2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Logical and probabilistic aspects of state estimation for Markovian systems

Dimitri Lefebvre, Carla Seatzu, Christoforos N. Hadjicostis, Alessandro Giua

***Abstract*— This paper is about state estimation in a class of labeled timed probabilistic automata. In detail, we consider continuous time Markov processes where the occurrence of some transitions produces observable events. Such observations can be used to update and refine the state estimation. In this setting, we discuss how a logical state estimation approach can be used to characterize the probabilistic state estimation whenever a new event is observed or when the system evolves without producing new observations (silent closure). The main results of the paper show that the final behaviour, as the silent closure goes to infinity, cannot be characterized only in terms of the graphical structure of the underlying automaton but also depends on the values of the firing rates.**

I. INTRODUCTION

In a standard Markov model there is no notion of observed output and the only measurable signal that can be used for the purpose of state estimation is the current time value t. Starting from a given initial state probability vector $\boldsymbol{\pi}_0$ that is assumed to be known, the current state probability vector $\boldsymbol{\pi}(t)$ can be computed from the knowledge of the transition rate matrix of the model. Vector $\boldsymbol{\pi}(t)$ allows one to estimate not only the set to which the current state belongs but also to obtain a probability measure associated with all possible values. Thus, a necessary and sufficient condition to ensure that the estimation error goes to zero in probability is the following: the system is ergodic — i.e., there exists a unique stationary distribution for the probability vector — and this distribution is *non-ambiguous* — i.e., it is a standard unit vector.[1] Furthermore, there exists a very elegant structural characterization of this property, namely the underlying graph of the Markov model must consist of a single absorbing component which contains a single state.

The usual way to include observations in Markov models is to associate them to the states according to nondeterministic or probabilistic mappings. Such approaches lead to hidden Markov processes or similar models [13], [23]. In this paper, we consider a different Markovian model, called labeled timed probabilistic automaton [9], [10], which can be seen as a continuous-time Markov process where some transitions are labeled with symbols from a given alphabet E of observable events. When such a transition occurs, an observation (e,t) is produced, where e is the observable event and t is the time of occurrence. This observation mechanism can be used to update and refine the state probability vector whenever a new event occurs or when time elapses with no observation (silent closure). In [9], [10] it has been shown that the conditional state probabilities are piecewise continuous signals: they are continuous when the silent closure increases, and (possibly) present discontinuities each time a new event is observed.

The goal of this paper is that of better characterizing this evolution, in particular as the silent closure, i.e., the time interval from the last observation to the current time, increases. To this aim we investigate the relationship that exists between the state estimation in terms of the conditional state probability vector and the corresponding logical observation in the underlying untimed automaton. Two main cases are considered: 1) the silent closure is finite, 2) the silent closure goes to infinity. A simple and quite intuitive result is provided in the first case, which applies to any labeled timed probabilistic automaton. On the contrary, in the second case the final evolution can be characterized in terms of the eigenstructure of the generator matrix relative to a special automaton that depends on the logical observation. We believe that such results are novel and, surprisingly, they show that the state probability when the silent closure goes to infinity, is not simply related to ergodicity properties of the graphical structure of the automaton as in the purely logical case.

In our opinion the proposed study has applications in numerous problems related to state estimation and detectability in a timed probabilistic setting as far as timed observations are captured. Vulnerability and privacy but also cyber attack detection are concerned at first. We notice that the results presented in this paper may be preliminary to further results in the framework of state estimation and detectability of labeled timed probabilistic automata. This is surely interesting because most of the contributions in the discrete event systems framework related to such problems either ignore probabilistic and timing aspects [1], [6], [14], [16], [18], [19], [21], [22] or consider a probabilistic but untimed setting [7], [8], [15].

This work was partially supported by the French National Research Agency under grant agreement ANR-22-CE10-0002 - MENACES, and by project SERICS (PE00000014) under the Italian MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU.

D. Lefebvre is with GREAH Laboratory, Normandy University, Le Havre, France, dimitri.lefebvre@univ-lehavre.fr

C. Seatzu and A. Giua are with DIEE, University of Cagliari, Cagliari, Italy, {seatzu,giua}@unica.it

C. N. Hadjicostis is with the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus, chadjic@ucy.ac.cy

[1]A standard unit vector is a vector with a unique nonzero component which must necessarily take a unitary value (since, in our case, we are dealing with a probability distribution).

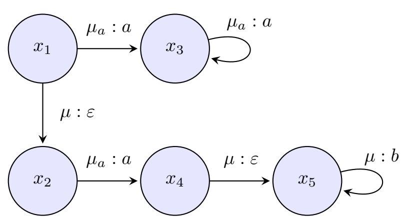

Fig. 1. LTPA in Example 1.

II. BACKGROUND

A. Labeled timed probabilistic automata

This section introduces the basic notions about the reference model used in this paper.

Definition 1 (**Labeled timed probabilistic automata**)*:* A (finite) *labeled timed probabilistic automaton* (LTPA) is a 4-tuple $G = (X, E, \Lambda, \boldsymbol{\pi}_0)$, where:

- $X = \{x_1, x_2, \ldots, x_n\}$ is a finite set of n *states*;
- E is an alphabet of observable events;
- $\Lambda \subseteq X \times E_\varepsilon \times \mathbb{R}_{>0} \times X$ is the *transition relation*, where $E_\varepsilon = E \cup \{\varepsilon\}$ and ε denotes the empty string on E, associated with events that are not observable;
- $\boldsymbol{\pi}_0 \in [0,1]^{1\times n}$ is an *initial probability vector*, with $\sum_{x_i \in X} \pi_{0,i} = 1$, where $\pi_{0,i}$ (the i-th entry of vector $\boldsymbol{\pi}_0$) refers to the initial probability of state x_i. ▲

The transition relation Λ specifies the dynamics of the LTPA: if $(x, e, \mu, x') \in \Lambda$, then a transition from state x to state x', which we call e-jump, may occur after a random delay θ, counted from the time when the system enters x. The delay θ follows an exponential distribution with probability density function $f(\theta) = \mu\ exp(-\mu\theta)$, where μ is the rate of the transition. An e-jump generates an observation e when $e \in E$, while no observation is generated when $e = \varepsilon$ (silent transition).

