Diagnosis of Stochastic Systems: Optimising Costs and Delays

Marie Duflot, Engel Lefaucheux, Isaline Plaid

▶ **To cite this version:**

Marie Duflot, Engel Lefaucheux, Isaline Plaid. Diagnosis of Stochastic Systems: Optimising Costs and Delays. QEST-FORMATS 2024, Sep 2024, Calgary, Alberta, Canada, Canada. hal-04617663

HAL Id: hal-04617663

https://inria.hal.science/hal-04617663

Submitted on 19 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Diagnosis of Stochastic Systems: Optimising Costs and Delays

Marie Duflot[1 [0000−0002−7538−8826]], Engel Lefaucheux[1 [0000−0003−0875−300X]], and Isaline Plaid[1,2 [0009−0008−3264−7892]]

[1] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
[2] ENS de Lyon

Abstract. Diagnosing a partially observable stochastic system consists in telling, based on the actions observed along a path, whether the system has encountered an (unobservable) event called fault. Several notions of diagnosability have been studied in this context depending on whether an (arbitrarily small) error probability is tolerated upon claiming a fault or not. Choosing which events can effectively be observed through sensors has a strong impact on diagnosis. With more observable events a given system is in general more likely to be diagnosable, and in case it is diagnosable it is more likely to detect a fault sooner. However, detecting events can have a cost, for instance in terms of sensors needed to detect those events. In this paper we tackle both diagnosability notions while restricting the set of observable events. First we try to balance costs and diagnosability, studying whether a system is diagnosable with less than N observable events. Then we study the latency of diagnosis, a measure that computes the mean time to wait between the occurrence of a fault and the moment when it is detected. We also consider two types of masks: static ones in which the set of observable events is fixed, and dynamic masks for which the set of observable events can evolve along an execution, based on what has been observed so far.

Keywords: Stochastic systems, Partial observation, Control, Fault diagnosis

1 Introduction

Diagnosis and diagnosability Diagnosis, as seen in the discrete event systems community, refers to the observation of a system prone to faults, and the swift detection of the aforementioned malfunctions. It involves both deciding whether there is a way to detect every faulty behaviour of the system, we then call the system *diagnosable*, and crafting the *diagnoser*, *i.e.* the process which will automatically detect whether the fault occurred based on observations. Diagnosers are often required to satisfy various criteria, such as memory size, detection speed, accuracy, etc. Diagnosers hold similarities to the monitors used in runtime verification [1], though specific to a problem (diagnosis) which goes beyond most logic considered for monitors, hence relying on different techniques.

Diagnosis of discrete event systems In [11], the authors proposed to model the systems by partially observable labelled transition systems (LTS). In their framework, the systems produce runs, associated to sequences of observable events, and diagnosability requires that the occurrence of faults within a run can be deduced from the sequence of observable events. Formally, an LTS is diagnosable if there exists a diagnoser that satisfies two properties: *reactivity* and *correctness*. Reactivity requires that whenever a fault occurred, the diagnoser eventually detects it. Correctness asks that the diagnoser only claims the existence of a fault when there actually was one. Checking diagnosability of an LTS was shown to be possible in polynomial time [8].

Diagnosis of stochastic systems In order to represent quantitative aspects of the model, LTS were extended to include probabilistic components. This extension was called probabilistic labelled transition systems (pLTS) [12, 3]. This model is similar to Markov chains, with labels on transitions. Due to the stochastic behaviour of the systems, the requirements on the diagnoser were relaxed: faults needs to be detected almost surely and some false positive can be tolerated. More precisely, concerning correctness, three natural specifications were considered. A first option, called FF-*diagnosability*, does not tolerate any error in the claims made by the diagnoser. In contrast, given an error threshold ε, $\varepsilon\mathsf{FF}$-*diagnosability* tolerates small errors, allowing to claim a fault if the conditional probability that no fault occurred does not exceed the threshold. Last, AFF-*diagnosability* requires the pLTS to be $\varepsilon\mathsf{FF}$-diagnosable for every $\varepsilon > 0$.

Remaining issues It is known how to decide whether a diagnoser satisfying the various notions of diagnosability exist for a pLTS. More recent works have focused on enhancing the diagnosis with additional properties or controlling a system in order to make it diagnosable [2, 9]. For instance, in [4], the authors combined the diagnosis objective with a safety condition in order to avoid faults as much as possible. Some works have also aimed at minimising the cost of actually diagnosing the system [14, 7]. To our knowledge however, this line of research was never tackled over probabilistic models.

Contributions In this paper, we are interested in optimising the cost, and reaction speed of the diagnosers for pLTS. More precisely, we show how to decide FF-diagnosability and AFF-diagnosability when the number of sensors (observable events) is bounded by a given number, or when an energy cost restriction on the number of sensors simultaneously turned on is imposed. We also introduce the latency of a diagnoser, a notion that represents how slow its detection of failures is on average. This is helpful for comparing multiple diagnosers with the same, or similar costs, and finding the more efficient one. We show how to compute the latency with respect to FF-diagnosers.

2 Preliminaries

2.1 Probabilistic Labelled Transition System

We consider stochastic discrete event systems where transitions are labelled with events.

Definition 1. *A* probabilistic labelled transition system *(pLTS) is a tuple* $\mathcal{A} = (Q, q_0, \Sigma, T, \mathbf{P})$ *where:*

- Q *is a set of states with* $q_0 \in Q$ *the initial state;*
- Σ *is a finite set of events, including a set* Σ_u *of unobservable events;*
- $T \subseteq Q \times \Sigma \times Q$ *is a set of transitions;*
- $\mathbf{P}$ *is a probability function from* T *to* $\mathbb{Q}_{>0}$ *fulfilling for all* $q \in Q$*:* $\sum_{a \in \Sigma, q' \in Q \mid (q,a,q') \in T} \mathbf{P}(q, a, q') = 1$.

Except if explicitly stated otherwise, we assume Q to be finite. We denote by $q \xrightarrow{a} q'$ the transition $(q, a, q') \in T$. By definition of the probability function, there exists at least one outgoing transition (q, a, q') in every state q of the pLTS.

A *run* of a pLTS $\mathcal{A}$ is a (finite or infinite) sequence $\rho = q_0 a_0 q_1 \dots$ such that for all i such that $a_i \in \rho$, we have $q_i \in Q$, $q_{i+1} \in Q$, $a_i \in \Sigma$ and $q_i \xrightarrow{a_i} q_{i+1}$. The length (finite or infinite) of a run ρ, denoted by $|\rho|$, is the number of events occurring in it. The event sequence associated with ρ is the word $\sigma_\rho = a_0 a_1 \dots$. We write $\Omega_\mathcal{A}$ for the set of all infinite runs in $\mathcal{A}$ starting from the initial state q_0, dropping the subscript if the pLTS is clear from context.

Given a finite run $\rho = q_0 a_0 q_1 \dots q_n$ and a (finite or infinite) run $\rho_2 = q_n a_n q_{n+1} \dots$ starting in the last state of ρ, we call concatenation of ρ and ρ_2 the run $\rho\rho_2 = q_0 a_0 q_1 \dots q_n a_n q_{n+1} \dots$. The run ρ is then a *prefix* of $\rho\rho_2$, which we denote by $\rho \preceq \rho\rho_2$ and we use the notation $\prec$ for a strict prefix. The *cylinder* generated by a finite run ρ consists of all the infinite runs that extend ρ: $Cyl(\rho) = \{\rho' \in \Omega_\mathcal{A} \mid \rho \preceq \rho'\}$.

Probabilities can be defined for cylinders of runs as usual by:

$$\mathbb{P}(Cyl(q_0 a_0 q_1 \dots q_n)) = \mathbf{P}(q_0, a_0, q_1) \times \dots \times \mathbf{P}(q_{n-1}, a_{n-1}, q_n) \ .$$

As usual, the probability measure is extended to measurable subsets of $\Omega_\mathcal{A}$ by Caratheodory's extension theorem. To simplify notations, given a finite run ρ, we will sometimes abuse notation and write $\mathbb{P}(\rho)$ for $\mathbb{P}(Cyl(\rho))$. Similarly if R is a denumerable set of finite runs such that no run is a prefix of another one, $\mathbb{P}(R) = \sum_{\rho \in R} \mathbb{P}(\rho)$ (which is consistent since all intersections of associated cylinders are empty).

2.2 Partial observation and ambiguity

Observation of the events in the system modelled by the pLTS is done with various sensors. As such, the sequence of events occurring within a run may not be completely observable. We represent this with the use of a *mask* function $\mathcal{M} : \Sigma^* \mapsto \mathsf{Dist}(\Sigma^*)$ that hides part of an event sequence to an observer. Intuitively, a mask will select a set $\Sigma^\bullet$ of actions to monitor and once one action from $\Sigma^\bullet$ is

produced by the run, it may change its choice for a new set. To represent this, the mask function is defined conjointly with a selection function $\mathcal{S}_\mathcal{M}$ associating to every sequence $\Sigma_1^\bullet a_1 \dots \Sigma_n^\bullet a_n$ where for all $i \leq n$ $a_i \in \Sigma_i^\bullet$, a distribution on the set of events 2^Σ such that any set $\Sigma^\bullet$ selected by $\mathcal{S}_\mathcal{M}$ satisfies $\Sigma^\bullet \cap \Sigma_u = \emptyset$. This latter condition is used to represent that some events, Σ_u, do not have any available sensor and thus can never be monitored.

