Amira Chouchane; Philippe Declerck
Diagnosis on a sliding window for partially observable Petri nets

Kybernetika, Vol. 58 (2022), No. 4, 479–497

Persistent URL: http://dml.cz/dmlcz/151160

KYBERNETIKA — VOLUME *58* (2022), NUMBER 4, PAGES 479–497

DIAGNOSIS ON A SLIDING WINDOW FOR PARTIALLY OBSERVABLE PETRI NETS

Amira Chouchane and Philippe Declerck

In this paper, we propose an algebraic approach to investigate the diagnosis of partially observable labeled Petri nets based on state estimation on a sliding window of a predefined length h. Given an observation, the resulting diagnosis state can be computed while solving integer linear programming problems with a reduced subset of basis markings. The proposed approach consists in exploiting a subset of h observations at each estimation step, which provides a partial diagnosis relevant to the current observation window. This technique allows a status update with a "forgetfulness" of past observations and enables distinguishing repetitive and punctual faults. The complete diagnosis state can be defined as a function of the partial diagnosis states interpreted on the sliding window. As the analysis shows that some basis markings can present an inconsistency with a future evolution, which possibly implies unnecessary computations of basis markings, a withdrawal procedure of these irrelevant basis markings based on linear programming is proposed.

Keywords: diagnosis, state estimation, partially observed Petri net, sliding window, integer linear programming

Classification: 93Axx,49Kxx

1. INTRODUCTION

The problem of estimating the state of a dynamic system is a fundamental issue in system theory. Observing a process and estimating its state consists in determining the values of system variables from a certain set of measurements given by a set of sensors. In a Petri Net (PN) framework, the state observer aims to give an on-line picture of the process and to provide the estimation of the system state based on the observation of a set of events. An important motivation is the fault diagnosis as the occurrence of faults can perturb or stop the evolution and production of the process. As it can contain transitions associated with events which are not observable or not available for the supervisory system, this PN is said to be partially observable. Unobservable events can also model faults, disturbances or noises that can affect the system. Indeed, the complete computerization of each process is not always possible due to the physical impossibility and the excessive cost. For instance, in a motorway with a payable access, the inputs and outputs at an unauthorized access are unobservable events if no sensor has

DOI: 10.14736/kyb-2022-4-0479

been added to detect the relevant events. These unobservable events can also represent faults.

In this article, we are interested in the fault diagnosis of a partially observed Labeled PN (LPN) which describes the fault model. We assume that the faults of the system are modeled by some unobservable transitions. Therefore, the occurrence of a fault is equivalent to the firing of the associated unobservable transition.

The literature is very rich of bibliographic work on the LPN diagnosis [1, 2, 3, 4, 5, 6, 7, 8, 9]. In what follows, we will cite the most relevant work, in relation with the proposed approach.

Diagnosis techniques based on the basis marking estimation have been subject to a number of interesting papers in the last ten years [1, 2, 10, 11, 12]. The set of basis markings is a reduced subset of the possible current markings which are coherent with a given observed sequence. It is based on the computation of a set of minimal explanation vectors associated with each transition of the observed sequence. In [1], a basis-marking based approach was suggested for the diagnosis of partially observable LPNs with an acyclic unobservable subnet. The faults are modeled by some unobservable transitions. In [10], a generalization of this approach assuming that faults could also be modeled by observable undistinguishable transitions was made. The determination of the set of minimal explanation vectors was based on a tabular algorithm that shows an exponential complexity in the worst case. To tackle this issue, it is shown in [10] that if the LPN is bounded, the off-line construction of an automaton called Basis Reachability Graph (BRG) will move off-line a burdensome part of the procedure. This BRG is a deterministic graph used for the on-line diagnosis that has as many nodes as possible basis markings. However, the number of basis markings can be huge even for a limited number of transitions, which will considerably affect the execution time of the approach [13, 14]. Assuming that all the system faults are observable, an Extended BRG (EBRG) was constructed in [4]. The EBRG corresponds to an automaton based on basis markings and has significantly fewer states than the reachability graph in most cases, but the number of nodes is still exponential.

To reduce the state space to be estimated, diagnostic techniques based on Integer Linear Programming (ILP) problems have been developed. In [6], the authors put forward an approach of fault diagnosis of a partially observed PN requiring an on-line computation of the set of possible fault events explaining the last observed event. The on-line computation consists in solving the ILP problems formulated on a net structure and based on g-markings. In [15], the authors developed a methodology for partially observed LPN diagnosis. This methodology calculates the least-cost transition firing sequences that are consistent with the observed word. It is assumed that each transition is associated with a nonnegative cost that will capture its likelihood. A recursive algorithm was developed, which may find the least-cost firing sequence while reconstructing only a finite number of transition firing sequences under the acyclicity assumption of the unobservable subnet. A fault is detected if it is included in a subset of unobservable sequences that are coherent with the observation verifying an ILP problem. In [7], a diagnosis approach of partially observed LPN, with undistinguishable observable transitions and with loops in the unobservable subnet, was presented. The faults of the system are modeled by unobservable transitions. A new system description on a reduced horizon

was described, which avoids the initial marking update for each observed transition of the observed word. A fault diagnosis approach was established based on the resolution of ILP problems. However, the resolution time of an ILP problem is exponential in the worst case and the diagnosis procedure can be time consuming. To further reduce the computation complexity, a relaxation of ILP problems allows solving these problems by an algorithm of linear programming, such as the ellipsoid algorithm of Khashiyan or the interior point algorithm of Karmarkar, which are polynomial in the worst case. The drawback of this approach is that the receding horizon can not be increased infinitely as it raises with the cardinality of the observed word. A solution to this problem is proposed in this paper.

In the present work, we introduce diagnosis indicators for fault diagnosis of partially observable LPNs under the form of ILP problems defined on a sliding horizon (window) instead of a receding one. The interest of the sliding window is to provide an updated interpretation focusing on the current status corresponding to the chosen horizon of observation and allowing the "forgetfulness" of the old events. This forgetfulness is necessary when the process evolves on a long operation time. We express any fault indicator under ILP problems for a given horizon length and a set of basis markings. Herein, the main lines of the classical technique [1, 10], which computes the sets of minimal explanation vectors and basis markings, are kept. We will see that at each step, the computation of the set of basis markings is necessary only at the beginning of the current window. Moreover, a diagnosis can be made even if the determination of the last sets of minimal explanation vectors and basis markings is not finished or cannot be finished due to a combinatorial explosion. This technique allows postponing the time-consuming computation of these sets for the current observation. In addition, we show that a minimal solution can lead to a basis marking, which cannot produce a sequence consistent with the following observations. In that situation, the generation of some minimal solutions and relevant basis markings is unsuccessful. The originality of this work is to present such a particular situation. Checking these unfruitful basis markings on a sliding horizon allows protectively preventing additional calculations that can be very costly from the point of view of time and space.

In this paper, the incidence matrices and the initial marking of the LPN are assumed to be known, along with the feasibility of the system. The occurrences of observable events are considered non-simultaneous. We take the hypothesis of acyclicity, ensuring that the solutions found by the algebraic method correspond to the firing sequences of unobservable transitions [16]. Moreover, the fault transitions are divided into various fault classes. For clarity, the case of undistinguishable observable transitions (the firings of several transitions can be observed but cannot be distinguished) is not considered in this paper, but relevant modeling can be found in the study of [7].

The remainder of this paper is organized as follows. In section 2, we present some preliminary notions and some useful definitions related to the proposed approach. In section 3, we expose the principle of our estimation approach for fault diagnosis. We build a polyhedron defined on a sliding window, which describes the estimation problem under an algebraic point of view. In section 4, we introduce different criteria which permit implementing the fault diagnosis procedure. In section 5, the computational complexity analysis and the comparative results are presented. In section 6, a bench-

mark is used to illustrate the numerical efficiency of the proposed technique. Section 7 concludes the paper and provides a number of perspectives for the present work.

2. PRELIMINARY

2.1. Background on LPN and notations

A Place/Transition (P/T) net is the structure $\mathcal{N} = (P, TR, W^+, W^-)$, where P is a set of $|P|$ places and T is a set of $|T|$ transitions. Matrices W^+ and W^- are respectively the $|P| \times |T|$ post and pre-incidence matrices over $\mathbb{N}$, where each row $l \in \{1, \ldots, |P|\}$ specifies the weight of the incoming and outgoing arcs of place $p_l \in P$. The incidence matrix is $W = W^+ - W^-$. The pre-set and post-set of node $z \in P \bigcup T$ are denoted by ${}^\bullet z$ and $z^\bullet$, respectively.