A *run* of the LTPA G is a trajectory

$$x_{j_0} \xrightarrow{e_1,\ \tau_1} x_{j_1} \xrightarrow{e_2,\ \tau_2} \ldots \xrightarrow{e_K,\ \tau_K} x_{j_K} \quad (1)$$

where, for $i = 1, \ldots, K$, $(x_{j_{i-1}}, e_i, \cdot, x_{j_i}) \in \Lambda$, τ_i denotes the time of occurrence of the i-th jump and $0 < \tau_1 < \tau_2 < \ldots < \tau_K$, where times τ_i are counted from the instant when the system enters x_{j_0}. Such a run determines a *timed sequence* $s_t = (e_1, \tau_1)(e_2, \tau_2)\ldots(e_K, \tau_K) \in (E_\varepsilon \times \mathbb{R}_{\geq 0})^*$, consisting of K pairs: s_t has *duration* $\tau_{last}(s_t) = \tau_K$ (time stamp of the last jump) and *length* $|s_t| = K$. The *empty sequence*, denoted by λ, has duration and length equal to 0.

A timed sequence s_t produces a *timed observation* denoted $P(s_t)$ and defined as $\sigma_t = P(s_t) = (e'_1, \tau'_1)(e'_2, \tau'_2)\ldots(e'_{K'}, \tau'_{K'}) \in (E \times \mathbb{R}_{\geq 0})^*$ obtained from s_t by projection P, which filters out all silent pairs. The observation has duration $\tau_{last}(\sigma_t) = \tau'_{K'}$ and length $|\sigma_t| = K'$. More specifically, $P : (E_\varepsilon \times \mathbb{R}_{\geq 0})^* \rightarrow (E \times \mathbb{R}_{\geq 0})^*$ is formally defined by (i) $P(\lambda) = \lambda$, (ii) $P((e, \tau)) = (e, \tau)$ for $e \in E$ and $P((\varepsilon, \tau)) = \lambda$, (iii) $P(s_t(e, \tau)) = P(s_t)P((e, \tau))$ for $s_t \in (E_\varepsilon \times \mathbb{R}_{\geq 0})^*$ and $(e, \tau) \in E_\varepsilon \times \mathbb{R}_{\geq 0}$.

We use $\sigma = H(\sigma_t) = e'_1 e'_2 \ldots e'_{K'} \in E^*$ to denote the *logical observation sequence* associated with σ_t, where H filters out the timing information.

A timed sequence s_t and a time $t_f \geq \tau_{last}(s_t)$ define a *timed evolution* $(s_t, t_f) \in (E_\varepsilon \times \mathbb{R}_{\geq 0})^* \times \mathbb{R}_{\geq 0}$ of duration t_f. Such a timed evolution includes a *silent closure* of duration $t_f - \tau_{last}(s_t)$ during which no further jump occurs. The *observed timed evolution* corresponding to (s_t, t_f) is $(\sigma_t, t_f) = (P(s_t), t_f)$, which also includes a silent closure of duration $t_\varepsilon = t_f - \tau_{last}(\sigma_t)$, during which no further observable jump occurs. We denote by $\overline{L}_s(G)$ (resp., $\overline{L}_\sigma(G)$) the *set of timed evolutions* (resp., the *set of observed timed evolutions*) corresponding to runs which start from an initial state, i.e., a state with nonzero initial probability.

Example 1: Figure 1 shows a graphical representation of an LTPA with $X = \{x_1, x_2, x_3, x_4, x_5\}$, alphabet $E = \{a, b\}$, $\boldsymbol{\pi}_0 = [1\ 0\ 0\ 0\ 0]$ and transition relation $\Lambda = \{(x_1, a, \mu_a, x_3), (x_1, \varepsilon, \mu, x_2), (x_3, a, \mu_a, x_3), (x_2, a, \mu_a, x_4), (x_4, \varepsilon, \mu, x_5), (x_5, b, \mu, x_5)\}$. A possible run starting from the initial state x_1 is

$$x_1 \xrightarrow{\varepsilon,\ 0.5} x_2 \xrightarrow{a,\ 2} x_4 \xrightarrow{\varepsilon,\ 4} x_5$$

which determines timed sequence $s_t = (\varepsilon, 0.5)(a, 2)(\varepsilon, 4)$ of duration $\tau_{last}(s_t) = 4$ and length $|s_t| = 3$. The corresponding observation $\sigma_t = P(s_t) = (a, 2)$ has duration $\tau_{last}(\sigma_t) = 2$, length $|\sigma_t| = 1$ and logical sequence $H(\sigma_t) = a$. At current time $t_f = 6$, the previous run determines a timed evolution $(s_t, t_f) = ((\varepsilon, 0.5)(a, 2)(\varepsilon, 4), 6)$ with a silent closure of duration $6 - 4 = 2$, and an observed evolution $(\sigma_t, t_f) = ((a, 2), 6)$ with a silent closure of duration $6 - 2 = 4$. ◇

B. Eigenstructure of matrices

This section contains a series of elementary definitions of linear algebra. Given a real matrix Q of order n, we denote by spec(Q) the set of its eigenvalues and by abs$(Q) = \max\{Re(\zeta) \mid \zeta \in \text{spec}(Q)\}$ the maximum among the real parts of the eigenvalues of Q In addition, for any eigenvalue ζ of Q, we use $\nu(\zeta)$ to denote the *algebraic multiplicity* of ζ and $\nu_{geo}(\zeta)$ the *geometric multiplicity* of ζ, i.e., the number of blocks associated to ζ in the Jordan form of Q [2], [20].

Assume matrix Q has a Jordan form consisting of k blocks. Given a block $i \in \{1, \ldots, k\}$ we can associate with it an eigenvalue ζ_i, a left eigenvector $\boldsymbol{v}_i^{(0)}$ and a chain of generalized left eigenvectors of length h_i

$$\boldsymbol{v}_i^{(h_i-1)} \longrightarrow \boldsymbol{v}_i^{(h_i-2)} \longrightarrow \ldots \longrightarrow \boldsymbol{v}_i^{(0)}.$$

Then, a basis[2] $\mathcal{V}$ of $\mathbb{R}^n$ is defined, consisting of the generalized left eigenvectors of the k chains

$$\mathcal{V} = \bigcup_{i=1}^{k} \left\{ \boldsymbol{v}_i^{(h_i-1)}, \boldsymbol{v}_i^{(h_i-2)}, \ldots, \boldsymbol{v}_i^{(0)} \right\}. \quad (2)$$

[2]A complex conjugate pair of eigenvalues ζ, ζ' can be associated with a complex conjugate pair of eigenvectors $\boldsymbol{v}, \boldsymbol{v}' = \boldsymbol{u} \pm j\boldsymbol{w}$. In $\mathcal{V}$, complex vectors $\boldsymbol{v}, \boldsymbol{v}'$ can be replaced by real vectors $\boldsymbol{u}, \boldsymbol{w}$ [5].

