Accurate Approximate Diagnosis of (Controllable) Stochastic Systems

Engel Lefaucheux[0000−0003−0875−300X]

Max Planck Institute for Software Systems, Saarland Informatics Campus, Saarbrücken, Germany elefauch@mpi-sws.org

Abstract. Diagnosis of partially observable stochastic systems prone to faults was introduced in the late nineties. Diagnosability may be specified in different ways: exact diagnosability requires that almost surely a fault is detected and that no fault is erroneously claimed; approximate diagnosability tolerates a small error probability when claiming a fault; last, accurate approximate diagnosability guarantees that the error probability can be chosen arbitrarily small. While all three notions were studied for passive systems such as observable Markov chains, only the exact notion was considered for systems equipped with a controller. As the approximate notion of diagnosability was shown to be undecidable in passive systems, in this article, we complete the picture by deciding the accurate approximate diagnosability for controllable observable Markov chains. More precisely, we show how to adapt the accurate approximate notion to the active setting and establish EXPTIME-completeness of the associated decision problem. We also show how to measure the set of faulty paths that are detected under the accurate approximate notion in the passive setting.

Keywords: Stochastic systems, Partial observation, Control, Diagnosis

1 Introduction

Diagnosis and diagnosability. There has been an increasing use of software systems for critical operations. When designing such systems, one aims at eliminating faults that could trigger unwanted behaviours. However, for embedded systems interacting with an unpredictable environment, the absence of faults is not a reasonable hypothesis. Thus diagnosis, whose goal consists to detect faults from the observations of the runs of the system, is a crucial task. One of the approaches frequently used to analyse *diagnosability* consists in modelling the system by a transition system whose states (depending on the internal part of the system) are unobservable and events may, depending on their nature, be observable or not. One of the proposed approaches consists in modelling these systems by partially observable labelled transition systems (poLTS) [24]. In such a framework, diagnosability requires that the occurrence of unobservable faults can be deduced accurately from the previous and subsequent observable events. In other words, defining the *disclosure set* of a system as the set of faulty paths

of the system that can be detected, a system is diagnosable if every faulty path belongs to the disclosure set. Diagnosability for poLTS was shown to be decidable in PTIME [18]. Diagnosis has since been extended to numerous models (Petri nets [12], pushdown systems [20], etc.) and settings (centralized, decentralized, distributed), and have had an impact on important application areas, *e.g.* for telecommunication network failure diagnosis.

Diagnosability for stochastic passive systems. In transition systems, the unpredictable behaviours of the environment are modelled by a nondeterministic choice between the possible events from the current state. However, in order to quantify the risks induced by the faults of the systems, the designer often substitutes the nondeterministic choice by a random choice or equivalently by a weighted one. Then the model becomes a discrete time observable Markov chain (oMC) in the *passive* case (*i.e.* without control). In these models, one can define a probability measure over infinite runs. In that context, the accuracy required to claim a path is faulty can be relaxed. There are three natural variants: (1) exact disclosure, which, as in the non-stochastic case, requires that every path sharing the given observation sequence is faulty in order to claim a fault occurred, (2) ε-disclosure for $\varepsilon > 0$ which tolerates small errors, allowing to claim the failure of a path if the conditional probability that the path is faulty exceeds $1 - \varepsilon$, and (3) Accurate Approximate disclosure (AA-disclosure) which is satisfied when the accuracy of the guess can be chosen arbitrarily high. Diagnosability with exact disclosure has been studied extensively for oMC [25, 6, 8]. In particular, various exact notions of diagnosability have been shown to be PSPACE-complete for oMC. Due to the quantitative requirement, diagnosability with ε-disclosure was shown to be undecidable while diagnosability with AA-disclosure was surprisingly shown to be in PTIME [7].

Active diagnosability. Embedded systems are often equipped with one (or more) controller(s) in order to maintain some functionalities of the system in case of a pathological behaviour of the environment. It is thus tempting to add to the controller a diagnosis task. Formally some of the observable events are controllable and considering its current observation, the controller chooses which subset of actions should be allowed to make the system diagnosable. As such, a controller only has access to the observations produced by the system to make his choice. This represents the idea that the control is realised by the same entity as the diagnosis. A system is said *actively diagnosable* if there exists a controller ensuring diagnosability [23, 26, 13, 14, 17]. In [17], the authors designed an exponential time algorithm and proved the optimality of this complexity. In stochastic systems, diagnosability has only been considered with exact disclosure and has been proven EXPTIME-complete [5].

Contribution In this paper, we study diagnosability in stochastic systems under AA-disclosure.

- we introduce an alternative definition of AA-disclosure and establish its equivalence with the notion introduced in [25] (Proposition 1)

– we show that measuring the set of AA-disclosing paths for oMC is PSPACE-complete (Theorem 3);
– we establish that diagnosability with AA-disclosure for controllable oMC is EXPTIME-complete (Theorem 4).

For space concerns, some technical proofs are deferred to the appendix.

2 Diagnosis of Markov Chains

2.1 Observable Markov Chains

For a finite alphabet Σ, we denote by Σ^* (resp. Σ^ω) the set of finite (resp. infinite) words over Σ, $\Sigma^\infty = \Sigma^* \cup \Sigma^\omega$ and ε the empty word. The length of a word w is denoted by $|w| \in \mathbb{N} \cup \{\infty\}$ and for $n \in \mathbb{N}$, Σ^n is the set of words of length n. A word $u \in \Sigma^*$ is a prefix of $v \in \Sigma^\infty$, written $u \leq v$, if $v = uw$ for some $w \in \Sigma^\infty$. The prefix is strict if $w \neq \varepsilon$. For $n \leq |w|$, we write $w_{\downarrow n}$ for the prefix of length n of w. Given a countable set S, a distribution on S is a mapping $\mu : S \to [0,1]$ such that $\sum_{s \in S} \mu(s) = 1$. The support of μ is $\mathsf{Supp}(\mu) = \{s \in S \mid \mu(s) > 0\}$. If $\mathsf{Supp}(\mu) = \{s\}$ is a single element, μ is a Dirac distribution on s written $\mathbf{1}_s$. We denote by $\mathsf{Dist}(S)$ the set of distributions on S.

For the purpose of partially observable problems, the model must be equipped with an *observation function* describing what an external observer can see. The observation function can be obtained via a labelling of states or transitions, both options being known to be equivalent. We thus define observable Markov chains (see Figure 1).

Definition 1 (Observable Markov chains). *An observable Markov chain (oMC) over alphabet Σ is a tuple $\mathcal{M} = (S, p, \mathsf{O})$ where S is a countable set of states, $p : S \to \mathsf{Dist}(S)$ is the transition function, and $\mathsf{O} : S \to \Sigma$ is the observation function.*

We write $p(s'|s)$ instead of $p(s)(s')$ to emphasise the probability of going to state s' conditioned by being in state s. Given a distribution $\mu_0 \in \mathsf{Dist}(S)$, we denote by $\mathcal{M}(\mu_0)$ the oMC with initial distribution μ_0. For decidability and complexity results, we assume that all probabilities occurring in the model (transition probabilities and initial distribution) are rationals. A (finite or infinite) path of $\mathcal{M}(\mu_0)$ is a sequence of states $\rho = s_0 s_1 \ldots \in S^\infty$ such that $\mu_0(s_0) > 0$ and for each $i \geq 0$, $p(s_{i+1}|s_i) > 0$. For a finite path, $\rho = s_0 s_1 \ldots s_n$, we call n its length and denote its ending state by $\mathsf{last}(\rho) = s_n$. A finite path ρ_1 prefixes a finite or infinite path ρ if there exists a path ρ_2 such that $\rho = \rho_1 \rho_2$. The set $\mathsf{Cyl}(\rho)$ represents the cylinder of infinite paths prefixed by ρ. We denote by $\mathsf{Path}(\mathcal{M}(\mu_0))$ (resp. $\mathsf{fPath}(\mathcal{M}(\mu_0))$) the set of infinite (finite) paths of $\mathcal{M}(\mu_0)$. The *observation sequence* of the path $\rho = s_0 s_1 \ldots$ is the word $\mathsf{O}(\rho) = \mathsf{O}(s_0)\mathsf{O}(s_1)... \in \Sigma^\infty$. For a set R of paths, $\mathsf{O}(R) = \{\mathsf{O}(\rho) \mid \rho \in R\}$ and for a set W of observation sequences, $\mathsf{O}^{-1}(W) = \{\rho \in \mathsf{Path}(\mathcal{M}(\mu_0)) \cup \mathsf{fPath}(\mathcal{M}(\mu_0)) \mid \mathsf{O}(\rho) \in W\}$.

Forgetting the labels, an oMC with an initial distribution μ_0 becomes a discrete time Markov chain (DTMC). In a DTMC, the set of infinite paths is the

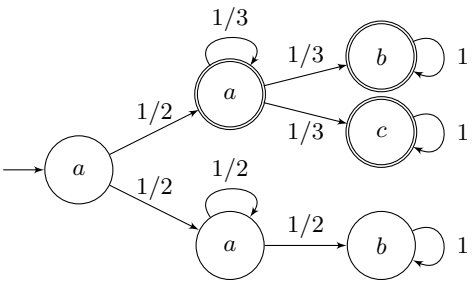

Fig. 1. An observable Markov chain. The arrow entering the leftmost state means that the initial disribution is a Dirac on this state. Faulty states are circled twice.

support of a probability measure extended from the probabilities of the cylinders by the Caratheodory's extension theorem:

$$\mathbf{P}_{\mathcal{M}(\mu_0)}(\mathsf{Cyl}(s_0 s_1 \dots s_n)) = \mu_0(s_0)p(s_1|s_0)\dots p(s_n|s_{n-1}) \ .$$

When $\mathcal{M}(\mu_0)$ is clear from context, we will sometimes omit the subscript, and write $\mathbf{P}$ for $\mathbf{P}_{\mathcal{M}(\mu_0)}$. Let $\rho \in \mathsf{fPath}(\mathcal{M})$, $w \in \Sigma^*$ and $E \subseteq \Sigma^\omega$, with a small abuse of notation we write $\mathbf{P}(\rho)$ for $\mathbf{P}(\mathsf{Cyl}(\rho))$, $\mathbf{P}(w)$ instead of $\mathbf{P}(\cup_{\rho \in \mathsf{O}^{-1}(w)} \mathsf{Cyl}(\rho))$ and $\mathbf{P}(E)$ instead of $\mathbf{P}(\{\rho \in \mathsf{Path}(\mathcal{M}(\mu_0)) \mid \rho \in \mathsf{O}^{-1}(E)\})$.

