On Approximate Opacity of Cyber-Physical Systems

Xiang Yin, *Member, IEEE*, Majid Zamani, *Senior Member, IEEE*, and Siyuan Liu, *Student Member, IEEE*

***Abstract*—Opacity is an important information-flow security property in the analysis of cyber-physical systems. It captures the plausible deniability of the system's secret behavior in the presence of an intruder that may access the information flow. Existing works on opacity only consider nonmetric systems by assuming that the intruder can always distinguish between two different outputs precisely. In this article, we extend the concept of opacity to systems whose output sets are equipped with metrics. Such systems are widely used in the modeling of many real-world systems whose measurements are physical signals. A new concept called approximate opacity is proposed in order to quantitatively evaluate the security guarantee level with respect to the measurement precision of the intruder. Then, we propose a new simulation-type relation, called approximate opacity-preserving simulation relation, which characterizes how close two systems are in terms of the satisfaction of approximate opacity. This allows us to verify approximate opacity for large-scale, or even infinite, systems using their abstractions. We also discuss how to construct approximate opacity-preserving symbolic models for a class of discrete-time control systems. Our results extend the definitions and analysis techniques for opacity from nonmetric systems to metric systems.**

***Index Terms*—Approximate simulation relations, finite abstractions, opacity, symbolic models.**

Manuscript received November 12, 2019; revised April 4, 2020; accepted May 25, 2020. Date of publication June 1, 2020; date of current version March 29, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61803259 and Grant 61833012, in part by the Shanghai Jiao Tong University Scientific and Technological Innovation Funds, in part by the German Research Foundation under Grant ZA 873/1-1, and in part by the H2020 ERC Starting Grant AutoCPS under Grant 804639. Recommended by Associate Editor K. Cai. *(Corresponding author: Xiang Yin.)*

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

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

Siyuan Liu is with the Department of Electrical and Computer Engineering, Technical University of Munich 80333, Munich, Germany (e-mail: sy.liu@tum.de).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2020.2998733

I. Introduction

A. Motivations

CYBER-physical systems (CPSs) are complex systems resulting from tight interactions of dynamical systems and computational devices. Such systems are generally very complex posing both continuous and discrete behaviors, which makes the verification and design of such systems significantly challenging. In particular, components in CPSs are usually connected via communication networks in order to acquire and exchange information so that some global functionality of the system can be achieved. However, this also brings new challenges for the verification and design of CPSs, since the communication between system components may release information that might compromise the security of the system. Therefore, how to analyze and enforce security for CPS is becoming an increasingly important issue and has drawn considerable attention in the literature in the past few years [1], [2].

In this article, we investigate an important information-flow security property called *opacity*. Roughly speaking, opacity is a confidentiality property that captures whether or not the "secret" of the system can be revealed to an intruder that can infer the system's actual behavior based on the information flow. A system is said to be opaque if it always has the plausible deniability for any of its secret behavior. The concept of opacity was originally proposed in the computer science literature as a unified notion for several security properties [3], [4]. Since then, opacity has been studied more extensively in the context of discrete-event systems (DESs), an important class of event-driven dynamical systems with discrete state spaces. For example, in [5]–[7], several state-based notions of opacity were proposed, which include current-state opacity, initial-state opacity, K-step opacity, and infinite-step opacity. In [8], Lin proposed two language-based opacity notions called strong opacity and weak opacity and investigated their relationships with some other properties. In [9], transformation algorithms among different notions of opacity were proposed. The aforementioned works mainly consider DES modeled by finite-state automata. More recently, the definitions and verification algorithms for different notions of opacity have been extended to other classes of (discrete) systems, including Petri nets [10]–[13], stochastic systems [14]–[16], recursive tile systems [17], and pushdown systems [18]. The interested readers are referred to recent surveys [19], [20] for more references and recent developments on this active research area.

Since opacity is an information-flow property, its definition strictly depends on the information model of the system. Most of the existing works in the literature formulate opacity by adopting the event-based observation model, i.e., some events of the system (either transition labels or state labels) are observable or distinguishable while some are not. This essentially assumes that the output of the system is symbolic in the sense that we can precisely distinguish between two outputs with different labels. Hereafter, we will also refer to opacity under this setting as *exact opacity*. Exact opacity is very meaningful for systems whose output sets are nonmetric, e.g., discrete systems whose outputs are logic events. However, for many real-world applications whose outputs are physical signals, instead of just saying that two events are distinguishable or indistinguishable, we may have a measurement to quantitatively evaluate how close two outputs are. Such systems are referred to as *metric systems*, where the output sets are equipped with appropriate metrics. For metric systems, if two signals are very close to each other, then it will be very hard to distinguish them unambiguously because of the measurement precision or potential measurement noises. A typical example of this scenario is linear or nonlinear discrete-time control systems with continuous state spaces and continuous output mappings. Therefore, existing definitions of opacity are too strong for metric systems, since they implicitly assume that the intruder can always distinguish between two output signals even when they are arbitrarily close to each other, which is not practical.

B. Our Contributions

In this article, we propose a new concept called *approximate opacity* that is more applicable to metric systems. In particular, we treat two outputs as "indistinguishable" outputs if their distance is smaller than a given threshold parameter $\delta \geq 0$. We consider three basic types of opacity: initial-state opacity, current-state opacity, and infinite-step opacity. For example, δ-approximate initial-state opacity requires that, for any state run starting from a secret state, there exists another state run starting from a nonsecret state, such that their corresponding output runs are δ-close to each other. Intuitively, δ-approximate initial-state opacity says that the intruder can never determine that the system is initiated from a secret state if it does not have an enough measurement precision, which is captured by parameter δ. In other words, instead of requiring that the system is exactly opaque, our new definitions essentially provide relaxed versions of opacity with a quantitative security guarantee level with respect to the measurement precision of the intruder.

For systems whose state spaces are very large or even infinite, it is desirable to construct abstract models that preserve opacity, to some extent, for the purpose of verification. To this end, we propose the concept of ε-approximate opacity-preserving simulation relation. We show that if there is an ε-approximate opacity-preserving simulation relation from system S_a to system S_b, then S_b being $(\delta - 2\varepsilon)$-approximate opaque implies that S_a is δ-approximate opaque. In particular, for a class of incrementally input-to-state stable discrete-time control systems with possibly infinite state spaces, we propose an effective approach to construct symbolic models (a.k.a. finite abstractions) that approximately simulate the original systems in the sense of opacity preserving and *vice versa*. The resulting symbolic model is finite if the state space of the original continuous system is within a bounded region. Therefore, the proposed abstraction technique together with the verification algorithm for the finite case provides a sound way for verifying opacity of discrete-time control systems with continuous state spaces.

The contributions of this article are summarized as follows.

1) New notions of δ-approximate opacity are proposed to quantitatively characterize the issue regarding the measurement precision of the intruder.
2) Effective algorithms are provided to verify different notions of approximate opacity.
3) New simulation relations termed as ε-approximate opacity-preserving simulation relations are proposed to characterize how close two systems are in terms of the satisfaction of approximate opacity.
4) For a class of discrete-time control systems, we show how to construct symbolic models that preserve opacity with given *a priori* precision.

C. Related Works

Our article is closely related to several works in the literature. First, several different approaches have been proposed in the literature to evaluate opacity more quantitatively rather than requiring that the system is opaque exactly [14], [21]–[23]. For example, in [22], the authors adopt the Jensen–Shannon divergence as the measurement to quantify secrecy loss. In [14], [21], and [23], stochastic DES models are used to study the probabilistic measurement of opacity. These approaches essentially aim to analyze how opaque a single system is, e.g., the probability of being opaque. However, they neither consider how close two systems are in terms of being opaque nor consider under what observation precision level, we can guarantee opacity.

There are also attempts in the literature that extend opacity from discrete systems to continuous systems. For example, in the recent results in [24]–[26], the authors extended the notion of opacity to (switched) linear systems. However, their definition of opacity is more related to an output reachability property rather than an information-flow property. Moreover, their formulation is mostly based on the setting of exact opacity, i.e., we can always distinguish between two different outputs precisely no matter how close they are, In [24], the authors mentioned the direction of using output metric to quantify opacity, and a property called strong ϵ-$\mathcal{K}$-initial-state opacity was proposed, which is closely related to our notions. However, no systematic study, e.g., verification and abstraction as we consider in this article, was provided for this property.

Regarding the techniques used in this article, first, our algorithms for the verification of approximate notions of opacity are motivated by the verification algorithms for exact opacity studied in [5] and [27]. In particular, we use the idea of constructing a new system, called the state estimator, that tracks all possible states consistent with the observation. However, our construction of state estimator is not exactly the same as the

existing one, as additional state information is needed in order to handle the issue of approximation.

Abstraction-based techniques have also been investigated in the literature for the verification and synthesis of opacity (see, e.g., [28]–[32]). In particular, in our recent work [28], we propose several notions of opacity-preserving (bi)simulation relations. However, these relations only preserve exact opacity for nonmetric systems. Our new relations extend the relations in [28] to metric systems by taking into account how close two systems are. Such an extension is motivated by the definition of approximate (bi)simulation relation originally proposed in [33]. However, the original definition of approximate (bi)simulation relation does not necessarily preserve approximate opacity. Constructing symbolic models for control systems is also an active research area (see, e.g., [34]–[37]). However, most of the existing works on the construction of symbolic models only consider the dynamics of the systems and are not taking into account the opacity property. In our approach, we need to consider both the dynamic and the secret of the system while constructing the symbolic model and guarantee the preservation of approximate opacity across related systems.

A related notion called differential privacy was introduced in [38] for database systems and has attracted significant attention in the past few years [39]–[41]. In particular, Jones *et al.* [40] extend the original notion of differential privacy to symbolic systems. Differential privacy requires that any two adjacent data should produce indistinguishable outputs in the probability sense. However, the essence of opacity is a confidentiality property that captures the plausible deniability of the system's secret behavior, while differential privacy captures whether or not any sensitive data can be learned under some side information. These two properties are incomparable in general. Note that there are also probabilistic versions of opacity studied in the literature for systems modeled as Markov chains [14], [21]–[23]. In those studies, the essence of probabilistic opacity is still plausible deniability but with a quantitative measure; the output at each state is still nonprobabilistic.

Finally, approximate notions of two related properties called diagnosability and predictability have recently been investigated in [42] and [43]. Their setting is very similar ours as we both consider a measurement uncertainty threshold. However, diagnosability and predictability can be preserved by the standard approximate simulation relation. We show that the standard approximate simulation relation does not preserve opacity. Therefore, the proposed approximate opacity-preserving simulation relation is different from the standard approximate simulation relation in the literature.

D. Organization

The rest of this article is organized as follows. In Section II, we first introduce some necessary preliminaries. Then, we propose the concept of approximate opacity in Section III. The verification procedures for approximate opacity are provided in Section IV. In Section V, approximate opacity-preserving simulation relations are proposed, and their properties are also discussed. In Section VI, we describe how to construct approximate opacity-preserving symbolic models for incrementally stable discrete-time control systems with continuous state spaces. Finally, we conclude this article by Section VII. Preliminary and partial version of this article is presented as an extended abstract in [44].

II. Preliminaries

A. Notation

The symbols $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{R}^+$, and $\mathbb{R}_0^+$ denote the set of natural, nonnegative integer, integer, real, positive, and nonnegative real numbers, respectively. Given a vector $x \in \mathbb{R}^n$, we denote by x_i the ith element of x, and by $\|x\|$ the infinity norm of x.

The closed ball centered at $u \in \mathbb{R}^m$ with radius λ is defined by $\mathcal{B}_\lambda(u) = \{v \in \mathbb{R}^m \mid \|u - v\| \leq \lambda\}$. We denote the closed ball centered at the origin in $\mathbb{R}^n$ and with radius λ by $\mathcal{B}_\lambda$. A set $B \subseteq \mathbb{R}^m$ is called a *box* if $B = \prod_{i=1}^m [c_i, d_i]$, where $c_i, d_i \in \mathbb{R}$ with $c_i < d_i$ for each $i \in \{1, \ldots, m\}$. The *span* of a box B is defined as $\text{span}(B) = \min\{|d_i - c_i| \mid i = 1, \ldots, m\}$. For a box $B \subseteq \mathbb{R}^m$ and $\mu \leq \text{span}(B)$, define the μ-approximation $[B]_\mu = [\mathbb{R}^m]_\mu \cap B$, where $[\mathbb{R}^m]_\mu = \{a \in \mathbb{R}^m \mid a_i = k_i\mu, k_i \in \mathbb{Z}, i = 1, \ldots, m\}$. Remark that $[B]_\mu \neq \varnothing$ for any $\mu \leq \text{span}(B)$. Geometrically, for any $\mu \in \mathbb{R}^+$ with $\mu \leq \text{span}(B)$ and $\lambda \geq \mu$, the collection of sets $\{\mathcal{B}_\lambda(p)\}_{p \in [B]_\mu}$ is a finite covering of B, i.e., $B \subseteq \bigcup_{p \in [B]_\mu} \mathcal{B}_\lambda(p)$. We extend the notions of span and *approximation* to finite unions of boxes as follows. Let $A = \bigcup_{j=1}^M A_j$, where each A_j is a box. Define $\text{span}(A) = \min\{\text{span}(A_j) \mid j = 1, \ldots, M\}$, and for any $\mu \leq \text{span}(A)$, define $[A]_\mu = \bigcup_{j=1}^M [A_j]_\mu$. The Minkowski sum of two sets $P, Q \subseteq \mathbb{R}^n$ is defined by $P \oplus Q = \{x \in \mathbb{R}^n | \exists_{p \in P, q \in Q}, x = p + q\}$. Given a set $\mathbb{S} \subseteq \mathbb{R}^n$ and a constant $\theta \in \mathbb{R}_{\geq 0}$, we define a new set $\mathbb{S}^\theta = \mathbb{S} \oplus \mathcal{B}_\theta$ as the inflated version of the set $\mathbb{S}$.