Note that such a basis always exists and $\sum_{i=1}^{k} h_i = n$. Multiple chains may be associated to the same eigenvalue, i.e., $i, i' \in \{1, \dots k\}$ and $i \neq i'$ does not necessarily imply $\zeta_i \neq \zeta_{i'}$.

III. STATE ESTIMATION FOR LTPA

In this section we focus on the problem of state estimation for LTPA. In particular, in the first subsection we consider the problem of state estimation only looking at the logical sequence that is generated during the system evolution. The solution is based on the notion of state observer, which corresponds to the deterministic finite automaton (DFA) equivalent to the original non deterministic finite automaton (NFA). In the second subsection we show how to compute the conditional state probability vector relative to a given observed timed evolution.

A. Logical state estimation via observer

Given an LTPA G, let us first define the support of a probability vector.

Definition 2 (**Support**)*:* Given an LTPA G with set of states X and state probability vector $\boldsymbol{\pi}$, the *support* of $\boldsymbol{\pi}$ is the subset of states $\mathcal{X}(\boldsymbol{\pi}) = \{x_i \in X \mid \pi_i > 0\}$, having nonzero probability. ▲

An LTPA G can be associated with an underlying NFA A_G defined as follows.

Definition 3 (**Underlying NFA associated to** G)*:* Let $G = (X, E, \Lambda, \boldsymbol{\pi}_0)$ be an LTPA. The *underlying nondeterministic finite automaton associated to* G is the 4-tuple $A_G = (X, E, \Delta_G, X_0)$, where

- $\Delta_G = \{(x, e, x') \mid (x, e, \cdot, x') \in \Lambda\} \subseteq X \times E_\varepsilon \times X$ is the *transition relation*;
- $X_0 = \mathcal{X}(\boldsymbol{\pi}_0)$ is the *set of initial states*. ▲

In simple words, A_G is obtained from G by disregarding the firing rates in the transition relation as well as the initial probabilities associated with the initial states.

In the literature about discrete event systems, a fundamental notion for the state estimation of an NFA is that of an *observer*, i.e., the DFA equivalent to the NFA [3]. Here we point out that the observer of the underlying NFA A_G can be used for state estimation ignoring the timing/probabilistic aspects: we call this automaton the *logical observer* of G and denote it by O_G. Each state of the logical observer is a subset of states of A_G, hence of states of G. Given $A_G = (X, E, \Delta, X_0)$, a subset $X' \subset X$ and an event $e \in E$ we first denote:

- $D_\varepsilon(X') \subseteq X$: the set of states reachable in A_G from states in X' by executing zero or more ε-transitions;
- $D_e(X') \subseteq X$: the set of states reachable in A_G from states in X' by executing exactly one e-transition.

The logical observer is formally defined as follows.

Definition 4 (**Logical observer of** G)*:* The *logical observer* of an LTPA G with underlying NFA $A_G = (X, E, \Delta, X_0)$ is defined as a DFA $O_G = (X_L, E, \delta_L, x_{L,0})$ where:

- $X_L \subseteq 2^X$ is the set of observer states;
- E is the alphabet;
- δ_L is the transition function defined for all $x_L \in X_L$ and $e \in E$ by $\delta_L(x_L, e) = D_\varepsilon(D_e(x_L))$ if $D_\varepsilon(D_e(x_L)) \neq \emptyset$; otherwise $\delta_L(x_L, e)$ is undefined;
- $x_{L,0} = D_\varepsilon(X_0)$ is the observer initial state. ▲

The initial state of O_G is defined as the set of states reachable from an initial state of A_G by executing zero or more ε-transitions. Then, all other states can be iteratively computed. By searching the observer states that have cardinality equal to 1, i.e., they are of the form $x_{L,k} = \{x_i\}$, one can provide the conditions to estimate exactly the LTPA state, based only on the logical information $H(\sigma_t)$ of a given timed observation.

B. Probabilistic state estimation via probability vector

In an LTPA, as in a classical Markov chain [12], it may be possible to compute, for $t_f \geq 0$, the *a priori* probabilities $\pi_i(t_f)$ that the system is in state $x_i \in X$ at time t_f, given an initial probability vector $\boldsymbol{\pi}_0$. In the next, we do not report the dependence to $\boldsymbol{\pi}_0$ when no confusion exists.

Definition 5 (**A priori state probability vector**)*:* Given a state $x_i \in X$, $\pi_i(t_f)$ is the probability to be in state x_i at time t_f ignoring the observation of the timed sequence σ_t. Consequently, $\boldsymbol{\pi}(t_f)$ is defined as the unconditional probability vector. ▲

If we denote by $\mu(x_i, x_j)$ the sum of the rates of the transitions from state x_i to state x_j,

$$\mu(x_i, x_j) = \sum_{(x_i, e, \mu, x_j) \in \Lambda} \mu, \tag{3}$$

the vector $\boldsymbol{\pi}(t_f)$ can be computed as [9], [12]:

$$\boldsymbol{\pi}(t_f) = \boldsymbol{\pi}(0) \cdot exp(Q t_f) \tag{4}$$

where the *transition rate matrix* (also known as generator matrix) $Q = \{q_{i,j}\}$ has elements: $q_{i,j} = \mu(x_i, x_j)$ for $j \neq i$ and $q_{i,i} = -\sum_{j \neq i} q_{i,j}$ for all i.

For an LTPA, however, we can exploit the additional information deriving from the observed evolution to update *a posteriori* the state probability vector.

Definition 6 (**Conditional state probability vector**)*:* Given an observed timed evolution (σ_t, t_f) and a state $x_i \in X$, $\pi_i(\sigma_t, t_f)$ is the probability to be in state x_i at time t_f *conditioned by the observation of timed sequence* σ_t. Consequently, $\boldsymbol{\pi}(\sigma_t, t_f)$ is defined as the conditional probability vector.

The maximal conditional state probability at time t_f is denoted by $\rho(\sigma_t, t_f) = \max\{\pi_i(\sigma_t, t_f) \mid x_i \in X\}$. ▲

The conditional probability vector $\boldsymbol{\pi}(\sigma_t, t_f)$ can be formally computed in an iterative way by considering the extended ε-sub chain of G and the set of e-transition matrices, $e \in E$ as described in [9], [10]. Note that when no event is observable, i.e., $\overline{L}_{\sigma_t}(G) = \{(\lambda, t_f) \mid t_f \in \mathbb{R}_{\geq 0}\}$ then the *a posteriori* probability vector $\boldsymbol{\pi}(\sigma_t, t_f)$ coincides with the *a priori* probability vector $\boldsymbol{\pi}(t)$ solution of Eq. (4) (where the entries of matrix Q are given by (3) with $e = \varepsilon$).