2.2 Faulty Paths and Notions of Disclosure

In this paper we are interested in the study of diagnosis, a problem in which one wants to detect whether the current path correspond to a faulty behaviour of the system. We focus on the particular case where the faulty behavior of the system is given by a subset of states $\mathsf{S}^\mathsf{F} \subseteq S$, called *faulty states*, of the model: a (finite or infinite) path $s_0 s_1 \dots$ is *faulty* if $s_i \in \mathsf{S}^\mathsf{F}$ for some i. The set of finite (resp. infinite) faulty paths is denoted F (resp. F_∞). A path that is not faulty is called *correct*. Remark that without loss of generality, we can assume that the set of faulty states is absorbing, *i.e.* if a path visits S^F, it forever remains in S^F.

In non-stochastic systems, a faulty path discloses its failure if it does not share its observation sequence with any correct path, *i.e.* given a path $\rho \in \mathsf{S}^\mathsf{F}$, it discloses its failure iff $\mathsf{O}^{-1}(\mathsf{O}(\rho)) \subseteq \mathsf{S}^\mathsf{F}$. When adding probabilities, one could keep the same definition of disclosure, this is what we call exact disclosure. Denoting $Disc^{exact}$ the set of infinite disclosing faulty paths, the exact diagnosability problem for oMC asks whether $\mathbb{P}(Disc^{exact}) = \mathbb{P}(\mathsf{F}_\infty)$. This problem is known to be PSPACE-complete for oMC [8].

However, one could also weaken the requirement by allowing potential false claims. In this case, a faulty path is disclosing if, based on its observation, the likelihood of the path to be faulty is high. To formalise this likelihood, we define the failure proportion as the conditional probability that a path is

faulty, given its observation sequence. Formally, given an oMC $\mathcal{M} = (S, p, \mathsf{O})$, an initial distribution μ_0, $\mathsf{S}^{\mathsf{F}} \subseteq S$ and an observation sequence $w \in \Sigma^*$, the failure proportion associated with the observation sequence w is:

$$\mathsf{Fprop}_{\mathcal{M}(\mu_0)}(w) = \frac{\mathbf{P}(\{\rho \in \mathsf{O}^{-1}(w) \mid \rho \in \mathsf{F}\})}{\mathbf{P}(w)}.$$

This proportion is undefined if $\mathbf{P}(w) = 0$.

Example 1. Consider the oMC of Figure 1 and the observation sequences a^k, $a^k b^n$ and $a^k c^m$. The observation sequence a^k, for $k > 1$, can be produced by a correct path with probability $1/2^{k-1}$ and by a faulty path with probability $1/2 \times 1/3^{k-2}$. Therefore, $\mathsf{Fprop}_{\mathcal{M}(\mu_0)}(a^k) = \frac{1/3^{k-2}}{1/2^{k-2}+1/3^{k-2}}$ which converges to 0 when k grows to infinity. The failure proportion of the observation $a^k b^n$ with $k > 1$ and $n \geq 1$ is similarly $\mathsf{Fprop}_{\mathcal{M}(\mu_0)}(a^k b^n) = \frac{1/3^{k-1}}{1/2^{k-1}+1/3^{k-1}}$ which remains constant for extensions of $a^k b^n$ as it does not depend on n. Finally, if $m \geq 1$, $\mathsf{Fprop}_{\mathcal{M}(\mu_0)}(a^k c^m) = 1$ as no correct path can produce a 'c'.

Let $\mathcal{M} = (S, p, \mathsf{O})$ be an oMC, μ_0 be an initial distribution and $\mathsf{S}^{\mathsf{F}} \subseteq S$. Given $\varepsilon > 0$ representing the confidence threshold expected for the detection, we can define the approximate notion of disclosure: an observation sequence $w \in \Sigma^*$ is called ε-disclosing if $\mathsf{Fprop}_{\mathcal{M}(\mu_0)}(w) > 1 - \varepsilon$. Moreover, it is ε-min-disclosing if it is ε-disclosing and no strict prefix of w is ε-disclosing. Writing $D^{\varepsilon}_{\min}$ for the set of ε-min-disclosing observation sequences, the ε-disclosure is defined by

$$Disc^{\varepsilon}(\mathcal{M}(\mu_0)) = \mathbf{P}(\{\rho \in \mathsf{F} \mid \exists \rho' \leq \rho, \mathsf{O}(\rho') \in D^{\varepsilon}_{\min}\}).$$

$Disc^{\varepsilon}$ is thus the probability that a path of the oMC will be faulty and disclose its failure with sufficiently low doubt. The ε-diagnosability problem consists then in deciding whether $Disc^{\varepsilon}(\mathcal{M}(\mu_0)) = \mathbf{P}(\mathsf{F}_{\infty})$. Unfortunately, it is known that this problem is undecidable for $\varepsilon \neq 0$:

Theorem 1 ([7]). *Given $0 < \varepsilon < 1$, the ε-diagnosability problem is undecidable for oMCs.*

In order to regain decidability one can consider a slightly more qualitative notion of approximate information control, that is called accurate approximate. Instead of deeming the failure of a path to be revealed when the proportion of faulty paths goes above a given threshold, an infinite observation sequence is AA-*disclosing* if this proportion converges toward 1. In other words, when observing an AA-disclosing observation sequence, one can get an arbitrarily high confidence that the path is faulty. Formally, an observation sequence $w \in \Sigma^{\omega}$ is AA-disclosing if $\lim_{n\to\infty} \mathsf{Fprop}_{\mathcal{M}(\mu_0)}(w_{\downarrow n}) = 1$. Writing D^{AA} for the set of AA-disclosing observation sequences, the AA-disclosure is defined by

$$Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0)) = \mathbf{P}(\{\rho \in \mathsf{F} \mid \mathsf{O}(\rho) \in D^{\mathsf{AA}}\})$$

As before, the AA-diagnosability problem consists in deciding if $Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0)) = \mathbf{P}(\mathsf{F}_\infty)$. When an oMC is not AA-diagnosable, it is interesting to measure the probability of undetected faulty paths. This motivates the AA-*disclosure problem* which consists in, given $\lambda \in [0;1]$ and $\bowtie \in \{>, \geq\}$, deciding whether $Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0)) \bowtie \lambda$.

AA-diagnosability was in fact initially defined in [25] slightly differently: a system was then called AA-diagnosable if it was ε-diagnosable for all $\varepsilon > 0$. However, the two definitions are in fact equivalent for oMC.

Proposition 1. *An oMC is AA-diagnosable iff it is ε-diagnosable for all $\varepsilon > 0$.*

Proof. Let $\mathcal{M}$ be an oMC and μ_0 an initial distribution.

Suppose that $\mathcal{M}(\mu_0)$ is AA-diagnosable. By definition, given an AA-disclosing observation sequence w, for all $\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that $w_{\downarrow n}$ is ε-disclosing. Therefore for all $\varepsilon > 0$, $Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0)) \leq Disc^{\varepsilon}(\mathcal{M}(\mu_0))$. Moreover, as $\mathcal{M}$ is AA-diagnosable, $Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0)) = \mathbf{P}(\mathsf{F})$. Thus, $Disc^{\varepsilon}(\mathcal{M}(\mu_0)) \geq \mathbf{P}(\mathsf{F})$. Finally, as only faulty paths are disclosing, for all $\varepsilon > 0$ $Disc^{\varepsilon}(\mathcal{M}(\mu_0)) \leq \mathbf{P}(\mathsf{F})$. Thus $Disc^{\varepsilon}(\mathcal{M}(\mu_0)) = \mathbf{P}(\mathsf{F})$ and $\mathcal{M}(\mu_0)$ is ε-diagnosable.

Conversely, suppose that $\mathcal{M}(\mu_0)$ is not AA-diagnosable. Let us consider the set of infinite words $D = \cap_{\varepsilon>0} D^{\varepsilon}_{min}\Sigma^{\omega} \setminus D^{\mathsf{AA}}$. Let us show that $\mathbf{P}(D) = 0$. Let $w \in D$, we have (1) for all $\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that $\mathsf{Fprop}(w_{\downarrow n}) > 1 - \varepsilon$ and (2) $(\mathsf{Fprop}(w_{\downarrow n}))_{n\in\mathbb{N}}$ does not converge toward 1. Given $\varepsilon > 0$, due to (1) we have

$$\begin{aligned}\mathbf{P}(\{\rho \in \mathsf{O}^{-1}(D) \setminus \mathsf{F}\}) &< \sum_{w\in D^{\varepsilon}_{min}} \mathbf{P}(\{\rho \in \mathsf{O}^{-1}(w) \setminus \mathsf{F}\}) \\ &< \sum_{w\in D^{\varepsilon}_{min}} \mathbf{P}(\{\rho \in \mathsf{O}^{-1}(w) \cap \mathsf{F}\})\frac{\varepsilon}{1-\varepsilon} \\ &< \frac{\varepsilon}{1-\varepsilon}.\end{aligned}$$

As this holds for all $\varepsilon > 0$, $\mathbf{P}(\{\rho \in \mathsf{O}^{-1}(D) \setminus \mathsf{F}\}) = 0$. Moreover, due to (2), there exists $\varepsilon > 0$ such that for infinitely many $n \in \mathbb{N}$ we have $\mathsf{Fprop}(w_{\downarrow n}) < 1 - \varepsilon$. For all $k \in \mathbb{N}$, we denote by E_k the set of prefixes w of words of D such that $\mathsf{Fprop}_{\mathcal{M}(\mu_0)}(w) < 1 - \varepsilon$ for the k'th time. We then have for all k:

$$\begin{aligned}\mathbf{P}(\{\rho \in \mathsf{O}^{-1}(E_k) \setminus \mathsf{F}\}) &= \sum_{w\in E_k} \mathbf{P}(\{\rho \in \mathsf{O}^{-1}(w) \setminus \mathsf{F}\}) \\ &> \sum_{w\in E_k} \mathbf{P}(\{\rho \in \mathsf{O}^{-1}(w) \cap \mathsf{F}\})\frac{\varepsilon}{1-\varepsilon} \\ &> \frac{\varepsilon}{1-\varepsilon}\mathbf{P}(\{\rho \in \mathsf{O}^{-1}(D) \cap \mathsf{F}\})\end{aligned}$$

As $(\mathbf{P}(\{\rho \in \mathsf{O}^{-1}(E_k) \setminus \mathsf{F}\}))_{k\in\mathbb{N}}$ converges toward $\mathbf{P}(\{\rho \in \mathsf{O}^{-1}(D) \setminus \mathsf{F}\})$ which is equal to 0, this implies that $\mathbf{P}(\{\rho \in \mathsf{O}^{-1}(D) \cap \mathsf{F}\}) = 0$ and thus that $\mathbf{P}(D) = 0$. As a consequence, $\lim_{\varepsilon\to 0} \mathbf{P}(D^{\varepsilon}_{min}) = \mathbf{P}(D^{\mathsf{AA}})$. As $\mathcal{M}(\mu_0)$ is not AA-diagnosable by assumption, there thus exists $\varepsilon > 0$ such that $\mathcal{M}(\mu_0)$ is not ε-diagnosable. □

The alternative definition of AA-diagnosability was introduced for two reasons. First, through Proposition 1 it helps build a better understanding of this notion, often misunderstood (see for instance the uniform / non-uniform discussion on AA-diagnosability in [8]). Second, it helps clarify and analyse the notion in a controllable framework: as we will see later, we aim to build a single strategy achieving arbitrary high confidence, not a family of strategies each achieving ε-diagnosability for increasingly small ε.

