Probability-Raising Causality
for Uncertain Parametric Markov Decision Processes with PAC Guarantees

Ryohei Oura[1]

Yuji Ito[2]

[1]Frontier Research Center, Toyota Motor Corporation, Japan
[2]Toyota Central R&D Labs., Inc., Japan

Abstract

Recent decision-making systems are increasingly complicated, making it crucial to verify and understand their behavior for a given specification. A promising approach is to comprehensively explain undesired behavior in the systems modeled by Markov decision processes (MDPs) through formal verification and causal reasoning. However, the reliable explanation using model-based probabilistic causal analysis has not been explored when the MDP's transition probabilities are uncertain. This paper proposes a method to identify potential causes of undesired behaviors in an uncertain parametric MDP (upMDP) using parameter sampling, model checking, and a set covering for the samples. A cause is defined as a subset of states based on a *probability-raising* principle. We show that the probability of each identified subset being a cause exceeds a specified threshold. Further, a lower bound of the probability that the undesired paths visit the subsets is maximized as much as possible while satisfying a nonredundancy condition. While computing these probabilities is complicated, this study derives probabilistically approximately correct lower bounds of both probabilities by the sampling. We demonstrate the effectiveness of the proposed method through a path-planning scenario.

1 INTRODUCTION

As modern decision-making systems have become more complex and exhibit various forms of stochasticity in many applications such as traffic networks and robotic delivery systems [Dudek et al., 1996, Dorri et al., 2018], it is increasingly important to understand their intricate behavior to answer *why* they violate or satisfy a given specification. For stochastic systems like Markov decision processes (MDPs),

model checking-based formal verification has proven effective in determining whether these systems satisfy the specifications [Baier and Katoen, 2008]. However, model checking typically provides only basic information: a satisfaction guarantee or a counterexample that violates the specification, rather than an in-depth explanation of the system's behavior.

As a promising approach to explain such behaviors, cause-effect reasoning has attracted a lot of attention in both machine learning and formal method communities [Baier et al., 2021a, Oberst and Sontag, 2019, Baier et al., 2022a, Coenen et al., 2022, Finkbeiner and Siber, 2023, Baier et al., 2024a, Finkbeiner et al., 2024, Dimitrova et al., 2020, Kazemi and Paoletti, 2022, Baier et al., 2024b, Oberst and Sontag, 2019, Tsirtsis et al., 2021, Triantafyllou et al., 2022, Baier et al., 2021c] (see Section 6.2 for more details). Recently, probabilistic causality for MDPs based on a *probability-raising* (PR) principle has been advocated [Baier et al., 2021b, 2022b, 2024a]. PR causes are defined as a subset of states of an MDP such that passing through the states increases the likelihood of undesired behavior compared to bypassing them, regardless of action selections. The authors have proposed a model checking-based algorithm to intensively identify ones that satisfy a path coverage criterion, offering an explainability advantage in verifying MDPs over other causal reasoning studies.

While this PR-causality framework has been established well for exact MDPs, reliably computing PR causes becomes a challenge when the transition probabilities themselves are epistemically uncertain. Although the algorithms in [Baier et al., 2022b, 2024a] have relied on fully known MDPs, this assumption is often unrealistic in complicated systems. Common approaches to handle the epistemically uncertain parameters of MDPs, such as transition probabilities, are using intervals [Givan et al., 2000, Delahaye et al., 2011, Puggelli et al., 2013, Jackson et al., 2020] and defining a class of uncertain (parametric) MDPs (upMDPs) [Nilim and El Ghaoui, 2005, Wiesemann et al., 2013, Wolff et al., 2012, Rickard et al., 2024]. Furthermore, *scenario op-*

timization program (SOP) [Campi and Garatti, 2008, 2011] has emerged as a powerful approach for handling such parameter uncertainty via sampling.

Although SOP-based formal verification and synthesis in up-MDPs well-investigated in [Cubuktepe et al., 2020, Badings et al., 2022a, Nilim and El Ghaoui, 2005, Wiesemann et al., 2013, Wolff et al., 2012, Rickard et al., 2024] (see Section 6.1 for details), the following non-trivial question remains in the causal reasoning context: Identifying *potential* PR causes *exhaustively* but *nonredundantly*. Cubuktepe et al. [2020], Badings et al. [2022a], which are most related to this paper, have provided probabilistically approximately correct (PAC) lower bounds for satisfaction probabilities by applying model checking to sampled MDPs and SOP formulations, but they do not offer causative states.

In this paper, based on parameter sampling, model checking, and a set covering strategy for the samples, we propose a method to exhaustively and nonredundantly identify subsets of states that act as potential PR causes in a upMDP. We show that the resulting collection of subsets satisfies the following three properties in a PAC manner: (i) Potential cause: Each identified subset of states becomes a PR cause with at least a specified probability. (ii) Exhaustiveness: A lower bound of the probability that the undesired paths traverse the collection as much as possible is maximized. (iii) Nonredundancy: Removing a subset of states from the collection decreases the lower bound.

Our contributions are summarized as follows:

1. We formalize the problem of exhaustively and nonredundantly identifying PR causes on a upMDP by introducing two probabilities that capture (a) the likelihood of a subset of states being a PR cause and (b) the likelihood that undesired paths maximally pass through such subsets. This addresses the limitation in [Baier et al., 2022b] when applying to uncertain models.

2. We provide lower bounds for the introduced probabilities over the parameter space based on a state set inclusion condition and a path condition. (Theorems 1 and 2). This result enables the computation of the PAC-guaranteed probabilities using sampled parameters.

3. We develop a set covering-based method combined with model checking-based PR cause synthesis to compute a solution to the problem formalized in 1 with PAC-guarantees, and its correctness is established in Theorem 3. This result paves the way towards compact and verifiable explanations under aleatoric and epistemic uncertainties.

4. We demonstrate the effectiveness of the proposed method with a path-planning scenario. The proposed method effectively finds exhaustive and nonredundant key waypoints that substantially influence undesired outcomes. However, two naïve approaches to be compared do not. This case study suggests a potential utility in control applications with epistemic and aleatoric uncertainties.

2 PRELIMINARIES

2.1 UNCERTAIN PARAMETRIC MARKOV DECISION PROCESSES

We first consider a parametric Markov decision process (pMDP). Then, we extend it toward an uncertain parametric MDP (upMDP) with probability distributions over the parameter space of the pMDP.

A pMDP is a tuple $\mathcal{M}_u = (S, A, s_0, u, P)$, where S is a finite set of states; A is a finite set of actions; $s_0 \in S$ is an initial state; $u = [p_0, \ldots, p_{m-1}]^\top$ is a real-valued parameter vector, where u takes a value in a subset of $\mathbb{R}^m$ denoted by $\mathcal{V} \subseteq \mathbb{R}^m$; and $P : S \times A \times S \times \mathcal{V} \to [0, 1]$ is a transition function, where $P(s, a, s', u)$ denotes the probability of transition from s to s' with the action a. For any state $s \in S$, we denote the set of enabled actions at s by $\mathcal{A}(s) = \{a \in A \mid \exists u, \exists s' \text{ s.t. } P(s, a, s', u) > 0\}$. We say that $s \in S$ is terminal if $\mathcal{A}(s) = \emptyset$. We denote the set of terminal states by E and assume $E \neq \emptyset$. A path is a sequence of states and actions $\rho = s_0 a_0 s_1 a_1 s_2 \ldots$ such that $P(s_i, a_i, s_{i+1}, u) > 0$ for some $u \in \mathcal{V}$ and all indices $i \geq 0$. A policy $\pi : \mathrm{FinPath} \times A \to [0, 1]$ is a function that assigns a probability to a pair of a finite path and an enabled action at the last state, where $\mathrm{FinPath}$ is the set of all finite paths. A path reaching E is interpreted as an undesired one.

A upMDP is a tuple $\mathcal{M}_{\mathbb{P}} = (\mathcal{M}_u, \mathbb{P})$, where $\mathcal{M}_u$ is a pMDP and $\mathbb{P}$ is a probability distribution over the parameter space $\mathcal{V}$ of $\mathcal{M}$. Intuitively, a upMDP is a pMDP combined with an associated probability distribution over possible parameter values. So, a parameter $u \in \mathcal{V}$ generated from $\mathbb{P}$ yields a concrete MDP $\mathcal{M}_u$ if the instantiated transition function $P_u = P(\cdot, \cdot, \cdot, u)$ is well-defined, that is, $\sum_{s' \in S} P_u(s, a, s') = 1$ for any $s \in S$ and any $a \in A$ if $a \in \mathcal{A}(s)$, otherwise 0. We assume that all parameter values in $\mathcal{V}$ yield well-defined MDPs. For any MDP $\mathcal{M}_u$, any state $s \in S$, and any policy π, we define the probability measure $\mathrm{Pr}^\pi_{\mathcal{M}_u, s}$ on measurable sets of paths starting from s under π in a standard way [Baier and Katoen, 2008]. For simplicity, we abbreviate $\mathrm{Pr}^\pi_{\mathcal{M}_u, s_0}$ as $\mathrm{Pr}^\pi_{\mathcal{M}_u}$.

Example 1. *We consider a upMDP depicted in Fig. 1 with the set of parameter variables $u = [p, q]^\top$, the set of actions $A = \{a, b\}$, and the terminal state $E = \{s_5\}$. The enabled action is only a at states other than s_0. Their transition probabilities are represented using parameters p and q. p and q are equal to $1/2$ with probability $1/10$ and, with probability $9/10$, p and q follow the uniform distributions over $[0.11, 0.51]$ and $[0.3, 0.7]$, respectively.*
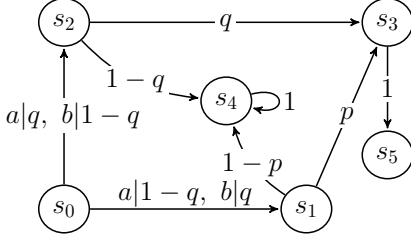
Figure 1: The upMDP with the parameter variables $V = \{p, q\}$, the set of actions $A = \{a, b\}$, and the terminal state $E = \{s_5\}$. The enabled action is only a at states other than s_0 and thus we omit it.

2.2 PROBABILITY-RAISING CAUSALITY

For convenience, we use linear temporal logic (LTL)-like notations such as $\neg$ (negation), $\wedge$ (and), $\diamond$ (eventually), and U (until) to represent some path properties. For any $X_0, X_1 \subseteq S$, the formula $X_0 U X_1$ is satisfied by a (state) path $\rho = s_0 s_1 \ldots$ such that there exists $j \geq 0$ such that, for any $i < j$, $s_i \in X_0$ and $s_j \in X_1$ hold, and we denote $\diamond X_1 = S U X_1$. Furthermore, $\neg \diamond X_1$ is satisfied by a path ρ such that all states in ρ are not contained by X_1. For simplicity, we abbreviate $\{s_0\} U \{s_1\}$ and $\diamond \{s_2\}$ as $s_0 U s_1$ and $\diamond s_2$, respectively, for any s_0, s_1, and $s_2 \in S$.

We consider a cause-effect relation over an MDP based on the probability-raising principle introduced in [Baier et al., 2022b, 2024a].

Definition 1 (SPR cause [Baier et al., 2022b]). *For any MDP $\mathcal{M}_u$ with a parameter $u \in \mathcal{V}$, the set of terminal states $E \subseteq S$, and a subset $C \subseteq S \setminus E$, we say that C is a strict probability-raising (SPR) cause for E on $\mathcal{M}_u$ if and only if the following conditions (M) and (S) hold:*

(M) $\forall c \in C, \exists \pi$ *such that* $\mathrm{Pr}^\pi_{\mathcal{M}_u}(\neg C U c) > 0$.
(S) $\forall c \in C, \forall \pi$ *where* $\mathrm{Pr}^\pi_{\mathcal{M}_u}(\neg C U c) > 0$,

$$\mathrm{Pr}^\pi_{\mathcal{M}_u}(\diamond E \mid \neg C U c) > \mathrm{Pr}^\pi_{\mathcal{M}_u}(\diamond E). \tag{1}$$

Intuitively, $\diamond E$ means that a path eventually reaches a state in E, and $\neg C U c$ denotes that a path eventually reaches c avoiding others in C. SPR cause is a minimal subset C of states and is interpreted as waypoints that raise the probability of reaching terminal states E. This probability-raising condition is defined by **(S)**, implying that visiting any state in C increases the probability of reaching E. The minimality of the subset is defined by **(M)**, removing states in C that can only be reached by passing through the others. We denote that C satisfies the above conditions **(M)** and **(S)** on an MDP $\mathcal{M}_u$ by $(\mathcal{M}_u, C) \models$ **(M)** and $(\mathcal{M}_u, C) \models$ **(S)**, respectively.

In Example 1, the set of states $C = \{s_2, s_3\}$ is an SPR cause on the MDP $\mathcal{M}_u$ with $p = 0.3$ and $q = 0.6$ for

the following reasons: First, there exist two paths $s_0 a s_2$ and $s_0 b s_1 a s_3$. Hence, C satisfies the condition **(M)** on $\mathcal{M}_u$. Second, the minimal probabilities over any policies $\mathrm{Pr}^{\min}_{\mathcal{M}_u}(\diamond s_5 \mid \neg C U s_2) = 0.6$ and $\mathrm{Pr}^{\min}_{\mathcal{M}_u}(\diamond s_5 \mid \neg C U s_3) = 1$ are greater than the maximal probability $\mathrm{Pr}^{\max}_{\mathcal{M}_u}(\diamond s_3) = 0.48$. Thus, C satisfies **(S)** on $\mathcal{M}$.

In the following, we do not mention the set of terminal states E for any upMDP when it is clear from the context.

Definition 2 (SPR-cause probability). *For any upMDP $\mathcal{M}_{\mathbb{P}}$ and any $C \subset S$, we define the probability that C is an SPR cause on $\mathcal{M}_{\mathbb{P}}$ as:*

$$F(\mathcal{M}_{\mathbb{P}}, C) = \int_{u \in \mathcal{V}} I(u, C) \mathrm{d}\mathbb{P}(u), \tag{2}$$

where $I : \mathcal{V} \times 2^S \to \{0, 1\}$ is an indicator function such that, for any $u \in \mathcal{V}$ and any $C \subset S$, $I(u, C) = 1$ if C is an SPR cause on $\mathcal{M}_u$ and otherwise 0. We call the probability defined by (2) the SPR-cause probability for C.

To evaluate how often an SPR cause covers paths reaching E on the upMDP, we consider the probability of paths passing through an SPR cause C conditioned on reaching E as a probabilistic coverage of SPR causes.

Definition 3 (Recall). *For any MDP $\mathcal{M}_u$, any policy π, and any $C \subseteq S$, we introduce the conditional probability called recall for C under π as*

$$\mathrm{Pr}^\pi_{\mathcal{M}_u}(\diamond C \mid \diamond E), \tag{3}$$

and $\mathrm{Pr}^\pi_{\mathcal{M}_u}(\diamond C \mid \diamond E) = 0$ if $\mathrm{Pr}^\pi_{\mathcal{M}_u}(\diamond E) = 0$.

