DATA-DRIVEN SAFETY VERIFICATION OF POMDPS: SHOWING LEARNABILITY OF SAFETY CERTIFICATES VIA SPECTRAL LEARNING OF WEIGHTED AUTOMATA

Anonymous authors
Paper under double-blind review

ABSTRACT

We develop a data-driven method for finite-horizon safety verification of a partially observable Markov decision process (POMDP) with unknown transition and observation kernels under a fixed finite-state controller. The method uses only observation data and covers two sampling regimes: multiple independent trajectories generated from a common initial distribution, and a single trajectory whose closed-loop latent process is stationary and geometrically mixing. We represent an observation-based safety specification by a deterministic finite automaton (DFA). Exploiting the weighted finite automaton (WFA) realization of the observation-word probability function of the closed-loop POMDP, we estimate empirical Hankel matrices and apply the Ho-Kalman algorithm to obtain a WFA estimate, which is used as a proxy for the unknown POMDP. More precisely, we estimate the probability with which the learned WFA generates words which satisfy the specification. This estimated probability, in turn, can be used to provide a PAC (Probably Approximately Correct) bound on the probability with which the words generated by the unknown POMDP satisfy the specification. This then allows us to provide PAC bounds on safety certificates for the original POMDP, showing that safety certificates of POMDPs are PAC learnable.

1 INTRODUCTION

Partially observable stochastic systems, usually modeled as partially observable Markov decision processes (POMDPs), arise when the underlying system states cannot be directly observed. In safety-critical systems, it is important to assess whether the resulting output satisfies a prescribed specification with sufficient probability. Existing verification methods for partially observable systems usually assume that the transition and observation kernels are known, enabling verification through belief-space representations and finite-state abstractions. In many data-driven settings, such a model may not be available, while trajectories can be observed. This motivates the question of whether formal verification guarantees can be obtained directly from observations.

In this work, we consider a POMDP under a fixed finite-state controller (FSC), where the latent transition and observation models are unknown. Given data from either independently generated observation trajectories or a single stationary, geometrically mixing trajectory, together with a specification automaton, our objective is to estimate the probability that the observation sequence satisfies the specification and to quantify the reliability of this estimate. Based on the observations, we develop a data-driven spectral verification method that learns a weighted finite automaton (WFA) of the word probability function and combines it with a deterministic finite automaton (DFA) to estimate the finite-horizon satisfaction probability with a finite-sample error guarantee.

The most closely related work addresses three components of our problem: model-based verification under partial observability, data-driven verification from sampled executions or learned probabilistic models, and spectral identification of latent sequential processes.

Related work: Model-based verification under partial observability. Model-based verification under partial observability differs primarily in how the information contained in observation histories is represented. Norman et al. (2017) constructed the continuous belief MDP and approximated it using a finite-grid abstraction, thereby obtaining lower and upper bounds on quantitative properties. Point-based methods, such as Bouton et al. (2020), approximated the belief-space solution at selected belief

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

points and used a product POMDP to synthesize policies satisfying temporal-logic specifications. An alternative is to represent observation histories through finite controller memory. Carr et al. (2021) trained an RNN-based policy, extracted a finite-state controller, and verified the Markov chain induced by composing the controller with the known POMDP. In this case, belief-related information is represented implicitly in the recurrent and controller memory rather than as an explicit belief state. At the level of a fixed hidden stochastic process, Zhang et al. (2005) introduced an observation-aware temporal logic for hidden Markov models (HMMs) and reduced its model-checking problem to probabilistic reachability in a product automaton.

These approaches differ in whether partial information is represented through beliefs or through controller memory, but their verification starts from a known transition or observation model. Our work formulates the verification problem at the level of the stochastic observation language induced by the closed loop. Rather than reconstructing or discretizing the belief space, we learn a finite-dimensional linear representation of the word-probability function and verify the DFA specification on the learned representation. This leads to a data-driven verification procedure for the case in which the POMDP model is unavailable.

Related work: Data-driven probabilistic verification and model checking for weighted automata. Data-driven verification methods can be organized according to whether they check a property directly from sampled executions, first learn a probabilistic model, or are given a weighted-automaton representation. Statistical model checking estimates or tests the satisfaction probability of a specified property directly from executions, without constructing an explicit system model. Younes & Simmons (2002) proposed a model-independent procedure for continuous stochastic logic properties using sequential acceptance sampling with prescribed error bounds. Sen et al. (2004) considered fixed traces obtained from passively observed black-box systems and allowed an inconclusive outcome when the statistical evidence was insufficient. An alternative line of work first learns a probabilistic model from execution data and subsequently applies conventional probabilistic model checking. Mao et al. (2016) studied the learning of deterministic probabilistic automata for verification and established convergence of the resulting temporal-property probabilities, while Wang et al. (2017) empirically compared this learn-then-check approach with direct statistical model checking. For quantitative estimation, Hérault et al. (2004) obtained an additive (ε, δ) approximation of positive-LTL satisfaction probabilities from sampled executions. Bazille et al. (2020) further derived finite-sample PAC bounds on the difference between temporal-property probabilities in an original discrete-time Markov chain (DTMC) and a learned one. In a complementary model-based direction, Buchholz & Kemper (2010) developed a general model-checking framework for explicitly given weighted automata using a semiring-based temporal logic and matrix computations.

Our method establishes a learn-then-check procedure tailored to partially observable closed-loop systems. The learned object is not a DTMC over observed system states, but a low-rank WFA realization of the generally non-Markovian observation process. Verification is then performed by synchronizing this realization with a DFA specification. Most importantly, the output is not merely a learned representation: our analysis provides an explicit finite-sample confidence bound for the resulting satisfaction-probability estimator.

Related work: Spectral learning for weighted automata. Spectral and moment-based methods provide non-iterative approaches for learning latent sequential models from observable statistics. Hsu et al. (2012) developed an SVD-based algorithm that approximates the joint distribution over observation sequences of finite length and the conditional distribution over a future observation, conditioned on some history of observations. In the automata setting, low-rank Hankel factorizations yield finite-dimensional weighted-automaton representations for stochastic languages (Balle & Mohri, 2012). Because an unconstrained spectral estimate may fail to define a proper probability distribution, Glaude & Pietquin (2016) introduced a constrained method-of-moments procedure that returns a proper probabilistic automaton with a PAC-style guarantee. Extensions to controlled and dependent-data settings have followed several related directions. Azizzadenesheli et al. (2016) used spectral methods to estimate the transition, observation, and reward parameters of a POMDP from trajectories generated under fixed memoryless policies. Balle & Maillard (2017) established spectral methods of moments for learning sequential models from a single trajectory, instead of using multiple i.i.d. trajectories. This approach leverages an SVD-based learning algorithm for weighted automata by using dependent data. Rácz et al. (2025) provided a closely related system-identification result for partially observed linear switched systems. They estimated Markov parameters from a single trajectory and applied a Ho-Kalman-type realization algorithm to obtain finite-sample bounds on the recovered system parameters.

The spectral-learning results above provide guarantees for observable sequence distributions, latent-

model parameters, Markov parameters, or recovered realization matrices. We use spectral realization as the estimation mechanism of a verification method and extend the analysis to the property level. Specifically, we propagate the errors of empirical word probabilities and the learned Hankel realization through the DFA-WFA product, yielding a finite-sample bound on the final satisfaction-probability error. The learned realization is therefore not used only for prediction or system reconstruction; it becomes the basis of a sound verification certificate for the original closed-loop process.

Summary of the contributions of the paper. With respect to the related work, the main contributions of the paper are: **(1)** Recognizing that for the purposes of safety verification with a fixed finite-state controller, unknown POMDPs can be modeled as WFAs; **(2)** Leveraging the existing finite-sample bounds for WFAs to provide finite-sample bounds on safety certificates of unknown POMDPs; **(3)** Showing that safety certificates of POMDPs are PAC learnable.

The paper is organized as follows. In Section 2, we define our problem. Next, we present our main result in Section 3 and discussion of the main theorem in Section 4. The proofs and numerical analysis are given in Appendices.

2 Problem Formulation

2.1 Notation

For a finite set $\mathcal{A}$, $|\mathcal{A}|$ denotes its cardinality and $\Delta(\mathcal{A})$ denotes the probability simplex over $\mathcal{A}$. We use boldface characters to denote random quantities, including random variables and stochastic processes. All vectors are column vectors unless explicitly transposed. We write I_n for the $n \times n$ identity matrix, e_i for the i-th standard basis vector, and $\mathbf{1}_n$ for the all-ones vector in $\mathbb{R}^n$. For an event E, $\mathbf{1}_E$ denotes its indicator.

For a matrix A, $\sigma_n(A)$ denotes its smallest singular value. The spectral and Frobenius norms are denoted by $\|A\|_2$ and $\|A\|_F$. The Kronecker product is denoted by $\otimes$.

Let $\mathcal{O}^T$ denote the set of words of length T over the observation alphabet $\mathcal{O}$, and let $\mathcal{O}^* := \bigcup_{T\geq 0} \mathcal{O}^T$ denote the set of all finite-length words over $\mathcal{O}$. The empty word is denoted by ϵ. Let $\mathbf{O}_{1:t} := (\mathbf{O}_1, \ldots, \mathbf{O}_t)$ denote the random variable sequence of observations from time step 1 to t. A word $w := o_1 \cdots o_k$ refers to the sequence of observations $(o_1, \ldots, o_k)$. The length of the word is denoted by $|w|$. For any words u and v, uv denotes their concatenation, and $u \preceq v$ denotes that u is a prefix of v.

Throughout this paper, all the random variables are understood on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and all stochastic processes are over discrete time $\mathbb{N}$.

