# On Approximate Opacity of Stochastic Control Systems

Siyuan Liu , *Graduate Student Member, IEEE*, Xiang Yin , *Member, IEEE*,
Dimos V. Dimarogonas , *Fellow, IEEE*, and Majid Zamani , *Senior Member, IEEE*

*Abstract*—**This article investigates an important class of information-flow security property called opacity for stochastic control systems. Opacity captures whether a system's *secret* behavior (a subset of the system's behavior that is considered to be critical) can be kept from outside observers. Existing works on opacity for control systems only provide a binary characterization of the system's security level by determining whether the system is opaque or not. In this work, we introduce a quantifiable measure of opacity that considers the likelihood of satisfying opacity for stochastic control systems modeled as general Markov decision processes (gMDPs). We also propose verification methods tailored to the new notions of opacity for finite gMDPs by using value iteration techniques. Then, a new notion called *approximate opacity-preserving stochastic simulation relation* is proposed, which captures the distance between two systems' behaviors in terms of preserving opacity. Based on this new system relation, we show that one can verify opacity for stochastic control systems using their abstractions (modeled as finite gMDPs). We also discuss how to construct such abstractions for a class of gMDPs under certain stability conditions.**

*Index Terms*—**Approximate simulation relations, finite abstractions, general Markov decision process (gMDP), opacity, stochastic control systems.**

## I. INTRODUCTION

**T**HE analysis and synthesis of security for cyber-physical systems (CPS) have attracted considerable attentions over the past decades due to the increasing information exchange in modern CPS. Various notions of security properties were considered in the literature for the protection of CPS against different types of attacks, which are typically classified into three categories along the CIA mnemonic: confidentiality, integrity, and availability [1], [2], [3].

This article is concerned with a confidentiality property called *opacity*. In this problem, we assume that the system is monitored by an outside eavesdropper (malicious intruder) that is modeled as a passive observer. Opacity essentially characterizes the system's capability to conceal its *secrets* information from being inferred by malicious intruders through the information flow [4], [5]. The concept of opacity is known to be originally introduced in [6] in the computer science literature to analyze security and privacy performances of cryptographic protocols. Since then, opacity has attracted extensive attentions in the discrete-event systems (DES) community. Depending on how the systems are modeled and what kind of secret the systems want to hide, different opacity notions have been proposed in the context of DES, including language-based notions [7] and state-based notions [4], [8]. Building on the various notions, verification and synthesis approaches were extensively developed for DES modeled as finite-state automata [9], [10], labeled transition systems [11], or Petri nets [12]. The applications of opacity can be found in many real-world scenarios, including location-based services, mobile robot systems, and web services; see [5] and [13] for more details on the historic developments of opacity.

In more recent years, there has been a growing interest in the analysis of opacity for dynamical systems with continuous state spaces [3], [14], [15], [16]. Particularly, different from the concepts of opacity appeared in DES literature, new notions of approximate opacity were proposed in [14], which are more applicable to continuous-space CPS whose outputs are physical signals. This new concept provides a quantitative evaluation of a system's security level by taking into account the imperfect measurements of external observations. After that, various approaches have been proposed for the verification or synthesis of controllers with respect to approximate opacity for continuous-space control systems [14], [17], [18], [19]. In all the aforementioned works, the definitions of opacity for dynamical systems only provide a binary characterization by simply reporting if the system is opaque or not. However, in many security-critical real-world applications, such as air traffic networks, infrastructure systems, robotic manufacturing, and biological systems, the systems inherently exhibit stochastic behavior due to the uncertainties and randomness in their operations and environments [20], [21]. Consequently, a binary characterization of whether a desired property is satisfied becomes inadequate and impractical. Given that the uncertainty affecting system evolution results in varying likelihoods of different

Siyuan Liu and Dimos V. Dimarogonas are with the Division of Decision and Control Systems, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden (e-mail: siyliu@kth.se; dimos@kth.se).

Xiang Yin is with the Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China, and also with the Key Lab of System Control and Information Processing, Ministry of Education, Shanghai 200240, China (e-mail: yinxiang@sjtu.edu.cn).

Majid Zamani is with the Computer Science Department, University of Colorado Boulder, Boulder, CO 80309 USA, and also with the Computer Science Department, Ludwig Maximilian University of Munich, 80539 München, Germany (e-mail: majid.zamani@colorado.edu).

Digital Object Identifier 10.1109/TAC.2024.3516202

system trajectories, it is more reasonable to evaluate the probability of a system being opaque. In fact, in many real-world scenarios, a small likelihood of violating opacity may be tolerable. Typical examples include the quantitative analysis of security in cloud computing systems [22] and applications addressing location privacy issues [23], both of which rely on probabilistic quantification of opacity properties. In addition, when a system is not secure in a nonstochastic sense, introducing a quantitative notion of opacity and developing corresponding probabilistic verification techniques allow us to assess the *degree of satisfaction or violation* of the security property of interest. This approach can be instrumental in the early stages of system design.

*Our contribution:* In this article, we propose a new concept of *approximate opacity* for stochastic control systems modeled as general Markov decision processes (gMDPs) over continuous state spaces. This new notion defines appropriate measures to quantify opacity in a probabilistic setting. We consider two basic types of opacity: initial-state and current-state opacity. In particular, a notion of *approximate initial-state opacity* (resp., *approximate current-state opacity*) is proposed for gMDPs that requires that any state run of the system should not violate initial-state opacity (resp., current-state opacity) with a probability greater than a threshold. Intuitively, the proposed new notions of opacity can be regarded as the probabilistic counterparts of the (qualitative) opacity notions in [14], which allows us to quantitatively evaluate opacity for gMDPs. Based on these new notions of opacity, we further propose a verification approach for gMDPs using abstraction-based techniques. In particular, new notions of opacity-preserving simulation relations are introduced to capture the distance between two gMDPs in terms of preserving approximate opacity. We show that one can verify approximate opacity of a (infinite) gMDP by using its finite abstractions (modeled as finite gMDPs) based on these relations. In addition, we provide an approach to construct such finite abstractions for a class of discrete-time gMDPs. For the finite abstractions (finite gMDPs), we show that the verification of approximate opacity can be effectively achieved by using value iteration techniques.

*Related works:* In the setting of stochastic DES, there have been several results in the literature that consider how to quantitatively evaluate opacity by using probabilistic models [24], [25], [26], [27], [28]. For example, different probabilistic notions of current-state opacity were introduced in [24] for probabilistic finite-state automata; this approach has been extended to handle infinite-step and K-step opacity in [28]. In [26], a Jensen-Shannon divergence-based measure was adopted to quantify secrecy loss in partially observed stochastic DES. Probabilistic measurement of opacity are also studied in [25] and [27] for (partially observed) Markov decision processes. Nevertheless, all the aforementioned works are restricted to stochastic DES that evolves over discrete (finite) state spaces. In real-world applications, however, the gMDPs are equipped with continuous, and thus infinite, state sets. A suitably defined notion of opacity together with efficient verification techniques are indeed desirable in order to quantitatively evaluate the security level of general stochastic control systems.

There has been an attempt to extend the notion of opacity to stochastic control systems in [29]. Our results here differ from those in [29] in the following directions. First and foremost, the notion of opacity presented in [29] still provides a binary (qualitative) characterization of opacity, whereas the one proposed in this article is a quantitative probabilistic notion. Second, we develop abstraction-based verification approaches for both

initial-state and current-state opacity here, whereas [29] only investigates initial-state opacity. Finally, effective verification algorithms are proposed here to verify different notions of opacity for finite gMDPs, which was not investigated in [29].

*Organization:* The rest of this article is organized as follows. In Section II, we introduce notations and necessary preliminaries on gMDPs. In Section III, we propose new notions of approximate opacity for gMDPs. Verification approaches tailored to approximate opacity for finite gMDPs are presented in Section IV. Section V introduces new notions of opacity-preserving stochastic simulation relations and discuss their properties. In Section VI, we describe how to construct finite abstractions for a class of stochastic control systems. A numerical case study is provided in Section VII to illustrate the effectiveness of the proposed results. Finally, Section VIII concludes this article.

## II. PRELIMINARIES

*Notations:* In this article, a probability space is written as $(\Omega, \mathcal{F}_\Omega, \mathbb{P}_\Omega)$, where $\Omega$ is the sample space, $\mathcal{F}_\Omega$ is a sigma-algebra on $\Omega$ representing the set of events, and $\mathbb{P}_\Omega : \mathcal{F}_\Omega \to [0,1]$ is a probability measure that assigns probabilities to events. It is assumed that the random variables discussed in this article are measurable functions of the form $X : (\Omega, \mathcal{F}_\Omega) \to (S_X, \mathcal{F}_X)$. Any random variable $X$ induces a probability measure on its measurable space $(S_X, \mathcal{F}_X)$ as $\text{Prob}\{A\} = \mathbb{P}_\Omega\{X^{-1}(A)\}$ for any $A \in \mathcal{F}_X$. For brevity, we often directly discuss the probability measure on $(S_X, \mathcal{F}_X)$ without explicitly mentioning the underlying probability space and the function $X$ itself. We further denote the set of all probability measures for a given measurable pair $(S_X, \mathcal{F}_X)$ as $\mathcal{P}(S_X, \mathcal{F}_X)$. We denote by $\mathbb{R}$ and $\mathbb{N}$ the set of real numbers and nonnegative integers, respectively. These symbols are annotated with subscripts to restrict them in the usual way, e.g., $\mathbb{R}_{>0}$ denotes the positive real numbers. For $a, b \in \mathbb{N}$ and $a \leq b$, we use $[a; b]$ to denote the corresponding intervals in $\mathbb{N}$. We use $\|x\|$ to denote the infinity norm of a vector $x \in \mathbb{R}^n$. Given any $a \in \mathbb{R}$, $|a|$ denotes the absolute value of $a$. The composition of functions $f$ and $g$ is denoted by $f \circ g$. We use notations $\mathcal{K}$ and $\mathcal{K}_\infty$ to denote the different classes of comparison functions, as follows: $\mathcal{K} = \{\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} : \gamma$ is continuous, strictly increasing and $\gamma(0) = 0\}$; $\mathcal{K}_\infty = \{\gamma \in \mathcal{K} : \lim_{r \to \infty} \gamma(r) = \infty\}$. For $\alpha, \gamma \in \mathcal{K}_\infty$, we write $\alpha < \gamma$ if $\alpha(s) < \gamma(s)$ for all $s > 0$, and $\mathcal{I}_d \in \mathcal{K}_\infty$ denotes the identity function. Given sets $X$ and $Y$ with $X \subset Y$, the complement of $X$ with respect to $Y$ is defined as $Y \setminus X = \{x \in Y : x \notin X\}$.

### A. General Markov Decision Processes

In this article, we focus on a class of discrete-time stochastic control systems, which can be formally modeled as gMDPs [30], [31].

*Definition 2.1:* A gMDP is a tuple

$$\Sigma = (X, U, T, Y, h) \tag{1}$$

where $X \subseteq \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$, and $Y \subseteq \mathbb{R}^q$ are Borel spaces denoting the state, external input, and output sets of the system, respectively. We use $\mathcal{B}(X)$ to denote the Borel sigma algebra on the state set $X$, thus $(X, \mathcal{B}(X))$ denotes the corresponding measurable space. We use $T : \mathcal{B}(X) \times X \times U \to [0,1]$ to denote a conditional stochastic kernel that assigns to any state $x \in X$ and any input $u \in U$, a probability measure $T(\cdot|x, u)$ over the

measurable space $(X, \mathcal{B}(X))$ so that for any set $\mathcal{A} \in \mathcal{B}(X)$

$$\mathbb{P}(x' \in \mathcal{A}|x, u) = \int_{\mathcal{A}} T(dx'|x, u). \tag{2}$$

The measurable function $h : X \to Y$ is the output map.

A system $\Sigma$ with a finite state set $X$ and finite input set $U$ is said to be a finite gMDP. The dynamical evolution of a gMDP can be equivalently described by

$$\Sigma : \begin{cases} \xi(k+1) = f(\xi(k), \nu(k), \varsigma(k)), \\ \zeta(k) = h(\xi(k)), \end{cases} \quad k \in \mathbb{N} \tag{3}$$

where $\xi(\cdot) : \mathbb{N} \to X$, $\zeta(\cdot) : \mathbb{N} \to Y$, and $\nu(\cdot) : \mathbb{N} \to U$ are the state, output, and input signals, respectively. In the probability space $(\Omega, \mathcal{F}_\Omega, \mathbb{P})$, we use $\varsigma = (\varsigma(0), \varsigma(1), \dots)$ to denote a sequence of independent and identically distributed (i.i.d.) random variables from $\Omega$ to the measurable set $V_\varsigma$, where $\varsigma(k) : \Omega \to V_\varsigma, k \in \mathbb{N}$. We use $\mathcal{U}$ to denote the collections of input sequences $\nu(\cdot) : \mathbb{N} \to U$, where $\nu(k)$ is independent of $\varsigma(t)$ for any $k, t \in \mathbb{N}$ and $t \geq k$. The map $f : X \times U \times V_\varsigma \to X$ is a measurable function serving as the state transition relation. Note that any gMDP endowed with the stochastic kernel $T$ as in Definition 2.1 can be equivalently represented by a gMDP as in (3) with pair $(f, \varsigma)$ [32]. Throughout this article, the system is assumed to be *total*, i.e., for all $x \in X$, and for all $u \in U$, $f(x, u, \varsigma)$ is defined. As described in Definition 2.1, a gMDP operates over continuous or uncountable state sets, thereby extending an MDP by incorporating a metric output set where properties of interest are defined [30]. This modeling approach is sufficiently versatile to include specific cases such as discrete-time stochastic control systems [33], labeled Markov processes [34], discrete-time stochastic hybrid games [35], stochastic switched systems [36], and other types of stochastic hybrid systems.

Given any initial state $x_0$ and input sequence $\nu(\cdot) \in \mathcal{U}$, the random sequences $\xi_{x_0\nu} : \Omega \times \mathbb{N} \to X$ and $\zeta_{x_0\nu} : \Omega \times \mathbb{N} \to Y$ satisfying (3) are, respectively, called the *state trajectory* and *output trajectory* of $\Sigma$ under input sequence $\nu$ from $x_0$. Given an input sequence $\nu = (\nu_0, \nu_1, \dots, \nu_{L-1})$ over finite time horizon $L \in \mathbb{N}$, $\xi_{x_0\nu} = (x_0, x_1, \dots, x_L)$ is a *finite state trajectory* of length $|\xi_{x_0\nu}| = L$ after initialization. For any $\xi_{x_0\nu} = (x_0, x_1, \dots, x_i, \dots, x_L)$, we use $\xi_{x_0\nu}(i) = x_i$ to denote the $i$th state of $\xi_{x_0\nu}$, and by $\xi_{x_0\nu}[0; i] := (x_0, x_1, \dots, x_i)$ the prefix of $\xi_{x_0\nu}$ with length $i$. In addition, we drop the subscripts of $\xi_{x_0\nu}$ and denote by $\xi$ a state trajectory of the system when the initial state $x_0$ and input sequence $\nu$ are not of interest in some contexts. In this article, we are interested in the state trajectories within a finite time horizon $n \in \mathbb{N}$. Thus, let us define $\mathcal{L}^n := \{\xi : |\xi| = n\}$ and $\mathcal{H}^n := \{\zeta : |\zeta| = n\}$, respectively, as the set of all finite state trajectories and output trajectories of length $n \in \mathbb{N}$. We also denote by $\mathcal{U}^n := \{\nu(\cdot) : [0; n-1] \to U\}$ the set of all input sequences over time horizon $n \in \mathbb{N}$.

## III. APPROXIMATE OPACITY FOR gMDPs

In this section, we introduce new notions of (bounded-horizon) approximate opacity quantifying the *probability* of a gMDP preserving security within a given time horizon $n \in \mathbb{N}$. Throughout this article, we adopt a state-based formulation of secrets, i.e., the "secrets" of the system as a set of states $X_S \subseteq X$. Hereafter, we slightly modify the formulation in Definition 2.1 to accommodate for initial states and secret states, as $\Sigma = (X, X_0, X_S, U, T, Y, h)$, where $X_0 \subseteq X$ is a set of initial states and $X_S \subseteq X$ is a set of secret states.

Here, we consider a passive *intruder* (outside observer) who knows the dynamics of the system and can observe the output sequences of the system without actively affecting the behavior of the system. Furthermore, we assume that due to imperfect measurement precision, the intruder is unable to distinguish two "$\varepsilon$-close" output trajectories $\zeta = (y_0, y_1, \dots, y_n)$ and $\zeta' = (y'_0, y'_1, \dots, y'_n)$ if their distance is smaller than a given threshold $\varepsilon \geq 0$, i.e., $\max_{i \in [0;n]} \|y_i - y'_i\| \leq \varepsilon$. The two $\varepsilon$-close trajectories $\zeta$ and $\zeta'$ are said to be $\varepsilon$-approximate output-equivalent in the eye of an intruder with measurement precision greater than $\varepsilon \geq 0$. Essentially, the intruder aims to infer certain secret information about the system using the output sequences observed online and the knowledge of the system model. Opacity essentially captures whether or not the system's secrets can be revealed to the intruder. For the sake of simplicity, we assume that the input sequences are internal information and unknown to the intruder. However, one could indeed drop this assumption to handle the setting where both input and output information are available to the intruder.

Before introducing the formal definitions of approximate opacity for gMDPs, let us define the sets of "secure" state trajectories, i.e., the state trajectories whose occurrences do not reveal secrets to the passive intruder. Depending on what kind of information are regarded as critical, we consider the following two types of secrets, one w.r.t. initial-state information, and the other w.r.t. current-state information of the system.

