Automatica 80 (2017) 127–134

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Initial-state detectability of stochastic discrete-event systems with probabilistic sensor failures☆

Xiang Yin [1]

Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China

ARTICLE INFO

Article history:
Received 4 April 2016
Received in revised form
23 January 2017
Accepted 27 January 2017

Keywords:
Discrete-event systems
State estimation
Initial-state detectability
PSPACE-completeness

ABSTRACT

Initial-state estimation is an important problem in discrete-event systems. In this problem, the initial-state of the system is unknown and one wants to determine the initial-state of the system based on its observation. In this paper, we investigate the problem of initial-state detection of partially-observed discrete-event systems in the stochastic setting. We consider two sources of randomness in this problem: stochastic dynamic of the system and probabilistic sensor failures. Specifically, we model a stochastic discrete-event system by a probabilistic finite-state automaton and we use a probabilistic projection function as the observation model. The notion of stochastic initial-state detectability (SI-detectability) is introduced in order to capture whether or not the probability of detecting the initial-state converges to one even in the presence of potential sensor failures. An approach for the verification of SI-detectability is proposed. We also investigate the complexity of the SI-detectability verification problem and we show that this problem is PSPACE-complete.

1. Introduction

State estimation is one of the most fundamental problems in Discrete-Event Systems (DES). In many applications, our information about the system is limited and we need to estimate the state of the system in order to make some decision. Due to its importance, the state estimation problem has received considerable attention in the DES literature; see, e.g., Cabasino, Hadjicostis, and Seatzu (2015), Özveren and Willsky (1990), Sears and Rudie (2014), Shu, Lin, and Ying (2007) and Yin and Lafortune (2015a). Recently, the state estimation problem has been studied systematically in the framework of detectability; see, e.g., Keroglou and Hadjicostis (2015), Shu and Lin (2013b) and Shu et al. (2007); Shu, Lin, Ying, and Chen (2008). Particularly, in Shu et al. (2007), the authors defined four types of detectability in order to capture different requirements in the current-state estimation problem. When the original system is not detectable, several approaches have also been proposed in order to actively enforce detectability, e.g., by sensor activations (Shu, Huang, & Lin, 2013; Yin & Lafortune, 2015b) and by supervisory control (Shu & Lin, 2013a; Yin & Lafortune, 2016).

One important class of state estimation problems is the *initial-state estimation problem*; see, e.g., Li and Hadjicostis (2013), Saboori and Hadjicostis (2013) and Shu and Lin (2013b). In this problem, we assume that the initial-state of the system is fully unknown or only partially known. Then, we want to *infer* the initial-state of the system by observing the output of the system. In particular, in Shu and Lin (2013b), the notion of I-detectability was proposed to capture whether or not we can always uniquely determine the initial-state of the system by observing a finite sequence of events. A polynomial-time algorithm for checking I-detectability was also provided in Shu and Lin (2013b). In Li and Hadjicostis (2013), an algorithm for recursively computing the minimum initial-marking of a Petri net was proposed. In Saboori and Hadjicostis (2013), the notion of initial-state opacity was investigated. This notion can be considered as the dual of I-detectability, since it requires that the intruder, which is modeled as an observer, can never determine the initial-state of the system.

Although the initial-state estimation problem and the notion of I-detectability have been studied in the literature, several important issues still remain. First, the notion of I-detectability is defined for logical DES. Specifically, it requires that we can always uniquely determine the initial-state of the system based on an arbitrarily long observation. However, this requirement sometimes is too strong when stochastic dynamic of the system

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Christoforos Hadjicostis under the direction of Editor Christos G. Cassandras.

E-mail address: xiangyin@umich.edu.

[1] Fax: +86 7347638041.

http://dx.doi.org/10.1016/j.automatica.2017.02.032
0005-1098/

is considered. For example, it is possible that the probability of detecting the initial-state of the system becomes arbitrarily close to one as we observe more and more events. Such an asymptomatic property is also very useful in practice, but I-detectability fails to capture this feature. Another issue is that, in all of the existing works on detectability, it is assumed that each observable event can be reliably observed. However, this assumption may not hold in practice, since sensors for observable events may fail. This leads to observation uncertainties and one has to handle this issue.

In order to address the above discussed issues, in this paper, we study the initial-state estimation problem for stochastic DES with probabilistic sensor failures. Specifically, this paper has the following contributions. First, we extend the initial-state estimation problem from logical DES to the stochastic setting. We consider a DES modeled by a probabilistic finite-state automaton in order to describe the stochastic dynamic of the system. Moreover, we model the observation of the system by a probabilistic projection function in order to address the issue of sensor failures. Namely, we assume that the occurrence of an observable event may not be observed with a given probability if its associated sensor is not reliable. Then, we define the notion of stochastic initial-state detectability (SI-detectability) with probabilistic sensor failures in order to capture whether or not the probability of detecting the initial-state goes to one. This notion is strictly weaker than logical I-detectability by taking the stochastic dynamic of the system into account. Then, we provide an approach to verify SI-detectability. This is based on the structure of robust initial-state estimator proposed in the paper. Finally, we also investigate the precise complexity of deciding SI-detectability. We show that, unlike I-detectability, which can be verified in polynomial-time, checking SI-detectability is PSPACE-complete.

Our work is related to several works in the literature; we discuss the differences between our work and these works. Property analysis of stochastic DES has been considered in many different works in the literature; see, e.g., Bertrand, Haddad, and Lefaucheux (2014), Chen and Kumar (2015a,b), Chen, Ibrahim, and Kumar (2016); Chen, Keroglou, Hadjicostis, and Kumar (2017), Keroglou and Hadjicostis (2013, 2015), Lunze and Schröder (2001), Saboori and Hadjicostis (2014), Shu et al. (2008) and Thorsley and Teneketzis (2005). The current-state detection problem in stochastic DES was studied in Keroglou and Hadjicostis (2015) and Shu et al. (2008) under the assumption that all sensors are reliable. In particular, the notion of A-detectability was proposed in Keroglou and Hadjicostis (2015). Our definition of SI-detectability is similar to A-detectability; both of them require to detect the (initial or current) state of the system *for sure* with probability one. However, we investigate the initial-state detection problem, which is different from the current-state detection problem. Moreover, we consider probabilistic sensor failures, which is also not considered in Keroglou and Hadjicostis (2015) and Shu et al. (2008). In Keroglou and Hadjicostis (2013), the notion of initial-state opacity was investigated in the stochastic setting. However, initial-state opacity and initial-state detectability are clearly incomparable; the former is an *always* property, while the latter is an *eventually* property. Moreover, Keroglou and Hadjicostis (2013) also assumes that the observation is always reliable. Regarding works on unreliable observations, Athanasopoulou, Li, and Hadjicostis (2010), Carvalho, Basilio, and Moreira (2012), Takai and Ushio (2012) and Thorsley, Yoo, and Garcia (2008) studied the sensor reliability issue in the fault diagnosis problem. In particular, Carvalho et al. (2012) and Takai and Ushio (2012) investigated the effect of intermittent sensor failure, which is stronger than the probabilistic sensor failure considered in this paper. The models used in Athanasopoulou et al. (2010) and Thorsley et al. (2008) are more related to our setting. Specifically, they also consider both stochastic dynamic of the system and probabilistic sensor failures. However, diagnosability is more related to the current-state estimation problem rather than the initial-state estimation problem. As a consequence, the verification procedure we propose in this paper is also very different from the approaches in Athanasopoulou et al. (2010) and Thorsley et al. (2008). Overall, all of the above mentioned works are clearly different from the problem considered in this paper. Our work provides a systematic study of the initial-state estimation problem under a fully stochastic framework with both stochastic system dynamic and probabilistic sensor outputs.

2. Initial-state detection with probabilistic sensor failures

2.1. System model

Let Σ be a finite set of events. A string is a finite sequence of events and Σ^* denotes the set of all finite strings over Σ, including the empty string ϵ. A language L is subset of Σ^*. For any string $s \in \Sigma^*$, $|s|$ denotes its length, where $|\epsilon| = 0$. We denote by $\bar{s}$ the set of prefixes of s, i.e., $\bar{s} = \{t \in \Sigma^* : \exists v \in \Sigma^* \text{ s.t. } tv = s\}$.