The function $\mathcal{M}$ associates to the event sequence $\sigma_\rho = a_1 \dots a_n$ of a run ρ a distribution over subwords of σ_ρ which we call the observed sequences of ρ under $\mathcal{M}$: for a subword $\sigma' = a_{i_1} \dots a_{i_m}$ of σ_ρ, denoting $i_0 = 1$, $i_{m+1} = n+1$, $H(k,r) = \{\Sigma^\bullet \in 2^{(\Sigma \setminus \Sigma_u)} \mid \text{ for all } k < j < r, a_j \notin \Sigma^\bullet\}$ the observable sets that contain no event strictly between a_k and a_r and $\Lambda = \{\Sigma_1^\bullet, \cdots, \Sigma_{m+1}^\bullet \in 2^{(\Sigma \setminus \Sigma_u)} \mid \forall j \leq m+1, \Sigma_j^\bullet \in H(i_{j-1}, i_j) \wedge \forall j \leq m, a_{i_j} \in \Sigma_j^\bullet\}$ we have

$$\mathcal{M}(\sigma_\rho)(\sigma') = \sum_{\Sigma_1^\bullet, \cdots, \Sigma_{m+1}^\bullet \in \Lambda} \mathcal{S}_\mathcal{M}(\epsilon)(\Sigma_1^\bullet) \times \mathcal{S}_\mathcal{M}(\Sigma_1^\bullet a_{i_1})(\Sigma_2^\bullet) \times \cdots \times \mathcal{S}_\mathcal{M}(\Sigma_1^\bullet a_{i_1} \dots \Sigma_m^\bullet a_{i_m})(\Sigma_{m+1}^\bullet)$$

By abuse of notation, for a run ρ, we write $\mathcal{M}(\rho)$ for $\mathcal{M}(\sigma_\rho)$.

As an example, partial observation of a system is often done by introducing a set Σ_o of *observable events*. The observed sequence of a run in this context is the sequence of observable events occurring within it. The mask achieving the same representation is a projection that can be produced by the selection function which returns to any sequence a Dirac distribution on Σ_o. Following this example, we say that a mask $\mathcal{M}$ is *static* if there exists a set $\Sigma_\mathcal{M} \subseteq \Sigma$ such that $\mathcal{M}$ is the projection as defined above on the events of $\Sigma_\mathcal{M}$. In this case, we say that $\Sigma_\mathcal{M}$ is the set observable through the static mask $\mathcal{M}$. The distribution produced by a static mask $\mathcal{M}$ on an event sequence σ is a Dirac. When using any mask which selection function only produces Dirac distributions (which we calls *deterministic* masks), we will identify the Dirac distribution of the mask with the event sequence selected by this distribution.

The *size of a mask* $\mathcal{M}$ denoted by $|\mathcal{M}| \in \mathbb{N}$ is the size of the largest set of event that can be selected by $\mathcal{S}_\mathcal{M}$ at a given point: $|\mathcal{M}| = \max\{|\Sigma^\bullet| \mid w \in (2^\Sigma \times \Sigma)^*, \mathcal{S}_\mathcal{M}(w)(\Sigma^\bullet) > 0\}$.

Example 1. Consider the pLTS represented in Figure 1 under the static mask $\mathcal{M}$ associated to $\{a, b\}$. In this pLTS, when probabilities are not specified, we assume a uniform distribution. The runs $\rho = q_0 u q_2 a q_2 \mathbf{f} f_1 b f_2 a f_2$ and $\rho' = q_0 u q_1 (a q_1)^\omega$ have the associated event sequences $\sigma_\rho = ua\mathbf{f}ba$ and $\sigma_{\rho'} = ua^\omega$ and the observed sequence $\mathcal{M}(\rho) = aba$ and $\mathcal{M}(\rho') = a^\omega$. The size of this mask is 2. ρ has probability $1/32$, ρ' being infinite the probability is not directly defined on it, but each of its finite prefixes has probability $3/4$.

Consider now the same pLTS, but under the mask function $\mathcal{M}'$ which initially observes $\{b\}$, and once a b occurs, switch sensor and becomes static, only observing a. Then $\mathcal{M}(\rho) = ba$ as the first a is not observable before a b occurs and $\mathcal{M}(\rho') = \epsilon$.

Having an infinite run of the system with only a finite observed sequence, like we had in the above example, is something we wish to avoid as a system should

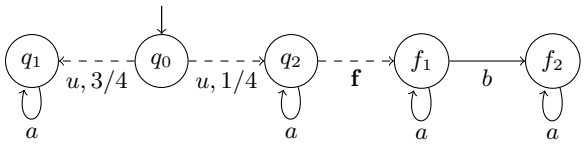

Fig. 1. pLTS under the static mask associated to the set $\{a, b\}$ (unobservable transitions are represented with a dotted line).

regularly give information to the user. As such, we require the pair of pLTS and mask we study to be *regular*: an infinite run will almost surely have an infinite observed sequence. In other words, if the pair $\mathcal{A}$ and $\mathcal{M}$ is regular, we have that $\mathbb{P}_{\mathcal{A}}(\{\rho \in \Omega_{\mathcal{A}} \mid \exists \sigma, \mathcal{M}(\rho)(\sigma) > 0 \wedge |\sigma| < \infty\}) = 0$.

Given a pair $\mathcal{A}$ and $\mathcal{M}$, and assuming $\mathcal{M}$ is static, one can verify that the pair is regular in polynomial time: we only need to check that from every state of the pLTS one can reach a transition that is observable under $\mathcal{M}$. From now on, we assume every pair $(\mathcal{A}, \mathcal{M})$ considered to be regular.

A *signalling run* with respect to $\mathcal{M}$ is a finite run $q_0 a_0 q_1 \cdots a_{n-1} q_n$ minimal for prefix relation, that allows to observe a given event sequence, *i.e.* such that there exists an event sequence σ such that $\mathcal{M}(q_0 a_0 q_1 \cdots q_{n-1})(\sigma) = 0$ and $\mathcal{M}(q_0 a_0 q_1 \cdots q_n)(\sigma) > 0$. We say that σ is a witness of the relevance of this run. Signalling runs are precisely the relevant runs w.r.t. partial observation issues since each new observable event provides additional information about the execution to an external observer. In the pLTS of Figure 1, $q_0 u q_2 a q_2 \mathbf{f} f_1$ is not a signalling run of witness a as its strict prefix $q_0 u q_2 a q_2$ shares the observed sequence a (and the latter is in fact a signalling run). In fact, if $\mathcal{M}$ is a static mask, signalling runs are exactly the runs ending with an event of $\Sigma_{\mathcal{M}}$ as defined in [3]. In the sequel, $\mathsf{SR}_{\mathcal{M}}$ denotes the set of signalling runs with respect to $\mathcal{M}$, $\mathsf{SR}_{\sigma,\mathcal{M}}$ the set of signalling runs with witness σ and $\mathsf{SR}_{n,\mathcal{M}}$ the set of signalling runs associated to a witness of length n. With a slight abuse of language, we will call the runs of $\mathsf{SR}_{n,\mathcal{M}}$ the signalling runs of observable length n.

2.3 Fault and diagnosis

We are interested in detecting failures within a system. To represent those failures, we rely on a special unobservable *fault* event $\mathbf{f} \in \Sigma$. A run ρ is *faulty* if its associated event sequence σ_ρ contains $\mathbf{f}$, otherwise it is *correct*. We write $F_{\mathcal{M}}$ (resp. $C_{\mathcal{M}}$) for the set of all finite faulty (resp. correct) signalling runs. We also write $F_{n,\mathcal{M}}$ (resp. $C_{n,\mathcal{M}}$) for their subsets of observable length $n \in \mathbb{N}$. W.l.o.g., by considering two copies of each state of the pLTS, we can assume that the state space Q of $\mathcal{A}$ is partitioned into correct states and faulty states: $Q = Q_f \uplus Q_c$ such that a run reaches a state of Q_f if and only if it is faulty. For a given mask $\mathcal{M}$, a finite event sequence $\sigma \in \Sigma^*$ is *ambiguous* if there exists a correct signalling run ρ and a faulty signalling run ρ' witnessed by σ. If $\mathcal{M}$ is deterministic, we say that a run is ambiguous if its observed sequence is ambiguous.

To every observed sequence $\sigma \in \Sigma^*$ under $\mathcal{M}$ associated to at least one run, we associate the *correctness proportion* w.r.t. a mask $\mathcal{M}$ representing the

likelihood that a run with observed sequence σ is correct:

$$\mathsf{CorP}_{\mathcal{M}}(\sigma) = \frac{\sum_{\rho \in C_{\mathcal{M}} \cap SR_{\sigma,\mathcal{M}}} \mathbb{P}(\rho)\mathcal{M}(\rho)(\sigma)}{\sum_{\rho \in SR_{\sigma,\mathcal{M}}} \mathbb{P}(\rho)\mathcal{M}(\rho)(\sigma)}.$$

This proportion can be used to select runs that are "sufficiently" ambiguous: for $\varepsilon \geq 0$, we define $\mathsf{FAmb}^{\varepsilon}_{n,\mathcal{M}}$ as the set of triples (ρ, p, σ) where ρ is a faulty signalling run with witness σ of length n satisfying $\mathsf{CorP}_{\mathcal{M}}(\sigma) > \varepsilon$, and $p = \mathcal{M}(\rho)(\sigma)$. Since a run ρ can have several witnesses, we keep track of the probability for ρ to produce an observation, to avoid counting multiple times the same probability in Definition 2.When $\mathcal{M}$ is deterministic, as the probability p is always 1 and σ is unique for a given run, we drop the second and third dimensions of $\mathsf{FAmb}^{\varepsilon}_{n,\mathcal{M}}$.

