On probability-raising causality in Markov decision processes ⋆

Christel Baier ✉, Florian Funke, Jakob Piribauer ✉, and Robin Ziemek ✉

Technische Universität Dresden
{christel.baier, florian.funke, jakob.piribauer,robin.ziemek}@tu-dresden.de

Abstract. The purpose of this paper is to introduce a notion of causality in Markov decision processes based on the probability-raising principle and to analyze its algorithmic properties. The latter includes algorithms for checking cause-effect relationships and the existence of probability-raising causes for given effect scenarios. Inspired by concepts of statistical analysis, we study quality measures (recall, coverage ratio and f-score) for causes and develop algorithms for their computation. Finally, the computational complexity for finding optimal causes with respect to these measures is analyzed.

1 Introduction

As modern software systems control more and more aspects of our everyday lives, they grow increasingly complex. Even small changes to a system might cause undesired or even disastrous behavior. Therefore, the goal of modern computer science does not only lie in the development of powerful and versatile systems, but also in providing comprehensive techniques to understand these systems. In the area of formal verification, counterexamples, invariants and related certificates are often used to provide a verifiable justification that a system does or does not behave according to a specification (see e.g., [30,16,32]). These, however, provide only elementary insights on the system behavior. Thus, there is a growing demand for a deeper understanding on *why* a system satisfies or violates a specification and *how* different components influence the performance. The analysis of causal relations between events occurring during the execution of a system can lead to such understanding. The majority of prior work in this direction relies on causality notions based on Lewis' counterfactual principle [29] stating the effect would not have occurred if the cause would not have happened. A prominent formalization of the counterfactual principle is given by Halpern and Pearl [21] via structural equation models. This inspired formal definitions of causality and related notions of blameworthiness and responsibility in Kripke and game structures (see, e.g., [15,11,14,40,19,41,7]).

In this work, we approach the concept of causality in a probabilistic setting, where we focus on the widely accepted *probability-raising principle* which has its roots in

⋆ This work was funded by DFG grant 389792660 as part of TRR 248, the Cluster of Excellence EXC 2050/1 (CeTI, project ID 390696704, as part of Germany's Excellence Strategy), DFG-projects BA-1679/11-1 and BA-1679/12-1,and the RTG QuantLA (GRK 1763).

P. Bouyer and L. Schröder (Eds.): FoSSaCS 2022, LNCS 13242, pp. 40–60, 2022.
https://doi.org/10.1007/978-3-030-99253-8_3

Table 1. Complexity results for MDPs and Markov chains (MC) with fixed effect set

	for fixed set Cause		find optimal cause	
	check PR condition	compute quality values (recall, covratio, f-score)	covratio-optimal = recall-optimal	f-score-optimal
SPR	$\in$ P	poly-time	poly-time	poly-space poly-time for MC threshold problem $\in$ NP $\cap$ coNP
GPR	$\in$ PSPACE and $\in$ P for MC	poly-time	poly-space threshold problems $\in$ PSPACE and NP-hard and NP-complete for MC	

philosophy [38,39,18,22] and has been refined by Pearl [35] for causal and probabilistic reasoning in intelligent systems. The different notions of probability-raising cause-effect relations discussed in the literature share the following two main principles:

(C1) Causes raise the probabilities for their effects, informally expressed by the requirement "$\Pr(\text{effect} \mid \text{cause}) > \Pr(\text{effect})$".
(C2) Causes must happen before their effects.

Despite the huge amount of work on probabilistic causation in other disciplines, research on probability-raising causes in the context of formal methods is comparably rare and has concentrated on Markov chains (see, e.g., [24,25,6] and the discussion of related work in Section 3.2). To the best of our knowledge, probabilistic causation for probabilistic operational models with nondeterminism has not been studied before.

We formalize the principles (C1) and (C2) for Markov decision processes (MDPs), a standard operational model combining probabilistic and non-deterministic behavior, and concentrate on reachability properties where both cause and effect are given as sets of states. Condition (C1) can be interpreted in two natural ways in this setting: On one hand, the probability-raising property can be locally required for each element of the cause. Such causes are called *strict probability-raising (SPR) causes* in our framework. This interpretation is especially suited when the task is to identify system states that have to be avoided for lowering the effect probability. On the other hand, one might want to treat the cause set globally as a unit in (C1) leading to the notion of *global probability-raising (GPR) cause*. Considering the cause set as a whole is better suited when further constraints are imposed on the candidates for cause set. This might apply, e.g., when the set of non-terminal states of the given MDP is partitioned into sets of states S_i under the control of an agent i, $1 \leqslant i \leqslant k$. For the task to identify which agent's decisions cause the effect only the subsets of $S_1, \ldots, S_k$ are candidates for causes. Furthermore, global causes are more appropriate when causes are used for monitoring purposes under partial observability constraints as then the cause candidates are sets of indistinguishable states.

Different causes for an effect according to our definition can differ substantially regarding how well they predict the effect and how well the executions exhibiting the cause cover the executions showing the effect. Taking inspiration from measures used in statistical analysis (see, e.g., [36]), we introduce quality measures that allow us to compare causes and to look for optimal causes: The *recall* captures the probability that the effect is indeed preceded by the cause. The *coverage-ratio* quantifies the fraction of the probability that cause and effect are observed and the probability that the effect but not the cause is observed. Finally, the *f-score*, a widely used quality measure for binary classifiers, is the harmonic mean of recall and precision, i.e., the probability that the cause is followed by the effect.

Contributions. The goal of this work are the mathematical and algorithmic foundations of probabilistic causation in MDPs based on (C1) and (C2). We introduce strict and global probability-raising causes in MDPs (Section 3). Algorithms are provided to check whether given cause and effect sets satisfy (one of) the probability-raising conditions (Section 4.1 and 4.2) and to check the existence of causes for a given effect (Section 4.1). In order to evaluate the coverage properties of a cause, we subsequently introduce the above-mentioned quality measures (Section 5.1). We give algorithms for computing these values for given cause-effect relations (Section 5.2) and characterize the computational complexity of finding optimal causes with respect to the different measures (Section 5.3). Table 1 summarizes our complexity results. An extended version of this paper containing the omitted proofs can be found in [8].

2 Preliminaries

Throughout the paper, we will assume some familiarity with basic concepts of Markov decision processes. Here, we only present a brief summary of the notations used in the paper. For more details, we refer to [37,9,23].

A *Markov decision process (MDP)* is a tuple $\mathcal{M} = (S, Act, P, \mathsf{init})$ where S is a finite set of states, Act a finite set of actions, $\mathsf{init} \in S$ the initial state and $P : S \times Act \times S \to [0,1]$ the transition probability function such that $\sum_{t \in S} P(s,\alpha,t) \in \{0,1\}$ for all states $s \in S$ and actions $\alpha \in Act$. An action α is *enabled* in state $s \in S$ if $\sum_{t \in S} P(s,\alpha,t) = 1$. We define $Act(s) = \{\alpha \mid \alpha \text{ is enabled in } s\}$. A state t is *terminal* if $Act(t) = \emptyset$. A Markov chain (MC) is a special case of an MDP where Act is a singleton (we then write $P(s,u)$ rather than $P(s,\alpha,u)$). A *path* in an MDP $\mathcal{M}$ is a (finite or infinite) alternating sequence $\pi = s_0\, \alpha_0\, s_1\, \alpha_1\, s_2 \cdots \in (S \times Act)^* \cup (S \times Act)^\omega$ such that $P(s_i, \alpha_i, s_{i+1}) > 0$ for all indices i. A path is called maximal if it is infinite or finite and ends in a terminal state. An MDP can be interpreted as a Kripke structure in which transitions go from states to probability distributions over states.

A *(randomized) scheduler* $\mathfrak{S}$ is a function that maps each finite non-maximal path $s_0\alpha_0 \ldots \alpha_{n-1}s_n$ to a distribution over $Act(s_n)$. $\mathfrak{S}$ is called deterministic if $\mathfrak{S}(\pi)$ is a Dirac distribution for all finite non-maximal paths π. If the chosen action only depends on the last state of the path, $\mathfrak{S}$ is called *memoryless*. We write MR for the class of memoryless (randomized) and MD for the class of memoryless deterministic schedulers. *Finite-memory* schedulers are those that are representable by a finite-state automaton.

The scheduler $\mathfrak{S}$ of $\mathcal{M}$ induces a (possibly infinite) Markov chain. We write $\Pr^{\mathfrak{S}}_{\mathcal{M},s}$ for the standard probability measure on measurable sets of maximal paths in the Markov chain induced by $\mathfrak{S}$ with initial state s. If φ is a measurable set of maximal paths, then $\Pr^{\max}_{\mathcal{M},s}(\varphi)$ and $\Pr^{\min}_{\mathcal{M},s}(\varphi)$ denote the supremum resp. infimum of the probabilities for φ under all schedulers. We use the abbreviation $\Pr^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}_{\mathcal{M},\mathsf{init}}$ and notations $\Pr^{\max}_{\mathcal{M}}$ and $\Pr^{\min}_{\mathcal{M}}$ for extremal probabilities. Analogous notations will be used for expectations. So, if f is a random variable, then, e.g., $\mathrm{E}^{\mathfrak{S}}_{\mathcal{M}}(f)$ denotes the expectation of f under $\mathfrak{S}$ and

$E^{max}_{\mathcal{M}}(f)$ its supremum over all schedulers. We use LTL-like temporal modalities such as $\Diamond$ (eventually) and U (until) to denote path properties. For $X, T \subseteq S$ the formula $X\,U\,T$ is satisfied by paths $\pi = s_0 s_1 \ldots$ such that there exists $j \geqslant 0$ such that for all $i < j : s_i \in X$ and $s_j \in T$ and $\Diamond T = S\,U\,T$. It is well-known that $\Pr^{min}_{\mathcal{M}}(X\,U\,T)$ and $\Pr^{max}_{\mathcal{M}}(X\,U\,T)$ and corresponding optimal MD-schedulers are computable in polynomial time.

If $s \in S$ and $\alpha \in Act(s)$, then (s, α) is said to be a state-action pair of $\mathcal{M}$. An *end component* (EC) of an MDP $\mathcal{M}$ is a strongly connected sub-MDP containing at least one state-action pair. ECs will be often identified with the set of their state-action pairs. An EC $\mathcal{E}$ is called maximal (abbreviated MEC) if there is no proper superset $\mathcal{E}'$ of (the set of state-action pairs of) $\mathcal{E}$ which is an EC.

3 Strict and global probability-raising causes

We now provide formal definitions for cause-effect relations in MDPs which rely on the probability-raising (PR) principle as stated by (C1) and (C2) in the introduction. We focus on the case where both causes and effects are state properties, i.e., sets of states.