Notation T^* represents the set of firing sequences, denoted x, consisting of transitions of T. We denote by $\pi_T : T^* \to \mathbb{N}^{|T|}$ the function that associates with x the $\overline{x}$ vector of dimension $|T|$ expressing the firing vector or count vector of sequence $x \in T^*$, where the ith component $\overline{x}_i$ is the firing number of the ith transition of T in sequence x. The marking of the set of places P is a vector $M \in \mathbb{N}^{|P|}$ that assigns to each place $p_i \in P$ a non-negative integer number of tokens M_i. The ith component M_i is denoted as $M(p_i)$. Marking M, reached from the initial marking M^{init} (which replaces the usual notation M_0) by firing sequence x, can be calculated by the fundamental relation: $M = M^{init} + W.\overline{x}$. Transition $x_i \in T$ is enabled at M if $M \geq W^-(., x_i)$ and may be fired yielding marking $M' = M + W(., x_i)$. We write $M[x \succ$ to denote that the sequence of transitions x is enabled at M, and we write $M[x \succ M'$ to denote that the firing of x yields M'.

A labeling function $\mathcal{L} : T \to E \cup \{\varepsilon\}$ assigns to each transition $x_i \in T$ either a symbol from a given alphabet E or the empty string ε. Without any loss of generality, mapping $\mathcal{L}$ is assumed to be surjective. In a partially observed LPN, we assume that the set of transitions T can be partitioned as $T = T_{ob} \bigcup T_{un}$, where set T_{ob} (resp. T_{un}) is the set of observable transitions (resp. unobservable transitions) associated with a label of E (resp. the empty string ε). Thus, $|T_{un}| \neq \emptyset$ as $\mathcal{L}$ is surjective. The restriction of $\mathcal{L}$ to T_{ob} and E is a function $\mathcal{L}' : T_{ob} \to E$ which is assumed to be injective (and so bijective): In other words, the case of undistinguishable observable transitions is not considered in this paper. The labeling functions $\mathcal{L}$ and $\mathcal{L}'$ can be also extended respectively to transition sequences, $\mathcal{L} : T^* \to \{E \bigcup \{\varepsilon\}\}^*$ and $\mathcal{L}' : T_{ob}^* \to E^*$.

The unobservable induced subnet of the PN $\mathcal{N}$ is defined as the new net $\mathcal{N}_{un} = (P, T_{un}, W_{un}^+, W_{un}^-)$, where W_{un}^+ and W_{un}^- are the restrictions of W^+ and W^- to $P \times T_{un}$. Therefore, $W_{un} = W_{un}^+ - W_{un}^-$. Likewise, the observed subnet of $\mathcal{N}$ is defined as the new net $\mathcal{N}_{ob} = (P, T_{ob}, W_{ob}^+, W_{ob}^-)$ where W_{ob}^+ and W_{ob}^- are the restrictions of W^+ and W^- to $P \times T_{ob}$. A reorganization of the columns as regards T_{un} and T_{ob} yields $W = \begin{pmatrix} W_{un} & W_{ob} \end{pmatrix}$. A sequence of unobservable transitions denoted $x_{un} \in T_{un}^*$ is of count vector $\overline{x_{un}} = \pi_{T_{un}}(x_{un})$ of dimension $|T_{un}|$ and a sequence of observable transitions denoted $x_{ob} \in T_{ob}^*$ is of count vector $\overline{x_{ob}} = \pi_{T_{ob}}(x_{ob})$ of dimension $|T_{ob}|$. The reorganization of the components of $\overline{x}$ yields $\overline{x} = \begin{pmatrix} \overline{x_{un}}^T & \overline{x_{ob}}^T \end{pmatrix}^T$.

Let us add the following notations. We consider a feasible firing sequence x from the initial marking M^{init}. Let $x_{ob} = \mathcal{L}'(x)$ be the observable projection of x. Assuming

that the occurrences of observable transitions are non-simultaneous, we associate an index $\langle i \rangle$ to each occurrence of an observable transition. We can then write x_{ob} as $x_{ob} = x_{ob}^{\langle 1 \rangle} x_{ob}^{\langle 2 \rangle} \cdots x_{ob}^{\langle k \rangle}$ where $x_{ob}^{\langle i \rangle} \in T_{ob}$ is the i^{th} observed transition of x_{ob} and $k = |x_{ob}|$. Let $x_{un}^{\langle 1 \rangle}, x_{un}^{\langle 2 \rangle}, \ldots, x_{un}^{\langle k \rangle}$ be the unobservable sequences that are coherent with transitions $x_{ob}^{\langle 1 \rangle}, x_{ob}^{\langle 2 \rangle}, \ldots, x_{ob}^{\langle k \rangle}$, respectively. That is, $x_{un}^{\langle 1 \rangle} x_{ob}^{\langle 1 \rangle} x_{un}^{\langle 2 \rangle} x_{ob}^{\langle 2 \rangle} \cdots x_{un}^{\langle k \rangle} x_{ob}^{\langle k \rangle}$ is a feasible firing sequence from M^{init}. Then, there exists a set of markings $M^{\langle 1 \rangle}, M^{\langle 2 \rangle}, \ldots, M^{\langle k+1 \rangle}$ such that $M^{\langle 1 \rangle}[x_{un}^{\langle 1 \rangle} x_{ob}^{\langle 1 \rangle} \succ M^{\langle 2 \rangle} \cdots M^{\langle k \rangle}[x_{un}^{\langle k \rangle} x_{ob}^{\langle k \rangle} \succ M^{\langle k+1 \rangle}$ with $M^{\langle 1 \rangle} = M^{init}$.

2.2. Definitions

The concept of explanation vectors [17] allows for focusing on specific evolutions of the PN which are consistent with the observations. The set of explanation vectors is defined for each observed transition $x_{ob}^{\langle i \rangle}$ as follows:

Definition 2.1. Let $x_{un}^{\langle i \rangle}$ be an unobservable sequence which leads to the firing of the observed transition $x_{ob}^{\langle i \rangle}$ from a starting marking $M^{\langle i \rangle}$. The relevant count vector is denoted $\overline{x_{un}}^{\langle i \rangle}$ and is named an explanation vector of $x_{ob}^{\langle i \rangle}$ from $M^{\langle i \rangle}$. The sets of possible unobservable sequences $x_{un}^{\langle i \rangle}$ and relevant explanation vectors $\overline{x_{un}}^{\langle i \rangle}$ for marking $M^{\langle i \rangle}$ and observation $x_{ob}^{\langle i \rangle}$ are denoted $SEQ(M^{\langle i \rangle}, x_{ob}^{\langle i \rangle})$ and $E(M^{\langle i \rangle}, x_{ob}^{\langle i \rangle})$, respectively, and defined as:

$$SEQ(M^{\langle i \rangle}, x_{ob}^{\langle i \rangle}) = \{x_{un}^{\langle i \rangle} \mid x_{un}^{\langle i \rangle} \in T_{un}^*,\ M^{\langle i \rangle}[x_{un}^{\langle i \rangle} \succ M' \text{ with } M'[x_{ob}^{\langle i \rangle} \succ\}$$
$$E(M^{\langle i \rangle}, x_{ob}^{\langle i \rangle}) = \{\overline{x_{un}}^{\langle i \rangle} \mid x_{un}^{\langle i \rangle} \in SEQ(M^{\langle i \rangle}, x_{ob}^{\langle i \rangle}\}.$$

The set of computed markings, which is produced following the firing of $x_{ob}^{\langle i \rangle}$ from $M^{\langle i \rangle}$, is used as known starting markings for $x_{ob}^{\langle i+1 \rangle}$, is denoted $\mathcal{M}^{\langle i+1 \rangle}$ and is determined by the consideration of all markings $M^{\langle i \rangle} \in \mathcal{M}^{\langle i \rangle}$ and the relevant computed sequences $x_{un}^{\langle i \rangle} x_{ob}^{\langle i \rangle}$. Formally, the set of starting markings is defined iteratively as follows:

$$\begin{aligned} \mathcal{M}^{\langle 1 \rangle} = &\quad \{\ M^{\langle 1 \rangle} \text{ with } M^{\langle 1 \rangle} = M^{init}\}, \\ \mathcal{M}^{\langle i+1 \rangle} = &\quad \{M^{\langle i+1 \rangle} \in \mathbb{N}^{|P|} \mid (\exists M^{\langle i \rangle} \in \mathcal{M}^{\langle i \rangle})(\exists x_{un}^{\langle i \rangle} \in SEQ(M^{\langle i \rangle}, x_{ob}^{\langle i \rangle})) : \\ &\quad M^{\langle i \rangle}[x_{un}^{\langle i \rangle} \succ M'[x_{ob}^{\langle i \rangle} \succ M^{\langle i+1 \rangle}\},\ \forall i \in \{1, \ldots, k\}. \end{aligned}$$

From $M^{\langle i \rangle}[x_{un}^{\langle i \rangle} \succ M'$ with $M'[x_{ob}^{\langle i \rangle} \succ$, we can easily deduce the following system with unknown $\overline{x_{un}}^{\langle i \rangle} \in \mathbb{N}^n$:

$$-W_{un}.\overline{x_{un}}^{\langle i \rangle} \leq M^{\langle i \rangle} - W_{ob}^{-}(., x_{ob}^{\langle i \rangle}). \tag{1}$$

Notation $W_{ob}^{-}(., x_{ob}^{\langle i \rangle})$ selects the column relevant to the observed transition $x_{ob}^{\langle i \rangle}$ in W_{ob}^{-}. Each vector $\overline{x_{un}}^{\langle i \rangle}$ satisfying (1) is an explanation vector if the unobservable subnet is acyclic [18].

