Automatica 131 (2021) 109713

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Marking diagnosability verification in labeled Petri nets☆

Ziyue Ma [a,*], Xiang Yin [b], Zhiwu Li [a,c]

[a] School of Electro-Mechanical Engineering, Xidian University, Xi'an 710071, China
[b] Department of Automation and the Key Laboratory of System Control and Information Processing, Shanghai Jiao Tong University, Shanghai 201108, China
[c] Institute of Systems Engineering, Macau University of Science and Technology, Taipa, Macau, China

ARTICLE INFO

Article history:
Received 16 January 2020
Received in revised form 13 April 2021
Accepted 18 April 2021
Available online 7 June 2021

Keywords:
Marking diagnosability
Petri net
Discrete event system
Basis reachability graph

ABSTRACT

This paper studies the marking diagnosability verification problem in labeled Petri nets. Marking diagnosability is a property implying the fact that a plant Petri net has ever reached a pre-defined set of *faulty markings* can be detected in a finite number of future steps. We first show that the conventional basis-reachability-graph-based methods cannot be used due to the existence of *partially faulty basis markings*. To overcome such a problem, we propose a transition partition rule to obtain two particular graphs called the *positive basis reachability graph* and the *negative basis reachability graph*. Then we develop an information structure, called a *dual verifier*, that is a parallel composition of the two basis reachability graphs and can be used to determine the marking diagnosability of a plant net. The proposed method has polynomial complexity in the number of basis markings.

1. Introduction

Fault diagnosis in *discrete event systems* (DESs) (Jiang, Huang, Chandra, & Kumar, 2001; Lafortune, Lin, & Hadjicostis, 2018; Sampath, Sengupta, Lafortune, Sinnamohideen, & Teneketzis, 1995) is to determine whether some particular events, called *faults*, have occurred according to the current observation. Since diagnosability is highly related to safety and robustness of cyber–physical systems, abundant work have been done in the last decades, e.g., its verification (Moreira, Jesus, & Basilio, 2011; Yin & Lafortune, 2015; Zad, Kwong, & Wonham, 2005), robust diagnosability (Carvalho, Moreira, & Basilio, 2017), and its enforcement (Yin & Lafortune, 2019).

Petri net is a widely-used model of DES which is more compact than automata. In Petri nets, structural techniques can be used to circumvent the need to enumerating the state space of a plant. Hence, many problems such as *supervisory control* (Basile, Cordone, & Piroddi, 2015) and *opacity verification* (Tong, Li, Seatzu, & Giua, 2017) can be solved with relatively low computational load. Fault diagnosis has also been extensively studied in *labeled Petri net* (LPN) in recent years. In Giua and Seatzu (2005) and Jiroveanu and Boel (2010), *minimal explanations* are introduced to abstract the firing of unobservable transitions. On the basis of such a notion, in Cabasino, Giua, and Seatzu (2010), a structure called the *basis reachability graph* is proposed to abstract the state space of a plant net and then is used for fault diagnosis. In Cabasino, Giua, Lafortune, and Seatzu (2012), it is shown that the diagnosability verification problem in an LPN can be converted to a reachability condition of the corresponding *verifier net*. In Basile, Chiacchio, and Tommasi (2012), integer linear programming is used to verify K-diagnosability of bounded nets. The decidability and complexity of diagnosability in general LPNs are studied in Yin and Lafortune (2017). There are also works on time Petri nets (Basile, Cabasino, & Seatzu, 2017), interpreted nets (Ramirez-Trevino, Ruiz-Beltran, Aramburo-Lizarraga, & Lopez-Mellado, 2012), and diagnosability enforcements (Ran, Giua, & Seatzu, 2019).

All the aforementioned methods consider faults as some particular events/transitions in a plant.[1] However, in many practical situations, instead of recognizing events, an operator of a plant may expect to know if the plant has or has ever reached some states of physical importance. For example, in an automated

☆ This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 61873342, 62061136004, and 61803259, 61703321, Shaanxi Provincial Natural Science Foundation, China under Grant No. 2019JQ-022, the Fundamental Research Funds for the Central Universities, China under Grant Nos. JB210413, and JB190407, and the Science and Technology Development Fund, MSAR, China, under Grant No. 0012/2019/A1. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Christoforos Hadjicostis under the direction of Editor Christos G. Cassandras.

* Corresponding author.
E-mail addresses: maziyue@xidian.edu.cn (Z. Ma), yinxiang@sjtu.edu.cn (X. Yin), zhwli@xidian.edu.cn (Z. Li).

https://doi.org/10.1016/j.automatica.2021.109713
0005-1098/

[1] Note that although the term "*fault*" usually means some undesired behavior in the usual sense, in the context of fault diagnosis, *faults* are not necessarily associated with practical failures or errors, but can be any events or behavior patterns that have particular interests.

production line, the content of a buffer may frequently exceed a threshold. An alarm is expected to be issued in finite steps after the threshold is exceeded in order to draw an operator's attention. Such a problem is referred to as the *marking diagnosis problem* if the plant is modeled by a Petri net. Briefly speaking, marking diagnosability is a property such that the fact that a plant net has ever reached a given set of *faulty markings* can be detected in a finite number of future steps.

In automata, the event and the state diagnosis problems are equivalent (Kumar & Takai, 2010; Zad et al., 2005). However, in Petri nets, methods for transition diagnosis such as Basile et al. (2012), Cabasino et al. (2010) are not applicable for marking diagnosis. The fact that the firing of a transition at a particular marking yields a faulty marking does not necessarily imply that the firing of this transition *always* yields a faulty marking. Thus, one cannot transform a marking diagnosis problem to a transition diagnosis problem by defining a set of faulty transitions in the original net. On the other hand, a marking diagnosis problem can be transformed into a state diagnosis problem of automata by constructing the entire reachability graph of the plant net and using the automaton-based methods (Yoo & Lafortune, 2002). However, the reachability graph of a net is in general extremely large even if the structure of a net is relatively small, and by doing so, we lose the advantages for analyzing diagnosability using the structural analysis techniques of Petri nets. Therefore, it is desirable to develop more efficient algorithms for the marking diagnosis problem in Petri nets.

In this paper, we formulate and systematically investigate the marking diagnosability verification problem in LPNs. Precisely speaking, marking diagnosability is a property such that any visit of a set of markings with particular interests, called *faulty markings*, can be inferred in a finite number of future steps thereafter. To our knowledge, this work is the first one that addresses marking diagnosis in the literature. Specifically, we assume that the set of faulty markings to be diagnosed is described by a linear marking specification called a *generalized mutual exclusion constraint* (Giua, DiCesare, & Silva, 1992). We develop a method based on the notion of *basis reachability graphs* (BRGs), an efficient tool for abstracting the state space of a Petri net (Ma, Tong, Li, & Giua, 2017; Ran, Su, Giua, & Seatzu, 2018; Tong et al., 2017) without constructing the entire reachability graph. However, unlike the case of transition diagnosability verification (Cabasino et al., 2010; Ran et al., 2018), the visit of faulty markings is in general not well preserved in an arbitrary built BRG. In fact, a BRG based on a blindly chosen T_E may not capture all marking trajectories of the original marking space for the purpose of marking diagnosis. As a result, we explore the reason behind and develop a method for marking diagnosability verification using a compact information structure.