In the sequel, let $\mathcal{M} = (S, Act, P, \mathsf{init})$ be an MDP and $\mathsf{Eff} \subseteq S \setminus \{\mathsf{init}\}$ a nonempty set of terminal states. (As the effect set is fixed, for the analysis of cause-effect relationships in $\mathcal{M}$ it suffices to assume all effect states are terminal by (C2).) Furthermore, we may assume that every state $s \in S$ is reachable from init.

We consider here two variants of the probability-raising condition: the global setting treats the set Cause as a unit, while the strict view requires the probability-raising condition for all states in Cause individually.

Definition 1 (Global and strict probability-raising cause (GPR/SPR cause)). *Let* $\mathcal{M}$ *and* Eff *be as above and* Cause *a nonempty subset of* $S \setminus \mathsf{Eff}$*. Then,* Cause *is said to be a* GPR cause *for* Eff *iff the following two conditions (G) and (M) hold:*

(G) For each scheduler $\mathfrak{S}$ *where* $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{Cause}) > 0$*:*

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\ \Diamond \mathsf{Eff} \mid \Diamond \mathsf{Cause}\ ) \ > \ \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{Eff}). \tag{GPR}$$

(M) For each $c \in \mathsf{Cause}$*, there is a scheduler* $\mathfrak{S}$ *with* $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg \mathsf{Cause})\,U\,c) > 0$.

Cause *is called an* SPR cause *for* Eff *iff (M) and the following condition (S) hold:*

(S) For each state $c \in \mathsf{Cause}$ *and each scheduler* $\mathfrak{S}$ *where* $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg \mathsf{Cause})\,U\,c) > 0$*:*

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\ \Diamond \mathsf{Eff} \mid (\neg \mathsf{Cause})\,U\,c\ ) \ > \ \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{Eff}). \tag{SPR}$$

Condition (M) can be seen as a minimality requirement as states $c \in \mathsf{Cause}$ which are not accessible from init without traversing other states in Cause could be omitted without affecting the true positives (events where an effect state is reached after visiting a cause state, "covered effects") or false negatives (events where an effect state is reached without visiting a cause state before, "uncovered effect"). More concretely, whenever a set $C \subseteq S \setminus \mathsf{Eff}$ satisfies conditions (G) or (S) then the set Cause of states $c \in C$ where $\mathcal{M}$ has a path from init satisfying $(\neg C)\,U\,c$ is a GPR resp. an SPR cause.

3.1 Examples and simple properties of probability-raising causes

We first observe that SPR/GPR causes cannot contain the initial state init, since otherwise an equality instead of an inequality would hold in (GPR) and (SPR). Furthermore as a direct consequence of the definitions and using the equivalence of the LTL formulas $\Diamond\mathsf{Cause}$ and $(\neg\mathsf{Cause})\,\mathrm{U}\,\mathsf{Cause}$ we obtain:

Lemma 1 (Singleton PR causes). *If* Cause *is a singleton then* Cause *is a SPR cause for* Eff *if and only if* Cause *is a GPR cause for* Eff.

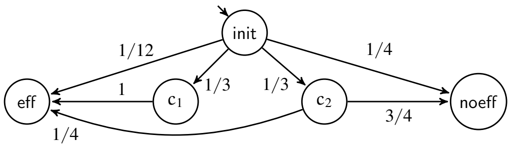
As the event $\Diamond\mathsf{Cause}$ is a disjoint union of all events $(\neg\mathsf{Cause})\,\mathrm{U}\,\mathsf{c}$ with $\mathsf{c} \in \mathsf{Cause}$, the probability for covered effects $\mathrm{Pr}^{\mathfrak{S}}_{\mathcal{M}}(\ \Diamond\mathsf{Eff} \mid \Diamond\mathsf{Cause}\ )$ is a weighted average of the probabilities $\mathrm{Pr}^{\mathfrak{S}}_{\mathcal{M}}(\ \Diamond\mathsf{Eff} \mid (\neg\mathsf{Cause})\,\mathrm{U}\,\mathsf{c}\ )$ for $\mathsf{c} \in \mathsf{Cause}$. This yields:

Lemma 2 (Strict implies global). *Every SPR cause for* Eff *is a GPR cause for* Eff.

Example 1 (Non-strict GPR cause). Consider the Markov chain $\mathcal{M}$ depicted below where the nodes represent states and the directed edges represent transitions labeled with their respective probabilities. Let $\mathsf{Eff} = \{\mathsf{eff}\}$. Then, $\mathrm{Pr}_{\mathcal{M}}(\Diamond\mathsf{Eff}) = \frac{1}{3} + \frac{1}{3}\cdot\frac{1}{4} + \frac{1}{12} = \frac{1}{2}$, $\mathrm{Pr}_{\mathcal{M}}(\Diamond\mathsf{Eff}|\Diamond\mathsf{c}_1) = \mathrm{Pr}_{\mathcal{M},\mathsf{c}_1}(\Diamond\mathsf{eff}) = 1$ and $\mathrm{Pr}_{\mathcal{M}}(\Diamond\mathsf{Eff}|\Diamond\mathsf{c}_2) = \mathrm{Pr}_{\mathcal{M},\mathsf{c}_2}(\Diamond\mathsf{eff}) = \frac{1}{4}$. Thus, $\{\mathsf{c}_1\}$ is both an SPR and a GPR cause for Eff, while $\{\mathsf{c}_2\}$ is not. The set $\mathsf{Cause} = \{\mathsf{c}_1, \mathsf{c}_2\}$ is a non-strict GPR cause for Eff as:

$$\mathrm{Pr}_{\mathcal{M}}(\ \Diamond\mathsf{Eff} \mid \Diamond\mathsf{Cause}\ ) = (\tfrac{1}{3} + \tfrac{1}{3}\cdot\tfrac{1}{4})/(\tfrac{1}{3} + \tfrac{1}{3}) = (\tfrac{5}{12})/(\tfrac{2}{3}) = \tfrac{5}{8} > \tfrac{1}{2} = \mathrm{Pr}_{\mathcal{M}}(\Diamond\mathsf{Eff}).$$

The second condition (M) is obviously fulfilled. Non-strictness follows from the fact that the SPR condition does not hold for state c_2. ◁



Example 2 (Probability-raising causes might not exist). PR causes might not exist, even if $\mathcal{M}$ is a Markov chain. This applies, e.g., to the Markov chain $\mathcal{M}$ with two states init and eff where $P(\mathsf{init}, \mathsf{eff}) = 1$ and the effect set $\mathsf{Eff} = \{\mathsf{eff}\}$. The only cause candidate is the singleton $\{\mathsf{init}\}$. However, the strict inequality in (GPR) or (SPR) does not hold for $\mathsf{Cause} = \{\mathsf{init}\}$. The same phenomenon occurs if all non-terminal states of a Markov chain reach the effect states with the same probability. In such cases, however, the non-existence of PR causes is well justified as the events $\Diamond\mathsf{Eff}$ and $\Diamond\mathsf{Cause}$ are stochastically independent for every set $\mathsf{Cause} \subseteq S \setminus \mathsf{Eff}$. ◁

Remark 1 (Memory needed for refuting PR condition). Let $\mathcal{M}$ be the MDP in Figure 1, where the notation is similar to Example 1 with the addition of actions α, β and γ. Let $\mathsf{Cause} = \{\mathsf{c}\}$ and $\mathsf{Eff} = \{\mathsf{eff}\}$. Only state s has a nondeterministic choice. Cause is not an PR cause. To see this, regard the deterministic scheduler $\mathfrak{T}$ that schedules β only for the first visit of s and α for the second visit of s. Then:

$$\mathrm{Pr}^{\mathfrak{T}}_{\mathcal{M}}(\Diamond\mathsf{eff}) \ = \ \tfrac{1}{2}\cdot\tfrac{1}{2} + \tfrac{1}{2}\cdot\tfrac{1}{2}\cdot 1\cdot\tfrac{1}{4} \ = \ \tfrac{5}{16} \ > \ \tfrac{1}{4} \ = \ \mathrm{Pr}^{\mathfrak{T}}_{\mathcal{M}}(\Diamond\mathsf{eff}|\Diamond\mathsf{c})$$

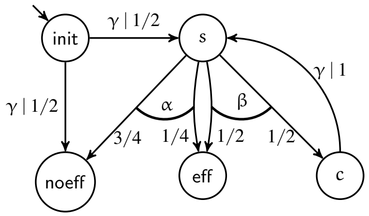


Fig. 1. MDP $\mathcal{M}$ from Remark 1

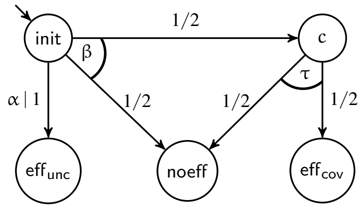


Fig. 2. MDP $\mathcal{M}$ from Remark 2

Denote the MR schedulers reaching c with positive probability as $\mathfrak{S}_\lambda$ with $\mathfrak{S}_\lambda(s)(\alpha) = \lambda$ and $\mathfrak{S}_\lambda(s)(\beta) = 1-\lambda$ for some $\lambda \in [0,1[$. Then, $\Pr^{\mathfrak{S}_\lambda}_{\mathcal{M},s}(\Diamond \mathsf{eff}) > 0$ and:

$$\Pr^{\mathfrak{S}_\lambda}_{\mathcal{M}}(\Diamond \mathsf{eff}) = \frac{1}{2} \cdot \Pr^{\mathfrak{S}_\lambda}_{\mathcal{M},s}(\Diamond \mathsf{eff}) < \Pr^{\mathfrak{S}_\lambda}_{\mathcal{M},s}(\Diamond \mathsf{eff}) = \Pr^{\mathfrak{S}_\lambda}_{\mathcal{M},c}(\Diamond \mathsf{eff}) = \Pr^{\mathfrak{S}_\lambda}_{\mathcal{M}}(\Diamond \mathsf{eff} | \Diamond c)$$

Thus, the SPR/GPR condition holds for Cause and Eff under all memoryless schedulers reaching Cause with positive probability, although Cause is not an PR cause. ◁

Remark 2 (Randomization needed for refuting PR condition). Consider the MDP $\mathcal{M}$ of Figure 2. Let $\mathsf{Eff} = \{\mathsf{eff}_{\mathsf{unc}}, \mathsf{eff}_{\mathsf{cov}}\}$ and $\mathsf{Cause} = \{c\}$. The two MD-schedulers $\mathfrak{S}_\alpha$ and $\mathfrak{S}_\beta$ that select α resp. β for the initial state init are the only deterministic schedulers. As $\mathfrak{S}_\alpha$ does not reach c, it is irrelevant for the SPR or GPR condition. $\mathfrak{S}_\beta$ satisfies (SPR) and (GPR) as $\Pr^{\mathfrak{S}_\beta}_{\mathcal{M}}(\Diamond \mathsf{Eff} | \Diamond c) = \frac{1}{2} > \frac{1}{4} = \Pr^{\mathfrak{S}_\beta}_{\mathcal{M}}(\Diamond \mathsf{Eff})$. The MR scheduler $\mathfrak{T}$ which selects α and β with probability $\frac{1}{2}$ in init reaches c with positive probability and violates (SPR) and (GPR) as $\Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond \mathsf{Eff} | \Diamond c) = \frac{1}{2} < \frac{5}{8} = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond \mathsf{Eff})$. ◁

Remark 3 (Cause-effect relations for regular classes of schedulers). The definitions of PR causes in MDPs impose constraints for all schedulers reaching a cause state. This condition is fairly strong and might lead to the phenomenon that no PR cause exists. However, replacing $\mathcal{M}$ with an MDP resulting from the synchronous parallel composition of $\mathcal{M}$ with a deterministic finite automaton representing a regular constraint on the scheduled state-action sequences (e.g., "alternate between actions α and β in state s" or "take α on every third visit to state s and actions β or γ otherwise") leads to a weaker notion of PR causality. This can be useful to obtain more detailed information on cause-effect relationships in special scenarios. For example at design time where multiple scenarios (regular classes of schedulers) are considered or for a post-hoc analysis. For the later, one seeks causes of an occurred effect and the information about the scheduled actions is either extractable from log files or gathered by a monitor. ◁

Remark 4 (Action causality and other forms of PR causality). Our notions of PR causes are purely state-based with conditions that compare probabilities under the same scheduler. However, in combination with model transformations, the proposed notions are also applicable for reasoning about other forms of PR causality.

