Abstract— Given a labeled Petri net with silent (unobservable) transitions, we are interested in performing initial marking estimation in a probabilistic setting. We assume a known finite set of initial markings, each with some a priori probability, and our goal is to obtain the conditional probabilities of initial markings of the Petri net, conditioned on an observed sequence of labels. Under a Markovian assumption on the probabilistic model, we develop a recursive algorithm that allows us to efficiently determine the conditional probabilities for each possible initial marking (conditioned on the sequence of observations seen so far). We illustrate the proposed methodology via an example and discuss potential applications in the context of initial state opacity for security applications.

I. INTRODUCTION

Petri nets (PNs) have emerged as one of the main formalisms for the modeling, analysis, and control of discrete event systems. Several interesting problems have been investigated, including current and initial state (marking) estimation in the presence of uncertainties in the initial state and in the system evolution. For instance in [1] the initial marking was assumed either completely unknown or known to belong to a given convex set, but the net evolution was perfectly known (i.e., all transitions were assumed to be observable). In other cases, the problem of current marking estimation has been studied under the assumption that the initial marking is perfectly known, but the net evolution can only be partially observed [2]. This led to the study of labeled PNs where only the labels associated with transitions may be observed. This introduces two different forms of nondeterminism: transitions labeled with the empty string \( \varepsilon \) are unobservable (silent), while transitions that share the same label with other transitions that may be simultaneously enabled are observable but indistinguishable.

M. P. Cabasino and C. Seatzu are with the Department of Electrical and Electronic Engineering at the University of Cagliari, Italy. E-mail: {cabasino, seatzu}@diee.unica.it

C. N. Hadjicostis is with the Department of Electrical and Computer Engineering at the University of Cyprus, Nicosia, Cyprus, and also with the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. E-mail: chadjic@ucy.ac.cy


Related problems were studied in [3] where it was shown that in a labeled PN without silent transitions, the set of possible markings following a sequence of \( k \) observations, is bounded by a polynomial function of \( k \); thus, one can potentially explicitly enumerate the set of possible markings. Assuming an acyclic unobservable subnet, the approach in [3] was extended to labeled PNs with unobservable transitions (by establishing that the number of possible markings following a sequence of \( k \) observations remains bounded by a polynomial function of \( k \)).

While Bayesian networks and safety/reliability analysis (see, for example, the survey paper in [4]) rely heavily on the use of probabilistic models, the majority of works on PN marking estimation from the discrete event systems community does not consider how likely or unlikely a certain PN state (current or initial, depending on the objective) may be. The work in this paper considers PN models where information about the likelihood (probability of occurrence) of the different transitions that can occur in the system is available, and can thus be taken into account to form a more informative estimate. More specifically, the proposed state estimation approach obtains the conditional probability associated with each initial state in the set of possible initial states, conditioned on a given sequence of observations. As we explain in more detail later in the paper, this problem formulation is directly related to the problem of initial state opacity, which has immediate implications to various types of privacy and security requirements in numerous applications.

The paper proposes an algorithm that works under a very general probabilistic setting. We assume that we are given: (i) a labeled Petri, possibly with silent (unobservable) transitions, (ii) a known set of possible initial markings, each with some a priori probability, and (iii) a description of the a priori probabilities with which different transitions occur from different markings. Our goal is to efficiently obtain the conditional probabilities of initial markings, conditioned on an observed sequence of labels. The end result is a recursive algorithm that resembles the forward algorithm in hidden Markov models (HMM’s) [5], with the main differences being the facts that (i) the focus is on initial states estimation, (ii) the algorithm can handle unobservable transitions (as also done in [6]), and (iii) the set (and number) of states of the HMM that are of interest at each observation is changing at each time step (unlike [6]). The proposed approach also extends our work in [7], [8] where, under similar assumptions, we obtained the set of possible current markings along with their a posteriori probabilities (given the observed evolution). More details about the connection to HMM’s can be found in [8].
The information provided by the proposed algorithm can be used in conjunction with a variety of supervisory control [9]–[11] and fault diagnosis or estimation algorithms [12], in order to relax stringent constraints imposed by existing methodologies that typically only provide binary information regarding the possibility or impossibility of current/initial states (but not their probabilities). As mentioned earlier (and discussed in more detail later), the probabilistic approach for initial states estimation allows us, for example, to assess initial state opacity guarantees.

