Data-driven Abstractions for Verification of Linear Systems

Coppola, Rudi; Peruffo, Andrea; Mazo, Manuel

DOI
10.1109/LCSYS.2023.3288731

Publication date
2023

Document Version
Final published version

Published in
IEEE Control Systems Letters

Citation (APA)
Coppola, R., Peruffo, A., & Mazo, M. (2023). Data-driven Abstractions for Verification of Linear Systems. *IEEE Control Systems Letters*, *7*, 2737-2742. https://doi.org/10.1109/LCSYS.2023.3288731

Data-Driven Abstractions for Verification of Linear Systems

Rudi Coppola, *Graduate Student Member, IEEE*, Andrea Peruffo, and Manuel Mazo Jr., *Fellow, IEEE*

***Abstract*—We introduce a novel approach for the construction of symbolic abstractions - simpler, finite-state models - which mimic the behaviour of a system of interest, and are commonly utilized to verify complex logic specifications. Such abstractions require an exhaustive knowledge of the concrete model, which can be difficult to obtain in real-world applications. To overcome this, we propose to sample finite length trajectories of an unknown system and build an abstraction based on the concept of ℓ-completeness. To this end, we introduce the notion of probabilistic behavioural inclusion. We provide probably approximately correct (PAC) guarantees that such an abstraction, constructed from experimental symbolic trajectories of finite length, includes all behaviours of the concrete system, for both finite and infinite time horizon. Finally, our method is displayed with numerical examples.**

***Index Terms*—Automata, modeling, statistical learning.**

I. Introduction

RECENT advances in data-driven modeling and analysis, enabled by machine-learning and artificial intelligence with unprecedented computing power, have led to a renaissance in the field of system's verification, which focuses on providing formal performance and safety guarantees. To this end, sample-based methods directly derive barrier functions to certify invariance [1], [16], or finite abstractions to verify and synthesize controllers [8], [12], [15]. Among others, scenario-based optimization techniques can be employed to provide probably approximately correct (PAC) guarantees on the desired performance metric.

In order to use these techniques, independent samples must be obtained from the probability distribution driving a system's uncertainty. Several works [8], [13], [19], [20] consider the special case of deterministic systems, where the sole uncertainty is the initial state, which is drawn from some probability distribution and dictates the entire trajectory of the system. In this case, independence of samples can be achieved by sampling the initial state independently, typically through a uniform distribution over compact initial sets. However, we need independent samples of transitions (typically) when data is used to construct finite abstractions or verify barrier functions: this implies collecting only one transition from each trajectory. As a result, when attempting to *learn* a barrier function for a large set, the number of samples needed to construct a meaningful abstraction grows significantly. The scenario-based approach, when it is fed with one-step transitions, only establishes PAC guarantees for one-step properties. This impairs the provision of guarantees for longer horizon properties, for instance the decrease of a function across several (or an infinite number of) steps. Further, typically we are interested in infinite horizon properties: a model where probabilities are placed upon the possible transitions can hardly be used to infer long or infinite horizon specifications, since the probability of satisfaction for these properties become trivial.

Contributions: Addressing this limitation is the main objective of this letter. We consider deterministic systems with unknown dynamics and uncertainty in their initialization. We present a data-driven construction of finite abstractions, relying on the notions of ℓ-complete behaviours and (probabilistic) behavioural inclusion, which captures the relation between a randomly sampled deterministic model and a transition system based on the collected system's behaviours. We leverage nonconvex scenario theory to provide PAC guarantees for the inclusion of the concrete system's finite behaviours in those of the abstractions. Our technique is significantly different from the related literature, as we apply the scenario approach to bound the probability of witnessing a previously unseen, new system behaviour, and leverage contraction properties to extend the results to behaviours of arbitrary length (longer than that of the experiments).

Related Work: When (part of) a model is unknown, one can perform verification by employing a two-step procedure: first, estimating a model with classical system identification techniques to then verify the identified model via traditional formal techniques. Recently, a collection of works propose the use of PAC guarantees to directly synthesise an abstraction from data, with guarantees of correctness, without the need to identify an underlying model. In [2], [6], [11] a sampled-based interval MDP is provided, employing the scenario approach to bound the transition probabilities of a stochastic dynamical

Manuscript received 17 March 2023; revised 20 May 2023; accepted 12 June 2023. Date of publication 22 June 2023; date of current version 12 July 2023. This work was supported by the European Research Council through the SENTIENT Project under Grant ERC-2017-STG 755953. Recommended by Senior Editor F. Dabbene. *(Corresponding author: Rudi Coppola.)*

The authors are with the Faculty of Mechanical, Maritime and Materials Engineering, TU Delft, 2628 CD Delft, The Netherlands (e-mail: r.coppola@tudelft.nl).

Digital Object Identifier 10.1109/LCSYS.2023.3288731

model. In [8], the authors define a PAC alternating simulation relationship between a symbolic abstraction and an underlying deterministic system, using one-step transitions. In [13], PAC over-approximations of monotone systems are computed, which are then used to build models for unknown monotone systems. Finally, [10] computes the growth rate of a system from data, which is then used to construct a model abstraction and synthesise a controller. The use of data-driven ℓ-complete models is briefly presented in [14] for linear PETC models. In [1], [16] the authors synthesise barrier certificates for unknown systems using template-based candidates, providing PAC bounds for their correctness, for stochastic and deterministic systems, respectively.

II. Preliminaries

A. Notation

$||A||_p$ denotes the induced p-norm of matrix A. We use a string notation for sequences, e.g., $\mathtt{r} = ab$ means $\mathtt{r}(1) = a$, $\mathtt{r}(2) = b$. We denote the length of a string with the subscript $\ell \in \mathbb{N}_+$, i.e., $\mathtt{r}_\ell$. Given two sequences $\mathtt{r}_m$ and $\mathtt{p}_n$ with $\infty \geq m > n$, we say that $\mathtt{r}$ exhibits $\mathtt{p}$ if there exists $k \geq 0$ such that $\mathtt{r}(k+i) = \mathtt{p}(i)$ for $i = \{1, \dots, n\}$, denoted $\mathtt{r} \models \Diamond \mathtt{p}$. Given a set of sequences S and a sequence $\mathtt{p}$ we say that S exhibits $\mathtt{p}$ if there exists $\mathtt{s} \in S$ such that $\mathtt{s} \models \Diamond \mathtt{p}$, denoted $S \models \Diamond \mathtt{p}$. We denote the uniform distribution supported on a domain $\mathcal{D} \subset \mathbb{R}^n$ by $\mathcal{U}_{\mathcal{D}}$.