Among all sequences in $SEQ(M^{\langle i \rangle}, x_{ob}^{\langle i \rangle})$, we distinguish the ones which present a behavior strictly necessary to this firing, i. e. sequences which correspond to minimal

count vectors. For an observed transition $x_{ob}^{\langle i\rangle}$ from $M^{\langle i\rangle}$, we get:

$$\begin{aligned} SEQ^{\min}(M^{\langle i\rangle}, x_{ob}^{\langle i\rangle}) &= \{x_{un}^{\langle i\rangle} \in SEQ(M^{\langle i\rangle}, x_{ob}^{\langle i\rangle}) \mid \nexists x'_{un}{}^{\langle i\rangle} \in SEQ(M^{\langle i\rangle}, x_{ob}^{\langle i\rangle}) : \\ &\quad \overline{x'_{un}}^{\langle i\rangle} < \overline{x_{un}}^{\langle i\rangle} \} \\ E^{\min}(M^{\langle i\rangle}, x_{ob}^{\langle i\rangle}) &= \{\overline{x_{un}}^{\langle i\rangle} \mid x_{un}^{\langle i\rangle} \in SEQ^{\min}(M^{\langle i\rangle}, x_{ob}^{\langle i\rangle}\}. \end{aligned}$$

Similarly, the set of basis markings [17] is a subset of $\mathcal{M}^{\langle i+1\rangle}$ defined as follows:

$$\begin{aligned} \mathcal{M}^{\min,\langle 1\rangle} = &\quad \{\ M^{\langle 1\rangle} \text{ with } M^{\langle 1\rangle} = M^{init}\}, \\ \mathcal{M}^{\min,\langle i+1\rangle} = &\quad \{M^{\langle i+1\rangle} \in \mathbb{N}^{|P|} \mid (\exists M^{\langle i\rangle} \in \mathcal{M}^{\min,\langle i\rangle}), (\exists x_{un}^{\langle i\rangle} \in SEQ^{\min}(M^{\langle i\rangle}, x_{ob}^{\langle i\rangle})): \\ &\quad M^{\langle i\rangle}[x_{un}^{\langle i\rangle} \succ M'[x_{ob}^{\langle i\rangle} \succ M^{\langle i+1\rangle}\},\ \forall i \in \{1, \ldots, k\}. \end{aligned}$$

The relevant minimal count vectors $\overline{x_{un}}^{\langle i\rangle}$ of $E^{\min}(M^{\langle i\rangle}, x_{ob}^{\langle i\rangle})$ with $M^{\langle i\rangle} \in \mathcal{M}^{\min,\langle i\rangle}$ are named *minimal* explanation vectors. They can be computed by the tabular approach (Algorithm 3.5 in [17]) which is a combinatorial technique treating (1) with $M^{\langle i\rangle} \in \mathcal{M}^{\min,\langle i\rangle}$.

Definition 2.2. (Cabasino et al. [17]) The set of j-vectors for an observed word w from M^{init} is the set of minimal count vectors $\overline{x_{un}}$ of unobservable transitions $x_{un} = x_{un}^{\langle 1\rangle} x_{un}^{\langle 2\rangle} \cdots x_{un}^{\langle k\rangle}$ interleaved with $x_{ob} = x_{ob}^{\langle 1\rangle} x_{ob}^{\langle 2\rangle} \cdots x_{ob}^{\langle k\rangle}$, where $L'(x_{ob}) = w$, whose firings enable x_{ob}.

If we adopt the definition of j-vectors presented in [17] for an LPN without undistinguishable observable transitions and with an acyclic unobservable subnet (the hypotheses considered in this work), the j-vectors of an observed sequence $x_{ob} = x_{ob}^{\langle 1\rangle} x_{ob}^{\langle 2\rangle} \cdots x_{ob}^{\langle k\rangle}$ from M^{init} can be formally defined as follows:

$$\begin{aligned} Y_{min}(M^{init}, x_{ob}) = &\quad \{\overline{x_{un}} = \sum_{i=1}^{k} \overline{x_{un}}^{\langle i\rangle} \mid \overline{x_{un}}^{\langle i\rangle} \in E^{\min}(M^{\langle i\rangle}, x_{ob}^{\langle i\rangle}) \text{ with} \\ &\quad M^{\langle 1\rangle}[x_{un}^{\langle 1\rangle} x_{ob}^{\langle 1\rangle} \succ M^{\langle 2\rangle} \cdots M^{\langle k\rangle}[x_{un}^{\langle k\rangle} x_{ob}^{\langle k\rangle} \succ M^{\langle k+1\rangle} \text{ and} \\ &\quad M^{\langle 1\rangle} = M^{init}\}. \end{aligned}$$

In this paper, we use the notion of the observation window, denoted $[x_{ob}^{\langle i\rangle}]_q^r$, which means that we observe the successive firings of transitions from $x_{ob}^{<q>}$ to $x_{ob}^{<r>}$ starting from $M^{<q>}$ (with $q \leq r$). The length of the observation window $[x_{ob}^{\langle i\rangle}]_q^r$ is defined as the number of the observed transitions in $[x_{ob}^{\langle i\rangle}]_q^r$, i. e. $r - q + 1$. In the following section, we present the estimation problem on a sliding window considered in this article.

3. ESTIMATION ON A SLIDING WINDOW COMPLETED WITH BASIS MARKINGS

Let us consider an LPN with a known structure and a known initial marking M^{init}. Given a sequence of labels $w \in E^*$ emitted by the firing of an observable sequence $x_{ob} = x_{ob}^{\langle 1\rangle} x_{ob}^{\langle 2\rangle} \cdots x_{ob}^{\langle |w|\rangle}$ generated by the LPN activity, $x_{un}^{\langle 1\rangle}, x_{un}^{\langle 2\rangle}, \ldots, x_{un}^{\langle |w|\rangle}$ are unobservable sequences coherent with transitions $x_{ob}^{\langle 1\rangle}, x_{ob}^{\langle 2\rangle}, \ldots, x_{ob}^{\langle |w|\rangle}$, respectively. A feasible firing sequence x from the initial marking M^{init} can be then written as $x =$

$x_{un}^{\langle 1\rangle} x_{ob}^{\langle 1\rangle} x_{un}^{\langle 2\rangle} x_{ob}^{\langle 2\rangle} \cdots x_{un}^{\langle |w|\rangle} x_{ob}^{\langle |w|\rangle}$. Then, we will algebraically describe the unobservable count vectors that are coherent with the observed labels, on a sliding observation window of fixed length h where $1 \leq h \leq |w|$, i. e. by considering for each estimation step h successive observed transitions while updating the starting marking successively.

3.1. Principle

For each estimation step k with $k \in [h \cdots |w|]$, we consider a window of observations $[x_{ob}^{\langle i\rangle}]_{k-h+1}^{k}$ where $h \in [1 \cdots |w|]$ is the fixed horizon length. At step k, observation $x_{ob}^{\langle k\rangle}$ is the last considered one, while observation $x_{ob}^{\langle k+1\rangle}$ is not available (not still available if we consider a usual behavior). To determine the diagnosis status at step k, the interpretation will not consider the sequences that interleave observations on horizon $\{1, \ldots, k-h\}$ starting from the initial marking $M^{\langle 1\rangle} = M^{init}$ and leading to a marking of $\mathcal{M}^{\min,\langle k-h+1\rangle}$. The starting markings at step k are then the basis markings $M^{\langle k-h+1\rangle} \in \mathcal{M}^{\min,\langle k-h+1\rangle}$. As explained in the introduction, the proposed approach keeps the standard approach [17] but adds a complementary technique based on a sliding window. The main lines of the procedure, for each step k with $k \in [h \cdots |w|]$, are presented as follows:

- The computation of the set of basis markings $\mathcal{M}^{\min,\langle k-h+1\rangle}$
- The checking of $\mathcal{M}^{\min,\langle k-h+1\rangle}$, based on the analysis of the coherence of markings in $\mathcal{M}^{\min,\langle k-h+1\rangle}$ w.r.t. the window of future observations $[x_{ob}^{\langle i\rangle}]_{k-h+1}^{k}$
- The diagnosis based on the window of observations $[x_{ob}^{\langle i\rangle}]_{k-h+1}^{k}$ with starting markings in $\mathcal{M}^{\min,\langle k-h+1\rangle}$.