Definition 4 (Recall optimality). *For any MDP $\mathcal{M}_u$, any $C \subseteq S$, and any $\mathcal{S} \subseteq 2^S$, we say that C is recall-optimal on $\mathcal{M}_u$ over $\mathcal{S}$ if and only if C is an SPR cause and satisfies*

$$\forall C' \in \mathcal{S}, C' \text{ is an SPR cause on } \mathcal{M}_u$$
$$\implies \min_\pi \mathrm{Pr}^\pi_{\mathcal{M}_u}(\diamond C | \diamond E) - \mathrm{Pr}^\pi_{\mathcal{M}_u}(\diamond C' | \diamond E) \geq 0, \tag{4}$$

Intuitively, when the conditional probability of passing through C before making undesired paths is maximal among all SPR causes, C is recall optimal.

Definition 5 (Recall-optimal probability). *For any MDP $\mathcal{M}_u$ and any $\mathcal{C}, \mathcal{S} \in 2^S$, we define the probability that $\mathcal{C} \subseteq 2^S$ covers recall-optimal SPR causes over $\mathcal{S} \subseteq 2^S$ under $\mathcal{M}_{\mathbb{P}}$ as:*

$$R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}, \mathcal{S}) = \int_{u \in \mathcal{V}} I'(u, \mathcal{C}, \mathcal{S}) \mathrm{d}\mathbb{P}(u), \tag{5}$$

where $I' : \mathcal{V} \times 2^{2^S} \times 2^{2^S} \to \{0, 1\}$ is an indicator function such that, for any $u \in \mathcal{V}$, any $\mathcal{C} \subseteq 2^S$, and any $\mathcal{S} \subseteq 2^S$, $I'(u, \mathcal{C}, \mathcal{S}) = 1$ if there exists $C \in \mathcal{C}$ such that C is recall-optimal over $\mathcal{S}$ under $\mathcal{M}_u$ and otherwise 0. We call the probability defined by (5) the recall-optimal probability for $\mathcal{C}$ over $\mathcal{S}$.

Recall optimal probability is interpreted as the probability over the parameter space of a upMDP that a given collection of subsets of states contains a recall optimal SPR cause.

Moreover, we define a collection of subsets of states whose SPR-cause probabilities exceed $\delta > 0$ as

$$\mathcal{S}_\delta = \{C \subset S \mid F(\mathcal{M}_\mathbb{P}, C) > \delta\}. \tag{6}$$

3 PROBLEM SETTING

Our goal is to find a nonredundant but exhaustive collection of potential SPR causes $\mathcal{C}^* \subset 2^S$ on the given upMDP. To reflect this intuition, we consider the following problem.

Problem 1. *For any δ, compute $\mathcal{C}^* \subset 2^S$ that satisfies the following two conditions:*

(C1) $\mathcal{C}^* \in \arg\max_{\mathcal{C} \subseteq \mathcal{S}_\delta} R(\mathcal{M}_\mathbb{P}, \mathcal{C}, \mathcal{S}_\delta)$,

(C2) $\forall \mathcal{C}' \subset \mathcal{C}^*, R(\mathcal{M}_\mathbb{P}, \mathcal{C}^*, \mathcal{S}_\delta) > R(\mathcal{M}_\mathbb{P}, \mathcal{C}', \mathcal{S}_\delta)$,

where $\mathcal{S}_\delta$ is defined by (6).

In Problem 1, **(C1)** and **(C2)** state an exhaustiveness and a nonredundancy conditions for $\mathcal{C}^*$, respectively, for the recall-optimal probability. Intuitively, **(C1)** requires that $\mathcal{C}^*$ maximally covers the paths reaching some terminal states in a probabilistic sense while the SPR-cause probability for each $C \in \mathcal{C}^*$ exceeds a given threshold δ. The maximality of recall-optimal probability corresponds to exhaustiveness, and the threshold for the SPR-cause probability reflects our intuition for potential causes. Plus, **(C2)** requires that the collection $\mathcal{C}^*$ is nonredundant in the sense that removing any member of $\mathcal{C}^*$ reduces the recall-optimal probability. We can analyticaly compute F defined by (2) and R defined by (5) if $\mathcal{V}$ is finite and $\mathbb{P}$ is known. However, in practice, $\mathcal{V}$ can be infinite, and $\mathbb{P}$ can be unknown. Hence, it is generally undecidable to compute the probabilities [Arming et al., 2018]. Thus, we relax Problem 1 by a probably approximately correct (PAC) style formulation. We consider the following two problems, assuming we can sample parameters identically and independently distributed from $\mathbb{P}$.

Problem 2. *For any $C \subset S$, any $\mathcal{C}, \mathcal{S} \subset 2^S$, any $N > 0$, and any $\beta \in (0, 1)$, compute $\eta_N(C, \beta)$ and $\zeta_N(\mathcal{C}, \mathcal{S}, \beta)$ that satisfy the following conditions:*

$$\mathbb{P}^N(F(\mathcal{M}_\mathbb{P}, C) \geq \eta_N(C, \beta)) \geq \beta, \tag{7}$$
$$\mathbb{P}^N(R(\mathcal{M}_\mathbb{P}, \mathcal{C}, \mathcal{S}) \geq \zeta_N(\mathcal{C}, \mathcal{S}, \beta)) \geq \beta, \tag{8}$$

where $\mathbb{P}^N$ is the N product measure of $\mathbb{P}$.

Problem 3. *For any $N > 0$ and any $\delta, \beta \in (0, 1)$, compute $\mathcal{C}^* \subset 2^S$ that satisfies the following two conditions:*

(PC1) $\mathcal{C}^* \in \arg\max_{\mathcal{C} \subseteq \mathcal{S}_{N,\delta,\beta}} \zeta_N(\mathcal{C}, \mathcal{S}_{N,\delta,\beta}, \beta)$,

(PC2) $\forall \mathcal{C}' \subset \mathcal{C}^*$,
$$\mathbb{P}^N(R(\mathcal{M}_\mathbb{P}, \mathcal{C}^*, \mathcal{S}_{N,\delta,\beta}) > R(\mathcal{M}_\mathbb{P}, \mathcal{C}', \mathcal{S}_{N,\delta,\beta})) > \beta,$$

where both η_N and ζ_N are the solutions to Problem 2, and $\mathcal{S}_{N,\delta,\beta}$ is defined as

$$\mathcal{S}_{N,\delta,\beta} = \{C \subset 2^S \mid \eta_N(C, \beta) > \delta\}. \tag{9}$$

In Problem 2, we compute probabilistically guaranteed lower bounds for the SPR-cause probability and recall-optimal probability. Problem 3 relaxes Problem 1 based on the PAC bounds.

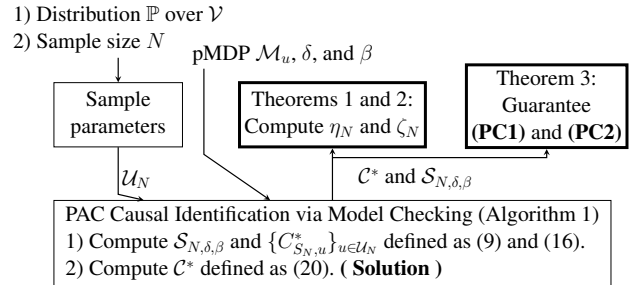## 4 PROBABILITY-RAISING CAUSAL IDENTIFICATION ON UNCERTAIN PARAMETRIC MDPS

Figure 2: Overview of our proposed method. We first sample the parameters of the upMDP. Then, we compute the solution to Problem 3 from the sampled MDPs based on model checking. We provide the probabilistic guarantees for the obtained result using Theorems 1, 2, and 3.

In this section, we describe how we obtain the solutions to Problems 2 and 3. In Section 4.1, we provide the solutions η_N and ζ_N to Problem 2 using parameter sampling. In Section 4.2, we construct a solution to Problem 3 based on a set covering for the sampled parameters. In Section 4.3, we provide a model checking-based algorithm to compute the solution to Problem 3. We show the overview of our approach in Fig. 2. In the following, we denote by $\mathcal{U}_N = (u_i)_{i=1}^N$ the sequence of N parameter vectors sampled from the parameter space $\mathcal{V}$ of the upMDP $\mathcal{M}_\mathbb{P}$.

4.1 PAC GUARANTEES FOR PROBABILITY-RAISING CAUSES ON UNCERTAIN PARAMETRIC MDPS

In this section, we provide a solution to Problem 2. We first give a lower bound of the SPR-cause probability for any subset of states $C \subseteq S \setminus E$ as a solution to (7). For any $u \in \mathcal{V}$ and any $S' \subseteq S$, we define $\mathscr{C}_{S',u} \subset S$ as

$$\mathscr{C}_{S',u} = \{c \in S' \mid (\mathcal{M}_u, \{c\}) \models (\mathbf{S})\}, \tag{10}$$

where **(S)** is defined in Def. 1. For any $C \subseteq S$, we define $n(C, \mathcal{U}_N)$ as the number of samples $u \in \mathcal{U}_N$ that satisfy the following condition:

$$C \subseteq \mathscr{C}_{S,u}, (\mathcal{M}_u, C) \models \textbf{(M)}. \qquad (11)$$

Note that $n(C, \mathcal{U}_N)$ is regarded as the number of times that each $c \in C$ is an SPR cause on $\mathcal{M}_u$ and C satisfies the condition **(M)** of SPR cause in Def. 1. Then, we define the function $\eta_N : 2^S \times (0, 1) \to \mathbb{R}$ as

$$\eta_N(C, \beta) = t^*(N - n(C, \mathcal{U}_N), \beta), \qquad (12)$$

where the function $t^* : [N] \times [0, 1] \to \mathbb{R}$ is defined such that, for any β,

$$t^*(0, \beta) = (1 - \beta)^{1/N}, \; t^*(N, \beta) = 0, \qquad (13)$$

and, for any $k = 1, \dots, N - 1$, $t^*(k, \beta)$ is the solution of

$$\frac{1 - \beta}{N} = \sum_{i=0}^{k} \binom{N}{i} (1 - t)^i t^{N-i}. \qquad (14)$$

Theorem 1. *For any upMDP $\mathcal{M}_{\mathbb{P}}$, any $N > 0$, any $C \subseteq S \setminus E$, and any $\beta \in (0, 1]$, we have*

$$\mathbb{P}^N \{ F(\mathcal{M}_{\mathbb{P}}, C) \geq \eta_N(C, \beta) \} \geq \beta, \qquad (15)$$

where η_N is defined as (12).

By Theorem 1, η_N defined as (12) is a solution to (7).

We provide the lower bound of the recall-optimal probability of any collection of subsets of states as a solution to (8). For any $u \in \mathcal{V}$ and any $S' \subset S$, we define a canonical SPR cause $C^*_{S',u}$ for S' as

$$C^*_{S',u} = \{ c \in \mathscr{C}_{S',u} \mid \exists \pi \text{ s.t. } \Pr^{\pi}_{\mathcal{M}_u}(\neg \mathscr{C}_{S',u} U c) > 0 \}, \quad (16)$$

Intuitively, $C^*_{S',u}$ is interpreted as the "front" of all singleton SPR causes over S'. For any $\mathcal{C}, \mathcal{S} \subseteq 2^S$, we define $m(\mathcal{C}, \mathcal{S}, \mathcal{U}_N)$ as the number of samples $u \in \mathcal{U}_N$ such that there exists $C \in \mathcal{C}$ that satisfies the following three conditions: (i) C covers all path reaching E that passes through $C^*_{S',u}$ for the union S' of $\mathcal{S}$, (ii) C is a subset of $\mathscr{C}_{S',u}$, and (iii) C satisfies **(M)** in Def. 1, that is,

$$\exists C \in \mathcal{C} \text{ s.t. } C \subseteq \mathscr{C}_{S',u}, (\mathcal{M}_u, C) \models \textbf{(M)},$$
$$\forall \rho \in \text{FinPath}, \rho \models \Diamond C^*_{S',u} \wedge \Diamond E \implies \Diamond C, \qquad (17)$$

where $\rho \models \varphi$ denotes that the path ρ on an MDP $\mathcal{M}_u$ satisfies the formula φ and $S' = \cup_{C' \in \mathcal{S}} C'$. We define the function $\zeta_N : 2^{2^S} \times 2^{2^S} \times (0, 1) \to \mathbb{R}$ as

$$\zeta_N(\mathcal{C}, \mathcal{S}, \beta) = t^*(N - m(\mathcal{C}, \mathcal{S}, \mathcal{U}_N), \beta), \qquad (18)$$

where t^* is defined as (14).

Theorem 2. *For any upMDP $\mathcal{M}_{\mathbb{P}}$, any $N > 0$, any $\mathcal{C}, \mathcal{S} \subseteq 2^S$, and any $\beta \in (0, 1]$, we have*

$$\mathbb{P}^N \{ R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}, \mathcal{S}) \geq \zeta_N(\mathcal{C}, \mathcal{S}, \beta) \} \geq \beta, \qquad (19)$$

where ζ_N is defined as (18).

By Theorem 2, ζ_N defined by (18) is a solution to (8).

Remark 1 (On Lower Bounds in Theorems 1 and 2). *The right-hand side of (14) is the cumulative distribution function of the binary distribution for k with the success probability $1 - t$. Clearly, the maximum and minimum values of t^* defined by (13) and (14) are $(1 - \beta)^{1/N}$ and 0. Thus, the values of $\eta_N(C, \beta)$ and $\zeta_N(\mathcal{C}, \mathcal{S}, \beta)$ defined by (12) and (18), respectively, are strictly decreasing for β, and they converge to 0 as β goes to 1. Moreover, t^* is strictly decreasing for k. Hence, $\eta_N(C, \beta)$ and $\zeta_N(\mathcal{C}, \mathcal{S}, \beta)$ are strictly increasing for $n(C, \mathcal{U}_N)$ and $m(\mathcal{C}, \mathcal{S}, \mathcal{U}_N)$, respectively, and they converge to $(1 - \beta)^{1/N}$ when $n(C, \mathcal{U}_N)$ and $m(\mathcal{C}, \mathcal{S}, \mathcal{U}_N)$ go to N. Please see Section 5.4 in [Badings et al., 2022a] for a more detailed discussion about the behavior of t^*.*

Remark 2 (Proof idea for Theorems 1 and 2). *Our probabilistic lower bounds for SPR-cause and recall-optimal probabilities are based on optimization formulations known as scenario optimization programs (SOPs) [Campi and Garatti, 2008, 2011]. We consider two SOPs whose constraints require that a given set of states $C \subseteq S$ is an SPR cause, and a given collection of sets of states $\mathcal{C} \subseteq 2^S$ contains a recall-optimal SPR cause, respectively, in some sampled MDPs. Please see Appendices A.1 and A.2 for details.*

4.2 PAC-GUARANTEED RECALL-OPTIMAL PROBABILITY-RAISING CAUSES ON UNCERTAIN PARAMETRIC MDPS

In this section, we provide the solution to Problem 3. For any $\beta, \delta \in (0, 1)$ and any $N > 0$, we define the collection $\mathcal{C}^*$ of subsets of states as

$$\mathcal{C}^* = \{ C^*_{S_N, u_i} \mid i \in \mathcal{I} \}, \qquad (20)$$

where $\mathcal{I}$ is a set of indices for sampled parameter vectors defined such that

$$\bigcup_{i \in \mathcal{I}} \mathcal{U}_N^i = \mathcal{U}_N \wedge \bigcup_{i \in \mathcal{I}'} \mathcal{U}_N^i \neq \mathcal{U}_N, \forall \mathcal{I}' \subset \mathcal{I}, \qquad (21)$$

$\mathcal{U}_N^i$ is defined as

$$\mathcal{U}_N^i = \{ u_j \in \mathcal{U}_N \mid C^*_{S_N, u_i} \subseteq \mathscr{C}_{u_j}$$
$$\wedge \forall \rho, \rho \models \Diamond C^*_{S_N, u_j} \wedge \Diamond E \implies \Diamond C^*_{S_N, u_i} \} \quad (22)$$

and $C^*_{S_N, u_i}$ is defined by (16) with $S_N = \cup_{C' \in \mathcal{S}_{N, \delta, \beta}} C'$.