2.2 POMDP under a Finite-State Controller

We consider the controlled dynamics of an unknown partially observable Markov decision process (POMDP). A POMDP is a tuple $\mathcal{P} = \langle \mathcal{X}, \mathcal{U}, \mathcal{O}, b_0, P, Z, \{\mathbf{X}_t\}_{t\geq 0}, \{\mathbf{U}_t\}_{t\geq 0}, \{\mathbf{O}_t\}_{t\geq 1} \rangle$, where $\mathcal{X}, \mathcal{U}$, $\mathcal{O}$ are finite state, action, observation sets respectively, and $\{\mathbf{X}_t\}_{t\geq 0}$, $\{\mathbf{U}_t\}_{t\geq 0}$, $\{\mathbf{O}_t\}_{t\geq 1}$ are discrete-time stochastic processes over probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $\mathcal{X}$, $\mathcal{U}$, $\mathcal{O}$ respectively. $b_0 : \mathcal{X} \to [0,1]$, $P : \mathcal{X} \times \mathcal{U} \times \mathcal{X} \to [0,1]$, $Z : \mathcal{X} \times \mathcal{U} \times \mathcal{O} \to [0,1]$ are maps such that the following conditions hold: 1) $\mathbb{P}(\mathbf{X}_0 = x) = b_0(x)$ for all $x \in \mathcal{X}$. 2) $\sum_{x' \in \mathcal{X}} P(x' \mid x, u) = 1$ for all (x, u), such that $\mathbb{P}(\mathbf{X}_{t+1} = x' \mid \sigma(\mathbf{X}_{0:t}, \mathbf{U}_{0:t}, \mathbf{O}_{1:t})) = P(x' \mid \mathbf{X}_t, \mathbf{U}_t)$ $\mathbb{P}$-a.s., $\forall\, t \geq 0$, $\forall\, x' \in \mathcal{X}$. 3) $\sum_{o \in \mathcal{O}} Z(o \mid x, u) = 1$ for all (x, u), such that $\mathbb{P}(\mathbf{O}_{t+1} = o' \mid \sigma(\mathbf{X}_{0:t}, \mathbf{U}_{0:t}, \mathbf{O}_{1:t}), \mathbf{X}_{t+1}) = Z(o' \mid \mathbf{X}_{t+1}, \mathbf{U}_t)$ $\mathbb{P}$-a.s., $\forall\, t \geq 0$, $\forall\, o' \in \mathcal{O}$. We call P the state transition kernel and Z the observation kernel, and they are assumed to be unknown. Control actions are generated by a probabilistic finite-state controller (FSC). A probabilistic FSC is a tuple $\mathcal{C} = \langle \mathcal{S}, \nu_0, \varphi, \eta \rangle$, where $\mathcal{S}$ is a finite controller state space, such that $\varphi : \mathcal{S} \to \Delta(\mathcal{U})$ is the action-selection kernel, $\eta : \mathcal{S} \times \mathcal{O} \to \Delta(\mathcal{S})$ is the controller-state update kernel, and $\nu_0 \in \Delta(\mathcal{S})$ is the initial distribution of FSC states. Intuitively, at time t, the FSC applies $\mathbf{U}_t \sim \varphi(\cdot \mid \mathbf{S}_t)$ and, after receiving observation $\mathbf{O}_{t+1}$, updates its internal state according to $\mathbf{S}_{t+1} \sim \eta(\cdot \mid \mathbf{S}_t, \mathbf{O}_{t+1})$. To specify the closed-loop dynamics, we call a POMDP $\mathcal{P}$ closed-loop under a probabilistic FSC $\mathcal{C}$ if there exists a process $\{\mathbf{S}_t\}_{t\geq 0}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in $\mathcal{S}$, such that $\{\mathbf{X}_t\}_{t\geq 0}$, $\{\mathbf{U}_t\}_{t\geq 0}$, $\{\mathbf{O}_t\}_{t\geq 1}$ satisfy the conditions in the definition of POMDPs, and in addition, $\mathbb{P}(\mathbf{S}_0 = s) = \nu_0(s)$, $\mathbb{P}(\mathbf{S}_{t+1} = s' \mid \sigma(\mathbf{S}_{0:t}, \mathbf{O}_{1:t+1})) = \eta(s' \mid \mathbf{S}_t, \mathbf{O}_{t+1})$ $\mathbb{P}$-a.s., $\mathbb{P}(\mathbf{U}_t = u \mid \sigma(\mathbf{S}_{0:t})) = \varphi(u \mid \mathbf{S}_t)$ $\mathbb{P}$-a.s..

For $\{\mathbf{X}_t\}_{t\geq 0}$ and $\{\mathbf{S}_t\}_{t\geq 0}$, define the closed-loop hidden process as $\{\mathbf{Q}_t\}_{t\geq 0} = \{(\mathbf{X}_t, \mathbf{S}_t)\}_{t\geq 0}$, and the product space for $\{\mathbf{Q}_t\}_{t\geq 0}$ is $\mathcal{Q} := \mathcal{X} \times \mathcal{S}$. Given an observation o', for $q = (x, s)$ and

$q' = (x', s')$, the observation-labeled transition kernel is

$$T^{\mathcal{Q}}_{o'}(q' \mid q) := \mathbb{P}(\mathbf{Q}_{t+1} = (x', s'), \mathbf{O}_{t+1} = o' \mid \mathbf{Q}_t = (x, s)) = \sum_{u \in \mathcal{U}} \varphi(u \mid s) P(x' \mid x, u) Z(o' \mid x', u) \eta(s' \mid s, o'). \tag{1}$$

Moreover, $T^{\mathcal{Q}} := \sum_{o \in \mathcal{O}} T^{\mathcal{Q}}_o$ is the transition kernel on the state space $\mathcal{Q}$. An edge-emitting hidden Markov model (edge-emitting HMM) (Travers, 2014) is a tuple $\langle \mathcal{Q}, \mathcal{O}, \{T^{\mathcal{Q}}_o\}_{o \in \mathcal{O}}, \mu_0, \{\mathbf{Q}_t\}_{t \geq 0}, \{\mathbf{O}_t\}_{t \geq 1} \rangle$, such that $\mathbb{P}(\mathbf{Q}_0 = q) = \mu_0(q), \forall q \in \mathcal{Q}$, and $\mathbb{P}(\mathbf{Q}_{t+1} = q', \mathbf{O}_{t+1} = o' \mid \sigma(\mathbf{Q}_{0:t})) = T^{\mathcal{Q}}_{o'}(q' \mid \mathbf{Q}_t),\ \mathbb{P}$-a.s., $\forall q' \in \mathcal{Q}, o' \in \mathcal{O}$.

2.3 Finite-Horizon Safety Specification

We define safety on the space of the output observation processes of the closed-loop POMDP, and a safety specification on it is given as follows.

Let the specification be represented by a deterministic finite automaton (DFA) $\mathcal{A}_s = \langle \mathcal{R}, \mathcal{O}, \delta_s, r_0, F \rangle$, where $\mathcal{R}$ is the finite automaton state set, $\mathcal{O}$ is the input symbol alphabet, $r_0 \in \mathcal{R}$ is the initial state, $\delta_s : \mathcal{R} \times \mathcal{O} \to \mathcal{R}$ is the transition function, and $F \subseteq \mathcal{R}$ is the set of accepting states.

By a slight abuse of notation, we also use δ_s to denote its natural extension from observation symbols to finite words. For a word $w = o_1 \cdots o_T$, the automaton evolves according to $r_t = \delta_s(r_{t-1}, o_t)$, $t = 1, \ldots, T$. Equivalently, $r_T = \delta_s(r_0, w)$.

Let $\mathcal{L}_{\text{safe}}$ be the safety set of observation sequences. The specification automaton $\mathcal{A}_s$ recognizes $\mathcal{L}_{\text{safe}}$ if and only if $\delta_s(r_0, w) \in F$, i.e., $\mathcal{L}_{\text{safe}} := \{w \in \mathcal{O}^* : \delta_s(r_0, w) \in F\}$.

We consider safety under two different regimes. The first is *transient finite-horizon safety*: starting from a certain initial distribution, we evaluate the probability that the system satisfies the safety specification over the next T time steps. The second is *stationary safety*: assuming that the closed-loop process admits a stationary distribution, we evaluate the same T step safety probability when the system is initialized according to that stationary distribution. For a fixed time horizon T, define

$$p_T := \mathbb{P}(\mathbf{O}_{1:T} \in \mathcal{L}_{\text{safe}}). \tag{2}$$

Definition 2.1 (Safety)**.** Given a prescribed risk tolerance $\delta_{\text{unsafe}} \in (0, 1)$, equivalently, a safety threshold $1 - \delta_{\text{unsafe}}$, the system is declared safe over horizon T if $p_T \geq 1 - \delta_{\text{unsafe}}$.

2.4 Data-Driven Verification Problem

In our setting, the transition and observation kernels of the POMDP are unknown, and the augmented hidden states $\{\mathbf{Q}_t\}_{t \geq 0}$ are not observable. Consequently, the verification procedure has access only to samples of output observation trajectories generated by the closed-loop system. Informally, we will learn either from N independent samples of length-L drawn from the joint distribution of the output process, or from one single sample trajectory of the output process. Formally, let us define the set of possible samples $\mathcal{D}^N_{\text{iid}} := \underbrace{\mathcal{O}^L \times \cdots \times \mathcal{O}^L}_{N \text{ times}}$ and $\mathcal{D}^N_{\text{st}} := \mathcal{O}^{N+L-1}$. The data used for learning are thus samples of the following random variables with values in $\mathcal{D}^N_*$, $* \in \{\text{iid}, \text{st}\}$:

$$\mathbf{D}^N_* = \begin{cases} \left(\mathbf{O}^{(1)}_{1:L}, \cdots, \mathbf{O}^{(N)}_{1:L}\right) & \text{if } * = \text{iid} \\ \mathbf{O}_{1:N+L-1} & \text{if } * = \text{st}, \end{cases} \tag{3}$$

where $\{\mathbf{O}^{(\ell)}_t\}_{t \geq 1}$ are processes with values in $\mathcal{O}$, such that σ-algebra generated by $\{\mathbf{O}^{(\ell_1)}_t\}_{t \geq 1}$ and $\{\mathbf{O}^{(\ell_2)}_t\}_{t \geq 1}$ are independent for any $\ell_1 \neq \ell_2$, and the joint probability distribution of $\{\mathbf{O}^{(\ell)}_t\}_{t \geq 1}$ is the same as $\{\mathbf{O}_t\}_{t \geq 1}$, $\forall \ell = 1, \ldots, N$. Then a sample used for learning is a sample of $\mathbf{D}^N_*$, that is $\mathbf{D}^N_*(\omega)$, for some $\omega \in \Omega$.

In order to learn from a single sample trajectory of $\{\mathbf{O}_t\}_{t \geq 1}$, we need to assume mixing of $\{\mathbf{O}_t\}_{t \geq 1}$. More precisely, we assume geometrically mixing of the underlying HMM of the closed-loop system.

Assumption 2.2 (Geometrically mixing in the single-trajectory setting)**.** Recall from Section 2.2, the definition of $T^{\mathcal{Q}}$, and the process $\{\mathbf{Q}_t\}_{t \geq 0} = \{(\mathbf{X}_t, \mathbf{S}_t)\}_{t \geq 0}$, we assume that there exist constants $\mathscr{C} > 0$ and $\theta \in (0, 1)$ such that for every $k > 0$, $\sup_{\substack{\nu, \nu' \in \Delta(\mathcal{Q}) \\ \nu \neq \nu'}} \frac{\|(T^{\mathcal{Q}})^k \nu - (T^{\mathcal{Q}})^k \nu'\|_1}{\|\nu - \nu'\|_1} \leq \mathscr{C} \theta^k$. We also

assume that $\{\mathbf{Q}_t\}_{t\geq 0}$ is stationary and its probability distribution $\nu(q) = \mathbb{P}(\mathbf{Q}_t = q)$ is the unique distribution such that $T^{\mathcal{Q}}\nu = \nu$.

It is well-known that Assumption 2.2 implies that $\{\mathbf{Q}_t\}_{t\geq 0}$ is an ergodic Markov chain.

Remark 2.3. If $\{\mathbf{O}_t\}_{t\geq 1}$ is generated by the closed-loop POMDP under the fixed FSC, and its augmented hidden state process satisfies Assumption 2.2, then there exists a stationary process with the same transition and observation kernels. Denotes its observation process by $\{\widetilde{\mathbf{O}}_t\}_{t\geq 1}$. For any $k \geq 1$, the joint probability distribution of $\mathbf{O}_{t+1}, \ldots, \mathbf{O}_{t+k}$ converges to the joint probability distribution of $\widetilde{\mathbf{O}}_1, \ldots, \widetilde{\mathbf{O}}_k$ as $t \to \infty$. The latter distribution is the same as the joint probability distribution due to stationarity. Hence, for large enough t, the probability distribution of $\mathbf{O}_{t+1}, \ldots, \mathbf{O}_{t+k}$ and $\widetilde{\mathbf{O}}_{t+1}, \ldots, \widetilde{\mathbf{O}}_{t+k}$ are almost indistinguishable. In particular, the safety probability of a late length-T window approaches the stationary safety p_T, which motivates our focus on the stationary regime. Unless stated otherwise, all subsequent results for the single-trajectory setting are understood to be under Assumption 2.2, which will not be restated explicitly.