The following notations relevant to step k and window length h are taken:

$$\overline{x_{un}} = \Big(\ (\overline{x_{un}}^{\langle k-h+1\rangle})^T \quad (\overline{x_{un}}^{\langle k-h+2\rangle>})^T \quad (\overline{x_{un}}^{\langle k-h+3\rangle>})^T \quad \ldots \quad \ldots \quad (\overline{x_{un}}^{\langle k\rangle})^T \ \Big)^T$$
$$\overline{x_{ob}} = \Big(\ (\overline{x_{ob}}^{\langle k-h+1\rangle})^T \quad (\overline{x_{ob}}^{\langle k-h+2\rangle>})^T \quad (\overline{x_{ob}}^{\langle k-h+3\rangle>})^T \quad \ldots \quad \ldots \quad (\overline{x_{ob}}^{\langle k\rangle})^T \ \Big)^T.$$

As $x_{ob}^{\langle i\rangle}$ is a unique observed transition, with $i \in \{k-h+1, \ldots, k\}$, each vector $\overline{x_{ob}}^{\langle i\rangle}$ is null except a unique component that is equal to 1 and which corresponds to the observed transition. Vector $\overline{x_{un}}$ includes the different explanation vectors $\overline{x_{un}}^{\langle i\rangle}$, $i \in \{k-h+1, \ldots, k\}$.

Remark 3.1. As the technique of state estimation on a sliding window considers a fixed horizon, it needs that a sufficient number of observation is available, that is, h observations and the first horizon considered by the estimation is $\{1, \ldots, h\}$, the second one is $\{2, \ldots, h+1\}$, and so on. This point implies that the approach starts at $k = h$. However, as the technique is flexible, it can easily be completed by a technique of state estimation on a receding horizon, proposed in [7], before its application. The form of the used system of relations is similar but must be adapted to the receding horizon h' increasing from 1 to $h-1$.

3.2. Algebraic modeling on a sliding window

In this part, we develop an algebraic model which describes the evolution of count vectors for the observation window $[x_{ob}^{\langle i\rangle}]_{k-h+1}^{k}$ beginning from a starting marking $M^{\langle k-h+1\rangle} \in \mathcal{M}^{\min,\langle k-h+1\rangle}$. We have $M^{\langle i\rangle}[x_{un}^{\langle i\rangle} \succ M'^{\langle i\rangle}$ with $M'^{\langle i\rangle}[x_{ob}^{\langle i\rangle} \succ M^{\langle i+1\rangle}$ for $i \in \{k-h+1,\ldots,k\}$.

3.2.1. Equations for first observed transition $x_{ob}^{\langle k-h+1\rangle}$ of the window

For the observed transition $x_{ob}^{\langle k-h+1\rangle}$ from $M^{\langle k-h+1\rangle}$, there exists marking $M'^{\langle k-h+1\rangle}$ with

$$M'^{\langle k-h+1\rangle} = M^{\langle k-h+1\rangle} + W_{un} \cdot \overline{x_{un}}^{\langle k-h+1\rangle}. \tag{2}$$

Marking $M'^{\langle k-h+1\rangle}$ satisfies $M'^{\langle k-h+1\rangle}[(x_{ob})^{\langle k-h+1\rangle} \succ$. Therefore, we get:

$$-W_{un} \cdot \overline{x_{un}}^{\langle k-h+1\rangle} + W_{ob}^{-}.\overline{x_{ob}}^{\langle k-h+1\rangle} \leq M^{\langle k-h+1\rangle}. \tag{3}$$

3.2.2. Equations for observations from $x_{ob}^{\langle k-h+2\rangle>}$ to $x_{ob}^{\langle k\rangle}$ of the window

For the observed transition $x_{ob}^{\langle k-h+2\rangle>}$ from $M^{\langle k-h+2\rangle>}$, there exists some marking $M'^{\langle k-h+2\rangle>}$ with

$$\begin{aligned} M'^{\langle k-h+2\rangle>} = \quad & M^{\langle k-h+2\rangle>} + W_{un} \cdot \overline{x_{un}}^{\langle k-h+2\rangle>} = M^{\langle k-h+1\rangle} + W_{un}.(\overline{x_{un}}^{\langle k-h+1\rangle} + \\ & \overline{x_{un}}^{\langle k-h+2\rangle>}) + W_{ob}.(\overline{x_{ob}}^{\langle k-h+1\rangle}). \end{aligned}$$

Marking $M'^{\langle k-h+2\rangle>}$ satisfies $M'^{\langle k-h+2\rangle>}[(x_{ob})^{\langle k-h+2\rangle>} \succ$. Hence, we have:

$$-W_{un}.(\overline{x_{un}}^{\langle k-h+1\rangle} + \overline{x_{un}}^{\langle k-h+2\rangle>}) - W_{ob}.(\overline{x_{ob}}^{\langle k-h+1\rangle}) + W_{ob}^{-}.\overline{x_{ob}}^{\langle k-h+2\rangle>} \leq M^{\langle k-h+1\rangle}.$$

More generally, we deduce for all $i \in \{k-h+2,\ldots,k\}$ the following relation:

$$-W_{un} \sum_{j=k-h+1}^{i} \overline{x_{un}}^{<j>} - W_{ob}. \sum_{j=k-h+1}^{i-1} \overline{x_{ob}}^{<j>} + W_{ob}^{-}.\overline{x_{ob}}^{\langle i\rangle} \leq M^{\langle k-h+1\rangle}. \tag{4}$$

3.2.3. Complete system

According to (3) and (4), the count vectors $\overline{x_{un}} \in \mathbb{N}^{h.|T_{un}|}$ and $\overline{x_{ob}} \in \mathbb{N}^{h.|T_{ob}|}$ fulfill the following polyhedron:

$$A.\overline{x_{un}} + B.\overline{x_{ob}} \leq b \tag{5}$$

where $A = \begin{pmatrix} -W_{un} & 0 & 0 & \ldots & 0 \\ -W_{un} & -W_{un} & 0 & \ldots & 0 \\ -W_{un} & -W_{un} & -W_{un} & & \vdots \\ \vdots & \vdots & & & 0 \\ -W_{un} & -W_{un} & \ldots & -W_{un} & -W_{un} \end{pmatrix}$,

$$B = \begin{pmatrix} W_{ob}^{-} & 0 & 0 & \dots & 0 \\ -W_{ob} & W_{ob}^{-} & 0 & \dots & 0 \\ -W_{ob} & -W_{ob} & W_{ob}^{-} & & \vdots \\ \vdots & \vdots & & & 0 \\ -W_{ob} & -W_{ob} & \dots & -W_{ob} & W_{ob}^{-} \end{pmatrix} \text{ and } b = \begin{pmatrix} M^{\langle k-h+1 \rangle} \\ M^{\langle k-h+1 \rangle} \\ M^{\langle k-h+1 \rangle} \\ \vdots \\ M^{\langle k-h+1 \rangle} \end{pmatrix}$$

with $M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min,\langle k-h+1 \rangle}$. As a function of h, the dimensions of matrices A and B and vector b are $(h.|P| \times h.|T_{un}|)$, $(h.|P| \times h.|T_{ob}|)$ and $h.|P| \times 1$, respectively. Matrices A and B depend on the structure of the PN, while vector b depends on the basis marking $M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min,\langle k-h+1 \rangle}$.

Remark 3.2. System (5) can easily be completed at step 1 to treat the case where the initial marking $M^{\langle 1 \rangle}$ is not completely known but belongs to a set of possible markings described by a system of form $\Gamma.M^{\langle 1 \rangle} \leq \Theta$.

3.3. Checking of starting basis markings

A count vector $x_{un}^{\langle k \rangle}$ that is consistent with a current observation $x_{ob}^{\langle k \rangle}$ must also be consistent with the following observations $x_{ob}^{\langle i \rangle}$ with $i \in \{k+1, \dots, |w|\}$. Formally, we have $M^{\langle i \rangle}[x_{un}^{\langle i \rangle} \succ M'^{\langle i \rangle}$ with $M'^{\langle i \rangle}[x_{ob}^{\langle i \rangle} \succ M^{\langle i+1 \rangle}$ for $i = k$, but we must also consider the same constraint for any $i \geq k+1$. In other words, the determination of $x_{un}^{\langle k \rangle}$ depends not only on observation $x_{ob}^{\langle k \rangle}$ but can also be refined according to the subsequent observations $x_{ob}^{\langle i \rangle}$, where $i \geq k+1$. In particular, it is implied that an unobservable sequence $x_{un}^{\langle i \rangle} \in SEQ(M^{\langle i \rangle}, x_{ob}^{\langle i \rangle})$ relevant to an explanation vector $\overline{x_{un}}^{\langle i \rangle} \in E(M^{\langle i \rangle}, x_{ob}^{\langle i \rangle})$ can lead to a basis marking which cannot satisfy $M^{<j>}[x_{un}^{<j>} \succ M'^{<j>}$ with $M'^{<j>}[x_{ob}^{<j>} \succ$ for a given $j \geq i+1$ and, therefore, cannot be the starting marking of a complete sequence containing the following observations. Accordingly, system (5) at a given step can be inconsistent for some basis markings but is consistent for at least one basis marking. Thus, the computation of some basis markings can be unfruitful for a sequence of observations. Practically, the standard approach [17] deduces the basis markings at a given step k from the successive observations $x_{ob}^{\langle i \rangle}$ for $i \in \{1, \dots, k\}$. However, some basis markings can become unreachable with the occurrences of new observations $x_{ob}^{\langle i \rangle}$, where $i \in \{k+1, \dots, |w|\}$, which are still unknown at step k. In other words, set $\mathcal{M}^{\min,\langle k+1 \rangle}$ computed at step k may contain some candidate basis markings that are possibly inconsistent w.r.t. the following observations at step $\langle i \rangle$, where $i \geq k+1$. Therefore, these irrelevant basis markings must be withdrawn from the procedure when they are detected.