Theorem 3. *For any upMDP $\mathcal{M}_{\mathbb{P}}$, any $\delta, \beta \in (0, 1)$, and any $N > 0$, $\mathcal{C}^*$ defined by (20) is a solution to Problem 3.*

In the following Example 2, we show how the condition in (22) eliminates redundant samples for the recall-optimal probability from all sampled parameters and how to construct the exhaustive solution $\mathcal{C}^*$ defined by (20).

Example 2. *We consider solving Problem 3 for the upMDP in Example 3. Let $\delta = 0$ and $\beta = 0.99$. We sampled $N = 1000$ parameters. We obtained a solution $\mathcal{C}^* = \{\{s_3\}\}$ with $\mathcal{S}_{N,\delta,\beta} = \{\{s_1\}, \{s_2\}, \{s_3\}, \{s_1, s_3\}, \{s_2, s_3\}\}$, where $\{s_3\}$, $\{s_1, s_3\}$, and $\{s_2, s_3\}$ are the canonical SPR causes on the sampled 7, 2, and 991 MDPs with $p = q$, $p > q$, and $p < q$, respectively. $\{s_3\}$ is a subset of $\{s_1, s_3\}$ and $\{s_2, s_3\}$, and all paths passing through $\{s_1, s_3\}$ or $\{s_2, s_3\}$ go through $\{s_3\}$. Thus, $\mathcal{U}_N^i$ defined by (22) that corresponds to $\{s_3\}$ is equal to $\mathcal{U}_N$. According to Theorem 1, the subsets of states $\{s_1\}$, $\{s_2\}$, $\{s_3\}$, $\{s_1, s_3\}$, and $\{s_2, s_3\}$ provide the lower bounds of SPR-cause probabilities 4.3×10^{-6}, 0.98, 0.995, 4.3×10^{-6}, and 0.98 with the confidence probability $\beta = 0.99$, respectively, and they exceed δ. For **(PC1)**, the lower bound $\zeta_N(\mathcal{C}^*, S_{N,\delta,\beta}, \beta)$ of recall-optimal probability is the maximum value $(1 - \beta)^{1/N} = 0.995$. For **(PC2)**, the subset $\mathcal{C} = \emptyset \subset \mathcal{C}^*$ clearly does not contain any SPR cause for all parameters. So, $R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}^*, S_{N,\delta,\beta}) > R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}, S_{N,\delta,\beta})$ holds.*

Remark 3 (Interpretation for (20)). *Intuitively, $\mathcal{U}_N^i$ defined by (22) collects all sampled parameters u_j such that $C^*_{S_N, u_i}$ for i-th sample u_i is recall optimal on $\mathcal{M}_{u_j}$ (See Lemma 5 in Appendix A.3). So, $\mathcal{C}^*$ is constructed to contain some recall-optimal SPR causes for each sampled MDP.*

An easier example is shown in Appendix D. Further, additional discussion about the interpretability of the resulting SPR causes is provided in Appendix C.

4.3 CAUSAL IDENTIFICATION ALGORITHM

We provide an algorithm to obtain the solution to Problem 3. The overall procedure is shown in Algorithm 1. In Lines from 1 to 6, we compute the canonical SPR cause $C^*_{S_N, u}$ defined as (16) for each sampled parameter vector, where $S_N = \cup_{C' \in \mathcal{S}_{N,\delta,\beta}} C'$, for each $\mathcal{M}_u$ with $u \in \mathcal{U}_N$ by applying Algorithm 2 to all states in $S \setminus E$. Then, in Line 8, we obtain $\mathcal{C}^*_{\mathcal{U}_N,\delta,\beta}$ defined by (20) as the solution to Problem 3. We compute the lower bounds of SPR-cause probabilities and recall-optimal probability through (12) and (18). Note that we can check the conditions of (16), (17), and (20) using a standard graph algorithm for path finding over $\mathcal{M}_u$. Algorithm 2 is identical to Algorithm 2 in [Baier et al., 2022b] and checks whether a given state $\{c\}$ is an SPR cause on a given MDP $\mathcal{M}_u$, that is, $(\mathcal{M}_u, \{c\}) \models$ (S). The computation of τ_u is sound and complete, i.e., $\tau_u(c) = 1$ if and only if $\{c\}$ is an SPR cause on $\mathcal{M}_u$, by Lemma 4 in [Baier et al., 2022b]. In Lines 1 and 2, we compute the minimum probability w_c of reaching E from c on $\mathcal{M}_u$ and the maximum probability q_s of reaching E from any $s \in S$

on a modified MDP $\mathcal{M}_u^{[c]}$. The computation is conducted by existing model checking algorithms such as value iteration. The modified MDP $\mathcal{M}_u^{[c]}$ is defined by removing all actions at c and adding a terminal state noeff that is not in E and an action γ that is enabled at c with the transition probabilities $P_u(c, \gamma, \text{eff}) = \text{Pr}^{\min}_{\mathcal{M}_u, c}(\Diamond E)$ and $P_u(c, \gamma, \text{noeff}) = 1 - \text{Pr}^{\min}_{\mathcal{M}_u, c}(\Diamond E)$, where eff is a fixed state in E and $\text{Pr}^{\min}_{\mathcal{M}_u, c}(\Diamond E) = \min_\pi \text{Pr}^\pi_{\mathcal{M}_u, c}(\Diamond E)$. $\tau_u(c)$ is calculated depending on the following 3 cases. When $w_c - q_{s_0} > 0$ and $w_c - q_{s_0} < 0$, we set $\tau_u(c) = 1$ and -1, respectively. In the corner case $w_c - q_{s_0} = 0$, we consider the set of actions $\mathcal{A}_{res}(s)\{a \in A | q_s = \sum_{s' \in S^{[c]}} P_u(s, a, s') q_{s'}\}$ for any state s, where $S^{[c]}$ is the set of states in $\mathcal{M}_u^{[c]}$. We construct a sub-MDP $\mathcal{M}_u^{\max, [c]}$ of $\mathcal{M}_u^{[c]}$ induced from $\mathcal{A}_{res}$. In Lines from 8 to 12, if c is reachable from s_0 in the sub-MDP $\mathcal{M}_u^{\max, [c]}$, then we set $\tau_u(c) = -1$, and otherwise $\tau_u(c) = 1$. The complexity of Algorithm 1 is polynomial for $|S|$ and N in practice. Please see Appendix B for details.

Algorithm 1 Probabilistically Guaranteed Recall-Optimal SPR Cause Identification for upMDP

Require: upMDP $\mathcal{M}_{\mathbb{P}}$, $N > 0$, $\delta \geq 0$, and $\beta \in [0, 1]$.
Ensure: $\mathcal{C}^*$, η_N, and ζ_N.
 1: Sample N parameters $\mathcal{U}_N$ from $\mathcal{V}$ according to $\mathbb{P}$.
 2: **for** each $u \in \mathcal{U}_N$ **do**
 3:    **for** each $c \in S \setminus E$ **do**
 4:       $\tau_u(c) \leftarrow$ Algorithm2$(\mathcal{M}_u, c)$.
 5:    **end for**
 6:    Compute $\mathscr{C}_{S_N, u}$ as $\mathscr{C}_{S_N, u} = \{c \in S | \tau_u(c) = 1\}$ and $C^*_{S_N, u}$ defined by (16), where $S_N = \cup_{C' \in \mathcal{S}_{N,\delta,\beta}} C'$.
 7: **end for**
 8: Initialize $\mathcal{I} = \emptyset$.
 9: **while** $\bigcup_{i \in \mathcal{I}} \mathcal{U}_N^i \neq \mathcal{U}_N$. **do**
10:    $\mathcal{I} \leftarrow \mathcal{I} \cup \{i\}$ such that $\mathcal{U}_N^i \setminus \bigcup_{j \in \mathcal{I}} \mathcal{U}_N^j \neq \emptyset$.
11: **end while**
12: Compute $\mathcal{C}^*$ from $\mathcal{I}$.
13: Compute η_N and ζ_N for $\mathcal{C}^*$ and $S_{N,\delta,\beta}$ using Theorems 1 and 2.

---

5 EXAMPLE

5.1 SETUP

We demonstrate that the proposed method exhaustively and nonredundantly finds SPR-causes on a 10×10 grid world with 3 obstacles, where the robot behaves as depicted in Figs. 3(a) and 3(b). The undesired behavior is reaching the terminal red cells. The blue arrows represent the enabled actions for each set of cells surrounded by the green dashed line. The actions the robot can take are Up, Right, Down, Left, and Stay. In Fig. 3(a), the cells $\{(4, i)\}_{i=6,7,8}$ is the one-way aisle where the robot is allowed to choose to go up. On the other hand, in Fig. 3(b),

Algorithm 2 SPR Cause Checking for Single State

Require: MDP $\mathcal{M}_u$, $c \in S$.
Ensure: $\tau_u(c) \in \{-1, 1\}$.
1: Compute $w_c = \min_\pi \mathrm{Pr}^\pi_{\mathcal{M}_u, c}(\Diamond E)$.
2: Compute $q_s = \max_\pi \mathrm{Pr}^\pi_{\mathcal{M}_u^{[c]}, s}(\Diamond E)$ for each $s \in S$.
3: **if** $w_c - q_{s_0} > 0$ **then**
4: $\quad \tau_u(c) = 1$
5: **else if** $w_c - q_{s_0} < 0$ **then**
6: $\quad \tau_u(c) = -1$.
7: **else if** $w_c - q_{s_0} = 0$ **then**
8: $\quad$ **if** c is reachable from s_0 in $\mathcal{M}_u^{\max, [c]}$ **then**
9: $\quad\quad \tau_u(c) = -1$.
10: $\quad$ **else**
11: $\quad\quad \tau_u(c) = 1$.
12: $\quad$ **end if**
13: **end if**

---

the robot can choose to move up and down the aisle. The robot transits in the intended direction with probability p_0 and moves perpendicularly with probability $(1 - p_0)/2$. The robot always stays in the same cell when it chooses Stay or is about to collide with the obstacles or go outside. The robot enters the red cell with probability p_1 (resp., p_2) regardless of the selected action when it is in the cell $(9, 5)$ (resp., $(8, 9)$) and moves in the intended way with the remaining probability. The robot moves carefully in the intended direction with probability 1 in the cells adjacent to the red cells, other than $(9, 5)$ and $(8, 9)$. We set the probability distributions followed by p_0, p_1, and p_2 as the uniform distributions over the intervals $[0.85, 0.9]$, $[0.45, 0.6]$, and $[0.5, 0.7]$. We sampled $N = 100$ parameter vectors from the parameter space. We conducted the experiments 10 times for each environment with $\beta = \beta_0 (= 0.9)$, $\beta_1 (= 0.99)$ and $\delta = \delta_0 (= 0.001)$, $\delta_1 (= 0.1)$. We implemented Algorithm 2 based on value iteration to compute w_c and q_s and to check some path conditions. We computed η_N and ζ_N defined in (12) and (18), respectively, using bi-section search.

SPR causes are cells whose traversal increases the probability of reaching red cells. All true SPR causes when $p_1 < p_2$ and $p_1 > p_2$ are shown as the yellow and orange/green cells, respectively. This is because crossing the yellow cells leads to entering the red cells by going through the top side when $p_1 < p_2$. Likewise, passing through the orange/green cells causes entry into the red cells from below when $p_1 > p_2$. Our method identifies and consolidates them through model checking for sampled MDPs and set covering. A detailed explanation for SPR causes in Example is described in Appendix E.

We compared our proposed method with two naïve approaches: NA1 and NA2. NA1 computes the canonical SPR cause $C^*_{S,u}$ defined by (16) based on Algorithm 2 in [Baier et al., 2022b] for the mean parameters. NA2 collects the canonical SPR causes $C^*_{S,u}$ for 8 parameter vectors

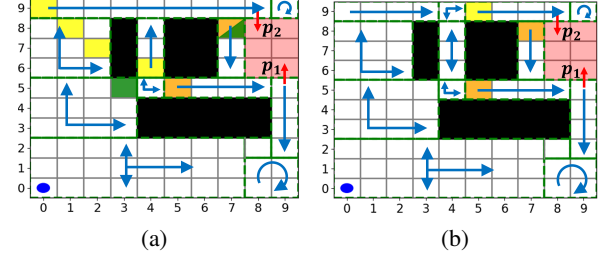corresponding to the vertices of the parameter region.

Figure 3: Two grid-worlds with three obstacles (black rectangles). The robot starts from $(0, 0)$. It enters the terminal red cells by the red arrow transitions with the indicated probabilities p_1 and p_2. The blue arrows denote the actions enabled in the regions. The yellow and orange/green cells denote the SPR cause when $p_1 < p_2$ and $p_1 > p_2$, respectively, obtained from our method with $\delta = 0.001$. The difference between the two environments is the enabled actions in the top middle cells.