1) *Initial-state secure trajectories:* For any state trajectory $\xi$ of length $|\xi| = L$, we say that $\xi$ *does not reveal the initial secret* if there exists $\xi'$ of length $L$ such that

$$\xi'(0) \in X_0 \setminus X_S \text{ and } \max_{i \in [0;L]} \|h(\xi(i)) - h(\xi'(i))\| \leq \varepsilon.$$

Essentially, the aforementioned indicates that for $\xi$, there always exists an $\varepsilon$-approximate output-equivalent state trajectory initiated from a nonsecret initial state. A state trajectory $\xi$ is said to be *initial-state secure* if it does not reveal the initial secret. We denote the set of initial-state secure state trajectories with finite time horizon $n \in \mathbb{N}$ as

$$\mathcal{L}^n_{\mathcal{I}_\varepsilon} := \{\xi \in \mathcal{L}^n : \xi \text{ is initial-state secure}\}. \tag{4}$$

2) *Current-state secure trajectories:* For any state trajectory $\xi$ of length $|\xi| = L$, we say that $\xi$ *does not reveal the current secret* if there exists $\xi'$ of length $L$ such that

$$\xi'(L) \in X \setminus X_S \text{ and } \max_{i \in [0;L]} \|h(\xi(i)) - h(\xi'(i))\| \leq \varepsilon.$$

Similarly, the aforementioned says that for $\xi$, there always exists an $\varepsilon$-approximate output-equivalent state trajectory arriving at a nonsecret current state. A state trajectory $\xi$ of length $L$ is said to be *current-state secure* if none of its prefix $\xi[0; k], k \leq L$ reveals the current secret. We denote the set of current-state secure state trajectories with finite time horizon $n \in \mathbb{N}$ as

$$\mathcal{L}^n_{\mathcal{C}_\varepsilon} := \{\xi \in \mathcal{L}^n : \xi \text{ is current-state secure}\}. \tag{5}$$

Essentially, upon observing any state trajectory in $\mathcal{L}^n_{\mathcal{I}_\varepsilon}$ (resp., $\mathcal{L}^n_{\mathcal{C}_\varepsilon}$), the intruder with imperfect measurement precision (captured by $\varepsilon$) is never certain that the initial (resp., current) state of the system is within the set of secret states. In this work, we are interested in establishing *a lower bound on the probability of the state trajectories*
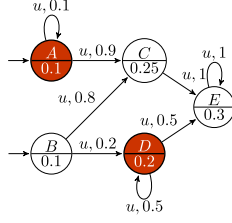
Fig. 1.    States marked by red denote secret states, states marked by sourceless arrows denote initial states, and the output map is specified by the value associated to each state. For each transition, the input and state transition probabilities are also specified on the arrows.

*of a gMDP being initial-state (or current-state) secure within a finite time horizon $n \in \mathbb{N}$.*

The formal definition of the new notions of approximate opacity for gMDPs is provided as follows.

*Definition 3.1 (*Approximate Opacity for gMDPs*):* Consider a gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and constants $\varepsilon \in \mathbb{R}_{\geq 0}$, $0 \leq \lambda \leq 1$, and $n \in \mathbb{N}$. System $\Sigma$ is said to be

1) *( $\varepsilon, \lambda$)-approximate initial-state opaque* over time horizon $n$ if for any $x_0 \in X_0$ and any $\nu \in \mathcal{U}^n$, $\mathbb{P}_{x_0}^\nu(\xi_{x_0\nu} \in L_{\mathcal{I}_\varepsilon}^n) \geq \lambda$;

2) *( $\varepsilon, \lambda$)-approximate current-state opaque* over time horizon $n$ if for any $x_0 \in X_0$ and any $\nu \in \mathcal{U}^n$, $\mathbb{P}_{x_0}^\nu(\xi_{x_0\nu} \in L_{\mathcal{C}_\varepsilon}^n) \geq \lambda$;

where $\mathbb{P}_{x_0}^\nu(\xi_{x_0\nu} \in L_{\mathcal{I}_\varepsilon}^n)$ (resp., $\mathbb{P}_{x_0}^\nu(\xi_{x_0\nu} \in L_{\mathcal{C}_\varepsilon}^n)$) denotes the probability that the state evolution of the system $\Sigma$ is initial-state (resp., current-state) secure over a finite time horizon $n$ starting from $x_0$ under input sequence $\nu$.

Essentially, Definition 3.1 requires that, for any initial state and any input sequence, the system is probable to produce an initial-state/current-state secure trajectory with a probability that is at least $\lambda$. The parameter $\varepsilon$ reflects the precision of measurements by potential external intruders, within which we can ensure the opacity of a gMDP. Without loss of generality, throughout the work, we assume

$$\forall x_0 \in X_0 : \{x \in X_0 : \|h(x_0) - h(x)\| \leq \varepsilon\} \nsubseteq X_S \quad (6)$$

for any gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$; otherwise $(\varepsilon, \lambda)$-approximate initial state (resp., current-state) opacity is trivially violated.

We illustrate the proposed notions of approximate opacity for gMDPs using the following example.

*Example 3.2:* Consider a finite MDP $\Sigma$ as shown in Fig. 1, where $X = \{A, B, C, D, E\}$, $X_0 = \{A, B\}$, $X_S = \{A, D\}$, $U = \{u\}$, $Y = \{0.1, 0.2, 0.25, 0.3\}$, and the output map is specified by the value associated with each state. The state transition probability equipped with each transition is also depicted on the arrows. First, we show that the system is $(\varepsilon, \lambda)$-approximate initial-state opaque within finite time horizon $n \in \mathbb{N}$ with $\varepsilon = 0, \lambda = 0.9$. By the definition of secure state trajectories in (4), we have for the secret initial state $A$, a finite state run in the form of $ACEE^*$ is initial-state secure, since there exists another state run starting from nonsecret initial state $B$, i.e., $BCEE^*$ that generates exactly the same output run; however, a finite state run in the form of $AAA^*CEE^*$ will generate an output run $(0.1, 0.1, 0.1^*, 0.25, 0.3, 0.3^*)$, for which there does not exist a finite run starting from a nonsecret state that generates the same output run. Thus, we have for $x_0 = A$,

$\mathbb{P}(\xi_{A\nu} \in L_{\mathcal{I}_0}^n) = 0.9 \geq \lambda$ [1] for any $n \in \mathbb{N}$ and $\nu = u^*$. Notice that for the nonsecret initial state $x_0 = B$, clearly we have $\mathbb{P}(\xi_{B\nu} \in L_{\mathcal{I}_0}^n) = 1 \geq \lambda$ for any $n \in \mathbb{N}$ and $\nu = u^*$. Therefore, the system is $(0, 0.9)$-approximate initial-state opaque within any finite time horizon $n \in \mathbb{N}$. Similarly, one can check that the system is $(0, 0.8)$-approximate current-state opaque within any finite time horizon $n \in \mathbb{N}$. Note that this example is for illustrative purposes. We will provide a formal procedure for verifying approximate opacity for gMDPs in the next sections.

*Remark 3.3:* Let us remark that a notion of approximate opacity was proposed in [14] for continuous-space control systems. The proposed notions in Definition 3.1 can be seen as the probabilistic counterparts of the notions in [14] to evaluate opacity for gMDPs quantitatively. Clearly, when analyzing nonstochastic control systems, Definition 3.1 boils down to [14, Definition III.1] with $\lambda = 1$. Note also that several probabilistic opacity notions were proposed in [29] to provide a quantifiable measure of opacity. However, these notions are tailored to stochastic DESs, which can be modeled as probabilistic finite automata, and thus, not suitable for general stochastic control systems with continuous state spaces as considered in this article.

*Remark 3.4:* In simple terms, *observability* [37] in control systems indicates that the system's states can be inferred from its external outputs, which is related to the concept of a system being nonopaque. The relationship between observability and opacity has been extensively explored in DES literature. Specifically, Lin [7] demonstrated that observability in DES can be redefined as language-based opacity by appropriately defining the languages and observation mappings. Nevertheless, understanding the link between observability and opacity in general nonlinear (stochastic) control systems remains a complex task. This is due to the abundant notions of observability in the literature, such as [37] and [38], and is thus a subject for future research.

*Remark 3.5:* Note that the proposed notion of approximate opacity requires that for any initial state and any possible input sequence $\nu \in \mathcal{U}^n$, the probability that the resulting state trajectory is initial-state (or current-state) opaque must exceed a given threshold. Given this, the results presented in this article are not limited by any specific control policy imposed on the system. Therefore, the proposed results can be directly applied to stochastic control systems under any predesigned control policy by restricting the set of possible control sequences, such as the discrete feedback or sampled-data feedback controllers developed in existing literature [39], [40], [41]. It would also be an interesting future research direction to explore the design of discrete or sampled-data feedback controllers to enforce opacity for gMDPs using approaches similar to those in these references.

## IV. VERIFICATION OF APPROXIMATE OPACITY FOR FINITE GMDPS

In this section, we show how to verify approximate opacity for finite gMDPs. The results here will provide a foundation for the verification of approximate opacity for infinite gMDPs through certain system relations as shown in Section V.

### A. Verification of Approximate Initial-State Opacity

In order to verify approximate initial-state opacity for finite gMDPs, we will first construct a new system that serves as an

---

[1]In the remainder, we omit the super- and subscripts of $\mathbb{P}_{x_0}^\nu$ when they are clear from the context.

approximate initial-state estimator. For the sake of brevity, for any $x \in X$, we define

$$B_\varepsilon(x) = \{x' \in X : \|h(x) - h(x')\| \le \varepsilon\}$$

as the set of states that are $\varepsilon$-close to $x$ in terms of output observations. The formal definition of an approximate initial-state estimator is provided as follows.

*Definition 4.1 (*Approximate initial-state estimator*):* Consider a finite gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and $\varepsilon \in \mathbb{R}_{\ge 0}$. The $\varepsilon$-approximate initial-state estimator is a finite gMDP (without outputs) $\Sigma_{I,\mathrm{obs}} = (X_I, X_{I0}, U, T_I)$, where

1) $X_I = \{(x, q) \in X \times 2^{X \times X} : (\bar{x}_1, \bar{x}_2) \in q \Rightarrow \bar{x}_2 \in B_\varepsilon(x)\}$ is the set of states;
2) $X_{I0} = \{(x, q) \in X_0 \times 2^{X_0 \times X_0} : (\bar{x}_1, \bar{x}_2) \in q \Leftrightarrow [\bar{x}_1 = \bar{x}_2 \wedge \bar{x}_2 \in B_\varepsilon(x)]\}$ is the set of initial states;
3) $U$ is the same set of inputs as in $\Sigma$;
4) $T_I : X_I \times U \times X_I \to [0, 1]$ is the transition probability function defined by: for any $x_I = (x, q) \in X_I$, any $x'_I = (x', q') \in X_I$, and any $u \in U$,
   a) $T_I(x'_I | x_I, u) \ne 0$ if
      i) $T(x' | x, u) > 0$;
      ii) $q' = \bigcup_{u' \in U} \{(\bar{x}_1, \bar{x}'_2) \in X \times X : (\bar{x}_1, \bar{x}_2) \in q \wedge T(\bar{x}'_2 | \bar{x}_2, u') > 0 \wedge \bar{x}'_2 \in B_\varepsilon(x')\}$;
   b) when $T_I(x'_I | x_I, u) \ne 0$, we have $T_I(x'_I | x_I, u) = T(x' | x, u)$.

In the aforementioned definition, for each $(x, q) \in X_I$, $x$ is the real system state, and $q$ indicates the initial-state estimate. Specifically, $q \in 2^{X \times X}$ is a set of state pairs, i.e., each state in $q$ is in the form of $(\bar{x}_1, \bar{x}_2)$, where $\bar{x}_1$ denotes the state where the trajectory comes from initially and $\bar{x}_2$ denotes a possible current state that is $\varepsilon$-close to its real current state. Thus, the first component of $q$ is the initial-state estimate. Note that from the construction of $\Sigma_{I,\mathrm{obs}}$, for any state trajectory $(x_0, \ldots, x_n)$ of $\Sigma$, there exists a unique initial-state estimation sequence $((x_0, q_0) \ldots (x_n, q_n))$ in its approximate initial-state estimator $\Sigma_{I,\mathrm{obs}}$. In addition, we denote by $\xi_{x_0}^I$ a state trajectory of $\Sigma_{I,\mathrm{obs}}$ generated from initial state $(x_0, q_0) \in X_{I0}$ under some input sequence.

Given any state trajectory $(x_0, \ldots, x_n)$ of $\Sigma$ and its output sequence $(h(x_0), \ldots, h(x_n))$, let us define the initial-state estimate with $\varepsilon$ precision by observing $(h(x_0), \ldots, h(x_n))$ as

$$E_I^\varepsilon(h(x_0), \ldots, h(x_n)) :=$$
$$\{\bar{x}_0 \in X_0 : \exists (\bar{x}_0, \ldots, \bar{x}_n) \in \mathcal{L}^n \text{ s.t. } \max_{i \in [0;n]} \|h(\bar{x}_i) - h(x_i)\| \le \varepsilon\}. \quad (7)$$

Consider the corresponding initial-state estimation sequence $((x_0, q_0) \ldots (x_n, q_n))$ in $\Sigma_{I,\mathrm{obs}}$. Since the first component of $q_n$ in $\Sigma_{I,\mathrm{obs}}$ is the initial-state estimate, we have $E_I^\varepsilon(h(x_0), \ldots, h(x_n)) = \{\bar{x}_1 \in X : (\bar{x}_1, \bar{x}_2) \in q_n\}$.

Accordingly, we can define the set of "bad" (secret revealing) states in $\Sigma_{I,\mathrm{obs}}$ as

$$X_I^{\mathrm{bad}} := \{(x, q) \in X_I : \{\bar{x}_1 \in X : (\bar{x}_1, \bar{x}_2) \in q\} \subseteq X_S\}. \quad (8)$$

Here, a "bad" state is a state in the initial-state estimator whose associated initial-state estimate lies entirely within the set of secret states.

Note that by fixing $\nu(t) \in U$ in $\Sigma_{I,\mathrm{obs}}$ at any time step $t \ge 0$, we obtain a transition probability matrix denoted by $\mathbf{P}_I^{\nu(t)} : X_I \times X_I \to [0, 1]$. Then, the next two lemmas are needed to prove the main result in this subsection.

*Lemma 4.2:* Consider a finite gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and its approximate initial-state estimator $\Sigma_{I,\mathrm{obs}}$ as in Definition 4.1. For any initial state $x_0$ and any input sequence $\nu = (\nu(0), \ldots, \nu(n-1))$, we have

$$1 - \mathbb{P}(\xi_{x_0\nu} \in L_{\mathcal{I}_\varepsilon}^n) = \mathbf{1}_{\mathrm{bad}}^I \mathbf{P}_I^{\nu(n-1)} \cdots \mathbf{P}_I^{\nu(1)} \mathbf{P}_I^{\nu(0)} \mathbf{1}_{x_0}^I$$

where $L_{\mathcal{I}_\varepsilon}^n$ is the set of initial-state secure state trajectories as in (4) within time horizon $n \in \mathbb{N}$, and $\mathbf{1}_{\mathrm{bad}}^I$ and $\mathbf{1}_{x_0}^I$ are the indicator vectors[2] for states in $X_I^{\mathrm{bad}}$ as defined in (8) and for initial states in $X_{I0}$ in $\Sigma_{I,\mathrm{obs}}$, respectively.

The proof of Lemma 4.2 is provided in the Appendix.

*Lemma 4.3:* Consider a finite gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and its approximate initial-state estimator $\Sigma_{I,\mathrm{obs}}$ as in Definition 4.1. If for any initial state $x_0$ and any input sequence $\nu = (\nu(0), \ldots, \nu(n-1))$, the following holds:

$$\mathbf{1}_{\mathrm{bad}}^I \mathbf{P}_I^{\nu(n-1)} \cdots \mathbf{P}_I^{\nu(1)} \mathbf{P}_I^{\nu(0)} \mathbf{1}_{x_0}^I \le 1 - \lambda \quad (9)$$

where $\mathbf{1}_{\mathrm{bad}}^I$ and $\mathbf{1}_{x_0}^I$ are the indicator vectors for the states in $\Sigma_{I,\mathrm{obs}}$ as in Lemma 4.2, then $\Sigma$ is $(\varepsilon, \lambda)$-approximate initial-state opaque over time horizon $n \in \mathbb{N}$.

The proof of Lemma 4.3 is provided in the Appendix. Next, we show how to verify approximate initial-state opacity for finite gMDPs using value iteration techniques.

*Bounded-step reachability:* For a finite gMDP $\Sigma_{I,\mathrm{obs}}$ as in Definition 4.1, the problem of checking whether (9) holds can be solved by computing the *maximal probability of reaching the "bad" states in $X_I^{\mathrm{bad}}$, as defined in (8), within finite time horizon $n \in \mathbb{N}$ starting from any initial state $(x_0, q_0) \in X_{I0}$ under any input sequence $\nu$*. This *bounded-step reachability* property is denoted by $\Diamond^{\le n} X_I^{\mathrm{bad}}$, where $\Diamond$ denotes the "eventually" operator. Let us define the maximal probabilities of the aforementioned reachability problem as

$$\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\le n} X_I^{\mathrm{bad}}) \quad (10)$$

where $\xi_{x_0}^I \models \Diamond^{\le n} X_I^{\mathrm{bad}}$ represents that the state trajectory $\xi_{x_0}^I$ of $\Sigma_{I,\mathrm{obs}}$ satisfies the property of "eventually reaching set $X_I^{\mathrm{bad}}$ within $n \in \mathbb{N}$ time steps." Let us further define the vector $(p_s)_{s \in X_I}$, with $p_s := \mathbb{P}^{\max}(\xi_s^I \models \Diamond^{\le n} X_I^{\mathrm{bad}})$ being the maximal probability of reaching $X_I^{\mathrm{bad}}$ within $n \in \mathbb{N}$ time steps starting from state $s \in X_I$ under any input sequence. We further denote by $X_I^{\mathrm{nb}} \subseteq X_I \setminus X_I^{\mathrm{bad}}$ the set of states from which $X_I^{\mathrm{bad}}$ is not reachable in the underlying directed graph of the finite gMDP $\Sigma_{I,\mathrm{obs}}$.

We show in the next lemma that the previously discussed maximal reachability probability of the finite gMDP $\Sigma_{I,\mathrm{obs}}$ can be computed by a value iteration approach. This lemma is inspired by [20, Thm.10.100] and provided without proof.