Given a function $f : \mathbb{N}_0^+ \to \mathbb{R}^n$, the (essential) supremum of f is denoted by $\|f\|_\infty := \text{(ess)sup}\{\|f(k)\|, k \geq 0\}$. A continuous function $\gamma : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and $\gamma(0) = 0$; γ is said to belong to class $\mathcal{K}_\infty$ if $\gamma \in \mathcal{K}$ and $\gamma(r) \to \infty$ as $r \to \infty$. A continuous function $\beta : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is said to belong to class $\mathcal{KL}$ if, for each fixed s, the map $\beta(r, s)$ belongs to class $\mathcal{K}$ with respect to r and, for each fixed nonzero r, the map $\beta(r, s)$ is decreasing with respect to s and $\beta(r, s) \to 0$ as $s \to \infty$. We identify a relation $R \subseteq A \times B$ with the map $R : A \to 2^B$ defined by $b \in R(a)$ iff $(a, b) \in R$. Given a relation $R \subseteq A \times B$, R^{-1} denotes the inverse relation defined by $R^{-1} = \{(b, a) \in B \times A : (a, b) \in R\}$.

B. System Model

In this article, we employ a notion of "*system*" introduced in [45] as the underlying model of CPS describing both continuous-space and finite control systems.

Definition II.1: A system S is a tuple

$$S = (X, X_0, U, \longrightarrow, Y, H) \tag{1}$$

where we have the following.

1) X is a (possibly infinite) set of states.
2) $X_0 \subseteq X$ is a (possibly infinite) set of initial states.
3) U is a (possibly infinite) set of inputs.

4) $\longrightarrow \subseteq X \times U \times X$ is a transition relation.
5) Y is a set of outputs.
6) $H : X \to Y$ is an output map.

A transition $(x, u, x') \in \longrightarrow$ is also denoted by $x \xrightarrow{u} x'$. For a transition $x \xrightarrow{u} x'$, state x' is called a u-successor, or simply a successor, of state x; state x is called a u-predecessor, or simply a predecessor, of state x'. We denote by $\mathbf{Post}_u(x)$ the set of all u-successors of state x and by $\mathbf{Pre}_u(x)$ the set of all u-predecessors of state x. For a set of states $q \in 2^X$, we define $\mathbf{Post}_u(q) = \cup_{x \in q} \mathbf{Post}_u(x)$ and $\mathbf{Pre}_u(q) = \cup_{x \in q} \mathbf{Pre}_u(x)$. A system S is said to be:

1) *metric*, if the output set Y is equipped with a metric $\mathbf{d} : Y \times Y \to \mathbb{R}_0^+$;
2) *finite* (or *symbolic*), if X and U are finite sets;
3) *deterministic*, if for any state $x \in X$ and any input $u \in U$, $|\mathbf{Post}_u(x)| \leq 1$ and *nondeterministic* otherwise.

Given a system $S = (X, X_0, U, \longrightarrow, Y, H)$ and any initial state $x_0 \in X_0$, a finite state run generated from x_0 is a finite sequence of transitions:

$$x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_{n-1}} x_{n-1} \xrightarrow{u_n} x_n \tag{2}$$

such that $x_i \xrightarrow{u_{i+1}} x_{i+1}$ for all $0 \leq i < n$. A finite output run is a sequence $y_0 y_1 \ldots y_n$ such that there exists a finite state run of the form (2) with $y_i = H(x_i)$, for $i = 0, \ldots, n$.

III. Exact and Approximate Opacity

In this section, we first review the notion of exact opacity. Then, we introduce the notion of approximate opacity.

A. Exact Opacity

In many applications, systems may have some "secrets" that do not want to be revealed to intruders that are potentially malicious. In this article, we adopt a state-based formulation of secrets. Specifically, we assume that $X_S \subseteq X$ is a set of *secret states*. Hereafter, we will always consider systems with secret states, and we write a system $S = (X, X_0, U, \longrightarrow, Y, H)$ with secret states X_S by a new tuple $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$.

In order to characterize whether or not a system is secure, the concept of opacity was proposed in the literature. We review three basic notions of opacity [9] as follows.

Definition III.1: Consider a system $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$. System S is said to be:

1) *initial-state opaque* if for any $x_0 \in X_0 \cap X_S$ and finite state run $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$, there exist $x_0' \in X_0 \setminus X_S$ and a finite state run $x_0' \xrightarrow{u_1'} x_1' \xrightarrow{u_2'} \cdots \xrightarrow{u_n'} x_n'$ such that $H(x_i) = H(x_i')$ for any $i = 0, 1, \ldots, n$;
2) *current-state opaque* if for any $x_0 \in X_0$ and finite state run $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$ such that $x_n \in X_S$, there exist $x_0' \in X_0$ and finite state run $x_0' \xrightarrow{u_1'} x_1' \xrightarrow{u_2'} \cdots \xrightarrow{u_n'} x_n'$ such that $x_n' \in X \setminus X_S$ and $H(x_i) = H(x_i')$ for any $i = 0, 1, \ldots, n$;
3) *infinite-step opaque* if for any $x_0 \in X_0$, any finite state run $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$ and any $k \in \{0, \ldots, n\}$, $x_k \in X_S$ implies that there exist $x_0' \in X_0$ and a finite state run $x_0' \xrightarrow{u_1'} x_1' \xrightarrow{u_2'} \cdots \xrightarrow{u_n'} x_n'$ such that $x_k' \in X \setminus X_S$ and $H(x_i) = H(x_i')$ for any $i = 0, 1, \ldots, n$.

The intuitions of the above definitions are as follows. Suppose that the output run of the system can be observed by a passive intruder that may use this information to infer the secret of the system. Then, initial-state opacity requires that the intruder should never know for sure that the system is initiated from a secret state no matter what output run is generated. Similarly, current-state opacity says that the intruder should never know for sure that the system is currently at a secret state no matter what output run is generated. Infinite-step opacity is stronger than both initial-state opacity and current-state opacity as it requires that the intruder should never know that the system is/was at a secret state for any specific instant k. For any system $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$, we assume without loss of generality that $\forall x_0 \in X_0 : \{x \in X_0 : H(x) = H(x_0)\} \not\subseteq X_S$. This assumption essentially requires that the secret of the system cannot be revealed initially; otherwise, the system is not opaque trivially.

Remark III.2: Definition III.1 implicitly considers the following model of the intruder: 1) the intruder knows the model of the system; and 2) it can only observe the output trajectory of the system. Therefore, the intruder essentially wants to use the output trajectory observed online and the knowledge of the system model to infer the internal behavior/state of the system. Note that, in our setting, the input information is assumed to be internal and the intruder does not know which input the system takes. This setting can be easily relaxed, and all results in this article can be extended to the case where both input and output information are available by the intruder. For example, we can simply refine the model of the system such that the output space of the refined system is a pair and the input leading to a state is also encoded in the output of this state.

Remark III.3: Our definition of infinite-step opacity requires that the intruder should never know for sure that the system is/was at a secret state *for any specific instant*. In some cases, the intruder may know that the system must have visited a secret state, although it cannot tell the precise instant. Such a requirement can be captured by the notion of strong (or trajectory-based) infinite-step opacity (see, e.g., [6, Remark 5]). This definition is stronger than ours, and which one to use is dependent on the applications. However, strong infinite-step opacity can be transformed to current-state opacity by augmenting the state space to encode whether a secret state has been visited or not.

B. Approximate Opacity

Note that Definition III.1 requires that for any secret behavior, there exists a nonsecret behavior such that they generate exactly the same output. Therefore, we will also refer to these definitions as *exact opacity*. Exact opacity essentially assumes that the intruder or the observer can always measure each output or distinguish between two different outputs precisely. This setting is reasonable for nonmetric systems where outputs are symbols or events. However, for metric systems, e.g., when the outputs are physical signals, this setting may be too restrictive. In particular,

owing to the imperfect measurement precision, which is almost the case for all physical systems, it is very difficult to distinguish between two observations if their difference is very small. Therefore, exact opacity may be too strong for metric systems, and it will be useful to define a weak and "robust" version of opacity by characterizing under which measurement precision the system is opaque. To this end, we define new notions of opacity called *approximate opacity* for metric systems.

Definition III.4: Let $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. System S is said to be:

1) *δ-approximate initial-state opaque* if for any $x_0 \in X_0 \cap X_S$ and finite state run $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$, there exist $x_0' \in X_0 \setminus X_S$ and a finite state run $x_0' \xrightarrow{u_1'} x_1' \xrightarrow{u_2'} \cdots \xrightarrow{u_n'} x_n'$ such that
$$\max_{i\in\{0,\ldots,n\}} \mathbf{d}(H(x_i), H(x_i')) \leq \delta$$

2) *δ-approximate current-state opaque* if for any $x_0 \in X_0$ and finite state run $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$ such that $x_n \in X_S$, there exist $x_0' \in X_0$ and finite state run $x_0' \xrightarrow{u_1} x_1' \xrightarrow{u_2'} \cdots \xrightarrow{u_n'} x_n'$ such that $x_n' \in X \setminus X_S$ and
$$\max_{i\in\{0,\ldots,n\}} \mathbf{d}(H(x_i), H(x_i')) \leq \delta$$

3) *δ-approximate infinite-step opaque* if for any $x_0 \in X_0$, any finite state run $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$ and any $k \in \{0, \ldots, n\}, x_k \in X_S$ implies that there exist $x_0' \in X_0$ and a finite state run $x_0' \xrightarrow{u_1} x_1' \xrightarrow{u_2'} \cdots \xrightarrow{u_n'} x_n'$ such that $x_k' \in X \setminus X_S$ and
$$\max_{i\in\{0,\ldots,n\}} \mathbf{d}(H(x_i), H(x_i')) \leq \delta.$$

The notions of δ-approximate initial-state, current-state opacity, and infinite-step opacity are very similar to their exact counterparts. The main difference is how we treat two outputs as indistinguishable outputs. Specifically, same as the exact case, we still assume that the intruder knows the system model and the output trajectory generated. However, we further assume that the intruder may not be able to distinguish an output trajectory from other δ-closed trajectories confidentially. Intuitively, the approximate version of opacity can be interpreted as "*the secret of the system cannot be revealed to an intruder that does not have an enough measurement precision related to parameter δ*." In other words, instead of providing an exact security guarantee, approximate opacity provides a relaxed and quantitative security guarantee with respect to the measurement precision of the intruder. Therefore, the value δ can be interpreted as either the measurement imprecision of the intruder or the security level the system can guarantee, i.e., under how powerful intruder the system is still secure. Clearly, when $\delta = 0$, each notion of δ-approximate opacity reduces to its exact version. Similar to the exact case, hereafter, we assume without loss of generality that
$$\forall x_0 \in X_0 : \{x \in X_0 : \mathbf{d}(H(x_0), H(x)) \leq \delta\} \not\subseteq X_S$$
for any system $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$. This assumption can be easily checked, and its nonsatisfaction means that δ-approximate initial-state opacity, δ-approximate current-state opacity, and δ-approximate infinite-step opacity are all violated trivially.

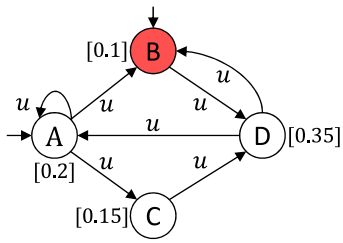


Fig. 1. Example for approximate opacity, where states marked by red denote secret states, states marked by input arrows denote initial states, and the output map is specified by the value associated to each state.

We illustrate exact opacity and approximate opacity by the following example.

Example III.5: Consider system $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ depicted in Fig. 1, where $X = \{A, B, C, D\}, X_0 = \{A, B\}, X_S = \{B\}, U = \{u\}, H = \{0.1, 0.15, 0.2, 0.35\} \subseteq \mathbb{R}$, and the output map is specified by the value associated with each state. Clearly, none of exact initial-state opacity, exact current-state opacity, and exact infinite-step opacity is satisfied, since we know immediately that the system is at secret state B when value 0.1 is observed.

Now, let us assume that the output set Y is equipped with metric $\mathbf{d}$ defined by $\mathbf{d}(y_1, y_2) = |y_1 - y_2|$. We claim that S is not 0.05-approximate current-state opaque. For example, let us consider finite run $B \xrightarrow{u} D \xrightarrow{u} B$ that generates output run [0.1][0.35][0.1]. However, there does not exist a finite run leading to a nonsecret state whose output run is 0.05-close to the above output run. To see this, in order to match the above output run, we must consider a run starting from state B, since for the initial state A, we have $\mathbf{d}(H(A), H(B)) = 0.1 \geq 0.05$, and the next state reached can only be D. From state D, we can reach states A and B, but $\mathbf{d}(H(A), 0.1) = 0.1 \geq 0.05 =: \delta$. Therefore, the only finite run that approximately matches the above output will end up with secret state B, i.e., we know unambiguously that the system is currently at a secret state even when we cannot measure the output precisely. In contrast, one can check that the system is 0.1-approximate current-state opaque.

Similarly, system S is not 0.1-approximate initial-state opaque, since for output run [0.1][0.35] starting from the secret state B, there is no run starting from a nonsecret initial state that can approximately match it. One can also check that the system is δ-approximate initial-state opaque only when $\delta \geq 0.15$. We will provide formal procedures for verifying approximate opacity later.

Remark III.6: Let $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ be a metric system. If the output map H is identity, i.e. $H(x) = x, \forall x \in X$, then S is trivially not exactly opaque as in Definition III.1 since we know the exact state of the system directly. However, this is not the case for the approximate notions of opacity as in Definition III.4, since the distance between a secret state and a nonsecret state can be very small even if their values are not exactly the same.

IV. Verification of Approximate Opacity for Finite Systems

In this section, we show how to verify approximate opacity for finite systems. This will also provide the basis for the verification of approximate opacity for infinite systems.

A. Verification of Approximate Initial-State Opacity

In order to verify δ-approximate initial-state opacity, we construct a new system called the *δ-approximate initial-state estimator* defined as follows.

Definition IV.1: Let $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. The δ-approximate initial-state estimator is a system (without outputs)

$$S_I = (X_I, X_{I0}, U, \underset{I}{\longrightarrow})$$

where we have the following.

1) $X_I \subseteq X \times 2^X$ is the set of states.
2) $X_{I0} = \{(x, q) \in X \times 2^X : x' \in q \Leftrightarrow \mathbf{d}(H(x), H(x')) \leq \delta\}$ is the set of initial states.
3) U is the set of inputs, which is the same as the one in S.
4) $\underset{I}{\longrightarrow} \subseteq X_I \times U \times X_I$ is the transition function defined by: for any $(x, q), (x', q') \in X \times 2^X$ and $u \in U$, $(x, q) \underset{I}{\overset{u}{\longrightarrow}} (x', q')$ if:
   1) $(x', u, x) \in \longrightarrow$;
   2) $q' = \cup_{\hat{u} \in U} \mathbf{Pre}_{\hat{u}}(q) \cap \{x'' \in X : \mathbf{d}(H(x'), H(x'')) \leq \delta\}$.

