Bounds on the Number of Markings Consistent with Label Observations in Petri Nets

Yu Ru, Student Member, IEEE, and Christoforos N. Hadjicostis, Senior Member, IEEE

Abstract—In this paper, we consider state estimation in Discrete Event Systems (DESs) modeled by labeled Petri nets and present upper bounds on the number of system states (markings) that are consistent with an observed sequence of labels. Our analysis is applicable to Petri nets that may have nondeterministic transitions (i.e., transitions that share the same label) and/or unobservable transitions (i.e., transitions that are associated with the null label). More specifically, given knowledge of the initial Petri net state, we show that the number of consistent markings in a Petri net with nondeterministic transitions is at most polynomial in the length of the observation sequence (i.e., polynomial in the number of labels observed); this is true despite the fact that the number of firing sequences can be exponential in the length of the observation sequence. The result applies to general Petri nets without any restrictions on the structure of the Petri net or the nature of the labeling that are associated with the null label). This polynomial dependency of the number of consistent markings on the length of the observation sequence also applies to Petri nets with unobservable transitions under the assumption that their unobservable subnets are structurally bounded. The bounds on the number of markings established in this paper imply that the state estimation problem can be solved with complexity that is polynomial in the length of the observation sequence.

Note to Practitioners—A variety of systems, such as manufacturing systems, computer networks, traffic systems, and others, can be modeled as discrete event systems at a higher level of abstraction. These models can then be used for the purpose of state estimation, supervisory control and fault diagnosis. Knowledge of the system state is always critical for controller design and fault diagnosis. In this paper, we study state estimation in DESs modeled by labeled Petri nets. This is a challenging problem because two kinds of uncertainty arise due to sensor limitations: a) occurrences of different state transitions may generate the same observation; b) occurrences of unobservable state transitions go unrecorded. We show that, under some reasonable assumptions on the nature of the given Petri net's unobservable transitions, the number of possible system states that are consistent with an observed sequence of events is upper bounded by a function polynomial in the length of the observation sequence. The implications are two-fold: a) the state estimation problem can be solved with complexity that is polynomial in the length of the observation sequence in a quite general setting; b) the polynomial bound can guide the design of systems, especially when configuring the state transition sensors, to reduce the uncertainty introduced in the state estimation stage. The work can also be extended to handle place sensors (e.g., sensors monitoring resource changes).

Index Terms—Discrete event systems, Petri nets, State estimation.

I. INTRODUCTION

A discrete event system (DES) is a dynamic system that evolves in accordance with the abrupt occurrence, at possibly unknown and irregular intervals, of physical events [1], [2]. Such systems arise in a variety of contexts, such as queueing systems, computer systems, communication systems, manufacturing systems, traffic systems, database systems and software systems. In many Discrete Event Systems (DESs) state information cannot be obtained directly due to limited sensor availability. As a result, in applications that require explicit state information (e.g., supervisory control [3]–[7]), the problem of state estimation becomes crucial. When the underlying DES model is a labeled Petri net, system states that can be reached after observing a sequence of labels are usually not unique due to the fact that transitions can be nondeterministic (i.e., transitions can share the same label) and/or unobservable (i.e., transitions can be associated with the null label) [8]–[10]. In this paper, following a worst case analysis of the state estimation problem, we find upper bounds on the number of system states that are consistent with the observation of sequence of labels. In particular, we obtain upper bounds on the number of consistent system states in labeled Petri nets that may possess nondeterministic and/or unobservable transitions. First, assuming that the given Petri net has no unobservable transitions, we show that the number of markings consistent with the observation of a sequence of labels is at most polynomial in the length of the sequence. Then, we show that these results can be extended to Petri nets with unobservable transitions under the assumption that the unobservable subnet is (deadlock) structurally bounded. Using these bounds, we show that the state estimation problem can be solved with complexity that is polynomial in the length of the observed sequence of labels; this does not require any assumption on nondeterministic transitions (such as the contact-free assumption which is key for [8]).

The paper is organized as follows. Section II presents basic definitions of Petri nets. In Section III, we formulate problems on finding upper bounds on the number of consistent markings. In Sections IV and V, we present upper bounds for these problems. State estimation problems for Petri nets with nondeterministic transitions and unobservable transitions are discussed in Section VI and are shown to have computational...
complexity that is polynomial in the length of the observation sequence. Conclusions are given in Section VII.

II. PRELIMINARIES ON PETRI NETS

In this section, we review basic definitions of Petri nets. For more details, refer to [11], [12].

Definition 1 A Petri net structure 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; $W : F \rightarrow \{1, 2, 3, \ldots\}$ is a weight function; $P \cap T = \emptyset$ and $P \cup T \neq \emptyset$.

The set of all input (or output) places of a transition $t \in T$ is defined as $\bullet t = \{p \in P | (p, t) \in F\}$ (or $\bullet^* = \{p \in P | (t, p) \in F\}$). Similarly, the set of all input (or output) transitions of a place $p \in P$ is defined as $\bullet p = \{t \in T | (t, p) \in F\}$ (or $\bullet^* = \{t \in T | (p, t) \in F\}$). A transition is a source transition if $\bullet t = \emptyset$.

A marking is a function $M : P \rightarrow \{0, 1, 2, \ldots\}$ that assigns to each place a nonnegative integer number of tokens. Pictorially, places are represented by circles, transitions by bars and tokens by black dots, as shown in Fig. 1. We denote by $M(p)$ the number of tokens in place $p$. A Petri net $G$ is a net structure $N$ with an initial marking $M_0$. A Petri net is acyclic if it has no directed circuits.

A transition $t$ is said to be enabled if each input place $p$ of $t$ is marked with at least $W(p, t)$ tokens, where $W(p, t)$ is the weight of the arc from $p$ to $t$. We use $M[t]$ to denote that $t$ is enabled at marking $M$. An enabled transition $t$ may fire; its firing removes $W(p, t)$ tokens from each input place $p$ of $t$ and adds $W(t, p)$ tokens to each output place $p$ of $t$.

In this paper, we assume that only one transition can fire at any instant. A $k$-length firing sequence from marking $M$ is a sequence of transitions $S = t_{s_1}t_{s_2}\cdots t_{s_k}, t_{s_1} \in T$, such that $M[t_{s_1}]M_1[t_{s_2}]M_2[\cdots t_{s_k}]M'$; this is denoted by $M[S]M'$ and we say $S$ is enabled at marking $M$. The marking $M'$ can also be obtained using the following state equation

$$M' = M + D\sigma$$  \hspace{1cm} (1)

where (i) $D$ is the $n \times m$ incidence matrix of $N$ with $-W(p_i, t_j) + W(t_j, p_i)$ as its $(i,j)$ entry (note that 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) and (ii) $\sigma$ is the $m \times 1$ firing vector of $S$ (i.e., a vector whose $i$th entry denotes the number of times transition $t_i$ has appeared in firing sequence $S$). We use $D[;t]$ (or $D(\cdot ; p)$) to denote the column (or row) of $D$ that corresponds to transition $t$ (or place $p$).

A source transition is always enabled as it has no input places.

A marking $M'$ is reachable from $M$ if there exists a firing sequence $S$ such that $M[S]M'$. Given a Petri net, the set of all reachable markings from $M_0$ is called the reachability set and is denoted by $R(G, M_0)$. If $\forall p \in P$ and $\forall M \in R(G, M_0)$, $M(p) \leq K$ for some positive integer $K$, then we say the Petri net is $K$-bounded or simply bounded. A Petri net is said to be structurally bounded if it is bounded for any finite initial marking $M_0$. Equivalently, a Petri net is structurally bounded if and only if there exists an $n$-dimensional vector $y$ with positive integer entries such that $y^T D \leq \mathbf{0}_m$, where $D$ is the incidence matrix of the Petri net, $y^T$ denotes the transpose of $y$ and $\mathbf{0}_m$ denotes an $m$-dimensional column vector with all entries being 0 (see Theorem 29 in [11]).

