Marking Predictability and Prediction in Labeled Petri Nets

Ziyue Ma, *Member, IEEE*, Xiang Yin, *Member, IEEE*, and Zhiwu Li, *Fellow, IEEE*

***Abstract*—This article studies the marking prediction problem in labeled Petri nets. Marking prediction aims to recognize *a priori* that the plant will inevitably reach a given set of *alert markings* in finite future steps. Specifically, we require that a marking prediction procedure should have the following properties: i) no missed alarm, i.e., an alarm can always be issued before reaching an alert marking; and ii) no false alarm, i.e., once an alarm is issued, the plant will eventually reach an alert marking in the future. To this end, the notion of *marking predictability* is proposed as a necessary and sufficient condition for the solvability of the marking prediction problem. A fundamental marking estimation problem in a labeled Petri net is first solved using *minimal explanations* and *basis reachability graphs*. Then, we propose two notions of basis markings called *boundary basis markings* and *basis indicators*, and prove that a plant is predictable with respect to a set of alert markings if all basis markings confusable with boundary basis markings are basis indicators. By properly selecting a set of explicit transitions, the set of basis indicators can be efficiently computed by structural analysis of the corresponding basis reachability graph. Our method has polynomial complexity in the number of basis markings. Finally, we present an effective algorithm for online marking prediction if the plant is predictable.**

***Index Terms*—Discrete event system, marking prediction, Petri net, state estimation.**

I. Introduction

PETRI nets have been proposed as a fundamental model for discrete event systems (DES) in a wide variety of applications. They have been an asset to reduce the computational load of many problems using structural analysis. Labeled Petri nets (LPNs) [1] are extensively used in modeling systems in which sensors are not fully deployed. In an LPN, some transitions are unobservable, i.e., their firings cannot be detected by an external agent, and some transitions are not distinguishable, i.e., an agent may not determine which one has fired among all those transitions sharing the same label. Due to the presence of unobservable and indistinguishable transitions, in general, one cannot determine the exact current marking in an LPN. Instead, one can only infer that an LPN is a set of possible markings called *consistent markings* with respect to a given observation. In the context of LPNs, much work on marking estimation has been done recently; (see, e.g., [1]–[5]). Marking estimation plays a key role in important applications such as *supervisory control* [6]–[8], *fault diagnosis* [9]–[11], *detectability* [12], [13], and *opacity* [14] of a DES.

Marking estimation is to determine the set of possible current markings that a plant may be at. However, in many practical situations, the operator of a plant may expect to know not only what the possible markings at which the plant may be, but also those that to be reached in the future evolution. For example, in an automated production line in which the content in a buffer may exceed a threshold, an alarm is expected to be issued before reaching the threshold such that some preventive measures are properly taken in advance. Therefore, instead of only estimating the current marking of a plant based on the observation history, one may also be interested in *predicting* possible reachable markings of the system, which is referred to as the *marking prediction problem*. The goal of marking prediction is to predict if the plant is going to reach a given set of markings of physical importance, namely *alert markings*, before actually reaching them. On the other hand, *marking predictability* is a property of a plant such that an alarm can always be raised before reaching the set of alert markings.[1]

The prediction problem in DESs has already drawn considerable attention over the past years. Particularly, in [15] and [16], the problem of *event prediction/prognosis* was studied for systems modeled by finite-state automata, where the notion of predictability was proposed. The goal in this framework is to predict the occurrence of some significant event, e.g., a *fault* event, before it actually occurs. Later, this work was extended to *decentralized systems* [17]–[22], *probabilistic automata* [23]–[26], and *robust prognosability* [27]. Several results on *event*

[1] The aim in *prediction problems* is to predict if something bad (e.g., failure) will happen in the future, while in the closely related *diagnosis problems*, the aim is to determine if something bad has occurred in the past.

Manuscript received July 14, 2019; revised February 27, 2020; accepted September 5, 2020. Date of publication September 15, 2020; date of current version July 28, 2021. This work was supported in part by the National Natural Science Foundation of China under Grants 61703321, 61803259, and 61873342, in part by the Shaanxi Provincial Natural Science Foundation under Grant 2019JQ-022, in part by the Fundamental Research Funds for the Central Universities under Grant JB190407, in part by the Shanghai Jiao Tong University Scientific Research fund, and in part by the Technological Innovation Funds, and the Science and Technology Development Fund under Grant 0012/2019/A1. Recommended by Associate Editor J. Komenda. *(Corresponding author: Ziyue Ma.)*

Ziyue Ma is with the School of Electro-Mechanical Engineering, Xidian University, Xi'an 710071, China (e-mail: maziyue@xidian.edu.cn).

Xiang Yin is with the Department of Automation, and the Key Laboratory of System Control and Information Processing, Shanghai Jiao Tong University, Shanghai 201108, China (e-mail: yinxiang@sjtu.edu.cn).

Zhiwu Li is with the School of Electro-Mechanical Engineering, Xidian University, Xi'an 710071, China, and also with Institute of Systems Engineering, Macau University of Science and Technology, Taipa 999078, Macau (e-mail: zhwli@xidian.edu.cn).

Digital Object Identifier 10.1109/TAC.2020.3024270

prognosability verification have been achieved using *model abstraction* [28] and *set-membership approach* [29].

The event prediction problem has also been studied for systems modeled by Petri nets more recently [30]–[35]. In [30], the set of all possible continuations of an observation with a limited size is worked out and used to decide whether a fault transition will or will not fire. In [31], event predictability is proved to be EXPSPACE-complete for unbounded Petri nets by reducing it to a Petri net model checking problem; the result has been further extended to the case of decentralized systems [32]. In [33], an algorithm is proposed to verify fault predictability in labeled Petri nets using *predictor graphs*. Moreover, the event prognosis problem has also been studied in stochastic Petri nets [34], [35].

The framework of event-based prediction can also be applied to the state or language prediction problem by refining the state-space of a plant automaton [17]. However, in Petri nets, methods for event prediction are not applicable for marking prediction due to the following reasons. First, the fact that the firing of a transition at one marking yields an alert marking does not necessarily mean that the firing of such a transition at any marking *always* leads to alert markings. Therefore, one cannot simply convert the marking prediction problem to an event prediction problem based on the original net. Such a transformation is only possible if the entire reachability graph of the system is constructed. However, the reachability graph is, in general, extremely large even when the plant net is bounded. For Petri nets, it is desirable to investigate the marking prediction problem in a more efficient way using structural analysis techniques.

In this article, we formulate and systematically investigate the marking prediction problem in LPNs. Specifically, we assume that the set of alert markings to be predicted is given by a linear constraint represented by a *generalized mutual exclusion constraint* (GMEC) [36], [37]. Given a plant LPN with a set of alert markings, our goal is to design a correct predictor in the sense that: i) "no missed alarm," i.e., an alarm can always be issued before the plant actually reaches an alert marking; ii) "no false alarm," i.e., the plant will eventually reach an alert marking within a finite number of steps once an alarm is issued. We propose a new notion called *marking predictability* as the necessary and sufficient condition for the existence of a predictor satisfying these two criteria.

To efficiently verify marking predictability and to solve the online prediction problem without constructing the entire reachability graph and using the automata-based approach, we develop a method based on the *basis reachability graphs* (BRGs), which have been proved to be an efficient tool for abstracting the state-space of a Petri net [14], [38], [39]. Particularly, one advantage of the BRG technique is that only part of the reachability space, namely *basis markings*, is enumerated; all other markings reachable from them by firing only *implicit transitions* can be characterized by a linear algebraic system. We prove that by choosing the set of explicit transitions as a superset of observable transitions, the set of consistent markings can always be represented by the union of a set of consistent basis markings.

We then introduce two important notions called *boundary basis markings* and *basis indicators* that are useful in marking prediction. A basis marking is called a basis indicator if from it some alert markings are necessarily reached after finite number of firings. The two notions extend the existing concepts of *boundary states* and *indicator states* (see, e.g., [17]) from the original state-space of a system to the abstracted basis marking space. We provide a characterization of marking predictability with respect to a set of alert markings in terms of the confusability of boundary basis markings and basis indicators. On the other hand, we show that, in general, it is difficult to compute basis indicators in an arbitrary BRG that can be used for marking estimation. This observation is quite different from other works: in many problems in Petri nets such as *fault diagnosis* [11], [39] and *opacity* [14], in which certain properties can be easily verified by inspecting those basis markings in a BRG that can be used for marking estimation. To overcome this problem, an additional condition for the selection of explicit transitions is proposed under which the set of basis indicators can be efficiently computed by structural analysis of the corresponding BRG. The complexity of the proposed approach for verifying marking predictability is polynomial in the number of basis markings. Since the number of basis markings is generally much smaller than the number of reachable markings [11], [38], the proposed method for marking predictability verification in LPNs is of efficiency. Finally, if a plant is predictable, a recursive algorithm is developed for online marking prediction based on basis markings.

This article is organized in nine sections. Basic notions of Petri nets and BRGs are recalled in Section II. In Section III, the problem of marking predictability in LPNs is formulated. In Section IV, a marking estimation method using BRGs is proposed. In Section V, the notions of boundary basis markings and basis indicators are introduced, and a sufficient and necessary condition for marking predictability in LPNs is developed. In Sections VI and VII, a condition for selecting explicit transitions is proposed, and a method to compute the set of basis indicators is developed, which is based on the structural analysis of a BRG. In Section VIII, an algorithm for online marking prediction is presented. Section IX draws the conclusion.

II. Preliminaries

A. Petri Nets

A Petri net is a four-tuple $N = (P, T, Pre, Post)$, where P is a set of m *places* represented by circles; T is a set of n *transitions* represented by bars; $Pre : P \times T \to \mathbb{N}$ and $Post : P \times T \to \mathbb{N}$ are the *pre-* and *post-incidence functions*, respectively, specifying the arcs in the net and can also be represented as matrices in $\mathbb{N}^{m\times n}$ (here $\mathbb{N} = \{0, 1, 2, \ldots\}$). The *incidence matrix* of a net is defined by $C = Post - Pre \in \mathbb{Z}^{m\times n}$ (here $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$).

For a transition $t \in T$, we define its *set of input places* as ${}^\bullet t = \{p \in P \mid Pre(p, t) > 0\}$ and its *set of output places* as $t^\bullet = \{p \in P \mid Post(p, t) > 0\}$. The notions for ${}^\bullet p$ and $p^\bullet$ are analogously defined.

A *marking* is a function $M : P \to \mathbb{N}$ that assigns each place of a Petri net a non-negative integer number of tokens, represented by black dots; a marking can also be represented as an m-component vector. We denote by $M(p)$ the number of tokens in place p at marking M. A *marked net* $G = \langle N, M_0\rangle$ is a net N with an initial marking M_0. Marking $[x_1, \ldots, x_m]^T$ is also denoted as $x_1p_1 + \cdots + x_mp_m$ for simplicity.

A transition t is said to be *enabled* at a marking M if $M \geq Pre(\cdot, t)$. If t is enabled at M, then it may fire and reach a new marking $M' = M_0 + C(\cdot, t)$. We write $M[t\rangle$ and $M[t\rangle M'$ to denote, respectively, that transition t is enabled at M and its occurrence yields M'. We denote by T^* the set of all finite sequences of transitions over T. Then, $M[\sigma\rangle M'$ analogously denotes that a sequence $\sigma = t_1t_2 \ldots t_k \in T^*$ is enabled (sequentially) at M and its occurrence finally yields M'. In this case, we say that M' is *reachable* from M. We denote by $R(N, M_0)$ the set of all markings reachable from the initial marking M_0. The *language* of $\langle N, M_0\rangle$ is defined as $L(N, M_0) = \{\sigma \in T^* \mid M_0[\sigma\rangle\}$. For any sequence $\sigma \in T^*$, $\mathbf{y}_\sigma$ denotes its *Parikh vector*, i.e., $y_\sigma(t) = k$ if transition t occurs k times in σ.

Given a sequence $\sigma \in T^*$, the *prefix-closure* of σ is defined as $Pr(\sigma) = \{\sigma' \in T^* \mid (\exists\sigma'' \in T^*)\sigma = \sigma'\sigma''\}$. A sequence $\bar\sigma \in T^*$ is said to be a *strict prefix* of σ if $\bar\sigma \in Pr(\sigma)$ and $\bar\sigma \neq \sigma$.

Given a Petri net $N = (P, T, Pre, Post)$, net $\hat N = (\hat P, \hat T, \hat{Pre}, \hat{Post})$ is a *subnet* of N if $\hat P \subseteq P$, $\hat T \subseteq T$ and $\hat{Pre}$ (resp., $\hat{Post}$) is the restriction of Pre (resp., $Post$) to $\hat P \times \hat T$. In particular, $\hat N$ is called the *$\hat T$-induced subnet* if $\hat N = (P, \hat T, \hat{Pre}, \hat{Post})$.

B. Labeled Petri Nets

A *labeled Petri net system*, or *labeled Petri net* (LPN) for the sake of simplicity, is a 4-tuple $G = (N, M_0, E, \ell)$, where $\langle N, M_0\rangle$ is a marked net, E is the *alphabet* (a set of labels), and $\ell : T \to E \cup \{\varepsilon\}$ is the *labeling function* that assigns each transition $t \in T$ either a symbol from E or the *silent* label ε. This naturally leads to a partition of the transition set as $T = T_o \dot\cup T_{uo}$, where $T_o = \{t \in T \mid \ell(t) \in E\}$ is the set of *observable* transitions and $T_{uo} = T \setminus T_o = \{t \in T \mid \ell(t) = \varepsilon\}$ is the set of *unobservable* transitions.

The labeling function can be extended to $\ell : T^* \to E^*$ recursively by i) $\ell(\varepsilon) = \varepsilon$ and ii) $\ell(\sigma t) = \ell(\sigma)\ell(t)$ with $\sigma \in T^*$ and $t \in T$. When a sequence $\sigma \in T^*$ fires, the *observation* of σ is $w = \ell(\sigma) \in E^*$. The *inverse projection* of an observation $w \in E^*$ with respect to $G = (N, M_0, E, \ell)$ is defined as $\ell^{-1}(w) = \{\sigma \in L(N, M_0) \mid \ell(\sigma) = w\}$. The *language* of an LPN $G = (N, M_0, E, \ell)$ is defined as $L(G) = \{\ell(\sigma) \mid \sigma \in L(N, M_0)\}$.

Given an LPN $G = (N, M_0, E, \ell)$, for an observation $w \in E^*$, we write $M_1[w\rangle M_2$ if there exists a sequence $\sigma \in T^*$ such that $\ell(\sigma) = w$ and $M_1[\sigma\rangle M_2$. Then, we define $\mathcal{C}(w)$ as the *set of consistent markings* of an observation $w \in L(G)$, i.e., $\mathcal{C}(w) = \{M \in R(N, M_0) \mid M_0[w\rangle M\}$.