For the sake of simplicity, we only consider the part of S_I that is reachable from initial states.

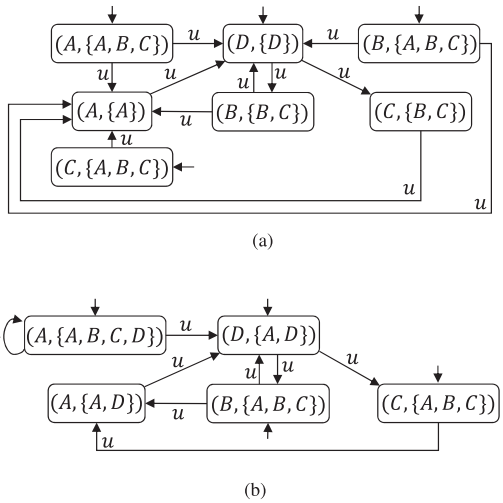
Intuitively, the δ-approximate initial-state estimator works as follows. Each initial state of S_I is a pair consisting of a system state and its δ-closed states; we consider all each pairs as the set of initial states. Then, from each state, we track *backwards* states that are consistent with the output information recursively. Our construction is motivated by the reversed-automaton-based initial-state estimator proposed in [9] but with the following differences. First, the way we defined information consistency is different. Here, we treat states whose outputs are δ-close to each other as consistent states. Moreover, the structure in [9] only requires a state space of 2^X, while our state space is $X \times 2^X$. The additional first component can be understood as the "reference trajectory" that is used to determine what is "δ-close" at each instant. We use the following result to show the main property of S_I.

Proposition IV.2: Let $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. Let $S_I = (X_I, X_{I0}, U, \underset{I}{\longrightarrow})$ be its δ-approximate initial-state estimator. Then, for any $(x_0, q_0) \in X_{I0}$ and any finite run

$$(x_0, q_0) \underset{I}{\overset{u_1}{\longrightarrow}} (x_1, q_1) \underset{I}{\overset{u_2}{\longrightarrow}} \cdots \underset{I}{\overset{u_n}{\longrightarrow}} (x_n, q_n)$$

we have

i) $x_n \overset{u_n}{\longrightarrow} x_{n-1} \overset{u_{n-1}}{\longrightarrow} \cdots \overset{u_1}{\longrightarrow} x_0$;



(a)

(b)

Fig. 2. Examples of δ-approximate initial-state estimators. (a) S_I when $\delta = 0.1$. (b) S_I when $\delta = 0.15$.

ii) $q_n = \{x_0' \in X : \exists x_0' \overset{u_n'}{\longrightarrow} x_1' \overset{u_{n-1}'}{\longrightarrow} \cdots \overset{u_1'}{\longrightarrow} x_n' \text{ s.t. } \max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x_{n-i}')) \leq \delta\}$.

Proof: See the Appendix. ■

The next theorem provides one of the main results of this section on the verification of δ-approximate initial-state opacity of finite metric systems.

Theorem IV.3: Let $S = (X, X_0, X_S, U, \underset{I}{\longrightarrow}, Y, H)$ be a finite metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. Let $S_I = (X_I, X_{I0}, U, \underset{I}{\longrightarrow})$ be its δ-approximate initial-state estimator. Then, S is δ-approximate initial-state opaque if and only if

$$\forall (x, q) \in X_I : x \in X_0 \cap X_S \Rightarrow q \cap X_0 \not\subseteq X_S. \tag{3}$$

Proof: See the Appendix. ■

We illustrate how to verify δ-approximate initial-state opacity by the following example.

Example IV.4: Let us still consider system S shown in Fig. 1. The δ-approximate initial-state estimator S_I when $\delta = 0.1$ is shown in Fig. 2(a). For example, for the initial state $(D, \{D\})$, we have $(D, \{D\}) \underset{I}{\overset{u}{\longrightarrow}} (B, \{B, C\})$ since $B \overset{u}{\longrightarrow} D$ and $\{B, C\} = \mathbf{Pre}_u(\{D\}) \cap \{x \in X : \mathbf{d}(0.1, H(x)) \leq 0.1\} = \{B, C\} \cap \{A, B, C\}$. However, for state $(B, \{B, C\}) \in X_I$, we have $B \in X_0 \cap X_S$ and $\{B, C\} \cap X_0 = \{B\} \subseteq X_S$. Therefore, by Theorem IV.3, we know that the system is not 0.1-approximate initial-state opaque. Similarly, we can also construct S_I for the case of $\delta = 0.15$, which is shown in Fig. 2(b). For state $(B, \{A, B, C\}) \in X_I$, which is the only state whose first component is in $X_0 \cap X_S$, we have $\{A, B, C\} \cap X_0 = \{A, B\} \not\subseteq X_S$. By Theorem IV.3, we know that the system is 0.15-approximate initial-state opaque.

B. Verification of Approximate Current-State Opacity

In order to verify δ-approximate current-state opacity, we also need to construct a new system called the *δ-approximate current-state estimator* defined as follows.

Definition IV.5: Let $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. The δ-approximate current-state estimator is a system (without outputs)

$$S_C = (X_C, X_{C0}, U, \underset{C}{\longrightarrow})$$

where we have the following.

1) $X_C \subseteq X \times 2^X$ is the set of states.
2) $X_{C0} = \{(x, q) \in X_0 \times 2^{X_0} : x' \in q \Leftrightarrow \mathbf{d}(H(x), H(x')) \leq \delta\}$ is the set of initial states.
3) U is the set of inputs, which is the same as the one in S.
4) $\underset{C}{\longrightarrow} \subseteq X_C \times U \times X_C$ is the transition function defined by: for any $(x, q), (x', q') \in X \times 2^X$ and $u \in U$, $(x, q) \underset{C}{\overset{u}{\longrightarrow}} (x', q')$ if:
   1) $(x, u, x') \in \longrightarrow$;
   2) $q' = \cup_{\hat{u} \in U} \mathbf{Post}_{\hat{u}}(x) \cap \{x'' \in X : \mathbf{d}(H(x'), H(x'')) \leq \delta\}$.

For the sake of simplicity, we only consider the part of S_C that is reachable from initial states.

The construction of S_C is similar to S_I. However, we need to track all forward runs from each pair of initial state and its information-consistent states. Still, we need the first component as the "reference state" to determine what are "δ-close" states. We use the following result to state the main properties of S_C.

Proposition IV.6: Let $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. Let $S_C = (X_C, X_{C0}, U, \underset{C}{\longrightarrow})$ be its δ-approximate current-state estimator. Then, for any $(x_0, q_0) \in X_{C0}$ and any finite run

$$(x_0, q_0) \underset{C}{\overset{u_1}{\longrightarrow}} (x_1, q_1) \underset{C}{\overset{u_2}{\longrightarrow}} \cdots \underset{C}{\overset{u_n}{\longrightarrow}} (x_n, q_n)$$

we have

i) $x_0 \overset{u_1}{\longrightarrow} x_1 \overset{u_2}{\longrightarrow} \cdots \overset{u_n}{\longrightarrow} x_n$;

ii) $q_n = \{x'_n \in X : \exists x'_0 \in X_0, \exists x'_0 \overset{u'_1}{\longrightarrow} x'_1 \overset{u'_2}{\longrightarrow} \cdots \overset{u'_n}{\longrightarrow} x'_n \text{ s.t. } \max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(H(x_i),\ H(x'_i)) \leq \delta\}$.

Proof: See the Appendix. ■

Now, we show the second main result of this section by providing a verification scheme for δ-approximate current-state opacity of finite metric systems.

Theorem IV.7: Let $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. Let $S_C = (X_C, X_{C0}, U, \underset{C}{\longrightarrow})$ be its δ-approximate current-state estimator. Then, S is δ-approximate current-state opaque if and only if

$$\forall (x, q) \in X_C : q \not\subseteq X_S. \tag{4}$$

Proof: See the Appendix. ■

C. Verification of Approximate Infinite-Step Opacity

Finally, we can combine the δ-approximate initial-state estimator S_I and the δ-approximate current-state estimator S_C to verify δ-approximate infinite-step opacity of finite metric systems. The verification scheme is provided by the following theorem.

Theorem IV.8: Let $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ be a finite metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. Let $S_I = (X_I, X_{I0}, U, \underset{I}{\longrightarrow})$ and $S_C = (X_C, X_{C0}, U, \underset{C}{\longrightarrow})$ be its δ-approximate initial-state estimator and δ-approximate current-state estimator, respectively. Then, S is δ-approximate infinite-step opaque if and only if

$$\forall (x, q) \in X_I, (x', q') \in X_C : x = x' \in X_S \Rightarrow q \cap q' \not\subseteq X_S. \tag{5}$$

Proof: See the Appendix. ■

Remark IV.9: We conclude this section by discussing the complexity of verifying approximate opacity. Let $S = (X, X_0, X_S, U, \longrightarrow, Y, H)$ be a finite metric system. The complexity of the verification algorithms for both approximate initial-state and current-state opacity is $O(|U| \times |X| \times 2^{|X|})$, which is the size of S_I or S_C. For approximate infinite-step opacity, we need to construct both S_I and S_C and compare each pair of states in S_I and S_C. Therefore, the complexity for verifying approximate infinite-step opacity using Theorem IV.8 is $O(|U| \times |X|^2 \times 4^{|X|})$. It is worth noting that the complexity of verifying exact opacity as in Definition III.1 is already known to be PSPACE complete [46]. Using a similar reduction, we can conclude that the complexity of verifying approximate opacity as in Definition III.4 is also PSPACE complete for $\delta > 0$. Finally, we note that the exponential complexity essentially comes from the subset construction to handle information uncertainty. In practice, the subset construction usually results in a quite small structure (see, e.g., [47] for detailed empirical studies on this issue).

V. Approximate Simulation Relations for Opacity

In the previous sections, we have introduced notions of approximate opacity and their verification procedures. However, when the system is very large or even infinite, verifying opacity based on the original system is not efficient or not even possible. Therefore, it will be beneficial if we can verify opacity based on an "equivalent" smaller or symbolic system. To this end, in this section, we study under what conditions two systems are equivalent and in what sense. Specifically, we introduce new notions of approximate opacity-preserving simulation relations, inspired by the one in [33]. The newly proposed simulation relations will provide the basis for abstraction-based verification of approximate opacity.

A. Approximate Initial-State Opacity-Preserving Simulation Relation

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

Definition V.1 (Approximate initial-state opacity-preserving simulation relation): Consider two metric systems $S_a = (X_a, X_{a0}, X_{aS}, U_a, \underset{a}{\longrightarrow}, Y_a, H_a)$ and $S_b = (X_b, X_{b0}, X_{bS}, U_b, \underset{b}{\longrightarrow}, Y_b, H_b)$ with the same output sets

$Y_a = Y_b$ and metric $\mathbf{d}$. For $\varepsilon \in \mathbb{R}_0^+$, a relation $R \subseteq X_a \times X_b$ is called an ε-approximate initial-state opacity-preserving simulation relation (ε-InitSOP simulation relation) from S_a to S_b if:

1)
   a) $\forall x_{a0} \in X_{a0} \cap X_{aS}, \exists x_{b0} \in X_{b0} \cap X_{bS}$: $(x_{a0}, x_{b0}) \in R$;
   b) $\forall x_{b0} \in X_{b0} \setminus X_{bS}, \exists x_{a0} \in X_{a0} \setminus X_{aS}$: $(x_{a0}, x_{b0}) \in R$;
2) $\forall (x_a, x_b) \in R : \mathbf{d}(H_a(x_a), H_b(x_b)) \leq \varepsilon$;
3) for any $(x_a, x_b) \in R$, we have
   a) $\forall x_a \xrightarrow[a]{u_a} x_a', \exists x_b \xrightarrow[b]{u_b} x_b' : (x_a', x_b') \in R$;
   b) $\forall x_b \xrightarrow[b]{u_b} x_b', \exists x_a \xrightarrow[a]{u_a} x_a' : (x_a', x_b') \in R$.

We say that S_a is ε-InitSOP simulated by S_b, denoted by $S_a \preceq_I^\varepsilon S_b$, if there exists an ε-InitSOP simulation relation R from S_a to S_b.

Note that although the above relation is similar to the approximate bisimulation relation proposed in [33], it is still a one sided relation here because condition 1) is not symmetric. We refer the interested readers to [28] to see why one needs strong condition 3) in Definition V.1 to show preservation of initial-state opacity in one direction when $\varepsilon = 0$.

The following main theorem provides a sufficient condition for δ-approximate initial-state opacity based on related systems as in Definition V.1.

Theorem V.2: Let $S_a = (X_a, X_{a0}, X_{aS}, U_a, \underset{a}{\longrightarrow}, Y_a, H_a)$ and $S_b = (X_b, X_{b0}, X_{bS}, U_b, \underset{b}{\longrightarrow}, Y_b, H_b)$ be two metric systems with the same output sets $Y_a = Y_b$ and metric $\mathbf{d}$, and let $\varepsilon, \delta \in \mathbb{R}_0^+$. If $S_a \preceq_I^\varepsilon S_b$ and $\varepsilon \leq \frac{\delta}{2}$, then the following implication holds:

$$S_b \text{ is } (\delta - 2\varepsilon)\text{-approximate initial-state opaque}$$
$$\Rightarrow S_a \text{ is } \delta\text{-approximate initial-state opaque.}$$

Proof: Consider an arbitrary secret initial state $x_0 \in X_{a0} \cap X_{aS}$ and a run $x_0 \xrightarrow[a]{u_1} x_1 \xrightarrow[a]{u_2} \cdots \xrightarrow[a]{u_n} x_n$ in S_a. Since $S_a \preceq_I^\varepsilon S_b$, by conditions 1)-a), 2), and 3)-a) in Definition V.1, there exist a secret initial state $x_0' \in X_{b0} \cap X_{bS}$ and a run $x_0' \xrightarrow[b]{u_1'} x_1' \xrightarrow[b]{u_2'} \cdots \xrightarrow[b]{u_n'} x_n'$ in S_b such that

$$\forall i \in \{0, 1, \ldots, n\} : \mathbf{d}(H_a(x_i), H_b(x_i')) \leq \varepsilon. \tag{6}$$

Since S_b is $(\delta - 2\varepsilon)$-approximate initial-state opaque, there exist a nonsecret initial state $x_0'' \in X_{b0} \setminus X_{bS}$ and a run $x_0'' \xrightarrow[b]{u_1''} x_1'' \xrightarrow[b]{u_2''} \cdots \xrightarrow[b]{u_n''} x_n''$ such that

$$\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(H_b(x_i'), H_b(x_i'')) \leq \delta - 2\varepsilon. \tag{7}$$

Again, since $S_a \preceq_I^\varepsilon S_b$, by conditions 1)-b), 2), and 3)-b) in Definition V.1, there exist an initial state $x_0''' \in X_{a0} \setminus X_{aS}$ and a run $x_0''' \xrightarrow[a]{u_1'''} x_1''' \xrightarrow[a]{u_2'''} \cdots \xrightarrow[a]{u_n'''} x_n'''$ such that

$$\forall i \in \{0, 1, \ldots, n\} : \mathbf{d}(H_a(x_i'''), H_b(x_i'')) \leq \varepsilon. \tag{8}$$

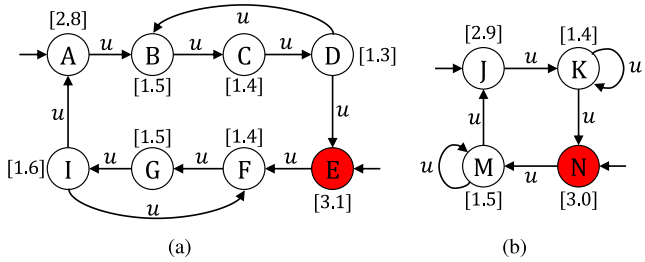


Fig. 3. Example of ε-InitSOP simulation relation.