B. Scenario Theory Background

Let $(\Delta, \mathcal{F}, \mathbb{P})$ be a probability space, where Δ is the sample space, endowed with a σ-algebra $\mathcal{F}$ and a probability measure $\mathbb{P}$; further, denote by Δ^N the N-Cartesian product of the sample space and with $\mathbb{P}^N$ its product measure. A point in $(\Delta^N, \mathcal{F}^N, \mathbb{P}^N)$ is thus a sample $(\delta_1, \dots, \delta_N)$ of N elements drawn independently from Δ according to the same probability $\mathbb{P}$. Each δ_i is regarded as an observation, or *scenario* [4], [9].[1] A set Θ, the decision space, contains the decisions, i.e., the optimization space – no particular structure is assumed for this set. To every $\delta \in \Delta$ there is associated a constraint set $\Theta_\delta \subseteq \Theta$ which identifies the decisions that are admissible for the situation represented by δ.

Typically, the scenario theory refers to an optimisation program, which computes θ^*_N, the solution of the optimisation program based on N samples. Once θ^*_N is computed, we are interested in assessing how it generalises to unseen scenarios $\delta \in \Delta$, or, rather, the probability of extracting a sample that violates the constraints defined by θ^*_N. We define:

Definition 1 (Violation [4]): The violation probability of a given $\theta \in \Theta$ is defined as

$$V(\theta) = \mathbb{P}[\delta \in \Delta | \theta \notin \Theta_\delta]. \tag{1}$$

$V(\theta)$ quantifies the probability with which a new randomly selected constraint Θ_δ is violated by θ. If $V(\theta) \leq \epsilon$, we say that θ is ϵ-robust against constraint violation.

Notice that in general $V(\theta)$ is not directly computable since $\mathbb{P}$ is not known. From [9], under mild assumptions, a confidence bound can be derived as follows:

Theorem 1 (PAC Bounds [9, Th. 1]): Given a confidence parameter $\beta \in (0, 1)$ and the solution θ^*_N, it holds

$$\mathbb{P}^N[V(\theta^*_N) \leq \epsilon(s^*_N, \beta, N)] \geq 1 - \beta, \tag{2}$$

where $\epsilon(\cdot)$ is the solution of a polynomial equation (omitted here for brevity) and s^*_N is the so-called complexity of the solution – it represents the minimum number of constraints ($m \leq N$) that yield the same solution θ^*_N.

Remark 1: In this letter the event space Δ is discrete, therefore we refer to scenario theory for degenerate problems, as per [4], [9].

C. Modeling Framework

Consider a time-invariant dynamical system described by

$$\Sigma_s := \begin{cases} x_{k+1} = f(x_k) = Ax_k + b, \\ y_k = h(x_k), \\ x_0 \sim \mathcal{P}(\mathcal{D}), \end{cases} \tag{3}$$

where $x_k \in \mathcal{D} \subseteq \mathbb{R}^{n_x}$ is the plant's state at time $k \in \mathbb{N}_+$, $A \in \mathbb{R}^{n \times n}$, and $b := (I - A)x_{\text{eq}}$ with x_{eq} equilibrium of f, i.e., $x_{\text{eq}} = f(x_{\text{eq}})$; the initial value x_0 is *sampled* from a probability distribution $\mathcal{P}$ with domain $\mathcal{D}$; $y_k \in \mathcal{Y}$ is the system output with $|\mathcal{Y}| < \infty$, and n_x is the state-space dimension. If the trajectory x_k exits $\mathcal{D}$ at time k, the output map returns a special symbol $y^\dagger$ for all $t \geq k$. We may think of the map $h(\cdot)$ as a *partitioning* map, that returns a partition label (or index). The matrix A and output map $h(\cdot)$ are unknown, but we assume that we can observe the output sequence $y_0, y_1, \dots,$ generated by $\Sigma_s(x_0)$. With $\mathcal{B}^\omega(\Sigma_s(x))$ and $\mathcal{B}_H(\Sigma_s(x))$ we denote the infinite and finite external behaviour, the output sequence, for the time interval $[0, H-1]$ of Σ_s starting from state x, respectively; we use the shorthand $\mathcal{B}^\omega(x)$ and $\mathcal{B}_H(x)$ when the system is clear from the context. Each behaviour inherits a probability of emerging, stemming from $\mathcal{P}$: we consider solely behaviours with *strictly positive* probability measure.

Let us now introduce the notion of *equivalence class* [18]:

$$[y] = \{x \in \mathcal{D} \mid y = h(x)\},$$

and similarly, we define the equivalence class for an output sequence $\mathtt{y}_{\ell_i} = y_{i_1} y_{i_2} \dots y_{i_\ell} \in \mathcal{Y}^\ell$ as

$$[\mathtt{y}_{\ell_i}] = \{x \in \mathcal{D} \mid y_{i_j} = h(f^{j-1}(x)) \text{ for } j = 1, \dots, \ell\}, \tag{4}$$

with $f^0(x) = x$. Equation (4) states that for $i = 1, \dots, |\mathcal{Y}|^\ell$, i.e., for every ℓ-sequence $\mathtt{y}_{\ell_i} \in \mathcal{Y}^\ell$, the output equivalence class $[\mathtt{y}_{\ell_i}]$ is the set of points x such that if the dynamical system is initialized at x, then the output sequence over the time interval $[0, \ell - 1]$ corresponds to $\mathtt{y}_{\ell_i}$. Further, for all $\ell \geq 1$, the set of all $[\mathtt{y}_{\ell_i}]$ forms a partition of the domain $\mathcal{D}$.