*Lemma 4.4:* (Value iteration for bounded-step reachability probabilities of finite gMDPs) Consider a finite gMDP $\Sigma_{I,\mathrm{obs}} =$

---

[2] Specifically, $\mathbf{1}_{\mathrm{bad}}^I$ is a binary row vector indexed by states in $X_I$ such that each element whose index is in $X_I^{\mathrm{bad}}$ is 1, otherwise it is 0. Similarly, given any $x_0$, $\mathbf{1}_{x_0}^I$ is a binary row vector indexed by states in $X_I$ such that each element with index $(x_0, q_0) \in X_{I0}$ (with the first component being $x_0$) is 1, otherwise it is 0. Note that $\mathbf{1}_{x_0}^I$ are *standard basis vectors* for any $x_0 \in X_0$.
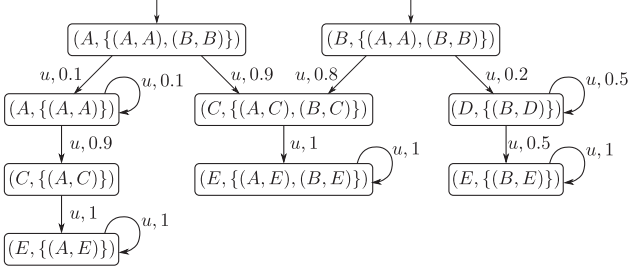
Fig. 2.   Example of approximate initial-state estimators.

$(X_I, X_{I0}, U, T_I)$ as in Definition 4.1 and set $X_I^{\text{bad}} \subseteq X_I$ as defined in (8). The vector $(p_s)_{s \in X_I}$ with $p_s = \mathbb{P}^{\max}(\xi_s^I \models \Diamond^{\leq n} X_I^{\text{bad}})$ can be computed by following the value iteration process:

1) $p_s = p_s^{(i)} = 1$, for any $i$ for all $s \in X_I^{\text{bad}}$;

2) $p_s = p_s^{(i)} = 0$, for any $i$ for all $s \in X_I^{\text{nb}}$;

3) $p_s = p_s^{(n)}$ for all $s \in X_I \setminus (X_I^{\text{bad}} \cup X_I^{\text{nb}})$,

with $p_s^{(0)} = 0, p_s^{(i+1)} = \max\{ \sum_{s' \in X_I} T_I(s'|s, u) p_s'^{(i)} \}$.

By using the value iteration approach in the above lemma, one can get the maximal probabilities $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}})$ for all initial states $x_0 \in X_0$. We discuss the complexity of computing $\mathbb{P}^{\max}(\xi_s^I \models \Diamond^{\leq n} X_I^{\text{bad}})$ later in Remark 4.9.

In the next theorem, we present the main result in this subsection on the verification of approximate initial-state opacity of finite gMDPs.

*Theorem 4.5:* Consider a finite gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and its approximate initial-state estimator $\Sigma_{I,\text{obs}}$ as in Definition 4.1. Then, $\Sigma$ is $(\varepsilon, \lambda)$-approximate initial-state opaque over time horizon $n \in \mathbb{N}$ if and only if for any $x_0 \in X_0$:

$$\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}}) \leq 1 - \lambda$$

with $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}})$ computed as in Lemma 4.4.

*Proof:* First recall from (10) that $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}})$ denotes the maximal probability of a trajectory $\xi_{x_0}^I$ in $\Sigma_{I,\text{obs}}$ eventually reaching $X_I^{\text{bad}}$ within time horizon $n \in \mathbb{N}$ under any input sequence. Thus, we have for any $x_0 \in X_0$

$$\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}}) = \max_\nu \mathbf{1}_{\text{bad}}^I \mathbf{P}_I^{\nu(n-1)} \cdots \mathbf{P}_I^{\nu(1)} \mathbf{P}_I^{\nu(0)} \mathbf{1}_{x_0}^I$$

where $\nu = (\nu(0), \ldots, \nu(n-1))$. Then, (9) holds if and only if $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}}) \leq 1 - \lambda$. The rest of the proof for this theorem follows readily by combining the results in Lemmas 4.2 and 4.3. ∎

*Example 4.6:* Let us consider the finite gMDP $\Sigma$ as shown in Fig. 1. In order to verify approximate opacity of $\Sigma$, we first build an $\varepsilon$-approximate initial-state estimator $\Sigma_{I,\text{obs}}$ following Definition 4.1, with $\varepsilon = 0$. The estimator is shown in Fig. 2, where $X_I = \{(A, \{(A, A), (B, B)\}), (B, \{(A, A), (B, B)\}), (A, \{(A, A)\}), (C, \{(A, C), (B, C)\}), (D, \{(B, D)\}), (C, \{(A, C)\}), (E, \{(A, E), (B, E)\}), (E, \{(B, E)\}), (E, \{(A, E)\})\}$, $U = \{u\}$, $X_{I0} = \{(A, \{(A, A), (B, B)\}), (B, \{(A, A), (B, B)\})\}$. By (8), we get $X_I^{\text{bad}} = \{(A, \{(A, A)\}), (C, \{(A, C)\}), (E, \{(A, E)\})\}$, from which one can readily have the indicator vectors $\mathbf{1}_{\text{bad}}^I = [0, 0, 1, 0, 0, 1, 0, 0, 1]$, $\mathbf{1}_A^I = [1, 0, 0, 0, 0, 0, 0, 0, 0]$,

and $\mathbf{1}_B^I = [0, 1, 0, 0, 0, 0, 0, 0, 0]$. Now, considering $\lambda = 0.9$, we show that $\Sigma$ is (0,0.9)-approximate initial-state opaque by applying Theorem 4.5. Let us first compute $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}})$ by following the value iteration process as in Lemma 4.4. Note that based on the underlying directed graph of this finite gMDP, we have $X_I^{\text{nb}} = \{(B, \{(A, A), (B, B)\}), (C, \{(A, C), (B, C)\}), (D, \{(B, D)\}), (E, \{(A, E), (B, E)\}), (E, \{(B, E)\})\}$ as the set of states from which $X_I^{\text{bad}}$ is not reachable. The equations of the value iteration process are as follows:

1) for the states in $X_I^{\text{bad}}$, we get $p_{(A, \{(A,A)\})}^{(i)} = p_{(C, \{(A,C)\})}^{(i)} = p_{(E, \{(A,E)\})}^{(i)} = 1$ for any $i$;

2) for the states in $X_I^{\text{nb}}$, we get $p_{(B, \{(A,A),(B,B)\})}^{(i)} = p_{(C, \{(A,C),(B,C)\})}^{(i)} = p_{(D, \{(B,D)\})}^{(i)} = p_{(E, \{(A,E),(B,E)\})}^{(i)} = p_{(E, \{(B,E)\})}^{(i)} = 0$ for any $i$;

3) for the other state $(A, \{(A, A), (B, B)\})$, which is neither in $X_I^{\text{bad}}$ nor in $X_I^{\text{nb}}$, we have $p_{(A, \{(A,A),(B,B)\})}^{(0)} = 0$, $p_{(A, \{(A,A),(B,B)\})}^{(i+1)} = \max\{0.1 p_{(A, \{(A,A),(B,B)\})}^{(i)}, 0.9 p_{(C, \{(A,C),(B,C)\})}^{(i)}\} = 0.1$.

Thus, the solution of the value iteration process in Lemma 4.4 yields the vector $(p_s)_{s \in X_I} = \{0.1, 1, 1, 1, 0, 0, 0, 0, 0\}$ with $n \geq 1$. Therefore, we obtain that $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}}) = 0.1$ for $x_0 = A$ and $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}}) = 0$ for $x_0 = B$, which shows that $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}}) \leq 1 - \lambda$ for any initial state $x_0$. By Theorem 4.5, we conclude that $\Sigma$ is (0,0.9)-approximate initial-state opaque within any time horizon $n \in \mathbb{N}$ with $n \geq 1$.

### B. Verification of Approximate Current-State Opacity

In this subsection, we show how to verify approximate current-state opacity for finite gMDPs. To achieve this, we first construct a new system that serves as an approximate current-state estimator as follows.

*Definition 4.7 (Approximate current-state estimator):* Consider a finite gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and $\varepsilon \in \mathbb{R}_{\geq 0}$. The $\varepsilon$-approximate current-state estimator is a finite gMDP (without outputs) $\Sigma_{C,\text{obs}} = (X_C, X_{C0}, U, T_I)$, where

1)  $X_C = \{(x, q, l) \in X \times 2^X \times \{0, 1\} : \bar{x} \in q \Rightarrow \bar{x} \in B_\varepsilon(x)\}$ is the set of states;

2)  $X_{C0} = \{(x, q, l) \in X_0 \times 2^{X_0} \times \{0, 1\} : [\bar{x} \in q \Leftrightarrow \bar{x} \in B_\varepsilon(x)] \wedge [l = 0]\}$ is the set of initial states;

3)  $U$ is the same set of inputs as in $\Sigma$;

4)  $T_C : X_C \times U \times X_C \to [0, 1]$ is the transition probability function defined by: for any $x_C = (x, q, l) \in X_C$, any $x_C' = (x', q', l') \in X_C$, and any $u \in U$,

a)  $T_C(x_C'|x_C, u) \neq 0$ if

i)  $T(x'|x, u) > 0$;

ii)  $q' = \bigcup_{u' \in U} \{\bar{x}' \in X : \bar{x} \in q \wedge T(\bar{x}'|\bar{x}, u') > 0 \wedge \bar{x}' \in B_\varepsilon(\bar{x}')\}$;

iii)  $l' = 0 \Leftrightarrow [l = 0] \wedge [\{\bar{x}' \in X : \bar{x}' \in q'\} \not\subseteq X_S]$; when $T_C(x_C'|x_C, u) \neq 0$, we have $T_C(x_C'|x_C, u) = T(x'|x, u)$.

In the aforementioned definition, for each $(x, q, l) \in X_C$, $x$ is the real system state, $q$ is the current-state estimate, and $l$ is a binary state taking values either 0 or 1. Note that $q \in 2^X$ is a set of states, with $\bar{x} \in q$ denoting its current state estimate.

According to Definition 4.7, for a state $(x', q', l') \in X_C$, $l' = 0$ if and only if:

1) $\{\bar{x}' \in X : \bar{x}' \in q'\} \nsubseteq X_S$, i.e., the current-state estimate does not lie entirely within the set of secret states;

2) $l = 0$ at the previous time step.

Otherwise, $l' = 1$.

Essentially, for any trajectory that is currently at state $(x, q, l)$, $l$ tracks whether or not the trajectory has revealed the current-state secret before or upon visiting the state. In particular, $l = 0$ indicates that none of the prefixes of the trajectory reveals the current-state secret; $l = 1$ indicates that the trajectory reveals the current-state secret at the current step or previous time steps.

For any state trajectory $(x_0, \ldots, x_n)$ of $\Sigma$, there exists a unique current-state estimation sequence $((x_0, q_0, l_0) \ldots (x_n, q_n, l_n))$ in its approximate current-state estimator $\Sigma_{C,\text{obs}}$. In addition, we denote by $\xi_{x_0}^C$ a state trajectory of $\Sigma_{C,\text{obs}}$ generated from initial state $(x_0, q_0, l_0) \in X_{C0}$ under some input sequence. Accordingly, we can define the set of "bad" (secret revealing) states as

$$X_C^{\text{bad}} = \{(x, q, l) \in X_C : l = 1\}. \tag{11}$$

By fixing an input $\nu(t) \in U$ in $\Sigma_{C,\text{obs}}$ at each time step $t \geq 0$, we obtain a transition probability matrix denoted by $\mathbf{P}_C^{\nu(t)} : X_C \times X_C \to [0, 1]$. Similar as stated in Lemmas 4.2 and 4.3, given any finite gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and its approximate current-state estimator $\Sigma_{C,\text{obs}}$ as in Definition 4.7, for any initial state $x_0$ and any input sequence $\nu = (\nu(0), \ldots, \nu(n-1))$, we have

$$1 - \mathbb{P}(\xi_{x_0\nu} \in L_{\mathcal{C}_\varepsilon}^n) = \mathbf{1}_{\text{bad}}^C \mathbf{P}_C^{\nu(n-1)} \cdots \mathbf{P}_C^{\nu(1)} \mathbf{P}_C^{\nu(0)} \mathbf{1}_{x_0}^C$$

where $L_{\mathcal{C}_\varepsilon}^n$ is the set of current-state secure state trajectories as in (5) within time horizon $n \in \mathbb{N}$, and $\mathbf{1}_{\text{bad}}^C$ and $\mathbf{1}_{x_0}^C$ are the indicator vectors for states in $X_C^{\text{bad}}$ as defined in (11) and for initial states in $X_{C0}$ in $\Sigma_{C,\text{obs}}$, respectively. Moreover, if any initial state $x_0$ and any input sequence $\nu = (\nu(0), \ldots, \nu(n-1))$, the following holds:

$$\mathbf{1}_{\text{bad}}^C \mathbf{P}_C^{\nu(n-1)} \cdots \mathbf{P}_C^{\nu(1)} \mathbf{P}_C^{\nu(0)} \mathbf{1}_{x_0}^C \leq 1 - \lambda \tag{12}$$

then $\Sigma$ is $(\varepsilon, \lambda)$-approximate current-state opaque over time horizon $n \in \mathbb{N}$. The verification of approximate current-state opacity for finite gMDPs follows similarly as in Section IV-A using value iteration techniques. Let us denote by $\mathbb{P}^{\max}(\xi_{x_0}^C \models \Diamond^{\leq n} X_C^{\text{bad}})$ the maximal probabilities of a trajectory $\xi_{x_0}^C$ in $\Sigma_{C,\text{obs}}$ reaching $X_C^{\text{bad}}$ within bounded time horizon $n \in \mathbb{N}$ under any input sequences. Note that by following the value iteration approach similar to Lemma 4.4, one can compute $\mathbb{P}^{\max}(\xi_s^C \models \Diamond^{\leq n} X_C^{\text{bad}})$ as well.

The next theorem summarizes the main result of the verification of approximate current-state opacity of finite gMDPs.

*Theorem 4.8:* Consider a finite gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and its approximate current-state estimator $\Sigma_{C,\text{obs}}$ as in Definition 4.7. Then, $\Sigma$ is $(\varepsilon, \lambda)$-approximate current-state opaque over time horizon $n \in \mathbb{N}$ if and only if for any $x_0 \in X_0$

$$\mathbb{P}^{\max}(\xi_{x_0}^C \models \Diamond^{\leq n} X_C^{\text{bad}}) \leq 1 - \lambda.$$

*Proof:* This theorem can be proved by the same reasoning for Theorem 4.5. ∎

*Remark 4.9:* The precise values of $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}})$ and $\mathbb{P}^{\max}(\xi_{x_0}^C \models \Diamond^{\leq n} X_C^{\text{bad}})$ can be computed by following the value iteration processes in Lemma 4.4, using existing computational tools such as `MDPtoolbox` [42]. Note that since the estimators $\Sigma_{I,\text{obs}}$ and $\Sigma_{C,\text{obs}}$ are *finite* gMDPs, the values of $\mathbb{P}^{\max}(\xi_{x_0}^I \models \Diamond^{\leq n} X_I^{\text{bad}})$ and $\mathbb{P}^{\max}(\xi_{x_0}^C \models \Diamond^{\leq n} X_C^{\text{bad}})$ can be computed in *polynomial time* in the size of the estimators. As a consequence, the question of whether a finite gMDP is approximate initial-state/current-state opaque for some $\lambda \in [0, 1)$ is *decidable* in polynomial time.

## V. APPROXIMATE OPACITY-PRESERVING STOCHASTIC SIMULATION RELATIONS

In the previous sections, we introduced notions of approximate opacity for gMDPs, and presented opacity verification approaches for finite gMDPs. However, the verification of approximate opacity directly on gMDPs with infinite state sets is infeasible. Therefore, in the next sections, we propose an abstraction-based approach for the opacity verification problem for infinite gMDPs by leveraging new notions of system relations between a gMDP and its finite abstraction.

To this end, in this section, we introduce new notions of approximate opacity-preserving stochastic simulation relations based on lifting for gMDPs, and then, show their usefulness in the verification of approximate opacity for gMDPs. Let us first recall the notion of $\delta$-lifted relation proposed in [30].

*Definition 5.1 ( $\delta$-lifted relation [30]):* Let $X$ and $\hat{X}$ be two sets associated with the corresponding measurable spaces $(X, \mathcal{B}(X))$ and $(\hat{X}, \mathcal{B}(\hat{X}))$. Consider a relation $R \subseteq X \times \hat{X}$ with $R \in \mathcal{B}(X \times \hat{X})$. We denote by $\bar{R}_\delta \subseteq \mathcal{P}(X, \mathcal{B}(X)) \times \mathcal{P}(\hat{X}, \mathcal{B}(\hat{X}))$ the corresponding $\delta$-*lifted relation* if there exists a probability measure $L$, referred to as a *lifting*, for the probability space $(X \times \hat{X}, \mathcal{B}(X \times \hat{X}), L)$ such that $(\Phi, \Theta) \in \bar{R}_\delta$, denoted by $\Phi \bar{R}_\delta \Theta$, if and only if

1) $\forall A \in \mathcal{B}(X), L(A \times \hat{X}) = \Phi(A)$;
2) $\forall \hat{A} \in \mathcal{B}(\hat{X}), L(X \times \hat{A}) = \Theta(\hat{A})$;
3) for the probability space $(X \times \hat{X}, \mathcal{B}(X \times \hat{X}), L)$, it holds that $x R \hat{x}$ with probability at least $1 - \delta$, equivalently, $L(R) \in [1 - \delta, 1]$.

Given a relation $R \subseteq X \times \hat{X}$, the aforementioned definition provides conditions under which the relation $R$ can be lifted to a relation $\bar{R}_\delta$ that relates probability measures over state sets $X$ and $\hat{X}$. The third condition in Definition 5.1 requires that the probability measure $L$ assigns a probability of at least $1 - \delta$ to the set of state pairs in the relation $R$.