A nondeterministic finite-state automaton (NFA) is a 4-tuple

$$G = (X, \Sigma, \delta, X_0) \tag{1}$$

where X is the finite set of states, Σ is the finite set of events, $\delta : X \times \Sigma \to 2^X$ is the partial nondeterministic transition function, and X_0 is the set of initial-states. The transition function δ is also extended to $X \times \Sigma^*$ in the usual manner; see, e.g., Cassandras and Lafortune (2008). The language generated by G from state $x \in X$ is $\mathcal{L}(G, x) = \{s \in \Sigma^* : \delta(x, s)!\}$, where "!" means "is defined". Then, the language generated by G is $\mathcal{L}(G) = \cup_{x \in X_0} \mathcal{L}(G, x)$.

A probabilistic finite-state automaton (PFA) is a 6-tuple

$$\mathcal{G} = (X, \Sigma, \delta, X_0, \pi_0, p) \tag{2}$$

where (X, Σ, δ, X_0) is a NFA and we call this NFA the *support* of $\mathcal{G}$. Hereafter, we use G to denote the support of $\mathcal{G}$. Also, $\pi_0 : X_0 \to [0, 1]$ is the initial-states distribution vector such that $\sum_{x \in X_0} \pi_0(x) = 1$ and each element of π_0 is nonnegative, and $p : X \times \Sigma \times X \to [0, 1]$ is the state transition probability function. For any $x, x' \in X, \sigma \in \Sigma$, we write $p(x', \sigma \mid x)$ as the probability that event σ occurs from state x and leads to state x'. We assume that p satisfies the following requirements

1. $\forall x, x' \in X, \sigma \in \Sigma : x' \in \delta(x, \sigma) \Leftrightarrow p(x', \sigma \mid x) > 0$;
2. $\forall x \in X : \sum_{\sigma \in \Sigma} \sum_{x' \in X} p(x', \sigma \mid x) = 1$.

Note that these two requirements together also implicitly implies that the system is *live*, i.e., $\forall x' \in X, \exists \sigma \in \Sigma : \delta(x, \sigma)!$. Function p is also extended to $X \times \Sigma^* \times X$ inductively as follows: for any $s \in \Sigma^*, \sigma \in \Sigma$, we have

$$p(x', s\sigma \mid x) = \sum_{x'' \in X} p(x'', s \mid x) p(x', \sigma \mid x'').$$

Then, the probability that string $s \in \Sigma^*$ occurs from state $x \in X$ is $p(s \mid x) = \sum_{x' \in X} p(x', s \mid x)$ and the probability that string $s \in \Sigma^*$ occurs from any initial-state is $p(s) = \sum_{x_0 \in X_0} p(s \mid x_0)\pi_0(x_0)$. We also write $\mathcal{L}(\mathcal{G}) = \mathcal{L}(G)$ and $\mathcal{L}(\mathcal{G}, x) = \mathcal{L}(G, x)$, where NFA G is the support of PFA $\mathcal{G}$.

2.2. Observation model

In this paper, we consider a probabilistic observation model. Specifically, the observation is specified by a *probabilistic projection function* $M : \Sigma \to [0, 1]$, where $M(\sigma)$ denotes the probability of observing event σ when it occurs. Note that this observation probability is independent from the probability that σ occurs at some state. For any event $\sigma \in \Sigma$, we say that σ is (1) *unobservable*,

if $M(\sigma) = 0$; and (2) *reliable*, if $M(\sigma) = 1$; and (3) *unreliable*, if $0 < M(\sigma) < 1$. We denote by Σ_{uo}, Σ_{re} and Σ_{ur} the sets of unobservable, reliable and unreliable events, respectively. We also define $\Sigma_o = \Sigma \setminus \Sigma_{uo} = \Sigma_{re} \cup \Sigma_{ur}$. Clearly, this observation model is more general than the standard natural projection (Cassandras & Lafortune, 2008) that is widely used in logical DES, where each event is either unobservable or reliable.

Let $s = \sigma_1\sigma_2\cdots\sigma_n \in \mathcal{L}(\mathcal{G})$ be a string. Let $I = \{i_1, i_2, \ldots, i_k\}$ be a set of integers such that $1 \leq i_1 < \cdots < i_k \leq n$. We say that I is a *realization index set* for s (w.r.t. M) if, for any $i \in \{1, \ldots, n\}$, we have (1) $M(\sigma_i) = 1 \Rightarrow i \in I$; and (2) $M(\sigma_i) = 0 \Rightarrow i \notin I$. We denote by $I_M(s)$ the set of realization index sets for string s w.r.t. M. Then, for any string $s = \sigma_1\sigma_2\cdots\sigma_n \in \mathcal{L}(\mathcal{G})$, we say that $\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_k}$ is an *output realization* of s under M, if $\{i_1, \ldots, i_k\} \in I_M(s)$. We denote by $Pr(\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_k} \mid s)$ the probability of this output realization given the occurrence of string s. Note that two different output realizations may yield the same *observation*. For example, for string abb, where $a \in \Sigma_{re}$ and $b \in \Sigma_{ur}$, $ab\epsilon$ and $a\epsilon b$ are two different output realizations which have the same observation ab. (We use ϵ to denote that the corresponding observation at that place is lost.) Then, the probability that $\alpha \in \Sigma_o^*$ is observed given the occurrence of s is

$$Pr(\alpha \mid s) = \sum_{\{i_1,\ldots,i_k\} \in I_M(s):\, \sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_k} = \alpha} Pr(\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_k} \mid s).$$

We define a mapping $O : \Sigma^* \to 2^{\Sigma_o^*}$ by, for any $s \in \Sigma^*$, $\alpha \in O(s) \Leftrightarrow Pr(\alpha \mid s) > 0$. Mapping O is also extended to $O : 2^{\Sigma^*} \to 2^{\Sigma_o^*}$ by, for any $L \subseteq \Sigma^* : O(L) = \{\alpha \in \Sigma_o^* : \exists s \in L \text{ s.t. } \alpha \in O(s)\}$. Therefore, $O(\mathcal{L}(\mathcal{G}))$ is the set of all possible observations of the system. Note that mapping O is essentially equivalent to the language dilation operator defined in Carvalho et al. (2012) or the communication loss operator defined in Lin (2014) if we do not consider the observation probability $Pr(\alpha \mid s)$. Finally, we denote by $Pr(\alpha \mid x)$ the probability that $\alpha \in \Sigma_o^*$ is observed starting from state $x \in X$.

2.3. Initial-state detection problem

In this paper, we investigate the initial-state detection problem. Initially, our knowledge about the initial-state distribution is given by π_0. However, this knowledge can be improved by observing more events generated by the system. Specifically, the posterior probability that the initial-state of the system is $x \in X_0$ given $\alpha \in \Sigma_o^*$ observed is

$$\hat{\pi}_0(x \mid \alpha) = \frac{Pr(\alpha \mid x)\pi_0(x)}{\sum_{x_0 \in X_0} Pr(\alpha \mid x_0)\pi_0(x_0)}. \tag{3}$$

We say that the initial-state of the system is *detected* after observing $\alpha \in \Sigma_o^*$ if $\exists x \in X_0 : \hat{\pi}_0(x \mid \alpha) = 1$, i.e., for all $x' \in X \setminus \{x\} : \hat{\pi}_0(x' \mid \alpha) = 0$. We define a detectability function $D : \Sigma_o^* \to \{0, 1\}$ by, for any $\alpha \in \Sigma_o^*$

$$D(\alpha) = \begin{cases} 1 & \text{if } \exists x \in X_0 : \hat{\pi}_0(x \mid \alpha) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

i.e., the initial-state is detected after observing $\alpha \in \Sigma_o^*$ if and only if $D(\alpha) = 1$.

