Discrete Event Dyn Syst (2018) 28:289–314
https://doi.org/10.1007/s10626-017-0263-8

Probabilistic system opacity in discrete event systems

Christoforos Keroglou[1,2] · **Christoforos N. Hadjicostis**[1]

Received: 15 September 2016 / Accepted: 16 October 2017 / Published online: 13 November 2017

Abstract In many emerging security applications, a system designer frequently needs to ensure that a certain property of a given system (that may reveal important details about the system's operation) be kept secret (opaque) to outside observers (eavesdroppers). Motivated by such applications, several researchers have formalized, analyzed, and described methods to verify notions of opacity in discrete event systems of interest. This paper introduces and analyzes a notion of opacity in systems that can be modeled as probabilistic finite automata or hidden Markov models. We consider a setting where a user needs to choose a specific hidden Markov model (HMM) out of m possible (different) HMMs, but would like to "hide" the true system from eavesdroppers, by not allowing them to have an arbitrary

This article belongs to the Topical Collection: *Special Issue on Diagnosis, Opacity and Supervisory Control of Discrete Event Systems*
Guest Editors: Christos G. Cassandras and Alessandro Giua

This work falls under the Cyprus Research Promotion Foundation (CRPF) Framework Programme for Research, Technological Development and Innovation 2009–2010 (CRPF's FP 2009–2010), co-funded by the Republic of Cyprus and the European Regional Development Fund, and specifically under Grant $T\Pi E/OPIZO/0609(BE)/08$. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of CRPF.

✉ Christoforos N. Hadjicostis
chadjic@ucy.ac.cy

Christoforos Keroglou
ckerog@umich.edu

[1] Department of Electrical and Computer Engineering, University of Cyprus, 1 Panepistimiou Avenue, 2109 Aglanzia, P.O.Box. 20537, 1678, Nicosia, Cyprus

[2] Present address: University of Michigan, EECS Building 1301 Beal Avenue, Ann Arbor, MI 48109-2122, USA

level of confidence as to which system has been chosen. We describe necessary and sufficient conditions (that can be checked with polynomial complexity), under which the intruder cannot distinguish the true HMM, namely, the intruder cannot achieve a level of certainty about its decision, which is above a certain threshold that we can *a priori* compute.

Keywords Privacy · Probabilistic finite automata · Opacity

1 Introduction and motivation

Motivated by the increased reliance of many applications on shared cyber-infrastructures (ranging from defense and banking to health care and power distribution systems), various notions of *security and privacy* have received considerable attention from researchers. A number of such notions focus on characterizing the *information flow* from the system to an eavesdropper (Focardi and Gorrieri 1994). *Opacity* falls in this category and aims at determining whether a given system's *secret* behavior (i.e., a subset of the behavior of the system that is considered critical and is usually represented by a predicate) is kept opaque to outsiders (Bryans et al. 2005; Saboori and Hadjicostis 2007). More specifically, this requires that the eavesdropper (modeled as a passive observer[1] of the system's behavior) never be able to establish the truth of the predicate.

Early works that studied notions of opacity in discrete event systems include (Bryans et al. 2005a, b; Badouel et al. 2006; Dubreil et al. 2008). The authors of Brayans et al. (2005a, b) focus on finite state Petri nets and define opacity with respect to state-based predicates. Multiple intruders that are modeled as observers with different observation capabilities are considered in Badouel et al. (2006), whereas the authors of Dubreil et al. (2008) consider a single intruder (that might observe different events than the ones observed/controlled by a supervisor that aims to control the system so as to avoid exposure of a property of interest) and establish that a minimally restrictive supervisor always exists, but might not be regular.

In Saboori and Hadjicostis (2007, 2013), the authors consider opacity with respect to state-based predicates in a discrete event system (DES) that can be modeled as a non-deterministic finite automaton with partial observation on its transitions. State-based notions of opacity exemplify the use of observers, and make more explicit the relationship between state-based notions of opacity and their verification with observers. Examples to motivate the study of current- and initial-state opacity in the context of sensor network coverage and encryption using pseudo-random number generators can be found in Saboori and Hadjicostis (2013, 2011). The authors in Wu and Lafortune (2013) showed that there exists a polynomial-time transformation between several notions of opacity, including language-based and state-based notions.

Motivated by the absence of likelihood information in most earlier work on opacity, the information-theoretic works in Millen (1987) and Wittbold and Johnson (1990), and more recently the works in Berard et al. (2010) and Brard et al. (2015), extend notions of opacity to probabilistic settings. In particular, state-based notions of opacity have been developed for probabilistic finite automata (PFA's) in Saboori and Hadjicostis (2014) by

[1] A passive observer is one that does not have any decision-making authority in the system (i.e., it cannot influence the operation of the system).

devising appropriate measures to quantify opacity. The following three notions were defined and analyzed in Saboori and Hadjicostis (2014):

(i) *Step-based almost current-state opacity* considers the *a priori* probability of violating current-state opacity, following any sequence of events of length k, and requires that this probability lies below a threshold for all possible sequences of length k (for all k).
(ii) *Almost current-state opacity* considers the *a priori* probability of violating current-state opacity following any sequence of events, and requires that this probability lies below a threshold.
(iii) *Probabilistic current-state opacity* requires that, for each possible sequence of observations, the following property holds: the increase in the *conditional* probability that the system current state lies in the set of secret states (conditioned on the given sequence of observations) compared to the prior probability (that the initial state lied in the set of secret states before any observation) is smaller than a given threshold.

The above ideas were extended in Keroglou and Hadjicostis (2013) to deal with corresponding notions of initial-state opacity in PFA's.

In this paper we study probabilistic system opacity. The setting we consider is as follows: a system is chosen at initialization among m known models, each of which is captured by a hidden Markov model (HMM). Our goal is to determine whether the true (chosen) system remains hidden from an intruder (eavesdropper). We assume that the eavesdropper observes (via a natural projection map) a subset of the events occurring in the system. We allow partial flow of information to the eavesdropper, as long as a strictly positive (nonzero) threshold of ambiguity holds, even in the worst case. In our setup, the worst case involves an eavesdropper who knows exactly the m HMMs and also knows the observation sequence that has been generated. The question is to determine whether the true (chosen) HMM remains hidden from the eavesdropper, for any observation sequence. Here, "remains hidden" should be interpreted in a probabilistic sense: the eavesdropper cannot have confidence above a certain threshold (bounded away from unity), even if she/he is willing to wait for an arbitrarily long sequence of observations.

The main contribution of this work is to provide a polynomial complexity verification algorithm for the setting of probabilistic system opacity (assuming that the HMMs can start from any initial state with nonzero probability). The probabilistic system opacity setting was introduced in our previous work (Keroglou and Hadjicostis 2016) and part of the material used to establish results in this work, was introduced in Keroglou and Hadjicostis (2014). In this paper we extend our previous works in two ways:

1) We validate the correctness of a polynomial complexity algorithm for probabilistic system opacity. Specifically, we provide complete proofs for Theorem 1 and Theorem 3, which were not provided in our previous papers (Keroglou and Hadjicostis 2016) and (Keroglou and Hadjicostis 2014).
2) We discuss (in Section 5) how probabilistic system opacity relates to (is actually a special case of) probabilistic current-state opacity in Saboori and Hadjicostis (2014). However, unlike probabilistic current-state opacity which is in general undecidable, we show that probabilistic system opacity can be verified with polynomial complexity.

The paper is organized as follows. Section 2 reviews the HMM model under consideration, as well as needed concepts and notation. Specifically, we are interested in classification among HMMs, i.e., the ability of the observer to distinguish between two (or more) HMMs. Sections 3 and 4 introduce the relevant notion of probabilistic system opacity and develop the verification algorithm. In some sense, classification can be seen as the opposite of

opacity. Indeed, we prove later in the paper, that, in our specific setup, classification and opacity are exactly opposite. Note here that classification characterizes the models, but, on the other hand, opacity is dependent on the specific sequence of observations; in other words, classification depends on unconditional probabilities whereas probabilistic system opacity depends on conditional probabilities (thus, *a priori*, their relationship is not clear). Section 5 makes connections with probabilistic current state opacity and Section 6 summarizes the contribution of this work and briefly discusses possible future extensions.

2 Notation and background

Definition 1 (HMM Model). An HMM is described by a five-tuple $S = (Q, E, \Delta, \Lambda, \pi_0)$, where $Q = \{q_1, q_2, ..., q_{|Q|}\}$ is the finite set of states; $E = \{e_1, e_2, ..., e_{|E|}\}$ is the finite set of outputs; $\Delta : Q \times Q \to [0\ 1]$ captures the state transition probabilities; $\Lambda : Q \times E \times Q \to [0\ 1]$ captures the output probabilities associated with transitions; and π_0 is the initial state probability distribution vector. Specifically, for $q, q' \in Q$ and $\sigma \in E$, the output probabilities associated with transitions are given by

$$\Lambda(q, \sigma, q') \equiv \Pr(q[t+1] = q', E[t+1] = \sigma \mid q[t] = q) ,$$

and the state transition probabilities are given by

$$\Delta(q, q') \equiv \Pr(q[t+1] = q' \mid q[t] = q) ,$$

where $q[t]$ ($E[t]$) is the state (output) of the HMM at time step (or epoch) t. The output function $\Lambda(q, \sigma, q')$ describes the conditional probability of observing the output σ associated with the transition to state q' from state q. The state transition function needs to satisfy

$$\Delta(q, q') = \sum_{\sigma \in E} \Lambda(q, \sigma, q'), \ \forall q, q' \in Q \tag{1}$$

and also

$$\sum_{i=1}^{|Q|} \Delta(q, q_i) = 1, \ \forall q \in Q.$$

Definition 2 (Markov chain). For any HMM model $S = (Q, E, \Delta, \Lambda, \pi_0)$, there exists an associated Markov chain $MC = (Q, \Delta, \pi_0)$, where $Q = \{q_1, q_2, ..., q_{|Q|}\}$ is the finite set of states; $\Delta : Q \times Q \to [0\ 1]$ captures the state transition probabilities; and π_0 is the initial state probability distribution vector. We also denote the Markov chain by $MC = (Q, A, \pi_0)$ where A is a $|Q| \times |Q|$ matrix such that $A(k, j) = \Delta(q_j, q_k)$.

Given an HMM model $S = (Q, E, \Delta, \Lambda, \pi_0)$, we also define for notational convenience the $|Q| \times |Q|$ matrix A_σ, associated with output $\sigma \in E$, as follows: the $(k, j)^{th}$ entry of A_σ captures the probability of a transition from state q_j to state q_k that produces output σ, i.e., $A_\sigma(k, j) = \Lambda(q_j, \sigma, q_k)$. Note that $A = \sum_{\sigma \in E} A_\sigma$ is a column stochastic matrix whose $(k, j)^{th}$ entry denotes the probability of taking a transition from state q_j to state q_k, without regard to the output produced, i.e., A is the transition matrix of the Markov chain $MC = (Q, A, \pi_0)$ that corresponds to the given HMM $S = (Q, E, \Delta, \Lambda, \pi_0)$. We denote an observation sequence of length n as $\omega = \omega[1]\omega[2]...\omega[n] \in E^*$, where $\omega[t] \in E$. We say that the observation sequence ω belongs to the language of HMM S ($L(S)$) iff there exists $q[0], q[1], q[2], ..., q[n]$ s.t. $\pi_0(q[0]) > 0$ and $\Delta(q[t], \omega[t+1], q[t+1]) > 0, \forall t = 0, 1, 2, ..., n-1$.