For the sake of illustration, let us consider the acyclic PN in Figure 1 where $T_{ob} = \{y_1, y_2\}$ and $T_{un} = \{x_1, x_2, x_3\}$. assume the following evolution $x_1 y_1 x_3 y_2$ starting from $M^{init} = M^{\langle 1 \rangle} = (1\ 1\ 0\ 0\ 0\ 0)^T$. Two successive observations y_1 and y_2 are generated. For observation y_1, the two minimal explanation vectors are:

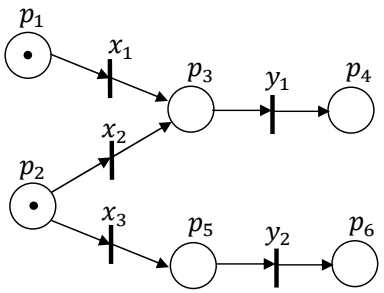

Fig. 1. Petri net for unfruitful basis marking analysis.

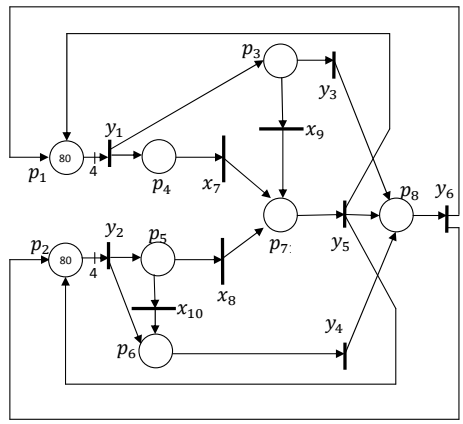

Fig. 2. Petri net of numerical analysis (Figure 4 in [19]).

- $\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T$ is relevant to x_1 which yields the basis marking $M^1 = (0\,1\,0\,1\,0\,0)^T \in \mathcal{M}^{\min,\langle 2\rangle}$.
- $\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T$ is relevant to x_2 which yields $M^2 = (1\,0\,0\,1\,0\,0)^T \in \mathcal{M}^{\min,\langle 2\rangle}$.

For observation y_2 from M^1, a minimal explanation vector is $\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T$ relevant to x_3 (consistent with sequence $x_1y_1x_3y_2$). For observation y_2 from M^2, no minimal explanation vector is possible as $M^2(p_2) = 0$. Therefore, M^2 is inconsistent w.r.t. observation y_1y_2.

Example 1 shows that the standard procedure produces a sequence leading to a basis marking M^2, which cannot be extended to a sequence containing a new observed transition. In other words, the observable transition y_2 is infeasible for the basis marking M^2; i. e., y_2 can never be fired in any firing sequence starting from this basis marking. However, the observable transition y_2 is L1-live for the other basis marking M^1; i. e., y_2 can be fired at least once in some firing sequences. In fact, the two observable transitions present a structural conflict described by a common place p_2 presenting two output unobservable transitions. This structural conflict leads to the loss of the liveness of the observable transition for the basis marking M^2, which contradicts the observation. For M^2, observation y_2 must use the same tokens already exploited, which yields an inconsistent system. To sum up, two minimal sequences are potentially necessary for observation y_1 but it can only be exploited in the following observation.

In what follows, we consider the case of a Backward Conflict-Free (BCF) unobservable subnet where the difficulty disappears.

Property 1. System (5) for a live BCF unobservable subnet is consistent for all basis markings.

P r o o f . Each system for a BCF unobservable subnet presents a unique minimal explanation vector, which leads to a basis marking. As the PN is assumed to be live, system (5) can follow a trajectory and each unique system for each observation is necessarily consistent. □

The checking of each marking $M^{\langle k-h+1\rangle} \in \mathcal{M}^{\min,\langle k-h+1\rangle}$ can now be made. The following result can then be stated:

Proposition 3.3. The basis marking $M^{\langle k-h+1\rangle} \in \mathcal{M}^{\min,\langle k-h+1\rangle}$ is unfruitful and must be withdrawn from $\mathcal{M}^{\min,\langle k-h+1\rangle}$ if the relevant polyhedron (5) has no solution.

P r o o f . We suppose that polyhedron (5) has no solution for a given starting marking $M^{\langle k-h+1\rangle} \in \mathcal{M}^{\min,\langle k-h+1\rangle}$ and a given observed sequence $x_{ob}^{\langle k-h+1\rangle}, \ldots, x_{ob}^{\langle k\rangle}$, we can then distinguish two cases. Case1: $\nexists x_{un}^{\langle k-h+1\rangle}, \ldots, x_{un}^{\langle k\rangle}$ are coherent respectively with $x_{ob}^{\langle k-h+1\rangle}, \ldots, x_{ob}^{\langle k\rangle}$ from $M^{\langle k-h+1\rangle}$. Case2: $M^{\langle k-h+1\rangle}$ can not be a starting marking for $x_{un}^{\langle k-h+1\rangle} x_{ob}^{\langle k-h+1\rangle} \ldots x_{un}^{\langle k\rangle} x_{ob}^{\langle k\rangle}$. Case1 can not be true as $x_{ob}^{\langle k-h+1\rangle} \cdots x_{ob}^{\langle k\rangle}$ is a feasible firing sequence, and then there exists at least one unobservable sequence $x_{un}^{\langle k-h+1\rangle} \cdots x_{un}^{\langle k\rangle}$ that is coherent with such an observed sequence. Consequently, case2 is true; i. e., $M^{\langle k-h+1\rangle}$ is an unfruitful basis marking. □

4. DIAGNOSIS ON A SLIDING WINDOW

We assume that each faulty behavior which may occur in the system is expressed by the firing of an unobservable transition. Naturally, there may also be other transitions that are unobservable whose firings correspond to normal behaviors. As a consequence, the set of unobservable transitions is partitioned in two subsets

$$T_{un} = T_f \cup T_n \tag{6}$$

where T_f is the set of unobservable transitions whose firings describe fault events, and T_n is the set of unobservable transitions whose firings correspond to regular unobservable events. Furthermore, the set of fault transitions T_f can also be partitioned into cl disjoint subsets that represent the different fault classes as follows:

$$T_f = T_f^1 \cup T_f^2 \cup \ldots \cup T_f^{cl}. \tag{7}$$

The objective is to detect a given fault class T_f^j using the technique of estimation on a sliding window.

Remark 4.1. In practice, faults can be modeled as a mixture of observable and unobservable transitions. In this case, if the observed word w includes an observable fault event, we can conclude immediately that the system is faulty. Otherwise, a state estimation procedure is necessary to generate the diagnosis verdict, and the proposed approach developed herein is then applicable.

First, let us consider the following definition of partial diagnosis states on a sliding window.

Definition 4.2. For an observed word w from M^{init}, we define the diagnosis function $\Delta_h^k(w, T_f^j)$ with respects to a given fault class T_f^j on a sliding observation window $[x_{ob}^{\langle i \rangle}]_{k-h+1}^k$ for all $k \in [h \cdots |w|]$ as follows:

$$\begin{cases} \Delta_h^k(w, T_f^j) = N_h : & \text{Fault } T_f^j \text{ does not occur on window } [x_{ob}^{\langle i \rangle}]_{k-h+1}^k \text{ from all} \\ & M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min, \langle k-h+1 \rangle}. \\ \Delta_h^k(w, T_f^j) = U_h : & \text{We can not conclude on the existence of } T_f^j \text{ on } [x_{ob}^{\langle i \rangle}]_{k-h+1}^k \\ & \text{from all } M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min, \langle k-h+1 \rangle}. \\ \Delta_h^k(w, T_f^j) = F_h : & \text{Fault } T_f^j \text{ occurs on window } [x_{ob}^{\langle i \rangle}]_{k-h+1}^k \text{ from all} \\ & M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min, \langle k-h+1 \rangle}. \end{cases}$$

For the current window of observations $[x_{ob}^{\langle i \rangle}]_{k-h+1}^k$, we can also define a first fault indicator $J^-(T_f^j, k)$ for each fault class $T_f^j \in T_f$, which describes the minimum number of occurrences of fault class T_f^j in the considered observation window from all markings $M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min, \langle k-h+1 \rangle}$. This first indicator is defined as follows:

$$\begin{cases} J^-(T_f^j, k) = min(c_f^j . \overline{x_{un}}) \\ \text{such that } (5), M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min, \langle k-h+1 \rangle} \end{cases} \tag{8}$$

with $c_f^j = \begin{pmatrix} c_j & \cdots & c_j & c_j \end{pmatrix} \in \{0,1\}^{h \times |T_{un}|}$ where c_j is a row vector of dimension $|T_{un}|$, for which all the elements are null, except the elements that are associated with fault transitions in T_f^j, which are equal to 1.