We conclude this section discussing how the conditional state probability vector can be used for the purpose of state estimation. Given an LTPA G, after observing evolution

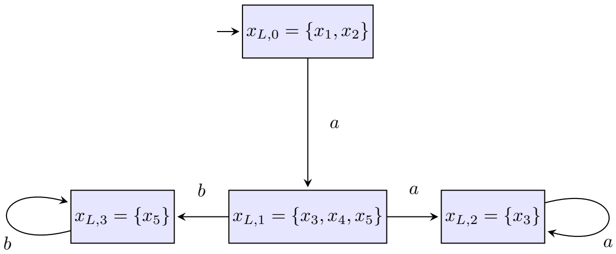

Fig. 2. The logical observer for the LTPA in Figure 1 for $\boldsymbol{\pi}_0 = [1\ 0\ 0\ 0\ 0]$.

(σ_t, t_f) one wants to estimate the *set of consistent states*, i.e., the set of states where G could be at time t_f. Given an observed evolution (σ_t, t_f), the set of states consistent with this observation is $\mathcal{X}(\boldsymbol{\pi}(\sigma_t, t_f))$, i.e., the support of the corresponding *a posteriori* probability vector. In addition, if the maximal state probability is $\rho(\sigma_t, t_f) = 1$, then necessarily there exists a state x_{i*} such that $\mathcal{X}(\boldsymbol{\pi}(\sigma_t, t_f)) = \{x_{i*}\}$ and the state can be correctly estimated at time t_f.

Example 2: Consider again the LTPA G in Figure 1 with initial distribution $\boldsymbol{\pi}_0 = [1\ 0\ 0\ 0\ 0]$. The logical observer is shown in Figure 2. Let $\mu_a = \mu = 1$. Let $\sigma_t = (a, 1)(a, 4)$ be a timed sequence of observations, and $t_f = 5$ be the final time instant of observation. The components $\pi_i(\sigma_t, t_f)$, $i = 1, 2, 3, 4, 5$ of the conditional probability vector vary with respect to time as shown in Figure 3. Finally, Figure 4 shows how the support of such probability vector changes with respect to time during the time intervals $(0, 1)$, $(1, 4)$ and $(4, 5]$. In particular, it shows how the support of the conditional probability vector in such time intervals is related to the states of the logical observer in Figure 2. Note that after the second observation of a the state is perfectly reconstructed; thus, the maximal state probability is equal to $\rho(\sigma_t, t) = 1$, $\forall t \in [4, 5]$. ⋄

IV. PROBABILISTIC VS. LOGICAL ESTIMATION

The relation between probabilistic and logical state estimation for LTPAs, which we have previously defined, is discussed in this section.

One can immediately verify that an LTPA G admits a timed observed evolution $(\sigma_t, t_f) \in \overline{L}_\sigma(G)$ with $\sigma_t = (e_1, \tau_1)(e_2, \tau_2)\ldots(e_K, \tau_K)$ if and only if its logical observer O_G admits an evolution[3]:

$$x_L = \delta_L^*(x_{L,0}, H(\sigma_t)) \in X_L,$$

where sequence $H(\sigma_t) = e_1 e_2 \ldots e_K \in E^*$ and x_L is some state in X_L. In the following, we discuss how the conditional probability vector $\boldsymbol{\pi}(\sigma_t, t_f)$ is related to such a state $x_L = \delta_L^*(x_{L,0}, H(\sigma_t))$, thus characterizing the evolution of the probabilistic state estimate.

For a given timed observed sequence σ_t, we will consider all possible timed evolutions (σ_t, t_f) for a finite final time $t_f \in [\tau_{last}(\sigma_t), \infty)$ or, equivalently, for an ε-closure $t_\varepsilon = t_f - \tau_{last}(\sigma_t) \in [0, \infty)$. The limit as $t_\varepsilon \to \infty$ will also be discussed.

[3]Here $\delta_L^* : X \times E^* \to X$ denotes the transitive and reflexive closure of transition function $\delta_L : X \times E \to X$.

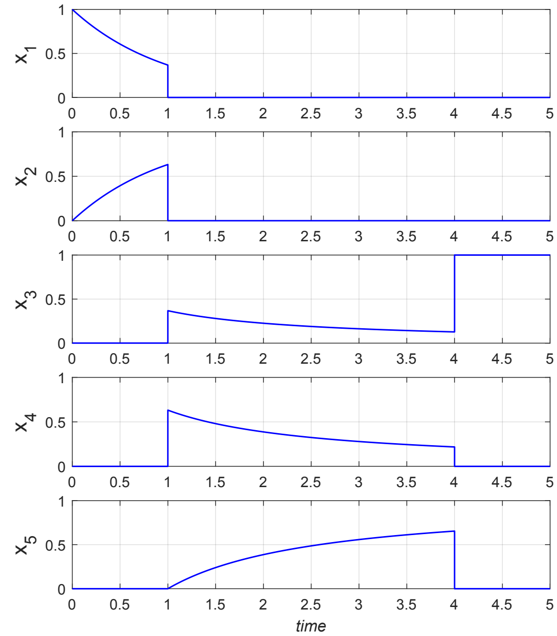

Fig. 3. Conditional probabilities relative to the LPTA in Figure 1, to the observation $\sigma_t = (a, 1)(a, 4)$ and to $t_f = 5$.

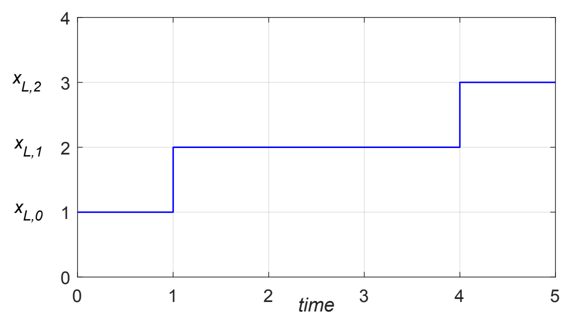

Fig. 4. The support of the conditional probabilities in Figure 3 as a function of the states of the logical observer in Figure 2 during the time intervals $(0, 1)$, $(1, 4)$ and $(4, 5]$.

A. Finite $t_\varepsilon \in [0, \infty)$

The following lemmata describe how the support of the conditional probability vector is related to the observer structure when no event has occurred yet (Lemma 1) and when a new event occurs (Lemma 2).