Remark 1 If $\pi[t]$ is the $|Q|$-dimensional vector whose jth entry denotes the probability of being in state q_j after t steps (or epochs), then we have $\pi[0] = \pi_0$ and $\pi[t+1] = A\pi[t] = A^{t+1}\pi_0$.

Next we recall the properties of irreducibility and aperiodicity for Markov chains.

Definition 3 (Irreducible or Strongly Connected Markov Chain) (Seneta 2006; Cassandras and Lafortune 2007). A Markov chain $MC = (Q, A, \pi_0)$ is irreducible if for all $q, q' \in Q$, there exists some $n \in \mathbb{N}$ such that $A^n(q', q) > 0$, where A^n is the n^{th} power of A. Equivalently, $\forall q' \in Q$, q' is reachable from any other state $q \in Q$. In such case, we observe that the graph[2] corresponding to MC is strongly connected.

Definition 4 (Periodic Markov chain) (Seneta 2006; Cassandras and Lafortune 2007). A state $q_i \in Q$ of a Markov chain $MC = (Q, A, \pi_0)$ is said to be periodic if the greatest common divisor d of the set $\{n > 0 : \Pr(q[n] = q_i \mid q[0] = q_i) > 0\}$ is $d \geq 2$. If $d = 1$, state q_i is said to be aperiodic. The Markov chain is said to be aperiodic if all states $q_i \in Q$ are aperiodic.

Remark 2 (Cassandras and Lafortune 2007) If a Markov chain $MC = (Q, A, \pi_0)$ is irreducible, then all its states have the same period. It follows that if $d = 1$ for any state of an irreducible Markov chain, then all states are aperiodic. On the other hand, if any state has period $d \geq 2$, then all states have the same period and the chain is said to be periodic with period $d \geq 2$.

Lemma 1 (Cassandras and Lafortune 2007) *(Stationary distribution of a Markov chain). If the Markov chain is irreducible and aperiodic, then* $\lim_{t\to\infty} \pi[t]$ *exists and is called the stationary distribution of the Markov chain denoted by* $\pi_s = [\pi_s(q_1), \pi_s(q_2), ..., \pi_s(q_{|Q|})]^T$, *where* ^{T} *denotes matrix/vector transposition.*

Definition 5 (Stationary Emission Probabilities of HMM). Given an HMM $S = (Q, E, \Delta, \Lambda, \pi_0)$, the stationary emission probability $\pi_e(e_i)$, $\forall e_i \in E$, can be expressed as

$$\pi_e(e_i) = \mathbf{1}^T(A_{e_i}\pi_s),$$

where $\mathbf{1}^T$ is the $|Q|$-dimensional row vector with all entries equal to unity. We denote the vector of stationary emission probabilities as $\pi_e = [\pi_e(e_1)...\pi_e(e_{|E|})]^T$.

Note that the stationary state probability vector π_s for an HMM S is the same as the stationary state probability vector of its associated Markov chain $MC = (Q, A, \pi_0)$.

2.1 Optimal decision rule (MAP rule)

Suppose that we are given two HMMs, captured by $S^{(1)} = (Q^{(1)}, E^{(1)}, \Delta^{(1)}, \Lambda^{(1)}, \pi_0^{(1)})$ and $S^{(2)} = (Q^{(2)}, E^{(2)}, \Delta^{(2)}, \Lambda^{(2)}, \pi_0^{(2)})$, with prior probabilities for each model given by P_1 and $P_2 = 1 - P_1$, respectively. Given $E^{(j)} = \left\{e_1^{(j)}, e_2^{(j)}, ..., e_{|E^{(j)}|}^{(j)}\right\}$, $j = \{1, 2\}$,

[2]The graph corresponding to the Markov chain (Q, A, π_0) is the graph $G = (Q, E)$ with vertices Q and edges $E \subseteq Q \times Q$ such that $(q_k, q_j) \in E$ iff $A(k, j) > 0$.

for the two HMMs, we define for notational convenience $E = E^{(1)} \cup E^{(2)}$ with $E = \{e_1, e_2, ..., e_{|E|}\}$. $A_{e_i}^{(j)}$ is the transition matrix for $S^{(j)}$, $j = \{1, 2\}$, under the output symbol $e_i \in E$. The matrix $A_{e_i}^{(j)}$, associated with output $e_i \in E^{(j)}$, is defined as follows: the $(k, l)^{th}$ entry of $A_{e_i}^{(j)}$ captures the probability of a transition from state q_l to state q_k that produces output e_i, i.e., $A_{e_i}^{(j)}(k, l) = \Lambda^{(j)}(q_l, e_i, q_k)$. We set $A_{e_i}^{(j)}$ to zero if $e_i \in E \setminus E^{(j)}$. If we observe a sequence of n outputs $\omega = \omega[1]\omega[2]...\omega[n]$, with $\omega[t] \in E$, that is generated by one of the two underlying HMMs, the classifier that minimizes the probability of error needs to implement the maximum *a posteriori* probability (MAP) rule. Specifically, the MAP classifier compares (as done in a classical hypothesis testing problem (Neyman and Pearson 1992))

$$\Pr(S^{(1)} \mid \omega) \underset{<}{\overset{>}{}} \Pr(S^{(2)} \mid \omega) \Rightarrow \frac{\Pr(\omega \mid S^{(1)})}{\Pr(\omega \mid S^{(2)})} \underset{<}{\overset{>}{}} \frac{P_2}{P_1},$$

and decides in favor of $S^{(1)}$ ($S^{(2)}$) if the left (right) quantity is larger. When we decide in favor of one or the other model, we incur a probability of error proportional to the probability of the model that was not selected; with some algebra, it can be shown that $\Pr(\text{error}, \omega) = \min\{P_1 \cdot \Pr(\omega \mid S^{(1)}), P_2 \cdot \Pr(\omega \mid S^{(2)})\}$.

2.2 Probability of misclassification between HMMs

To calculate the *a priori* probability of error before any sequence of observations of length n is observed, we need to consider all possible observation sequences of length n:

$$\Pr(\text{error at } n) = \sum_{\omega \in E^n} \Pr(\text{error}, \omega), \tag{2}$$

where E^n is the set of all sequences of length n with outputs from E. We arbitrarily index each of the d^n ($d = |E|$) sequences of observations via $\omega(i)$, $i \in \{1, 2, ..., d^n\}$, and use $P_i^{(j)}$ to denote $P_i^{(j)} = \Pr(\omega(i)|S^{(j)})$, $j \in \{1, 2\}$. Note that some of these sequences may have zero probability under one of the two models (or even both models). The probability of misclassification between the two systems, after n steps, can then be expressed as

$$\begin{aligned} \Pr(\text{error at } n) &= \sum_{i=1}^{d^n} \Pr(\text{error}, \omega(i)) \\ &= \sum_{i=1}^{d^n} \min\left\{P_1 \cdot P_i^{(1)}, P_2 \cdot P_i^{(2)}\right\}. \end{aligned} \tag{3}$$

We can calculate $P_i^{(j)} = \Pr(\omega(i)|S^{(j)})$ with an iterative algorithm, a detailed description of which can be found in Athanasopoulou and Hadjicostis (2008) and Fu (1982). Specifically, given sequence $\omega = \omega[1]\omega[2]...\omega[n]$ we calculate

$$\rho_n^{(j)} = A_{\omega[n]}^{(j)} A_{\omega[n-1]}^{(j)} ... A_{\omega[1]}^{(j)} \pi_0^{(j)},$$

which is essentially a vector whose k^{th} entry captures the probability of reaching state $q_k \in Q^{(j)}$ while generating the sequence of outputs ω (i.e., $\rho_n^{(j)}(k) = \Pr(q[n] = q_k, \omega|S^{(j)})$). If we sum up the entries of $\rho_n^{(j)}$ we obtain $P_\omega^{(j)} = \Pr(\omega \mid S^{(j)}) = \sum_{k=1}^{|Q^{(j)}|} \rho_n^{(j)}(k)$.

Utilizing the above algorithm, we can certainly compute the probability of error at n by explicitly calculating $P_j \cdot P_i^{(j)}$ for each sequence $\omega(i)$. However, the calculation of the *a priori* probability of error is computationally difficult for large values of n due to the

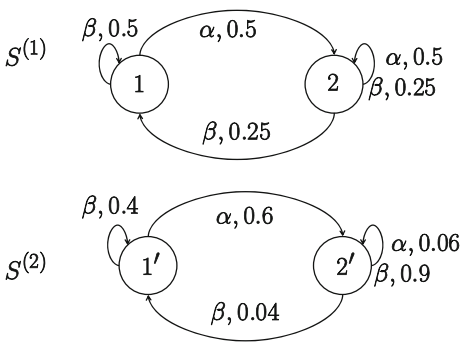


Fig. 1 $S^{(1)}$ (top) and $S^{(2)}$ (bottom) used in Example 1

exponential number of the possible sequences of observations ω; thus, in this paper, we are interested in obtaining easily calculable bounds for the *a priori* probability of error or misclassification. It is well-known that the MAP classifier described here minimizes the probability of error (misclassification); thus, any other rule, will be suboptimal in terms of minimizing the probability of error and can be used to obtain a bound on the probability of error. The following example is borrowed from Keroglou and Hadjicostis (2014).

Example 1 Suppose we are given the HMMs, $S^{(1)} = (Q^{(1)}, E^{(1)}, \Delta^{(1)}, \Lambda^{(1)}, \pi_0^{(1)})$ and $S^{(2)} = (Q^{(2)}, E^{(2)}, \Delta^{(2)}, \Lambda^{(2)}, \pi_0^{(2)})$ shown in Fig. 1, with $E^{(1)} = E^{(2)} = E = \{\alpha, \beta\}$, $\pi_0^{(1)} = \pi_0^{(2)} = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$, and $P_1 = P_2 = 0.5$. The corresponding $A_\alpha^{(1)}$, $A_\beta^{(1)}$, $A_\alpha^{(2)}$, $A_\beta^{(2)}$ are as follows:

$$A_\alpha^{(1)} = \begin{bmatrix} 0 & 0 \\ 0.5 & 0.5 \end{bmatrix}, A_\beta^{(1)} = \begin{bmatrix} 0.5 & 0.25 \\ 0 & 0.25 \end{bmatrix},$$

$$A_\alpha^{(2)} = \begin{bmatrix} 0 & 0 \\ 0.6 & 0.06 \end{bmatrix}, A_\beta^{(2)} = \begin{bmatrix} 0.4 & 0.04 \\ 0 & 0.9 \end{bmatrix}.$$

If the sequence $\omega(\ell) = \beta\alpha\beta\alpha$ is observed, we have $P_\ell^{(1)} = \sum_{k=1}^{|Q^{(1)}|} \rho_4^{(1)}(k) = 0.0625$, where $\rho_4^{(1)} = A_\alpha^{(1)} A_\beta^{(1)} A_\alpha^{(1)} A_\beta^{(1)} \pi_0^{(1)}$, and $P_\ell^{(2)} = \sum_{k=1}^{|Q^{(2)}|} \rho_4^{(2)}(k) = 0.0187$, where $\rho_4^{(2)} = A_\alpha^{(2)} A_\beta^{(2)} A_\alpha^{(2)} A_\beta^{(2)} \pi_0^{(2)}$. Thus, the probability of error between the two models when this specific sequence is observed is $\Pr(\text{error}, \omega(\ell)) = 0.0094$ (i.e., $S^{(1)}$ will be selected and $\Pr(\text{error}, \omega(\ell)) = P_2 \cdot P_\ell^{(2)}$).