Let us introduce the notion of contraction, which is convenient for the discussions in Section IV.

Definition 2 (Contraction Map [3]): Given a metric space $(\mathcal{D}, d)$, a map $f : \mathcal{D} \to \mathcal{D}$ is a contraction if it is Lipschitz with constant $\mathscr{C} < 1$, i.e.,

$$d(f(x), f(y)) \leq \mathscr{C} \cdot d(x, y) \quad \forall x, y \in \mathcal{D},\ 0 \leq \mathscr{C} < 1. \tag{5}$$

[1] As indicated in footnote 1 on [4] one could equivalently consider δ_i as independent random elements of a probability space.

In this letter, we solve the following problem.

Problem Statement: Given an unknown affine system, build an abstraction such that, with high confidence, the probability of witnessing a behaviour that is not included in the abstraction's behaviours is below a threshold value.

III. Sampling and Abstractions

A. Abstractions via Transition Systems

In this letter, we adopt the framework of finite-state abstractions in the form of transition systems (TS).

Definition 3 (Transition System [18]): A transition system $\mathcal{S}$ is a tuple $(\mathcal{X}, \mathcal{X}_0, \mathcal{E}, \mathcal{Y}, \mathcal{H})$ where:

- $\mathcal{X}$ is the (possibly infinite) set of states,
- $\mathcal{X}_0 \subseteq \mathcal{X}$ is the set of initial states,
- $\mathcal{E} \subseteq \mathcal{X} \times \mathcal{X}$ is the set of edges, or transitions,
- $\mathcal{Y}$ is the set of outputs, and
- $\mathcal{H} : \mathcal{X} \to \mathcal{Y}$ is the output map.

We consider *non-blocking* transition systems, that is, every state is equipped with at least one outgoing transition. Notice that model (3) can be described equivalently as a non-blocking and deterministic transition system where the (initial) states belong to probability distributions derived by $x_0 \sim \mathcal{P}$; we denote such transition system by $\mathcal{S}_{\Sigma_s}$.

In order to construct the embedding of (3) as a TS, we need full knowledge of the system dynamics to compute the transitions $\mathcal{E}$. As we only have access to the output y_k, we recall the notion of *behavioural inclusion*:

Definition 4 [(Probabilistic) Behavioural Inclusion]: Consider two systems $\mathcal{S}_a$ and $\mathcal{S}_b$ with $\mathcal{Y}_a = \mathcal{Y}_b$. We say that $\mathcal{S}_b$ behaviourally includes $\mathcal{S}_a$, denoted by $\mathcal{S}_a \preceq_{\mathcal{B}} \mathcal{S}_b$, if $\mathcal{B}^{\omega}(\mathcal{S}_a) \subseteq \mathcal{B}^{\omega}(\mathcal{S}_b)$. We say that $\mathcal{S}_b$ behaviourally includes $\mathcal{S}_a$ until horizon H if $\mathcal{B}_H(\mathcal{S}_a) \subseteq \mathcal{B}_H(\mathcal{S}_b)$, denoted $\mathcal{S}_a \preceq_{\mathcal{B}_H} \mathcal{S}_b$. Further, we say that $\mathcal{S}_a$ is behaviourally included in $\mathcal{S}_b$ with probability greater or equal than $1-\epsilon$, denoted by $\mathbb{P}[\mathcal{S}_a \preceq_{\mathcal{B}} \mathcal{S}_b] \geq 1-\epsilon$, if for $x_0 \sim \mathcal{P}$ it holds that:

$$\mathbb{P}\big[\mathcal{B}^{\omega}(\mathcal{S}_a(x_0)) \subseteq \mathcal{B}^{\omega}(\mathcal{S}_b) | x_0 \sim \mathcal{P}\big] \geq 1-\epsilon, \qquad (6)$$

We denote $\mathbb{P}[\mathcal{S}_a \preceq_{\mathcal{B}_H} \mathcal{S}_b] \geq 1-\epsilon$, if the previous relationship holds until time horizon H.

A natural way of building a behaviourally inclusive abstraction is by fixing a length ℓ, and mapping a concrete state to an abstract one sharing the future ℓ outputs. We may elaborate this intuition through a so-called ℓ-complete model:

Definition 5 [(Strongest Asynchronous) ℓ-Complete Abstraction [7], [17]]: Let $\mathcal{S} := (\mathcal{X}, \mathcal{X}_0, \mathcal{E}, \mathcal{Y}, \mathcal{H})$ be a transition system, and let $\mathcal{X}_\ell \subseteq \mathcal{Y}^\ell$ be the set of all ℓ-long subsequences of all behaviours in $\mathcal{S}$. Then, the system $\mathcal{S}_\ell = (\mathcal{X}_\ell, \mathcal{B}_\ell(\mathcal{S}), \mathcal{E}_\ell, \mathcal{Y}^\ell, \mathcal{H})$ is called the (strongest asynchronous) ℓ-complete abstraction (SAℓ-CA) of $\mathcal{S}$, where

- $\mathcal{E}_\ell = \{(k\sigma, \sigma k') | k, k' \in \mathcal{Y}, \sigma \in \mathcal{Y}^{\ell-1}, k\sigma, \sigma k' \in \mathcal{X}_\ell\}$,
- $\mathcal{H}(k\sigma) = k$,

where $\mathcal{B}_\ell(\mathcal{S})$ denotes all the external traces of system $\mathcal{S}_\ell$ and $\mathcal{Y}^\ell$ is the cartesian product $\mathcal{Y} \times \ldots \times \mathcal{Y}$ repeated ℓ times.