With the accurate approximate approach to diagnosability, one regains decidability. Indeed, the AA-diagnosability problem for finite oMC was shown to be in PTIME in [7]. This result relies on the notion of distance between two oMC introduced in [16] and defined in the following way: the distance between two oMC $\mathcal{M}_1$ and $\mathcal{M}_2$ with initial distribution μ_1 and μ_2 is

$$d(\mathcal{M}_1(\mu_1), \mathcal{M}_2(\mu_2)) = \max_{E \subseteq \Sigma^\omega} |\mathbf{P}_{\mathcal{M}_1(\mu_1)}(E) - \mathbf{P}_{\mathcal{M}_2(\mu_2)}(E)|^1.$$

The authors of [16] show how to decide in PTIME if the distance between two oMC is 1 thanks to the following characterisation.

Proposition 2 ([16]). *Given two oMC $\mathcal{M}_1$ and $\mathcal{M}_2$ and two initial distributions μ_1 and μ_2, $d(\mathcal{M}_1(\mu_1), \mathcal{M}_2(\mu_2)) < 1$ iff there exists $w \in \Sigma^*$ and two distributions π_1 and π_2 such that, denoting for $i \in \{1,2\}, \mu_i^w(s) = \mathbf{P}_{\mathcal{M}_i(\mu_i)}(\{\rho s \in S^* \mid \mathsf{O}(\rho s) = w\})$, we have, $\mathsf{Supp}(\pi_i) \subseteq \mathsf{Supp}(\mu_i^w)$ and $d(\mathcal{M}_1(\pi_1), \mathcal{M}_2(\pi_2)) = 0$ (*i.e. $\forall w' \in \Sigma^*, \mathbf{P}_{\mathcal{M}_1(\pi_1)}(w') = \mathbf{P}_{\mathcal{M}_2(\pi_2)}(w')$*).*

Finally, the link between the distance 1 of two oMC and AA-diagnosability was established in [7], giving the PTIME algorithm:

Theorem 2 ([7]). *Let $\mathcal{M}$ be a finite oMC and μ_0 be an initial distribution. $\mathcal{M}(\mu_0)$ is not AA-diagnosable iff there exist two states $s \in \mathsf{S^F}$ and $s' \in S \setminus \mathsf{S^F}$ with s' belonging to a bottom strongly connected component (BSCC)[2] of $\mathcal{M}$ and there exist two finite paths ρ and ρ' of $\mathsf{fPath}(\mathcal{M}(\mu_0))$ such that $\mathsf{last}(\rho) = s$, $\mathsf{last}(\rho') = s'$, $\mathsf{O}(\rho) = \mathsf{O}(\rho')$ and $d(\mathcal{M}(\mathbf{1}_s), \mathcal{M}(\mathbf{1}_{s'})) < 1$.*

From the above theorem, one deduces that AA-diagnosability can be tested by checking the distance 1 of an at most quadratic number of oMC, leading to the PTIME algorithm. The results of this paper also study AA-diagnosability by establishing links to the distance 1 problem. These results however go farther than the characterisation of Theorem 2. In particular, when studying controllable systems, we will need to consider infinite oMC. To that end, we can already note that, speaking of the sufficiency condition only, a more general result was in fact proven in [7]:

Proposition 3 ([7]). *Let $\mathcal{M}$ be an oMC, μ_0 be an initial distribution, two states $s \in \mathsf{S^F}$ and $s' \in S \setminus \mathsf{S^F}$ with s' such that no faulty state can be reached from s' and two finite paths ρ and ρ' of $\mathsf{fPath}(\mathcal{M}(\mu_0))$ such that $\mathsf{last}(\rho) = s$, $\mathsf{last}(\rho') = s'$,*

[1] Note that the absolute values are technically not necessary as $\mathbf{P}_{\mathcal{M}_1(\mu_1)}(E) = 1 - \mathbf{P}_{\mathcal{M}_1(\mu_1)}(\Sigma^\omega \setminus E)$

[2] A BSCC is a strongly connected component that cannot be escaped from.

$\mathsf{O}(\rho) = \mathsf{O}(\rho')$. *Then* $\mathcal{M}(\mu_0)$ *is* AA*-diagnosable implies that* $d(\mathcal{M}(\mathbf{1}_q), \mathcal{M}(\mathbf{1}_{q'})) = 1$.

While AA-diagnosability can be decided in polynomial time, the AA-disclosure problem is a bit more complicated. This is not surprising as AA-diagnosability consists in testing whether $Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0))$ is equal to $\mathbf{P}(\mathsf{F}_\infty)$ (the latter being easy to compute as it is solely a reachability property) while the AA-disclosure requires to measure precisely $Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0))$.

Theorem 3. *The* AA*-disclosure problem for finite oMC is* PSPACE*-complete.*

Proof (Sketch of proof). In order to solve the AA-disclosure problem in PSPACE. We first build an exponential size oMC which contains additional information compared to the original one. Then we show that there are two kinds of BSCC in this new oMC: the ones that are reached by paths that almost surely have an AA-disclosing observation sequence, and the ones that are reached by paths that almost surely do not correspond to AA-disclosing observation sequences. We then use the existing results for the AA-diagnosability problem to determine the status of each BSCC. Finally, computing the AA-disclosure of the oMC is equivalent to computing the probability to reach the "AA-disclosing" BSCC, which can be done in NC in the size of the oMC, thus giving an overall PSPACE algorithm.

The hardness is obtained by reduction from the universality problem for non-deterministic finite automaton (NFA), which is known to be PSPACE-complete [19]. □

3 Diagnosis of Controllable Systems

3.1 Controllable Observable Markov chains

An extension of the oMC formalism allowing us to express control requires us to fix at least two features of this formalism: the nature of the control and the distribution of probabilities of the controlled system. *Controllable weighted Observable Markov chains* (CoMC) are an extension of oMC equivalent to the model of controllable weighted labelled transition systems (CLTS) which were introduced for diagnosis in [5] (the difference between the two models lies in whether the states or the transitions are labelled by an observation). CoMC can also be compared to partially observable Markov decision processes (POMDP): the two classes of models are as expressive, but CoMC can be exponentially more succinct.

In order to specify the control in a CoMC, a subset of observable events is considered as controllable. The control strategy forbids a subset of controllable events depending on the sequence of observations it has received so far. The transitions of the system are no longer labelled by (rational) probabilities but rather by (integer) weights which represent their relative probabilities. Given a state and a set of allowed events, in order to obtain a probability distribution on the allowed transitions, the weights of the outgoing transitions labelled by uncontrollable or allowed controllable actions are normalised. Provided that the control strategy does not create any deadlock, the controlled CoMC is an oMC.

Definition 2 (CoMC). *A Controllable weighted Observable Markov chains (CoMC) over alphabet Σ is a tuple $\mathbb{M} = (S, T, \mathsf{O})$ where S is a finite set of states, $T : S \times S \to \mathbb{N}$ is the transition function labelling transitions with integer weights and $\mathsf{O} : S \to \Sigma$ is the observation function.*

The alphabet is partitioned into controllable and uncontrollable events $\Sigma = \Sigma_c \uplus \Sigma_e$. A set $\Sigma_s \subseteq \Sigma$ of *allowed events* in a state $s \in S$ is a set of observations such that $\Sigma_e \subseteq \Sigma_s$ and $\{s' \in S \mid T(s,s') > 0 \wedge \mathsf{O}(s') \in \Sigma_s\} \neq \emptyset$. Given a state s and a set of allowed events Σ_s, we define the transition probability $p(s, \Sigma_s)$ such that for all s' with $\mathsf{O}(s') \in \Sigma_s$, $p(s, \Sigma_s)(s') = \frac{T(s,s')}{\sum_{s'', \mathsf{O}(s'') \in \Sigma_s} T(s,s'')}$. As before, we write $p(s'|s, \Sigma_s)$ instead of $p(s, \Sigma_s)(s')$. Given an initial distribution μ_0, an infinite path of a CoMC $\mathbb{M}(\mu_0)$ is a sequence $\rho = s_0\Sigma_0 s_1 \Sigma_1 \dots$ where $\mu_0(s_0) > 0$ and $p(s_{i+1}|s_i, \Sigma_i) > 0$, for $s_i \in S$ and Σ_i is a set of allowed events in s_i, for all $i \geq 0$. As for oMC, we define finite paths, and we use similar notations for the various sets of paths. A sequence of observations and sets of allowed events $b \in (\Sigma \times 2^\Sigma)^*\Sigma$ is called a *knowledge sequence*. The knowledge sequence of a path of a CoMC $\rho = s_0\Sigma_0 s_1 \Sigma_1 \dots s_i$ is $K(\rho) = \mathsf{O}(s_0)\Sigma_0\mathsf{O}(s_1)\Sigma_1 \dots \mathsf{O}(s_i)$.

The nondeterministic choice of the set of allowed events is resolved by strategies.

Definition 3 (Strategy for CoMC). *A* strategy *of CoMC $\mathbb{M}$ with initial distribution μ_0 is a mapping $\sigma : (\Sigma \times 2^\Sigma)^*\Sigma \to \mathsf{Dist}(2^\Sigma)$ associating to any knowledge sequence a distribution on sets of events.*

We will only consider here strategies that do not generate a deadlock, i.e. strategies σ such that for all state s reached after a knowledge b, $\sigma(b)$ is a distribution on sets of allowed events for s. Given a strategy σ, a path $\rho = s_0\Sigma_0 s_1 \Sigma_1 \dots$ of $\mathbb{M}(\mu_0)$ is *σ-compatible* if for all i, $\Sigma_i \in \mathsf{Supp}(\sigma(K(s_0\Sigma_0 s_1 \Sigma_1 \dots s_i))$. A strategy σ is *deterministic* if $\sigma(b)$ is a Dirac distribution for each knowledge sequence b. In this case, we denote by $\sigma(b)$ the set of allowed actions $\Sigma_a \in 2^\Sigma$ such that $\sigma(b) = \mathbf{1}_{\Sigma_a}$. Let b be a knowledge sequence. We define $B_{\mathbb{M}(\mu_0)}(b)$ the *belief* about states corresponding to b as follows:

$$B_{\mathbb{M}(\mu_0)}(b) = \{s \mid \exists \rho \in \mathsf{fPath}(\mathbb{M}(\mu_0)),\ K(\rho) = b \wedge s = \mathsf{last}(\rho)\}$$

A strategy σ is *belief-based* if for all b, $\sigma(b)$ only depends on its belief $B_{\mathbb{M}(\mu_0)}(b)$ (*i.e.* given two knowledge sequence b and b' if $B_{\mathbb{M}(\mu_0)}(b) = B_{\mathbb{M}(\mu_0)}(b')$ then $\sigma(b) = \sigma(b')$). For belief-based strategies, we will sometimes write $\sigma(B)$ for the choice of the strategy made for knowledge sequences producing the belief B.