Suppose, the task is to check whether taking action α in state s raises the effect probabilities compared to never scheduling α in state s. Let $\mathcal{M}_0$ and $\mathcal{M}_1$ be copies of $\mathcal{M}$ with the following modifications: In $\mathcal{M}_0$, the only enabled action of state s is α, while in $\mathcal{M}_1$ the enabled actions of state s are the elements of $Act_{\mathcal{M}}(s) \setminus \{\alpha\}$. Let now $\mathcal{N}$ be the MDP whose initial state has a single enabled action and moves with probability $1/2$ to $\mathcal{M}_0$ and $\mathcal{M}_1$. Then, action α raises the effect probability in $\mathcal{M}$ iff the initial state of $\mathcal{M}_0$ consitutes an SPR cause in $\mathcal{N}$. This idea can be generalized to check whether scheduler classes satisfying a regular constraint have higher effect probability compared to all other schedulers. In this case, we can deal with an MDP $\mathcal{N}$ as above where $\mathcal{M}_0$ and $\mathcal{M}_1$ are defined as the synchronous product of deterministic finite automata and $\mathcal{M}$. $\lhd$

3.2 Related work

Previous work in the direction of probabilistic causation in stochastic operational models has mainly concentrated on Markov chains. Kleinberg [24,25] introduced *prima facie causes* in finite Markov chains where both causes and effects are formalized as PCTL state formulae, and thus they can be seen as sets of states as in our approach. The correspondence of Kleinberg's PCTL constraints for prima facie causes and the strict probability-raising condition formalized using conditional probabilities has been worked out in the survey article [5]. Our notion of SPR causes corresponds to Kleinberg's prima facie causes, except for the minimality condition (M). Ábrahám et al [1] introduces a hyperlogic for Markov chains and gives a formalization of probabilistic causation in Markov chains as a hyperproperty, which is consistent with Kleinberg's prima facie causes, and with SPR causes up to minimality. Cause-effect relations in Markov chains where effects are ω-regular properties have been introduced in [6]. The notion relies on the strict probability-raising condition, but requires completeness in the sense that every path where the effect occurs has a prefix in the cause set. The paper [6] permits a non-strict inequality in the SPR condition with the consequence that causes always exist, which is not the case for our notions.

The survey article [5] introduces notions of global probability-raising causes for Markov chains where causes and effects can be path properties. [5]'s notion of *reachability causes* in Markov chains directly corresponds to our notion GPR causes, the only difference being that [5] deals with a relaxed minimality condition and requires that the cause set is reachable without visiting an effect state before. The latter is inherent in our approach as we suppose that all states are reachable and the effect states are terminal.

To the best of our knowledge, probabilistic causation in MDPs has not been studied before. The only work in this direction we are aware of is the recent paper by Dimitrova et al [17] on a hyperlogic, called PHL, for MDPs. While the paper focuses on the foundation of PHL, it contains an example illustrating how action causality can be formalized as a PHL formula. Roughly, the presented formula expresses that taking a specific action α increases the probability for reaching effect states. Thus, it also relies on the probability-raising principle, but compares the "effect probabilities" under different schedulers (which either schedule α or not) rather than comparing probabilities under the same scheduler as in our PR condition. However, as Remark 4 argues, to some extent our notions of PR causes can reason about action causality as well.

There has also been work on causality-based explanations of counterexamples in probabilistic models [27,28]. The underlying causality notion of this work, however, relies on the non-probabilistic counterfactual principle rather than the probability-raising condition. The same applies to the notions of forward and backward responsibility in stochastic games in extensive form introduced in the recent work [7].

4 Checking the existence of PR causes and the PR conditions

We now turn to algorithms for checking whether a given set Cause is an SPR or GPR cause for Eff. As condition (M) of SPR and GPR causes is verifiable by standard model checking techniques in polynomial time, we concentrate on checking the probability-raising conditions (SPR) and (GPR). For Markov chains, both (SPR) and (GPR) can be checked in polynomial time by computing the corresponding probabilities. So, the interesting case is checking the PR conditions in MDPs.

We start by stating that for the SPR and GPR condition, it suffices to consider schedulers minimizing the probability to reach an effect state from every cause state.

Notation 1 (MDP with minimal effect probabilities from cause candidates). If $C \subseteq S$ then we write $\mathcal{M}_{[C]}$ for the MDP resulting from $\mathcal{M}$ by removing all enabled actions of the states in C. Instead, $\mathcal{M}_{[C]}$ has a new action γ that is enabled exactly in the states $s \in C$ with the transition probabilities $P_{\mathcal{M}_{[C]}}(s,\gamma,\mathsf{eff}) = \Pr^{\min}_{\mathcal{M},s}(\Diamond \mathsf{Eff})$ and $P_{\mathcal{M}_{[C]}}(s,\gamma,\mathsf{noeff}) = 1 - \Pr^{\min}_{\mathcal{M},s}(\Diamond \mathsf{Eff})$. Here, eff is a fixed state in Eff and noeff a (possibly fresh) terminal state not in Eff. We write $\mathcal{M}_{[c]}$ if $C = \{c\}$ is a singleton.

Lemma 3. *Let $\mathcal{M} = (S, Act, P, \mathsf{init})$ be an MDP and $\mathsf{Eff} \subseteq S$ a set of terminal states. Let $\mathsf{Cause} \subseteq S \setminus \mathsf{Eff}$. Then, Cause is an SPR cause (resp. a GPR cause) for Eff in $\mathcal{M}$ if and only if Cause is an SPR cause (resp. a GPR cause) for Eff in $\mathcal{M}_{[\mathsf{Cause}]}$.*

4.1 Checking the strict probability-raising condition and the existence of causes

The basis of both checking the existence of PR causes or checking the SPR condition for a given cause candidate is the following polynomial time algorithm to check whether the SPR condition holds in a given state c of $\mathcal{M}$ for all schedulers $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond c) > 0$:

Algorithm 2. Input: state $c \in S$, set of terminal states $\mathsf{Eff} \subseteq S$.
Task: Decide whether (SPR) holds in c for all schedulers $\mathfrak{S}$.
Compute $w_c = \Pr^{\min}_{\mathcal{M},c}(\Diamond \mathsf{Eff})$ and $q_s = \Pr^{\max}_{\mathcal{M}_{[c]},s}(\Diamond \mathsf{Eff})$ for each state s in $\mathcal{M}_{[c]}$.

1. If $q_{\mathsf{init}} < w_c$, then return "yes, (SPR) holds for c".
2. If $q_{\mathsf{init}} > w_c$, then return "no, (SPR) does not hold for c".
3. Suppose $q_{\mathsf{init}} = w_c$. Let $A(s) = \{\alpha \in Act_{\mathcal{M}_{[c]}}(s) \mid q_s = \sum_{t \in S_{[c]}} P_{\mathcal{M}_{[c]}}(s,\alpha,t) \cdot q_t\}$ for each non-terminal state s. Let $\mathcal{M}^{\max}_{[c]}$ denote the sub-MDP of $\mathcal{M}_{[c]}$ induced by the state-action pairs (s,α) where $\alpha \in A(s)$.
   3.1 If c is reachable from init in $\mathcal{M}^{\max}_{[c]}$, then return "no, (SPR) does not hold for c".
   3.2 If c is not reachable from init in $\mathcal{M}^{\max}_{[c]}$, then return "yes, (SPR) holds for c".

Lemma 4. *Algorithm 2 is sound and runs in polynomial time.*

Soundness. Let $\mathcal{N} = \mathcal{M}_{[c]}$. Soundness is obvious in case 1. For case 2, consider a real number λ with $1 > \lambda > \frac{w_c}{q_{init}}$ and MD-schedulers $\mathfrak{T}$ and $\mathfrak{S}$ realizing $\Pr^{\mathfrak{T}}_{\mathcal{N},s}(\Diamond \mathsf{Eff}) = q_s$ and $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond c) > 0$ for all states s. We can combine $\mathfrak{T}$ and $\mathfrak{S}$ to a new MR-scheduler $\mathfrak{U}$ with the property that $\Pr^{\mathfrak{U}}_{\mathcal{N}}(\Diamond t) = \lambda \Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond t) + (1-\lambda)\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond t)$ for all terminal states t and for $t = c$. Then, $\mathfrak{U}$ witnesses a violation of (SPR). For case 3.1 consider an MD-scheduler $\mathfrak{S}$ of $\mathcal{M}^{max}_{[c]}$ where c is reachable from init via a $\mathfrak{S}$-path and $\Pr^{\mathfrak{S}}_{\mathcal{N},s}(\Diamond \mathsf{Eff}) = q_s$ for all states s. Then, (SPR) does not hold for c in the scheduler $\mathfrak{S}$. In case 3.2 we have $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond c) = 0$ for all schedulers $\mathfrak{S}$ for $\mathcal{N}$ with $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond \mathsf{Eff}) = q_{init} = w_c$. But then $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond c) > 0$ implies $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond \mathsf{Eff}) < w_c$ as required in (SPR). □

By applying Algorithm 2 to all states $c \in \mathsf{Cause}$ and standard algorithms to check the existence of a path satisfying $(\neg \mathsf{Cause})\, \mathrm{U}\, c$ for every state $c \in \mathsf{Cause}$, we obtain:

Theorem 3 (Checking SPR causes). *The problem "given $\mathcal{M}$, Cause and Eff, check whether Cause is a SPR cause for Eff in $\mathcal{M}$" is solvable in polynomial-time.*

Remark 5 (Memory requirements for refuting the SPR property). As the soundness proof for Algorithm 2 shows: If Cause does not satisfy the SPR condition, then there is an MR-scheduler $\mathfrak{S}$ for $\mathcal{M}_{[\mathsf{Cause}]}$ witnessing the violation of (SPR). Scheduler $\mathfrak{S}$ corresponds to a finite-memory (randomized) scheduler $\mathfrak{T}$ with two memory cells for $\mathcal{M}$: "before Cause" (where $\mathfrak{T}$ behaves as $\mathfrak{S}$) and "after Cause" (where $\mathfrak{T}$ behaves as an MD-scheduler minimizing the effect probability form every state). ◁

Lemma 5 (Criterion for the existence of probability-raising causes). *Let $\mathcal{M}$ be an MDP and Eff a nonempty set of states. Then Eff has an SPR cause in $\mathcal{M}$ iff Eff has a GPR cause in $\mathcal{M}$ iff there is a state $c_0 \in S \setminus \mathsf{Eff}$ such that the singleton $\{c_0\}$ is an SPR cause (and therefore a GRP cause) for Eff in $\mathcal{M}$. In particular, the existence of SPR/GPR causes can be checked with Algorithm 2 in polynomial time.*

4.2 Checking the global probability-raising condition

Theorem 4. *The problem "given $\mathcal{M}$, Cause and Eff, check whether Cause is a GPR cause for Eff in $\mathcal{M}$" is solvable in polynomial space.*

In order to provide an algorithm, we perform a model transformation after which the violation of (GPR) by a scheduler $\mathfrak{S}$ can be expressed solely in terms of the expected frequencies of the state-action pairs of the transformed MDP under $\mathfrak{S}$. This allows us to express the existence of a scheduler witnessing the non-causality of Cause in terms of the satisfiability of a quadratic constraint system. Then we can restrict the quantification in (G) to MR-schedulers in the transformed model. We trace back the memory requirements to $\mathcal{M}_{[\mathsf{Cause}]}$ and to the original MDP $\mathcal{M}$ yielding the second main result. Still, memory can be necessary to witness non-causality (Remark 1).

Theorem 5. *Let $\mathcal{M}$ be an MDP with effect set Eff as before and Cause a set of non-effect states such that condition (M) holds. If Cause is not a GPR cause for Eff, then there is an MR-scheduler for $\mathcal{M}_{[\mathsf{Cause}]}$ refuting the GPR condition for Cause in $\mathcal{M}_{[\mathsf{Cause}]}$ and a finite-memory scheduler for $\mathcal{M}$ with two memory cells refuting the GPR condition for Cause in $\mathcal{M}$.*

The remainder of this section is concerned with the proofs of Theorem 4 and Theorem 5. We suppose that both the effect set Eff and the cause candidate Cause are fixed disjoint subsets of the state space of the MDP $\mathcal{M}$ and that Cause satisfies (M).

Checking the GPR condition (Proof of Theorem 4). The first step is a polynomial-time model transformation which permits to make the following assumptions when checking the GPR condition of Cause for Eff.

- **(A1)** $\mathsf{Eff} = \{\mathsf{eff}_{\mathsf{unc}}, \mathsf{eff}_{\mathsf{cov}}\}$ consists of two terminal states.
- **(A2)** For every state $\mathsf{c} \in \mathsf{Cause}$, there is only a single enabled action, say $Act(\mathsf{c}) = \{\gamma\}$, and there exists $w_{\mathsf{c}} \in [0,1] \cap \mathbb{Q}$ such that $\mathsf{P}(\mathsf{c},\gamma,\mathsf{eff}_{\mathsf{cov}}) = w_{\mathsf{c}}$ and $\mathsf{P}(\mathsf{c},\gamma,\mathsf{noeff}_{\mathsf{fp}}) = 1-w_{\mathsf{c}}$ where $\mathsf{noeff}_{\mathsf{fp}}$ is a terminal non-effect state and $\mathsf{noeff}_{\mathsf{fp}}$ and $\mathsf{eff}_{\mathsf{cov}}$ are only accessible via the γ-transition from the states $\mathsf{c} \in \mathsf{Cause}$.
- **(A3)** $\mathcal{M}$ has no end components and there is a further terminal state $\mathsf{noeff}_{\mathsf{tn}}$ and an action τ such that $\tau \in Act(\mathsf{s})$ implies $\mathsf{P}(\mathsf{s},\tau,\mathsf{noeff}_{\mathsf{tn}}) = 1$.

Intuitively, $\mathsf{eff}_{\mathsf{cov}}$ stands for covered effects ("Eff after Cause") and can be seen as a true positive, while $\mathsf{eff}_{\mathsf{unc}}$ represents the uncovered effects ("Eff without preceding Cause") and corresponds to a false negative. Let $\mathfrak{S}$ be a scheduler in $\mathcal{M}$. Note that $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg\mathsf{Cause})\,\mathrm{U}\,\mathsf{Eff}) = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{eff}_{\mathsf{unc}})$ and $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond(\mathsf{Cause} \wedge \Diamond\mathsf{Eff})) = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{eff}_{\mathsf{cov}})$. As the cause states can not reach each other we also have $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg\mathsf{Cause})\,\mathrm{U}\,\mathsf{c}) = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{c})$ for each $\mathsf{c} \in \mathsf{Cause}$. The intuitive meaning of $\mathsf{noeff}_{\mathsf{fp}}$ is a false positive ("no effect after Cause"), while $\mathsf{noeff}_{\mathsf{tn}}$ stands for true negatives where neither the effect nor the cause is observed. Note that $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond(\mathsf{Cause} \wedge \neg\Diamond\mathsf{Eff})) = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{noeff}_{\mathsf{fp}})$ and $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond\mathsf{Cause} \wedge \neg\Diamond\mathsf{Eff})) = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{noeff}_{\mathsf{tn}})$.

Justification of assumptions (A1)-(A3): We justify the assumptions as we can transform $\mathcal{M}$ into a new MDP of the same asymptotic size satisfying the above assumptions. Thanks to Lemma 3, we may suppose that $\mathcal{M} = \mathcal{M}_{[\mathsf{Cause}]}$ (see Notation 1) without changing the satisfaction of the GPR condition. We then may rename the effect state eff and the non-effect state noeff reachable from Cause into $\mathsf{eff}_{\mathsf{cov}}$ and $\mathsf{noeff}_{\mathsf{fp}}$, respectively. Furthermore, we collapse all other effect states into a single state $\mathsf{eff}_{\mathsf{unc}}$ and all true negative states into $\mathsf{noeff}_{\mathsf{tn}}$. Similarly, by renaming and possibly duplicating terminal states we also suppose that $\mathsf{noeff}_{\mathsf{fp}}$ has no other incoming transitions than the γ-transitions from the states in Cause. This ensures (A1) and (A2). For (A3) consider the set T of terminal states in the MDP obtained so far. We remove all end components by switching to the MEC-quotient [2], i.e., we collapse all states that belong to the same MEC $\mathcal{E}$ into a single state $\mathsf{s}_{\mathcal{E}}$ while ignoring the actions inside $\mathcal{E}$. Additionally, we add a fresh τ-transition from the states $\mathsf{s}_{\mathcal{E}}$ to $\mathsf{noeff}_{\mathsf{tn}}$ (i.e., $\mathsf{P}(\mathsf{s}_{\mathcal{E}},\tau,\mathsf{noeff}_{\mathsf{tn}}) = 1$). The τ-transitions from states $\mathsf{s}_{\mathcal{E}}$ to $\mathsf{noeff}_{\mathsf{tn}}$ mimic cases where schedulers of the original MDP eventually enter an end component and stay there forever with positive probability.

With assumptions (A1)-(A3), the GPR condition can be reformulated as follows:

Lemma 6. *Under assumptions (A1)-(A3),* Cause *satisfies the GPR condition if and only if for each scheduler* $\mathfrak{S}$ *with* $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond\mathsf{Cause}) > 0$ *the following condition holds:*

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond\mathsf{Cause}) \cdot \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond\mathsf{eff}_{\mathsf{unc}}) \;<\; \big(1-\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond\mathsf{Cause})\big) \cdot \sum_{\mathsf{c}\in\mathsf{Cause}} \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond\mathsf{c}) \cdot w_{\mathsf{c}} \qquad \text{(GPR-1)}$$

With assumptions (A1)-(A3), a terminal state of $\mathcal{M}$ is reached almost surely under any scheduler after finitely many steps in expectation. Given a scheduler $\mathfrak{S}$ for $\mathcal{M}$, the expected frequencies (i.e., expected number of occurrences in maximal paths) of state action-pairs (s, α), states $s \in S$ and state-sets $T \subseteq S$ under $\mathfrak{S}$ are defined by:

$$freq_{\mathfrak{S}}(s,\alpha) \stackrel{\text{def}}{=} \mathrm{E}^{\mathfrak{S}}_{\mathcal{M}}(\text{number of visits to } s \text{ in which } \alpha \text{ is taken})$$
$$freq_{\mathfrak{S}}(s) \stackrel{\text{def}}{=} \sum\nolimits_{\alpha\in Act(s)} freq_{\mathfrak{S}}(s,\alpha), \qquad freq_{\mathfrak{S}}(T) \stackrel{\text{def}}{=} \sum\nolimits_{s\in T} freq_{\mathfrak{S}}(s).$$

Let T be one of the sets $\{\mathsf{eff}_{\mathsf{cov}}\}$, $\{\mathsf{eff}_{\mathsf{unc}}\}$, Cause, or a singleton $\{c\}$ with $c \in \mathsf{Cause}$. As T is visited at most once during each run of $\mathcal{M}$ (assumptions (A1) and (A2)), we have $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond T) = freq_{\mathfrak{S}}(T)$ for each scheduler $\mathfrak{S}$. This allows us to express the violation of the GPR condition in terms of a quadratic constraint system over variables for the expected frequencies of state-action pairs in the following way:

Let *StAct* denote the set of state-action pairs in $\mathcal{M}$. We consider the following constraint system over the variables $x_{s,\alpha}$ for each $(s,\alpha) \in StAct$ where we use the short form notation $x_s = \sum_{\alpha\in Act(s)} x_{s,\alpha}$:

$$x_{s,\alpha} \geqslant 0 \qquad \text{for all } (s,\alpha) \in StAct \tag{1}$$
$$x_{\mathsf{init}} = 1 + \sum_{(t,\alpha)\in StAct} x_{t,\alpha} \cdot P(t,\alpha,\mathsf{init}) \tag{2}$$
$$x_s = \sum_{(t,\alpha)\in StAct} x_{t,\alpha} \cdot P(t,\alpha,s) \qquad \text{for all } s \in S\setminus\{\mathsf{init}\} \tag{3}$$

Using well-known results for MDPs without ECs (see, e.g., [23, Theorem 9.16]), given a vector $x \in \mathbb{R}^{StAct}$, then x is a solution to (1) and the balance equations (2) and (3) if and only if there is a (possibly history-dependent) scheduler $\mathfrak{S}$ for $\mathcal{M}$ with $x_{s,\alpha} = freq_{\mathfrak{S}}(s,\alpha)$ for all $(s,\alpha) \in StAct$ if and only if there is an MR-scheduler $\mathfrak{S}$ for $\mathcal{M}$ with $x_{s,\alpha} = freq_{\mathfrak{S}}(s,\alpha)$ for all $(s,\alpha) \in StAct$.