Combining (6)–(8), and using the triangle inequality, we have

$$\max_{i \in \{0,1,\ldots,n\}} : \mathbf{d}(H_a(x_i), H_a(x_i''')) \leq \delta. \tag{9}$$

Since $x_0 \in X_{a0} \cap X_{aS}$ and $x_0 \xrightarrow[a]{u_1} x_1 \xrightarrow[a]{u_2} \cdots \xrightarrow[a]{u_n} x_n$ are arbitrary, we conclude that S_a is δ-approximate initial-state opaque. ■

The following corollary is a simple consequence of the result in Theorem V.2 but for the lack of δ-approximate initial-state opacity.

Corollary V.3: Let $S_a = (X_a, X_{a0}, X_{aS}, U_a, \underset{a}{\longrightarrow}, Y_a, H_a)$ and $S_b = (X_b, X_{b0}, X_{bS}, U_b, \underset{b}{\longrightarrow}, Y_b, H_b)$ be two metric systems with the same output sets $Y_a = Y_b$ and metric $\mathbf{d}$ and let $\varepsilon, \delta \in \mathbb{R}_0^+$. If $S_b \preceq_I^\varepsilon S_a$, then the following implication holds:

$$S_b \text{ is not } (\delta + 2\varepsilon)\text{-approximate initial-state opaque}$$
$$\Rightarrow S_a \text{ is not } \delta\text{-approximate initial-state opaque.}$$

Proof: Since $S_b \preceq_I^\varepsilon S_a$, by Theorem V.2, we know that S_a being δ-approximate initial-state opaque implies that S_b is $(\delta + 2\varepsilon)$-approximate initial-state opaque. Hence, S_b not being $(\delta + 2\varepsilon)$-approximate initial-state opaque implies that S_a is not δ-approximate initial-state opaque. ■

Remark V.4: It is worth remarking that δ and ε are parameters specifying two different types of precision. Parameter δ is used to specify the measurement precision under which we can guarantee opacity for a single system, while parameter ε is used to characterize the "distance" between two systems in terms of being approximate opaque. The reader should not be confused by the different roles of these two parameters.

We illustrate ε-InitSOP simulation relation by the following example.

Example V.5: Let us consider systems S_a and S_b shown in Fig. 3 (a) and (b), respectively. We mark all secret states by red, and the output map is specified by the value associated with each state. Let us consider the following relation: $R = \{(A, J), (B, K), (C, K), (D, K), (E, N), (F, M), (G, M), (I, M)\}$. We claim that R is an ε-InitSOP simulation relation from S_a to S_b when $\varepsilon = 0.1$. We check item by item following Definition V.1. First, for $E \in X_{a0} \cap X_{aS}$, we have $N \in X_{b0} \cap X_{bS}$ such that $(E, N) \in R$. Similarly, for $J \in X_{b0} \setminus X_{bS}$, we have $A \in X_{a0} \setminus X_{aS}$ such that $(A, J) \in R$. Therefore, condition 1) in Definition V.1 holds. In addition, for any $(x_a, x_b) \in R$, we have $\mathbf{d}(H_a(x_a), H_a(x_b)) \leq 0.1$, e.g., $\mathbf{d}(H_a(A), H_b(J)) = 0.1$ and $\mathbf{d}(H_a(C), H_b(K)) = 0$.

Therefore, condition 2) in Definition V.1 holds. Finally, we can also check that condition 3) in Definition V.1 holds. For example, for $(D,K)\in R$ and $D\xrightarrow[a]{u}B$, we can choose $K\xrightarrow[b]{u}K$ such that $(B,K)\in R$; for $(E,M)\in R$ and $N\xrightarrow[b]{u}M$, we can choose $E\xrightarrow[b]{u}F$ such that $(F,M)\in R$. Therefore, we know that R is an ε-InitSOP simulation relation from S_a to S_b, i.e., $S_a\preceq_I^{\varepsilon}S_b$.

Then, by applying the verification algorithm in Section IV, we can check that S_b is δ-approximate initial-state opaque for $\delta=0.1$. Therefore, according to Theorem V.2, we conclude that S_a is 0.3-approximate initial-state opaque, where $0.3=\delta+2\varepsilon$, without applying the verification algorithm to S_a directly.

B. Approximate Current-State Opacity-Preserving Simulation Relation

Now, we provide a notion of approximate simulation relation for preserving current-state opacity.

Definition V.6 (Approximate current-state opacity-preserving simulation relation): Let $S_a=(X_a,X_{a0},X_{aS},U_a,\xrightarrow[a]{},Y_a,H_a)$ and $S_b=(X_b,X_{b0},X_{bS},U_b,\xrightarrow[b]{},Y_b,H_b)$ be two metric systems with the same output sets $Y_a=Y_b$ and metric $\mathbf{d}$. For $\varepsilon\in\mathbb{R}_0^+$, a relation $R\subseteq X_a\times X_b$ is called an ε-approximate current-state opacity-preserving simulation relation (ε-CurSOP simulation relation) from S_a to S_b if:

1) $\forall x_{a0}\in X_{a0},\exists x_{b0}\in X_{b0}:(x_{a0},x_{b0})\in R$;
2) $\forall(x_a,x_b)\in R:\mathbf{d}(H_a(x_a),H_b(x_b))\leq\varepsilon$;
3) for any $(x_a,x_b)\in R$, we have
   a) $\forall x_a\xrightarrow[a]{u_a}x_a',\exists x_b\xrightarrow[b]{u_b}x_b':(x_a',x_b')\in R$;
   b) $\forall x_a\xrightarrow[a]{u_a}x_a'\in X_{aS},\exists x_b\xrightarrow[b]{u_b}x_b'\in X_{bS}:(x_a',x_b')\in R$;
   c) $\forall x_b\xrightarrow[b]{u_b}x_b',\exists x_a\xrightarrow[a]{u_a}x_a':(x_a',x_b')\in R$;
   d) $\forall x_b\xrightarrow[b]{u_b}x_b'\in X_b\setminus X_{bS},\exists x_a\xrightarrow[a]{u_a}x_a'\in X_a\setminus X_{aS}:(x_a',x_b')\in R$.

We say that S_a is ε-CurSOP simulated by S_b, denoted by $S_a\preceq_C^{\varepsilon}S_b$, if there exists an ε-CurSOP simulation relation R from S_a to S_b.

The following theorem provides a sufficient condition for δ-approximate current-state opacity based on related systems, as in Definition V.6.

Theorem V.7: Let $S_a=(X_a,X_{a0},X_{aS},U_a,\xrightarrow[a]{},Y_a,H_a)$ and $S_b=(X_b,X_{b0},X_{bS},U_b,\xrightarrow[b]{},Y_b,H_b)$ be two metric systems with the same output sets $Y_a=Y_b$ and metric $\mathbf{d}$, and let $\varepsilon,\delta\in\mathbb{R}_0^+$. If $S_a\preceq_C^{\varepsilon}S_b$ and $\varepsilon\leq\frac{\delta}{2}$, then the following implication holds:

$$S_b \text{ is } (\delta-2\varepsilon)\text{-approximate current-state opaque}$$
$$\Rightarrow S_a \text{ is } \delta\text{-approximate current-state opaque.}$$

Proof: Let us consider an arbitrary initial state $x_0\in X_{a0}$ and finite run $x_0\xrightarrow[a]{u_1}x_1\xrightarrow[a]{u_2}\cdots\xrightarrow[a]{u_n}x_n$ in S_a such that $x_n\in X_{aS}$. We consider the following two cases: $n=0$ and $n\neq 0$. If $n=0$, we know that $x_0\in X_{aS}$. Since we assume that $\{x\in X_0:(H_a(x_0),H_a(x))\leq\delta\}\not\subseteq X_{aS}$, we observe immediately that there exists $x_0'\in X_{a0}\setminus X_{aS}$ such that $\mathbf{d}(H_a(x_0),H_a(x_0'))\leq\delta$. Then, we consider the case of $n\geq 1$. Since $S_a\preceq_C^{\varepsilon}S_b$, by conditions 1), 2), 3)-a), and 3)-b) in Definition V.6, there exist an initial state $x_0'\in X_{b0}$ and a finite run $x_0'\xrightarrow[b]{u_1'}x_1'\xrightarrow[b]{u_2'}\cdots\xrightarrow[b]{u_n'}x_n'$ in S_b such that $x_n'\in X_{bS}$ and

$$\forall i\in\{0,1,\ldots,n\}:\mathbf{d}(H_a(x_i),H_b(x_i'))\leq\varepsilon. \tag{10}$$

Since S_b is $(\delta-2\varepsilon)$-approximate current-state opaque, there exist an initial state $x_0''\in X_{b0}$ and a finite run $x_0''\xrightarrow[b]{u_1''}x_1''\xrightarrow[b]{u_2''}\cdots\xrightarrow[b]{u_n''}x_n''$ such that $x_n''\in X_b\setminus X_{bS}$ and

$$\max_{i\in\{0,1,\ldots,n\}}\mathbf{d}(H_b(x_i'),H_b(x_i''))\leq\delta-2\varepsilon. \tag{11}$$

Again, since $S_a\preceq_C^{\varepsilon}S_b$, by conditions 1), 2), 3)-c), and 3)-d) in Definition V.6, there exist an initial state $x_0'''\in X_{a0}$ and a finite run $x_0'''\xrightarrow[a]{u_1'''}x_1'''\xrightarrow[a]{u_2'''}\cdots\xrightarrow[a]{u_n'''}x_n'''$ such that $x_n'''\in X_a\setminus X_{aS}$ and

$$\forall i\in\{0,1,\ldots,n\}:\mathbf{d}(H_a(x_i'''),H_b(x_i''))\leq\varepsilon. \tag{12}$$

Combining (10)–(12), and using the triangle inequality, we have

$$\max_{i\in\{0,1,\ldots,n\}}\mathbf{d}(H_a(x_i),H_a(x_i'''))\leq\delta. \tag{13}$$

Since $x_0\in X_{a0}$ and $x_0\xrightarrow[a]{u_1}x_1\xrightarrow[a]{u_2}\cdots\xrightarrow[a]{u_n}x_n$ are arbitrary, we conclude that S_a is δ-approximate current-state opaque. ■

C. Approximate Infinite-Step Opacity-Preserving Simulation Relation

Finally, by combining the ε-CurSOP simulation relation and the ε-InitSOP simulation relation, we provide a notion of approximate simulation relation for preserving infinite-step opacity.

Definition V.8 (Approximate infinite-step opacity-preserving simulation relation): Let $S_a=(X_a,X_{a0},X_{aS},U_a,\xrightarrow[a]{},Y_a,H_a)$ and $S_b=(X_b,X_{b0},X_{bS},U_b,\xrightarrow[b]{},Y_b,H_b)$ be two metric systems with the same output sets $Y_a=Y_b$ and metric $\mathbf{d}$. For $\varepsilon\in\mathbb{R}_0^+$, a relation $R\subseteq X_a\times X_b$ is called an ε-approximate infinite-step opacity-preserving simulation relation (ε-InfSOP simulation relation) from S_a to S_b if it is both an ε-CurSOP simulation relation from S_a to S_b and an ε-InitSOP simulation relation from S_a to S_b, i.e.,

1)
   a) $\forall x_{a0}\in X_{a0},\exists x_{b0}\in X_{b0}:(x_{a0},x_{b0})\in R$;
   b) $\forall x_{a0}\in X_{a0}\cap X_{aS},\exists x_{b0}\in X_{b0}\cap X_{bS}:(x_{a0},x_{b0})\in R$;
   c) $\forall x_{b0}\in X_{b0}\setminus X_{bS},\exists x_{a0}\in X_{a0}\setminus X_{aS}:(x_{a0},x_{b0})\in R$;
2) $\forall(x_a,x_b)\in R:\mathbf{d}(H_a(x_a),H_b(x_b))\leq\varepsilon$;
3) for any $(x_a,x_b)\in R$, we have
   a) $\forall x_a\xrightarrow[a]{u_a}x_a',\exists x_b\xrightarrow[b]{u_b}x_b':(x_a',x_b')\in R$;
   b) $\forall x_a\xrightarrow[a]{u_a}x_a'\in X_{aS},\exists x_b\xrightarrow[b]{u_b}x_b'\in X_{bS}:(x_a',x_b')\in R$;
   c) $\forall x_b\xrightarrow[b]{u_b}x_b',\exists x_a\xrightarrow[a]{u_a}x_a':(x_a',x_b')\in R$.

d) $\forall x_b \xrightarrow[b]{u_b} x_b' \in X_b \setminus X_{bS}, \exists x_a \xrightarrow[a]{u_a} x_a' \in X_a \setminus X_{aS} : (x_a', x_b') \in R$.