II. PRELIMINARIES

Definition 1: A Petri net structure \( N \) is a 4-tuple \( N = (P,T,F,W) \) where \( P = \{p_1,p_2,\ldots,p_n\} \) is a finite set of \( n \) places; \( T = \{t_1,t_2,\ldots,t_m\} \) is a finite set of \( m \) transitions; \( F \subseteq (P \times T) \cup (T \times P) \) is a set of arcs; \( W : F \to \mathbb{N} \) is a weight function where \( \mathbb{N} \) is the set of natural numbers; \( P \cap T = \emptyset \) and \( P \cup T \neq \emptyset \).

Pictorially, places are represented by circles, transitions by bars, and tokens by black dots, as shown on the left of Fig. 1. The set of all input (or output) places of a transition \( t \in T \) is defined as \( \ast t = \{p \in P \mid (p,t) \in F\} \) (or \( \ast t = \{p \in P \mid (p,t) \in F\} \)). The \( n \times m \) matrix \( D \) satisfying \( D(i,j) = -W(p_i,t_j) + W(t_j,p_i) \) if \( W(p_i,t_j) \) or \( W(t_j,p_i) \) is not defined for a specific place \( p_i \) and transition \( t_j \), it is taken to be 0) is called the incidence matrix of Petri net \( N \).

For convenience, we will sometimes abuse notation and write \( D(p_i,t_j) \) (instead of \( D(i,j) \)) to denote the \( (i,j) \)th entry of matrix \( D \).

A marking is a function \( M : P \to \mathbb{N} \cup \{0\} \) that assigns to each place a nonnegative integer number of tokens. We use \( M(p) \) to denote the number of tokens in place \( p \). A Petri net \( G = (N,M_0) \) is a net structure \( N \) with an initial marking \( M_0 \). Given a PN, a transition \( t \) is enabled at marking \( M \) if \( \forall p \in \ast t, \ M(p) \geq W(p,t) \); this is denoted by \( M(t) \). An enabled transition \( t \) may fire, and its firing removes \( W(p,t) \) tokens from each input place \( p \), \( p \in \ast t \), and adds \( W(t,p) \) tokens to each output place \( p' \), \( p' \in \ast t \), resulting in a marking \( M' \); this is denoted by \( M(t)M' \) and the marking \( M' \) also satisfies \( M' = M + D(t) \), where \( D(t) \) denotes the column of the incidence matrix of the \( N \) corresponding to transition \( t \).

A \( k \)-length firing sequence \( S \) from marking \( M_0 \) is a sequence of transitions \( S = t_{s_1}t_{s_2}\cdots t_{s_k}, \ t_{s_j} \in T, \) such that \( M_0[t_{s_1}M_1[t_{s_2}M_2\cdots t_{s_k}]M; \) this is denoted by \( M_0[S]M \) (note that in this paper we assume that only one transition can fire at each time instant). The final marking \( M \) can also be written as \( M = M_0 + D\delta \), where \( \delta \) is the incidence matrix of \( N \), and \( i \) \( \delta \) is the \( m \times 1 \) firing vector of \( S \) with its \( i \)th entry representing the number of times transition \( t_i \) appears in \( S \). [More specifically, given \( S = t_{s_1}t_{s_2}\cdots t_{s_k}, \ t_{s_j} \in T \), we have \( \delta(i) = |T_i| \), where \( T_i = \{j \mid t_{s_j} = t_i\} \) and \( |T| \) denotes the cardinality of the set \( T \).]

The set of all finite-length firing sequences, including the empty or null sequence \( S = \emptyset \) is denoted by \( T^* \). Given a sequence \( S \in T^* \), we say \( S' \in T^* \) is a prefix of \( S \) if \( S = S'S'' \) for some \( S'' \in T^* \) (i.e., \( S \) is the concatenation of the sequence of transitions in \( S'' \)) followed by the sequence of transitions \( S'' \). The set of all prefixes of a transition firing sequence \( S \) is \( S = \{S' \mid \exists S'' \in T^* \text{ such that } S = S'S''\} \). Note that \( S = \emptyset \) by definition (since \( S = \emptyset \in T^* \)).

A marking \( M \) is reachable from \( M_0 \) if there exists a firing sequence \( S \) such that \( M_0[S]M \). Given a PN \( G = (N,M_0) \), the set of all markings reachable from \( M_0 \) is called the reachability set and is denoted by \( R(N,M_0) \). If \( \forall p \in P \) and \( \forall M \in R(N,M_0), \ M(p) \leq K \) for some finite positive integer \( K \), then we say the PN is bounded.