A labeling function $L : T \rightarrow \Sigma \cup \{\lambda\}$ assigns to each transition $t \in T$ a label from a given alphabet $\Sigma \cup \{\lambda\}$. The labeling function of the Petri net in Fig. 1 is $L(t_1) = L(t_2) = a$ and $L(t_3) = b$. Note that the same label may be associated with more than one transition and unobservable transitions are labeled with the null label $\lambda$. If transition $t$ is not associated with the null label, we say that $t$ is nondeterministic when its label is shared with other transitions or we say $t$ is deterministic otherwise. Similarly, a label $e \in \Sigma$ is nondeterministic if there exists more than one transition $t$ such that $L(t) = e$, otherwise $e$ is deterministic. We use $T_e$ to denote the set of transitions associated with the label $e \in \Sigma \cup \{\lambda\}$. Note that $T_e \cap T_{\lambda} = \emptyset$ if $e \neq \lambda$. If all transitions are observable, i.e., if $T_{\lambda} = \emptyset$, we say that the labeling function is $\lambda$-free and the Petri net is a $\lambda$-free labeled Petri net.

To handle unobservable transitions, we define the unobservable subnet $N_\lambda$ of a Petri net $N$.

Definition 2 Given a Petri net $N = (P,T,F,W)$ and $T_\lambda \subseteq T$, we define the unobservable subnet as a net $N_\lambda = (P_\lambda, T_\lambda, F_\lambda, W_\lambda)$, where $P_\lambda = \{p \in P | \exists t \in T_\lambda, p \in \bullet \}$, $F_\lambda$ is the restriction of $F$ to $(P_\lambda \times T_\lambda) \cup (T_\lambda \times P_\lambda)$ and $W_\lambda$ is the restriction of $W$ to $F_\lambda$.

Note that the above definition of unobservable subnet is essentially the $T_{\lambda}$-induced subnet in [13] after removing isolated places.

Given a firing sequence $S = t_{s_1}t_{s_2}\cdots t_{s_k}$, the observation sequence is $\omega = L(S) := L(t_{s_1})L(t_{s_2})\cdots L(t_{s_k})$, i.e., a string in $\Sigma^*$ (the set of all possible strings generated from the alphabet $\Sigma$). Note that the null label $\lambda$ does not appear in an observation sequence and therefore, the occurrence of unobservable transitions in an execution of a Petri net goes unrecorded. In an observation sequence, a label is also called an event. Now we introduce the definition of the set of consistent markings [8] (namely, the set of possible system states consistent with an observed sequence of labels) and the set of consistent firing vectors.

Definition 3 Given a Petri net $G$ with initial marking $M_0$ and observed sequence of labels $\omega$, the set of consistent markings is $C(\omega) = \{M \mid \exists S \in T^* : M_0[S]M \text{ and } L(S) = \omega\}$; the set of consistent firing vectors is $FV(\omega) = \{\sigma \mid \exists S \in T^* : M_0[S]M, L(S) = \omega, \text{ and } \sigma \text{ is the firing vector of } S\}$. 

\[ \text{Fig. 1. Simple Petri net used to demonstrate basic concepts.} \]
III. PROBLEM FORMULATION

The general problem we consider is the following: given a labeled Petri net with a known initial marking and after observing a sequence of labels (generated by transition activities in the Petri net), we want to find an upper bound on the number of markings that are consistent with the observation sequence\(^2\). To simplify the problem, we first consider \(\lambda\)-free labeled Petri nets.

**Problem 1** Consider a \(\lambda\)-free labeled Petri net \(G\) with a known initial marking \(M_0\). Given an observed sequence of labels \(\omega\) with length \(k\) (which corresponds to an (unknown) underlying firing sequence \(S = t_1 t_2 \cdots t_{2k}\) such that \(\omega = L(S)\)), find a tight upper bound on the number of consistent markings \(|C(\omega)|\), where \(|C(\omega)|\) denotes the cardinality of the set \(C(\omega)\).

 Depending on the Petri net and the observation, the exact number of consistent markings may increase, decrease or remain unchanged as we observe more and more labels. However, there are extreme cases in which the number of consistent markings always increases. One such case is discussed in the following example.

**Example 1** Consider the Petri net in Fig. 2. There are three source transitions \(t_1, t_2\) and \(t_3\), all of which are associated with label \(e\) and are always enabled. If we observe the sequence of labels \(e e e\), the computation of consistent markings by enumerating all firing sequences proceeds as illustrated in Fig. 3.(a). The root of the tree is the initial marking \(M_0 = (0 0 0)^T\). The number of leafs in the tree is equal to the number of firing sequences and increases exponentially. Nodes with the same color (or the same number) at the same level represent the same consistent marking. The enumeration of consistent markings shown in Fig. 3.(b) is obtained after merging identical markings at each level.

In the next section, we show that the number of consistent markings in a \(\lambda\)-free labeled Petri net increases at most polynomially in the length of the observation sequence despite the fact that the number of firing sequences may increase exponentially (as was the case in the Petri net in Example 1). For Petri nets with unobservable transitions, we consider the following restricted problem.

**Problem 2** Consider a labeled Petri net \(G\) with a known initial marking \(M_0\) and a structurally bounded unobservable subnet. Given an observed sequence of labels \(\omega\) with length \(k\) (which corresponds to an (unknown) underlying firing sequence \(S\) that may include unobservable transitions and satisfies \(\omega = L(S)\)), find a tight upper bound on the number of consistent markings \(|C(\omega)|\).

IV. UPPER BOUNDS ON \(|C(\omega)|\) FOR \(\lambda\)-FREE LABELED PETRI NETS

In this section, we consider Problem 1 in Section III and upper bound the number of consistent markings in a \(\lambda\)-free labeled Petri net. In particular, we obtain upper bounds that are polynomial in the length of the observed sequence of labels.

One way to bound the number of consistent markings is to first bound the number of markings consistent with the observation sequence of labels. From the state equation (1), we have for every place \(p\)

\[ M(p) = M_0(p) + D(p, \cdot)\sigma. \]

Define \(a_1\) (or \(a_2\)) to be the maximum entry of \(M_0\) (or \(D\)); then, after observing a sequence \(\omega\) of \(k\) labels,

\[ M(p) \leq a_1 + a_2(1 1 \cdots 1)^T \sigma \leq a_1 + a_2 k, \]

because for any firing vector \(\sigma\) that corresponds to a firing sequence \(S\) that can be mapped to \(\omega\), we have \(\sum_{t \in S} \sigma(t) = k\) (where \(\sigma(t)\) denotes the number of occurrences of transition \(t\) in the firing sequence \(S\)). This implies that a straightforward upper bound on the number of consistent markings is \((1 + a_1 + a_2 k)^n\).

**Lemma 1** Consider a \(\lambda\)-free labeled Petri net with a known initial marking \(M_0\). After observing a sequence of labels \(\omega\) with length \(k\), the number of consistent markings can be upper bounded by \((1 + a_1 + a_2 k)^n\), where \(a_1\) (or \(a_2\)) is the maximum entry of \(M_0\) (or \(D\)).

In general, the above straightforward bound is not tight (especially for Petri nets with a relatively large number of places) because (i) many markings that satisfy \(M(p) \leq a_1 + a_2 k\) are not reachable given the observed sequence of labels \(\omega\), and (ii) the bound depends on the maximal entry of the initial marking \(M_0\) while, intuitively, the number of consistent markings will only depend on \(\omega\) when \(M_0\) is large enough.