In the next subsections, we introduce the new notions of *approximate opacity-preserving stochastic simulation relations* based on $\delta$-lifted relation as in the aforementioned definition. The stochastic simulation relations can be used to capture the similarities between a gMDP and its finite abstraction.

### A. Initial-State Opacity-Preserving Stochastic Simulation Relation

First, we introduce a new notion of stochastic simulation relations for preserving approximate initial-state opacity.

*Definition 5.2 (Approximate initial-state opacity-preserving stochastic simulation relation):* Consider two gMDPs $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{T}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$. System $\hat{\Sigma}$ is $(\epsilon, \delta)$-*initial-state opacity-preservingly ( $(\epsilon, \delta)$-InitSOP) simulated by $\Sigma$, i.e., $\hat{\Sigma} \preceq_{\mathcal{I}_\delta^\epsilon} \Sigma$, if there exists a relation $R_x \subseteq X \times \hat{X}$ and Borel measurable stochastic kernels $L_T(\cdot|x, \hat{x}, \hat{u})$ and $L_{\hat{T}}(\cdot|x, \hat{x}, u)$ on $X \times \hat{X}$ such that

1) a) $\forall x_0 \in X_0 \cap X_S$, $\exists \hat{x}_0 \in \hat{X}_0 \cap \hat{X}_S$, s.t. $(x_0, \hat{x}_0) \in R_x$;

    b) $\forall \hat{x}_0 \in \hat{X}_0 \setminus \hat{X}_S$, $\exists x_0 \in X_0 \setminus X_S$, s.t. $(x_0, \hat{x}_0) \in R_x$;

a) $\forall (x, \hat{x}) \in R_x$, $\|h(x) - \hat{h}(\hat{x})\| \leq \epsilon$;

b) $\forall (x, \hat{x}) \in R_x$, we have

    a) $\forall \hat{u} \in \hat{U}$, $\exists u \in U$, s.t. $T(\cdot|x, u)$ $\bar{R}_\delta$ $\hat{T}(\cdot|\hat{x}, \hat{u})$ with lifting $L_T(\cdot|x, \hat{x}, \hat{u})$;

    b) $\forall u \in U$, $\exists \hat{u} \in \hat{U}$, s.t. $T(\cdot|x, u)$ $\bar{R}_\delta$ $\hat{T}(\cdot|\hat{x}, \hat{u})$ with lifting $L_{\hat{T}}(\cdot|x, \hat{x}, u)$.

The third condition of Definition 5.2 implies that for any $\hat{u} \in \hat{U}$, there exists $u \in U$, and vice versa, such that the state probability measures are in the lifted relation after one state transition for any $(x, \hat{x}) \in R_x$. In other words, the $\delta$-lifted relation $\bar{R}_\delta$ requires that the state pairs remain in the relation $R_x$ in the next time step with a probability of at least $1 - \delta$ if they are in the relation $R_x$ at the current time step. This new notion of stochastic simulation relation can be used to relate two gMDPs in terms of preserving approximate initial-state opacity. In an $(\epsilon, \delta)$-InitSOP stochastic simulation relation, the parameter $\delta$ quantifies the distance in probability between gMDPs, and $\epsilon$ is used to capture the closeness of output trajectories of two gMDPs. If $\hat{\Sigma} \preceq_{\mathcal{I}_\delta^\epsilon} \Sigma$ with $\hat{\Sigma}$ being a (finite) gMDP, $\hat{\Sigma}$ is called an InitSOP (finite) abstraction of system $\Sigma$.

Before showing the usefulness of this relation, let us denote for any measurable set $E$ over the output trajectories of a gMDP $\Sigma$ with $E \subseteq \mathcal{H}^n$, the $\epsilon$-*expansion* and $\epsilon$-*contraction* of set $E$, denoted by $E^\epsilon$ and $E^{-\epsilon}$, respectively. In particular, for any $\epsilon > 0$, $E^\epsilon$ and $E^{-\epsilon}$ are the $\epsilon$ neighborhoods of set $E$, where $E^\epsilon$ and $E^{-\epsilon}$ are, respectively, the smallest measurable set in $\mathcal{H}^n$ containing $E$, and the largest measurable set contained in $E$ satisfying

$$E^\epsilon = \{(y_0, \ldots, y_n) \in \mathcal{H}^n : \exists (\bar{y}_0, \ldots, \bar{y}_n) \in E$$
$$\max_{i \in [0;n]} \|\bar{y}_i - y_i\| \leq \epsilon\} \quad (13)$$

$$E^{-\epsilon} = \{(y_0, \ldots, y_n) \in \mathcal{H}^n : \forall (\bar{y}_0, \ldots, \bar{y}_n) \in \mathcal{H}^n \setminus E$$
$$\min_{i \in [0;n]} \|\bar{y}_i - y_i\| \geq \epsilon\}. \quad (14)$$

The following proposition provides us a probabilistic closeness guarantee between a gMDP and its InitSOP abstraction.

*Proposition 5.3:* Consider two gMDPs $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{T}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$. If $\hat{\Sigma} \preceq_{\mathcal{I}_\delta^\epsilon} \Sigma$, then for any $(x_0, \hat{x}_0) \in R_x$ and any input sequence $\hat{\nu}$ on $\hat{\Sigma}$, there exists an input sequence $\nu$ for $\Sigma$, and vice versa, such that for all measurable events $E \subseteq \mathcal{H}^n$ and the $\epsilon$-neighborhoods $E^\epsilon$ and $E^{-\epsilon}$ as defined in (13) and (14)

$$\mathbb{P}(\hat{\zeta}_{\hat{x}_0 \hat{\nu}} \in E^{-\epsilon}) - \gamma_\delta \leq \mathbb{P}(\zeta_{x_0 \nu} \in E) \leq \mathbb{P}(\hat{\zeta}_{\hat{x}_0 \hat{\nu}} \in E^\epsilon) + \gamma_\delta \quad (15)$$

where $|\zeta_{x_0 \nu}| = |\hat{\zeta}_{\hat{x}_0 \hat{\nu}}| = n$ with constant $\gamma_\delta := 1 - (1 - \delta)^n$.

The proof of Proposition 5.3 is provided in the Appendix. Before showing the main result of this section, let us define the sets of "secure output sequences" with respect to the secure state trajectories $\mathcal{L}_{\mathcal{I}_\epsilon}^n$ and $\mathcal{L}_{\mathcal{C}_\epsilon}^n$ as defined in (4) and (5), respectively,

$$\mathcal{H}_{\mathcal{I}_\epsilon}^n := \{(y_0, \ldots, y_n) \in \mathcal{H}^n : \exists (x_0, \ldots, x_n) \in \mathcal{L}_{\mathcal{I}_\epsilon}^n, \text{s.t.}$$
$$y_i = h(x_i) \, \forall i \in [0;n]\}; \quad (16)$$



Fig. 3. Example of $(\epsilon, \delta)$-InitSOP from $\hat{\Sigma}$ to $\Sigma$, i.e., $\hat{\Sigma} \preceq_{\mathcal{I}_\delta^\epsilon} \Sigma$.

$$\mathcal{H}_{\mathcal{C}_\epsilon}^n := \{(y_0, \ldots, y_n) \in \mathcal{H}^n : \exists (x_0, \ldots, x_n) \in \mathcal{L}_{\mathcal{C}_\epsilon}^n, \text{s.t.}$$
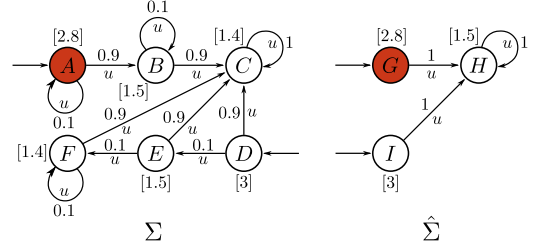$$y_i = h(x_i) \, \forall i \in [0;n]\} \quad (17)$$

where $\mathcal{H}_{\mathcal{I}_\epsilon}^n$ and $\mathcal{H}_{\mathcal{C}_\epsilon}^n$ denotes the set of secure output sequences w.r.t. approximate initial-state opacity and approximate current-state opacity, respectively, over time horizon $n \in \mathbb{N}$. The following lemma will be used later to prove our main results.

*Lemma 5.4:* Consider a gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$, the associated sets of secure state trajectories $L_{\mathcal{I}_\epsilon}$ and $L_{\mathcal{C}_\epsilon}$ in (4) and (5), and the sets of secure output sequences in (16) and (17). For any initial state $x_0$ and input sequence $\nu$, we have

$$\mathbb{P}(\xi_{x_0 \nu} \in L_{\mathcal{I}_\epsilon}^n) \geq \lambda \iff \mathbb{P}(\zeta_{x_0 \nu} \in \mathcal{H}_{\mathcal{I}_\epsilon}^n) \geq \lambda \quad (18)$$

$$\mathbb{P}(\xi_{x_0 \nu} \in L_{\mathcal{C}_\epsilon}^n) \geq \lambda \iff \mathbb{P}(\zeta_{x_0 \nu} \in \mathcal{H}_{\mathcal{C}_\epsilon}^n) \geq \lambda. \quad (19)$$

The proof of Lemma 5.4 is provided in the Appendix. The next theorem presents one of the main results of this article, which shows that the newly introduced stochastic simulation relation in Definition 5.2 can be used for the preservation of approximate initial-state opacity between two gMDPs.

*Theorem 5.5:* Consider two gMDPs $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{T}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$. Let $\epsilon, \delta \in \mathbb{R}_{\geq 0}$. If $\hat{\Sigma} \preceq_{\mathcal{I}_\delta^\epsilon} \Sigma$ and $\gamma_\delta = 1 - (1 - \delta)^n \leq \lambda$, then the following holds:

$\hat{\Sigma}$ is $(\varepsilon, \lambda)$-approximate initial-state opaque

$\Rightarrow \Sigma$ is $(\varepsilon + 2\epsilon, \lambda_\delta)$-approximate initial-state opaque $\quad (20)$

where the opacity satisfaction holds within time horizon $n \in \mathbb{N}$ and $\lambda_\delta = (1 - \gamma_\delta)(\lambda - \gamma_\delta)$.

The proof of Theorem 5.5 is provided in the Appendix. This result provides the foundation for verifying approximate initial-state opacity of gMDPs using abstraction-based techniques. As gMDPs are equipped with uncountable state spaces as well as metric output spaces, the verification/analysis of properties over gMDPs is generally infeasible. It is thus promising to approximate these models by simpler ones that are prone to be verified, such as finite gMDPs [21], [30]. In this work, by constructing a finite abstraction $\hat{\Sigma}$ of the concrete system $\Sigma$ via the InitSOP stochastic simulation relation between them, one can first efficiently verify opacity over the simpler system $\hat{\Sigma}$ by applying the verification approach in Section IV, and then, carry back the results to the concrete system $\Sigma$ by using Theorem 5.5.

The following example illustrates the usefulness of the new notion of InitSOP stochastic simulation relation as presented in Theorem 5.5.

*Example 5.6:* Consider two gMDPs $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{T}, \hat{Y}, \hat{h})$ as depicted in Fig. 3,

where $X = \{A, B, C, D, E, F\}$, $X_0 = \{A, D\}$, $X_S = \{A\}$; $\hat{X} = \{G, H, I\}$, $\hat{X}_0 = \{G, I\}$, $\hat{X}_S = \{G\}$, and $U = \hat{U} = \{u\}$. All secret states are marked in red, and the output of each state is specified by the value associated with it. Let us consider the following relation: $R_x = \{(A, G), (B, H), (C, H), (E, H), (F, H), (D, I)\}$. We claim that $R_x$ is an $(\epsilon, \delta)$-initial-state opacity-preserving stochastic simulation relation from $\hat{\Sigma}$ to $\Sigma$ with $\epsilon = 0.1$ and $\delta = 0.1$. This is shown by checking the conditions in Definition 5.2 as follows. First, for the secret initial state $A \in X_0 \cap X_S$, there exists $G \in \hat{X}_0 \cap \hat{X}_S$, s.t. $(A, G) \in R_x$. Similarly, for the nonsecret initial state $I \in \hat{X}_0 \setminus \hat{X}_S$, there exists $D \in X_0 \setminus X_S$ s.t. $(D, I) \in R_x$. Thus, condition 1 in Definition 5.2 holds. Also, one can check that $\|h(x) - \hat{h}(\hat{x})\| \le \epsilon$ holds, with $\epsilon = 0.1$ for all $(x, \hat{x}) \in R_x$, which implies that condition 2 in Definition 5.2 also holds. Moreover, one can check that conditions 3a and 3b hold for any $(x, \hat{x}) \in R_x$ as well. As an example, for $(A, G) \in R_x$ and $u \in U$, there exists $u \in \hat{U}$, s.t., with probability $1 - \delta = 0.9$, the succeeding states are $B$ and $H$ with $(B, H)$ being in relation $R_x$. One can conclude that $R_x$ is an $(\epsilon, \delta)$-initial-state opacity-preserving stochastic simulation relation from $\hat{\Sigma}$ to $\Sigma$ with $\epsilon = 0.1$ and $\delta = 0.1$, i.e., $\hat{\Sigma} \preceq_{\mathcal{I}_\delta^{0.1}} \Sigma$. By applying the verification approach in Section IV-A, we can easily check that $\hat{\Sigma}$ is $(0.2, 1)$-approximate initial-state opaque within any finite time horizon $n \in \mathbb{N}$. Therefore, according to Theorem 5.5, we conclude that $\Sigma$ is $(0.4, 0.81^n)$-approximate initial-state opaque within finite time horizon $n \in \mathbb{N}$, without needing to apply the verification approach to $\Sigma$ directly. ◇

## B. Current-State Opacity-Preserving Stochastic Simulation Relation

Here, we introduce a new notion of approximate current-state opacity-preserving stochastic simulation relations.

*Definition 5.7 (*Approximate current-state opacity-preserving stochastic simulation relation*):* Consider two gMDPs $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{T}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$. System $\hat{\Sigma}$ is $(\epsilon, \delta)$-*current-state opacity-preservingly (* $(\epsilon, \delta)$-*CurSOP) simulated by* $\Sigma$, i.e., $\hat{\Sigma} \preceq_{\mathcal{C}_\delta^\epsilon} \Sigma$, if there exists a relation $R_x \subseteq X \times \hat{X}$ and Borel measurable stochastic kernels $L_T(\cdot|x, \hat{x}, \hat{u})$ and $L_{\hat{T}}(\cdot|x, \hat{x}, u)$ on $X \times \hat{X}$ such that

1) a) $\forall x_0 \in X_0$, $\exists \hat{x}_0 \in \hat{X}_0$, s.t. $(x_0, \hat{x}_0) \in R_x$;
   b) $\forall \hat{x}_0 \in \hat{X}_0$, $\exists x_0 \in X_0$, s.t. $(x_0, \hat{x}_0) \in R_x$;
   a) $\forall (x, \hat{x}) \in R_x$, $\|h(x) - \hat{h}(\hat{x})\| \le \epsilon$;
   b) $\forall (x, \hat{x}) \in R_x$, we have[3]
      a) $\forall u \in U$, $\exists \hat{u} \in \hat{U}$, s.t. $T(\cdot|x, u) \ \bar{R}_\delta \ \hat{T}(\cdot|\hat{x}, \hat{u})$ with lifting $L_{\hat{T}}(\cdot|x, \hat{x}, u)$;
      b) $\forall u \in U$, s.t. $x^+ \in X_S$, $\exists \hat{u} \in \hat{U}$, s.t. $\hat{x}^+ \in \hat{X}_S$ and $T(\cdot|x, u) \ \bar{R}_\delta \ \hat{T}(\cdot|\hat{x}, \hat{u})$ with lifting $L_{\hat{T}}(\cdot|x, \hat{x}, u)$;
      c) $\forall \hat{u} \in \hat{U}$, $\exists u \in U$, s.t. $T(\cdot|x, u) \ \bar{R}_\delta \ \hat{T}(\cdot|\hat{x}, \hat{u})$ with lifting $L_T(\cdot|x, \hat{x}, \hat{u})$;
      d) $\forall \hat{u} \in \hat{U}$, s.t. $\hat{x}^+ \in \hat{X} \setminus \hat{X}_S$, $\exists u \in U$, s.t. $x^+ \in X \setminus X_S$ and $T(\cdot|x, u) \ \bar{R}_\delta \ \hat{T}(\cdot|\hat{x}, \hat{u})$ with lifting $L_T(\cdot|x, \hat{x}, \hat{u})$.

---

[3] We denote $x^+ := f(x, u, \varsigma)$ and $\hat{x}^+ := \hat{f}(\hat{x}, \hat{u}, \varsigma)$ for the sake of brevity.

This new notion of stochastic simulation relation can be used to relate two gMDPs in terms of preserving approximate current-state opacity. If $\hat{\Sigma} \preceq_{\mathcal{C}_\delta^\epsilon} \Sigma$ with $\hat{\Sigma}$ being a (finite) gMDP, $\hat{\Sigma}$ is called a CurSOP (finite) abstraction of $\Sigma$.

The next proposition provides us with a probabilistic closeness guarantee between a gMDP and its CurSOP abstraction.

*Proposition 5.8:* Consider two gMDPs $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{T}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$. If $\hat{\Sigma} \preceq_{\mathcal{C}_\delta^\epsilon} \Sigma$, then for any $(x_0, \hat{x}_0) \in R_x$ and any input sequence $\hat{\nu}$ on $\hat{\Sigma}$, there exists an input sequence $\nu$ for $\Sigma$, and vice versa, such that for all measurable events $E \subseteq \mathcal{H}^n$ and the $\epsilon$-neighborhoods $E^\epsilon$ and $E^{-\epsilon}$ as defined in (13)–(14), the inequality (15) holds with constant $\gamma_\delta = 1 - (1 - \delta)^n$.

The proof of this proposition follows similar reasoning as that of Proposition 5.3 and is omitted here. By employing the aforementioned proposition, we provide the following theorem that shows that the newly introduced stochastic simulation relation in Definition 5.7 can be used for the preservation of approximate current-state opacity between two gMDPs.