The main content of this paper is summarized as follows. First, by introducing two notions pertaining to basis markings called *faulty basis markings* and *partially faulty basis markings*, we reveal that the information loss may happen if a basis trajectory passes a partially faulty basis marking. To overcome such a problem, we propose a transition partition rule to obtain two particular BRGs called the *positive basis reachability graph* (positive BRG) and the *negative basis reachability graph* (negative BRG). Based on positive and negative BRGs an information structure called a *dual verifier* is developed to verify marking diagnosability. The structural complexity of the dual verifier is polynomial in the number of basis markings in the positive and the negative BRGs. Since the number of basis markings is usually much smaller than that of reachable markings (Cabasino et al., 2010; Ma et al., 2017), our method is in general more efficient than using methods for state diagnosis (Kumar & Takai, 2010; Zad et al., 2005) on the reachability automata of LPNs. Therefore, our proposed approach can reduce the computational complexity if a discrete event system is modeled as a Petri net and the state relations (i.e., the reachability automaton) are unknown.

2. Preliminaries

2.1. 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 from places to transitions and from transitions to places 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 $\mathbb{Z} = \{0, \pm1, \pm2, \ldots\}$).

For a transition $t \in T$, we define the *set of its input places* as ${}^\bullet t = \{p \in P \mid Pre(p, t) > 0\}$ and the *set of its 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* $\langle N, M_0\rangle$ 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 transition 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 *firing 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 said to be 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 $\hat N = (P, \hat T, \hat Pre, \hat Post)$.

2.2. Labeled Petri nets

A *labeled Petri net* (LPN) is a 4-tuple $G = (N, M_0, E, \ell)$, where $\langle N, M_0\rangle$ is a marked net, E is the *alphabet*, 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 leads to a partition $T = T_o \dot\cup T_{uo}$, 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 $\mathscr{C}(w)$ as the *set of consistent markings* of an observation $w \in L(G)$, i.e., $\mathscr{C}(w) = \{M \in R(N, M_0) \mid M_0[w\rangle M\}$.

2.3. Basis reachability graph

Definition 2.1 (*Ma et al., 2017*)**.** Given a Petri net $N = (P, T, Pre, Post)$, a pair $\pi = (T_E, T_I)$ is called a *basis partition* of T if (1) $T_I \subseteq T$, $T_E = T \setminus T_I$; and (2) 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.

Note that the acyclicity on the T_I-induced subnet is *not* an assumption on plants. In fact, one can always subjectively choose a set T_E such that the implicit subset is acyclic — based on which the BRG-based marking abstraction technique can be used.

Definition 2.2. Given Petri net $N = (P, T, Pre, Post)$, basis partition $\pi = (T_E, T_I)$, marking M, and 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 transition t at marking M;
- $Y(M, t) = \{\mathbf{y}_\sigma \in \mathbb{N}^{|T|} \mid \sigma \in \Sigma(M, t)\}$ as the set of *explanation vectors* of transition t at marking M;
- $Y_{min}(M, t)$ denotes the set of all minimal elements of $Y(M, t)$, i.e., the *minimal explanation vectors*.

Definition 2.3. Given a Petri net $N = (P, T, Pre, Post)$ with an initial marking M_0 and a basis partition $\pi = (T_E, T_I)$, the set of its *basis markings* $\mathscr{M}$ is recursively defined as: (i) $M_0 \in \mathscr{M}$; (ii) if $M \in \mathscr{M}$, then for all $t \in T_E$, for all $\mathbf{y} \in Y_{min}(M, t)$:

$$M' = M + C \cdot \mathbf{y} + C(\cdot, t) \quad \Rightarrow \quad M' \in \mathscr{M}.$$

The *basis reachability graph* (BRG) is a finite state automaton $\mathscr{B} = (\mathscr{M}, Tr, \Delta, M_0)$, where: (i) the state set $\mathscr{M}$ is the set of *basis markings*; (ii) the event set Tr is the set of pairs $(t, \mathbf{y}) \in T_E \times \mathbb{N}^{|T|}$; (iii) the transition relation $\Delta \subseteq \mathscr{M} \times Tr \times \mathscr{M}$ is:

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

(iv) the initial state is the initial marking $M_0 \in \mathscr{M}$.

In the worst case, the size of a basis reachability space $\mathscr{M}$ is the same as that of the reachability set $R(N, M_0)$ (e.g., when $T_E = T$ and $T_I = \emptyset$). However, in practice we can usually find a basis partition $T_E \subset T$, under which $\mathscr{M}$ is much smaller than the reachability set, i.e., $|\mathscr{M}| \ll |R(N, M_0)|$. Detailed discussions and numerical results can be found in Cabasino et al. (2010), Ma et al. (2017), Ran et al. (2018) and Tong et al. (2017).

For the convenience of presentation, in the sequel of this paper, 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 $(t_{i_1}, \mathbf{y}_{i_1}), (t_{i_2}, \mathbf{y}_{i_2}), \ldots, (t_{i_n}, \mathbf{y}_{i_n})$. We write $\ell(\phi) = \ell(t_{i_1} t_{i_2} \cdots t_{i_n})$.

Definition 2.4 (*Ma et al., 2017*)**.** Given a marked net $\langle N, M_0\rangle$ with $\pi = (T_E, T_I)$, the *implicit reach* of a marking M is a set of markings: $R_I(M) = \{M' \in \mathbb{N}^{|P|} \mid M[\sigma\rangle M', \sigma \in T_I^*\}$.

Proposition 2.1 (*Ma et al., 2017*)**.** *Given a marked net $\langle N, M_0\rangle$ where $N = (P, T, Pre, Post)$ and its BRG $\mathscr{B} = (\mathscr{M}, Tr, \Delta, M_0)$ with respect to $\pi = (T_E, T_I)$, the following two statements are equivalent:*

1. *there exists 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 $\sigma_j \in T_I^*$ (for all $j \in \{1, \ldots, n+1\}$), $t_{i_j} \in T_E$ (for all $j \in \{1, \ldots, n\}$), such that $M_0[\sigma\rangle M$;*
2. *there is a path in the BRG $\mathscr{B}$*

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

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

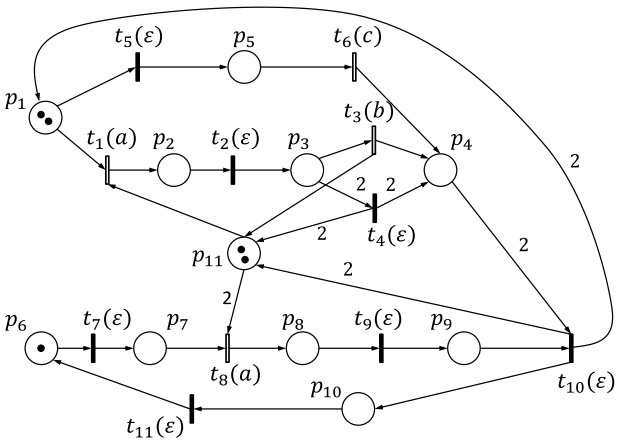

Fig. 1. A labeled Petri net.

The following proposition is a direct generalization of the result in Cabasino et al. (2010).

Proposition 2.2. *Given an LPN $G = (N, M_0, E, \ell)$ with $N = (P, T, Pre, Post)$, let $\mathscr{B} = (\mathscr{M}, Tr, \Delta, M_0)$ 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:*

$$\mathscr{C}(w) = \bigcup_{M_b \in \mathscr{M}(w)} R_I(M_b) = \bigcup_{M_b \in \mathscr{M}(w)} \{M \mid M = M_b + C_I \cdot \mathbf{y}\}$$

where $\mathscr{M}(w) = \{M_b \in \mathscr{M} \mid M_0 \xrightarrow{\phi} M_b, \ell(\phi) = w\}$.