As for oMC, the failure of a path is defined by the reachability of a set $\mathsf{S}^\mathsf{F} \in S$ of faulty states of the CoMC and we assume again that this set is absorbing.

A strategy σ on $\mathbb{M}(\mu_0)$ defines an infinite oMC $\mathbb{M}_\sigma(\mu_0)$ where the set of states is the finite σ-compatible paths, the observation function associates $\Sigma_{n-1}\mathsf{O}(s_n)$ with the state corresponding to the finite path $\rho = s_0\Sigma_0 \dots \Sigma_{n-1}s_n$ (Σ_{n-1} being omitted if $n = 0$) and the transition function p_σ is defined for ρ a σ-compatible path and $\rho' = \rho\Sigma_a s'$ by $p_\sigma(\rho'|\rho) = \sigma(K(\rho))(\Sigma_a)p(s'|s, \Sigma_a)$. We denote by $\mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}$ the probability measure induced by this oMC. When the strategy possesses some

good regularity properties, this oMC is equivalent to a finite one (*i.e.* there is a one-to-one correspondence between the paths of each oMC, it preserves the knowledge sequence and the probability. The two oMC have therefore the same disclosure properties). For instance given a deterministic belief based strategy σ, one can define the oMC $\mathbb{M}'_\sigma$ with set of states $S \times 2^\Sigma \times 2^S$, observation $\mathsf{O}'_\sigma(s, \Sigma^\bullet, B) = (\mathsf{O}(s), \Sigma^\bullet)$, initial distribution $\mu_0^\sigma(s, \emptyset, \mathsf{Supp}(\mu_0) \cap \mathsf{O}^{-1}(\mathsf{O}(s))) = \mu_0(s)$ and transition function $p'_\sigma((s_1, \Sigma_1, B_1) \mid (s_2, \Sigma_2, B_2)) = p(s_1 \mid s_2, \Sigma_2)$ if $\sigma(B_1) = \Sigma_2$ and $B_2 = B_{\mathbb{M}(\mu_1)}(\mathsf{O}(s_2))$ for μ_1 a distribution of support B_1, $p'_\sigma((s_1, \Sigma_1, B_1) \mid (s_2, \Sigma_2, B_2)) = 0$ otherwise. The oMC $\mathbb{M}'_\sigma$ is exponential in the size of $\mathbb{M}$ and is equivalent to $\mathbb{M}_\sigma$. When considering belief-based strategies, we will call $\mathbb{M}_\sigma$ the finite equivalent oMC.

Writing $\mathcal{V}_{\mathbb{M}_\sigma(\mu_0)}$ for the set of infinite paths corresponding to AA-disclosing observation sequences in $\mathbb{M}_\sigma(\mu_0)$, we have $Disc^{\mathsf{AA}}(\mathbb{M}_\sigma(\mu_0)) = \mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(\mathcal{V}_{\mathbb{M}_\sigma(\mu_0)})$. The control of the system is assumed to support the diagnosis. Therefore, the AA-*diagnosability problem* for CoMC consists in, given a CoMC $\mathbb{M}$ and an initial distribution μ_0, deciding whether there exists a strategy σ such that $\mathbb{M}_\sigma(\mu_0)$ is AA-diagnosable (aka, such that $\mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(\mathcal{V}_{\mathbb{M}_\sigma(\mu_0)}) = \mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(\mathsf{F}_\infty)$).

Example 2. Consider the CoMC on the left of Figure 2. Without any control (*i.e.* with a strategy permanently allowing every event), one obtains the oMC of Figure 1, which is not AA-diagnosable. However, assuming 'b' is a controllable event, the strategy that always forbids it induces the oMC on the right of Figure 2 which is AA-diagnosable: every faulty path almost surely contains a 'c' that can not be generated by a correct path. This oMC is in fact exactly diagnosable as once a 'c' occurs the failure proportion becomes equal to 1.

Remark that an observation sequence of the oMC induced by a CoMC and a strategy contains both the observation of the state of the CoMC and the choices of allowed events done by the strategy. The observation sequence of a path in the induced oMC is therefore equal to the knowledge sequence of the corresponding path in the CoMC and as such, we will only speak of observation sequences in the following. This choice of observation was done to express that the choices made by the strategy are known to the observer. An important consequence of this decision is that the strategy does not modify which observation sequences are AA-disclosing.

Lemma 1. *Given $\mathbb{M}$ a CoMC, μ_0 an initial distribution, $\mathsf{S}^\mathsf{F} \subseteq S$, σ, σ' two strategies and w an observation sequence produced by at least one path of $\mathbb{M}_\sigma(\mu_0)$ and at least one path of $\mathbb{M}_{\sigma'}(\mu_0)$, then $\mathsf{Fprop}_{\mathbb{M}_{\sigma'}(\mu_0)}(w) = \mathsf{Fprop}_{\mathbb{M}_\sigma(\mu_0)}(w)$.*

Proof. Let $\mathbb{M}$ be a CoMC, μ_0 be an initial distribution, σ be a strategy and $w = o_0\Sigma_0 \dots \Sigma_{n-1}o_n$ be an observation sequence produced by at least one path of $\mathbb{M}_\sigma(\mu_0)$. By definition of w, $\mathsf{Fprop}_{\mathbb{M}_\sigma(\mu_0)}(w)$ is defined and in particular

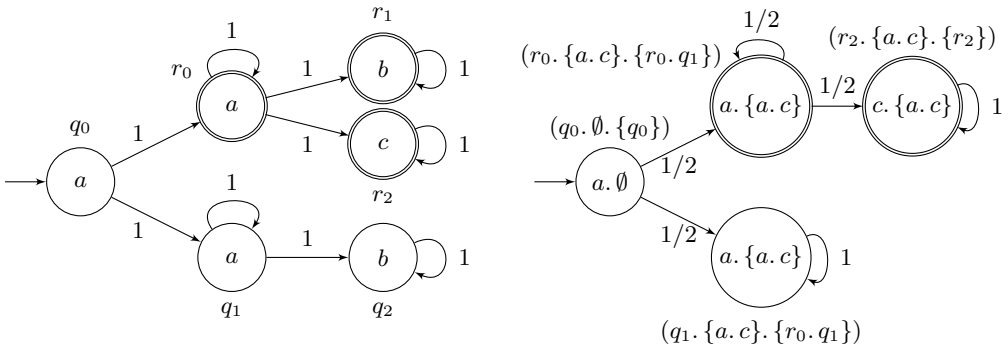

Fig. 2. A CoMC (left) and the finite oMC (right) induced by this CoMC and the strategy that always allow $\{a, c\}$. The observation of a state of the oMC is the pair composed of the observation of the associated state in the CoMC and of the set of allowed events that lead to it. Its name is the triple composed of the associated state in the CoMC, the set of allowed event leading to it and the belief about states that hold in the CoMC when entering this state. The probability in the induced oMC to loop on $(r_0.\{a.c\}.\{r_0.q_1\})$ is obtained by dividing the weight $T(r_0, r_0)$ by the weights $T(r_0, r_0)$ and $T(r_0, r_2)$, thus 1/2. The weight $T(r_0, r_1)$ is ignored as b is forbidden.

$\prod_{i=0}^{n-1} \sigma(\mathsf{O}(w_{\downarrow 2i+1}))(\Sigma_i) \neq 0$. We have

$$\begin{aligned}
\mathsf{Fprop}_{\mathbb{M}_\sigma(\mu_0)}(w) &= \frac{\mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(\{\rho \in \mathsf{O}^{-1}(w) \mid \rho \in \mathsf{F}\})}{\mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(w)} \\
&= \frac{\sum_{\rho \in \mathsf{O}^{-1}(w) | \rho \in \mathsf{F}} \mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(\rho)}{\sum_{\rho \in \mathsf{O}^{-1}(w)} \mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(\rho)} \\
&= \frac{\sum_{\rho = s_0\Sigma_0 \dots s_n \in \mathsf{O}^{-1}(w) | \rho \in \mathsf{F}} \prod_{i=0}^{n-1} \sigma(\mathsf{O}(w_{\downarrow 2i+1}))(\Sigma_i) p(s_{i+1} \mid s_i, \Sigma_i)}{\sum_{\rho = s_0\Sigma_0 \dots s_n \in \mathsf{O}^{-1}(w)} \prod_{i=0}^{n-1} \sigma(\mathsf{O}(w_{\downarrow 2i+1}))(\Sigma_i) p(s_{i+1} \mid s_i, \Sigma_i)} \\
&= \frac{\sum_{\rho = s_0\Sigma_0 \dots s_n \in \mathsf{O}^{-1}(w) | \rho \in \mathsf{F}} \prod_{i=0}^{n-1} p(s_{i+1} \mid s_i, \Sigma_i)}{\sum_{\rho = s_0\Sigma_0 \dots s_n \in \mathsf{O}^{-1}(w)} \prod_{i=0}^{n-1} p(s_{i+1} \mid s_i, \Sigma_i)}
\end{aligned}$$

which is independent of σ, therefore for any strategy σ' such that at least one path of $\mathbb{M}_{\sigma'}(\mu_0)$ produces w, $\mathsf{Fprop}_{\mathbb{M}_{\sigma'}(\mu_0)}(w) = \mathsf{Fprop}_{\mathbb{M}_\sigma(\mu_0)}(w)$. □

3.2 Solving AA-diagnosability for CoMCs

While accurate approximate diagnosability is simpler than exact diagnosability for oMC (PTIME vs PSPACE)[7, 6], for CoMCs this difference disappears and both problems are EXPTIME-complete. The EXPTIME-completeness of exact

diagnosis for CoMC was established in [5]. We will devote this section to the proof of the following theorem:

Theorem 4. *The* AA*-diagnosability problem over CoMCs is* EXPTIME*-complete.*

First, the hardness is obtained directly by applying the proof of Proposition 3 of [5]. This proof relies on a reduction from safety games with imperfect information [9] to establish EXPTIME-hardness of an exact notion of diagnosability. Their proof also applies to AA-diagnosability as, in the system they build, a faulty path is exactly diagnosable iff it is AA-diagnosable.

Proposition 4. *The* AA*-diagnosability problem over CoMCs is* EXPTIME*-hard.*

The most important step to solve AA-diagnosability for CoMC is to develop a good understanding on the strategies optimising AA-disclosure. For starters, with a straightforward adaptation of a proof of [15], we show that one can consider deterministic strategies only.