3 Probabilistic system opacity

Probabilistic system opacity considers the following setting: we are given m HMMs, denoted by $S^{(i)}$ for $i \in \{1, 2, ..., m\}$. The prior probability of $S^{(i)}$ is P_i, $P_i > 0$, with the

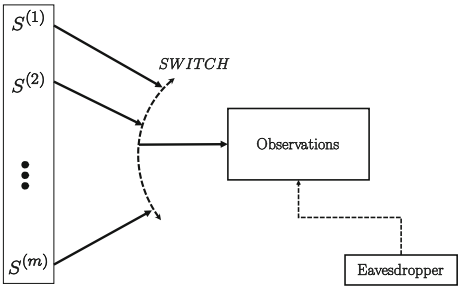


Fig. 2 An HMM is chosen out of m different HMMs, $S^{(1)}$, $S^{(2)}$, ..., $S^{(m)}$; an Eavesdropper knows the exact structure of these HMMs and also observes the observation sequence that is generated; probabilistic system opacity holds if the Eavesdropper is kept confused about which is the true HMM that generates the observation

prior probabilities satisfying $\sum_{i=1}^{m} P_i = 1$. A user is supposed to choose one of these models, say $S^{(i)}$, and would like to keep an observer (eavesdropper) confused about the chosen HMM, for any behavior that might occur in the chosen HMM, regardless of the sequence of observations generated by it and regardless of how long the observer is willing to wait (refer to Fig. 2). This means that for any (arbitrarily long) observation sequence that can be generated by the chosen HMM, the observer must *not* be able to determine the chosen HMM, at least not with absolute certainty or with certainty that tends asymptotically to unity.

Definition 6 (Probabilistic System Opacity). Consider a set of m HMMs, $S^{(i)} = \left(Q^{(i)}, E^{(i)}, \Delta^{(i)}, \Lambda^{(i)}, \pi_0^{(i)}\right)$, for $i \in \{1, ..., m\}$, with corresponding Markov chains $MC^{(i)} = \left(Q^{(i)}, A^{(i)}, \pi_0^{(i)}\right)$ that are irreducible and aperiodic and with initial probability distribution $\pi_0^{(i)} > 0$. Probabilistic system opacity holds if there exists an $\alpha_0 > 0$, such that for any chosen $S^{(i)}$ and for any observation sequence ω that could be generated by $S^{(i)}$, we have

$$\alpha(\omega) := \frac{\sum_{k=1, k\neq i}^{m} P_k P_\omega^{(k)}}{\sum_{k=1}^{m} P_k P_\omega^{(k)}} \geq \alpha_0 \ .$$

Remark 3 Note that in Definition 6, we assume that the initial probability distribution $\pi_0^{(i)}$ is a strictly positive vector (i.e., all initial states are possible among all m HMMs). If this is the case, we will argue that probabilistic system opacity can be verified with polynomial complexity. The complexity and the verification algorithm in the more general case, where $\pi_0^{(i)}$ is not necessarily strictly positive, remains an open problem.

4 Polynomial verification of probabilistic opacity

In the following definition, we discuss the problem of probabilistic opacity for two HMMs. It will become obvious from the discussions in this section that the conditions for m HMMs, to be probabilistically opaque are based on the conditions for a pair of HMMs to be probabilistically opaque, which is defined next.

Definition 7 (Pairwise Probabilistic Opacity). Two HMMs, $S^{(i)} = (Q^{(i)}, E, \Delta^{(i)}, \Lambda^{(i)}, \pi_0^{(i)})$ for $i \in \{1, 2\}$ and prior probabilities[3] P_1 and P_2 are probabilistically opaque if there exists α_0 $(0 < \alpha_0 < 1/2)$ such that for any observation sequence ω

$$(\forall \omega \in L(S^{(1)}) \cup L(S^{(2)})) \text{ we have } \alpha(\omega) \geq \alpha_0 \ ,$$

with

$$\alpha(\omega) = \min \left\{ \frac{P_1 P_\omega^{(1)}}{P_1 P_\omega^{(1)} + P_2 P_\omega^{(2)}}, \frac{P_2 P_\omega^{(2)}}{P_1 P_\omega^{(1)} + P_2 P_\omega^{(2)}} \right\} .$$

[Recall that $P_i P_\omega^{(i)}$, $i \in \{1, 2\}$, is the probability that observation ω is generated by HMM $S^{(i)}$.]

To determine whether two HMMs are probabilistically opaque, we will employ tools from probabilistic equivalence and HMM classification; we introduce some relevant definitions next.

Definition 8 (Probabilistic Equivalence for HMMs (Tzeng 1989)). Two HMMs, $S^{(i)} = (Q^{(i)}, E^{(i)}, \Delta^{(i)}, \Lambda^{(i)}, \pi_0^{(i)})$, $i \in \{1, 2\}$ with $E = E^{(1)} = E^{(2)}$ are probabilistically equivalent iff for any string $\omega \in L(S)$ $(L(S) = L(S^{(1)}) \cup L(S^{(2)}))$ the two HMMs, accept the string with equal probability.

Remark 4 Two HMMs can be tested for probabilistic equivalence with polynomial complexity (Tzeng 1989).

Remark 5 We say that two HMMs are *probabilistically equivalent from steady–state* iff the two HMMs $S^{(i)} = (Q^{(i)}, E^{(i)}, \Delta^{(i)}, \Lambda^{(i)}, \pi_s^{(i)})$, for $i \in \{1, 2\}$, where $\pi_s^{(i)}$ is their respective steady–state probabilities (Definition 1), are probabilistically equivalent.

Theorem 1 (Probability of Error Among Two HMMs Tending to Zero) *Consider two HMMs,* $S^{(i)} = (Q^{(i)}, E^{(i)}, \Delta^{(i)}, \Lambda^{(i)}, \pi_0^{(i)})$, $i \in \{1, 2\}$, *with corresponding Markov chains* $MC^{(i)} = (Q^{(i)}, A^{(i)}, \pi_0^{(i)})$ *that are irreducible and aperiodic. If* $S^{(1)}$ *and* $S^{(2)}$ *are not probabilistically equivalent from steady–state, then*

$$(\forall \epsilon > 0)(\exists n_0 \in \mathbb{N}) \text{ such that for } n \geq n_0 \ \Pr(\text{error at } n) < \epsilon,$$

where $\Pr(\text{error at } n)$ *is the probability of misclassification for the two HMMs (defined in* Eq. 2*).*

Proof The proof is provided in Appendix. □

In other words, if the two HMMs are not probabilistically equivalent from steady–state, then the probability of error among the two HMMs tends, when n goes to infinity, to zero. The proof in Appendix relies on the fact that we are able to discriminate between the two HMM models using a suboptimal decision rule (Definition 12) based on the empirical frequencies of output symbols, as long as the two systems are characterized, at steady-state,

[3]Usually $P_1 + P_2 = 1$, but in our case we keep the priors as if the two HMMs, were part of a setting with m HMMs, as it is described in Definition 6. This helps us to avoid notational overhead involving renormalizations of priors (namely, $P_i' = P_i/(P_1 + P_2)$ for $i = 1, 2$).

by different statistical properties for the occurrence of output symbols or different statistical properties of finite sequences of consecutive output symbols (this occurs if and only if the two HMMs are not probabilistically equivalent from steady–state). The theoretical analysis in Appendix establishes an upper bound on the misclassification probability, which is described by a function that decreases exponentially with the length of the observation sequence (as long as the two systems are characterized, at steady-state, by different statistical properties for the stationary emission probabilities (Definition 5) or stationary emission probabilities for a finite number of consecutive output symbols).

Now, we revisit the following theorem, which was introduced and proved in Keroglou and Hadjicostis (2016). This theorem establishes the necessary and sufficient conditions needed for two HMMs to be probabilistically opaque.

Theorem 2 (Conditions for Pairwise Probabilistic Opacity (Definition 7)) *Consider two HMMs* $S^{(i)} = \left(Q^{(i)}, E^{(i)}, \Delta^{(i)}, \Lambda^{(i)}, \pi_0^{(i)}\right)$, $i = 1, 2$, *with corresponding Markov chains* $MC^{(i)} = \left(Q^{(i)}, A^{(i)}, \pi_0^{(i)}\right)$ *that are irreducible and aperiodic. These two HMMs are pairwise probabilistically opaque iff they are probabilistically equivalent from steady-state.*

Proof Let us use the following notation:

- $\omega = \omega[1]\omega[2]...\omega[n]$, where $\omega[t] \in E$ for $t \in \{1, ..., n\}$;
- $\mathbf{1}^T = [1...1]$ is a row vector with n identical elements equal to 1;
- $A_\omega^{(1)} = A_{\omega[n]}^{(1)} \cdots A_{\omega[1]}^{(1)}$ and $A_\omega^{(2)} = A_{\omega[n]}^{(2)} \cdots A_{\omega[1]}^{(2)}$;
- For a vector π, $\min\{\pi\}$ is the minimum element of the vector and $\max\{\pi\}$ is the maximum element of the vector;
- $\alpha(\omega) = \min\left\{\frac{P_1 P_\omega^{(1)}}{P_1 P_\omega^{(1)} + P_2 P_\omega^{(2)}}, \frac{P_2 P_\omega^{(2)}}{P_1 P_\omega^{(1)} + P_2 P_\omega^{(2)}}\right\}$, where $P_i P_\omega^{(i)}$, $i = 1, 2$, is the probability that observation ω is generated by HMM $S^{(i)}$.

($\rightarrow$) Suppose that the two HMMs are probabilistically opaque; we need to show that the two HMMs are probabilistically equivalent from steady-state. We know that if the probability of error does not tend to zero, then the two HMMs are probabilistically equivalent from steady–state according to the contraposition of Theorem 1. It remains to prove that if the two HMMs are probabilistically opaque, then the probability of error among the two HMMs does not tend to zero. If the two HMMs are probabilistically opaque, we argue that the probability of error when trying to classify between $S^{(1)}$ and $S^{(2)}$ based on a sequence of observations satisfies

$$(\exists\, 0 < \alpha_0' < 1\,)(\forall n \in \mathbb{N})\{\Pr(\text{error at } n) \geq \alpha_0'\}\,.$$

This is proved easily because we know that $(\exists \alpha_0)(\forall \omega \in L(S^{(1)}) \cup L(S^{(2)}))$, we have $\min\left\{\frac{P_1 P_\omega^{(1)}}{P_1 P_\omega^{(1)} + P_2 P_\omega^{(2)}}, \frac{P_2 P_\omega^{(2)}}{P_1 P_\omega^{(1)} + P_2 P_\omega^{(2)}}\right\} \geq \alpha_0$. Therefore, for each $n \in \mathbb{N}$

$$\begin{aligned}
\Pr(\text{error at } n) &= \sum_{\omega:|\omega|=n} \left(\min\left\{P_1 P_\omega^{(1)}, P_2 P_\omega^{(2)}\right\}\right) \\
&\geq \sum_{\omega:|\omega|=n} \alpha_0 \left(P_1 P_\omega^{(1)} + P_2 P_\omega^{(2)}\right) \\
&= \alpha_0 \left(P_1 \sum_{\omega:|\omega|=n} P_\omega^{(1)} + P_2 \sum_{\omega:|\omega|=n} P_\omega^{(2)}\right) = \alpha_0(P_1 + P_2) = a_0'\,.
\end{aligned}$$

This proves that the probability of error does not tend to zero; therefore, the two HMMs are probabilistically equivalent from steady–state.