3. Stochastic initial-state detectability with probabilistic sensor failures

In Shu and Lin (2013b), the notion of I-detectability was introduced for logical (non-stochastic) DES in order to capture whether or not the initial-state of the system can be detected within a finite delay. Let $G = (X, \Sigma, \delta, X_0)$ be a NFA and suppose that all observable events are reliable, i.e., $\Sigma = \Sigma_{re} \dot{\cup} \Sigma_{uo}$. First, we recall the definition of I-detectability.

Definition 1 (*I-Detectability Shu & Lin, 2013b*)**.** NFA $G = (X, \Sigma, \delta, X_0)$ is said to be I-detectable w.r.t. $\Sigma_{re} \subseteq \Sigma$ if

$$(\exists n \in \mathbb{N})(\forall x \in X_0)(\forall \alpha \in P(\mathcal{L}(G, x)) : |\alpha| \geq n)[|\,\mathrm{In}(\alpha)| = 1]$$

where $\mathrm{In}(\alpha) = \{x \in X_0 : \exists s \in \mathcal{L}(G, x) \text{ s.t. } P_{re}(s) = \alpha\}$ and $P_{re} : \Sigma^* \to \Sigma_{re}^*$ is the natural projection.

Intuitively, I-detectability requires that there do not exist two arbitrarily long strings starting from two distinct initial-states such that the observations of these two strings are equivalent. The following example illustrates I-detectability and its drawback in the stochastic setting.

Example 1. Let us consider NFA G_1, which is the support of PFA $\mathcal{G}_1$ shown in Fig. 1(a). The initial-states are $X_0 = \{1, 2\}$. Suppose that $\Sigma_{re} = \Sigma = \{a, b\}$, i.e., all events can be reliably observed. Then, we know that G_1 is not I-detectable, since string $(ba)^n$ is defined at both initial-states 1 and 2 for any $n \in \mathbb{N}$. Now, let us assume that the transition probability function p for $\mathcal{G}_1$ is specified by the number associated with each transition in Fig. 1(a), e.g., we have $p(2, a \mid 1) = 0.1$. We also assume that $\pi_0(1) = \pi_0(2) = 0.5$. Then, we know that the probability that string $(ba)^n$ occurs is $p((ba)^n) = p((ba)^n \mid 1)\pi_0(1) + p((ba)^n \mid 2)\pi_0(2) = 0.5 \times (0.9 \times 1)^n + 0.5 \times (1 \times 0.9)^n = 0.9^n$, which goes to zero as n increases. Once two a (respectively, two b) are observed *in succession*, we know immediately that the initial-state of the system is state 1 (respectively, state 2). In other words, the probability of detecting the initial-state of $\mathcal{G}_1$ goes to one when the system executes infinite number of steps. □

The above example illustrates that I-detectability for logical DES may not be adequate for the initial-state detection problem in the stochastic setting even without considering the issue of probabilistic sensor failures. In order to resolve this issue, we introduce the notion of stochastic initial-state detectability (SI-detectability) with probabilistic sensor failures.

Definition 2 (*SI-Detectability*)**.** Let $\mathcal{G} = (X, \Sigma, \delta, X_0, \pi_0, p)$ be a PFA and $M : \Sigma \to [0, 1]$ be a probabilistic projection function. We say that $\mathcal{G}$ is SI-detectable w.r.t. M if

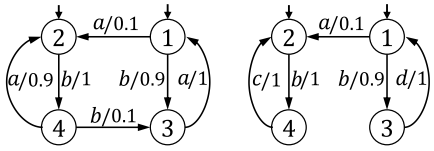
$$(\forall \Delta > 0)(\exists n \in \mathbb{N}) \quad \text{s.t. } Pr[s \in \mathcal{L}(\mathcal{G}) : ND \wedge |s| = n] < \Delta$$

where we have $Pr[s \in \mathcal{L}(\mathcal{G}) : ND \wedge |s| = n] = \sum_{s \in \Sigma^* : |s| = n} \sum_{\alpha \in O(s)} (1 - D(\alpha)) Pr(\alpha \mid s) p(s)$.

We make several comments on SI-detectability.

Remark 1. Intuitively, SI-detectability requires that the probability of detecting the initial-state of the system will converge to one as the length of the string generated by the system increases. One can easily verify that SI-detectability for a PFA is strictly weaker than I-detectability for its support even without considering probabilistic observations. For example, we have shown in Example 1 that PFA $\mathcal{G}_1$ is SI-detectable although its support G_1 is not I-detectable. □

Remark 2. In Keroglou and Hadjicostis (2015), the notion of A-detectability is defined for the *current-state* detection problem. Specifically, under the assumption that all sensors are reliable, A-detectability requires $\forall \Delta > 0, \exists N \in \mathbb{N} : Pr(\{s \in \Sigma^* : |s| \geq N, |R(X_0, O(s))| > 1\}) < \Delta$, where $R(X_0, O(s))$ is the current-state estimate of the system. Comparing SI-detectability with A-detectability, one can easily verify that these two notions are incomparable; none of them implies the other. On the other hand, these two notions do bear some similarities, since in both of these two problems, we need to know the (current or initial) state of the system for sure with probability one. Similar criteria in the form of $\forall \Delta > 0, \exists N \in \mathbb{N}$ are also used in the fault diagnosis (Thorsley &



(a) $\mathcal{G}_1$. (b) $\mathcal{G}_2$.

Fig. 1. Examples for stochastic initial-state detectability. The number associated with each transition denotes the probability of this transition rather than the observation probability.

Teneketzis, 2005) and the fault prognosis (Chen & Kumar, 2015b) of stochastic DES. However, our definition focuses on the initial-state detection problem rather than the current-state detection problem. Moreover, we consider the issue of observation uncertainty, which is not considered in Keroglou and Hadjicostis (2015).

Remark 3. We note that I-detectability is defined in terms of the length of the *observed* string, while our SI-detectability is defined in terms of the length of the *generated* string. The reason why we choose the generated string to defined SI-detectability is that, for any string s, the length of its observation may not be unique under the probabilistic projection function.

Remark 4. The probabilistic sensor failure model we use in this paper is also different from the intermittent sensor failure model used in Carvalho et al. (2012) and Takai and Ushio (2012). Specifically, in the intermittent sensor model, we need to consider the worst-case where failure always occurs in each unreliable sensor. However, such a worst-case analysis may be too strong in practice. For example, let us consider PFA $\mathcal{G}_2$ shown in Fig. 1(b) and assume that $\pi_0(1) = \pi_0(2) = 0.5, M(a) = M(b) = 1, M(c) = 0.5$ and $M(d) = 0$. By using the intermittent sensor failure model, we need to consider the case where we always cannot observe the occurrence of event c. Therefore, we cannot detect the initial state since string b^n can be observed from both initial-states 1 and 2 for any $n \in \mathbb{N}$. However, $\mathcal{G}_2$ is SI-detectable, since the probability that the sensor for c always fails goes to zero as the length of the string increases. Specifically, for initial-state 1, reliable event a will eventually occur due to the stochastic dynamic of the system and its occurrence will reveal this initial-state. For initial-state 2, we know that c will eventually be observed and its occurrence will also reveal this initial-state. This example also illustrates that the stochastic dynamic of the system and the probabilistic observation play different roles in our problem.

4. Verification of SI-detectability

4.1. Robust initial-state estimator

In order to check SI-detectability, the first question is how to determine whether or not $D(\alpha) = 1$ after observing $\alpha \in O(\mathcal{L}(\mathcal{G}))$. Note that this cannot be done by directly using the initial-state estimator proposed in Saboori and Hadjicostis (2013) and Shu and Lin (2013b) since we also need to consider the observation uncertainties. To resolve this issue, we propose the structure of *robust initial-state estimator* that estimates all possible initial-states in the presence of unreliable sensors. First, we introduce some necessary definitions.