Lemma 1: Given an LTPA $G = (X, E, \Lambda, \boldsymbol{\pi}_0)$ with logical observer $O_G = (X_L, E, \delta_L, x_{L,0})$ and an observed timed evolution $(\lambda, t_f) \in \overline{L}_\sigma(G)$, it holds:

(i) $t_f = 0 \Longrightarrow \mathcal{X}(\boldsymbol{\pi}(\lambda, 0)) = \mathcal{X}(\boldsymbol{\pi}_0) \subseteq x_{L,0}$;
(ii) $t_f > 0 \Longrightarrow \mathcal{X}(\boldsymbol{\pi}(\lambda, t_f)) = x_{L,0}$.

Proof. If $t_f = 0$ then $\boldsymbol{\pi}(\lambda, 0) = \boldsymbol{\pi}_0$ and $\mathcal{X}(\boldsymbol{\pi}(\lambda, 0)) = \mathcal{X}(\boldsymbol{\pi}_0) = X_0 \subseteq D_\varepsilon(X_0) = x_{L,0}$ according to the definition of logical observer. If $t_f > 0$ then in the interval $[0, t_f]$ any arbitrary sequence of unobservable jumps may have occurred (due to the exponential distribution of the delays). Thus the states of G with a nonzero probability are exactly the states of the underling NFA A_G associated with G that are reachable from a state in X_0 with zero or more ε-transitions. Thus, $\mathcal{X}(\boldsymbol{\pi}(\lambda, t_f)) = D_\varepsilon(X_0) = x_{L,0}$. □

Lemma 1 claims that for an empty observation (λ, t_f) the set of states with nonzero probabilities is a subset of the initial state of the logical observer $x_{L,0}$ at $t_f = 0$ and is equal to $x_{L,0}$ for $t_f > 0$.

Lemma 2: Given an LTPA $G = (X, E, \Lambda, \boldsymbol{\pi}_0)$ with logical observer $O_G = (X_L, E, \delta_L, x_{L,0})$, consider an observed timed evolution $(\sigma_t, t_f) \in \overline{L}_\sigma(G)$ with $\sigma_t = \sigma_t'(e, \tau)$. If one defines $x_L' = \delta_L^*(x_{L,0}, H(\sigma_t'))$ and $x_L = \delta_L^*(x_{L,0}, H(\sigma_t))$ it holds:

(i) $t_f = \tau_{last}(\sigma_t) \Longrightarrow \mathcal{X}(\boldsymbol{\pi}(\sigma_t, t_f) = D_e(x_L') \subseteq x_L$;
(ii) $t_f > \tau_{last}(\sigma_t) \Longrightarrow \mathcal{X}(\boldsymbol{\pi}(\sigma_t, t_f)) = x_L$.

Proof. We first consider the particular case $\sigma_t' = \lambda$. The occurrence time of event e is $\tau > \tau_{last}(\lambda)$ $(= 0)$ according to Eq. (1). This means that just before event e occurs, the probability vector has support $\mathcal{X}(\boldsymbol{\pi}(\sigma_t', \tau^-)) = x_L'$, according to Lemma 1.$(ii)$. Now at time τ a single transition labeled e occurs, hence: $\mathcal{X}(\boldsymbol{\pi}(\sigma, \tau)) = D_e(x_L') \subseteq D_\varepsilon(D_e(x_L')) = x_L$ according to the definition of logical observer, thus proving (i). When $t_f > \tau$ in the interval $(\tau, t_f]$ any arbitrary sequence of unobservable jumps may have occurred and, as in the proof of the previous lemma, we can claim that $\mathcal{X}(\boldsymbol{\pi}(\sigma_t, t_f)) = D_\varepsilon(D_e(x_L')) = x_L$, thus proving also (ii). Iterating on the length of σ_t', we can prove Lemma 2 for observation sequences of arbitrary length. □

Lemma 2 claims that each time a new event e is observed after a previous sequence σ_t', the set of states with nonzero probabilities is the set of states that can be reached by the occurrence of an e-transition from states in the observer state consistent with the logical sequence $H(\sigma_t')$. Immediately after, however, as time progresses without any new event observation, the set of states with nonzero probabilities coincides with the observer state consistent with the observed logical sequence $H(\sigma_t) = H(\sigma_t' e)$.

Example 3: Consider again the LTPA G in Figure 1 with initial distribution $\boldsymbol{\pi}_0 = [1\ 0\ 0\ 0\ 0]$ and let $\sigma_t = (a, 1)$ be a timed sequence of observations. At time $t = 1$ the state probability vector switches to $\boldsymbol{\pi}(\sigma_t, 1) = [0\ 0\ \pi_3\ \pi_4\ 0]$ with $\pi_3, \pi_4 > 0$ and $\pi_3 + \pi_4 = 1$ and whose support satisfies $\mathcal{X}(\boldsymbol{\pi}(\sigma_t, 1)) = \{x_3\ x_4\}$ and is included in the observer state $x_{L,1} = \{x_3, x_4, x_5\}$. Then, an arbitrarily small amount of time dt later, as shown in Figure 3, the probability vector $\boldsymbol{\pi}(\sigma_t, 1 + dt)$ has support $\mathcal{X}(\boldsymbol{\pi}(\sigma_t, 1 + dt)) = \{x_3, x_4, x_5\}$ that coincides with the observer state $x_{L,1}$. ⋄

B. Limit as $t_\varepsilon \to \infty$

Let x_L be a state of the logical observer O_G and let $\boldsymbol{\pi}$ be an arbitrary probability vector of G such that $\mathcal{X}(\boldsymbol{\pi}) = x_L$. Let $\boldsymbol{\pi}'$ be the vector of dimension $|x_L|$ obtained by projecting $\boldsymbol{\pi}$ on its support x_L. We define $M_{x_L} \in \{0, 1\}^{|X| \times |x_L|}$ to be the matrix of binary entries such that $\boldsymbol{\pi}' = \boldsymbol{\pi} \times M_{x_L}$. In detail, we first order the states in x_L according to the enumeration used for $X = \{x_1, x_2, \ldots, x_n\}$. Then, $m_{i,j}$, i.e., the element of M_{x_L} at row i and column j, equals 1 if the jth state in x_L is x_i, and equals 0 otherwise. Observe that this also implies that $\boldsymbol{\pi} = \boldsymbol{\pi}' \times (M_{x_L})^T$.