*Theorem 5.9:* Consider two gMDPs $\Sigma = (X, X_0, X_S, U, T, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{T}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$. Let $\epsilon, \delta \in \mathbb{R}_{\ge 0}$. If $\hat{\Sigma} \preceq_{\mathcal{C}_\delta^\epsilon} \Sigma$ and $\gamma_\delta = 1 - (1 - \delta)^n \le \lambda$, then the following holds:

$\hat{\Sigma}$ is $(\varepsilon, \lambda)$-approximate current-state opaque

$\Rightarrow \Sigma$ is $(\varepsilon + 2\epsilon, \lambda_\delta)$-approximate current-state opaque    (21)

where the opacity satisfaction holds within time horizon $n \in \mathbb{N}$ and $\lambda_\delta = (1 - \gamma_\delta)(\lambda - \gamma_\delta)$.

The proof of Theorem 5.9 is provided in the Appendix.

*Remark 5.10:* Let us remark that the parameters $\epsilon$ and $\delta$ that appeared in Theorems 5.5 and 5.9 are related to the opacity-preserving stochastic simulation relations between two gMDPs as in Definitions 5.2 and 5.7, respectively. They are used to specify two different types of precision: $\epsilon$ characterizes the "closeness" between the output trajectories of two gMDPs in terms of being approximate opaque, while $\delta$ is used to quantify the probability of the pair of state trajectories of two gMDPs remaining in the simulation relation after each time step.

## VI. FINITE ABSTRACTIONS FOR GMDPS

In the previous section, we introduced notions of approximate opacity-preserving stochastic simulation relations and discussed their usefulness in verifying approximate opacity for gMDPs using their abstractions. In this section, we will discuss how to establish the proposed simulation relations for gMDPs and their finite abstractions (finite gMDPs). To be specific, we provide conditions under which one can always construct finite abstractions under the desired $(\epsilon, \delta)$-InitSOP (resp., CurSOP) stochastic simulation relations for a (possibly infinite) gMDP.

In the following, we consider a gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ with $X = X_0$, and assume its output map $h$ satisfies the following general Lipschitz assumption:

$$\|h(x) - h(x')\| \le \ell(\|x - x'\|)    (22)$$

for all $x, x' \in X$, where $\ell \in \mathcal{K}_\infty$. In addition, the existence of opacity-preserving stochastic simulation relations between $\Sigma$ and its finite abstraction is established under the assumption that $\Sigma$ is *incrementally input-to-state stable* [43] as in the following definition.

*Definition 6.1:* A gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$ is called *incrementally input-to-state stable* ($\delta$-ISS) if there exists a function $V : X \times X \to \mathbb{R}_{\geq 0}$ such that $\forall x, x' \in X$, $\forall u, u' \in U$, the following conditions hold:

$$\underline{\alpha}(\|x - x'\|) \leq V(x, x') \leq \overline{\alpha}(\|x - x'\|) \quad (23)$$

$$\mathbb{E}\left[V(f(x, u, \varsigma), f(x', u', \varsigma))\big|x, x', u, u'\right] - V(x, x') \leq$$
$$- \bar{\kappa}(V(x, x')) + \bar{\rho}(\|u - u'\|) \quad (24)$$

for some $\underline{\alpha}, \overline{\alpha}, \bar{\kappa} \in \mathcal{K}_\infty$, and $\bar{\rho} \in \mathcal{K}_\infty \cup \{0\}$.

The aforementioned definition can be seen as a stochastic counterpart of the $\delta$-ISS Lyapunov functions defined in [44] for discrete-time nonstochastic systems.

In addition, we assume that there exists a function $\gamma \in \mathcal{K}_\infty$ such that $V$ satisfies

$$V(x, x') - V(x, x'') \leq \gamma(\|x' - x''\|) \, \forall x, x', x'' \in X. \quad (25)$$

This assumption is not restrictive in the sense that condition (25) can be satisfied for any differentiable function $V$ restricted to a compact subset of $X \times X$ by applying the mean value theorem [43].

## A. Finite gMDPs

In this subsection, we approximate a gMDP $\Sigma$ with a finite abstraction (finite gMDP) $\hat{\Sigma}$. The construction follows a similar procedure as in [45, Algorithm 1] with some modifications to incorporate the secret information of the system.

Consider a concrete gMDP $\Sigma = (X, X_0, X_S, U, T, Y, h)$, we assume that for the rest of this article, sets $X$, $X_S$, and $U$ are of the form of finite unions of boxes and $X_0 = X$. Consider a tuple $q = (\eta, \theta, \mu)$ of parameters, where $0 < \eta \leq \min\{\text{span}(X_S), \text{span}(X \setminus X_S)\}$ is the state set quantization, $0 < \mu < \text{span}(U)$ is the external input set quantization, and $\theta \in \mathbb{R}_{\geq 0}$ is a design parameter.

The finite abstraction of $\Sigma$ is a finite gMDP, represented by the tuple $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{T}, \hat{Y}, \hat{h})$, where $\hat{X} = \hat{X}_0 = [X]_\eta$, $\hat{X}_S = [X_S^\theta]_\eta$, $\hat{U} = [U]_\mu$, $\hat{Y} = \{h(\hat{x})|\hat{x} \in \hat{X}\}$, $\hat{h}(\hat{x}) = h(\hat{x})$ $\forall \hat{x} \in \hat{X}$. Intuitively, the state and input sets of $\hat{\Sigma}$ consist of finitely many representative points, each of which represents an aggregate of continuous states or inputs of the concrete system. We denote by $\Pi_x : X \to \hat{X}$ the abstraction map that assigns to any $x \in X$, the representative point $\hat{x} \in \hat{X}$ of $x$. Similarly, $\Pi_u : U \to \hat{U}$ is used to denote the abstraction maps for external inputs. Given the quantization parameters, the abstraction maps $\Pi_x$ and $\Pi_u$ satisfy the following inequalities:

$$\|\Pi_x(x) - x\| \leq \eta \, \forall x \in X; \|\Pi_u(u) - u\| \leq \mu \, \forall u \in U. \quad (26)$$

The dynamics $\hat{f} : \hat{X} \times \hat{U} \times V_\varsigma \to \hat{X}$ of the abstraction is defined as $\hat{f}(\hat{x}, \hat{u}, \varsigma) = \Pi_x(f(\hat{x}, \hat{u}, \varsigma))$. We use $\Xi : \hat{X} \to 2^X$ to denote the map that assigns any representative point $\hat{x} \in X$ to the corresponding partition set represented by $\hat{x}$, i.e. $\Xi(\hat{x}) = \{x \in X : \Pi_x(x) = \hat{x}\}$. Consequently, one can compute the transition probability matrix $\hat{T}$ for $\hat{\Sigma}$ as $\hat{T}(\hat{x}'|\hat{x}, \hat{u}) = T(\Xi(\hat{x}')|\hat{x}, \hat{u})$, for all $\hat{x}, \hat{x}' \in \hat{X}$, and all $\hat{u} \in \hat{U}$.

## B. Establishing Opacity-Preserving Stochastic Simulation Relations

In the following theorem, we show that for a $\delta$-ISS gMDP $\Sigma$, if the tuple $q = (\eta, \theta, \mu)$ of quantization parameters satisfies certain conditions, then the finite gMDP $\hat{\Sigma}$ constructed in Section VI-A is an $(\epsilon, \delta)$-InitSOP finite abstraction of $\Sigma$.

*Theorem 6.2:* Consider a $\delta$-ISS gMDP as in Definition 6.1, associated with function $V$ satisfying (23)–(25) with functions $\underline{\alpha}$, $\overline{\alpha}$, $\bar{\kappa}$ $\bar{\rho}$. For any desired parameters $\epsilon \in \mathbb{R}_{>0}$, $\delta \in (0, 1]$, let $\hat{\Sigma}$ be a finite abstraction of $\Sigma$ as constructed in Section VI-A, with the quantization parameters $q = (\eta, 0, \mu)$ satisfying

$$\eta \leq \min\{\gamma^{-1}(\underline{\alpha} \circ \ell^{-1}(\epsilon)\delta - (\mathcal{I}_d - \bar{\kappa}) \circ \underline{\alpha} \circ \ell^{-1}(\epsilon) - \bar{\rho}(\mu)),$$
$$\overline{\alpha}^{-1} \circ \underline{\alpha} \circ \ell^{-1}(\epsilon)\} \quad (27)$$

then, $\hat{\Sigma}$ is $(\epsilon, \delta)$-InitSOP simulated by $\Sigma$ via the relation

$$R_x = \{(x, \hat{x}) \in X \times \hat{X} | V(x, \hat{x}) \leq \underline{\alpha} \circ \ell^{-1}(\epsilon)\}. \quad (28)$$

The proof of Theorem 6.2 is provided in the Appendix. The following theorem provides another main result of this section, showing that for a $\delta$-ISS gMDP $\Sigma$, if the tuple $q = (\eta, \theta, \mu)$ of quantization parameters satisfies certain conditions, then the finite gMDP $\hat{\Sigma}$ constructed in Section VI-A is an $(\epsilon, \delta)$-CurSOP finite abstraction of $\Sigma$.

*Theorem 6.3:* Consider a $\delta$-ISS gMDP as in Definition 6.1, associated with function $V$ satisfying (23)–(25) with functions $\underline{\alpha}$, $\overline{\alpha}$, $\bar{\kappa}$, and $\bar{\rho}$. For any desired parameters $\epsilon \in \mathbb{R}_{>0}$, $\delta \in (0, 1]$, let $\hat{\Sigma}$ be a finite abstraction of $\Sigma$ as constructed in Section VI-A, with the quantization parameters $q = (\eta, \theta, \mu)$ satisfying

$$\eta \leq \min\{\gamma^{-1}(\underline{\alpha} \circ \ell^{-1}(\epsilon)\delta - (\mathcal{I}_d - \bar{\kappa}) \circ \underline{\alpha} \circ \ell^{-1}(\epsilon) - \bar{\rho}(\mu)),$$
$$\overline{\alpha}^{-1} \circ \underline{\alpha} \circ \ell^{-1}(\epsilon)\} \quad (29)$$
$$\ell^{-1}(\epsilon) \leq \theta \quad (30)$$

then, $\hat{\Sigma}$ is $(\epsilon, \delta)$-CurSOP simulated by $\Sigma$ via the relation

$$R_x = \{(x, \hat{x}) \in X \times \hat{X} | V(x, \hat{x}) \leq \underline{\alpha} \circ \ell^{-1}(\epsilon)\}. \quad (31)$$

The proof of Theorem 6.3 is provided in Appendix.

*Remark VI.4:* As seen from Theorems 6.2 and 6.3, for any desired parameters $\epsilon \in \mathbb{R}_{>0}$, $\delta \in (0, 1]$, one can always find suitable quantization parameters $\eta$ and $\mu$ to construct a finite abstraction $\hat{\Sigma}$ that opacity-preservingly simulates the original gMDP $\Sigma$. Recall from Propositions 5.3 and 5.8 that the smaller $\delta$ is, the more precise the abstraction $\hat{\Sigma}$ is in terms of generating similar output trajectories as the original gMDP $\Sigma$. However, note that according to conditions (27) and (29), a smaller $\delta$ will lead to smaller quantization parameters $\eta$ and $\mu$ for the construction of abstraction $\hat{\Sigma}$. This will further result in building a "fine" abstraction with a state set that is too large, and thus, too expensive to compute or store. Therefore, from a practical point of view, it is advisable to choose the value of $\delta$ as small as possible, given the available computing power.

*Remark 6.5:* Observe that the proposed opacity-preserving stochastic simulation relations for the initial state (or the current state) are one-sided because condition 1 in Definition 5.2 (and conditions 3b and 3d in Definition 5.7) lacks symmetry. On the other hand, condition 3 in Definition 5.2 (along with conditions 3a and 3c in Definition 5.7) exhibits symmetry (two-sided). Thus, in order to establish suitable opacity-preserving stochastic

Fig. 4. Road traffic model observed by an outside intruder.

simulation relations, the $\delta$-ISS assumption remains essential for the original gMDPs as detailed in [43].

## VII. CASE STUDY

In this section, we illustrate our results on a road traffic model adapted from [46], as shown in Fig. 4. Here, the traffic flow model is considered as one cell of a road traffic network. The cell is assumed to be equipped with at least one measurable entry and one exit. The traffic flow dynamics of the cell is given by

$$\Sigma : \begin{cases} \mathbf{x}(k+1) = & (1 - \frac{\tau v}{l} - e)\mathbf{x}(k) + b\nu(k) + d\omega(k) \\ \mathbf{y}(k) = & c\mathbf{x}(k) \end{cases}$$

where $\tau$ is the sampling time in hour, $l$ is the length of each cell in kilometers, and $v$ is the traffic flow speed in kilometers per hour. The state $\mathbf{x}(k)$ of $\Sigma$ represents the density of the traffic in the vehicle per cell at a specific time instant indexed by $k$. The scalar $b$ denotes the number of vehicles that are allowed to enter the cell during each sampling time controlled by the input signals $\nu(\cdot) \in \{0, 1\}$, where $\nu(\cdot) = 1$ (resp. $\nu(\cdot) = 0$) corresponds to green (resp., red) traffic light. The constant $e$ denotes the percentage of vehicles that leave the cell during each sampling time through exits. The system is considered to be affected by additive noise where $\omega(k)$ is a sequence of independent random variables with standard normal distribution that models environmental uncertainties. The values of the parameters in the system are: $\tau = \frac{30}{60 \times 60}$ h, $l = 1$ km, $v = 60$km/h, $b = 0.5$, $c = 0.1$, $d = 0.1$, and $e = 0.4$.

In this example, we consider that the initial density of the traffic model contains critical information that needs to be kept secret (e.g., a high traffic density might indicate activities happening with vehicles gathering in that road cell). It is implicitly assumed that there is a malicious intruder (with imprecise measurements captured by $\varepsilon \in \mathbb{R}_{\geq 0}$) who is remotely observing the output behavior of the road cell, intending to carry out an attack. Therefore, the system is interested in verifying its security level in terms of the probability guarantee of maintaining plausible deniability for secret initial conditions. This can be formulated as an ($\varepsilon, \lambda$)-approximate initial-state opacity verification problem as in Definition 3.1 with $\lambda$ characterizing the probability guarantee on the occurrence of secure executions of the system. Recall that a secure execution of the system is a state trajectory $(x_0, \dots, x_n)$, for which there exists a state trajectory generated from a nonsecret state such that their output trajectories are indistinguishable for an intruder with measurement precision $\varepsilon$. An example of such a pair of trajectories generated by the road traffic model is depicted in Fig. 5. Note that the verification of opacity directly on continuous-space control systems is generally infeasible.



Fig. 5. State trajectory (Traj.1) initiated from a secret initial condition and its ($\varepsilon$-close) output equivalent trajectory (Traj.2) started from a nonsecret initial condition. The shaded area in light blue represents the $\varepsilon$-band around Traj.1.



Fig. 6. Constructed finite abstraction $\hat{\Sigma}$ for the system $\Sigma$ with $\hat{\Sigma} \preceq_{\mathcal{I}_\delta^\epsilon} \Sigma$. The states marked by red denote secret states; the states marked by sourceless arrows denote initial states; the outputs are specified by the value associated with each state. For each transition, the input and state transition probabilities are also specified on the arrows.

Next, we show the approximate initial-state opacity of $\Sigma$ by applying the abstraction-based verification approach as proposed in this article. We are interested in the state set $X := [0, 3]$, initial set $X_0 := [2, 3]$, and secret set $X_S := [2.5, 3] \cup [0, 0.5]$. Considering precision parameters $\delta = 0.15$ and $\epsilon = 1$, let us first construct a finite gMDP as the approximate initial-state opacity-preserving abstraction of $\Sigma$. Consider function $V(x, x') = \|x - x'\|$. One can readily verify that function $V$ satisfies (23)–(25) with functions $\underline{\alpha}(s) = \overline{\alpha}(s) = s$, $\bar{\kappa}(s) = (\frac{\tau v}{l} - e)s = 0.9s$, $\bar{\rho}(s) = bs = 0.5s$, and $\ell(s) = 0.1s$, which implies that $\Sigma$ is $\delta$-ISS as in Definition 6.1. Next, by leveraging Theorem 6.2, let us choose $q = (0.5, 0, 0)$ as the quantization parameters that satisfies (27), and then, construct a finite gMDP following the steps in Section VI-A. The obtained finite gMDP $\hat{\Sigma}$ is depicted as in Fig. 6. Note that the states $q_1$ to $q_6$ are the representative points in $\hat{\Sigma}$ after discretizing the state set using quantization parameter $\eta = 0.5$. Once obtaining this opacity-preserving abstraction $\hat{\Sigma}$, we apply the verification approach as presented in Section IV, summarized as in the following steps:

1) constructing an $\varepsilon$-approximate initial-state estimator $\Sigma_{I,\text{obs}}$ of $\hat{\Sigma}$;
2) solving the value iteration process for the maximal probabilities of $\Sigma_{I,\text{obs}}$ reaching the bad region $X_I^{\text{bad}}$ [cf. (8)];
3) leveraging Theorem 4.5 to conclude the approximate initial-state opacity of $\hat{\Sigma}$.

Following these steps, we obtain that $\hat{\Sigma}$ is $(0.05, 1)$-approximate initial-state opaque over any finite time horizon $n \in \mathbb{N}$. Note that the verification of initial-state opacity for $\hat{\Sigma}$ follows exactly the same process as already explained in Example 4.6, and thus, we omit here the detailed process due to lack of space. Moreover, since $\hat{\Sigma}$ is $(\epsilon, \delta)$-InitSOP simulated by $\Sigma$ with $\epsilon = 1$ and $\delta = 0.15$, by further leveraging Theorem 5.5, we get that $\Sigma$ is $(2.05, 0.7225^n)$-approximate initial-state opaque over any time horizon $n \in \mathbb{N}$.

Let us remark that this example serves as an intuitive illustration of the proposed results. We intended to establish a coarse discretization on the state set by choosing a large quantization parameter $\eta$ for the sake of a visualization of the constructed finite gMDP as in Fig. 6. Instead, one can always resort to existing software tools, e.g., Faust [47], to generate and export a finer abstraction that better approximates the concrete gMDP using smaller quantization parameters. Once the exported abstract model (a finite gMDP) is obtained, one can verify opacity on the finite gMDP by employing the well-known probabilistic model checkers, such as PRISM [48], together with the verification approach presented in Section IV.