Proof. This proposition follows from Theorem 3.8 in Cabasino et al. (2010) by changing observable/unobservable transitions as explicit/implicit ones. □

3. Marking diagnosability problem formulation

Marking diagnosability is a property such that the fact that a plant Petri net has ever reached a pre-defined set of markings with particular interests, called *faulty markings*, can be detected in a finite number of future steps. Before proceeding formally, we first present an example to illustrate the notion of marking diagnosability in LPNs.

Example 3.1. Consider the LPN in Fig. 1 in which $\ell(t_1) = \ell(t_8) = a$, $\ell(t_3) = b$, $\ell(t_6) = c$, and $\ell(t_i) = \varepsilon$ for all t_i where $i \neq 1, 3, 6, 8$. This net has 57 reachable markings. Suppose that an operator observes a sequence of events $w = aa$ and wants to determine if the plant has reached a set of markings:

$$F = \{M \mid M(p_2) + 2M(p_3) \geq 3\}.$$

We use $\sigma \oplus t$ with $\sigma \in T^*$ and $t \in T$ to denote the set of all sequences generated by inserting transition t in σ. For example, $t_1 t_2 \oplus t_3 = \{t_1 t_2 t_3, t_1 t_3 t_2, t_3 t_1 t_2\}$. Hence, from the prospective of the operator, the set of consistent sequences of observation $w = aa$ is:

$$(t_1 t_1 \oplus t_7) \cup (t_1 t_1 t_2 \oplus t_7) \cup (t_1 t_1 t_2 t_2 \oplus t_7) \cup (t_1 t_2 t_1 t_2 \oplus t_7) \\ \cup (t_1 t_2 t_1 t_2 t_4 \oplus t_7) \cup (t_1 t_2 t_1 t_2 t_4 \oplus t_7).$$

Since $M_0[t_1 t_1\rangle 2p_2 + p_6 \notin F$ and $M_0[t_1 t_1 t_2\rangle p_2 + p_3 + p_6 \in F$, the operator knows that the plant may have reached F but cannot be sure.

However, by observing (any of) the next event the operator immediately knows that the plant must have ever reached F. In fact, if the next observed event is "a", all sequences consistent

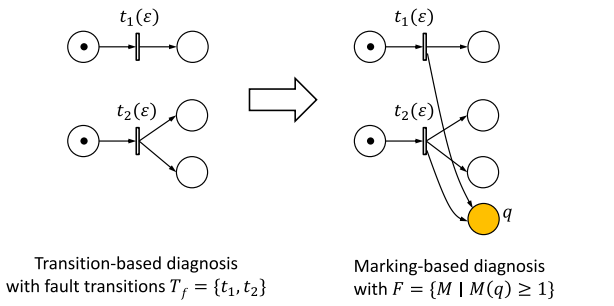

Fig. 2. Illustration of the transformation from a transition diagnosis problem to its equivalent marking diagnosis problem.

with $w = aaa$ must contain t_4, which implies that the plant necessarily passes a faulty marking at which place p_3 holds two tokens. On the other hand, if the next event is "b", the set of consistent sequences of $w = aab$ is:

$$(t_1t_1t_2t_3 \oplus t_7) \cup (t_1t_2t_1t_3 \oplus t_7) \cup (t_1t_1t_2t_3t_2 \oplus t_7) \cup (t_1t_2t_1t_3t_2 \oplus t_7)$$

all of which pass F. In fact, for this LPN, the operator can always know that a faulty marking in F has ever been reached after observing a finite number of events in the future.

Remark 1. Note that, one cannot transform a marking diagnosis problem to a transition diagnosis problem by simply defining a set of faulty transitions in the original net, since that the firing of a transition at one marking yields a faulty marking does not necessarily imply that the firing of such a transition *always* leads to faulty markings. For instance, in the LPN in Fig. 1 of Example 3.1, although firing transition t_1 twice at M_0 yields a faulty marking, we cannot define t_1 as a fault transition and use transition diagnosis method for this net.

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

- A1: G is *deadlock-free*, i.e., at any reachable marking at least one transition can fire;
- A2: G is bounded;

Assumption A1 on deadlock-freeness is a common assumption in the analysis of Petri nets with partial observation (Cabasino et al., 2010; Ran et al., 2018). Assumption A2 guarantees that any BRG of a plant is bounded (Ma et al., 2017) 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. Moreover, we also assume that the set of *faulty markings* to be diagnosed, denoted as $F \subseteq \mathbb{N}^m$, is represented by a *generalized mutual exclusion constraint* (GMEC).

Definition 3.1 (*Giua et al., 1992*)**.** A GMEC is a pair $(\mathbf{w}, k)$, where $\mathbf{w} \in \mathbb{Z}^m$ and $k \in \mathbb{Z}$, that defines a set of markings:

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

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

Remark 2. Although defining the set of faulty markings by a single GMEC seems restrictive, we believe that using GMECs to represent faults is general enough to capture existing notions of transition faults and the case of multiple faults. In fact, the marking diagnosis problem whose set of fault markings is defined by a GMEC is more general than the conventional transition diagnosis problem in the literature. By adding a dummy place q with $Pre(q, \cdot) = \mathbf{0}$ and $Post(q, t) = 1$ (resp., $Post(q, t) = 0$) for all faulty transitions (resp., non-faulty transitions) t, a transition diagnosis problem can be transformed to a marking diagnosis problem of the transformed net with $F = \{M \mid M(q) \geq 1\}$. Such a procedure is illustrated in Fig. 2. Moreover, our method can be extended to cases where F is defined by multiple GMECs (see Remark 4 in Section 5).

We define the *fault language* of a marking M with respect to a given set of faulty markings F as the set of all sequences having at least one prefix that reaches set F, i.e.,

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

Note that this definition does not require that the final marking reached by the firing of σ belongs to set F. The fault language of an initial marking M_0 is denoted by $L_{M_0,F}$. We say "trajectory $M_0[\sigma\rangle$ passes F" or "sequence σ passes F" if $\sigma \in L_{M_0,F}$.

Definition 3.2 (*Marking Diagnosability*)**.** Given an LPN $G = (N, E, \ell, M_0)$ with $N = (P, T, Pre, Post)$ and a set of faulty markings F, plant G is *diagnosable* (with respect to F) if for all σ such that $M_0[\sigma\rangle M \in F$, there exists an integer $K_\sigma \in \mathbb{N}$ such that the following condition holds:

$$\begin{aligned} &\forall \sigma' \in \ell^{-1}(\ell(\sigma)), \forall \sigma'' \in T^* : \\ &\sigma'\sigma'' \in L(N, M_0) \wedge |\sigma''| \geq K_\sigma \;\Rightarrow\; \sigma'\sigma'' \in L_{M_0,F}. \end{aligned} \tag{1}$$

In words, an LPN G is diagnosable if for any $M_0[\sigma\rangle M$ where $M \in F$, there exists an integer K_σ such that for all $\sigma' \in L(N, M_0)$ that looks like σ, for all $\sigma'' \in T^*$ that is a feasible continuation of σ' with length $|\sigma''| \geq K_\sigma$, trajectory $M_0[\sigma'\sigma''\rangle$ necessarily passes F. Note that, by Definition 3.2, G is diagnosable when F is not reachable (since Eq. (1) is satisfied). However, we do not need to check the reachability of F in advance (see Remark 3 in Section 5).