Symmetrically, we can determine the maximum number of occurrences of faults in class T_f^j.

$$\begin{cases} J^+(T_f^j, k) = max(c_f^j . \overline{x_{un}}) \\ \text{such that } (5), M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min, \langle k-h+1 \rangle} \end{cases} \tag{9}$$

Based on the two fault indicators mentioned above, we get the following results for partial diagnosis on a sliding window:

Proposition 4.3. Let us consider an observed word w from M^{init}. The partial diagnosis states of fault class T_f^j, associated with an observation window $[x_{ob}^{\langle i \rangle}]_{k-h+1}^k$ of length h for each step k with $k \in [h \cdots |w|]$, is defined as follows:

- If $J^-(T_f^j, k) \geq 1$, then $\Delta_h^k(w, T_f^j) = F_h$.
- If $J^+(T_f^j, k) = 0$, then $\Delta_h^k(w, T_f^j) = N_h$.
- If $J^-(T_f^j, k) = 0$ and $J^+(T_f^j, k) \geq 1$, then $\Delta_h^k(w, T_f^j) = U_h$.

Indeed, if $J^-(T_f^j, k) \geq 1$, we can say that at least a fault relevant to T_f^j has been detected at step k. Moreover, situation $J^-(T_f^j, k) \geq 1$ guarantees a repeated characteristic of fault class T_f^j.

By considering only set $\mathcal{M}^{\min,\langle k-h+1\rangle}$, the proposed partial diagnosis at a given step k does not need the computation of the set of minimal explanation vectors $E^{\min}(M^{\langle i\rangle}, x_{ob}^{\langle i\rangle})$ and basis markings $\mathcal{M}^{\min,\langle i\rangle}$ for $i \in \{k-h+2,\ldots,k\}$, which may show to be time consuming. Finally, it leads to a sliding procedure: The next step needs the computation of $E^{\min\langle k-h+2\rangle>}$ and $\mathcal{M}^{\min,\langle k-h+2\rangle>}$, and so on.

Remark 4.4. Note that the detection of a fault with certainty can be unsuccessful for a given horizon length h, while a greater window which exploits more information can lead to detection even if the exact occurrences can remain unknown. Symmetrically, $J^{-}(T_f^j,k) \geq 1$ does not give information on iterations $\langle k-h+j\rangle$ except that the detected faults of class T_f^j are relevant to the current window $[x_{ob}^{\langle i\rangle}]_{k-h+1}^{k}$.

Now, we establish the relationships between the partial diagnosis status $\Delta_h^k(w,T_f^j)$ obtained for each window $[x_{ob}^{\langle i\rangle}]_{k-h+1}^{k}$ with $k \in [h\cdots|w|]$ and the complete diagnosis state $\Delta(w,T_f^j)$ associated with the complete observation window $[x_{ob}^{\langle i\rangle}]_1^{|w|}$ from M^{init}.

Definition 4.5. For an observed word w from M^{init}, we define the diagnosis function $\Delta(w,T_f^j)$ associated with T_f^j as follows:

$$\begin{cases} \Delta(w,T_f^j)=N: & \text{Fault class } T_f^j \text{ does not occur for the observed word } w \text{ from } M^{init}.\\ \Delta(w,T_f^j)=U: & \text{We can not conclude on the existence of } T_f^j \text{ for the observed word } w \text{ from } M^{init}.\\ \Delta(w,T_f^j)=F: & \text{Fault } T_f^j \text{ occurs for the observed word } w \text{ from } M^{init}. \end{cases}$$

Proposition 4.6. Let us consider an observed word w from M^{init}. The diagnosis state $\Delta(w,T_f^j)$ associated with T_f^j can be deduced from the partial diagnosis states on the sliding window of length h as follows:

1. If $\forall k \in [h\cdots|w|]$, $\Delta_h^k(w,T_f^j)=N_h$, then $\Delta(w,T_f^j)\in\{U,N\}$.
2. If $\exists k \in [h\cdots|w|]$, $\Delta_h^k(w,T_f^j)=F_h$, then $\Delta(w,T_f^j)=F$.
3. If $\exists k \in [h\cdots|w|]$, $\Delta_h^k(w,T_f^j)=U_h$, then $\Delta(w,T_f^j)=U$.

Proof.

1) If for all steps $k \in [h\cdots|w|]$, $\Delta_h^k(w,T_f^j)=N_h$, then no fault of class T_f^j is contained in any minimal sequence consistent with an observation window of length h, but there may exist one non minimal sequence that is consistent with w which contains T_f^j, so $\Delta(w,T_f^j)=\{U,N\}$.

2) We suppose that there exists step $k \in [h\cdots|w|]$ such that $\Delta_h^k(w,T_f^j)=F_h$, then fault class T_f^j surely occurs on window $[x_{ob}^{\langle i\rangle}]_{k-h+1}^{k}$ from all $M^{\langle k-h+1\rangle} \in \mathcal{M}^{\min,\langle k-h+1\rangle}$. Therefore, all firable sequences which are consistent with $x_{ob}^{\langle k-h+1\rangle}$

$\cdots x_{ob}^{\langle k \rangle}$ from any $M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\langle k-h+1 \rangle}$ necessarily contain at least one fault transition of T_f^j. Consequently, all firable sequences consistent with $x_{ob} = x_{ob}^{\langle 1 \rangle} \cdots x_{ob}^{\langle k-h+1 \rangle} \cdots x_{ob}^{\langle k \rangle} \cdots x_{ob}^{\langle |w| \rangle}$ from $\mathcal{M}^{\langle 1 \rangle} = M^{init}$ contain at least one fault transition of T_f^j, and then $\Delta(w, T_f^j) = F$.

3) We suppose that there exists some step $k \in [h \cdots |w|]$ such that $\Delta_h^k(w, T_f^j) = U_h$. Hence, there exists $M^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min, \langle k-h+1 \rangle}$ from which there exists a firable sequence consistent with $x_{ob}^{\langle k-h+1 \rangle} \cdots x_{ob}^{\langle k \rangle}$ which contains some fault in T_f^j. There exists also $M'^{\langle k-h+1 \rangle} \in \mathcal{M}^{\min, \langle k-h+1 \rangle}$ from which there exists a firable sequence which is consistent with $x_{ob}^{\langle k-h+1 \rangle} \cdots x_{ob}^{\langle k \rangle}$, and which does not contain any fault in T_f^j. As a result, there exists a firable sequence that is consistent with x_{ob} from M^{init}, and which contains some fault in T_f^j. However, we can not conclude on the existence of a sequence consistent with x_{ob} from M^{init}, which does not include a fault transition of T_f^j. Hence, $\Delta(w, T_f^j) = U$.

□

We can now make the following connections with four levels of diagnosis which express an increasing level of alarm from 0 to 3 described in [17] and mainly based on the notions of justification defined in section 2.2.

- $\Delta(w, T_f^j) \in \{0, 1\}$ if all justifications consistent with x_{ob} from M^{init} do not contain a fault transition. Level 0 expresses the absence of faults in a minimal or non minimal sequence consistent with x_{ob}, which corresponds to status N. Level 1 expresses the absence of faults in minimal sequences consistent with x_{ob} and the presence of fault in a non-minimal sequence consistent with x_{ob}, which corresponds to status U. Therefore, this diagnostic state is equivalent to $\Delta(w, T_f^j) = \{N, U\}$ defined in proposition 4.6.

- $\Delta(w, T_f^j) = 2$ if there is a pair of justifications where an element of the pair contains a fault transition, whereas the other one does not, which corresponds to status U.

- $\Delta(w, T_f^j) = 3$ if all justifications contain a fault transition. This diagnostic state is equivalent to $\Delta(w, T_f^j) = F$ defined in proposition 4.6.

Based on the diagnosis approach defined in [17], it is possible to distinguish between diagnosis states 0 and 1 based on an ILP problem defined as a function of basis marking $\mathcal{M}^{\min, \langle |w|+1 \rangle}$. This distinction is possible by the present sliding window approach only for the two particular cases $h = 1$ and $h = |w|$.