## VIII. CONCLUSION

In this article, we proposed new notions of approximate initial-state and current-state opacity for continuous-space stochastic control systems. This new notion provides a quantifiable measure of opacity, allowing us to quantitatively evaluate the probability of a stochastic system satisfying opacity during each execution. We presented novel methods to verify approximate opacity for stochastic systems modeled as finite gMDPs by using value iteration techniques on their approximate initial-state (or current-state) estimators. Using newly proposed notions of opacity-preserving simulation relations, an abstraction-based verification approach is then developed to show approximate opacity for infinite gMDPs through their finite abstractions (modeled as finite gMDPs). We also discussed how to construct finite abstractions that preserve approximate opacity for a class of stochastic control systems.

For future works, we plan to investigate the synthesis of control policies to enforce opacity for gMDPs. Moreover, we would like to extend the results to other notions of opacity including infinite-step opacity and preopacity. Note that the verification of infinite-step opacity or preopacity is much more challenging since we also need to consider how future information can affect our knowledge about the current status of the system. Thus, a new information structure is needed for constructing the estimator that involves the computation of delayed state estimates. The challenge in extending our results to verify opacity over infinite time horizons lies in the limitation of the abstraction-based verification technique for *continuous-space* gMDPs. Notice that from Propositions 5.3 and 5.8, the guarantee on the probabilistic distance between a gMDP and its abstraction becomes vacuous when $n$ goes to infinity. In fact, the problem of building abstractions for stochastic systems for infinite-horizon properties is known to be challenging [21] due to the conservative nature of Lyapunov-like techniques (simulation relations) and associated error bounds. As a potential direction to extend the results to an infinite time horizon, one may consider using similar ideas as proposed in [49], which builds bisimilar finite abstractions for infinite-horizon properties (safety and reachability) by assuming and exploiting some additional structural features, e.g., in terms of the existence of *absorbing sets* [49] of the original stochastic system.

*Proof of Lemma 4.2:* Consider an initial state $x_0$ in $\Sigma$ and an input sequence $\nu = (\nu(0), \ldots, \nu(n-1))$. For any state trajectory $\xi_{x_0\nu} = (x_0, \ldots, x_n)$ in $\Sigma$, by Definition 4.1 of the approximate initial-state estimator $\Sigma_{I,\text{obs}}$, there exists a unique initial-state estimation sequence $((x_0, q_0) \ldots (x_n, q_n))$ in $\Sigma_{I,\text{obs}}$ generated from $(x_0, q_0)$ under $\nu$. Recall that under $\nu = (\nu(0), \ldots, \nu(k-1))$, the transition probability function $T_I$ in $\Sigma_{I,\text{obs}}$ boils down to $\mathbf{P}_I^{\nu(i)} : X_I \times X_I \to [0, 1]$ at each time step $i \in [0; n-1]$. Thus, $\mathbf{1}_{\text{bad}}^I \mathbf{P}_I^{\nu(n-1)} \cdots \mathbf{P}_I^{\nu(1)} \mathbf{P}_I^{\nu(0)} \mathbf{1}_{x_0}^I$ gives us the probability of a trajectory of $\Sigma_{I,\text{obs}}$ starting from initial state $(x_0, q_0)$ under input sequence $\nu$ entering into states in $X_I^{\text{bad}}$. Now consider that $((x_0, q_0) \ldots (x_n, q_n))$ reaches $X_I^{\text{bad}}$ at any time step $k \in [0; n]$, then we have for all $i \in [k; n]$, $(x_i, q_i) \in X_I^{\text{bad}}$, and $\{\bar{x}_1 \in X : (\bar{x}_1, \bar{x}_2) \in q_i\} \subseteq X_S$ holds by (8). This means that from time step $k$, the initial-state estimate is a subset of $X_S$, which further implies that when observing the output sequence $(h(x_0), \ldots, h(x_n))$ of $\Sigma$, we get $\{x_0' \in X_0 : \exists (x_0', \ldots, x_n') \in \mathcal{L} \text{ s.t. } \max_{i \in [0;n]} \|h(x_i) - h(x_i')\| \leq \varepsilon\} \subseteq X_S$ by (7). Thus, $\xi_{x_0\nu}$ is not initial-state secure, i.e., $\xi_{x_0\nu} \notin \mathcal{L}_{\mathcal{I}_\varepsilon}^n$, from the definition of $\mathcal{L}_{\mathcal{I}_\varepsilon}^n$ in (4). Recall that $T_I(x_I'|x_I, u) = T(x'|x, u)$ also holds by Definition 4.1. Therefore, we get $\mathbf{1}_{\text{bad}}^I \mathbf{P}_I^{\nu(n-1)} \cdots \mathbf{P}_I^{\nu(1)} \mathbf{P}_I^{\nu(0)} \mathbf{1}_{x_0}^I = \mathbb{P}(\xi_{x_0\nu} \notin \mathcal{L}_{\mathcal{I}_\varepsilon}^n) = 1 - \mathbb{P}(\xi_{x_0\nu} \in \mathcal{L}_{\mathcal{I}_\varepsilon}^n)$. ∎

*Proof of Lemma 4.3:* Consider an initial state $x_0$ in $\Sigma$ and an input sequence $\nu = (\nu(0), \ldots, \nu(n-1))$. Let $\xi_{x_0\nu} = (x_0, \ldots, x_n)$ be an arbitrary state trajectory in $\Sigma$ and $((x_0, q_0) \ldots (x_n, q_n))$ be its corresponding initial-state estimation sequence in $\Sigma_{I,\text{obs}}$. If (9) holds, then by leveraging Lemma 4.2, we get $1 - \mathbb{P}(\xi_{x_0\nu} \in L_{\mathcal{I}_\varepsilon}^n) \leq 1 - \lambda$, which implies $\mathbb{P}(\xi_{x_0\nu} \in L_{\mathcal{I}_\varepsilon}^n) \geq \lambda$. Thus, we conclude that $\Sigma$ is $(\varepsilon, \lambda)$-approximate initial-state opaque over time horizon $n \in \mathbb{N}$ by Definition 3.1. ∎

*Proof of Proposition 5.3:* First note that condition 3 in Definition 5.2 intuitively implies that, for any pair of states $(x, \hat{x})$ in relation $R_x$, there always exist inputs such that the succeeding states $(x', \hat{x}')$ are also in relation $R_x$ with probability at least $1 - \delta$. Consequently, by combining the initial conditions in condition 1 of Definition 5.2, for any pair of initial states that are in relation $(x_0, \hat{x}_0) \in R_x$ and any control strategy $\nu$, there exists $\hat{\nu}$ (and vice versa), such that the corresponding state trajectories $\xi_{x_0\nu}$ and $\hat{\xi}_{\hat{x}_0\hat{\nu}}$ satisfy

$$\mathbb{P}\{(\xi_{x_0\nu}(k), \hat{\xi}_{\hat{x}_0\hat{\nu}}(k)) \in R_x \ \forall k \in [0; n]\} \geq (1-\delta)^n. \quad (32)$$

Now, let us consider the following events for any set $E \subseteq \mathcal{H}^n$: $\mathcal{E}_1 := \{\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in E^{-\epsilon}\}$, $\mathcal{E}_2 := \{\zeta_{x_0\nu} \in E\}$, $\mathcal{E}_3 := \{(\xi_{x_0\nu}(k), \hat{\xi}_{\hat{x}_0\hat{\nu}}(k)) \in R_x \ \forall k \in [0; n]\}$, where $\zeta_{x_0\nu}$ and $\hat{\zeta}_{\hat{x}_0\hat{\nu}}$ are the corresponding output trajectories, and $E^\epsilon$ and $E^{-\epsilon}$ are defined in (13) and (14), respectively. From condition 2 of Definition 5.2, we get $\mathcal{E}_1 \cap \mathcal{E}_3 \implies \mathcal{E}_2$, and thus, $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_3) \leq \mathbb{P}(\mathcal{E}_2)$. This further implies that the following inequality holds $\mathbb{P}(\bar{\mathcal{E}}_2) = 1 - \mathbb{P}(\mathcal{E}_2) \leq 1 - \mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_3) = \mathbb{P}(\bar{\mathcal{E}}_1 \cup \bar{\mathcal{E}}_3) \leq \mathbb{P}(\bar{\mathcal{E}}_1) + \mathbb{P}(\bar{\mathcal{E}}_3)$, where $\bar{\mathcal{E}}_1$, $\bar{\mathcal{E}}_2$, and $\bar{\mathcal{E}}_3$ denote the complements of events $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$, respectively. Since we have, by (32), $\mathbb{P}(\bar{\mathcal{E}}_3) \leq 1 - (1-\delta)^n$, then we get $\mathbb{P}(\bar{\mathcal{E}}_2) \leq \mathbb{P}(\bar{\mathcal{E}}_1) + 1 - (1-\delta)^n \implies 1 - \mathbb{P}(\mathcal{E}_2) \leq 1 -$

$\mathbb{P}(\mathcal{E}_1) + 1 - (1-\delta)^n \Longrightarrow \mathbb{P}(\mathcal{E}_1) \leq \mathbb{P}(\mathcal{E}_2) + 1 - (1-\delta)^n$, which gives us $\mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in E^{-\epsilon}) - \gamma_\delta \leq \mathbb{P}(\zeta_{x_0\nu} \in E)$. The proof of $\mathbb{P}(\zeta_{x_0\nu} \in E) \leq \mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in E^\epsilon) + \gamma_\delta$ uses similar reasoning and is omitted here. ∎

*Proof of Lemma 5.4:* We provide the proof for (18). The proof of (19) can be derived similarly. Let us consider events $\mathcal{E}_1 := \{\xi_{x_0\nu} \in L_{\mathcal{I}_\varepsilon}^n\}$ and $\mathcal{E}_2 := \{\zeta_{x_0\nu} \in \mathcal{H}_{\mathcal{I}_\varepsilon}^n\}$. Note that the proof of (18) can be done by showing $\mathcal{E}_1 \Longleftrightarrow \mathcal{E}_2$. Consider an arbitrary initial state $x_0$, an arbitrary input sequence $\nu$, a finite state run $\xi_{x_0\nu} = (x_0, \ldots, x_n)$ and the corresponding output trajectory $\zeta_{x_0\nu} = (y_0, \ldots, y_n)$ with $y_i = h(x_i)$.

First, let us prove that $\mathcal{E}_1 \Longrightarrow \mathcal{E}_2$. Suppose $\xi_{x_0\nu} = (x_0, \ldots, x_n) \in L_{\mathcal{I}_\varepsilon}^n$. Then, by (16), we get $\zeta_{x_0\nu} \in \mathcal{H}_{\mathcal{I}_\varepsilon}^n$ holds readily, which gives us $\mathcal{E}_1 \Longrightarrow \mathcal{E}_2$. Next, we show $\mathcal{E}_2 \Longrightarrow \mathcal{E}_1$. Suppose $\zeta_{x_0\nu} = (y_0, \ldots, y_n) \in \mathcal{H}_{\mathcal{I}_\varepsilon}^n$, we have by (16)

$$\exists \xi_{x_0'\nu'} = (x_0', \ldots, x_n') \in L_{\mathcal{I}_\varepsilon}^n, y_i = h(x_i') \ \forall i \in [0;n]. \quad (33)$$

In addition, by (4), there exists $(x_0'', \ldots, x_n'') \in \mathcal{L}^n$ s.t.

$$x_0'' \in X_0 \backslash X_S, \text{ and } \max_{i \in [0;n]} \|h(x_i') - h(x_i'')\| \leq \varepsilon. \quad (34)$$

Combining (34) with (33), we have $\|y_i - h(x_i'')\| \leq \varepsilon \ \forall i \in [0;n]$. Therefore, we have for $\xi_{x_0\nu} = (x_0, \ldots, x_n)$, $\exists x_0'' \in X_0 \backslash X_S$, s.t. $\max_{i \in [0;n]} \|y_i - h(x_i'')\| \leq \varepsilon$. Thus, it follows that $\xi_{x_0\nu} \in L_{\mathcal{I}_\varepsilon}^n$ holds as well, which further implies $\mathcal{E}_2 \Longrightarrow \mathcal{E}_1$. ∎

*Proof of Theorem 5.5:* First note that for a nonsecret initial-state $x_0 \in X_0 \backslash X_S$, any trajectory from $x_0$ satisfies $\xi_{x_0\nu} \in \mathcal{L}_{\mathcal{I}_\varepsilon}^n$, which directly implies that $\mathbb{P}(\xi_{x_0\nu} \in \mathcal{L}_{\mathcal{I}_\varepsilon}^n) = 1$ for any $\varepsilon \in \mathbb{R}_{\geq 0}$. Now, consider an arbitrary initial state $x_0 \in X_0 \cap X_S$ and a random state trajectory $\xi_{x_0\nu} = (x_0, \ldots, x_n)$ with length $n$ generated from $x_0$ under an input sequence $\nu$ in system $\Sigma$. Since $\hat{\Sigma} \preceq_{\mathcal{I}_\delta^\epsilon} \Sigma$, by condition 1a in Definition 5.2, for $x_0 \in X_0 \cap X_S$ in $\Sigma$, there exists $\hat{x}_0 \in \hat{X}_0 \cap \hat{X}_S$ in $\hat{\Sigma}$, such that $(x_0, \hat{x}_0) \in R_x$. Consider this pair of initial states $(x_0, \hat{x}_0) \in R_x$, by Proposition 5.3 together with condition 3a in Definition 5.2, there exists an input sequence $\hat{\nu}$ in $\hat{\Sigma}$, s.t. given any $E \subseteq \mathcal{H}^n$ and the $\epsilon$-neighborhoods $E^\epsilon$ and $E^{-\epsilon}$ as in (13) and (14), we have

$$\mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in E^{-\epsilon}) - \gamma_\delta \leq \mathbb{P}(\zeta_{x_0\nu} \in E) \leq \mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in E^\epsilon) + \gamma_\delta \quad (35)$$

with constant $\gamma_\delta = 1 - (1-\delta)^n$, where $\zeta_{x_0\nu}$ and $\hat{\zeta}_{\hat{x}_0\hat{\nu}}$ denotes the output sequence corresponding to the state trajectory $\xi_{x_0\nu}$ in $\Sigma$ and the state trajectory $\hat{\xi}_{\hat{x}_0\hat{\nu}}$ in $\hat{\Sigma}$, respectively. Now notice that since $\hat{\Sigma}$ is $(\varepsilon, \lambda)$-approximate initial-state opaque within time horizon $n \in \mathbb{N}$, we have from Definition 3.1 that $\mathbb{P}(\hat{\xi}_{\hat{x}_0\hat{\nu}} \in \hat{\mathcal{L}}_{\mathcal{I}_\varepsilon}^n) \geq \lambda$, where $\hat{\mathcal{L}}_{\mathcal{I}_\varepsilon}^n$ is the set of initial-state secure state trajectories in $\hat{\Sigma}$ as defined in (4). Thus, by Lemma 5.4, the corresponding output run $\hat{\zeta}_{\hat{x}_0\hat{\nu}} = (\hat{y}_0, \ldots, \hat{y}_n)$ satisfies

$$\mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in \hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^n) \geq \lambda \quad (36)$$

where $\hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^n$ is the set of secure output sequences in $\hat{\Sigma}$ as defined in (16), i.e.,

$$\hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^n = \{(\hat{y}_0, \ldots, \hat{y}_n) \in \hat{\mathcal{H}}^n : \exists (\hat{x}_0', \ldots, \hat{x}_n') \in \hat{\mathcal{L}}_{\mathcal{I}_\varepsilon}^n, \text{ s.t.}$$
$$\hat{y}_i = \hat{h}(\hat{x}_i') \ \forall i \in [0;n]\}. \quad (37)$$

Combining (37) and (36), we have

$$\mathbb{P}(\zeta_{x_0\nu} \in \hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^\epsilon) \geq \mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in \hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^n) - \gamma_\delta \geq \lambda - \gamma_\delta \quad (38)$$

where $\hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^\epsilon$ is $\epsilon$-expansion of $\hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^n$ as defined in (13), which can be further rewritten as

$$\hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^\epsilon \overset{(13)}{=} \{(\tilde{\hat{y}}_0, \ldots, \tilde{\hat{y}}_n) \in \hat{\mathcal{H}}^n : \exists (\hat{y}_0, \ldots, \hat{y}_n) \in \hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^n, \text{ s.t.}$$
$$\max_{i \in [0;n]} \|\hat{y}_i - \tilde{\hat{y}}_i\| \leq \epsilon\}$$

$$\overset{(38)}{=} \{(\tilde{\hat{y}}_0, \ldots, \tilde{\hat{y}}_n) \in \hat{\mathcal{H}}^n : \exists (\hat{x}_0', \ldots, \hat{x}_n') \in \hat{\mathcal{L}}_{\mathcal{I}_\varepsilon}^n, \text{ s.t.}$$
$$\max_{i \in [0;n]} \|\tilde{\hat{y}}_i - \hat{h}(\hat{x}_i')\| \leq \epsilon\}. \quad (39)$$

Next, in order to show the approximate initial-state opacity of system $\Sigma$, let us consider the set of secure output runs within time horizon $n \in \mathbb{N}$ in $\Sigma$ w.r.t. the constant $(\varepsilon + 2\epsilon) \geq 0$ as follows:

$$\mathcal{H}_{\mathcal{I}_{(\varepsilon+2\epsilon)}}^n = \{(y_0, \ldots, y_n) \in \mathcal{H}^n : \exists (x_0, \ldots, x_n) \in \mathcal{L}_{\mathcal{I}_{(\varepsilon+2\epsilon)}}^n,$$
$$\text{s.t. } y_i = h(x_i) \ \forall i \in [0;n]\}. \quad (40)$$