Definition 2: A labeled PN is a 3-tuple \( G_L = (N,\Sigma,L) \), where \( N \) is a PN structure, \( \Sigma = \{\epsilon_1,\epsilon_2,\ldots,\epsilon_{|\Sigma|}\} \) is the set of labels (also called alphabet), and \( L : T \to \Sigma \cup \{\varepsilon\} \) is a labeling function that assigns a label (which can be the null label \( \varepsilon \)) to each transition.

If \( L(t_1) = L(t_2) = \epsilon \in \Sigma \), the firings of \( t_1 \) and \( t_2 \) are not distinguishable solely based on the observed label \( \epsilon \). Similarly, if \( L(t) = \varepsilon \), then the firing of transition \( t \) is not observed. In both cases the firing of the transition/transition may can perhaps be indirectly inferred by the firing of subsequent transitions and/or knowledge of the current and subsequent markings. For each \( \epsilon \in \Sigma \cup \{\varepsilon\} \), we define \( T_\epsilon = \{t \in T \mid \ L(t) = \epsilon\} \). In particular, \( T_\varepsilon \) is the set of silent or unobservable transitions. The transition in \( T_\varepsilon \) with \( |T_\varepsilon| = 1 \) for \( \epsilon \in \Sigma \) is said to be deterministic, and the transitions in \( T_\varepsilon \) with \( |T_\varepsilon| \geq 2 \) for \( \epsilon \in \Sigma \) are said to be indistinguishable. Given a labeled PN \( G_L = (N,\Sigma,L) \), a transition \( t \) is observable if \( L(t) \in \Sigma; \) \( T_0 = \{t \in T \mid \ L(t) \neq \varepsilon\} \) is used to denote the set of observable transitions. Clearly, we have \( T = T_\varepsilon \cup T_0 \) and \( T_\varepsilon \cap T_0 \emptyset \).

Given a firing sequence \( S = t_{s_1}t_{s_2}\cdots t_{s_k} \) in a labeled PN, the corresponding observation sequence is \( \omega = L(S) := L(t_{s_1})L(t_{s_2})\cdots L(t_{s_k}) \), i.e., a string in \( \{\varepsilon\} \cup \Sigma \cup \Sigma^2 \cup \ldots \cup \Sigma^k \) (depending on how many of the transitions in the sequence \( S \) are unobservable). Given an observation sequence \( \omega \) generated by a labeled PN, there can be multiple firing sequences that can be mapped to \( \omega \), as well as multiple consistent sequences of markings (system states). In particular, we will be interested in firing sequences that are consistent with the given sequence of observations and also end with an observable event (as well as their corresponding current and initial marking estimates). As we explain later, sequences that end with observable events allow us to infer (if necessary) information about all possible firing sequences that are consistent with the given sequence of observations, while allowing for easier tracking of the probabilities with which different consistent firing sequences occur.

Definition 3: Given a labeled PN \( G_L = (N,\Sigma,L) \) with initial marking in the finite set \( M_0 = \{M_0(1),M_0(2),\ldots,M_0(|M_0|)\} \) and an observed label sequence \( \omega \in \Sigma^* \) (i.e., \( \omega \in \Sigma^*, \omega \neq \varepsilon \)), the set of reduced consistent firing sequences, the set of reduced consistent markings, and the set of consistent initial markings are defined respectively.
as
\[ S_r(\omega) = \{ S \in T^*T_0 \mid \exists M_0 \in M_0 : \{ M_0[S]\text{ and } L(S) = \omega \} \}, \]
\[ C_r(\omega) = \{ M \in (N \cup \{0\})^n \mid \exists S \in T^*T_0, \exists M_0 \in M_0 : \{ M_0[S\rangle M \text{ and } L(S) = \omega \} \}. \]
\[ I(\omega) = \{ M_0 \in M_0 \mid \exists S \in T^*T_0 : \{ M_0[S]\text{ and } L(S) = \omega \} \}. \]

For \( \omega = \varepsilon \), we take \( S_r(\varepsilon) = \{ \varepsilon \}, C_r(\varepsilon) = M_0 \), and \( I(\varepsilon) = M_0 \).