The SAℓ-CA encodes each state as an ℓ-long external trace, as shown in Fig. 1. When $\ell = 3$ and $\mathcal{Y} = \{y_0, y_1, y_2\}$, assume that the three-step trajectory $\hat{x}_0 x_1 x_2$ yields the output $\mathcal{B}_3(\mathcal{S}(\hat{x}_0)) = y_0 y_1 y_2$; then, the SA$\ell$-CA has the abstract state $\mathrm{y}_3 := y_0 y_1 y_2$. Moreover, $\hat{x}_0$ belongs to the equivalence class $[y_0 y_1 y_2]$. The transitions of an ℓ-complete model obey the so-called "domino rule", meaning that starting from the trace $y_0 y_1 y_2$, the next ℓ-trace must begin with $y_1 y_2$; e.g., state $y_0 y_1 y_2$ can transition to $y_1 y_2 y_0$, $y_1 y_2 y_1$, and $y_1 y_2 y_2$. The output of a state is its first element, so $\mathcal{H}(y_0 y_1 y_2) = y_0$.

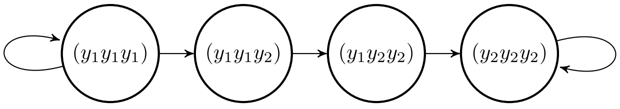

Fig. 1. Example of SAℓ-CA, with $\ell = 3$.

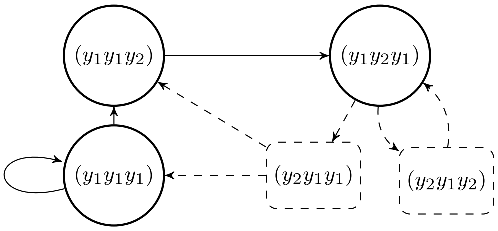

Fig. 2. Construction of a non-blocking automaton. Dashed lines indicate artificial states and transitions, added by the domino completion.

An autonomous transition system $\mathcal{S}$ is behaviourally included in its corresponding SAℓ-CA model (see [17]). Constructing such a model requires only external behaviours of $\mathcal{S}$; considering x_0 as an input, the SAℓ-CA represents a symbolic input/output model. The value of ℓ defines the trajectory horizon, embedding the future of each concrete state x_k. A transition from an abstract state y_{ℓ_i} to y_{ℓ_j} exists if and only if the last $\ell - 1$ output symbols of the former match the first $\ell - 1$ symbols of the latter; consequently, for increasing values of ℓ the domino rule is more stringent, or in other words, the partition defined by the set of all $[\mathrm{y}_{\ell_i}]$ becomes finer. We only need to collect the possible ℓ-behaviours of a system to construct an ℓ-complete model.

B. Data-Driven Abstraction: Finite-Time Guarantees

Given system (3), let us collect N output sequences of length $H \geq \ell$, and we construct $\mathcal{X}_\ell^N$, the set of ℓ-sequences y_ℓ, that acts as the state set of the data-driven SAℓ-CA.

Definition 6 (Data-Driven SAℓ-CA): The ℓ-complete abstraction $\mathcal{S}_\ell^N = (\mathcal{X}_\ell^N, \mathcal{X}_\ell^N, \mathcal{E}_\ell, \mathcal{Y}^\ell, \mathcal{H})$ is called the data-driven SAℓ-CA of $\mathcal{S}$, where

- $\mathcal{X}_\ell^N$ is the state space built from the ℓ-sequences collected from N trajectories of an underlying concrete system.

The transitions, output space, and output map follow Definition 5.

Remark 2 (Domino Completion): A *blocking* TS may arise after collecting N samples, as depicted in Fig. 2, where the state corresponding to $y_1 y_2 y_1$ has no outgoing transitions. The existence of $y_1 y_2 y_1$ *implies* the existence of at least one sequence starting with $y_2 y_1$. Thus, we may add artificially all states corresponding to sequences $y_2 y_1 *$. To simplify this process, by analysing the collected states, we add a *minimal* set of states that complete the transitions: in this example, we may add only state $y_2 y_1 y_1$. We repeat the procedure until we obtain a non-blocking transition system.

From here on we assume that the data-driven SAℓ-CA is non-blocking. Once we collect N trajectories from (3) and construct the corresponding data-driven SAℓ-CA, we leverage the scenario theory to provide bounds on the probabilistic behavioural inclusion between $\mathcal{S}_\ell^N$ and $\mathcal{S}_{\Sigma_s}$.

Let us sample N i.i.d. initial conditions $\{x_{0,i}\}_{i=1}^N$ of the dynamical system, and consider the resulting H-long behaviours, denoted by $\{\mathcal{B}_H(x_{0,i})\}_{i=1}^N$. We define the N scenarios $\{\delta_i\}_{i=1}^N$ as binary vectors indicating which ℓ-sequences are included in a single behaviour. Formally,

$$\delta_i(j) = \begin{cases} 1 & \text{if } \mathcal{B}_H(x_{0,i}) \models \Diamond \mathrm{y}_{\ell_j}, \\ 0 & \text{else}, \end{cases}$$

for $j \in \{1, \dots, |\mathcal{Y}|^\ell\}$. The scenario program results

$$\begin{aligned} \min_{\theta\in\Theta} \quad & \mathbf{1}^{\mathrm{T}}_{|\mathcal{Y}|^\ell} \cdot \theta \\ \text{s.t.} \quad & \mathbf{1}^{\mathrm{T}}_{|\mathcal{Y}|^\ell}(\theta - \delta_i) \geq 0, \quad i = 1, \dots, N, \end{aligned} \tag{7}$$

where the search space of θ is defined as $\Theta := [0,1]^{|\mathcal{Y}|^\ell}$. The solution θ_N^* in practice indicates which ℓ-sequences were witnessed in the samples collected; the solution changes solely when we collect a new value for δ_i, previously unseen. The complexity of the solution s_N^* is equal to the cardinality of the smallest subset of the N H-sequences collected that yield the same solution to θ_N^*.

Remark 3: We interpret program (7) as the collection of labels (the ℓ-sequences of partitions) from a discrete probability distribution of unknown support size. The scenario theory provides a bound to the probability of collecting a new, unseen, label from the unknown distribution.