Multiple notions of diagnosis have been defined for stochastic systems. In this paper we focus on FF-diagnosability and AFF-diagnosability which were first introduced in [13]. The first notion aims at detecting almost surely faults without any ambiguity, while the latter allows for an arbitrarily small ambiguity.

Definition 2. *Let $\mathcal{A}$ be a pLTS. Let $\mathcal{M}$ be a mask.*

- *for a given $\varepsilon \geq 0$, $\mathcal{A}$ is εFF-diagnosable under a mask $\mathcal{M}$ if*

$$\lim_{n \to +\infty} \sum_{(\rho,p,\sigma) \in \mathsf{FAmb}^{\varepsilon}_{n,\mathcal{M}}} p.\mathbb{P}(\rho) = 0;$$

- *$\mathcal{A}$ is FF-diagnosable under a mask $\mathcal{M}$ if it is 0FF-diagnosable under $\mathcal{M}$;*
- *$\mathcal{A}$ is AFF-diagnosable under a mask $\mathcal{M}$ if it is εFF-diagnosable under $\mathcal{M}$ for all $\varepsilon > 0$.*

Example 2. Consider the pLTS of Figure 1, this pLTS is FF-diagnosable under the static mask associated to the set $\{a, b\}$. Indeed, the event b can only occur in a faulty run, thus the correctness proportion of any event sequence containing b is 0 and therefore the set $\mathsf{FAmb}^{0}_{n,\mathcal{M}}$ does not contain any run with the event b. Moreover, if a fault occurs in a run, with probability 1 the event b will occur later in this same run. Thus the probability of the sets $\mathsf{FAmb}^{0}_{n,\mathcal{M}}$ eventually converges to 0.

Consider now the pLTS of Figure 2 under the same mask. We can never be sure that a fault occurred as for every finite faulty run, there exists a finite correct run with the same observed sequence. Thus $\lim_{n \to +\infty} \mathbb{P}(\mathsf{FAmb}^{0}_{n,\mathcal{M}}) = 1/2$ (a run is faulty with probability 1/2 and all faulty runs are ambiguous) and the pLTS is not FF-diagnosable. However, in a faulty run, the event a will almost surely (due to the law of large numbers) occur less than the event b while the opposite holds for correct runs. As such, the longer we observe the run, the higher our confidence in whether this run is faulty or correct, allowing one to detect faulty runs with arbitrary accuracy. It is thus AFF-diagnosable.

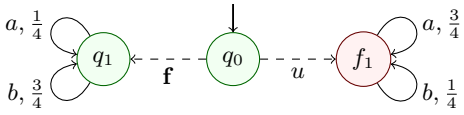

Fig. 2. This pLTS is not FF-diagnosable, but it is AFF-diagnosable under the static mask $\mathcal{M}$ associated to the set $\{a, b\}$.

2.4 Diagnosability under a given static mask

The Z-*diagnosability* problem (with $\mathbf{Z} \in \{\mathsf{FF}, \varepsilon\mathsf{FF}, \mathsf{AFF}\}$) consists in deciding whether a given pLTS is Z-*diagnosable*. The authors of [3] studied the diagnosability problems associated to the notions given in Definition 2 for a given static mask. This section aims at recalling their results. To make the static property of the mask clear, we add an S in all our notations of diagnosability.

S-FF-*diagnosability:* For a given pLTS $\mathcal{A} = (Q, q_0, \Sigma, T, \mathbf{P})$ and a static mask associated to a set $\Sigma_o \subseteq \Sigma$ of events, the authors of [3] first build a deterministic automaton $Obs(\mathcal{A})$ which accepts the non-ambiguous faulty observed sequences (see Figure 3). The construction is akin to a determinisation of the underlying automaton of $\mathcal{A}$, where the events of $\Sigma \setminus \Sigma_o$ are considered ϵ-transitions. Formally, $Obs(\mathcal{A}) = (S = 2^{Q_c}, s_0 = \{q_0\}, \Sigma_o, T', F = \emptyset)$ where $(U, a, U') \in T'$ iff $U' = \{q' \in Q \mid$ there exists $q \in U$ and ρ a correct signalling run starting in q and ending in q' such that $\mathcal{M}(\rho) = a\}$.

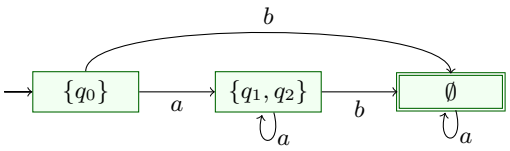

Fig. 3. $Obs(\mathcal{A})$ for the pLTS of figure 1

They then define the pLTS $\mathcal{A}_{FF} = \mathcal{A} \times Obs(\mathcal{A})$ as the product of $\mathcal{A}$ and $Obs(\mathcal{A})$ synchronised over observed events on which they establish the following characterisation and through it the decidability of S-FF-diagnosability:

Proposition 1 ([3]). *Let $\mathcal{A}$ be a finite pLTS and $\mathcal{M}$ be a static mask associated to a set $\Sigma_o \subseteq \Sigma$ of events. Then $\mathcal{A}$ is S-FF-diagnosable under $\mathcal{M}$ if and only if $\mathcal{A}_{FF}$ has no bottom strongly connected component (BSCC) containing a state (q, U) with $q \in Q_f$ and $U \neq \emptyset$.*

Theorem 1 ([3]). *The S-FF-diagnosability problem for a given static mask $\mathcal{M}$ is* PSPACE*-complete.*

S-εFF-diagnosability and S-AFF-diagnosability : The following undecidability result is established in [3], motivating the focus on S-AFF-diagnosability.

Theorem 2 ([3]). *The S-εFF-diagnosability problem for a given static mask $\mathcal{M}$ is undecidable.*

While a seemingly harder notion, allowing for mistakes in the detection of faults, [3] showed that S-AFF-diagnosability was in fact simpler than S-FF-diagnosability and in particular could be tested directly on the pLTS instead of requiring an exponential determinisation.

Theorem 3 ([3]). *The S-AFF-diagnosability problem for a given static mask $\mathcal{M}$ is decidable in* PTIME.

2.5 Cost and latency of diagnosis

The ability to diagnose faults within a system highly depends on the mask. Intuitively, for a diagnosable system, the more we can observe, the easier and quicker the diagnosis, but the higher the cost to set and keep functioning the corresponding sensors. In this paper, we are interested in the balancing act between cost and speed of detection. More precisely, we separately study two distinct problems. First we focus on how to find a small mask that still allows to diagnose the system. Then we introduce and compute an efficiency measure for diagnosis.

When limiting oneself to static masks, reducing cost means reducing the number of sensors and thus the number of events that can be observed through the mask, which is given by the size of the mask. Moreover, there might not exist sensors for some events, in particular for the fault itself. We thus want to include a set of events that are forbidden to observe through the mask.

Definition 3. *Let $\mathcal{A}$ be a pLTS, N an integer and Σ_u a set of unobservable events. The N-S-FF-diagnosability (resp. N-S-AFF-diagnosability) decision problem consists in deciding whether there exists a static mask $\mathcal{M}$ such that $\mathcal{A}$ is S-FF-diagnosable (resp. S-AFF-diagnosable) under $\mathcal{M}$ and $|\mathcal{M}| \leq N$.*

The same problem can be defined lifting the static restriction on the mask. In this case, reducing the size of the mask does not automatically mean reducing the number of sensors that are needed, but how many need to be activated simultaneously.

Definition 4. *Let $\mathcal{A}$ be a pLTS, N an integer and Σ_u a set of unobservable events. The N-FF-diagnosability (resp. N-AFF-diagnosability) decision problem consists in deciding whether there exists a mask $\mathcal{M}$ such that $\mathcal{A}$ is FF-diagnosable (resp. AFF-diagnosable) under $\mathcal{M}$ and $|\mathcal{M}| \leq N$.*

We now move to defining a latency notion for diagnosis, focusing on FF-diagnosability. Given a pLTS $\mathcal{A}$ and a static mask $\mathcal{M}$, the latency of the diagnosis $L_{\mathcal{A},\mathcal{M}}$ is the expectancy on the number of steps between the occurrence of a fault and its detection. Formally, for $n \in \mathbb{N}$, we define the sets:

- $FD_{n,\mathcal{M}} = \{\rho \in F_{k,\mathcal{M}} \mid k \leq n \wedge \rho \notin \mathsf{FAmb}^0_{k,\mathcal{M}} \wedge \forall \rho' \prec \rho, \mathsf{CorP}(\mathcal{M}(\rho')) > 0\}$ for the set of signalling faulty runs with observable length at most n and such that the fault was detected thanks to the last observation;
- $S_{n,\mathcal{M}} = FD_{n,\mathcal{M}} \cup \{\rho \in F_{n,\mathcal{M}} \mid \rho \in \mathsf{FAmb}^0_{k,\mathcal{M}}\}$.

This later set contains every signalling faulty run of observable length at most n where either the fault was just detected, or the fault is yet to be identified. The latency $L_{\mathcal{A},\mathcal{M}}$ can then be defined as follows :

$$L_{\mathcal{A},\mathcal{M}} = \lim_{n \to +\infty} \sum_{\rho = \rho_1 \mathbf{f} \rho_2 \in S_{n,\mathcal{M}} \wedge \rho_1 \text{ correct}} \mathbb{P}(\rho)|\rho_2|$$

3 Optimising static diagnosis

3.1 Reducing costs

In this subsection, we are interested in reducing the costs of using static masks, and thus in the N-S-FF-diagnosability and N-S-AFF-diagnosability decision problems. We start with N-S-FF-diagnosability. The following result is mainly a consequence of [3] and in particular of Theorem 1.