**Assumption A1** Given a labeled PN \( G_L = (N, \Sigma, L) \) and a finite set of possible initial markings \( M_0 = \{ M_0^{(1)}, M_0^{(2)}, ..., M_0^{(|M_0|)} \} \), we assume that \( |R(N, M_0)| < \infty \), \( \forall M_0 \in M_0 \). Equivalently (since \( M_0 \) is a finite set), we have \( |R(N, M_0)| < \infty \), where \( R(N, M_0) \equiv \cup_{M_0 \in M_0} R(N, M_0) \).

**Remark 1:** In order to deal with unobservable transitions, researchers (see, for example, [13]) typically require the unobservable subnet (\( T_z \)-induced subnet) of a labeled PN to be acyclic or deadlock structurally bounded. Among other implications, either of these two assumptions means that the length of any firing sequence of unobservable transitions from any bounded initial marking \( M \) is finite (bounded in length [3]) and thus the set of markings reachable from any initial marking \( M \) via unobservable transitions is also finite. In this paper, we relax the restriction (e.g. in [3]) that the unobservable subnet be deadlock structurally bounded; in particular, we allow firing sequences of unobservable transitions to include sequences of arbitrary length (as long as the resulting reachable set contains a bounded number of markings).

### III. Problem Formulation and General Solution

We consider a labeled PN \( G_L = (N, \Sigma, L) \) with initial marking in the finite set \( M_0 = \{ M_0^{(1)}, M_0^{(2)}, ..., M_0^{(|M_0|)} \} \) with known a priori probability for each initial marking, i.e., \( \Pr(M_0^{(i)}) = p_0^{(i)}, i = 1, 2, ..., |M_0|, \) where \( \sum_i p_0^{(i)} = 1 \). Given a sequence of observations \( \omega \), we would like to obtain the sets of reduced consistent markings and consistent initial markings in Definition 3, along with their a posteriori probabilities (given \( \omega \)), i.e., we would like to obtain
\[ I_p(\omega) = \{ \{ M_0^{(i)} \mid p_0^{(i)}(\omega) \} \mid M_0^{(i)} \in I(\omega) \}, \]
where \( p_0^{(i)}(\omega) = \Pr(M_0^{(i)} \mid \omega) \) is the a posteriori probability (given \( \omega \)) of the consistent initial marking \( M_0^{(i)}, M_0^{(i)} \in I(\omega) \). Note that \( p_0^{(i)}(\omega) \) are conditional probabilities, and we clearly have \( \sum_{i \in I(\omega)} p_0^{(i)}(\omega) = 1 \) (we already assumed that \( \sum_{i \in I(\omega)} p_0^{(i)}(\omega) = 1 \)). The problem of determining the set of reduced consistent (current) markings and their associated posterior probabilities has been addressed in [7].

For the above estimation problem to make sense, we would need to know (or be able to compute based on some probabilistic model) the a priori probability of occurrence for each firing sequence \( s^{(\kappa)} \in T^* \) (where \( \kappa \) is an integer used to index all firing sequences, in some arbitrary ordering) from any given initial marking \( M_0^{(j)} \in M_0 \); we will denote these probabilities by
\[ \Pr(\kappa|j) \equiv \Pr(s^{(\kappa)} \mid M_0^{(j)}), \forall M_0^{(j)} \in M_0, s^{(\kappa)} \in T^*. \]
Similarly, we use
\[ \Pr(\kappa, j) \equiv \Pr(s^{(\kappa)}, M_0^{(j)}) = \Pr(\kappa|j)p_0^{(j)}(\omega) \]
to denote the joint probability that the PN started at initial marking \( M_0^{(j)} \) and the sequence of transitions \( s^{(\kappa)} \) occurred. Note that some of these probabilities might be zero; in particular, when sequence \( s^{(\kappa)} \) is not possible from initial marking \( M_0^{(j)} \), we have \( \Pr(\kappa|j) = 0 \).

Given a sequence of observations \( \omega \), we are interested in obtaining the conditional probabilities \( p_0^{(i)}(\omega) \) (where \( i \) is an arbitrary index of each possible initial marking in \( I(\omega) \)). Given the sequence of observations \( \omega \), the only possible firing sequences are sequences in the set \( S_r(\omega) \). Since the sequences in the set \( S_r(\omega) \) are not prefixes of each other, we have that a priori (i.e., before any observations are made), the probability of observing \( \omega \) is
\[ \Pr(\omega) = \sum_{j : M_0^{(j)} \in M_0} \sum_{\kappa : s^{(\kappa)} \in S_r(\omega)} \Pr(\kappa, j) \]
\[ = \sum_{j : M_0^{(j)} \in M_0} p_0^{(j)}(\omega) \sum_{\kappa : s^{(\kappa)} \in S_r(\omega)} \Pr(\kappa|j), \quad (2) \]
i.e., we need to consider all possible initial markings \( M_0^{(j)} \) and, for each such initial marking, consider all sequences of transitions \( s^{(\kappa)} \) that satisfy \( s^{(\kappa)} \subseteq S_r(\omega) \) (i.e., \( s^{(\kappa)} \) generates the observation sequence \( \omega \) and ends with an observable event). The probability of observing \( \omega \) is then simply the sum of the joint probabilities \( \Pr(s^{(\kappa)}, M_0^{(j)}) \).