5.2 RESULTS

We show that the proposed method finds solutions to Problem 3 well for various sample sizes. We show the results for δ_0 and β_1 (See Appendix F for more results). Fig. 4(a) shows the recall-optimal probabilities R and their lower bounds ζ of the solution $\mathcal{C}^*$ obtained from the proposed method as blue and magenta lines, respectively, for each sample size N with $\mathcal{S}_{N,\delta,\beta}$ defined by (9). $R(\mathcal{M}_\mathbb{P}, \mathcal{C}^*, \mathcal{S}_{N,\delta,\beta})$ is approximately computed by Monte Carlo method in (5) with 1000 samples. We observe that ζ bounds R from below and exceeds δ_0 for all sample sizes, and the error between them decreases towards 0. Moreover, the lower bounds are always obtained as the maximal values $(1 - \beta)^{1/N}$. This implies that the proposed method achieves **(PC1)** in Problem 3. Then, we compare R with the maximum recall-optimal probabilities R_{sub} over all proper subsets of the solution $\mathcal{C}^*$ shown in Fig. 4(a) (green lines). R_{sub} are always lower than R, and thus the obtained solution achieves **(PC2)** well. Fig. 4(b) (resp., 4(c)) shows the maximum (resp., minimum) SPR-cause probabilities $F_{\max}$ (resp., $F_{\min}$) and the maximum (resp., minimum) lower bounds $\eta_{\max}$ (resp., $\eta_{\min}$) over all state sets C in the obtained solution $\mathcal{C}^*$ as blue and green lines, respectively, for each sample size N. Likewise R, $F_{\max}$ and $F_{\min}$ are approximately calculated with 1000 samples. $\eta_{\max}$ (resp., $\eta_{\min}$) bounds $F_{\max}$ (resp., $F_{\min}$) from blow for each sample size. The error between η and F decreases with sample size.

To investigate the relations between obtained SPR causes and the system behavior, we depict $\mathcal{C}^*$ obtained from the proposed method with $\delta = 0.001$ and $N = 100$ for the last experiment in Figs. 3(a) and 3(b), and they correspond to the yellow, orange, and green cells. The sets of the yellow and
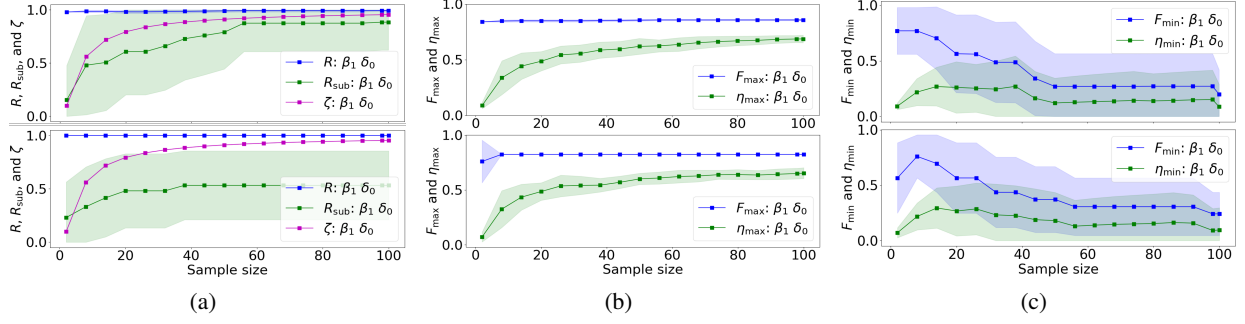
Figure 4: Results for δ_0 and β_1. (a): The means and standard deviations of the recall-optimal probabilities R and their lower bounds ζ for the solutions obtained from the proposed method against each sample size (blue and magenta). We further show the maximum recall-optimal probabilities R_{sub} for all proper subsets of the solutions (green). (b) (resp., (c)): The means and standard deviations of the maximum (resp., minimum) SPR-cause probabilities and their lower bounds over all potential SPR causes in the solutions obtained from the proposed method against each sample size. The results of the upper and the lower figures in (a)-(c) correspond to two environments depicted in Fig 3(a) and 3(b), respectively.

orange/green cells correspond to SPR causes with $p_1 < p_2$ and $p_1 > p_2$, respectively. In Fig. 3(a), we observed that our method found all possible SPR causes. Moreover, by modifying the one-way aisle in the top middle into a two-way aisle in Fig. 3(b), the size of the potential SPR causes is reduced.

We compare the proposed method with the naïve approaches. NA1 only found SPR causes $\{(0, 8), (1, 8), (2, 7), (4, 6)\}$ in Fig. 3(a), and could not find SPR causes associated with $p_1 > p_2$. Thus, the recall-optimal probabilities obtained from NA1 are lower than 1 when $\delta = \delta_0$. In contrast, the proposed method further found $\{(3, 5), (7, 8)\}$ and $\{(5, 5), (7, 8)\}$, and hence achieves (**PC1**). NA2 additionally found SPR causes $\{(3, 5), (7, 8)\}$ and $\{(5, 5), (7, 8)\}$ in the environment in Fig. 3(a), but removing them does not reduce the recall-optimal probability when $\delta = \delta_1$. Hence, the solution provided by NA2 is redundant in the sense of (**PC2**). The proposed method only provides $\{(0, 9), (1, 8), (2, 7), (4, 6)\}$ when $\delta = \delta_1$ and thus achieves the nonredundancy.

6 RELATED WORKS

6.1 FORMAL VERIFICATION AND SYNTHESIS FOR UNCERTAIN SYSTEMS

Formal verification and synthesis for epistemically uncertain systems have been successfully advocated to guarantee that the system satisfies a given specification [Cubuktepe et al., 2020, Badings et al., 2022a,b, Nilim and El Ghaoui, 2005, Wiesemann et al., 2013, Wolff et al., 2012, Rickard et al., 2024]. The existing works, however, have not considered causal analysis. Cubuktepe et al. [2020], Badings et al. [2022a,b] have developed the PAC-guaranteed verification methods for upMDPs and upMCs based on SOP formulation using samples of model parameters. The synthesis of

probabilistically guaranteed control policies on upMDPs has also been investigated using optimization and learning based on system behavior samples in [Nilim and El Ghaoui, 2005, Wiesemann et al., 2013, Wolff et al., 2012, Rickard et al., 2024]. Identifications of uncertain MDPs also have been deeply studied using Bayesian learning [Dearden et al., 1999, Ross et al., 2007, Suilen et al., 2022] and data-driven abstractions [Jackson et al., 2020, Lavaei et al., 2022].

6.2 CAUSAL ANALYSIS FOR DETERMINISTIC AND STOCHASTIC SYSTEMS

Causal analysis for deterministic and stochastic systems has been intensively investigated [Baier et al., 2024a, Coenen et al., 2022, Finkbeiner et al., 2024, Dimitrova et al., 2020, Finkbeiner and Siber, 2023, Kazemi and Paoletti, 2022, Baier et al., 2024b, Oberst and Sontag, 2019, Tsirtsis et al., 2021, Triantafyllou et al., 2022, Baier et al., 2021c]. However, existing works have not considered the epistemic uncertainty of the system model itself. Baier et al. [2024a] have extended the PR cause-effect relations in [Baier et al., 2021b, 2022b] based on linear temporal logic. The PR causality and their model checking-based identification have the advantage of an explainable verification for probabilistic systems. Coenen et al. [2022], Finkbeiner et al. [2024] have considered temporal causality on transition systems using omega-regular properties. An action-level PR causality has been considered on MDPs in [Dimitrova et al., 2020] based on hyperproperty [Clarkson et al., 2014]. Aside from the PR principle, another well-known causality is based on *counterfactuality*. Counterfactual reasoning operators have recently been incorporated into some temporal logic frameworks in [Finkbeiner and Siber, 2023, Kazemi and Paoletti, 2022]. To perform counterfactual reasoning on MDPs, Oberst and Sontag [2019], Tsirtsis et al. [2021], Triantafyllou et al. [2024, 2022] have conducted encoding of MDPs using structural

causal models. As a complementary concept for causality, responsibility attribution has been investigated in [Baier et al., 2024b, Triantafyllou et al., 2022, Baier et al., 2021c].

7 CONCLUSION

This paper presented a novel approach to exhaustively and nonredundantly identify subsets of states as potential probability-raising (PR) causes on an uncertain parametric Markov decision process. Based on model checking and a set covering strategy for the sampled parameters, the identified potential PR causes satisfy the following three properties in a probabilistically approximately correct manner: (i) The probability of each identified subset being a PR cause is at least a specified value. (ii) A lower bound of the probability that the undesired paths pass through the subsets as much as possible is maximized. (iii) Eliminating any identified subset degrades the lower bound. We plan to extend the proposed method to concurrent systems.

References

Sebastian Arming, Ezio Bartocci, Krishnendu Chatterjee, Joost-Pieter Katoen, and Ana Sokolova. Parameter-independent strategies for pmdps via pomdps. In *Proceedings of the 15th International Conference on Quantitative Evaluation of Systems*, pages 53–70. Springer, 2018.

Thom Badings, Murat Cubuktepe, Nils Jansen, Sebastian Junges, Joost-Pieter Katoen, and Ufuk Topcu. Scenario-based verification of uncertain parametric mdps. *International Journal on Software Tools for Technology Transfer*, 24(5):803–819, 2022a.

Thom S Badings, Nils Jansen, Sebastian Junges, Marielle Stoelinga, and Matthias Volk. Sampling-based verification of ctmcs with uncertain rates. In *Proceedings of the 34th International Conference on Computer Aided Verification*, pages 26–47. Springer, 2022b.

Christel Baier and Joost-Pieter Katoen. *Principles of Model Checking*. MIT press, 2008.

Christel Baier, Clemens Dubslaff, Florian Funke, Simon Jantsch, Rupak Majumdar, Jakob Piribauer, and Robin Ziemek. From verification to causality-based explications. In *Proceeding of the 48th International Colloquium on Automata, Languages, and Programming*, pages 1–1, 2021a.

Christel Baier, Florian Funke, Simon Jantsch, Jakob Piribauer, and Robin Ziemek. Probabilistic causes in Markov chains. In *Proceedings of the 19th International Symposium on Automated Technology for Verification and Analysis*, pages 205–221, 2021b.

Christel Baier, Florian Funke, and Rupak Majumdar. A game-theoretic account of responsibility allocation. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence*, pages 1773–1779. IJCAI, 2021c.

Christel Baier, Clemens Dubslaff, Florian Funke, Simon Jantsch, Jakob Piribauer, and Robin Ziemek. Operational causality–necessarily sufficient and sufficiently necessary. In *A Journey from Process Algebra via Timed Automata to Model Learning: Essays Dedicated to Frits Vaandrager on the Occasion of His 60th Birthday*, pages 27–45. Springer, 2022a.

Christel Baier, Florian Funke, Jakob Piribauer, and Robin Ziemek. On probability-raising causality in Markov decision processes. *arXiv preprint arXiv:2201.08768*, 2022b.

Christel Baier, Jakob Piribauer, and Robin Ziemek. Foundations of probability-raising causality in Markov decision processes. *Logical Methods in Computer Science*, 20, 2024a.

Christel Baier, Roxane van den Bossche, Sascha Klüppelholz, Johannes Lehmann, and Jakob Piribauer. Backward responsibility in transition systems using general power indices. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 20320–20327, 2024b.

Marco C Campi and Simone Garatti. The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, 19(3):1211–1230, 2008.

Marco C Campi and Simone Garatti. A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality. *Journal of Optimization Theory and Applications*, 148(2):257–280, 2011.

Michael R Clarkson, Bernd Finkbeiner, Masoud Koleini, Kristopher K Micinski, Markus N Rabe, and César Sánchez. Temporal logics for hyperproperties. In *Proceedings of the 3rd International Conference on Principles of Security and Trust*, pages 265–284. Springer Verlag, 2014.

Norine Coenen, Bernd Finkbeiner, Hadar Frenkel, Christopher Hahn, Niklas Metzger, and Julian Siber. Temporal causality in reactive systems. In *Proceedings of the 20th International Symposium on Automated Technology for Verification and Analysis*, pages 208–224. Springer, 2022.

Murat Cubuktepe, Nils Jansen, Sebastian Junges, Joost-Pieter Katoen, and Ufuk Topcu. Scenario-based verification of uncertain mdps. In *Proceedings of the 26th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 287–305. Springer, 2020.

Richard Dearden, Nir Friedman, and David Andre. Model based Bayesian exploration. In *Proceedings of the*

15th Conference on Uncertainty in Artificial Intelligence, pages 150–159, 1999.

Benoit Delahaye, Kim G Larsen, Axel Legay, Mikkel L Pedersen, and Andrzej Wasowski. Decision problems for interval Markov chains. In *Proceedings of the 5th International Conference on Language and Automata Theory and Applications*, pages 274–285. Springer, 2011.

Rayna Dimitrova, Bernd Finkbeiner, and Hazem Torfah. Probabilistic hyperproperties of Markov decision processes. In *Proceedings of the 18th International Symposium on Automated Technology for Verification and Analysis*, pages 484–500. Springer, 2020.

Ali Dorri, Salil S Kanhere, and Raja Jurdak. Multi-agent systems: A survey. *IEEE Access*, 6:28573–28593, 2018.

Gregory Dudek, Michael RM Jenkin, Evangelos Milios, and David Wilkes. A taxonomy for multi-agent robotics. *Autonomous Robots*, 3:375–397, 1996.

Bernd Finkbeiner and Julian Siber. Counterfactuals modulo temporal logics. In *Proceedings of the 24th International Conference on Logic*, volume 94, pages 181–204, 2023.

Bernd Finkbeiner, Hadar Frenkel, Niklas Metzger, and Julian Siber. Synthesis of temporal causality. In *Proceedings of the 36th International Conference on Computer Aided Verification*, pages 87–111. Springer, 2024.

Robert Givan, Sonia Leach, and Thomas Dean. Bounded-parameter Markov decision processes. *Artificial Intelligence*, 122(1-2):71–109, 2000.

John Jackson, Luca Laurenti, Eric Frew, and Morteza Lahijanian. Safety verification of unknown dynamical systems via Gaussian process regression. In *Proceedings of the 59th Conference on Decision and Control*, pages 860–866. IEEE, 2020.

Milad Kazemi and Nicola Paoletti. Causal temporal reasoning for Markov decision processes. *arXiv preprint arXiv:2212.08712*, 2022.

Abolfazl Lavaei, Sadegh Soudjani, Emilio Frazzoli, and Majid Zamani. Constructing mdp abstractions using data with formal guarantees. *IEEE Control Systems Letters*, 7:460–465, 2022.

Raha Moraffah, Mansooreh Karami, Ruocheng Guo, Adrienne Raglin, and Huan Liu. Causal interpretability for machine learning-problems, methods and evaluation. *ACM SIGKDD Explor. Newslet.*, 22(1):18–33, 2020.

Arnab Nilim and Laurent El Ghaoui. Robust control of Markov decision processes with uncertain transition matrices. *Operations Research*, 53(5):780–798, 2005.

Michael Oberst and David Sontag. Counterfactual off-policy evaluation with gumbel-max structural causal models. In *Proceedings of the 15th International Conference on Machine Learning*, pages 4881–4890. PMLR, 2019.

Alberto Puggelli, Wenchao Li, Alberto L Sangiovanni-Vincentelli, and Sanjit A Seshia. Polynomial-time verification of pctl properties of mdps with convex uncertainties. In *Proceedings of the 25th International Conference on Computer Aided Verification*, pages 527–542. Springer, 2013.

Luke Rickard, Alessandro Abate, and Kostas Margellos. Learning robust policies for uncertain parametric Markov decision processes. In *Proceedings of the 6th Annual Learning for Dynamics & Control Conference*, pages 876–889. PMLR, 2024.