Another way to bound the number of consistent markings is to first bound the number of firing vectors. The basic idea is that every consistent marking must be associated with at least one\(^3\) firing vector regardless of the exact ordering of transitions in the corresponding firing sequence; therefore, if we bound the number of possible firing vectors for the observed sequence of labels, we automatically have an upper bound on the number of consistent markings. This should be contrasted with the number of possible firing sequences which, as we will see, is generally much larger (because many firing sequences can share the same firing vector). In the remainder of this section we obtain an upper bound on the number of consistent markings.

\(^2\)Note that there will be at least one consistent marking since we assume that the observed sequence of labels is generated by activities in the Petri net.

\(^3\)Note that given a firing vector \(\sigma\), there exists a unique final marking \(M = M_0 + D\sigma\). However, given a marking \(M\), there may exist more than one corresponding firing vectors. For example, if two transitions have the same input and output places and identical arc weights, then the firing of any one of them results in the same marking but in different firing vectors. It is easy to verify that a sufficient condition for a given reachable consistent marking to correspond to a unique firing vector is that the incidence matrix \(D\) has full column rank.
consistent markings by bounding the number of possible firing vectors; we also relax this bound a bit so as to be able to express it using only a few structural parameters of the λ-free labeled Petri net and the length of the observation sequence.

Before considering the general case, we demonstrate our result for the extreme case of the net in Example 1: there is a single nondeterministic label e associated with three source transitions $t_1, t_2$ and $t_3$, all of which remain enabled all the time. Clearly, $|T_e| = 3$ where $T_e = \{t_1, t_2, t_3\}$. If the observation sequence is $\omega = e^k$, i.e., $\omega = e \ldots e$ with $k$ instances of label $e$ (where $k \geq 0$), then

$$\sigma(t_1) + \sigma(t_2) + \sigma(t_3) = k,$$

(2)

with $\sigma(t_i) \geq 0$ for $i = 1, 2, 3$. Obviously, every valid firing vector $\sigma$ must satisfy Eq. (2). Therefore, the number of all solutions of Eq. (2) is upper bound to valid firing vectors and, thus, an upper bound on the number of markings that are consistent with the observation sequence $e^k$.

Since $\sigma(t_3) = k - (\sigma(t_1) + \sigma(t_2))$, the number of solutions to Eq. (2) is equal to the number of all possible combinations of $\sigma(t_1)$ and $\sigma(t_2)$, which satisfy

$$\sigma(t_1) + \sigma(t_2) \leq k.$$

Equivalently, the above inequality can be expressed as a disjunction of the following equalities:

$$\sigma(t_1) + \sigma(t_2) = 0, \sigma(t_1) + \sigma(t_2) = 1, \ldots, \sigma(t_1) + \sigma(t_2) = k.$$  

Therefore, the number of solutions is

$$1 + 2 + \ldots + k + (k + 1) = \frac{(k + 1)(k + 2)}{2} = C_{k+2}^2,$$

where $C_n^r$ is the binomial coefficient "n choose r", and this is an upper bound on the number of markings that are consistent with the observation sequence $e^k$.

Similarly, if $T_e = \{t_1, t_2, t_3, t_4\}$ and the observation sequence is still $e^k$, an upper bound on the number of consistent markings is $C_3^4$. Based on these observations, one might

**Proposition 1** Consider a λ-free labeled Petri net with a known initial marking $M_0$ and all transitions $t_1, \ldots, t_l$ labeled $e$. If the observation sequence is $e^k$, then the number of consistent markings is upper bounded by $C_{k+l-1}^l$, where $l \geq 1$ and $k \geq 0$.

**Proof:** If $l = 1$, then $e$ is deterministic and the number of consistent markings is 1. Note that $C_{k+l-1}^l = 1$ when $l = 1$, which implies that the upper bound $C_{k+l-1}^l$ holds for $l = 1$.

In the case $l \geq 2$, since $\sigma(t_1) + \cdots + \sigma(t_l) = k$, where $\sigma(t_i) \geq 0$ for $i = 1, \ldots, l$, we have

$$(\sigma(t_1) + 1) + \cdots + (\sigma(t_l) + 1) = k + l.$$  

Let $\beta(t_i) = \sigma(t_i) + 1$; we have

$$\beta(t_1) + \cdots + \beta(t_l) = k + l,$$

where $\beta(t_i) \geq 1$ for $i = 1, \ldots, l$. Now we can think of the problem as dividing $k + l$ balls into $l$ groups while there should be at least one ball in every group. To arrange the $k + l$ balls into $l$ groups (each with at least one ball), we need to choose $l - 1$ separators among $k + l - 1$ possible separators (see Fig. 4). It is easy to verify that there are $C_{k+l-1}^{l-1}$ possible combinations, that is to say, there are $C_{k+l-1}^{l-1}$ possible combinations of $\sigma(t_1), \ldots, \sigma(t_l)$ such that

$$\sigma(t_1) + \cdots + \sigma(t_l) = k.$$  

Therefore, the number of firing vectors (and thus the number of consistent markings) is upper bounded by $C_{k+l-1}^{l-1}$. ■

Now we turn to the general case. To simplify the representation, we assume that

- there are $d$ labels $e_1, e_2, \ldots, e_d$ where $1 \leq d \leq m$;
- $l_{e_i} = |T_{e_i}| \geq 2$ for $i = 1, \ldots, j$ while $l_{e_i} = |T_{e_i}| = 1$ for $i = j + 1, \ldots, d$;
- the observation is a sequence of labels $\omega$ with length $k$, in which $e_i$ appears $k_{e_i}$ times and $k_{e_1} + \cdots + k_{e_d} = k$.

Furthermore, $k_{e_i} \geq 0$ and $k \geq 0$.

![Fig. 3. Computing consistent markings for the net in Fig. 2.](image)

![Fig. 4. Illustration of dividing balls.](image)
Note that if $j = 0$, all labels are deterministic; if $j = d$, all labels are nondeterministic.

**Proposition 2** Consider a $\lambda$-free labeled Petri net with a known initial marking $M_0$ as described above. If the observation is a sequence of labels $\omega$ with length $k$ in which label $e_i$ appears $k_{e_i}$ times and $k_{e_1} + \ldots + k_{e_d} = k$, then the number of consistent markings is upper bounded by $\prod_{i=1}^j C_{k_{e_i} + l_{e_i} - 1}^{l_{e_i} - 1}$, which is defined to be 1 if $j = 0$.

**Proof:** Given an observed sequence of labels $\omega$, a firing vector $\sigma$ satisfies the following set of equations:

$$
\sum_{t \in T_{e_1}} \sigma(t) = k_{e_1} \\
\vdots \\
\sum_{t \in T_{e_j}} \sigma(t) = k_{e_j} \\
\sigma(t_{e_{j+1}}) = k_{e_{j+1}} \\
\vdots \\
\sigma(t_{e_d}) = k_{e_d}
$$

where $T_{e_i}$ is the set of transitions that are associated with label $e_i$ for $i = 1, 2, \ldots, j$ and $T_{e_j} = \{ t_{e_j} \}$ for $i = j + 1, \ldots, d$. For nondeterministic label $e_i$, $\sum_{t \in T_{e_i}} \sigma(t) = k_{e_i}$ and there are $C_{k_{e_i} + l_{e_i} - 1}^{l_{e_i} - 1}$ possible solutions of $\sigma(t)$ for $t \in T_{e_i}$ following Proposition 1. Therefore, the number of solutions to Eq. (3) (namely, the upper bound on the number of consistent markings) is the product of all $C_{k_{e_i} + l_{e_i} - 1}^{l_{e_i} - 1}$, i.e., $\prod_{i=1}^j C_{k_{e_i} + l_{e_i} - 1}^{l_{e_i} - 1}$, because every pair of sets in $T_{e_1}, T_{e_2}, \ldots, T_{e_j}$ are disjoint. If $j = 0$, then all labels are deterministic and there is only one consistent marking.