Using the total rule of probability and Bayes’ rule [14], we can arrive at the following: for each marking \( M_0^{(i)} \in M_0 \), we can calculate
\[ p_0^{(i)}(\omega) \equiv \Pr(M_0^{(i)} \mid \omega) = \frac{\Pr(M_0^{(i)}, \omega)}{\Pr(\omega)} \]
\[ = \frac{\sum_{\kappa : s^{(\kappa)} \in S_r(\omega)} \Pr(\kappa, i)}{\sum_{j : M_0^{(j)} \in M_0} \sum_{\kappa : s^{(\kappa)} \in S_r(\omega)} \Pr(\kappa, j)} \]
\[ = \frac{p_0^{(i)}(\omega) \sum_{\kappa : s^{(\kappa)} \in S_r(\omega)} \Pr(\kappa|j)}{\sum_{j : M_0^{(j)} \in M_0} p_0^{(j)}(\omega) \sum_{\kappa : s^{(\kappa)} \in S_r(\omega)} \Pr(\kappa|j)}. \] (3)
Specifically, given observation sequence \( \omega \), the probability that we have started from a particular initial state \( M_0^{(i)} \) depends on the probability of all sequences that start from \( M_0^{(i)} \) and generate observation \( \omega \), normalized by the probability of all sequences (starting from any initial state in \( M_0 \)) that generate observation \( \omega \). Note that this probability will be zero if marking \( M_0^{(i)} \notin I(\omega) \).

**Remark 2:** State-based opacity formulations typically assume that a certain subset \( S \) of system states are secret, and aim at understanding whether an intruder (that knows the system model and observes certain behavior generated by the
system) is able to determine that the system is in states in the set \( S \) (current state opacity, e.g., [15]) or started from states in the set \( S \) (initial state opacity, e.g., [16]). Specifically, probabilistic current-state opacity in [17] requires that, for each possible sequence of observations \( \omega \) (that can be generated by underlying activity in a given probabilistic finite automaton), the following property holds: the increase in the conditional probability \( \Pr(S|\omega) \) that the system current state lies in the set of secret states \( S \) (conditioned on the given sequence of observations \( \omega \)) compared to the prior probability \( \Pr(S) \) (that the initial state lied in the set of secret states \( S \) before any observation became available) is smaller than a given threshold.

It is clear that the techniques proposed here can certainly be used to analyze probabilistic initial state opacity in PNs by comparing the \textit{a priori} probability \( \Pr(S) = \sum_{i:M^{(i)}_0 \in S} t^{(i)}_0 \) to the \textit{a posteriori} probability \( \Pr(S|\omega) = \sum_{i:M^{(i)}_0 \in S} p_0^{(i)}(\omega) \). However, the verification of initial state opacity (i.e., the ability to determine whether the increase in \( \Pr(S|\omega) - \Pr(S) \) remains below a threshold for all sequences of observations \( \omega \) that might be generated by the given Petri net) will require additional work.

Performing the calculations in (3) might be difficult for a variety of reasons. For instance, the set \( S_r(\omega) \) might have infinite cardinality (due to the firing of unobservable transitions, which could also include the firing of an arbitrary number of unobservable transitions in a cyclic sequence).

IV. RECURSIVE ALGORITHM FOR INITIAL STATE ESTIMATION

A. Model and Assumptions

\textbf{Assumption A2:} When the PN \( N = (P,T,F,W) \) is at a particular marking \( M \), the following holds: for each transition that is enabled at \( M \) (i.e., for each \( t \in T \) such that \( M[t] \)) we have an associated probability \( p_M(t) \) that indicates the \textit{a priori} probability that \( t \) fires at \( M \). In addition,

$$\sum_{t \in T:M[t]} p_M(t) < 1,$$

where \( 1 - \sum_{t \in T:M[t]} p_M(t) \geq 0 \) is the probability that no transition fires at marking \( M \).

