that the original behaviour, on which the experiment is conducted, exhibits certain structural properties and if we then exploit these properties when recovering the strongest model. A natural candidate for a structural property here is time invariance and this matches the situation discussed in Moor et al. (2006).

Definition 5. A behaviour $\mathfrak{B} \subseteq W^{\mathbb{N}_0}$ is *time invariant* if $\sigma\mathfrak{B} \subseteq \mathfrak{B}$; see also Willems (1991). □

Time-invariant linear systems satisfy the above condition as well as automata with unrestricted initial state. Transition systems with non-trivial initial condition do in general not comply with the above notion of time invariance; this is followed up further below in this section. We adapt Definition 3 to obtain models from experiments on time-invariant behaviours.

Definition 6. Given an experiment $S \subseteq W^*$ on some time-invariant behaviour over W, we say that $\mathfrak{B}_S \subseteq W^{\mathbb{N}_0}$ is a *model obtained from S under the assumption of time-invariance* if

$$\forall\, \mathfrak{B}' \subseteq W^{\mathbb{N}_0} \;: \quad S \text{ is an exp. on } \mathfrak{B}' \text{ and } \sigma\mathfrak{B}' \subseteq \mathfrak{B}' \quad \Rightarrow \quad \mathfrak{B}' \subseteq \mathfrak{B}_S\,. \quad (6)$$

If, in addition, $\mathfrak{B}_S \subseteq \mathfrak{B}'_S$ holds for all likewise obtained models $\mathfrak{B}'_S$, we say that $\mathfrak{B}_S$ is a *strongest model obtained from S under the assumption of time invariance*. □

Again, the strongest model $\mathfrak{B}_S$ exists uniquely, referred to as the *abstraction obtained from S under the assumption of time invariance*, and we obtain the following characterisation.

Proposition 7. Given an experiment $S \subseteq W^*$ on some time-invariant behaviour over W, then

$$\mathfrak{B}_S := \{\, w \in W^{\mathbb{N}_0} \,|\, \forall\, k \in \mathbb{N}_0\, \exists\, l \in \mathbb{N}_0 :\; w|_{[k,k+l)} \in S \,\} \quad (7)$$

is the unique strongest model obtained from S under the assumption of time invariance.

Proof. As a preliminary observation, we note by Eq. (7) that $\sigma\mathfrak{B}_S \subseteq \mathfrak{B}_S$ and $\mathfrak{B}_S \subseteq \langle S, W^{\mathbb{N}_0}\rangle$. Hence, S is an experiment on the time invariant behaviour $\mathfrak{B}_S$. Thus, we have $\mathfrak{B}_S \subseteq \mathfrak{B}'_S$ for any model $\mathfrak{B}'_S$ obtained from S under the assumption of time-invariance. We now show that $\mathfrak{B}_S$ is a model obtained from S under the assumption of time invariance. Consider an arbitrary time-invariant behaviour $\mathfrak{B}'$ such that S is an experiment on $\mathfrak{B}'$, i.e, $\mathfrak{B}' \subseteq \langle S, W^{\mathbb{N}_0}\rangle$. Now pick arbitrary $w \in \mathfrak{B}'$ and $k \in \mathbb{N}_0$. By time invariance, we have that $\sigma^k w \in \sigma^k\mathfrak{B}' \subseteq \mathfrak{B}' \subseteq \langle S, W^{\mathbb{N}_0}\rangle$, and, hence, there exists $s \in S$ and $v \in W^{\mathbb{N}_0}$ such that $\sigma^k w = \langle s, v\rangle$, This implies $w|_{[k,l)} = s \in S$ for $l = |s|$. By the arbitrary choice of w and k, we conclude $\mathfrak{B}' \subseteq \mathfrak{B}_S$, and, hence, $\mathfrak{B}_S$ is indeed a model obtained from S under the assumption of time invariance. Together with the preliminary observation, this implies that $\mathfrak{B}_S$ is a strongest model. By definition, strongest models include each other and this implies uniqueness. □

The abstraction $\mathfrak{B}_S$ obtained from an experiment effectively tracks a sequence within pre S until it reaches S, but instead of allowing for any arbitrary behaviour thereafter, it will drop some prefix of the tracked sequence to continue tracking within pre S. This is illustrated in Figure 2 and we will discuss this mechanism in more detail in Section 3. To this end, we note from the characterisation Eq. (7) that $\mathfrak{B}_S$ itself is time invariant and that for the special case of $S = \mathfrak{B}|_{[0,l]}$ with $l \in \mathbb{N}_0$, the model $\mathfrak{B}_S$ matches the *strongest l-complete approximation* from Moor and Raisch (1999).