Definition 7 (**x_L-equivalent LTPA**)*:* Given an LTPA $G = (X, E, \Lambda, \boldsymbol{\pi}_0)$ and an observed timed sequence σ_t, let $x_L = \delta_L^*(x_{L,0}, H(\sigma_t))$ be the state of the logical observer O_G consistent with σ_t. The *x_L-equivalent* LTPA is defined by $G' = (x_L, E, \Lambda', \boldsymbol{\pi}_0')$ where:

- $\Lambda' = \{(x, \varepsilon, \mu, \bar{x}) \in \Lambda \mid x, \bar{x} \in x_L\} \cup \{(x, e, \mu, x) \mid x \in x_L, e \in E, (x, e, \mu, \bar{x}) \in \Lambda\}$;
- $\boldsymbol{\pi}_0' = \boldsymbol{\pi}(\sigma_t, \tau_{last}(\sigma_t)) \times M_{x_L}$. ▲

In other words, the structure of G' is obtained from G by i) changing the arrival state of any observable transition emanating from a state $x \in x_L$ so that it is self-looped on x; ii) removing all states in $X \setminus \{x_L\}$ and their input and output transitions. The initial probability vector of G' is the projection on x_L of the vector $\boldsymbol{\pi}(\sigma_t, \tau_{last}(\sigma_t))$ of G.

To compute $\boldsymbol{\pi}'(\lambda, t_f)$, we adapt here the method initially proposed in [9], [10]. For this purpose, we define the x_L-equivalent LTPA generator as the $\mid x_L \mid \times \mid x_L \mid$ real matrix $Q_{x_L} = \{q_{i,j}\}$ where

- each off-diagonal element $q_{i,j}$ is equal to the sum of the rates of ε-transitions in G' from x_i to x_j, or is equal to 0 if no such a transition exists:
$$q_{i,j} = \sum_{(x_i, \varepsilon, \mu, x_j) \in \Lambda'} \mu, \qquad i, j \in \{1, \ldots, \mid x_L \mid\},\ i \neq j;$$
- each diagonal element is equal to the negative of the sum of the rates of all transitions in G' emanating from x_i, or is equal to 0 if no such a transition exists
$$q_{i,i} = -\sum_{(x_i, e, \mu, x) \in \Lambda'} \mu, \qquad i \in \{1, \ldots, \mid x_L \mid\}.$$

Lemma 3: Consider an x_L-equivalent LTPA G' with initial probability vector $\boldsymbol{\pi}_0'$ and generator Q_{x_L}. Let $\mathcal{V}$ be a basis of left generalized eigenvectors of Q_{x_L} composed by k chains as detailed in Eq. (2). The state probability vector at time t_f assuming no event is observed in $[0, t_f]$ is

$$\boldsymbol{\pi}'(\lambda, t_f) = \frac{\displaystyle\sum_{i=1}^{k} \sum_{j=0}^{h_i - 1} \beta_{i,j} \left(\sum_{p=0}^{j} \frac{(t_f)^p}{p!} exp(\zeta_i t) \boldsymbol{v}_i^{(j-p)} \right)}{\left\| \displaystyle\sum_{i=1}^{k} \sum_{j=0}^{h_i - 1} \beta_{i,j} \left(\sum_{p=0}^{j} \frac{(t_f)^p}{p!} exp(\zeta_i t) \boldsymbol{v}_i^{(j-p)} \right) \right\|_1}, \tag{5}$$

where parameters $\beta_{i,j} \in \mathbb{R}, i = 1, ..., k, j = 0, ..., h_i - 1$ are the components of the initial probability vector $\boldsymbol{\pi}_0'$ expressed in basis $\mathcal{V}$:

$$\boldsymbol{\pi}_0' = \sum_{i=1}^{k} \sum_{j=0}^{h_i - 1} \beta_{i,j} \boldsymbol{v}_i^{(j)}. \tag{6}$$

Proof. The state probability vector $\boldsymbol{\pi}'(\lambda, t_f)$ can be computed thanks to the x_L-*equivalent* LTPA G' [9], [10]

$$\boldsymbol{\pi}'(\lambda, t_f) = \frac{\boldsymbol{\pi}'_0 \; exp(Q_{x_L} t_f)}{||\boldsymbol{\pi}'_0 \; exp(Q_{x_L} t_f)||_1}.$$

Using the notations introduced in Section II.B, for any generalized left eigenvector $v_i^{(j)}$, $j = 0, \ldots, h_i - 1$, of Q_{x_L}, it holds:

$$\boldsymbol{v}_i^{(j)} exp(Q_{x_L} t_f) = \sum_{p=0}^{j} \frac{(t_f)^p}{p!} exp(\zeta_i t) \boldsymbol{v}_i^{(j-p)} \qquad (7)$$

i.e., any evolution that starts from a generalized eigenvector of the chain of rank j will contain (and only contain) components along all the generalized eigenvectors of the chain of rank j or lower, i.e., $\boldsymbol{v}_i^{(j)}, \boldsymbol{v}_i^{(j-1)}, \ldots, \boldsymbol{v}_i^{(0)}$. Then, replacing in $\boldsymbol{\pi}'_0 \; exp(Q_{x_L} t_f)$ the vector $\boldsymbol{\pi}'_0$ by Eq. (6) and using in addition Eq. (7), it holds,

$$\begin{aligned} \boldsymbol{\pi}'_0 exp(Q_{x_L} t_f) &= \textstyle\sum_{i=1}^{k} \sum_{j=0}^{h_i-1} \beta_{i,j} \boldsymbol{v}_i^{(j)} exp(Q_{x_L} t_f) \\ &= \textstyle\sum_{i=1}^{k} \sum_{j=0}^{h_i-1} \beta_{i,j} \left(\sum_{p=0}^{j} \frac{(t_f)^p}{p!} exp(\zeta_i t_f) \boldsymbol{v}_i^{(j-p)} \right). \end{aligned}$$

Equation (5) results consequently. □

In addition, matrix Q_{x_L} has interesting properties that are summed up in Lemma 4.

Lemma 4: Matrix Q_{x_L} satisfies the following properties:

(a) Q_{x_L} has a real and non-positive eigenvalue ζ_F = abs(Q_{x_L}), called *Frobenius eigenvalue*.

(b) For any other eigenvalue $\zeta \neq \zeta_F$ it holds that $Re(\zeta) < \zeta_F$. Note however that ζ_F may have multiplicity greater than one.

(c) The left and right eigenvectors associated to ζ_F can be chosen non-negative.

(d) If Q_{x_L} is *irreducible* then ζ_F is a simple eigenvalue and these eigenvectors can be chosen positive: they are called *dominant eigenvectors*.

Proof. By construction, the generator Q_{x_L} of the x_L-equivalent LTPA is a diagonally dominant Metzler[4] matrix with non-positive diagonal elements. There exist a non-negative matrix P and a real $\alpha \in \mathbb{R}$ such that $Q_{x_L} = P + \alpha I$. This implies that the eigenstructures of Q_{x_L} and P are closely related: $\boldsymbol{v}$ is an eigenvector of Q_{x_L} associated to eigenvalue ζ if and only if $\boldsymbol{v}$ is an eigenvector of P associated to eigenvalue $\zeta - \alpha$. Based on this observation, it is not difficult to show that properties (a), (b), (c) and (d) follow from Perron-Frobenius theorem [2], [20]. □

To determine the final probability vector as $t_f \to \infty$ we need to identify the dominant terms in Eq. (5), which may depend on the initial probability vector.