The proposed diagnosis can be seen as an additional module that completes the initial approach of estimation [17] which alternates the computation of the basis markings and the minimal explanation vectors. The two procedures are synchronized by the end of the computation of the set of basis markings. Hence, in practice and for the sake of efficiency, the operations can be distributed in two units which work in parallel. The additional module presents some numerical advantages with respect to the diagnosis of the standard

approach in [17]. In fact, the computation of sets $\mathcal{M}^{\min,\langle i\rangle}$ and $E^{\min}(M^{\langle i\rangle}, x_{obs}^{\langle i\rangle})$ for $i \in \{k-h+2,\ldots,k\}$ is not necessary for the diagnosis relevant to a given step k and the corresponding window $[x_{ob}^{\langle i\rangle}]_{k-h+1}^{k}$ even if these sets are necessary at the following steps.

5. DISCUSSION

In this section, we present a comparison between the diagnostic approach proposed in this article with the approach in [17] based on the computation of basis markings and also with the algebraic approach proposed in [7] based on the resolution of ILP problems on a receding horizon.

In [7], a diagnosis approach of partially observed LPN was based on the estimation of a receding horizon, which corresponds to the sliding horizon observer described in this paper for the particular case where $h = |w|$, i. e. by considering all the observations in the same window (in this case $N = N_h$, $U = U_h$ and $F = F_h$). For fault detection based on an observed word w from M^{init}, two ILP problems are defined with each $|w|.|T_{un}|$ variables. However, one limitation of this technique is that the horizon cannot be increased exhaustively, as the resolution of an ILP problem is exponential in the worst case. Therefore, diagnosis on a sliding horizon which replaces the receding horizon allows for avoiding the problem of increasing matrices especially for large networks, using a limited number of observations in each estimation step. In fact, the modification of the fault indicators related to the optimization problems based on estimation on a receding horizon makes it possible to use the partial measurements of the state to update the fault indicators. Partial diagnostic states over a sliding horizon are then provided, which makes it possible to interpret the diagnosis state for the observed word w from M^{init}. Fault detection is then based on the resolution of $2 * (|w| - h + 1)$ ILP problems of $h.|T_{un}|$ variables each with $h \leq |w|$. This reduces the complexity of the system in the worst case, and allows a reduction in the length of the observation window. However, the decrease in the value of h leads to an increase in the number of calculation windows and consequently leads to the calculation of start markings (normally, basis markings) for each additional window. Nevertheless, it's will known that the computation of basis markings suffers from some drawbacks.

The diagnostic approach proposed in [17], based on basis marking estimation, can be considered as a diagnosis approach on a sliding horizon for the particular case $h = 1$ (by considering a single observation at each estimation step). In section 3.3, we have shown that unfruitful basis markings would lead to inconsistent sequences. It is worth noting that the inconsistency of a marking can appear not only at the next estimation step as for the PN of Figure 1 but also after many estimation steps (if we consider one observed transition for each estimation step as in [17]): If $M(p_5) = 100$, then 101 firings of y_2 are necessary, and so are 101 estimation steps. To reduce the presence of these unfruitful basis markings, we consider a set of data taken together and not separated at each step. In other words, the checking of the basis markings with a consistency analysis is made not on an elementary window based on a unique observation $x_{ob}^{\langle k\rangle}$ (as in [17]) but on a window of observations in order to anticipate the future behavior of the process, from its current state, over a finite sliding horizon.

In fact, the previous drawback is not the only one. Indeed, the approach using the basis markings suffers also from a state explosion problem already for very small PNs (see Table 6 associated with the LPN of Figure 2 for $M_0 = [80\ 80\ 0\ 0\ 0\ 0\ 0]^T$). In this paper, the proposed technique is based on a diagnostic module which uses a set of basis markings but does not need the last computed sets.

To conclude, the diagnostic approach developed in this paper makes the best possible trade off between the approaches suggested in [7] and [17] by choosing the appropriate value of h. Moreover, the proposed approach is adaptive as the CPU time of the additional module depends on the horizon length which can be adapted to the available time. However, an increase in the speed of diagnosis on the sliding window may come at the expense of a decrease in the accuracy of the results (uncertain verdicts).

6. EXAMPLE 3

We consider the LPN of Figure 2 proposed in [19] which models a part of a large manufacturing system. Let $T_{ob} = \{y_1, y_2, y_3, y_4, y_5, y_6\}$, $E = \{a, b, c, d, e, f\}$, $T_{un} = \{x_7, x_8, x_9, x_{10}\}$ and one single fault class $T_f = \{x_9, x_{10}\}$. The observable transitions are labeled as follows: $L(y_1) = a$, $L(y_2) = b$, $L(y_3) = c$, $L(y_4) = d$, $L(y_5) = e$, $L(y_6) = f$.

Let us consider an observed word $w = abeeedac$ associated with firing sequence $x_{ob} = x_1x_2x_5x_5x_5x_4x_1x_3$ from initial marking $M^{init} = \begin{pmatrix} 20 & 20 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T$. Table 1 represents the different diagnosis results exploiting the sliding window of length $h = 3$, $h = 4$ and $h = 5$. We denote by Δ_1 and Δ_2 the diagnosis status of fault transitions x_9 and x_{10}, respectively. The calculations are performed on an Intel PC with a clock of 1.80 Ghz as in [19]. The ILP problems use the glpk function of the software GNU module Octave which is mostly compatible with Matlab and close to Scilab.

The detection of fault x_9 occurs at the same step 5 for $h = 3$, 4 and 5. At step 5 with $h = 3$, starting from the three basis markings $\begin{pmatrix} 17 & 17 & 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}^T$, $\begin{pmatrix} 17 & 17 & 1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}^T$ and $\begin{pmatrix} 17 & 17 & 1 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}^T$ computed at step 3, the firing of x_9 is necessary for the observations $y_5y_5y_5$. Therefore, we determine the window of observations where the fault has occurred. The firing date is possible between the firing of y_2 and the last firing of y_5 as the occurrences of observable events are assumed to be non-simultaneous.

The next steps show the "forgetfulness" of past faults. Particularly, the status is N_h for x_9 at step 8 with $h = 3$. Starting from the basis marking that is unique here, and equal to $\begin{pmatrix} 19 & 19 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}^T$ the firing of x_9 cannot occur for observations $y_4y_1y_3$. As a consequence, a new occurrence of fault x_9 can be detected and can be distinguished from the first occurrence.

The computation at step k must be repeated for all basis markings of set $\mathcal{M}^{\min,\langle k-h+1\rangle}$ computed at step $k - h$. If $k = 7$, the technique depends on the basis markings of $\mathcal{M}^{\min,<5>}$ computed at step 4 for $h = 3$ with $|\mathcal{M}^{\min,<5>}| = 6$. The computation at step k does not depend on sets $\mathcal{M}^{\min,\langle k-h+2\rangle>}, \mathcal{M}^{\min,\langle k-h+3\rangle>}, \ldots, \mathcal{M}^{\min,\langle k\rangle}$ (in the example, not $\mathcal{M}^{\min,<6>}, \mathcal{M}^{\min,<7>}$) and on their computations. Based on an increasing horizon technique [7], as well as the classical approach based on the basis markings [17], we obtain the same result for fault x_9 ($\Delta_1 = F$), whereas the diagnosis state for x_{10} is $\Delta_2 = N$ which is a more accurate result than ours (U).

Step k	h=3 Observations - (Δ_1, Δ_2) - time(s)	h=4 Observations - (Δ_1, Δ_2) - time (s)	h=5 Observations - (Δ_1, Δ_2) - time (s)
1	-	-	-
2	-	-	-
3	$y_1y_2y_5 - (U_h, U_h) - 0.0560$	-	-
4	$y_2y_5y_5 - (U_h, U_h) - 0.0948$	$y_1y_2y_5y_5 - (U_h, U_h) - 0.0580$	-
5	$y_5y_5y_5 - (F_h, N_h) - 0.0906$	$y_2y_5y_5y_5 - (F_h, N_h) - 0.0958$	$y_1y_2y_5y_5y_5 - (F_h, N_h) - 0.0500$
6	$y_5y_5y_4 - (U_h, N_h) - 0.2516$	$y_5y_5y_5y_4 - (F_h, N_h) - 0.1156$	$y_2y_5y_5y_5y_4 - (F_h, N_h) - 0.0888$
7	$y_5y_4y_1 - (U_h, N_h) - 0.1039$	$y_5y_5y_4y_1 - (U_h, N_h) - 0.2536$	$y_5y_5y_5y_4y_1 - (F_h, N_h) - 0.0916$
8	$y_4y_1y_3 - (N_h, N_h) - 0.0683$	$y_5y_4y_1y_3 - (U_h, N_h) - 0.0959$	$y_5y_5y_4y_1y_3 - (U_h, N_h) - 0.2156$
Result	$x_{ob} - (F, U) - 0.6652$	$x_{ob} - (F, U) - 0.6188$	$x_{ob} - (F, N/U) - 0.4460$

Tab. 1. Example 3: Diagnostic approach based on sliding horizon for $M^{init} = (20 \;\; 20 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0)^T$.