The goal is to find a decision rule to determine based on data if the system is safe. This decision rule has to be a so-called safety certificate, defined below:

Definition 2.4 (Data-driven safety certificate)**.** Let $\mathfrak{D}_{\text{iid}}^N$ be the space of datasets containing N i.i.d. sampling trajectories, and $\mathfrak{D}_{\text{st}}^N$ be the space of datasets containing single trajectory with N windows. A data-driven safety certificate (safety certificate for short) is a $(\delta, \delta_{\text{unsafe}})$-measurable map $\Gamma_N : \mathfrak{D}_*^N \to \{0, 1\}$, $* \in \{\text{iid}, \text{st}\}$ such that for a dataset $\mathcal{D}_*^N \in \mathfrak{D}_*^N$, 1) if $p_T < 1 - \delta_{\text{unsafe}}$ (the underlying POMDP is unsafe), then $\mathbb{P}(\Gamma_N(\mathbf{D}_*^N) = 1) < \delta$, 2) if $p_T > 1 - \delta_{\text{unsafe}}$ (the underlying POMDP is safe), then $\mathbb{P}(\Gamma_N(\mathbf{D}_*^N) = 0) < \delta$. Intuitively, Γ_N is a rule based on data to decide if the underlying system is safe: $\Gamma_N(\mathbf{D}_*^N) = 1$ means that the rule predicts that the system is safe, and $\Gamma_N(\mathbf{D}_*^N) = 0$ means that the rule predicts that the system is unsafe.

The safety certificate can be viewed as hypothesis testing procedure with significance δ, and power $1 - \delta$, with the null hypothesis that the system is unsafe, i.e., it is a balanced test. This means that we can also use the test to reject safety (by checking of $\Gamma_N = 0$) with significance δ and power $1 - \delta$.

3 Main Result

The main result of this paper is stated as follows.

Theorem 3.1 (PAC learnability of safety certificates)**.** *Let $\delta, \delta_{\text{unsafe}} \in (0, 1)$. For a fixed verification horizon T, there exists a finite sample threshold $m_{\text{cert}}(\delta, \delta_{\text{unsafe}})$ such that, for every $N \geq m_{\text{cert}}(\delta, \delta_{\text{unsafe}})$, there exists a data-driven safety certificate Γ_N at confidence level $1 - \delta$. In fact, we can take $m_{\text{cert}}(\delta, \delta_{\text{unsafe}}) = \frac{c(1+\sqrt{\ln(1/\delta)})^2}{(p_T - (1-\delta_{\text{unsafe}}))^2}$ for a suitable constant c.*

That is, the theorem says that we can always find a safety certificate given enough data points, and this safety certificate arises via spectral learning for a suitable choice of δ and δ_{unsafe}. The sample complexity of $m_{\text{cert}}(\delta, \delta_{\text{unsafe}})$ is $O(\ln(1/\delta))$ and $O\left(1/(p_T - (1 - \delta_{\text{unsafe}}))^2\right)$. Therefore, achieving a higher confidence level, corresponding to a smaller δ, or certifying a system with a higher safety threshold requires more samples.

The proof of the main theorem is constructive, as it relies on proposing a particular decision rule Γ_N, and a particular constant c, such that for $N \geq \frac{c(1+\sqrt{\ln(1/\delta)})^2}{(p_T - (1-\delta_{\text{unsafe}}))^2}$, Γ_N is a safety certificate with significance $1 - \delta$. Γ_N will rely on estimating p_T from data, by learning a WFA which generates the output distribution of the closed-loop POMDP.

3.1 WFA Representation of POMDPs

In this section, we show that the safety probability in Equation 2 admits a linear representation.

We define the word probability function $M : \mathcal{O}^* \to [0, 1]$ by $M(w) := \mathbb{P}(\mathbf{O}_{1:|w|} = w)$, with $M(\epsilon) := 1$. If $\{\mathbf{O}_t\}_{t\geq 1}$ is stationary, then $M(w) = \mathbb{P}(\mathbf{O}_{s+1:s+|w|} = w)$ for any $s \in \mathbb{N}$. Then p_T defined in Equation 2 can be written as $p_T = \sum_{w \in \mathcal{L}_{\text{safe}} \cap \mathcal{O}^T} M(w)$.

For the stochastic process $\{\mathbf{Q}_t\}_{t\geq 0}$, let $d := |\mathcal{Q}|$, and enumerate the augmented hidden-state space

as $\mathcal{Q} = \{q_1, \ldots, q_d\}$. For each observation symbol $o \in \mathcal{O}$, define the observation-labeled transition matrix $A_o^{\mathcal{Q}} \in \mathbb{R}^{d \times d}$ by $[A_o^{\mathcal{Q}}]_{ji} := T_o^{\mathcal{Q}}(q_j \mid q_i)$. For a word $w = o_1 \cdots o_k$ define $A_w^{\mathcal{Q}} := A_{o_k}^{\mathcal{Q}} \cdots A_{o_1}^{\mathcal{Q}}$, $A_\epsilon := I_d$, then we have $M(w) = \mathbf{1}_d^\top A_w^{\mathcal{Q}} \mu_0$.

The matrix representation of word probability can be expressed as a weighted finite automaton (WFA). A weighted finite automaton of dimension r over a finite alphabet $\mathcal{O}$ is a tuple $\mathcal{A} = \langle C, \{A_o\}_{o \in \mathcal{O}}, B \rangle$, where $B \in \mathbb{R}^{r \times 1}$, $C \in \mathbb{R}^{1 \times r}$, and $A_o \in \mathbb{R}^{r \times r}$. It realizes the function $M_{\mathcal{A}} : \mathcal{O}^* \to \mathbb{R}$ defined by $M_{\mathcal{A}}(w) = C A_w B$. Consequently, $\mathcal{A}_{\mathcal{Q}} := \langle \mathbf{1}_d^\top, \{A_o^{\mathcal{Q}}\}_{o \in \mathcal{O}}, \mu_0 \rangle$ is a d-dimensional WFA realization of M.

Although this realization follows directly from the closed-loop HMM, it is expressed in the unknown hidden state coordinates and need not be minimal. A WFA realization of M is called minimal if no WFA of smaller dimension realizes the same function. To identify such a realization without referring to the hidden states, we introduce the Hankel matrix $H_M \in \mathbb{R}^{\mathcal{O}^* \times \mathcal{O}^*}$, whose rows are indexed by suffix words and columns by prefix words: $[H_M]_{v,u} := M(uv),\ u, v \in \mathcal{O}^*$.

By the Hankel-rank characterization of weighted automata, the dimension of a minimal WFA realization of M is $n := \operatorname{rank}(H_M)$. Since M admits the d-dimensional realization $\mathcal{A}_{\mathcal{Q}}$, we have $n \le d$. Choose suffix and prefix sets

$$\alpha = \{v_1, \ldots, v_n\}, \qquad \beta = \{u_1, \ldots, u_n\}, \tag{4}$$

such that the finite Hankel block $[H_{\alpha,\beta}]_{ij} = M(u_j v_i)$ has rank n. For each $o \in \mathcal{O}$, define the shifted Hankel block $H_{\alpha,o,\beta} \in \mathbb{R}^{n \times n}$ and boundary vectors $h_{\alpha,\epsilon} \in \mathbb{R}^{n \times 1}, h_{\epsilon,\beta} \in \mathbb{R}^{1 \times n}$ by $[H_{\alpha,o,\beta}]_{ij} := M(u_j o v_i),\ [h_{\alpha,\epsilon}]_i := M(v_i),\ [h_{\epsilon,\beta}]_j := M(u_j)$. A square full-rank form of the Hankel realization construction gives $A_o = H_{\alpha,\beta}^{-1} H_{\alpha,o,\beta},\ B = H_{\alpha,\beta}^{-1} h_{\alpha,\epsilon},\ C = h_{\epsilon,\beta}$. The resulting WFA $\mathcal{A} := \langle C, \{A_o\}_{o \in \mathcal{O}}, B \rangle$ is a minimal realization of M and satisfies $M(w) = C A_w B$ for a word $w = o_1 \cdots o_k$, with $A_w := A_{o_k} \cdots A_{o_1}$ and $A_\epsilon := I_n$. Different full-rank choices of (α, β) produce isomorphic minimal realizations and therefore give the same word probability and the same value of p_T (see Proposition C.1). Hence, we fix one choice of pair (α, β). Only a finite collection of word probabilities is required to construct this realization. Let $\mathcal{W}_{\alpha,\beta} := \{u_j v_i, u_j o v_i, u_j, v_i : i, j \in \{1, \ldots, n\}, o \in \mathcal{O}\}$, and define

$$L_H := \max_{w \in \mathcal{W}_{\alpha,\beta}} |w|. \tag{5}$$

3.2 Learning WFA

In this section, we construct an empirical WFA realization from observation data under the two sampling regimes. To this end, we first define empirical estimates of the word probabilities required to form the Hankel matrices. For every non-empty $w \in \mathcal{W}_{\alpha,\beta}$, its empirical probability is defined by

$$\widehat{\mathbf{M}}(w) = \begin{cases} \frac{1}{N} \sum_{\ell=1}^{N} \mathbf{1}_{\left\{\mathbf{O}_{1:|w|}^{(\ell)} = w\right\}}, & \text{for } N \text{ independent trajectories,} \\ \frac{1}{N} \sum_{s=0}^{N-1} \mathbf{1}_{\left\{\mathbf{O}_{s+1:s+|w|} = w\right\}}, & \text{for a single trajectory.} \end{cases}$$

We set $\widehat{\mathbf{M}}(\epsilon) := 1$. Both sampling regimes require $L \ge L_H$.

The empirical quantities $\widehat{\mathbf{H}}_{\alpha,\beta},\ \widehat{\mathbf{H}}_{\alpha,o,\beta},\ \widehat{\mathbf{h}}_{\alpha,\epsilon},\ \widehat{\mathbf{h}}_{\epsilon,\beta}$ are obtained by replacing every occurrence of $M(w)$ in the corresponding quantities by $\widehat{\mathbf{M}}(w)$. Whenever $\widehat{\mathbf{H}}_{\alpha,\beta}$ is nonsingular, the empirical WFA is defined by $\widehat{\mathbf{A}}_o = \widehat{\mathbf{H}}_{\alpha,\beta}^{-1} \widehat{\mathbf{H}}_{\alpha,o,\beta},\ \widehat{\mathbf{B}} = \widehat{\mathbf{H}}_{\alpha,\beta}^{-1} \widehat{\mathbf{h}}_{\alpha,\epsilon},\ \widehat{\mathbf{C}} = \widehat{\mathbf{h}}_{\epsilon,\beta}$. Thus, $\widehat{\mathcal{A}} = \langle \widehat{\mathbf{C}}, \{\widehat{\mathbf{A}}_o\}_{o \in \mathcal{O}}, \widehat{\mathbf{B}} \rangle$ provides a data-driven estimate of the word probability function. For a word $w = o_1 \cdots o_k$, define $\widehat{\mathbf{A}}_w := \widehat{\mathbf{A}}_{o_k} \cdots \widehat{\mathbf{A}}_{o_1}$, and let $\widehat{\mathbf{A}}_\epsilon := I_n$. The word probability estimated by the empirical WFA realization is $\widehat{\mathbf{C}} \widehat{\mathbf{A}}_w \widehat{\mathbf{B}}$.