($\leftarrow$) Suppose that the two HMMs are probabilistically equivalent from steady–state; then, for any ω, we have

$$\mathbf{1}^T A_\omega^{(1)} \pi_s^{(1)} = \mathbf{1}^T A_\omega^{(2)} \pi_s^{(2)} =: \pi_{\omega,s} \ .$$

We next prove that the two HMMs are Probabilistically Opaque. Four useful inequalities for $i \in \{1, 2\}$ are the following:

$$\begin{aligned} P_\omega^{(i)} &= \mathbf{1}^T A_\omega^{(i)} \pi_0^{(i)} \\ &\geq \mathbf{1}^T A_\omega^{(i)} \min\{\pi_0^{(i)}\}\mathbf{1} \\ &\geq \min\{\pi_0^{(i)}\}\pi_{\omega,s} \ , \end{aligned}$$

$$\begin{aligned} P_\omega^{(i)} &= \mathbf{1}^T A_\omega^{(i)} \pi_0^{(i)} \\ &\leq \mathbf{1}^T A_\omega^{(i)} \max\left\{\pi_0^{(i)}\right\}\mathbf{1} \\ &\leq \frac{\max\left\{\pi_0^{(i)}\right\}}{\min\left\{\pi_s^{(i)}\right\}} \mathbf{1}^T A_\omega^{(i)} \pi_s^{(i)} \\ &\leq \frac{\max\left\{\pi_0^{(i)}\right\}}{\min\left\{\pi_s^{(i)}\right\}} \pi_{\omega,s} \ . \end{aligned}$$

In summary, we have

$$\min\left\{\pi_0^{(i)}\right\}\pi_{\omega,s} \leq P_\omega^{(i)} \leq \frac{\max\left\{\pi_0^{(i)}\right\}}{\min\left\{\pi_s^{(i)}\right\}}\pi_{\omega,s} \ .$$

From the previous inequalities we can rewrite $\alpha(\omega) = \min\left\{\frac{P_1 P_\omega^{(1)}}{P_1 P_\omega^{(1)} + P_2 P_\omega^{(2)}}, \frac{P_2 P_\omega^{(2)}}{P_1 P_\omega^{(1)} + P_2 P_\omega^{(2)}}\right\} \geq \min\{c_1, c_2\}$, where $c_1 < 1$ and $c_2 < 1$, with

$$c_i = \frac{P_i \min\{\pi_0^{(i)}\}}{P_1 \frac{\max\left\{\pi_0^{(1)}\right\}}{\min\{\pi_s^{(1)}\}} + P_2 \frac{\max\left\{\pi_0^{(2)}\right\}}{\min\{\pi_s^{(2)}\}}},$$

which proves that for any ω, of any length n, the observer is uncertain with a threshold of at least $\alpha_0 = \min\{c_1, c_2\}$. □

Theorem 3 discussed and proven in the remainder of this section, was presented without proof in Keroglou and Hadjicostis (2016). The following lemmas are useful in proving Theorem 3.

Lemma 1 *If probabilistic system opacity holds then the probability of error among m HMMs with $S^{(i)}$ as the chosen system, does not tend to zero* ($(\exists 0 < \epsilon < 1)(\forall n_0)(\exists n \geq n_0)$ ($\Pr(error\ at\ n,\ S^{(i)}) \geq \epsilon$).

Proof We have the following statements:

1. From Probabilistic System Opacity we have for any $\omega \in L(S)$:
$$\frac{\sum_{k=1,k\neq i}^{m} P_k P_\omega^{(k)}}{\sum_{k=1}^{m} P_k P_\omega^{(k)}} \geq \alpha_0$$
2. The set of decisions is $\{H_0, H_1\}$, where H_0 : { We accept $S^{(i)}$} and H_1 : { We reject $S^{(i)}$}
3. Using the MAP rule, we decide in favor of H_0 when $P_i P_\omega^{(i)} > \sum_{k\neq i} P_k P_\omega^{(k)}$ and in favor of H_1 when $P_i P_\omega^{(i)} \leq \sum_{k\neq i} P_k P_\omega^{(k)}$
4. $\Omega_n = \{\omega : |\omega| = n\}$
5. $\Omega_a^n = \left\{ \omega : (|\omega| = n) \wedge \left(P_i P_\omega^{(i)} > \sum_{k\neq i} P_k P_\omega^{(k)} \right) \right\}$
6. $\Omega_r^n = \left\{ \omega : (|\omega| = n) \wedge \left(P_i P_\omega^{(i)} \leq \sum_{k\neq i} P_k P_\omega^{(k)} \right) \right\}$
7. $(\forall \omega \in \Omega_n)\left(\sum_{k\neq i} P_k P_\omega^{(k)} \geq \alpha_0 \sum P_k P_\omega^{(k)} \geq \alpha_0 P_i P_\omega^{(i)} \right.$ (Definition 6)
8. $\Pr(\text{error at } n,\ S^{(i)}) = \sum_{\omega\in\Omega_a^n} \sum_{k\neq i} P_k P_\omega^{(k)} + \sum_{\omega\in\Omega_r^n} P_i P_\omega^{(i)}$

From 1–8 we have

$$\begin{aligned}\Pr(\text{error at } n,\ S^{(i)}) &\geq \alpha_0 \left(\sum_{\omega\in\Omega_a^n} P_i P_\omega^{(i)} + \sum_{\omega\in\Omega_r^n} P_i P_\omega^{(i)} \right) \\ &\geq \alpha_0 \sum_{\omega\in\Omega_n} P_i P_\omega^{(i)} = \alpha_0 P_i = \epsilon\end{aligned}$$

The proof is concluded. □

Lemma 2 *For $a, b_1, b_2, ..., b_m$ nonnegative real numbers, we have that*

$$\min\{a, b_1 + b_2 + ...b_m\} \leq \sum_{i=1}^{m} \min\{a, b_i\}$$

Proof We use mathematical induction:

1. For $m = 2$, we have to prove that $\min\{a, b_1 + b_2\} \leq \min\{a, b_1\} + \min\{a, b_2\}$. We take three cases i) $a > b_1 \geq 0$ and $a > b_2 \geq 0$, ii) $0 \leq a \leq b_1$ and $0 \leq a \leq b_2$, iii) $0 \leq a \leq b_1$ and $a > b_2 \geq 0$.

   We prove only the first case, the proofs for the other cases are left to the interested reader. For the first case, $\min\{a, b_1\} + \min\{a, b_2\} = b_1 + b_2$ which implies that $\min\{a, b_1 + b_2\} \leq (b_1 + b_2) = \min\{a, b_1\} + \min\{a, b_2\}$.

2. Let us suppose that for $k < m$, $\min\{a, b_1 + b_2 + ...b_k\} \leq \sum_{i=1}^{k} \min\{a, b_i\}$.
3. We want to prove that $\min\{a, b_1 + b_2 + ... + b_k + b_{k+1}\} \leq \sum_{i=1}^{k+1} \min\{a, b_i\}$. Indeed, with $B_k = b_1 + ...b_k$ we have from 1) and 2) that $\min\{a, B_k + b_{k+1}\} \leq \min\{a, B_k\} + \min\{a, b_{k+1}\} \leq \sum_{i=1}^{k} \min\{a, b_i\} + \min\{a, b_{k+1}\} = \sum_{i=1}^{k+1} \min\{a, b_i\}$. The proof is concluded. □

Lemma 3 *If the probability of error among m HMMs with $S^{(i)}$ as the chosen system, does not tend to zero, then there exists at least one $S^{(j)}$, such that $S^{(i)}$ and $S^{(j)}$ are pairwise probabilistic opaque.*

Proof We have the following statements:

1. The pairwise probability of error for $S^{(i)}$ and $S^{(j)}$ is defined below:
$$\Pr(\text{pairwise error at } n,\ S^{(i)}, S^{(j)}) = \sum_{\omega:|\omega|=n} \min\{P_i P_\omega^{(i)}, P_j P_\omega^{(j)}\}$$
2. The probability of error among m HMMs with $S^{(i)}$ as the chosen system can be formulated as given below:
$$\Pr(\text{error at } n,\ S^{(i)}) = \sum_{\omega:|\omega|=n} \min\left\{P_i P_\omega^{(i)}, \sum_{k\neq i} P_k P_\omega^{(k)}\right\}$$

We prove this lemma by contraposition. If there does not exist $S^{(j)}$ such that $S^{(i)}$ and $S^{(j)}$ are pairwise probabilistic opaque, then according to Definition 7 (and from Theorems 1 and 2) for any $S^{(j)}$ the pairwise probability of error for $S^{(i)}$ and $S^{(j)}$ tends to zero. This can be formulated as:

$$(\forall 0 < \epsilon < 1)(\exists n_{ij})(\forall n \geq n_{ij})$$
$$\Pr(\text{pairwise error at } n,\ S^{(i)}, S^{(j)}) < \epsilon.$$

$$\begin{aligned}\Pr(\text{error at } n,\ S^{(i)}) &\overset{2)}{=} \sum_{\omega:|\omega|=n} \min\left\{P_i P_\omega^{(i)}, \sum_{k\neq i} P_k P_\omega^{(k)}\right\} \\ &\overset{\text{Lemma 2}}{\leq} \sum_{\omega:|\omega|=n} \sum_{j\neq i} \min\left\{P_i P_\omega^{(i)}, P_j P_\omega^{(j)}\right\} \\ &\overset{1)}{=} \sum_{j\neq i} \Pr(\text{pairwise error at } n,\ S^{(i)}, S^{(j)}).\end{aligned}$$

From 1), 2) and from Lemma 2 and if we pick $n_0 = \max\{n_{ij}\}$ we have that

$$(\forall 0 < \epsilon' < 1)(\exists n_0)(\forall n \geq n_0)$$
$$\Pr(\text{error at } n,\ S^{(i)}) \leq \Pr(\text{pairwise error at } n,\ S^{(i)}, S^{(j)}) \leq (k-1)\epsilon = \epsilon'.$$

The proof is concluded. □

Theorem 3 (Necessary and Sufficient Conditions for Probabilistic System Opacity) *Consider a set of m HMMs, $S^{(i)} = (Q^{(i)}, E^{(i)}, \Delta^{(i)}, \Lambda^{(i)}, \pi_0^{(i)})$, $i \in \{1, ..., m\}$, with corresponding Markov chains $MC^{(i)} = (Q^{(i)}, A^{(i)}, \pi_0^{(i)})$ that are irreducible and aperiodic and with initial probability distribution $\pi_0^{(i)} > 0$. The following statements are true:*

($\rightarrow$) If the property of probabilistic system opacity as described in Definition 6 holds, then, for any i there exists at least one HMM $S^{(j)}$, $j \neq i$, such that $S^{(i)}$ and $S^{(j)}$ are pairwise probabilistically opaque (Definition 7).

($\leftarrow$) If, for any i, there exists at least one HMM $S^{(j)}$, $j \neq i$, such that $S^{(i)}$ and $S^{(j)}$ are pairwise probabilistically opaque, then the property of probabilistic system opacity holds.

Proof ($\rightarrow$) We prove the statement by combining Lemma 1 and Lemma 3.