The violation of (GPR-1) in Lemma 6 and the condition $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{Cause}) > 0$ can be reformulated in terms of the frequency-variables as follows where x_{Cause} is an abbreviation for $\sum_{c\in\mathsf{Cause}} x_c$:

$$x_{\mathsf{Cause}} \cdot x_{\mathsf{eff}_{\mathsf{unc}}} \geqslant (1 - x_{\mathsf{Cause}}) \cdot \sum_{c\in\mathsf{Cause}} x_c \cdot w_c \tag{4}$$
$$x_{\mathsf{Cause}} > 0 \tag{5}$$

Lemma 7. *Under assumptions (A1)-(A3), the set* Cause *is not a GPR cause for* Eff *in* $\mathcal{M}$ *iff the constructed quadratic system of inequalities (1)-(5) has a solution.*

Proof of Theorem 4. The existence of a solution to the quadratic system of inequalities (Lemma 7) can straight-forwardly be formulated as a sentence in the language of the existential theory of the reals. The system of inequalities can be constructed from $\mathcal{M}$, Cause, and Eff in polynomial time. Its solvability is decidable in polynomial space as the decision problem of the existential theory of the reals is in PSPACE [13]. □

Memory requirements of schedulers in the original MDP (Proof of Theorem 5). As stated above, every solution to the linear system of inequalities (1), (2), and (3) corresponds to expected frequencies of state-action pairs of an MR-scheduler in the transformed model satisfying (A1)-(A3). Hence:

Corollary 1. *Under assumptions (A1)-(A3),* Cause *is no GPR cause for* Eff *iff there exists an MR-scheduler* $\mathfrak{T}$ *with* $\Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond \mathsf{Cause}) > 0$ *violating the GPR condition.*

The model transformation we used for assumptions (A1)-(A3), however, does affect the memory requirements of schedulers. We may further restrict the MR-schedulers necessary to witness non-causality under assumptions (A1)-(A3). For the following lemma, recall that τ is the action of the MEC quotient used for the extra transition from states representing MECs to a new trap state (see also assumption (A3)).

Lemma 8. *Assume (A1)-(A3). Given an MR-scheduler* $\mathfrak{U}$ *with* $\Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond \mathsf{Cause}) > 0$ *that violates* (GPR)*, an MR-scheduler* $\mathfrak{T}$ *with* $\mathfrak{T}(s)(\tau) \in \{0,1\}$ *for each state* s *with* $\tau \in Act(s)$ *that satifies* $\Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond \mathsf{Cause}) > 0$ *and violates* (GPR) *is computable in polynomial time.*

The condition that τ only has to be scheduled with probability 0 or 1 in each state is the key to transfer the sufficiency of MR-schedulers to the MDP $\mathcal{M}_{[\mathsf{Cause}]}$. This fact is of general interest as well and stated in the following theorem where τ again is the action added to move from a state $s_{\mathcal{E}}$ to the new trap state in the MEC-quotient.

Theorem 6. *Let* $\mathcal{M}$ *be an MDP with pairwise disjoint action sets for all states. Then, for each MR-scheduler* $\mathfrak{S}$ *for the MEC-quotient of* $\mathcal{M}$ *with* $\mathfrak{S}(s_{\mathcal{E}})(\tau) \in \{0,1\}$ *for each MEC* $\mathcal{E}$ *of* $\mathcal{M}$ *there is an MR-scheduler* $\mathfrak{T}$ *for* $\mathcal{M}$ *such that every action* α *of* $\mathcal{M}$ *that does not belong to an MEC of* $\mathcal{M}$*, has the same expected frequency under* $\mathfrak{S}$ *and* $\mathfrak{T}$*.*

Proof sketch. The crux are cases where $\mathfrak{S}(s_{\mathcal{E}})(\tau) = 0$, which requires to traverse the MEC $\mathcal{E}$ of $\mathcal{M}$ in a memoryless way such that all actions leaving $\mathcal{E}$ have the same expected frequency under $\mathfrak{T}$ and $\mathfrak{S}$. First, we construct a finite-memory scheduler $\mathfrak{T}'$ that always leaves each such end component according to the distribution given by $\mathfrak{S}(s_{\mathcal{E}})$. By [23, Theorem 9.16], we then conclude that there is an MR-scheduler $\mathfrak{T}$ under which the expected frequencies of all state-action pairs are the same as under $\mathfrak{T}'$. □

Proof of Theorem 5. The model transformation establishing assumptions (A1)-(A3) results in the MEC-quotient of $\mathcal{M}_{[\mathsf{Cause}]}$ up to the renaming and collapsing of terminal states. By Corollary 1 and Theorem 6, we conclude that Cause is not a GPR cause for Eff in $\mathcal{M}$ iff there is a MR-scheduler $\mathfrak{S}$ for $\mathcal{M}_{[\mathsf{Cause}]}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}_{[\mathsf{Cause}]}}(\Diamond \mathsf{Cause}) > 0$ that violates (GPR). As in Remark 5, $\mathfrak{S}$ can be extended to a finite-memory randomized scheduler $\mathfrak{T}$ for $\mathcal{M}$ with two memory cells. □

Remark 6 (On lower bounds on GPR checking). Solving systems of quadratic inequalities with linear side constraints is NP-hard in general (see, e.g., [20]). For convex problems, in which the associated symmetric matrix in the quadratic inequality has only non-negative eigenvalues, the problem is, however, solvable in polynomial time [26]. Unfortunately, the quadratic constraint system given by (1)-(5) is not of this form. Even if Cause is a singleton $\{c\}$ and the variable $x_{\mathsf{eff}_{\mathsf{unc}}}$ is forced to take a constant value y by (1)-(3), i.e., by the structure of the MDP, the inequality (4) takes the form:

$$x_c \cdot w_c - x_c^2 \cdot (w_c + y) \leqslant 0 \qquad (*)$$

Here, the 1×1-matrix $(-w_c - y)$ has a negative eigenvalue. Although it is not ruled out that (1)-(5) belongs to another class of efficiently solvable constraint systems, the NP-hardness result in [33] for the solvability of quadratic inequalities of the form (*) with linear side constraints might be an indication for the computational difficulty. $\lhd$

5 Quality and optimality of causes

The goal of this section is to identify notions that measure how "good" causes are and to present algorithms to determine good causes according to proposed quality measures. We have seen so far that small (singleton) causes are easy to determine (see Section 4.1). Moreover, it is easy to see that the proposed existence-checking algorithm can be transformed such that it returns a singleton (strict or global) probability-raising cause $\{c_0\}$ with maximal *precision*, i.e., a state c_0 where $\inf_{\mathfrak{S}} \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{Eff} | \Diamond c_0) = \Pr^{\min}_{\mathcal{M},c_0}(\Diamond \mathsf{Eff})$ is maximal. On the other hand, singleton or small cause sets might have poor coverage in the sense that the probability of paths which reach an effect state without visiting a cause state before ("uncovered effects") can be large. This motivates the consideration of quality notions for causes that incorporate how well effect scenarios are covered. We take inspiration of quality measures that are considered in statistical analysis (see e.g. [36]). This includes the *recall* as a measure for the relative coverage (proportion of covered effects among all effect scenarios), the *coverage ratio* (quotient of covered and uncovered effects) as well as the *f-score*. The f-score is a standard measure for classifiers defined by the harmonic mean of precision and recall. It can be seen as a compromise to achieve both good precision and good recall.

Throughout this section, we assume as before an MDP $\mathcal{M} = (S, Act, \mathsf{P}, \mathsf{init})$ and a set $\mathsf{Eff} \subseteq S$ are given where all effect states are terminal. Furthermore, we suppose that all states $s \in S$ are reachable from init.

5.1 Quality measures for causes

In statistical analysis, the precision of a classifier with binary outcomes ("positive" or "negative") is defined as the ratio of all true positives among all positively classified elements, while its recall is defined as the ratio of all true positives among all actual positive elements. Translated to our setting, we consider classifiers induced by a given cause set Cause that return "positive" for sample paths in case that a cause state is visited and "negative" otherwise. The intuitive meaning of true positives and false negatives is as explained after Definition 1. The meaning of true negatives and false positives is analogous. We use $\mathsf{tp}^{\mathfrak{S}}$ for the probability for true positives under $\mathfrak{S}$. The notations $\mathsf{fp}^{\mathfrak{S}}$, $\mathsf{fn}^{\mathfrak{S}}$, $\mathsf{tn}^{\mathfrak{S}}$ have analogous meanings.

With this interpretation of causes as binary classifiers in mind, the recall and precision and coverage ratio of a cause set Cause *under a scheduler* $\mathfrak{S}$ is defined as follows (assuming $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{Eff}) > 0$ resp. $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{Cause}) > 0$ resp. $\Pr^{\mathfrak{S}}_{\mathcal{M}}\big((\neg \mathsf{Cause})\, \mathrm{U}\, \mathsf{Eff}\big) > 0$):

$$\begin{aligned} precision^{\mathfrak{S}}(\mathsf{Cause}) &= \Pr^{\mathfrak{S}}_{\mathcal{M}}(\, \Diamond \mathsf{Eff} \mid \Diamond \mathsf{Cause}\,) = \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}} + \mathsf{fp}^{\mathfrak{S}}} \\ recall^{\mathfrak{S}}(\mathsf{Cause}) &= \Pr^{\mathfrak{S}}_{\mathcal{M}}(\, \Diamond \mathsf{Cause} \mid \Diamond \mathsf{Eff}\,) = \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}} + \mathsf{fn}^{\mathfrak{S}}} \end{aligned}$$

$$covrat^{\mathfrak{S}}(\mathsf{Cause}) = \frac{\Pr^{\mathfrak{S}}_{\mathcal{M}}\big(\lozenge(\mathsf{Cause} \wedge \lozenge \mathsf{Eff})\big)}{\Pr^{\mathfrak{S}}_{\mathcal{M}}\big((\neg \mathsf{Cause})\,\mathrm{U}\,\mathsf{Eff}\big)} = \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{fn}^{\mathfrak{S}}}.$$

For the coverage ratio, if $\Pr^{\mathfrak{S}}_{\mathcal{M}}\big((\neg \mathsf{Cause})\,\mathrm{U}\,\mathsf{Eff}\big) = 0$ and $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{Cause}) > 0$ we define $covrat^{\mathfrak{S}}(\mathsf{Cause}) = +\infty$. Finally, the f-score of Cause *under a scheduler* $\mathfrak{S}$ is defined as the harmonic mean of the precision and recall (assuming $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{Cause}) > 0$, which implies $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{Eff}) > 0$ as Cause is a PR cause):

$$fscore^{\mathfrak{S}}(\mathsf{Cause}) \stackrel{\text{def}}{=} 2 \cdot \frac{precision^{\mathfrak{S}}(\mathsf{Cause}) \cdot recall^{\mathfrak{S}}(\mathsf{Cause})}{precision^{\mathfrak{S}}(\mathsf{Cause}) + recall^{\mathfrak{S}}(\mathsf{Cause})}$$

If, however, $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{Eff}) > 0$ and $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{Cause}) = 0$ we define $fscore^{\mathfrak{S}}(\mathsf{Cause}) = 0$.