Proposition 1: Consider a confidence β, and N trajectories of length H collected from (3), and the corresponding data-driven SAℓ-CA $\mathcal{S}_\ell^N$ based on the observed ℓ-sequences. For a new sampled initial condition $x_0 \sim \mathcal{P}(\mathcal{D})$ it holds that

$$\mathbb{P}^N[\mathbb{P}[\mathcal{S}_{\Sigma_s} \preceq_{\mathcal{B}_H} \mathcal{S}_\ell^N] \geq 1 - \epsilon(s_N^*, N, \beta)] \geq 1 - \beta. \tag{8}$$

Proof (Sketch): The scenario theory guarantees that the probability of sampling an initial condition x_0 that generates an unseen ℓ-sequence, over the time horizon H, is bounded by ϵ; hence, (8) holds. ■

C. Role of ℓ

Let us fix N, β, and H and let us sample the set of trajectories $\{\mathcal{B}_H(x_{0,i})\}_{i=1}^N$. Consider two different values $\ell_1 < \ell_2 \leq H$, for each solve program (7), and construct two SAℓ-CAs, namely $\mathcal{S}_{\ell_1}^N$ and $\mathcal{S}_{\ell_2}^N$. We denote by s_{N,ℓ_1}^* and s_{N,ℓ_2}^* the complexity of the respective solutions. It is possible to show that $s_{N,\ell_1}^* \leq s_{N,\ell_2}^*$, which implies $\epsilon(s_{N,\ell_1}^*, N, \beta) \leq \epsilon(s_{N,\ell_2}^*, N, \beta)$. On the other hand, we have $\mathcal{B}_H(\mathcal{S}_{\ell_2}^N) \subseteq \mathcal{B}_H(\mathcal{S}_{\ell_1}^N)$. While a natural choice for ℓ is H, since it generates an abstraction containing *only* the behaviours that were sampled, a smaller value for ℓ allows us to build an abstraction richer in behaviours (spurious or not), with tighter PAC guarantees and smaller state set.

IV. Infinite Behaviours

Proposition 1 states that we can construct an abstraction that behaviourally includes the concrete system, with PAC guarantees, up to the horizon H. Let us now discuss how to extend the guarantees to infinite horizon properties.

We denote the probability of sampling an initial condition and thereafter visiting an arbitrary set S within k steps

$$\mu_0^k(S) := \mathbb{P}[x_0 : \mathcal{B}_k(x_0) \models \Diamond S] = \mathbb{P}\left[\bigcup_{i=0}^k \mathrm{Pre}_{\mathcal{D}}^i(S)\right], \tag{9}$$

where

$$\mathrm{Pre}_{\mathcal{D}}^i(S) := \{x' \in \mathcal{D} | f^i(x') \in S, f^j(x') \in \mathcal{D}, 0 \leq j \leq i\},$$

the points in the domain whose trajectory remains within $\mathcal{D}$ for all steps $j \leq i$, and the i-th step is within S. We specialize (9) for arbitrary equivalence classes $[\mathrm{y}_\ell]$ as

$$\mathbb{P}[x_0 : \mathcal{B}^\omega(x_0) \models \Diamond \mathrm{y}_\ell] = \mu_0^\infty([\mathrm{y}_\ell]) \geq \mu_0^k([\mathrm{y}_\ell]), \quad \forall k,$$

the probability of sampling an initial state x_0 that eventually leads to witness the trace y_ℓ corresponds to the probability of the equivalence class of y_ℓ together with all the sets *eventually* leading to it, i.e., the $\mathrm{Pre}_{\mathcal{D}}^k([\mathrm{y}_\ell])$. The quantity $\mu_0^k([\mathrm{y}_\ell])$ is a monotonically non-decreasing function with k: for a small k the probability of sampling y_ℓ may be negligible, for $k \to \infty$, it may reach a large value. Further, $\mu_0^k([\mathrm{y}_\ell])$ describes the accumulated probability of visiting $[\mathrm{y}_\ell]$, and how this changes – due to the system dynamics – with k. Notice also that if $\mu_0^\tau([\mathrm{y}_\ell]) = \mu_0^{\tau+1}([\mathrm{y}_\ell])$ for some $\tau \in \mathbb{N}$, it holds that $\mu_0^\tau([\mathrm{y}_\ell]) = \mu_0^\infty([\mathrm{y}_\ell])$; equivalently $\mu_0^\tau([\mathrm{y}_\ell])$ measures the largest set that can *ever* visit $[\mathrm{y}_\ell]$. The following assumption pivots on the observations above.

Assumption 1: Given a system (3), assume that a monotonically non-decreasing function φ is known such that for some $k \in \mathbb{N}$ and for all sets S corresponding to arbitrary unions of equivalence classes, $S = \bigcup_{j\in J} [\mathrm{y}_{\ell_j}]$ with $J \subseteq \{1, 2, \dots, |\mathcal{Y}|^\ell\}$, the following holds

$$\mu_0^k(S) \geq \varphi(k) \cdot \mu_0^\infty(S). \tag{10}$$

This trivially implies that

$$\mu_0^k(S) < \epsilon \implies \mu_0^\infty(S) < \overline{\gamma} := (\varphi(k))^{-1} \cdot \epsilon.$$

In practice, Assumption 1 allows to link the probability measure of visiting any ℓ-sequence's equivalence class S in k steps with the probability of visiting it in an infinite number of steps. Importantly, the function $\varphi(k)$ describes how the ratio $\mu_0^k(S)/\mu_0^\infty(S)$ changes over time, due to the dynamics of the system. Function φ is monotonically non-decreasing, and either tends to 1 as $k \to \infty$ or attains $\varphi(k) = 1$ for all $k \geq \tau$ for some finite τ. Assumption 1 allows us to consider behaviours of arbitrary length.

Proposition 2: Consider N trajectories of length H collected from (3), the corresponding data-driven SAℓ-CA $\mathcal{S}_\ell^N$, and, for a given β, the bound $\overline{\epsilon} := \epsilon(s^*, N, \beta)$ resulting from solving program (7). Let Assumption 1 hold for $k = H - \ell$. Then, for $x_0 \sim \mathcal{P}(\mathcal{D})$,

$$\mathbb{P}^N[\mathbb{P}[\mathcal{S}_{\Sigma_s} \preceq_{\mathcal{B}} \mathcal{S}_\ell^N] \geq 1 - \overline{\gamma}] \geq 1 - \beta, \tag{11}$$

i.e., with confidence $1-\beta$, $\mathcal{S}_\ell^N$ probably behaviourally includes the model (3) with probability not smaller than $1-\overline{\gamma}$.