Theorem 4. *The N-S-FF-diagnosability problem is* PSPACE*-complete.*

Proof. Let us start by showing that the N-S-FF-diagnosability problem is PSPACE-hard by reduction from the S-FF-diagnosability problem which is known to be PSPACE-complete (Theorem 1). Let $\mathcal{A}$ be a finite pLTS over an alphabet Σ and $\mathcal{M}$ be a static mask. Let $\Sigma_{\mathcal{M}}$ be the set observable through $\mathcal{M}$. We define the set $\Sigma_u = \Sigma \setminus \Sigma_{\mathcal{M}}$ and claim that $\mathcal{A}$ is S-FF-diagnosable for $\mathcal{M}$ iff $\mathcal{A}$ is $|\mathcal{M}|$-S-FF-diagnosable with unobservables Σ_u. Indeed, if $\mathcal{A}$ is S-FF-diagnosable for $\mathcal{M}$ then $\mathcal{M}$ is a witness of $\mathcal{A}$ being $|\mathcal{M}|$-S-FF-diagnosable with unobservable Σ_u. Conversely, if $\mathcal{A}$ is $|\mathcal{M}|$-S-FF-diagnosable, assume there is a static mask $\mathcal{M}'$ such that $\mathcal{A}$ is S-FF-diagnosable for $\mathcal{M}'$, $|\mathcal{M}'| \leq |\mathcal{M}|$ and $\Sigma_{\mathcal{M}'} \cap \Sigma_u = \emptyset$. Then, since $\mathcal{M}'$ observes a subset of the events observed through $\mathcal{M}$, it implies that if a run is ambiguous with respect to $\mathcal{M}$ then it is ambiguous with respect to $\mathcal{M}'$. Thus $\forall n$ $\mathsf{FAmb}^0_{n,\mathcal{M}} \subseteq \mathsf{FAmb}^0_{n,\mathcal{M}'}$, and since the probability of the righthand side of this inclusion converges towards zero, so does the leftand side. Our system $\mathcal{A}$ is thus S-FF-diagnosable for $\mathcal{M}$, which concludes the hardness proof.

For PSPACE easiness, let $\mathcal{A}$ be a finite pLTS over an alphabet Σ, N an integer and Σ_u a set of unobservable events. We non-deterministically guess a static mask $\mathcal{M}$ with observable set $\Sigma_{\mathcal{M}}$ such that $|\Sigma_{\mathcal{M}}| \leq N$, $\Sigma_{\mathcal{M}} \cap \Sigma_u = \emptyset$ and $\mathcal{A}$ is S-FF-diagnosable under $\mathcal{M}$. $\mathcal{M}$ is an object of size at most N, and verifying that the guessed mask satisfy the list of properties can be done in PSPACE thanks to Theorem 1. Thus, the N-S-FF-diagnosability problem is in NPSPACE. Moreover, from Savitch's theorem, we know that PSPACE =NPSPACE. Therefore the N-S-FF-diagnosability problem is in PSPACE. □

We now turn to N-S-AFF-diagnosability which once again is partially inspired from the results of [3] and in particular of Theorem 3, though N-S-AFF-diagnosability happens to be harder than S-AFF-diagnosability.

Theorem 5. *The N-S-AFF-diagnosability is* NP*-complete.*

Proof. Proving that N-S-AFF-diagnosability is in NP can be done similarly to the previous proof: we guess a mask of polynomial size with the right properties, those properties, including regularity, can be checked in PTIME thanks to Theorem 3, thus producing an NP algorithm for N-S-AFF-diagnosability.

For the hardness proof, we use a reduction from the 3-SAT problem. For a given formula ϕ, we create a $\mathcal{A}_\phi$ as in figure 4 that has one state for each variable (right) and one for each disjunction (left). Our formula is then satisfiable iff $\mathcal{A}_\phi$ is N-S-AFF-diagnosable with N equal to the number of variables. Indeed, finding a mask with our regularity condition implies that at least one event is observable in each state (except from q_0). Choosing those events on the right-hand states, we already have N observable events through our mask, which is the maximum and the system is thus N-S-AFF-diagosable iff those observable events are enough to ensure the regularity on the lefthand states and hence the satisfaction of the formula. Conversely if the formula is satisfiable, the valuation that satisfies it gives a set of N observable events (x_i or $\neg x_i$ for every i) that make the system regular and diagnosable (as in this system any observation allows to detect the fault). The full proof is given in Appendix A. □

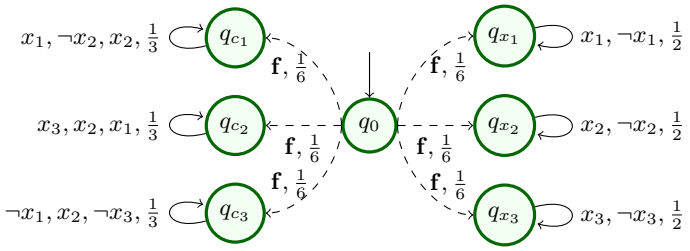

Fig. 4. Example of $\mathcal{A}_\phi$ for the formula $\phi = (x_1 \vee \neg x_2 \vee x_2) \wedge (x_3 \vee x_2 \vee x_1) \wedge (\neg x_1 \vee x_2 \vee \neg x_3)$.

Remark 1. A natural extension of the above cost consideration is to distribute weights to each observation, and to require the selected mask to have a total weight lower than a given threshold. In this framework, the unobservable events of Σ_u can be given infinite weight for instance. The proofs of this subsection would apply exactly as well in this case, thus we chose to abstain from the additional formalism.

3.2 Computing the latency of S-FF-diagnosis

In Subsection 2.5, given a pLTS $\mathcal{A}$ and a static mask $\mathcal{M}$, we defined the latency $L_{\mathcal{A},\mathcal{M}}$ of $\mathcal{M}$ for $\mathcal{A}$ in the context of S-FF-diagnosis. The goal of this section is to (i) show how to compute $L_{\mathcal{A},\mathcal{M}}$ and (ii) link the value of $L_{\mathcal{A},\mathcal{M}}$ and the S-FF-diagnosability of $\mathcal{A}$ under $\mathcal{M}$. We start by this latter point thanks to the following result:

Theorem 6. *Let $\mathcal{A}$ be a pLTS and let $\mathcal{M}$ be a mask. $L_{\mathcal{A},\mathcal{M}}$ is finite if and only if $\mathcal{A}$ is S-FF-diagnosable for $\mathcal{M}$.*

The proof of this theorem is decomposed into Lemma 1 and Lemma 2.

Lemma 1. *Let $\mathcal{A}$ be a pLTS and let $\mathcal{M}$ be a mask. If $L_{\mathcal{A},\mathcal{M}}$ is finite, then $\mathcal{A}$ is S-FF-diagnosable for $\mathcal{M}$.*

Proof. Let $\mathcal{A}$ be a pLTS and let $\mathcal{M}$ be a mask. Assume by contrapositive that $\mathcal{A}$ is not S-FF-diagnosable.

As $\lim_{n\to+\infty} \mathbb{P}(\mathsf{FAmb}^0_{n,\mathcal{M}}) \neq 0$ and as the measure of faulty infinite runs which prefix of length n is not faulty decreases to 0, there exists a finite faulty run ρ_f and $\varepsilon > 0$ such that for all $n \geq |\rho_f|$, denoting $F_{\rho_f,n} = Cyl(\rho_f) \cap \mathsf{FAmb}^0_{n,\mathcal{M}}$, we have $\mathbb{P}(F_{\rho_f,n}) \geq \varepsilon$. Given $H \geq 0$, for all $n \geq |\rho_f| + H/\varepsilon$, we define $L_{\mathcal{A},\mathcal{M},n} = \sum_{\rho=\rho_1\mathbf{f}\rho_2\in S_{n,\mathcal{M}}\wedge f\notin\rho_1} \mathbb{P}(\rho)|\rho_2|$, the latency of the mask restricted to runs of length n. We have:

$$\begin{aligned} L_{\mathcal{A},\mathcal{M},n} &= \sum_{\rho=\rho_1\mathbf{f}\rho_2\in S_{n,\mathcal{M}}\wedge f\notin\rho_1} \mathbb{P}(\rho)|\rho_2| \geq \sum_{\rho=\rho_f\rho'\in S_{n,\mathcal{M}}} \mathbb{P}(\rho)|\rho'| \\ &\geq \frac{H}{\varepsilon} \sum_{\rho=\rho_f\rho'\in\mathsf{FAmb}^0_{n,\mathcal{M}}} \mathbb{P}(\rho) = \frac{H}{\varepsilon} \sum_{\rho\in F_{\rho_f,n}} \mathbb{P}(\rho) \geq H \end{aligned}$$

The first inequality holds as 1) we restrict the set of paths and, 2) since the path is not split precisely after the first fault, it reduces both the probability of the prefix and the length of the suffix. As this holds for all $n \geq |\rho_f| + H/\varepsilon$, $L_{\mathcal{A},\mathcal{M}} \geq H$ for all $H \geq 0$, thus $L_{\mathcal{A},\mathcal{M}}$ is infinite. □

We now turn to the more involved converse implication.

