Robustly Complete Finite-State Abstractions for Verification of Stochastic Systems

Yiming Meng(✉) and Jun Liu(✉)

University of Waterloo, Waterloo, ON N2L 3G1, Canada
{yiming.meng,j.liu}@uwaterloo.ca

Abstract. In this paper, we focus on discrete-time stochastic systems modelled by nonlinear stochastic difference equations and propose robust abstractions for verifying probabilistic linear temporal specifications. The current literature focuses on developing sound abstraction techniques for stochastic dynamics without perturbations. However, soundness thus far has only been shown for preserving the satisfaction probability of certain types of temporal-logic specification. We present constructive finite-state abstractions for verifying probabilistic satisfaction of general ω-regular linear-time properties of more general nonlinear stochastic systems. Instead of imposing stability assumptions, we analyze the probabilistic properties from the topological view of metrizable space of probability measures. Such abstractions are both sound and approximately complete. That is, given a concrete discrete-time stochastic system and an arbitrarily small $\mathcal{L}_1$-perturbation of this system, there exists a family of finite-state Markov chains whose set of satisfaction probabilities contains that of the original system and meanwhile is contained by that of the slightly perturbed system. A direct consequence is that, given a probabilistic linear-time specification, initializing within the winning/losing region of the abstraction system can guarantee a satisfaction/dissatisfaction for the original system. We make an interesting observation that, unlike the deterministic case, point-mass (Dirac) perturbations cannot fulfill the purpose of robust completeness.

Keywords: Verification of stochastic systems · Finite-state abstraction · Robustness · Soundness · Completeness · $\mathcal{L}_1$-perturbation · Linear temporal logic · Metrizable space of probability measures

1 Introduction

Formal verification is a rigorous mathematical technique for verifying system properties using formal analysis or model checking [4]. So far, abstraction-based formal verification for deterministic systems has gained its maturity [5]. Whilst bisimilar (equivalent) symbolic models exist for linear (control) systems [17,30], sound and approximately complete finite abstractions can be achieved via stability assumptions [14,25] or robustness (in terms of Dirac perturbations) [19–21].

S. Bogomolov and D. Parker (Eds.): FORMATS 2022, LNCS 13465, pp. 80–97, 2022.
https://doi.org/10.1007/978-3-031-15839-1_5

There is a recent surge of interest in studying formal verification for stochastic systems. The verification of temporal logics for discrete-state homogeneous Markov chains can be solved by existing tools [4,6,9,24].

In terms of verification for general discrete-time continuous-state Markov systems, a common theme is to construct abstractions to approximate the probability of satisfaction in proper ways. First attempts [2,3,26,28,29] were to relate the verification of fundamental probabilistic computation tree logic (PCTL) formulas to the computation of corresponding value functions. The authors [31,32] developed alternative techniques to deal with the potential error blow-up in infinite-horizon problems. The same authors [34] investigated the necessity of absorbing sets on the uniqueness of the solutions of corresponding Bellman equations. The related PCTL verification problem can be then precisely captured by finite-horizon ones. They also proposed abstractions for verifying general bounded linear-time (LT) properties [33], and extended them to infinite-horizon reach-avoid and repeated reachability problems [33,35].

Markov set-chains are also constructive to be abstractions. The authors [1] showed that the error is finite under strong assumptions on stability (ergodicity). A closely related approach is to apply interval-valued Markov chains (IMCs), a family of finite-state Markov chains with uncertain transitions, as abstractions for the continuous-state Markov systems with certain transition kernel. The authors [18] argued without proof that for every PCTL formula, the probability of (path) satisfaction of the IMC abstractions forms a compact interval, which contains the real probability of the original system. They further developed 'O'-maximizing/minimizing algorithms based on [15,38] to obtain the upper/lower bound of the satisfaction probability of 'next', 'bounded until', and 'until' properties. The algorithm provides a fundamental view of computing the bounds of satisfaction probability given IMCs. However, the intuitive reasoning for soundness seems inaccurate based on our observation (readers who are interested in the details are referred to Remark 6 of this paper). Inspired by [18], the work in [7] formulated IMC abstraction for verifying bounded-LTL specifications; the work in [11,12] constructed IMC abstractions for verifying general ω-regular properties of mixed-monotone systems, and provided a novel automata-based approach in obtaining the bounds of satisfaction probability. Both [11,12, Fact 1] and [10] claimed the soundness of verifying general ω-regular properties using IMC abstractions, but a proof is not provided.

Motivated by these issues, our first contribution is to provide a formal mathematical proof of the soundness of IMC abstractions for verifying ω-regular linear-time properties. We show that, for any discrete-time stochastic dynamical systems modelled by a stochastic difference equation and any linear-time property, an IMC abstraction returns a compact interval of probability of (path) satisfaction which contains the satisfaction probability of the original system. A direct consequence is that starting within the winning/losing region computed by the abstraction can guarantee a satisfaction/dissatisfaction for the original system. The second contribution of this paper is to deal with stochastic systems with extra uncertain perturbations (due to, e.g., measurement errors or modeling uncertainties). Under mild assumptions, we show that, in verifying probabilistic satisfaction

of general ω-regular linear-time properties, IMC abstractions that are both sound and approximately complete are constructible for nonlinear stochastic systems. That is, given a concrete discrete-time continuous-state Markov system $\mathbb{X}$, and an arbitrarily small $\mathcal{L}_1$-bounded perturbation of this system, there exists an IMC abstraction whose set of satisfaction probability contains that of $\mathbb{X}$, and meanwhile is contained by that of the slightly perturbed system. We argue in Sect. 4 that to make the IMC abstraction robustly complete, the perturbation is generally necessary to be $\mathcal{L}_1$-bounded rather than only bounded in terms of point mass. We analyze the probabilistic properties based on the topology of metrizable space of (uncertain) probability measures, and show that the technique proves more powerful than purely discussing the value of probabilities. We also would like to clarify that the main purpose of this paper is not on providing more efficient algorithms for computing abstractions. We aim to provide a theoretical foundation of IMC abstractions for verifying continuous-state stochastic systems with perturbations and hope to shed some light on designing more powerful robust verification algorithms.

The rest of the paper is organized as follows. Section 2 presents some preliminaries on probability spaces and Markov systems. Section 3 presents the soundness of abstractions in verifying ω-regular linear-time properties for discrete-time nonlinear stochastic systems. Section 4 presents the constructive robust abstractions with soundness and approximate completeness guarantees. We discuss the differences of robustness between deterministic and stochastic systems. The paper is concluded in Sect. 5.

Notation: We denote by $\prod$ the product of ordinary sets, spaces, or function values. Denote by $\otimes$ the product of collections of sets, or sigma algebras, or measures. The n-times repeated product of any kind is denoted by $(\cdot)^n$ for simplification. Denote by $\pi_j : \prod_{i=0}^{\infty}(\cdot)_i \to (\cdot)_j$ the projection to the j^{th} component. We denote by $\mathscr{B}(\cdot)$ the Borel σ-algebra of a set.

Let $|\cdot|$ denote the infinity norm in $\mathbb{R}^n$, and let $\mathbb{B} := \{x \in \mathbb{R}^n : |x| \leq 1\}$. We denote by $\|\cdot\|_1 := \mathcal{E}|\cdot|$ the $\mathcal{L}_1$-norm for $\mathbb{R}^n$-valued random variables, and let $\mathbb{B}_1 := \{X : \mathbb{R}^n\text{-valued random variable with } \|X\|_1 \leq 1\}$. Given a matrix M, we denote by M_i its i^{th} row and by M_{ij} its entry at i^{th} row and j^{th} column.

Given a general state space $\mathcal{X}$, we denote by $\mathfrak{P}(\mathcal{X})$ the space of probability measures. The space of bounded and continuous functions on $\mathcal{X}$ is denoted by $C_b(\mathcal{X})$. For any stochastic processes $\{X_t\}_{t\geq 0}$, we use the shorthand notation $X := \{X_t\}_{t\geq 0}$. For any stopped process $\{X_{t\wedge\tau}\}_{t\geq 0}$, where τ is a stopping time, we use the shorthand notation X^τ.

2 Preliminaries

We consider $\mathbb{N} = \{0, 1, \cdots\}$ as the discrete time index set, and a general Polish (complete and separable metric) space $\mathcal{X}$ as the state space. For any discrete-time $\mathcal{X}^\infty$-valued stochastic process X, we introduce some standard concepts.