($\leftarrow$) We need to show that for any system $S^{(i)}$ and for any observation sequence ω that can be generated by $S^{(i)}$, we have

$$\alpha(\omega) = \frac{\sum_{k=1,k\neq i}^{m} P_k P_\omega^{(k)}}{\sum_{k=1}^{m} P_k P_\omega^{(k)}} \geq \alpha_0 \ .$$

Suppose $S^{(j)}$ is the HMM that is pairwise probabilistically opaque with $S^{(i)}$. Then, from Definition 7, there exists an α_{ij}, for any observation sequence ω, such that $\min\left\{\frac{P_i P_\omega^{(i)}}{P_i P_\omega^{(i)} + P_j P_\omega^{(j)}}, \frac{P_j P_\omega^{(j)}}{P_i P_\omega^{(i)} + P_j P_\omega^{(j)}}\right\} \geq \alpha_{ij}$. Thus, for any observation sequence ω that could be generated by $S^{(i)}$, we have

$$\begin{aligned}\alpha(\omega) &= \frac{\sum_{k=1,k\neq i}^{m} P_k P_\omega^{(k)}}{\sum_{k=1}^{m} P_k P_\omega^{(k)}} = 1 - \frac{P_i P_\omega^{(i)}}{\sum_{k=1}^{m} P_k P_\omega^{(k)}} \\ &\geq 1 - \frac{P_i P_\omega^{(i)}}{P_i P_\omega^{(i)} + P_j P_\omega^{(j)}} = \frac{P_j P_\omega^{(j)}}{P_i P_\omega^{(i)} + P_j P_\omega^{(j)}} \geq \alpha_{ij} = \alpha_0 \ .\end{aligned}$$

Therefore, probabilistic system opacity holds if, for any chosen $S^{(i)}$, there exists another system $S^{(j)}$, such that $S^{(i)}$ and $S^{(j)}$ are pairwise probabilistically opaque (Definition 7). □

Remark 6 Note that in the definition of probabilistic system opacity (Definition 7) nothing is stated about the need to have, for each system $S^{(i)}$ another system $S^{(j)}$, such that $S^{(i)}$ and $S^{(j)}$ are pairwise opaque. This is somewhat surprising, as one might think that we do not need probabilistically opaque HMMs in order to have probabilistic system opacity (e.g., we could hide some of the observation sequences generated by $S^{(i)}$, with an HMM $S^{(k)}$ and other sequences in another HMM $S^{(k')}$, without $S^{(k)}$ and $S^{(k')}$ needing to be probabilistically opaque with $S^{(i)}$). This line of thought is not correct: if two HMMs are not probabilistically opaque, then the probability of error tends to zero eventually for all observation sequences that can be generated by $S^{(i)}$ (this is part of the proof of the verification of two probabilistically opaque HMMs). Thus, if there is no HMM that is probabilistically

opaque with $S^{(i)}$, then one can always distinguish that the observed sequence is generated by $S^{(i)}$, with certainty that tends, at least asymptotically, to unity.

Remark 7 Probabilistic system opacity can be verified with polynomial complexity in terms of the size of the state space of the system. Indeed, according to Theorem 3, to verify the notion of probabilistic system opacity, we need to check for pairwise probabilistic opacity for all HMM pairs $S^{(i)}, S^{(j)}$, where $i, j \in \{1, ..., m\}$ (m^2 pairs). According to Theorem 2 two HMMs are pairwise probabilistically opaque iff they are probabilistically equivalent from steady-state. Thus, we need to compute the steady-state, which can be done with polynomial complexity, and then we need to check for probabilistic equivalence from steady-state, which can also be done with polynomial complexity (Tzeng 1989).

5 Connections with probabilistic current–state opacity

Probabilistic system opacity can be seen as a special case of probabilistic current–state opacity, as introduced in Saboori and Hadjicostis (2014) for a given HMM $(Q, E, \Delta, \Lambda, \pi_0)$. In Saboori and Hadjicostis (2014) the authors defined probabilistic current–state opacity in a general setup, where there exists a secret set of states, $Q_s \subseteq Q$, that need to remain "hidden" from an eavesdropper in the following (probabilistic) sense: the confidence of the eavesdropper, captured by the conditional probability that the system state lies in the set of secret states Q_s after the execution of observation sequence ω should not exceed a threshold θ. In general, the problem is proven undecidable in Saboori and Hadjicostis (2014). Probabilistic system opacity (which, as we argue below, is a special case of probabilistic current–state opacity) is shown not only to be decidable, but also verifiable with polynomial complexity.

To establish the connection between probabilistic system opacity and current–state opacity, we first reproduce the definition of probabilistic current–state opacity, with a little modification, in order to match the notation used in this paper.

Definition 9 Given an HMM $S = (Q, E, \Delta, \Lambda, \pi_0)$ and a set of secret states $Q_s \subseteq Q$, HMM S is probabilistically current–state opaque with respect to Q_s, and a parameter θ or (Q_s, θ)–probabilistically current–state opaque, if

$$\forall \omega \in L(S) : \sum_{\forall q_i \in Q_s} \rho_\omega(i) - \sum_{\forall q_i \in Q_s} \pi_0(i) \leq \theta,$$

where $\rho_\omega = A_{\omega[n]} A_{\omega[n-1]} ... A_{\omega[1]} \pi_0$, with $\omega = \omega[1]\omega[2]...\omega[n]$.

Our setup (assumed here to have m=2 HMMs with prior probabilities P_1 and P_2) can be seen as a special case of probabilistic current–state opacity as follows:[4]

Let $S = (Q, E, \Delta, \Lambda, \pi_0)$, be the union of two given irreducible HMMs $S^{(1)} = (Q^{(1)}, E^{(1)}, \Delta^{(1)}, \Lambda^{(1)}, \pi_0^{(1)})$ and $S^{(2)} = (Q^{(2)}, E^{(2)}, \Delta^{(2)}, \Lambda^{(2)}, \pi_0^{(2)})$, where

- $\pi_0 = \begin{bmatrix} P_1 \pi_0^{(1)} \\ P_2 \pi_0^{(2)} \end{bmatrix}$ (note $\pi_0 > 0$)
- any $q_k \in Q^{(i)} : \frac{\pi_0(q_k)}{P_i} = \pi_0^{(i)}(q_k)$

[4]For simplicity we describe a case with only two HMMs, but the connection with probabilistic current-state opacity, can be easily generalised to the case of m HMMs.

- $E = E^{(1)} \cup E^{(2)}$
- $Q = Q^{(1)} \cup Q^{(2)}$
- The functions Δ and Λ are defined as follows:
  - {For any $q_1 \in Q^{(1)}$ and $q_2 \in Q^{(2)}$}, we have $\{\Delta(q_1, q_2) = \Delta(q_2, q_1) = 0\}$
  - {For any $i \in \{1, 2\}$ and $q_k^{(i)}, q_{k'}^{(i)} \in Q^{(i)}$}, we have
$$\left\{\Delta\left(q_k^{(i)}, q_{k'}^{(i)}\right) = \Delta^{(i)}\left(q_k^{(i)}, q_{k'}^{(i)}\right)\right\}$$
  - {For any $i \in \{1, 2\}$, $q_k^{(i)}, q_{k'}^{(i)} \in Q^{(i)}$} and $\sigma \in E$}, we have
$$\left\{\Lambda\left(q_k^{(i)}, \sigma, q_{k'}^{(i)}\right) = \Lambda^{(i)}\left(q_k^{(i)}, \sigma, q_{k'}^{(i)}\right)\right\}.$$

In the above setup, the problem can be decomposed, from the system's perspective, into two probabilistic current-state opacity problems. This is because we need to protect both systems ($S^{(1)}$ and $S^{(2)}$). Thus, we need to take into consideration two cases depending on the chosen HMM. If we chose $S^{(1)}$ (or $S^{(2)}$) then the set of secret states is $Q_s = Q^{(1)}$ (or $Q_s = Q^{(2)}$) respectively.

Case 1. If we chose $S^{(1)}$, then $Q_s = Q^{(1)}$ and probabilistic current-state opacity implies that $\forall\omega \in L(S) : \alpha_1(\omega) - P_1 \leq \theta$;

Case 2. If we chose $S^{(2)}$, then $Q_s = Q^{(2)}$ and probabilistic current-state opacity implies that $\forall\omega \in L(S) : \alpha_2(\omega) - P_2 \leq \theta$.

We can easily prove the following for all $\omega \in L(S)$:

1. $\alpha_1(\omega) = 1 - \alpha_2(\omega)$
2. $\alpha_2(\omega) \overset{1,\text{Case 1}}{\geq} (1 - P_1) - \theta = P_2 - \theta$
3. $\alpha_1(\omega) \overset{1,\text{Case 2}}{\geq} (1 - P_2) - \theta = P_1 - \theta$

1. If both Case 1 and Case 2 hold with $\theta < \min\{P_1, P_2\}$, then for $i \in \{1, 2\}$, we have that $\alpha_i(\omega) \overset{2,3}{\geq} \alpha_0$, where $\alpha_0 = \min\{P_1 - \theta, P_2 - \theta\}$ $(0 < \alpha_0 < 1)$. In that case the system is also probabilistically opaque.
2. If the system is probabilistically opaque, then there exists some $1 > \alpha_0 > 0$ such that for all $\omega \in L(S)$, $\alpha_1(\omega) \geq \alpha_0 \overset{1}{\Rightarrow} \alpha_2(\omega) - P_2 \leq 1 - (\alpha_0 + P_2)$ and $\alpha_2(\omega) \geq \alpha_0 \overset{1}{\Rightarrow} \alpha_1(\omega) - P_1 \leq 1 - (\alpha_0 + P_1)$. This implies that S is (Q_s, θ)–probabilistically current–state opaque (both Case 1 and Case 2 hold), for $\theta \geq \max\{1-(\alpha_0+P_1), 1-(\alpha_0+P_2)\}$. In order to have a meaningful θ we want $\alpha_0 + P_1 < 1$ and $\alpha_0 + P_2 < 1$ equivalently $\alpha_0 < 1 - P_2 = P_1$ and $\alpha_0 < 1 - P_1 = P_2$ or $\alpha_0 < \min\{P_1, P_2\}$. This is always valid, because if a system is probabilistically opaque for a threshold α_0' then the system is also probabilistically opaque for all $\alpha_0 \leq \alpha_0'$.

6 Conclusions and future work

In this work, we analyzed and verified a notion of probabilistic opacity related to distinguishing the true system among a set of possible systems, based on a sequence of observations. We established necessary and sufficient conditions under which this notion of probabilistic system opacity can be verified with polynomial complexity. In order to establish polynomial verification algorithms, we analyzed the specific case of HMMs with all

initial states possible. There is an interesting feature in this case: despite the fact that the notion of probabilistic system opacity is concerned with probabilities conditioned on the observation sequence, probabilistic system opacity turns out to be the exact opposite of the notion that captures the ability to classify the system (i.e., distinguishing the correct HMM), which is a notion that relies on the ensemble of observation sequences. An open problem is to solve the general case starting with an arbitrary initial state distribution in each HMM. In the described setup, we choose an HMM out of m possible HMMs, which is essentially a multiple hypothesis testing problem, applied to the HMM setup we have. An interesting extension of this setup would be to involve a change in the behavior of the chosen HMM that occurs at an *a priori* unknown instant of time. In addition to detecting the underlying HMM (the one that was chosen at the switch time), an additional challenge here is the fact that the instant at which the switch occurs is also unknown. An interesting approach to overcome this challenge would be to use sequential methods i.e., repeated sequential probability ratio test (SPRT) as in Chen and Willett (2000) and Cardenas et al. (2004).