Lemma 2. *Let $\mathcal{A}$ be a pLTS and let $\mathcal{M}$ be a mask. If $\mathcal{A}$ is S-FF-diagnosable for $\mathcal{M}$, then $L_{\mathcal{A},\mathcal{M}}$ is finite.*

Proof. Assume that $\mathcal{A}$ is S-FF-diagnosable for $\mathcal{M}$. We build the pLTS $\mathcal{A}_{FF}$ as defined in Subsection 2.4. When a faulty run reaches a state with $U = \emptyset$, the fault is immediately detected. We build the new pLTS $\mathcal{A}^Q_{FF}$ where every state of the form (q, U) with $q \in Q_f$ and $U = \emptyset$ is fused in a single state $q_{d,f}$ (redirecting every transition to (q, U) toward $q_{d,f}$). The latency $L_{\mathcal{A},\mathcal{M}}$ can thus be reinterpretated as the expectation on the number of steps between the occurrence of the fault and the arrival in $q_{d,f}$.

For every state q of $\mathcal{A}^Q_{FF}$, we define X_q to be the expectation on the number of steps for runs starting in q between reaching a faulty state, and reaching $q_{d,f}$. In particular, as noted above, $X_{q_0} = L_{\mathcal{A},\mathcal{M}}$, but also $X_{q_{d,f}} = 0$ and if q is a correct state belonging to a BSCC of the system, as no fault can be triggered in any run starting in q, $X_q = 0$. Moreover, for every state q of $\mathcal{A}^Q_{FF}$, we have the following linear equality:

$$X_q = u(q) + \sum_{(q,a,q')\in T} \mathbf{P}(q, a, q')X_{q'} \tag{1}$$

where $u(q)$ is equal to 1 if q is faulty and 0 otherwise.

We simplify the system of equations induced by (1) by removing every occurrence of the terms X_q where q is neither q_0 nor belong to a BSCC. This is achieved by replacing successively every occurrence of X_q by

$$\frac{u(q) + \sum_{(q,a,q') \in T \wedge q' \neq q} \mathbf{P}(q,a,q') X_{q'}}{1 - \sum_{(q,a,q) \in T} \mathbf{P}(q,a,q)}.$$

Note that this value is always defined as, q not belonging to a BSCC, it must have a transition that does not go to itself, thus $\sum_{(q,a,q) \in T} \mathbf{P}(q,a,q) < 1$. Once this transformation is achieved, we obtain an equation of the form

$$X_{q_0} = c + \sum_{q' \in U} p_{q'} X_{q'}$$

where $c \in \mathbb{R}$, $p_{q'} < 1$ and U is a set of states belonging to BSCC. As $\mathcal{A}$ (or equivalently $\mathcal{A}_{FF}$) is S-FF-diagnosable, by Proposition 1, $\mathcal{A}_{FF}$ has no BSCC containing a state (q, U) with $q \in Q_f$ and $U \neq \emptyset$. Every state $q' \in U$ is thus either correct (in which case $X_{q'} = 0$) or is $q_{d,f}$ (in which case $X_{q'} = 1$). Therefore we have $L_{\mathcal{A},\mathcal{M}} = X_{q_0} = c + p_{q_{d,f}} < \infty$. □

Note that the system of linear equations produced in the above proof is nonsingular (it has a single solution) assuming the system is S-FF-diagnosable. Indeed, the value of the variables associated to a state belonging to a BSCC is directly fixed, and the transformation used for computing q_0 can be done for any state outside a BSCC to obtain a single possible solution. Nonsingular linear systems have nice properties that can be used to compute efficiently their solution.

Theorem 7. *Let $\mathcal{A}$ be a pLTS and let $\mathcal{M}$ be a mask. Computing $L_{\mathcal{A},\mathcal{M}}$ can be done in* PSPACE.

Proof. During the proof of Lemma 2, we built a system of linear equations (see Equation 1) involving $L_{\mathcal{A},\mathcal{M}}$. The transformations realised on the system within the lemma would allow one to compute $L_{\mathcal{A},\mathcal{M}}$, though this naive algorithm lies in EXPTIME. However, the resolution of linear equations is a field that has been heavily studied, allowing for quicker algorithm.

In [6], the authors shows that the resolution of a nonsingular linear system lies in NC^2 with respect to the number of equations. This means that, with n equations, this problem can be computed in time $\mathcal{O}(log^2(n))$ using a polynomial number of processors in parallel. Moreover, from the results of [10], we know that applying an NC^2 algorithm on an exponential entry produces a PSPACE algorithm. Therefore, assuming our system of equations is nonsingular, and noting that the equations can individually be computed in PSPACE, the solution can be computed in PSPACE.

Therefore, in order to compute the latency of this system, the following algorithm lie in PSPACE: First test whether $\mathcal{A}$ is S-FF-diagnosable for $\mathcal{M}$, which can be done in PSPACE thanks to Theorem 1. If it is not, then $L_{\mathcal{A},\mathcal{M}} = \infty$

according to Theorem 6. If it is, then the system of equations induced by (1) is nonsingular and thus can be solved in PSPACE as discussed above, producing the value of $L_{\mathcal{A},\mathcal{M}}$ as a byproduct of the resolution. □

4 Dynamic diagnosis

In this section, we remove the previous static assumption on the mask and consider the N-FF-diagnosability and N-AFF-diagnosability decision problem. In both cases, we reduce our problem to diagnosis of controllable systems as studied in [2] and [9]. Let us first quickly recall the relevant definitions and results from these papers (note that the two papers use slightly different formalism, that are in fact equivalent for all intent and purpose. Hence here we state the results in a single framework).

Definition 5. *A Controllable weighted LTS (CwLTS) over alphabet Σ is a tuple $\mathbb{M} = (Q, q_0, \Sigma, T)$ where Q is a finite set of states with $q_0 \in Q$ being the initial state, the alphabet of event Σ is partitioned into observable Σ_o and unobservable Σ_u as well as in controllable Σ_c and uncontrollable Σ_e events (with $\Sigma_u \subseteq \Sigma_e$), and $T : S \times \Sigma \times S \to \mathbb{R}$ is the transition function labelling transitions with weights.*

A set $\Sigma^\bullet \subseteq \Sigma$ of *allowed events* is a set of events such that $\Sigma_e \subseteq \Sigma^\bullet$. Given two states $q, q' \in Q$, a set of allowed events $\Sigma^\bullet$ and $a \in \Sigma^\bullet$, we define the transition probability $p(q, \Sigma^\bullet, a)(q') = \frac{T(q,a,q')}{\sum_{q'' \in Q, b \in \Sigma^\bullet} T(q,b,q'')}$. Note that this definition requires that the set $\{q' \in Q \mid \exists a \in \Sigma^\bullet, T(q, a, q') \neq 0\}$ is not empty, and thus the choice of allowed events must ensure this property. This choice is fixed by a strategy.

Definition 6. *A* strategy *of CwLTS $\mathbb{M}$ is a mapping $\alpha : \Sigma_o^* \to \mathsf{Dist}(2^\Sigma)$ associating to any sequence of observable events a distribution on sets of allowed events.*

A run of a CwLTS $\mathbb{M}$ is a (finite or infinite) sequence $\rho = q_0 a_0 q_1 a_1 \dots$ where for all i, $T(q_i, a_i, q_{i+1}) > 0$. Given a strategy α, ρ is α-compatible if, denoting $\mathcal{M}_{\Sigma_o}$ the static mask projecting an event sequence onto the set Σ_o, for all i, there exists $\Sigma_i^\bullet \in 2^\Sigma$ such that $\alpha(\mathcal{M}_{\Sigma_o}(a_0 \dots a_{i-1}))(\Sigma_i^\bullet) > 0$ and $a_i \in \Sigma_i^\bullet$. We only consider strategies that produce sets of allowed events that ensure the transition probability is well defined, *i.e.* such that for all finite run $\rho = q_0 a_0 q_1 a_1 \dots q_n$, we have for all $\Sigma_i^\bullet$ such that $\alpha(\mathcal{M}_{\Sigma_o}(a_0 \dots a_{n-1}))(\Sigma_i^\bullet) > 0$, $\sum_{q', b \in \Sigma_i^\bullet} T(q_n, b, q') > 0$.

A strategy α on $\mathbb{M}$ defines an infinite pLTS $\mathbb{M}_\alpha$ where the set of states is the α-compatible runs, the initial state is q_0, for all run ρ ending in a state q, event a and state q', $(\rho, a, \rho a q') \in T$ and $\mathbf{P}(\rho, a, \rho a q') = \sum_{\Sigma_i^\bullet \in 2^\Sigma} \alpha(\mathcal{M}_{\Sigma_o}(\rho))(\Sigma_i^\bullet) p(\rho, \Sigma_i^\bullet, a)(q')$.

The FF-diagnosability (resp. AFF-diagnosability) problem over CwLTS consists in deciding whether there exists a strategy α such that $\mathbb{M}_\alpha$ is FF-diagnosable (resp. AFF-diagnosable) under the static mask associated to Σ_o.

Theorem 8 ([2, 9]). *The FF-diagnosability (resp. AFF-diagnosability) problem over CwLTS is EXPTIME-complete.*

The main cause of the exponential complexity in Theorem 8 is the exponential size of the belief (*aka*, the set of states that could be reached after a given observed sequence, which is represented by a deterministic automaton akin to *Obs* defined in Subsection 2.4), which both papers need to compute. The N-FF-diagnosability and N-AFF-diagnosability decision problems can be reduced to the FF-diagnosability and AFF-diagnosability problems over CwLTS. While our reduction produces an exponentially bigger CwLTS, the set of beliefs of this CwLTS remains "only" exponential, thus achieving the same complexity.

Theorem 9. *The N-FF-diagnosability (resp. N-AFF-diagnosability) problem is* EXPTIME*-complete.*

Proof. The main idea of this proof is to show the similarities between finding a mask ensuring diagnosability, and building a strategy that ensures diagnosability under a static fixed mask.