2.1 Canonical Sample Space

Given a stochastic process X defined on some (most likely unknown) probability space $(\Omega^\dagger, \mathscr{F}^\dagger, \mathbb{P}^\dagger)$. Since we only care about the probabilistic behavior of trajectories in the state space, we prefer to work on the canonical probability spaces $(\Omega, \mathcal{F}, \mathcal{P}) := (\mathcal{X}^\infty,\ \mathscr{B}(\mathcal{X}^\infty),\ \mathbb{P}^\dagger \circ X^{-1})$ and regard events as sets of sample paths (see details in [23, Section 2.1]). We denote by $\mathcal{E}$ the associated expectation. In the context of discrete state spaces $\mathcal{X}$, we specifically use the boldface notation $(\mathbf{\Omega}, \mathbf{F}, \mathbf{P})$ for the canonical spaces of some discrete-state processes.

Remark 1. *We usually denote by ν_i the marginal distribution of $\mathcal{P}$ at some $i \in \mathbb{N}$. We can informally write the n-dimensional distribution (on n-dimensional cylinder set) as $\mathcal{P}(\cdot) = \otimes_{i=1}^n \nu_i(\cdot)$ regardless of the dependence.*

2.2 Markov Transition Systems

For any discrete-time stochastic process X, we set $\mathcal{F}_t = \sigma\{X_0, X_1, \cdots, X_t\}$ to be the natural filtration.

Definition 1 (Markov process). *A stochastic process X is said to be a Markov process if each X_t is $\mathcal{F}_t$-adapted and, for any $\Gamma \in \mathscr{B}(\mathcal{X})$ and $t > s$, we have*

$$\mathcal{P}[X_t \in \Gamma \mid \mathcal{F}_s] = \mathcal{P}[X_t \in \Gamma \mid \sigma\{X_s\}], \quad a.s. \tag{1}$$

Correspondingly, for every t, we define the transition probability as

$$\Theta_t(x, \Gamma) := \mathcal{P}[X_{t+1} \in \Gamma \mid X_t = x], \quad \Gamma \in \mathscr{B}(\mathcal{X}). \tag{2}$$

We denote $\Theta_t := \{\Theta_t(x, \Gamma) :\ x \in \mathcal{X},\ \Gamma \in \mathscr{B}(\mathcal{X})\}$ as the family of transition probabilities at time t. Note that homogeneous Markov processes are special cases such that $\Theta_t = \Theta_s$ for all $t, s \in \mathbb{N}$.

We are interested in Markov processes with discrete observations of states, which is done by assigning abstract labels over a finite set of atomic propositions. We define an abstract family of labelled Markov processes as follows.

Definition 2 (Markov system). *A Markov system is a tuple $\mathbb{X} = (\mathcal{X}, [\![\Theta]\!], \Pi, L)$, where*

- *$\mathcal{X} = \mathcal{W} \cup \Delta$, where $\mathcal{W}$ is a bounded working space, $\Delta := \mathcal{W}^c$ represents all the out-of-domain states;*
- *$[\![\Theta]\!]$ is a collection of transition probabilities from which Θ_t is chosen for every t;*
- *Π is the finite set of atomic propositions;*
- *$L : \mathcal{X} \to 2^\Pi$ is the (Borel-measurable) labelling function.*

For $X \in \mathbb{X}$ with $X_0 = x_0$ a.s., we denote by $\mathcal{P}_X^{x_0}$ the law, and $\{\mathcal{P}_X^{x_0}\}_{X \in \mathbb{X}}$ by its collection. Similarly, for any initial distribution $\nu_0 \in \mathfrak{P}(\mathcal{X})$, we define the law by $\mathcal{P}_X^{\nu_0}(\cdot) = \int_{\mathcal{X}} \mathcal{P}_X^x(\cdot)\nu_0(dx)$, and denote $\{\mathcal{P}_X^{\nu_0}\}_{X \in \mathbb{X}}$ by its collection. We denote by $\{\mathcal{P}_n^{q_0}\}_{n=0}^\infty$ (resp. $\{\mathcal{P}_n^{\nu_0}\}_{n=0}^\infty$) a sequence of $\{\mathcal{P}_X^{x_0}\}_{X \in \mathbb{X}}$ (resp. $\{\mathcal{P}_X^{\nu_0}\}_{X \in \mathbb{X}}$). We simply use $\mathcal{P}_X$ (resp. $\{\mathcal{P}_X\}_{X \in \mathbb{X}}$) if we do not emphasize the initial condition.

For a path $\varpi := \varpi_0\varpi_1\varpi_2\cdots \in \Omega$, define by $L_\varpi := L(\varpi_0)L(\varpi_1)L(\varpi_2)\cdots$ its trace. The space of infinite words is denoted by

$$(2^\Pi)^\omega = \{A_0A_1A_2\cdots : A_i \in 2^\Pi,\ i = 0, 1, 2\cdots\}.$$

A linear-time (LT) property is a subset of $(2^\Pi)^\omega$. We are only interested in LT properties Ψ such that $\Psi \in \mathscr{B}((2^\Pi)^\omega)$, i.e., those are Borel-measurable.

Remark 2. *Note that, by [35] and [36, Proposition 2.3], any ω-regular language of labelled Markov processes is measurable. It follows that, for any Markov process X of the given $\mathbb{X}$, the traces L_ϖ generated by measurable labelling functions are also measurable. For each $\Psi \in \mathscr{B}((2^\Pi)^\omega)$, we have the event $L_\varpi^{-1}(\Psi) \in \mathcal{F}$.*

A particular subclass of LT properties can be specified by linear temporal logic (LTL)[1]. To connect with LTL specifications, we introduce the semantics of path satisfaction as well as probabilistic satisfaction as follows.

Definition 3. *For the syntax of LTL formulae Ψ and the semantics of satisfaction of Ψ on infinite words, we refer readers to [20, Section 2.4].*

For a given labelled Markov process X from $\mathbb{X}$ with initial distribution ν_0, we formulate the canonical space $(\Omega, \mathcal{F}, \mathcal{P}_X^{\nu_0})$. For a path $\varpi \in \Omega$, we define the path satisfaction as

$$\varpi \vDash \Psi \iff L_\varpi \vDash \Psi.$$

We denote by $\{X \vDash \Psi\} := \{\varpi :\ \varpi \vDash \Psi\} \in \mathcal{F}$ the events of path satisfaction. Given a specified probability $\rho \in [0, 1]$, we define the probabilistic satisfaction of Ψ as

$$X \vDash \mathcal{P}_{\bowtie\rho}^{\nu_0}[\Psi] \iff \mathcal{P}_X^{\nu_0}\{X \vDash \Psi\} \bowtie \rho,$$

where $\bowtie \in \{\leq, <, \geq, >\}$.

2.3 Weak Convergence and Prokhorov's Theorem

We consider the set of possible uncertain measures within the topological space of probability measures. The following concepts are frequently used later.

Definition 4 (Tightness of set of measures). *Let $\mathcal{X}$ be any topological state space and $M \subseteq \mathfrak{P}(\mathcal{X})$ be a set of probability measures on $\mathcal{X}$. We say that M is tight if, for every $\varepsilon > 0$ there exists a compact set $K \subset \mathcal{X}$ such that $\mu(K) \geq 1-\varepsilon$ for every $\mu \in M$.*

Definition 5 (Weak convergence). *A sequence $\{\mu_n\}_{n=0}^\infty \subseteq \mathfrak{P}(\mathcal{X})$ is said to converge weakly to a probability measure μ, denoted by $\mu_n \Rightarrow \mu$, if*

$$\int_{\mathcal{X}} h(x)\mu_n(dx) \to \int_{\mathcal{X}} h(x)\mu(dx),\ \ \forall h \in C_b(\mathcal{X}). \tag{3}$$

We frequently use the following alternative condition [8, Proposition 2.2]:

$$\mu_n(A) \to \mu(A),\ \ \forall A \in \mathscr{B}(\mathcal{X})\ s.t.\ \mu(\partial A) = 0. \tag{4}$$

[1] While we consider LTL due to our interest, it can be easily seen that all results of this paper in fact hold for any measurable LT property, including ω-regular specifications.

It is straightforward from Definition 5 that weak convergence of measures also describes the convergence of probabilistic properties. We refer readers to [27] and [23, Remark 3] for more details on the weak topology.

