Verification of Robust Multi-Agent Systems

AAAI Track

Raphaël Berthon
RWTH Aachen University & Université Paris-Saclay,
CNRS, ENS Paris-Saclay
Aachen, Germany
rberthon@lmf.cnrs.fr

Joost-Pieter Katoen
RWTH Aachen University
Aachen, Germany
katoen@cs.rwth-aachen.de

Munyque Mittelmann
LIPN, CNRS, Université Sorbonne Paris Nord
Villetaneuse, France
mittelmann@lipn.univ-paris13.fr

Aniello Murano
University of Naples Federico II
Naples, Italy
aniello.murano@unina.it

ABSTRACT

Stochastic multi-agent systems are a central modeling framework for autonomous controllers, communication protocols, and cyber-physical infrastructures. In many such systems, however, transition probabilities are only estimated from data and may therefore be partially unknown or subject to perturbations. In this paper, we study the verification of robust strategies in stochastic multi-agent systems with imperfect information, in which coalitions must satisfy a temporal specification while dealing with uncertain system transitions, partial observation, and adversarial agents. By focusing on bounded-memory strategies, we introduce a robust variant of the model-checking problem for a probabilistic, observation-based extension of Alternating-time Temporal Logic. We characterize the complexity of this problem under different notions of perturbation, thereby clarifying the computational cost of robustness in stochastic multi-agent verification and supporting the use of bounded-memory strategies in uncertain environments.

KEYWORDS

Model Checking; Multi-Agent Systems; Strategic Reasoning

ACM Reference Format:
Raphaël Berthon, Joost-Pieter Katoen, Munyque Mittelmann, and Aniello Murano. 2026. Verification of Robust Multi-Agent Systems: AAAI Track. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 9 pages. https://doi.org/10.65109/DHRR7601

1 INTRODUCTION

AI-based technologies have been widely adopted in a variety of fields, including robotics, autonomous agents, and computer vision. The ubiquity of such technologies highlights the challenges of ensuring the trustworthiness, correctness, and robustness of computational systems. Formal verification provides principled approaches to address these challenges. In particular, model checking [19] offers fully automated techniques in which systems are represented as labeled transition models, and their correct behaviors are specified by temporal logic formulas.

Multi-Agent Systems (MAS) are systems composed of multiple autonomous components that interact in a shared environment [47]. The evolution of a MAS is determined by the behavior of the agents. Alternating-time Temporal Logic (ATL) [2] is a formalism to reason about strategies in MAS. Besides temporal modalities, ATL contains strategic modalities to express how coalitions of agents can cooperate or compete to achieve their objectives. Several aspects of MAS are inherently uncertain, due to both random events and the unpredictable behavior of agents. Such uncertainty can be quantified from experiments or past observations and captured by stochastic models, such as Markov decision processes (MDPs) and stochastic MAS. Probabilistic ATL (PATL) [17] extends ATL to the probabilistic setting, allowing reasoning about randomized strategic abilities of agents operating in a system with stochastic transitions. On the other hand, uncertainty in MAS may also originate from agents' partial observability of the system. In the imperfect-information setting, model checking strategic abilities with perfect recall is undecidable, even for deterministic MAS [23], which motivates a careful choice of restricted yet expressive classes of strategies [8, 13].

Although the frequency of random events can be measured, the precise probabilities are often unknown, and the system may face perturbations. For example, different weather centers often provide varying precipitation probabilities for the same region and time. Another example is model-based reinforcement learning, where agents estimate the agent-environment interaction model (e.g., an MDP [36]) from data. Since the model is learned from their interaction with the environment, its transitions are susceptible to errors. Strategic behavior in such settings is difficult, as it requires dealing with uncertainty in the system transitions while interacting with other agents, who may be cooperative or adversarial.

In this paper, we investigate verification methods for *robust strategies*, i.e., strategies ensuring a temporal specification despite perturbations in transition probabilities. We also consider whether *coalitional* strategies exist that satisfy PATL-like specifications. Instead of fixed transition probabilities after joint actions, we consider

This work is licensed under a Creative Commons Attribution International 4.0 License.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/DHRR7601

stochastic MAS with probability ranges within a given perturbation bound. The resulting setting captures both coalitional strategic uncertainty and uncertainty on exact transition probabilities.

We consider three definitions of robustness. The first, using ε-perturbations, follows the classical setting in which every transition probability in the stochastic system may deviate by at most ε from a nominal value [16]. The other two notions are more involved and based on parametric systems, where transition probabilities are represented as rational functions over a finite set of *parameters*. We study both the case of a fixed number of parameters and the case of an arbitrary (unbounded) number of parameters. This separation is crucial because the arbitrary parameter case typically leads to significant jumps in complexity.

In our setting, agents' strategies depend on their observations of the system, which captures imperfect information. Although recent work has provided decidable model-checking results for PATL under imperfect information and memoryless strategies [6, 7], we also allow for strategies with bounded memory. We consider strategies based on automata, which are widely used in verification [5] and offer more expressivity in relation to standard bounded memory considered in ATL. We illustrate formally this improvement in Proposition 4.2, where we show that our automata-based strategies strictly generalize classical k-recall strategies. At the same time, this additional expressiveness does not increase the worst-case complexity of model checking in our setting.

This setting is interesting from a practical point of view, as full memory is not always feasible. In fact, in real-world applications, such as economics, robotics, or AI, agents are often limited by their computational resources, processing power, and time constraints. Analogous limitations may occur with humans who interact with the system, who can only retain and process a finite amount of information [27, 29]. Observation-based strategies with bounded memory allow handling uncertainty while being computationally feasible and avoiding the extremes of either oversimplification (memoryless strategies) or impracticality (unbounded memory). Having a richer representation for memory is particularly important under partial observation, where bounded memory is one of the few ways to obtain decidability [14]. Therefore, bounded memory strikes a balance between flexibility and efficiency, which is important in practical scenarios.

	Universal reachability (memoryful)	Observation-based PATL (bounded-memory)
ε-perturb.	in **P** (Thm. 7.1*) [16]	in $\Sigma_2^{\mathbf{P}}$ (Thm. 7.6**)
Fix param.	in **NP** ∩ **co-NP** (Thm. 7.2*)	in $\Sigma_2^{\mathbf{P}}$ (Thm. 7.8**)
Unbounded	in $\forall\mathbb{R}$ (Thm. 7.3*)	in $\Sigma_3^{\mathbb{R}}$ (Thm. 7.9*)

Table 1: Summary of the results. Results labeled with * allow randomized strategies. Results labeled with ** consider deterministic strategies for the coalition but allow randomized strategies outside the coalition.

Contribution. This paper provides complexity results for obtaining robust strategies with bounded memory in uncertain stochastic MAS. We show our definition of bounded memory is strictly more general than recalling the b last states visited. We consider the logic PATL for reasoning about observation-based strategies with bounded memory. We then define a variant of the model-checking problem that allows for possible perturbations and determine its complexity for specifications in PATL. Our approach has two main advantages. On one side, bounded memory is a trade-off between computational cost and retaining information. On the other hand, we can guarantee that strategies are resilient to (bounded) perturbations of the models. We provide novel results with reasonable complexity for model checking, going from "simple" reachability objectives to more complex ones describing coalitional behavior. Whereas our reachability results (Theorems 7.1, 7.2, 7.3) are close to the existing literature, they serve as essential building blocks for our PATL results (Theorems 7.6, 7.8, 7.9). These latter results and their proofs are entirely new and technically demanding because they require combining MAS and parametric model checking. Omitted proofs and technical details can be found in [12].

Instead of representing the memory of strategies with the n last states visited, we use automata with n states, increasing the strategic capabilities. While randomized strategies are even more expressive, we obtain better complexity results for deterministic strategies. We consider different kinds of perturbations, namely perturbations represented by parameters that can be common to different actions and thus model dependencies, and two subcases: (i) a fixed number of parameters, and (ii) possible perturbations that can take any value within an interval of ε. Table 1 summarizes our main results. We mainly focus on membership results. Some hardness results may be difficult to obtain even for basic cases. In particular, universal reachability with unbounded parameters is the dual of existential reachability for which $\exists\mathbb{R}$-hardness is a known open problem [31]. Going further to $\Sigma_3^{\mathbb{R}}$, hardness results in the hierarchy of the reals is known to be particularly challenging [41]. We also discuss how different approaches fail at giving a $\Sigma_2^{\mathbf{P}}$-hardness result in [12]. Concerning the case with a fixed number of parameters, no **NP** ∩ **co-NP** -hard problem is known, and the existence of such a hard problem would imply **P** ≠ **NP**.