Mapping $\Xi : \Sigma_o^* \to 2^{X\times X}$ is defined by: for any $\alpha \in \Sigma_o^*$,

$$\Xi(\alpha) = \{(x, x') \in 2^{X\times X} : \exists s \in \Sigma^* \text{ s.t. } \alpha \in O(s) \wedge x' \in \delta(x, s)\}.$$

Composition operator $\circ : 2^{X\times X} \times 2^{X\times X} \to 2^{X\times X}$ is defined by: for any $q_1, q_2 \in 2^{X\times X}$, we have $q_1 \circ q_2 = \{(x_1, x_3) \in 2^{X\times X} : \exists x_2 \in X \text{ s.t. } (x_1, x_2) \in q_1 \wedge (x_2, x_3) \in q_2\}$. We are now ready to introduce the robust initial-state estimator.

Definition 3 (*Robust Initial-State Estimator*)**.** Let $\mathcal{G}$ be a PFA and M be a probabilistic projection function. Let G be the support of $\mathcal{G}$ and $\Sigma = \Sigma_{re} \dot\cup \Sigma_{ur} \dot\cup \Sigma_{uo}$ Then, the robust initial-state estimator for $\mathcal{G}$ and M is the deterministic finite-state automaton (DFA) $G_{obs} = (Q_{obs}, \Sigma_o, \delta_{obs}, q_{obs,0})$, where $Q_{obs} \subseteq 2^{X\times X}$ is the set of states, $\delta_{obs} : Q_{obs} \times \Sigma_o \to Q_{obs}$ is the transition function such that, $\forall q \in Q_{obs}, \sigma \in \Sigma_o : \delta_{obs}(q, \sigma) = q \circ \Xi(\sigma)$ and the initial-state is defined by $q_{obs,0} = \Xi(\epsilon) \cap (X_0 \times X)$. For the sake of simplicity, we only consider the reachable part of G_{obs}.

Remark 5. For each $(x, x') \in 2^{X\times X}$, we call x the starting state and call x' the ending state. Intuitively, G_{obs} tracks all possible pairs of starting state and ending state that are consistent with the observation. The difference between the robust initial-state estimator and the initial-state estimator without observation uncertainty (Saboori & Hadjicostis, 2013; Shu & Lin, 2013b) is that, here we need to treat an unreliable event σ as both observable event and unobservable event. Specifically, the possibility that the sensors for events in Σ_{ur} fail is considered in $\Xi(\sigma)$, where mapping O is used. On the other hand, the transition function δ_{obs} is defined for all events in $\Sigma_{re} \cup \Sigma_{ur}$. This essentially allows us to "encode" the sensor reliability issue into the plant model. A similar feature also exists in the stochastic fault diagnosis problem; see, e.g., Thorsley et al. (2008).

In order to compute $\Xi(\sigma)$, we construct a new NFA $\hat{G} = (X, \hat\Sigma, \hat\delta, X_0)$ as follows. The event set is $\hat\Sigma = \Sigma \cup \hat\Sigma_{ur}$, where $\hat\Sigma_{ur} = \{\hat\sigma : \sigma \in \Sigma_{ur}\}$ is a set of new events. The transition function $\hat\delta$ is obtained as follows. First, we copy all transitions in δ. Then, for any $\sigma \in \Sigma_{ur}$ and any $x' \in \delta(x, \sigma)$, we add a new transition $x' \in \hat\delta(x, \hat\sigma)$. We define $Re : \hat\Sigma \to \Sigma$ as the function that renames events in $\hat\Sigma$ back to Σ, i.e., $Re(\hat\sigma) = \sigma$ if $\hat\sigma \in \hat\Sigma_{ur}$ and $Re(\sigma) = \sigma$ if $\sigma \in \Sigma$. We denote by $P_o : \hat\Sigma^* \to \Sigma_o^*$ the natural projection. Then, we have $\forall s \in \mathcal{L}(\hat G) : P_o(s) \in O(Re(s))$. This is because that the possibility that an unreliable event $\sigma \in \Sigma$ may fail has been taken care by the corresponding event $\hat\sigma \in \hat\Sigma_{ur}$, which is added in parallel with σ. Then, for any x, we have

$$\{x' \in X : \exists s \in \Sigma^* \text{ s.t. } \alpha \in O(s) \wedge x' \in \delta(x, s)\}$$
$$= \{x' \in X : \exists s \in \hat\Sigma^* \text{ s.t. } \alpha \in P_o(s) \wedge x' \in \hat\delta(x, s)\}.$$

Therefore, $\Xi(\sigma)$ can be computed by taking the standard unobservable reach in $\hat G$ under natural projection P_o. Similar constructions are also used in Carvalho et al. (2012) and Lin (2014) for different purposes.

Example 2. Let us consider PFA $\mathcal{G}_2$ shown in Fig. 1(b), where $\pi_0(1) = \pi_0(2) = 0.5$. Suppose that the projection function M is defined by $M(a) = M(b) = 1$ and $M(c) = M(d) = 0.5$, i.e., $\Sigma_{re} = \{a, b\}$ and $\Sigma_{ur} = \{c, d\}$. Then, the robust initial-state estimator G_{obs} is shown in Fig. 2. Initially, since all events defined at X_0 are reliable, we have that $q_{obs,0} = \Xi(\epsilon) \cap (X_0 \times X) = \{(1, 1), (2, 2)\}$. Once event b is observed, we know that $\Xi(b) = \{(1, 1), (1, 3), (2, 2), (2, 4)\}$ and we have $\delta_{obs}(q_{obs,0}, b) = \{(1, 1), (2, 2)\} \circ \{(1, 1), (1, 3), (2, 2), (2, 4)\} = \{(1, 1), (1, 3), (2, 2), (2, 4)\}$. □

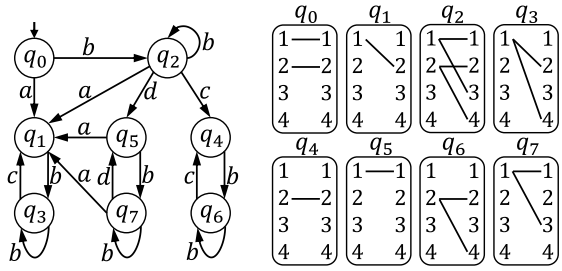


Fig. 2. The robust initial-state estimator G_{obs} for $\mathcal{G}_2$. For each $q \in 2^{X\times X}$, we connect state x on the LHS with state x' on the RHS to denote that $(x, x') \in q$.

For any state $q \in Q_{obs}$, we denote by $S(q)$ the set of starting states in q, i.e., $S(q) = \{x_1 \in X : \exists x_2 \in X \text{ s.t. } (x_1, x_2) \in q\}$. By definition, we know that $\mathcal{L}(G_{obs}) = O(\mathcal{L}(\mathcal{G}))$, i.e., any potential observation is defined in G_{obs}. The following result reveals that G_{obs} correctly estimates the initial-state of the system.

Proposition 1. *For any $\alpha \in O(\mathcal{L}(\mathcal{G}))$, we have*

$$S(\delta_{obs}(q_{obs,0}, \alpha)) = \{x \in X_0 : \hat{\pi}_0(x \mid \alpha) > 0\}. \tag{5}$$

4.2. Verification algorithm

Let $G_{obs} = (Q_{obs}, \Sigma_o, \delta_{obs}, q_{obs,0})$ be the robust initial-state estimator. We construct the DFA $\hat{G}_{obs} = (Q_{obs}, \hat{\Sigma}, \hat{\delta}_{obs}, q_{obs,0})$ by adding a self-loop at each state in Q_{obs} for each event in $\hat{\Sigma}_{ur} \cup \Sigma_{uo}$. Next, we construct the product of $\hat{G}$ and $\hat{G}_{obs}$ denoted by $G_{aug} = (Q_{aug}, \hat{\Sigma}, \delta_{aug}, Q_{aug,0}) := \hat{G} \times \hat{G}_{obs}$, where "$\times$" is the usual product composition operation of automata; see, e.g., Cassandras and Lafortune (2008). The product automaton G_{aug} has the following property. Suppose that a string $s \in \mathcal{L}(G_{aug})$ leads to $(x, q) \in Q_{aug}$. Then, it implies that the string generated by the system is $Re(s)$, which leads to x in G, and the observation of $Re(s)$ is $P_o(s)$, which leads to q in G_{obs}. We call a state $(x, q) \in Q_{aug}$ in G_{aug} a *certain* state if $S(q)$ is a singleton; otherwise, we call it an uncertain state. We denote by $Q_{cer} \subseteq Q_{aug}$ the set of certain states, i.e., $Q_{cer} = \{(x, q) \in Q_{aug} : |S(q)| = 1\}$.