Since a plant LPN G is assumed to be bounded, Eq. (1) in Definition 3.2 can be verified by using the *reachability analysis*. Specifically, one can construct the entire reachability graph, and then reduce the marking diagnosability verification problem to an event diagnosability verification problem in it. However, this method is very exhaustive as the size of the reachability graph of a net is generally extremely large even if the net structure is relatively limited. On the other hand, BRGs have been proved to be an efficient information structure to abstract the state space of Petri nets (Cabasino et al., 2010; Ma et al., 2017; Ran et al., 2018; Tong et al., 2017). Therefore, we aim to develop a method to efficiently solve Problem 1 by leveraging the structure property of the Petri net without computing the entire reachability graph.

4. Marking diagnosability verification in LPNs using dual verifiers

4.1. Fault-ness of basis markings

In the diagnosability verification of transition faults (Ran et al., 2018), a particular type of BRG was used where fault transitions are added to the explicit set T_E such that their firings can be explicitly tracked. On the other hand, as shown in Section 3, an arbitrary BRG may not capture the information of faulty markings for the purpose of marking diagnosis. To encode the information of faulty markings into a BRG, the following definition is necessary.

Definition 4.1. Given an LPN $G = (N, M_0, E, \ell)$, its BRG $\mathscr{B} = (\mathscr{M}, Tr, \Delta, M_0)$ whose basis partition satisfies $T_E \supseteq T_o$, and a set of faulty markings $F = L_{(\mathbf{w},k)}$:

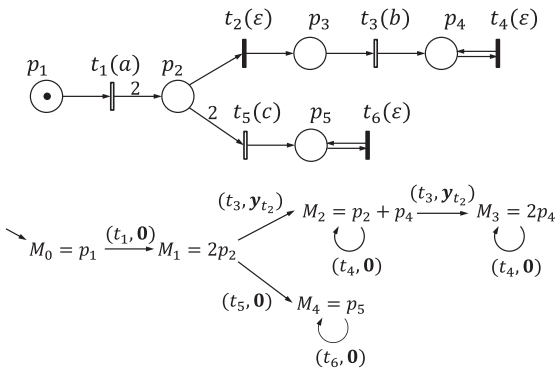

Fig. 3. The LPNs used in Example 4.1.

- the set of *faulty basis markings* is defined as: $\mathscr{F} = \{M_b \in \mathscr{M} \mid R_l(M_b) \subseteq F\}$;
- the set of *partially faulty basis markings* is defined as: $\mathscr{P} = \{M_b \in \mathscr{M} \mid [R_l(M_b) \cap F \neq \emptyset] \wedge [R_l(M_b) \setminus F \neq \emptyset]\}$;
- a basis marking $M_b \in \mathscr{M} \setminus (\mathscr{F} \cup \mathscr{P})$ is said to be a *non-faulty basis marking*.

In plain words, a basis marking M_b is a faulty basis marking (resp., partially faulty basis marking) if all markings (resp., some but not all markings) in its implicit reach belong to F. Since the implicit subnet is acyclic, checking if a marking is faulty or partially faulty can be done by using the state equation of Petri nets.

According to Proposition 2.1, each trajectory in an LPN is associated to a path in a corresponding BRG and vice versa. Hence, one may intuitively conjecture that Eq. (1) can be verified by inspecting the paths in a basis reachability graph, analogous to what was done in *event diagnosability verification* (Ran et al., 2018). Unfortunately, such a conjecture is false.

Example 4.1. Consider the LPN G shown in Fig. 3. Its BRG w.r.t. $T_E = \{t_1, t_3, t_4, t_5, t_6\}$, shown in the same figure, has five basis markings. Let us first consider set $F_1 = \{M \mid M(p_3) \geq 2\}$. Then $\mathscr{F}_1 = \emptyset$ and $\mathscr{P}_1 = \{M_1\}$. For path $M_0 \to M_1 \to M_2 \to M_3$ whose observation is *abb*, there exist two trajectories $M_0[t_1t_2t_2t_3t_3\rangle$ and $M_0[t_1t_2t_3t_2t_3\rangle$, both of which coincide with *abb*, such that the former passes F_1 while the latter does not.

As another example, consider $F_2 = \{M \mid M(p_3) \geq 1\}$ such that $\mathscr{F}_2 = \emptyset$ and $\mathscr{P}_2 = \{M_1, M_2\}$. Consider the same path $M_0 \to M_1 \to M_2 \to M_3$ whose observation is *abb*. In this case, however, one can readily verify that all trajectories coinciding with *abb* necessarily pass F_2 due to the firing of t_3.

As the last example, consider the same $F_2 = \{M \mid M(p_3) \geq 1\}$ but a different path $M_0 \to M_1 \to M_4$ whose observation is *ac*. In this case, all trajectories coinciding with *ac* are $t_1t_5t_6^*$ that do not pass F_2.

Example 4.1 shows that, for the purpose of marking diagnosis, tracking basis markings in the abstracted basis space of a BRG with $T_E \supseteq T_o$ may not capture all trajectories of the original marking space. This is quite different from the previous BRG-based methods in the literature where many important properties of Petri nets including *marking estimation* (Cabasino et al., 2010), *transition diagnosability* (Ran et al., 2018), and *opacity* (Tong et al., 2017) can be verified by simply inspecting the basis markings in an arbitrary BRG with $T_E \supseteq T_o$. Therefore, additional augmentations have to be done to encode the faultness of trajectories into classical BRGs. We point out that if there are no partially faulty basis markings, then the methods in Kumar and Takai (2010) and Zad et al. (2005) can be used by treating the BRG as an automaton in which all fully faulty basis markings are faulty states (due to the limit of space we do not address this in detail). On the other hand, there is no efficient method to guarantee the non-existence of partially faulty basis markings except adding all transitions with non-zero influences into set T_E. However, in Ma et al. (2017) we have proved that elaborating set T_E usually increases the size of the resulting BRG simultaneously. In the next section we develop a verifier-like structure called a *dual verifier* to solve the problem without adding all transitions with non-zero influences into set T_E.

4.2. Positive and negative BRGs

In this subsection, we show that the faulty and the non-faulty trajectory can be separately encoded into two BRGs that are then composed to obtain a compact information structure. We first define two particular BRGs, namely the *positive BRG* and the *negative BRG*.