The upper bound in Proposition 2 requires specific information about the observation sequence (namely, the $k_{e_i}$’s) and the distribution of transitions with respect to labels (namely, the $l_{e_i}$’s). The following theorem relaxes this upper bound and expresses it in terms of structural parameters of the given labeled Petri net and the length of the observation sequence.

**Theorem 1** Consider a $\lambda$-free labeled Petri net with a known initial marking $M_0$. If the observation sequence $\omega$ has length $k$, then the number of consistent markings is upper bounded by

$$
\frac{(\frac{k_{e_1} + 1}{2})^j (\frac{k_{e_j} + 7 - 1}{2})^j}{((\frac{k_{e_1} + 1}{2})^j)}
$$

where $j \geq 1$ is the number of nondeterministic labels, $\bar{I} = \max\{l_{e_1}, l_{e_2}, \ldots, l_{e_d}\}$ and $\underline{I} = \min\{l_{e_1}, l_{e_2}, \ldots, l_{e_d}\}$. If $j = 0$, we define the value of Eq. (4) to be 1.

**Proof:** We first prove that the result holds for $j \geq 1$.

Assume that $k_{e_1} + \ldots + k_{e_d} = q$; obviously $q \leq k$. Since

$$
\prod_{i=1}^j C_{k_{e_i} + l_{e_i} - 1}^{l_{e_i} - 1} = \prod_{i=1}^j \frac{(k_{e_i} + 1) \cdots (k_{e_i} + l_{e_i} - 1)}{(l_{e_i} - 1)!}
$$

and

$$
\frac{(k_{e_1} + 1) \cdots (k_{e_j} + l_{e_j} - 1)}{(l_{e_j} - 1)!} \leq \frac{(k_{e_1} + 1) \cdots (k_{e_j} + \bar{I} - 1)}{(\underline{I} - 1)!},
$$

we get

$$
\prod_{i=1}^j C_{k_{e_i} + l_{e_i} - 1}^{l_{e_i} - 1} \leq \prod_{i=1}^j \frac{(k_{e_1} + 1) \cdots (k_{e_j} + \bar{I} - 1)}{(\underline{I} - 1)!}.
$$

Note that for $1 \leq r \leq \bar{I} - 1$, we have

$$
(k_{e_1} + r) + \ldots + (k_{e_j} + r) = q + j r
$$

and

$$
(k_{e_1} + r) + \ldots + (k_{e_j} + r) \geq j \sqrt[4]{(k_{e_1} + r) \cdots (k_{e_j} + r)}
$$

by the arithmetic-mean/geometric-mean inequality [14]. Therefore,

$$
(k_{e_1} + r) \cdots (k_{e_j} + r) \leq \left( \frac{q + j r}{j} \right)^j
$$

and

$$
\prod_{i=1}^j C_{k_{e_i} + l_{e_i} - 1}^{l_{e_i} - 1} \leq \prod_{i=1}^j \frac{(k_{e_1} + 1) \cdots (k_{e_j} + \bar{I} - 1)}{(\underline{I} - 1)!}
$$

follows from the arithmetic-mean/geometric-mean inequality. Note that if $j = 0$, then all labels are deterministic and the number of consistent markings is 1, which is consistent with the limit of Eq. (4) as $j$ goes to 0.

**Remark 1** In the expression for the upper bound in Eq. (4), parameters $j, \bar{I}$ and $\underline{I}$ depend solely on the labeled Petri net structure. The upper bound is polynomial in the length of the observation sequence, i.e., it is $O(k^j(\bar{I}-1))$ which might come as a surprise since the number of possible firing sequences can increase exponentially with the length of the observation sequence. The reason is that there are many firing sequences that correspond to the same firing vector and thus, the same marking.

If we treat Eq. (4) as a continuous function of $j$, its limit using L’Hôpital’s Rule is also 1 as $j$ goes to 0.
V. UPPER BOUNDS ON $|\mathcal{C}(\omega)|$ FOR PETRI NETS WITH UNOBSERVABLE TRANSITIONS

In this section, we restrict our attention to Petri nets whose unobservable subnet is structurally bounded and consider Problem 2 as defined in Section III.

One way to bound the number of consistent markings is to first bound the number of tokens in every place. For any $M \in \mathcal{C}(\omega)$, $M = M_0 + D\sigma \geq 0$, where $\sigma$ is a firing vector corresponding to $M$. Using the partition of observable transitions and unobservable transitions, the state equation of the unobservable subnet can be written as

$$M^\lambda = M_0^\lambda + D_{\lambda}^\lambda \sigma_o + D_{\lambda uo}^\lambda \sigma_{uo} \geq 0_{n_1},$$

(9)

where $M^\lambda$ (or $M_0^\lambda$) is the restriction of $M$ (or $M_0$) to the set of places $P_\lambda$ in the unobservable subnet $N_\lambda$; $D_{\lambda}^\lambda$ (or $D_{\lambda uo}^\lambda$) is the submatrix of the incidence matrix $D$ that has rows that correspond to the places in $P_\lambda$ and columns that correspond to the transitions in $T_\lambda = T \setminus T_{\lambda}$ (or $T_{\lambda}$); $\sigma_o$ (or $\sigma_{uo}$) is the restriction of $\sigma$ to $T_o$ (or $T_{uo}$); and $n_1 = |P_\lambda| \leq n$.

Since the unobservable subnet is structurally bounded, there exists an $n_1$-dimensional vector $y_\lambda$ with positive integer entries such that $y_\lambda^T D_{\lambda uo}^\lambda \leq 0_{n_1}$, where $l_\lambda = |T_\lambda|$. Multiplying both sides of Eq. (9) by $y_\lambda^T$ on the left, we get

$$y_\lambda^T M^\lambda = y_\lambda^T M_0^\lambda + y_\lambda^T D_{\lambda}^\lambda \sigma_o + y_\lambda^T D_{\lambda uo}^\lambda \sigma_{uo} \geq 0.$$  

(10)

As $y_\lambda^T D_{\lambda}^\lambda \sigma_o \leq 0$, $y_\lambda^T M^\lambda \leq y_\lambda^T M_0^\lambda + y_\lambda^T D_{\lambda uo}^\lambda \sigma_{uo}$. Let $c_1 = y_\lambda^T M_0^\lambda$ and $c_2$ be the maximal entry of $y_\lambda^T D_{\lambda uo}^\lambda$. Then $y_\lambda^T M^\lambda \leq c_1 + c_2 (1 \ldots 1) \sigma_o \leq c_1 + c_2 k$, where $k$ is the length of the observation sequence $\omega$. As $y_\lambda$ is a vector with positive integer entries, $M^\lambda (p) \leq y_\lambda^T M^\lambda \leq c_1 + c_2 k$ for every $p \in P_\lambda$. Combining the above with the result in Lemma 1, we get the following straightforward upper bound on the number of consistent markings.