2 RELATED WORK

Our work is related to the research on probabilistic logics for MAS, interval and parametric Markov Decision Processes (MDPs), and synthesis of robust strategies.

Verification of Stochastic MAS. The verification of stochastic MAS against specifications given in probabilistic logics has been widely studied, including with specifications in probabilistic ATL (PATL) [17], Probabilistic Alternating-Time μ-Calculus [42], and Probabilistic Strategy Logic (PSL) [3]. Verification of concurrent stochastic games has also been implemented using the PRISM model checker [34, 35]. Recent work studied the model-checking problem for PATL under imperfect information and memoryless strategies for the proponent coalition, for both deterministic [6] and mixed strategies [7]. A timed extension of this setting was considered in [28]. Also related is [37], which investigated an extension of PSL for analyzing information transparency in partially observable stochastic MAS. Finally, [11] considered a variant of PATL with probabilistic natural strategies. While natural strategies also capture bounded memory, directly applying the methods proposed in this paper would require representing them as automata, leading to an exponential blow-up in their representation.

Interval MDPs. Interval Markov chains (IMCs) generalize Markov chains with interval-valued transition probabilities: strategies must hold for any system whose transition set is within the interval. Similar to our approach, IMCs are a modeling tool for probabilistic systems with uncertainty of the exact transition probabilities [24]. The model checking of formulas in Linear Temporal Logic (LTL) over IMCs is in **EXPSPACE** and **PSPACE**-hard [9]. In the generalization to Interval MDPs (IMDPs), model checking Probabilistic Computation Tree Logic (PCTL) has been proved polynomial [16] by using the ellipsoid method. However, this approach cannot be easily and directly generalized to observation-based strategies. The problem of finding robust randomized strategies for multiple objectives in IMDPs was shown to be **PSPACE**-hard [24]. Another approach for the verification of PCTL properties in uncertain MDPs considers convex uncertainty sets as a generalization of intervals [39]. Similarly, [15] consider MDPs where each transition has an uncertainty set, with mean-payoff objectives, and gives an **NP**∩**co-NP** membership result.

Parametric MDPs. A way to generalize uncertainty and robustness is by considering parameters. In particular, transition probabilities can be represented as equations over a given set of parameters. While IMCs consider intervals of transition probabilities, the transition probabilities of Parametric MCs are given by polynomials with rational coefficients over a fixed set of real-valued parameters. This generalizes IMCs, as dependencies between transition probabilities can occur. The model checking problem for parametric Markov chains with specifications in PCTL was studied in [4]. In that work, the authors have shown that, in the univariate case, the existential PCTL model checking problem is **NP**-complete, and identified fragments of PCTL where model checking is solvable in polynomial time. On parametric MDPs, instead, even the simplest cases become challenging [31]: with a fixed number of parameters, any reachability objective can be done in **NP**, but with an arbitrary number of parameters, most problems become $\exists\mathbb{R}$-complete. Our separation between fixed and unbounded perturbations takes its root in these results. A related problem is to compute an instantiation of the unspecified parameters in a parametric MDP such that the resulting MDP satisfies a given temporal logic specification. This problem was studied in [21] using convex optimization techniques.

In general, the problem of separating satisfying from violating regions is hard on parametric MDPs, since these are bounded by non-linear functions [30]. Parametric MDPs have been used to synthesize strategies in a two-player partially observable stochastic game framework [1]. Such a framework generalizes partially observable MDPs (POMDP) to two agents, where only one of them has full observability of the system. The synthesis approach involves a reduction to a series of synthesis problems for parametric MDPs combined with mixed-integer linear programming.

Robust Strategies. The synthesis of robust memoryless policies for uncertain POMDPs with reachability and expected cost objectives has been investigated in [43]. Similar to IMDPs, uncertain POMDPs consider a set of probability intervals. The proposed solution uses convex optimization and the authors have shown that the problem is **NP**-hard for memoryless behavioral strategies. This work was extended with finite-state policies, which maintain the same lower bound complexity [22]. The problem of synthesizing robust strategies for MDPs with ellipsoidal uncertainty has been studied alongside PCTL specifications and applied to compute risk-limiting strategies for energy pricing [40]. Experimental evaluations show promising results on s-rectangular robust MDPs [33, 46], but it is still too early to reasonably implement these techniques for multi-agent systems. Recently, [26] considered the verification of interval stochastic MAS under worst-case assumptions about transition uncertainty. They focused on memoryless and non-observation-based strategies and provided verification algorithms for the 2-player setting, together with an implementation using PRISM.

The robustness of MAS has also been investigated from other perspectives. Robust ATL [38] is an extension of ATL to express the ability of a coalition of agents to tolerate the violation of an environment assumption in a purely deterministic setting. The environment assumption is described as a temporal logic property. A violation means the assumption's truth value may be "false" in some moments of the system execution. Recently, this notion was extended to the probabilistic setting [50]. While we consider verification under model perturbations, Yu et al. [48] have studied the case in which the agent's actions may fail with a certain probability.

3 PRELIMINARIES

In this paper, we fix finite non-empty sets of agents Ag, actions Ac, and atomic propositions AP. We write $\boldsymbol{c}$ for a tuple of actions $(c_a)_{a\in\mathrm{Ag}}$, one for each agent, and such tuples are called *action profiles.* Given an action profile $\boldsymbol{c}$ and $C \subseteq \mathrm{Ag}$, we let c_C be the actions of agents in C, and $\boldsymbol{c}_{-C}$ be $(c_b)_{b\notin C}$. Similarly, we let $\mathrm{Ag}_{-C} = \mathrm{Ag} \setminus C$. We denote by $\mathbb{N}_+$ the set $\{1, 2, \ldots\}$.

Distributions. Let X be a finite non-empty set. A *(probability) distribution* over X is a function $\mathrm{d} : X \to [0,1]$ such that $\sum_{x\in X} \mathrm{d}(x) = 1$. Let $\mathrm{Dist}(X)$ be the set of distributions over X. We write $x \in \mathrm{d}$ for $\mathrm{d}(x) > 0$. If $\mathrm{d}(x) = 1$ for some element $x \in X$, then d is a *point (a.k.a. Dirac) distribution.* If, for $i \in I$, d_i is a distribution over X_i, then, writing $X = \prod_{i\in I} X_i$, the *product distribution* of the d_i is the distribution $\mathrm{d} : X \to [0,1]$ defined by $\mathrm{d}(x) = \prod_{i\in I} \mathrm{d}_i(x_i)$.

Stochastic Systems. A *stochastic concurrent multi-agent system* (or simply *system*) $\mathcal{G}$ is a tuple $(\mathrm{St}, \mathrm{L}, \delta, v, (\mathrm{obs}_a)_{a\in\mathrm{Ag}})$ where (i) St is a finite non-empty set of *states*; (ii) $\mathrm{L} : \mathrm{St} \times \mathrm{Ag} \to 2^{\mathrm{Ac}} \setminus \{\emptyset\}$ is a *legality function* defining the available actions for each agent in each state, we write $\mathrm{L}(\boldsymbol{s})$ for the tuple $(\mathrm{L}(s,a))_{a\in\mathrm{Ag}}$ and let $|\mathrm{L}(s)| = \sum_{a\in\mathrm{Ag}} |\mathrm{L}(s,a)|$; (iii) for each state $s \in \mathrm{St}$ and each action $\boldsymbol{c} \in \mathrm{L}(\boldsymbol{s})$, the *stochastic transition function* δ gives the probability $\delta(s,\boldsymbol{c})(s')$ of a transition from state s for all $s' \in \mathrm{St}$ if each player $a \in \mathrm{Ag}$ plays the action $\boldsymbol{c}_a$, and remark that $\delta(s,\boldsymbol{c}) \in \mathrm{Dist}(\mathrm{St})$; and (iv) $v : \mathrm{St} \to 2^{\mathrm{AP}}$ is a *labeling function.* (v) for each $a \in \mathrm{Ag}$, obs_a is an observation function $\mathrm{obs} : \mathrm{St} \to O_a$ for some observation set O_a. Throughout this paper, we assume that systems are uniform, that is, if two states have the same observation for an agent a, then a has the same available actions in both states. Formally, if $O_a(s) = O_a(s')$ then $\mathrm{L}(s,a) = \mathrm{L}(s',a)$, for any $s, s' \in \mathrm{St}$.