Quality measures for cause sets. Let Cause be a PR cause. The recall of Cause measures the relative coverage in terms of the worst-case conditional probability for covered effects (true positives) among all scenarios where the effect occurs.

$$recall(\mathsf{Cause}) = \inf_{\mathfrak{S}} recall^{\mathfrak{S}}(\mathsf{Cause}) = \Pr^{\min}_{\mathcal{M}}(\ \lozenge \mathsf{Cause} \mid \lozenge \mathsf{Eff}\ )$$

when ranging over all schedulers $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{Eff}) > 0$. Likewise, the coverage ratio and f-score of Cause are defined by the worst-case coverage ratio resp. f-score (when ranging over schedulers for which $covrat^{\mathfrak{S}}(\mathsf{Cause})$ resp. $fscore^{\mathfrak{S}}(\mathsf{Cause})$ is defined):

$$covrat(\mathsf{Cause}) = \inf_{\mathfrak{S}} covrat^{\mathfrak{S}}(\mathsf{Cause}), \quad fscore(\mathsf{Cause}) = \inf_{\mathfrak{S}} fscore^{\mathfrak{S}}(\mathsf{Cause})$$

5.2 Computation schemes for the quality measures for fixed cause set

For this section, we assume a fixed PR cause Cause is given and address the problem to compute its quality values. Since all quality measures are preserved by the switch from $\mathcal{M}$ to $\mathcal{M}_{[\mathsf{Cause}]}$ as well as the transformations of $\mathcal{M}_{[\mathsf{Cause}]}$ to an MDP that satisfies conditions (A1)-(A3) of Section 4.2, we may assume that $\mathcal{M}$ satisfies (A1)-(A3).

While efficient computation methods for $recall(\mathsf{Cause})$ are known from literature (see [10,31] for poly-time algorithms to compute conditional reachability probabilities), we are not aware of known concepts that are applicable for computing the coverage ratio or the f-score. Indeed, both are efficiently computable:

Theorem 7. *The values* $covrat(\mathsf{Cause})$ *and* $fscore(\mathsf{Cause})$ *and corresponding worst-case schedulers are computable in polynomial time.*

By definition, the value $covrat(\mathsf{Cause})$ is the infimum over a quotient of reachability probabilities for disjoint sets of terminal states. While this is not the case for the f-score, we can express $fscore(\mathsf{Cause})$ in terms of the supremum of such a quotient. More precisely, under assumptions (A1)-(A3) and assuming $fscore(\mathsf{Cause}) > 0$, we have:

$$fscore(\mathsf{Cause}) = \frac{2}{\mathsf{X}+2} \quad \text{where} \quad \mathsf{X} = \sup_{\mathfrak{S}} \frac{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{noeff}_{\mathsf{fp}}) + \Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{eff}_{\mathsf{unc}})}{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{eff}_{\mathsf{cov}})}$$

where $\mathfrak{S}$ ranges over all schedulers with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{eff}_{\mathsf{cov}}) > 0$. Furthermore, we have $fscore(\mathsf{Cause}) = 0$ if and only if $recall(\mathsf{Cause}) = 0$ if and only if there exists a scheduler $\mathfrak{S}$ satisfying $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{Eff}) > 0$ and $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathsf{Cause}) = 0$.

So, the remaining task to prove Theorem 7 is a generally applicable technique for computing extremal ratios of reachability probabilities in MDPs without ECs.

Max/min ratios of reachability probabilities for disjoint sets of terminal states. Suppose we are given an MDP $\mathcal{M} = (S, Act, P, \mathsf{init})$ without ECs and disjoint subsets $\mathsf{U}, \mathsf{V} \subseteq S$ of terminal states. Given a scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{V}) > 0$ we define:

$$ratio^{\mathfrak{S}}_{\mathcal{M}}(\mathsf{U},\mathsf{V}) \;=\; \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{U}) \,/\, \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{V})$$

The goal is to compute the extremal values: $ratio^{\min}_{\mathcal{M}}(\mathsf{U},\mathsf{V}) = \inf_{\mathfrak{S}} ratio^{\mathfrak{S}}_{\mathcal{M}}(\mathsf{U},\mathsf{V})$ and $ratio^{\max}_{\mathcal{M}}(\mathsf{U},\mathsf{V}) = \sup_{\mathfrak{S}} ratio^{\mathfrak{S}}_{\mathcal{M}}(\mathsf{U},\mathsf{V})$ where $\mathfrak{S}$ ranges over all schedulers such that $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{V}) > 0$. For their computation, we rely on a polynomial reduction to the classical *stochastic shortest path problem* [12]. For this, consider the MDP $\mathcal{N}$ arising from $\mathcal{M}$ by adding reset transitions from all terminal states $\mathsf{t} \in S \setminus \mathsf{V}$ to init. Thus, exactly the V-states are terminal in $\mathcal{N}$. The MDP $\mathcal{N}$ might contain ECs, which, however, do not intersect with V. We equip $\mathcal{N}$ with the weight function that assigns 1 to all states in U and 0 to all other states. For a scheduler $\mathfrak{T}$ with $\Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond \mathsf{V}) = 1$, let $\mathrm{E}^{\mathfrak{T}}_{\mathcal{N}}(\boxplus \mathsf{V})$ be the expected accumulated weight until reaching V under $\mathfrak{T}$. Let $\mathrm{E}^{\min}_{\mathcal{N}}(\boxplus \mathsf{V}) = \inf_{\mathfrak{T}} \mathrm{E}^{\mathfrak{T}}_{\mathcal{N}}(\boxplus \mathsf{V})$ and $\mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathsf{V}) = \sup_{\mathfrak{T}} \mathrm{E}^{\mathfrak{T}}_{\mathcal{N}}(\boxplus \mathsf{V})$, where $\mathfrak{T}$ ranges over all schedulers with $\Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond \mathsf{V}) = 1$. We can rely on known results [12,3,4] to obtain that both $\mathrm{E}^{\min}_{\mathcal{N}}(\boxplus \mathsf{V})$ and $\mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathsf{V})$ are computable in polynomial time. As $\mathcal{N}$ has only non-negative weights, $\mathrm{E}^{\min}_{\mathcal{N}}(\boxplus \mathsf{V})$ is finite and a corresponding MD-scheduler with minimal expectation exists. If $\mathcal{N}$ has an EC containing at least one U-state, which is the case iff $\mathcal{M}$ has a scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{U}) > 0$ and $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{V}) = 0$, then $\mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathsf{V}) = +\infty$. Otherwise, $\mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathsf{V})$ is finite and the maximum is achieved by an MD-scheduler as well.

Theorem 8. *Let $\mathcal{M}$ be an MDP without ECs and U, V disjoint sets of terminal states in $\mathcal{M}$, and let $\mathcal{N}$ be the constructed MDP as above. Then, $ratio^{\min}_{\mathcal{M}}(\mathsf{U},\mathsf{V}) = \mathrm{E}^{\min}_{\mathcal{N}}(\boxplus \mathsf{V})$ and $ratio^{\max}_{\mathcal{M}}(\mathsf{U},\mathsf{V}) = \mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathsf{V})$. Thus, both values are computable in polynomial time, and there is an MD-scheduler minimizing $ratio^{\mathfrak{S}}_{\mathcal{M}}(\mathsf{U},\mathsf{V})$, and an MD-scheduler maximizing $ratio^{\mathfrak{S}}_{\mathcal{M}}(\mathsf{U},\mathsf{V})$ if $ratio^{\max}_{\mathcal{M}}(\mathsf{U},\mathsf{V})$ is finite.*

Proof of Theorem 7. Using assumptions (A1)-(A3), we obtain that $covrat(\mathsf{Cause}) = ratio^{\min}_{\mathcal{M}}(\mathsf{U},\mathsf{V})$ where $\mathsf{U} = \{\mathsf{eff}_{\mathsf{cov}}\}$, $\mathsf{V} = \{\mathsf{eff}_{\mathsf{unc}}\}$. Similarly, with $\mathsf{U} = \{\mathsf{noeff}_{\mathsf{fp}}, \mathsf{eff}_{\mathsf{unc}}\}$, $\mathsf{V} = \{\mathsf{eff}_{\mathsf{cov}}\}$, we get $fscore(\mathsf{Cause}) = 0$ if $ratio^{\max}_{\mathcal{M}}(\mathsf{U},\mathsf{V}) = +\infty$ and $fscore(\mathsf{Cause}) = 2/(ratio^{\max}_{\mathcal{M}}(\mathsf{U},\mathsf{V}) + 2)$ otherwise. Thus, the claim follows from Theorem 8. □

5.3 Quality-optimal probability-raising causes

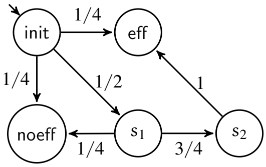
An SPR cause Cause is called *recall-optimal* if $recall(\mathsf{Cause}) = \max_{\mathsf{C}} recall(\mathsf{C})$ where C ranges over all SPR causes. Likewise, *ratio-optimality* resp. *f-score-optimality* of Cause means maximality of $covrat(\mathsf{Cause})$ resp. $fscore(\mathsf{Cause})$ among all SPR causes. Recall-, ratio- and f-score-optimality for GPR causes are defined accordingly.

Lemma 9. *Let* Cause *be an SPR or a GPR cause. Then,* Cause *is recall-optimal if and only if* Cause *is ratio-optimal.*

Recall- and ratio-optimal SPR causes. The techniques of Section 4.1 yield an algorithm for generating a canonical SPR cause with optimal recall and ratio. To see this, let $\mathcal{C}$ denote the set of states that constitute a singleton SPR cause. The canonical cause $\mathsf{CanCause}$ is defined as the set of states $\mathsf{c} \in \mathcal{C}$ such that there is a scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg \mathcal{C})\,\mathrm{U}\,\mathsf{c}) > 0$. Obviously, $\mathcal{C}$ and $\mathsf{CanCause}$ are computable in polynomial time.