A *strongly connected component* (SCC) in G_{aug} is a maximal set of states $v \subseteq Q_{aug}$ such that $\forall q_1, q_2 \in v, \exists s \in \hat{\Sigma}^* : q_2 \in \delta_{aug}(q_1, s)$. We denote by $\{v_1, \ldots, v_m\}$ the set of SCCs in G_{aug}. Then, we construct a new NFA

$$T = (V, \hat{\Sigma}, \delta_T, V_0) \tag{6}$$

where

- $V = \{v_1, \ldots, v_m\}$ is the set of SCCs in G_{aug};
- $\delta_T : V \times \hat{\Sigma} \to 2^V$ is the nondeterministic transition function defined by: for any $v_1, v_2 \in V, \sigma \in \hat{\Sigma}$, we have $v_2 \in \delta_T(v_1, \sigma)$ if

  $$\exists q_1 \in v_1, \exists q_2 \in v_2 : [q_2 \in \delta_{aug}(q_1, \sigma)] \wedge [v_1 \neq v_2].$$

- The set of initial-states V_0 is the set of SCCs in G_{aug} that contain a state in $Q_{aug,0}$, i.e., $V_0 = \{v \in V : v \cap Q_{aug,0} \neq \emptyset\}$.

Moreover, for each SCC $v \in V$, we say that

- v is a *certain SCC*, if $v \subseteq Q_{cer}$, and we denote by $V_{cer} \subseteq V$ the set of certain SCCs.
- v is an *uncertain SCC* if, $v \not\subseteq Q_{cer}$, and we denote by $V_{unc} \subseteq V$ the set of uncertain SCCs.
- v is a *terminal SCC* if, $\forall \sigma \in \hat{\Sigma} : \delta_T(v, \sigma) \neg !$, where "$\neg !$" means "is not defined", and we denote by $V_{ter} \subseteq V$ the set of terminal SCCs.

Example 3. We still consider PFA $\mathcal{G}_2$ shown in Fig. 1(b) and function M defined in Example 2. The robust initial-state estimator G_{obs} has been shown in Fig. 2. The corresponding $\hat{G}$, $\hat{G}_{obs}$ are shown in Fig. 3(a) and (b), respectively. Specifically, $\hat{G}$ is obtained by adding new transitions labeled with $\hat{c}$ and $\hat{d}$ to G in parallel with transitions label with unreliable events c and d, respectively. The product NFA G_{aug} is also shown in Fig. 3(c), which has seven SCCs; each set of states in a dashed rectangular in Fig. 3(c) represents a SCC. Then, NFA T is just the NFA by considering each dashed rectangular in Fig. 3(c) as a single state. For example, we know that state $(4, q_2) \in G_{aug}$ is not a certain state, since $S(q_2) = \{1, 2\}$. Therefore, SCC v_3 is an uncertain SCC. However, SCCs v_6 and v_7 are certain and both of them are also terminal SCCs. □

Remark 6. Note that NFA T is acyclic, i.e., there is no cycle in T, since states in the same cycle are merged into the same SCC. We also note that, if a SCC $v \in V$ contains one certain state (respectively, uncertain state), then we know that all states in this SCC are certain (respectively, uncertain). Namely, $V = V_{cer} \dot{\cup} V_{unc}$, and, for each SCC $v \in V$, $v \not\subseteq Q_{cer}$ and $v \cap Q_{cer} = \emptyset$ are equivalent. This is because that, once we detect the initial-state, we will detect the initial-state forever. □

Based on NFA T, we are now ready to present the main result to verify SI-detectability.

Theorem 1. *Let $\mathcal{G}$ be a PFA and M be a probabilistic projection function. Let T be the acyclic NFA constructed from $G_{aug} = \hat{G} \times \hat{G}_{obs}$. Then, $\mathcal{G}$ is SI-detectable w.r.t. M, if and only if, $V_{ter} \cap V_{unc} = \emptyset$, i.e., any terminal SCC is certain.*

The following example illustrates how to use Theorem 1 to verify SI-detectability.

Example 4. Let us still consider PFA $\mathcal{G}_2$ shown in Fig. 1(b), where $\pi_0(1) = \pi_0(2) = 0.5$ and M is defined by $M(a) = M(b) = 1$ and $M(c) = M(d) = 0.5$. The NFA T has been shown in Fig. 3(c). Since the only two terminal SCCs v_6 and v_7 are certain, by Theorem 1, $\mathcal{G}_2$ is SI-detectable w.r.t. M. □

Remark 7. In Keroglou and Hadjicostis (2015), a Markov-chain-based approach was proposed for the verification of A-detectability. Specifically, the approach in Keroglou and Hadjicostis (2015) evaluates some properties on the set of *recurrent states* in a Markov-chain. Moreover, the trellis-based initial-state estimation is also related to the approach in Shu and Lin (2013b) that augments the system model by tracking the initial states; both of them essentially require to record a pair of states. One may also use similar ideas to verify SI-detectability by constructing a Markov-chain based on G_{aug} and estimating the initial-states based on the augmented model. In fact, computing all recurrent states in a Markov-chain is equivalent to computing the set of terminal SCCs. Hence, this alternative approach should be equivalent to our approach, which verifies SI-detectability *directly* based on the *structural analysis* of G_{aug}.

Remark 8. Theorem 1 also reveals a *structural property* of SI-detectability. In particular, we see that SI-detectability does not depend on the specific transition probability of $\mathcal{G}$ or the specific value of M. Instead, it only depends on the support of $\mathcal{G}$, i.e., G, and the partition on Σ induced by M, i.e., which events are reliable, unreliable or unobservable. In other words, given a threshold Δ, without changing the support of $\mathcal{G}$ and the partition on Σ induced by M, modifying the values of p and M will only affect the value of the corresponding integer n. However, it will not affect the *existence* of such an integer. Namely, the probability of detecting the initial-state will still converge to one but with a different

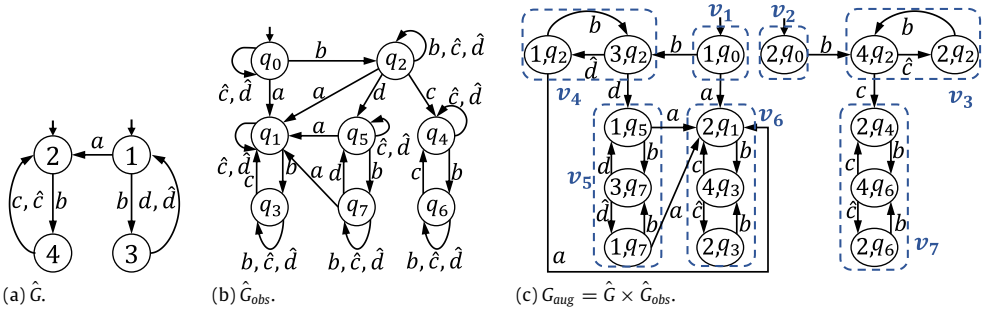


Fig. 3. Examples for the verification of SI-detectability, where $\Sigma_{re} = \{a, b\}$ and $\Sigma_{ur} = \{c, d\}$.

rate of convergence. This structural property is very useful in practice, since in many cases, we only know that a transition or an observation is possible, but it may be very hard to obtain the precise value of p or M.