Given an LPN $G = (N, M_0, E, \ell)$, net $\hat G = (\hat N, \hat M_0, E, \hat\ell)$ is said to be a *subnet* of G if $\hat N$ is a subnet of N, and $\hat M_0$ and $\hat\ell$ are M_0 and ℓ restricted to $\hat N$, respectively.

C. Basis Reachability Graph

Definition 2.1: [38] Given a Petri net $N = (P, T, Pre, Post)$, a pair $\pi = (T_E, T_I)$ is called a *basis partition* of T if i) $T_I \subseteq T$, $T_E = T \setminus T_I$ and ii) the T_I-induced subnet is acyclic. The sets T_E and T_I are called the set of *explicit transitions* and the set of *implicit transitions*, respectively. □

Definition 2.2: Given a Petri net $N = (P, T, Pre, Post)$, a basis partition $\pi = (T_E, T_I)$, a marking M, and a transition $t \in T_E$, we define

$$\Sigma(M, t) = \{\sigma \in T_I^* \mid M[\sigma\rangle M', M' \geq Pre(\cdot, t)\}$$

as the set of *explanations* of t at M, and we define

$$Y(M, t) = \{\mathbf{y}_\sigma \in \mathbb{N}^{|T_I|} \mid \sigma \in \Sigma(M, t)\}$$

as the set of *explanation vectors*. □

Definition 2.3: Given a Petri net $N = (P, T, Pre, Post)$, a basis partition $\pi = (T_E, T_I)$, a marking M, and a transition $t \in T_E$, we define

$$\Sigma_{\min}(M, t) = \{\sigma \in \Sigma(M, t) \mid \nexists\sigma' \in \Sigma(M, t) : \mathbf{y}_{\sigma'} \lneq \mathbf{y}_\sigma\}$$

as the set of *minimal explanations* of t at M, and we define

$$Y_{\min}(M, t) = \{\mathbf{y}_\sigma \in \mathbb{N}^{|T_I|} \mid \sigma \in \Sigma_{\min}(M, t)\}$$

as the corresponding set of *minimal explanation vectors*. □

Definition 2.4: Given a Petri net $N = (P, T, Pre, Post)$ with an initial marking M_0 and a basis partition $\pi = (T_E, T_I)$, its *basis marking set* $\mathcal{M}$ is recursively defined as follows:

1) $M_0 \in \mathcal{M}$;
2) If $M \in \mathcal{M}$, then $\forall t \in T_E$, $\forall \mathbf{y} \in Y_{\min}(M, t)$, we have

$$M + C_I \cdot \mathbf{y} + C(\cdot, t) \in \mathcal{M}$$

where C_I is the incidence matrix of the T_I-induced subnet. A marking M in $\mathcal{M}$ is called a *basis marking* (with respect to basis partition $\pi = (T_E, T_I)$). □

The BRG of a marked net $\langle N, M_0\rangle$, with respect to basis partition $\pi = (T_E, T_I)$, is a deterministic finite state automaton (DFA) $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ [38], where

1) the state set $\mathcal{M}$ is the set of *basis markings*;
2) the event set $Tr = T_E \times \mathbb{N}^{|T_I|}$ is the set of pairs in the form of $(t, \mathbf{y}) \in T_E \times \mathbb{N}^{|T_I|}$;
3) the transition relation Δ is

$$\Delta = \{(M_1, (t, \mathbf{y}), M_2) \mid t \in T_E, \mathbf{y} \in Y_{\min}(M_1, t)$$
$$M_2 = M_1 + C_I \cdot \mathbf{y} + C(\cdot, t)\}$$

4) the initial state is the initial marking M_0.

For the convenience of presentation, in the sequel of this article, we use $\phi = (t_{i_1}, \mathbf{y}_{i_1})(t_{i_2}, \mathbf{y}_{i_2}) \cdots (t_{i_n}, \mathbf{y}_{i_n})$ to denote a sequence of labels of arcs in a BRG, i.e., a path $M_{b,1} \to M_{b,2} \to \cdots \to M_{b,n}$, where the arcs on this path are sequentially labeled by $(t_{i_1}, \mathbf{y}_{i_1}), (t_{i_2}, \mathbf{y}_{i_2}), \ldots, (t_{i_n}, \mathbf{y}_{i_n})$. Let $\ell(\phi) = \ell(t_{i_1}t_{i_2} \cdots t_{i_n})$ and $\phi_{\uparrow T_E} = t_{i_1}t_{i_2} \cdots t_{i_n}$ ($\uparrow$ is the *natural projection* operator). The prefix-closure of ϕ is denoted as $Pr(\phi)$.

Definition 2.5: [38] Given a net $G = \langle N, M_0\rangle$ with $\pi = (T_E, T_I)$, the *implicit reach* of a marking M is a set of markings: $R_I(M) = \{M' \mid M[\sigma\rangle M', \sigma \in T_I^*\}$. □

Proposition 2.1: [38] Given a Petri net $N = (P, T, Pre, Post)$, let $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ be the BRG with respect to $\pi = (T_E, T_I)$. The following two statements are equivalent:

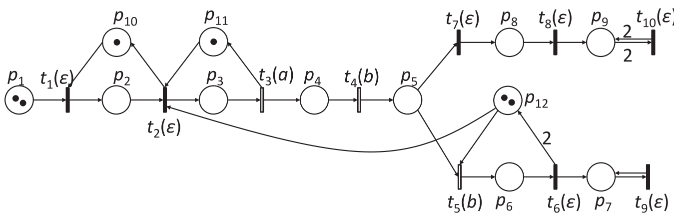

Fig. 1. Labeled Petri net.

1) there exist a marking M and a firing sequence in the form of $\sigma = \sigma_1 t_{i_1} \cdots \sigma_n t_{i_n} \sigma_{n+1}$, where $\sigma_j \in T_I^*, t_{i_j} \in T_E$ for all $j \in \{1, \ldots, n+1\}$, such that $M_0[\sigma\rangle M$;
2) there is a following path in the BRG $\mathcal{B}$:

$$M_0 \xrightarrow{(t_{i_1}, \mathbf{y}_1)} M_{b,1} \xrightarrow{(t_{i_2}, \mathbf{y}_2)} \cdots \xrightarrow{(t_{i_n}, \mathbf{y}_n)} M_{b,n}$$

such that $M \in R_I(M_{b,n})$.

III. Marking Prediction Problem Formulation

In this section, we formulate the marking prediction problem and propose the notion of marking predictability as the necessary and sufficient condition for the solvability of this problem.

Before proceeding formally, we first present an example to illustrate the notion of marking prediction in LPNs.

Example 3.1: Consider the LPN in Fig. 1 in which $\ell(t_3) = a, \ell(t_4) = \ell(t_5) = b$ while the labels of all other transitions are the silent label ε. Suppose that the set of alert markings is $S_1 = \{M \mid M(p_8) + M(p_9) \geq 2\}$, i.e., the operator of the plant needs to be pre-alerted if the plant will inevitably reach a marking at which places p_8 and p_9 hold at least two tokens. By inspection, if we observe aab or aba, we can conclude that one of the two tokens originally in p_1 is now in p_4 while the other is in either p_5, p_8, or p_9. Since transition t_5 is disabled, the two tokens will inevitably reach p_9. Hence, by observing aab or aba, we can correctly issue an alarm to warn that the plant will inevitably reach a marking in S.

Let us consider another set of alert markings $S' = \{M \mid M(p_7) \geq 2\}$. In this case, we cannot be pre-alerted before the plant reaching S'. Only by observing $abbabb$, we can confirm that the two tokens originally in p_1 will inevitably go to place p_7. However, the markings consistent with $abbabb$ are $p_6 + p_7 + p_{10} + p_{11}$ and $2p_7 + p_{10} + p_{11} + 2p_{12}$, which indicates that when an alarm is issued after observing $abbabb$, the plant may have already reached the marking $2p_7 + p_{10} + p_{11} + 2p_{12} \in S'$. Hence, in this case, G is not predictable with respect to S'. □

In the sequel of this article, we assume that a plant LPN $G = (N, M_0, E, \ell)$ with $N = (P, T, Pre, Post)$ satisfies the following two assumptions:

1) A1: G is deadlock-free;
2) A2: G is bounded.

Assumption A1 on deadlock-freeness is a common assumption in the analysis of partially observed Petri nets [11], [39]. This assumption will be used when computing the so-called *basis indicators* (which will be introduced shortly) in Sections VI and VIII. Assumption A2 guarantees that any BRG of a plant is bounded [38] regardless of the basis partition. On the other hand, here we do not require the acyclicity of the unobservable subnet, which is often needed in the literature.

In this article, we also assume that the set of *alert markings* to be predicted, denoted as $S \subseteq \mathbb{N}^{|P|}$, is represented by a *generalized mutual exclusion constraint* (GMEC).

Definition 3.1: [36] A *generalized mutual exclusion constraint* is a pair $(\mathbf{w}, k)$, where $\mathbf{w} \in \mathbb{Z}^m$ and $k \in \mathbb{Z}$, that defines a set of markings

$$\mathcal{L}_{(\mathbf{w},k)} = \{M \in \mathbb{N}^m \mid \mathbf{w}^T \cdot M \leq k\}.$$

The *token count* of GMEC $(\mathbf{w}, k)$ at a marking M is the value of $\mathbf{w}^T \cdot M$. □

Remark 1: Set S defined by a GMEC is in the "$\leq$" form. If S is given in the "$\mathbf{w}^T \cdot M \geq k$" form, e.g., as in Example 3.1, it can be equivalently converted to $-\mathbf{w}^T \cdot M \leq -k$ and is associated to GMEC $(-\mathbf{w}, -k)$. □

Remark 2: We assume that S is defined by a single GMEC $(\mathbf{w}, k)$ to make this article concise. Our approach can be straightforwardly generalized to cases where S is defined by a conjunction of multiple GMECs. □

To simplify the notation, we define the *alert language* of a marking M with respect to a given alert set S, which consists of all such sequences σ's, each of which has at least one prefix that reaches S. Precisely speaking:

$$L_{M,S} = \{\sigma \in L(N, M) \mid \exists \bar{\sigma} \in Pr(\sigma) : M[\bar{\sigma}\rangle M' \in S\}.$$

Note that this definition does not mean that the marking reached by the firing σ is in S.[2] Following this notation, the alert language of an initial marking M_0 is denoted by $L_{M_0,S}$.

Given a plant LPN G and a set of alert markings S, a *predictor* is a mechanism that issues prediction alarms for reaching set S based on the observation history. Formally, a predictor is defined as a function

$$\mathcal{A} : L(G) \rightarrow \{0, 1\}$$

where "1" means that an alarm is issued and "0" means that no alarm is issued. We say that a predictor $\mathcal{A}$ is *correct* if it satisfies the following two criteria.

1) "No missed alarm": an alarm is always issued *before* the plant actually reaching S. Precisely speaking, for all $\sigma \in L_{M_0,S}$ whose observation is $w = \ell(\sigma)$, there exists $\bar{w} \in Pr(w), \bar{w} \neq w$ such that $\mathcal{A}(\bar{w}) = 1$.
2) "No false alarm": once an alarm is issued, the plant will *eventually* reach S by firing a finite number of transitions. Precisely speaking, if $\mathcal{A}(w) = 1$ for $w \in L(G)$, then there exists $K_w \in \mathbb{N}$ such that for all $\sigma \in \ell^{-1}(w)$, for all $\sigma' \in T^*$ such that $|\sigma'| \geq K_w$ and $\sigma\sigma' \in L(N, M_0)$, $\sigma\sigma' \in L_{M_0,S}$ holds.

In this article, we focus on solving the following two problems related to marking prediction:

[2] We point out that set $L_{M,S}$ is *right-closed*, i.e., $\sigma \in L_{M,S}$ implies that $\sigma\sigma' \in L_{M,S}$ for any valid continuation σ' of σ. Such a property brings mathematical convenience in the subsequential deductions in this article.

1) under what conditions, there exists a correct predictor satisfying the above two criteria;
2) if so, how to design such a predictor and implement it online.

First, we propose the notion of marking predictability in LPNs as the existence condition for a predictor as follows.

Definition 3.2 (Marking predictability): Given an LPN $G = (N, E, \ell, M_0)$ with $N = (P, T, Pre, Post)$ and a set of alert markings S, plant G is said to be *predictable* with respect to S if for all $\sigma \in L_{M_0,S}, w = \ell(\sigma)$, there exist $\bar{w} \in Pr(w), \bar{w} \neq w$ and $K_w \in \mathbb{N} \setminus \{0\}$ such that the following condition is satisfied:

$$\forall \sigma' \in \ell^{-1}(\bar{w}) \setminus L_{M_0,S}, \forall \sigma'' \in T^*, \sigma'\sigma'' \in L(N, M_0) : \quad |\sigma''| \geq K_w \Rightarrow \sigma'\sigma'' \in L_{M_0,S}. \tag{1}$$

□

In plain words, an LPN G is said to be predictable with respect to S if for all sequences σ's such that $M_0[\sigma\rangle M \in S$,[3] then there always exists a strict prefix of the observation of σ, i.e., $\bar{w} \in Pr(\ell(\sigma))$ and $\bar{w} \neq \ell(\sigma)$, such that 1) for all σ' that looks like $\bar{w}$, and 2) for all $\sigma'' \in T^*$ that is a valid continuation of σ', and the length of σ'' is greater than or equal to a certain bound K_w, trajectory $M_0[\sigma'\sigma''\rangle$ necessarily passes S.

We will show next that Definition 3.2 indeed provides the necessary and sufficient condition for the existence of a correct predictor. Before doing so, we first introduce some necessary notations. Given a marking M, we use $d_{\max}(M)$ to denote the maximal length of sequences firable at M and do not reach S, i.e.,

$$d_{\max}(M) = \max_{\sigma \in L(N,M) \setminus L_{M,S}} |\sigma|.$$

Clearly, $d_{\max}(M) = \infty$ if and only if there exists an arbitrarily long sequence σ firable at M, which does not belong to $L_{M,S}$. We denote by

$$H = \{M \in R(N, M_0) \mid d_{\max}(M) \neq \infty\}$$

the set of all reachable markings that will inevitably reach S. Since the plant G is bounded, set H is finite. Now, we are ready to present the main result of this section.

Theorem 3.1: There exists a correct predictor $\mathcal{A}$ if and only if G is predictable with respect to S.