Theorem 9. *If* $\mathcal{C} \neq \varnothing$ *then* CanCause *is a ratio- and recall-optimal SPR cause.*

This is not true for the f-score. To see this, Consider the Markov chain on the right hand side. We have $\mathsf{CanCause} = \{s_1\}$, which has *precision*(CanCause) $= \frac{3}{4}$ and *recall*(CanCause) $= \frac{3}{8}/(\frac{1}{4}+\frac{3}{8}) = \frac{3}{5}$. But the SPR cause $\{s_2\}$ has better f-score as its precision is 1 and it has the same recall as CanCause.



F-score-optimal SPR cause. From Section 5.2, we see that f-score-optimal SPR causes in MDPs can be computed in polynomial space by computing the f-score for all potential SPR causes one by one in polynomial time (Theorem 7). As the space can be reused after each computation, this results in polynomial space. For Markov chains, we can do better and compute an f-score-optimal SPR cause in polynomial time via a polynomial reduction to the stochastic shortest path problem:

Theorem 10. *In Markov chains that have SPR causes, an f-score-optimal SPR cause can be computed in polynomial time.*

Proof. We regard the given Markov chain $\mathcal{M}$ as an MDP with a singleton action set $Act = \{\alpha\}$. As $\mathcal{M}$ has SPR causes, the set $\mathcal{C}$ of states that constitute a singleton SPR cause is nonempty. We may assume that $\mathcal{M}$ has no non-trivial (i.e., cyclic) bottom strongly connected components as we may collapse them. Let $w_c = \Pr_{\mathcal{M},c}(\Diamond \mathsf{Eff})$. We switch from $\mathcal{M}$ to a new MDP $\mathcal{K}$ with state space $S_{\mathcal{K}} = S \cup \{\mathsf{eff}_{cov}, \mathsf{noeff}_{fp}\}$ with fresh states eff_{cov} and noeff_{fp} and the action set $Act_{\mathcal{K}} = \{\alpha, \gamma\}$. The MDP $\mathcal{K}$ arises from $\mathcal{M}$ by adding (i) for each state $c \in \mathcal{C}$ a fresh state-action pair (c,γ) with $P_{\mathcal{K}}(c,\gamma,\mathsf{eff}_{cov}) = w_c$ and $P_{\mathcal{K}}(c,\gamma,\mathsf{noeff}_{fp}) = 1 - w_c$ and (ii) reset transitions to init with action label α from the new state noeff_{fp} and all terminal states of $\mathcal{M}$, i.e., $P_{\mathcal{K}}(\mathsf{noeff}_{fp},\alpha,\mathsf{init}) = 1$ and $P_{\mathcal{K}}(s,\alpha,\mathsf{init}) = 1$ for $s \in \mathsf{Eff}$ or if s is a terminal non-effect state of $\mathcal{M}$. So, exactly eff_{cov} is terminal in $\mathcal{K}$, and $Act_{\mathcal{K}}(c) = \{\alpha,\gamma\}$ for $c \in \mathcal{C}$, while $Act_{\mathcal{K}}(s) = \{\alpha\}$ for all other states s. Intuitively, taking action γ in state $c \in \mathcal{C}$ selects c to be a cause state. The states in Eff represent uncovered effects in $\mathcal{K}$, while eff_{cov} stands for covered effects.

We assign weight 1 to all states in $U = \mathsf{Eff} \cup \{\mathsf{noeff}_{fp}\}$ and weight 0 to all other states of $\mathcal{K}$. Let $V = \{\mathsf{eff}_{cov}\}$. Then, $f = E^{\min}_{\mathcal{K}}(\boxplus V)$ and an MD-scheduler $\mathfrak{S}$ for $\mathcal{K}$ such that $E^{\mathfrak{S}}_{\mathcal{K}}(\boxplus V) = f$ are computable in polynomial time. Let $\mathcal{C}_\gamma$ denote the set of states $c \in \mathcal{C}$ where $\mathfrak{S}(c) = \gamma$ and let Cause be the set of states $c \in \mathcal{C}_\gamma$ where $\mathcal{M}$ has a path satisfying $(\neg \mathcal{C}_\gamma)\,\mathrm{U}\, c$. Then, Cause is an SPR cause of $\mathcal{M}$. With arguments as in Section 5.2 we obtain *fscore*(Cause) $= 2/(f+2)$. It remains to show that Cause is f-score-optimal. Let C be an arbitrary SPR cause. Then, $C \subseteq \mathcal{C}$. Let $\mathfrak{T}$ be the MD-scheduler for $\mathcal{K}$ that schedules γ in C and α for all other states of $\mathcal{K}$. Then, *fscore*$(C) = 2/(f^{\mathfrak{T}}+2)$ where $f^{\mathfrak{T}} = E^{\mathfrak{T}}_{\mathcal{K}}(\boxplus V)$. Hence, $f \leqslant f^{\mathfrak{T}}$, which yields *fscore*(Cause) $\geqslant$ *fscore*(C). □

The naïve adaption of the construction presented in the proof of Theorem 10 for MDPs would yield a stochastic game structure where the objective of one player is to minimize the expected accumulated weight until reaching a target state. Although algorithms for *stochastic shortest path (SSP) games* are known [34], they rely on assumptions on the game structure which would not be satisfied here. However, for the threshold problem *SPR-f-score* where inputs are an MDP $\mathcal{M}$, Eff and $\vartheta \in \mathbb{Q}_{\geqslant 0}$ and the task is to decide the existence of an SPR cause whose f-score exceeds ϑ, we can establish a polynomial reduction to SSP games, which yields an NP ∩ coNP upper bound:

Theorem 11. *The decision problem SPR-f-score is in* NP ∩ coNP.

Proof sketch. Given an MDP $\mathcal{M}$, an effect set Eff, and $\vartheta \in \mathbb{Q}$, we construct an SSP game [34] after a series of model transformations ensuring (i) that terminal states are reached almost surely and (ii) that Eff is reached with positive probability under all schedulers. Condition (i) is established by a standard MEC-quotient construction. To establish condition (ii), we provide a construction that forces schedulers to leave an initial sub-MDP in which the minimal probability to reach Eff is 0. This construction – unlike the MEC-quotient – affects the possible combinations of probability values with which terminal states and potential cause states can be reached, but the existence of an SPR cause satisfying the f-score-threshold condition is not affected.

The underlying idea of the construction of the game shares similarities with the MDP constructed in the proof of Theorem 10: Player 0 takes the role to select potential cause states while player 1 takes the role of a scheduler in the transformed MDP. Using the observation that for each cause C, $\mathit{fscore}(\mathsf{C}) > \vartheta$ iff

$$2(1-\vartheta)\mathrm{Pr}^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{C} \wedge \Diamond \mathsf{Eff}) - \vartheta \mathrm{Pr}^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond \mathsf{C} \wedge \Diamond \mathsf{Eff}) - \vartheta \mathrm{Pr}^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{C} \wedge \neg\Diamond \mathsf{Eff}) > 0 \qquad (\times)$$

for all schedulers $\mathfrak{S}$ for $\mathcal{M}$ with $\mathrm{Pr}^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathsf{Eff}) > 0$, weights are assigned to Eff-states and other terminal states depending on whether player 0 has chosen to include a state to the cause beforehand. In the resulting SSP game, both players have optimal MD-strategies [34]. Given such strategies ζ for player 0 and $\mathfrak{S}$ for player 1, the resulting expected accumulated weight agrees with the left-hand side of $(\times)$ when considering $\mathfrak{S}$ as a scheduler for the transformed MDP and the cause C induced by the states that ζ chooses to belong to the cause. Thus, player 0 wins the constructed game iff an SPR cause with f-score above the threshold ϑ exists. The existence of optimal MD-strategies for both players allows us to decide this threshold problem in NP and coNP. □

Optimality and threshold constraints for GPR causes. Computing optimal GPR causes for either quality measure can be done in polynomial space by considering all cause candidates, checking the GPR condition in polynomial space (Theorem 4) and computing the corresponding quality measure in polynomial time (Section 5.2). However, we show that no polynomial-time algorithms can be expected as the corresponding threshold problems are NP-hard. Let GPR-covratio (resp. GPR-recall, GPR-f-score) denote the decision problems: Given $\mathcal{M}$, Eff and $\vartheta \in \mathbb{Q}$, decide whether there exists a GPR cause with coverage ratio (resp. recall, f-score) at least ϑ.

Theorem 12. *The problems GPR-covratio, GPR-recall and GPR-f-score are NP-hard and belong to PSPACE. For Markov chains, all three problems are NP-complete. NP-hardness even holds for tree-like Markov chains.*

Proof sketch. NP-hardness is established via a polynomial reduction from the knapsack problem. Membership to NP for Markov chains resp. to PSPACE = NPSPACE for MDPs is obvious as we can guess nondeterministically a cause candidate and then check (i) the GPR condition in polynomial time (Markov chains) resp. polynomial space (MDPs) and (ii) the threshold condition in polynomial time (see Section 5.2). □

6 Conclusion

The goal of the paper was to formalize the probability-raising principle in MDPs and related quality notions for PR causes as well as studying fundamental algorithmic problems for them. We considered the strict (local) and the global view. Our results indicate that GPR causes are more general and leave more flexibility to achieve better accuracy, while algorithmic reasoning about SPR causes is simpler.

Existential definition of SPR/GPR causes. The proposed definition of PR causes relies on a universal quantification over all relevant schedulers. However, another approach could be via existential quantification, i.e. there is a scheduler $\mathfrak{S}$ such that (GPR) or resp. (SPR) hold. The resulting notion of causality yields fairly the same results (up to $\Pr^{\max}_{\mathcal{M},c}(\Diamond \mathsf{Eff})$ instead of $\Pr^{\min}_{\mathcal{M},c}(\Diamond \mathsf{Eff})$ etc). A canonical existential SPR cause can be defined in analogy to the universal case and shown to be recall- and ratio-optimal (cf. Theorem 9). The problem to find an existential f-score-optimal SPR cause is even simpler and solvable in polynomial time as the construction presented in the proof of Theorem 10 can be adapted for MDPs (thanks to the simpler nature of $\max_{\mathsf{C}} \sup_{\mathfrak{S}} \mathit{fscore}^{\mathfrak{S}}(\mathsf{C})$ compared to $\max_{\mathsf{C}} \inf_{\mathfrak{S}} \mathit{fscore}^{\mathfrak{S}}(\mathsf{C})$). However, NP-hardness for the existence of GPR causes with threshold constraints for the quality carries over to the existential definition (as NP-hardness holds for Markov chains, Theorem 12).