We conclude this section by discussing the complexity of the proposed approach for verifying SI-detectability. First, we need to construct G_{obs}, which has $2^{|X_0|\times|X|}$ states and $|\Sigma|2^{|X_0|\times|X|}$ transitions in the worst-case. Constructing $\hat{G}$ and $\hat{G}_{obs}$ are linear in the sizes of G and G_{obs}, respectively. Therefore, in the worst case, $G_{aug} = \hat{G} \times \hat{G}_{obs}$ has $|X|2^{|X_0|\times|X|}$ states and $|\Sigma|\,|X|2^{|X_0|\times|X|}$ transitions. Computing all SCCs for G_{obs} is linear in the size of G_{obs}. Therefore, the overall complexity is $O(|\Sigma|\,|X|2^{|X_0|\times|X|})$, which is exponential in the size of $\mathcal{G}$. However, we will show later that this exponential complexity is unavoidable.

5. The complexity of SI-detectability

So far, we have provided an approach for verifying SI-detectability. However, the complexity of our approach is exponential in the size of $\mathcal{G}$. It was shown in Shu and Lin (2013b) that, for logical DES, the notion of I-detectability can be checked in polynomial-time. Unfortunately, we show in this section that checking SI-detectability is PSPACE-complete, which means that it is highly unlikely that such a polynomial-time algorithm exists. Note that, in Keroglou and Hadjicostis (2015), the authors also show that verifying A-detectability is PSPACE-hard by reducing the *Universality Problem for NFA* to the A-detectability verification problem. However, the universality problem is more related to the current-state estimation problem rather than the initial-state estimation problem. Hereafter, we provide a different approach for establishing the complexity result.

It is well-known that the *Language Equivalence Problem for NFAs* are PSPACE-complete, which is stated as follows.

Theorem 2 (*Stockmeyer & Meyer, 1973*)**.** *Let A and B be two NFAs with unique initial-states $x_{A,0}$ and $x_{B,0}$, respectively. Deciding whether or not $\mathcal{L}(A) = \mathcal{L}(B)$ is PSPACE-complete.*

We will not directly use the language equivalence problem for NFAs to prove the PSPACE-hardness of the SI-detectability verification problem. Instead, we consider a variation of this problem. Note that, to test whether or not $\mathcal{L}(A) = \mathcal{L}(B)$, it suffices to test whether or not $\mathcal{L}(A) \subseteq \mathcal{L}(B)$ and $\mathcal{L}(B) \subseteq \mathcal{L}(A)$. First, we have the following corollary.

Corollary 1. *Let A and B be two NFAs with unique initial-states $x_{A,0}$ and $x_{B,0}$, respectively. Deciding whether or not $\mathcal{L}(A) \subseteq \mathcal{L}(B)$ is PSPACE-hard.*

For any two languages $L_1, L_2 \subseteq \Sigma^*$, we say that L_1 and L_2 are comparable if $L_1 \subseteq L_2$ or $L_2 \subseteq L_1$. Then, we also have the following result.

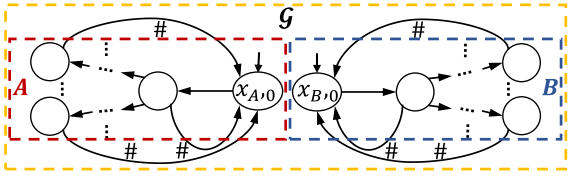


Fig. 4. Conceptual illustration of how to construct $\mathcal{G}$ from A and B.

Corollary 2. *Let A and B be two NFAs with unique initial-states $x_{A,0}$ and $x_{B,0}$, respectively. Deciding whether or not $\mathcal{L}(A)$ and $\mathcal{L}(B)$ are comparable is PSPACE-hard.*

Proof. We reduce the language inclusion problem to the language comparison problem. Let A and B be two NFAs, where $\mathcal{L}(A), \mathcal{L}(B) \subseteq \Sigma^*$ and $x_{0,A}$ and $x_{0,B}$ are the unique initial-states of A and B, respectively. We construct a new NFA B' by adding a self-loop transition labeled with # at the initial-state $x_{0,B}$ in B, where $\# \notin \Sigma$ is a new event. Then, we claim that $\mathcal{L}(A) \subseteq \mathcal{L}(B)$ if and only if $\mathcal{L}(A)$ and $\mathcal{L}(B')$ are comparable.

($\Rightarrow$) By construction, we have that $\mathcal{L}(B) \subset \mathcal{L}(B')$. Therefore, $\mathcal{L}(A) \subseteq \mathcal{L}(B)$ implies that $\mathcal{L}(A) \subset \mathcal{L}(B')$, which means that $\mathcal{L}(A)$ and $\mathcal{L}(B')$ are comparable.

($\Leftarrow$) Suppose that $\mathcal{L}(A)$ and $\mathcal{L}(B')$ are comparable, i.e., $\mathcal{L}(B') \subseteq \mathcal{L}(A)$ or $\mathcal{L}(A) \subseteq \mathcal{L}(B')$. Since $\#^* \in \mathcal{L}(B') \setminus \mathcal{L}(A)$, we know that the only possibility is $\mathcal{L}(A) \subseteq \mathcal{L}(B')$. This implies that $\mathcal{L}(A) \subseteq \mathcal{L}(B') \cap \Sigma^* = \mathcal{L}(B)$. □

Using the above corollary, we are now ready to show that checking SI-detectability is PSPACE-complete.

Theorem 3. *Let $\mathcal{G}$ be a PFA and M be a probabilistic projection function. Deciding whether or not $\mathcal{G}$ is SI-detectable w.r.t. M is PSPACE-complete.*

Proof. It is in PSPACE, since we can check the condition in Theorem 1 by constructing G_{aug} and T on the fly in a nondeterministic manner, which only requires polynomial space. Then, by the Savitch's theorem (Savitch, 1970), we know that it is in PSPACE.

Next, we show that this problem is PSPACE-hard by reducing the language comparison problem to the SI-detectability verification problem. Let $A = (X_A, \Sigma, \delta_A, x_{0,A})$ and $B = (X_B, \Sigma, \delta_B, x_{0,B})$ be two NFAs with unique initial-states $x_{A,0}$ and $x_{B,0}$, respectively. We assume that A and B are both live. Otherwise, we can add a self-loop with a new symbol at each state in A and B; this will not affect the result. We construct a PFA $\mathcal{G} = (X, \Sigma \cup \{\#\}, \delta, X_0, \pi_0, p)$, where $\# \notin \Sigma$ is a new event. Its support $G = (X, \Sigma \cup \{\#\}, \delta, X_0)$ is obtained as follows. First, we take the union of A and B, i.e., $X = X_A \cup X_B$, $X_0 = \{x_{A,0}, x_{B,0}\}$ and δ is consistent with δ_A and δ_B. Then, we add a transition labeled with # from each state in $X_A \setminus \{x_{A,0}\}$ (respectively, $X_B \setminus \{x_{B,0}\}$) to initial-state $x_{A,0}$ (respectively, $x_{B,0}$). A conceptual illustration of this construction is shown in Fig. 4. The initial-state distribution is $\pi_0(x_{A,0}) = \pi_0(x_{A,0}) = 0.5$. We set the transition probability by uniform distribution at each state, i.e.,

$\forall x, x' \in X, \sigma \in \Sigma \cup \{\#\} : p(x', \sigma \mid x) = \frac{1}{|Tran(x)|}$, where $Tran(x) = \{(x, \sigma, x') \in X \times (\Sigma \cup \{\#\}) \times X : x' \in \delta(x, \sigma)\}$ is the set of transitions defined at x. The projection function M is defined by $\forall \sigma \in \Sigma \cup \{\#\} : M(\sigma) = 1$, i.e., all events are reliable. Note that the size of $\mathcal{G}$ is linear in the sizes of A and B. Hereafter, we show that $\mathcal{L}(A)$ and $\mathcal{L}(B)$ are comparable if and only if $\mathcal{G}$ is not SI-detectable w.r.t. M.