Lemma 2. *Given $\mathbb{M}$ a CoMC, μ_0 an initial distribution, $\mathsf{S}^{\mathsf{F}} \subseteq S$ and σ a strategy, there exists a deterministic strategy σ' such that $Disc^{\mathsf{AA}}(\mathbb{M}_\sigma(\mu_0)) = \mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(\mathsf{F}_\infty)$ implies $Disc^{\mathsf{AA}}(\mathbb{M}_{\sigma'}(\mu_0)) = \mathbf{P}_{\mathbb{M}_{\sigma'}(\mu_0)}(\mathsf{F}_\infty)$.*

Proof. In the proof of Lemma 1 of [15], the authors show that a randomised 'observation based' strategy can be seen as an average over a family of deterministic 'observation based' strategies[3]. A consequence of their equation (2) in our framework is the following: given a strategy σ, for every set of path E, there exists a deterministic strategy σ_{det} such that (a) $\mathsf{Path}(\mathbb{M}_{\sigma_{det}}(\mu_0)) \subseteq \mathsf{Path}(\mathbb{M}_\sigma(\mu_0))$ and (b) $\mathbf{P}_{\mathbb{M}_{\sigma_{det}}(\mu_0)}(E) \geq \mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(E)$. Using this result with the appropriate set E we will show that if $\mathbb{M}_\sigma(\mu_0)$ is AA-diagnosable then $\mathbb{M}_{\sigma_{det}}(\mu_0)$ is AA-diagnosable.

We define $E_\sigma = \mathcal{V}_{\mathbb{M}_\sigma(\mu_0)} \cup (\mathsf{Path}(\mathbb{M}_\sigma(\mu_0)) \setminus \mathsf{F}_\infty)$ which are the set of infinite σ-compatible paths that are either correct or AA-disclosing. Let σ_{det} be the strategy obtained by applying the result of [15] on the set E_σ. Suppose $\mathbb{M}_\sigma(\mu_0)$ is AA-diagnosable. By definition, this is equivalent to $\mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(E_\sigma) = 1$. Due to (b), this implies that $\mathbf{P}_{\mathbb{M}_{\sigma_{det}}(\mu_0)}(E_\sigma) = 1$ too. Moreover $\mathcal{V}_{\mathbb{M}_{\sigma_{det}}(\mu_0)} = \mathcal{V}_{\mathbb{M}_\sigma(\mu_0)} \cap \mathsf{Path}(\mathbb{M}_{\sigma_{det}}(\mu_0))$, thanks to Lemma 1 and (a). Thus

$$\begin{aligned} E_\sigma &= \mathcal{V}_{\mathbb{M}_{\sigma_{det}}(\mu_0)} \cup (\mathcal{V}_{\mathbb{M}_\sigma(\mu_0)} \setminus \mathsf{Path}(\mathbb{M}_{\sigma_{det}}(\mu_0)) \cup (\mathsf{Path}(\mathbb{M}_\sigma(\mu_0)) \setminus \mathsf{F}_\infty) \\ &= E_{\sigma_{det}} \cup (\mathcal{V}_{\mathbb{M}_\sigma(\mu_0)} \cup (\mathsf{Path}(\mathbb{M}_\sigma(\mu_0)) \setminus \mathsf{F}_\infty) \setminus \mathsf{Path}(\mathbb{M}_{\sigma_{det}}(\mu_0)) \end{aligned}$$

where $E_{\sigma_{det}} = \mathcal{V}_{\mathbb{M}_{\sigma_{det}}(\mu_0)} \cup (\mathsf{Path}(\mathbb{M}_{\sigma_{det}}(\mu_0)) \setminus \mathsf{F}_\infty)$.

Finally, $\mathbf{P}_{\mathbb{M}_{\sigma_{det}}(\mu_0)}(\mathcal{V}_{\mathbb{M}_\sigma(\mu_0)} \cup (\mathsf{Path}(\mathbb{M}_\sigma(\mu_0)) \setminus \mathsf{F}_\infty) \setminus \mathsf{Path}(\mathbb{M}_{\sigma_{det}}(\mu_0)) = 0$ as no path of this set is σ_{det}-compatible. Therefore $\mathbf{P}_{\mathbb{M}_{\sigma_{det}}(\mu_0)}(E_{\sigma_{det}}) = 1$ which implies that $\mathbb{M}_{\sigma_{det}}(\mu_0)$ is AA-diagnosable. □

We can further restrict the strategies by limiting ourselves to belief-based strategy. This is far from an intuitive result. Indeed, while the AA-diagnosability

[3] In our framework, by definition, every strategy is 'observation based'.

of an oMC depends heavily on the exact values of the probabilities in the oMC, this result implies that the control only needs to remember the set of states potentially reached with a given observation sequence, not the probabilities with which one is in each state. Remark though that the choice made by the strategy in each belief depends on the probabilities.

Lemma 3. *Given $\mathbb{M}$ a CoMC, μ_0 an initial distribution, $\mathsf{S}^{\mathsf{F}} \subseteq S$ and σ a deterministic strategy, there exists a deterministic belief based strategy σ' such that $Disc^{\mathsf{AA}}(\mathbb{M}_\sigma(\mu_0)) = \mathbf{P}_{\mathbb{M}_\sigma(\mu_0)}(\mathsf{F}_\infty)$ implies $Disc^{\mathsf{AA}}(\mathbb{M}_{\sigma'}(\mu_0)) = \mathbf{P}_{\mathbb{M}_{\sigma'}(\mu_0)}(\mathsf{F}_\infty)$.*

Proof. Let $\mathbb{M}$ be a CoMC, μ_0 be an initial distribution and σ be a deterministic strategy such that $\mathbb{M}_\sigma(\mu_0)$ is AA-diagnosable. We define a belief based strategy σ' from σ in the following way. Let $\rho \in \mathsf{fPath}(\mathbb{M}_\sigma(\mu_0))$. We define by E_ρ the set of finite path producing the same belief as ρ, *i.e.* $E_\rho = \{\rho' \in \mathsf{fPath}(\mathbb{M}_\sigma(\mu_0)) \mid B_{\mathbb{M}(\mu_0)}(\mathsf{O}(\rho')) = B_{\mathbb{M}(\mu_0)}(\mathsf{O}(\rho))\}$. We define $\sigma'(B_{\mathbb{M}(\mu_0)}(\mathsf{O}(\rho))) = \bigcup_{\rho' \in E_\rho} \sigma(\mathsf{O}(\rho'))$. In other words, in a given belief, σ' allows anything that σ allowed at least once in this belief. Let us show that $\mathbb{M}_{\sigma'}(\mu_0)$ is AA-diagnosable.

Let two states $q = (s, \Sigma^\bullet, B) \in \mathsf{S}^{\mathsf{F}}$ and $q' = (s', \Sigma^\bullet, B) \in S \setminus \mathsf{S}^{\mathsf{F}}$ belonging to a BSCC of $\mathbb{M}_{\sigma'}(\mu_0)$ and reached by two finite paths ρ and ρ' of $\mathsf{fPath}(\mathbb{M}_{\sigma'}(\mu_0))$ with $\mathsf{O}(\rho) = \mathsf{O}(\rho')$. We will show that $d(\mathbb{M}_{\sigma'}(\mathbf{1}_q), \mathbb{M}_{\sigma'}(\mathbf{1}_{q'})) = 1$ using the characterisation given in Proposition 2. More precisely, for any observations sequence $w \in \Sigma^*$, and any pair of distributions on the set of states reached from q and from q' after observing w, we consider the probabilistic language generated by similar distributions in $\mathbb{M}_\sigma$ (*i.e.* distributions giving the same weight to the states of the original CoMC $\mathbb{M}$) and rely on the fact that $\mathbb{M}_\sigma$ is AA-diagnosable to show that the generated languages are different. This implies the distance is 1 thanks to Proposition 2.

Let $w \in \Sigma^*$ such that $\mathbb{P}_{\mathbb{M}_{\sigma'}(\mathbf{1}_q)}(w) > 0$ and $\mathbb{P}_{\mathbb{M}_{\sigma'}(\mathbf{1}_{q'})}(w) > 0$, we denote by B_w, B_q and $B_{q'}$ the beliefs reached after observing w from the beliefs B, $\{q\}$ and $\{q'\}$ respectively, let two distributions μ'_1 and μ'_2 such that $\mathsf{Supp}(\mu'_1) \subseteq B_q$, $\mathsf{Supp}(\mu'_2) \subseteq B_{q'}$. As σ' does not allow events that are never allowed by σ in the same belief, there exists an observation sequence $w_\sigma \in \Sigma^*$ such that $\mathbb{P}_{\mathbb{M}_\sigma(\mu_0)}(w_\sigma) > 0$ and the belief reached in $\mathbb{M}(\mu_0)$ after a path of observation w_σ from the initial distribution is B_w, *i.e.* $B_{\mathbb{M}(\mu_0)}(w_\sigma) = B_w$.

We can thus define initial distributions μ_1 and μ_2 on the set of states reached after observing w_σ in $\mathbb{M}_\sigma$ mimicking the distributions μ'_1 and μ'_2 (*i.e.* for every state $q_0 = (s_0, \Sigma_0, B_w)$ of $\mathbb{M}_{\sigma'}(\mu_0)$, we select some q_1, state of $\mathbb{M}_\sigma(\mu_0)$ associated to a σ-compatible paths ρ that ends in s_0 and such that $\mathsf{O}(\rho) = w_\sigma$, and we set for $i \in \{1, 2\}, \mu'_i(q_0) = \mu_1(q_1)$). From Proposition 3 and Proposition 2, there exists a word w_d such that $\mathbb{P}_{\mathbb{M}_\sigma(\mu_1)}(w_d) \neq \mathbb{P}_{\mathbb{M}_\sigma(\mu_2)}(w_d)$. This implies that there exists a word w'_d such that $\mathbb{P}_{\mathbb{M}_{\sigma'}(\mu'_1)}(w'_d) \neq \mathbb{P}_{\mathbb{M}_{\sigma'}(\mu'_2)}(w'_d)$. Indeed, let E be the set of observation sequences of the form $w'a$ where w' is a strict prefix of w_d, $a \in \Sigma$, $\mathbb{P}_{\mathbb{M}_{\sigma'}(\mu'_1)}(w'a) > 0$ and $\mathbb{P}_{\mathbb{M}_\sigma(\mu_1)}(w'a) = 0$. If $\mathbb{P}_{\mathbb{M}_{\sigma'}(\mu'_1)}(E) \neq \mathbb{P}_{\mathbb{M}_{\sigma'}(\mu'_2)}(E)$,

this implies our result. Otherwise, by construction of the strategy σ' we have:

$$\begin{aligned}\mathbb{P}_{\mathbb{M}_{\sigma'}(\mu_1')}(w_d) =& \mathbb{P}_{\mathbb{M}_{\sigma}(\mu_1)}(w_d) \times (1 - \mathbb{P}_{\mathbb{M}_{\sigma'}(\mu_1')}(E)) \\ \neq& \mathbb{P}_{\mathbb{M}_{\sigma}(\mu_2)}(w_d) \times (1 - \mathbb{P}_{\mathbb{M}_{\sigma'}(\mu_1')}(E)) \\ =& \mathbb{P}_{\mathbb{M}_{\sigma}(\mu_2)}(w_d) \times (1 - \mathbb{P}_{\mathbb{M}_{\sigma'}(\mu_2')}(E)) \\ =& \mathbb{P}_{\mathbb{M}_{\sigma'}(\mu_2')}(w_d),\end{aligned}$$

in which case we can choose $w_d' = w_d$. As this holds for any $w \in \Sigma^*$ and pair of distributions μ_1' and μ_2', according to Proposition 2 we have $d(\mathbb{M}_{\sigma'}(\mathbf{1}_q), \mathbb{M}_{\sigma'}(\mathbf{1}_{q'})) = 1$. From Theorem 2, we can thus deduce that $\mathbb{M}_{\sigma'}(\mu_0)$ is AA-diagnosable. Therefore belief-based strategies are sufficient to decide AA-diagnosability. □

A naive NEXPTIME algorithm can be obtained from these two lemmas: we guess a deterministic belief-based strategy then verify AA-diagnosability of the exponential oMC generated by the CoMC and the strategy. In the following proposition, we show how to efficiently build a good belief-based strategy, which gives us an EXPTIME algorithm.