For instance, to show that the problems are in EXPTIME, from a given pLTS, we build a CwLTS containing many copies of the initial pLTS, each representing the behaviour of this pLTS under a choice of actions to observe through a mask. The strategy is then responsible for moving between copies, emulating choices of the selection function. From a strategy, we can thus build a mask.

Note that the strategies developed in [2, 9] relies on a finite amount of memory (and can thus be used in practice). Our approach ensures that the masks we build enjoy the same property.

The full proof is given in Appendix B. □

5 Conclusion

In this article, we considered two improvements over the classical notions of FF-diagnosis and AFF-diagnosis by aiming to improve the diagnoser of the system. First, we considered how to reduce the cost of diagnosing the system by limiting the number of sensors needed to be installed (in the case of a static mask), or when sensors need to be turned on (dynamic mask). Finding a static mask that requires a number of sensor below a given threshold was shown to be PSPACE-complete for FF-diagnosability and NP-complete for AFF-diagnosability while finding a dynamic mask turning at most a fixed number of sensors simultaneously was show to be EXPTIME-complete. Secondly, in order to compare the efficiency of two masks, so as to better judge whether using more sensors bring a real benefit, we introduced the notion of latency of a diagnoser. This notion represents an expectancy on the delay to detect a faulty behaviour of the system. We showed how to compute the latency of a static mask in PSPACE.

Our study of latency is however mostly limited to FF-diagnosis: while we can define a similar notion for AFF-diagnosis, the reliance on the correctness proportion of the run, and thus on the probabilities of a set of run, within that definition prohibits the current computing approach. If computing the exact value of this latency appears to be a very hard problem, in a future work we intend to produce efficient approximation algorithms through bounding the speed of decrease of the correctness proportion over faulty runs.

References

1. A. Bauer, M. Leucker, and C. Schallhart. Runtime verification for ltl and tltl. *ACM Trans. Softw. Eng. Methodol.*, 20(4), 2011.
2. N. Bertrand, E. Fabre, S. Haar, S. Haddad, and L. Hélouët. Active diagnosis for probabilistic systems. In *Proceedings of FoSSaCS'14*, volume 8412 of *LNCS*, pages 29–42. Springer, 2014.
3. N. Bertrand, S. Haddad, and E. Lefaucheux. A Tale of Two Diagnoses in Probabilistic Systems. *Information and Computation*, page 104441, 2019.
4. N. Bertrand, S. Haddad, and E. Lefaucheux. Diagnosis and degradation control for probabilistic systems. *Discret. Event Dyn. Syst.*, 30(4):695–723, 2020.
5. D. Berwanger and L. Doyen. On the power of imperfect information. In *Proceedings of FSTTCS'08*, volume 2 of *LIPIcs*, pages 73–82. Leibniz-Zentrum für Informatik, 2008.
6. A. Borodin, J. [von zur Gathen], and J. Hopcroft. Fast parallel matrix and gcd computations. *Information and Control*, 52(3):241 – 256, 1982.
7. F. Cassez, S. Tripakis, and K. Altisen. Synthesis of optimal-cost dynamic observers for fault diagnosis of discrete-event systems. In *First Joint IEEE/IFIP Symposium on Theoretical Aspects of Software Engineering (TASE '07)*, pages 316–325, 2007.
8. S. Jiang, Z. Huang, V. Chandra, and R. Kumar. A polynomial algorithm for testing diagnosability of discrete-event systems. *IEEE Transactions on Automatic Control*, 46(8):1318–1321, 2001.
9. Engel Lefaucheux. Accurate approximate diagnosis of (controllable) stochastic systems. In *Quantitative Evaluation of Systems (QEST)*, pages 413–434. Springer, 2021.
10. Walter L. Ruzzo. On uniform circuit complexity. *Journal of Computer and System Sciences*, 22(3):365–383, 1981.
11. M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis. Diagnosability of discrete-event systems. *IEEE Trans. Aut. Cont.*, 40(9):1555–1575, 1995.
12. D. Thorsley and D. Teneketzis. Diagnosability of stochastic discrete-event systems. *IEEE Transactions on Automatic Control*, 50(4):476–492, 2005.
13. D. Thorsley and D. Teneketzis. Diagnosability of stochastic discrete-event systems. *Transactions on Automatic Control*, 50(4):476–492, 2005.
14. Tae-Sic Y. and S. Lafortune. On the computational complexity of some problems arising in partially-observed discrete-event systems. In *Proceedings of the 2001 American Control Conference*, volume 1, pages 307–312 vol.1, 2001.

A *NP*-hardness of *N*-*S*-AFF-diagnosability.

Theorem 5. *The N-S-*AFF*-diagnosability is* NP*-complete.*

Proof. Here, we establish the hardness part of the proof which is obtained by reduction from the 3-SAT problem. It is interesting to note that AFF-diagnosability barely appears in the hardness proof below. In fact, what is shown here is that selecting a static mask of bounded size which satisfy the regularity condition on observations is already NP-hard.
As a reminder, given a finite formula ϕ of the form $\bigvee \phi_i$ where each ϕ_i is a conjunction of at most three literals (*aka* a variable or its negation), the 3-SAT problem asks whether there exists a valuation of the variables satisfying the formula. This problem is well known to be NP-complete.

Let ϕ be a formula in the form given above, with $c_1, \dots c_k$ clauses containing 3 literals: $\phi = \bigwedge_{i=1}^{k} c_i = \bigwedge_{i=1}^{k} (l_{i,1} \vee l_{i,2} \vee l_{i,3})$ and N the number of variables $(x_1, x_2, ..., x_N)$ appearing in ϕ (in other words each $l_{i,j}$ is either a x_m or a $\neg x_m$). We build the pLTS $\mathcal{A}_\phi = (Q, q_0, \Sigma, T, \mathbf{P})$ (see example in Figure 5) where:

- $Q = \{q_0, q_{x_1}, \dots, q_{x_N}, q_{c_1}, \dots, q_{c_k}\}$;
- $\Sigma = \{u, \mathbf{f}, x_1, \neg x_1, x_2, \neg x_2, ..., x_N, \neg x_N\}$;
- $T = \{(q_0, \mathbf{f}, q_{x_i}), (q_{x_i}, \neg x_i, q_{x_i}), (q_{x_i}, x_i, q_{x_i}) \mid i \in \{1, 2, \dots, N\}\} \cup \{(q_0, \mathbf{f}, q_{c_i}), (q_{c_i}, l_{i,j}, q_{c_i}) \mid i \in \{1, 2, \dots, k\} \wedge j \in \{1, 2, 3\}\}$;
- $\mathbf{P}$ is a uniform probability distribution.

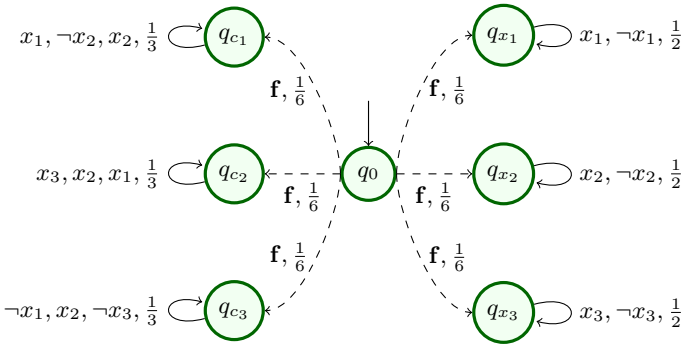

Fig. 5. Example of $\mathcal{A}_\phi$ for the formula $\phi = (x_1 \vee \neg x_2 \vee x_2) \wedge (x_3 \vee x_2 \vee x_1) \wedge (\neg x_1 \vee x_2 \vee \neg x_3)$.

Let us show that ϕ is satisfiable if and only if $\mathcal{A}_\phi$ is N-S-AFF-diagnosable with unobservable set $\Sigma_u = \{\mathbf{f}\}$.

Remark first that in $\mathcal{A}_\phi$, every run becomes faulty after exiting q_0, hence we have $\mathbb{P}(\mathsf{FAmb}^{\varepsilon}_{n,\mathcal{M}}) = 0$ for all $n \geq 1$ and all $\varepsilon > 0$ and the fault can be claimed exactly as soon as a single observation occurs. Proving that this system is N-S-AFF-diagnosable thus boils down to finding a regular static mask (any

such mask will do). Moreover, every state beside q_0 only contains self-loops. A mask $\mathcal{M}$ is regular for our system iff in every state beside q_0 there is a transition labelled by an event observable through $\mathcal{M}$.

$\Rightarrow$ Assume that ϕ is satisfiable, let ν be a valuation which satisfies ϕ (*i.e.* ν gives the values that the x_i must take to satisfy ϕ, $\nu(x_i) \in \{\top, \bot\}$). We build the static mask $\mathcal{M}$ such that

$$\Sigma_{\mathcal{M}} = \{x_i \mid \nu(x_i) = \top\} \cup \{\neg x_i \mid \nu(x_i) = \bot\}.$$

By definition, for every i, $\{x_i, \neg x_i\} \cap \Sigma_{\mathcal{M}} \neq \emptyset$ and we have an observable event from every state q_{x_i}. Moreover, as ϕ is satisfied by ν, every clause c_i must contain a literal accepted by ν, and thus appearing in $\Sigma_{\mathcal{M}}$. Thus, from the earlier remark, $\mathcal{A}_\phi$ is S-AFF-diagnosable under $\Sigma_{\mathcal{M}}$. As $\mathbf{f} \notin \Sigma_{\mathcal{M}}$ and $|\Sigma_{\mathcal{M}}| = N$, we produced a witness that $\mathcal{A}_\phi$ is N-S-AFF-diagnosable.