Theorem 1 (Prokhorov). *Let $\mathcal{X}$ be a complete separable metric space. A family $\Lambda \subseteq \mathfrak{P}(\mathcal{X})$ is relatively compact if and only if it is tight. Consequently, for each sequence $\{\mu_n\}$ of tight Λ, there exists a $\mu \in \bar{\Lambda}$ and a subsequence $\{\mu_{n_k}\}$ such that $\mu_{n_k} \Rightarrow \mu$.*

Remark 3. *The first part of Prokhorov's theorem provides an alternative criterion for verifying the compactness of family of measures w.r.t. the corresponding metric space using tightness. On a compact metric space $\mathcal{X}$, every family of probability measures is tight.*

2.4 Discrete-Time Continuous-State Stochastic Systems

We define Markov processes determined by the difference equation

$$X_{t+1} = f(X_t) + b(X_t)w_t + \vartheta\xi_t \tag{5}$$

where the state $X_t(\varpi) \in \mathcal{X} \subseteq \mathbb{R}^n$ for all $t \in \mathbb{N}$, the stochastic inputs $\{w_t\}_{t\in\mathbb{N}}$ are i.i.d. Gaussian random variables with covariance $I_{k\times k}$ without loss of generality. Mappings $f : \mathbb{R}^n \to \mathbb{R}^n$ and $b : \mathbb{R}^n \to \mathbb{R}^{n\times k}$ are locally Lipschitz continuous. The memoryless perturbation $\xi_t \in \mathbb{B}_1$ are independent random variables with intensity $\vartheta \geq 0$ and unknown distributions.

For $\vartheta \neq 0$, (5) defines a family $\mathbb{X}$ of Markov processes X. A special case of (5) is such that ξ has Dirac (point-mass) distributions $\{\delta_x : x \in \mathbb{B}\}$ centered at some uncertain points within a unit ball.

Remark 4. *Gaussian random variables are naturally selected to simulate Brownian motions at discrete times. Note that in [11], random variables are used with known unimodal symmetric density with an interval as the support. Their choice is in favor of the mixed-monotone models to provide a more accurate approximation of transition probabilities. Other than the precision issue, such a choice does not bring us more of the other $\mathcal{L}_1$ properties. Since we focus on formal analysis based on $\mathcal{L}_1$ properties rather than providing accurate approximation, using Gaussian randomnesses as a realization does not lose any generality.*

We only care about the behaviors in the bounded working space $\mathcal{W}$. By defining stopping time $\tau := \inf\{t \in \mathbb{N} : X \notin \mathcal{W}\}$ for each X, we are able to study the probability law of the corresponding stopped (killed) process X^τ for any initial condition x_0 (resp. ν_0), which coincides with $\mathcal{P}_X^{x_0}$ (resp. $\mathcal{P}_X^{\nu_0}$) on $\mathcal{W}$. To avoid any complexity, we use the same notation X and $\mathcal{P}_X^{x_0}$ (resp. $\mathcal{P}_X^{\nu_0}$) to denote the stopped processes and the associated laws. Such processes driven by (5) can be written as a Markov system

$$\mathbb{X} = (\mathcal{X}, [\![\mathcal{T}]\!], \Pi, L), \tag{6}$$

where for all $x \in \mathcal{X} \setminus \mathcal{W}$, the transition probability should satisfy $\mathcal{T}(x, \Gamma) = 0$ for all $\Gamma \cap \mathcal{W} \neq \emptyset$; $[\![\mathcal{T}]\!]$ is the collection of transition probabilities. For ξ having Dirac distributions, the transition $\mathcal{T}$ is of the following form:

$$\mathcal{T}(x, \cdot) \in \begin{cases} \{\mu \sim \mathcal{N}(f(x) + \vartheta\xi,\ \ b(x)b^T(x)),\ \xi \in \mathbb{B}\}, & \forall x \in \mathcal{W}, \\ \{\mu :\ \mu(\Gamma) = 0,\ \ \forall \Gamma \cap \mathcal{W} \neq \emptyset\}, & \forall x \in \mathcal{X} \setminus \mathcal{W}. \end{cases} \tag{7}$$

Assumption 1. *We assume that* $\boldsymbol{in} \in L(x)$ *for any* $x \notin \Delta$ *and* $\boldsymbol{in} \notin L(\Delta)$*. We can also include 'always* $(\boldsymbol{in})$*' in the specifications to observe sample paths for 'inside-domain' behaviors, which is equivalent to verifying* $\{\tau = \infty\}$*.*

2.5 Robust Abstractions

We define a notion of abstraction between continuous-state and finite-state Markov systems via state-level relations and measure-level relations.

Definition 6. *A (binary) relation* γ *from* A *to* B *is a subset of* $A \times B$ *satisfying (i) for each* $a \in A$, $\gamma(a) := \{b \in B : (a, b) \in \gamma\}$*; (ii) for each* $b \in B$, $\gamma^{-1}(b) := \{a \in A : (a, b) \in \gamma\}$*; (iii) for* $A' \subseteq A$, $\gamma(A') = \cup_{a \in A'} \gamma(a)$*; (iv) and for* $B' \subseteq B$, $\gamma^{-1}(B') = \cup_{b \in B'} \gamma^{-1}(b)$*.*

Definition 7. *Given a continuous-state Markov system*

$$\mathbb{X} = (\mathcal{X}, [\![\mathcal{T}]\!], \Pi, L)$$

and a finite-state Markov system

$$\mathbb{I} = (\mathcal{Q}, [\![\Theta]\!], \Pi, L_{\mathbb{I}}),$$

where $\mathcal{Q} = (q_1, \cdots, q_n)^T$ *and* $[\![\Theta]\!]$ *stands for a collection of* $n \times n$ *stochastic matrices. A state-level relation* $\alpha \subseteq \mathcal{X} \times \mathcal{Q}$ *is said to be an abstraction from* $\mathbb{X}$ *to* $\mathbb{I}$ *if (i) for all* $x \in \mathcal{X}$*, there exists* $q \in \mathcal{Q}$ *such that* $(x, q) \in \alpha$*; (ii) for all* $(x, q) \in \alpha$*,* $L_{\mathbb{I}}(q) = L(x)$*.*

A measure-level relation $\gamma_\alpha \subseteq \mathfrak{P}(\mathcal{X}) \times \mathfrak{P}(\mathcal{Q})$ *is said to be an abstraction from* $\mathbb{X}$ *to* $\mathbb{I}$ *if for all* $i \in \{1, 2, \cdots, n\}$*, all* $\mathcal{T} \in [\![\mathcal{T}]\!]$ *and all* $x \in \alpha^{-1}(q_i)$*, there exists* $\Theta \in [\![\Theta]\!]$ *such that* $(\mathcal{T}(x, \cdot), \Theta_i) \in \gamma_\alpha$ *and that* $\mathcal{T}(x, \alpha^{-1}(q_j)) = \Theta_{ij}$ *for all* $j \in \{1, 2, \cdots, n\}$*.*

Similarly, $\gamma_\alpha \subseteq \mathfrak{P}(\mathcal{Q}) \times \mathfrak{P}(\mathcal{X})$ *is said to be an abstraction from* $\mathbb{I}$ *to* $\mathbb{X}$ *if for all* $i \in \{1, 2, \cdots, n\}$*, all* $\Theta \in [\![\Theta]\!]$ *and all* $x \in \alpha^{-1}(q_i)$*, there exists* $\mathcal{T} \in [\![\mathcal{T}]\!]$ *such that* $(\Theta_i, \mathcal{T}(x, \cdot)) \in \gamma_\alpha$ *and that* $\mathcal{T}(x, \alpha^{-1}(q_j)) = \Theta_{ij}$ *for all* $j \in \{1, 2, \cdots, n\}$*.*

If such relations α *and* γ_α *exist, we say that* $\mathbb{I}$ *abstracts* $\mathbb{X}$ *(resp.* $\mathbb{X}$ *abstracts* $\mathbb{I}$*) and write* $\mathbb{X} \preceq_{\gamma_\alpha} \mathbb{I}$ *(resp.* $\mathbb{I} \preceq_{\gamma_\alpha} \mathbb{X}$*).*

Assumption 2. *Without loss of generality, we assume that the labelling function is amenable to a rectangular partition*[2]*. In other words, a state-level abstraction can be obtained from a rectangular partition.*

[2] See e.g. [11, Definition 1].