Proposition 5. *The* AA*-diagnosability problem over CoMCs is in* EXPTIME*.*

Proof. Let $\mathbb{M}$ be a CoMC and μ_0 be an initial distribution. This proof is done in two steps.

1. We show that, given two deterministic belief based strategies σ_1 and σ_2 such that σ_1 is less restrictive than σ_2 and a state q belonging to a BSCC of both $\mathbb{M}_{\sigma_1}(\mu_0)$ and $\mathbb{M}_{\sigma_2}(\mu_0)$, then if the paths of $\mathbb{M}_{\sigma_2}(\mu_0)$ that visits q are almost surely AA-disclosing then so are the paths of $\mathbb{M}_{\sigma_1}(\mu_0)$ that visits q. In other words, within a BSCC, the least restrictive a strategy is, the better it is for the purpose of diagnosis.
2. Thanks to the result obtained in the first step, we efficiently build a strategy in the form of a greatest fixed point: we start by the most permissive strategy and iteratively restrict it to prune the BSCC that cause the strategy not to achieve AA-diagnosability.

Let σ and σ' be two deterministic belief-based strategies such that for any belief B of $\mathbb{M}$ $\sigma(B) \subseteq \sigma'(B)$. Let q be a faulty state associated to a belief B and belonging to a BSCC of both $\mathbb{M}_{\sigma}(\mu_0)$ and $\mathbb{M}_{\sigma'}(\mu_0)$. Assume that there exists a positive measure of paths in $\mathbb{M}_{\sigma'}(\mu_0)$ that visit q and that are not associated to an AA-disclosing observation sequence. Defining $B' = (B \setminus \mathsf{S}^{\mathsf{F}}) \cup \{q\}$, this is equivalent to saying that the CoMC $\mathbb{M}_{\sigma'}(\mu_1)$, where μ_1 is an initial distribution of support B', is not AA-diagnosable. Therefore we can use the characterisation of Theorem 2. Without loss of generality, as q belongs to a BSCC, we can assume the pair of states given by the characterisation is (q, q') where $q' \notin \mathsf{S}^{\mathsf{F}}$, is associated to the belief B, belongs to a BSCC of $\mathbb{M}_{\sigma'}(\mu_1)$ and is such that $d(\mathbb{M}_{\sigma'}(\mathbf{1}_q), \mathbb{M}_{\sigma'}(\mathbf{1}_{q'})) < 1$. Let w, π_1 and π_2 be the observation sequence and the two distributions obtained by applying Proposition 2 on the pair of CoMC $(\mathbb{M}_{\sigma'}(\mathbf{1}_q), \mathbb{M}_{\sigma'}(\mathbf{1}_{q'}))$. Let $q'' \notin \mathsf{S}^{\mathsf{F}}$ be a state belonging to a BSCC of $\mathbb{M}_{\sigma}(\mu_1)$

reachable from q' by a σ-compatible path with observation sequence ww'. Let π'_1 and π'_2 be the distribution obtained after observing w' starting in π_1 and π_2. As $\forall v \in \Sigma^*, \mathbf{P}_{\mathbb{M}_{\sigma'}(\pi_1)}(v) = \mathbf{P}_{\mathbb{M}_{\sigma'}(\pi_2)}(v)$, we also have $\forall v \in \Sigma^*, \mathbf{P}_{\mathbb{M}_{\sigma'}(\pi'_1)}(v) = \mathbf{P}_{\mathbb{M}_{\sigma'}(\pi'_2)}(v)$. This implies that $\forall v \in \Sigma^*, \mathbf{P}_{\mathbb{M}_{\sigma}(\pi'_1)}(v) = \mathbf{P}_{\mathbb{M}_{\sigma}(\pi'_2)}(v)$. Indeed, given $v \in \Sigma^*$, we have

$$
\begin{aligned}
\mathbf{P}_{\mathbb{M}_\sigma(\pi'_1)}(v) &= \sum_{\rho \in \mathsf{O}^{-1}(v)} \mathbf{P}_{\mathbb{M}_\sigma(\pi'_1)}(\rho) \\
&= \sum_{\rho = s_0 \Sigma_0 \dots s_n \in \mathsf{O}^{-1}(v)} \pi'_1(s_0) \prod_{i=0}^{n-1} \sigma(\mathsf{O}(v_{\downarrow 2i+1}))(\Sigma_i) p(s_{i+1} \mid s_i, \Sigma_i) \\
&= \left(\prod_{i=0}^{n-1} \sigma(\mathsf{O}(v_{\downarrow 2i+1}))(\Sigma_i) \right) \sum_{\rho = s_0 \Sigma_0 \dots s_n \in \mathsf{O}^{-1}(v)} \pi'_1(s_0) \prod_{i=0}^{n-1} \frac{T(s_i, s_{i+1})}{\sum_{s'', \mathsf{O}(s'') \in \Sigma_i} T(s_i, s'')} \\
&= \left(\prod_{i=0}^{n-1} \sigma(\mathsf{O}(v_{\downarrow 2i+1}))(\Sigma_i) \right) \sum_{\rho = s_0 \Sigma_0 \dots s_n \in \mathsf{O}^{-1}(v)} \pi'_2(s_0) \prod_{i=0}^{n-1} \frac{T(s_i, s_{i+1})}{\sum_{s'', \mathsf{O}(s'') \in \Sigma_i} T(s_i, s'')} \\
&= \mathbf{P}_{\mathbb{M}_\sigma(\pi'_2)}(v).
\end{aligned}
$$

As a consequence, $d(\mathbb{M}_\sigma(\mathbf{1}_q), \mathbb{M}_\sigma(\mathbf{1}_{q'})) < 1$. From Theorem 2, this implies that $\mathbb{M}_\sigma(\mu_1)$ is not AA-diagnosable and thus there exists a positive measure of paths in $\mathbb{M}_\sigma(\mu_0)$ that visit q and that are not associated to an AA-disclosing observation sequence. Therefore, having restricted the strategy σ' did not allow to regain AA-diagnosability of the paths visiting q. This means that a strategy achieving AA-diagnosability of the CoMC must ensure that q cannot be reached.

Using this result, we build iteratively the most permissive strategy ensuring AA-diagnosability. We start with the strategy σ_0 allowing everything. Assume we built the strategy σ_k such that any less permissive strategy do not ensure AA-diagnosability. If $\mathbb{M}_{\sigma_k}(\mu_0)$ is not AA-diagnosable, there exists two states s and s' associated to the same belief B that satisfies the characterisation of Theorem 2. W.l.o.g one can assume that both of these states belong to BSCCs of $\mathbb{M}_{\sigma_k}(\mu_0)$. From our preliminary result, we know that any strategy that contains the states s and s' in a BSCC does not ensure AA-diagnosability. As any strategy less permissive than σ_k does not ensure AA-diagnosability, we need to restrict the strategy so that the belief B is not reachable, or that B is not associated to states belonging to a BSCC anymore. The latter is in fact not sufficient as Theorem 2 would still apply on the pair of states (s, s'). Thus we build σ_{k+1} as the most permissive strategy such that $\mathbb{M}_{\sigma_{k+1}}(\mu_0)$ does not contain the belief B, which can easily be done with belief based strategies. This procedure ends when the strategy σ_n that is created either is the most permissive strategy ensuring AA-diagnosability or if one cannot build a strategy removing the problematic belief. This algorithm is in EXPTIME as every step of the procedure can be done in exponential time (verification of AA-diagnosability, identification of the pair of problematic states and creation of the new strategy are all steps that can be done in EXPTIME) and there is at most exponentially many steps as each one of them

removes at least one belief from the system, and there are exponentially many beliefs. Therefore, the AA-diagnosability problem can be solved in EXPTIME. □

Remark that the above proof builds the strategy ensuring AA-diagnosability when it exists.

4 Conclusion

This paper considers the accurate approximate notion of disclosure for diagnosability. We establish how to decide AA-diagnosability in CoMC and how to measure the AA-disclosure in oMC. Measuring the AA-disclosure was not developed for CoMC here as the notion is undecidable (straightforward application of the undecidability of the emptiness problem for probabilistic automata).

Opacity is a notion that intuitively appears as some kind of dual to diagnosability. The goal of opacity is to make sure some secret paths of the system are not detected by an observer. Following the idea that some small amount of revealed secret information is not problematic, this line of research favors a quantitative approach to the problem, thus closer to the AA-disclosure problem we studied for oMC. In this endeavour, various measures for the disclosure set have been introduced [22, 1, 4, 3]. Opacity has been studied in an active framework called observable Markov decision processes (oMDP) where the controller is deemed internal to the system and thus makes its choice with more information than just the observation sequence. This framework is thus not equivalent to the CoMC model presented in this paper; the strategy is more powerful. As such, while measuring the disclosure is undecidable (for any disclosure notion) in CoMC, some positive results were established in oMDP [2]. However, as this work only considered the exact notion of disclosure, it would be interesting to see if the approximate approach pushed here could also be applied for oMDP. Moreover, this framework also makes sense for a study of diagnosability as the control defined in oMDP can correspond to designing choices of the system.

References

1. B. Bérard, K. Chatterjee, and N. Sznajder. Probabilistic opacity for Markov decision processes. *Information Processing Letters*, 115(1):52–59, 2015.
2. B. Bérard, S. Haddad, and E. Lefaucheux. Probabilistic disclosure: Maximisation vs. minimisation. In *Proceedings of FSTTCS'17*, volume 93 of *LIPIcs*, pages 13:1–13:14. Leibniz-Zentrum für Informatik, 2017.
3. B. Bérard, O. Kouchnarenko, J. Mullins, and M. Sassolas. Preserving opacity on interval Markov chains under simulation. In *Proceedings of WODES'16*, pages 319–324. IEEE, 2016.
4. B. Bérard, J. Mullins, and M. Sassolas. Quantifying opacity. *Mathematical Structures in Computer Science*, 25(2):361–403, 2015.
5. 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.