$\Leftarrow$ Assume now that there exists a static mask $\mathcal{M}$, such that $|\Sigma_{\mathcal{M}}| \leq N$ and $\mathbf{f} \notin \Sigma_{\mathcal{M}}$ and $\mathcal{A}_\phi$ is S-AFF-diagnosable under $\mathcal{M}$. From our earlier remark, we know that for each state q_{x_i}, one of the exiting transition is labelled by an event in $\Sigma_{\mathcal{M}}$. As there are N such states and the events exiting them are disjoint, there is in fact for each q_{x_i} exactly one exiting transition labelled in $\Sigma_{\mathcal{M}}$. Denote $z_i \in \{x_i, \neg x_i\}$ this event. We define the valuation ν such that $\nu(x_i) = \top$ iff $z_i = x_i$. Since our mask is regular, at least one transition from each q_{c_j} is labelled with an observable event, and since with the z_i we have already N events in $\mathcal{M}$, each clause c_j must contain a z_i and thus be satisfied by ν. □

B Dynamic diagnosis is *EXPTIME*-complete

Theorem 9. *The N-FF-diagnosability (resp. N-AFF-diagnosability) problem is* EXPTIME*-complete.*

We separate this proof in two propositions.

Proposition 2. *The N-FF-diagnosability (resp. N-AFF-diagnosability) problem is in* EXPTIME*.*

Proof. Let $\mathcal{A} = (Q, q_0, \Sigma, T, \mathbf{P})$ be a pLTS, $N \in \mathbb{N}$ and $\Sigma_u \subseteq \Sigma$ be the unobservable events. We denote $\Sigma_o = \Sigma \setminus \Sigma_u$. We will build a CwLTS $\mathbb{M}$ such that there exists a strategy diagnosing $\mathbb{M}$ iff there exists a mask of size $\leq N$ allowing to diagnose $\mathcal{A}$.

The idea of the construction is the following: we build multiple copies of the original $\mathcal{A}$, one for each possible choice of set of observable events by the mask, as well as a "control" copy (where the states are associated to the letter c). In the control copy, the available actions are the sets of observable events that can be selected by the mask. We start in the control copy, from where the strategy thus selects one potential mask choice $\Sigma^\bullet$, moving to the copy of the system associated to $\Sigma^\bullet$. Within this copy, the behaviour of the system mimics what happens within

the initial pLTS when $\Sigma^\bullet$ is selected by the mask. Once a transition is observed, the system goes back to the control copy to select a new set of observations. This CwLTS is exponential in size as we copied it once for each potential selection of the mask, but the belief is not doubly exponential as an observer always knows in which copy of the system the run is. More precisely, each location of the belief is associated to a set of states, all of which belongs to the same copy. Thus there are at most an exponential number of beliefs associated to each copy.

Formally, we first define the set of potential mask choices of size $\leq N$: $H_\Sigma = \{\Sigma^\bullet \mid \Sigma^\bullet \subseteq \Sigma \setminus \Sigma_u \wedge |\Sigma^\bullet| \leq N\}$. Then, we define the CwLTS $\mathbb{M} = (Q', (q_0, c), \Sigma', T')$ such that

- $Q' = \{(q, Z) \mid q \in Q \wedge Z \in H_\Sigma \cup \{c\}\}$,
- $\Sigma' = \Sigma \cup H_\Sigma \cup \{u\}$ where $\Sigma_u \cup \{u\}$ is unobservable (u is an additional unobservable event added to represent actions that could be observable, but are currently hidden by the mask) and H_Σ is controllable (*i.e.* $\Sigma'_e = \Sigma' \setminus H_\Sigma$),
- the transition function T' is defined as follows
  - $T'((q, c), \Sigma^\bullet, (q, \Sigma^\bullet)) = 1$ for $\Sigma^\bullet \in H_\Sigma$. This represents the selection of a set of actions to observe, going from the controlled copy to the copy associated to the selected set,
  - $T'((q_1, \Sigma^\bullet), u, (q_2, \Sigma^\bullet)) = \mathbf{P}(q_1, b, q_2)$ for $\Sigma^\bullet \in H_\Sigma$ and $b \in \Sigma \setminus (\Sigma_u \cup \Sigma^\bullet)$ such that $(q_1, b, q_2) \in T$. This represents taking a transition that could have been observable given another mask, but the current choice of observables, $\Sigma^\bullet$, makes it unobservable. As no observation is received, the selection of the mask does not change.
  - $T'((q_1, \Sigma^\bullet), a, (q_2, \Sigma^\bullet)) = \mathbf{P}(q_1, a, q_2)$ for $\Sigma^\bullet \in H_\Sigma$ and $a \in \Sigma_u$ such that $(q_1, a, q_2) \in T$. This represent taking an action which is inherently unobservable, such as the fault.
  - $T'((q_1, \Sigma^\bullet), a, (q_2, c)) = \mathbf{P}(q_1, a, q_2)$ for $\Sigma^\bullet \in H_\Sigma$ and $a \in \Sigma^\bullet$ such that $(q_1, a, q_2) \in T$. This represents taking an action that is currently observed by the mask, and thus going back to the control copy of the system to make a new selection.
  - if none of the above, T' return 0.

Let us show that there exists a mask $\mathcal{M}$ of size $\leq N$ with Σ_u unobservable such that $\mathcal{A}$ is FF-diagnosable (resp. AFF-diagnosable) under $\mathcal{M}$ iff there exists a strategy α for $\mathbb{M}$ such that $\mathbb{M}_\alpha$ is FF-diagnosable (resp. AFF-diagnosable) under the static mask $\mathcal{M}_{\Sigma'_o}$ associated to $\Sigma'_o = (\Sigma \setminus \Sigma_u) \cup H_\Sigma$.

Let $\mathcal{M}$ be a mask of size $\leq N$ with Σ_u unobservable, produced by the selection function $\mathcal{S}_\mathcal{M}$, such that $\mathcal{A}$ is FF-diagnosable (resp. AFF-diagnosable) under $\mathcal{M}$. We define the strategy α for $\mathbb{M}$ which, when a run ends in the control copy, emulates the selection function $\mathcal{S}_\mathcal{M}$ in order to decide which actions, beside the uncontrollable actions (noting that actions from Σ'_e are not available in the control copy) are allowed, and blocks every non-controllable action for runs outside the control copy (noting that only actions from Σ'_e are available anyway). Formally, for any sequence $\sigma = \Sigma_1^\bullet a_1 \dots \Sigma_n^\bullet a_n \in {\Sigma'_o}^*$, we select for any set $E \subseteq \Sigma'_o$, $\alpha(\sigma)(\Sigma'_e \cup \{E\}) = \mathcal{S}_\mathcal{M}(\sigma)(E)$ and for any event sequence

σ' ending with an element of H_Σ, $\alpha(\sigma')$ is a Dirac distribution on the set of uncontrollable events Σ'_e. Then $\mathbb{M}_\alpha$ is FF-diagnosable (resp. AFF-diagnosable) under the static mask $\mathcal{M}_o$ associated to Σ_o. Indeed, let ρ be a finite run of $\mathbb{M}_\alpha$, let $\sigma_\rho = \Sigma_1^\bullet a_{i_1^1} \dots a_{i_{m_1}^1} \dots \Sigma_n^\bullet a_{i_1^n} \dots a_{i_{m_n}^n}$ be its event sequence and σ_ρ^o be the observed sequence of ρ under $\mathcal{M}_o$. Let ρ' be the run of $\mathcal{A}$ following the same states and transitions (ignoring the information on the copies contained in the second component of the states of ρ). Observe that its event sequence $\sigma_{\rho'}$ can be obtained from σ_ρ by removing the elements of H_Σ and replacing occurrences of the unobservable u by the action of Σ it replaced in $\mathbb{M}$. Denote p the probability that $\mathcal{S}_\mathcal{M}$ selected the sets $\Sigma_1^\bullet, \dots, \Sigma_n^\bullet$: $p = \prod_{i=1}^n \mathcal{S}_\mathcal{M}(\Sigma_1^\bullet a_{i_{m_1}^1} \dots \Sigma_{i-1}^\bullet a_{i_{m_{i-1}}^{i-1}})(\Sigma_i^\bullet)$. We have that $\mathbb{P}(\rho) = p\mathbb{P}(\rho')$ (since ρ and ρ' follow the same transitions, except for the choices of masks which total probability is p). Moreover, $\mathsf{CorP}_{\mathcal{M}_{\Sigma_o}}(\sigma_\rho) = \mathsf{CorP}_\mathcal{M}(\sigma_{\rho'})$ as the probability of the choice of the mask appears in every run sharing the same observation, and thus does not affect the correctness proportion. Hence, noting that σ_ρ^o is the only observed sequence of ρ under $\mathcal{M}_o$ and that ρ' will produce the observed sequence σ_ρ^o under $\mathcal{M}$ with probability p, we have that $(\rho, 1, \sigma_\rho^o) \in \mathsf{FAmb}^\varepsilon_{2n,\mathcal{M}_o}$ iff $(\rho', p, \sigma_\rho^o) \in \mathsf{FAmb}^\varepsilon_{n,\mathcal{M}}$ Observing that, since selecting a mask in $\mathbb{M}_\alpha$ does not modify which run is faulty, nor their correctness proportion, then $\mathbb{P}(\mathsf{FAmb}^\varepsilon_{2n,\mathcal{M}_o}) = \mathbb{P}(\mathsf{FAmb}^\varepsilon_{2n+1,\mathcal{M}_o})$, we thus have that, for all $\varepsilon \geq 0$ εFF-diagnosability of $\mathcal{A}$ under $\mathcal{M}$ is equivalent to εFF-diagnosability of $\mathbb{M}_\alpha$ under $\mathcal{M}_o$.