We say that S_a is ε-InfSOP simulated by S_b, denoted by $S_a \preceq^{\varepsilon}_{IF} S_b$, if there exists an ε-InfSOP simulation relation R from S_a to S_b.

Similar to the cases of initial-state opacity and current-state opacity, we have the following theorem as a sufficient condition for δ-approximate infinite-step opacity based on related systems as in Definition V.8.

Theorem V.9: Let $S_a = (X_a, X_{a0}, X_{aS}, U_a, \xrightarrow[a]{}, Y_a, H_a)$ and $S_b = (X_b, X_{b0}, X_{bS}, U_b, \xrightarrow[b]{}, Y_b, H_b)$ be two metric systems with the same output sets $Y_a = Y_b$ and metric $\mathbf{d}$, and let $\varepsilon, \delta \in \mathbb{R}_0^+$. If $S_a \preceq^{\varepsilon}_{IF} S_b$ and $\varepsilon \leq \frac{\delta}{2}$, then the following implication holds:

$$S_b \text{ is } (\delta - 2\varepsilon)\text{-approximate infinite-step opaque}$$
$$\Rightarrow S_a \text{ is } \delta\text{-approximate infinite-step opaque.}$$

Proof: Let us consider an arbitrary initial state $x_0 \in X_{a0}$ and finite run $x_0 \xrightarrow[a]{u_1} x_1 \xrightarrow[a]{u_2} \cdots \xrightarrow[a]{u_n} x_n$ in S_a such that $x_k \in X_{aS}$ for some $k = 0, \ldots, n$. We consider the following two cases:

If $k = 0$, then we have $x_0 \in X_{aS}$. Since $S_a \preceq^{\varepsilon}_{IF} S_b$ implies $S_a \preceq^{\varepsilon}_{I} S_b$, by the proof of Theorem V.2, we know that there exist an initial state $x_0' \in X_{a0} \setminus X_{aS}$ and a run $x_0' \xrightarrow[a]{u_1'} x_1' \xrightarrow[a]{u_2'} \cdots \xrightarrow[a]{u_n'} x_n'$ such that $\max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H_a(x_i), H_a(x_i')) \leq \delta$.

If $k \geq 1$, then similar to the proof of Theorem V.7, by conditions 1)-a), 2), 3)-a), 3)-b), 3)-c), and 3)-d) in Definition V.8 and the fact S_b is $(\delta - 2\varepsilon)$-approximate infinite-step opaque, there exist an initial state $x_0' \in X_{a0}$ and a finite run $x_0' \xrightarrow[a]{u_1'} x_1' \xrightarrow[a]{u_2'} \cdots \xrightarrow[a]{u_n'} x_n'$ such that $x_k' \in X_a \setminus X_{aS}$ and $\max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H_a(x_i), H_a(x_i')) \leq \delta$.

Since $x_0 \in X_{a0}$, $x_0 \xrightarrow[a]{u_1} x_1 \xrightarrow[a]{u_2} \cdots \xrightarrow[a]{u_n} x_n$ and index k are arbitrary, we conclude that S_a is δ-approximate infinite-step opaque. ■

VI. Opacity of Control Systems

In the previous section, we have introduced notions of approximate opacity-preserving simulation relation and discussed their properties. This allows us to verify approximate opacity for infinite systems, e.g., continuous dynamic systems, based on their finite abstractions. Then, the following question arises naturally: How can we construct such an abstraction for a given system possibly with infinite number of states? In general, how to construct finite abstractions is system-dependent and not all systems admit symbolic models. In this section, we show that a class of discrete-time control systems do admit symbolic models for the purpose of verifying approximate opacity under certain stability assumption.

To be more specific, we consider a class of discrete-time control systems of the following form.

Definition VI.1: A discrete-time control system Σ is defined by the tuple $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$, where $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{Y}$ are the state, input, and output sets, respectively, and are subsets of normed vector spaces with appropriate finite dimensions. Set $\mathbb{S} \subseteq \mathbb{X}$ is a set of secret states. The map $f : \mathbb{X} \times \mathbb{U} \to \mathbb{X}$ is called the transition function, and $h : \mathbb{X} \to \mathbb{Y}$ is the output map and assumed to satisfy the following Lipschitz condition: $\|h(x) - h(y)\| \leq \alpha(\|x - y\|)$ for some $\alpha \in \mathcal{K}_\infty$ and all $x, y \in \mathbb{X}$. The discrete-time control system Σ is described by difference equations of the form

$$\Sigma : \begin{cases} \xi(k+1) = f(\xi(k), \upsilon(k)) \\ \zeta(k) \quad\quad = h(\xi(k)) \end{cases} \tag{14}$$

where $\xi : \mathbb{N}_0 \to \mathbb{X}, \zeta : \mathbb{N}_0 \to \mathbb{Y}$, and $\upsilon : \mathbb{N}_0 \to \mathbb{U}$ are the state, output, and input signals, respectively.

We write $\xi_{x\upsilon}(k)$ to denote the point reached at time k under the input signal υ from initial condition $x = \xi_{x\upsilon}(0)$. Similarly, we denote by $\zeta_{x\upsilon}(k)$ the output corresponding to state $\xi_{x\upsilon}(k)$, i.e., $\zeta_{x\upsilon}(k) = h(\xi_{x\upsilon}(k))$. In the above definition, we implicitly assumed that set $\mathbb{X}$ is positively invariant.[1]

Now, we introduce the notion of incremental input-to-state stability (δ-ISS) leveraged later to show some of the main results of the article.

Definition VI.2 (see [48]): System $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$ is called incrementally input-to-state stable (δ-ISS) if there exist a $\mathcal{KL}$ function β and $\mathcal{K}_\infty$ function γ such that $\forall x, x' \in \mathbb{X}$ and $\forall \upsilon, \upsilon' : \mathbb{N}_0 \to \mathbb{U}$, the following inequality holds for any $k \in \mathbb{N}$:

$$\|\xi_{x\upsilon}(k) - \xi_{x'\upsilon'}(k)\| \leq \beta(\|x - x'\|, k) + \gamma(\|\upsilon - \upsilon'\|_\infty). \tag{15}$$

Example VI.3: As an example, for a linear control system

$$\xi(k+1) = A\xi(k) + B\upsilon(k), \quad \zeta(k) = C\xi(k) \tag{16}$$

where all eigenvalues of A are inside the unit circle, the functions β and γ can be chosen as

$$\beta(r, k) = \|A^k\| r; \quad \gamma(r) = \|B\| \left(\sum_{m=0}^{\infty} \|A^m\| \right) r. \tag{17}$$

In general, it is difficult to check inequality (15) directly for nonlinear systems. Fortunately, δ-ISS can be characterized using Lyapunov functions.

Definition VI.4 (see [48]): Consider a control system Σ and a continuous function $V : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_0^+$. Function V is called a δ-ISS Lyapunov function for Σ if there exist $\mathcal{K}_\infty$ functions α_1, α_2, ρ and $\mathcal{K}$ function σ such that:

1) for any $x, x' \in \mathbb{X}$, $\alpha_1(\|x - x'\|) \leq V(x, x') \leq \alpha_2(\|x - x'\|)$;
2) for any $x, x' \in \mathbb{X}$ and $u, u' \in \mathbb{U}$, $V(f(x,u), f(x',u')) - V(x,x') \leq -\rho(V(x,x')) + \sigma(\|u - u'\|)$.

The following result characterizes δ-ISS in terms of existence of δ-ISS Lyapunov functions.

Theorem VI.5 (see [48]): Consider a control system Σ.

1) Σ is δ-ISS if it admits a δ-ISS Lyapunov function.
2) If $\mathbb{U}$ is compact and convex and $\mathbb{X}$ is compact, then the existence of a δ-ISS Lyapunov function is equivalent to δ-ISS.

The following technical lemma will be used later to show some of the main results of this section.

[1] Set $\mathbb{X}$ is called positively invariant under (14) if $\xi_{x\upsilon}(k) \in \mathbb{X}$ for any $k \in \mathbb{N}$, any $x \in \mathbb{X}$, and any $\upsilon : \mathbb{N}_0 \to \mathbb{U}$.

Lemma VI.6: Consider a control system Σ. Suppose V is a δ-ISS Lyapunov function for Σ. Then, there exist $\kappa, \lambda \in \mathcal{K}_\infty$, where $\kappa(s) < s$ for any $s \in \mathbb{R}^+$, such that

$$V(f(x,u), f(x',u')) \leq \max\{\kappa(V(x,x')), \lambda(\|u-u'\|)\} \quad (18)$$

for any $x, x' \in \mathbb{X}$ and any $u, u' \in \mathbb{U}$.

The proof is similar to that of [49, Th. 1] and is omitted here due to lack of space.

In order to provide the main results of this section, we first describe control systems in Definition VI.1 as metric systems as in Definition II.1. More precisely, given a control system $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$, we define an associated metric system

$$S(\Sigma) = (X, X_0, X_S, U, \longrightarrow, Y, H) \quad (19)$$

where $X = \mathbb{X}$, $X_0 = \mathbb{X}$, $X_S = \mathbb{S}$, $U = \mathbb{U}$, $Y = \mathbb{Y}$, $H = h$, and $x \xrightarrow{u} x'$ if and only if $x' = f(x,u)$. We assume that the output set Y is equipped with the infinity norm: $\mathbf{d}(y_1, y_2) = \|y_1 - y_2\|$, $\forall y_1, y_2 \in Y$. We have a similar assumption for the state set X.

Now, we introduce a symbolic system for the control system $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$. To do so, from now on, we assume that sets $\mathbb{X}, \mathbb{S}$ and $\mathbb{U}$ are of the form of finite union of boxes. Consider a concrete control system Σ and a tuple $\mathsf{q} = (\eta, \mu, \theta)$ of parameters, where $0 < \eta \leq \min\{\text{span}(\mathbb{S}), \text{span}(\mathbb{X} \setminus \mathbb{S})\}$ is the state set quantization, $0 < \mu \leq \text{span}(\mathbb{U})$ is the input set quantization, and θ is a design parameter. Now, let us introduce the symbolic system

$$S_{\mathsf{q}}(\Sigma) = (X_{\mathsf{q}}, X_{\mathsf{q}0}, X_{\mathsf{q}S}, U_{\mathsf{q}}, \underset{\mathsf{q}}{\longrightarrow}, Y_{\mathsf{q}}, H_{\mathsf{q}}) \quad (20)$$

where $X_{\mathsf{q}} = X_{\mathsf{q}0} = [\mathbb{X}]_\eta$, $X_{\mathsf{q}S} = [\mathbb{S}^\theta]_\eta$, $U_{\mathsf{q}} = [\mathbb{U}]_\mu$, $Y_{\mathsf{q}} = \{h(x_{\mathsf{q}}) \mid x_{\mathsf{q}} \in X_{\mathsf{q}}\}$, $H_{\mathsf{q}}(x_{\mathsf{q}}) = h(x_{\mathsf{q}})$, $\forall x_{\mathsf{q}} \in X_{\mathsf{q}}$, and $x_{\mathsf{q}} \underset{\mathsf{q}}{\xrightarrow{u_{\mathsf{q}}}} x'_{\mathsf{q}}$ if and only if $\|x'_{\mathsf{q}} - f(x_{\mathsf{q}}, u_{\mathsf{q}})\| \leq \eta$.

We can now state the first main result of this section showing that, under some condition over the quantization parameters η and μ, $S_{\mathsf{q}}(\Sigma)$ and $S(\Sigma)$ are related under an approximate initial-state opacity-preserving simulation relation.

Theorem VI.7: Let $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$ be a δ-ISS control system. For any desired precision $\varepsilon > 0$, and any tuple $\mathsf{q} = (\eta, \mu, 0)$ of parameters satisfying

$$\beta\left(\alpha^{-1}(\varepsilon), 1\right) + \gamma(\mu) + \eta \leq \alpha^{-1}(\varepsilon) \quad (21)$$

we have $S(\Sigma) \preceq^\varepsilon_I S_{\mathsf{q}}(\Sigma) \preceq^\varepsilon_I S(\Sigma)$.

Proof: We start by proving $S(\Sigma) \preceq^\varepsilon_I S_{\mathsf{q}}(\Sigma)$. Consider the relation $R \subseteq X \times X_{\mathsf{q}}$ defined by $(x, x_{\mathsf{q}}) \in R$ if and only if $\|x - x_{\mathsf{q}}\| \leq \alpha^{-1}(\varepsilon)$. Since $\eta \leq \text{span}(\mathbb{S})$, $X_S \subseteq \bigcup_{p \in [\mathbb{S}]_\eta} \mathcal{B}_\eta(p)$, and by (21), $\forall x \in X_S$, $\exists x_{\mathsf{q}} \in X_{\mathsf{q}S}$ such that

$$\|x - x_{\mathsf{q}}\| \leq \eta \leq \alpha^{-1}(\varepsilon). \quad (22)$$

Hence, $(x, x_{\mathsf{q}}) \in R$ and condition 1)-a) in Definition V.1 is satisfied. For every $x_{\mathsf{q}} \in X_{\mathsf{q}} \setminus X_{\mathsf{q}S}$, by choosing $x = x_{\mathsf{q}}$, which is also inside set $X \setminus X_S$, one gets $(x, x_{\mathsf{q}}) \in R$, and hence, condition 1)-b) in Definition V.1 holds as well. Now, consider any $(x, x_{\mathsf{q}}) \in R$. Condition 2) in Definition V.1 is satisfied by the definition of R and the Lipschitz assumption:

$$\|H(x) - H_{\mathsf{q}}(x_{\mathsf{q}})\| = \|h(x) - h(x_{\mathsf{q}})\| \leq \alpha(\|x - x_{\mathsf{q}}\|) \leq \varepsilon.$$

Let us now show that condition 3) in Definition V.1 holds.