3.3 Estimating p_T

We now use the learned WFA to estimate the safety probability p_T. Recall the safety specification automaton $\mathcal{A}_s$ introduced in Section 2.3. Assume that it has D_s states, and fix an enumeration $\mathcal{R} = \{1, 2, \ldots, D_s\}$. For each $r \in \mathcal{R}$, let $e_r \in \mathbb{R}^{D_s}$ denote the r-th standard basis vector. For each observation symbol o, define the specification-automaton transition matrix $E_o \in \{0,1\}^{D_s \times D_s}$ by $[E_o]_{ij} := \mathbf{1}_{\{\delta_s(j,o)=i\}}$, hence $E_o e_r = e_{\delta_s(r,o)}$. Define the accepting-state indicator vector $\chi_F \in \{0,1\}^{D_s}$ by $[\chi_F]_r := \mathbf{1}_{\{r \in F\}}$.

Consider the exact WFA realization $\langle C, \{A_o\}_{o\in\mathcal{O}}, B\rangle$ of dimension n. Define the Kronecker product realization $\langle \widetilde{C}, \{\widetilde{A}_o\}_{o\in\mathcal{O}}, \widetilde{B}\rangle$ by $\widetilde{A}_o := E_o \otimes A_o,\ \forall o \in \mathcal{O},\ \widetilde{B} := e_{r_0} \otimes B,\ \widetilde{C} := \chi_F^\top \otimes C$.
The product realization admits a direct block interpretation. If the specification automaton is in state r and the WFA state is $z \in \mathbb{R}^n$, the corresponding product state is $\widetilde{z} = e_r \otimes z$. Upon reading an observation symbol o, $\widetilde{A}_o\widetilde{z} = (E_o \otimes A_o)(e_r \otimes z) = (E_o e_r) \otimes (A_o z) = e_{\delta_s(r,o)} \otimes A_o z$. Thus, the symbol o simultaneously updates the WFA state through A_o and moves the resulting linear state to the block associated with the next specification automaton state $\delta_s(r, o)$.
The following proposition shows that the DFA-WFA product provides a linear representation of the probability of each word satisfying the safety specification. For the proof, see Appendix B.

Proposition 3.2. *For every word $w = o_1 \cdots o_k \in \mathcal{O}^*$, define $E_w := E_{o_k} \cdots E_{o_1}$ and $\widetilde{A}_w := \widetilde{A}_{o_k} \cdots \widetilde{A}_{o_1}$. Then $\widetilde{C}\widetilde{A}_w\widetilde{B} = \mathbf{1}_{\{w\in\mathcal{L}_{\text{safe}}\}} M(w)$.*

Consequently, for a fixed verification horizon T, $p_T = \sum_{w\in\mathcal{L}_{\text{safe}}\cap\mathcal{O}^T} M(w) = \sum_{w\in\mathcal{O}^T} \widetilde{C}\widetilde{A}_w\widetilde{B}$. Define an aggregated product operator $\widetilde{G} := \sum_{o\in\mathcal{O}} \widetilde{A}_o$. Then $\widetilde{G}^T = \sum_{w\in\mathcal{O}^T} \widetilde{A}_w$, and therefore $p_T = \widetilde{C}\widetilde{G}^T\widetilde{B}$. Replacing the exact Hankel realization by the empirical realization $\widehat{\mathcal{A}} := \langle \widehat{\mathbf{C}}, \{\widehat{\mathbf{A}}_o\}_{o\in\mathcal{O}}, \widehat{\mathbf{B}}\rangle$, define $\widehat{\widetilde{\mathbf{A}}}_o := E_o \otimes \widehat{\mathbf{A}}_o$, for $o \in \mathcal{O}$, $\widehat{\widetilde{\mathbf{B}}} := e_{r_0} \otimes \widehat{\mathbf{B}}$, $\widehat{\widetilde{\mathbf{C}}} := \chi_F^\top \otimes \widehat{\mathbf{C}}$. Let $\widehat{\widetilde{\mathbf{G}}} := \sum_{o\in\mathcal{O}} \widehat{\widetilde{\mathbf{A}}}_o$. The estimated finite-horizon satisfaction probability is

$$\widehat{\mathbf{p}}_T = \widehat{\widetilde{\mathbf{C}}}\,\widehat{\widetilde{\mathbf{G}}}^{\,T}\,\widehat{\widetilde{\mathbf{B}}}. \tag{6}$$

3.4 Proof of Theorem 3.1: Construction of Γ_N

We now construct a data-driven safety certificate by $\widehat{\mathbf{p}}_T$ and a finite-sample error bound. In particular, for a suitable choice of a function $\varepsilon_T : (0,1) \times \mathbb{N} \to (0,\infty)$, the estimated $\widehat{\mathbf{p}}_T$ in Equation 6 can be used to construct a measurable map $\widehat{\Gamma}^{\text{Spect}}_{N,\varepsilon_T} : \mathfrak{D}_*^N \to \{0,1\}$ defined as follows,

$$\widehat{\Gamma}^{\text{Spect}}_{N,\varepsilon_T}(\mathbf{D}_*^N) = \begin{cases} 1, & \text{if } \widehat{\mathbf{p}}_T - \varepsilon_T(\delta, N) > 1 - \delta_{\text{unsafe}}, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

We will show that for a suitable choice of ε_T, for large enough N, $\widehat{\Gamma}^{\text{Spect}}_{N,\varepsilon_T}$ is a $(\delta, \delta_{\text{unsafe}})$-safety certificate by finding $\varepsilon_T(\delta, N)$ such that $\mathbb{P}(|\widehat{\mathbf{p}}_T - p_T| \le \varepsilon_T(\delta, N)) \ge 1 - \delta$. To this end, we need the following quantities.
Recall the suffix and prefix sets introduced in Equation 4, and define the enlarged Hankel block $H_{\alpha',\beta'}$, for $\alpha' := \alpha \cup \{\epsilon\},\ \beta' := \{\epsilon\} \cup \beta \cup \left(\bigcup_{o\in\mathcal{O}} \beta o\right)$, where βo denotes the concatenation of $u_i \in \beta$ and o. Notice that $H_{\alpha,\beta}, \{H_{\alpha,o,\beta}\}_{o\in\mathcal{O}}, h_{\alpha,\epsilon}, h_{\epsilon,\beta}$ are sub-blocks of $H_{\alpha',\beta'}$, for the selected α and β. Moreover, $\text{rank}(H_{\alpha',\beta'}) = n$. Define

$$c_{\alpha',\beta'} := \max_{w\in\mathcal{O}^L} \sum_{v'_i\in\alpha'} \sum_{u'_j\in\beta'} \mathbf{1}_{\{u'_j v'_i \preceq w\}}, \quad n_{\alpha'} := \#\{|v'| : v' \in \alpha'\},\ n_{\beta'} := \#\{|u'| : u' \in \beta'\}, \tag{8}$$

$$d' := \min\left\{|\alpha'||\beta'|, 2n_{\alpha'}n_{\beta'}\right\}, \quad m := \sum_{v'_i\in\alpha'} \sum_{u'_j\in\beta'} M(u'_j v'_i). \tag{9}$$

Here, $c_{\alpha',\beta'}$ is the maximum number of selected concatenations $u'_j v'_i$ that can simultaneously be prefixes of a length-L word. The quantities $n_{\alpha'}$ and $n_{\beta'}$ count the distinct word lengths represented in α' and β', respectively, while d' is an effective dimension parameter that accounts for both the number of selected Hankel entries and their length structure. m is the total probability mass of the selected concatenations.
Let us define the constant $K_*(\delta, N)$ for two different sampling regimes. For the i.i.d setting, define

$$K_{\text{iid}}(\delta, N) := \sqrt{\frac{c_{\alpha',\beta'}}{N}} \left(1 + \sqrt{\ln\frac{1}{\delta}}\right). \tag{10}$$

For the single-trajectory setting, define

$$K_{\text{st}}(\delta, N) := \sqrt{n}\left[\left(\sqrt{L_H} + \sqrt{\frac{2\mathscr{C}}{1-\theta}}\right)\sqrt{\frac{2m}{N}} + \frac{2L_H\mathscr{C}}{\theta(1-\theta)}\sqrt{\left(1 + \frac{L_H - 1}{N}\right)\frac{d'\ln(1/\delta)}{2N}}\right]. \tag{11}$$

We then write the error bound for the two sampling regimes uniformly as

$$K_*(\delta, N) := \begin{cases} K_{\mathrm{iid}}(\delta, N), & \text{for } N \text{ i.i.d. trajectories,} \\ K_{\mathrm{st}}(\delta, N), & \text{for a single trajectory of } N \text{ sampling windows.} \end{cases} \tag{12}$$

Define the constants

$$K_2 := \max\left\{1, \frac{2\sigma_n(H_{\alpha,\beta}) + 2n}{\sigma_n^2(H_{\alpha,\beta})}\right\},$$

$$k_T(\delta, N) := \sqrt{|F|}\Bigg[(1 + \sup_{o\in\mathcal{O}} \|\widetilde{A}_o\|_F)^T \left(K_2 K_*(\delta, N) + \frac{\sqrt{n}}{\sigma_n(H_{\alpha,\beta})}\right) + \Bigg((1 + \sup_{o\in\mathcal{O}} \|\widetilde{A}_o\|_F)^T - (\sup_{o\in\mathcal{O}} \|\widetilde{A}_o\|_F)^T\Bigg)\sqrt{nD_s}\left(K_2 K_*(\delta, N) + \frac{\sqrt{n}}{\sigma_n(H_{\alpha,\beta})}\right) + \sqrt{n}(\sup_{o\in\mathcal{O}} \|\widetilde{A}_o\|_F)^T\Bigg].$$

To prove the main theorem, we require the following assumptions:

Assumption 3.3. The word probability function M has finite Hankel rank n, and the data-independent suffix and prefix sets $\alpha = \{v_1, \ldots, v_n\}$, $\beta = \{u_1, \ldots, u_n\}$ satisfy $\mathrm{rank}(H_M) = \mathrm{rank}(H_{\alpha,\beta}) = n$.

Assumption 3.4. There exists $N_0 \in \mathbb{N}$, such that for $N \geq N_0$, $K_*(\delta, N) < \frac{\sigma_n(H_{\alpha,\beta})}{2}$ and $\varepsilon_A(\delta, N) := \sqrt{D_s} K_2 K_*(\delta, N) \leq 1$.

Note that these assumptions are always true for some α, β (Proposition C.1), and for some N_0 as both $K_*(\delta, N)$ and $\varepsilon_A(\delta, N)$ converge to 0 as $N \to \infty$.