($\Rightarrow$) Suppose that $\mathcal{L}(A)$ and $\mathcal{L}(B)$ are comparable. We assume without loss of generality that $\mathcal{L}(A) \subseteq \mathcal{L}(B)$. Then, we know that, for any $n \in \mathbb{N}$, we have $\sum_{s \in \mathcal{L}(\mathcal{G}, x_{A,0}) : |s| = n} (1 - D(s)) Pr(s \mid x_{A,0}) = 1$. This is because, for any $s \in \mathcal{L}(\mathcal{G}, x_{A,0})$, there exists $s \in \mathcal{L}(\mathcal{G}, x_{B,0})$. Therefore, by choosing $\Delta = \pi_0(x_{A,0}) - \varepsilon$, where ε is an arbitrarily small number, we know that $\mathcal{G}$ is not SI-detectable w.r.t. M.

($\Leftarrow$) By contraposition. Suppose that $\mathcal{L}(A)$ and $\mathcal{L}(B')$ are not comparable, i.e., $\mathcal{L}(B) \not\subset \mathcal{L}(A)$ or $\mathcal{L}(A) \not\subset \mathcal{L}(B)$. Let $s \in \mathcal{L}(A) \setminus \mathcal{L}(B)$ and $t \in \mathcal{L}(B) \setminus \mathcal{L}(A)$. Clearly, whenever s or $w\#s$ occurs, where $w \in (\Sigma \cup \{\#\})^*$, we know immediately that the initial-state is $x_{A,0}$; similarly, whenever t or $w\#t$ occurs, where $w \in (\Sigma \cup \{\#\})^*$, we know immediately that the initial-state is $x_{B,0}$. Due to the presence of event #, we can always reset to initial-states infinite often. Therefore, if the system starts from initial-state $x_{A,0}$, then the probability that string s or $w\#s$ occurs goes to one as the length of the generated string increases. Similarly, if the system starts from initial-state $x_{B,0}$, then the probability that string t or $w\#t$ occurs goes to one as the length of the generated string increases. Therefore, the probability of detecting the initial-state goes to one, i.e., $\mathcal{G}$ is SI-detectable w.r.t. M. □

6. Conclusion

In this paper, we investigated the initial-state detectability problem in the stochastic setting. Both stochastic dynamic of the system and probabilistic sensor failures were considered. The notion of SI-detectability was introduced in order to capture whether or not the probability of detecting the initial-state converges to one. An algorithm for the verification of SI-detectability was provided. Finally, we proved that checking SI-detectability is PSPACE-complete.

Appendix. Proofs not contained in main body

Proof of Proposition 1. First, we claim that

$$\delta_{obs}(q_{obs,0}, \alpha) = \{(x_1, x_2) \in 2^{X \times X} : \exists x_1 \in X_0, x_2 \in X, \exists s \in \Sigma^* \text{ s.t. } \alpha \in O(s) \wedge x_2 \in \delta(x_1, s)\}. \quad (A.1)$$

We prove this claim by induction on the length of $\alpha \in \Sigma_o^*$.

For $|\alpha| = 0$, i.e., $\alpha = \epsilon$, we know that Eq. (A.1) holds by the definition of $q_{obs,0}$. Let us assume that Eq. (A.1) holds for $|\alpha| = k$. Then, we need to show that Eq. (A.1) still holds for $\alpha' = \alpha\sigma$, where $|\alpha| = k$ and $\sigma \in \Sigma_o$. We have that

$$\begin{aligned}
&\delta_{obs}(q_{obs,0}, \alpha\sigma) = \delta_{obs}(q_{obs,0}, \alpha) \circ \Xi(\sigma) \\
&= \{(x_1, x_2) \in 2^{X \times X} : \exists x_1 \in X_0, x_2 \in X, \exists s \in \Sigma^* \\
&\quad \text{s.t. } \alpha \in O(s) \wedge x_2 \in \delta(x_1, s)\} \\
&\circ \{(x_2, x_3) \in 2^{X \times X} : \exists s \in \Sigma^* \text{ s.t. } \sigma \in O(s) \wedge x_3 \in \delta(x_2, s)\} \\
&= \{(x_1, x_3) \in 2^{X \times X} : \exists x_1 \in X_0, x_3 \in X, \exists s \in \Sigma^* \\
&\quad \text{s.t. } \alpha\sigma \in O(s) \wedge x_3 \in \delta(x_1, s)\}.
\end{aligned}$$

Therefore, we know that Eq. (A.1) always holds. Then, by the definition of $\hat{\pi}_0(x \mid \alpha)$, we know that $\hat{\pi}_0(x \mid \alpha) > 0$ iff $Pr(\alpha \mid x) > 0$ and $\pi_0(x) > 0$. Moreover, $Pr(\alpha \mid x) > 0$ iff $\exists s \in \mathcal{L}(\mathcal{G}, x) : \alpha \in O(s)$. Therefore, $\{x \in X_0 : \hat{\pi}_0(x \mid \alpha) > 0\} = \{x \in X_0 : \exists s \in \mathcal{L}(\mathcal{G}, x) \text{ s.t. } \alpha \in O(s)\} = S(\delta_{obs}(q_{obs,0}, \alpha))$. □

Proof of Theorem 1. ($\Rightarrow$) By contraposition. Suppose that there exists an uncertain terminal SCC in T; say $v \in V$. Let $(x, q) \in v$ be an uncertain state in v and let $s \in \mathcal{L}(G_{aug})$ be a string that reaches (x, q) in G_{aug}. We define $s_R = Re(s)$ and $\alpha = P_o(s)$. Then, we know that $x \in \delta(x_0, s_R)$, $\delta_{obs}(q_{obs,0}, \alpha) = q$ and $\alpha \in O(s_R)$. We choose $\Delta = Pr(\alpha \mid s_R) p(s_R)$ and we claim that, for any integer $n \in \mathbb{N}$, we have that $Pr[s \in \mathcal{L}(\mathcal{G}) : ND \wedge |s| = n] \geq \Delta$. To see this, we consider the following two cases for n.

Case 1: $n \leq |s_R|$. Since (x, q) is uncertain, we know that $D(\alpha) = 0$, which implies that, for any prefix $\beta \in \overline{\alpha}$, we have $D(\beta) = 0$. Therefore, for any $n \leq |s_R|$, we can choose $t \in \overline{s_R}$ such that $|t| = n$. Then, we know that $Pr[s \in \mathcal{L}(\mathcal{G}) : ND \wedge |s| = n] = \sum_{s \in \mathcal{L}(\mathcal{G}) : |s| = n} \sum_{\alpha \in O(s)} (1 - D(\alpha)) Pr(\alpha \mid s) p(s) \geq \sum_{\beta \in \overline{\alpha}} Pr(\beta \mid t) p(t) \geq Pr(\alpha \mid s_R) p(s_R) = \Delta$

Case 2: $n > |w|$. Since $(x, q) \in V_{ter} \cap V_{unc}$, we know that $\forall t \in \mathcal{L}(G, x), \forall \beta \in O(t) : D(\alpha\beta) = 0$. Therefore, for any $n > |w|$, Then, we know that $Pr[s \in \mathcal{L}(\mathcal{G}) : ND \wedge |s| = n] = \sum_{s \in \mathcal{L}(\mathcal{G}) : |s| = n} \sum_{\alpha \in O(s)} (1 - D(\alpha)) Pr(\alpha \mid s) p(s) \geq \sum_{s_R t \in \mathcal{L}(\mathcal{G}) : |s_R t| = n} Pr(\alpha \mid s_R) Pr(s_R t) = Pr(\alpha \mid s_R) p(s_R) = \Delta$.