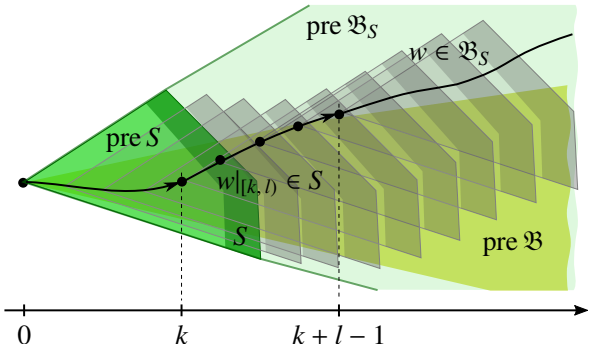

Fig. 2. Abstraction $\mathfrak{B}_S \subseteq W^{\mathbb{N}_0}$ obtained from $S \subseteq W^*$ under the assumption of time-invariance

As with the naive abstractions, the characterisation by Eq. (7) implies that we may without loss of generality assume experiments to be prefix-free, and that for trivial experiments $\varepsilon \in S$ we obtain the trivial abstraction $\mathfrak{B}_S = W^{\mathbb{N}_0}$. Also, for refinements by Eq. (5) we again obtain $\mathfrak{B}_{S'} \subseteq \mathfrak{B}_S$, i.e., refined experiments lead to potentially better abstractions.

Moreover, certain "false positives" of the underlying test can be identified and eliminated. More specifically, provided an experiment $S \subseteq W^*$ we may remove all sequences s that do not contribute to the model, i.e. $s \notin \operatorname{pre}\mathfrak{B}_S$, since we know by $\mathfrak{B} \subseteq \mathfrak{B}_S$ that such sequences can not occur within the prefix pre $\mathfrak{B}$ of the original behaviour. Technically, let

$$S' := S \,\cap\, (\operatorname{pre}\mathfrak{B}_S)\,, \quad (8)$$

to observe $\mathfrak{B}_{S'} = \mathfrak{B}_S$ for the abstraction $\mathfrak{B}_{S'}$ obtained from S'. As long as the derived abstractions are the only concern, one may without loss of generality restrict considerations to *trim experiments* in the sense of $S \subseteq \operatorname{pre}\mathfrak{B}_S$.

1.3 Abstractions of time-variant behaviours

In the case that the underlying phenomenon is considered time invariant except for restricted initial conditions, the naive abstraction from Section 1.1 and its more advanced time-invariant variation from Section 1.2 can be combined to abstract the start-up behaviour and the long-term behaviour individually. For the latter, we propose the following *time-invariant abstraction*.

Definition 8. Given $\mathfrak{B} \subseteq W^{\mathbb{N}_0}$, a *time-invariant abstraction* is a time-invariant behaviour $\mathfrak{B}'$ with $\mathfrak{B} \subseteq \mathfrak{B}' \subseteq W^{\mathbb{N}_0}$. A *strongest time-invariant abstraction* is a time-invariant abstraction $\mathfrak{B}_{\text{tia}}$ with $\mathfrak{B}_{\text{tia}} \subseteq \mathfrak{B}'$ for any time-invariant abstraction $\mathfrak{B}'$. □

The strongest time-invariant abstraction exists uniquely and is characterised by the following proposition.

Proposition 9. Given a behaviour $\mathfrak{B} \subseteq W^{\mathbb{N}_o}$, the strongest time-invariant abstraction $\mathfrak{B}_{tia}$ exists uniquely with

$$\mathfrak{B}_{tia} = \cup\{\, \sigma^k\mathfrak{B} \,|\, k \in \mathbb{N}_o \,\}. \tag{9}$$

Proof. With $k = 0$ we observe that $\mathfrak{B} \subseteq \mathfrak{B}_{tia}$. Moreover, $\sigma\mathfrak{B}_{tia} = \cup\{\, \sigma(\sigma^k\mathfrak{B}) \,|\, k \in \mathbb{N}_o \,\} = \cup\{\, \sigma^k\mathfrak{B} \,|\, k \in \mathbb{N} \,\} \subseteq \mathfrak{B}_{tia}$. Thus, $\mathfrak{B}_{tia}$ is a time-invariant abstraction of $\mathfrak{B}$. Let $\mathfrak{B}'$ denote any time-invariant abstraction of $\mathfrak{B}$ and pick an arbitrary $w \in \mathfrak{B}_{tia}$. Then $w \in \sigma^k\mathfrak{B}$ for some $k \in \mathbb{N}_o$, and, hence, $w \in \sigma^k\mathfrak{B}'$. By time-invariance of $\mathfrak{B}'$ we obtain $w \in \sigma^k\mathfrak{B}' \subseteq \mathfrak{B}'$. We conclude $\mathfrak{B}_{tia} \subseteq \mathfrak{B}'$ and have established that $\mathfrak{B}_{tia}$ is a strongest time-invariant abstraction. Uniqueness follows from strongest time-invariant abstractions to include each other. □

We conclude this section by proposing a behavioural abstraction $\mathfrak{B}_{abs}$ of a not-necessarily time-invariant behaviour $\mathfrak{B}$.