Consider any $u \in U$. Choose an input $u_{\mathsf{q}} \in U_{\mathsf{q}}$ satisfying

$$\|u - u_{\mathsf{q}}\| \leq \mu. \quad (23)$$

Note that the existence of such u_{q} is guaranteed by the inequality $\mu \leq \text{span}(\mathbb{U})$ which guarantees that $\mathbb{U} \subseteq \bigcup_{p \in [\mathbb{U}]_\mu} \mathcal{B}_\mu(p)$. Consider the unique transition $x \xrightarrow{u} x' = f(x,u)$ in $S(\Sigma)$. It follows from the δ-ISS assumption on Σ and (23) that the distance between x' and $f(x_{\mathsf{q}}, u_{\mathsf{q}})$ is bounded as

$$\begin{aligned}\|x' - f(x_{\mathsf{q}}, u_{\mathsf{q}})\| &\leq \beta\left(\|x - x_{\mathsf{q}}\|, 1\right) + \gamma\left(\|u - u_{\mathsf{q}}\|\right) \\ &\leq \beta\left(\alpha^{-1}(\varepsilon), 1\right) + \gamma(\mu). \end{aligned} \quad (24)$$

Since $X \subseteq \bigcup_{p \in [\mathbb{X}]_\eta} \mathcal{B}_\eta(p)$, there exists $x'_{\mathsf{q}} \in X_{\mathsf{q}}$ such that

$$\|f(x_{\mathsf{q}}, u_{\mathsf{q}}) - x'_{\mathsf{q}}\| \leq \eta \quad (25)$$

which, by the definition of $S_{\mathsf{q}}(\Sigma)$, implies the existence of $x_{\mathsf{q}} \underset{\mathsf{q}}{\xrightarrow{u_{\mathsf{q}}}} x'_{\mathsf{q}}$ in $S_{\mathsf{q}}(\Sigma)$. Using the inequalities (21), (24), (25), and triangle inequality, we obtain

$$\begin{aligned}\|x' - x'_{\mathsf{q}}\| &\leq \|x' - f(x_{\mathsf{q}}, u_{\mathsf{q}}) + f(x_{\mathsf{q}}, u_{\mathsf{q}}) - x'_{\mathsf{q}}\| \\ &\leq \|x' - f(x_{\mathsf{q}}, u_{\mathsf{q}})\| + \|f(x_{\mathsf{q}}, u_{\mathsf{q}}) - x'_{\mathsf{q}}\| \\ &\leq \beta\left(\alpha^{-1}(\varepsilon), 1\right) + \gamma(\mu) + \eta \leq \alpha^{-1}(\varepsilon).\end{aligned}$$

Therefore, we conclude that $(x', x'_{\mathsf{q}}) \in R$, and condition 3)-a) in Definition V.1 holds. Let us now show that condition 3)-b) in Definition V.1 also holds.

Now, consider any $(x, x_{\mathsf{q}}) \in R$ and any $u_{\mathsf{q}} \in U_{\mathsf{q}}$. Choose the input $u = u_{\mathsf{q}}$ and consider the unique $x' = f(x,u)$ in $S(\Sigma)$. Using δ-ISS assumption for Σ, we bound the distance between x' and $f(x_{\mathsf{q}}, u_{\mathsf{q}})$ as

$$\|x' - f(x_{\mathsf{q}}, u_{\mathsf{q}})\| \leq \beta\left(\|x - x_{\mathsf{q}}\|, 1\right) \leq \beta\left(\alpha^{-1}(\varepsilon), 1\right). \quad (26)$$

Using the definition of $S_{\mathsf{q}}(\Sigma)$, inequalities (21) and (26), and the triangle inequality, we obtain

$$\begin{aligned}\|x' - x'_{\mathsf{q}}\| &\leq \|x' - f(x_{\mathsf{q}}, u_{\mathsf{q}}) + f(x_{\mathsf{q}}, u_{\mathsf{q}}) - x'_{\mathsf{q}}\| \\ &\leq \|x' - f(x_{\mathsf{q}}, u_{\mathsf{q}})\| + \|f(x_{\mathsf{q}}, u_{\mathsf{q}}) - x'_{\mathsf{q}}\| \\ &\leq \beta\left(\alpha^{-1}(\varepsilon), 1\right) + \eta \leq \alpha^{-1}(\varepsilon).\end{aligned}$$

Therefore, we conclude that $(x', x'_{\mathsf{q}}) \in R$, and condition 3)-b) in Definition V.1 holds.

In a similar way, one can prove that $S_{\mathsf{q}}(\Sigma) \preceq^\varepsilon_I S(\Sigma)$. ■

Remark VI.8: Note that there always exist quantization parameters q such that inequality (21) holds as long as $\beta(\alpha^{-1}(\varepsilon), 1) < \alpha^{-1}(\varepsilon)$. By assuming that the discrete-time control system Σ is a sampled-data version of an original continuous-time one with the sampling time τ, one can ensure the latter inequality by choosing the sampling time large enough given that $\beta(r, 1) = \hat{\beta}(r, \tau) < r$ for some $\mathcal{KL}$ function $\hat{\beta}$ establishing the incremental stability of the original continuous-time system. For example, for the function in (17), one has $\beta(r, 1) = \|A\| r = \|\mathrm{e}^{\hat{A}\tau}\| r$, where $\hat{A}$ is the state matrix of the original continuous-time linear control system.

The following example illustrates how to use Theorem VI.7 to verify approximate opacity for an infinite system based on its finite abstraction.

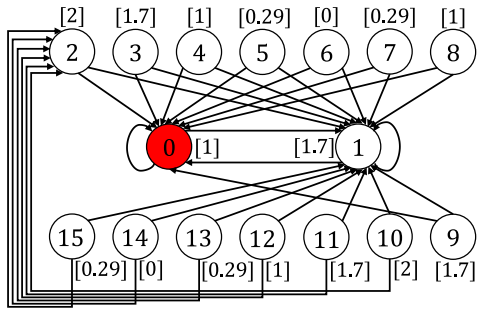


Fig. 4. Symbolic model $S_{\mathsf{q}}(\Sigma)$ associated with control systems Σ of (27) with $\eta = 0.1, \mu = 0.001$, and $\varepsilon = 0.9$.

Example VI.9: Let us consider the following simple system:

$$\Sigma : \begin{cases} \xi(k+1) = 0.1\xi(k) + \upsilon(k) \\ \zeta(k) = \sin(2.5\pi\xi(k)) + 1 \end{cases} \tag{27}$$

where $\mathbb{X} = [0, 1.6[, \mathbb{S} = [0, 0.1[$ and $\mathbb{U} = \{0.001\}$. This system is clearly δ-ISS, and according to (17), we have $\beta(r,k) = 0.1^k r$ and $\gamma(r) = \sum_{m=0}^{\infty} 0.1^m r$. In addition, function h satisfies the Lipschitz condition with $\alpha(r) = 2.5\pi r$. By (21), the parameters $\mathsf{q} = (\eta, \mu, 0)$ and the abstract precision ε should satisfy $\frac{0.04}{\pi}\varepsilon + \frac{10}{9}\mu + \eta \leq \frac{0.4}{\pi}\varepsilon$. Let us consider desired abstract precision $\varepsilon = 0.9$ and quantization parameters $\mathsf{q} = (\eta, \mu, 0) = (0.1, 0.001, 0)$ satisfying the inequality. Then, we obtain symbolic system $S_{\mathsf{q}}(\Sigma)$ shown in Fig. 4 , and by Theorem VI.7, we have $S(\Sigma) \preceq_I^{0.9} S_{\mathsf{q}}(\Sigma) \preceq_I^{0.9} S(\Sigma)$. Essentially, we discretize the state space of [0,1.6[into 16 discrete states based on parameter η. One can easily check that $S_{\mathsf{q}}(\Sigma)$ is 0-approximate initial-state opaque since for any run from secret initial state 0, there exists a run from nonsecret state 8 such that their outputs are exactly the same. Therefore, by Theorem V.2, we can conclude that Σ is 1.8-approximate initial-state opaque.

The next theorem provides similar results as in Theorem VI.7 but by leveraging δ-ISS Lyapunov functions. To show the next result, we will make the following supplementary assumption on the δ-ISS Lyapunov functions: there exists a function $\hat{\gamma} \in \mathcal{K}_\infty$ such that

$$\forall x, x', x'' \in \mathbb{X}, \quad V(x,x') - V(x',x'') \leq \hat{\gamma}(\|x - x''\|). \tag{28}$$

Inequality (28) is not restrictive at all, provided that we are interested in the dynamics of the control system on a compact subset of the state set $\mathbb{X}$ (see the discussion in [34]).

Theorem VI.10: Let $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$ admit a δ-ISS Lyapunov function V satisfying (28). For any desired precision $\varepsilon > 0$, and any tuple $\mathsf{q} = (\eta, \mu, 0)$ satisfying

$$\alpha_2(\eta) \leq \alpha_1(\alpha^{-1}(\varepsilon)) \tag{29}$$

$$\max\{\kappa(\alpha_1(\alpha^{-1}(\varepsilon))), \lambda(\mu)\} + \hat{\gamma}(\eta) \leq \alpha_1(\alpha^{-1}(\varepsilon)) \tag{30}$$

we have $S(\Sigma) \preceq_I^{\varepsilon} S_{\mathsf{q}}(\Sigma) \preceq_I^{\varepsilon} S(\Sigma)$.

Proof: We start by proving $S(\Sigma) \preceq_I^{\varepsilon} S_{\mathsf{q}}(\Sigma)$. Consider the relation $R \subseteq X \times X_{\mathsf{q}}$ defined by $(x, x_{\mathsf{q}}) \in R$ if and only if $V(x, x_{\mathsf{q}}) \leq \alpha_1(\alpha^{-1}(\varepsilon))$. Since $\eta \leq \text{span}(\mathbb{S})$ and $X_S \subseteq \bigcup_{p \in [\mathbb{S}]_\eta} \mathcal{B}_\eta(p)$, for every $x \in X_S$, there always exists $x_{\mathsf{q}} \in X_{\mathsf{q}S}$ such that $\|x - x_{\mathsf{q}}\| \leq \eta$. Then, we have

$$V(x, x_{\mathsf{q}}) \leq \alpha_2(\|x - x_{\mathsf{q}}\|) \leq \alpha_2(\eta) \leq \alpha_1(\alpha^{-1}(\varepsilon))$$

because of (29) and α_2 being a $\mathcal{K}_\infty$ function. Hence, $(x, x_{\mathsf{q}}) \in R$, and condition 1)-a) in Definition V.1 is satisfied. For every $x_{\mathsf{q}} \in X_{\mathsf{q}} \setminus X_{\mathsf{q}S}$, by choosing $x = x_{\mathsf{q}}$, which is also inside set $X \setminus X_S$, one gets trivially $(x, x_{\mathsf{q}}) \in R$, and hence, condition 1)-b) in Definition V.1 holds as well. Now, consider any $(x, x_{\mathsf{q}}) \in R$. Condition 2) in Definition V.1 is satisfied by the definition of R and the Lipschitz assumption on map h as in Definition VI.1

$$\begin{aligned} \|H(x) - H_{\mathsf{q}}(x_{\mathsf{q}})\| &= \|h(x) - h(x_{\mathsf{q}})\| \leq \alpha(\|x - x_{\mathsf{q}}\|) \\ &\leq \alpha(\alpha_1^{-1}(V(x, x_{\mathsf{q}})) \leq \varepsilon. \end{aligned}$$

Let us now show that condition 3) in Definition V.1 holds.

Consider any $u \in U$. Choose an input $u_{\mathsf{q}} \in U_{\mathsf{q}}$ satisfying

$$\|u - u_{\mathsf{q}}\| \leq \mu. \tag{31}$$

Note that the existence of such u_{q} is guaranteed by the inequality $\mu \leq \text{span}(\mathbb{U})$ which guarantees that $\mathbb{U} \subseteq \bigcup_{p \in [\mathbb{U}]_\mu} \mathcal{B}_\mu(p)$. Consider the unique transition $x \xrightarrow{u} x' = f(x,u)$ in $S(\Sigma)$. Given δ-ISS Lyapunov function V for Σ, inequality (18), and (31), one obtains

$$\begin{aligned} V(x', f(x_{\mathsf{q}}, u_{\mathsf{q}})) &\leq \max\{\kappa(V(x, x_{\mathsf{q}})), \lambda(\|u - u_{\mathsf{q}}\|)\} \\ &\leq \max\{\kappa(\alpha_1(\alpha^{-1}(\varepsilon))), \lambda(\mu)\}. \end{aligned} \tag{32}$$

Since $X \subseteq \bigcup_{p \in [\mathbb{X}]_\eta} \mathcal{B}_\eta(p)$, there exists $x'_{\mathsf{q}} \in X_{\mathsf{q}}$ such that

$$\|f(x_{\mathsf{q}}, u_{\mathsf{q}}) - x'_{\mathsf{q}}\| \leq \eta \tag{33}$$

which, by the definition of $S_{\mathsf{q}}(\Sigma)$, implies the existence of $x_{\mathsf{q}} \xrightarrow{u_{\mathsf{q}}}_{\mathsf{q}} x'_{\mathsf{q}}$ in $S_{\mathsf{q}}(\Sigma)$. Using the inequalities (28), (30), (32), and (33), we obtain

$$\begin{aligned} V(x', x'_{\mathsf{q}}) &\leq V(x', f(x_{\mathsf{q}}, u_{\mathsf{q}})) + \hat{\gamma}(\|f(x_{\mathsf{q}}, u_{\mathsf{q}}) - x'_{\mathsf{q}}\|) \\ &\leq \max\{\kappa(\alpha_1(\alpha^{-1}(\varepsilon))), \lambda(\mu)\} + \hat{\gamma}(\eta) \\ &\leq \alpha_1(\alpha^{-1}(\varepsilon)). \end{aligned}$$

Therefore, we conclude that $(x', x'_{\mathsf{q}}) \in R$, and condition 3)-a) in Definition V.1 holds. Let us now show that condition 3)-b) in Definition V.1 also holds.

Now, consider any $(x, x_{\mathsf{q}}) \in R$. Consider any $u_{\mathsf{q}} \in U_{\mathsf{q}}$. Choose the input $u = u_{\mathsf{q}}$ and consider the unique $x' = f(x,u)$ in $S(\Sigma)$. Given δ-ISS Lyapunov function V for Σ and inequality (18), one gets

$$V(x', f(x_{\mathsf{q}}, u_{\mathsf{q}})) \leq \kappa(V(x, x_{\mathsf{q}})) \leq \kappa(\alpha_1(\alpha^{-1}(\varepsilon))). \tag{34}$$

Using the definition of $S_{\mathsf{q}}(\Sigma)$, and inequalities (28), (30), and (34), we obtain

$$\begin{aligned} V(x', x'_{\mathsf{q}}) &\leq V(x', f(x_{\mathsf{q}}, u_{\mathsf{q}})) + \hat{\gamma}(\|f(x_{\mathsf{q}}, u_{\mathsf{q}}) - x'_{\mathsf{q}}\|) \\ &\leq \kappa(\alpha_1(\alpha^{-1}(\varepsilon))) + \hat{\gamma}(\eta) \leq \alpha_1(\alpha^{-1}(\varepsilon)). \end{aligned}$$

Therefore, we conclude that $(x', x'_{\mathsf{q}}) \in R$, and condition 3)-b) in Definition V.1 holds.