($\Leftarrow$) Suppose $s \in \mathcal{L}(G)$ is generated and $\alpha \in O(s)$ is observed. Let $x \in \delta(X_0, s)$ and $q = \delta_{obs}(q_{obs,0}, \alpha)$. We know that $(x, q) \in Q_{aug}$. We denote by $p(x, q)$ the probability that the initial-state can be detected in the future given that the current state is x and the current estimator state is q. Note that $p(x, q)$ is non-zero, since $V_{ter} \cap V_{unc} = \emptyset$ and (x, q) can always reach a certain state, i.e., $(\exists t \in \mathcal{L}(G, x))(\exists \beta \in O(t))[|S(\delta_{obs}(q, \beta))| = 1]$. We denote by $t_{\min}(x, q)$ the shortest t satisfying the above condition and define $n_{\min}(x, q) := |t_{\min}(x, q)|$. Then, we know that, given x and q, the probability that the initial-state is detected in the next $n_{\min}(x, q)$-steps is non-zero and we denote by $p_D(x, q) > 0$ this probability. Then, we define $n_{\max} := \max_{x \in X, q \in Q_{obs}} n_{\min}(x, q)$ and $p_{\min} := \min_{x \in X, q \in Q_{obs}} p_D(x, q)$. We know that, for any instant, the probability that the initial-state can be detected in the next $n_{\max}$-steps is greater than or equal to $p_{\min}$. Therefore, for any $k \in \mathbb{N}$, we have that $Pr[s \in \mathcal{L}(\mathcal{G}) : ND \wedge |s| = k n_{\max}] \leq (1 - p_{\min})^k$. Then, for any $\Delta > 0$, by taking $n \in \mathbb{N}$ such that $n \geq n_{\max} \lceil \log \frac{\Delta}{1 - p_{\min}} \rceil$, where $\lceil k \rceil$ denotes the smallest integer greater than or equal to k, we have that $Pr[s \in \mathcal{L}(\mathcal{G}) : ND \wedge |s| = n] \leq \Delta$, i.e., $\mathcal{G}$ is SI-detectable w.r.t. M. □

References

Athanasopoulou, E., Li, L., & Hadjicostis, C. N. (2010). Maximum likelihood failure diagnosis in finite state machines under unreliable observations. *IEEE Transactions on Automatic Control, 3*(55), 579–593.

Bertrand, N., Haddad, S., & Lefaucheux, E. (2014). Foundation of diagnosis and predictability in probabilistic systems. In *34th IARCS conf. foundations of soft. tech. theo. computer sci.* (pp. 417–429).

Cabasino, M. P., Hadjicostis, C. N., & Seatzu, C. (2015). Probabilistic marking estimation in labeled petri nets. *IEEE Transactions on Automatic Control, 60*(2), 528–533.

Carvalho, L. K., Basilio, J. C., & Moreira, M. V. (2012). Robust diagnosis of discrete event systems against intermittent loss of observations. *Automatica, 48*(9), 2068–2078.

Cassandras, C., & Lafortune, S. (2008). *Introduction to discrete event systems* (2nd ed.). Springer.

Chen, J., Ibrahim, M., & Kumar, R. (2016). Quantification of secrecy in partially observed stochastic discrete event systems. *IEEE Transactions on Automation Science and Engineering, 14*(1), 185–195.

Chen, J., Keroglou, C., Hadjicostis, C. N., & Kumar, R. (2017). Revised test for stochastic diagnosability of discrete-event systems. *IEEE Transactions on Automation Science and Engineering*,.

Chen, J., & Kumar, R. (2015a). Failure detection framework for stochastic discrete event systems with guaranteed error bounds. *IEEE Transactions on Automatic Control, 60*(6), 1542–1553.

Chen, J., & Kumar, R. (2015b). Stochastic failure prognosability of discrete event systems. *IEEE Transactions on Automatic Control, 60*(6), 1570–1581.

Keroglou, C., & Hadjicostis, C.N. (2013). Initial state opacity in stochastic DES. In *18th IEEE conf. emerging tech. factory automation* (pp. 1–8).

Keroglou, C., & Hadjicostis, C. N. (2015). Detectability in stochastic discrete event systems. *Systems & Control Letters, 84*, 21–26.

Li, L., & Hadjicostis, C. N. (2013). Minimum initial marking estimation in labeled petri nets. *IEEE Transactions on Automatic Control*, *58*(1), 198–203.
Lin, F. (2014). Control of networked discrete event systems: dealing with communication delays and losses. *SIAM Journal on Control and Optimization*, *52*(2), 1276–1298.
Lunze, J., & Schröder, J. (2001). State observation and diagnosis of discrete-event systems described by stochastic automata. *Discrete Event Dynamic Systems*, *11*(4), 319–369.
Özveren, C. M., & Willsky, A. S. (1990). Observability of discrete event dynamic systems. *IEEE Transactions on Automatic Control*, *35*(7), 797–806.
Saboori, A., & Hadjicostis, C. N. (2013). Verification of initial-state opacity in security applications of discrete event systems. *Information Sciences*, *246*, 115–132.
Saboori, A., & Hadjicostis, C. N. (2014). Current-state opacity formulations in probabilistic finite automata. *IEEE Transactions on Automatic Control*, *59*(1), 120–133.
Savitch, W. J. (1970). Relationships between nondeterministic and deterministic tape complexities. *Journal of Computer and System Sciences*, *4*(2), 177–192.
Sears, D., & Rudie, K. (2014). On computing indistinguishable states of nondeterministic finite automata with partially observable transitions. In *53rd IEEE conf. decision and control* (pp. 6731–6736).
Shu, S., Huang, Z., & Lin, F. (2013). Online sensor activation for detectability of discrete event systems. *IEEE Transactions on Automation Science and Engineering*, *10*(2), 457–461.
Shu, S., & Lin, F. (2013a). Enforcing detectability in controlled discrete event systems. *IEEE Transactions on Automatic Control*, *58*(8), 2125–2130.
Shu, S., & Lin, F. (2013b). I-detectability of discrete-event systems. *IEEE Transactions on Automation Science and Engineering*, *10*(1), 187–196.
Shu, S., Lin, F., & Ying, H. (2007). Detectability of discrete event systems. *IEEE Transactions on Automatic Control*, *52*(12), 2356–2359.
Shu, S., Lin, F., Ying, H., & Chen, X. (2008). State estimation and detectability of probabilistic discrete event systems. *Automatica*, *44*(12), 3054–3060.
Stockmeyer, L.J., & Meyer, A.R. (1973). Word problems requiring exponential time. In *Proc. 5th ACM symp. theory of computing* (pp. 1–9).
Takai, S., & Ushio, T. (2012). Verification of codiagnosability for discrete event systems modeled by mealy automata with nondeterministic output functions. *IEEE Transactions on Automatic Control*, *57*(3), 798–804.
Thorsley, D., & Teneketzis, D. (2005). Diagnosability of stochastic discrete-event systems. *IEEE Transactions on Automatic Control*, *50*(4), 476–492.
Thorsley, D., Yoo, T.-S., & Garcia, H.E. (2008). Diagnosability of stochastic discrete-event systems under unreliable observations. In *American control conference* (pp. 1158–1165).
Yin, X., & Lafortune, S. (2015a). Codiagnosability and coobservability under dynamic observations: Transformation and verification. *Automatica*, *61*, 241–252.
Yin, X., & Lafortune, S. (2015b). A general approach for solving dynamic sensor activation problems for a class of properties. In *54th IEEE conf. decision and control* (pp. 3610–3615).
Yin, X., & Lafortune, S. (2016). A uniform approach for synthesizing property-enforcing supervisors for partially-observed discrete-event systems. *IEEE Transactions on Automatic Control*, *61*(8), 2140–2154.

Xiang Yin was born in Anhui, China, in 1991. He received the B.Eng. degree from Zhejiang University in 2012, the M.S. degree from the University of Michigan, Ann Arbor, in 2013, and the Ph.D. degree from the University of Michigan, Ann Arbor, in 2017, all in electrical engineering. His research interests include supervisory control of discrete-event systems, model-based fault diagnosis, formal methods, security and their applications to cyber and cyber–physical systems. He received the Outstanding Reviewer Award from Automatica in 2016, the Outstanding Reviewer Award from IEEE Transactions on Automatic Control in 2017 and the IEEE Conference on Decision and Control (CDC) Best Student Paper Award Finalist in 2016. He is the co-chair of the IEEE CSS Technical Committee on Discrete Event Systems.