For each state $s \in \mathrm{St}$ and joint action $\boldsymbol{c} \in \prod_{a\in\mathrm{Ag}} \mathrm{L}(s,a)$, we assume that there is a state $s' \in \mathrm{St}$ such that $\delta(s,\boldsymbol{c})(s')$ is non-zero, that is every state has a successor state from a legal action, hence $\boldsymbol{c} \in \mathrm{L}(s,a)$.

Example 3.1 (River). Let us consider a system $\mathcal{G}_{river}$ with two companies sharing the usage of a river. At every step, each company has two available actions: discharge wastewater directly into the river (action d) or treat it before discharging it into the river (action t). Atomic propositions state whether the river's water quality has reached low (proposition low) or high levels (proposition $high$). The system is shown in Figure 1. The propositions low and $high$ are true only in state q_{low} and q_{high}, resp.

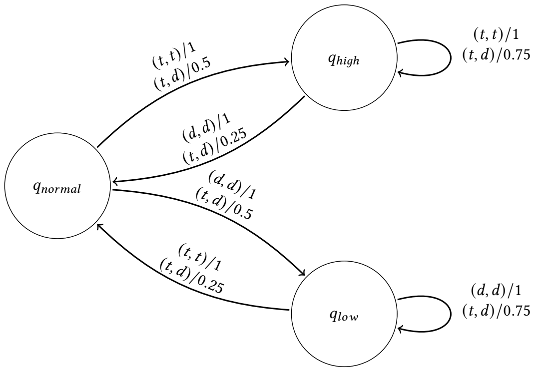

Figure 1: System $\mathcal{G}_{river}$ representing the interaction between two companies. The transition function is written in terms of labels on the arrows. Each transition from state q to q' is annotated with one or more labels of the form $(x, y)/z$ where (x, y) denotes the joint action of the companies in state q, and z denotes the probability of arriving at the state q'. The transition probabilities for the joint action (d, t) are the same as for (t, d), and are thus omitted.

In state q_{normal}, no proposition is true, representing that the water quality is normal. If both companies discharge the wastewater, the water quality is guaranteed to decrease (from high to normal, and normal to low). Similarly, if both companies treat the water, the quality will increase (from low to normal, and normal to high). When only one company treats the water while the other discharges the wastewater, the effect is not deterministic. If the level was normal, the quality may increase or decrease with a probability of 50%. If the quality is high (similarly, low), the quality may be maintained with a probability of 75% or decrease (resp. increase) with a probability of 25%.

Markov Models. A *partially observable Markov decision process* (POMDP) is a one-player stochastic system, given with an observation function for every agent. When all observation functions are injective, the system is fully observable and we have a *Markov decision process* (MDP) where we omit the observations.

A *Markov chain* is an MDP with $|\mathrm{L}(s)| = 1$ for every $s \in \mathrm{St}$. We often represent a Markov chain M as a tuple (St, p) where St is a countable non-empty set of states and $p \in \mathrm{St} \to \mathrm{Dist}(\mathrm{St})$ is the transition matrix.

Plays. A *play* or path in a system $\mathcal{G}$ is an infinite sequence $\pi = s_0 s_1 \cdots \in \mathrm{St}^\omega$ of states such that there exists a sequence $\boldsymbol{c}_0 \boldsymbol{c}_1 \cdots$ of joint-actions such that $\boldsymbol{c}_i \in \mathrm{L}(s_i)$ and $s_{i+1} \in \delta(s_i, \boldsymbol{c}_i)$ (i.e., $\delta(s_i, \boldsymbol{c}_i)(s_{i+1}) > 0$) for every $i \geq 0$. We write π_i for s_i, $\pi_{\geq i}$ for the suffix of π starting at position i. Finite paths are called *histories*, and the set of all histories is denoted Hist. We write $\mathrm{last}(h)$ for the last state of a history h and $len(h)$ for the size of h. We extend observation functions $\mathrm{obs} : \mathrm{St} \to O_a$ to paths with $\mathrm{obs} : \mathrm{St}^\omega \to O_a^\omega$.

4 GENERAL STRATEGIES

In what follows, we introduce a general class of observation-based strategies, both for infinite and bounded memory.

Definition 4.1 (General Bounded-Memory Strategies). Let $\mathcal{G} = (\mathrm{St}, \mathrm{L}, \delta, v, (\mathrm{obs}_a)_{a \in \mathrm{Ag}})$ be a stochastic system with set of actions Ac and set of observations O. In the most general case strategies for agent a are unbounded functions $\sigma : \mathrm{obs}_a(\mathrm{Hist}) \to \mathrm{d}(\mathrm{Ac})$.

We write Str_a for the set of unbounded strategies for agent a. We use Definition 10.97 of [5] and extend it as in [20] with randomization and partial observations. Strategy σ is a *general randomized strategy* with *bounded recall* if it can be represented as a scheduler $(Q, act, \Delta, start)$ where Q is a possibly infinite set of modes (or memory states), $\Delta : Q \times O \to \mathrm{Dist}(Q)$ is a randomized transition function, $act : Q \times O \to \mathrm{Dist}(\mathrm{Ac})$ randomly selects the next action, and $start : O \to \mathrm{Dist}(Q)$ randomly selects a starting mode for the observation o. As in [20], a strategy has infinite memory if $b = \infty$, finite memory $b \in \mathbb{N}_+$ if $|Q| = b$ and is *memoryless* if $|Q| = 1$, it is *behavioral* if act is randomized but all other distributions are singletons, it is *mixed* if $start$ is randomized but all other distributions are singletons, and it is *pure* (or *deterministic*) if all distributions are singletons. We call a strategy *observation-based* if obs isn't injective. Our definition of b-bounded recall strategies uses a finite-state scheduler. This is more general than recalling the last b states, as we show next (see Prop. 4.2).

First, we point out differences in our representation of MAS with imperfect information in relation to the approach used in Belardinelli et al. [6, 7]. While we define stochastic systems with imperfect information as $\mathcal{G} = (\mathrm{St}, \mathrm{L}, \delta, v, (\mathrm{obs}_a)_{a \in \mathrm{Ag}})$, they instead consider $\mathcal{G} = (\mathrm{St}, \mathrm{L}, \delta, v, (\sim_a)_{a \in \mathrm{Ag}})$ where $\sim_a$ is an equivalence relation of $\mathrm{St} \times \mathrm{St}$[1]. These two definitions of imperfect information are equivalent, since for $s, s' \in \mathrm{St}$, we can take $\mathrm{obs}_a(s) = \mathrm{obs}_a(s')$ iff $s \sim_a s'$. However, our definitions make it simpler to define our observation-based strategies using schedulers, in particular, functions $\Delta : Q \times O \to \mathrm{Dist}(Q)$, $act : Q \times O \to \mathrm{Dist}(\mathrm{Ac})$ and $start : O \to \mathrm{Dist}(Q)$.

Our definition of finite-memory strategies differs from the usual one in logics for strategic reasoning, as in [13], which considers *k-sequential strategies*, defined as $\sigma : \mathrm{obs}_a(O^k) \to \mathrm{d}(\mathrm{Ac})$ for $k \in \mathbb{N}$. Our scheduler-based finite-memory strategies are strictly more general than sequential strategies, formally:

Proposition 4.2. *The following holds:*

- *For every k-sequential strategy, there exists a scheduler-based finite-memory strategy playing the same actions given the same history.*
- *There exists an MDP $\mathcal{G}_g$ and a* LTL *formula φ_g such that some scheduler-based finite-memory strategy satisfies φ_g on $\mathcal{G}_g$ with probability* 1*, but no k-sequential strategy satisfies it, regardless of k.*

---

[1] Notice that the uniformity assumptions in [6, 7] are analogous to ours.