(A1) Conduct an experiment S_{sux} on $\mathfrak{B}$ to obtain the naive abstraction $\mathfrak{B}_{sua}$ to address the start-up behaviour.

(A2) Conduct an experiment S_{ltx} on the strongest time-invariant abstraction $\mathfrak{B}_{tia}$ of $\mathfrak{B}$ to obtain the abstraction $\mathfrak{B}_{lta}$ under the assumption of time-invariance to address the long-time behaviour.

(A3) Report $\mathfrak{B}_{abs} := \mathfrak{B}_{sua} \cap \mathfrak{B}_{lta}$ as abstraction of $\mathfrak{B}$.

By construction, we have $\mathfrak{B} \subseteq \mathfrak{B}_{tia} \subseteq \mathfrak{B}_{lta}$ and $\mathfrak{B} \subseteq \mathfrak{B}_{sua}$ and, hence, $\mathfrak{B}_{abs}$ indeed is a behavioural abstraction of $\mathfrak{B}$ in that it satisfies the inclusion $\mathfrak{B} \subseteq \mathfrak{B}_{abs}$. It can be seen that for the special case of $S_{sux} = \mathfrak{B}|_{[0,l]}$ and $S_{ltx} = \mathfrak{B}_{tia}|_{[0,l]}$, the behavioural abstraction $\mathfrak{B}_{abs}$ matches the *strongest asynchronous l-complete approximation* proposed in Schmuck and Raisch (2014).

2. EXPERIMENTS ON STATE MACHINES

We recall from Moor and Raisch (1999) how, when conducting an experiment on $\mathfrak{B} \subseteq W^{\mathbb{N}_o}$, the test for $s \in \operatorname{pre}\mathfrak{B}$ can be organised for the case that $\mathfrak{B}$ is realised by a *state machine*.

Definition 10. A *state machine* is a tuple $P = (X, W, \delta, X_o)$ with the (not necessarily finite) *state set* X, the *external signal space* W, the *transition relation* $\delta \subseteq X \times W \times X$ and the *initial states* X_o. With a state machine P, we associate the *full behaviour*

$$\mathfrak{B}_{full} := \{\, (w, x) \in W^{\mathbb{N}_o} \times X^{\mathbb{N}_o} \,|\, \forall\, k \in \mathbb{N}_o : \\ x(0) \in X_o \text{ and } (x(k), w(k), x(k+1)) \in \delta \,\}, \tag{10}$$

and we say that P realises the *external behaviour*

$$\mathfrak{B}_{ex} := \{\, w \in W^{\mathbb{N}_o} \,|\, \exists\, x \in X^{\mathbb{N}_o} :\ (w, x) \in \mathfrak{B}_{full} \,\}. \tag{11}$$

We write $Q \cong \mathfrak{B}$ to indicate that a state machine Q with external signal space W realises an external behaviour $\mathfrak{B}$ over W. □

For a state machine $P = (X, W, \delta, X_o)$, a state $\xi \in X$ can only contribute to the full behaviour $\mathfrak{B}_{full}$ if it can be reached from an initial state and if it can be continued to an infinite-time trajectory, i.e., if there exists $(w, x) \in \mathfrak{B}_{full}$ and $k \in \mathbb{N}_o$ such that $x(k) = \xi$. If every state exhibits this property, P is said to be *trim*. In this section, we focus attention on trim state machines.

For a state machine $P \cong \mathfrak{B}_{ex}$ with full behaviour $\mathfrak{B}_{full}$ and a finite external sequence $s \in W^*$, the *set of compatible states* is defined

$$\mathcal{X}_s := \{\, \xi \in X \,|\, \exists\, (w, x) \in \mathfrak{B}_{full} :\ w|_{[0,|s|)} = s,\ \xi = x(|s|) \,\}, \tag{12}$$

i.e., the set of states in which the state machine can possibly reside after the occurrence of s. In particular, for any state $\xi \in \mathcal{X}_s$, there exists $(w, x) \in \mathfrak{B}_{full}$ with $s < w$, and, thus, $s \in \operatorname{pre}\mathfrak{B}_{ex}$. Vice versa, if $s \in \operatorname{pre}\mathfrak{B}_{ex}$, we choose $v \in W^{\mathbb{N}_o}$ such that $w := \langle s, v\rangle \in \mathfrak{B}_{ex}$, and, hence, there exists $x \in X^{\mathbb{N}_o}$ with $(w, x) \in \mathfrak{B}_{full}$. This implies $x(|s|) \in \mathcal{X}_s$, and we conclude that $s \in \operatorname{pre}\mathfrak{B}_{ex}$ if and only if $\mathcal{X}_s \neq \emptyset$.