3 Soundness of Robust IMC Abstractions

IMCs[3] are quasi-Markov systems on a discrete state space with upper/under approximations ($\hat{\Theta}/\check{\Theta}$) of the real transition matrices. To abstract the transition probabilities of continuous-state Markov systems (6), $\hat{\Theta}$ and $\check{\Theta}$ are obtained from over/under approximations of $\mathcal{T}$ based on the state space partition. Throughout this section, we assume that $\hat{\Theta}$ and $\check{\Theta}$ have been correspondingly constructed.

Given an IMC, we recast it to a true finite-state Markov system

$$\mathbb{I} = (\mathcal{Q}, [\![\Theta]\!], \Pi, L_{\mathbb{I}}), \tag{8}$$

where

- $\mathcal{Q}$ is the finite state-space partition with dimension $N+1$ containing $\{\Delta\}$, i.e., $\mathcal{Q} = (q_1, q_2, \cdots, q_N, \Delta)^T$;
- $[\![\Theta]\!]$[4] is a set of stochastic matrices satisfying

$$[\![\Theta]\!] = \{\Theta : \text{stochastic matrices with } \check{\Theta} \leq \Theta \leq \hat{\Theta} \text{ componentwisely}\}; \tag{9}$$

- $\Pi, L_{\mathbb{I}}$ are as before.

To make $\mathbb{I}$ an abstraction for (8), we need the approximation to be such that $\check{\Theta}_{ij} \leq \int_{q_j} \mathcal{T}(x, dy) \leq \hat{\Theta}_{ij}$ for all $x \in q_i$ and $i, j = 1, \cdots, N$, as well as $\Theta_{N+1} = (0, 0, \cdots, 1)$. We further require that the partition should respect the boundaries induced by the labeling function, i.e., for any $q \in \mathcal{Q}$,

$$L_{\mathbb{I}}(q) := L(x), \ \forall x \in q.$$

Clearly, the above connections on the state and transition probabilities satisfy Definition 7.

The Markov system $\mathbb{I}$ is understood as a family of 'perturbed' Markov chains generated by the uncertain choice of Θ for each t. The n-step transition matrices are derived based on $[\![\Theta]\!]$ as

$$\begin{aligned} [\![\Theta^{(2)}]\!] &= \{\Theta_0\Theta_1 : \ \ \Theta_0, \Theta_1 \in [\![\Theta]\!]\}, \\ &\cdots \\ [\![\Theta^{(n)}]\!] &= \{\Theta_0\Theta_1\cdots\Theta_n : \ \ \Theta_i \in [\![\Theta]\!], \ i = 0, 1, \cdots, n\}. \end{aligned}$$

Given an initial distribution $\mu_0 \in \mathfrak{P}(\mathcal{Q})$, the marginal probability measure at each t forms a set

$$\mathfrak{P}(\mathcal{Q}) \supseteq \mathscr{M}_t^{\mu_0} := \{\mu_t = (\Theta^{(t)})^T \mu_0 : \ \ \Theta^{(t)} \in [\![\Theta^{(t)}]\!]\}. \tag{10}$$

If we do not emphasize the initial distribution μ_0, we also use $\mathscr{M}_t$ to denote the marginals for short.

We aim to show the soundness of robust IMC abstractions in this section. The proofs in this section are completed in [23].

[3] We omit the definition from this paper due to the limitation of space. For a formal definition see e.g. [18, Definition 3].

[4] This is a necessary step to guarantee proper probability measures in (10). Algorithms can be found in [16] or [18, Section V-A].

3.1 Weak Compactness of Marginal Space $\mathscr{M}_t$ of Probabilities

The following lemma is rephrased from [37, Theorem 2] and shows the structure of the $\mathscr{M}_t$ for each $t \in \mathbb{N}$ and any initial distribution μ_0.

Lemma 1. *Let $\mathbb{I}$ be a Markov system of the form (8) that is derived from an IMC. Then the set $\mathscr{M}_t$ of all possible probability measures at each time $t \in \mathbb{N}$ is a convex polytope, and immediately is compact. The vertices of $\mathscr{M}_t$ are of the form*

$$(V_{i_t})^T \cdots (V_{i_2})^T (V_{i_1})^T \mu_0 \tag{11}$$

for some vertices V_{i_j} of $[\![\Theta]\!]$.

An illustrative example is provided in [23, Example 1]. Now we introduce the total variation distance $\|\cdot\|_{\mathrm{TV}}$ and see how $(\mathscr{M}_t, \|\cdot\|_{\mathrm{TV}})$ (at each t) implies the weak topology.

Definition 8 (Total variation distance). *Given two probability measures μ and ν on $\mathcal{X}$, the total variation distance is defined as*

$$\|\mu - \nu\|_{TV} = 2 \sup_{\Gamma \in \mathscr{B}(\mathcal{X})} |\mu(\Gamma) - \nu(\Gamma)|. \tag{12}$$

In particular, if $\mathcal{X}$ is a discrete space, $\|\mu - \nu\|_{TV} = \sum_{q \in \mathcal{X}} |\mu(q) - \nu(q)|$ (1-norm).

Corollary 1. *Let $\mathbb{I}$ be a Markov system of the form (8) that is derived from an IMC. Then at each time $t \in \mathbb{N}$, for each $\{\mu_n\} \subseteq \mathscr{M}_t$, there exists a $\mu \in \mathscr{M}_t$ and a subsequence $\{\mu_{n_k}\}$ such that $\mu_{n_k} \Rightarrow \mu$. In addition, for each $h \in C_b(\mathcal{X})$ and $t \in \mathbb{N}$, the set $H = \{\sum_{\mathcal{X}} h(x)\mu(x),\ \mu \in \mathscr{M}_t\}$ forms a convex and compact subset in $\mathbb{R}$.*

Remark 5. *Note that since $\mathcal{Q}$ is bounded and finite, any metrizable family of measures on $\mathcal{Q}$ is compact. However, the convexity does not hold in general (see [23, Remark 6] for details).*

3.2 Weak Compactness of Probability Laws of $\mathbb{I}$ on Infinite Horizon

We focus on the case where $I_0 = q_0$ a.s. for any $q_0 \in \mathcal{Q} \setminus \{\Delta\}$. The cases for arbitrary initial distribution should be similar. We formally denote $\mathscr{M}^{q_0} := \{\mathbf{P}_I^{q_0}\}_{I \in \mathbb{I}}$ by the set of probability laws of every discrete-state Markov processes $I \in \mathbb{I}$ with initial state $q_0 \in \mathcal{Q}$. We denote $\mathscr{M}_t^{q_0}$ by the set of marginals at t.

Proposition 1. *For any $q_0 \in \mathcal{Q}$, every sequence $\{\mathbf{P}_n^{q_0}\}_{n=0}^{\infty}$ of $\mathscr{M}^{q_0}$ has a weakly convergent subsequence.*

The above property is an extension of the marginal weak compactness relying on the (countable) product topology. The following result demonstrates the probabilistic regularity of general IMC abstractions.