Theorem 3.5. *Let us define*

$$\varepsilon_T(\delta, N) := |\mathcal{O}|^T K_2 K_*(\delta, N) k_T(\delta, N). \tag{13}$$

Under Assumptions 3.3 and 3.4, for $N \geq N_0$, and every fixed verification horizon $T \geq 1$,

$$\mathbb{P}\left(|\widehat{\mathbf{p}}_T - p_T| \leq \varepsilon_T(\delta, N)\right) \geq 1 - \delta, \tag{14}$$

Note that for fixed T, $\varepsilon_T(\delta, N) = O\left(\frac{1+\sqrt{\ln(1/\delta)}}{\sqrt{N}}\right)$. That is, the difference between the true and estimated safety probabilities decreases in a way which is consistent with classical PAC bounds. Discussions of the bound will be presented in Section 4. Informally, the main steps behind the proof of Theorem 3.5 are as follows. Using (Balle Pigem, 2013, Theorem 6.2.1) for the N i.i.d. trajectories setting and (Balle & Maillard, 2017, Theorem 3) for the single-trajectory setting, we show that, in either case, with probability at least $1 - \delta$, all Hankel estimation errors required for spectral realization are bounded by $K_*(\delta, N)$. Under Assumption 3.4, the estimated spectral realization is well defined. The perturbation bound for spectral realization then gives $\max\left\{\max_{o\in\mathcal{O}} \|\widehat{\mathbf{A}}_o - A_o\|_F, \|\widehat{\mathbf{B}} - B\|_F, \|\widehat{\mathbf{C}} - C\|_F\right\} \leq K_2 K_*(\delta, N)$. The safety specification automaton is fixed and therefore introduces no additional statistical error. Using the Kronecker product of the DFA-WFA realization, we obtain $\max_{o\in\mathcal{O}} \|\widehat{\widetilde{\mathbf{A}}}_o - \widetilde{A}_o\|_F \leq \sqrt{D_s} K_2 K_*(\delta, N)$, $\|\widehat{\widetilde{\mathbf{B}}} - \widetilde{B}\|_F \leq K_2 K_*(\delta, N)$, $\|\widehat{\widetilde{\mathbf{C}}} - \widetilde{C}\|_F \leq \sqrt{|F|} K_2 K_*(\delta, N)$. Finally, both p_T and $\widehat{\mathbf{p}}_T$ can be expressed as sums of the weights assigned by the corresponding DFA-WFA realizations to accepted words of length T. Using standard matrix-norm inequalities, the error bound for the estimated safety probability is $|\widehat{\mathbf{p}}_T - p_T| \leq \varepsilon_T(\delta, N)$.

All the subsequent estimates are consequences of the initial Hankel error bound. Therefore, the final inequality holds with probability at least $1 - \delta$. The complete derivation will be given in Appendix A.

We next derive the certification consequences of Equation 14. Let $\mathbf{D}_*^N$, $* \in \{\mathrm{iid}, \mathrm{st}\}$ be the random dataset define in Equation 3, taking values in $\mathcal{D}_*^N$. The following two corollaries show that the certification rule given in Equation 7 controls the probability of false certification.

Corollary 3.6. *If Assumption 3.3 and 3.4 hold, then $\Gamma_N = \widehat{\Gamma}^{\mathrm{Spect}}_{N,\varepsilon_T}$ is well-defined and if $p_T < 1 - \delta_{\mathrm{unsafe}}$, then $\mathbb{P}\left(\Gamma_N(\mathbf{D}_*^N) = 1\right) \leq \delta$, for $* \in \{\mathrm{iid}, \mathrm{st}\}$.*

Corollary 3.7. *Define the safety margin* $\Delta_T := p_T - (1 - \delta_{\text{unsafe}})$. *Suppose that*

$$\Delta_T > 2\varepsilon_T(\delta, N). \tag{15}$$

Under the Assumption 3.3 and 3.4, $\Gamma_N = \widehat{\Gamma}^{\text{Spect}}_{N,\varepsilon_T}$ *is a data-driven safety certificate at confidence level* $1 - \delta$ *and risk tolerance* δ_{unsafe} *in the sense of Definition 2.4. In particular, if* $p_T > 1 - \delta_{\text{unsafe}}$, *then* $\mathbb{P}\left(\Gamma_N(\mathbf{D}_*^N) = 0\right) \leq \delta$, *for* $* \in \{\text{iid}, \text{st}\}$.

The proof of the two corollaries follows from a hypothesis testing argument and will be given in Appendix D. The first corollary controls false declarations of safety: an unsafe system is certified with probability at most δ. The second controls false non-certification when the true satisfaction probability is separated from the safety threshold.

Given Theorem 3.5, Corollaries 3.6 and 3.7, Theorem 3.1 can be proved as follows.

Proof of Theorem 3.1. For fixed T, Equation 10, Equation 11 and Equation 12 imply that there is a constant $C_K > 0$, such that $K_*(\delta, N) \leq C_K \frac{1+\sqrt{\ln(1/\delta)}}{\sqrt{N}}$. Consequently, the two inequalities in Assumption 3.4 holds whenever $N \geq C_0(1 + \sqrt{\ln(1/\delta)})^2$, for some suitable C_0. The factors in $k_T(\delta, N)$ are bounded for fixed T. Hence, there exists a constant $C_T > 0$, such that $\varepsilon_T(\delta, N) \leq C_T \frac{1+\sqrt{\ln(1/\delta)}}{\sqrt{N}}$. Corollary 3.6 says that if $p_T < 1 - \delta_{\text{unsafe}}$, then $\mathbb{P}(\Gamma_N(\mathbf{D}_*^N) = 1) \leq \delta$. Suppose $p_T > 1 - \delta_{\text{unsafe}}$. Choose $c > \max\{C_0, 4C_T^2\}$, and let $m_{\text{cert}}(\delta, \delta_{\text{unsafe}}) := \left\lceil \frac{c(1+\sqrt{\ln(1/\delta)})^2}{(p_T - (1-\delta_{\text{unsafe}}))^2} \right\rceil$. Since $0 < p_T - (1 - \delta_{\text{unsafe}}) \leq 1$, every $N \geq m_{\text{cert}}(\delta, \delta_{\text{unsafe}})$ satisfies Assumption 3.4 and $2\varepsilon_T(\delta, N) \leq 2C_T \frac{1+\sqrt{\ln(1/\delta)}}{\sqrt{N}} \leq 2C_T \frac{p_T - (1-\delta_{\text{unsafe}})}{\sqrt{c}} < \Delta_T$. Corollary 3.7 therefore gives $\mathbb{P}(\Gamma_N(\mathbf{D}_*^N) = 0) \leq \delta$. This proves the claim. □

4 Discussion and conclusions

Theorem 3.5 shows that 1) For every fixed verification horizon T and fixed δ, the estimation error bound decreases at the rate of $N^{-1/2}$. Therefore, $\widehat{\mathbf{p}}_T$ is a high-probability consistent estimator of p_T. 2) The failure probability δ affects the bound through $\sqrt{\ln(1/\delta)}$. A smaller δ provides a stronger confidence guarantee but results in a wider error bound. 3) Both the i.i.d.-trajectory and single-trajectory regimes achieve the same decreasing rate $N^{-1/2}$. However, the constants in the finite-sample error bound are different. In the single-trajectory setting, the constants additionally depend on the geometric-mixing parameters and the length of required words to construct finite Hankel matrices. 4) A smaller value of $\varepsilon_T(\delta, N)$ produces a tighter confidence interval for p_T and hence a less conservative safety certificate. Systems whose true safety probabilities are close to the prescribed threshold generally require more data before they can be successfully certified.

The following constants also affect the bounds: 1) The selection of α, β affects the error bound through $H_{\alpha,\beta}$. The constant K_2 becomes large when $\sigma_n(H_{\alpha,\beta})$ is small, because an ill-conditioned Hankel block makes the spectral reconstruction more sensitive to estimation errors. A small value of $\sigma_n(H_{\alpha,\beta})$ may also increase the sample size threshold N_0. 2) Although different full-rank selections of α and β produce isomorphic minimal WFA realizations and therefore the same exact safety probability p_T, their finite-sample error bounds need not be identical. Different selections can yield different safety certificates.

In Appendix E, we illustrate our result through a small but non-trivial example. Although the current finite-sample bound requires a large number of data points, the experiment demonstrates that the system can be certified once the sample size is sufficiently large. The large sample requirement reflects the conservativeness introduced by compounding several worst-case concentration and perturbation bounds, and improving these constants is an important direction for future work.

AI use statement

In this work, we used OpenAI Codex to generate synthetic data sets, design and provide feedback on research methodology or experiments, and interpret numerical results. We did not used generative AI tools to help develop theoretical models and conceptual frameworks, formulate mathematical claims,

provide critical ingredients for proving mathematical claims, assist in the writing of proofs, propose or refine hypotheses, implement methods. We did not use generative AI to assist with translation, clean and reformat dataset, support qualitative and thematic data analysis, as these tasks are not applicable to this work. Additionally, we used OpenAI Codex to create and edit software code, summarize and analyze existing literature. We have reviewed all AI-assisted work. LLM-generated code was verified and tested for correctness by the authors. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

Kamyar Azizzadenesheli, Alessandro Lazaric, and Animashree Anandkumar. Reinforcement learning of pomdps using spectral methods. In *Conference on Learning Theory*, pp. 193–256. PMLR, 2016.

Borja Balle and Odalric-Ambrym Maillard. Spectral learning from a single trajectory under finite-state policies. In *International Conference on Machine Learning*, pp. 361–370. PMLR, 2017.

Borja Balle and Mehryar Mohri. Spectral learning of general weighted automata via constrained matrix completion. *Advances in neural information processing systems*, 25, 2012.

Borja Balle, Xavier Carreras, Franco M Luque, and Ariadna Quattoni. Spectral learning of weighted automata: A forward-backward perspective. *Machine learning*, 96(1):33–63, 2014.

Borja de Balle Pigem. *Learning finite-state machines: statistical and algorithmic aspects*. PhD thesis, Universitat Politècnica de Catalunya, 2013.

Hugo Bazille, Blaise Genest, Cyrille Jégourel, and Jun Sun. Global pac bounds for learning discrete time markov chains. In *International Conference on Computer Aided Verification*, pp. 304–326. Springer, 2020.

Maxime Bouton, Jana Tumova, and Mykel J Kochenderfer. Point-based methods for model checking in partially observable markov decision processes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 10061–10068, 2020.

Peter Buchholz and Peter Kemper. Model checking for a class of weighted automata. *Discrete Event Dynamic Systems*, 20(1):103–137, 2010.

Steven Carr, Nils Jansen, and Ufuk Topcu. Task-aware verifiable rnn-based policies for partially observable markov decision processes. *Journal of Artificial Intelligence Research*, 72:819–847, 2021.

Hadrien Glaude and Olivier Pietquin. Pac learning of probabilistic automaton based on the method of moments. In *International Conference on Machine Learning*, pp. 820–829. PMLR, 2016.

Thomas Hérault, Richard Lassaigne, Frédéric Magniette, and Sylvain Peyronnet. Approximate probabilistic model checking. In *International Workshop on Verification, Model Checking, and Abstract Interpretation*, pp. 73–84. Springer, 2004.