Provided that P is trim, sets of compatible states can be computed by a recursion with the *one-step forward reachability operator* $\operatorname{post}_{(\cdot)}$:

$$\mathcal{X}_{\langle s,\omega\rangle} = \operatorname{post}_\omega \mathcal{X}_s := \{\, \xi' \,|\, \exists\, \xi \in \mathcal{X}_s :\ (\xi, \omega, \xi') \in \delta \,\}, \tag{13}$$

where $s \in W^*$, $\omega \in W$, and $\mathcal{X}_\varepsilon = X_o$. If P is not trim, we can still run the above iteration, however, due to blocking states we may then obtain "false positives" if we test for $s \in \operatorname{pre}\mathfrak{B}_{ex}$ by $\mathcal{X}_s \neq \emptyset$.

When we have $X_o = X$ for a state machine $P = (X, W, \delta, X_o)$, then the external behaviour $\mathfrak{B}_{ex}$, $P \cong \mathfrak{B}_{ex}$, is time invariant. If, on the other hand, $X_o \neq X$, $\mathfrak{B}_{ex}$ may fail to be time invariant. In this case, Section 1 proposes to perform experiments also on the time-invariant abstraction $\mathfrak{B}_{tia} = \cup\{\, \sigma^k\mathfrak{B}_{ex} \,|\, k \in \mathbb{N}_o \,\}$. Now consider $v \in \sigma^k\mathfrak{B}_{ex}$ for some $k \in \mathbb{N}_o$. Then $v = \sigma^k w$ for some $w \in \mathfrak{B}_{ex}$ and we can choose x such that $(w, x) \in \mathfrak{B}_{full}$. Here, we observe that $\sigma^l(w, x)$ with any $l \in \mathbb{N}_o$ is in the full behaviour of $P_{tia} := (X, W, \delta, X)$. In particular $v = \sigma^k w$ is in the external behaviour of P_{tia}. Vice versa, consider a signal w in the external behaviour of P_{tia} and denote $x \in X^{\mathbb{N}_o}$ a corresponding state trajectory. Provided that P is trim, there exists $(v, z) \in \mathfrak{B}_{full}$, such that $z(l) = x(0)$ for some $l \in \mathbb{N}_o$. This implies that $(\langle v|_{[0,l)}, w\rangle, \langle z|_{[0,l)}, x\rangle) \in \mathfrak{B}_{full}$, and, hence, $w = \sigma^l\langle v|_{[0,l)}, w\rangle \in \sigma^l\mathfrak{B}_{ex} \subseteq \mathfrak{B}_{tia}$. We conclude that $P_{tia} \cong \mathfrak{B}_{tia}$. In other words, a realisation P_{tia} of the strongest time-invariant abstraction $\mathfrak{B}_{tia}$ can be obtained by dropping the initial condition. Therefore, experiments on the time-invariant abstraction of an external behaviour $\mathfrak{B}_{ex}$ realised by P amount to the same recursion Eq. (13), however, now initialised by $\mathcal{X}_\varepsilon = X$.

As a special case, consider $W = U \times Y$ with U the *input range* and Y the *output range*, a *set-valued transition function* $f : X \times U \rightsquigarrow X$ and a *set-valued output function* $g : X \rightsquigarrow Y$. To obtain a transition relation, we let $(\xi, (\mu, \nu), \xi') \in \delta \subseteq X \times W \times X$ if and only if $\xi' \in f(\xi, \mu)$ and $\nu \in g(\xi')$. The resulting state machine $P = (X, W, \delta, X_o)$ is referred to as *input-output state machine* and the recursion Eq. (13) can be rewritten as

$$\mathcal{X}_{\langle s,(\mu,\nu)\rangle} = \operatorname{post}_{(\mu,\nu)} \mathcal{X}_s = f(\mathcal{X}_s, \mu) \cap g^{-1}(\nu), \tag{14}$$