Theorem 2. *Let $\mathbb{I}$ be a Markov system of the form (8) that is derived from an IMC. Then for any LTL formula Ψ, the set $S^{q_0} = \{\mathbf{P}_I^{q_0}(I \vDash \Psi)\}_{I \in \mathbb{I}}$ is a convex and compact subset in $\mathbb{R}$, i.e., a compact interval.*

3.3 Soundness of IMC Abstractions

Proposition 2. *Let $\mathbb{X}$ be a Markov system driven by (6). Then every sequence $\{\mathcal{P}_n^{x_0}\}_{n=0}^{\infty}$ of $\{\mathcal{P}_X^{x_0}\}_{X\in\mathbb{X}}$ has a weakly convergent subsequence. Consequently, for any LTL formula Ψ, the set $\{\mathcal{P}_X^{x_0}(X \vDash \Psi)\}_{X\in\mathbb{X}}$ is a compact subset in $\mathbb{R}$.*

Lemma 2. *Let $X \in \mathbb{X}$ be any Markov process driven by (6) and $\mathbb{I}$ be the finite-state IMC abstraction of $\mathbb{X}$. Suppose the initial distribution ν_0 of X is such that $\nu_0(q_0) = 1$. Then, there exists a unique law $\mathbf{P}_I^{q_0}$ of some $I \in \mathbb{I}$ such that, for any LTL formula Ψ,*

$$\mathcal{P}_X^{\nu_0}(X \vDash \Psi) = \mathbf{P}_I^{q_0}(I \vDash \Psi).$$

Theorem 3. *Assume the settings in Lemma 2. For any LTL formula Ψ, we have*

$$\mathcal{P}_X^{\nu_0}(X \vDash \Psi) \in \{\mathbf{P}_I^{q_0}(I \vDash \Psi)\}_{I\in\mathbb{I}},$$

Proof. The conclusion is obtained by combining Lemma 2 and Theorem 2. ■

Corollary 2. *Let $\mathbb{X}$, its IMC abstraction $\mathbb{I}$, an LTL formula Ψ, and a constant $\rho \in [0, 1]$ be given. Suppose $I \vDash \mathbf{P}_{\bowtie\rho}^{q_0}[\Psi]$ for all $I \in \mathbb{I}$, we have $X \vDash \mathcal{P}_{\bowtie\rho}^{\nu_0}[\Psi]$ for all $X \in \mathbb{X}$ with $\nu_0(q_0) = 1$.*

Remark 6. *Note that we do not have $\mathcal{P}_X^{\nu_0} \in \{\mathbf{P}_I^{q_0}\}_{I\in\mathbb{I}}$ since each $\mathbf{P}_I^{q_0}$ is a discrete measure whereas $\mathcal{P}_X^{\nu_0}$ is not. They only coincide when measuring Borel subset of $\mathbf{F}$ (recall notation in Sect. 2.1). It would be more accurate to state that $\mathcal{P}_X^{\nu_0}(X \vDash \Psi)$ is a member of $\{\mathbf{P}_I^{q_0}(I \vDash \Psi)\}_{I\in\mathbb{I}}$ rather than say "the true distribution (the law as what we usually call) of the original system is a member of the distribution set represented by the abstraction model" [18].*

Proposition 3. *Let $\varepsilon := \max_i \|\hat{\Theta}_i - \check{\Theta}_i\|_{TV}$. Then for each LTL formula Ψ, as $\varepsilon \to 0$, the length $\lambda(S^{q_0}) \to 0$.*

By Lemma 2, for each $X \in \mathbb{X}$, there exists exactly one $\mathbf{P}_I$ of some $I \in \mathbb{I}$ by which satisfaction probability equals to that of X. The precision of $\hat{\Theta}$ and $\check{\Theta}$ determines the size of S^{q_0}. Once we are able to calculate the exact law of X, the S^{q_0} becomes a singleton by Proposition 3. Special cases are provided in [23, Remark 10].

4 Robust Completeness of IMC Abstractions

In this section, we are given a Markov system $\mathbb{X}_1$ driven by (5) with point-mass perturbations of strength $\vartheta_1 \geq 0$. Based on $\mathbb{X}_1$, we first construct an IMC abstraction $\mathbb{I}$. We then show that $\mathbb{I}$ can be abstracted by a system $\mathbb{X}_2$ with more general $\mathcal{L}_1$-bounded noise of any arbitrary strength $\vartheta_2 > \vartheta_1$.

Recalling the soundness analysis of IMC abstractions in Sect. 3, the relation of satisfaction probability is induced by a relation between the continuous and discrete transitions. To capture the probabilistic properties of stochastic processes, reachable set of probability measures is the analogue of the reachable set

in deterministic cases. We rely on a similar technique in this section to discuss how transition probabilities of different uncertain Markov systems are related. To metricize sets of Gaussian measures and to connect them with discrete measures, we prefer to use Wasserstein metric.

Definition 9. *Let $\mu, \nu \in \mathfrak{P}(\mathcal{X})$ for $(\mathcal{X}, |\cdot|)$, the Wasserstein distance*[5] *is defined by $\|\mu - \nu\|_W = \inf \mathcal{E}|X - Y|$, where the infimum is taken over all joint distributions of the random variables X and Y with marginals μ and ν respectively. We frequently use the following duality form of definition*[6]*,*

$$\|\mu - \nu\|_W := \sup \left\{ \left| \int_{\mathcal{X}} h(x) d\mu(x) - \int_{\mathcal{X}} h(x) d\nu(x) \right|, \ h \in C(\mathcal{X}), \mathrm{Lip}(h) \leq 1 \right\}.$$

The discrete case, $\|\cdot\|_W^d$, is nothing but to change the integral to summation. Let $B_W = \{\mu \in \mathfrak{P}(\mathcal{X}) : \|\mu - \delta_0\|_W \leq 1\}$. Given a set $\mathfrak{G} \subseteq \mathfrak{P}(\mathcal{X})$, we denote $\|\mu\|_{\mathfrak{G}} = \inf_{\nu \in \mathfrak{G}} \|\mu - \nu\|_W$ by the distance from μ to $\mathfrak{G}$, and $\mathfrak{G} + r\mathbb{B}_W := \{\mu : \ \|\mu\|_{\mathfrak{G}} \leq r\}$[7] *by the r-neighborhood of $\mathfrak{G}$.*

Note that $\mathbb{B}_W$ is dual to $\mathbb{B}_1$. For any $\mu \in \mathbb{B}_W$, the associated random variable X should satisfy $\mathcal{E}|X| \leq 1$, and vice versa. The following well-known result estimates the Wasserstein distance between two Gaussians.

Proposition 4. *Let $\mu \sim \mathcal{N}(m_1, \Sigma_1)$ and $\nu \sim \mathcal{N}(m_2, \Sigma_2)$ be two Gaussian measures on $\mathbb{R}^n$. Then*

$$|m_1 - m_2| \leq \|\mu - \nu\|_W \leq \left(\|m_1 - m_2\|_2^2 + \|\Sigma_1^{1/2} - \Sigma_2^{1/2}\|_F^2 \right)^{1/2}, \tag{13}$$

where $\|\cdot\|_F$ is the Frobenius norm.

Proposition 5. *[13] For any μ, ν on some discrete and finite space $\mathcal{Q}$, we have*

$$\|\mu - \nu\|_W^d \leq \mathrm{diam}(\mathcal{Q}) \cdot \|\mu - \nu\|_{TV}. \tag{14}$$

Before proceeding, we define the set of transition probabilities of $\mathbb{X}_i$ from any box $[x] \subseteq \mathbb{R}^n$ as

$$\mathbb{T}_i([x]) = \{\mathcal{T}(x, \cdot) : \ \mathcal{T} \in [\![\mathcal{T}]\!]_i, \ x \in [x]\}, \ i = 1, 2,$$

and use the following lemma to approximate $\mathbb{T}_1([x])$.

Lemma 3. *Fix any $\varepsilon > 0$, any box $[x] \subseteq \mathbb{R}^n$. For all $\kappa > 0$, there exists a finitely terminated algorithm to compute an over-approximation of the set of (Gaussian) transition probabilities from $[x]$, such that*

$$\mathbb{T}_1([x]) \subseteq \widehat{\mathbb{T}_1([x])} \subseteq \mathbb{T}_1([x]) + \kappa \mathbb{B}_W,$$

where $\widehat{\mathbb{T}_1([x])}$ is the computed over-approximation set of Gaussian measures.

[5] This is formally termed as 1st-Wasserstein metric. We choose 1st-Wasserstein metric due to the convexity and nice property of test functions.

[6] $\mathrm{Lip}(h)$ is the Lipschitz constant of h such that $|h(x_2) - h(x_1)| \leq \mathrm{Lip}(h)|x_2 - x_1|$.

[7] This is valid by definition.

Remark 7. *The lemma renders the inclusions with larger Wasserstein distance to ensure no missing information about the covariances. The proof is provided in [23].*

Definition 10. *For $i = 1, 2$, we introduce the modified transition probabilities for $\mathbb{X}_i = (\mathcal{X}, [\![\mathcal{T}]\!]_i, x_0, \Pi, L)$ based on (7). For all $\mathcal{T}_i \in [\![\mathcal{T}]\!]_i$, let*

$$\tilde{\mathcal{T}}_i(x, \Gamma) = \begin{cases} \mathcal{T}_i(x, \Gamma), \ \forall \Gamma \subseteq \mathcal{W}, \ \forall x \in W, \\ \mathcal{T}_i(x, \mathcal{W}^c), \ \Gamma = \partial\mathcal{W}, \ \forall x \in W, \\ 1, \ \Gamma = \partial\mathcal{W}, \ x \in \partial\mathcal{W}. \end{cases} \tag{15}$$

Correspondingly, let $[\![\tilde{\mathcal{T}}]\!]$ denote the collection. Likewise, we also use $\tilde{(\cdot)}$ to denote the induced quantities of any other types w.r.t. such a modification.