Proof: (If) If G is predictable with respect to S, then we can construct the predictor $\mathcal{A}$ as the following:

$$(\forall w \in L(G))\ \mathcal{A}(w) = 1 \quad \Leftrightarrow \quad \mathcal{C}(w) \subseteq H.$$

By the definition of set H, clearly $\mathcal{A}$ satisfies the "no false alarm" criterion, since whenever an alarm is issued, S will be reached by firing at most $\max_{M \in H} d_{\max}(M)$ transitions. Now we prove that $\mathcal{A}$ also satisfies "no missed alarm" criterion by contradiction. Suppose that there exists a sequence $\sigma \in L_{M_0,S}$ with $\ell(\sigma) = w$ such that for all $\tilde{w} \in Pr(w), \tilde{w} \neq w$, $\mathcal{C}(\tilde{w}) \nsubseteq H$. Since G is predictable with respect to S, there exist $\bar{w} \in Pr(w), \bar{w} \neq w$ and $K_w > 0$ such that for all $\sigma' \in \ell^{-1}(\bar{w}) \setminus L_{M_0,S}$, for all $\sigma'\sigma'' \in L(N, M_0)$, $|\sigma''| \geq K_w \Rightarrow \sigma'\sigma'' \in L_{M_0,S}$. This indicates that for all markings $M \in \mathcal{C}(\bar{w})$, M will inevitably reach S by firing at most K_w transitions, implying that $\mathcal{C}(\bar{w}) \subseteq H$ and $\mathcal{A}(\bar{w}) = 1$. This contradicts the assumption $\mathcal{C}(\tilde{w}) \nsubseteq H$ for all $\tilde{w} \in Pr(w), \tilde{w} \neq w$.

(Only If) If G is not predictable with respect to S, there exists a sequence $\sigma \in L_{M_0,S}$ such that $M_0[\sigma\rangle M \in S$ and σ does not satisfy the condition in Definition 3.2. To guarantee "no missed alarm," a predictor must issue an alarm ($\mathcal{A}(\bar{w}) = 1$) for at least one observation $\bar{w}$ that is a strict prefix of $\ell(\sigma)$. However, since σ does not satisfy the condition in Definition 3.2, for any $\bar{w} \in Pr(\ell(\sigma)), \bar{w} \neq \ell(\sigma)$, there necessarily exist a sequence $\sigma' \in \ell^{-1}(\bar{w}) \setminus L_{M_0,S}$ and an infinite long sequence σ'' such that $\sigma'\sigma'' \in L(N, M_0)$ and $\sigma'\sigma'' \notin L_{M_0,S}$. This indicates that the alarm $\mathcal{A}(\bar{w}) = 1$ may be a false alarm. Hence, a correct predictor does not exist. ■

Then, the marking predictability verification problem is formulated as follows.

Problem 1: Given a plant LPN $G = (N, E, \ell, M_0)$ with $N = (P, T, Pre, Post)$ and a set of alert markings $S = \mathcal{L}_{(\mathbf{w},k)}$, determine if G is predictable with respect to S.

Remark 3: As we have mentioned in the introductory section, since the plant LPN G is bounded, the condition in Definition 3.2 can be verified by using the *reachability analysis*. Specifically, one can construct the entire reachability graph, and then reduce the problem to an event prediction in the reachability graph, which can be solved using automaton-based methods (e.g., [16]). However, this approach is very exhaustive as the size of the reachability graph of a net is generally extremely large even if the net is bounded. On the other hand, BRGs have been proved to be an efficient tool to abstract the state-space of Petri nets [11], [14], [38]–[40]. Therefore, this article aims to develop a method to efficiently solve Problem 1 by leveraging the structural property of the Petri net without computing the entire reachability graph.

The purpose of marking prediction is to raise an alarm to inform an operator that a plant will inevitably reach a given set of markings in the future. To correctly issue an alarm for an observation, the predictor necessarily has the full knowledge of the current consistent markings. Hence, in the next section, we study the marking estimation problem in LPNs and develop a method to perform marking estimation using BRGs, which is a fundamental step to establish our method for marking prediction exposed in the rest of this article.

IV. Marking Estimation in LPNs Using Basis Reachability Graphs

In the literature, BRGs, as defined in Section II, have been used in various contexts to abstract the state-space of a plant net [11], [14], [38]–[40]. A useful property of BRG is as follows: if the *unobservable subnet* is acyclic, the set of consistent markings can be represented as the union of unobservable reach of the consistent basis markings defined as follows.

Definition 4.1: Given an LPN $G = (N, M_0, E, \ell)$ with $N = (P, T, Pre, Post)$, let $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ be the BRG with respect to $\pi = (T_E, T_I)$. The set of *consistent basis markings* of an observation $w \in L(G)$, denoted as $\mathcal{M}(w)$, is the set of

[3]Although $L_{M_0,S}$ consists of all sequences σ's whose firing at M_0 pass but may not eventually yield a marking in S, for the purpose of marking prediction, we only need to consider σ's such that $M_0[\sigma\rangle M \in S$.

basis markings M_i's such that there exists a path labeled by $(t_{i_1}, \mathbf{y}_{i_1}), \ldots, (t_{i_n}, \mathbf{y}_{i_n})$ from M_0 to M_i and $\ell(t_{i_1} \cdots t_{i_n}) = w$. □

Proposition 4.1: [40] Given an LPN $G = (N, M_0, E, \ell)$ with $N = (P, T, Pre, Post)$, let $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ be the BRG with respect to $\pi = (T_o, T_{uo})$. For an observation $w \in T_o^*$, it holds

$$\begin{aligned} \mathcal{C}(w) &= \bigcup_{M_b \in \mathcal{M}(w)} R_{uo}(M_b) \\ &= \bigcup_{M_b \in \mathcal{M}(w)} \{M \mid M = M_b + C_{uo} \cdot \mathbf{y}\}. \end{aligned} \tag{2}$$

Proposition 4.1 provides an efficient way for computing consistent markings if the unobservable subnet is acyclic. Specifically, we can compute consistent basis markings $\mathcal{M}(w)$ of an observation w recursively based on $\mathcal{M}(\bar{w})$, where $\bar{w}$ is a prefix of w on a BRG, staring from the empty observation ε. However, if the unobservable subnet is not acyclic, then partition $\pi = (T_o, T_{uo})$ is not a valid basis partition, and Proposition 4.1 cannot be directly used. Moreover, since the basis partition of an LPN is in general not unique [38], an LPN may have various different BRGs depending on different basis partitions. In the following, we show that for the purpose of marking estimation, it is sufficient and necessary[4] to select a basis partition $\pi = (T_E, T_I)$ such that $T_E \supseteq T_o$, i.e., all observable transitions are explicit.

Condition 1: The basis partition $\pi = (T_E, T_I)$ satisfies $T_E \supseteq T_o$.

If Condition 1 is satisfied, the set of consistent markings $\mathcal{C}(w)$ can be described by a set of linear algebraic equations based on the consistent basis markings, as stated in the following Proposition 4.2 that is generalized from the results in [40].

Proposition 4.2: Given an LPN $G = (N, M_0, E, \ell)$ with $N = (P, T, Pre, Post)$, let $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ be the BRG with respect to $\pi = (T_E, T_I)$ where $T_E \supseteq T_o$. For an observation $w \in L(G)$, it holds

$$\begin{aligned} \mathcal{C}(w) &= \bigcup_{M_b \in \mathcal{M}(w)} R_I(M_b) \\ &= \bigcup_{M_b \in \mathcal{M}(w)} \{M \mid M = M_b + C_I \cdot \mathbf{y}\}. \end{aligned} \tag{3}$$

Proof: This theorem follows the same arguments in the proof of Theorems 4.9 and 4.13 in [40], by treating the *implicit transitions* here as the *unobservable transitions*. ■

A BRG, according to its definition, is a deterministic automaton in which $(t, \mathbf{y})$'s on arcs carry information of minimal explanations which are used for purposes such as *reachability* [38] and *diagnosability analysis* [41]. On the other hand, one may have noticed that to compute $\mathcal{M}(w)$ using Definition 4.1 does not require the knowledge of $\mathbf{y}$ in $(t, \mathbf{y})$ but only the transition t and its label $\ell(t)$.

Now we are ready to present Algorithm 1 that performs the marking estimation using BRGs. Algorithm 1 iteratively computes the set of basis markings consistent with the current observation. The offline stage of Algorithm 1 requires to compute a BRG whose structural complexity is $|\mathcal{M}|$. In general, there is no indicator for the size of $\mathcal{M}$ as far as we know—$|R(N, M_0)|$ is the only known upper bound. However, numerical results in [14], [38] show that, in general, $|\mathcal{M}| \ll |R(N, M_0)|$ holds, i.e., the size of a BRG is much smaller than the corresponding reachability graph. On the other hand, the online stages of Algorithm 1 is a one-step-look-ahead procedure whose complexity is negligible. We use Example 4.1 to illustrate it.

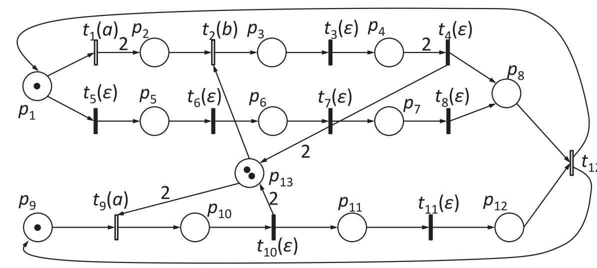

Fig. 2. Labeled Petri net used in Example 4.1.

TABLE I
LIST OF BASIS MARKINGS IN FIG. 2 WITH $T_E = \{t_1, t_2, t_9, t_{12}\}$

M_0	[1 0 0 0 0 0 0 0 1 0 0 0 2]	M_5	[0 0 2 0 0 0 0 0 1 0 0 0 0]
M_1	[0 2 0 0 0 0 0 0 1 0 0 0 2]	M_6	[0 1 1 0 0 0 0 0 0 0 1 0 1]
M_2	[1 0 0 0 0 0 0 0 0 1 0 0 0]	M_7	[0 0 0 0 0 0 0 1 0 1 0 0 0]
M_3	[0 1 1 0 0 0 0 0 1 0 0 0 1]	M_8	[0 0 2 0 0 0 0 0 0 0 1 0 0]
M_4	[0 2 0 0 0 0 0 0 0 1 0 0 0]		

Algorithm 1: BRG-Based Marking Estimation.

Input: An LPN $G = (N, M_0, E, \ell)$

Offline Stage:

1: Find a basis partition $\pi = (T_E, T_I)$ with $T_E \supseteq T_o$;
2: Compute BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$;

Online Stage:

3: Let $w = \varepsilon$;
4: Compute $\mathcal{M}(w)$;
5: Let $\mathcal{C}(w) = \bigcup_{M_b \in \mathcal{M}(w)} R_I(M_b)$;
6: Wait until some event $e \in E$ occurs;
7: Updated the observation as $w := we$, go to Step 4.

Example 4.1: Consider the LPN in Fig. 2 that represents an assembly line. In this LPN, $T_o = \{t_1, t_2, t_9, t_{12}\}$, where $\ell(t_1) = \ell(t_9) = a$, $\ell(t_2) = b$, $\ell(t_{12}) = c$, and $\ell(t) = \varepsilon$ for all $t \in T_{uo}$. This net has 39 reachable markings. Since the T_{uo}-induced subnet is acyclic, partition $\pi = (T_E, T_I)$ with $T_E = T_o = \{t_1, t_2, t_9, t_{12}\}$ and $T_I = T_{uo} = \{t_3, t_4, t_5, t_6, t_7, t_8, t_{10}, t_{11}\}$ is a valid basis partition that satisfies Condition 1. The corresponding BRG is depicted in Fig. 3, which contains nine basis markings listed in Table I.

Suppose that $w = aca$ is observed. By sequentially computing consistent basis markings of ε, a, ac, and aca, we have the set $\mathcal{M}(w)$

$$\{M_0\} \xrightarrow{a} \{M_1, M_2\} \xrightarrow{c} \{M_0\} \xrightarrow{a} \{M_1, M_2\}.$$

Hence, $\mathcal{C}(aca) = R_I(M_1) \cup R_I(M_2)$. □

[4]For marking prediction, condition $T_E \supseteq T_o$ is only necessary but not sufficient, as will be shown in Section VI.

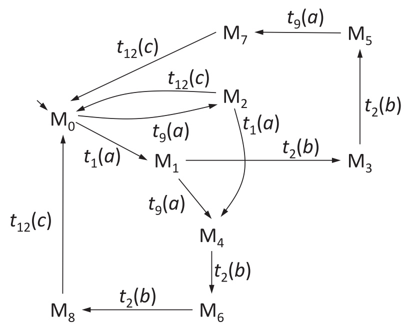

Fig. 3. BRG of the net in Fig. 2 with $T_E = \{t_1, t_2, t_9, t_{12}\}$.

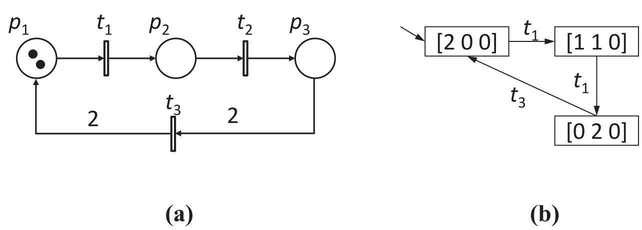

Fig. 4. Petri net used in Example 4.2.

Remark 4: A marking estimation algorithm using BRGs was also proposed in our previous work [42] in unlabeled Petri nets with unobservable transitions. In [42], some unobservable transitions were chosen as *pseudo-observable transitions*, which are in fact those in T_E/T_{uo}. However, the method in [42] is primitive and much less efficient, which requires to compute multiple BRGs online. □

Proposition 4.2 shows that Condition 1 is a sufficient condition to develop a marking estimation method in LPNs using BRGs. Now we show that this condition is also necessary, i.e., a BRG with $T_E \not\supseteq T_o$ may not be used to perform marking estimation as the equation in Proposition 4.2 no longer holds due to the overabstraction of the firings of observable transitions.

Fact 1: If $T_E \not\supseteq T_o$, then $\mathcal{C}(w) \subseteq \bigcup_{M\in\mathcal{M}(w)} R_I(M)$, where the inclusion may be proper (i.e., "$\subsetneq$" may hold).

Example 4.2: Consider the net on the left of Fig. 4 where $T_o = \{t_1, t_2, t_3\}$ and each transition is labeled itself, i.e, $\ell(t_i) = t_i$. For a basis partition $T_E = \{t_1, t_3\}$ and $T_I = \{t_2\}$, the corresponding BRG consists of three basis markings shown on the right. For observation t_1t_2, clearly, $\mathcal{C}(t_1t_2)$ consists of a unique marking $M = [1, 0, 1]^T$, and hence t_1 should be permitted to fire. However, marking M is not a consistent basis marking in the BRG. If we consider $w_{\uparrow T_E} = t_1$, by the fact that $\mathcal{M}(w_{\uparrow T_E}) = \{[1, 1, 0]^T\}$, we may erroneously conclude that the plant is possibly at marking $[1, 1, 0]^T$ and $[1, 0, 1]^T$. □

V. Alertness and Confusability of Basis Markings

In this section, we first introduce several notions of basis markings regarding their *alert* properties. Then, we present a characteristic of marking predictability in terms of the alertness and the confusability of basis markings.