Let us introduce some notations.

Definition 8: Consider an x_L-equivalent LTPA G' whose initial probability vector $\boldsymbol{\pi}'_0$ is expressed as in Eq. (5). We define the set

$$\begin{aligned} B(\boldsymbol{\pi}'_0) = \{(i,j) \in \mathbb{N}^2 \mid \; \beta_{i,j} \neq 0 \; \wedge \nexists (i',j') \in \mathbb{N}^2 \\ \text{with } \beta_{i',j'} \neq 0 \text{ and } \; Re(\zeta_{i'}) > Re(\zeta_i)\}, \end{aligned}$$

containing the indices of non-null coefficients β's in Eq. (6) associated with the dominant abscissa eigenvalues. We also define

$$j_{sup} = \max \; \{j \in \mathbb{N} \;\mid (\exists i \in \mathbb{N}) \; (i,j) \in B(\boldsymbol{\pi}'_0)\},$$

the rank of generalized eigenvectors associated with a dominant term in Eq. (5) and

$$I = \{i \in \mathbb{N} \mid (i, j_{sup}) \in B(\boldsymbol{\pi}'_0)\},$$

the set of indices of chains associated with a dominant term in Eq. (5). ▲

Note that in the previously defined set I, for all $i \in I$, it holds that eigenvalues ζ_i have the same real part.

The following propositions provide sufficient conditions for the existence of a final probability vector as $t_f \to \infty$.

Proposition 1: Assume there exists a coefficient $\beta_{i,j} > 0$ with $\zeta_i = \zeta_F$ in Eq. (6), i.e., the initial probability vector has a non-null component along one of the generalized eigenvectors associated to the Frobenius eigenvector. Then for all $i \in I$ it holds that $\zeta_i = \zeta_F$ and

$$\lim_{t_f \to \infty} \boldsymbol{\pi'}(\lambda, t_f) \;=\; \frac{\displaystyle\sum_{i \in I} \beta_{i,j_{sup}} \; \boldsymbol{v}_i^{(0)}}{\left\| \displaystyle\sum_{i \in I} \beta_{i,j_{sup}} \; \boldsymbol{v}_i^{(0)} \right\|_1}. \qquad (8)$$

where I and j_{sup} are given in Definition 8.

Proof: The Frobenius eigenvalue is the unique abscissa dominant eigenvalue and since by assumption there exists $i^* \in I$ with $\zeta_{i^*} = \zeta_F$, it holds that $\zeta_i = \zeta_F$ for all $i \in I$. Being ζ_F real, there are no dominant complex eigenvalues in (5), hence its limit as $t_f \to \infty$ exists and is given by (8). □

Proposition 2: Assume there exists a coefficient $\beta_{i,j} > 0$ with $\zeta_i = \zeta_F$ in Eq. (6). Assume eigenvector ζ_F has geometric multiplicity $\nu_{geo} = 1$. Then it admits a unique[5] left eigenvector $\boldsymbol{v}_F$ and

$$\lim_{t_f \to \infty} \boldsymbol{\pi'}(\lambda, t_f) \;=\; \frac{\boldsymbol{v}_F}{||\boldsymbol{v}_F||_1}.$$

Proof: Follows from Eq. (8), because in this case $|I| = 1$. □

Example 4: Consider again the LTPA G in Figure 1, its logical observer in Figure 2 and the $x_{L,1}$-equivalent LTPA detailed in Figure 5. The generator matrix is

$$Q_{x_L} = \begin{bmatrix} -\mu_a & 0 & 0 \\ 0 & -\mu & \mu \\ 0 & 0 & -\mu \end{bmatrix}$$

with eigenvalues $\zeta_1 = -\mu_a$ and $\zeta_2 = -\mu$. Eigenvalue ζ_1 has eigenvector $\boldsymbol{v}_1^{(0)} = [\,1\,0\,0\,]$. A chain of length 2 is associated with eigenvalue ζ_2, with eigenvector $\boldsymbol{v}_2^{(0)} = [\,0\,0\,1\,]$ and generalized eigenvector $\boldsymbol{v}_2^{(1)} = [\,0\,1\,0\,]$.

Observer state $x_{L,1}$ is only reachable from observer

[4]A matrix is *Metzler* if all its non-diagonal elements are non-negative.

[5]Modulo a multiplicative constant.

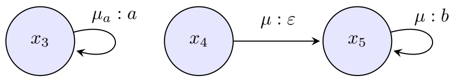

Fig. 5. $x_{L,1}$-equivalent LTPA for Example 1.

state $x_{L,0}$ upon the occurrence of event a. Thus the $x_{L,1}$-equivalent LTPA has initial state $\boldsymbol{\pi}'_0 = [\ \pi_{3,0}\ \pi_{4,0}\ \pi_{5,0}\ ]$ with $\pi_{3,0}, \pi_{4,0} > 0$ and $\pi_{5,0} = 0$, since $D_a(x_{L,0}) = \{x_3, x_4\}$. This implies that

$$\boldsymbol{\pi}'_0 = \beta_{1,0}\boldsymbol{v}_1^{(0)} + \beta_{2,1}\boldsymbol{v}_2^{(1)} \tag{9}$$

with $\beta_{1,0}, \beta_{2,1} > 0$. We need to discuss three possible cases.