where $s \in W^*$, $\mu \in U$, $\nu \in Y$, and $g^{-1}(\nu) := \{\, \xi' \,|\, \nu \in g(\xi') \,\}$. This class of transition systems can be used to represent hybrid dynamics with a discrete-event external interface, i.e., $X = V \times \mathbb{R}^n$ and V, U and Y being finite sets. Note that this setting accounts for physical time obtained by a regular sampling period as well as logic time obtained by triggering events via thresholds or mode invariants and guard regions. For either case, reachability has been extensively studied and the literature provides effective procedures for the evaluation of Eq. (14), including exact evaluation for restricted classes of continuous dynamics, e.g. Alur et al. (1996, 2000); Lafferriere et al. (2000), as well as safe over-approximations for richer classes of continuous dynamics, e.g., Althoff et al. (2010); Chutinan and Krogh (1998); Frehse (2008); Henzinger et al. (2000); Maler and Dang (1998); Mitchell et al. (2005); Reissig (2011). For finite-state abstractions based on experiments, we may via Eq. (14) utilise

any exact or safe over-approximation method as an underlying computational procedure.

3. REALISATIONS OF STRONGEST MODELS

Provided that the signal space W is finite and provided that an experiment $S \subseteq W^*$ is *bounded-in-time* by an integer maximum length over all sequences, then S is a finite set. In this case, a deterministic finite-automaton realisation of the strongest model $\mathfrak{B}_S$ can be constructed directly from S.

3.1 Realisations of naive abstractions

Given a prefix-free experiment $S \subseteq W^*$, we first seek for a realisation of the naive abstraction $\mathfrak{B}_S$ obtained from S, as characterised by Eq. (2). For our candidate $Q = (Z, W, \eta, Z_o)$, we choose the state set $Z = \operatorname{pre} S$ and organise transitions such that the state records past values of the external signal until it passes S. More precisely, we define η as the set of all transitions $(\zeta, \omega, \zeta') \in Z \times W \times Z$ that satisfy

$$\zeta' = \begin{cases} \zeta\omega & \text{if } \zeta \in \operatorname{pre} S \text{ and } \zeta \notin S\,, \\ \zeta & \text{if } \zeta \in S\,, \end{cases} \tag{15}$$

and we initialise the record with $Z_o := \{\varepsilon\}$, i.e., no past signal values so far. Note that the proposed candidate is *deterministic*, i.e., it exhibits one initial state and for every state ζ and every external symbol ω, there is at most one successor state ζ' with $(\zeta, \omega, \zeta') \in \eta$.

Lemma 11. For a prefix-free and bounded-in-time experiment S over W with the strongest model $\mathfrak{B}_S$ obtained from S, consider $Q = (Z, W, \eta, Z_o)$ with $Z = \operatorname{pre} S$, η as in Eq. (15) and $Z_o = \{\varepsilon\}$. Then we have $Q \cong \mathfrak{B}_S$.

Proof. Let $\mathfrak{B}_{ex}$ and $\mathfrak{B}_{full}$ denote the external behaviour and the full behaviour associated with Q, respectively. Pick an arbitrary $w \in \mathfrak{B}_{ex}$ and choose $z \in Z^{\mathbb{N}_o}$ such that $(w, z) \in \mathfrak{B}_{full}$. By construction, we have $z(0) = \varepsilon$ and $z(k+1) = \langle w(0), \ldots, w(k)\rangle$ for all k with $z(k) \in (\operatorname{pre} S) \setminus S$. Since the length of the sequences of S is bounded, this implies the existence of l such that $s =: z(l) \in S$. Observe that $s < w$, and, hence, $w \in \mathfrak{B}_S$ by Eq. (2). To this end, we conclude that $\mathfrak{B}_{ex} \subseteq \mathfrak{B}_S$. Now pick an arbitrary $w \in \mathfrak{B}_S$ to establish the converse inclusion. By Eq. (2), we pick $s < w$ such that $s \in S$, and let $l := |s|$. Consider the candidate state trajectory z defined by $z(0) := \varepsilon$, $z(k) := \langle w(0), \ldots, w(k-1)\rangle$ for all $k \in \mathbb{N}_o$, $0 < k \leq l$, and $z(k) := z(l)$ for all $k \in \mathbb{N}_o$, $l < k$. Note that the range of z, by construction, consists of all prefixes of s and we have indeed $z \in Z^{\mathbb{N}_o}$. Moreover, for $k < l$, we have that $z(k)$ is a strict prefix of s, and, since S is prefix-free, we have that $z(k) \notin S$ for all $k < l$. By Eq (15), first case, this amounts to $(z(k), w(k), z(k+1)) \in \eta$ for all $k \in \mathbb{N}_o$, $0 \leq k < l$. For $k \in \mathbb{N}_o$, $k \geq l$, we have by definition $z(k) = z(l) = s$ and we refer to Eq (15), second case, in order to conclude $(z(k), w(k), z(k+1)) \in \eta$. We therefore have that $(w, z) \in \mathfrak{B}_{full}$, and, hence, $w \in \mathfrak{B}_{ex}$. This concludes the proof of $\mathfrak{B}_S \subseteq \mathfrak{B}_{ex}$. □