Though Assumption A2 is quite general, it is nevertheless restrictive: from a given marking \( M \), the probabilities with which the firings of transitions occur do not depend on how one arrives at marking \( M \). Note that the description of the probabilistic model requires knowledge of \( p_M(t) \) for all markings \( M \) (that are reachable from an initial marking in the finite set \( M_0 \)) and all transitions \( t \) (that are enabled at \( M \)), which can result in a complicated description.

Given the model described in Assumption A2, we can easily calculate the \textit{a priori} probabilities

$$\Pr(\kappa, j) = \frac{\sum_{j:M^{(j)}_0 \in S} t^{(i)}_0 \cdot p_M(t)}{\Pr(M^{(j)})},$$

where \( R(N,M_0) = \cup_{M_0 \in M_0} R(N,M_0) \). Specifically, given any sequence \( s^{(\kappa)} = t_{s_1}t_{s_2} \cdots t_{s_i}, t_{s_i} \in T \), such that \( M^{(j)}[t_{s_1}]M_1[t_{s_2}]M_2 \cdots [t_{s_i}]M_i \), we have

$$\Pr(\kappa, j) = \frac{\sum_{j:M^{(j)}_0 \in S} t^{(i)}_0 \cdot p_M(t)}{\Pr(M^{(j)})},$$

B. A recursive way to compute \( C_r(\omega) \)

In [7] we showed that the set \( C_r(e_{i_1}e_{i_2} \cdots e_{i_k}) \) can be obtained as a function of the observation \( e_{i_k} \) and the previous \( C_r(e_{i_1}e_{i_2} \cdots e_{i_{k-1}}) \). More specifically, we have

$$C_r(e_{i_1}e_{i_2} \cdots e_{i_k}) = \{ M \in (\mathbb{N} \cup \{0\})^n \mid \exists S \in T^* T_0, \exists M' \in C_r(e_{i_1}e_{i_2} \cdots e_{i_{k-1}}) \text{ s.t. } L(S) = e_{i_k}, M'[S]M \},$$

which provides us with a recursive way of maintaining the set of current state estimates \( C_r \) as observations become available.

C. Recursive Algorithm for Initial State Estimation

In this subsection we present the recursive algorithm for initial state estimation. Let us first note that the denominator for the expression in (3) is simply a normalizing factor, thus the quantity we focus our attention to is the numerator of (3). The easiest thing to do is to keep track of the auxiliary variables

$$p^{(i;\ell)}(\omega) = \sum_{s^{(\kappa)} \in S_r(\omega):M^{(i)}_0 \in S^{(\kappa)} \in M^{(i)}} \Pr(\kappa|i),$$

where \( i \) is an index of the set of possible initial states (in the finite set \( M_0 \)) and \( \ell \) is an index of the set of possible current states (in the set \( C_r(\omega) \)). Specifically, once the above quantities are available for all \( M^{(i)}_0 \in M_0 \) and all \( M^{(i)} \in C_r(\omega) \), we can easily obtain the needed probabilities, for all \( M^{(i)}_0 \in M_0 \) as

$$p_0^{(i)}(\omega) = \frac{\Pr(M^{(i)}_0 | \omega)}{\sum_{j:M^{(j)}_0 \in S_r(\omega)} p^{(j;\ell)}(\omega)},$$

where \( p^{(j;\ell)}(\omega) \) for \( \omega = e_{i_1}e_{i_2} \cdots e_{i_k} \), we argue that we can recursively obtain the probabilities \( p^{(j;\ell)}(e_{i_1}e_{i_2} \cdots e_{i_k}), \) for \( \ell = 1, 2, \ldots, k \). We make the following observations:

(i) Given \( p^{(i;\ell)}(\omega) \) (for all valid pairs of an initial state \( M^{(i)}_0 \in M_0 \) and a current state \( M^{(i)} \in C_r(\omega) \)), we can obtain the numerator of (3) as

$$p_0^{(i)} \sum_{\kappa,s^{(\kappa)} \in S_r(\omega)} \Pr(\kappa|i) = \sum_{t:M^{(i)}_0 \in C_r(\omega)} p^{(i;\ell)}(\omega).$$

Since the denominator of the expression in (3) is simply a normalizing coefficient, these auxiliary variables are sufficient for obtaining the initial state probabilities in (3).