To evaluate the effectiveness of the proposed fault diagnosis approach based on the sliding window with respect to the discrete approach in [17], we reconsider the above example for the initial marking $M^{init} = \begin{pmatrix} 80 & 80 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T$ for observed sequences $Y5 = y_1y_1y_1y_1y_2y_2y_2y_2y_5y_5y_5y_5y_5y_5y_5y_5y_5y_5$, $Y6 = y_1y_1y_1y_1y_2y_2y_2y_2y_5y_5$ $y_5y_5y_5y_5y_5y_5y_5y_5y_5y_5y_4y_4y_4y_4$ and $Y7 = y_1y_1y_1y_1y_2y_2y_2y_2y_5y_5y_5y_5y_5y_5y_5y_5y_5y_5y_4$ $y_4y_4y_4y_1y_1y_1y_1$, for which there exists an explosion of the basis marking number. Table 6 shows the effectiveness of the proposed approach compared to the discrete approach in [17] in terms of computation time. We also note that for $h = 20$, we obtain a lower computation time and also better precision than that for $h = 12$ and $h = 16$ ("U" becomes "U/N" for x_{10}) for all the considered observed sequences $Y5, Y6$ and $Y7$.

Observed sequence x_{ob}	**Sliding window**			**Basis marking approach**
	$h = 12$	$h = 16$	$h = 20$	
	(Δ_1, Δ_2) - time(s)	(Δ_1, Δ_2) - time(s)	(Δ_1, Δ_2) - time(s)	(Δ_1, Δ_2) - time(s)
$Y5$	$(F, U) - 4$	$(F, U) - 3$	$(F, U/N) - 2$	(F,N)-365
$Y6$	$(F, U) - 30$	$(F, U) - 5$	$(F, U/N) - 3$	(F,N)-593
$Y7$	$(F, U) - 333$	$(F, U) - 36$	$(F, U/N) - 5$	(F,N)-1332

Tab. 2. Numerical comparison.

7. CONCLUSION

In this paper, we have presented an algebraic approach for fault class diagnosis in an LPN system. The proposed approach has been based on the state estimation on a sliding window of length h and the fault detection has been performed by the ILP problem resolution. For an observed word w, a set of $|w| - h + 1$ observation windows have been defined such that each window would include exactly h observed transitions. For each observation window i, a polyedron of form $A_i.x_i \leq b_i$ has been defined, where A_i depends on a sub-structure of the LPN. On the other hand, the data vector b_i depends on the starting markings of the considered window i which is a set of basis markings computed at step i by the classical approach developed in [17]. We have considered the extreme values of h. For $h = 1$, the present approach is equivalent to the discrete

approach [17] for which we have to compute all the basis markings associated respectively with each observed transition of w. If we consider the extreme case $h = |w|$, then fault detection is based on a single observation window that includes all the transitions of the observation. In this case, our approach is equivalent to the diagnosis on a receding horizon, as discussed in [7]. The drawback of the last mentioned approach is that the system $A_i.x_i \leq b_i$ becomes wide and its resolution complex when w is long since the resolution of an ILP problem is exponential (w.r.t. the number of system variables) in the worst case. If we reduce the value of h, we decrease the size of the sub-systems to be solved, but in return we increase the number of starting basis markings that we must calculate, which may lead to a combinatorial explosion.

The numerical tests show that the information flows can be compatible as the technique is numerically efficient. It is worth noting that the choice of the horizon length which is connected to the accuracy of the diagnosis is not fixed and can be adapted to the used available computation facilities.

A perspective is to avoid the time-costly computations of minimal explanation vectors and basis markings. For this, we will show in some future work that the modeling of time allows considering the sequences directly. An other advantage is to process the LPNs with cyclic unobservable subnet.

(Received November 27, 2021)

REFERENCES

[1] M. P. Cabasino, A. Giua, M. Pocci, and C. Seatzu: Discrete event diagnosis using labeled Petri nets. An application to manufacturing systems. Control Engrg. Practice *19* (2011), 9, 989–1001. DOI:10.1016/j.conengprac.2010.12.010

[2] G. Jiroveanu and K. B. René: The diagnosability of Petri net models using minimal explanations. IEEE Trans. Automat. Control *55* (2010), 7, 1663–1668. DOI:10.1109/TAC.2010.2046106

[3] A. Chouchane and P. Declerck: Diagnostic de réseaux de Petri partiellement observables avec indicateurs algébriques. Génie industriel et productique *2* (2019), 1, 11–25.

[4] N. Ran, S. Wang, H. Su, and C. Wang: Fault diagnosis for discrete event systems modeled by bounded Petri nets. Asian J. Control *19* (2017), 4, 1532–1541. DOI:10.1002/asjc.1500

[5] D. Lefebvre: On-line fault diagnosis with partially observed Petri nets. IEEE Trans. Automat. Control *59* (2013), 7, 1919–1924. DOI:10.1109/TAC.2013.2294617

[6] F. Basile, C. Pasquale, and D. T. Gianmaria: An efficient approach for online diagnosis of discrete event systems. IEEE Trans. Automat- Control *54* (2009), 4, 748–759. DOI:10.1109/TAC.2009.2014932

[7] A. Chouchane, P. Declerck, A. Khedher, and A. Kamoun: Diagnostic based on estimation using linear programming for partially observable Petri nets with indistinguishable events. Int. J. Systems Science: Operations and Logistics *7* (2020), 2, 192–205. DOI:10.1080/23302674.2018.1554169

[8] A. Chouchane: Analytical redundancy relationship generation on a progressive horizon for fault diagnosis of a labelled Petri net. IMA J. Math. Control Inform. *38* (2021), 3, 908–928. DOI:10.1093/imamci/dnab015

[9] A. Chouchane, A. Khedher, O. Nasri, and A. Kamoun: Diagnosis of partially observed Petri net based on analytical redundancy relationships. Asian J. Control *21* (2019), 5, 2218–2231. DOI:10.1002/asjc.1832

[10] M. P. Cabasino, G. Alessandro, and C. Seatzu: Diagnosis using labeled Petri nets with silent or undistinguishable fault events. IEEE Trans. Systems Man Cybernet.: Systems *43* (2012), 2, 345–355. DOI:10.1109/TSMCA.2012.2199307

[11] G. Jiroveanu, K. B. René, and B. Behzad: On-line monitoring of large Petri net models under partial observation. Discrete Event Dynamic Systems *18* (2008), 3, 323–354. DOI:10.1007/s10626-007-0036-x

[12] Y. Tong, Z. Li, C. Seatzu, and A. Giua: Verification of state-based opacity using Petri nets. IEEE Trans. Automat. Control *62* (2016), 6, 2823–2837. DOI:10.1109/TAC.2016.2620429

[13] A. Boussif, L. Baisi, and M. Ghazel: An experimental comparison of three diagnosis techniques for discrete event systems. In: DX'17-28th International Workshop on Principles of Diagnosis, 2017.

[14] A. Chouchane: Estimation et diagnostic de réseaux de Petri partiellement observables. Diss. Université d'Angers; École nationale d'ingénieurs de Sfax 2018.

[15] L. Li and C. N. Hadjicostis: Least-cost firing sequence estimation in labeled Petri nets with unobservable transitions. American Control 2007 DOI:10.1109/ACC.2007.4282814

[16] G. Stremersch and K. B. René: Structuring acyclic Petri nets for reachability analysis and control. Discrete Event Dynamic Systems *12* (2002)1, 7–41. Conference, IEEE, 2007. DOI:10.1023/A:1013331703036

[17] M. P. Cabasino, G. Alessandro, and C. Seatzu: Fault detection for discrete event systems using Petri nets with unobservable transitions. Automatica *46* (2010), 9, 1531–1539. DOI:10.1016/j.automatica.2010.06.013

[18] T. Murata: Petri nets: Properties, analysis and applications. Proc. IEEE *77* (1989), 4, 541–580. DOI:10.1109/5.24143

[19] C. Mahulea, C. Seatzu, M. P. Cabasino, and M. Silva: Fault diagnosis of discrete-event systems using continuous Petri nets. IEEE Trans. Systems Man Cybernetics – Part A: Systems and Humans *42* (2012), 4, 970–984. DOI:10.1109/TSMCA.2012.2183358

Amira Chouchane, COSYS-ESTAS, Univ. Gustave Eiffel, Lille, F-59650 Villeneuve d'Ascq. France.
e-mail: amira.chouchane@univ-eiffel.fr

Philippe Declerck, LARIS, Université d'Angers, 49000 Angers. France.
e-mail: philippe.declerck@univ-angers.fr