3.2 Realisations of abstractions assuming time-invariance

We now turn to the abstraction $\mathfrak{B}_S$ under the assumption of time-invariance obtained from a prefix-free experiment $S \subseteq W^*$. Here, we additionally assume that S is non-trivial, i.e., $\varepsilon \notin S$; see Eq. (7). Our candidate $Q = (Z, W, \eta, Z_o)$ to realise $\mathfrak{B}_S$ is defined by the same state space $Z = \operatorname{pre} S$ as for the naive abstraction, however, we now define η as the set of all transitions $(\zeta, \omega, \zeta') \in Z \times W \times Z$ that satisfy

$$\zeta' = \langle \zeta^\triangleleft, \omega\rangle \text{ for the longest suffix } \zeta^\triangleleft \text{ of } \zeta \text{ with } \zeta^\triangleleft \notin S. \tag{16}$$

Indeed, if $\zeta \notin S$ the longest qualifying suffix $\zeta^\triangleleft \notin S$ is identified $\zeta^\triangleleft = \zeta$, and Eq. (16) effectively collapses to the first case in Eq. (15) and encodes the state to record one more external symbol ω. If, on the other hand, $\zeta^\triangleleft \in S$, the transition relation encodes that first a minimum number of symbols are dropped from ζ to become a strict prefix $\zeta^\triangleleft$ in S, and that then appending ω complies with S in the sense of $\zeta' = \langle \zeta^\triangleleft, \omega\rangle \in \operatorname{pre} S$. In particular, our candidate Q is again deterministic. We establish that Q in fact realises $\mathfrak{B}_S$ by two technical propositions.

Proposition 12. For a non-trivial and prefix-free experiment S over W and the strongest model $\mathfrak{B}_S$ obtained from S under the assumption of time-invariance, let $Q = (Z, W, \eta, Z_o)$ with $Z = \operatorname{pre} S$, η as in Eq. (16) and $Z_o = \{\varepsilon\}$. Then the external behaviour $\mathfrak{B}_{ex}$ satisfies $\mathfrak{B}_S \subseteq \mathfrak{B}_{ex}$.

Proof. We pick an arbitrary $w \in \mathfrak{B}_S$ and refer to Eq. (7) to make two preliminary observations. First, given $k \in \mathbb{N}_o$, we can choose $H(k) \in \mathbb{N}_o$ such that

$$w|_{[k,\,H(k))} \in S\,; \tag{17}$$

since S is prefix-free, the choice is unique. Second, given k let

$$L(k) := \min\{\, t \in \mathbb{N}_o \mid H(t) \geq k \,\}\,; \tag{18}$$

since k itself is in the former set, it is non-empty and the minimum is well defined. Note also, that $L(\,\cdot\,)$ is monotone. We are now in the position to define our candidate state trajectory $z \in Z^{\mathbb{N}_o}$ by $z(k) := w|_{[L(k),\,k)}$ for all $k \in \mathbb{N}_o$; see Fig. 3 for the construction so far.

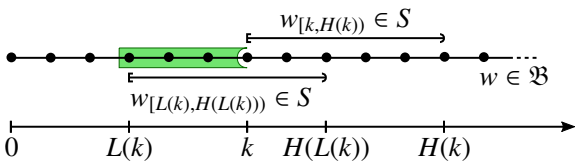

Fig. 3. Candidate state $z(k) := w|_{[L(k),\,k)}$ (green colour)