Proof: We denote θ_N^* as the optimal solution of the scenario program and $V(\theta_N^*)$ represents the probability of drawing a new initial condition x_0 which results in a H-long behaviour exhibiting one (or more) previously unseen ℓ-sequence:

$$V(\theta_N^*) := \mathbb{P}[x_0 : \mathcal{B}_H(\mathcal{S}_{\Sigma_s}(x_0)) \models \Diamond \mathtt{y}_\ell \wedge \mathtt{y}_\ell \notin \mathcal{X}_\ell^N].$$

The scenario theory assures that $V(\theta_N^*) < \overline{\epsilon}$, with confidence not smaller than $1-\beta$. Let us denote with $\tilde{S}$ the set of unseen ℓ-sequences $\tilde{\mathtt{y}}_{\ell_j}$, such that $\mathcal{B}^\omega(\mathcal{S}_{\Sigma_s})$ exhibits $\tilde{\mathtt{y}}_{\ell_j}$ but $\tilde{\mathtt{y}}_{\ell_j} \notin \mathcal{X}_\ell^N$; with $\mathcal{W}$ the union of the corresponding equivalence classes $\bigcup_{\tilde{\mathtt{y}}_{\ell_j} \in \tilde{S}} [\tilde{\mathtt{y}}_{\ell_j}]$. The scenario theory ensures that

$$V(\theta_N^*) = \mathbb{P}[\mathcal{B}_H(x_0) \models \Diamond \tilde{\mathtt{y}}_{\ell_j}, \tilde{\mathtt{y}}_{\ell_j} \in \tilde{S}] = \mu_0^k(\mathcal{W}),$$

where $\mu_0^k(\cdot)$ is the probability measure of all initial conditions $x_0 \in \mathcal{D}$ which exhibit any $\tilde{\mathtt{y}}_{\ell_j} \in \tilde{S}$ in at most H steps (recall that $k = H - \ell$). Let us apply Assumption 1; then

$$\mu_0^\infty(\mathcal{W}) \leq \frac{1}{\varphi(k)} \cdot \mu_0^k(\mathcal{W}) < \frac{1}{\varphi(k)} \cdot \overline{\epsilon} = \overline{\gamma}.$$

It follows that with confidence $1-\beta$, all unseen sequences $\tilde{\mathtt{y}}_{\ell_j}$ have a total probability measure of being exhibited by an infinite behaviour upper-bounded by $\overline{\gamma}$, hence the behavioural relationship holds with PAC bound $\overline{\gamma}$. ■

One can think of the value $\overline{\epsilon}$ as a bound on the probability measure of initialising the system such that an unseen ℓ-sequence is produced in less than H steps (with confidence $1-\beta$). We then leverage the knowledge of φ to extend the scenario guarantees from the finite to the infinite horizon.

A. Uncertain Affine Stable Systems

Let us consider the class of affine stable systems described by (3), where both A and x_{eq} are unknown. We introduce the following proposition as a stepping stone for providing infinite time guarantees.

Proposition 3 (Finite Exit Time): Given an affine system

$$x_{k+1} = f(x_k) = Ax_k + b, \quad x_k, x_{\mathrm{eq}} \in \mathcal{D} \subset \mathbb{R}^n, \tag{12}$$

with $b := (I-A)x_{\mathrm{eq}}$, if $||A||_2 \leq \alpha < 1$, then for any $d > 0$ the sets $S_d := \{x \in \mathbb{R}^n | |x - x_{\mathrm{eq}}|_2 \leq d\}$ and $Q \subseteq \mathcal{D} \setminus S_d$ satisfy

$$\mathcal{D} \subseteq A^{-k} S_d, \qquad \mathcal{D} \cap A^{-k} Q = \emptyset$$

for $k \geq \kappa(d, \hat{d}, \alpha) := \lceil \log_\alpha(\frac{d}{\hat{d}}) \rceil$, where $\hat{d}$ is an upper bound on the radius of the smallest ball containing the domain, i.e., $\hat{d} \geq \inf\{r > 0 | |x - x_{\mathrm{eq}}|_2 > r \implies x \notin \mathcal{D}\}$.

The proposition exploits the fact that the ball of radius $\alpha^{-1} d$ is contained in the preimage of the set S_d, i.e., $S_{\alpha^{-i} d} \subseteq f^{-i}(S_d)$. Further, disjoint sets have disjoint preimages through the affine map f: $S_d \cap Q = \emptyset \implies f^{-i}(S_d) \cap f^{-i}(Q) = \emptyset$ for $i \in \mathbb{N}$. These arguments are easily adapted even if the equilibrium point $x_{\mathrm{eq}} \notin \mathcal{D}$.

We compute the φ function exploiting a coarse knowledge of the dynamics, as bounds on the eigenvalues' magnitude.

Proposition 4: For systems defined by (3), with $0 < ||A||_2 \leq \alpha < 1$, $\rho \geq |\det(A^{-1})| > 1$, and $x_0 \sim \mathcal{U}_\mathcal{D}$, assume there exists an equivalence class $[y^*]$ containing a ball of radius $\check{d} > 0$ around x_{eq}, then we can construct a function φ satisfying Assumption 1 for every k, with

$$\varphi(k) = \begin{cases} \left(1 + \rho^{\bar{k}-k} \sum_{i=0}^{z(k)-1} \rho^{-i(k+1)}\right)^{-1} & \text{for } k \leq \bar{k} \\ 1 & \text{for } k > \bar{k} \end{cases} \tag{13}$$

where: $z(k) = \lceil (\bar{k}+1)/(k+1) \rceil - 1$, and $\bar{k} = \kappa(\check{d}, \hat{d}, \alpha)$ (see Proposition 3), with $\hat{d} \geq \inf\{r > 0 | |x - x_{\mathrm{eq}}|_2 > r \implies x \notin \mathcal{D}\}$.