In a similar way, one can prove that $S_{\mathsf{q}}(\Sigma) \preceq_I^{\varepsilon} S(\Sigma)$. ■

Remark VI.11: One can readily verify that there always exists a choice of quantization parameter $\mathsf{q} = (\eta, \mu)$ such that inequalities (29) and (30) hold simultaneously. Although the result in Theorem VI.10 seems more general than that of Theorem VI.7 in terms of the existence of quantization parameter q, the symbolic model $S_{\mathsf{q}}(\Sigma)$, computed by using the quantization parameters q provided in Theorem VI.7 whenever existing, is likely to have fewer states than those of the model computed by using the quantization parameters provided in Theorem VI.10 owing to the conservative nature of δ-ISS Lyapunov functions.

The following theorems illustrate the other main results of this section showing that, under similar conditions over the quantization parameters η and μ, $S_{\mathsf{q}}(\Sigma)$ and $S(\Sigma)$ are related under an approximate current-state opacity-preserving simulation relation.

Theorem VI.12: Let $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$ be a δ-ISS control system. For any desired precision $\varepsilon > 0$, and any tuple $\mathsf{q} = (\eta, \mu, \theta)$ of parameters satisfying

$$\beta\left(\alpha^{-1}(\varepsilon), 1\right) + \gamma(\mu) + \eta \leq \alpha^{-1}(\varepsilon)$$
$$\alpha^{-1}(\varepsilon) \leq \theta$$

we have $S(\Sigma) \preceq^{\varepsilon}_{C} S_{\mathsf{q}}(\Sigma)$.

Proof: Consider the relation $R \subseteq X \times X_{\mathsf{q}}$ defined by $(x, x_{\mathsf{q}}) \in R$ if and only if $\|x - x_{\mathsf{q}}\| \leq \alpha^{-1}(\varepsilon)$. Note that conditions 1), 2), 3)-a), and 3)-c) of ε-CurSOP simulation relation in Definition V.6 are similar to that of ε-InitSOP simulation relation; therefore, their proofs are similar to that in Theorem VI.7 and are omitted here. Here, we show that conditions 3)-b) and 3)-d) in Definition V.6 hold.

Let us consider an arbitrary transition $x \xrightarrow{u} x' = f(x, u)$ with $x' \in X_S$ in $S(\Sigma)$. Similar to the proof of condition 3)-a), we can show the existence of a transition $x_{\mathsf{q}} \xrightarrow[\mathsf{q}]{u_{\mathsf{q}}} x'_{\mathsf{q}}$ in $S_{\mathsf{q}}(\Sigma)$ where $(x', x'_{\mathsf{q}}) \in R$ holds, where the input $u_{\mathsf{q}} \in U_{\mathsf{q}}$ satisfies $\|u - u_{\mathsf{q}}\| \leq \mu$. By the construction of the secret set in the symbolic system, one has $X_{\mathsf{q}S} = [\mathbb{S}^{\theta}]_{\eta}$ with $\theta \geq \alpha^{-1}(\varepsilon)$ and $0 < \eta \leq \min\{\text{span}(\mathbb{S}), \text{span}(\mathbb{X} \setminus \mathbb{S})\}$. Therefore, since $(x', x'_{\mathsf{q}}) \in R$, which implies that $\|x' - x'_{\mathsf{q}}\| \leq \alpha^{-1}(\varepsilon)$, we obtain that $x'_{\mathsf{q}} \in X_{\mathsf{q}S}$. Thus, we conclude that condition 3)-b) in Definition V.6 holds. In a similar way, we can show that condition 3)-d) in Definition V.6 holds as well, which completes the proof. ■

Theorem VI.13: Let $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$ admits a δ-ISS Lyapunov function V satisfying (28). For any desired precision $\varepsilon > 0$, and any tuple $\mathsf{q} = (\eta, \mu, \theta)$ satisfying

$$\alpha_2(\eta) \leq \alpha_1(\alpha^{-1}(\varepsilon))$$
$$\max\{\kappa(\alpha_1(\alpha^{-1}(\varepsilon))), \lambda(\mu)\} + \hat{\gamma}(\eta) \leq \alpha_1(\alpha^{-1}(\varepsilon))$$
$$\alpha^{-1}(\varepsilon) \leq \theta$$

we have $S(\Sigma) \preceq^{\varepsilon}_{C} S_{\mathsf{q}}(\Sigma)$.

Proof: The proof is similar to that of Theorems VI.10 and VI.12 and is omitted here due to lack of space. ■

Since we show $S(\Sigma) \preceq^{\varepsilon}_{I} S_{\mathsf{q}}(\Sigma)$ and $S(\Sigma) \preceq^{\varepsilon}_{C} S_{\mathsf{q}}(\Sigma)$ under the *same* relation in Theorems VI.7 and VI.12 (respectively, Theorems VI.10 and VI.13), by the definition of approximate infinite-state opacity-preserving simulation relation, we consequently get the following results.

Theorem VI.14: Let $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$ be a δ-ISS control system. For any desired precision $\varepsilon > 0$, and any tuple $\mathsf{q} = (\eta, \mu, \theta)$ of parameters satisfying

$$\beta\left(\alpha^{-1}(\varepsilon), 1\right) + \gamma(\mu) + \eta \leq \alpha^{-1}(\varepsilon)$$
$$\alpha^{-1}(\varepsilon) \leq \theta$$

we have $S(\Sigma) \preceq^{\varepsilon}_{IF} S_{\mathsf{q}}(\Sigma)$.

Theorem VI.15: Let $\Sigma = (\mathbb{X}, \mathbb{S}, \mathbb{U}, f, \mathbb{Y}, h)$ admit a δ-ISS Lyapunov function V satisfying (28). For any desired precision $\varepsilon > 0$, and any tuple $\mathsf{q} = (\eta, \mu, \theta)$ satisfying

$$\alpha_2(\eta) \leq \alpha_1(\alpha^{-1}(\varepsilon))$$
$$\max\{\kappa(\alpha_1(\alpha^{-1}(\varepsilon))), \lambda(\mu)\} + \hat{\gamma}(\eta) \leq \alpha_1(\alpha^{-1}(\varepsilon))$$
$$\alpha^{-1}(\varepsilon) \leq \theta$$

we have $S(\Sigma) \preceq^{\varepsilon}_{IF} S_{\mathsf{q}}(\Sigma)$.

VII. Conclusion

In this article, we extended the concept of opacity to metric systems by proposing the notion of approximate opacity. Verification algorithms and approximate relations that preserve approximate opacity were also provided. We also discussed how to construct finite abstractions that approximately simulate a class of control systems in terms of opacity preservation. Our result bridges the gap between the opacity analysis of finite discrete systems and continuous control systems.

Among the many possible directions for future work that will be built based on the proposed framework, we mention several directions of immediate interest. One direction is to extend our framework to the stochastic setting for almost opacity [14], [21]–[23]. In addition, we are interested in constructing approximate opacity-preserving symbolic models for more classes of systems. Finally, we plan to extend approximate opacity-preserving simulation relation to approximate opacity-preserving *alternating* simulation relation [45] and solve the problem of controller synthesis enforcing approximate opacity [46], [50]–[54].

Appendix

A. Proofs not Contained in the Main Body

Proof of Proposition IV.2

Proof: It is straightforward to show (i). Hereafter, we prove (ii) by induction on the length of input sequence.

When $n = 0$, i.e., there is no input sequence, we have that $(x_0, q_0) \in X_{I0}$. By the definition of X_{I0}, we know that

$$q_0 = \{x'_0 \in X : \mathbf{d}(H(x_0), H(x'_0)) \leq \delta\}$$

which implies (ii) immediately.

To proceed the induction, we assume that (ii) holds when $n = k$. Now, we need to show that (ii) also holds when $n = k + 1$. Consider arbitrary pair $(x_0, q_0) \in X_{I0}$ and finite run

$$(x_0, q_0) \xrightarrow[I]{u_1} (x_1, q_1) \xrightarrow[I]{u_2} \cdots \xrightarrow[I]{u_n} (x_n, q_n) \xrightarrow[I]{u_{n+1}} (x_{n+1}, q_{n+1}).$$

Then, we have

$$\begin{aligned} q_{n+1} &= \cup_{\hat{u}\in U}\, \mathbf{Pre}_{\hat{u}}(q_n) \cap \{x \in X : \mathbf{d}(H(x_{n+1}), H(x)) \leq \delta\} \\ &= \{x \in X : \exists x' \in q_n, u'_{n+1} \in U \text{ s.t. } (x, u'_{n+1}, x') \in \longrightarrow\} \\ &\quad \cap \{x \in X : \mathbf{d}(H(x_{n+1}), H(x)) \leq \delta\} \\ &= \left\{x \in X : \begin{array}{c}[\exists x' \in q_n, u'_{n+1} \in U \text{ s.t. } (x, u'_{n+1}, x') \in \longrightarrow] \\ \wedge [\mathbf{d}(H(x_{n+1}), H(x)) \leq \delta]\end{array}\right\}. \end{aligned}$$

By the induction hypothesis, we know that

$$q_n = \left\{x'_0 \in X : \begin{array}{c}\exists x'_0 \xrightarrow{u'_n} x'_1 \xrightarrow{u'_{n-1}} \cdots \xrightarrow{u'_1} x'_n \text{ s.t.} \\ \max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_{n-i})) \leq \delta\end{array}\right\}.$$

Therefore, by combining the above two equations, one gets

$$\begin{aligned} q_{n+1} &= \left\{x \in X : \begin{array}{c}\exists x \xrightarrow{u'_{n+1}} x'_0 \xrightarrow{u'_n} x'_1 \xrightarrow{u_{n-1}} \cdots \xrightarrow{u'_1} x'_n \\ \text{s.t. } \max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_{n-i})) \leq \delta \\ \wedge \mathbf{d}(H(x_{n+1}), H(x)) \leq \delta\end{array}\right\} \\ &= \left\{x \in X : \begin{array}{c}\exists x''_0 \xrightarrow{u'_{n+1}} x''_1 \xrightarrow{u'_n} \cdots \xrightarrow{u'_1} x''_{n+1} \text{ s.t.} \\ \max_{i\in\{0,1,\ldots,n+1\}} \mathbf{d}(H(x_i), H(x''_{n+1-i})) \leq \delta\end{array}\right\}. \end{aligned}$$

Therefore, one obtains that the induction step holds. ■

Proof of Theorem IV.3

Proof: ($\Rightarrow$) By contraposition: suppose that there exists a state $(x, q) \in X_I$ such that $x \in X_0 \cap X_S$ and $q \cap X_0 \subseteq X_S$. Let

$$(x_0, q_0) \xrightarrow[I]{u_1} (x_1, q_1) \xrightarrow[I]{u_1} \cdots \xrightarrow[I]{u_n} (x_n, q_n)$$

be a run reaching $(x, q) =: (x_n, q_n)$. By Proposition IV.2, we have $x_n \xrightarrow{u_n} x_{n-1} \xrightarrow{u_{n-1}} \cdots \xrightarrow{u_1} x_1$, which is well-defined in S as $x_n \in X_0$. Moreover, by Proposition IV.2, we have

$$q_n = \left\{x'_0 \in X : \begin{array}{c}\exists x'_0 \xrightarrow{u'_n} x'_1 \xrightarrow{u'_{n-1}} \cdots \xrightarrow{u'_1} x'_n \text{ s.t.} \\ \max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_{n-i})) \leq \delta\end{array}\right\}.$$

However, since $q_n \cap X_0 \subseteq X_S$, we know that there does not exist $x'_0 \in X_0 \setminus X_S$ and $x'_0 \xrightarrow{u'_n} x'_1 \xrightarrow{u'_{n-1}} \cdots \xrightarrow{u'_1} x'_n$ such that $\max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_{n-i})) \leq \delta$. Therefore, by considering $x_n \in X_0 \cap X_S$ and $x_n \xrightarrow{u_n} x_{n-1} \xrightarrow{u_{n-1}} \cdots \xrightarrow{u_1} x_1$, we know the system is not δ-approximate initial-state opaque.

($\Leftarrow$) By contradiction: suppose that (3) holds and assume that S is not δ-approximate initial-state opaque. Then, there exists a secret initial state $x_0 \in X_0 \cap X_S$ and a sequence of transitions $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$ such that there does not exist a nonsecret initial state $x'_0 \in X_0 \setminus X_S$ and a sequence of transitions $x'_0 \xrightarrow{u'_1} x'_1 \xrightarrow{u'_2} \cdots \xrightarrow{u'_n} x'_n$ such that $\max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_i)) \leq \delta$. Let us consider the following sequence of transitions in S_I

$$(x_n, q_0) \xrightarrow[I]{u_n} (x_{n-1}, q_1) \xrightarrow[I]{u_{n-1}} \cdots \xrightarrow[I]{u_1} (x_0, q_n).$$

By Proposition IV.2, we know that

$$q_n = \left\{x'_0 \in X : \begin{array}{c}\exists x'_0 \xrightarrow{u'_n} x'_1 \xrightarrow{u'_{n-1}} \cdots \xrightarrow{u'_1} x'_n \text{ s.t.} \\ \max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_i)) \leq \delta\end{array}\right\}.$$

By (3), we have $q_n \cap X_0 \not\subseteq X_S$. Therefore, there exist a nonsecret initial state $x'_0 \in X_0 \setminus X_S$ and a sequence $x'_0 \xrightarrow{u'_1} x'_1 \xrightarrow{u'_2} \cdots \xrightarrow{u'_n} x'_n$ such that $\max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_i)) \leq \delta$. This is a contradiction, i.e., S has to be δ-approximate initial-state opaque. ■

Proof of Proposition IV.6

Proof: The proof is similar to that of Proposition IV.2, which can be done by induction on the length of the sequence. ■

Proof of Theorem IV.7

Proof: By Proposition IV.6, for each state (x, q) encountered, the second component is exactly the set of all possible current states consistent with the observation. Then, the proof is similar to that of Theorem IV.3. ■

Proof of Theorem IV.8

Proof: By contraposition: suppose that there exist two states $(x_n, q'_n) \in X_I, (x_n, q_n) \in X_C$ such that $x_n \in X_S$ and $q_n \cap q'_n \subseteq X_S$. Let

$$\begin{aligned} &(x_0, q_0) \xrightarrow[C]{u_1} (x_1, q_1) \xrightarrow[C]{u_2} \cdots \xrightarrow[C]{u_n} (x_n, q_n) \\ &(x_{n+m}, q_{n+m}) \xrightarrow[I]{u_{n+m}} (x_{n+m-1}, q_{n+m-1}) \\ &\cdots \xrightarrow[I]{u_{n+1}} (x_n, q'_n) \end{aligned}$$

be two runs reaching (x, q) and (x, q'), respectively. By Propositions IV.2 and IV.6, we have $x_0 \in X_0$ and

$$\begin{aligned} x_0 \xrightarrow{u_1} \cdots \xrightarrow{u_{n-1}} x_{n-1} \xrightarrow{u_n} x_n \xrightarrow{u_{n+1}} x_{n+1} \xrightarrow{u_{n+2}} \\ \cdots \xrightarrow{u_{n+m}} x_{n+m}. \end{aligned}$$

Moreover, one has

$$q_n \cap q'_n = \left\{x'_n \in X : \begin{array}{c}\exists x'_0 \in X_0, \exists x'_0 \xrightarrow{u'_1} \cdots \xrightarrow{u'_{n+m}} x'_{n+m} \\ \text{s.t. } \max_{i\in\{0,1,\ldots,n+m\}} \mathbf{d}(H(x_i), H(x'_i)) \leq \delta\end{array}\right\}.$$

Since $q_n \cap q'_n \subseteq X_S$, we know that there does not exist $x'_0 \in X_0$ and $x'_0 \xrightarrow{u'_1} \cdots \xrightarrow{u'_{n+m}} x'_{n+m}$ such that $x'_n \in X \setminus X_S$ and $\max_{i\in\{0,1,\ldots,n+m\}} \mathbf{d}(H(x_i), H(x'_i)) \leq \delta$. Therefore, the system is not δ-approximate infinite-step opaque.