Daniel Hsu, Sham M Kakade, and Tong Zhang. A spectral algorithm for learning hidden markov models. *Journal of Computer and System Sciences*, 78(5):1460–1480, 2012.

Hua Mao, Yingke Chen, Manfred Jaeger, Thomas D Nielsen, Kim G Larsen, and Brian Nielsen. Learning deterministic probabilistic automata from a model checking perspective. *Machine Learning*, 105(2):255–299, 2016.

Gethin Norman, David Parker, and Xueyi Zou. Verification and control of partially observable probabilistic systems. *Real-Time Systems*, 53(3):354–402, 2017.

Dániel Rácz, Mihály Petreczky, and Bálint Daróczy. A finite-sample bound for identifying partially observed linear switched systems from a single trajectory. *IEEE Control Systems Letters*, 2025.

Koushik Sen, Mahesh Viswanathan, and Gul Agha. Statistical model checking of black-box probabilistic systems. In *International Conference on Computer Aided Verification*, pp. 202–215. Springer, 2004.

Nicholas F Travers. Exponential bounds for convergence of entropy rate approximations in hidden markov models satisfying a path-mergeability condition. *Stochastic Processes and their Applications*, 124(12):4149–4170, 2014.

Jingyi Wang, Jun Sun, Qixia Yuan, and Jun Pang. Should we learn probabilistic models for model checking? a new approach and an empirical study. In *International Conference on Fundamental Approaches to Software Engineering*, pp. 3–21. Springer, 2017.

Håkan LS Younes and Reid G Simmons. Probabilistic verification of discrete event systems using acceptance sampling. In *International Conference on Computer Aided Verification*, pp. 223–235. Springer, 2002.

Lijun Zhang, Holger Hermanns, and David N Jansen. Logic and model checking for hidden markov models. In *International Conference on Formal Techniques for Networked and Distributed Systems*, pp. 98–112. Springer, 2005.

A Error Analysis

This section derives a finite-sample bound on the error between the true and empirical satisfaction probabilities. The analysis proceeds in three steps. First, we establish a uniform bound for the empirical word probabilities required by the Hankel realization. Second, we propagate this error to the empirical WFA parameters. Finally, we propagate the WFA realization error through the specification-automaton product to obtain a probability bound on $|\widehat{\mathbf{p}}_T - p_T|$.

A.1 Sample Bounds for Hankel Matrix Estimation

We derive Hankel matrix error bounds for two sampling regimes separately:

For N i.i.d. trajectories, we have the following lemma:

Lemma A.1. *(Balle Pigem, 2013, Theorem 6.2.1) Let $\widehat{\mathbf{H}}_{\alpha',\beta'}$ be constructed as described in Section 3.2. With probability at least $1-\delta$,*

$$\|H_{\alpha',\beta'} - \widehat{\mathbf{H}}_{\alpha',\beta'}\|_F \leq \sqrt{\frac{c_{\alpha',\beta'}}{N}} \left(1 + \sqrt{\ln \frac{1}{\delta}}\right), \tag{16}$$

with $c_{\alpha',\beta'}$ given by Equation 8.

For the single-trajectory setting, we assume that the augmented hidden-state process is initialized from its invariant distribution.

Lemma A.2. *(Balle & Maillard, 2017, Theorem 3). Suppose that the closed-loop hidden-state process is $(\mathscr{C}, \theta)$-geometrically mixing. Then, with probability at least $1-\delta$,*

$$\|\widehat{\mathbf{H}}_{\alpha',\beta'} - H_{\alpha',\beta'}\|_F \leq \sqrt{n}\Bigg[\left(\sqrt{L_H} + \sqrt{\frac{2\mathscr{C}}{1-\theta}}\right)\sqrt{\frac{2m}{N}} + \frac{2L_H\mathscr{C}}{\theta(1-\theta)}\sqrt{\left(1 + \frac{L_H - 1}{N}\right)\frac{d'\ln(1/\delta)}{2N}}\Bigg],$$

with L_H given in Equation 5, d', m are given in Equation 9.

Because the required Hankel blocks and boundary vectors are sub-blocks of $H_{\alpha',\beta'}$, with probability at least $1-\delta$,

$$\max\left\{\begin{array}{c}\|H_{\alpha,\beta} - \widehat{\mathbf{H}}_{\alpha,\beta}\|_F, \quad \max\limits_{o\in\mathcal{O}} \|H_{\alpha,o,\beta} - \widehat{\mathbf{H}}_{\alpha,o,\beta}\|_F, \\ \|h_{\alpha,\epsilon} - \widehat{\mathbf{h}}_{\alpha,\epsilon}\|_F, \quad \|h_{\epsilon,\beta} - \widehat{\mathbf{h}}_{\epsilon,\beta}\|_F\end{array}\right\} \leq K_*(\delta, N). \tag{17}$$

A.2 WFA Realization Error Bound

For simplicity, let $H_1 := H_{\alpha,\beta}$, $\widehat{\mathbf{H}}_1 := \widehat{\mathbf{H}}_{\alpha,\beta}$, for a fixed $o \in \mathcal{O}$, let $H_2 := H_{\alpha,o,\beta}$, $\widehat{\mathbf{H}}_2 := \widehat{\mathbf{H}}_{\alpha,o,\beta}$, $h := h_{\alpha,\epsilon}$, $\widehat{\mathbf{h}} := \widehat{\mathbf{h}}_{\alpha,\epsilon}$.

Since

$$|\sigma_n(\widehat{\mathbf{H}}_1) - \sigma_n(H_1)| \le \|\widehat{\mathbf{H}}_1 - H_1\|_2 \le \|\widehat{\mathbf{H}}_1 - H_1\|_F,$$

with probability at least $1-\delta$,

$$|\sigma_n(\widehat{\mathbf{H}}_1) - \sigma_n(H_1)| \le K_*(\delta, N),$$

where $* \in \{\mathrm{iid}, \mathrm{st}\}$. Then with probability at least $1-\delta$,

$$\sigma_n(\widehat{\mathbf{H}}_1) \ge \sigma_n(H_1) - K_*(\delta, N).$$

Under Assumption 3.4, for large enough N, $K_*(\delta, N) < \frac{\sigma_n(H_1)}{2}$. Then with probability at least $1-\delta$, that $\sigma_n(\widehat{\mathbf{H}}_1) \ge \frac{\sigma_n(H_1)}{2} > 0$ i.e., $\widehat{\mathbf{H}}_1$ is invertible. Since $\|\widehat{\mathbf{H}}_1^{-1}\|_2 = \frac{1}{\sigma_n(\widehat{\mathbf{H}}_1)}$, we have $\|\widehat{\mathbf{H}}_1^{-1}\|_2 \le \frac{2}{\sigma_n(H_1)}$.

We have the following bounds:

$$\|H_1\|_F \le n, \quad \|H_2\|_F \le n, \quad \|h\|_F \le \sqrt{n}.$$

If Equation 17 holds and under Assumption 3.4, then since $\widehat{\mathbf{H}}_1^{-1} - H_1^{-1} = \widehat{\mathbf{H}}_1^{-1}(H_1 - \widehat{\mathbf{H}}_1)H_1^{-1}$,

$$\|\widehat{\mathbf{H}}_1^{-1} - H_1^{-1}\|_F \le \frac{2}{\sigma_n^2(H_1)} K_*(\delta, N).$$

We have the following error for WFA parameters:

For each $o \in \mathcal{O}$,

$$\begin{aligned}
\|\widehat{\mathbf{A}}_o - A_o\|_F &= \|\widehat{\mathbf{H}}_1^{-1}\widehat{\mathbf{H}}_2 - H_1^{-1}H_2\|_F \\
&\le \|\widehat{\mathbf{H}}_1^{-1}\|_2 \|\widehat{\mathbf{H}}_2 - H_2\|_F + \|H_2\|_F \|\widehat{\mathbf{H}}_1^{-1} - H_1^{-1}\|_F \\
&\le \frac{2\sigma_n(H_1) + 2n}{\sigma_n^2(H_1)} K_*(\delta, N).
\end{aligned}$$

Replacing H_2 by h and $\widehat{\mathbf{H}}_2$ by $\widehat{\mathbf{h}}$ gives

$$\|\widehat{\mathbf{B}} - B\|_F \le \frac{2\sigma_n(H_1) + 2\sqrt{n}}{\sigma_n^2(H_1)} K_*(\delta, N).$$

Then, with probability at least $1-\delta$,

$$\max\left\{\max_{o\in\mathcal{O}} \|\widehat{\mathbf{A}}_o - A_o\|_F, \|\widehat{\mathbf{B}} - B\|_F, \|\widehat{\mathbf{C}} - C\|_F\right\} \le K_2 K_*(\delta, N), \tag{18}$$

where $K_2 = \max\left\{1, \frac{2\sigma_n(H_{\alpha,\beta})+2n}{\sigma_n^2(H_{\alpha,\beta})}\right\}$.

A.3 Error Propagation through the Product Realization

We now propagate the realization error in Equation 18 through the Kronecker-product construction with the specification automaton. Recall the exact product realization $\widetilde{A}_o = E_o \otimes A_o$, $\widetilde{B} = e_{r_0} \otimes B$, $\widetilde{C} = \chi_F^\top \otimes C$, and the empirical counterpart $\widehat{\widetilde{\mathbf{A}}}_o = E_o \otimes \widehat{\mathbf{A}}_o$, $\widehat{\widetilde{\mathbf{B}}} = e_{r_0} \otimes \widehat{\mathbf{B}}$, $\widehat{\widetilde{\mathbf{C}}} = \chi_F^\top \otimes \widehat{\mathbf{C}}$.

Corollary A.3. *Under Assumptions 3.3 and 3.4, for $N \ge N_0$, with probability at least $1-\delta$, the following inequalities hold simultaneously:*

$$\begin{aligned}
\max_{o\in\mathcal{O}} \|\widehat{\widetilde{\mathbf{A}}}_o - \widetilde{A}_o\|_F &\le \sqrt{D_s} K_2 K_*(\delta, N), \\
\|\widehat{\widetilde{\mathbf{B}}} - \widetilde{B}\|_F &\le K_2 K_*(\delta, N), \\
\|\widehat{\widetilde{\mathbf{C}}} - \widetilde{C}\|_F &\le \sqrt{|F|} K_2 K_*(\delta, N).
\end{aligned} \tag{19}$$

Proof. By multiplicativity of the Frobenius norm with respect to the Kronecker product, for every $o \in \mathcal{O}$,

$$\|\widehat{\widetilde{\mathbf{A}}}_o - \widetilde{A}_o\|_F = \|E_o \otimes (\widehat{\mathbf{A}}_o - A_o)\|_F = \sqrt{D_s}\|\widehat{\mathbf{A}}_o - A_o\|_F,$$
$$\|\widehat{\widetilde{\mathbf{B}}} - \widetilde{B}\|_F = \|e_{r_0} \otimes (\widehat{\mathbf{B}} - B)\|_F = \|\widehat{\mathbf{B}} - B\|_F,$$
$$\|\widehat{\widetilde{\mathbf{C}}} - \widetilde{C}\|_F = \|\chi_F^\top \otimes (\widehat{\mathbf{C}} - C)\|_F = \sqrt{|F|}\|\widehat{\mathbf{C}} - C\|_F.$$

Since $\|E_o\|_F = \sqrt{D_s}$, $\|e_{r_0}\|_F = 1$, and $\|\chi_F\|_F = \sqrt{|F|}$, the result follows directly from Equation 18. □

A.4 Finite-Sample Satisfaction Probability

Recall that the exact satisfaction probability and its empirical estimate are $p_T = \widetilde{C}\widetilde{G}^T\widetilde{B}$ and $\widehat{\mathbf{p}}_T = \widehat{\widetilde{\mathbf{C}}}\widehat{\widetilde{\mathbf{G}}}^T\widehat{\widetilde{\mathbf{B}}}$.