Non-strict inequality in the PR conditions. Our notions of PR causes are in line with the classical approach of probability-raising causality in literature with strict inequality in the PR condition. This has the consequence that causes might not exist (see Example 2). The switch to a relaxed definition of PR causes with non-strict inequality seems to be a minor change that identifies more sets as causes. Indeed, the proposed algorithms for checking the SPR and GPR condition (Section 4) can easily be modified for the relaxed definition. While this leads to a questionable notion of causality (e.g., $\{\mathsf{init}\}$ would always be a recall- and ratio-optimal SPR cause under the relaxed definition), it could be useful in combination with other side constraints. E.g., requiring the relaxed PR condition for all schedulers which reach a cause state with positive probability and requiring the existence of a scheduler where the PR condition with strict inequality holds might be a useful alternative definition that agrees with Def. 1 for Markov chains.

Relaxing the minimality condition (M). As many causality notions of the literature include some minimality constraint, we included condition (M). However, (M) could be dropped without affecting the algorithmic results presented here. This can be useful when the task is to identify components or agents that are responsible for the occurrences of undesired effects. In these cases the cause candidates are fixed (e.g., for each agent i, the set of states controlled by agent i), but some of them might violate (M).

Future directions include PR causality when causes and effects are path properties and the investigation of other quality measures for PR causes inspired by other indices for binary classifiers used in machine learning or customized for applications of cause-effect reasoning in MDPs. More sophisticated notions of probabilistic backward causality and considerations on PR causality with external interventions as in Pearl's do-calculus [35] are left for future work.

Acknowledgments We would like to thank Simon Jantsch and Clemens Dubslaff for their helpful comments and feedback on the topic of causality in MDPs.

References

1. Ábrahám, E., Bonakdarpour, B.: HyperPCTL: A temporal logic for probabilistic hyperproperties. In: McIver, A., Horváth, A. (eds.) 15th International Conference on Quantitative Evaluation of Systems (QEST). Lecture Notes in Computer Science, vol. 11024, pp. 20–35. Springer (2018), https://doi.org/10.1007/978-3-319-99154-2_2
2. de Alfaro, L.: Formal Verification of Probabilistic Systems. Phd thesis, Stanford University, Stanford, USA (1997), https://wcl.cs.rpi.edu/pilots/library/papers/TAGGED/4375-deAlfaro(1997)-FormalVerificationofProbabilisticSystems.pdf
3. de Alfaro, L.: Computing minimum and maximum reachability times in probabilistic systems. In: Baeten, J.C.M., Mauw, S. (eds.) 10th International Conference on Concurrency Theory (CONCUR). Lecture Notes in Computer Science, vol. 1664, pp. 66–81. Springer (1999), https://doi.org/10.1007/3-540-48320-9_7
4. Baier, C., Bertrand, N., Dubslaff, C., Gburek, D., Sankur, O.: Stochastic shortest paths and weight-bounded properties in Markov decision processes. In: Dawar, A., Grädel, E. (eds.) 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018. pp. 86–94. ACM (2018), https://doi.org/10.1145/3209108.3209184
5. Baier, C., Dubslaff, C., Funke, F., Jantsch, S., Majumdar, R., Piribauer, J., Ziemek, R.: From verification to causality-based explications (invited talk). In: Bansal, N., Merelli, E., Worrell, J. (eds.) 48th International Colloquium on Automata, Languages, and Programming, (ICALP). LIPIcs, vol. 198, pp. 1:1–1:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2021), https://doi.org/10.4230/LIPIcs.ICALP.2021.1
6. Baier, C., Funke, F., Jantsch, S., Piribauer, J., Ziemek, R.: Probabilistic causes in Markov chains. CoRR **abs/2104.13604** (2021), https://arxiv.org/abs/2104.13604, accepted for publication at ATVA'21.
7. Baier, C., Funke, F., Majumdar, R.: A game-theoretic account of responsibility allocation. In: Zhou, Z. (ed.) 30th International Joint Conference on Artificial Intelligence (IJCAI). pp. 1773–1779. ijcai.org (2021), https://doi.org/10.24963/ijcai.2021/244
8. Baier, C., Funke, F., Piribauer, J., Ziemek, R.: On probability-raising causality in markov decision processes (2022), https://arxiv.org/abs/2201.08768
9. Baier, C., Katoen, J.P.: Principles of Model Checking (Representation and Mind Series). The MIT Press, Cambridge, MA (2008)
10. Baier, C., Klein, J., Klüppelholz, S., Märcker, S.: Computing conditional probabilities in Markovian models efficiently. In: Ábrahám, E., Havelund, K. (eds.) 20th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS). Lecture Notes in Computer Science, vol. 8413, pp. 515–530. Springer (2014), https://doi.org/10.1007/978-3-642-54862-8_43
11. Beer, I., Ben-David, S., Chockler, H., Orni, A., Trefler, R.J.: Explaining counterexamples using causality. Formal Methods in System Design **40**(1), 20–40 (2012), https://doi.org/10.1007/s10703-011-0132-2
12. Bertsekas, D.P., Tsitsiklis, J.N.: An analysis of stochastic shortest path problems. Mathematics of Operations Research **16**(3), 580–595 (1991)
13. Canny, J.F.: Some algebraic and geometric computations in PSPACE. In: 20th Annual ACM Symposium on Theory of Computing (STOC). pp. 460–467. ACM (1988)
14. Chockler, H.: Causality and responsibility for formal verification and beyond. In: First Workshop on Causal Reasoning for Embedded and safety-critical Systems Technologies (CREST). EPTCS, vol. 224, pp. 1–8 (2016), https://doi.org/10.4204/EPTCS.224.1
15. Chockler, H., Halpern, J.Y., Kupferman, O.: What causes a system to satisfy a specification? ACM Transactions on Computational Logic **9**(3), 20:1–20:26 (2008)
16. Clarke, E.M., Grumberg, O., Peled, D.: Model Checking. MIT Press (1999)

17. Dimitrova, R., Finkbeiner, B., Torfah, H.: Probabilistic hyperproperties of Markov decision processes. In: Hung, D.V., Sokolsky, O. (eds.) 18th International Symposium on Automated Technology for Verification and Analysis (ATVA). Lecture Notes in Computer Science, vol. 12302, pp. 484–500. Springer (2020), https://doi.org/10.1007/978-3-030-59152-6_27
18. Eells, E.: Probabilistic Causality. Cambridge Studies in Probability, Induction and Decision Theory, Cambridge University Press (1991)
19. Friedenberg, M., Halpern, J.Y.: Blameworthiness in multi-agent settings. In: 33rd Conference on Artificial Intelligence (AAAI). pp. 525–532. AAAI Press (2019), https://doi.org/10.1609/aaai.v33i01.3301525
20. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman (1979)
21. Halpern, J.Y., Pearl, J.: Causes and explanations: A structural-model approach: Part 1: Causes. In: 17th Conference in Uncertainty in Artificial Intelligence (UAI). pp. 194–202 (2001)
22. Hitchcock, C.: Probabilistic causation. In: Hájek, A., Hitchcock, C. (eds.) The Oxford Handbook of Probability and Philosophy, pp. 815–832. Oxford University Press (2016)
23. Kallenberg, L.: Lecture Notes Markov Decision Problems - version 2020 (02 2020)
24. Kleinberg, S., Mishra, B.: The temporal logic of causal structures. In: 25th Conference on Uncertainty in Artificial Intelligence (UAI). pp. 303–312 (2009)
25. Kleinberg, S.: Causality, Probability and Time. Cambridge University Press (2012)
26. Kozlov, M.K., Tarasov, S.P., Khachiyan, L.G.: The polynomial solvability of convex quadratic programming. USSR Computational Mathematics and Mathematical Physics **20**(5), 223–228 (1980)
27. Kuntz, M., Leitner-Fischer, F., Leue, S.: From probabilistic counterexamples via causality to fault trees. In: Flammini, F., Bologna, S., Vittorini, V. (eds.) 30th International Conference on Computer Safety, Reliability, and Security (SAFECOMP). Lecture Notes in Computer Science, vol. 6894, pp. 71–84. Springer (2011), https://doi.org/10.1007/978-3-642-24270-0_6
28. Leitner-Fischer, F.: Causality Checking of Safety-Critical Software and Systems. Ph.D. thesis, University of Konstanz, Germany (2015), http://kops.uni-konstanz.de/handle/123456789/30778
29. Lewis, D.: Counterfactuals and comparative possibility. Journal of Philosophical Logic **2**(4), 418–446 (1973)
30. Manna, Z., Pnueli, A.: The Temporal Logic of Reactive and Concurrent Systems: Safety. Springer-Verlag (1995)
31. Märcker, S.: Model checking techniques for design and analysis of future hardware and software systems. Ph.D. thesis, TU Dresden, Germany (2020), https://d-nb.info/1232958204
32. Namjoshi, K.S.: Certifying model checkers. In: 13th International Conference on Computer Aided Verification (CAV). Lecture Notes in Computer Science, vol. 2102, pp. 2–13. Springer (2001), https://doi.org/10.1007/3-540-44585-4_2
33. Pardalos, P.M., Vavasis, S.A.: Quadratic programming with one negative eigenvalue is np-hard. Journal of Global optimization **1**(1), 15–22 (1991)
34. Patek, S.D., Bertsekas, D.P.: Stochastic shortest path games. SIAM Journal on Control and Optimization **37**(3), 804–824 (1999)
35. Pearl, J.: Causality. Cambridge University Press, 2nd edn. (2009)
36. Powers, D.: Evaluation: From precision, recall and f-factor to ROC, informedness, markedness & correlation. Mach. Learn. Technol. **2** (01 2008)
37. Puterman, M.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons, Inc., New York, NY (1994)
38. Reichenbach, H.: The Direction of Time. Dover Publications (1956)

39. Suppes, P.: A Probabilistic Theory of Causality. Amsterdam: North-Holland Pub. Co. (1970)
40. Yazdanpanah, V., Dastani, M.: Distant group responsibility in multi-agent systems. In: Baldoni, M., Chopra, A.K., Son, T.C., Hirayama, K., Torroni, P. (eds.) 19th International Conference on Princiles and Practice of Multi-Agent Systems (PRIMA). Lecture Notes in Computer Science, vol. 9862, pp. 261–278. Springer (2016), https://doi.org/10.1007/978-3-319-44832-9_16
41. Yazdanpanah, V., Dastani, M., Jamroga, W., Alechina, N., Logan, B.: Strategic responsibility under imperfect information. In: Elkind, E., Veloso, M., Agmon, N., Taylor, M.E. (eds.) 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS). pp. 592–600. International Foundation for Autonomous Agents and Multiagent Systems (2019), http://dl.acm.org/citation.cfm?id=3331745