Observe that we have indeed $z(k) \in Z = \operatorname{pre} S$ for all $k \in \mathbb{N}_o$, and, for $k = 0$, we also have $z(0) = \varepsilon \in Z_o$. To establish $(z(k), w(k), z(k+1)) \in \eta$, we distinguish two cases. For case (a), we assume that $H(L(k)) > k$. This implies by Eq. (18) that $L(k+1) = L(k)$ and, hence, $z(k+1) = w|_{[L(k),\,k+1)} = \langle z(k), w(k)\rangle$. Also by the case hypothesis, we have $z(k) \notin S$ and, hence, $z(k)$ is the longest prefix of itself with $z(k) \notin S$. This demonstrates that $(z(k), w(k), z(k+1)) \in \eta$ and we proceed with case (b). Here, we formally assume that $H(L(k)) \leq k$, which by Eq. (11) collapses to $H(L(k)) = k$. From this we conclude $L(k+1) > L(k)$ and $z(k) = w|_{[L(k),\,k)} = w|_{[L(k),H(L(k)))} \in S$. Then $z(k+1) = w|_{[L(k+1),\,k+1)} = \langle w|_{[L(k+1),\,k)}, w(k+1)\rangle$ with $w|_{[L(k+1),\,k)}$ a strict suffix of $z(k)$, and, hence $w|_{(L(k+1),\,k)} \notin S$. If there was a longer suffix $w|_{(t,\,k)}$ of $z(k)$ with $w|_{(t,\,k)} \notin S$ this would contradict minimality of $L(k+1)$. Thus, $w|_{(L(k+1),\,k)}$ is the longest qualifying suffix of $z(k)$ in Eq. (16) and we again obtain $(z(k), w(k), z(k+1)) \in \eta$. This concludes both cases (a) and (b). Therefore, (w, z) is in the full behaviour of Q and we finally obtain $w \in \mathfrak{B}_{ex}$. □

Proposition 13. For a non-trivial, prefix-free, trim and bounded-in-time experiment S over W and the strongest model $\mathfrak{B}_S$ obtained from S under the assumption of time-invariance, let

$Q = (Z, W, \eta, Z_o)$ with $Z = \operatorname{pre} S$, η as in Eq. (16) and $Z_o = \{\varepsilon\}$. Then the external behaviour $\mathfrak{B}_{ex}$ satisfies $\mathfrak{B}_{ex} \subseteq \mathfrak{B}_S$.

Proof. Let $\mathfrak{B}_{full}$ denote the full behaviour of Q, pick an arbitrary $w \in \mathfrak{B}_{ex}$ and choose z such that $(w, z) \in \mathfrak{B}_{full}$. Inspecting the construct in Eq. (16), at any time k the state $z(k) \in Z = \operatorname{pre} S$ matches consecutive external symbols from the strict past of k, i.e., we may write $z(k) = w|_{[L(k),k)}$ with $L(k) \leq k$. Then there exists a unique time $H(k)$, $L(k) \leq k \leq H(k)$, such that

$$z(H(k)) = w|_{[L(k),H(k))} \in S \ . \tag{19}$$

To establish $w \in \mathfrak{B}_S$, we fix an arbitrary $k \in \mathbb{N}_o$ and show by a repeated argument that there exists $l \in \mathbb{N}_o$ such that $w|_{[k,k+l)} \in S$. For this purpose, consider any $k' \geq k$ with $L(k') \leq k$, where we know that such k' exists by the witness $k = k'$. Referring to our notation introduced above, we obtain $s = z(H(k')) = w|_{[L(k'),H(k'))} \in S$. Since S is trim, we can extend to $v \in \mathfrak{B}_S$, $s < v$. Since $\mathfrak{B}_S$ is time-invariant, we have that $\sigma^{k-L(k')} v \in \sigma^{k-L(k')} \mathfrak{B}_S \subseteq \mathfrak{B}_S$. As an immediate consequence of Eq. (7), we can uniquely choose a prefix $r < \sigma^{k-L(k')} v$ with $r \in S$; see Fig. 4 for an illustration of the construction of r.

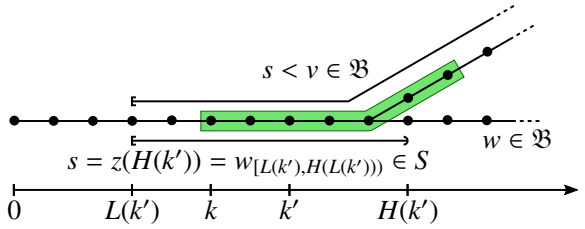

Fig. 4. Construction of $r \in S$ (green colour), case (b)

The first $H(k') - k$ symbols of $\sigma^{k-L(k')} v$ are $w|_{[k,H(k'))}$. We distinguish two cases. For case (a), we assume that $|r| \leq H(k') - k$. Then we can use $l := H(k') - k$ and obtain $w|_{[k,k+l)} = r \in S$ as required. For case (b), we must have $|r| > H(k') - k$. Here, we observe that $w|_{[k,H(k'))} \in (\operatorname{pre} S) \setminus S$. Referring to Eq. (16), $z(H(k') + 1)$ is constructed form the longest suffix of $z(H(k'))$ not in S and by appending $w(H(k'))$ to that suffix. We write $k'' := H(k') + 1$ and $z(k'') = w|_{[L(k''),k'')}$ to observe that $L(k'') \leq k$ for the maximal choice of the suffix. We then substitute k' by k'' and repeat the argument. Now assume that the argument went on indefinitely by branching into case (b). Then $H(k') - L(k')$ would grow arbitrarily large to contradict the boundedness of S. Hence, the argument terminates with case (a). □

Our main result on the realisation of models obtained under the assumption of time invariance is an immediate consequence of Propositions 12 and 13.