Definition 4.2. Given an LPN $G = (N, M_0, E, \ell)$ and a set of faulty markings $F = \mathscr{L}_{(\mathbf{w},k)}$, the *positive BRG*, denoted by $\mathscr{B}^+ = (\mathscr{M}^+, Tr^+, \Delta^+, M_0)$, is the BRG with respect to basis partition $\pi^+ = (T_E^+, T_l^+)$ where

$$T_E^+ = T_o \cup \{t \in T \mid \mathbf{w}^T \cdot C(\cdot, t) > 0\}.$$

The *negative BRG*, denoted by $\mathscr{B}^- = (\mathscr{M}^-, Tr^-, \Delta^-, M_0)$, is the BRG with respect to basis partition $\pi^- = (T_E^-, T_l^-)$ where

$$T_E^- = T_o \cup \{t \in T \mid \mathbf{w}^T \cdot C(\cdot, t) < 0\}.$$

In plain words, a positive (resp., negative) BRG is a BRG with respect to the basis partition in which the set of explicit transitions is the union of T_o and the set of transitions with positive (resp., negative) influence. Since the basis partitions π^+ and π^- are different, the sets of basis markings in a positive BRG and of a negative BRG are in general different (i.e., $\mathscr{M}^+ \neq \mathscr{M}^-$). To avoid confusion, a basis marking in a positive BRG (resp., negative BRG) is denoted as M_b^+ (resp., M_b^-). The sets of faulty basis markings and of partially faulty basis markings in $\mathscr{B}^+$ (resp., in $\mathscr{B}^-$) are denoted as $\mathscr{F}^+$ and $\mathscr{P}^+$ (resp., $\mathscr{F}^-$ and $\mathscr{P}^-$) accordingly. The T_l^+-induced subnet (resp., the T_l^--induced subnet) is denoted as C_l^+ (resp., C_l^-). The implicit reach of a basis marking M_b^+ in $\mathscr{B}^+$ (resp., M_b^- in $\mathscr{B}^-$) is denoted as $R_{l^+}(M_b^+)$ (resp., $R_{l^-}(M_b^-)$).

In the following we show how the non-faulty and faulty trajectories are associated to paths in positive and negative BRGs, respectively. First, the following proposition shows that any trajectory that passes set F is associated to a path in the negative BRG $\mathscr{B}^-$ which passes at least one basis marking $M_b^- \in F$ (note that $M_b^- \in F$ does not require $M_b^- \in \mathscr{F}$).

Proposition 4.1. *Given an LPN $G = (N, M_0, E, \ell)$, a set of faulty markings $F = \mathscr{L}_{(\mathbf{w},k)}$, and a negative BRG $\mathscr{B}^-$, there exists a sequence $\sigma \in T^*$ such that $M_0[\sigma\rangle M \in F$ if and only if there exists a path ϕ in $\mathscr{B}^-$ such that $M_0 \xrightarrow{\phi} M_b^- \in F$ and $\ell(\sigma) = \ell(\phi)$.*

Proof. ($\Leftarrow$) is trivial. For ($\Rightarrow$), since $M_0[\sigma\rangle M \in F$, by Proposition 2.1 in $\mathscr{B}^-$ there exists a path labeled by ϕ such that $M_0 \xrightarrow{\phi} M_b^-$, $M \in R_{l^-}(M_b^-)$, and $\ell(\sigma) = \ell(\phi)$. Since $\mathbf{w}^T \cdot C(\cdot, t) \geq 0$ holds for all $t \in T_l^-$, for all markings $M \in R_{l^-}(M_b^-)$, $\mathbf{w}^T \cdot M \geq \mathbf{w}^T \cdot M_b^-$ holds. Hence $M \in F$ implies $M_b^- \in F$. □

Second, we introduce the notion of *faulty arcs* in a positive BRG $\mathscr{B}^+$. We show that any trajectory that does not pass F is associated to a path that does not pass any faulty arcs in $\mathscr{B}^+$.

Definition 4.3. Given a set of faulty markings $F = \mathscr{L}_{(\mathbf{w},k)}$ and a positive BRG $\mathscr{B}^+$, an arc $M_b^+ \xrightarrow{(t,\mathbf{y})}$ is a *faulty arc* if $M_b^+ + C_l \cdot \mathbf{y} \leq k$.

Proposition 4.2. *Given an LPN $G = (N, M_0, E, \ell)$, a set of faulty markings $F = \mathscr{L}_{(\mathbf{w},k)}$, and a positive BRG $\mathscr{B}^+$, there exists a sequence $\sigma \in L(N, M_0) \setminus L_{M_0,F}$ if and only if in $\mathscr{B}^+$ there exists a path*

$$M_{b_0} \xrightarrow{(t_{i_1},\mathbf{y}_1)} M_{b_1}^+ \xrightarrow{(t_{i_2},\mathbf{y}_2)} \cdots M_{b_{n-1}}^+ \xrightarrow{(t_{i_n},\mathbf{y}_n)} M_{b_n}^+$$

such that $\ell(\sigma) = \ell(\phi)$, $\phi = (t_{i_1}, \mathbf{y}_1)\cdots(t_{i_n}, \mathbf{y}_n)$ and all arcs on this path are not faulty arcs.

Proof. The ($\Leftarrow$) is trivial. To prove ($\Rightarrow$), let $\sigma = \sigma_1 t_1 \cdots \sigma_n t_n \sigma_{n+1}$ where $\sigma_i \in T_I^*$, $t_i \in T_E$. The following trajectory does not pass F:

$$M_0[\sigma_1\rangle[t_1\rangle M_1[\sigma_2\rangle[t_2\rangle \cdots [\sigma_n\rangle[t_n\rangle M_n[\sigma_{n+1}\rangle M.$$

From σ_1 we can extract a sequence $\sigma_{min,1}$ that is a minimal explanation of t_1 and move the transitions that do not belong to $\sigma_{min,1}$ after the firing of t_1, i.e.,

$$M_0[\sigma_{min,1}\rangle[t_1\rangle M_{b_1}^+[\sigma_1'\rangle M_1[\sigma_2\rangle[t_2\rangle \cdots [\sigma_n\rangle[t_n\rangle M_n[\sigma_{n+1}\rangle M.$$

According to Cabasino et al. (2010), this operation is feasible since the implicit subnet is acyclic. Hence, in the rearranged trajectory, marking $M_{b_1}^+$ is a basis marking in $\mathscr{B}^+$. Since for all $t \in T_I^+$, $\mathbf{w}^T \cdot C(\cdot, t) \leq 0$ holds, which indicates $\mathbf{w}^T \cdot C_I^+ \cdot \mathbf{y} \leq \mathbf{w}^T \cdot C_I^+ \cdot \hat{\mathbf{y}}$ for all $\mathbf{y} \geq \hat{\mathbf{y}}$. Then $M_0[\sigma_{min,1}\rangle[t_1\rangle M_{b_1}^+[\sigma_1'\rangle$ does not pass F, i.e., $M_{b_1}^+[\sigma_1'\rangle[\sigma_2\rangle[t_2\rangle \cdots$ does not pass F. Let $M_{b_1}^+$ be the new initial basis marking and such reasoning can be repeatedly applied. Finally we reach a trajectory:

$$M_0[\sigma_{min,1}t_1\rangle M_{b_1}^+[\sigma_{min,2}t_2\rangle \cdots [\sigma_{min,n}t_n\rangle M_{b_n}^+[\sigma_{n+1}'\rangle M$$

such that all $M_{b_i}^+ + C_I^+ \cdot \mathbf{y}_{\sigma_{min,i}} \notin F$. This completes the proof. □

4.3. Marking diagnosability verification using dual verifiers

Now we are ready to introduce the *dual verifier* that will be used for the verification of marking diagnosability. Roughly speaking, a dual verifier is an automaton parallelly composed by a positive BRG and a negative BRG, in which each state is a tuple $((M_b^+, \gamma^+), (M_b^-, \gamma^-))$.