Lemma A.4. *Let*

$$S_{\widetilde{A}} := \sup_{o\in\mathcal{O}} \|\widetilde{A}_o\|_F, \qquad \rho_T := (1 + S_{\widetilde{A}})^T - S_{\widetilde{A}}^T.$$

Then, under Assumptions 3.3 and 3.4, for $N \geq N_0$, for every $T \geq 1$, with probability at least $1 - \delta$,

$$\|\widehat{\widetilde{\mathbf{G}}}^T - \widetilde{G}^T\|_F \leq |\mathcal{O}|^T\sqrt{D_s}K_2K_*(\delta, N)\rho_T.$$

Proof. Let $S_{\widehat{\widetilde{\mathbf{A}}}} := \sup_{o\in\mathcal{O}} \|\widehat{\widetilde{\mathbf{A}}}_o\|_F$. For every $o \in \mathcal{O}$,

$$\begin{aligned}\|\widehat{\widetilde{\mathbf{A}}}_o\|_F &\leq \|\widehat{\widetilde{\mathbf{A}}}_o - \widetilde{A}_o\|_F + \|\widetilde{A}_o\|_F \\ &\leq \sqrt{D_s}K_2K_*(\delta, N) + \sup_{o\in\mathcal{O}} \|\widetilde{A}_o\|_F \\ &\leq \max\{1, \sqrt{D_s}K_2K_*(\delta, N)\}(1 + S_{\widetilde{A}}).\end{aligned}$$

If Assumption 3.4 holds, $\sqrt{D_s}K_2K_*(\delta, N) \leq 1$, then $S_{\widehat{\widetilde{\mathbf{A}}}} \leq 1 + S_{\widetilde{A}}$.

If Equation 19 holds, then

$$\|\widehat{\widetilde{\mathbf{G}}} - \widetilde{G}\|_F \leq \sum_{o\in\mathcal{O}} \|\widehat{\widetilde{\mathbf{A}}}_o - \widetilde{A}_o\|_F \leq |\mathcal{O}|\sqrt{D_s}K_2K_*(\delta, N).$$

Since

$$\|\widetilde{G}\|_F \leq |\mathcal{O}|S_{\widetilde{A}}, \qquad \|\widehat{\widetilde{\mathbf{G}}}\|_F \leq |\mathcal{O}|S_{\widehat{\widetilde{\mathbf{A}}}}, \tag{20}$$

and

$$\widehat{\widetilde{\mathbf{G}}}^T - \widetilde{G}^T = \sum_{k=1}^{T} \widehat{\widetilde{\mathbf{G}}}^{T-k}(\widehat{\widetilde{\mathbf{G}}} - \widetilde{G})\widetilde{G}^{k-1},$$

we have

$$\begin{aligned}\|\widehat{\widetilde{\mathbf{G}}}^T - \widetilde{G}^T\|_F &\leq \sum_{k=1}^{T} \|\widehat{\widetilde{\mathbf{G}}}\|_F^{T-k}\|\widehat{\widetilde{\mathbf{G}}} - \widetilde{G}\|_F\|\widetilde{G}\|_F^{k-1} \\ &\leq |\mathcal{O}|^T\sqrt{D_s}K_2K_*(\delta, N)\sum_{k=1}^{T}(1 + S_{\widetilde{A}})^{T-k}S_{\widetilde{A}}^{k-1} \\ &= |\mathcal{O}|^T\sqrt{D_s}K_2K_*(\delta, N)\rho_T.\end{aligned}$$

□

We also obtain the bound

$$\|\widehat{\widetilde{\mathbf{G}}}^T\|_F \leq \left(\sum_{o\in\mathcal{O}} \|\widehat{\widetilde{\mathbf{A}}}_o\|_F\right)^T \leq |\mathcal{O}|^T(1 + S_{\widetilde{A}})^T. \tag{21}$$

A.5 Proof of Equation 14 in Theorem 3.5

Proof. Suppose that Equation 19 holds. The satisfaction probability error can be written as

$$\begin{aligned}|\widehat{\mathbf{p}}_T - p_T| &= \left|\widehat{\widetilde{\mathbf{C}}}\widehat{\widetilde{\mathbf{G}}}^T\widehat{\widetilde{\mathbf{B}}} - \widetilde{C}\widetilde{G}^T\widetilde{B}\right| \\ &\le \|\widehat{\widetilde{\mathbf{C}}} - \widetilde{C}\|_F\|\widehat{\widetilde{\mathbf{G}}}^T\|_F\|\widehat{\widetilde{\mathbf{B}}}\|_F \\ &\quad + \|\widetilde{C}\|_F\|\widehat{\widetilde{\mathbf{G}}}^T - \widetilde{G}^T\|_F\|\widehat{\widetilde{\mathbf{B}}}\|_F \\ &\quad + \|\widetilde{C}\|_F\|\widetilde{G}^T\|_F\|\widehat{\widetilde{\mathbf{B}}} - \widetilde{B}\|_F.\end{aligned}$$

For $\widehat{\widetilde{\mathbf{B}}}$,

$$\|\widehat{\widetilde{\mathbf{B}}}\|_F \le K_2K_*(\delta, N) + \frac{\sqrt{n}}{\sigma_n(H_{\alpha,\beta})}. \tag{22}$$

For $\widetilde{C}$,

$$\|\widetilde{C}\|_F \le \sqrt{|F|n}. \tag{23}$$

By Equation 19, Equation 20, Equation 21, Equation 22, and Equation 23,

$$\begin{aligned}|\widehat{\mathbf{p}}_T - p_T| &\le \sqrt{|F|}|\mathcal{O}|^T K_2K_*(\delta, N) \\ &\quad \times \left[(1 + S_{\widetilde{A}})^T\left(K_2K_*(\delta, N) + \frac{\sqrt{n}}{\sigma_n(H_{\alpha,\beta})}\right)\right. \\ &\quad \left. + \rho_T\sqrt{nD_s}\left(K_2K_*(\delta, N) + \frac{\sqrt{n}}{\sigma_n(H_{\alpha,\beta})}\right) + \sqrt{n}S_{\widetilde{A}}^T\right].\end{aligned}$$

Let

$$\begin{aligned}k_T(\delta, N) &:= \sqrt{|F|}\left[(1 + S_{\widetilde{A}})^T\left(K_2K_*(\delta, N) + \frac{\sqrt{n}}{\sigma_n(H_{\alpha,\beta})}\right)\right. \\ &\quad \left. + \rho_T\sqrt{nD_s}\left(K_2K_*(\delta, N) + \frac{\sqrt{n}}{\sigma_n(H_{\alpha,\beta})}\right) + \sqrt{n}S_{\widetilde{A}}^T\right].\end{aligned}$$

Then, with probability at least $1 - \delta$,

$$|\widehat{\mathbf{p}}_T - p_T| \le |\mathcal{O}|^T K_2K_*(\delta, N)k_T(\delta, N).$$

For fixed T, with $\varepsilon_T(\delta, N)$ defined in Equation 13, this proves Equation 14. □

B Proof of Proposition 3.2

Proof. By the mixed-product property of the Kronecker product,

$$\begin{aligned}\widetilde{C}\widetilde{A}_w\widetilde{B} &= (\chi_F^\top \otimes C)(E_w \otimes A_w)(e_{r_0} \otimes B) \\ &= (\chi_F^\top E_w e_{r_0}) \otimes (CA_wB) \\ &= (\chi_F^\top e_{\delta_s(r_0,w)})(CA_wB).\end{aligned}$$

We have $\chi_F^\top e_{\delta_s(r_0,w)} = \mathbf{1}_{\{\delta_s(r_0,w)\in F\}} = \mathbf{1}_{\{w\in\mathcal{L}_{\text{safe}}\}}$. Hence, $\widetilde{C}\widetilde{A}_w\widetilde{B} = \mathbf{1}_{\{w\in\mathcal{L}_{\text{safe}}\}}M(w)$. □

C Invariance of p_T

Proposition C.1. *Let $M : \mathcal{O}^* \to [0, 1]$ be a word probability function and (α, β), $(\widehat{\alpha}, \widehat{\beta})$ be two pairs of n-row selections and n-column selections with $\operatorname{rank}(H_{\alpha,\beta}) = \operatorname{rank}(H_{\widehat{\alpha},\widehat{\beta}}) = \operatorname{rank}(H_M) = n$. Let $\Sigma = \langle C, \{A_o\}_{o\in\mathcal{O}}, B\rangle$ and $\Sigma' = \langle C', \{A'_o\}_{o\in\mathcal{O}}, B'\rangle$ be their corresponding realizations. Then there exists $S \in \mathbb{R}^{n\times n}$ such that $\det(S) \neq 0$, and for all $o \in \mathcal{O}$,*

$$A'_o = SA_oS^{-1}, \qquad B' = SB, \qquad C' = CS^{-1}.$$

Proof. Both realizations have dimension n and realize the same function M, hence they are minimal realizations. The result follows from the uniqueness of minimal realizations up to isomorphism (Balle et al., 2014, Corollary 2). □

D Proof of Corollaries 3.6 and 3.7

Proof. Under Assumptions 3.3 and 3.4, $\mathbb{P}(|\widehat{\mathbf{p}}_T - p_T| > \varepsilon_T(\delta, N)) \leq \delta$. Suppose that $p_T \leq 1 - \delta_{\text{unsafe}}$ holds. Then

$$\begin{aligned}\{\Gamma_N(\mathbf{D}_*^N) = 1\} &= \{\widehat{\mathbf{p}}_T > 1 - \delta_{\text{unsafe}} + \varepsilon_T(\delta, N)\} \\ &\subseteq \{\widehat{\mathbf{p}}_T - p_T > \varepsilon_T(\delta, N)\} \\ &\subseteq \{|\widehat{\mathbf{p}}_T - p_T| > \varepsilon_T(\delta, N)\}.\end{aligned}$$

Consequently, $\mathbb{P}(\Gamma_N(\mathbf{D}_*^N) = 1) \leq \delta$.

Suppose that $p_T > 1 - \delta_{\text{unsafe}}$ holds. Then

$$\begin{aligned}\{\Gamma_N(\mathbf{D}_*^N) = 0\} &= \{\widehat{\mathbf{p}}_T \leq 1 - \delta_{\text{unsafe}} + \varepsilon_T(\delta, N)\} \\ &\subseteq \{p_T - \widehat{\mathbf{p}}_T \geq \Delta_T - \varepsilon_T(\delta, N) > \varepsilon_T(\delta, N)\} \\ &\subseteq \{|\widehat{\mathbf{p}}_T - p_T| > \varepsilon_T(\delta, N)\}.\end{aligned}$$

It follows that $\mathbb{P}(\Gamma_N(\mathbf{D}_*^N) = 0) \leq \delta$, equivalently $\mathbb{P}(\Gamma_N(\mathbf{D}_*^N) = 1) \geq 1 - \delta$. □

E Numerical Analysis

We present numerical results for two sampling regimes. Both cases use the same POMDP state, action and observation spaces, defined as follows.