($\Leftarrow$) By contradiction: suppose that (5) holds and assume, for the sake of contradiction, that S is not δ-approximate infinite-step opaque. Then, we know that there exists an initial state $x_0 \in X_0$, a sequence of transitions $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$ and an index $k \in \{0, \ldots, n\}$ such that $x_k \in X_S$ and there does not exist an initial state $x'_0 \in X_0$ and a sequence of transitions $x'_0 \xrightarrow{u'_1} x'_1 \xrightarrow{u'_2} \cdots \xrightarrow{u'_n} x'_n$ such that $x'_k \in X \setminus X_S$ and $\max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_i)) \leq \delta$. Let us consider the following sequence of transitions in S_C:

$$(x_0, q_0) \xrightarrow[C]{u_1} (x_1, q_1) \xrightarrow[C]{u_2} \cdots \xrightarrow[C]{u_k} (x_k, q_k)$$

and the following sequence of transitions in S_I:

$$(x_n, q'_n) \xrightarrow[I]{u_n} (x_{n-1}, q'_{n-1}) \xrightarrow[I]{u_{n-1}} \cdots \xrightarrow[I]{u_{k+1}} (x_k, q'_k).$$

By Propositions IV.2 and IV.6, we know that

$$q_n \cap q'_n = \left\{ x'_k \in X : \begin{array}{c} \exists x'_0 \in X_0, \exists x'_0 \xrightarrow{u'_1} \cdots \xrightarrow{u'_n} x'_n \text{ s.t.} \\ \max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_i)) \leq \delta \end{array} \right\}.$$

Since (5) holds, we know that $q_n \cap q'_n \not\subseteq X_S$. Therefore, there exists $x'_0 \in X_0$ and a sequence of transitions $x'_0 \xrightarrow{u'_1} \cdots \xrightarrow{u'_n} x'_n$ such that $x_k \in X \setminus X_S$ and $\max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(H(x_i), H(x'_i)) \leq \delta$, which is a contradiction, i.e., S has to be δ-approximate infinite-step opaque. ■

REFERENCES

[1] K.-D. Kim and P. Kumar, "Cyber-physical systems: A perspective at the centennial," *Proc. IEEE*, vol. 100, no. Special Centennial Issue, pp. 1287–1308, May 2012.

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

[3] L. Mazaré, "Using unification for opacity properties," in *Proc. Workshop Issues Theory Secur.*, 2004, vol. 7, pp. 165–176.

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

[5] A. Saboori and C. Hadjicostis, "Verification of K-step opacity and analysis of its complexity," *IEEE Trans. Autom. Sci. Eng.*, vol. 8, no. 3, pp. 549–559, Jul. 2011.

[6] 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.

[7] A. Saboori and C. Hadjicostis, "Verification of initial-state opacity in security applications of discrete event systems," *Inf. Sci.*, vol. 246, pp. 115–132, 2013.

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

[9] Y. Wu and S. Lafortune, "Comparative analysis of related notions of opacity in centralized and coordinated architectures," *Discrete Event Dyn. Syst.*, vol. 23, no. 3, pp. 307–339, Sep. 2013.

[10] Y. Tong, Z. Li, C. Seatzu, and A. Giua, "Decidability of opacity verification problems in labeled Petri net systems," *Automatica*, vol. 80, pp. 48–53, 2017.

[11] Y. Tong, Z. Li, C. Seatzu, and A. Giua, "Verification of state-based opacity using Petri nets," *IEEE Trans. Autom. Control*, vol. 62, no. 6, pp. 2823–2837, Jun. 2017.

[12] X. Cong, M. Fanti, A. Mangini, and Z. Li, "On-line verification of current-state opacity by Petri nets and integer linear programming," *Automatica*, vol. 94, pp. 205–213, 2018.

[13] F. Basile and G. De Tommasi, "An algebraic characterization of language-based opacity in labeled Petri nets," in *Proc. 14th Int. Workshop Discrete Event Syst.*, 2018, pp. 329–336.

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

[15] C. Keroglou and C. Hadjicostis, "Probabilistic system opacity in discrete event systems," *Discrete Event Dyn. Syst.*, vol. 28, no. 2, pp. 289–314, 2018.

[16] B. Wu, Z. Liu, and H. Lin, "Parameter and insertion function co-synthesis for opacity enhancement in parametric stochastic discrete event systems," in *Proc. Amer. Control Conf.*, 2018, pp. 3032–3037.

[17] S. Chédor, C. Morvan, S. Pinchinat, and H. Marchand, "Diagnosis and opacity problems for infinite state systems modeled by recursive tile systems," *Discrete Event Dyn. Syst.*, vol. 25, nos. 1/2, pp. 271–294, 2014.

[18] K. Kobayashi and K. Hiraishi, "Verification of opacity and diagnosability for pushdown systems," *J. Appl. Math.*, vol. 2013, 2013, Art. no. 654059.

[19] 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.

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

[21] B. Bérard, J. Mullins, and M. Sassolas, "Quantifying opacity," *Math. Struct. Comput. Sci.*, vol. 25, no. 2, pp. 361–403, 2015.

[22] 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.

[23] 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.

[24] B. Ramasubramanian, R. Cleaveland, and S. Marcus, "A framework for opacity in linear systems," in *Proc. Amer. Control Conf.*, 2016, pp. 6337–6344.

[25] B. Ramasubramanian, R. Cleaveland, and S. Marcus, "A framework for decentralized opacity in linear systems," in *Proc. 54th Annu. Allerton Conf. Commun., Control, Comput.*, 2016, pp. 274–280.

[26] B. Ramasubramanian, R. Cleaveland, and S. Marcus, "Opacity for switched linear systems: Notions and characterization," in *Proc. 56th IEEE Conf. Decis. Control*, 2017, pp. 5310–5315.

[27] 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.

[28] K. Zhang, X. Yin, and M. Zamani, "Opacity of nondeterministic transition systems: A (bi)simulation relation approach," *IEEE Trans. Autom. Control*, vol. 64, no. 12, pp. 5116–5123, Dec. 2019.

[29] M. Noori-Hosseini, B. Lennartson, and C. Hadjicostis, "Compositional visible bisimulation abstraction applied to opacity verification," in *Proc. 14th Int. Workshop Discrete Event Syst.*, 2018, pp. 434–441.

[30] M. Noori-Hosseini, B. Lennartson, and C. Hadjicostis, "Incremental observer reduction applied to opacity verification and synthesis," 2018, *arXiv:1812.08083v3*.

[31] B. Wu and H. Lin, "Privacy verification and enforcement via belief abstraction," *IEEE Control Syst. Lett.*, vol. 2, no. 4, pp. 815–820, Oct. 2018.

[32] S. Mohajerani, Y. Ji, and S. Lafortune, "Compositional and abstraction-based approach for synthesis of edit functions for opacity enforcement," *IEEE Trans. Autom. Control*, to be published.

[33] A. Girard and G. J. Pappas, "Approximation metrics for discrete and continuous systems," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 782–798, May 2007.

[34] A. Girard, G. Pola, and P. Tabuada, "Approximately bisimilar symbolic models for incrementally stable switched systems," *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 116–126, Jan. 2010.

[35] G. Reissig, "Computing abstractions of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 56, no. 11, pp. 2583–2598, Nov. 2011.

[36] M. Zamani, G. Pola, M. Mazo, and P. Tabuada, "Symbolic models for nonlinear control systems without stability assumptions," *IEEE Trans. Autom. Control*, vol. 57, no. 7, pp. 1804–1809, Jul. 2012.

[37] M. Zamani, A. Abate, and A. Girard, "Symbolic models for stochastic switched systems: A discretization and a discretization-free approach," *Automatica*, vol. 55, pp. 183–196, 2015.

[38] C. Dwork, "Differential privacy," in *Encyclopedia of Cryptography and Security*. New York, NY, USA: Springer, 2011, pp. 338–340.

[39] J. Le Ny and G. Pappas, "Differentially private filtering," *IEEE Trans. Autom. Control*, vol. 59, no. 2, pp. 341–354, Feb. 2014.

[40] A. Jones, K. Leahy, and M. Hale, "Towards differential privacy for symbolic systems," in *Proc. Amer. Control Conf.*, 2019, pp. 372–377.

[41] K. Yazdani, A. Jones, K. Leahy, and M. Hale, "Differentially private LQ control," 2018, *arXiv:1807.05082v4*.

[42] G. Pola, E. De Santis, and M. Di Benedetto, "Approximate diagnosis of metric systems," *IEEE Control Syst. Lett.*, vol. 2, no. 1, pp. 115–120, Jan. 2018.

[43] G. Fiore, E. De Santis, G. Pola, and M. Di Benedetto, "On approximate predictability of metric systems," in *Proc. 6th IFAC Conf. Anal. Des. Hybrid Syst.*, 2018, pp. 169–174.

[44] X. Yin and M. Zamani, "Towards approximate opacity of cyber-physical systems," in *Proc. 10th ACM/IEEE Int. Conf. Cyber-Phys. Syst.*, 2019, pp. 310–311.

[45] P. Tabuada, *Verification and Control of Hybrid Systems: A Symbolic Approach*, 1st ed. New york, NY, USA: Springer, 2009.

[46] F. Cassez, J. Dubreil, and H. Marchand, "Synthesis of opaque systems with static and dynamic masks," *Formal Methods Syst. Des.*, vol. 40, no. 1, pp. 88–115, 2012.

[47] L. Clavijo and J. Basilio, "Empirical studies in the size of diagnosers and verifiers for diagnosability analysis," *Discrete Event Dyn. Syst.*, vol. 27, no. 4, pp. 701–739, 2017.

[48] D. N. Tran, "Advances in stability analysis for nonlinear discrete-time dynamical systems," Ph.D. dissertation, Dept. Elect. Eng., Univ. Newcastle, Callaghan, NSW, Australia, 2018.

[49] A. Swikir, A. Girard, and M. Zamani, "From dissipativity theory to compositional synthesis of symbolic models," in *Proc. Indian Control Conf.*, 2018, pp. 30–35.

[50] J. Dubreil, P. Darondeau, and H. Marchand, "Supervisory control for opacity," *IEEE Trans. Autom. Control*, vol. 55, no. 5, pp. 1089–1100, May 2010.
[51] B. Zhang, S. Shu, and F. Lin, "Maximum information release while ensuring opacity in discrete event systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 4, pp. 1067–1079, Jul. 2015.
[52] X. Yin and S. Lafortune, "A uniform approach for synthesizing property-enforcing supervisors for partially-observed discrete-event systems," *IEEE Trans. Autom. Control*, vol. 61, no. 8, pp. 2140–2154, Aug. 2016.
[53] Y. Tong, Z. Li, C. Seatzu, and A. Giua, "Current-state opacity enforcement in discrete event systems under incomparable observations," *Discrete Event Dyn. Syst.*, vol. 28, no. 2, pp. 161–182, 2018.
[54] Y. Ji, X. Yin, and S. Lafortune, "Enforcing opacity by insertion functions under multiple energy constraints," *Automatica*, vol. 108, 2019, Art. no. 108476.

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

In 2017, he joined the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, where he is currently an Associate Professor. His research interests include control and diagnosis of discrete-event systems, formal methods, security, and their applications to cyber and cyber-physical systems.

Dr. Yin received the Outstanding Reviewer Awards from the *Automatica*, the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and the *Journal of Discrete Event Dynamic Systems*. He also received the IEEE Conference on Decision and Control Best Student Paper Award Finalist in 2016. He is the Co-Chair of the IEEE Control Systems Society Technical Committee on Discrete Event Systems. He is also a member of the IEEE Control Systems Society Conference Editorial Board.

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, 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.

He is currently an Assistant Professor with the Department of Computer Science, University of Colorado Boulder, Boulder, CO, USA. Between September 2012 and December 2013, he was a Postdoctoral Researcher with the Delft Centre for Systems and Control, Delft University of Technology, Delft, The Netherlands. From 2014 to 2019, he was an Assistant Professor with the Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany. From 2013 to 2014, he was an Assistant Professor with the Design Engineering Department, Delft University of Technology. 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 received an ERC Starting Grant Award from the European Research Council in 2018.

Siyuan Liu (Student Member, IEEE) received the B.Sc. degree in automation science and the M.Eng. degree in control engineering from Beihang University, Beijing, China, in 2014 and 2017, respectively. She is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany.

Her current research interests include formal methods, security properties, and compositional methods for verification and control of large-scale cyber-physical systems.