Remark 8. *We introduce the concept only for analysis. The above modification does not affect the law of the stopped processes since we do not care about the 'out-of-domain' transitions. We use a weighted point mass to represent the measures at the boundary, and the mean should remain the same. It can be easily shown that the Wasserstein distance between any two measures in $[\![\tilde{\mathcal{T}}]\!](x, \cdot)$ is upper bounded by that of the non-modified ones.*

Theorem 4. *For any $0 \leq \vartheta_1 < \vartheta_2$, we set $\mathbb{X}_i = (\mathcal{X}, [\![\tilde{\mathcal{T}}]\!]_i, x_0, \Pi, L)$, $i = 1, 2$, where $\mathbb{X}_1$ is perturbed by point masses with intensity ϑ_1, and $\mathbb{X}_2$ is perturbed by general L_1-perturbation with intensity ϑ_2. Then, under Assumption 2, there exists a rectangular partition $\mathcal{Q}$ (state-level relation $\alpha \subseteq \mathcal{X} \times \mathcal{Q}$), a measure-level relation γ_α and a collection of transition matrices $[\![\Theta]\!]$, such that the system $\mathbb{I} = (\mathcal{Q}, [\![\Theta]\!], q_0, \Pi, L)$ abstracts $\mathbb{X}_1$ and is abstracted by $\mathbb{X}_2$ by the following relation:*

$$\mathbb{X}_1 \preceq_{\gamma_\alpha} \mathbb{I}, \ \ \mathbb{I} \preceq_{\gamma_\alpha^{-1}} \mathbb{X}_2. \tag{16}$$

Proof. We construct a finite-state IMC with partition $\mathcal{Q}$ and an inclusion of transition matrices $[\![\Theta]\!]$ as follows. By Assumption 2, we use uniform rectangular partition on $\mathcal{W}$ and set $\alpha = \{(x, q) : q = \eta \lfloor \frac{x}{\eta} \rfloor\} \cup \{(\Delta, \Delta)\}$, where $\lfloor \cdot \rfloor$ is the floor function and η is to be chosen later. Denote the number of discrete nodes by $N + 1$.

Note that any family of (modified) Gaussian measures $[\![\tilde{\mathcal{T}}]\!]_1$ is induced from $[\![\mathcal{T}]\!]_1$ and should contain its information. For any $\tilde{\mathcal{T}} \in [\![\tilde{\mathcal{T}}]\!]_1$ and $q \in \mathcal{Q}$,

(i) for all $\tilde{\nu} \sim \tilde{\mathcal{N}}(m, s^2) \in \tilde{\mathbb{T}}_1(\alpha^{-1}(q), \cdot)$, store $\{(m_l, \Sigma_l) = (\eta \lfloor \frac{m}{\eta} \rfloor, \eta^2 \lfloor \frac{s^2}{\eta^2} \rfloor)\}_l$;
(ii) for each l, define $\tilde{\nu}_l^{\text{ref}} \sim \tilde{\mathcal{N}}(m_l, \Sigma_l)$ (implicitly, we need to compute $\nu_l^{\text{ref}}(\Delta)$); compute $\tilde{\nu}_l^{\text{ref}}(\alpha^{-1}(q_j))$ for each $q_j \in \mathcal{Q} \setminus \Delta$;
(iii) for each l, define $\mu_l^{\text{ref}} = [\tilde{\nu}_l^{\text{ref}}(\alpha^{-1}(q_1)), \cdots, \tilde{\nu}_l^{\text{ref}}(\alpha^{-1}(q_N)), \tilde{\nu}_l^{\text{ref}}(\Delta)]$;
(iv) compute $\mathbf{ws} := (\sqrt{2N} + 2)\eta$ and $\mathbf{tv} := N\eta \cdot \mathbf{ws}$;
(v) construct $[\![\mu]\!] = \bigcup_l \{\mu : \left\|\mu - \mu_l^{\text{ref}}\right\|_{\text{TV}} \leq \mathbf{tv}(\eta), \ \mu(\Delta) + \sum_j^N \mu(q_j) = 1\}$;
(vi) Let $\gamma_\alpha = \{(\tilde{\nu}, \mu), \ \mu \in [\![\mu]\!]\}$ be a relation between $\tilde{\nu} \in \tilde{\mathbb{T}}(\alpha^{-1}(q))$ and the generated $[\![\mu]\!]$.

Repeat the above step for all q, the relation γ_α is obtained. The rest of the proof falls in the following steps. For $i \leq N$, we simply denote $\mathfrak{G}_i := \tilde{\mathbb{T}}_1(\alpha^{-1}(q_i))$ and $\hat{\mathfrak{G}}_i := \widehat{\tilde{\mathbb{T}}_1(\alpha^{-1}(q_i))}$.

Claim 1: For $i \leq N$, let $[\![\Theta_i]\!] = \gamma_\alpha(\hat{\mathfrak{G}}_i)$. Then the finite-state IMC $\mathbb{I}$ with transition collection $[\![\Theta]\!]$ abstracts $\mathbb{X}_1$.

Indeed, for each $i = 1, \cdots, N$ and each $\tilde{T}$, we have $\gamma_\alpha(\mathfrak{G}_i) \subseteq \gamma_\alpha(\widehat{\mathfrak{G}}_i)$. We pick any modified Gaussian $\tilde{\nu} \in \hat{\mathfrak{G}}_i$, there exists a $\tilde{\nu}^{\text{ref}}$ such that (by Proposition 4) $\left\|\tilde{\nu} - \tilde{\nu}^{\text{ref}}\right\|_{\text{W}} \leq \left\|\nu - \nu^{\text{ref}}\right\|_{\text{W}} \leq \sqrt{2N}\eta$. We aim to find all discrete measures μ induced from $\tilde{\nu}$ (such that their probabilities match on discrete nodes as requirement by Definition 7). All such μ should satisfy[8],

$$\begin{aligned} \left\|\mu - \mu^{\text{ref}}\right\|_{\text{W}}^d &= \left\|\mu - \mu^{\text{ref}}\right\|_{\text{W}} \\ &\leq \|\mu - \tilde{\nu}\|_{\text{W}} + \left\|\tilde{\nu} - \tilde{\nu}^{\text{ref}}\right\|_{\text{W}} + \left\|\tilde{\nu}^{\text{ref}} - \mu^{\text{ref}}\right\|_{\text{W}} \\ &\leq (2 + \sqrt{2N})\eta, \end{aligned} \tag{17}$$

where the first term of line 2 is bounded by,

$$\begin{aligned} \|\mu - \tilde{\nu}\|_{\text{W}} &= \sup_{h \in C(\mathcal{X}), \text{Lip}(h) \leq 1} \left| \int_{\mathcal{X}} h(x) d\mu(x) - \int_{\mathcal{X}} h(x) d\tilde{\nu}(x) \right| \\ &\leq \sup_{h \in C(\mathcal{X}), \text{Lip}(h) \leq 1} \sum_{j=1}^{n} \int_{\alpha^{-1}(q_j)} |h(x) - h(q_j)| d\tilde{\nu}(x) \\ &\leq \eta \sum_{j=1}^{n} \int_{\alpha^{-1}(q_j)} d\tilde{\nu}(x) \leq \eta, \end{aligned} \tag{18}$$

and the third term of line 2 is bounded in a similar way. By step (v)(vi) and Proposition 5, all possible discrete measures μ induced from $\tilde{\nu}$ should be included in $\gamma_\alpha(\hat{\mathfrak{G}}_i)$. Combining the above, for any $\tilde{\nu} \in \mathfrak{G}_i$ and hence in $\hat{\mathfrak{G}}_i$, there exists a discrete measures in $\Theta_i \in \gamma_\alpha(\hat{\mathfrak{G}}_i)$ such that for all q_j we have $\tilde{\nu}(\alpha^{-1}(q_j)) = \Theta_{ij}$. This satisfies the definition of abstraction.