Appendix A: Proof of Theorem 1

In order to simplify the notation, we present a proof that is appropriate for a decision rule that we introduce. We name this rule, "empirical rule" (A.1) which is based on the total number of single events. The empirical rule is useful only when we can distinguish statistically the two HMMs, counting only single events. This rule is illustrated for the case of single events, but can also be applied for (finite) event sequences that can be produced by at least one of the two HMMs. This is equivalent to the statement "$S^{(1)}$ and $S^{(2)}$ are not probabilistically equivalent from steady–state" in Theorem 1. The statistical measure that we use is the distance in variation, which compares the expected frequency of an event, against the measured frequency. The expected frequency of an event can be computed easily and is equivalent to what we call "steady-state emission probability" for a single event. In order to prove the asymptotical tightness of the upper bound on the probability of misclassification, we use a generalisation of "Hoeffding's inequality" (Glynn and Ormoneit 2002), for functions of Markov chains. We apply the generalised Hoeffding's inequality, to distinguish two enhanced HMM models defined in 1, and we prove that these enhanced models, are also irreducible and aperiodic, as long as the given HMMs are irreducible and aperiodic. The expected frequency for event sequences can be computed without any state explosion in the enhanced models (compared to the initial HMMs) something that is established in Lemma 6. Finally, we use the generalised Hoeffding's inequality, combined with the empirical rule, in order to establish Theorem 1. Most of the material in this appendix was presented in our previous work in Keroglou and Hadjicostis (2014) where we explored an emprirical frequency rule for stochastic fault diagnosis. The new material in this paper is related to important proofs that were omitted in our previous work due to space limitations. Specifically, we provide a complete proof in three important Lemmas (Lemmas 4, 5, and 6), and in Section A.5 that uses Hoeffding's inequality to obtain an upper bound on the probability of misclassification.

A.1 Empirical rule

We define a suboptimal rule for HMM classification, which compares the total number of occurrences of each event (see Definition 10) against their frequencies (Definition 11) expected in each of the two HMMs.

Definition 10 (Fraction of times event e_i appears ($m_n(e_i)$)). Suppose we are given an observation sequence of length n ($\omega = \omega[1]\cdots\omega[n]$). We define $m_n(e_i) = \frac{1}{n}\sum_{t=1}^{n} g_{e_i}(\omega[t])$, where

$$g_{e_i}(\omega[t]) = \begin{cases} 1, & \text{if } \omega[t] = e_i, \\ 0, & \text{otherwise.} \end{cases}$$

In other words, $m_n(e_i)$ is the fraction of times event e_i appears in observation sequence ω.

Definition 11 (Distance in Variation $d_V(v, v')$ Between Two Probability Vectors v, v'). The distance in variation (Dembo and Zeitouni 1998) between two $|E|$-dimensional probability vectors v, v' is defined as

$$d_V(v, v') = \frac{1}{2}\sum_{j=1}^{|E|} |v(j) - v'(j)| \geq 0 \ ,$$

where $v(j)$ ($v'(j)$) is the jth entry of vector v (v').

Let the stationary emission probabilities in Definition 6 for HMM $S^{(1)}$ ($S^{(2)}$) be denoted by the $|E|$-dimensional vector $\pi_e^{(1)} = \left[\pi_{e_1}^{(1)}, ..., \pi_{e_{|E|}}^{(1)}\right]^T$ (respectively, by $\pi_e^{(2)} = [\pi_{e_1}^{(2)}, ..., \pi_{e_{|E|}}^{(2)}]^T$). Then, we have $d_V(\pi_e^{(1)}, \pi_e^{(2)}) = \frac{1}{2}\sum_{j=1}^{|E|} |\pi_{e_j}^{(1)} - \pi_{e_j}^{(2)}|$.

Definition 12 (Empirical Rule). Given two irreducible and aperiodic HMMs, $S^{(1)}$ and $S^{(2)}$, and a sequence of observations $\omega = \omega[1]\omega[2]\cdots\omega[n]$, we perform classification using the following suboptimal rule.

- We first compute $m_n = [m_n(e_1), m_n(e_2), \cdots, m_n(e_{|E|})]^T$ as in Definition 10.
- We then set $\theta = \frac{1}{2}d_V(\pi_e^{(1)}, \pi_e^{(2)})$, where $\pi_e^{(j)}$, $j \in \{1, 2\}$, is the stationary emission probability vector for $S^{(j)}$, and compare

$$d_V(m_n, \pi_e^{(1)}) \gtrless \theta \ . \tag{4}$$

- We decide in favor of $S^{(1)}$ ($S^{(2)}$) if the right (left) quantity is larger.

A.2 Enhanced HMM model

In this section we define a function of the states of the underlying Markov chain of the two HMMs $S^{(1)}$ and $S^{(2)}$, that counts the occurrences of each event $e_i \in E$, with which we arrive at that state. This is not necessarily possible in $S^{(j)}$, $j \in \{1, 2\}$, because in general we can reach a state via different events. The reason we need to define a function of the states is so that we can analyze the empirical rule (Definition 12) using existing techniques for Markov chain analysis.

First, we obtain, for each of the given HMMs, an enhanced construction that allows us to discriminate the transition to the same state but via different events. We prove that our enhanced construction inherits the properties of irreducibility and aperiodicity (the two conditions needed to apply Theorem 2) from the corresponding original HMM. The two enhanced HMM models are denoted by $\widetilde{S}^{(j)} = \{\widetilde{Q}^{(j)}, E, \widetilde{\Delta}^{(j)}, \widetilde{\Lambda}^{(j)}, \widetilde{\pi}_0{}^{(j)}\}$, $j \in \{1, 2\}$. The enhanced construction creates replicas of each state, depending on the event via which one reaches this state. Thus, for each state $q_h \in Q^{(j)}$, we create states $q_{h,e_i} \in \widetilde{Q^{(j)}}$, $e_i \in E$, to represent that we reach state $q_h \in Q^{(j)}$ under the output symbol e_i. Clearly, we end up with at most $|\widetilde{Q}^{(j)}| = |Q| \times |E|$ states.

The following discussion applies to each original HMM and its enhanced model (we drop j, $j \in \{1, 2\}$, to simplify notation). In the state probability vectors $\pi[k]$, $\widetilde{\pi}[t]$, where t is the current state epoch, states are indexed in the order shown below

$$\pi[t] = \begin{bmatrix} \pi[t](q_1) \\ \pi[t](q_2) \\ \vdots \\ \pi[t](q_{|Q|}) \end{bmatrix}, \quad \widetilde{\pi}[t] = \begin{bmatrix} \widetilde{\pi}[t](q_{1,e_1}) \\ \widetilde{\pi}[t](q_{1,e_2}) \\ \vdots \\ \widetilde{\pi}[t](q_{1,e_{|E|}}) \\ \widetilde{\pi}[t](q_{2,e_1}) \\ \vdots \\ \widetilde{\pi}[t](q_{|Q|,e_{|E|}}) \end{bmatrix}.$$

The matrix $\widetilde{A}_{e_i}$, $e_i \in E$, satisfies $\widetilde{A}_{e_i}(q_{h,e_i}, q'_{h,e'_i}) = A_{e_i}(q_h, q'_h)$, $\forall e'_i \in E$ and $\forall q_h, q'_h \in Q$ (zero otherwise). We also have for, $e'_{i'} \in E$ and $q_{h,e_i}, q'_{h,e'_i} \in \widetilde{Q}$, $\widetilde{\Lambda}(q'_{h,e'_i}, e_i, q_{h,e_i}) = \widetilde{A}_{e_i}(q_{h,e_i}, e_i, q'_{h,e'_i})$ (zero otherwise). We observe that matrix $\widetilde{A}_{e_i}$ is constructed by blocks of matrix A_{e_i}. If we define row-vector $R_{|E|} = \underbrace{[11\cdots1]}_{|E|-times}$ and let

$$R_{i,|E|} = \underbrace{[0\cdots0\,1\,0\cdots0]}_{\text{single one at } i \text{ th position}},$$

then the state transition matrix $\widetilde{A}^{(j)}_{e_i}$ for the enhanced model $\widetilde{S}^{(j)}$ can be written as[5]

$$\widetilde{A}^{(j)}_{e_i} = A^{(j)}_{e_i} \otimes \left(R^T_{i,|E|} \otimes R_{|E|}\right).$$

Example 2 We create the enhanced HMM models $\widetilde{S}^{(1)}$ (shown in Fig. 3) and $\widetilde{S}^{(2)}$ for $S^{(1)}$ and $S^{(2)}$ respectively (shown in Fig. 1). We note that the underlying state transition matrix, for each enhanced model, is irreducible and aperiodic (as we will see,

[5]The Kronecker product (Brewer 1978) $A \otimes B$ of a $p \times q$ matrix A with an $m \times n$ matrix B is the $pm \times qn$ matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1q}B \\ a_{21}B & a_{22}B & \cdots & a_{2q}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1}B & a_{p2}B & \cdots & a_{pq}B \end{bmatrix}.$$

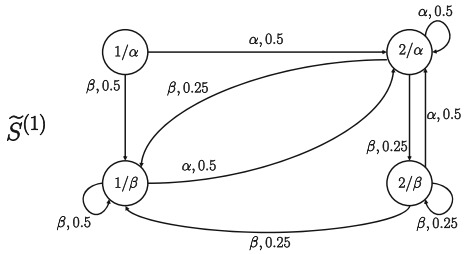


Fig. 3 Enhanced model $\widetilde{S}^{(1)}$ for HMM model $S^{(1)}$ in Fig. 1

$\widetilde{S}^{(j)}$ will be irreducible and aperiodic as long as $S^{(j)}$ is irreducible and aperiodic). The corresponding $\widetilde{A}_{\alpha}^{(1)}$, $\widetilde{A}_{\beta}^{(1)}$, $\widetilde{A}_{\alpha}^{(2)}$, $\widetilde{A}_{\beta}^{(2)}$ are as follows:

$$\widetilde{A}_{\alpha}^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.50 & 0.50 & 0.50 & 0.50 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \widetilde{A}_{\beta}^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.50 & 0.50 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.25 & 0.25 \end{bmatrix},$$

$$\widetilde{A}_{\alpha}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.60 & 0.60 & 0.06 & 0.06 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \widetilde{A}_{\beta}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.40 & 0.40 & 0.04 & 0.04 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.90 & 0.90 \end{bmatrix}.$$

A.3 Required conditions for using Hoeffding's inequality on enhanced models

Proposition 1 (Hoeffding's Inequality on Enhanced HMM Model) *Consider enhanced HMMs,* $\widetilde{S}^{(j)} = \{\widetilde{Q}^{(j)}, E, \widetilde{\Delta}^{(j)}, \widetilde{A}^{(j)}, \widetilde{\pi}_0^{(j)}\}$, $j \in \{1, 2\}$, *with* $|E|$ *events and transition matrix* $\widetilde{A}^{(j)}$. *Assuming the Markov chains that correspond to the enhanced models* $\widetilde{S}^{(j)}$, $j \in \{1, 2\}$, *are irreducible and aperiodic, we denote their stationary distributions by* $\widetilde{\pi}^{(j)} > 0$ *and stationary emission distribution for events* $e_i \in E$ *by* $\widetilde{\pi}_e^{(j)} > 0$.