Definition 4.4. The *dual-next* function $\Omega_d : (\mathscr{M}^+ \times \Gamma) \times (\mathscr{M}^- \times \Gamma) \times (E \cup \{\varepsilon\}) \to (\mathscr{M}^+ \times \Gamma) \times (\mathscr{M}^- \times \Gamma)$, where $\Gamma = \{0, 1\}$, is defined as:

$$\Omega_d((M_b^+, \gamma^+), (M_b^-, \gamma^-), e) = \{(\hat{M}_b^+, \hat{\gamma}^+), (\hat{M}_b^-, \hat{\gamma}^-)\}$$

where

$$\begin{cases} (M_b^+, (t^+, \mathbf{y}^+), \bar{M}_b^+) \in \Delta^+, (M_b^-, (t^-, \mathbf{y}^-), \hat{M}_b^-) \in \Delta^-, \\ \ell(t^+) = \ell(t^-) = e, \\ \hat{\gamma}^+ = \begin{cases} 0, & \text{if } \gamma^+ = 0 \wedge M_b^+ \xrightarrow{(t^+,\mathbf{y}^+)} \text{ is not a faulty arc,} \\ 1, & \text{otherwise} \end{cases} \\ \hat{\gamma}^- = \begin{cases} 0, & \text{if } \gamma^- = 0 \wedge \hat{M}_b^- \notin F, \\ 1, & \text{otherwise} \end{cases} \end{cases}$$

Definition 4.5 (*Dual Verifier*)**.** Given an LPN $G = (N, M_0, E, \ell)$ and a set of faulty markings $F = \mathscr{L}_{(\mathbf{w},k)}$, let $\mathscr{B}^+ = (\mathscr{M}^+, Tr^+, \Delta^+, M_0)$ and $\mathscr{B}^- = (\mathscr{M}^-, Tr^-, \Delta^-, M_0)$ be a positive BRG and a negative BRG, respectively. A *dual verifier* is an automaton $\mathscr{D} = (D, E, \delta, d_0)$ where (1) the state set $D \subseteq (\mathscr{M}^+ \times \Gamma) \times (\mathscr{M}^- \times \Gamma)$ where $\Gamma = \{0, 1\}$; (2) the initial state $d_0 = ((M_0, 0), (M_0, 0))$; (3) the nondeterministic transition rule δ is defined as: for each $e \in E \cup \{\varepsilon\}$:

$$\delta(((M_b^+, \gamma^+), (M_b^-, \gamma^-)), e) = \Omega_d((M_b^+, \gamma^+), (M_b^-, \gamma^-), e).$$

The main features of a dual verifier are the following:

1. in each state of a dual verifier, components (M_b^+, γ^+) and (M_b^-, γ^-) track the evolutions in the positive BRG and in the negative BRG, respectively;
2. in the positive BRG component (M_b^+, γ^+), flag γ^+ is switched from 0 to 1 when and only when a faulty arc in the positive BRG is passed;
3. in the negative BRG component (M_b^-, γ^-), flag γ^- is switched from 0 to 1 when and only when a basis marking $M_b^- \in F$ in the negative BRG is passed.

Definition 4.6. A cycle in a dual verifier $d_1 \to d_2 \to \cdots \to d_n \to d_1$ is said to be a *confused cycle* if $d_i = ((M_{b_i}^+, 0), (M_{b_i}^-, 1))$ for all $i = 1, \ldots, n$.

Finally, we state the following result showing that marking diagnosability in an LPN can be verified by checking the (non-) existence of confused cycles in the corresponding dual verifier.

Theorem 4.1. *Given an LPN $G = (N, M_0, E, \ell)$ and a set of faulty markings $F = \mathscr{L}_{(\mathbf{w},k)}$, G is diagnosable with respect to F if and only if the corresponding dual verifier $\mathscr{D} = (D, E, \delta, d_0)$ does not contain confused cycles.*

Proof. (Only If) By contradiction, suppose that G is diagnosable w.r.t. F and there exists a confused cycle in $\mathscr{D}$: $d_1 \to d_2 \to \cdots \to d_n \to d_1$, where $d_i = ((M_{b_i}^+, 0), (M_{b_i}^-, 1))$ for $i = 1, \ldots, n$. Following the same argument in Yoo and Lafortune (2002), in the positive BRG and in the negative BRG there exist two infinite paths $M_0 \xrightarrow{\phi_0^+} M_{b_1}^+ \xrightarrow{(\phi^+)^n}$ and $M_0 \xrightarrow{\phi_0^-} M_{b_1}^- \xrightarrow{(\phi^-)^n}$, respectively, such that $\ell(\phi_0^+(\phi^+)^n) = \ell(\phi_0^-(\phi^-)^n)$ and the first trajectory does not pass F while the second trajectory passes F. Then, according to Propositions 4.2 and 4.1, there exist two trajectories $M_0[\sigma\rangle$ and $M_0[\sigma'\rangle$ that are associated to $M_0 \xrightarrow{\phi_0^+} M_{b_1}^+ \xrightarrow{(\phi^+)^n}$ and $M_0 \xrightarrow{\phi_0^-} M_{b_1}^- \xrightarrow{(\phi^-)^n}$, respectively. This implies that the required K_σ in Definition 3.2 does not exist, which contradicts that assumption that G is diagnosable.

(If) Suppose that G is not diagnosable w.r.t. F. This implies that there exist two sequences σ, σ' such that $M_0[\sigma\rangle M \in F$, $M_0[\sigma'\rangle M' \notin F$ with $\ell(\sigma) = \ell(\sigma')$, and there exist two sequences $\hat{\sigma}, \hat{\sigma}'$ where $\ell(\hat{\sigma}') = \ell(\hat{\sigma})$ and $\sigma(\hat{\sigma})^n, \sigma'(\hat{\sigma}')^n \in L(N, M_0)$ for arbitrary $n \in \mathbb{N}$. According to Propositions 4.2 and 4.1, in the positive and the negative BRG there exist two infinite paths $M_0 \xrightarrow{\phi_0^+} M_{b_1}^+ \xrightarrow{(\phi^+)^n} M_{b_1}^+$ and $M_0 \xrightarrow{\phi_0^-} M_{b_1}^- \xrightarrow{(\phi^-)^n} M_{b_1}^-$, respectively, such that (i) $\ell(\phi_0^+) = \ell(\phi_0^-) = \ell(\sigma)$, $\ell(\phi^+) = \ell(\phi^-) = \ell(\sigma')$; (ii) in $\mathscr{B}^+$ path $M_0 \xrightarrow{\phi_0^+} M_{b_1}^+$ does not pass any faulty arcs; (iii) in $\mathscr{B}^-$ path $M_0 \xrightarrow{\phi_0^-} M_{b_1}^-$ passes at least one basis marking $M_b^- \in F$. Hence, by the construction of Definition 4.5, in $\mathscr{D}$ there necessarily exists a path from $((M_0, 0), (M_0, 0))$ to $((M_{b_1}^+, 0), (M_{b_1}^-, 1))$. Moreover, since the size of $\mathscr{D}$ is finite, in $\mathscr{D}$ there necessarily exists a cycle $((M_{b_1}^+, 0), (M_{b_1}^-, 1)) \to \cdots \to ((M_{b_1}^+, 0), (M_{b_1}^-, 1))$ that is a confused cycle in $\mathscr{D}$. □

Remark 3. If F is not reachable, then $\mathscr{B}^+$ does not contain any fault marking while $\mathscr{B}^-$ does not contain any faulty arc. In such a case, the flags in the corresponding dual verifier are all zero, which tells that the plant is diagnosable.