Claim 2: $\gamma_\alpha^{-1}(\gamma_\alpha(\mathfrak{G}_i)) \subseteq \mathfrak{G}_i + (2\eta + N\eta \cdot \mathbf{tv}(\eta)) \cdot \mathbb{B}_W$. This is to recover all possible (modified) measures $\tilde{\nu}$ from the constructed $\gamma_\alpha(\mathfrak{G}_i)$, such that their discrete probabilities coincide. Note that, the 'ref' information is recorded when computing $\gamma_\alpha(\mathfrak{G}_i)$ in the inner parentheses. Therefore, for any $\mu \in \gamma_\alpha(\mathfrak{G}_i)$ there exists a μ^{ref} within a total variation radius $\mathbf{tv}(\eta)$. We aim to find corresponding measure $\tilde{\nu}$ that matches μ by their probabilities on discrete nodes. All such $\tilde{\nu}$ should satisfy,

$$\begin{aligned} \left\|\tilde{\nu} - \tilde{\nu}^{\text{ref}}\right\|_{\text{W}} &\leq \|\tilde{\nu} - \mu\|_{\text{W}} + \left\|\mu - \mu^{\text{ref}}\right\|_{\text{W}}^d + \left\|\mu^{\text{ref}} - \tilde{\nu}^{\text{ref}}\right\|_{\text{W}} \\ &\leq 2\eta + N\eta \cdot \mathbf{tv}(\eta), \end{aligned} \tag{19}$$

[8] Note that we also have $\left\|\mu - \mu^{\text{ref}}\right\|_{\text{W}}^d \leq \|\mu - \tilde{\nu}\|_{\text{W}}^d + \left\|\tilde{\nu} - \tilde{\nu}^{\text{ref}}\right\|_{\text{W}}^d + \left\|\tilde{\nu}^{\text{ref}} - \mu^{\text{ref}}\right\|_{\text{W}}^d = \left\|\tilde{\nu} - \tilde{\nu}^{\text{ref}}\right\|_{\text{W}}^d$, but it is hard to connect $\left\|\tilde{\nu} - \tilde{\nu}^{\text{ref}}\right\|_{\text{W}}^d$ with $\left\|\tilde{\nu} - \tilde{\nu}^{\text{ref}}\right\|_{\text{W}}$ for general measures. This connection can be done if we only compare Dirac or discrete measures.

where the bounds for the first and third terms are obtained in the same way as (18). The second term is again by a rough comparison in Proposition 5. Note that $\tilde{\nu}^{\text{ref}}$ is already recorded in $\mathfrak{G}_i$. The inequality in (19) provides an upper bound of Wasserstein deviation between any possible satisfactory measure and some $\tilde{\nu}^{\text{ref}} \in \mathfrak{G}_i$.

Claim 3: If we can choose η and κ sufficiently small such that $2\eta + N\eta \cdot \mathbf{tv}(\eta) + \kappa \leq \vartheta_2 - \vartheta_1$, then $\mathbb{I} \preceq_{\gamma_\alpha^{-1}} \mathbb{X}_2$. Indeed, the $[\![\Theta]\!]$ is obtained by $\gamma_\alpha(\hat{\mathfrak{G}}_i)$ for each i. By Claim 2 and Lemma 3, we have that for each i

$$\gamma_\alpha^{-1}(\gamma_\alpha(\hat{\mathfrak{G}}_i)) \subseteq \hat{\mathfrak{G}}_i + (2\eta + N\eta \cdot \mathbf{tv}(\eta)) \cdot \mathbb{B}_W \subseteq \mathfrak{G}_i + (2\eta + N\eta \cdot \mathbf{tv}(\eta) + \kappa) \cdot \mathbb{B}_W.$$

By the construction, we can verify that $\tilde{\mathbb{T}}_2(\alpha^{-1}(q_i)) = \mathfrak{G}_i + (\vartheta_2 - \vartheta_1) \cdot \mathbb{B}_W$. The selection of η makes $\gamma_\alpha^{-1}(\gamma_\alpha(\hat{\mathfrak{G}}_i)) \subseteq \tilde{\mathbb{T}}_2(\alpha^{-1}(q_i))$, which completes the proof. ■

Remark 9. *The relation γ_α (resp. γ_α^{-1}) provides a procedure to include all proper (continuous, discrete) measures that connect with the discrete probabilities. The key point is to record $\tilde{\nu}^{ref}$, μ^{ref}, and the corresponding radius. These are nothing but finite coverings of the space of measures. This also explains the reason why we use 'finite-state' rather than 'finite' abstraction. The latter has a meaning of using finite numbers of representative measures to be the abstraction.*

To guarantee a sufficient inclusion, conservative estimations are made. These estimations can be done more accurately given more assumptions (see details in [23, Remark 14]).

Remark 10. *To guarantee the second abstraction based on γ_α^{-1}, we search all possible measures that has the same discrete probabilities as $\mu \in \gamma_\alpha(\hat{\mathfrak{G}}_i)$, not only those Gaussians with the same covariances as $\mathfrak{G}_i$ (or $\hat{\mathfrak{G}}_i$). Such a set of measures provide a convex set w.r.t. Wasserstein distance. Recall that in the forward step of creating $\mathbb{I}$, we have used both Wasserstein and total variation distance to find a convex inclusion of all Gaussian or Gaussian related measures. There ought to be some measures that are 'non-recoverable' to Gaussians, unless we extract some 'Gaussian recoverable' discrete measures in $[\![\Theta_i]\!]$, but this loses the point of over-approximation. In this view, IMC abstractions provide unnecessarily larger inclusions than needed.*

For the deterministic case [20], the above mentioned 'extraction' is possible, since the transition measures do not have diffusion, the convex inclusion becomes a collection of vertices themselves (also see [23, Remark 6]). Based on these vertices, we are able to use γ_α to find the δ measures within a convex ball w.r.t. Wasserstein distance. In contrast to this special case [20], where the uncertainties are bounded w.r.t. the infinity norm, for stochastic systems, we can only guarantee the approximated completeness via a robust $\mathcal{L}_1$-bounded perturbation with strictly larger intensity than the original point-mass perturbation. However, this indeed describes a general type of uncertainties for the stochastic systems to guarantee $\mathcal{L}_1$-related properties, including probabilistic properties. Unless higher-moment specifications are of interests, uncertain $\mathcal{L}_1$-random variables are what we need to be the analogue of perturbations in [20].

Corollary 3. *Given an LTL formula Ψ, let $S_i^{\nu_0} = \{\mathcal{P}_X^{\nu_0}(X \vDash \Psi)\}_{X \in \mathbb{X}_i}$ $(i = 1, 2)$ and $S_{\mathbb{I}}^{q_0} = \{\mathbf{P}_I^{q_0}(I \vDash \Psi)\}_{I \in \mathbb{I}}$, where the initial conditions are such that $\nu_0(\alpha^{-1}(q_0)) = 1$. Then all the above sets are compact and $S_1^{\nu_0} \subseteq S_{\mathbb{I}}^{q_0} \subseteq S_2^{\nu_0}$.*

The proof is shown in [23].

5 Conclusion

In this paper, we constructed an IMC abstraction for continuous-state stochastic systems with possibly bounded point-mass (Dirac) perturbations. We showed that such abstractions are not only sound, in the sense that the set of satisfaction probability of linear-time properties contains that of the original system, but also approximately complete in the sense that the constructed IMC can be abstracted by another system with stronger but more general $\mathcal{L}_1$-bounded perturbations. Consequently, the winning set of the probabilistic specifications for a more perturbed continuous-state stochastic system contains that of the less Dirac perturbed system. Similar to most of the existing converse theorems, e.g. converse Lyapunov functions, the purpose is not to provide an efficient approach for finding them, but rather to characterize the theoretical possibilities of having such existence.

It is interesting to compare with robust deterministic systems, where no random variables are involved. In [20], both perturbed systems are w.r.t. bounded point masses. More heavily perturbed systems abstract less perturbed ones and hence preserve robust satisfaction of linear-time properties. However, when we try to obtain the approximated completeness via uncertainties in stochastic system, the uncertainties should be modelled by more general $\mathcal{L}_1$ random variables. Note that the probabilistic properties is dual to the weak topology of measures, we study the laws of processes instead of the state space *per se*. The state-space topology is not sufficient to quantify the regularity of IMC abstractions. In contrast, $\mathcal{L}_1$ uncertain random variables are perfect analogue of the uncertain point masses (in $|\cdot|$) for deterministic systems. If we insist on using point masses as the only type of uncertainties for stochastic systems, the IMC type abstractions would possibly fail to guarantee the completeness. For example, suppose the point-mass perturbations represent less precision of deterministic control inputs [22, Definition 2.3], the winning set decided by the ϑ_2-precision stationary policies is not enough to cover that of the IMC abstraction, which fails to ensure an approximated bi-similarity of IMCs compared to [20].