A. Alertness of Basis Markings

Definition 5.1: Given an LPN $G = (N, M_0, E, \ell)$, a BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ that satisfies Condition 1, and a set of alert markings $S = \mathcal{L}_{(\mathbf{w},k)}$.

1) The set of *fully alert basis markings* is defined as $\mathcal{F} = \{M_b \in \mathcal{M} \mid R_I(M_b) \subseteq S\}$.
2) The set of *partially alert basis markings* is defined as $\mathcal{P} = \{M_b \in \mathcal{M} \mid R_I(M_b) \cap S \neq \emptyset \wedge R_I(M_b) \setminus S \neq \emptyset\}$.
3) The set of *weakly alert basis markings* is defined as $\mathcal{W} = \{M_b \in \mathcal{M} \mid R_I(M_b) \cap S = \emptyset \wedge (\exists (M_b, \phi, M_b') \in \Delta)\, M_b' \in \mathcal{F} \cup \mathcal{P}, \ell(\phi) = \varepsilon\}$. □

In other words, a basis marking is fully alert ($M_b \in \mathcal{F}$) if all markings in its implicit reach belong to S, and a basis marking is partially alert ($M_b \in \mathcal{P}$) if not all but some markings in its implicit reach belong to S. For a weakly alert basis marking $M_b \in \mathcal{W}$, its implicit reach does not contain any alert markings, but in the BRG, there exists a path ϕ with $\ell(\phi) = \varepsilon$ from M_b that reaches some fully or partially alert basis markings. The three sets $\mathcal{F}, \mathcal{P}, \mathcal{W}$ are subsets of $\mathcal{M}$ and are mutually disjoint by Definition 5.1. Given a set of basis markings $\mathcal{M}$ and a set of alert markings $S = \mathcal{L}_{(\mathbf{w},k)}$, the following proposition can be used to compute the sets of fully and partially alert basis markings.

Proposition 5.1: Given an LPN $G = (N, M_0, E, \ell)$, a BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ that satisfies Condition 1, and a set of alert markings $S = \mathcal{L}_{(\mathbf{w},k)}$, a basis marking M_b is fully alert if and only if the following integer constraint is NOT feasible:

$$\begin{cases} M_b + C_I \cdot \mathbf{y} \geq \mathbf{0} \\ \mathbf{w}^T \cdot (M_b + C_I \cdot \mathbf{y}) \geq k + 1 \\ \mathbf{y} \geq \mathbf{0}. \end{cases} \tag{4}$$

A basis marking M_b is partially alert if and only if the following integer constraint is feasible:

$$\begin{cases} M_b + C_I \cdot \mathbf{y} \geq \mathbf{0} \\ \mathbf{w}^T \cdot (M_b + C_I \cdot \mathbf{y}) \leq k \\ M_b + C_I \cdot \mathbf{y}' \geq \mathbf{0} \\ \mathbf{w}^T \cdot (M_b + C_I \cdot \mathbf{y}') \geq k + 1 \\ \mathbf{y}, \mathbf{y}' \geq \mathbf{0}. \end{cases} \tag{5}$$

Proof: The acyclicity of the implicit subnet ensures that the state equation gives a sufficient and necessary condition for reachability. On the one hand, a basis marking $M_b \in \mathcal{F}$ if and only if (4) is not feasible, i.e., there is no marking $M \in R_I(M_b)$ such that $M \notin S$. On the other hand, a basis marking $M_b \in \mathcal{P}$ if and only if (5) is feasible, i.e., there exists two markings $M, M' \in R_I(M_b)$ such that $M \in S$ and $M' \notin S$. ■

To determine the sets $\mathcal{F}$ and $\mathcal{P}$, $2|\mathcal{M}|$ ILPs of (4) and (5) need to be solved. Numerical results [38] show that the total time to solve such ILPs is quite small ($< 10\%$) compared with the time consumption of constructing a BRG.

Moreover, set $\mathcal{W}$ can be computed by first determining sets $\mathcal{F}$ and $\mathcal{P}$ followed by a backward search via arcs labeled $(t, \mathbf{y})$ with $\ell(t) = \varepsilon$ from $\mathcal{F} \cup \mathcal{P}$. The following proposition relates the consistent alert markings with the sets of $\mathcal{F}, \mathcal{P}$, and $\mathcal{W}$.

Proposition 5.2: Given an LPN $G=(N,M_0,E,\ell)$ and a BRG $\mathcal{B}=(\mathcal{M},Tr,\Delta,M_0)$ that satisfies Condition 1, for any observation $w\in L(G)$, the following condition holds:

$$\mathcal{C}(w)\cap S\neq\emptyset \Leftrightarrow \mathcal{M}(w)\cap(\mathcal{F}\cup\mathcal{P}\cup\mathcal{W})\neq\emptyset.$$

Proof: By Proposition 4.2, $\mathcal{C}(w)=\bigcup_{M_b\in\mathcal{M}(w)}R_I(M_b)$ holds. If $\mathcal{C}(w)\cap S\neq\emptyset$, among all $M_b\in\mathcal{M}(w)$, there exists at least one basis marking M_b such that $R_I(M_b)\cap S\neq\emptyset$, which indicates that $M_b\in\mathcal{F}\cup\mathcal{P}$.

On the other hand, if $\mathcal{C}(w)\cap S=\emptyset$, all basis markings $M_b\in\mathcal{M}(w)$ necessarily satisfy $R_I(M_b)\cap S=\emptyset$, which indicates that $\mathcal{M}(w)\cap(\mathcal{F}\cup\mathcal{P})=\emptyset$. Suppose, by contradiction, that there exists $M_b\in\mathcal{M}(w)\cap\mathcal{W}$. By Definition 5.1, there exists $M_b'\in\mathcal{F}\cup\mathcal{W}$ and $(t,\mathbf{y})$ with $\ell(t)=\varepsilon$ such that $(M_b,(t,\mathbf{y}),M_b')\in\Delta$, which implies $\mathcal{M}(w)\cap(\mathcal{F}\cup\mathcal{P})\neq\emptyset$. Hence, $\mathcal{M}(w)\cap\mathcal{W}=\emptyset$ holds. ■

Hereafter, we assume that $\mathcal{M}(\varepsilon)\cap(\mathcal{F}\cup\mathcal{P}\cup\mathcal{W})=\emptyset$; otherwise, a plant may reach S without firing any observable transitions and the system is trivially not predictable. Since, by Proposition 5.2, $\mathcal{M}(w)\cap(\mathcal{F}\cup\mathcal{P}\cup\mathcal{W})\neq\emptyset$ implies that a plant may have reached the set S, to guarantee the criterion "no missed alarm," an alarm should be issued when (or earlier than when) observing w such that i) $\mathcal{M}(w)\cap(\mathcal{F}\cup\mathcal{P}\cup\mathcal{W})=\emptyset$; and ii) there exists an event e such that $\mathcal{M}(we)\cap(\mathcal{F}\cup\mathcal{P}\cup\mathcal{W})\neq\emptyset$. This motivates the following notion of *boundary basis markings*.

Definition 5.2: Given an LPN $G=(N,M_0,E,\ell)$, a BRG $\mathcal{B}=(\mathcal{M},Tr,\Delta,M_0)$ that satisfies Condition 1, and a set of alert markings S, a basis marking $M_b\in\mathcal{M}$ is said to be a *boundary basis marking* if

1) there exists $\phi\in Tr^*$ such that $(M_0,\phi,M_b)\in\Delta$ and for all $\bar{\phi}\in Pr(\phi)$ such that $(M_0,\bar{\phi},M_b')\in\Delta$, $M_b'\notin\mathcal{F}\cup\mathcal{P}\cup\mathcal{W}$ holds;
2) there exists $(t,\mathbf{y})$ with $t\in T_o$ such that $(M_b,(t,\mathbf{y}),M_b')\in\Delta$ and $M_b'\in\mathcal{F}\cup\mathcal{P}\cup\mathcal{W}$.

We denote by $\mathcal{U}$ the set of all boundary basis markings. □

Intuitively, a boundary basis marking M_b is a basis marking such that i) it can be reached from initial basis marking M_0 without passing any basis markings in $\mathcal{F}\cup\mathcal{P}\cup\mathcal{W}$; and ii) it may reach some basis marking $M_b'\in\mathcal{F}\cup\mathcal{P}\cup\mathcal{W}$ via only one arc labeled by an observable transition, i.e., $M_b\xrightarrow{(t,\mathbf{y})}M_b'$ where $\ell(t)\in E$. Now, suppose that the set of consistent basis markings $\mathcal{M}(w)$ of an observation w contains some boundary marking $M_b\in\mathcal{U}$. To guarantee "no missed alarm," an alarm has to be issued if $\mathcal{M}(w)$ contains any boundary basis marking. However, it may also happen that, from some consistent marking, there exists an infinite-long trajectory along which a plant will not reach S, which implies that such an alarm may be a false alarm. Therefore, to guarantee "no false alarm," we need to inspect all markings $M\in\mathcal{C}(w)$ whether $d_{\max}(M)\neq\infty$ holds or not. Since $\mathcal{C}(w)=\bigcup_{M_b\in\mathcal{M}(w)}R_I(M_b)$, it is sufficient to inspect all consistent basis markings $\mathcal{M}(w)$ instead of all consistent markings $\mathcal{C}(w)$. To characterize this, we introduce the following notion of *basis indicator*.

Definition 5.3: [Basis indicator] Given a plant LPN G, a set of alert markings S, and a BRG $\mathcal{B}$ that satisfies Condition 1, a basis marking M_b is called a *basis indicator* if $M_b\notin\mathcal{F}\cup\mathcal{P}\cup\mathcal{W}$ and there exists an integer $K\in\mathbb{N}$ such that

$$(\forall\sigma\in T^*)|\sigma|\geq K, M_b[\sigma\rangle \quad\Rightarrow\quad \sigma\in L_{M_b,S}. \tag{6}$$

The set of all basis indicators is denoted as $\mathcal{I}$. □

In other words, a basis marking M_b is a basis indicator if i) M_b is not in $\mathcal{F}\cup\mathcal{P}\cup\mathcal{W}$, which implies that from M_b, the plant will not reach S by firing only unobservable transitions; and ii) from M_b, the plant will inevitably reach S. The following theorem provides a characteristic of predictability in Definition 3.2 using basis indicators and boundary basis markings. Intuitively speaking, a plant is predictable with respect to S if and only if for any observation w such that $\mathcal{M}(w)$ contains at least one boundary basis marking, all consistent basis markings in $\mathcal{M}(w)$ necessarily be basis indicators.

Theorem 5.1: An LPN $G=(N,M_0,E,\ell)$ is predictable with respect to S if and only if for any observation $w\in L(G)$, the following condition holds:

$$\mathcal{M}(w)\cap\mathcal{U}\neq\emptyset\Rightarrow\mathcal{M}(w)\subseteq\mathcal{I}.$$

Proof: (Only if) If G is predictable with respect to S, then for all w such that $\mathcal{C}(w)\cap S\neq\emptyset$, there exists $\bar{w}$, which is a strict prefix of w, such that for all markings $M\in\mathcal{C}(\bar{w})$ there exists an integer $K=d_{\max}(M)+1$ such that for all sequences σ with length $\geq K$, $\sigma\in L_{M,S}$. Let $K_w=1+\max_{M\in\mathcal{C}(\bar{w})}d_{\max}(M)$. Then, (6) holds.

(If) If G is not predictable with respect to S, there exists a sequence $\sigma\in L_{M_0,S}$ such that $M_0[\sigma\rangle M\in S$ and σ does not satisfy the condition in Definition 3.2. We rewrite σ as $\sigma=\sigma_1 t\sigma_2$, where $\sigma_2\in T_{uo}^*$ and $\ell(t)\in E$. By the definition of boundary basis marking, the consistent basis marking $\mathcal{M}(w)$ of observation $w=\ell(\sigma_1)$ necessarily contains a boundary basis marking. Since σ does not satisfy the condition in Definition 3.2, there necessarily exist a sequence $\sigma'\in\ell^{-1}(\ell(\sigma_1))\setminus L_{M_0,S}$ and an infinite long sequence σ'' such that $M_0[\sigma'\rangle M'\sigma''$, $M'\in\mathcal{C}(w)$, and $\sigma'\sigma''\notin L_{M_0,S}$. This indicates the nonexistence of K_w in (6). ■

To verify (6) in Theorem 5.1, one needs to inspect all basis markings in $\mathcal{M}(w)$ for any observation w. This motivates the notion of *marking confusability* that will be introduced in the next subsection.

B. Confusability of Basis Markings

Definition 5.4: Given an LPN $G=(N,M_0,E,\ell)$ and a BRG $\mathcal{B}=(\mathcal{M},Tr,\Delta,M_0)$ that satisfies Condition 1, two basis markings M_b and M_b' are said to be *confusable* if there exist two paths $\phi_1,\phi_2\in Tr^*$ such that

1) $(M_0,\phi_1,M_{b,1}),(M_0,\phi_2,M_{b,2})\in\Delta$ and $\ell(\phi_1)=\ell(\phi_2)$;
2) for all $\bar{\phi}_i\in Pr(\phi_i), i\in\{1,2\}$, $(M_0,\bar{\phi}_i,M_b'')\in\Delta$ implies $M_b''\notin\mathcal{F}\cup\mathcal{P}\cup\mathcal{W}$.

□

In other words, two basis markings $M_{b,1},M_{b,2}$ are confusable if they can be reached from M_0 via two paths in the BRG having the same observation, and all basis markings reached along on the paths are not alert. According to such a definition, each boundary basis marking is confusable with itself (by letting

$\phi_1 = \phi_2$). With the notion of confusability of basis markings, Theorem 5.1 can be rewritten as the following.

Theorem 5.2: An LPN $G = (N, M_0, E, \ell)$ is predictable with respect to S if and only if, for any two confusable basis markings $M_b, M_b' \in \mathcal{M}$, that M_b is a boundary basis marking implies that M_b' is a basis indicator.

Proof: ($\Rightarrow$) This proof is by contrapositive. Suppose that there exists two basis markings $M_b, M_b' \in \mathcal{M}$ are confusable, M_b is a boundary basis marking, and M_b' is not a basis indicator. We can conclude that there exists an observation w such that $M_b, M_b' \in \mathcal{M}(w)$, which means that $\mathcal{M}(w) \cap \mathcal{U} \neq \emptyset$ and $\mathcal{M}(w) \nsubseteq \mathcal{I}$. Hence, by Theorem 5.1, G is not predictable with respect to S.