**Theorem 2** Consider a labeled Petri net with a known initial marking $M_0$ and a structurally bounded unobservable subnet, i.e., there exists an $n_1$-dimensional column vector $y_\lambda$ with positive integer entries that satisfies $y_\lambda^T D_{\lambda uo}^\lambda \leq 0_{n_1}$. If the observation sequence $\omega$ has length $k$, then the number of consistent markings is upper bounded by $(1 + a_1 + a_2 k)^{n-n_1} (1 + c_1 + c_2 k)^{n_1}$, where $a_1 = \max_{p \in P_\lambda} M_0(p)$, $a_2 = \max_{p \in P \setminus P_\lambda, t \in T} D(p, t)$, $n_1 = |P_\lambda|$, $n = |P|$, $c_1 = y_\lambda^T M_0^\lambda$, and $c_2$ is the maximal entry of $y_\lambda^T D_{\lambda uo}^\lambda$.

Note that if there is no unobservable transition (namely, if $l_\lambda = 0$), then the upper bound is exactly the bound in Lemma 1.

**Remark 2** Note that [15] computes a bound on the number of tokens in a place $p$ along all markings uncontrollably reachable from a given marking by constructing an abstract syntax tree. Here, we compute a bound on the number of tokens in any place $p$ for all markings that are consistent with an observation sequence $\omega$.

---

Fig. 5. A simple Petri net which is structurally bounded but not deadlock structurally bounded.

**Remark 3** Note that the bound in Theorem 2 depends on the value of $y_\lambda$ which is not unique (e.g., $c y_\lambda$ for any positive integer $c$ also satisfies $c y_\lambda^T D_{\lambda uo}^\lambda \leq 0$); also this bound may be too loose in a Petri net with a large number of places. To get a better bound, we may want to optimize $y_\lambda$ such that $1 + c_1 + c_2 k$ is minimized for values of $k$ of interest; however, even if we know a Petri net is structurally bounded, it may not be easy to find $y_\lambda$. For some Petri nets, structural boundedness cannot be established (e.g., input dominant Petri nets\(^8\) are shown to be structurally bounded with $y = 1_n$ and are easy to identify by checking conditions on each transition separately [16]). For more results on how to determine if a Petri net is structurally bounded, refer to [17]. The important observation for us here is that for Petri nets with structurally bounded unobservable subnets, $c_1$ and $c_2$ are constants.

Now we consider a special case of structurally bounded unobservable subnets, namely, the class of unobservable subnets for which there exists an $n_1$-dimensional column vector $y_\lambda$ with positive integer entries such that $y_\lambda^T D_{\lambda uo}^\lambda < 0_{n_1}$. In general, a Petri net with $n$ places and $m$ transitions is said to be deadlock structurally bounded if there exists an $n$-dimensional column vector $y$ with positive integer entries such that $y^T D < 0_n$. Note that there exist Petri nets that are structurally bounded but not deadlock structurally bounded (an example is shown in Fig. 5).

An important class of deadlock structurally bounded Petri nets is the class of acyclic Petri nets without source transitions. We discuss this result in the following proposition.

**Proposition 3** If a Petri net is acyclic and there are no source transitions, then the Petri net is deadlock structurally bounded.

**Proof:** Refer to the Appendix.

Now we rewrite Eq. (10) as

$$y_\lambda^T M_0^\lambda + y_\lambda^T D_{\lambda}^\lambda \sigma_o \geq - y_\lambda^T D_{\lambda uo}^\lambda \sigma_{uo}.$$  

If the unobservable subnet is deadlock structurally bounded, then

$$c_1 + c_2 k \geq y_\lambda^T M_0^\lambda + y_\lambda^T D_{\lambda}^\lambda \sigma_o \geq - y_\lambda^T D_{\lambda uo}^\lambda \sigma_{uo} \geq 1_{l_\lambda}^T \sigma_{uo} = \sum_{t \in T_{\lambda}} \sigma(t).$$

Therefore, every valid $\sigma$ must satisfy $\sum_{t \in T_{\lambda}} \sigma(t) \leq c_1 + c_2 k$.

---

\(^8\)A Petri net $G$ is called an input dominant Petri net if for each transition $t \in T$, the sum of the input arc weights is larger than or equal to the sum of the output arc weights.
Now we can generalize the result in Proposition 2 to handle unobservable transitions. To simplify the representation, we assume that

- the null label is denoted by \( \lambda \) and there are \( d \) other labels \( e_1, e_2, \ldots, e_d \) where \( 1 \leq d < m \);
- \( l_i = |T_i| \), \( l_s = |T_s| \geq 2 \) for \( i = 1, \ldots, j \) while \( l_e = |T_e| = 1 \) for \( i = j + 1, \ldots, d \);
- the observation is a sequence of labels \( \omega \) with length \( k \), in which \( e_i \) appears \( k_{e_i} \) times and \( k_{e_1} + \ldots + k_{e_d} = k \).

Furthermore, \( k_{e_i} \geq 0 \) and \( k \geq 0 \).

Proposition 4 Consider a labeled Petri net with a known initial marking \( M_0 \) and a deadlock structurally bounded unobservable subnet, i.e., there exists an \( n_1 \) dimensional column vector \( y_\lambda \) with positive integer entries that satisfies \( y_\lambda^T D_{uo} < 0 \).

If the observation is a sequence of labels \( \omega \) with length \( k \) in which \( e_i \) appears \( k_{e_i} \) times and \( k_{e_1} + \ldots + k_{e_d} = k \), then the number of consistent markings is upper bounded by

\[
\prod_{i=1}^j c_{l_i+e_i-1}^{l_i-1} \times C_{c_1+c_2}^{l_e},
\]

where \( c_1 = y_\lambda^T M_0^\lambda \) and \( c_2 \) is the maximal entry of \( y_\lambda^T D_0^\lambda \). The upper bound is defined to be \( C_{c_1+c_2}^{l_e} \) if \( j = 0 \).

Proof: A valid firing vector \( \sigma \) satisfies the equation array in Eq. (3) and \( \sum_{t \in T_s} \sigma(t) \leq c_1 + c_2 k \). Following the same steps as in the proof for Proposition 1, we can show that there are \( C_{c_1+c_2}^{l_e} \) possible combinations of \( \{ \sigma(t) \} \) such that \( \sum_{t \in T_s} \sigma(t) \leq c_1 + c_2 k \). Then, the number of valid firing vectors (and therefore, the number of consistent markings) is upper bounded by \( \prod_{i=1}^j c_{l_i+e_i-1}^{l_i-1} \times C_{c_1+c_2}^{l_e} \) because the sets \( T_{e_1}, T_{e_2}, \ldots, T_{e_j}, T_\lambda \) are disjoint. Note that if there is no unobservable transition (namely, if \( l_\lambda = 0 \), then the upper bound is exactly the bound in Proposition 2.

Corollary 1 Consider a labeled Petri net with a known initial marking \( M_0 \) and a deadlock structurally bounded unobservable subnet, i.e., there exists an \( n_1 \) dimensional column vector \( y_\lambda \) with positive integer entries that satisfies \( y_\lambda^T D_{uo} < 0 \). If the observation sequence \( \omega \) has length \( k \), then the number of consistent markings is upper bounded by

\[
\frac{(l_e + 1)}{j} \times C_{c_1+c_2}^{l_e},
\]

where \( j \geq 1 \) is the number of nondeterministic labels, \( \bar{l} = \max\{l_{e_1}, l_{e_2}, \ldots, l_{e_j}\} \), \( l = \min\{l_{e_1}, l_{e_2}, \ldots, l_{e_j}\} \), \( c_1 = y_\lambda^T M_0^\lambda \) and \( c_2 \) is the maximal entry of \( y_\lambda^T D_0^\lambda \). If \( j = 0 \), we define the value of Eq. (12) to be \( C_{c_1+c_2}^{l_e} \).

Proof: Direct application of Theorem 1.