Consider the trajectory $(\hat{x}_0', \ldots, \hat{x}_n') \in \hat{\mathcal{L}}_{\mathcal{I}_\varepsilon}^n$ in (39). By (4), we get that there exists $(\hat{x}_0'', \ldots, \hat{x}_n'') \in \hat{\mathcal{L}}^n$ such that

$$\hat{x}_0'' \in \hat{X}_0 \backslash \hat{X}_S, \text{ and } \max_{i \in [0;n]} \|\hat{h}(\hat{x}_i') - \hat{h}(\hat{x}_i'')\| \leq \varepsilon. \quad (41)$$

Note that by conditions 1b and 3b in Definition 5.2, for $\hat{x}_0'' \in \hat{X}_0 \backslash \hat{X}_S$ in $\hat{\Sigma}$, there exists $x_0' \in X_0 \backslash X_S$ in $\Sigma$ with $(x_0', \hat{x}_0'') \in R_x$, such that the trajectory $(x_0', \ldots, x_n')$ from $x_0'$ satisfies $(x_i', \hat{x}_i'') \in R_x \ \forall i \in [0;n]$ with probability at least $(1-\delta)^n$, i.e., $\mathbb{P}\{(x_i', \hat{x}_i'') \in R_x \ \forall i \in [0;n]\} \geq (1-\delta)^n$. Let us consider again the random trajectory $\xi_{x_0\nu}$ in $\Sigma$ and its corresponding output trajectory $\zeta_{x_0\nu}$, and define the events: $\mathcal{E}_1 := \{(x_i', \hat{x}_i'') \in R_x \ \forall i \in [0;n]\}$; $\mathcal{E}_2 := \{\zeta_{x_0\nu} \in \hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^\epsilon\}$; and $\mathcal{E}_3 := \{\zeta_{x_0\nu} \in \mathcal{H}_{\mathcal{I}_{(\varepsilon+2\epsilon)}}^n\}$. Note that $\mathcal{E}_1$ further implies that

$$\max_{i \in [0;n]} \|h(x_i') - \hat{h}(\hat{x}_i'')\| \leq \epsilon \quad (42)$$

by condition 2 in Definition 5.2. Now consider $\zeta_{x_0\nu} = (y_0, \ldots, y_n)$ with $\mathcal{E}_2 := \{\zeta_{x_0\nu} \in \hat{\mathcal{H}}_{\mathcal{I}_\varepsilon}^\epsilon\}$ holds, by leveraging the aforementioned implications from $\mathcal{E}_2$, we get that there exists $(x_0', \ldots, x_n')$ with $x_0' \in X_0 \backslash X_S$ in $\Sigma$, such that $\max_{i \in [0;n]} \|h(x_i') - y_i\| \leq \varepsilon + 2\epsilon$, where the inequality is obtained by combining (39), (41), and (42). This implies that $\zeta_{x_0\nu} \in \mathcal{H}_{\mathcal{I}_{(\varepsilon+2\epsilon)}}^n$ holds. Therefore, we obtain that $\mathcal{E}_1 \cap \mathcal{E}_2 \Longrightarrow \mathcal{E}_3$ holds, and thus, $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \mathbb{P}(\mathcal{E}_3)$. Therefore, we get $\mathbb{P}(\zeta_{x_0\nu} \in \mathcal{H}_{\mathcal{I}_{(\varepsilon+2\epsilon)}}^n) \overset{(39)}{\geq} (1-\delta)^n(\lambda - \gamma_\delta) = (1-\gamma_\delta)(\lambda - \gamma_\delta)$. By leveraging again (18), we have $\mathbb{P}(\xi_{x_0\nu} \in L_{\mathcal{I}_{(\varepsilon+2\epsilon)}}^n) \geq (1-\gamma_\delta)(\lambda - \gamma_\delta)$. Thus, we can conclude that $\Sigma$ is $(\varepsilon + 2\epsilon, \lambda_\delta)$-approximate initial-state opaque within time horizon $n \in \mathbb{N}$, with $\lambda_\delta = (1-\gamma_\delta)(\lambda - \gamma_\delta)$. ∎

*Proof of Theorem 5.9:* Consider an arbitrary initial state $x_0$ and a random state trajectory $\xi_{x_0\nu} = (x_0, \ldots, x_n)$ with length $n$ generated from $x_0$ under a control sequence $\nu$ in system $\Sigma$. Since $\hat{\Sigma} \preceq_{\mathcal{C}_\delta^\epsilon} \Sigma$, by condition 1a in Definition 5.7, for $x_0 \in X_0$ in $\Sigma$, there exists $\hat{x}_0 \in \hat{X}_0$ in $\hat{\Sigma}$, such that $(x_0, \hat{x}_0) \in R_x$. Now

consider this pair of initial states $(x_0, \hat{x}_0) \in R_x$, by Proposition 5.8 together with condition 3a in Definition 5.7, there exists an input sequence $\hat{\nu}$ in $\hat{\Sigma}$, such that given any set $E \subseteq \mathcal{H}^n$ and the $\epsilon$-neighborhoods $E^\epsilon$ and $E^{-\epsilon}$ as defined in (13) and (14), we have

$$\mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in E^{-\epsilon}) - \gamma_\delta \leq \mathbb{P}(\zeta_{x_0\nu} \in E) \leq \mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in E^\epsilon) + \gamma_\delta \quad (43)$$

with constant $\gamma_\delta = 1 - (1 - \delta)^n$, where $\zeta_{x_0\nu}$ and $\hat{\zeta}_{\hat{x}_0\hat{\nu}}$ denotes the output sequence corresponding to the state trajectory $\xi_{x_0\nu}$ in $\Sigma$ and the state trajectory $\hat{\xi}_{\hat{x}_0\hat{\nu}}$ in $\hat{\Sigma}$, respectively. Now notice that since $\hat{\Sigma}$ is $(\varepsilon, \lambda)$-approximate current-state opaque within time horizon $n \in \mathbb{N}$, we have from Definition 3.1 that $\mathbb{P}(\hat{\xi}_{\hat{x}_0\hat{\nu}} \in \hat{L}^n_{\mathcal{C}_\varepsilon}) \geq \lambda$, where $\hat{\mathcal{L}}^n_{\mathcal{C}_\varepsilon}$ is the set of current-state secure state trajectories in $\hat{\Sigma}$ as defined in (5). Thus, by Lemma 5.4, the corresponding output run $\hat{\zeta}_{\hat{x}_0\hat{\nu}} = (\hat{y}_0, \ldots, \hat{y}_n)$ satisfies

$$\mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in \hat{\mathcal{H}}^n_{\mathcal{C}_\varepsilon}) \geq \lambda \quad (44)$$

where $\hat{\mathcal{H}}^n_{\mathcal{C}_\varepsilon}$ is the set of secure output sequences in $\hat{\Sigma}$ as defined in (17), i.e.,

$$\hat{\mathcal{H}}^n_{\mathcal{C}_\varepsilon} = \{(\hat{y}_0, \ldots, \hat{y}_n) \in \hat{\mathcal{H}}^n : \exists (\hat{x}'_0, \ldots, \hat{x}'_n) \in \hat{\mathcal{L}}^n_{\mathcal{C}_\varepsilon}, \text{ s.t.}$$
$$\hat{y}_i = \hat{h}(\hat{x}'_i) \, \forall i \in [0; n]\}. \quad (45)$$

Combining (43) and (44), we have $\mathbb{P}(\zeta_{x_0\nu} \in \hat{\mathcal{H}}^\epsilon_{\mathcal{C}_\varepsilon}) \geq \mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in \hat{\mathcal{H}}^n_{\mathcal{C}_\varepsilon}) - \gamma_\delta \geq \lambda - \gamma_\delta$, where $\hat{\mathcal{H}}^\epsilon_{\mathcal{C}_\varepsilon}$ is $\epsilon$-expansion of $\hat{\mathcal{H}}^n_{\mathcal{C}_\varepsilon}$ as defined in (13), which can be further rewritten as

$$\hat{\mathcal{H}}^\epsilon_{\mathcal{C}_\varepsilon} \overset{(13)}{=} \{(\hat{\bar{y}}_0, \ldots, \hat{\bar{y}}_n) \in \hat{\mathcal{H}}^n : \exists (\hat{y}_0, \ldots, \hat{y}_n) \in \hat{\mathcal{H}}^n_{\mathcal{C}_\varepsilon}, \text{ s.t.}$$
$$\max_{i \in [0; n]} \|\hat{y}_i - \hat{\bar{y}}_i\| \leq \epsilon\}$$
$$\overset{(46)}{=} \{(\hat{\bar{y}}_0, \ldots, \hat{\bar{y}}_n) \in \hat{\mathcal{H}}^n : \exists (\hat{x}'_0, \ldots, \hat{x}'_n) \in \hat{\mathcal{L}}^n_{\mathcal{C}_\varepsilon}, \text{ s.t.}$$
$$\max_{i \in [0; n]} \|\hat{\bar{y}}_i - \hat{h}(\hat{x}'_i)\| \leq \epsilon\}. \quad (46)$$

Next, in order to show the approximate current-state opacity of system $\Sigma$, let us consider the set of secure output runs in $\Sigma$ within time horizon $n \in \mathbb{N}$ w.r.t. the constant $\varepsilon + 2\epsilon \geq 0$

$$\mathcal{H}^n_{\mathcal{C}_{(\varepsilon+2\epsilon)}} = \{(y_0, \ldots, y_n) \in \mathcal{H}^n : \exists (x_0, \ldots, x_n) \in \mathcal{L}^n_{\mathcal{C}_{(\varepsilon+2\epsilon)}},$$
$$\text{s.t. } y_i = h(x_i) \, \forall i \in [0; n]\}. \quad (47)$$

Consider the trajectory $(\hat{x}'_0, \ldots, \hat{x}'_n) \in \hat{\mathcal{L}}^n_{\mathcal{C}_\varepsilon}$ in (46). By (5), we get that for any of its prefix $(\hat{x}'_0, \ldots, \hat{x}'_k) \, \forall k \in [0; n]$, there exists $(\hat{x}''_0, \ldots, \hat{x}''_k) \in \hat{\mathcal{L}}^n$ such that

$$\hat{x}''_k \in \hat{X} \setminus \hat{X}_S, \text{ and } \max_{i \in [0; k]} \|\hat{h}(\hat{x}'_i) - \hat{h}(\hat{x}''_i)\| \leq \varepsilon. \quad (48)$$

For any of the prefix $(\hat{x}'_0, \ldots, \hat{x}'_k)$, by conditions 1b, 3c, and 3d in Definition 5.7, there exists $x'_0 \in X_0$ with $(x'_0, \hat{x}'_0) \in R_x$, s.t. the trajectory $(x'_0, \ldots, x'_k)$ satisfies $(x'_i, \hat{x}''_i) \in R_x \, \forall i \in [0; k]$ with probability at least $(1 - \delta)^n$, and $x'_k \in X \setminus X_S$, i.e., $\mathbb{P}\{(x''_i, \hat{x}''_i) \in R_x \, \forall i \in [0; k]\} \geq (1 - \delta)^k$.

Let us consider again the random trajectory $\xi_{x_0\nu}$ in $\Sigma$ and its corresponding output trajectory $\zeta_{x_0\nu}$, and define the events: $\mathcal{E}_1 := \{(x'_i, \hat{x}''_i) \in R_x \, \forall i \in [0; n]\}$; $\mathcal{E}_2 := \{\zeta_{x_0\nu} \in \hat{\mathcal{H}}^\epsilon_{\mathcal{C}_\varepsilon}\}$; and $\mathcal{E}_3 := \{\zeta_{x_0\nu} \in \mathcal{H}^n_{\mathcal{C}_{(\varepsilon+2\epsilon)}}\}$. Note that $\mathcal{E}_1$ further implies that

$$\max_{i \in [0; n]} \|h(x'_i) - \hat{h}(\hat{x}''_i)\| \leq \epsilon \quad (49)$$

by condition 2 in Definition 5.7. Now consider $\zeta_{x_0\nu} = (y_0, \ldots, y_n)$ with $\mathcal{E}_2 := \{\zeta_{x_0\nu} \in \hat{\mathcal{H}}^\epsilon_{\mathcal{C}_\varepsilon}\}$ holds, by leveraging the aforementioned implications from $\mathcal{E}_2$, we get that there exists $(x'_0, \ldots, x'_n)$ in $\Sigma$ being current-state secure with none of its prefix revealing current-state secret as in (48), such that $\max_{i \in [0; n]} \|h(x'_i) - y_i\| \leq \varepsilon + 2\epsilon$, where the inequality is obtained by combining (46), (48), and (49). This implies that $\zeta_{x_0\nu} \in \mathcal{H}^n_{\mathcal{C}_{(\varepsilon+2\epsilon)}}$ holds. Therefore, we obtain that $\mathcal{E}_1 \cap \mathcal{E}_2 \Longrightarrow \mathcal{E}_3$ holds, and thus, $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \mathbb{P}(\mathcal{E}_3)$. Therefore, we get $\mathbb{P}(\zeta_{x_0\nu} \in \mathcal{H}^n_{\mathcal{C}_{(\varepsilon+2\epsilon)}}) \overset{(39)}{\geq} (1 - \delta)^n (\lambda - \gamma_\delta) = (1 - \gamma_\delta)(\lambda - \gamma_\delta)$. By leveraging again (19), we have $\mathbb{P}(\xi_{x_0\nu} \in L^n_{\mathcal{C}_{(\varepsilon+2\epsilon)}}) \geq (1 - \gamma_\delta)(\lambda - \gamma_\delta)$. Thus, we can conclude that $\Sigma$ is $(\varepsilon + 2\epsilon, \lambda_\delta)$-approximate current-state opaque within time horizon $n \in \mathbb{N}$, with $\lambda_\delta = (1 - \gamma_\delta)(\lambda - \gamma_\delta)$. $\blacksquare$

*Proof of Theorem 6.2:* We start with showing condition 1 in Definition 5.2. Consider any secret initial state $x_0 \in X_0 \cap X_S$ in $\Sigma$, since $\eta \leq \text{span}(X_S)$, there always exists a representative point $\hat{x}_0 = \Pi_x(x_0)$ in $\hat{\Sigma}$ inside the set $\hat{X}_0 \cap \hat{X}_S$ by the construction of sets $\hat{X}_0 = [X]_\eta$ and $\hat{X}_S = [X^\theta_S]_\eta$. Note that $\|x_0 - \hat{x}_0\| \leq \eta$ holds by (28). By combining (23) and (27), we have $V(x_0, \hat{x}_0) \leq \overline{\alpha}(\|x_0 - \hat{x}_0\|) \leq \overline{\alpha}(\eta) \leq \underline{\alpha} \circ \ell^{-1}(\varepsilon)$, which implies $(x_0, \hat{x}_0) \in R_x$ by the definition of $R_x$ in (28). Hence, condition 1a is satisfied readily. For condition 1b, for every $\hat{x}_0 \in \hat{X}_0 \setminus \hat{X}_S$, by choosing $x_0 = \hat{x}_0$ which is also inside $X_0 \setminus X_S$, we again get $V(x_0, \hat{x}_0) = 0 \leq \underline{\alpha} \circ \ell^{-1}(\varepsilon)$, which implies that $(x_0, \hat{x}_0) \in R_x$ holds. We proceed with showing condition 2. Combining the Lipschitz assumption (22) and (23), one gets for any $(x, \hat{x}) \in R_x$, $\|h(x) - h(\hat{x})\| \overset{(23)}{\leq} \ell(\|x - \hat{x}\|) \overset{(24)}{\leq} \ell \circ \underline{\alpha}^{-1}(V(x, \hat{x})) \overset{(29)}{\leq} \epsilon$, which shows that condition 2 is satisfied.