Proof: (Sketch. Please refer to the online more detailed version of this letter [5] for details). By definition of $\bar{k}$, and since the dynamics are linear, for any set $S \subseteq \mathcal{D}$:

$$\mu_0^{\bar{k}}(S) = \mu_0^\infty(S), \tag{14}$$

$$\mu_0^0(\mathrm{Pre}_\mathcal{D}(S)) \leq \rho \cdot \mu_0^0(S). \tag{15}$$

From: $\bigcup_{i=0}^q \mathrm{Pre}_\mathcal{D}^i(Q) = \bigcup_{i=0}^{q-1} \mathrm{Pre}_\mathcal{D}^i(Q) \cup \bigcup_{i=1}^q \mathrm{Pre}_\mathcal{D}^i(Q)$, it follows that

$$\mu_0^q(Q) \leq (1+\rho) \cdot \mu_0^{q-1}(Q).$$

Let $q, k \in \mathbb{N}^+$ with $q \geq k$ and denote by $z(k) := \lceil (q+1)/(k+1) \rceil - 1$. For any arbitrary set Q the set $\bigcup_{i=0}^q \mathrm{Pre}_\mathcal{D}^i(Q)$ can be recast as a union of sets in the form $\bigcup_{i=t}^{t+k} \mathrm{Pre}_\mathcal{D}^i(Q)$ which results in:

$$\mu_0^q(Q) \leq \sum_{i=0}^{z(k)-1} \mu_{q-i(k+1)-k}^{q-i\cdot(k+1)}(Q) + \mu_0^k(Q),$$

and thus, by virtue of (15):

$$\mu_0^q(Q) \leq \mu_0^k(Q) \left(1 + \rho^{q-k} \sum_{i=0}^{z(k)-1} \rho^{-i(k+1)}\right). \tag{16}$$

Combining (14) with (16) concludes the proof. ■

In practical terms, the value $\bar{k}$ from Proposition 4 ensures that $\mu_0^{\bar{k}}([\mathtt{y}_\ell]) = \mu_0^\infty([\mathtt{y}_\ell])$ for all $[\mathtt{y}_\ell]$. In other words, the system reaches a steady behaviour after $\bar{k}$ steps.

Remark 4: For a discrete-time linear system, the asymptotic stability condition $||A||_2 < 1$ implies that the system's flow is a contraction map with respect to the Euclidean norm.

Remark 5: The proof above can be adapted for nonlinear systems as long as the map $f(x)$ in (3) is a contraction (see Definition 2), and equation (15) holds. The former ensures that there exists a finite exit time for every set of interest. The latter bounds the growth of sets, when considering the inverse dynamics. A contraction that is also a *lipeomorphism* satisfies these conditions. This extension is ongoing work.

V. Experimental Evaluation

A. Linear Stable System

Let us consider the system

$$x_{k+1} = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ -1.8 & 1 \end{bmatrix} x_k. \tag{17}$$

The state space $\mathcal{D} = [-1, 1]^2$ is partitioned into 81 regions by a uniform grid. We sample $N = 10^5$ initial conditions x_0 from the uniform distribution $\mathcal{U}_\mathcal{D}$, we collect trajectories

TABLE I
RESULTS WITH 81^2 PARTITIONS, FOR TWO VALUES OF ℓ

ℓ	β	# sequences	s_N^*	$\overline{\epsilon}$	$\hat{V}$
3	10^{-12}	33541	17221	0.019	$1.4 \cdot 10^{-3}$
9	10^{-12}	67099	67099	0.069	$5.2 \cdot 10^{-3}$

of length $H = 9$, and we consider $\ell = 3$. We collect 454 ℓ-sequences and construct the corresponding abstraction. Setting $\beta = 10^{-12}$, we compute the scenario bounds (8),

$$\overline{\epsilon} = \epsilon(s^*, \beta, N) = 3.54 \cdot 10^{-3}.$$

To verify these bounds empirically, we sample in addition $M = 10^6$ initial conditions, and get an empirical violation probability $\hat{V} \simeq 6 \cdot 10^{-6}$, a value well below the bounds. In order to extend the guarantees from horizon $H = 9$ to the infinite horizon we employ Assumption 1. We leverage Proposition 4, and take conservative[2] values for the parameters, as $\rho = 2$, $\alpha = 0.8$, $d_{\max} = 1$, and $d_{\min} = 0.12$, which give $\overline{k} = 10$. We compute φ using (13), and by setting $k = H - \ell = 6$, we can upper bound the measure of $\mu_0^\infty(S)$

$$\mu_0^\infty(S) = \mu_0^{\overline{k}}(S) < \bar{\gamma} = \varphi(k)^{-1} \cdot \overline{\epsilon} = 6.01 \cdot 10^{-2}.$$

Our abstraction holds for infinite horizon properties with PAC guarantees

$$\mathbb{P}^N[\mathbb{P}[\mathcal{S} \preceq_{\mathcal{B}} \mathcal{S}_\ell^N] \geq 1 - 6.01 \cdot 10^{-2}] \geq 1 - 10^{-12}.$$

Empirically, after collecting additional $M = 10^6$ trajectories, we obtain $\hat{V} = 2.5 \cdot 10^{-4}$, abiding our guarantees.