Proof Sketch. For the second point, consider the system $\mathcal{G}_g$ in Figure 2 and the formula $\varphi_g = \mathbf{G}((blue \Rightarrow \mathbf{GF}blue) \wedge (red \Rightarrow \mathbf{GF}red))$. After any history on $\mathcal{G}_g$, the scheduler of Figure 3 is in q_b (resp. q_r) iff q_1 (resp. q_2) was visited, and then plays the matching action, ensuring φ_g almost surely. In contrast, any k-sequential strategy must, after k steps in q_3, play the same action in q_4 regardless of whether q_1 or q_2 occurred, so φ_g cannot hold almost surely. The full proof is in [12]. □

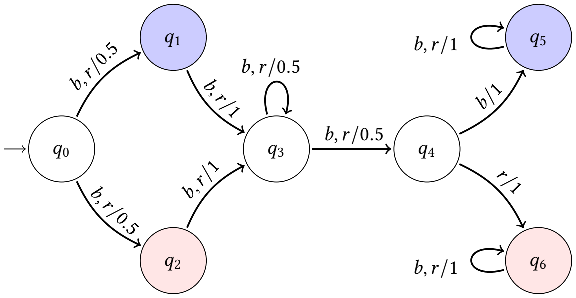

Figure 2: An MDP $\mathcal{G}_g$ where scheduler-based strategies are more general than sequential strategies.

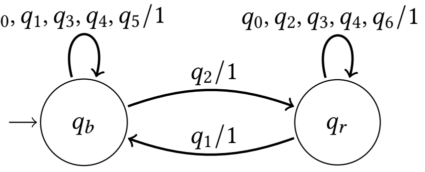

Figure 3: A two-state scheduler enforcing φ_g on $\mathcal{G}_g$.

We highlight that our example uses a deterministic and perfect-information scheduler-based strategy. Cristau et al. [20] details how randomization leads to even more general strategies, hence our choice to call our strategies "general", or to simply refer to them as "strategies". We note that the two definitions coincide on infinite-memory strategies.

Observation-based and bounded recall strategies are not mutually exclusive. We take $Str_{a,n}$ as the set of observation-based strategies with bounded recall n for agent a and $Str_n = \cup_{a \in \mathrm{Ag}} Str_{a,n}$.

Remark 1. The techniques used later in this paper to establish the complexity results can be applied to both bounded recall and memoryless strategies, yielding identical complexity bounds. Since bounded recall captures the case of memoryless strategies and is more general, we will focus on *bounded recall strategies.*

5 PROBABILISTIC ATL

Next, we recall the syntax of Probabilistic Alternating-time Temporal Logic (PATL) [6], defined as follows:

Definition 5.1. The syntax of PATL is given by the grammar

$$\varphi ::= p \mid \varphi \vee \varphi \mid \neg\varphi \mid \langle\langle C \rangle\rangle^{\bowtie d}(\mathbf{X}\varphi) \mid \langle\langle C \rangle\rangle^{\bowtie d}(\varphi \mathbf{U} \varphi) \mid \langle\langle C \rangle\rangle^{\bowtie d}(\varphi \mathbf{R} \varphi)$$

where $p \in \mathrm{AP}$, $C \subseteq \mathrm{Ag}$, d is a rational constant in $[0, 1]$, and $\bowtie \in \{\leq, <, >, \geq\}$.

The intuitive reading of the operators is as follows: "next" $\mathbf{X}$, "release" $\mathbf{R}$ and "until" $\mathbf{U}$ are the standard temporal operators. We make use of the usual syntactic sugar $\mathbf{F}\varphi := \top \mathbf{U} \varphi$ and $\mathbf{G}\varphi := \bot \mathbf{R} \varphi$ for temporal operators. $\langle\langle C \rangle\rangle^{\bowtie d}\varphi$ asserts that there exists an observation-based strategy for the coalition C to collaboratively enforce φ with a probability in relation $\bowtie$ with constant d.

Before presenting the semantics, we show how to define the probability space on outcomes.

Probability Space on Outcomes. We define a strategy profile as $\boldsymbol{\sigma} = (\sigma_a)_{a \in \mathrm{Ag}}$, and its *output* from state s is a play π that starts in state s and is extended by letting each agent follow their strategies in $\boldsymbol{\sigma}$, formally $\pi_0 = s$, and for every $k \geq 0$ there exists $\boldsymbol{c}_k \in (\sigma_a(\pi_{\leq k}))_{a \in \mathrm{Ag}}$ such that $\pi_{k+1} \in \delta(\pi_k, \boldsymbol{c}_k)$.

The (potentially infinite) set of outputs of a profile of strategies $\boldsymbol{\sigma}$ and state s is denoted $Output(\boldsymbol{\sigma}, s)$.

A given system $\mathcal{G}$, strategy profile $\boldsymbol{\sigma}$, and state s induce an infinite-state Markov chain $M_{\boldsymbol{\sigma},s}$ whose states are finite prefixes of plays (i.e., histories) in $Output(\boldsymbol{\sigma}, s)$. Given two histories h and hs' in $Output(\boldsymbol{\sigma}, s)$, the transition probabilities in $M_{\boldsymbol{\sigma},s}$ are defined as $p(h, hs') = \sum_{\prod_{a \in \mathrm{Ag}} \mathrm{Ac}} \boldsymbol{\sigma}(\boldsymbol{c})(h) \times \delta(\mathrm{last}(h), \boldsymbol{c})(s')$. The Markov chain $M_{\boldsymbol{\sigma},s}$ induces a canonical probability space on its set of infinite paths [32], and thus also in $Output(\boldsymbol{\sigma}, s)$. [2]

Given a coalition strategy $\boldsymbol{\sigma}_C \in \prod_{a \in C} Str_{a,n}$ with memory bound n, the set of possible outcomes of $\boldsymbol{\sigma}_C$ from a state $s \in \mathrm{St}$ is the set $out_C(\boldsymbol{\sigma}_C, s) = \{Output((\boldsymbol{\sigma}_C, \boldsymbol{\sigma}_{-C}), s) : \boldsymbol{\sigma}_{-C} \in \prod_{a \in \mathrm{Ag}_{-C}} Str_a\}$ of paths that the players in C enforce when they follow the strategy $\boldsymbol{\sigma}_C$, namely, for each $a \in \mathrm{Ag}$, player a follows strategy σ_a in $\boldsymbol{\sigma}_C$. We use $\mu_s^{\boldsymbol{\sigma}_C}$ to range over the measures induced by $out_C(\boldsymbol{\sigma}_C, s)$.

Definition 5.2. PATL formulas are interpreted in a stochastic system $\mathcal{G}$ and a path π. Let $b \in \mathbb{N}_+ \cup \{\infty\}$. The semantics of PATL with b-bounded memory, denoted by the satisfaction relation $\models_b$, is defined as follows:

$$\begin{array}{ll}
\mathcal{G}, \pi \models_b p & \text{iff } p \in v(\pi_0) \\
\mathcal{G}, \pi \models_b \neg\varphi & \text{iff } \mathcal{G}, \pi \not\models_b \varphi \\
\mathcal{G}, \pi \models_b \varphi_1 \vee \varphi_2 & \text{iff } \mathcal{G}, \pi \models_b \varphi_1 \text{ or } \mathcal{G}, \pi \models_b \varphi_2 \\
\mathcal{G}, \pi \models_b \langle\langle C \rangle\rangle^{\bowtie d}\varphi & \text{iff } \exists \boldsymbol{\sigma}_C \in \prod_{a \in C} Str_{a,b} \\
& \quad \text{s.t. } \forall \mu_{\pi_0}^{\boldsymbol{\sigma}_C} \in out_C(\boldsymbol{\sigma}_C, \pi_0), \\
& \quad \mu_{\pi_0}^{\boldsymbol{\sigma}_C}(\{\pi' : \mathcal{G}, \pi' \models_b \varphi\}) \bowtie d \\
\mathcal{G}, \pi \models_b \mathbf{X}\varphi & \text{iff } \mathcal{G}, \pi_{\geq 1} \models_b \varphi \\
\mathcal{G}, \pi \models_b \varphi_1 \mathbf{U} \varphi_2 & \text{iff } \exists k \geq 0 \text{ s.t. } \mathcal{G}, \pi_{\geq k} \models_b \varphi_2 \text{ and} \\
& \quad \forall j \in [0, k).\ \mathcal{G}, \pi_{\geq j} \models_b \varphi_1 \\
\mathcal{G}, \pi \models_b \varphi_1 \mathbf{R} \varphi_2 & \text{iff } \forall k \geq 0, \mathcal{G}, \pi_{\geq k} \models_b \varphi_2 \text{ or} \\
& \quad \exists j \in [0, k).\ \mathcal{G}, \pi_{\geq j} \models_b \varphi_1
\end{array}$$

Remark 2. PATL defined over systems with exactly one agent corresponds to PCTL over POMDPs with bounded-memory policies, and

---

[2] This is a classic construction, see for instance [10, 18].

the literature we build on focuses almost exclusively on the fragment of PCTL restricted to reachability objectives. Our objective is to handle all PATL with bounded-memory in the MAS setting.