Next, we show condition 3 also holds. By taking the conditional expectation from (25), for any $(x, \hat{x}) \in R_x$, for $u \in U$, there exists $\hat{u} \in \hat{U}$, and vice versa, such that $\mathbb{E}[V(f(x, u, \varsigma), \hat{f}(\hat{x}, \hat{u}, \varsigma)) \mid x, \hat{x}, u, \hat{u}] - \mathbb{E}[V(f(x, u, \varsigma), f(\hat{x}, \hat{u}, \varsigma)) \mid x, \hat{x}, u, \hat{u}] \leq \mathbb{E}[\gamma(\|\hat{f}(\hat{x}, \hat{u}, \varsigma) - f(\hat{x}, \hat{u}, \varsigma)\|) \mid x, \hat{x}, u, \hat{u}]$, where $\hat{f}(\hat{x}, \hat{u}, \varsigma) = \Pi_x(f(\hat{x}, \hat{u}, \varsigma))$. Using (26), we further get

$$\mathbb{E}\left[V(f(x, u, \varsigma), \hat{f}(\hat{x}, \hat{u}, \varsigma)) \mid x, \hat{x}, u, \hat{u}\right]$$
$$- \mathbb{E}\left[V(f(x, u, \varsigma), f(\hat{x}, \hat{u}, \varsigma)) \mid x, \hat{x}, u, \hat{u}\right] \leq \gamma(\eta). \quad (50)$$

By (24), we have $\mathbb{E}[V(f(x, u, \varsigma), f(\hat{x}, \hat{u}, \varsigma)) \mid x, \hat{x}, u, \hat{u}] \leq (\mathcal{I}_d - \bar{\kappa})(V(x, \hat{x})) + \bar{\rho}(\|u - u'\|)$. By further using (26) and (50), it follows that

$$\mathbb{E}\left[V(f(x, u, \varsigma), \hat{f}(\hat{x}, \hat{u}, \varsigma)) \mid x, \hat{x}, u, \hat{u}\right]$$
$$\leq (\mathcal{I}_d - \bar{\kappa})(V(x, \hat{x})) + \bar{\rho}(\mu) + \gamma(\eta). \quad (51)$$

Applying Markov's inequality [50] to (51), we get the following inequality chain:

$$\mathbb{P}\{V(f(x, u, \varsigma), \hat{f}(\hat{x}, \hat{u}, \varsigma)) \leq \underline{\alpha} \circ \ell^{-1}(\epsilon) \mid x, \hat{x}, u, \hat{u}\}$$
$$= 1 - \mathbb{P}\{V(f(x, u, \varsigma), \hat{f}(\hat{x}, \hat{u}, \varsigma)) \geq \underline{\alpha} \circ \ell^{-1}(\epsilon) \mid x, \hat{x}, u, \hat{u}\}$$
$$\geq 1 - \frac{1}{\underline{\alpha} \circ \ell^{-1}(\epsilon)} \mathbb{E}\left[V(f(x, u, \varsigma), \hat{f}(\hat{x}, \hat{u}, \varsigma)) \mid x, \hat{x}, u, \hat{u}\right]$$

$$\stackrel{(29)}{\geq} 1 - \frac{1}{\underline{\alpha} \circ \ell^{-1}(\epsilon)} \left[ (\mathcal{I}_d - \bar{\kappa}) \circ \underline{\alpha} \circ \ell^{-1}(\epsilon) + \bar{\rho}(\mu) + \gamma(\eta) \right]$$

$$\stackrel{(28)}{\geq} 1 - \delta. \tag{52}$$

Note that the last inequality in (52) is achieved from (27), where one has $\frac{1}{\underline{\alpha} \circ \ell^{-1}(\epsilon)} \{ (\mathcal{I}_d - \bar{\kappa}) \circ \underline{\alpha} \circ \ell^{-1}(\epsilon) + \bar{\rho}(\mu) + \gamma(\eta) \} \leq \delta$. Hence, condition 3 of Definition 5.2 is also satisfied, which completes the proof. ∎

*Proof of Theorem 6.3:* We start with showing condition 1 in Definition 5.7. Consider any initial state $x_0 \in X_0$ in $\Sigma$, since $\eta \leq \text{span}(X_0)$, there always exists a representative point $\hat{x}_0 = \Pi_x(x_0)$ in $\hat{\Sigma}$ by the construction of $\hat{X}_0$ such that $\|x_0 - \hat{x}_0\| \leq \eta$ holds by (26). By (23) and (29), we have $V(x_0, \hat{x}_0) \leq \overline{\alpha}(\|x_0 - \hat{x}_0\|) \leq \overline{\alpha}(\eta) \leq \underline{\alpha} \circ \ell^{-1}(\epsilon)$, which implies $(x_0, \hat{x}_0) \in R_x$ by the definition of $R_x$ in (31). For condition 1b, for every $\hat{x}_0 \in \hat{X}_0$, by choosing $x_0 = \hat{x}_0$, which is also inside $X_0$, we again get $V(x_0, \hat{x}_0) = 0 \leq \underline{\alpha} \circ \ell^{-1}(\varepsilon)$, which implies that $(x_0, \hat{x}_0) \in R_x$ holds. Hence, condition 1 is satisfied readily. Conditions 2, 3a, and 3c can be shown following the same reasoning as in Theorem 6.2, thus is omitted here. Finally, let us show that conditions 3b and 3d in Definition 5.7 hold as well. Consider any pair of states $(x, \hat{x}) \in R_x$, and suppose that under an input $u \in U$, we get $x^+ = f(x, u, \varsigma) \in X_S$. Similar to the proof of condition 3a, we can show that by choosing $\hat{u} = u$, we have $(x^+, \hat{x}^+) \in R_x$ with probability greater than $1 - \delta$, where $\hat{x}^+ = \hat{f}(\hat{x}, \hat{u}, \varsigma)$. Since $(x^+, \hat{x}^+) \in R_x$, by combining (31) with (23) and (30), we have $\|x^+ - \hat{x}^+\| \leq \underline{\alpha}^{-1}(V(x^+, \hat{x}^+)) \leq \ell^{-1}(\epsilon) \leq \theta$. Moreover, by the construction of the secret set in the finite abstraction, one has $\hat{X}_S = [X_S^\theta]_\eta$ with the inflation parameter $\theta$ satisfying $\ell^{-1}(\epsilon) \leq \theta$ and $0 < \eta \leq \min\{\text{span}(X_S), \text{span}(X \setminus X_S)\}$. This further implies that $\hat{x}^+ \in \hat{X}_S$ holds as well. Thus, condition 3b is satisfied. Condition 3d can be proved in a similar way. Therefore, we conclude that condition 3 of Definition 5.7 is satisfied, which completes the proof. ∎

## REFERENCES

[1] H. Sandberg, S. Amin, and K. H. Johansson, "Cyberphysical security in networked control systems: An introduction to the issue," *IEEE Control Syst. Mag.*, vol. 35, no. 1, pp. 20–23, Feb. 2015.

[2] S. M. Dibaji, M. Pirani, D. B. Flamholz, A. M. Annaswamy, K. H. Johansson, and A. Chakrabortty, "A systems and control perspective of CPS security," *Annu. Rev. Control*, vol. 47, pp. 394–411, 2019.

[3] S. Liu, A. Trivedi, X. Yin, and M. Zamani, "Secure-by-construction synthesis of cyber-physical systems," *Annu. Rev. Control*, vol. 53, pp. 30–50, 2022.

[4] A. Saboori and C. N. Hadjicostis, "Notions of security and opacity in discrete event systems," in *Proc. IEEE Conf. Decis. Control*, 2007, pp. 5056–5061.

[5] S. Lafortune, F. Lin, and C. N. Hadjicostis, "On the history of diagnosability and opacity in discrete event systems," *Annu. Rev. Control*, vol. 45, pp. 257–266, 2018.

[6] L. Mazaré, "Using unification for opacity properties," in *Proc. Workshop Issues Theory Secur.*, vol. 4, 2004, pp. 165–176.

[7] F. Lin, "Opacity of discrete event systems and its applications," *Automatica*, vol. 47, no. 3, pp. 496–503, 2011.

[8] S. Yang and X. Yin, "Secure your intention: On notions of pre-opacity in discrete-event systems," *IEEE Trans. Autom. Control*, vol. 68, no. 8, pp. 4754–4766, Aug. 2023.

[9] A. Saboori and C. Hadjicostis, "Verification of infinite-step opacity and complexity considerations," *IEEE Trans. Autom. Control*, vol. 57, no. 5, pp. 1265–1269, May 2012.

[10] X. Yin and S. Lafortune, "A new approach for the verification of infinite-step and k-step opacity using two-way observers," *Automatica*, vol. 80, pp. 162–171, 2017.

[11] J. W. Bryans, M. Koutny, L. Mazaré, and P. Y. Ryan, "Opacity generalised to transition systems," *Int. J. Inf. Secur.*, vol. 7, no. 6, pp. 421–435, 2008.

[12] Y. Tong, H. Lan, and C. Seatzu, "Verification of k-step and infinite-step opacity of bounded labeled petri nets," *Automatica*, vol. 140, 2022, Art. no. 110221.

[13] R. Jacob, J.-J. Lesage, and J.-M. Faure, "Overview of discrete event systems opacity: Models, validation, and quantification," *Annu. Rev. Control*, vol. 41, pp. 135–146, 2016.

[14] X. Yin, M. Zamani, and S. Liu, "On approximate opacity of cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 66, no. 4, pp. 1630–1645, Apr. 2020.

[15] B. Ramasubramanian, W. Cleaveland, and S. Marcus, "Notions of centralized and decentralized opacity in linear systems," *IEEE Trans. Autom. Control*, vol. 265, no. 4, pp. 1442–1455, 2020.

[16] L. An and G.-H. Yang, "Opacity enforcement for confidential robust control in linear cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 1234–1241, Mar. 2020.

[17] S. Liu and M. Zamani, "Verification of approximate opacity via barrier certificates," *IEEE Control Syst. Lett.*, vol. 5, no. 4, pp. 1369–1374, Apr. 2020.

[18] M. Mizoguchi and T. Ushio, "Abstraction-based control under quantized observation with approximate opacity using symbolic control barrier functions," *IEEE Control Syst. Lett.*, vol. 6, pp. 2222–2227, Dec. 2021.

[19] S. Liu and M. Zamani, "Compositional synthesis of opacity-preserving finite abstractions for interconnected systems," *Automatica*, vol. 131, 2021, Art. no. 109745.

[20] C. Baier, J.-P. Katoen, and K. G. Larsen, *Principles of Model Checking*. Cambridge, MA, USA: MIT Press, 2008.

[21] A. Lavaei, S. Soudjani, A. Abate, and M. Zamani, "Automated verification and synthesis of stochastic hybrid systems: A survey," *Automatica*, vol. 146, 2022, Art. no. 110617.

[22] W. Zeng and M. Koutny, "Quantitative analysis of opacity in cloud computing systems," *IEEE Trans. Cloud Comput.*, vol. 9, no. 3, pp. 1210–1219, Jul.–Sep. 2021.

[23] C. Mu and D. Clark, "Verifying opacity properties in security systems," *IEEE Trans. Dependable Secure Comput.*, vol. 20, no. 2, pp. 1450–1460, Feb. 2022.

[24] A. Saboori and C. N. Hadjicostis, "Current-state opacity formulations in probabilistic finite automata," *IEEE Trans. Autom. Control*, vol. 59, no. 1, pp. 120–133, Jan. 2014.

[25] B. Bérard, K. Chatterjee, and N. Sznajder, "Probabilistic opacity for Markov decision processes," *Inf. Process. Lett.*, vol. 115, no. 1, pp. 52–59, 2015.

[26] J. Chen, M. Ibrahim, and R. Kumar, "Quantification of secrecy in partially observed stochastic discrete event systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 1, pp. 185–195, Jan. 2017.

[27] M. Ahmadi, B. Wu, H. Lin, and U. Topcu, "Privacy verification in POMDPs via barrier certificates," in *Proc. IEEE Conf. Decis. Control*, 2018, pp. 5610–5615.

[28] X. Yin, Z. Li, W. Wang, and S. Li, "Infinite-step opacity and k-step opacity of stochastic discrete-event systems," *Automatica*, vol. 99, pp. 266–274, 2019.

[29] S. Liu, X. Yin, and M. Zamani, "On a notion of approximate opacity for discrete-time stochastic control systems," in *Proc. Amer Control Conf.*, 2020, pp. 5413–5418.

[30] S. Haesaert, S. E. Zadeh Soudjani, and A. Abate, "Verification of general Markov processes by approximate similarity relations and policy refinement," *SIAM J. Control Optim*, vol. 55, no. 4, pp. 2333–2367, 2017.

[31] A. Lavaei, S. Soudjani, and M. Zamani, "Compositional abstraction-based synthesis of general MDPS via approximate probabilistic relations," *Nonlinear Anal., Hybrid Syst.*, vol. 39, 2021, Art. no. 100991.

[32] O. Kallenberg and O. Kallenberg, *Foundations of Modern Probability*, vol. 2. Berlin, Germany: Springer, 1997.

[33] J. Lu and R. E. Skelton, "Mean-square small gain theorem for stochastic control: Discrete-time case," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 490–494, Mar. 2002.

[34] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden, "Metrics for labelled Markov processes," *Theor. Comput. Sci.*, vol. 318, no. 3, pp. 323–354, 2004.

[35] J. Ding, M. Kamgarpour, S. Summers, A. Abate, J. Lygeros, and C. Tomlin, "A Stochastic games framework for verification and control of discrete time stochastic hybrid systems," *Automatica*, vol. 49, no. 9, pp. 2665–2674, 2013.

[36] J.-W. Lee and G. E. Dullerud, "Uniform stabilization of discrete-time switched and Markovian jump linear systems," *Automatica*, vol. 42, no. 2, pp. 205–218, 2006.

[37] R. E. Kalman, "On the general theory of control systems," in *Proc 1st Int. Conf. Autom. Control*, Moscow, Russia, 1960, pp. 481–492.

[38] H.-F. Chen, "On stochastic observability and controllability," *Automatica*, vol. 16, no. 2, pp. 179–190, 1980.

[39] G. Chen, C. Fan, J. Sun, and J. Xia, "Mean square exponential stability analysis for Itô stochastic systems with aperiodic sampling and multiple time-delays," *IEEE Trans. Autom. Control*, vol. 67, no. 5, pp. 2473–2480, May 2021.

[40] G. Chen, G. Du, J. Xia, X. Xie, and J. H. Park, "Controller synthesis of aperiodic sampled-data networked control system with application to interleaved flyback module integrated converter," *IEEE Trans. Circuits Syst. I, Reg. Papers.*, vol. 70, no. 11, pp. 4570–4580, Nov. 2023.

[41] G. Chen, J. Xia, J. H. Park, H. Shen, and G. Zhuang, "Sampled-data synchronization of stochastic Markovian jump neural networks with time-varying delay," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 8, pp. 3829–3841, Aug. 2021.

[42] I. Chadès, G. Chapron, M.-J. Cros, F. Garcia, and R. Sabbadin, "MDPtoolbox: A multi-platform toolbox to solve stochastic dynamic programming problems," *Ecography*, vol. 37, no. 9, pp. 916–920, 2014.

[43] M. Zamani, P. Mohajerin, R. Esfahani, A. M. Abate, and J. Lygeros, "Symbolic control of stochastic systems via approximately bisimilar finite abstractions," *IEEE Trans. Autom. Control*, vol. 59, no. 12, pp. 3135–3150, Dec. 2014.

[44] D. N. Tran, B. S. Rüffer, and C. M. Kellett, "Convergence properties for discrete-time nonlinear systems," *IEEE Trans. Autom. Control*, vol. 64, no. 8, pp. 3415–3422, Aug. 2018.

[45] A. Lavaei, S. Soudjani, and M. Zamani, "From dissipativity theory to compositional construction of finite Markov decision processes," in *Proc. 21st Int. Conf. Hybrid Syst., Comput. Control*, 2018, pp. 21–30.

[46] C. Canudas-de Wit, L. L. Ojeda, and A. Y. Kibangou, "Graph constrained-CTM observer design for the Grenoble south ring," *IFAC Proc. Vols.*, vol. 45, no. 24, pp. 197–202, 2012.

[47] S. Soudjani, C. Gevaerts, and A. Abate, "FAUST$^2$: Formal abstractions of uncountable-state stochastic processes," in *Tools and Algorithms for the Construction and Analysis of Systems* (Lecture Notes in Computer Science), vol. 9035. Berlin, Germany: Springer, 2015, pp. 272–286.

[48] M. Kwiatkowska, G. Norman, and D. Parker, "Prism: Probabilistic symbolic model checker," in *TOOLS*. Berlin, Germany: Springer, 2002, pp. 200–204.

[49] I. Tkachev and A. Abate, "On infinite-horizon probabilistic properties and stochastic bisimulation functions," in *Proc IEEE Conf Decis. Control*, 2011, pp. 526–531.

[50] R. Durrett, *Probability: Theory and Examples*, vol. 49. Cambridge, U.K.: Cambridge Univ Press, 2019.

**Xiang Yin (Member, IEEE)** was born in Anhui, China, in 1991. He received the B.Eng. degree in electrical engineering from Zhejiang University, Hangzhou, China, in 2012, and the M.S. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2013 and 2017, respectively.

Since 2017, he has been with the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, where he is currently an Associate Professor. His research interests include formal methods, discrete-event systems, and cyber-physical systems.

Dr. Yin was the IEEE Conference on Decision and Control (CDC) Best Student Paper Award Finalist in 2016. He is serving as the Cochair of the IEEE CSS Technical Committee on Discrete Event Systems, an Associate Editor for the *Journal of Discrete Event Dynamic Systems: Theory & Applications*, and a Member of the IEEE CSS Conference Editorial Board.

**Dimos V. Dimarogonas (Fellow, IEEE)** was born in Athens, Greece. He received the diploma degree in electrical and computer engineering and the Ph.D. degree in mechanical engineering from the National Technical University of Athens, Athens, Greece, in 2001 and 2007, respectively.

Between 2007 and 2010, he held Postdoctoral positions with the Department of Automatic Control, KTH Royal Institute of Technology, Stockholm, Sweden, and the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA. He is currently a Professor with the Division of Decision and Control Systems, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology. His current research interests include multiagent systems, hybrid systems and control, robot navigation and manipulation, human–robot interaction, and networked control.

Dr. Dimarogonas serves as an Associate Editor for *Automatica* and a Senior Editor for IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS. He was the recipient of the ERC starting Grant in 2014, the ERC Consolidator Grant in 2019, and the Knut och Alice Wallenberg Academy Fellowship in 2015.

**Majid Zamani (Senior Member, IEEE)** received the B.Sc. degree in electrical engineering from the Isfahan University of Technology, Isfahan, Iran, in 2005, the M.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2007, and the M.A. degree in mathematics and the Ph.D. degree in electrical engineering from the University of California, Los Angeles, CA, USA, both in 2012.

Between September 2012 and December 2013, he was a Postdoctoral Researcher with the Delft Center for Systems and Control, Delft University of Technology, Delft, Netherlands. From 2013 to 2014, he was an Assistant Professor with the Design Engineering Department, Delft University of Technology. From 2014 to 2019, he was an Assistant Professor with the Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany. He is currently an Associate Professor with the Computer Science Department, University of Colorado Boulder, Boulder, CO, USA. He is also a guest Professor with the Computer Science Department, Ludwig Maximilian University of Munich, Munich, Germany. His research interests include verification and control of hybrid systems, embedded control software synthesis, networked control systems, and incremental properties of nonlinear control systems.

Dr. Zamani was the recipient of the George S. Axelby Outstanding Paper Award from the IEEE Control Systems Society in 2023, the NSF Career award in 2022, and the ERC Starting Grant and Proof of Concept Grant from the European Research Council in 2018 and 2023, respectively.

**Siyuan Liu** (Graduate Student Member, IEEE) received the B.E. degree in automation science and the M.E. degree in control engineering from Beihang University, Beijing, China, in 2014 and 2017, respectively, and the Ph.D. degree in electrical and computer engineering from the Technical University of Munich, Munich, Germany, in 2022.

She was a Research Assistant with the Hybrid Control Systems Lab, Ludwig-Maximilian University of Munich, Munich. In 2021, she was a visiting Researcher with the KTH Royal Institute of Technology, Stockholm, Sweden, where she is currently a Postdoctoral Research Fellow with the Division of Decision and Control Systems. Her current research interests include formal verification and control of hybrid systems, multi-agent systems, and compositional analysis of large-scale cyber-physical systems.