(ii) Given the auxiliary variables \( p^{(j;\ell')}(\omega') \) for \( \omega' = e_{i_1}e_{i_2} \cdots e_{i_{k-1}} \) (where \( \ell' \) is an index for the markings in \( C_r(\omega') \)), we can obtain the auxiliary variables for \( p^{(j;\ell)}(\omega) \) for \( \omega = e_{i_1}e_{i_2} \cdots e_{i_{k-1}}e_{i_k} = \omega' e_{i_k} \) (where \( \ell \) is an index for the markings in \( C_r(\omega') \)) as follows: we first realize that any string \( s^{(\kappa)} \in S_r(\omega) \) that leads from marking \( M^{(i)}_0 \in M_0 \)
to marking $M^{(i)} \in C_\tau(\omega)$ can always be written as the concatenation of a prefix string $s(\kappa') \in S_\tau(\omega')$ and some suffix string $s \in T^*T$ that satisfy

$$M_0^{(i)}[s(\kappa')]M^{(i')}[s]M^{(i)}, \quad L(s) = e_{ik}$$

for some intermediate marking $M^{(i')} \in C_\tau(\omega')$. Now, we have

$$p^{(i,l)}(\omega) = \sum_{\kappa' : s(\kappa') \in S_\tau(\omega'), M_0^{(i)}[s(\kappa')]M^{(i)'}} \Pr(\kappa'|i),$$

$$\sum_{s \in T^*T, L(s) = e_{ik}: M^{(i)}[s]M^{(i)}} \Pr(s|T'),$$

thus, using (6), we can prove that

$$p^{(i,l)}(\omega) = \sum_{\omega' : M^{(i)} \in C_\tau(\omega')} \sum_{T' : \omega' \in C_\tau(\omega')} \Pr(s|T'),$$

which establishes the recursive calculation of the auxiliary variables $p^{(i,l)}(\omega)$ (and therefore the probabilities of the initial states).

The recursive algorithm for initial state estimation is summarized below. Its complexity is linear in the size of $M_0$ (denoted by $m_0 = |M_0|$) and depends on the number of markings that are possible after each observation. Following our assumptions, the number of markings after each observation is finite but could be changing; if $Q$ is the maximum (or an upper bound) on the number of markings after each observation, the complexity of the recursive step of the algorithm is $O(m_0Q^2)$. Following similar steps as the analysis of the algorithm for current state estimation in [8], we can show that the overall complexity of the recursive algorithm for initial state estimation following $k$ observations is $O(m_0(k + m)Q^2 + Q^3).

**V. Example**

Consider the labeled PN $G_L = (N, \Sigma, L)$ on the left of Fig. 1 (taken from [7]) where $N = (P, T, F, W)$ (with $P = \{p_1, p_2, p_3, p_4, p_5\}$, $T = \{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8\}$, $F$ and $W$ are as shown in the figure), $\Sigma = \{\alpha, \beta\}$, and the labeling function $L$ is defined as $L(t_1) = L(t_2) = L(t_3) = \varepsilon$, $L(t_4) = L(t_5) = L(t_7) = \alpha$, $L(t_6) = L(t_8) = \beta$. Note that $T_\alpha = \{t_1, t_2, t_3\}$ is the set of unobservable transitions. We have $T_\alpha = \{t_4, t_5, t_6, t_7\}$ and $T_\beta = \{t_6, t_8\}$.

We assume that the set of possible initial markings is $M_0 = \{M_0^{(1)}, M_0^{(2)}, M_0^{(3)}\}$, where $M_0^{(1)} = [1 \ 0 \ 0 \ 0 \ 0]^T$, $M_0^{(2)} = [0 \ 1 \ 0 \ 0 \ 0]^T$, $M_0^{(3)} = [0 \ 0 \ 1 \ 0 \ 0]^T$, with corresponding probabilities $p_0^{(1)} = 1/2$, $p_0^{(2)} = 1/3$, and $p_0^{(3)} = 1/6$.

The reachable set $R(N, M_0) \equiv R(N, M_0^{(1)}) \cup R(N, M_0^{(2)}) \cup R(N, M_0^{(3)})$ is identical to the one shown on the right of Fig. 1 (and satisfies Assumption A1) with the following reachable markings: $M^{(1)} = [1 \ 0 \ 0 \ 0 \ 0]^T = M_0^{(1)}$, $M^{(2)} = [0 \ 1 \ 0 \ 0 \ 0]^T = M_0^{(2)}$, $M^{(3)} = [0 \ 0 \ 1 \ 0 \ 0]^T = M_0^{(3)}$.