Stephane Ross, Brahim Chaib-draa, and Joelle Pineau. Bayes-adaptive pomdps. *Advances in neural information processing systems*, 20:1225–1232, 2007.

Philippe Schnoebelen. The complexity of temporal logic model checking. *Adv. Modal Logic*, 4(35):393–436, 2002.

Marnix Suilen, Thiago D Simão, David Parker, and Nils Jansen. Robust anytime learning of Markov decision processes. *Advances in Neural Information Processing Systems*, 35:28790–28802, 2022.

Stelios Triantafyllou, Adish Singla, and Goran Radanovic. Actual causality and responsibility attribution in decentralized partially observable Markov decision processes. In *Proceedings of the 5th AAAI/ACM Conference on AI, Ethics, and Society*, pages 739–752, 2022.

Stelios Triantafyllou, Aleksa Sukovic, Debmalya Mandal, and Goran Radanovic. Agent-specific effects: A causal effect propagation analysis in multi-agent mdps. In *Proceedings of 41st International Conference on Machine Learning*, 2024.

Stratis Tsirtsis, Abir De, and Manuel Rodriguez. Counterfactual explanations in sequential decision making under uncertainty. *Advances in Neural Information Processing Systems*, 34:30127–30139, 2021.

Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem. Robust Markov decision processes. *Mathematics of Operations Research*, 38(1):153–183, 2013.

Eric M Wolff, Ufuk Topcu, and Richard M Murray. Robust control of uncertain Markov decision processes with temporal logic specifications. In *Proceedings of the 51st Conference on Decision and Control*, pages 3372–3379. IEEE, 2012.

Supplementary Material

Ryohei Oura[1] **Yuji Ito**[2]

[1]Frontier Research Center, Toyota Motor Corporation, Japan
[2]Toyota Central R&D Labs., Inc., Japan

A DERIVATION OF MAIN THEOREMS

A.1 SCENARIO OPTIMIZATION

In this section, we briefly review the scenario optimization programs (SOPs) [Campi and Garatti, 2008, 2011], and how the obtained solutions are guaranteed with PAC bounds. We consider a family of constraints $K^u \subseteq \mathbb{R}^d$ with a natural number d parameterized by $u \in \mathcal{V}$, where $\mathcal{V}$ is an uncountable (and possibly continuous) parameter space, and K^u is convex and closed for each u. We assume that $\mathcal{V}$ is endowed with a σ-algebra $\mathcal{B}(\mathcal{V})$ and a probability measure $\mathbb{P}$. Suppose that N i.i.d. samples sequence $\mathcal{U}_N = (u_i)_{i=1,\ldots,N}$ are available.

For any $\mathcal{U}'_N \subseteq \mathcal{U}_N$, we consider the following optimization program called SOP:

$$\max_{\gamma} \quad \gamma$$
$$\text{s.t.} \quad \gamma \in \bigcap_{u \in \mathcal{U}'_N} K^u, \tag{23}$$

We introduce the following assumptions.

Assumption 1 (Assumption 2.1 in [Campi and Garatti, 2011]). *Every optimization problem subject to a finite set $F \subset \mathcal{V}$, i.e.,*

$$\max_{\gamma} \quad \gamma$$
$$\text{s.t.} \quad \gamma \in \bigcap_{u \in F} K^u, \tag{24}$$

is feasible, and its feasibility domain has a nonempty interior. Moreover, the solution of (24) exists and is unique.

Assumption 2 (Assumption 2.2 in [Campi and Garatti, 2011]). *Almost surely with respect to multi-sample $\mathcal{U}_N$, the solution γ^* of the optimization program in (23) violates constraints for all $u \in \mathcal{U}_N \setminus \mathcal{U}'_N$.*

Then, we have the following arguments.

Theorem 4 (Theorem 1 in [Campi and Garatti, 2008]). *Under Assumption 1, if $|\mathcal{U}'_N| = 0$, then we have*

$$\mathbb{P}^N \left\{ \mathbb{P}(\gamma^* \notin K^u) > \varepsilon \right\} < \sum_{i=0}^{d-1} \varepsilon^i (1-\varepsilon)^{N-i}, \tag{25}$$

where γ^* is the solution to (23).

Theorem 5 (Theorem 2.1 in [Campi and Garatti, 2011]). *Under Assumptions 1 and 2, we have that*

$$\mathbb{P}^N \{\mathbb{P}(\gamma^* \notin K^u) > \varepsilon\} < \binom{k+d-1}{k} \sum_{i=0}^{k+d-1} \varepsilon^i (1-\varepsilon)^{N-i}, \tag{26}$$

where γ^* is the solution to (23).

A.2 DERIVATION OF THEOREMS 1 AND 2

In this section, we prove Theorems 1 and 2 based on scenario optimization program (SOP) formulations. We first review the following lemma to show the necessary and sufficient condition of SPR causes.

Lemma 1 (Lemma 21 in [Baier et al., 2022b]). *For any pMDP $\mathcal{M}_u$, any $C \subset S$, and any $u_j \in \mathcal{V}$, C is an SPR cause on $\mathcal{M}_{u_j}$ if and only if $C \subseteq \mathscr{C}_{S,u_j}$ and C satisfies (M) in Def.1, where $\mathscr{C}_{S,u_j}$ is defined by (10).*

For any MDP $\mathcal{M}_u$ and any $C \subset S$, we define γ_C^u as

$$\gamma_C^u = \begin{cases} 1 & \text{if } C \text{ is an SPR cause on } \mathcal{M}_u, \\ 0 & \text{otherwise,} \end{cases} \tag{27}$$

For any $C \subseteq S$, let $K_C^u = [0, \gamma_C^u]$. For any sequence of N sampled parameters $\mathcal{U}_N \in \mathcal{V}^N$, we define $\mathcal{U}_{N,1}^C = \{u \in \mathcal{U}_N \mid C \subseteq \mathscr{C}_{S,u}, (\mathcal{M}_{\mathbb{P}}, C) \models (M)\}$, where we treat the same elements in $\mathcal{U}_N$ as duplicates. Then, we define $\mathcal{U}_{N,0}^C = \mathcal{U}_N \setminus \mathcal{U}_{N,1}^C$, where $\mathscr{C}_{S,u}$ is defined as (10). Then, we consider the following SOP:

$$\begin{aligned} \max_{\gamma} \quad & \gamma \\ \text{s.t.} \quad & \gamma \in \bigcap_{u \in \mathcal{U}_N'} K_C^u, \end{aligned} \tag{28}$$

where $\mathcal{U}_N' \subseteq \mathcal{U}_N$. Let γ_C^* be the solution of (28). Note that $\gamma_C^* = \min_{u \in \mathcal{U}_N'} \gamma_C^u$.

proof of Theorem 1. For any $C \subset S$, we consider the two cases (1) $\mathcal{U}_{N,0}^C = \emptyset$ and (2) $\mathcal{U}_{N,0}^C \neq \emptyset$.

(1): We set $\mathcal{U}_N' = \mathcal{U}_{N,1}^C$. By Lemma 1, for any $u \in \mathcal{U}_{N,1}^C$, C is an SPR cause on $\mathcal{M}_u$. Thus, the intersection of all K_C^u with $u \in \mathcal{U}_{N,1}^C$ is $[0,1]$ and hence there exists a unique solution $\gamma_C^* = 1$ to (28). Thus, for any $\varepsilon > 0$, by Theorem 4, we have

$$\mathbb{P}^N \{\mathbb{P}(\gamma_C^* \notin K_C^u) > \varepsilon\} < (1-\varepsilon)^N, \tag{29}$$

and hence

$$\mathbb{P}^N \{\mathbb{P}(\gamma_C^* \in K_C^u) \geq 1 - \varepsilon\} \geq 1 - (1-\varepsilon)^N, \tag{30}$$

where the internal probability $\mathbb{P}$ measures the frequency of u for K_C^u. We choose $\beta = 1 - (1-\varepsilon)^N$ and hence

$$\mathbb{P}^N \left\{\mathbb{P}(\gamma_C^u \geq \gamma_C^*) \geq (1-\beta)^{1/N}\right\} \geq \beta. \tag{31}$$

We have that $\gamma_C^* = 1$ and thus

$$\mathbb{P}^N \left\{\mathbb{P}(\gamma_C^u = 1) \geq (1-\beta)^{1/N}\right\} \geq \beta. \tag{32}$$

Therefore, we have

$$\mathbb{P}^N \left\{F(\mathcal{M}_{\mathbb{P}}, C) \geq (1-\beta)^N\right\} \geq \beta. \tag{33}$$

(2): For any $\mathcal{U}'_N \neq \emptyset$ in (28), the intersection of all K^u_C with $u \in \mathcal{U}'_N$ is nonempty and there exists a unique solution γ^*_C to (28). Thus, for any $\varepsilon > 0$ and any a-priori number of discarded sampled parameters $q = |\mathcal{U}_N \setminus \mathcal{U}'_N|$, by Theorem 5, we have

$$
\mathbb{P}^N \{\mathbb{P}(\gamma^*_C(q) \notin K^u_C) \leq \varepsilon\} = \mathbb{P}^N \{\mathbb{P}(\gamma^*_C(q) \in K^u_C) \geq 1 - \varepsilon\}
$$
$$
\geq 1 - \sum_{i=0}^{q} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i}. \tag{34}
$$

We set $t = 1 - \varepsilon$ and equate the right-hand side of (34) to $1 - (1-\beta)/N$. Then, we have

$$
\mathbb{P}^N \{\mathbb{P}(\gamma^u_C \geq \gamma^*_C(q)) \geq t^*(q)\} \geq 1 - \frac{1-\beta}{N}, \tag{35}
$$

where $t^*(q)$ is the solution of

$$
\frac{1-\beta}{N} = \sum_{i=0}^{q} \binom{N}{i} (1-t)^i t^{N-i}. \tag{36}
$$

We consider the case $q = |\mathcal{U}^C_{N,0}|$ to derive the lower bound of $F(\mathcal{M}_\mathbb{P}, C)$ for $\mathcal{U}'_N = \mathcal{U}^C_{N,1}$. Since the number $|\mathcal{U}^C_{N,0}|$ is not known before observing actual samples, we simultaneously consider the bound for all $q = 1, \ldots, N-1$. By Bool's inequality and (35), we have

$$
\mathbb{P}^N \{\cap_{q=1}^N \mathcal{E}_q\} = 1 - \mathbb{P}^N \{\cup_{q=1}^N \mathcal{E}^c_q\}
$$
$$
\geq 1 - \sum_{q=1}^N \mathbb{P}^N \{\mathcal{E}^c_q\}
$$
$$
\geq \beta, \tag{37}
$$

where $\mathcal{E}_q$ denotes the event $\mathbb{P}(\gamma^u_C \geq \gamma^*_C(q)) \geq t^*(q)$ and $\mathcal{E}^c_q$ is its complement event. Clearly, we have

$$
\mathbb{P}^N \{\mathbb{P}(\gamma^u_C \geq \gamma^*_C) \geq t^*(|\mathcal{U}^C_{N,0}|)\} \geq \mathbb{P}^N \{\forall q' = 1, \ldots, N, \mathbb{P}(\gamma^u_C \geq \gamma^*_C) \geq t^*(q')\}
$$
$$
\geq \beta. \tag{38}
$$

Note that we can take $\mathcal{U}'_N = \mathcal{U}^C_{N,1}$ when $q = |\mathcal{U}^C_{N,0}|$. By Lemma 1, for any $u \in \mathcal{U}^C_{N,1}$, C is an SPR cause on $\mathcal{M}_u$. Thus, the intersection of all K^u_C with $u \in \mathcal{U}^C_{N,1}$ is $[0, 1]$ and hence there exists a unique solution $\gamma^*_C = 1$ to (28) when $\mathcal{U}'_N = \mathcal{U}^C_{N,1}$. Thus, we have

$$
\mathbb{P}^N \{\mathbb{P}(\gamma^u_C \geq \gamma^*_C) = t^*(|\mathcal{U}^C_{N,0}|)\} = \mathbb{P}^N \{\mathbb{P}(\gamma^u_C = 1) \geq t^*(|\mathcal{U}^C_{N,0}|)\}
$$
$$
= \mathbb{P}^N \{F(\mathcal{M}_\mathbb{P}, C) \geq t^*(|\mathcal{U}^C_{N,0}|)\}. \tag{39}
$$

Therefore, by (38), we have

$$
\mathbb{P}^N \{F(\mathcal{M}_\mathbb{P}, C) \geq t^*(|\mathcal{U}^C_{N,0}|)\} \geq \beta. \tag{40}
$$

$\square$

We show the following lemmas before proving Theorem 2. For any path property φ and any path ρ of an MDP, we denote $\rho \models \varphi$ if and only if ρ satisfies φ. Moreover, recall that we denote all finite paths on the pMDP as $\mathrm{FinPath}$.

Lemma 2. *For any MDP $\mathcal{M}_u$, any $S' \subseteq S$, any SPR cause C' on $\mathcal{M}_u$ that is a subset of S', we have that*

$$
\forall \rho \in \mathrm{FinPath}, \rho \models \Diamond C' \implies \Diamond C^*_{S',u}, \tag{41}
$$

where $C^*_{S',u}$ is defined by (16).

Proof. Suppose that $\rho \models \neg\Diamond C^*_{S',u} \wedge \Diamond C'$. This implies that there exists $s \in C'$ such that $\rho \models \neg C^*_{S',u} \mathrm{U} s$. By (16), for any $c \in \mathscr{C}_{S',u} \setminus C^*_{S',u}$ and for any $\rho \in \mathrm{FinPath}$, we have $\rho \not\models \neg\mathscr{C}_{S',u} \mathrm{U} c$. Thus, there exist $c' \in \mathscr{C}_{S',u}$ with $c' \neq c$ and a path ρ' such that $\rho' \models \Diamond(c' \wedge \Diamond c)$ and $\rho' \models \neg\mathscr{C}_{S',u} \mathrm{U} c'$. This implies that, for any path ρ, we have $\rho \not\models \neg C^*_{S',u} \mathrm{U} c$ since $C^*_{S',u}$ is constructed from (16). However, this contradicts the assumption. $\square$

Lemma 3. *For any MDP $\mathcal{M}_u$, any $S' \subset S$, and any SPR cause $C \subseteq S'$, $C^*_{S',u}$ is recall optimal over $\mathcal{S}$ for any $\mathcal{S} \subseteq 2^{S'}$, where $C^*_{S',u}$ is defined as (16).*

Proof. By Lemma 2, for any policy π and any SPR cause $C' \in \mathcal{S}$ on $\mathcal{M}_u$, we have that

$$\mathrm{Pr}^\pi_{\mathcal{M}_u}(\Diamond C' \wedge \Diamond E) \leq \mathrm{Pr}^\pi_{\mathcal{M}_u}(\Diamond C^*_{S',u} \wedge \Diamond E). \tag{42}$$

This yields the claim. $\qquad\square$

Lemma 4. *For any MDP $\mathcal{M}_u$, any $S' \subseteq S$, and any $C \subseteq \mathscr{C}_{S',u}$ that satisfies **(M)** in Def. 1, if (43) holds, then C is recall-optimal on $\mathcal{M}_u$ over $\mathcal{S} = 2^{S'}$.*

$$\forall \rho \in \mathrm{FinPath}, \rho \models \Diamond C^*_{S',u} \wedge \Diamond E \implies \Diamond C, \tag{43}$$

where $\mathscr{C}_{S',u}$ is defined by (10), and $C^*_{S',u}$ is defined by (16).