($\Leftarrow$) This proof is by contrapositive. Suppose that G is not predictable with respect to S. By Theorem 5.1, there exists an observation $w \in L(G)$ such that $\mathcal{M}(w) \cap \mathcal{U} \neq \emptyset$ and $\mathcal{M}(w) \nsubseteq \mathcal{I}$. This implies that there exist two confusable basis markings $M_b, M_b' \in \mathcal{M}(w)$ such that $M_b \in \mathcal{M}(w) \cap \mathcal{U}$ and $M_b' \in \mathcal{M}(w) \setminus \mathcal{I}$ (note that it may happen that $M_b = M_b'$), i.e., M_b is a boundary basis marking while M_b' is not a basis indicator. ∎

The confusability of two basis markings can be verified using the *BRG-observer* [14] whose size is exponential in the size of the BRG $\mathcal{M}$. To further mitigate the computational complexity, we propose to check confusability using a different structure called the *twin-BRG*.

Definition 5.5: Given an LPN $G = (N, M_0, E, \ell)$ and its BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$, the twin-BRG of $\mathcal{B}$ is a nondeterministic automaton $B = (X, E, \delta, x_0)$, where $X \subseteq \mathcal{M} \times \mathcal{M}$, $x_0 = (M_0, M_0)$, and the transition relation δ is iteratively defined as

1) $X = \{(M_0, M_0)\}$
2) for all $(M', M'') \in X$

$$\begin{cases} (M', (t', \cdot), \bar{M}') \in \Delta \\ (M'', (t'', \cdot), \bar{M}'') \in \Delta \\ \ell(t') = \ell(t'') = e \in E \end{cases} \Rightarrow \begin{cases} \bar{x} \in X \\ ((M', M''), e, \bar{x}) \in \delta \end{cases} \tag{7}$$

where $\bar{x} = (\bar{M}', \bar{M}'')$. □

The construction of twin-BRGs is analogous to the *verifier* automaton in [43] using *parallel composition* of automata, and hence it is omitted here due to the limit of space. Intuitively, a twin-BRG captures all pairs of sequences having the same observation in its single structure. We use the following example to illustrate it.

Example 5.1: Consider the BRG $\mathcal{B}$ in Fig. 5 (a) where $\ell(t_1) = \ell(t_2) = a$ and $\ell(t_3) = \varepsilon$. Its twin-BRG B is depicted in Fig. 5(b) as follows. Initially, set X contains one state $X = \{(M_0, M_0)\}$. Since $(M_0, (t_1, \mathbf{y}_1), M_1), (M_0, (t_2, \mathbf{y}_1), M_3) \in \Delta$, and $\ell(t_1) = \ell(t_2) = a$, by (7), four states $(M_1, M_1), (M_1, M_3), (M_3, M_1)$, and (M_3, M_3) are added to state set X, and four arcs labeled a are also added. For state (M_1, M_1), by $(M_1, (t_3, \mathbf{y}_3), M_2)$ with $\ell(t_3) = \varepsilon$, three states $(M_1, M_2), (M_2, M_1)$, and (M_2, M_2) are added. Similarly, another two states (M_2, M_3) and (M_3, M_2) are added to states (M_1, M_3) and (M_3, M_1), respectively. □

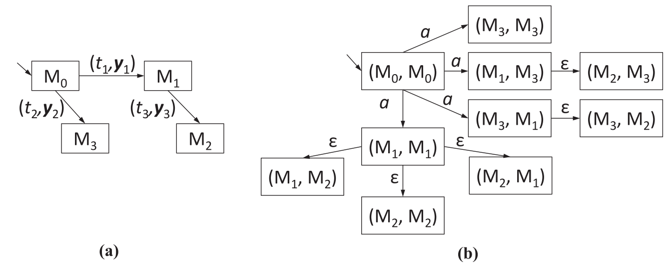

Fig. 5. (a) A BRG $\mathcal{B}$ and (b) its twin-BRG B.

The following proposition provides a way to verify confusability of two basis markings using the twin-BRG.

Proposition 5.3: Given a BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ and its twin-BRG $B = (X, E, \delta, (M_0, M_0))$, two basis markings M_b and M_b' are confusable if and only if there exists a path from (M_0, M_0) to (M_b, M_b') such that all states $(\bar{M}_b, \bar{M}_b')$ reached along the path satisfy $\bar{M}_b, \bar{M}_b' \notin \mathcal{F} \cup \mathcal{P} \cup \mathcal{W}$. □

Proof: For the "if" part, suppose that in B, there exists a path from (M_0, M_0) to (M_b, M_b') and all states on the path do not contain basis markings in $\mathcal{F} \cup \mathcal{P} \cup \mathcal{W}$. By the construction of the twin-BRG, from the path, we can extract two sequences $\phi, \phi' \in Tr^*$ such that in the BRG $(M_0, \phi, M_b), (M_0, \phi', M_b') \in \Delta$ holds, and all basis markings on the two paths $M_0 \xrightarrow{\phi} M_b$, $M_0 \xrightarrow{\phi'} M_b'$ are not in $\mathcal{F} \cup \mathcal{P} \cup \mathcal{W}$. According to Definition 5.4, M_b and M_b' are confusable. The "only if" part can be analogously proved.

The complexity of computing a twin-BRG is of quadratic in the number of basis markings in the corresponding BRG, i.e., $O(|\mathcal{M}|^2)$. To check if two basis markings are confusable, it suffices to remove all states in $\mathcal{F} \cup \mathcal{P} \cup \mathcal{W}$ from the twin-BRG and search if the pair considered is reachable in the remaining structure. Therefore, checking whether or not two basis markings are confusable can be done with quadratic complexity in the size of the BRG.

VI. Computing Basis Indicators: A Special Case

In the previous sections, we have shown that checking marking predictability is equivalent to check if boundary basis markings are only confusable with basis indicators. The computation of boundary basis markings is rather straightforward according to their definition. However, deciding how to compute the set of basis indicators is much more challenging. Specifically, for the case of indicator states in finite-state automaton [17], an indicator state can be checked by searching the existence of a cycle without alert states. However, the BRG is an abstracted structure with some information of markings implicitly stored. Particularly, a path may pass a partially alert basis marking without actually hitting S. Therefore, new techniques are needed to handle indicator markings in the BRG.

In this section, we present our first main result on how to compute basis indicators in a special case where there is no partially alert basis marking. Although this special case looks restrictive, it provides useful implications that help us establish a method applicable to general cases in the next section.

A. Computing Basis Indicators With No Partially Alert Basis Marking

The following theorem shows that if there does not exist any partially alert basis marking in the BRG (i.e., $\mathcal{P} = \emptyset$), then the set of basis indicators $\mathcal{I}$ can be computed by structural analysis of the BRG.

Theorem 6.1: Given an LPN $G = (N, M_0, E, \ell)$, its BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ that satisfies Condition 1, and a set of alert markings S, suppose that $\mathcal{P} = \emptyset$. A basis marking $M_b \notin \mathcal{F} \cup \mathcal{W}$ is a basis indicator if and only if there exists an integer K such that all path $M_b \xrightarrow{\phi}$ with $|\phi| \geq K$ necessarily contains at least one basis marking in $\mathcal{F}$.

Proof: (Only if) Suppose that M_b is a basis indicator, i.e., there exists an integer K' such that for all σ such that $M_b[\sigma\rangle$ and $|\sigma| \geq K', \sigma \in L_{M_b,S}$ holds. Now we prove that there necessarily exists an integer K such that all path $M_b \xrightarrow{\phi}$ with $|\phi| \geq K$ necessarily contains at least one basis marking in $\mathcal{F}$.

By contradiction, suppose that such an integer K does not exist. Since the net is assumed to be deadlock-free (Assumption A1), from M_b, there necessarily exists an infinite-long path $M_b \to M_{b,1} \to M_{b,2} \to \cdots$ such that all M_{b_i} satisfy $R_I(M_{b,i}) \cap S = \emptyset$. From the path, we can construct the following infinite-long trajectory:

$$M_{b,0}[\sigma_1\rangle M_1[t_{i_1}\rangle M_{b,1}[\sigma_2\rangle M_2[t_{i_2}\rangle M_{b,2} \cdots$$

that does not pass any marking in S. This contradicts the fact that M_b is a basis indicator. Therefore, there necessarily exists an integer K such that all paths $M_b \xrightarrow{\phi}$ with $|\phi| \geq K$ contain at least one basis marking in $\mathcal{F}$.

(If) Suppose that there exists an integer K such that all paths $M_b \xrightarrow{\phi}$ with $|\phi| \geq K$ necessarily contain at least one basis marking in $\mathcal{F}$. We prove that M_b is a basis indicator by contradiction.

Suppose that $\sigma = \sigma_1 t_{i_1} \cdots \sigma_K t_{i_K} \sigma_{K+1}$, where $t_{i_j} \in T_E$ and $\sigma_j \in T_I^*$ for $1 \leq j \leq K+1$, is a sequence such that trajectory $M_b[\sigma\rangle M$ does not pass S. According to Proposition 2.1, in the BRG, there exists a path

$$M_b \xrightarrow{(t_{i_1},\cdot)} M_{b,1} \xrightarrow{(t_{i_2},\cdot)} \cdots \xrightarrow{(t_{i_K},\cdot)} M_{b,K}$$

such that all $M_{b,j} \notin \mathcal{F}$, which obviously contradicts the fact that from M_b, all paths whose lengths are no less than K pass some fully alert basis marking. Hence, all sequences σ such that $M_b[\sigma\rangle$ with $|\sigma_{\uparrow T_E}| \geq K$ necessarily pass S. Since the T_I-induced subnet is acyclic, the length of all sequences σ with $|\sigma_{\uparrow T_E}| = K$ is also bounded. Therefore, M_b is a basis indicator.

Theorem 6.1 can be used to test if a basis marking M_b is a basis indicator: the existence of the integer K can be verified by first removing all fully alert basis markings from the BRG followed by checking if the remaining part of the BRG accessible from M_b contains a cycle. Then, if there does not exist any partially alert basis marking in the BRG, according to Theorem 5.2, the predictability of G can be verified by checking confusability using the twin-BRG of a plant. We use the following example to illustrate this. ■

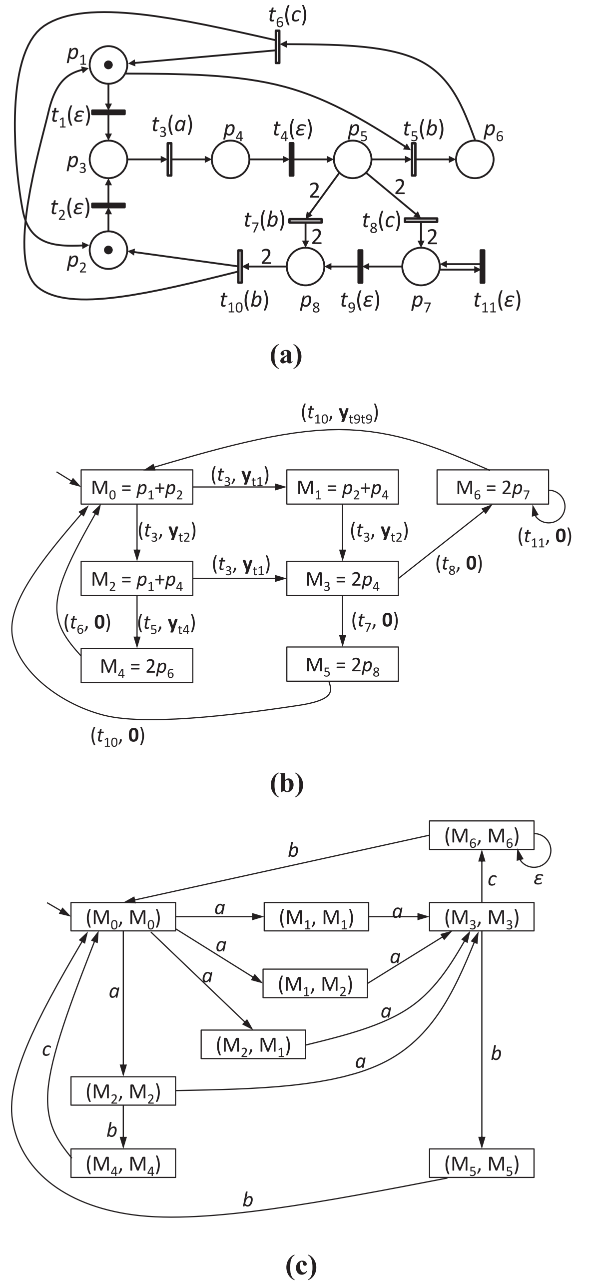

Fig. 6. (a) LPN net of Example 6.1. (b) BRG with $T_E = \{t_3, t_5, t_7, t_{10}, t_{11}\}$. (c) Twin-BRG.

Example 6.1: Consider the LPN in Fig. 6 (a) with a set of alert markings $S = \{M \mid M(p_7) + 2M(p_8) \geq 2\}$. Its BRG with $T_E = \{t_3, t_5, t_7, t_{10}, t_{11}\}$ is depicted in Fig. 6(b). With Proposition 5.1, in the BRG, there are one boundary basis marking $\mathcal{U} = \{M_3\} = \{2p_4\}$, two fully alert basis markings $\mathcal{F} = \{M_5, M_6\} = \{2p_8, 2p_7\}$, and no partially or weakly alert basis markings, i.e., $\mathcal{P} = \mathcal{W} = \emptyset$. Since $\mathcal{P} = \emptyset$, Theorem 6.1 can be applied, and hence by inspecting the BRG, we have $\mathcal{I} = \{M_1, M_3\}$. By constructing the twin-BRG in Fig. 6(c), we see that there exists a unique boundary basis marking M_3

that is only confusable with itself. Hence, according to Theorem 5.2, G is predictable with respect to S. A correct predictor is $(abc)^l aa, l \geq 0 \Leftrightarrow \mathcal{A} = 1$. □

Theorem 6.1 can be used to compute the set of basis indicators $\mathcal{I}$ if there is no partially alert basis markings, based on which Theorem 5.2 can be applied to verify the marking predictability problem. However, the condition "$\mathcal{P} = \emptyset$" is very strong that can hardly be satisfied in practice. In the next subsection, we will explain why the condition in Theorem 6.1 no longer holds when $\mathcal{P} \neq \emptyset$ and the difficulty for a general case. Furthermore, in the next section, we will develop a more general method for marking prediction without requiring the nonexistence of partially alert basis markings.

B. Difficulties due to the Existence of Partially Alert Basis Markings

In this subsection, we present observations on the existence of partially alert basis markings that provide us some useful conceptual intuitions. In the literature, many important properties of Petri nets such as *diagnosability* [11], [39] and *opacity* [14] can be verified by simply inspecting the basis markings in an arbitrary BRG that can be used for marking estimation. However, due to the existence of partially alert basis markings, not all BRGs can be used to solve the marking prediction problem. In general, the properties of alertness of basis markings provide neither a sufficient nor a necessary condition for marking predictability.