The opposite direction of the proof is obtained through the same construction. Given a strategy α for $\mathbb{M}$ such that $\mathbb{M}_\alpha$ is FF-diagnosable (resp. AFF-diagnosable) under the static mask associated to Σ_o, we define the mask $\mathcal{M}$ and its selection function $\mathcal{S}_\mathcal{M}$ such that for any event sequence $\sigma = \Sigma_1^\bullet a_1 \dots \Sigma_n^\bullet a_n \in \Sigma_o^*$ and for any set $E \subseteq \Sigma_o$, $\mathcal{S}_\mathcal{M}(\sigma)(E)$ is the probability that the transition labelled by E is taken in $\mathbb{M}_\alpha$ after observing σ. As this is similar to the previous direction, the relations between runs of $\mathcal{A}$ and $\mathbb{M}_\alpha$ hold again, hence $\mathcal{A}$ is FF-diagnosable (resp. AFF-diagnosable) under $\mathcal{M}$. □

Proposition 3. *The N-FF-diagnosability (resp. N-AFF-diagnosability) problem is* EXPTIME*-hard.*

Proof. This hardness is established by reduction from the resolution of safety games with imperfect information which are known to be EXPTIME-complete [5].

A safety game with imperfect information $G = (L, l_0, \Sigma, \Delta, \mathcal{O}, F, obs)$ is a two player game where L is a set of locations with l_0 the initial location, Σ is an alphabet, $\Delta \subseteq L \times \Sigma \times L$ is the transition relation such that for every $l \in L$ and $a \in \Sigma$ there exists at least one location $l' \in L$ such that $(l, a, l') \in \Delta$, $\mathcal{O}$ is an alphabet of observation, with $F \subseteq \mathcal{O}$ being final observations, and $obs : L \to \mathcal{O}$ is an observation function mapping every location to an observation.

These games are realised between two players, A and B. Starting in location l_0, A first selects a letter $a_0 \in \Sigma$, and then B chooses a location l_1 such that $(l_0, a_0, l_1) \in \Delta$. A is not directly informed that the game reached l_1, he instead only observes $o_1 = obs(l_1)$. The next rounds are played similarly from the new location. Player A wins if for all i, $o_i \notin F$.

Given a safety game $G = (L, l_0, \Sigma, \Delta, \mathcal{O}, F, obs)$ we define the pLTS $\mathcal{A}_G = (Q, q_0, \Sigma', T, \mathbf{P})$, represented in Figure 6, where:

- $Q = \{q_0, q_1, q_2, f_\sharp\} \uplus L \uplus (L \times \Sigma) \uplus \{q_o \mid o \in \mathcal{O}\}$;
- $\Sigma' = \{\mathbf{f}, u, \sharp\} \uplus \mathcal{O} \uplus \Sigma$;
- $(q_0, u, l_0) \in T$, $(q_0, u, q_1) \in T$, for $o \in \mathcal{O}, (q_1, u, q_o) \in T$, $(q_o, o, q_o) \in T$, $(q_1, o, q_1), (q_1, u, q_2) \in T$ and $(q_2, \sharp, q_2) \in T$. For $(l, a, l') \in \Delta$, $(l, a, (l', a)) \in T$ and $((l', a), obs(l'), l') \in T$. For $l_f \in obs^{-1}(F)$, $(l_f, \mathbf{f}, f_\sharp) \in T$, $(f_\sharp, \sharp, f_\sharp) \in T$.
- $\mathbf{P}$ is the uniform probability distribution, assigning the same probability to every transition exiting a state.

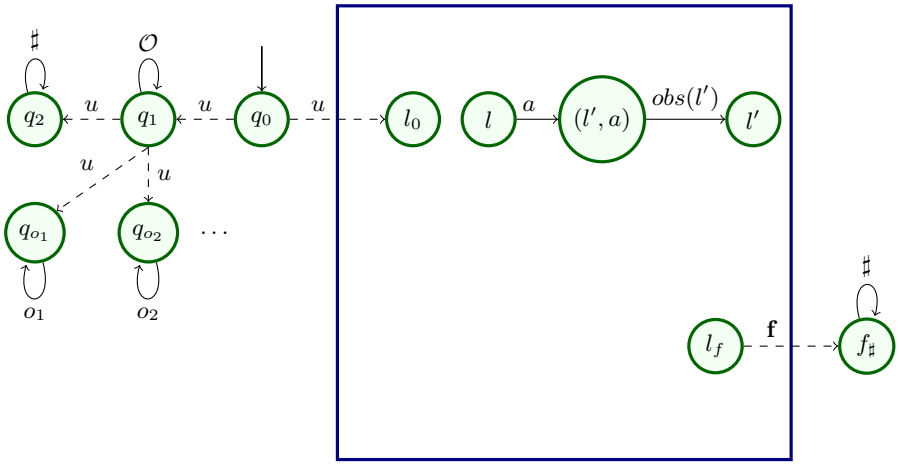

Fig. 6. Illustration of the transformation of G

Set $N = |\Sigma| + |\mathcal{O}|$ and $\Sigma_u = \{\mathbf{f}, u\}$. Let us show that there exists a winning strategy for player A in G if and only if $\mathcal{A}_G$ is N-FF-diagnosable if and only if $\mathcal{A}_G$ is N-AFF-diagnosable.

The first point to observe in this construction is that, as long as no element of $\Sigma \cup \{\sharp\}$ has been observed, the system may be in the states q_2 or q_o for any $o \in \mathcal{O}$. As a consequence, due to the regularity condition, as long as one may be in those states, any selection function of a mask of $\mathcal{A}_G$ can only attribute probabilities to set containing every element of $\mathcal{O} \cup \{\sharp\}$. As $N = |\Sigma| + |\mathcal{O}|$, this means that, during that period, the mask can observe all but one element of Σ. The main idea of our construction is to represent the selection of an action in the safety game, by the choice of not observing that very action in the pLTS.

Another thing to note is that given a mask $\mathcal{M}$, a faulty run ρ which looped at least once on $f_\sharp$, and σ an observed sequence of ρ (note that due to the loop of $f_\sharp$ and the regularity of the mask, σ ends with $\sharp$) if σ contains an element of Σ, then $\mathsf{CorP}(\sigma) = 0$, otherwise $\mathsf{CorP}(\sigma) > 0$. In other words, a mask ensures

FF-diagnosability and AFF-diagnosability iff one observes at least one element of Σ along every faulty run.

Assume that player A has a winning strategy $\alpha : \mathcal{O}^* \to \Sigma$ in G. Then we define the mask $\mathcal{M}$ such that after an observed sequence w, if $w \in \mathcal{O}^*$, $\mathcal{M}$ observes the elements of $\mathcal{O} \cup \{\sharp\} \cup (\Sigma \setminus \alpha(w))$[3], otherwise, an element of $\Sigma \cup \{\sharp\}$ having been observed, asking the mask to observe $\{\sharp\}$ is enough to know when (and if) $f_\sharp$ is reached, and to satisfy the regularity condition. Let $\rho = q_0 u l_0 a_0 (l_1, a_0) obs(l_1) l_1 \dots l_n \mathbf{f} f_\sharp (\sharp f_\sharp)^k$ be a faulty signalling run of $\mathcal{A}_G$, the path $p = l_0 a_0 \dots a_{n-1} l_n$ represent a possible play in the game G where player A selected $a_0 \dots a_{n-1}$ and observed $obs(l_1) \dots obs(l_n)$. As $obs(l_n) \in F$ and α is a winning strategy, this play does not follow α. Thus by definition of $\mathcal{M}$, every observed sequence of ρ contains an element of Σ. As this holds for every faulty signalling run, from our earlier remark, we deduce that $\mathcal{M}$ ensures FF-diagnosability and AFF-diagnosability of $\mathcal{A}_G$.

Conversely, given a mask $\mathcal{M}$ of $\mathcal{A}_G$ ensuring FF-diagnosability and AFF-diagnosability of $\mathcal{A}_G$, we define the strategy α as follows:

given a sequence $w \in \mathcal{O}^*$, as pointed out before, $\mathcal{M}$ cannot observe every element of Σ after w, pick $a \in \Sigma$ that $\mathcal{M}$ does not observe after w, we set $\alpha(w) = a$. Given a play $p = l_0 a_0 l_1 \dots l_n$ in G that is compatible with the strategy α we build the run $\rho = q_0 u l_0 a_0 (l_1, a_0) obs(l_1) l_1 \dots l_n$. By definition of α, there exists an observed sequence w of ρ under $\mathcal{M}$ such that $w \in \mathcal{O}^*$. Due to our previous remarks, as $\mathcal{M}$ ensures FF-diagnosability and AFF-diagnosability of $\mathcal{A}_G$, this means that $obs(l_n) \notin F$. As this holds for every play compatible with α, α is a winning strategy for A. □

[3] *i.e.* if $w = a_1 \dots a_m$, for any sequence of sets $\Sigma_1^\bullet, \dots \Sigma_m^\bullet \in 2^{\Sigma'}$, $\mathcal{S}_\mathcal{M}(\Sigma_1^\bullet a_1 \dots \Sigma_m^\bullet a_m)(\mathcal{O} \cup \{\sharp\} \cup (\Sigma \setminus \alpha(w))) = 1$.