6. N. Bertrand, S. Haddad, and E. Lefaucheux. Foundation of diagnosis and predictability in probabilistic systems. In *Proceedings of FSTTCS'14*, volume 29 of *LIPIcs*, pages 417–429. Leibniz-Zentrum für Informatik, 2014.
7. N. Bertrand, S. Haddad, and E. Lefaucheux. Accurate approximate diagnosability of stochastic systems. In *Proceedings of LATA'16*, volume 9618 of *LNCS*, pages 549–561. Springer, 2016.
8. N. Bertrand, S. Haddad, and E. Lefaucheux. A Tale of Two Diagnoses in Probabilistic Systems. *Information and Computation*, page 104441, 2019.
9. 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.
10. A. Borodin. On relating time and space to size and depth. *SIAM J. Comput.*, 6:733–744, 12 1977.
11. A. Borodin, J. [von zur Gathen], and J. Hopcroft. Fast parallel matrix and gcd computations. *Information and Control*, 52(3):241 – 256, 1982.
12. M. P. Cabasino, A. Giua, S. Lafortune, and C. Seatzu. A new approach for diagnosability analysis of petri nets using verifier nets. *Transactions on Automatic Control*, 57(12):3104–3117, 2012.
13. F. Cassez and S. Tripakis. Fault diagnosis with static and dynamic observers. *Fundamenta Informaticae*, 88:497–540, 2008.
14. E. Chanthery and Y. Pencolé. Monitoring and active diagnosis for discrete-event systems. *IFAC Proceedings Volumes*, 42(8):1545 – 1550, 2009.
15. K. Chatterjee, L. Doyen, H. Gimbert, and T. A. Henzinger. Randomness for free. In *Proceedings of MFCS'10*, volume 6281 of *LNCS*, pages 246–257. Springer, 2010.
16. T. Chen and S. Kiefer. On the total variation distance of labelled Markov chains. In *Proceedings of CSL-LICS'14*, pages 33:1–33:10. ACM, 2014.
17. S. Haar, S. Haddad, T. Melliti, and S. Schwoon. Optimal constructions for active diagnosis. *Journal of Computer and System Sciences*, 83(1):101–120, 2017.
18. S. Jiang, Z. Huang, V. Chandra, and R. Kumar. A polynomial algorithm for testing diagnosability of discrete-event systems. *Transactions on Automatic Control*, 46(8):1318–1321, 2001.
19. A. R. Meyer and L. J. Stockmeyer. The equivalence problem for regular expressions with squaring requires exponential space. In *SWAT'72*, pages 125–129. IEEE, 1972.
20. C. Morvan and S. Pinchinat. Diagnosability of pushdown systems. In *Proceedings of HVC'09*, volume 6405 of *LNCS*, pages 21–33. Springer, 2009.
21. K Mulmuley. A fast parallel algorithm to compute the rank of a matrix over an arbitrary field. In *STOC 86*, page 338339, 1986.
22. A. Saboori and Ch. N. Hadjicostis. Current-state opacity formulations in probabilistic finite automata. *Transactions on Automatic Control*, 59(1):120–133, 2014.
23. M. Sampath, S. Lafortune, and D. Teneketzis. Active diagnosis of discrete-event systems. *Transactions on Automatic Control*, 43(7):908–929, 1998.
24. M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis. Diagnosability of discrete-event systems. *Transactions on Automatic Control*, 40(9):1555–1575, 1995.
25. D. Thorsley and D. Teneketzis. Diagnosability of stochastic discrete-event systems. *Transactions on Automatic Control*, 50(4):476–492, 2005.
26. D. Thorsley and D. Teneketzis. Active acquisition of information for diagnosis and supervisory control of discrete-event systems. *Discrete Event Dynamic Systems*, 17:531–583, 2007.

A AA-disclosure problem for oMC

Theorem 3. *The* AA*-disclosure problem for finite oMC is* PSPACE*-complete.*

We decompose the proof of the theorem in the two following proposition, each establishing one direction.

Proposition 6. *The* AA*-disclosure problem for finite oMC is in* PSPACE*.*

Proof. To establish this result, we first build an exponential size oMC which contains additional information: the set of states the system could be in after the observation sequence. Then we show that there are two kinds of BSCC in this new oMC: the ones that are reached by paths that almost surely have an AA-disclosing observation sequence, and the ones that are reached by paths that do not correspond to AA-disclosing observation sequences. We can then use the existing results for the AA-diagnosability problem to determine the status of each BSCC. Therefore, computing the AA-disclosure of the oMC is equivalent to computing the probability to reach the "AA-disclosing" BSCC, which can be done in NC in the size of the oMC, thus giving an overall PSPACE algorithm.

Let $\mathcal{M} = (S, p, \mathsf{O})$ be a finite oMC and μ_0 be an initial distribution. We build a new oMC $\mathcal{M}' = (S', p', \mathsf{O}')$ which has the same behaviour as $\mathcal{M}$ but where the states are enriched with an additional information (the set of states the system can be in, given the produced observation sequence):

- $S' = S \times 2^S$;
- For $(s, B), (s', B') \in S'$, $p'((s', B') \mid (s, B)) = p(s' \mid s)$ if $B' = \cup_{q \in B} \mathsf{Supp}(p(q)) \cap \mathsf{O}^{-1}(\mathsf{O}(s'))$ else, $p'((s', B') \mid (s, B)) = 0$;
- For $(s, B) \in S'$, $\mathsf{O}'(s, B) = \mathsf{O}(s)$.

We define the initial distribution μ_0' for $\mathcal{M}'$ by $\mu_0'(s, \mathsf{Supp}(\mu_0) \cap \mathsf{O}^{-1}(\mathsf{O}(s))) = \mu_0(s)$ for all $s \in S$. There is a one-to-one correspondence between the paths of $\mathcal{M}(\mu_0)$ and $\mathcal{M}'(\mu_0')$: every path $\rho = s_0 s_1 \cdots s_n$ of $\mathcal{M}(\mu_0)$ is associated to an unique path $\rho' = (s_0, B_0)(s_1, B_1) \cdots (s_n, B_n)$ with $\mathsf{O}(\rho) = \mathsf{O}(\rho')$, $\mathbf{P}_{\mathcal{M}(\mu_0)}(\rho) = \mathbf{P}_{\mathcal{M}'(\mu_0')}(\rho')$ and B_n contains the set of states of S that can be reached with a path of observation $\mathsf{O}(\rho)$. Due to the latter property, B_n only depends on $\mathsf{O}(\rho)$ and is called the *belief* associated to $\mathsf{O}(\rho)$.

Let $(s, B) \in S'$ such that $s \in \mathsf{S}^{\mathsf{F}}$ and (s, B) belongs to a BSCC of $\mathcal{M}'$. We claim that either for every path ρ ending in (s, B), $\mathbf{P}(\{\rho' \in \mathsf{Path}(\mathcal{M}'(\mu_0')) \mid \rho \preceq \rho' \wedge \mathsf{O}(\rho') \in D^{\mathsf{AA}}\}) = 0$ or for every path ρ ending in (s, B), $\mathbf{P}(\{\rho' \in \mathsf{Path}(\mathcal{M}'(\mu_0')) \mid \rho \preceq \rho' \wedge \mathsf{O}(\rho') \in D^{\mathsf{AA}}\}) = \mathbf{P}(\rho)$. In other words, there are two categories of BSCC composed of faulty states: the good ones, that almost surely accurate approximately disclose the fault, and the bad ones that do not accurate approximately disclose the fault at all. Moreover, a BSCC containing the state (s, B) do not disclose the fault at all iff there exists a state $s' \in B$ such that s' belongs to a BSCC of $\mathcal{M}$, $s' \notin \mathsf{S}^{\mathsf{F}}$ and $d(\mathcal{M}(\mathbf{1}_s), \mathcal{M}(\mathbf{1}_{s'})) < 1$.

Let (s, B) be a state belonging to a BSCC of $\mathcal{M}'$. Assume that for all $s' \in B$ such that s' belongs to a BSCC of $\mathcal{M}$ and $s' \notin \mathsf{S}^{\mathsf{F}}$ we have $d(\mathcal{M}(\mathbf{1}_s), \mathcal{M}(\mathbf{1}_{s'})) = 1$.

We denote $B' = (B \setminus \mathsf{S}^{\mathsf{F}}) \cup \{s\}$, and define $\mathcal{M}''$ by removing the path leading to a faulty state (aka, a path either starts faulty or forever remain correct). Then as s belongs to a BSCC of $\mathcal{M}$, we can directly use Theorem 2 to obtain that for any initial distribution μ_1 of support B', we have that $\mathcal{M}''(\mu_1)$ is AA-diagnosable. As the limitation to the states of $B \setminus B'$ and the transformation from $\mathcal{M}$ to $\mathcal{M}''$ can only increase the failure proportion, this ensures that $\mathbf{P}(\{\rho' \in \mathsf{Path}(\mathcal{M}'(\mu_0')) \mid \rho \preceq \rho' \wedge \mathsf{O}(\rho') \in Disc^{\mathsf{AA}}\}) = \mathbf{P}(\rho)$.

Conversely, if there exists a state $s' \in B$ such that s' belongs to a BSCC of B, $s' \notin \mathsf{S}^{\mathsf{F}}$ and $d(\mathcal{M}(\mathbf{1}_s), \mathcal{M}(\mathbf{1}_{s'})) < 1$, then one can rely on the proof of Lemma A of [8] to obtain the result. For the sake of pedagogy, we present the proof here in the simpler case where B does not contain any faulty state beside s. Using Proposition 2 and the correspondence between $\mathcal{M}$ and $\mathcal{M}'$, one deduces that there exists $\rho_{(s,B)} \in \mathsf{fPath}(\mathcal{M}(\mathbf{1}_{(s,B)}))$ and $\alpha > 0$ such that for all $w \in \Sigma^*$ with $\mathsf{O}(\rho) \leq w$

$$\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s,B)})}(\{\rho' \in \mathsf{fPath}(\mathcal{M}'(\mathbf{1}_{(s,B)})) \mid \rho_{(s,B)} \preceq \rho' \wedge \mathsf{O}(\rho') = w\}) \tag{1}$$

$$\leq \alpha \mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s',B)})}(\{\rho' \in \mathsf{fPath}(\mathcal{M}'(\mathbf{1}_{(s',B)})) \mid \mathsf{O}(\rho') = w\}). \tag{2}$$