Remark 4. When F is defined by a conjunction of GMECs $(\mathbf{W}, \mathbf{k}) \in \mathbb{N}^{m \times r} \times \mathbb{N}^{m \times 1}$, $r \in \mathbb{N}$, i.e., $F = \{M \mid \mathbf{W}^T \cdot M \leq \mathbf{k}\}$, by defining the positive and negative BRGs on $T_E^+ = T_o \cup \{t \in T \mid$

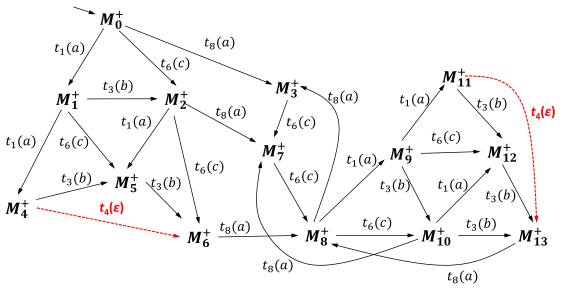

Fig. 4. The positive BRG $\mathscr{B}^+$ of the LPN in Fig. 1 with $T_E^+ = \{t_1, t_3, t_4, t_6, t_8\}$. Dashed arcs denote the faulty arcs.

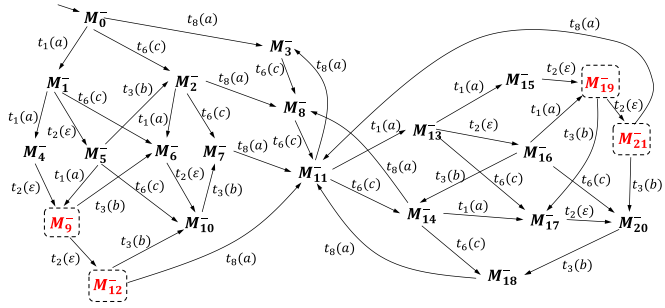

Fig. 5. The negative BRG $\mathscr{B}^-$ of the LPN in Fig. 1 with $T_E^- = \{t_1, t_2, t_3, t_6, t_8\}$. Basis markings $M_{b_i}^- \in F$ are marked by dashed boxes.

$\mathbf{W}^T \cdot C(\cdot, t) \not\leq \mathbf{0}\}$ and $T_E^- = T_o \cup \{t \in T \mid \mathbf{W}^T \cdot C(\cdot, t) \not\geq \mathbf{0}\}$, all results in this section analogously hold. On the other hand, when F is defined by a disjunction of GMECs, i.e., $F = \bigcup_{i=1}^{r} \mathscr{L}_{(\mathbf{w}_i, k_i)}$, the diagnosability of a plant can be verified by checking the diagnosability of each $(\mathbf{w}_i, k_i)$, since the plant is diagnosable if and only if all $(\mathbf{w}_i, k_i)$'s are diagnosable.

In practice, an operator of a plant is often interested in whether or not the resources in some places are overflowed or underflowed, and thus set F is usually either right-closed (i.e., $M \in F$ implies $\forall M' \geq M, M' \in F$) or left-closed ($M \in F$ implies $\forall M' \leq M, M' \in F$). A right- or left-closed set F can always be characterized by a set of conjunctive or disjunctive GMECs that can be easily established when modeling a system.

Example 4.2. Let us consider again the LPN in Fig. 1 and a set of faulty markings $F = \{M \mid M(p_2) + 2M(p_3) \geq 3\}$, i.e., $F = \mathscr{L}_{(\mathbf{w},k)}$ where $\mathbf{w} = [0, -1, -2, 0, 0, 0, 0, 0, 0, 0, 0]^T$, $k = -3$. Since $\mathbf{w}^T \cdot C(\cdot, t_4) = 2$ and $\mathbf{w}^T \cdot C(\cdot, t_2) = -1$, we choose $T_E^+ = \{t_1, t_3, t_4, t_6, t_8\}$, $T_E^- = \{t_1, t_2, t_3, t_6, t_8\}$, and compute the corresponding positive BRG $\mathscr{B}^+$ and negative BRG $\mathscr{B}^-$, as depicted in Figs. 4 and 5, respectively. For better readability, in these figures minimal explanation vectors are not drawn. The positive BRG $\mathscr{B}^+$ (Fig. 4) contains 14 basis markings while two arcs in it are faulty: $M_4 \xrightarrow{(t_4, \mathbf{y}_{t_2 t_2})}$ and $M_{11} \xrightarrow{(t_4, \mathbf{y}_{t_2 t_2})}$ where $M_4 = 2p_2 + p_6$, $M_{11} = 2p_2 + p_{10}$. The negative BRG (Fig. 5) contains 22 basis markings among which four basis markings $M_9 = p_2 + p_3 + p_6$, $M_{12} = 2p_3 + p_6$, $M_{19} = p_2 + p_3 + p_{10}$, and $M_{21} = 2p_3 + p_{10}$ belong to set F.

The corresponding dual verifier consists of 111 states, and hence only part of it is graphically presented in Fig. 6. This dual verifier does not contain any confused cycles, and hence by Theorem 4.1 the plant LPN is diagnosable w.r.t. F. For instance, for trajectory $M_0[t_1 t_1 t_2\rangle p_2 + p_3 + p_6 \in F$ whose observation is aa, the corresponding state in the dual verifier is $(M_4^+, 0, M_9^-, 1)$. Then, from $(M_4^+, 0, M_9^-, 1)$, the dual verifier will eventually move to a state that double-flagged 1 such that the visit of some faulty markings can be determined.

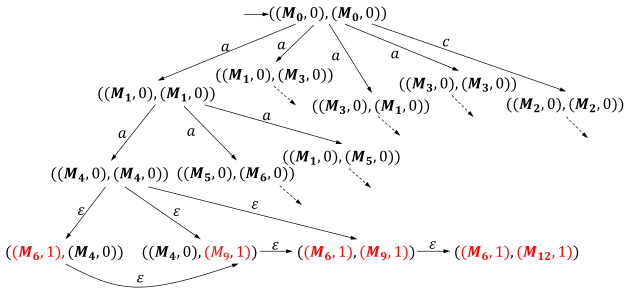

Fig. 6. (Part of) the dual verifier constructed from $\mathscr{B}^+$ in Fig. 4 and $\mathscr{B}^-$ in Fig. 5.

For a positive BRG and a negative BRG with $|\mathscr{M}^+|$ and $|\mathscr{M}^-|$ basis markings, respectively, the corresponding dual verifier has at most $2 \cdot |\mathscr{M}^+| \cdot |\mathscr{M}^-|$ states. Hence, the complexity of our approach is $O(|\mathscr{M}^+| \cdot |\mathscr{M}^-|)$. In comparison, the complexity of first establishing an event-diagnosis problem in the reachability automata of a plant LPN followed by using the conventional *verifier-based algorithm* (Yoo & Lafortune, 2002) is $O(|R(N, M_0)|^2)$. Since many works (Cabasino et al., 2010; Ma et al., 2017; Tong et al., 2017) in the literature have shown that $|\mathscr{M}^+|, |\mathscr{M}^-| \ll |R(N, M_0)|$ in general, our method is practically more efficient than applying conventional methods on the manipulated reachability automaton. For instance, the LPN in Fig. 1 has 57 reachable markings, while the sets $\mathscr{M}^+$ and $\mathscr{M}^-$ in Example 4.2 only have 14 and 22 basis markings, respectively. Note that in some extreme cases (e.g., all transitions are observable/all unobservable transitions are with self-loops) it may happen that the only valid T_E^+ and T_E^- are $T_E^+ = T_E^- = T$. In such cases our method (in which $\mathscr{M}^+ = \mathscr{M}^- = R(N, M_0)$) has the same complexity of the conventional methods by manipulating the reachability automaton.