Proof. By (43), we have

$$\forall \pi, \mathrm{Pr}^\pi_{\mathcal{M}_{u_j}}(\Diamond C^*_{S',u_j} \mid \Diamond E) \leq \mathrm{Pr}^\pi_{\mathcal{M}_{u_j}}(\Diamond C \mid \Diamond E). \tag{44}$$

By Lemma 3, C^*_{S',u_j} is recall optimal on $\mathcal{M}_{u_j}$ over $\mathcal{S}$. Thus, by (44), for any SPR cause $C' \subseteq S'$, we have

$$\forall \pi, \mathrm{Pr}^\pi_{\mathcal{M}_{u_j}}(\Diamond C' \mid \Diamond E) \leq \mathrm{Pr}^\pi_{\mathcal{M}_{u_j}}(\Diamond C^*_{S',u_j} \mid \Diamond E) \leq \mathrm{Pr}^\pi_{\mathcal{M}_{u_j}}(\Diamond C \mid \Diamond E). \tag{45}$$

By Lemma 1, C is SPR cause on $\mathcal{M}_{u_j}$. Therefore, C is recall optimal on $\mathcal{M}_{u_j}$ over $\mathcal{S}$. $\qquad\square$

For any MDP $\mathcal{M}_u$ and any $\mathcal{C}, \mathcal{S} \subseteq 2^S$, we define $\xi^u_{\mathcal{C},\mathcal{S}}$ as

$$\xi^u_{\mathcal{C},\mathcal{S}} = \begin{cases} 1 & \text{if } \exists C \in \mathcal{C} \text{ s.t. } C \text{ is recall-optimal on } \mathcal{M}_u \text{ over } \mathcal{S}, \\ 0 & \text{otherwise.} \end{cases} \tag{46}$$

For any $\mathcal{C}, \mathcal{S} \subseteq 2^S$, we define the subset of sampled parameters $\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,1}$ as

$$\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,1} = \{u \in \mathcal{U}_N \mid \exists C \in \mathcal{C} \text{ s.t. } C \subseteq \mathscr{C}_{S',u}, (\mathcal{M}_u, C) \models \textbf{(M)},$$
$$\forall \rho \in \mathrm{FinPath}, \rho \models \Diamond C^*_{S',u} \wedge \Diamond E \implies \Diamond C\}, \tag{47}$$

where, we treat the same elements in $\mathcal{U}_N$ as duplicates, and $\mathscr{C}_{S',u}$ and $C^*_{S',u}$ are defined by (10) and (16) with $S' = \cup_{C' \in \mathcal{S}} C'$, respectively. Then, we define $\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,0}$ as $\mathcal{U}_N \setminus \mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,1}$. To prove Theorem 2, we consider the following SOP:

$$\begin{aligned} \max_{\xi} \quad & \xi \\ \text{s.t.} \quad & \xi \in \bigcap_{u \in \mathcal{U}'_N} J^u_{\mathcal{C},\mathcal{S}}, \end{aligned} \tag{48}$$

where $J^u_{\mathcal{C},\mathcal{S}} = [0, \xi^u_{\mathcal{C},\mathcal{S}}]$ and $\mathcal{U}'_N \subseteq \mathcal{U}_N$. We denote the solution of (48) by $\xi^*_{\mathcal{C},\mathcal{S}}$. Note that $\xi^*_{\mathcal{C},\mathcal{S}} = \min_{u \in \mathcal{U}'_N} \xi^u_{\mathcal{C},\mathcal{S}}$.

Proof of Theorem 2. For any $\mathcal{C}, \mathcal{S} \subset 2^S$, we consider the two case (1) $\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,0} = \emptyset$ and (2) $\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,0} \neq \emptyset$.

(1): We set $\mathcal{U}'_N = \mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,1}$. By Lemma 4, for any $u \in \mathcal{U}'_N$, there exists $C \in \mathcal{C}$ such that C is recall-optimal on $\mathcal{M}_u$ over $\mathcal{S}$. Thus, the intersection of all $J^u_{\mathcal{C},\mathcal{S}}$ with $u \in \mathcal{U}'_N$ is $[0,1]$ and hence there exists a unique solution $\xi^*_{\mathcal{C},\mathcal{S}} = 1$ to (28). Thus, for any $\varepsilon > 0$, by Theorem 4, we have

$$\mathbb{P}^N\left\{\mathbb{P}(\xi^*_{\mathcal{C},\mathcal{S}} \in J^u_{\mathcal{C},\mathcal{S}}) \geq 1 - \varepsilon\right\} \geq 1 - (1 - \varepsilon)^N, \tag{49}$$

where the internal probability $\mathbb{P}$ measures the frequency of u for $J^u_{\mathcal{C},\mathcal{S}}$. We choose $\beta = 1 - (1 - \varepsilon)^N$ and hence

$$\mathbb{P}^N\left\{\mathbb{P}(\xi^u_{\mathcal{C},\mathcal{S}} \geq \xi^*_{\mathcal{C},\mathcal{S}}) \geq (1 - \beta)^{1/N}\right\} \geq \beta. \tag{50}$$

We have that $\xi^*_{\mathcal{C},\mathcal{S}} = 1$ and thus

$$\mathbb{P}^N \left\{ \mathbb{P}(\xi^u_{\mathcal{C},\mathcal{S}} = 1) \geq (1 - \beta)^{1/N} \right\} \geq \beta. \tag{51}$$

Therefore, we have

$$\mathbb{P}^N \left\{ R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}, \mathcal{S}) \geq (1 - \beta)^{1/N} \right\} \geq \beta. \tag{52}$$

(2): For any $\mathcal{U}'_N \neq \emptyset$ in (48), the intersection of all $J^u_{\mathcal{C},\mathcal{S}}$ with $u \in \mathcal{U}'_N$ is nonempty and there exists a unique solution $\xi^*_{\mathcal{C},\mathcal{S}}$ to (28). Thus, for any $\varepsilon > 0$ and any a-priori number of discarded sampled parameters $q = |\mathcal{U}_N \setminus \mathcal{U}'_N|$, by Theorem 5, we have

$$\mathbb{P}^N \left\{ \mathbb{P}(\xi^*_{\mathcal{C},\mathcal{S}}(q) \in J^u_{\mathcal{C},\mathcal{S}}) \geq 1 - \varepsilon \right\} \geq 1 - \sum_{i=0}^{q} \binom{N}{i} \varepsilon^i (1 - \varepsilon)^{N-i}. \tag{53}$$

We set $t = 1 - \varepsilon$ and equate the right-hand side of (53) to $1 - (1 - \beta)/N$, where $\beta \in (0, 1)$. Then, we have

$$\mathbb{P}^N \left\{ \mathbb{P}(\xi^u_{\mathcal{C},\mathcal{S}} \geq \xi^*_{\mathcal{C},\mathcal{S}}(q)) \geq t^*(q) \right\} \geq 1 - \frac{1 - \beta}{N}, \tag{54}$$

where $t^*(q)$ is the solution of

$$\frac{1 - \beta}{N} = \sum_{i=0}^{q} \binom{N}{i} (1 - t)^i t^{N-i}. \tag{55}$$

We consider the case $q = |\mathcal{U}^C_{N,0}|$ to derive the lower bound of $R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}, \mathcal{S})$ for $\mathcal{U}'_N = \mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,1}$. Since the number $|\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,0}|$ is not known before observing actual samples, we consider simultaneously the bound for all $q = 1, \ldots, N$. By Bool's inequality and (54), we have

$$\mathbb{P}^N \left\{ \cap_{q=1}^N \mathcal{E}_q \right\} = 1 - \mathbb{P}^N \left\{ \cup_{q=1}^N \mathcal{E}^c_q \right\}$$

$$\geq 1 - \sum_{q=1}^N \mathbb{P}^N \left\{ \mathcal{E}^c_q \right\}$$

$$\geq \beta, \tag{56}$$

where $\mathcal{E}_q$ denotes the event $\mathbb{P}(\xi^u_{\mathcal{C},\mathcal{S}} \geq \xi^*_{\mathcal{C},\mathcal{S}}) \geq t^*(q)$ and $\mathcal{E}^c_q$ is its complement event. Clearly, we have

$$\mathbb{P}^N \left\{ \mathbb{P}(\xi^u_{\mathcal{C},\mathcal{S}} \geq \xi^*_{\mathcal{C},\mathcal{S}}) \geq t^*(|\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,0}|) \right\} \geq \mathbb{P}^N \left\{ \forall q' = 1, \ldots, N, \mathbb{P}(\xi^u_{\mathcal{C},\mathcal{S}} \geq \xi^*_{\mathcal{C},\mathcal{S}}) \geq t^*(q') \right\}$$

$$\geq \beta. \tag{57}$$

Note that we can take $\mathcal{U}'_N = \mathcal{U}^C_{N,1}$ when $q = |\mathcal{U}^C_{N,0}|$. As in the case (1), by Lemma 21 in [Baier et al., 2022b] and Lemma 3, $\xi^*_{\mathcal{C},\mathcal{S}} = 1$ holds. Thus, we have

$$\mathbb{P}^N \left\{ \mathbb{P}(\xi^u_{\mathcal{C},\mathcal{S}} \geq \xi^*_{\mathcal{C},\mathcal{S}}) = t^*(|\mathcal{U}^C_{N,0}|) \right\} = \mathbb{P}^N \left\{ \mathbb{P}(\xi^u_{\mathcal{C},\mathcal{S}} = 1) \geq t^*(|\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,0}|) \right\}$$

$$= \mathbb{P}^N \left\{ R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}, \mathcal{S}) \geq t^*(|\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,0}|) \right\}. \tag{58}$$

Therefore, by (57), we have

$$\mathbb{P}^N \left\{ R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}, \mathcal{S}) \geq t^*(|\mathcal{U}^{\mathcal{C},\mathcal{S}}_{N,0}|) \right\} \geq \beta. \tag{59}$$

$\square$

A.3 DERIVATION OF THEOREM 3

In this section, we prove Theorem 3. First, we show the following lemmas. Lemma 5 states that the necessary and sufficient condition for any subset of states to be recall-optimal on any MDP $\mathcal{M}_u$. Then, Lemma 6 states that the solution defined by (20) accomplishes the maximum lower bound for the recall-optimal probability, that is the solution satisfies **(PC1)** in Problem 3, based on Lemma 5. Proposition 1 with Lemmas 6 and 7 assures that $\mathcal{C}^*$ defined by (20) fulfills **(PC2)** in Problem 3.

For simplicity, we abbreviate $C^*_{S_N,u_i}$ defined by (16) and $\mathscr{C}_{S_N,u_i}$ defined by (10) with $S_N = \cup_{C' \in \mathcal{S}_{N,\delta,\beta}} C'$ as C^*_i and $\mathscr{C}_i$, respectively. Moreover, recall that we denote all finite paths on the pMDP as $\mathrm{FinPath}$. For any $C, C', \mathscr{C} \subseteq S$, we consider the following condition:

$$(C \subseteq \mathscr{C}) \wedge (\forall \rho \in \mathrm{FinPath}, \rho \models \Diamond C' \wedge \Diamond E \implies \Diamond C). \tag{60}$$

We denote the first operand in (60) as $\varphi_{\mathrm{cause}}(C, \mathscr{C})$ and the second operand in (60) as $\varphi_{\mathrm{recall}}(C, C')$.

Lemma 5. *For any pMDP $\mathcal{M}_u$, any $S' \subset S$, any $C \subseteq S'$ that satisfies **(M)** in Def. 1, and any $u_j \in \mathcal{V}$, C is recall-optimal on $\mathcal{M}_{u_j}$ over $\mathcal{S} = 2^{S'}$ if and only if $\varphi_{\mathrm{cause}}(C, \mathscr{C}_{S',u_j}) \wedge \varphi_{\mathrm{recall}}(C, C^*_{S',u_j})$ holds, where $\mathscr{C}_{S',u_j}$ is defined by (10) and C^*_{S',u_j} is defined by (16).*

Proof. ($\Leftarrow$) This immediately follows from Lemma 4.

($\Rightarrow$) Suppose that $\neg\varphi_{\mathrm{cause}}(C, \mathscr{C}_{S',u_j}) \vee \neg\varphi_{\mathrm{recall}}(C, C^*_{S',u_j})$ holds. Then, we consider two cases (i) $\varphi_{\mathrm{cause}}(C, \mathscr{C}_{S',u_j})$ does not hold and (ii) $\varphi_{\mathrm{cause}}(C, \mathscr{C}_{S',u_j})$ holds but $\varphi_{\mathrm{recall}}(C, C^*_{S',u_j})$ does not hold.

(i) By Lemma 1, C is not SPR cause on $\mathcal{M}_{u_j}$, and hence C is not recall optimal $\mathcal{M}_{u_j}$ over $\mathcal{S}$.

(ii) By $\neg\varphi_{\mathrm{recall}}(C, C^*_{S',u_j})$, we have

$$\exists \rho \text{ s.t. } \rho \models \Diamond C^*_{S',u_j} \wedge \Diamond E \wedge \neg\Diamond C. \tag{61}$$

By $\varphi_{\mathrm{cause}}(C, \mathscr{C}_{S',u_j})$ and Lemma 1, C is an SPR cause on $\mathcal{M}_u$, and hence, by Lemma 4,

$$\{\rho \in \mathrm{FinPath} \mid \rho \models \Diamond C^*_{S',u_j} \wedge \Diamond E\} \supseteq \{\rho \in \mathrm{FinPath} \mid \rho \models \Diamond C \wedge \Diamond E\}. \tag{62}$$

Thus, by (62) and (61), we have

$$\{\rho \in \mathrm{FinPath} \mid \rho \models \Diamond C^*_{S',u_j} \wedge \Diamond E\} \supset \{\rho \in \mathrm{FinPath} \mid \rho \models \Diamond C \wedge \Diamond E\}. \tag{63}$$

Hence, we have

$$\exists \pi \text{ s.t. } \mathrm{Pr}^\pi_{\mathcal{M}_{u_j}}(\Diamond C^*_{S',u_j} \mid \Diamond E) > \mathrm{Pr}^\pi_{\mathcal{M}_{u_j}}(\Diamond C \mid \Diamond E). \tag{64}$$

Therefore, C is not recall optimal $\mathcal{M}_{u_j}$ over $\mathcal{S}$.