Finer Partitioning: Consider again system (17), where $\mathcal{D}$ is uniformly partitioned into 81^2 regions. For this partition we have $\overline{k} = 20$. We sample $N = 10^6$ initial conditions $x_0 \sim \mathcal{U}_{\mathcal{D}}$, we collect trajectories of length $H = 9$, and we consider $\ell_1 = 3$ and $\ell_2 = H$, whose results are reported in Table I, with the latter needing a domino completion adding 4857 sequences. We highlight the trade off between ℓ and $\overline{\epsilon}$: a smaller ℓ provides tighter $\overline{\epsilon}$, but the corresponding abstraction generates more spurious behaviours. Finally, observe that equation (16) holds generally for any $q \geq k$, not only for $q = \overline{k}$: this allows us to relate guarantees for experiments of different finite horizons. For example, suppose that we consider the probabilistic behavioural inclusion of $\mathcal{S}$ in $\mathcal{S}_{\ell_1}$ until horizon $H' = 12$. Following similar steps as in the proof of Proposition 2, leveraging (16), we can conclude

$$\mathbb{P}^N[\mathbb{P}[\mathcal{S} \preceq_{\mathcal{B}_{12}} \mathcal{S}_3^N] \geq 1 - 0.16] \geq 1 - 10^{-12}.$$

We verify these bounds numerically by sampling $M = 10^6$ trajectories of length $H' = 12$: we obtain $\hat{V} = 1.49 \cdot 10^{-3}$.

VI. CONCLUSION AND FUTURE WORK

We have presented a method to construct a finite, data-driven abstraction of an unknown affine deterministic system under uniform random sampling of a set of initial conditions. Note that, with little effort, this can be generalised to other classes of distributions, e.g., piecewise constant. We introduce the notion of probabilistic behavioural inclusion, and use it to bound the probability of unseen behaviours of the concrete system. We then build an ℓ-complete automaton that generates behaviours of the concrete system, based on trajectories up to time H; under additional assumptions, our construction holds for infinite time. Future work includes the extension to control synthesis and stochastic systems.

ACKNOWLEDGMENT

The authors would like to thank Gabriel de Albuquerque Gleizer and Giannis Delimpaltadakis for their helpful discussions throughout the development of this letter.

REFERENCES

[1] P. Akella and A. D. Ames, "A barrier-based scenario approach to verifying safety-critical systems," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 11062–11069, Oct. 2022.

[2] T. S. Badings, A. Abate, N. Jansen, D. Parker, H. A. Poonawala, and M. Stoelinga, "Sampling-based robust control of autonomous systems with non-gaussian noise," in *Proc. Workshops 36th AAAI Conf. Artif. Intell.*, 2022, p. 9.

[3] F. Bullo, *Contraction Theory for Dynamical Systems*. London, U.K.: Kindle Direct, 2023. [Online]. Available: https://fbullo.github.io/ctds

[4] M. C. Campi, S. Garatti, and F. A. Ramponi, "A general scenario theory for nonconvex optimization and decision making," *IEEE Trans. Autom. Control*, vol. 63, no. 12, pp. 4067–4078, Dec. 2018.

[5] R. Coppola, A. Peruffo, and M. Mazo, "Data-driven abstractions for verification of deterministic systems," 2022, *arXiv:2211.01793*.

[6] M. Cubuktepe, N. Jansen, S. Junges, J.-P. Katoen, and U. Topcu, "Scenario-based verification of uncertain MDPS," in *Proc. Int. Conf. Tools Algorithms Construction Anal. Syst.*, 2020, pp. 287–305.

[7] G. A. De Gleizer and M. Mazo, "Computing the sampling performance of event-triggered control," in *24th ACM Int. Conf. Hybrid Syst. Comput. Control (HSCC)*, 2021, pp. 1–7.

[8] A. Devonport, A. Saoud, and M. Arcak, "Symbolic abstractions from data: A PAC learning approach," in *Proc. 60th IEEE Conf. Decis. Control (CDC)*, 2021, pp. 599–604.

[9] S. Garatti and M. C. Campi, "The risk of making decisions from data through the lens of the scenario approach," *IFAC-PapersOnLine*, vol. 54, no. 7, pp. 607–612, 2021.

[10] M. Kazemi, R. Majumdar, M. Salamati, S. Soudjani, and B. Wooding, "Data-driven abstraction-based control synthesis," 2022, *arXiv:2206.08069*.

[11] A. Lavaei, S. Soudjani, E. Frazzoli, and M. Zamani, "Constructing MDP abstractions using data with formal guarantees," *IEEE Control Syst. Lett.*, vol. 7, pp. 460–465, 2022.

[12] R. Majumdar, N. Ozay, and A.-K. Schmuck, "On abstraction-based controller design with output feedback," in *Proc. 23rd Int. Conf. Hybrid Syst. Comput. Control*, 2020, pp. 1–11.

[13] A. Makdesi, A. Girard, and L. Fribourg. "Data-driven models of monotone systems." 2022. [Online]. Available: https://hal.science/hal-03709123v1

[14] A. Peruffo and M. Mazo, "Data-driven abstractions with probabilistic guarantees for linear PETC systems," *IEEE Control Syst. Lett.*, vol. 7, pp. 115–120, 2022.

[15] S. Sadraddini and C. Belta, "Formal guarantees in data-driven model identification and control synthesis," in *Proc. 21st Int. Conf. Hybrid Syst. Comput. Control (CPS Week)*, 2018, pp. 147–156.

[16] A. Salamati, A. Lavaei, S. Soudjani, and M. Zamani, "Data-driven safety verification of stochastic systems via barrier certificates," *IFAC PapersOnLine*, vol. 54, no. 5, pp. 7–12, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896321012416

[17] A.-K. Schmuck, P. Tabuada, and J. Raisch, "Comparing asynchronous L-complete approximations and quotient based abstractions," in *Proc. 54th IEEE Conf. Decis. Control (CDC)*, 2015, pp. 6823–6829.

[18] P. Tabuada, *Verification and Control of Hybrid Systems: A Symbolic Approach*. New York, NY, USA: Springer, 2009.

[19] Z. Wang and R. M. Jungers, "A data-driven method for computing polyhedral invariant sets of black-box switched linear systems," *IEEE Control Syst. Lett.*, vol. 5, no. 5, pp. 1843–1848, Nov. 2021.

[20] Z. Wang and R. M. Jungers, "Scenario-based set invariance verification for black-box nonlinear systems," *IEEE Control Syst. Lett.*, vol. 5, no. 1, pp. 193–198, Jan. 2021.

[2] True values are $|\det(A^{-1})| = 1.96$, $||A||_2 = 0.75$, and $d_{\min} = 0.11$.