Definition 5.3. Given a system $\mathcal{G}$, state $s \in \mathrm{St}$, a memory bound $b \in \mathbb{N}_+ \cup \{\infty\}$ and a formula φ in PATL, the *model checking problem* for PATL consists of deciding whether $\mathcal{G}, s \models_b \varphi$.

Proposition 5.4 (From [25]). *The model checking problem for* PATL *using infinite-memory strategies is undecidable.*

Proof. With infinite memory, the definition of scheduler-based strategies and classical sequential strategies coincide as unbounded functions $\sigma : \mathrm{obs}_a(\mathrm{Hist}) \to \mathrm{d}(\mathrm{Ac})$. Hence we can directly apply the known result that PATL with infinite memory and under partial observation is undecidable [25]. ◻

For the rest of the paper, we will assume $b \in \mathbb{N}_+$.

6 ROBUSTNESS

Parametric Systems. In a parametric system, transition probabilities are replaced with equations over a finite set of variables X. We can thus consider whether a property holds for all valuations over X that yield a well-defined set of probabilities. We use a definition similar to [31].

A *parametric stochastic system* (or param-*System*) $\mathcal{G}$ is a tuple $(\mathrm{St}, \mathrm{L}, X, \delta, v, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, \kappa)$ where

- St, L, v, and $(\mathrm{obs}_a)_{a\in\mathrm{Ag}}$ are defined as for stochastic systems;
- X is a finite set of real-valued *parameters*;
- for each state $s \in \mathrm{St}$ and each action $\boldsymbol{c} \in \mathrm{L}(\boldsymbol{s})$, the *parametric stochastic transition function* δ gives the (conditional) probability of a transition from state s for all $s' \in \mathrm{St}$ if each player $a \in \mathrm{Ag}$ plays the action $\boldsymbol{c}_a$. This probability is a rational function over X, formally $\delta(s, \boldsymbol{c})(s') \in \mathbb{Q}[X]$; and
- κ is a set of polynomial constraints of the form $A \bowtie B$ where $A, B \in \mathbb{Q}[X]$ and $\bowtie \in \{<, >, \leq, \geq, =\}$. Later on, we use these constraints to capture different kinds of perturbation.

On a param-system $\mathcal{G}$, a valuation $\mathrm{Val} : X \to \mathbb{R}$ is *well-defined* if replacing every $x \in X$ by $\mathrm{Val}(x)$ yields a stochastic system $\mathcal{G}[\mathrm{Val}]$. By extending the co-domain of Val to rational functions over X, we can formally define well-defined valuations as requiring (i) probabilities to be non-negative: for each $s, s' \in \mathrm{St}$ and $\boldsymbol{c} \in \mathrm{L}(\boldsymbol{s})$, we have $\mathrm{Val}(\delta(s, \boldsymbol{c}, s')) \geq 0$, and (ii) transitions to induce distributions: for each $s, \in \mathrm{St}$ and $\boldsymbol{c} \in \mathrm{L}(\boldsymbol{s})$ we have $\Sigma_{s' \in \mathrm{St}} \mathrm{Val}(\delta(s, \boldsymbol{c}, s')) = 1$.

As above, *parametric POMDPs* or param-POMDPs are defined as param-Systems with only one agent, *parametric MDPs* or param-MDPs as param-POMDPs with perfect information, and *parametric MCs* or param-MCs as param-MDPs with $|\mathrm{L}(s)| = 1$ for every $s \in \mathrm{St}$.

Example 6.1 (River (cont.)). Suppose now that the river's water quality is less stable than previously assumed. We remark on some dependency in this instability and represent it with parameters. Some transition probabilities now involve a perturbation x, and all these probabilities change together. We take the following family of perturbations parameterized by x in the system $\mathcal{G}_{river}$ from Example 3.1, that we call $\mathcal{G}^p_{river}$ and illustrated in Figure 4.

When in state q_{normal}, the probability that the water quality increases when only one company treats the water may be slightly higher than planned, and so the probability of going to q_{high} is $0.5 + x$. As a consequence, the probability to go to q_{low} under this situation is $0.5 - x$, to compensate. At the same time, the probability of staying in q_{high} with only one company treating water may be slightly lower than expected. but follows the same trend, hence it is $0.75 - x$. Still, we can check that, as long as x stays within 0 and 0.25, a company alone can always make sure to have at least probability 0.375 of having a good water quality within two time steps by treating the water.

Following prior work [15, 22, 24, 43], the notion of robust strategies we consider requires dealing with uncertainty about the precise probabilities. Such uncertainty is modeled in terms of *perturbations* that the system may face. Our first goal is to verify whether a strategy achieves a *goal*, even though the system may face perturbations of $\varepsilon \in [0, 1]$. We start by introducing ε-approximated systems.

Given a stochastic system $\mathcal{G} = (\mathrm{St}, \mathrm{L}, \delta, \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}})$, and $\varepsilon \in [0, 1]$, we define the *ε-approximated* parametric system $\mathcal{G}_\varepsilon = (\mathrm{St}, \mathrm{L}, X, \delta', \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, \kappa)$ where for every $s, s' \in \mathrm{St}$ and $\boldsymbol{c} \in \mathrm{L}(\boldsymbol{s})$, every transition probability $\delta'(s, \boldsymbol{c}, s') \in X$ is a new variable. We add to κ the constraints that $\delta'(s, \boldsymbol{c}, s') - \delta(s, \boldsymbol{c}, s') \leq \varepsilon$ and $\delta(s, \boldsymbol{c}, s') - \delta'(s, \boldsymbol{c}, s') \leq \varepsilon$, or equivalently $|\delta(s, \boldsymbol{c}, s') - \delta'(s, \boldsymbol{c}, s')| \leq \varepsilon$. Hence, compared with $\mathcal{G}$, every probability in $\mathcal{G}_\varepsilon$ may have an error of up to ε. IMDPs are a subcase of ε-approximated parametric systems, more precisely IMDPs correspond to ε-approximated MDPs [16].

Remark 3. In general, parameters can introduce discontinuities in the structure of the graph, since edges can disappear when they probability become 0. While it could change the complexity of the problem, this is not the case for the two results we use as building blocks in [16, 31]. Hence we have no need to check whether the graph structure is modified by the parameter valuations.

Example 6.2 (River (cont.)). We now consider an interval perturbation $\epsilon = 0.05$ in the system $\mathcal{G}_{river}$ from Example 3.1. For each possible states q, q' and joint action (x, y), the 0.05-approximated parametric system $\mathcal{G}_{0.05,river}$ allows for variations in the transitions probability that differ at most 0.05 from the transitions in $\mathcal{G}_{river}$. Hence probability 0.5 may be replaced by any value in $[0.45, 0.55]$. One possible realisability of $\mathcal{G}_{0.05,river}$ is illustrated in Figure 5. The obtained system, which we shall denote $\mathcal{G}'_{river}$, has some notable differences in relation to $\mathcal{G}_{river}$. First, the transition from q_{high} when

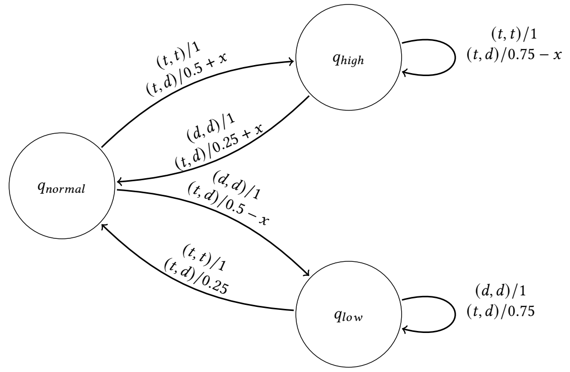

Figure 4: The parametric system $\mathcal{G}^p_{river}$ shows the case where we are unsure of some transitions, that may change together depending on $x \in [0, 0.25]$.

both companies discharge wastewater is no longer deterministic. With high probability, the transition still leads to state q_{normal} (that is, the water quality is decreased to normal levels), but there is a small chance that the water quality decreases to low levels, moving to state q_{low} instead. Second, when only one company treats the water at state q_{normal}, it is more likely that the water quality will decrease (moving to state q_{low}) than to increase (moving to state q_{high}). Clearly, it is more probable to reach state q_{low} from the other two states in $\mathcal{G}'_{river}$ than in the original system $\mathcal{G}_{river}$. Notice this is not the case for every realisability of $\mathcal{G}_{0.05,river}$, as it admits other modifications in the transition probabilities.