Remark 4 Lemma 1 (or Theorem 2 and Corollary 1) can be relaxed to handle a set of possible initial markings, i.e., cases where we do not know exactly what the initial marking is but we know that it belongs to a given set of markings \( M_0 \). Naturally, in that case, we only need to redefine \( a_1 \) (or \( c_1 \)) as the maximal entry of all possible initial markings (or \( c_1 = \max_{M_0 \in M_0} y_\lambda^T M_0^\lambda \)). However, the bound in Theorem 1 can be directly used as long as the initial marking is unique even though we may not know what it is.

VI. IMPLICATIONS TO STATE ESTIMATION

In this section, we first review earlier work on state estimation in the context of DESs modeled by Petri nets and then apply the results in Section IV to establish that the computational complexity of the state estimation problem is polynomial in the length of the observation sequence when the underlying model is a labeled Petri net.

State estimation in DESs has been considered by many researchers [6], [8]–[10], [18], [19]. For instance, using the theory of generalized state space systems, the authors of [18] proposed an extended reduced Luenberger observer to reconstruct the Petri net marking and the firing vector based on partially measured places and transitions. In [6], [19], Giua et al. studied the problem of estimating the marking of a Petri net based on event observations, under the assumption that the net structure is known while the initial marking is totally or partially unknown; they were able to obtain a marking estimate that is a lower bound of the actual marking. In [8], Giua et al. studied state estimation in \( \lambda \)-free labeled Petri nets in which transitions can share labels (i.e., transitions are not necessarily deterministic); they showed that, as long as nondeterministic transitions are contact-free,\(^4\) the set of possible markings can be represented by a linear system with a fixed structure which does not depend on the length of the observation sequence. In [9], the authors considered Petri nets with deterministic and unobservable transitions under the assumption that the unobservable subnet is acyclic and backward conflict-free,\(^10\) and used a linear system with a fixed structure to represent markings that are consistent with a given observation sequence. The assumption that the unobservable subnet is backward conflict-free is relaxed in [10] and, as a result, the representation of consistent markings becomes more complicated.

The problem we consider here is the estimation of consistent markings when we observe a sequence of labels in labeled Petri nets that may have nondeterministic transitions and/or unobservable transitions. We first focus on Petri nets without unobservable transitions (\( \lambda \)-free labeled Petri nets).

A. State Estimation: \( \lambda \)-free Case

In this subsection we consider state estimation for a \( \lambda \)-free labeled Petri net \( G \). Given the observation of a sequence of labels \( \omega \) with length \( k \) (corresponding to an (unknown) underlying firing sequence \( t_{s_1}, t_{s_2}, \ldots, t_{s_l} \)), we are interested in computing all consistent markings. We assume that

A1) The structure of the Petri net \( G \) is known;
A2) The initial marking \( M_0 \) is known;
A3) The labeling function is \( \lambda \)-free (i.e., all transition firings are associated with labels that can be observed).

\(^4\)Nondeterministic transitions are contact-free if any pair of nondeterministic transitions \( t_i \) and \( t_j \) does not share input or output places, i.e., \( \omega_i \cap \omega_j = \emptyset \).

\(^10\)The backward conflict-free assumption means that all unobservable transitions have no common output places.
We first recall Algorithm 5 in [20] as shown below. This is a recursive online algorithm that computes the set of consistent markings as labels are observed. The idea is straightforward: for each observed label, one considers all transitions that share the label and are enabled; then, one simply enumerates all consistent markings.

Algorithm 1
1. \( \omega_0 = \lambda, C(\omega_0) = \{M_0\} \).
2. Let \( i = 0 \).
3. Wait until a new event \( e \) is observed.
4. Let \( i = i + 1, \omega_i = \omega_{i-1} \cdot e, C(\omega_i) = \emptyset \).
5. For all \( M \in C(\omega_{i-1}) \)
   
   For all \( t \) such that \( L(t) = e \) and \( M[t] \)
   
   Compute \( M' = M + D(t) \).
   
   Set \( C(\omega_i) = C(\omega_i) \cup \{M'\} \).

In the above algorithm, \( \omega_0 = \lambda \) is an initialization of the observation sequence; it does not mean we observe the null label \( \lambda \). Note that at Step 6, \( C(\omega_i) \) gives the set of all markings that are consistent with the observation of label sequence \( \omega_i \). To analyze the computational complexity of the above algorithm, assume that the number of consistent markings is \( N_{k-1} \) (or \( N_k \)) when the observation sequence has length \( k-1 \) (or \( k \)). If a new event \( e \) is observed after the observation sequence \( \omega_{k-1} \), then we need to (i) consider every transition associated with \( e \) for every consistent marking in \( C(\omega_{k-1}) \), (ii) obtain the next marking if it is enabled, (iii) compare the new marking with other consistent markings computed at stage \( k \), and (iv) add it to \( C(\omega_k) \) if it is not already included. Roughly, the complexity is \( N_{k-1} \times |T_k| \times (n + n + nN_k) \) in terms of scalar comparisons and additions, where the first \( n \) is the number of comparisons to determine whether some transition associated with \( e \) is enabled, the second \( n \) is the number of additions to compute the next marking and \( nN_k \) is the number of comparisons to check whether the next marking has been added into \( C(\omega_k) \). Following Theorem 1, we know that the number of consistent markings is at most polynomial in the length \( k \) of the observation sequence. Therefore, the complexity of one step computation from \( k-1 \) to \( k \) is \( O(nk^{2j}T^{-1}) \). Clearly, the complexity of computing \( C(\omega_k) \) starting from \( M_0 \) using Algorithm 1 is \( O(nk^{2j}T^{-1} + 1) \). This means that we can compute all consistent markings with complexity that is polynomial in the length of the observation sequence, even for general Petri nets without need for any assumptions (such as the contact-free assumption of [8]). The reason why we only focus on the length of the observation sequence is that, in the state estimation problem, the plant is given and only observations change with time.

Remark 5 The computational complexity of Algorithm 1 is an exponential function of the structural parameters \( j \) and \( 7 \), that is to say, the time required for computing consistent markings increases exponentially in the structural parameters (though it does increase polynomially in the length of an observation sequence). Note that this is also true for marking estimation in Petri nets under the contact-free assumption on nondeterministic transitions (see Proposition 10 of [8]). Also note that the bound on the computational complexity of Algorithm 1 is not exponential in the number of transitions but exponential in \( j \) and \( 7 \) which characterize the distribution of transitions with respect to labels.

Remark 6 Using similar arguments as above, the least-cost firing sequence estimation problem (more specifically, the estimation of the firing sequence with the least total cost given a sequence of observed labels and known costs of transition firings) was shown to also have complexity that is polynomial in the length of the observation sequence [21].

Example 2 Consider the Petri net in Fig. 6 with initial marking \( M_0 = (100 100 100 2)^T \). The labeling function is \( L(t_1) = L(t_2) = L(t_3) = a, L(t_4) = b \) and \( L(t_5) = c \). We randomly generate\(^{11} \) a sequence of labels \( \omega = caababaababaaabbbbaababbaabaaa \) of length 30 and then compute consistent markings using Algorithm 1. Typical values of the number of consistent markings, the number of firing vectors and the upper bound obtained in Proposition 2 are plotted against the length of the observation sequence in Fig. 7. Note that a bound on the number of possible (length 30) firing sequences that could correspond to the observed label sequence \( \omega = [T_c] \cdot [T_a] \cdot [T_a] \cdots [T_c] \cdot [T_a] \cdot [T_b] \) which turns out to be \( 316 \times 2^{12} = 1.7632 \times 10^{11} \). The exact number of firing sequences is \( 1.4599 \times 10^{11} \), which is much larger than 1989, i.e., the upper bound in Proposition 2. Note that the number of consistent markings is even lower at 1027. Fig. 8 shows a plot of the maximal, average and minimal number of consistent markings over 100 randomly generated observation sequences. Fig. 9 shows a plot of the maximal number of consistent markings over 100 random observation sequences together with the upper bound obtained from Theorem 1. Note that in this example, since \( j = 2, \ell = 3 \), \( k \) is 2, the bound on the number of consistent markings is \( \left( \frac{2}{4} + \frac{2}{4} \right)^4 \), i.e., \( O(k^4) \).