Using the enhanced models ($\widetilde{S}^{(1)}$ *and* $\widetilde{S}^{(2)}$*) for each* $e_i \in E$, *we define the indicator functions* $f_{e_i}(q_{h,e_j})$, $\forall q_{h,e_j} \in \widetilde{Q}$, *as*

$$f_{e_i}(q_{h,e_j}) = \begin{cases} 1, & \text{if } e_j = e_i, \\ 0, & \text{otherwise.} \end{cases}$$

Let $m_n(e_i) = \frac{1}{n} \sum_{t=1}^{n} f_{e_i}(q[t])$, *i.e., the* $|E|$*-dimensional vector* $m_n = [m_n(e_1), m_n(e_2), ..., m_n(e_{|E|})]^T$ *denotes the empirical frequencies with which each event appears in the given observation window of length n. Let* M_j, *be the smallest integer such that* $(\widetilde{A}^{(j)})^{M_j} > 0$, *element-wise, and* $\lambda_j = \min_{l,l'} \left\{ \frac{(\widetilde{A}^{(j)})^{M_j}(l,l')}{\widetilde{\pi}^{(j)}(l)} \right\}$, *where* $\widetilde{\pi}^{(j)}(l)$ *is the stationary distribution of* $\widetilde{S}^{(j)}$. *As long as the enhanced model* $\widetilde{S}^{(j)}$ *is irreducible and aperiodic, it*

can be shown Glynn and Ormoneit (2002) and Hadjicostis (2005) that the following is true for $n > \frac{2M_j}{\lambda_j \epsilon}$, *for each event* e_i $(1 \leq i \leq |E|)$, *and for* $F^{(j)}(n) = \exp\left(-\frac{\lambda_j{}^2\left(n\epsilon - \frac{2M_j}{\lambda_j}\right)^2}{2nM_j{}^2}\right)$:

$$Pr(|m_n(e_i) - \widetilde{\pi}_e^{(j)}(e_i)| \geq \epsilon) \leq F^{(j)}(n). \tag{5}$$

In order to use Eq. 5, we need $\widetilde{S}^{(j)}$ to correspond to an irreducible (Definition 3) and aperiodic (Definition 4) Markov chain. We now show that $\widetilde{S}^{(j)}$ is irreducible and aperiodic if $S^{(j)}$ is irreducible and aperiodic. Also, we establish that $\widetilde{\pi}_e^{(j)} = \pi_e^{(j)}$.

Lemma 4 *If HMM* $S^{(j)} = (Q^{(j)}, E^{(j)}, \Delta^{(j)}, \Lambda^{(j)}, \pi_0^{(j)})$ *is irreducible (Definition 3), then the enhanced HMM* $\widetilde{S}^{(j)} = \{\widetilde{Q}^{(j)}, E, \widetilde{\Delta}^{(j)}, \widetilde{\Lambda}^{(j)}, \widetilde{\pi}_0{}^{(j)}\}$ *is also irreducible.*

Proof We prove irreducibility by establishing the property that any state $q_{h,e_i} \in \widetilde{Q}^{(j)}$ that does not belong to a set of strongly connected states (Definition 3), may exhibit outgoing transitions but will have no incoming transition. Consider in the enhanced model the set of states

$$\widetilde{Q}_{ss} = \{q_{m,e} \in \widetilde{Q} | \exists q_{m'} \in Q, \exists e \in E \text{ s.t. } \Lambda(q_{m'}, e, q_m) > 0\}.$$

Since the set of states Q in the original system is strongly connected, we can easily show that the states in $\widetilde{Q}_{ss}$ are strongly connected: given $q_{m,e}, q_{m',e'} \in \widetilde{Q}$ we can find a path to connect them as follows: Let $q_{m''}$ be such that $\Lambda(q_{m''}, e', q_{m'}) > 0$. Then, we can find a path

$$q_m \xrightarrow{e_{i_1}} q_{i_1} \xrightarrow{e_{i_2}} q_{i_2} \to \cdots \xrightarrow{e_{i_t}} q_{i_t} = q_{m''} \ .$$

(because the original HMM is irreducible). Therefore

$$q_{m,e} \xrightarrow{e_{i_1}} q_{i_1,e_1} \xrightarrow{e_{i_2}} q_{i_2,e_{i_2}} \to \cdots \xrightarrow{e_{i_t}} q_{i_t} = q_{m''} \xrightarrow{e'} q_{m',e'}$$

is a path that connects $q_{m,e} \in \widetilde{Q}$ to $q_{m',e'} \in \widetilde{Q}$. We finally conclude that the states that do not belong to the set of strongly connected states, have only outgoing transitions. Therefore, by choosing an appropriate initial distribution function that excludes all these transient states,[6] we can ensure that all of these transient states will never be visited. □

Lemma 5 *If HMM* $S^{(j)} = (Q^{(j)}, E^{(j)}, \Delta^{(j)}, \Lambda^{(j)}, \pi_0^{(j)})$ *is aperiodic (Definition 4), then the enhanced HMM* $\widetilde{S}^{(j)} = \{\widetilde{Q}^{(j)}, E, \widetilde{\Delta}^{(j)}, \widetilde{\Lambda}^{(j)}, \widetilde{\pi}_0{}^{(j)}\}$ *is also aperiodic.*

Proof We show that if the enhanced model $\widetilde{S}^{(j)}$ is periodic with period k, this contradicts the fact that $S^{(j)}$ is aperiodic. Suppose that $\widetilde{S}^{(j)}$ is periodic with period k (Definition 4 and Lemma 2). This means we can group all possible states of $\widetilde{S}^{(j)}$ to k groups $(\widetilde{C}_1, \widetilde{C}_2, \cdots, \widetilde{C}_k)$ such that for a state $q_{l,e} \in \widetilde{C}_m$, there exist one-step transitions only to states in $\widetilde{C}_{m'}$, where $m' = m + 1 \mod k$.

[6] We can always do this since subsequent behavior of the enhanced model does not depend on whether we start from state $q_{h,e}$ or $q_{h,e'}$.

Due to the construction of enhanced models, the outgoing behaviour of $q_{l,e}$ states $\forall e \in E$ are copies of the outgoing behaviour of $q_l \in Q$. We can easily see that if there exists $q_{l,e} \in \widetilde{Q}$, that belongs to $\widetilde{C}_m$, then also $q_{l,e'} \in \widetilde{Q}$ belongs to $\widetilde{C}_m$, for all $e, e' \in E$ (due to the same outgoing behaviour). Thus, we can also group $q \in Q$ into $C_i, i \in \{1, 2, ..., k\}$, classes. Thus, $S^{(j)}$ is periodic, with period k, which is a contradiction. □

A.4 Consistency of stationary emission probabilities for $S^{(j)}$ and $\widetilde{S}^{(j)}$

We now show that in the enhanced model $\widetilde{S}^{(j)}$, the stationary emission probabilities of each event are consistent with the original model $S^{(j)}$ for $j = 1, 2$.

Lemma 6 *The computed stationary emission probabilities for symbols in the enhanced model $\widetilde{S}^{(j)}$, $j \in \{1, 2\}$ which is denoted respectively by $\widetilde{\pi}_e^{(j)}$, is identical to $\pi_e^{(j)}$ corresponding to $S^{(j)}$.*

Proof Let $\widetilde{\pi}_s^{(j)}$ denote the steady-state distribution vector in the enhanced model j. Then, we have that under each model

$$\pi_s^{(j)} = (I_n \otimes R_{|E|}) \times \widetilde{\pi}_s^{(j)}, \ j \in \{1, 2\}.$$

For $S^{(j)}$ and $\forall e_i \in E$, the stationary emission probability $\pi_e^{(j)}(e_i)$ can be expressed as

$$\pi_e^{(j)}(e_i) = R_n \times \left(A_{e_i}^{(j)} \times \pi_s^{(j)}\right),$$

whereas for the enhanced model[7] $\widetilde{S}^{(j)}$.

$$\begin{aligned}
\widetilde{\pi}_e^{(j)}(e_i) &= R_{n|E|} \times \widetilde{A}_{e_i}^{(j)} \times \widetilde{\pi}_s^{(j)} \\
&= R_{n|E|} \times \left(A_{e_i}^{(j)} \otimes \left(R_{i,|E|}^T \otimes R_{|E|}\right)\right) \times \widetilde{\pi}_s^{(j)} \\
&= (R_n \otimes R_{|E|}) \times \left(A_{e_i}^{(j)} \otimes \left(R_{i,|E|}^T \otimes R_{|E|}\right)\right) \times \widetilde{\pi}_s^{(j)} \\
&= (R_n \times A_{e_i}^{(j)}) \otimes \left(R_{|E|} \times \left(R_{i,|E|}^T \otimes R_{|E|}\right)\right) \times \widetilde{\pi}_s^{(j)} \\
&= \left(\left(R_n \times A_{e_i}^{(j)}\right) \otimes R_{|E|}\right) \times \widetilde{\pi}_s^{(j)} \\
&= \left(R_n \times A_{e_i}^{(j)}\right) \otimes (R_1 \times R_{|E|}) \times \widetilde{\pi}_s^{(j)} \\
&= R_n \times \left(A_{e_i}^{(j)} \otimes R_{|E|}\right) \times \widetilde{\pi}_s^{(j)}.
\end{aligned}$$

[7] In performing this analysis, we use the following well known properties of the Kronecker product: 1. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$; 2. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ for matrices A, B, C, D of appropriate dimensions (Brewer 1978).

Moreover, we have

$$\begin{aligned}
\pi_e^{(j)}(e_i) &= R_n \times \left(A_{e_i}^{(j)} \times \pi_s^{(j)}\right) \\
&= R_n \times \left(A_{e_i}^{(j)} \times (I_n \otimes R_{|E|}) \times \widetilde{\pi}_s^{(j)}\right) \\
&= R_n \times \left(A_{e_i}^{(j)} \otimes R_1\right) \times (I_n \otimes R_{|E|}) \times \widetilde{\pi}_s^{(j)} \\
&= R_n \times \left(A_{e_i}^{(j)} \times I_n\right) \otimes (R_1 \times R_{|E|}) \times \widetilde{\pi}_s^{(j)} \\
&= R_n \times \left(A_{e_i}^{(j)} \otimes R_{|E|}\right) \times \widetilde{\pi}_s^{(j)} \\
&= \widetilde{\pi}_e^{(j)}(e_i),
\end{aligned}$$

which allows us to conclude that $\pi_e^{(j)}(e_i) = \widetilde{\pi}_e^{(j)}(e_i), \forall e_i \in E$. □

A.5 Upper bound on the probability of error

Given two HMMs $S^{(1)}$ and $S^{(2)}$ (each irreducible and aperiodic), we construct the corresponding enhanced HMM models $(\widetilde{S}^{(1)}, \widetilde{S}^{(2)})$ with underlying irreducible and aperiodic Markov chains $\widetilde{MC}^{(1)} = (\widetilde{Q}^{(1)}, \widetilde{A}^{(1)}, \widetilde{\pi}_0{}^{(1)})$ and $\widetilde{MC}^{(2)} = (\widetilde{Q}^{(2)}, \widetilde{A}^{(2)}, \widetilde{\pi}_0{}^{(2)})$ (i.e., this means that $\widetilde{A}^{(1)}$ and $\widetilde{A}^{(2)}$ are primitive matrices). Suppose we have[8] $d_V(\widetilde{\pi}_e^{(1)}, \widetilde{\pi}_e^{(2)}) > 0$. Then, if we apply the empirical rule and use Hoeffding's inequality (Proposition 1), we obtain the upper bound on the probability of error using the empirical rule where $F(n)$ is given by

$$F(n) = \max\{F^{(1)}(n), F^{(2)}(n)\}\ . \tag{6}$$

Proof We consider two error cases :

Case 1: Decide $S^{(1)}$ when the system is $S^{(2)}$;
Case 2: Decide $S^{(2)}$ when the system is $S^{(1)}$.