For future work, it would be useful to extend the current approach to robust stochastic control systems. It would be interesting to design algorithms to construct IMC (resp. bounded-parameter Markov decision processes) abstractions for more general robust stochastic (resp. control) systems with $\mathcal{L}_1$ perturbations based on metrizable space of measures and weak topology. The size of state discretization can be refined given more specific assumptions on system dynamics

and linear-time objectives. For verification or control synthesis w.r.t. probabilistic safety or reachability problems, comparisons can be made with stochastic Lyapunov-barrier function approaches.

References

1. Abate, A., D'Innocenzo, A., Di Benedetto, M.D., Sastry, S.S.: Markov set-chains as abstractions of stochastic hybrid systems. In: Egerstedt, M., Mishra, B. (eds.) HSCC 2008. LNCS, vol. 4981, pp. 1–15. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-78929-1_1
2. Abate, A., Katoen, J.P., Mereacre, A.: Quantitative automata model checking of autonomous stochastic hybrid systems. In: Proceedings of Hybrid Systems: Computation and Control (HSCC), pp. 83–92 (2011)
3. Abate, A., Prandini, M., Lygeros, J., Sastry, S.: Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. Automatica **44**(11), 2724–2734 (2008)
4. Baier, C., Katoen, J.P.: Principles of Model Checking. MIT Press, Cambridge (2008)
5. Belta, C., Yordanov, B., Aydin Gol, E.: Formal Methods for Discrete-Time Dynamical Systems. SSDC, vol. 89. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-50763-7
6. Bustan, D., Rubin, S., Vardi, M.Y.: Verifying ω-regular properties of Markov chains. In: Alur, R., Peled, D.A. (eds.) CAV 2004. LNCS, vol. 3114, pp. 189–201. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-27813-9_15
7. Cauchi, N., Laurenti, L., Lahijanian, M., Abate, A., Kwiatkowska, M., Cardelli, L.: Efficiency through uncertainty: scalable formal synthesis for stochastic hybrid systems. In: Proceedings of Hybrid Systems: Computation and Control (HSCC), pp. 240–251 (2019)
8. Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Cambridge University Press, Cambridge (2014)
9. Dehnert, C., Junges, S., Katoen, J.-P., Volk, M.: A storm is coming: a modern probabilistic model checker. In: Majumdar, R., Kunčak, V. (eds.) CAV 2017. LNCS, vol. 10427, pp. 592–600. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-63390-9_31
10. Delimpaltadakis, G., Laurenti, L., Mazo Jr., M.: Abstracting the sampling behaviour of stochastic linear periodic event-triggered control systems. arXiv preprint arXiv:2103.13839 (2021)
11. Dutreix, M., Coogan, S.: Specification-guided verification and abstraction refinement of mixed monotone stochastic systems. IEEE Trans. Autom. Control **66**(7), 2975–2990 (2020)
12. Dutreix, M.D.H.: Verification and synthesis for stochastic systems with temporal logic specifications. Ph.D. thesis, Georgia Institute of Technology (2020)
13. Gibbs, A.L., Su, F.E.: On choosing and bounding probability metrics. Int. Stat. Rev. **70**(3), 419–435 (2002)
14. Girard, A., Pola, G., Tabuada, P.: Approximately bisimilar symbolic models for incrementally stable switched systems. IEEE Trans. Autom. Control **55**(1), 116–126 (2009)
15. Givan, R., Leach, S., Dean, T.: Bounded-parameter Markov decision processes. Artif. Intell. **122**(1–2), 71–109 (2000)

16. Hartfiel, D.J.: Markov Set-Chains. Springer, Heidelberg (2006)
17. Kloetzer, M., Belta, C.: A fully automated framework for control of linear systems from temporal logic specifications. IEEE Trans. Autom. Control **53**(1), 287–297 (2008)
18. Lahijanian, M., Andersson, S.B., Belta, C.: Formal verification and synthesis for discrete-time stochastic systems. IEEE Trans. Autom. Control **60**(8), 2031–2045 (2015)
19. Li, Y., Liu, J.: Robustly complete synthesis of memoryless controllers for nonlinear systems with reach-and-stay specifications. IEEE Trans. Autom. Control **66**(3), 1199–1206 (2020)
20. Liu, J.: Robust abstractions for control synthesis: completeness via robustness for linear-time properties. In: Proceedings of Hybrid Systems: Computation and Control (HSCC), pp. 101–110 (2017)
21. Liu, J.: Closing the gap between discrete abstractions and continuous control: completeness via robustness and controllability. In: Dima, C., Shirmohammadi, M. (eds.) FORMATS 2021. LNCS, vol. 12860, pp. 67–83. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-85037-1_5
22. Majumdar, R., Mallik, K., Soudjani, S.: Symbolic controller synthesis for büchi specifications on stochastic systems. In: Proceedings of the 23rd International Conference on Hybrid Systems: Computation and Control, pp. 1–11 (2020)
23. Meng, Y., Liu, J.: Robustly complete finite-state abstractions for verification of stochastic systems. arXiv preprint arXiv:2205.01854 (2022)
24. Parker, D.: Verification of probabilistic real-time systems. In: Proceedings of 2013 Real-Time Systems Summer School (ETR 2013) (2013)
25. Pola, G., Girard, A., Tabuada, P.: Approximately bisimilar symbolic models for nonlinear control systems. Automatica **44**(10), 2508–2516 (2008)
26. Ramponi, F., Chatterjee, D., Summers, S., Lygeros, J.: On the connections between PCTL and dynamic programming. In: Proceedings of Hybrid Systems: Computation and Control (HSCC), pp. 253–262 (2010)
27. Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes and Martingales, Volume 1: Foundations. Cambridge Mathematical Library (2000)
28. Soudjani, S.E.Z., Abate, A.: Adaptive gridding for abstraction and verification of stochastic hybrid systems. In: 2011 Eighth International Conference on Quantitative Evaluation of SysTems, pp. 59–68. IEEE (2011)
29. Summers, S., Lygeros, J.: Verification of discrete time stochastic hybrid systems: a stochastic reach-avoid decision problem. Automatica **46**(12), 1951–1961 (2010)
30. Tabuada, P., Pappas, G.J.: Linear time logic control of discrete-time linear systems. IEEE Trans. Autom. Control **51**(12), 1862–1877 (2006)
31. Tkachev, I., Abate, A.: On infinite-horizon probabilistic properties and stochastic bisimulation functions. In: 2011 50th IEEE Conference on Decision and Control and European Control Conference, pp. 526–531. IEEE (2011)
32. Tkachev, I., Abate, A.: Regularization of bellman equations for infinite-horizon probabilistic properties. In: Proceedings of Hybrid Systems: Computation and Control (HSCC), pp. 227–236 (2012)
33. Tkachev, I., Abate, A.: Formula-free finite abstractions for linear temporal verification of stochastic hybrid systems. In: Proceedings of Hybrid Systems: Computation and Control (HSCC), pp. 283–292 (2013)
34. Tkachev, I., Abate, A.: Characterization and computation of infinite-horizon specifications over Markov processes. Theoret. Comput. Sci. **515**, 1–18 (2014)
35. Tkachev, I., Mereacre, A., Katoen, J.P., Abate, A.: Quantitative model-checking of controlled discrete-time Markov processes. Inf. Comput. **253**, 1–35 (2017)

36. Vardi, M.Y.: Automatic verification of probabilistic concurrent finite state programs. In: 26th Annual Symposium on Foundations of Computer Science (FOCS), pp. 327–338. IEEE (1985)
37. Vassiliou, P.C.: Non-homogeneous Markov set systems. Mathematics **9**(5), 471 (2021)
38. Wu, D., Koutsoukos, X.: Reachability analysis of uncertain systems using bounded-parameter Markov decision processes. Artif. Intell. **172**(8–9), 945–954 (2008)