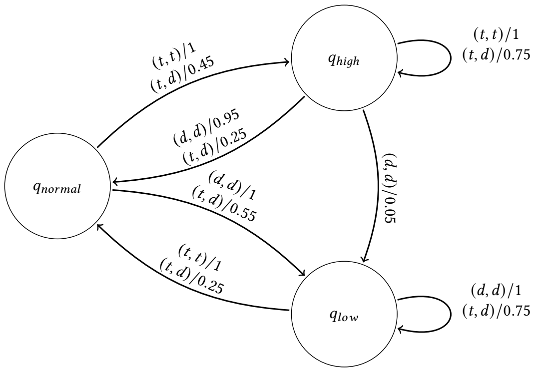

Figure 5: The system $\mathcal{G}'_{river}$, obtained from a realisability of the 0.05-approximated system $\mathcal{G}_{0.05,river}$.

We now introduce the robust model checking problem for PATL.

Definition 6.3. Given a parametric system $\mathcal{G}$, state $s \in \mathrm{St}$, and a formula φ in PATL, a memory-bound b, *the parametric model checking problem* for PATL consists of deciding whether $\mathcal{G}[Val], s \models_b \varphi$ for each well-defined valuation Val.

In the case where $\mathcal{G}$ is an ε-approximated system, we call it the ε-robust model checking problem for PATL.

We also consider the case where we assume the number of perturbed transitions is fixed. Indeed, in many cases, we can consider that only a few critical components have an uncertain behavior. As we will see, this assumption greatly helps to reduce the model checking complexity.

7 ROBUST MODEL CHECKING OF PATL

We start by the case in which the specification φ defines a reachability objective. More precisely, we are interested in the case where a reachability specification φ holds for all (arbitrary) strategies of the agents, formally: $\varphi = \langle\langle\emptyset\rangle\rangle^{\bowtie d}\mathbf{F}R$, where d is a rational constant in $[0,1]$, $\bowtie \in \{\leq, <, >, \geq\}$, and R is a Boolean combination of atomic propositions. This is the same as considering whether a formula holds for all possible strategies in a parametric MDP. To obtain results on existential reachability, duality can be used, since there exists a strategy reaching a target R with probability $> d$ iff we can prove that it is not the case that all strategies reach R with probability $\leq d$. For more results on parametric reachability, see [31].

We start with ε-approximated MDPs. Since they are a specific case of interval MDPs, we can use the following result:

Theorem 7.1. *[From [16]] The following problem is polynomial: Given an ε-approximated MDP $\mathcal{G} = (\mathrm{St}, \mathrm{L}, X, \delta, \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, \kappa)$, and a reachability objective R, determine whether for all b-bounded strategies and all well-defined parameters X, R can be reached with probability $\bowtie d$ with $\bowtie \in \{<, >, \geq, \leq\}$.*

For the more general case of parametric systems, we first consider the case in which we have a bound on the number of parameters, extending results on parametric MDPs from [31].

Theorem 7.2. *The following problem is in* **NP** $\cap$ **co-NP***: Given a* param*-MDP $\mathcal{G} = (\mathrm{St}, \mathrm{L}, X, \delta, \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, C)$ where the number of parameters is less than a fixed m, and a reachability objective R, determine whether for all b-bounded strategies and all well-defined parameters X, R can be reached with probability $\bowtie d$.*

Proof. We first show that this problem is in **co-NP**. The dual problem, that is deciding if there exists a parameter and a strategy that does satisfy a reachability probability with probability $\bowtie d$, is in **NP** by Proposition 15 of [31], hence our problem is in **co-NP**.

To show that our problem is in **NP**, we first remark that we are trying to find a strategy satisfying a reachability property in an MDP: we restrict ourselves to memoryless strategies (by Lemma 10.102 of [5]). We use the same technique as Theorem 12 of [31]: We guess a strategy, representing the worst possible case, either minimizing (for $\geq d$ and $> d$) or maximizing (for $\leq d$ and $< d$) the probability to reach R, and check that this worst-case strategy is still optimal. Using the Bellman equations, we translate this optimality constraint into an $\exists\mathbb{R}$ formula φ with a fixed number of parameters. Since we are trying to decide whether for all well-defined valuations on X, φ holds, we have a $\forall\exists\mathbb{R}$ formula with a fixed amount of parameters, which can be checked in **P**. ☐

Finally, we consider the case where we have an arbitrary number of parameters. From now on, we start making use of complexity classes involving the *theory of the reals*. While the polynomial hierarchy can be understood as finding values for Boolean-valued formulas, the theory of the reals considers real-valued formulas, yielding its own complexity classes. In particular, $\forall\mathbb{R}$ consists of deciding if a formula is true for all possible real valuations of its variables. For an integer k, the class $\Pi_k^{\mathbb{R}}$ designates the problem of deciding formulas starting with $\forall$ and with $k-1$ quantifier alternations. We give more details on this theory in [12].

Theorem 7.3. *The following problem is in $\forall\mathbb{R}$: Given a* param*-MDP $\mathcal{G} = (\mathrm{St}, \mathrm{L}, X, \delta, \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, \kappa)$, and a reachability objective R, determine whether for all b-bounded strategies and all well-defined parameters X, R can be reached with probability $\bowtie d$.*

We can now take a system $\mathcal{G}$, and consider the model checking of PATL under different definitions of robustness. We start with an example, before considering the case of ε-robustness, using a subcase we later build upon.

Example 7.4 (River (cont.)). Going back to our running example, let b denote a memory bound. The following formula expresses that a coalition C has a strategy to guarantee, with probability at least 0.5, that the river quality will be high in the next state: $\langle\langle C\rangle\rangle^{\geq 0.5}\mathbf{X}high$.

In the system $\mathcal{G}_{river}$ and starting from state q_{normal}, this formula is true with memory at most b for any non-empty coalition, that is, $\mathcal{G}_{river}, q_{normal} \models_b \bigwedge_{C\neq\emptyset}\langle\langle C\rangle\rangle^{\geq 0.5}\mathbf{X}high$. However, this is not the case in the modified system $\mathcal{G}'_{river}$ (Figure 5), since a company alone cannot guarantee to reach q_{high} in the next state with enough probability. Thus,

$$\mathcal{G}'_{river}, q_{normal} \not\models_b \bigwedge_{C\neq\emptyset}\langle\langle C\rangle\rangle^{\geq 0.5}\mathbf{X}high$$

In all realisability of $\mathcal{G}_{0.05,river}$, both companies should cooperate to ensure the river will be clean in the next state with a probability at least 0.95. Thus,

$$\mathcal{G}_{0.05,river}[\mathrm{Val}], q_{normal} \models_b \langle\langle \mathrm{Ag}\rangle\rangle^{\geq 0.95}\mathbf{X}high$$

for any valuation Val.

In the following, a formula φ has a chaining of coalition operators if it contains a sub-formula $\langle\langle C\rangle\rangle^{\bowtie d}\psi$, where ψ itself contains a coalition operator. We use such formulas as a base case on which we build our decision procedure for the complete logic.

Lemma 7.5. *The following problem is in* **NP**: *Given an ε-approximated system $\mathcal{G} = (\mathrm{St}, \mathrm{L}, X, \delta, \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, \kappa)$, model check ε-robustly a* PATL *formula φ without chaining of coalition operators, and using deterministic b-bounded strategies for the coalitions.*

We then conclude on the robust model checking problem:

Theorem 7.6. *The following problem is in* $\mathbf{NP}^{\mathbf{NP}} = \Sigma_2^{\mathbf{P}}$: *Given an ε-approximated system $\mathcal{G} = (\mathrm{St}, \mathrm{L}, X, \delta, \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, \kappa)$, model check ε-robustly a* PATL *formula φ using deterministic b-bounded strategies for the coalitions.*

A result similar to Lemma 7.5 holds for parametric systems with a fixed number of parameters, but is in **NP** ∩ **co-NP** :

Lemma 7.7. *The following problem is in* **NP**: *Given a* param-*system $\mathcal{G} = (\mathrm{St}, \mathrm{L}, X, \delta, \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, C)$, where the number of parameters is less than a fixed m, model check parametrically a* PATL *formula φ with no chaining of coalition operators, and using deterministic b-bounded strategies for the coalitions.*

We can then conclude as in Theorem 7.6.