$\square$

Lemma 6. *For any upMDP $\mathcal{M}_\mathbb{P}$, any $\delta, \beta \in [0, 1]$, and any $N > 0$, $\mathcal{C}^*$ defined by (20) satisfies the following two conditions:*

$$\forall u \in \mathcal{U}_N, \exists C \in \mathcal{C}^* \text{ s.t. } C \text{ is recall-optimal on } \mathcal{M}_u \text{ over } \mathcal{S}_{N,\delta,\beta}, \tag{65}$$

and

$$\zeta_N(\mathcal{C}^*, \mathcal{S}_{N,\delta,\beta}, \beta) = (1 - \beta)^{1/N}, \tag{66}$$

where ζ_N is defined as (18).

Proof. By (20), for any $u_j \in \mathcal{U}_N$, there exists $C_i \in \mathcal{C}^*$ such that $C_i \subseteq \cup_{C' \in \mathcal{S}_{N,\delta,\beta}} C'$, C satisfies **(M)**, and $\varphi_{\mathrm{cause}}(C_i, \mathscr{C}_j) \wedge \varphi_{\mathrm{recall}}(C_i, C^*_j)$ holds. Hence, by Lemma 5, C_i is recall-optimal on $\mathcal{M}_{u_j}$ over $\mathcal{S}_{N,\delta,\beta}$. Thus, we have (65) and

$$m(\mathcal{C}^*, \mathcal{S}_{N,\delta,\beta}, \mathcal{U}_N) = |\mathcal{U}_N| = N, \tag{67}$$

where m is he number of samples in $\mathcal{U}_N$ such that there exists $C \in \mathcal{C}^*$ that satisfies condition defined by(17). Therefore, since the maximum value of t^* defined by (14) is $t^*(0, \beta) = (1 - \beta)^{1/N}$ for any β and any N, (66) holds by (18) and (65).

$\square$

By Lemma 6, $\mathcal{C}^*$ defined by (20) achieves the condition **(PC1)** in Problem 3 since the maximum value of ζ_N defined in (18) is $(1 - \beta)^{1/N}$.

Lemma 7. *For any upMDP $\mathcal{M}_{\mathbb{P}}$, any $\mathcal{C}, \mathcal{S} \subseteq 2^S$, and any $\mathcal{C}' \subseteq \mathcal{C}$, we have the following equation:*

$$R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}, \mathcal{S}) - R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}', \mathcal{S}) = \int_{\mathcal{V}} \xi_{\mathrm{diff}}(u, \mathcal{C}, \mathcal{C}', \mathcal{S}) \mathrm{d}\mathbb{P}(u), \tag{68}$$

where $\xi^{\mathrm{diff}} : \mathcal{V} \times 2^{2^S} \times 2^{2^S} \times 2^{2^S}$ is defined as

$$\xi_{\mathrm{diff}}(u, \mathcal{C}, \mathcal{C}', \mathcal{S}) = \begin{cases} 1 & \textit{if } \xi^u_{\mathcal{C}, \mathcal{S}} = 1 \wedge \xi^u_{\mathcal{C}', \mathcal{S}} = 0, \\ 0 & \textit{otherwise,} \end{cases} \tag{69}$$

where $\xi^u_{\mathcal{C}, \mathcal{S}}$ is defined by (46).

Proof.

$$R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}, \mathcal{S}) - R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}', \mathcal{S}) = \int_{\mathcal{V}} I'(u, \mathcal{C}, \mathcal{S}) - I'(u, \mathcal{C}', \mathcal{S}) \mathrm{d}\mathbb{P}(u), \tag{70}$$

and $I'(u, \mathcal{C}, \mathcal{S}) - I'(u, \mathcal{C}', \mathcal{S})$ is nonnegative for all $u \in \mathcal{V}$ since $\mathcal{C}' \subseteq \mathcal{C}$. Thus,

$$= \int_{\mathcal{V}} \xi_{\mathrm{diff}}(u, \mathcal{C}, \mathcal{C}', \mathcal{S}) \mathrm{d}\mathbb{P}(u). \tag{71}$$

$\square$

Based on Lemmas 6, 5, and 7, we show that $\mathcal{C}^*$ defined by (20) achieves the condition **(PC2)** in Problem 3. In the following, we abbreviate $\mathcal{S}_{N,\delta,\beta}$ as $\mathcal{S}$, for simplicity. For any $\mathcal{C}' \subset \mathcal{C}^*$, we define the subset of sampled parameters $\mathcal{U}^{\mathcal{C}^*, \mathcal{C}', \mathcal{S}}_{N,1}$ as

$$\mathcal{U}^{\mathcal{C}^*, \mathcal{C}', \mathcal{S}}_{N,1} = \mathcal{U}^{\mathcal{C}^*, \mathcal{S}}_{N,1} \setminus \mathcal{U}^{\mathcal{C}', \mathcal{S}}_{N,1}, \tag{72}$$

where $\mathcal{U}^{\mathcal{C}, \mathcal{S}}_{N,1}$ is defined by (47). Then, we define $\mathcal{U}^{\mathcal{C}^*, \mathcal{C}', \mathcal{S}}_{N,0}$ as $\mathcal{U}_N \setminus \mathcal{U}^{\mathcal{C}^*, \mathcal{C}', \mathcal{S}}_{N,1}$. We consider the following SOP:

$$\begin{aligned} \max_{\xi} \quad & \xi \\ \text{s.t.} \quad & \xi \in \bigcap_{u \in \mathcal{U}'_N} J^u_{\mathcal{C}^*, \mathcal{C}', \mathcal{S}}, \end{aligned} \tag{73}$$

where $J^u_{\mathcal{C}^*, \mathcal{C}', \mathcal{S}} = [0, \xi_{\mathrm{diff}}(u, \mathcal{C}^*, \mathcal{C}', \mathcal{S})], \mathcal{U}'_N \subset \mathcal{U}_N$, and ξ_{diff} and $\xi^u_{\mathcal{C}, \mathcal{S}}$ are defined by (69) and (46), respectively. We denote the solution of (73) by ξ^*. Note that $\xi^* = \min_{u \in \mathcal{U}'_N} \xi_{\mathrm{diff}}(u, \mathcal{C}^*, \mathcal{C}', \mathcal{S})$. For simplicity, we denote $\xi_{\mathrm{diff}}(u, \mathcal{C}^*, \mathcal{C}', \mathcal{S})$ as ξ^u_{diff}.

Proposition 1. *For any upMDP $\mathcal{M}_{\mathbb{P}}$, any $N > 0$, any $\beta \in (0, 1)$, if $\mathcal{C}^* \neq \emptyset$, then, for any proper subset $\mathcal{C}' \subset \mathcal{C}^*$, we have*

$$\mathbb{P}^N \{R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}^*, \mathcal{S}) > R(\mathcal{M}_{\mathbb{P}}, \mathcal{C}', \mathcal{S})\} > \beta, \tag{74}$$

where $\mathcal{C}^*$ is defined by (20), and $\mathcal{S}$ is defined by (9).

Proof of Proposition 1. For any $\mathcal{U}'_N \neq \emptyset$ in (73), the intersection of all $J^u_{\mathcal{C}^*, \mathcal{C}', \mathcal{S}}$ with $u \in \mathcal{U}'_N$ is nonempty and there exists a unique solution ξ^* to (73). Thus, for any $\varepsilon > 0$ and any a-priori number of discarded sampled parameters $q = |\mathcal{U}_N \setminus \mathcal{U}'_N| < N$, by Theorem 5, we have

$$\mathbb{P}^N \{\mathbb{P}(\xi^u_{\mathrm{diff}} \geq \xi^*) \geq 1 - \varepsilon\} \geq 1 - \sum_{i=0}^{q} \binom{N}{i} \varepsilon^i (1 - \varepsilon)^{N-i}. \tag{75}$$

We set $t = 1 - \varepsilon$ and equate the right-hand side of (75) to $1 - (1 - \beta)/N$, where $\beta \in (0, 1)$. Then, we have

$$\mathbb{P}^N \{\mathbb{P}(\xi^u_{\mathrm{diff}} \geq \xi^*) \geq t^*(q)\} \geq 1 - \frac{1 - \beta}{N}, \tag{76}$$

where $t^*(q)$ is the solution of

$$\frac{1-\beta}{N} = \sum_{i=0}^{q} \binom{N}{i} (1-t)^i t^{N-i}, \tag{77}$$

By Bool's inequality, we have

$$\mathbb{P}^N \left\{ \cap_{q=1}^N \mathcal{E}_q \right\} \geq 1 - \sum_{q=1}^{N} \mathbb{P}^N \left\{ \mathcal{E}_q^c \right\}$$
$$\geq \beta, \tag{78}$$

where $\mathcal{E}_q$ denotes the event $\mathbb{P}(\xi_{\text{diff}}^u \geq \xi^*) \geq t(q)$ and $\mathcal{E}_q^c$ is its complement, and clearly

$$\mathbb{P}^N \left\{ \mathcal{E}_{|\mathcal{U}_{N,0}^{\mathcal{C}^*,\mathcal{C}',\mathcal{S}}|} \right\} \geq \mathbb{P}^N \left\{ \cap_{q=1}^N \mathcal{E}_q \right\}. \tag{79}$$

Note that we can take $\mathcal{U}'_N = \mathcal{U}_{N,1}^{\mathcal{C}^*,\mathcal{C}',\mathcal{S}}$ when $q = |\mathcal{U}_{N,0}^{\mathcal{C}^*,\mathcal{C}',\mathcal{S}}|$, and $\mathcal{U}'_N \neq \emptyset$ by (20). Then, we have $\xi_{\mathcal{C}^*,\mathcal{S}}^u = 1$ and $\xi_{\mathcal{C}',\mathcal{S}}^u = 0$ for all $u \in \mathcal{U}_{N,1}^{\mathcal{C}^*,\mathcal{C}',\mathcal{S}}$ by Lemma 5 and hence $\xi^* = 1$. Thus, by (78) and (79), we have

$$\mathbb{P}^N \left\{ \mathbb{P}(\xi_{\text{diff}}^u = 1) \geq t(|\mathcal{U}_{N,0}^{\mathcal{C}^*,\mathcal{C}',\mathcal{S}}|) \right\} \geq \beta, \tag{80}$$

and $t(|\mathcal{U}_{N,0}^{\mathcal{C}^*,\mathcal{C}',\mathcal{S}}|) > 0$ holds since $\beta < 1$ and $|\mathcal{U}_{N,0}^{\mathcal{C}^*,\mathcal{C}',\mathcal{S}}| < N$, thus,

$$\mathbb{P}^N \left\{ \mathbb{P}(\xi_{\text{diff}}^u = 1) > 0 \right\} \geq \beta, \tag{81}$$

Therefore , by Lemma 7, we have (74). □

Proof of Theorem 3. By Theorems 1 and 2, η_N defined by (12) and ζ_N defined by (18) are the solutions to Problem 2. Moreover, $\mathcal{C}^*$ fulfills **(PC1)** in Problem 3 by (66) in Lemma 6 and fulfills **(PC2)** in Problem 3 by Proposition 1. Thus, we have the claim. □

B COMPUTATIONAL COMPLEXITY OF ALGORITHM 1

The computational complexity of Algorithm 1 is in

$$\mathcal{O}\left(N \cdot \text{poly}(|S|) + N^2 |S| + 2^{|S|} \log \frac{1}{\varepsilon} \right), \tag{82}$$

where N is the number of sampled parameters, $|S|$ denotes the number of states, $\text{poly}(n)$ denotes a polynomial function of n, and ε is the approximation error for computing η_N, that is, ε denotes the length of each sub-interval when dividing the unit interval $[0,1]$ with equal length to conduct bisection search for η_N.

The breakdown of the computational complexity is as follows. The complexity of computing the solution $\mathcal{C}^*$ to Problem 3 is in $\mathcal{O}(N \cdot \text{poly}(|S|) + N^2 |S|)$. The complexity of computing the lower bounds η_N for all members $C \in \mathcal{C}^*$ is in $\mathcal{O}(2^{|S|} \cdot \log \frac{1}{\varepsilon})$ because we compute η_N for all subsets $C \in 2^S$ in the worst case, and the complexity of bisection search to approximately compute η_N is in $\mathcal{O}(\log \frac{1}{\varepsilon})$. However, in general, the computation of η_N for all $C \in \mathcal{C}^*$ is not intensive. This is because the number of obtained potential causes is much smaller than $2^{|S|}$ in practice. For example, the number of potential causes obtained in Example is at most three. The complexity of computing ζ_N is in $\mathcal{O}(1)$ because ζ_N is immediately given as $(1-\beta)^{1/N}$ by Theorem 3.

In what follows, we intuitively describe the details for the complexity to obtain the solution $\mathcal{C}^*$.

- In Lines from 1 to 7:
  The complexity in Lines from 1 to 7 is in $\mathcal{O}(N \cdot \text{poly}(|S|))$. The complexity of Algorithm 2 is in $\mathcal{O}(\text{poly}(|S|))$ by Lemma 4 in [Baier et al., 2022b]. Also, please see [Baier and Katoen, 2008] if you are not familiar with temporal

model checking and its complexity. The complexity of probabilistic verifications for reachability properties in MDPs is in $\mathcal{O}(\text{poly}(|S|))$ and is described by Corollary 10.107 in [Baier and Katoen, 2008], on which the result of Lemma 4 in [Baier et al., 2022b] relies. S_N can be computed by $S_N = \{c \in S \mid \eta_N(\{c\}, \beta) \geq \delta\}$. This is because, for any $C \subset S$ and any $l \in \mathbb{R}$, satisfaction of $\eta_N(\{c\}, \beta) \geq l$ is a necessary condition for $\eta_N(C, \beta) \geq l$ by Definition 1, and hence S_N can be constructed by gathering all states which satisfy the condition $\eta_N(\{c\}, \beta) \geq l$. So, the complexity for computing S_N is in $\mathcal{O}(N \cdot |S|)$. Obviously, the computational complexity of obtaining $\mathscr{C}_{S_N,u}$ using τ is in $\mathcal{O}(|S|)$. The complexity of computing $C^*_{S_N,u}$ is in $\mathcal{O}(2^{|\phi_1|} \cdot |S|)$, where $\phi_1 = \mathscr{C}_{S_N,u} U c$ and $|\phi_1|$ denote the number of operators in ϕ_1 and hence $|\phi_1| = 1$. This is because the complexity of model checking for a transition system (directed graph) and an LTL formula ϕ is in $\mathcal{O}(2^{|\phi|} \cdot |S|)$ by Theorem 3.3 in [Shonoebelen2002]. Thus, the computational complexity for $C^*_{S_N,u}$ is in $\mathcal{O}(|S|)$.

Hence, the complexity for Lines from 1 to 7 is in $\mathcal{O}(N \cdot \text{poly}(|S|))$.

- In Lines from 9 to 11:
  The complexity in Lines from 9 to 11 is in $\mathcal{O}(N^2 \cdot |S|)$. For each $u_i \in \mathcal{U}_N$, the complexity of checking the condition in (22) for any $u_j \in \mathcal{U}_N$ is in $\mathcal{O}(|S|)$ by Theorem 3.3 in [Schnoebelen, 2002]. Hence, the complexity to obtain $\mathcal{U}_N^i$ is in $\mathcal{O}(N \cdot |S|)$. So, the complexity of computing $\mathcal{U}_N^i$ for all $u_i \in \mathcal{U}_N$ is in $\mathcal{O}(N^2 \cdot |S|)$.