We consider a partially observable system with two hidden states $x_1 = \text{stable}$ and $x_2 = \text{degraded}$, two actions $u_1 = \text{hold}$ and $u_2 = \text{recover}$, and two observations $o_1 = \text{normal}$ and $o_2 = \text{alert}$.

The same FSC and DFA are used in both sampling regimes. The FSC has one internal state s. At each step, it selects the actions according to $\varphi(u_1 \mid s) = 0.75,\ \varphi(u_2 \mid s) = 0.25$, and its internal state update is deterministic, $\eta(s \mid s, o) = 1$ for every $o \in \mathcal{O}$.

The specification automaton is represented by the following 3-state DFA:

Table 1: Transition table of the safety specification automaton.

r_0	o_1 (normal) has not been observed yet	Accepted
r_1	o_1 (normal) has been observed	Accepted
r_2	o_2 (alert) is observed after having observed o_1 (normal)	Rejected

The initial state is r_0, the accepting states are r_0 and r_1, and r_2 is an absorbing rejecting state. Throughout both experiments, the verification horizon $T = 3$, and $L = 3$. We set $\delta = 0.05$ and $\delta_{\text{unsafe}} = 0.15$.

The numerical implementation was developed with OpenAI Codex assistance. See Section 4 for details.

E.1 I.i.d. Sampling Regime

The controlled transition matrices are

$$P_{u_1} = \begin{bmatrix} 0.97 & 0.03 \\ 0.03 & 0.97 \end{bmatrix}, \qquad P_{u_2} = \begin{bmatrix} 0.97 & 0.03 \\ 0.57 & 0.43 \end{bmatrix}.$$

Both actions use the same observation kernel,

$$Z_{u_1} = Z_{u_2} = \begin{bmatrix} 0.97 & 0.03 \\ 0.03 & 0.97 \end{bmatrix},$$

The initial hidden-state distribution is $\mu_0 = [0.3 \quad 0.7]^\top$. We select $\alpha = \beta = \{o_1, o_2\}$. The resulting exact Hankel block is

$$H_{\alpha,\beta} = \begin{bmatrix} 0.3746521545 & 0.1180463955 \\ 0.0374578455 & 0.4698436045 \end{bmatrix}.$$

The corresponding shifted Hankel blocks are

$$H_{\alpha,o_1,\beta} = \begin{bmatrix} 0.352241365834275 & 0.100255298538225 \\ 0.022410788665725 & 0.017791096961775 \end{bmatrix},$$

$$H_{\alpha,o_2,\beta} = \begin{bmatrix} 0.015682723038225 & 0.089392945339275 \\ 0.021775122461775 & 0.380450659160725 \end{bmatrix}.$$

The boundary vectors are

$$h_{\alpha,\epsilon} = \begin{bmatrix} 0.41211 \\ 0.58789 \end{bmatrix}, \qquad h_{\epsilon,\beta} = [0.41211 \quad 0.58789].$$

The resulting minimal WFA realization $\langle C, A_{o_1}, A_{o_2}, B \rangle$ is numerically given by

$$A_{o_1} = \begin{bmatrix} 0.948991835979 & 0.262252457970 \\ -0.027959092731 & 0.016958164021 \end{bmatrix},$$

$$A_{o_2} = \begin{bmatrix} 0.027959092731 & -0.016958164021 \\ 0.044116456811 & 0.811090907269 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.723918978063 \\ 1.193532378422 \end{bmatrix}, \qquad C = [0.41211 \quad 0.58789].$$

The singular values of $H_{\alpha,\beta}$ are

$$\sigma_1 = 0.515329412623, \qquad \sigma_2 = 0.333002834412.$$

Consequently,

$$\operatorname{rank}(H_{\alpha,\beta}) = 2, \qquad \sigma_n(H_{\alpha,\beta}) = 0.333002834412,$$

We use $N = 1.3 \times 10^{14}$ independent observation trajectories, each of length 3. The implementation uses one multinomial count draw over the eight length-3 words. This has exactly the same distribution as drawing and aggregating N independent trajectories but does not materialize N individual rows.

The extended bases are

$$\alpha' = \{\epsilon, o_1, o_2\},$$
$$\beta' = \{\epsilon, o_1, o_2, o_1o_1, o_2o_1, o_1o_2, o_2o_2\}.$$

Thus $H_{\alpha',\beta'} \in \mathbb{R}^{3\times 7}$, and exhaustive enumeration of the length-three words gives $c_{\alpha',\beta'} = 6$.

At $\delta = 0.05$ and the stated sample count,

$$K_{\mathrm{iid}}(0.05, N) = 5.86673898625 \times 10^{-7}.$$

The theoretical constants and applicability conditions are given as follows. For this experiment,

$$n = 2, \quad D_s = 3, \quad |F| = 2, \quad |\mathcal{O}| = 2, \quad T = 3, \quad \delta = 0.05.$$

We use the conservative values

$$\underline{\sigma} = 0.333 \le \sigma_n(H_{\alpha,\beta}), \qquad S_{\widetilde{A}} \le \overline{S}_{\widetilde{A}} = 1.71,$$

These values give

$$K_2 = 42.0781141502, \qquad \rho_3 = 14.9023, \qquad k_3 = 348.774415187,$$

and the propagated finite-horizon radius is

$$\varepsilon_3(0.05, N) = 0.0688791279901.$$

Table 2: Effect of the Hankel basis on the i.i.d. finite-sample safety certificate.

α	β	$\sigma_n(H_{\alpha,\beta})$	K_2	$c_{\alpha',\beta'}$	N_0	m_{cert}	$\widehat{p}_T - \varepsilon_T$	Status
$\{o_1, o_2\}$	$\{o_1, o_2\}$	0.3330	42.08	6	2.38e+05	4.71e+14	0.8535	Certified
$\{\epsilon, o_1\}$	$\{o_1, o_2\}$	0.2175	94.16	6	1.19e+06	5.45e+15	0.6882	Not certified
$\{o_1, o_2\}$	$\{\epsilon, o_1\}$	0.2677	63.60	6	5.43e+05	5.61e+15	0.6848	Not certified
$\{\epsilon, o_2\}$	$\{o_1, o_2\}$	0.2058	104.94	6	1.48e+06	7.57e+15	0.6464	Not certified
$\{\epsilon, o_2\}$	$\{\epsilon, o_1\}$	0.1679	155.40	5	2.70e+06	6.94e+16	0.0864	Not certified
$\{\epsilon, o_1\}$	$\{\epsilon, o_1\}$	0.1700	150.17	6	3.03e+06	7.51e+16	0.0527	Not certified
$\{o_1, o_2\}$	$\{\epsilon, o_2\}$	0.1455	204.04	6	5.59e+06	1.19e+18	-2.5388	Not certified
$\{\epsilon, o_1\}$	$\{\epsilon, o_2\}$	0.1008	420.00	5	1.97e+07	8.66e+18	-8.4154	Not certified
$\{\epsilon, o_2\}$	$\{\epsilon, o_2\}$	0.0900	516.05	6	3.57e+07	1.93e+19	-13.0118	Not certified

The two finite-sample conditions in Assumption 3.4 are satisfied by

$$K_{\text{iid}}(0.05, N) = 5.86674 \times 10^{-7} < \frac{\sigma}{2} = 0.1665,$$

and

$$\sqrt{D_s} K_2 K_{\text{iid}}(0.05, N) = 4.27576 \times 10^{-5} \leq 1.$$

The individual minimum sample counts for these two conditions are 1615 and 237668, respectively. Hence both conditions are guaranteed for every $N \geq N_0 = 237668$. Notice that theorem applicability alone does not imply a useful safety certificate: at N_0 the propagated safety radius is still much larger than the safety margin.

Table 2 shows the effect of the selection of α, β on the finite-sample certificate. We compare all nine pairs of 2-word basis constructed from $\{\epsilon, o_1, o_2\}$. Every candidate Hankel block has rank 2, and therefore produces an isomorphic minimal WFA realization with the same exact p_T. However, the finite-sample bounds differ because the selection of α and β changes the associated constants.

For the realized i.i.d. sample, the safety certificate gives

$$\widehat{\mathbf{p}}_3 = 0.922340254404, \qquad \varepsilon_3 = 0.0688791279901$$

$$\widehat{\mathbf{p}}_3 - \varepsilon_3 = 0.8534611264137488 > 0.85$$

The system is successfully certified as safe.

E.2 Single-Trajectory Sampling Regime

The controlled transition matrices are

$$P_{u_1} = \begin{bmatrix} 0.97 & 0.03 \\ 0.03 & 0.97 \end{bmatrix}, \qquad P_{u_2} = \begin{bmatrix} 0.97 & 0.03 \\ 0.21 & 0.79 \end{bmatrix}.$$

Both actions use the same observation kernel,

$$Z_{u_1} = Z_{u_2} = \begin{bmatrix} 0.97 & 0.03 \\ 0.03 & 0.97 \end{bmatrix},$$

The resulting closed-loop chain has stationary distribution $\mu_\infty = [5/7 \quad 2/7]^\top$. The single trajectory is initialized according to μ_∞. For this 2-state chain, the geometric-mixing condition holds with constants $\mathscr{C} = 1, \theta = 0.895$. We select $\alpha = \beta = \{o_1, o_2\}$.

Exact enumeration of the eight length-three observation words gives

$$p_3 = 0.9039314285714287.$$

For the prescribed threshold $1 - \delta_{\text{unsafe}} = 0.85$, the exact safety margin is

$$\Delta_3 = p_3 - 0.85 = 0.0539314285714287.$$

Table 3: Single-trajectory sample thresholds.

Condition	Minimum N
$K_{\mathrm{st}}(\delta, N) < \frac{\sigma_n(H_{\alpha,\beta})}{2}$	6,030,443
$\sqrt{D_s} K_2 K_{\mathrm{st}}(\delta, N) \leq 1$	1,530,158,302
$\varepsilon_3(\delta, N) < \Delta_3$	2,177,771,750,775,615,077
$2\varepsilon_3(\delta, N) < \Delta_3$	8,711,064,462,746,255,295

Table 3 shows the sample size required for theorem applicability and required for successful safety certification. The first two rows correspond to the finite-sample perturbation conditions in Assumption 3.4. The third row reports an oracle diagnostic at which the error radius becomes smaller than the exact safety margin. The final row enforces the two-radius condition in Corollary 3.7 and therefore gives the pre-data threshold that guarantees successful certification with probability at least $1 - \delta = 0.95$. Here, N denotes the number of overlapping length-3 windows, so the corresponding trajectory contains $N + 2$ observations.

These thresholds are evaluated analytically from the concentration and perturbation bounds; no trajectory of length 10^{18} was explicitly generated. Therefore, the experiment establishes a theoretical sample-planning result, rather than reporting an observed empirical certificate. This indicates the conservativeness of our theoretical bound in single trajectory setting. This results from combining the geometric-mixing concentration inequality with several worst-case perturbation bounds. It should be interpreted as a sufficient pre-data guarantee rather than as an estimate of the trajectory length required in practice.