Lemma 14. For a non-trivial, prefix-free, trim and bounded-in-time experiment S over W the strongest model $\mathfrak{B}_S$ obtained from S under the assumption of time-invariance is realised by $Q = (Z, W, \eta, Z_o)$ with $Z = \operatorname{pre} S$, η as in Eq. (16) and $Z_o = \{\varepsilon\}$, i.e., we have $Q \cong \mathfrak{B}_S$. □

CONCLUSION

We have revisited a fairly general scheme of behavioural abstractions based on experiments, originally proposed in Moor et al. (2006) to address supervisory controller synthesis for hybrid systems with a prescribed discrete-event external interface. Our present extension addresses time-variant behaviours and accounts for a conservative reachability analysis in the underlying computational procedures, i.e., tolerates "false positives" when testing whether a prescribed finite sequence is within the prefix of the given behaviour. We also complement Moor et al. (2006) in providing a detailed construction of a deterministic finite automaton realisation for the abstraction.

REFERENCES

M. Althoff, O. Stursberg, and M. Buss. Computing reachable sets of hybrid systems using a combination of zonotopes and polytopes. *Nonlinear Analysis: Hybrid Systems*, 4:233–249, 2010.

R. Alur, T.A. Henzinger, and P.-H. Ho. Automatic symbolic verification of embedded systems. *IEEE Trans. Softw. Eng.*, 22:181–201, 1996.

R. Alur, T. A. Henzinger, G. Lafferriere, and G. Pappas. Discrete abstractions of hybrid systems. *Proceedings of the IEEE*, 88:971–984, July 2000.

A. Chutinan and B. H. Krogh. Computing polyhedral approximations to flow pipes for dynamic systems. In *IEEE 37th International Conference on Decision and Control*, 1998.

G. Frehse. PHAVer: algorithmic verification of hybrid systems past HyTech. *International Journal on Software Tools for Technology Transfer*, 10:263–279, 2008.

T. A. Henzinger, B. Horowitz, R. Majumdar, and H. Wong-Toi. Beyond HyTech: Hybrid systems analysis using interval numerical methods. In *Hybrid Systems: Computation and Control*, LNCS 1790, 2000.

G. Lafferriere, G. J. Pappas, and S. Sastry. O-minimal hybrid systems. *Mathematics of Control, Signals and Systems*, 13: 1–21, 2000.

O. Maler and T. Dang. Reachability analysis via face lifting. In *Hybrid Systems: Computation and Control*, LNCS 1386, pages 96–109, 1998.

I. M. Mitchell, A. M. Bayen, and C. J. Tomlin. A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games. *IEEE Trans. Autom. Control*, 50: 947–957, 2005.

T. Moor and J. Raisch. Supervisory control of hybrid systems within a behavioural framework. *Systems and Control Letters*, 38:157–166, 1999.

T. Moor, J. M. Davoren, and J. Raisch. Learning by doing: systematic abstraction refinement for hybrid control synthesis. *IEE Proceedings – Control Theory and Applications*, pages 591–599, 2006.

G. Pola, A. Girard, and P. Tabuada. Approximately bisimilar symbolic models for nonlinear control systems. *Automatica*, 44(10):2508–2516, 2008.

J. Raisch and S. D. O'Young. Discrete approximation and supervisory control of continuous systems. *IEEE Trans. Autom. Control*, 43:569–573, 1998.

G. Reissig. Computing abstractions of nonlinear systems. *IEEE Trans. Autom. Control*, 56(11):2583–2598, 2011.

A.-K. Schmuck and J. Raisch. Asynchronous l-complete approximations. *Systems and Control Letters*, 73:67–75, 2014.

A.-K. Schmuck, P. Tabuada, and J. Raisch. Comparing asynchronous l-complete approximations and quotient based abstractions. In *54th IEEE Conference on Decision and Control*, pages 6823–6829, 2015.

P. Tabuada. *Verification and Control of Hybrid Systems: A Symbolic Approach*. Springer-Verlag, 2009.

J. C. Willems. Paradigms and puzzles in the theory of dynamic systems. *IEEE Trans. Autom. Control*, 36:258–294, 1991.