5. Conclusion

We propose a framework for marking diagnosis of labeled Petri nets using basis reachability graph. An effective algorithm is developed to verify marking diagnosability of a plant based on the dual verifier. 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

Basile, F., Cabasino, M. P., & Seatzu, C. (2017). Diagnosability analysis of labeled time Petri net systems. *IEEE Transactions on Automatic Control*, *62*(3), 1384–1396.

Basile, F., Chiacchio, P., & Tommasi, G. De (2012). On K-diagnosability of Petri nets via integer linear programming. *Automatica*, *48*(9), 2047–2058.

Basile, F., Cordone, R., & Piroddi, L. (2015). A branch and bound approach for the design of decentralized supervisors in Petri net models. *Automatica*, *52*, 322–333.

Cabasino, M., Giua, A., Lafortune, S., & Seatzu, C. (2012). A new approach for diagnosability analysis of Petri nets using verifier nets. *IEEE Transactions on Automatic Control*, *57*(12), 3104–3117.

Cabasino, M., Giua, A., & Seatzu, C. (2010). Fault detection for discrete event systems using Petri nets with unobservable transitions. *Automatica*, *46*(9), 1531–1539.

Carvalho, L. K., Moreira, M. V., & Basilio, J. C. (2017). Diagnosability of intermittent sensor faults in discrete event systems. *Automatica*, *79*(C), 315–325.

Giua, A., DiCesare, F., & Silva, M. (1992). Generalized mutual exclusion constraints for Petri nets with uncontrollable transitions. In *Proceedings of the IEEE int. conf. on systems, man, and cybernetics*, Chicago, USA (pp. 947–949).

Giua, A., & Seatzu, C. (2005). Fault detection for discrete event systems using Petri nets with unobservable transitions. In *Proceedings of Joint 44th conference on decision and control and 2005 European control conference*, Seville, Spain (pp. 6323–6328).

Jiang, S., Huang, Z., Chandra, V., & Kumar, R. (2001). A polynomial algorithm for testing diagnosability of discrete-event systems. *IEEE Transactions on Automatic Control*, *46*(8), 1318–1321.
Jiroveanu, George, & Boel, René K. (2010). The diagnosability of Petri net models using minimal explanations. *IEEE Transactions on Automatic Control*, *55*(7), 1663–1668.
Kumar, R., & Takai, S. (2010). Decentralized prognosis of failures in discrete event systems. *IEEE Transactions on Automatic Control*, *55*(1), 48–59.
Lafortune, S., Lin, F., & Hadjicostis, C. N. (2018). On the history of diagnosability and opacity in discrete event systems. *Annual Reviews in Control*, *45*, 257–266.
Ma, Z., Tong, Y., Li, Z., & Giua, A. (2017). Basis marking representation of Petri net reachability spaces and its application to the reachability problem. *IEEE Transactions on Automatic Control*, *62*(3), 1078–1093.
Moreira, M. V., Jesus, T. C., & Basilio, J. C. (2011). Polynomial time verification of decentralized diagnosability of discrete event systems. *IEEE Transactions on Automatic Control*, *56*(7), 1679–1684.
Ramirez-Trevino, A., Ruiz-Beltran, E., Aramburo-Lizarraga, J., & Lopez-Mellado, E. (2012). Structural diagnosability of DES and design of reduced Petri net diagnosers. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *42*(2), 416–429.
Ran, N., Giua, A., & Seatzu, C. (2019). Enforcement of diagnosability in labeled Petri nets via optimal sensor selection. *IEEE Transactions on Automatic Control*, *64*(7), 2997–3004.
Ran, N., Su, H., Giua, A., & Seatzu, C. (2018). Codiagnosability analysis of bounded Petri nets. *IEEE Transactions on Automatic Control*, *63*(8), 1192–1199.
Sampath, M., Sengupta, R., Lafortune, S., Sinnamohideen, K., & Teneketzis, D. (1995). Diagnosability of discrete event systems. *IEEE Transactions on Automatic Control*, *40*(9), 1555–1575.
Tong, Y., Li, Z. W., Seatzu, C., & Giua, A. (2017). Verification of state-based opacity using Petri nets. *IEEE Transactions on Automatic Control*, *62*(6), 2823–2837.
Yin, X., & Lafortune, S. (2015). Codiagnosability and coobservability under dynamic observations: Transformation and verification. *Automatica*, *61*, 241–252.
Yin, X., & Lafortune, S. (2017). Synthesis of maximally-permissive supervisors for the range control problem. *IEEE Transactions on Automatic Control*, *62*(8), 3914–3929.
Yin, X., & Lafortune, S. (2019). A general approach for optimizing dynamic sensor activations for discrete event systems. *Automatica*, *105*, 376–383.
Yoo, T. S., & Lafortune, S. (2002). Polynomial-time verification of diagnosability of partially observed discrete-event systems. *IEEE Transactions on Automatic Control*, *47*(9), 1491–1495.
Zad, S. H., Kwong, R. H., & Wonham, W. M. (2005). Fault diagnosis in discrete-event systems: Incorporating timing information. *IEEE Transactions on Automatic Control*, *50*(7), 1010–1015.

Ziyue Ma received the B.S. degree and the M.S. degree from Peking University, China, in 2007 and 2011, respectively. In 2017 he got the Ph.D. degree in cotutorship between the School of Electro-Mechanical Engineering of Xidian University, China (in Mechatronic Engineering), and the Department of Electrical and Electronic Engineering of University of Cagliari, Italy (in Electronics and Computer Engineering). He joined Xidian University in 2011, where he is currently an Associate Professor in 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 member of 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–'21), European Control Conference (ECC'19–'21), and IEEE International Conference on Systems, Man, and Cybernetics (SMC'19–'21). He is/was the Track Committee Member of the International Conference on Emerging Technologies and Factory Automation (ETFA'17–'21).

Xiang Yin was born in Anhui, China, in 1991. He received the B.Eng. degree from Zhejiang University in 2012, the M.S. degree from the University of Michigan, Ann Arbor, in 2013, and the Ph.D. degree from the University of Michigan, Ann Arbor, in 2017, all in electrical engineering. Since 2017, he has been with the Department of Automation, Shanghai Jiao Tong University, where he is an Associate Professor. His research interests include formal methods, discrete-event systems, and cyber–physical systems.

Dr. Yin is serving as the co-chair of the *IEEE CSS Technical Committee on Discrete Event Systems*, an Associate Editor for the *Journal of Discrete Event Dynamic Systems: Theory & Applications*, and a member of the *IEEE CSS Conference Editorial Board*. Dr. Yin 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 received the B.S., M.S., and Ph.D. degrees from Xidian University in 1989, 1992, and 1995, respectively. He joined Xidian University in 1992. His interests include discrete event systems and Petri nets. He published two monographs in Springer and CRC Press and 150+ papers in Automatica and IEEE Transactions (mostly regular). He was a Visiting Professor at 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. Now, he is also with the Institute of Systems Engineering, Macau University of Science and Technology, Taipa, Macau.

Dr. Li serves (served) an Associate Editor of the IEEE Trans. Automation Science and Engineering, IEEE Trans. Systems, Man, and Cybernetics, Part A: Systems and Human Beings, IEEE Trans. Systems, Man, and Cybernetics: Systems, IEEE Access (also a Senior Editor), and Information Sciences (Elsevier). He is a recipient of Alexander von Humboldt Research Grant and Research in Paris.