Therefore, for all $w \in \Sigma^*$ and initial distribution μ_1 of support B we have:

$$\mathsf{Fprop}_{\mathcal{M}'(\mu_1)}(w) \leq \frac{\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s,B)})}(w)}{\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s,B)})}(w) + \frac{\mu_1(s')}{\mu_1(s)}\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s',B)})}(w)} \tag{3}$$

$$\leq \frac{\varepsilon_w + \sum\limits_{\rho \mid \mathsf{O}(\rho\rho_{(s,B)}) \leq w} \frac{\alpha \mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s,B)})}(\rho)}{\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s,B)})}(\rho_{(s,B)})} \mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s',B)})}(w^\rho)}{\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s,B)})}(w) + \frac{\mu_1(s')}{\mu_1(s)}\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s',B)})}(w)} \tag{4}$$

where w^ρ is such that $w = \mathsf{O}(\rho)w^\rho$, the first term $\varepsilon_w = \mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s,B)})}(\{\rho \in \mathsf{fPath}(\mathcal{M}(\mathbf{1}_{(s,B)}) \mid \nexists \rho_1, \rho_2, \rho = \rho_1\rho_{(s,B)}\rho_2 \wedge \mathsf{O}(\rho) = w\})$ is the probability of the set of paths with observation w that do not contain the infix $\rho_{(s,B)}$ and the second term relies on the bound from Equation 2 to bound the probability of every other paths. As with probability 1, a path of $\mathcal{M}'(\mathbf{1}_{(s,B)})$ visits (s, B) infinitely often, it will almost surely contain a $\rho_{(s,B)}$ subpath, more precisely: the value $\frac{\varepsilon_w}{\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s,B)})}(w)}$ almost surely converges to 0 when $|w|$ diverges to ∞. Let $w \in \Sigma^\omega$, if $\mathsf{Fprop}_{\mathcal{M}'(\mu_1)}(w_{\downarrow n}) \xrightarrow{n\to\infty} 1$ then, for all ρ such that $\mathsf{O}(\rho\rho_{(s,B)}) \leq w$ we have that $\frac{\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s',B)})}(w^\rho_{\downarrow n})}{\mathbf{P}_{\mathcal{M}'(\mathbf{1}_{(s,B)})}(w_{\downarrow n})}$ converges to 0, thus, due to Equation 4, $\varepsilon_{w_{\downarrow n}}$ does not converge to 0, which can only happen with probability 0. Therefore $\mathsf{Fprop}_{\mathcal{M}'(\mu_1)}(w_{\downarrow n})$ almost surely does not converge to 1. This implies that $\mathbf{P}\{\rho' \in \mathsf{Path}(\mathcal{M}'(\mu_0')) \mid \rho \preceq \rho' \wedge \mathsf{O}(\rho') \in D^{AA}\} = 0$.

This result establishes that the BSCC of $\mathcal{M}'$ are partitionned between the good ones that accurately approximately and almost surely disclose the fault and the bad ones that do not accurately approximately disclose it at all. Moreover, given a state (s_0, B_0) belonging to a BSCC of $\mathcal{M}'$, if there exists a state $s_0' \in B_0$ such that s_0' belongs to a BSCC of B, $s_0' \notin \mathsf{S}^{\mathsf{F}}$ and $d(\mathcal{M}(\mathbf{1}_{s_0}), \mathcal{M}(\mathbf{1}_{s_0'})) < 1$,

then for any state (s_1, B_1) belonging to the same BSCC, one can find a state $s_1' \in B_1$ satisfying a similar property with respect to s_1. In other words, for every BSCC of $\mathcal{M}'$, we only need to check a single state (s, B) of the BSCC to identify whether the BSCC is disclosing or not. Furthermore, this check can be done by testing the distance 1 between copies of $\mathcal{M}$ starting in s and copies starting in some of the states in B. There is thus at most linearly many tests to do, each of which can be done in polynomial time in the size of $\mathcal{M}$.

Therefore, one can obtain the value of $Disc^{AA}(\mathcal{M}'(\mu_0'))$ by computing the probability to reach the good BSCC, which is known to be possible in PTIME in the size of $\mathcal{M}'$. In fact, as computing this probability amount to solve a linear system of equations, this can even be done in NC [11, 21]. The oMC $\mathcal{M}'$ being exponential in the size of $\mathcal{M}$, and as NC blown up to the exponential is equal to PSPACE [10], this yields a PSPACE algorithm. As $Disc^{AA}(\mathcal{M}(\mu_0)) = Disc^{AA}(\mathcal{M}'(\mu_0'))$, this allows us to solve the AA-disclosure problem. □

Proposition 7. *The* AA*-disclosure problem for finite oMC is* PSPACE*-hard.*

Proof. We now establish the hardness by reducing the universality problem for non-deterministic finite automaton (NFA), which is known to be PSPACE-complete [19].

An NFA is a tuple $\mathcal{A} = (Q, \Sigma, T, q_0, F)$ where Q is the set of states, q_0 is the initial state, F is the set of accepting states, Σ is the alphabet and $T \in Q \times \Sigma \times Q$ is the transition function. An NFA is universal if for all $w = a_1 a_2 \dots a_n \in \Sigma^n$, there exists a path $q_0 a_1 q_1 a_2 \dots q_n$ such that $q_n \in F$ and for all $1 \leq i \leq n, (q_{i-1}, a_i, q_i) \in T$.

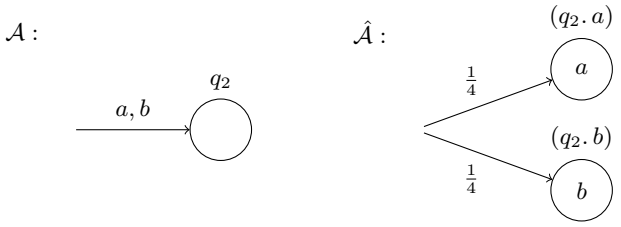

Fig. 3. From NFA $\mathcal{A}$ to incomplete oMC $\hat{\mathcal{A}}$. The label next to the state is its name. We will not always display the state's name so as not to overload the figure.

Let $\mathcal{A} = (Q, \Sigma, T, q_0, F)$ be an NFA. W.l.o.g. we can assume that $F = Q$ and $\Sigma = \{a, b\}$. Our first step is to push the observations onto the states (as shown in Figure 3). From $\mathcal{A}$ we define the incomplete oMC $\hat{\mathcal{A}} = (S_A, p_A, O_A)$ and the initial distribution μ_0^A such that:

- $S_A = Q \times \Sigma$;
- for $(q, c), (q', d) \in S_A$, if $(q, d, q') \in T$, then $p_A((q', d) \mid (q, c)) = \frac{1}{|S_A|+1}$, else $p_A((q', d) \mid (q, c)) = 0$;

– for $(q, c) \in S_A, O_A(q, c) = c$;
– for $(q', d) \in S_A$, if $(q_0, d, q') \in T$, then $\mu_0^A(q', d) = \frac{1}{|S_A|+1}$, else $\mu_0^A(q', d) = 0$.

This oMC is incomplete as none of the distributions μ_0^A and $p_A(\cdot \mid s)$ (for $s \in S_A$) sum to 1. We now build the oMC $\mathcal{M} = (S, p, O)$ represented in Figure 4 where

– $S = S_A \cup \{s_\sharp, f_a, f_b, f_\sharp\}$;
– given $s, s' \in S_A$, $p(s' \mid s) = p_A(s' \mid s)$, $p(s_\sharp \mid s) = 1 - \sum_{s' \in S_A} p(s' \mid s)$, for $h \in \{f_a, f_b\}$ and $g \in \{f_a, f_b, f_\sharp\}$, $p(g \mid h) = 1/3$ and $p(f_\sharp \mid f_\sharp) = p(s_\sharp \mid s_\sharp) = 1$;
– for $s \in S_A$, $O(s) = O_A(s)$, $O(s_\sharp) = O(f_\sharp) = \sharp, O(f_a) = a$ and $O(f_b) = b$.

We also define μ_0 as $\mu_0(s) = \mu_0^A(s)$ for $s \in S_A$ and $\mu_0(f_a) = \mu_0(f_b) = \frac{1-\sum_{s \in S_A} \mu_0(s)}{2}$.

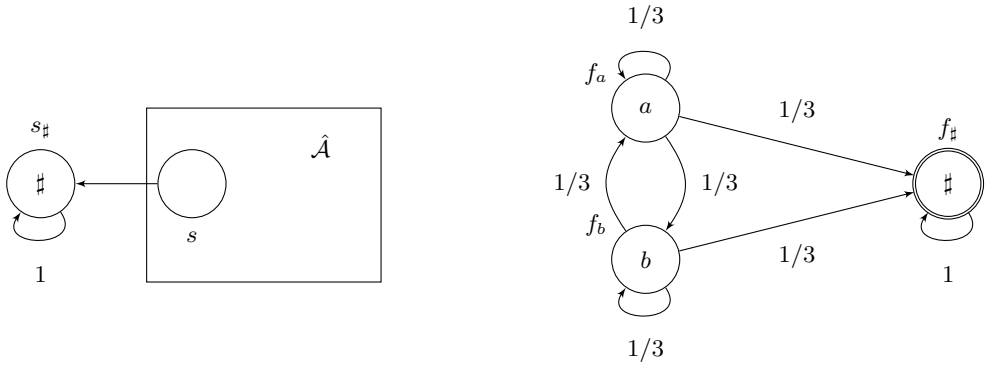

Fig. 4. A reduction for PSPACE-hardness of the AA-disclosure problem.

Choosing $\mathsf{S}^{\mathsf{F}} = \{f_\sharp\}$, let us show that $\mathcal{A}$ is not universal iff $Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0)) > 0$.

Suppose first that $\mathcal{A}$ is not universal. There thus exists a word $w \in \Sigma^*$ such that no path starting in S_A has observation sequence w. As there exists one faulty path ρ (starting in either f_a or f_b) associated to $w\sharp$, we have $\mathsf{Fprop}_{\mathcal{M}(\mu_0)}(w\sharp) = 1$. Therefore $Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0)) \geq \mathbf{P}_{\mathcal{M}(\mu_0)}(\rho) > 0$.

Conversely, assume that $\mathcal{A}$ is universal. Let ρ be a path ending in $f_\sharp$ with observation sequence $\mathsf{O}(\rho) = w\sharp$ for some $w \in \Sigma^*$. As $\mathcal{A}$ is universal, there exists a finite path ρ' in $\hat{\mathcal{A}}$ with observation sequence w. As for every state s of $\hat{\mathcal{A}}$, $p(s_\sharp \mid s) > 0$, ρ' can be extended into a finite path ρ'' ending in $s_\sharp$ with observation $w\sharp$. Thus, $\mathsf{Fprop}_{\mathcal{M}(\mu_0)}(w\sharp) < 1$. Moreover, every path ending with a $\sharp$ remains with probability 1 in either $s_\sharp$ or $f_\sharp$, due to this for every $k \geq 2$, $\mathsf{Fprop}_{\mathcal{M}(\mu_0)}(w\sharp^k) = \mathsf{Fprop}_{\mathcal{M}(\mu_0)}(w\sharp)$. Therefore, $w\sharp^\omega \notin D^{\mathsf{AA}}$. This implies that no infinite path visiting $f_\sharp$ corresponds to an AA-disclosing observation sequence. $f_\sharp$ being the only faulty state, $Disc^{\mathsf{AA}}(\mathcal{M}(\mu_0)) = 0$. □