**Algorithm for Initial State Estimation**

1. **Given**: Labeled PN $G_L = (N, \Sigma, L)$ with initial markings in the finite set $M_0 = \{M_0^{(1)}, M_0^{(2)}, \ldots, M_0^{(M_0)}\}$, each with probability $p_0^{(i)}$.

2. **Given**: Streaming sequence of observations $\omega = e_{i1}e_{i2}\ldots e_{ik}e_{ik}$.

3. **Initialize**: $\omega = \varepsilon$, $C_\tau(\omega) = M_0$ with $p^{(i,l)}(\omega) = p_0^{(i)}$ for $M^{(i)} \in C_\tau(\omega)$ (zero otherwise).

4. **For** $\ell = 1, 2, \ldots, k$ **do**
   4.1. Compute $C_\tau(\omega e_{i\ell})$ using $C_\tau(\omega)$ via Eq. (5).
   4.2. For all $M^{(i)} \in M_0$ and for all $M^{(i)} \in C_\tau(\omega e_{i\ell})$, calculate $p^{(i)}(\omega e_{i\ell})$ using Eq. (7).
   4.3. For each $M^{(i)} \in M_0$ set the normalized probability to 
   4.4. $p^{(i)}(\omega e_{i\ell}) = \frac{1}{\sum_{M^{(i)} \in M_0} \sum_{M^{(i)} \in C_\tau(\omega e_{i\ell})} p^{(i)}(\omega e_{i\ell})}$.
   5. Set $\omega = \omega e_{i\ell}$.
   6. **Output** $I_p(\omega) = \{(M^{(i)}, p^{(i)}(\omega)) \mid M^{(i)} \in I(\omega)\}$.
which lead to $M^{(5)}$.

(3) From marking $M^{(3)}$, observation $\alpha$ is not possible.

Putting the above information together and using (7), we obtain the following probabilities: $p^{(1,3)}(\alpha) = 1/2 \times 4/26 = 2/26$, $p^{(1,5)}(\alpha) = 1/2 \times 9/26 = 9/52$, $p^{(2,3)}(\alpha) = 1/3 \times 12/26 = 4/26$, $p^{(2,5)}(\alpha) = 1/3 \times 1/26 = 1/78$ (all other $p^{(i,j)}(\alpha)$ are zero). After renormalization, we have $\Pr(M^{(1)} | \alpha) = 3/5$ and $\Pr(M^{(2)} | \alpha) = 2/5$.

When the second observation ($\beta$) is made, we realize the following (in the analysis below, we only consider the possible starting markings, i.e., the markings $M^{(3)}$ and $M^{(5)}$ which are the possible markings following the observation $\alpha$). From marking $M^{(3)}$, observation $\beta$ is possible via the transition $t_6$ which leads to $M^{(4)}$, whereas from marking $M^{(5)}$, observation $\beta$ is possible via the transition $t_8$ which leads to $M^{(4)}$. Putting the above information together, we obtain the following: $p^{(1,3)}(\alpha \beta) = 1/2 \times p^{(1,5)}(\alpha \beta) = 9/104$, $p^{(1,4)}(\alpha \beta) = 1/2 \times p^{(1,3)}(\alpha \beta) = 1/26$, $p^{(2,3)}(\alpha \beta) = 1/2 \times p^{(2,5)}(\alpha \beta) = 1/156$, $p^{(2,4)}(\alpha \beta) = 1/2 \times p^{(2,3)}(\alpha \beta) = 2/26$ (all other $p^{(i,j)}(\alpha \beta)$ are zero). After renormalization, we have $\Pr(M^{(3)} | \alpha \beta) = 3/5$ and $\Pr(M^{(2)} | \alpha \beta) = 2/5$.

Similarly, we can obtain the conditional probabilities for the initial markings after the third observation is made.

VI. CONCLUSIONS AND FUTURE WORK

In this paper we have considered a Markovian PN setting and developed a recursive algorithm for initial marking estimation given a sequence of observed labels, possibly in the presence of silent transitions. As a future research in this framework we plan to consider applications of such an approach to fault diagnosis, supervisory control, and opacity. We also plan to characterize conditions under which the complexity of the recursive algorithm can be reduced (e.g., when separate tracking of each consistent state and its corresponding probability is not required).

REFERENCES