Fact 2: Given an LPN $G = (N, M_0, E, \ell)$ with BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ that satisfies Condition 1 and a set of alert markings S, if there exists a path $M_b \xrightarrow{\phi} M_b' \xrightarrow{\phi'} M_b''$ in the BRG such that $M_b' \in \mathcal{P}$, it

1) neither implies that all sequences that coincide with $\ell(\phi\phi')$ pass S;
2) nor implies the existence of a sequence that coincides with $\ell(\phi\phi')$ does not pass S.

□

The first condition in Fact 2 is not difficult to be understood: since $M_b' \in \mathcal{P}$ implies $R_I(M_b') \setminus S \neq \emptyset$, some infinite firing sequences that coincide with $\ell(\phi\phi')$ may pass some markings in $R_I(M_b') \setminus S$ without passing S. On the other hand, it may also happen that all firing sequences passing $R_I(M_b') \setminus S$ will eventually pass some markings in S. We show the correctness of Fact 2 using the following example.

Example 6.2: Consider the LPN G with $S = \{M \mid M(p_2) \geq 2\}$ shown in Fig. 7 (a). Its BRG with respect to $T_E = \{t_1, t_3, t_4\}$, shown in the same figure, has six basis markings among which there is one partially alert basis marking: $M_2 = 2p_2 \in \mathcal{P}$. Although $M_2 \in \mathcal{P}$ (and $M_2 \in S$), we cannot conclude that any firing sequence passing $R_I(M_2)$ necessarily passes S. For instance, consider the path $M_0 \xrightarrow{(t_1,\mathbf{0})} M_1 \xrightarrow{(t_1,\mathbf{0})} M_2 \xrightarrow{(t_3,\mathbf{y}_{t_2})} M_4$, where $\phi = (t_1, \mathbf{0})(t_1, \mathbf{0})$, $\phi' = (t_3, \mathbf{y}_{t_2})$, and $\ell(\phi\phi') = aab$. Observing aab does not imply that the plant has passed S: sequence $t_1t_2t_1t_3$ satisfies $\ell(t_1t_2t_1t_3) = aab$, but trajectory $M_0[t_1t_2t_1t_3\rangle$ does not pass S.

As another example, consider the LPN G with $S = \{M \mid M(p_3) \geq 2\}$ shown in Fig. 7(b). Its BRG with respect to $T_E = \{t_1, t_3, t_4\}$, shown in the same figure, has four basis markings, one of which is a partially alert basis marking, i.e., $M_2 = 2p_2 \in \mathcal{P}$. Consider the path $M_0 \xrightarrow{(t_1,\mathbf{0})} M_1 \xrightarrow{(t_1,\mathbf{0})} M_2 \xrightarrow{(t_3,\mathbf{y}_{t_2t_2})} M_3$, where $\phi = (t_1, \mathbf{0})(t_1, \mathbf{0})$, $\phi' = (t_3, \mathbf{y}_{t_2t_2})$, and $\ell(\phi\phi') = aab$. Although $M_2 \notin S$, one can verify that any firing sequence that coincides with aab necessarily passes S. In fact, this system is predictable—an alarm can be issued for $w = \varepsilon$. □

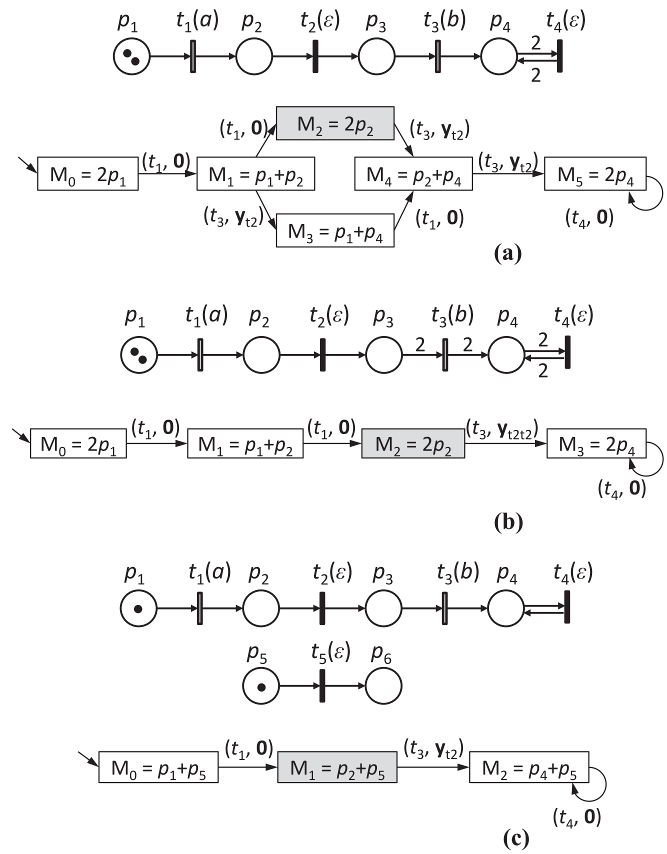

Fig. 7. LPNs used in Examples 6.2 and 6.3.

On the other hand, one may intuitively conjuncture that, for a partially alert basis marking M_b, the existence of a sequence that passes $R_I(M_b)$ without passing S can be determined by inspecting the minimal explanations of explicit transitions firable from M_b. Unfortunately, this conjecture is also false.

Fact 3: Given an LPN $G = (N, M_0, E, \ell)$ with BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ and a set of alert markings S, let $M_b \in \mathcal{P}$ and $\sigma_{\min}$ be a minimal explanation of $t \in T_E$. That all trajectories coinciding with $M_b[\sigma_{\min}t\rangle$ pass S does not necessarily imply that for all σ such that $\mathbf{y}_\sigma \geq \mathbf{y}_{\sigma_{\min}}$, $M_b[\sigma t\rangle$ passes S. □

We use the following example to illustrate the above fact.

Example 6.3: Consider the LPN G with $S = \{M \mid M(p_3) + M(p_5) \geq 2\}$ shown in Fig. 7(c). Its BRG with respect to $T_E = \{t_1, t_3, t_4\}$, shown in the same figure, has three basis markings among which $M_1 = p_2 + p_5 \in \mathcal{P}$. At $M_1 = p_2 + p_5$, the unique minimal explanation to fire t_3 is $\sigma_{\min} = t_2$, and the trajectory $M_1[t_2\rangle M[t_3\rangle M_2$ passes marking $M = p_3 + p_5 \in S$. However, observing the firing of t_3 (i.e., event b) at M_1 does not necessarily imply that a marking in S has been reached, since the trajectory $M_1[t_5t_2t_3\rangle$ satisfies $\ell(t_5t_2t_3) = b$ but does not pass S. □

Facts 2 and 3 show that the condition in Theorem 6.1 is not a sufficient or a necessary condition for marking predictability if $\mathcal{P} \neq \emptyset$. On the one hand, one needs to enumeratively examine the unobservable reach of a partially alert basis marking to determine if S will be necessarily passed or not. On the other hand, as we have mentioned before, there is no efficient way to find a basis partition that guarantees $\mathcal{P} = \emptyset$. To overcome this problem, in the next section, we propose another condition on the selection of explicit transitions such that the computation of set $\mathcal{I}$ can be done by structural analysis.

VII. Computing Basis Indicators: General Cases

In this section, we discuss the computation of basis indicators for general cases. The general idea is to add an additional requirement on the basis partition before constructing the BRG to circumvent the need of enumerating all explanations of basis markings. Specifically, besides Condition 1, we require that the basis partition used to compute the BRG should also satisfy the following condition.

Condition 2: The basis partition $\pi = (T_E, T_I)$ satisfies

$$\mathbf{w}^T \cdot C(\cdot, t) > 0 \quad \Rightarrow \quad t \in T_E.$$

Remark 5: We remind that in some practical cases (e.g., Example 3.1), set S is given in "$\geq$" form, i.e., $\mathbf{w}^T \cdot M \geq k$. Since such an inequality can be equivalently rewritten as $-\mathbf{w}^T \cdot M \leq -k$, in such a case, Condition 2 becomes "$\mathbf{w}^T \cdot C(\cdot, t) < 0 \Rightarrow t \in T_E$." □

Remark 6: We point out that Conditions 1 and 2 are not assumptions on plants. Instead, the two conditions provide a guideline to the plant operator for choosing a suitable set of explicit transitions—based on which the BRG-based state-abstraction technique can be used for marking prediction. □

The quantity $\mathbf{w}^T \cdot C(\cdot, t)$ is called the *influence* of a transition [37]. Therefore, Condition 2 essentially requires that all transitions with positive influence should be selected as explicit transitions. The following proposition shows that if the BRG satisfies both Conditions 1 and 2, then the computation of the set of basis indicators $\mathcal{I}$ can be done by inspecting the minimal explanation vectors.

Proposition 7.1: Given an LPN $G = (N, M_0, E, \ell)$ with BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ that satisfies Condition 2 and a set of alert markings S, let $M_b \in \mathcal{P}$ be a partial-alert basis marking and $M_b \xrightarrow{(t,\mathbf{y})}$ be an outgoing arc from M_b where $\mathbf{y}$ is a minimal explanation vector of $t \in T_E$. Then, $M_b + C_I \cdot \mathbf{y} = M \in S$ if and only if all trajectories $M_b[\sigma t\rangle$ pass S, where $\sigma \in T_I^*, \mathbf{y}_\sigma \geq \mathbf{y}$.

Proof: For "$\Leftarrow$," suppose that all trajectories $M_b[\sigma t\rangle$ pass S. There necessarily exists a minimal explanation σ' whose corresponding vector is $\mathbf{y}_{\sigma'} = \mathbf{y}$, that passes S, i.e., $M_b[\sigma'\rangle M$ passes S. Suppose that $\sigma' = \sigma_1'\sigma_2'$ such that $M_b[\sigma_1'\rangle M'[\sigma_2'\rangle M$ and $M' \in S$, i.e., $\mathbf{w}^T \cdot M' \leq k$. By Condition 2, $\mathbf{w}^T \cdot C(\cdot, t) \leq 0$ holds for any $t \in T_I$, which indicates that $\mathbf{w}^T \cdot C_I \cdot \mathbf{y}_{\sigma''} \leq 0$. Therefore, $\mathbf{w}^T \cdot M = \mathbf{w}^T \cdot M' + \mathbf{w}^T \cdot C_I \cdot \mathbf{y}_\sigma \leq k$, which indicates that $M' \in S$. ■

For the "$\Rightarrow$" part, since $M \in S$, $\mathbf{w}^T \cdot M = \mathbf{w}^T \cdot (M_b + C_I \cdot \mathbf{y}) = \mathbf{w}^T \cdot M_b + \mathbf{w}^T \cdot C_I \cdot \mathbf{y} \leq k$ holds. By Condition 2, $\mathbf{w}^T \cdot C(\cdot, t) \leq 0$ holds for any $t \in T_I$. Hence, $\mathbf{y}_\sigma \geq \mathbf{y}$ indicates that $\mathbf{w}^T \cdot C_I \cdot \mathbf{y}_\sigma \leq \mathbf{w}^T \cdot C_I \cdot \mathbf{y}$. Therefore, we have $\mathbf{w}^T \cdot M_b + \mathbf{w}^T \cdot C_I \cdot \mathbf{y}_\sigma \leq k$, which indicates that $M_b[\sigma\rangle M \in S$, i.e., trajectory $M_b[\sigma t\rangle$ necessarily passes S.

According to Proposition 7.1, if Condition 2 is satisfied, and if an arc $M_b \xrightarrow{(t,\mathbf{y})}$ satisfies $M_b + C_I \cdot \mathbf{y} = M \in S$, then any sequence whose Parikh vector is equal or larger than $\mathbf{y}$ necessarily passes S. Then, if all outgoing arcs of M_b labeled by $(t_{i_j}, \mathbf{y}_j)$ satisfy $M_b + C_I \cdot \mathbf{y}_j \in S$, all trajectories from M_b containing any transition $t \in T_E$ necessarily pass S. Such a type of basis markings is called *pseudo-partially alert basis markings*.

Definition 7.1: Given an LPN $G = (N, M_0, E, \ell)$, its BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$ that satisfies Conditions 1 and 2, and a set of alert markings $S = \mathcal{L}_{(\mathbf{w},k)}$, a basis marking M_b is a *pseudo-partially alert basis marking* if $M_b \in \mathcal{P}$ and for all $t \in T_E$ such that $M_b \xrightarrow{(t,\mathbf{y})}$, $\mathbf{w}^T \cdot (M_b + C_I \cdot \mathbf{y}) \leq k$ holds. The set of pseudo-partially alert basis markings is denoted as $\mathcal{P}_F$. □

The computation of pseudo-partially alert basis markings $\mathcal{P}_F$ can be done by simply checking all basis markings in $\mathcal{P}$ if $\mathbf{w}^T \cdot (M_b + C_I \cdot \mathbf{y}) \leq k$ holds for all output arcs $M_b \xrightarrow{(t,\mathbf{y})}$ from M_b. Then, we naturally deliver the following result.

Proposition 7.2: Given a pseudo-partially alert basis marking $M_b \in \mathcal{P}_F$, for all σt such that $M_b[\sigma t\rangle$ where $\sigma \in T_I^*$ and $t \in T_E$, $\sigma t \in L_{M_b,S}$ holds.

Proof: Directly from Proposition 7.1. ■

Now we can expose the main result of this section. That is, if a BRG satisfies both Conditions 1 and 2, then the basis indicators in the BRG can be computed by inspecting the fully and pseudo-partially alert basis markings in the BRG.

Theorem 7.1: Given an LPN G, a set of alert markings S, and a BRG satisfying both Conditions 1 and 2, a basis marking $M_b \notin \mathcal{F} \cup \mathcal{P}$ is a basis indicator if and only if there exists an integer K such that all paths $M_b \xrightarrow{\phi}$ with $|\phi| \geq K$ necessarily contain at least one basis marking in $\mathcal{F} \cup \mathcal{P}_F$.

Proof: (Only if) By contradiction, suppose that M_b is a basis indicator and such an integer K does not exist. Since M_b is a basis indicator, there exists an integer K' such that for all σ such that $M_b[\sigma\rangle$ and $|\sigma| \geq K'$, $\sigma \in L_{M_b,S}$ holds.

Now, for all basis markings M_b' in the BRG $\mathcal{B}$, we remove all outgoing arcs $M_b' \xrightarrow{(t,\mathbf{y})}$ such that $M_b' + C_I \cdot \mathbf{y} \leq k$ to obtain a sub-BRG $\mathcal{B}_{\text{sub}}$. Then, there are two cases both of which eventually reach contradictions.