Case 1:
The decision of $S^{(1)}$ is equivalent to the event

$$H^{(1)} : d_V(m_n, \pi_e^{(1)}) < \theta,$$

which necessarily implies $d_V(m_n, \pi_e^{(2)}) \geq \theta$ $\left(\text{for } \theta = \frac{1}{2} d_V\left(\pi_e^{(1)}, \pi_e^{(2)}\right)\right)$. Otherwise, we reach a contradiction, because $d_V(m_n, \pi_e^{(2)}) < \theta$ and $d_V(m_n, \pi_e^{(1)}) < \theta$ imply

$$\begin{aligned}
d_V\left(\pi_e^{(1)}, \pi_e^{(2)}\right) &< d_V\left(m_n, \pi_e^{(1)}\right) + d_V\left(m_n, \pi_e^{(2)}\right) \\
&= 2\theta \\
&= d_V\left(\pi_e^{(1)}, \pi_e^{(2)}\right)\ .
\end{aligned}$$

[8] Equivalently, $d_V(\pi_e^{(1)}, \pi_e^{(2)}) > 0$ (Lemma 6).

Thus, $H^{(1)}$ implies $d_V(m_n, \pi_e^{(1)}) \geq \theta$, which implies

$$H_k^{(1)} : \left\{ \exists e_k \in E \text{ such that } \left| m_n(e_k) - \pi_e^{(2)}(e_k) \right| > \frac{\theta}{|E|} \right\}.$$

Therefore, we have to consider two different subcases:

a) $m_n(e_k) - \pi_e^{(2)}(e_k) > 0$,
b) $m_n(e_k) - \pi_e^{(2)}(e_k) < 0$.

The probability of error for Case 1 and subcase a), after n observations, is

$$\begin{aligned} \Pr(\text{error at } n,\ \text{Case 1}) &= \Pr(H^{(1)}|S^{(2)})P(S^{(2)}) \\ &\leq \Pr\left(H_k^{(1)}|S^{(2)}\right) P(S^{(2)}) \\ &\leq F^{(2)}(n)P(S^{(2)}) , \end{aligned}$$

where $\Pr\left(H_k^{(1)}|S^{(2)}\right) \equiv \Pr(m_n(e_k) - \pi_e^{(2)}(e_k)) > \frac{\theta}{|E|}) \leq F_n^{(2)}$, for $\epsilon = \frac{\theta}{|E|}$.

In Case 1a) we can immediately apply Eq. 5, but in Case 1b) in order to find a positive measure we choose to count the number of appearances of all elements in $k^c = \{e \in E \mid \text{s.t. } e \neq e_k\}$, i.e., all possible events except $e_k \in E$. Then $m_n(e_{k^c}) - \pi_e^{(2)}(e_{k^c}) = (1 - m_n(e_k)) - (1 - \pi_e^{(2)}(e_k)) > 0$, and we can apply (5), which leads us to the same bound.

Case 2:

With the same reasoning as in Case 1, we establish the following inequality

$$\begin{aligned} \Pr(\text{error at } n,\ \text{Case 2}) &= \Pr(H^{(2)}|S^{(1)})P(S^{(1)}) \\ &\leq \Pr\left(H_{k'}^{(2)}|S^{(1)}\right) P(S^{(1)})) \\ &\leq F^{(1)}(n)P(S^{(1)}), \end{aligned}$$

where $H^{(2)} : d_V(m_n, \pi_e^{(1)}) > \theta$, which implies that $H_{k'}^{(2)}$: {There exists at least one $e_{k'}$ such that $|m_n(e_{k'}) - \pi_e^{(1)}(e_{k'})| > \frac{\theta}{|E|}$, where $e_{k'} \in E$ }. The claim follows using similar arguments as in Case 1.

Finally, we prove that

$$\begin{aligned} \Pr(\text{error at } n) &= \Pr(\text{error at } n,\ \text{Case 1}) + \Pr(\text{error at } n,\ \text{Case 2}) \\ &= \Pr(H^{(1)}|S^{(2)})P(S^{(2)}) + \Pr(H^{(2)}|S^{(1)})P(S^{(1)})) \\ &\leq \Pr\left(H_k^{(1)}|S^{(2)}\right) P(S^{(2)}) + \Pr\left(H_{k'}^{(2)}|S^{(1)}\right) P(S^{(1)})) \\ &\leq F(n)(P(S^{(1)}) + P(S^{(2)})) \equiv F(n) \end{aligned}$$

In other words, if two HMMs have different expected frequencies for at least one single event, this allows to asymptotically distinguish the two HMMs using the empirical rule. Without loss of generality, we can extend this result to event sequences, that have different expected frequencies. Knowing that the expected frequency of an event sequence can be computed as the probability of occurrence of an event sequence, generated by an HMM that starts at the steady-state, then if there is at least one event sequence with different expected frequencies, is equivalent to the fact that the two HMMs, are not probabilistically equivalent from the steady-state. This concludes the proof. □

References

Athanasopoulou E, Hadjicostis CN (2008) Probability of error bounds for failure diagnosis and classification in hidden Markov models. In: Proceedings of IEEE conference on decision and control, pp 1477–1482

Badouel E, Bednarczyk M, Borzyszkowski A, Caillaud B, Darondeau P (2006) Concurrent secrets. In: Proceedings of the 8th international workshop on discrete event systems, pp 51–57

Berard B, Mullins J, Sassolas M (2010) Quantifying opacity. In: Proceedings of 7th international conference on the quantitative evaluation of systems (QEST), pp 263–272

Brard B, Mullins J, Sassolas M (2015) Quantifying opacity. Math Struct Comput Sci 25(2):361–403

Brewer J (1978) Kronecker products and matrix calculus in system theory. IEEE Trans Circuits Syst 25(9):772–781

Bryans JW, Koutny M, Ryan P (2005a) Modelling dynamic opacity using Petri nets with silent actions. ser. Formal Aspects in Security and Trust. Springer 173:159–172

Bryans JW, Koutny M, Ryan P (2005b) Modelling opacity using Petri nets. Electron Notes Theor Comput Sci 121:101–115

Bryans JW, Koutny M, Mazare L, Ryan P (2005) Opacity generalised to transition systems. In: Proceedings of the 3rd international workshop on formal aspects in security and trust, pp 81–95

Cardenas AA, Baras JS, Ramezani V (2004) Distributed change detection for worms, DDos and other network attacks. In: Proceedings of the 2004 American control conference, vol 2, pp 1008–1013

Cassandras CG, Lafortune S (2007) Introduction to discrete event systems. Springer, Berlin

Chen B, Willett P (2000) Detection of hidden Markov model transient signals. IEEE Trans Aerosp Electron Syst 36(4):1253–1268

Dembo A, Zeitouni O (1998) Large deviations techniques and applications. Springer, New York

Dubreil J, Darondeau P, Marchand H (2008) Opacity enforcing control synthesis. In: Proceedings of 9th international workshop on discrete event systems, pp 28–35

Focardi R, Gorrieri R (1994) A taxonomy of trace–based security properties for CCS. In: Proceedings of the 7th workshop on computer security foundations, pp 126–136

Fu KS (1982) Syntactic pattern recognition and applications. Prentice-Hall, Upper Saddle River

Glynn P, Ormoneit D (2002) Hoeffding's inequality for uniformly ergodic Markov chains. Stat Prob Lett 56:143–146

Hadjicostis CN (2005) Probabilistic detection of FSM single state-transition faults based on state occupancy measurements. IEEE Trans Autom Control 50(12):2078–2083

Keroglou C, Hadjicostis CN (2013) Initial state opacity in stochastic DES. In: Proceedings of 18th conference on emerging technologies factory automation (ETFA), pp 1–8

Keroglou C, Hadjicostis CN (2014) Hidden Markov model classification based on empirical frequencies of observed symbols. In: Proceedings of 12th international workshop on discrete event systems (WODES), pp 7–12

Keroglou C, Hadjicostis CN (2016) Probabilistic system opacity in discrete event systems. In: Proceedings of 13th international workshop on discrete event systems (WODES), pp 379–384

Millen JK (1987) Covert channel capacity. In: Proceedings of IEEE symposium on security and privacy, pp 60–66

Neyman J, Pearson ES (1992) On the problem of the most efficient tests of statistical hypotheses. Springer, New York, pp 73–108

Saboori A, Hadjicostis CN (2007) Notions of security and opacity in discrete event systems. In: Proceedings of 46th IEEE conference on decision and control, pp 5056–5061

Saboori A, Hadjicostis CN (2011) Coverage analysis of mobile agent trajectory via state-based opacity formulations. Control Engineering Practice (Special Issue on Selected Papers from 2nd International Workshop on Dependable Control of Discrete Systems) 19(9):967–977

Saboori A, Hadjicostis CN (2013) Verification of initial-state opacity in security applications of DES. Inf Sci 246:115–132

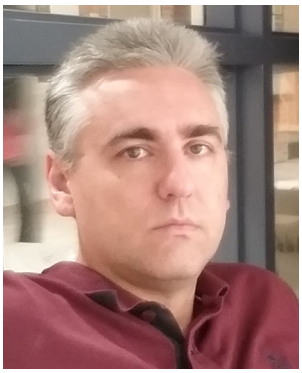
Saboori A, Hadjicostis CN (2014) Current-state opacity formulations in probabilistic finite automata. IEEE Trans Autom Control 59(1):120–133

Seneta E (2006) Non-negative matrices and Markov chains. Springer Series in Statistics, Berlin

Tzeng W-G (1989) The equivalence and learning of probabilistic automata. In: Proceedings of 30th annual symposium on foundations of computer science, pp 268–273

Wittbold JT, Johnson DM (1990) Information flow in nondeterministic systems. In: Proceedings of IEEE computer society symposium on research in security and privacy, pp 144–161

Wu Y-C, Lafortune S (2013) Comparative analysis of related notions of opacity in centralized and coordinated architectures. Discrete Event Dynamic Systems 23(3):307–339



Christoforos Keroglou received the diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki, and the Master and the Ph.D. degree in electrical and computer engineering from the University of Cyprus. He is currently a Post-Doctoral Fellow at the EECS Department, University of Michigan, Ann Arbor. He is interested in contributing to a deeper understanding of optimal decision-making in communication problems. His research interests include state estimation-based problems modeled by logical and stochastic discrete event systems. His main focus is on applications in privacy and fault diagnosis.

Christoforos N. Hadjicostis received the S.B. degrees in electrical engineering, in computer science and engineering, and in mathematics, the M.Eng. degree in electrical engineering and computer science in 1995, and the Ph.D. degree in electrical engineering and computer science in 1999, all from the Massachusetts Institute of Technology, Cambridge. In 1999, he joined the Faculty at the University of Illinois at Urbana-Champaign, where he served as Assistant and then Associate Professor with the Department of Electrical and Computer Engineering, the Coordinated Science Laboratory, and the Information Trust Institute. Since 2007, he has been with the Department of Electrical and Computer Engineering. University of Cyprus, where he is currently Professor and Dean of Engineering. His research focuses on fault diagnosis and tolerance in distributed dynamic systems, error control coding, monitoring, diagnosis and control of large-scale discrete-event systems, and applications to network security, anomaly detection, energy distribution systems, medical diagnosis, biosequencing, and genetic regulatory models. He currently serves as Associate Editor of IEEE Transactions on Automatic Control, IEEE Transactions on Automation Science and Engineering, Automatica, Nonlinear Analysis: Hybrid Systems, and the Journal of Discrete Event Dynamic Systems; he has also served as Associate Editor of IEEE Transactions on Control Systems Technology, and IEEE Transactions on Circuits and Systems I.