Theorem 7.8. *The following problem is in* $\mathbf{NP}^{\mathbf{NP}} = \Sigma_2^{\mathbf{P}}$: *Given a* param-*system $\mathcal{G} = (\mathrm{St}, \mathrm{L}, X, \delta, \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, C)$, where the number of parameters is fixed and less than m, model check parametrically a* PATL *formula using deterministic b-bounded strategies for the coalitions.*

We now consider the case of arbitrary parameters:

Theorem 7.9. *The following problem is in* $\exists\forall\exists\mathbb{R} = \Sigma_3^{\mathbb{R}}$: *Given system $\mathcal{G} = (\mathrm{St}, \mathrm{L}, X, \delta, \ell, (\mathrm{obs}_a)_{a\in\mathrm{Ag}}, C)$ and a memory bound $b \in \mathbb{N}_+$, model check parametrically a* PATL *formula φ using randomized b-bounded strategies.*

8 DISCUSSION ON APPLICATION

Cyber-physical systems provide a natural domain where robust strategic reasoning under partial observability is indispensable [45]. Examples include smart cities, transportation infrastructures, and industrial automation, where software controllers interact with a physical environment whose dynamics are only approximately known, and where failures or attacks may perturb the expected behaviour of the system. In such settings, systems need to be robust, despite noise and disturbances, to function efficiently and avoid the exploitation of potential risks and threats from attackers [49].

A representative scenario is that of green buildings in a smart city. Buildings communicate and coordinate to share resources such as energy and water efficiently [44]. Each building has private information (e.g., resource levels or user preferences) and local objectives (e.g., minimising costs or contributing resources to the community), while the overall system should maintain global quality-of-service and safety constraints [45]. This cooperative behaviour can be captured by PATL specifications. For instance, the formula $\langle\langle C\rangle\rangle^{\geq 0.8}\mathbf{G}\Big(\bigwedge_{a\in C}\neg energyLow_a\Big)$ expresses that a coalition C of buildings has an observation-based, bounded-memory strategy to ensure, with probability at least 0.8, that none of its members reaches an unacceptably low energy level. Our robust semantics then ask whether this guarantee holds for *all* instantiations of the transition probabilities within the prescribed uncertainty set. This captures, for example, deviations in the efficiency of solar panels, inaccuracies in predicted heating demand, or model updates obtained from new data.

The same framework also supports security-oriented specifications for cyber-physical infrastructures [49]. Modeling the interaction between attackers and defenders as a stochastic MAS, one can express properties such as $\neg\langle\langle\{a\}\rangle\rangle^{\geq 0.2}\mathbf{F}(access_a \wedge \neg securityCheck_a)$, which states that an attacker a does *not* have a strategy to gain unauthorized access with probability at least 0.2. Robust model checking then verifies that this remains true even if the probabilities of successful attacks or detection change.

9 CONCLUSION

This paper addresses the problem of verifying the robustness of strategies for agents acting in stochastic MAS whose transition probabilities are uncertain. Focusing on observation-based, bounded-memory strategies represented by finite automata, we have introduced a robust variant of the model-checking problem for PATL and established upper bounds on its complexity under three different notions of perturbation. On the quantitative side, our results clarify the computational cost of verifiable robustness in MAS. In particular, we show that robust strategic reasoning remains decidable even in the presence of coalitions, imperfect information, and parametric uncertainty, while highlighting the sharp complexity jump induced by unboundedly many parameters. Conceptually, our work connects several previously separate lines of research: verification of stochastic MAS in strategic logics, interval and parametric models for uncertain probabilities, and robust strategy synthesis.

We left open the case of lower bounds; while some are unrealistically high, we think it would be possible to get lower bounds on the robust model checking of PATL. As future work, we aim to consider the problem of computing the robustness level of a given strategy, that is, the maximal perturbation under which the strategy is still successful. This would open the question of finding the most robust strategy in a stochastic MAS.

ACKNOWLEDGMENTS

This research has been supported by the DFG Project POMPOM (KA 1462/6-1) and ANR-22-CE48-0012.

REFERENCES

[1] Mohamadreza Ahmadi, Murat Cubuktepe, Nils Jansen, Sebastian Junges, Joost-Pieter Katoen, and Ufuk Topcu. 2018. The Partially Observable Games We Play for Cyber Deception. *CoRR* abs/1810.00092 (2018).

[2] Rajeev Alur, Thomas A. Henzinger, and Orna Kupferman. 2002. Alternating-time temporal logic. *Journal of the ACM* 49, 5 (2002), 672–713. https://doi.org/10.1145/585265.585270

[3] Benjamin Aminof, Marta Kwiatkowska, Bastien Maubert, Aniello Murano, and Sasha Rubin. 2019. Probabilistic Strategy Logic. In *Proc. of IJCAI 2019*. 32–38.

[4] Christel Baier, Christian Hensel, Lisa Hutschenreiter, Sebastian Junges, Joost-Pieter Katoen, and Joachim Klein. 2020. Parametric Markov chains: PCTL complexity and fraction-free Gaussian elimination. *Information and Computation* 272 (2020), 104504.

[5] Christel Baier and Joost-Pieter Katoen. 2008. *Principles of Model Checking*. MIT Press.

[6] Francesco Belardinelli, Wojciech Jamroga, Munyque Mittelmann, and Aniello Murano. 2023. Strategic Abilities of Forgetful Agents in Stochastic Environments. In *Proc. of KR 2023*. 726–731.

[7] Francesco Belardinelli, Wojtek Jamroga, Munyque Mittelmann, and Aniello Murano. 2024. Verification of Stochastic Multi-Agent Systems with Forgetful Strategies. In *Proc. of AAMAS 2024*. 160–169.

[8] Francesco Belardinelli, Alessio Lomuscio, Aniello Murano, and Sasha Rubin. 2020. Verification of multi-agent systems with public actions against strategy logic. *Artif. Intell.* 285 (2020).

[9] Michael Benedikt, Rastislav Lenhardt, and James Worrell. 2013. LTL Model Checking of Interval Markov Chains. In *Proc. of TACAS 2013*, Vol. 7795. Springer, 32–46. https://doi.org/10.1007/978-3-642-36742-7_3

[10] Raphaël Berthon, Nathanaël Fijalkow, Emmanuel Filiot, Shibashis Guha, Bastien Maubert, Aniello Murano, Laureline Pinault, Sophie Pinchinat, Sasha Rubin, and Olivier Serre. 2020. Alternating Tree Automata with Qualitative Semantics. *ACM Trans. Comput. Logic* 22, 1 (2020), 1–24.

[11] Raphaël Berthon, Joost-Pieter Katoen, Munyque Mittelmann, and Aniello Murano. 2024. Natural Strategic Ability in Stochastic Multi-Agent Systems. In *Proc. of AAAI 2024*. 17308–17316.

[12] Raphaël Berthon, Joost-Pieter Katoen, Munyque Mittelmann, and Aniello Murano. 2026. Verification of Robust Multi-Agent Systems. arXiv:2602.13405 [cs.LO] https://arxiv.org/abs/2602.13405

[13] Raphaël Berthon, Bastien Maubert, Aniello Murano, Sasha Rubin, and Moshe Y Vardi. 2021. Strategy Logic with Imperfect Information. *ACM Trans. Comput. Logic* 22, 1 (2021).

[14] Petr Cermák, Alessio Lomuscio, Fabio Mogavero, and Aniello Murano. 2018. Practical verification of multi-agent systems against SLK specifications. *Inf. Comput.* 261 (2018), 588–614.

[15] Krishnendu Chatterjee, Ehsan Kafshdar Goharshady, Mehrdad Karrabi, Petr Novotný, and Dorde Zikelic. 2024. Solving Long-run Average Reward Robust MDPs via Stochastic Games. In *Proc. of IJCAI 2024*. 6707–6715. https://doi.org/10.24963/ijcai.2024/741

[16] Taolue Chen, Tingting Han, and Marta Z. Kwiatkowska. 2013. On the complexity of model checking interval-valued discrete time Markov chains. *Inf. Process. Lett.* 113, 7 (2013), 210–216. https://doi.org/10.1016/J.IPL.2013.01.004

[17] Taolue Chen and Jian Lu. 2007. Probabilistic alternating-time temporal logic and model checking algorithm. In *Proc. of FSKD 2007*. 35–39.