Case I: The accessible part of $\mathcal{B}_{\text{sub}}$ from M_b does not contain cycles. Then, from M_b, by passing finite arcs, some basis marking $\bar{M}_b$ with no outgoing arc is reached. Since the net is assumed to be deadlock-free (Assumption A1), in the original BRG, each basis marking has at least one output arc. This indicates that all original outgoing arcs of $\bar{M}_b$ labeled $(t_{i_j}, \mathbf{y}_j)$ satisfy $M_b' + C_I \cdot \mathbf{y}_j \leq k$ and hence are removed, which indicates $\bar{M}_b \in \mathcal{F} \cup \mathcal{P}_F$. Let K be the length of the longest path from M_b in $\mathcal{B}$, and all paths $M_b \xrightarrow{\phi}$ with $|\phi| \geq K$ necessarily pass $\mathcal{F} \cup \mathcal{P}_F$. This contradicts the assumption that the bound K does not exist.

Case II: The accessible part of $\mathcal{B}_{\text{sub}}$ from M_b contains at least one cycle. Then, from M_b, there exists an infinite-long path $M_b \to M_{b,1} \to M_{b,2} \to \cdots$. From the path, we can construct the following infinite-long trajectory:

$$M_b[\sigma_1\rangle M_1[t_{i_1}\rangle M_{b,1}[\sigma_2\rangle M_2[t_{i_2}\rangle M_{b,2}\cdots.$$

Since $\mathbf{w}^T \cdot C(\cdot,t) \leq 0$ for all $t \in T_I$, $M_{b,j}[\bar{\sigma}\rangle \bar{M} \notin S$ holds for all $\bar{\sigma} \in Pr(\sigma_{j+1})$. Therefore, such a trajectory does not pass S. This contradicts the assumption that M_b is a basis indicator.

(If) Suppose that there exists an integer K such that all paths $M_b \xrightarrow{\phi}$ with $|\phi| \geq K$ necessarily contain at least one basis marking in $\mathcal{F} \cup \mathcal{P}_F$. We claim that for all sequences σ such that $|\sigma_{\uparrow T_E}| \geq K$ and σ is firable at M_b, trajectory $M_b[\sigma\rangle$ necessarily passes S.

By contradiction, suppose that there exists a sequence $\sigma = \sigma_1 t_{i_1} \cdots \sigma_K t_{i_K} \sigma_{K+1}$, where $t_{i_j} \in T_E$ and $\sigma_j \in T_I^*$, $1 \leq j \leq K+1$, such that $M_b[\sigma\rangle M$ does not pass S. According to Proposition 2.1, in the BRG, there exists a path

$$M_b \xrightarrow{(t_{i_1},\cdot)} M_{b,1} \xrightarrow{(t_{i_2},\cdot)} \cdots \xrightarrow{(t_{i_K},\cdot)} M_{b,K}$$

such that all $M_{b,j} \notin \mathcal{F} \cup \mathcal{P}_F$. This contradicts the fact that from M_b, all paths with length $\geq K$ pass some basis marking in $\mathcal{F} \cup \mathcal{P}_F$.

As a result, all trajectories $M_b[\sigma\rangle$ with $|\sigma_{\uparrow T_E}| \geq K$ necessarily pass S. Since the T_I-induced subnet is acyclic, the length of all sequences σ with $|\sigma_{\uparrow T_E}| = K$ is also bounded. Therefore, M_b is a basis indicator. ■

Theorem 7.1 can be used to compute basis indicators $\mathcal{I}$ in general cases regardless of the emptiness of $\mathcal{P}$. To test if a basis marking M_b is a basis indicator, one can first remove all basis markings in $\mathcal{F} \cup \mathcal{P}_F$ from the BRG followed by testing if the remaining part of the BRG accessible from M_b' is acyclic, which can be done in linear time to the scale of the BRG, i.e., $|\mathcal{M}|$. Hence, to compute all basis indicators (i.e., the set $\mathcal{I}$) requires $O(|\mathcal{M}|^2)$ complexity. Once all basis indicators are obtained, Theorem 5.2 can be applied to verify the marking predictability problem. Finally, based on results obtained so far, we present Algorithm 2 for marking predictability verification in LPNs.

In Algorithm 2, a basis partition satisfying Conditions 1 and 2 is found by step 1. We note that a partition that satisfies both conditions always exists (e.g., $T_E = T$), and Algorithm 3 in [38] can be augmented to find a suitable partition π. Similar to the argument in [38], such a procedure has polynomial complexity $O(|P| \cdot |T|^2)$. Then, steps 2 and 3 compute the BRG $\mathcal{B}$ and the twin-BRG B, respectively. Step 4 computes the sets of fully, partially, and weakly alert basis markings using Proposition 5.1 and Definition 7.1. Steps 5–7 compute the set of boundary basis markings $\mathcal{U}$, the set of pseudo-partially alert basis markings $\mathcal{P}_F$, and the set of basis indicators $\mathcal{I}$, respectively. In the loop of steps 8–12, for each boundary marking $M_b \in \mathcal{U}$, all basis markings confusable with it are examined. If there exists some basis marking M_b' that is not a basis indicator and confusable with M_b, the algorithm outputs NO and exit, meaning that the plant G is not predictable with respect to S. Otherwise, it output YES meaning that the plant is predictable.

Now let us discuss the complexity of Algorithm 2. As we have discussed in the previous sections, the complexity of Steps 1–7 of Algorithm 2 are as follows: $O(|P| \cdot |T|^2)$, $O(|\mathcal{M}|)$, $O(|\mathcal{M}|^2)$, $O(|\mathcal{M}|^2)$, $O(|\mathcal{M}|)$, $O(|\mathcal{M}|)$, and $O(|\mathcal{M}|^2)$, respectively. Hence, the complexity of Steps 1–7 is $O(|\mathcal{M}|^2 + |P| \cdot |T|^2)$. The complexity of the loop of Steps 8–12 is $O(|\mathcal{M}|^2)$. Therefore, the total complexity of Algorithm 2 is $O(|\mathcal{M}|^2 + |P| \cdot |T|^2) \approx O(|\mathcal{M}|^2)$. Since the number of basis markings is in general much smaller than the number of reachable markings [11], [38], our method for marking predictability verification in LPNs is of efficiency.

Algorithm 2: Predictability Verification.

Input: An LPN $G = (N, M_0, E, \ell)$ and a set of alert markings $S = \mathcal{L}_{(\mathbf{w},k)}$
Output: YES (predictable) / NO (not predictable)

1: Find a basis partition $\pi = (T_E, T_I)$ that satisfies Conditions 1 and 2;
2: Compute BRG $\mathcal{B} = (\mathcal{M}, Tr, \Delta, M_0)$;
3: Compute twin-BRG $B = (X, E, \delta, (M_0, M_0))$;
4: Compute the sets $\mathcal{F}, \mathcal{P}$, and $\mathcal{W}$;
5: Compute the set of boundary basis markings $\mathcal{U}$;
6: Compute the set of pseudo-partially alert basis markings $\mathcal{P}_F$;
7: Compute the set of basis indicators $\mathcal{I}$;
8: **for all** $M_b \in \mathcal{U}$, **do**;
9:     **for all** M_b' that is confusable with M_b, **do** ;
10:         If $M_b' \notin \mathcal{I}$, output NO, exit;
11:     **end for**
12: **end for**
13: Output YES, exit.

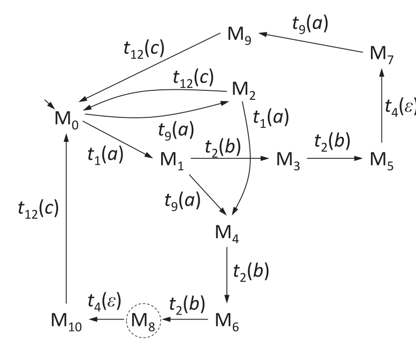

Fig. 8. BRG of the LPN in Fig. 2 with $T_E = \{t_1, t_2, t_4, t_9, t_{12}\}$.

Theorem 7.2: Algorithm 2 is correct.

Proof: Directly from Theorem 7.1. ■

Example 7.1: Let us consider again the LPN in Fig. 2. Now we use Algorithm 2 to verify the marking predictability with respect to the set of alert markings $S = \{M \mid M(p_4) + M(p_{11}) + 2M(p_{12}) \geq 3\}$, i.e., $\mathbf{w} = [0,0,0,-1,0,0,0,0,0,0,-1,-2,0]^T, k = -3$.

Although the T_{uo}-induced subnet is acyclic, the basis partition $\pi = (T_E, T_I)$ with $T_E = T_o$ does not satisfy Condition 2. Since there are two transitions t_4 and t_{12} that satisfy $\mathbf{w}^T \cdot C(\cdot,t) > 0$, let $T_E = T_o \cup \{t_4, t_{12}\} = \{t_1, t_2, t_4, t_9, t_{12}\}$. The corresponding BRG contains 11 basis markings depicted in Fig. 8 , as listed in Table II . Among all basis markings, only marking M_8 is weakly alert, i.e., $\mathcal{F} = \mathcal{W} = \emptyset$ and $\mathcal{P} = \{M_8\}$, which

TABLE II
BASIS MARKINGS IN FIG. 8 WITH $T_E = \{t_1, t_2, t_4, t_9, t_{12}\}$

M_0	[1 0 0 0 0 0 0 0 1 0 0 0 2]	M_6	[0 1 1 0 0 0 0 0 0 0 1 0 1]
M_1	[0 2 0 0 0 0 0 0 1 0 0 0 2]	M_7	[0 0 0 0 0 0 0 1 1 0 0 0 2]
M_2	[1 0 0 0 0 0 0 0 0 1 0 0 0]	M_8	[0 0 2 0 0 0 0 0 0 0 1 0 0]
M_3	[0 1 1 0 0 0 0 0 1 0 0 0 1]	M_9	[0 0 0 0 0 0 0 1 0 1 0 0 0]
M_4	[0 2 0 0 0 0 0 0 0 1 0 0 0]	M_{10}	[0 0 0 0 0 0 0 1 0 0 1 0 2]
M_5	[0 0 2 0 0 0 0 0 1 0 0 0 0]		

indicates the set of boundary basis markings is $\mathcal{U} = \{M_6\}$. Since from M_8 there is a unique outgoing arc $M_8 \xrightarrow{(t_4, \mathbf{y}_{t_3 t_3})}$ such that $\mathbf{w}^T \cdot (M_8 + C_I \cdot \mathbf{y}_{t_3 t_3}) = -3 \leq k$, M_8 is a pseudo-partially alert basis marking, i.e., $\mathcal{P}_F = \{M_8\}$. By Theorem 7.1, we have two basis indicators $\mathcal{I} = \{M_4, M_6\}$. The corresponding twin-BRG has 13 states (not drawn) such that $\mathcal{U}(M_6) = \{M_6\}$, i.e., boundary basis marking M_6 is only confusable with itself. By Theorem 5.2, the plant LPN is predictable with respect to S. □

VIII. ONLINE MARKING PREDICTION ALGORITHM

In this section, we address the problem of online marking prediction by showing how to design a predictor that correctly issues an alarm when necessary, according to an observation w. Specifically, we provide two approaches for designing a predictor: one follows the idea of "predict when it has to" and the other follows the idea of "predict whenever the information is sufficient." Both design approaches work correctly if a plant is marking predictable.

As discussed in Section IV, one direct approach to design a predictor is to issue an alarm when $\mathcal{M}(w)$ contains boundary basis markings. Hence, we have the following proposition.

Proposition 8.1: If a plant G is predictable with respect to S, the following predictor $\mathcal{A}$ is correct:

$$\mathcal{A}(w) = 1 \quad \Leftrightarrow \quad \mathcal{M}(w) \cap \mathcal{U} \neq \emptyset.$$

Proof: Suppose that $\mathcal{M}(w) \cap \mathcal{U} \neq \emptyset$. Since from $M_b \in \mathcal{U}$, there exists a sequence $\sigma_1 t \sigma_2$ such that $\sigma_1, \sigma_2 \in T_{uo}^*, t \in T_o$. Hence, $\mathcal{A}$ satisfies "no missed alarm." On the other hand, since G is predictable with respect to S, all $M_b \in \mathcal{M}(w)$ are basis indicators, which indicates that for all $M \in \mathcal{C}(w)$, there exists an integer K such that $M[\sigma\rangle, |\sigma| \geq K$ implies $\sigma \in L_{M,S}$. Hence, $\mathcal{A}$ satisfies "no false alarm." ■

The predictor designed in Proposition 8.1 issues the alarm when one or more boundary basis markings are consistent with the current observation. However, the alarm is issued only one step before S is reached, i.e., the plant may reach S by observing only one addition observable event. In fact, an alarm may be issued earlier when all consistent basis markings are basis indicators. This idea is explored by the following result.

Proposition 8.2: If a plant G is predictable with respect to S, the following predictor $\mathcal{A}$ is correct:

$$\mathcal{A}(w) = 1 \quad \Leftrightarrow \quad \mathcal{M}(w) \subseteq \mathcal{I}.$$

Proof: Since all $M_b \in \mathcal{M}(w)$ are basis indicators when an alarm is issued, $\mathcal{A}$ satisfies "no false alarm." On the other hand, since G is predictable with respect to S, $\mathcal{M}(w) \cap \mathcal{U} \neq \emptyset$ implies $\mathcal{M}(w) \subseteq \mathcal{I}$, which indicates that $\mathcal{A}$ satisfies "no missed alarm." ■

TABLE III
ILLUSTRATION OF OBSERVATION $w = abbacaa$ IN EXAMPLE 8.1

w	$\mathcal{M}(w)$	$\mathcal{A}(w)$	w	$\mathcal{M}(w)$	$\mathcal{A}(w)$
ε	M_0	0	$abba$	M_9	0
a	M_1, M_2	0	$abbac$	M_0	0
ab	M_3	0	$abbaca$	M_1, M_2	0
abb	M_5, M_7	0	$abbacaa$	M_4	1

Algorithm 3: Online Marking Prediction.

Input: An LPN $G = (N, M_0, E, \ell)$ and a set of alert markings $S = \mathcal{L}_{(\mathbf{w},k)}$

Offline Stage:

1: Call Algorithm 2. If return NO, exit;
2: Derive the BRG $\mathcal{B}$ and the set of basis indicators $\mathcal{I}$ from Algorithm 2;

Online Stage:

3: Let $w = \varepsilon$ and $\mathcal{A}(w) = 0$;
4: Compute $\mathcal{M}(w)$;
5: **if** $\mathcal{M}(w) \subseteq \mathcal{I}$, **then**
6: $\quad\mathcal{A}(w) = 1$;
7: **end if**
8: Wait until some event $e \in E$ fires, updated $w := we$, goto Step 4.

Algorithm 3 summarizes the design of a predictor for online marking prediction using Proposition 8.2. In the offline stage, Step 1 calls Algorithm 2 to verify the marking predictability of G with respect to S. If G is predictable with respect to S, the BRG $\mathcal{B}$, the set of boundary basis markings $\mathcal{U}$, and the set of basis indicators $\mathcal{I}$ are derived from Algorithm 2. In the online stage, the predictor monitors the events generated by the plant and computes the consistent basis markings accordingly. It issues an alarm if all consistent basis markings are basis indicators. Since the complexity of Algorithm 2 is $|\mathcal{M}|^2$, the offline stage of Algorithm 3 is also $|\mathcal{M}|^2$. Moreover, the online stage of Algorithm 3 is a one-step-look-ahead procedure whose complexity is negligible.