Remark 7 Note that the Petri net in Fig. 6 is not contact-free because \( p_1 \) is both the input place of \( t_4 \) and the output place of \( t_1 \), and \( t_1, t_4 \) share the same label \( a \). Therefore, we cannot compute the set of consistent markings using Algorithm 9 in [8], in which consistent markings are represented by a linear system with a fixed structure under the contact-free assumption on nondeterministic transitions. Also note that, since the Petri net is easily verified to be 212-bounded, the upper bound in Theorem 1 may exceed the number of all possible markings (i.e., \( 2^{12^4} = 2.02 \times 10^9 \)) for large \( k \) (in the scenario we considered, the upper bound at length 30 is \( 7.412 \times 10^4 \), which is much smaller than the number of \(^{12} \)Labels \( a, b \) and \( c \) are chosen with probabilities \( P(a) = \frac{3}{6}, P(b) = \frac{2}{6} \) and \( P(c) = \frac{1}{6} \). Note that the risk of generating a sequence of observed labels whose set of consistent markings is empty is low in this case since the initial marking has been chosen large enough. Also note that in reality the sequence of labels observed will be the result of transition activity in the Petri net.
Fig. 6. The Petri net analyzed in Example 2.

Fig. 7. Plot of three quantities of interest against the length of the observed sequence of labels.

Fig. 8. Plot of the maximal, average and minimal number of consistent markings over 100 randomly generated observation sequences.

Fig. 9. Plot of the maximal number of consistent markings over 100 randomly generated observation sequences together with the polynomial upper bound.

all possible markings). Therefore, the polynomial bound is useful in applications where the state of the system becomes periodically known or is periodically reset.

**Remark 8** Fig. 7 and Fig. 9 show that the upper bound obtained in Proposition 2 and the upper bound obtained in Theorem 1 are much larger than the number of consistent markings. The tightness of the bounds, however, depends on the structure of the labeled Petri net and the observation sequence. In fact, the upper bound in Proposition 2 can be reached in Petri nets in which (i) the incident matrix has full column rank (because in such cases, given an initial marking, the current marking corresponds to a unique firing vector), and (ii) the initial marking is large enough or there are enough source transitions (so that there exists at least one firing sequence for every possible firing vector). The upper bound obtained in Theorem 1 is close if, in addition to the conditions mentioned above to reach the upper bound in Proposition 2, the following conditions hold: a) there are no deterministic transitions (so that the inequality in Equation (8) becomes equality); b) \( T = L \) (so that the inequality in Equation (5) becomes equality); c) the number of observed labels for every nondeterministic label is the same (so that the inequality in Equation (6) becomes equality). For example, the number of markings that are consistent with observation sequence \( \omega = e^k \) for the Petri net in Fig. 2 is \( C^2_{k+2} = \frac{k^2 + 3k + 2}{2} \) since the incidence matrix has full column rank and there are enough source transitions; in this case, the bound in Theorem 1 (namely, \( (k+1)^2 = \frac{k^2 + 3k + 4}{2} = C^2_{k+2} + \frac{1}{2} \)) is very close as conditions a), b) and c) are all satisfied.

Now we modify Algorithm 1 slightly to express consistent markings using one marking and valid firing vectors. Define elementary firing vectors \( \sigma_0 = [0 \ldots 0]^T \) and \( \sigma_j = [0 \ldots 1 \ldots 0]^T \) (with the \( j \)th entry being 1 for \( 1 \leq j \leq m \)) be \( m \)-dimensional column vectors.

**Algorithm 2**

1. Let \( \omega_0 = \lambda \), \( M_b = M_0 \), \( FV(\omega_0) = \{\sigma_0\} \).
2. Let \( i = 0 \).
3. Wait until a new event \( e \) is observed.
4. Let \( i = i + 1 \), \( \omega_i = \omega_{i-1} e \), \( FV(\omega_i) = \emptyset \).
5. If \( e \) is deterministic
   
   For all \( \sigma \in FV(\omega_{i-1}) \)
   
   If \( (M_b + D \sigma)[T_e] \) and \( \sigma \not\in FV(\omega_i) \)
   
   \( FV(\omega_i) = FV(\omega_i) \cup \{\sigma\} \).


In Algorithm 1, all markings consistent with the observation sequence \( \omega \) are stored, while in Algorithm 2, we store only one marking \( M_b \) and all valid firing vectors. Generally speaking, the former is better in terms of saving storage space if there...
are more transitions than places and the latter is better if there are more places than transitions.

Note that here $M_b$ plays a similar role to $M_{b,\omega}$ in [8]. The main difference is that we allow $M_b$ to have negative entries. Using $M_b$, we can represent $C(\omega_i)$, i.e., markings consistent with the observation sequence $\omega_i$, as

$$C(\omega_i) = \{ M | M = M_b + D \sigma; \sigma \in FV(\omega_i) \} .$$

Since we update $M_b$ only when $e$ is deterministic, we can restrict the firing vector $\sigma$ to having entries only associated with nondeterministic transitions so as to save space when storing it. Also note that the representation using firing vectors may have redundancy since there may exist two or more firing vectors corresponding to the same consistent marking.

B. State Estimation: Unobservable Transition Case

In this subsection we relax assumption A3 and allow the existence of unobservable transitions (i.e., $\lambda$-labeled transitions) in the Petri net. More formally, we keep assumptions A1 and A2 but replace assumption A3 of Section VI-A with the assumption below.

A3’ Labels associated with firings of transitions in $T \setminus T_\lambda$ can be observed but transitions in $T_\lambda$ cannot (they are associated with the empty label).

First we recall the definition of the unobservable reach from a marking $M$ [22].

Definition 4 Given a Petri net, the unobservable reach $\mathbb{U}(M)$ from $M$ is $\{M' | 3S \in T_\lambda, M[S]M'\}$. To compute the set of consistent markings when unobservable transitions are present, we can modify Algorithm 1 slightly using the notion of the unobservable reach. Notice that the algorithm below assumes that the unobservable reach from a reachable marking $M$ can be computed (e.g., it is finite).

Algorithm 3

1. $\omega_0 = \lambda$, $C(\omega_0) = \{ \mathbb{U}(M_0) \}$.
2. Let $i = 0$.
3. Wait until a new event $e$ is observed.
4. Let $i = i + 1$, $\omega_i = \omega_{i-1} + e$, $C(\omega_i) = \emptyset$.
5. For all $M \in C(\omega_{i-1})$
   For all $t$ such that $L(t) = e$ and $M[t]$
   Compute $M' = M + D(\cdot, t)$.
   Set $C(\omega_i) = C(\omega_i) \cup \{M'\}$.
6. Let $C' = \bigcup_{M \in C(\omega_i)} \mathbb{U}(M)$, $C(\omega_i) = C'$.

In general, the computational complexity of Algorithm 3 can be high since the method essentially amounts to reachability analysis. However, if the unobservable subnet is structurally bounded, the computational complexity of Algorithm 3 is also polynomial in the length of the observation sequence; this can be easily argued using the result in Theorem 2. The following three examples are used to demonstrate the bounds in Theorem 2 and Corollary 1.