  So, the overall computational complexity in Lines from 9 to 11 is in $\mathcal{O}(N^2 \cdot |S|)$.

Thus, the complexity of computing solution $\mathcal{C}^*$ is in $\mathcal{O}\left(N \cdot \text{poly}(|S|) + N^2|S|\right)$.

C INTERPRETABILITY

Obtained SPR causes are interpreted through a probability-raising perspective and recall optimality as follows: They are subsets of states whose traversal increases the probability of undesired behavior in a certain portion of the parameter space of the upMDP, Moreover, the obtained collection of them cover the undesired paths maximally in most of the parameter space of the upMDP.

We can also give an interpretation similar to causal interventional interpretability [Moraffah et al., 2020]. For example, considering the following differences in conditional probabilities:

$$\max_\pi \Pr^\pi_{\mathcal{M}_u}(\Diamond E \mid \Diamond C) - \Pr^\pi_{\mathcal{M}_u}(\Diamond E \mid \neg\Diamond C),$$

and

$$\min_\pi \Pr^\pi_{\mathcal{M}_u}(\Diamond E \mid \Diamond C) - \Pr^\pi_{\mathcal{M}_u}(\Diamond E \mid \neg\Diamond C).$$

Intuitively, the differences in probability, respectively, represent how much passage of the cause C affects the probability values of undesired behavior pessimistically and optimistically. These can be seen as variants of "what if" scenarios usually considered in a causal analysis context [Moraffah et al., 2020]. However, the above comparisons are not for an actual path but for the overall behavior of the system.

As for larger systems modeled by upMDPs, such as the number of states that they have can reach millions, a possible difficulty for interpretability is the case where we have massive potential SPR causes, and hence we may fall into the situation where we struggle to grasp how much each potential cause affect the undesired behavior. In this case, we take various approaches to enhance ease of understanding for humans as follows:

1. We calculate the aforementioned differences in probability for the union of potential causes rather than each potential cause. Then, we can interpret the differences in probability as how much the passage of any potential causes affects the probability of undesired behavior.

2. We pick up the obtained potential causes to be interpreted. For example, we can choose k potential causes from all of them in order of the values of lower bounds η_N for SPR cause probability. Then, we can interpret the differences in probability as how much the passage of each top-k potential cause affects the probability of undesired behavior.

D ANOTHER EXAMPLE TO COMPREHEND HOW (22) WORKS

Example 3. *We consider the upMDP depicted in Fig. 5 with the set of parameter variables $u = [p, q]^\top$, the set of actions $A = \{a, b\}$, and the terminal bad state $E = \{s_3\}$. p and q follow the uniform distributions over $[0.11, 0.51]$ and $[0.3, 0.7]$, respectively.*
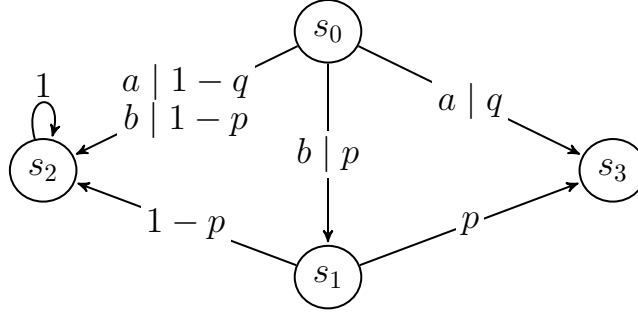
Figure 5: The upMDP with the parameter variables $V = \{p, q\}$, the set of actions $A = \{a, b\}$, and the terminal state $E = \{s_3\}$. The enabled action is only a at states other than s_0 and thus we omit it.

In Fig. 5, we have no SPR cause when $q > p$. This is because there exists a policy π such that $\pi(s_0, a) > (p - p^2\pi(s_0, b))/q$, and hence we have

$$
\begin{aligned}
\mathrm{Pr}^\pi_{\mathcal{M}_u}(\Diamond s_3) &= \pi(s_0, a)q + \pi(s_0, b)p^2 \\
&> p - p^2\pi(s_0, b) + p^2\pi(s_0, b) \\
&= \mathrm{Pr}^\pi_{\mathcal{M}_u}(\Diamond s_3 \mid \Diamond s_1)
\end{aligned}
\tag{83}
$$

However, when $p > q$, $\{s_1\}$ is the unique SPR cause. Let us explain the reason for this in two cases.

(1) When $q \geq p^2$, for any policy π, the probability of reaching s_3 from s_0 is $\mathrm{Pr}^\pi_{\mathcal{M}_u}(\Diamond s_3) = \pi(s_0, a)q + \pi(s_0, b)p^2$ and, since $q > p^2$, we have $\pi(s_0, a)q + \pi(s_0, b)p^2 < q$. So, we have $\mathrm{Pr}^\pi_{\mathcal{M}_u}(\Diamond s_3) < \mathrm{Pr}^\pi_{\mathcal{M}_u}(\Diamond s_3 \mid \Diamond s_1)$ since $q < p$. Thus, $\{s_1\}$ is the SPR cause.

(2) When $q < p^2$, the maximal probability of reaching s_3 from s_1 is p^2 and obviously $p^2 < p$. So, we have $\{s_1\}$ as the SPR cause.

We consider solving Problem 3 for the upMDP in Example 3. Let $\delta = 0.1$ and $\beta = 0.99$. We sampled $N = 1000$ parameters. We obtained a solution $\mathcal{C}^* = \{\{s_1\}\}$ with $\mathcal{S}_{N,\delta,\beta} = \{\{s_1\}\}$, where $\{s_1\}$ is the canonical SPR causes on the sampled 992 MDPs with $p > q$. Obviously, $\mathcal{U}_N^i$ that corresponds to $\{s_1\}$ is equal to $\mathcal{U}_N$. According to Theorem 1, $\{s_1\}$ provides the lower bounds of SPR-cause probabilities 0.98 with the confidence probability $\beta = 0.99$, respectively, and they exceed δ. For **(PC1)**, the lower bound $\zeta_N(\mathcal{C}^*, \mathcal{S}_{N,\delta,\beta}, \beta)$ of recall-optimal probability is the maximum value $(1 - \beta)^{1/N} = 0.995$. For **(PC2)**, the subset $\mathcal{C} = \emptyset \subset \mathcal{C}^*$ clearly does not contain any SPR cause for all parameters. So, $R(\mathcal{M}_\mathbb{P}, \mathcal{C}^*, \mathcal{S}_{N,\delta,\beta}) > R(\mathcal{M}_\mathbb{P}, \mathcal{C}, \mathcal{S}_{N,\delta,\beta})$.

E   INTUITIVE EXPLANATION FOR SPR CAUSES ON EXAMPLE SECTION

The undesired behavior of the two environments depicted in 3(a) and 3(b) in Example is to enter the red cells. That is, the red cells are "undesired" terminal states. Our demonstration has aimed to nonredundantly and exhaustively find the causes of the undesired behavior. In this example, causes are seen as the states such that passing through the states raises the probability of the undesired behavior. Yellow, orange, and green cells correspond to the causes for some parameters of the upMDP. Our proposed method finds these potential causes by sampling upMDP parameters, model checking-based cause synthesis for each sampled MDP, and consolidating them by the proposed set covering technique.

We can observe that the three causes, depending on the parameter values in these environments, have been found well. Intuitively, in both environments, when the probability of entering the red cells from above is greater than that of entering the red cells from below, that is $p_1 < p_2$, the robot passing through the yellow cells is more likely to enter the red cells from above than other behavior that does not pass through the yellow cells. Likewise, when $p_1 > p_2$, the probability of entering the red cells from below rises after passing through the orange/green cells. The difference between the orange cells and green cells as the SPR causes is subtle. When p_0 is sufficiently large, that is p_0 is close to the upper limit 0.9, the chance to unintentionally go up from the cell $(4, 5)$ is sufficiently small. Hence, the chance to visit the red cells from above is sufficiently small. So, in this case, the set of orange cells is an SPR cause. On the contrary, when p_0 is relatively small, that is p_0 is close to the lower limit 0.8, the set of green cells is an SPR cause. Moreover, we can intuitively find that these three potential causes cover all undesired paths, and removing any potential cause makes an uncovered undesired path appear.

Our algorithm has effectively found these intuitive potential causes by combining three ingredients: (i) sampling parameters of the environment and PAC-bound we introduced, (ii) model checking-based SPR cause synthesis for each sampled MDP based on [Baier et al., 2022b], and (iii) consolidation of these synthesized causes by the proposed set covering based on (22). In detail, when we have tried to obtain subsets of states that become SPR causes with positive probability, we have obtained all three SPR causes. On the other hand, when we have tried to obtain subsets of states that have become SPR causes with the probability threshold $\delta = 0.1$, we obtain only yellow cells. This is because the probability of $p_1 > p_2$ is sufficiently small.

F   DETAILED RESULTS FOR EXAMPLE

Fig. 6(a) and Fig. 6(b) show the results obtained from our proposed method for two environments depicted in Figs. 3(a) and 3(b), respectively, with all experimental parameters $\beta = \beta_0, \beta_1$ and $\delta = \delta_0, \delta_1$. We depict the results with blue, green, magenta, and red lines for the parameters (δ_0, β_0), (δ_0, β_1), (δ_1, β_0), and (δ_1, β_1), respectively. The top row in Fig. 6 shows the recall-optimal probabilities R, their lower bounds ζ, and the minimum recall-optimal probabilities over all proper subsets of the obtained solution $\mathcal{C}^*$ under each sample size N as dotted, solid, and dashed lines, respectively. We find that ζ and R_{sub} bound R from below for all sample sizes and parameters. Moreover, ζ takes a higher value when $\beta = \beta_0$ than when $\beta = \beta_1$. The middle row (resp., the bottom row) shows the maximum (resp., minimum) SPR-cause probabilities $F_{\max}$ (resp., $F_{\min}$) and the maximum (resp., minimum) lower bounds $\eta_{\max}$ (resp., $\eta_{\min}$) for all state sets C in the obtained solution $\mathcal{C}^*$ as dotted and solid lines, respectively, for each sample size N. We observe that $\eta_{\max}$ (resp., $\eta_{\min}$) bounds $F_{\max}$ (resp., $\eta_{\min}$) from below and exceeds δ for all sample sizes and all experimental parameters.
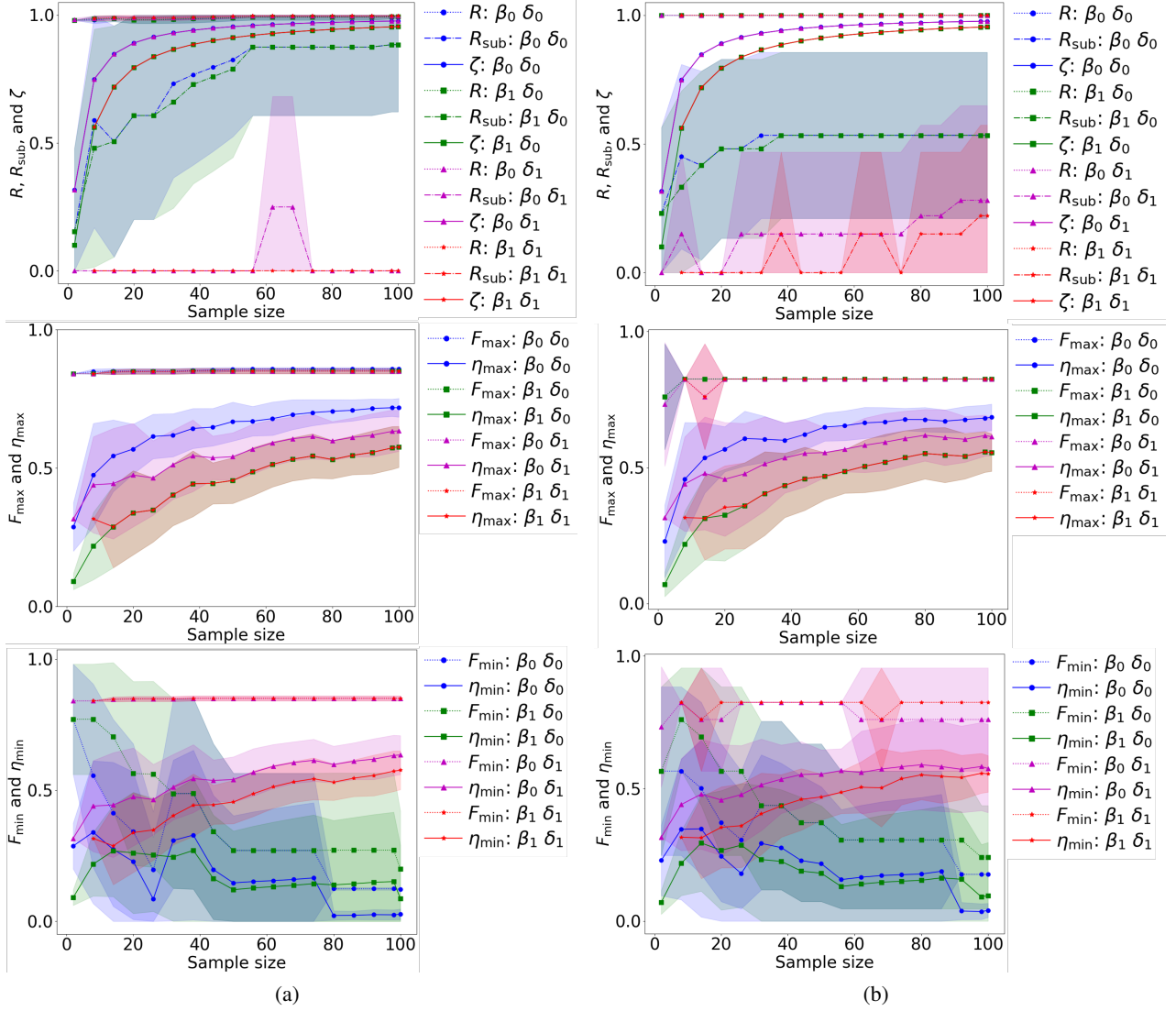
Figure 6: (a) and (b) show the results with all experimental parameters for two environments depicted in Fig. 3(a) and Fig. 3(b), respectively. (Top row): The means and standard deviations of the recall-optimal probabilities R and their lower bounds ζ for the solutions obtained from the proposed method against each sample size (blue and magenta). We further show the maximum recall-optimal probabilities R_{sub} for all proper subsets of the solutions (green). (Middle row) (resp., (Bottom row)): The means and standard deviations of the maximum (resp., minimum) SPR-cause probabilities and their lower bounds over all potential SPR causes in the solutions obtained from the proposed method against each sample size. The results of the upper and the lower figures in (a)-(c) correspond to two environments depicted in Fig 3(a) and 3(b), respectively.