Example 8.1: Again consider the LPN in Fig. 2 where the set of alert markings is $S = \{M \mid M(p_4) + M(p_{11}) + 2M(p_{12}) \geq 3\}$, and the corresponding BRG with $T_E = T_o \cup \{t_4, t_{12}\} = \{t_1, t_2, t_4, t_9, t_{12}\}$ is depicted in Fig. 8. According to Example 7.1, G is predictable with respect to S, and the set of basis indicators is $\mathcal{I} = \{M_4, M_6\}$. According to Proposition 8.2, an alarm is issued if $\mathcal{M}(w) \subseteq \{M_4, M_6\}$.

Considering the observation $w = abbacaa$, the consistent basis markings $\mathcal{M}(w)$ of each prefix of w are listed in Table III . For all $\bar{w}$ that are strict prefixes of w, $\mathcal{M}(w) \nsubseteq \mathcal{I}$, and hence no alarm is issued. Finally, when $w = abbacaa$, we have $\mathcal{M}(abbacaa) = \{M_4\} \subseteq \mathcal{I}$, and an alarm is issued. One can readily verify that from $M_4 = 2 \cdot p_1 + p_{10}$, the plant will inevitably reaching S by further observing bb. □

IX. Conclusion

In this article, we proposed a framework for marking prediction of labeled Petri nets using BRG. The condition of marking predictability was proposed as the necessary and sufficient condition for the existence of a correct predictor with no missed alarm and no false alarm. We provided characteristics of marking predictability in terms of basis markings. An effective algorithm was then developed to verify marking predictability of a plant based on the notion of minimal explanations and BRGs. The complexity of the proposed algorithms is quadratic with respect to the number of basis markings. Finally, an online marking prediction algorithm for labeled Petri nets was also proposed. In the future, we would like to extend our framework to the stochastic setting as well as the decentralized setting for labeled Petri nets.

References

[1] A. Giua, D. Corona, and C. Seatzu, "State estimation of lambda-free labeled Petri nets with contact-free nondeterministic transitions," *Discrete Event Dyn. Syst.*, vol. 15, no. 1, pp. 85–108, 2005.

[2] D. Corona, A. Giua, and C. Seatzu, "Marking estimation of Petri nets with silent transitions," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1695–1699, Sep. 2007.

[3] M. Cabasino, C. N. Hadjicostis, and C. Seatzu, "Probabilistic marking estimation in labeled Petri nets," *IEEE Trans. Autom. Control*, vol. 60, pp. 528–533, Feb. 2015.

[4] F. Basile, M. P. Cabasino, and C. Seatzu, "State estimation and fault diagnosis of labeled time Petri net systems with unobservable transitions," *IEEE Trans. Autom. Control*, vol. 60, no. 4, pp. 997–1009, Apr. 2015.

[5] Z. Ma, Z. Li, and A. Giua, "Marking estimation in a class of time labeled Petri nets," *IEEE Trans. Autom. Control*, vol. 65, no. 2, pp. 493–506, Feb. 2020.

[6] J. Moody and P. Antsaklis, "Petri net supervisors for DES with uncontrollable and unobservable transitions," *IEEE Trans. Autom. Control*, vol. 45, no. 3, pp. 462–476, Mar. 2000.

[7] F. Basile, R. Cordone, and L. Piroddi, "A branch and bound approach for the design of decentralized supervisors in Petri net models," *Automatica*, vol. 52, pp. 322–333, 2015.

[8] Z. Ma, G. Zhu, Z. Li, and A. Giua, "Computation of admissible marking sets in weighted synchronization-free petri nets by dynamic programming," *IEEE Trans. Autom. Control*, vol. 65, no. 6, pp. 2662–2669, Jun. 2020.

[9] A. Ramirez-Trevino, E. Ruiz-Beltran, I. Rivera-Rangel, and E. López-Mellado, "Online fault diagnosis of discrete event systems: A Petri net-based approach," *IEEE Trans. Autom. Sci. Eng.*, vol. 4, no. 1, pp. 31–39, Jan. 2007.

[10] G. Jiroveanu and R. K. Boel, "The diagnosability of Petri net models using minimal explanations," *IEEE Trans. Autom. Control*, vol. 55, no. 7, pp. 1663–1668, Jul. 2010.

[11] M. Cabasino, A. Giua, and C. Seatzu, "Fault detection for discrete event systems using Petri nets with unobservable transitions," *Automatica*, vol. 46, no. 9, pp. 1531–1539, 2010.

[12] S. Shu and F. Lin, "Generalized detectability for discrete event systems," *Syst. Control Lett.*, vol. 60, no. 5, pp. 310–317, 2011.

[13] H. Lan, Y. Tong, J. Guo, and C. Seatzu, "Verification of C-detectability using Petri nets," *Inform. Sci.*, vol. 528, pp. 294–310, 2020.

[14] Y. Tong, Z. W. Li, C. Seatzu, and A. Giua, "Verification of state-based opacity using Petri nets," *IEEE Trans. Autom. Control*, vol. 62, no. 6, pp. 2823–2837, Jun. 2017.

[15] T. Jéron, H. Marchand, S. Genc, and S. Lafortune, "Predictability of sequence patterns in discrete event systems," in *Proc. 17th IFAC World Congr.*, 2008, pp. 537–543.

[16] S. Genc and S. Lafortune, "Predictability of event occurrences in partially-observed discrete-event systems," *Automatica*, vol. 45, no. 2, pp. 301–311, 2009.

[17] R. Kumar and S. Takai, "Decentralized prognosis of failures in discrete event systems," *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 48–59, Jan. 2010.

[18] S. Takai and R. Kumar, "Distributed failure prognosis of discrete event systems with bounded-delay communications," *IEEE Trans. Autom. Control*, vol. 57, no. 5, pp. 1259–1265, May 2012.

[19] X. Yin and Z. J. Li, "Decentralized fault prognosis of discrete event systems with guaranteed performance bound," *Automatica*, vol. 69, pp. 375–379, 2016.

[20] X. Yin and Z. Li, "Reliable decentralized fault prognosis of discrete-event systems," *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 46, no. 11, pp. 1598–1603, Nov. 2016.

[21] F. Liu, "Predictability of failure event occurrences in decentralized discrete-event systems and polynomial-time verification," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 1, pp. 498–504, Jan. 2018.

[22] X. Yin and Z. Li, "Decentralized fault prognosis of discrete-event systems using state-estimate-based protocols," *IEEE Trans. Cybern.*, vol. 49, no. 4, pp. 1302–1313, Apr. 2019.

[23] N. Bertrand, S. Haddad, and E. Lefaucheux, "Foundation of diagnosis and predictability in probabilistic systems," in *Proc. 34th IARCS Annu. Conf. Found. Softw. Technol. Theor. Comput. Sci.*, 2014, pp. 417–429.

[24] J. Chen and R. Kumar, "Stochastic failure prognosability of discrete event systems," *IEEE Trans. Autom. Control*, vol. 60, no. 6, pp. 1570–1581, Jun. 2015.

[25] F. Nouioua, P. Dague, and L. Ye, "Predictability in probabilistic discrete event systems," in *Soft Methods for Data Science*, (Advances in Intelligent Systems and Computing), M. Ferraro *et al.* Eds., vol. 456. Cham. Switzerland: Springer, 2017.

[26] B. Benmessahel, M. Touahria, and F. Nouioua, "Predictability of fuzzy discrete event systems," *Discrete Event Dyn. Syst.*, vol. 27, no. 4, pp. 641–673, 2017.

[27] S. Takai, "Robust prognosability for a set of partially observed discrete event systems," *Automatica*, vol. 51, pp. 123–130, 2015.

[28] M. Yokotani, T. Kondo, and S. Takai, "Abstraction-based verification and synthesis for prognosis of discrete event systems," *Asian J. Control*, vol. 18, no. 4, pp. 1279–1288, 2016.

[29] G. Fiore, E. De Santis, and M. D. Di Benedetto, "Predictability for finite state machines: A set-membership approach," in *Proc. 14th IFAC Workshop Discrete Event Syst.*, Sorrento, Italy, 2018, pp. 355–360.

[30] D. Lefebvre, "Fault diagnosis and prognosis with partially observed Petri nets," *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 44, no. 10, pp. 1413–1424, Oct. 2014.

[31] X. Yin, "Verification of prognosability for labeled Petri nets," *IEEE Trans. Autom. Control*, vol. 63, no. 6, pp. 1828–1834, Jun. 2018.

[32] W. Wu, X. Yin, and S. Li, "Verification of coprognosability in decentralized fault prognosis of labeled Petri nets," in *Proc. 57th IEEE Conf. Decis. Control*, Miami, FL, USA, 2018, pp. 4845–4850.

[33] D. You, S. Wang, and C. Seatzu, "Verification of fault-predictability in labeled Petri nets using predictor graphs," *IEEE Trans. Autom. Control*, vol. 64, no. 10, pp. 4353–4360, Oct. 2019.

[34] R. Ammour, E. Leclercq, E. Sanlaville, and D. Lefebvre, "Faults prognosis using partially observed stochastic Petri-nets: An incremental approach," *Discrete Event Dyn. Syst.*, vol. 28, no. 2, pp. 247–267, 2017.

[35] R. Ammour, E. Leclercq, E. Sanlaville, and D. Lefebvre, "Fault prognosis of timed stochastic discrete event systems with bounded estimation error," *Automatica*, vol. 82, pp. 35–41, 2017.

[36] A. Giua, F. DiCesare, and M. Silva, "Generalized mutual exclusion constraints for Petri nets with uncontrollable transitions," in *Proc. IEEE Int. Conf. Syst. Man Cybern.*, Chicago, IL, USA, 1992, pp. 947–949.

[37] Z. Ma, Z. Li, and A. Giua, "Design of optimal Petri net controllers for disjunctive generalized mutual exclusion constraints," *IEEE Trans. Autom. Control*, vol. 60, no. 7, pp. 1774–1785, Jul. 2015.

[38] Z. Ma, Y. Tong, Z. Li, and A. Giua, "Basis marking representation of Petri net reachability spaces and its application to the reachability problem," *IEEE Trans. Autom. Control*, vol. 62, no. 3, pp. 1078–1093, Mar. 2017.

[39] N. Ran, H. Su, A. Giua, and C. Seatzu, "Codiagnosability analysis of bounded Petri nets," *IEEE Trans. Autom. Control*, vol. 63, no. 8, pp. 1192–1199, Apr. 2018.

[40] M. Cabasino, A. Giua, M. Pocci, and C. Seatzu, "Discrete event diagnosis using labeled Petri nets: An application to manufacturing systems," *Control Eng. Pract.*, vol. 19, no. 9, pp. 989–1001, 2011.

[41] M. P. Cabasino, S. Lafortune, and C. Seatzu, "Optimal sensor selection for ensuring diagnosability in labeled Petri nets," *Automatica*, vol. 49, no. 8, pp. 2373–2383, 2013.

[42] Z. Ma, Y. Tong, Z. Li, and A. Giua, "Marking estimation in labelled Petri nets by the representative marking graph," in *Proc. 20th IFAC World Congr.*, 2017, pp. 11175–11181.

[43] T. S. Yoo and S. Lafortune, "Polynomial-time verification of diagnosability of partially observed discrete-event systems," *IEEE Trans. Autom. Control*, vol. 47, no. 9, pp. 1491–1495, Sep. 2002.

Ziyue Ma (Member, IEEE) received the B.S. and M.S. degrees in chemistry from Peking University, Beijing, China, in 2007 and 2011, respectively, and the Ph.D. Diploma in cotutorship between the School of Electro-Mechanical Engineering of Xidian University, China (in mechatronic engineering) and the Department of Electrical and Electronic Engineering, University of Cagliari, Italy (in electronics and computer engineering), in 2017.

He joined Xidian University, Xi'an, China, in 2011, where he is currently an Associate Professor with the School of Electro-Mechanical Engineering. His research interests include control theory in discrete event systems, automaton and Petri net theories, fault diagnosis/prognosis, resource optimization, and information security.

Dr. Ma is a Technical Committee Member of IEEE Control System Society (IEEE-CSS) on Discrete Event Systems. He is serving/has served as the Associate Editor of the IEEE Conference on Automation Science and Engineering (CASE'17–'20), European Control Conference (ECC'19, '20), and IEEE International Conference on Systems, Man, and Cybernetics (SMC'19, '20). He is/was the Track Committee Member of the International Conference on Emerging Technologies and Factory Automation (ETFA'17–'20). In 2016, he received the Outstanding Reviewer Award from the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Xiang Yin (Member, IEEE) was born in Anhui, China, in 1991. He received the B.Eng. degree from Zhejiang University, Hangzhou, China, in 2012, the M.S. degree from the University of Michigan, Ann Arbor, MI, USA, in 2013, and the Ph.D degree from the University of Michigan, in 2017, all in electrical engineering.

Since 2017, he has been with the Department of Automation, Shanghai Jiao Tong University, China, where he is an Associate Professor. His research interests include formal methods, control of discrete-event systems, model-based fault diagnosis, security, and their applications to cyber and cyber-physical systems.

Dr. Yin received the Outstanding Reviewer awards from *Automatica*, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and the *Journal of Discrete Event Dynamic Systems*. He also received the IEEE Conference on Decision and Control (CDC) Best Student Paper Award Finalist in 2016. He is the Co-Chair of the IEEE CSS Technical Committee on Discrete Event Systems.

Zhiwu Li (Fellow, IEEE) received the B.S. degree in mechanical engineering, the M.S. degree in automatic control, and the Ph.D. degree in manufacturing engineering from Xidian University, Xi'an, China, in 1989, 1992, and 1995, respectively.

He joined Xidian University in 1992. He was a Visiting Professor with the University of Toronto, Technion (Israel Institute of Technology), Martin-Luther University, Conservatoire National des Arts et Métiers (Cnam), Meliksah Universitesi, and King Saud University. His work was cited by engineers and researchers from more than 50 countries and areas, including prestigious R&D institutes such as IBM, Volvo, HP, GE, GM, ABB, and Huawei. Currently, he is with the Institute of Systems Engineering, Macau University of Science and Technology, Taipa, Macau. He has authored/coauthored two monographs in Springer and CRC Press and 100+ papers in *Automatica* and IEEE Transactions (mostly regular). His research interests include discrete event systems and Petri nets.

Dr. Li serves (served) as an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, PART A: SYSTEMS AND HUMAN BEINGS, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, IEEE ACCESS, and *Information Sciences* (Elsevier). He has received Alexander von Humboldt Research Grant and Research in Paris.