Example 3 We first look at the Petri net in Fig. 10 with initial marking $M_0 = (0 0)^T$ and labeling function $L(t_1) = L(t_4) = a$ and $L(t_2) = L(t_3) = \lambda$. As the unobservable subnet is structurally bounded but not deadlock structurally bounded, only the bound in Theorem 2 can be applied: Since the unobservable subnet is an input dominant Petri net, we can choose $y_\lambda = (1 1)^T$ as $y_\lambda^T D_{uo} = (0 0)$; therefore, $c_1 = y_\lambda^T M_0 = 0$ and $c_2 = 1$. As $P_\lambda = P$, $a_1 = a_2 = 0$. Therefore, the bound in Theorem 2 is $(1+c_1+c_2k)^n = (1+k)^2$ where $k$ is the length of the observation sequence. In Fig. 11 we plot both this bound and the actual number of consistent markings against $k$. The particular sequence of observations that was used to generate the plot in Fig. 11 was a sequence of 15 consecutive instances of label $a$.

Example 4 Now we look at the Petri net in Fig. 12 with initial marking $M_0 = (1 1 0 0 0)^T$ and labeling function $L(t_1) = a$, $L(t_2) = b$ and $L(t_3) = L(t_4) = L(t_5) = L(t_6) = \lambda$. Note that the unobservable subnet is acyclic and there is no source transition in the subnet; therefore, we can use Algorithm 4 in the Appendix to compute the vector $y_\lambda$. First, we use Eq. (13) to get the following partition of the unobservable subnet: $T_\lambda = T_1 \cup T_2 \cup T_3$, where $T_1 = \emptyset$, $T_2 = \{t_3, t_4\}$ and $T_3 = \{t_5, t_6\}$; $P_\lambda = P_1 \cup P_2 \cup P_3$, where $P_1 = \{p_1, p_2\}$, $P_2 = \{p_4\}$ and $P_3 = \{p_3, p_5\}$. Using the partition, we can rearrange the incidence matrix $D_{uo}^\lambda$ (that corresponds to the unobservable subnet) by exchanging the row corresponding to $p_3$ for the row corresponding to $p_4$; this results in the following matrix

$$D' = \begin{pmatrix} -F_{1,2} & -F_{1,3} \\ F_{2,2} & -F_{2,3} \\ F_{3,2} & F_{3,3} \end{pmatrix} .$$
In this paper, we presented upper bounds on the number of consistent markings in DESs when the underlying model is a labeled Petri net. For λ-free labeled Petri nets, we showed that the number of consistent markings is at most polynomial in the length of the observation sequence. Polynomial bounds on the number of consistent markings were also obtained for Petri nets with structurally bounded unobservable subnet. These bounds were used to show that the state estimation problem can be solved with complexity that is polynomial in the length of the observation sequence. They can also be used to guide the design of systems, especially when configuring transition sensors, to reduce the uncertainty introduced in the state estimation stage.

In the future, we will study bounds on the number of consistent markings in a more general setting (e.g., Petri nets with place sensors and transition sensors [23]) and apply state estimation schemes to supervisory controller synthesis.

APPENDIX

PROOF OF PROPOSITION 3

We show that an acyclic Petri net $G$ without source transitions is deadlock structurally bounded by giving an algorithm for computing a vector $y$ with positive integer entries such that $y^T D < 0^T_m$, where $D$ is the incident matrix of $G$ and $m$ is the number of transitions. This algorithm is useful as the polynomial bound in Corollary 1 depends on the vector $y$.

For an acyclic Petri net $G$ without source transition, we can define $T_1 = \emptyset$ and

$$
P_k := \{ p \in P \setminus \bigcup_{i=1}^{k-1} P_i | \bullet p \subseteq \bigcup_{i=1}^{k-1} T_i \},
$$

$$
T_{k+1} := \{ t \in T \setminus \bigcup_{i=1}^{k-1} T_i | \bullet t \subseteq \bigcup_{i=1}^{k} P_i \}
$$

(13) for $k \in \{1, 2, 3, \ldots \}$ (c.f. Chapter 5 in [24]). We mention the following properties of this partition. Since the number of transitions (or the number of places) is finite, there exists a positive integer $\mu$ (or $\mu'$) such that $T_k \neq \emptyset$ if $1 < k \leq \mu$ and $T_k = \emptyset$ if $k > \mu$ (or $P_k \neq \emptyset$ if $1 \leq k \leq \mu'$ and $P_k = \emptyset$ if $k > \mu'$); therefore, the set of transitions $T$ is partitioned into $\mu - 1$ nonempty sets and the set of places $P$ is partitioned into $\mu'$ sets, where $\mu = \mu'$ or $\mu = \mu' + 1$. For simplicity, we assume that $\mu = \mu'$; if $\mu = \mu' + 1$, it is straightforward to modify our algorithm to accommodate this case. With this partition, we can order rows (or columns) of the incidence matrix $D$ such that places (or transitions) in $P_1$ (or $T_2$) come first and then places (or transitions) in $P_2$ (or $T_3$), and so forth. Now $D$ has the following block structure

$$
D' = \begin{pmatrix}
-F_{1,2} & -F_{1,3} & \cdots & -F_{1,\mu} \\
F_{2,2} & -F_{2,3} & \cdots & -F_{2,\mu} \\
\vdots & \vdots & \ddots & \vdots \\
F_{\mu-1,2} & F_{\mu-1,3} & \cdots & -F_{\mu-1,\mu} \\
F_{\mu,2} & F_{\mu,3} & \cdots & F_{\mu,\mu}
\end{pmatrix}
$$

(14)
where $F_{i,j}$ is a $|P_1| \times |T_1|$ matrix with nonnegative integer entries for $i = 2, \ldots, \mu$ and $j = 1, \ldots, \mu$. These properties and the block structure of $D$ are proved in [24] for the general case without the assumption that there are no source transitions.

One observation is that as long as there is a block $-F(i, j)$ for some $2 \leq i \leq \mu$ and $1 \leq j \leq \mu$, then all the blocks in the right of $-F(i, j)$ have the form $-F(i, j')$ for $j' = j+1, \ldots, \mu$.

Now we give the following algorithm to compute $y$.

**Algorithm 4**

Input: A Petri net $G$ with incidence matrix $D$.

Output: A vector $y$ with positive integer entries such that $y^T D < 0^T_m$.

1. Rearrange the rows and columns of $D$ to obtain $D'$ and keep the mapping of rows.
2. Let $y' = 1^n$.
3. For $j = n, n - 1, \ldots, 1$
   
   If $y'^T D'(:, j) < 0$, go to the next $j$; else, choose the smallest index $i$ such that $D'(i, j) < 0$ and increase the value $y(i)$ to the smallest integer such that $y^T D'(:, j) < 0$.
4. Rearrange the vector $y'$ using the mapping of rows to get the $y$ corresponding to $D$.

In the algorithm, we can always find the smallest index $i$ such that $D'(i, j) < 0$ because there is no source transition in the Petri net. The key idea of this algorithm is that when we increase the $i$-th entry of $y'$ for some $j$, $y'^T D(:, j') < 0$ holds for $j' = j+1, \ldots, n$. If $j$ and $j'$ belongs to the same column block, it is obvious. If $j$ belongs to a column block with smaller index than $j'$, the result holds because in $D'$, $D'(i, j') \leq 0$ as long as $D'(i, j) < 0$ (due to the structure of $D'$). This implies that the vector $y$ obtained from Algorithm 4 indeed satisfies $y^T D < 0^T_m$.

**References**