- Case 1: $\mu_a < \mu$. This means $\zeta_F = \zeta_1$, and this eigenvalue has geometric multiplicity $\nu_{geo} = 1$ (only one chain is associated with it). By Proposition 2, it follows that
$$\lim_{t_f \to \infty} \boldsymbol{\pi'}(\lambda, t_f) \;=\; \frac{\boldsymbol{v}_1^{(0)}}{\left|\left|\boldsymbol{v}_1^{(0)}\right|\right|_1} \;=\; \boldsymbol{v}_1^{(0)}.$$
- Case 2: $\mu_a > \mu$. This means $\zeta_F = \zeta_2$, and again the Frobenius eigenvalue has geometric multiplicity $\nu_{geo} = 1$ (only one chain). By Proposition 2, it follows that
$$\lim_{t_f \to \infty} \boldsymbol{\pi'}(\lambda, t_f) \;=\; \frac{\boldsymbol{v}_2^{(0)}}{\left|\left|\boldsymbol{v}_2^{(0)}\right|\right|_1} \;=\; \boldsymbol{v}_2^{(0)}.$$
- Case 3: $\mu_a = \mu$. This means $\zeta_F = \zeta_1 = \zeta_2$ and this eigenvalue has geometric multiplicity $\nu_{geo} = 2$, thus two chains are associated with it: $\{\boldsymbol{v}_1^{(0)}; \boldsymbol{v}_2^{(1)} \longrightarrow \boldsymbol{v}_2^{(0)}\}$. From Eq. (9), we get $B(\boldsymbol{\pi}'_0) = \{(1,0), (2,1)\}$, $I = \{2\}$ and $j_{sup} = 1$. This means that the unique dominant mode is $t \cdot exp(-\mu t)$. By Proposition 1, it follows that
$$\lim_{t_f \to \infty} \boldsymbol{\pi'}(\lambda, t_f) \;=\; \frac{\beta_{2,1}\ \boldsymbol{v}_2^{(0)}}{\left|\left|\beta_{2,1}\ \boldsymbol{v}_2^{(0)}\right|\right|_1} \;=\; \boldsymbol{v}_2^{(0)}.$$

These results are consistent with the state probability evolution shown in Fig. 3, corresponding to rates $\mu_a = \mu = 1$. After $(a, 1)$ has been observed and before the occurrence of observation $(a, 4)$, the logical observer is in state $x_{L,1} = \{x_3, x_4, x_5\}$. Hence during the interval $t \in [1, 4)$, we expect that the probabilities of all states $x \notin x_{L,1}$ be null, while according to Eq. (9) it holds that $\pi_5(1) = 0$. Fig. 3 also shows, as discussed in Case 3 above, that when the silent closure increases, the probability vector $\boldsymbol{\pi'}(\lambda, t) = [\ \pi_3(t)\ \pi_4(t)\ \pi_5(t)\ ]$ tends to $\boldsymbol{v}_2^{(0)} = [\ 0\ 0\ 1\ ]$. $\diamond$

In this example, the computation of the final probability vector does not depend on the initial probability vector $\boldsymbol{\pi}'_0$ and is fully determined by the eigenstructure of the x_L-equivalent LTPA.

V. CONCLUSIONS AND FUTURE WORK

This paper has discussed logical and probabilistic aspects of state estimation for a class of labeled timed probabilistic automata. In particular, some results have been proposed to characterize the evolution of the conditional state probability in two situations: immediately after an observation or when no additional observation is collected in the long run.

In our further work, we will improve such conditions and introduce timed detectability notions for timed probabilistic automata. In particular, we are interested in conditions which imply that the state probability vector reaches a non-ambiguous stationary distribution at some observations or tends to such a distribution when no observation occurs during a sufficiently long duration.

REFERENCES

[1] M. V. S. Alves and J. C. Basilio, State estimation and detectability of networked discrete event systems with multi-channel communication networks, Proc. American Control Conference, pp. 5602–5607, Philadelphia, PA, USA, 2019.

[2] L. Benvenuti and L. Farina, Eigenvalue regions for positive systems, Systems & Control Letters, 51(3–4), pp. 325–330, 2004.

[3] C. G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*, Springer, 2008.

[4] P. A. Gagniuc, *Markov Chains: From Theory to Implementation and Experimentation*, USA, NJ: John Wiley and Sons, 2017.

[5] A. Giua, and C. Seatzu, *Analysis of Dynamical Systems*, Springer Italia, 2009 (in Italian).

[6] C. N. Hadjicostis and C. Seatzu, K-detectability in discrete event systems, Proc. 55th IEEE Conference on Decision and Control, pp. 420–425, Las Vegas, NV, 2016.

[7] C. Keroglou and C. N. Hadjicostis, Detectability in stochastic discrete event systems, Systems & Control Letters, 84, pp. 21–26, 2015.

[8] C. Keroglou and C. N. Hadjicostis, Verification of detectability in probabilistic finite automata, Automatica, 86, pp. 192–198, 2017.

[9] D. Lefebvre, C. Seatzu, C. N. Hadjicostis, and A. Giua, Probabilistic state estimation for labeled continuous time Markov models with applications to attack detection, Discrete Event Dynamic Systems, 32(1), pp. 65–88, 2022.

[10] D. Lefebvre, C. Seatzu, C. N. Hadjicostis, and A. Giua, Corrections to "Probabilistic state estimation for labeled continuous time Markov models with applications to attack detection", Discrete Event Dynamic Systems, 32(3), pp. 539–544, 2022.

[11] S. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer, 1993.

[12] J. R. Norris, *Markov Chains*, Cambridge Press, 1997.

[13] L. R. Rabiner, A tutorial on Hidden Markov Models and selected applications in speech recognition, Proceedings of the IEEE, 77(2), pp. 257–286,1989.

[14] S. Shu, F. Lin, and H. Ying, Detectability of discrete event systems. IEEE Transactions on Automatic Control, 52(12), pp. 2356–2359, 2007.

[15] S. Shu, F. Lin, H. Ying, and X. Chen, State estimation and detectability of probabilistic discrete event systems, Automatica, 44, pp. 3054–3060, 2008.

[16] S. Shu and F. Lin. Generalized detectability for discrete event systems, Systems & Control Letters, 60(5), pp. 310–317, 2011.

[17] S. Shu and F. Lin, I-detectability of discrete-event systems, IEEE Transactions on Automation Science and Engineering, 10(1), pp. 187–196, 2013.

[18] S. Shu, Delayed detectability of discrete event systems, IEEE Transactions on Automatic Control, 58(4), pp. 862–875, 2013.

[19] S. Shu, Z. Huang, and F. Lin, Online sensor activation for detectability of discrete event systems, IEEE Transactions on Automation Science and Engineering, 10(2), pp. 457–461, 2013.

[20] M. E. Valcher and P. Santesso, Asymptotic exponential cones of Metzler matrices and their use in the solution of an algebraic problem, Linear Algebra and its Applications, 432(4), pp. 980–1006, 2010.

[21] D. Wang, X. Wang, and Z. Li, State-based fault diagnosis of discrete-event systems with partially observable outputs, Information Sciences, 529, pp. 87–100, 2020.

[22] K. Zhang and A. Giua, K-delayed strong detectability of discrete-event systems, Proc. 58th IEEE Conference on Decision and Control, pp. 7647–7652, Nice, France, 2019.

[23] L. Zhou, S. Shu, and F. Lin, Detectability of discrete-event systems under nondeterministic observations, IEEE Transactions on Automation Science and Engineering, 18(3), pp. 1315–1327, 2021.