[18] Edmund M. Clarke, Orna Grumberg, Daniel Kroening, Doron Peled, and Helmut Veith. 2018. *Model Checking*. MIT press.

[19] Edmund M. Clarke, Thomas A. Henzinger, Helmut Veith, and Roderick Bloem. 2018. *Handbook of Model Checking* (1st ed.). Springer Publishing Company, Incorporated.

[20] Julien Cristau, Claire David, and Florian Horn. 2010. How do we remember the past in randomised strategies?. In *Proc. of GANDALF 2010 (EPTCS, Vol. 25)*. 30–39.

[21] Murat Cubuktepe, Nils Jansen, Sebastian Junges, Joost-Pieter Katoen, and Ufuk Topcu. 2022. Convex Optimization for Parameter Synthesis in MDPs. *IEEE Trans. Autom. Control.* 67, 12 (2022), 6333–6348.

[22] Murat Cubuktepe, Nils Jansen, Sebastian Junges, Ahmadreza Marandi, Marnix Suilen, and Ufuk Topcu. 2021. Robust finite-state controllers for uncertain POMDPs. In *Proc. of the AAAI Conf. on Artificial Intelligence*.

[23] Catalin Dima and Ferucio Laurentiu Tiplea. 2011. Model-checking ATL under Imperfect Information and Perfect Recall Semantics is Undecidable. *CoRR* abs/1102.4225 (2011).

[24] Ernst Moritz Hahn, Vahid Hashemi, Holger Hermanns, Morteza Lahijanian, and Andrea Turrini. 2019. Interval Markov decision processes with multiple objectives: from robust strategies to Pareto curves. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 29, 4 (2019), 1–31.

[25] Jianye Hao, Songzheng Song, Yang Liu, Jun Sun, Lin Gui, Jin Song Dong, and Ho-fung Leung. 2012. Probabilistic Model Checking Multi-agent Behaviors in Dispersion Games Using Counter Abstraction. In *PRIMA 2012 (Lecture Notes in Computer Science, Vol. 7455)*. Springer, 16–30.

[26] Angel Y He and David Parker. 2026. Robust verification of concurrent stochastic games. *arXiv preprint arXiv:2601.12003* (2026).

[27] Wojciech Jamroga, Damian Kurpiewski, and Vadim Malvone. 2020. Natural strategic abilities in voting protocols. In *Int. Workshop on Socio-Technical Aspects in Security and Trust*. 45–62.

[28] Wojciech Jamroga, Marta Kwiatkowska, Wojciech Penczek, Laure Petrucci, and Teofil Sidoruk. 2025. Probabilistic Timed ATL. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems*. 1051–1059.

[29] Wojciech Jamroga, Vadim Malvone, and Aniello Murano. 2019. Natural strategic ability. *Artificial Intelligence* 277 (2019), 103170. https://doi.org/10.1016/j.artint.2019.103170

[30] Sebastian Junges, Erika Ábrahám, Christian Hensel, Nils Jansen, Joost-Pieter Katoen, Tim Quatmann, and Matthias Volk. 2024. Parameter synthesis for Markov models: covering the parameter space. *Formal Methods Syst. Des.* 62, 1 (2024), 181–259. https://doi.org/10.1007/S10703-023-00442-X

[31] Sebastian Junges, Joost-Pieter Katoen, Guillermo A. Pérez, and Tobias Winkler. 2021. The complexity of reachability in parametric Markov decision processes. *J. Comput. Syst. Sci.* 119 (2021), 183–210. https://doi.org/10.1016/J.JCSS.2021.02.006

[32] John G. Kemeny, J. Laurie Snell, and Anthony W. Knapp. 1976. Stochastic Processes. In *Denumerable Markov Chains*. Springer, 40–57.

[33] Navdeep Kumar, Kaixin Wang, Kfir Yehuda Levy, and Shie Mannor. 2024. Efficient Value Iteration for s-rectangular Robust Markov Decision Processes. In *ICML 2024, Vienna, Austria*. https://openreview.net/forum?id=J4LTDgwAZq

[34] Marta Kwiatkowska, Gethin Norman, and David Parker. 2022. Probabilistic model checking and autonomy. *Annual Review of Control, Robotics, and Autonomous Systems* 5 (2022), 385–410.

[35] Marta Kwiatkowska, Gethin Norman, David Parker, Gabriel Santos, and Rui Yan. 2022. Probabilistic Model Checking for Strategic Equilibria-based Decision Making: Advances and Challenges. In *Proc. of MFCS 2022*. 4–22.

[36] Thomas M. Moerland, Joost Broekens, Aske Plaat, and Catholijn M. Jonker. 2023. Model-based reinforcement learning: A survey. *Foundations and Trends® in Machine Learning* 16, 1 (2023), 1–118.

[37] Chunyan Mu, Nima Motamed, Natasha Alechina, and Brian Logan. 2025. Probabilistic Strategy Logic with Degrees of Observability. *Proceedings of the AAAI Conference on Artificial Intelligence* 39, 14 (Apr. 2025), 15092–15099. https://doi.org/10.1609/aaai.v39i14.33655

[38] Aniello Murano, Daniel Neider, and Martin Zimmermann. 2023. Robust Alternating-Time Temporal Logic. In *Proc. of JELIA 2023*. Springer, 796–813.

[39] Alberto Puggelli, Wenchao Li, Alberto L. Sangiovanni-Vincentelli, and Sanjit A. Seshia. 2013. Polynomial-Time Verification of PCTL Properties of MDPs with Convex Uncertainties. In *Proc. of CAV 2013*. 527–542.

[40] Alberto Puggelli, Alberto L. Sangiovanni-Vincentelli, and Sanjit A. Seshia. 2014. Robust strategy synthesis for probabilistic systems applied to risk-limiting renewable-energy pricing. In *Proc. of EMSOFT 2014*. ACM, 13:1–13:10.

[41] Marcus Schaefer and Daniel Stefankovic. 2024. Beyond the Existential Theory of the Reals. *Theory Comput. Syst.* 68, 2 (2024), 195–226. https://doi.org/10.1007/S00224-023-10151-X

[42] Fu Song, Yedi Zhang, Taolue Chen, Yu Tang, and Zhiwu Xu. 2019. Probabilistic alternating-time μ-calculus. In *Proc. of AAAI*. 6179–6186.

[43] Marnix Suilen, Nils Jansen, Murat Cubuktepe, and Ufuk Topcu. 2020. Robust Policy Synthesis for Uncertain POMDPs via Convex Optimization. In *Proc. of IJCAI 2020*. 4113–4120.

[44] Wayes Tushar, Chau Yuen, Tapan K Saha, Thomas Morstyn, Archie C Chapman, M Jan E Alam, Sarmad Hanif, and H Vincent Poor. 2021. Peer-to-peer energy systems for connected communities: A review of recent advances and emerging challenges. *Applied Energy* 282 (2021), 116131.

[45] Wayes Tushar, Chau Yuen, Tapan K Saha, Sohrab Nizami, Mollah Rezaul Alam, David B Smith, and H Vincent Poor. 2023. A survey of cyber-physical systems from a game-theoretic perspective. *IEEE Access* 11 (2023), 9799–9834.

[46] Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem. 2013. Robust Markov Decision Processes. *Math. Oper. Res.* 38, 1 (2013), 153–183. https://doi.org/10.1287/MOOR.1120.0566

[47] Michael Wooldridge. 2009. *An Introduction to Multiagent Systems*. John Wiley & Sons.

[48] Pian Yu, Shufang Zhu, Giuseppe De Giacomo, Marta Kwiatkowska, and Moshe Y. Vardi. 2024. The Trembling-Hand Problem for LTLf Planning. In *Proc. of IJCAI (2024)*.

[49] Quanyan Zhu and Tamer Basar. 2015. Game-Theoretic Methods for Robustness, Security, and Resilience of Cyberphysical Control Systems: Games-in-Games Principle for Optimal Cross-Layer Resilient Control Systems. *IEEE Control Systems Magazine* 35, 1 (2015), 46–65. https://doi.org/10.1109/MCS.2014.2364710

[50] Martin Zimmermann. 2025. Robust probabilistic temporal logics. *Inform. Process. Lett.* 188 (2025), 106522.