

# Rational Verification for Probabilistic Systems

Julian Gutierrez<sup>1</sup>, Lewis Hammond<sup>2</sup>, Anthony W. Lin<sup>3</sup>, Muhammad Najib<sup>3</sup>, Michael Wooldridge<sup>2</sup>

<sup>1</sup>Monash University

<sup>2</sup>University of Oxford

<sup>3</sup>University of Kaiserslautern

julian.gutierrez@monash.edu, {lewis.hammond, mjw}@cs.ox.ac.uk, {lin, najib}@cs.uni-kl.de

## Abstract

Rational verification is the problem of determining which temporal logic properties will hold in a multi-agent system, under the assumption that agents in the system act rationally, by choosing strategies that collectively form a game-theoretic equilibrium. Previous work in this area has largely focussed on deterministic systems. In this paper, we develop the theory and algorithms for rational verification in probabilistic systems. We focus on *concurrent stochastic games* (CSGs), which can be used to model uncertainty and randomness in complex multi-agent environments. We study the rational verification problem for both non-cooperative games and cooperative games in the qualitative probabilistic setting. In the former case, we consider LTL properties satisfied by the Nash equilibria of the game and in the latter case LTL properties satisfied by the core. In both cases, we show that the problem is 2EXPTIME-complete, thus not harder than the much simpler verification problem of model checking LTL properties of systems modelled as Markov decision processes (MDPs).

## 1 Introduction

*Rational verification* is the problem of determining which temporal logic properties will hold in a multi-agent system, under the assumption that agents in the system act rationally, by choosing strategies/policies for acting which collectively form a game-theoretic equilibrium (Gutierrez, Harrenstein, and Wooldridge 2017; Wooldridge et al. 2016). Rational verification has been studied for a range of models: typically, each agent is modelled as a non-deterministic reactive program, where non-determinism captures the choices available to each agent at each time step – the strategies available to a player correspond to each possible way that an agent can resolve its non-determinism. To be able to reason about game-theoretic equilibria, the model also needs to capture the preferences that players have, and a common approach for this is to associate with each agent a temporal logic “goal” formula that the player desires to be satisfied. For example, in the Reactive Modules reasoning framework, agents are modelled using the Reactive Modules language (Alur and Henzinger 1999), and agent preferences are modelled with goals expressed in linear temporal logic (LTL) (Gutierrez et al. 2020; Gutierrez et al. 2018).

The most basic decision problem in rational verification is as follows: *Given a system  $M$ , and a temporal logic for-*

*mula  $\varphi$ , does there exist a Nash equilibrium profile of strategies  $\vec{\sigma} = (\sigma_1, \dots, \sigma_n)$  for the players in  $M$  such that  $\varphi$  will be satisfied under the assumption that players act according to  $\vec{\sigma}$ . This decision problem is known as E-NASH, and is 2EXPTIME-complete for Reactive Modules games (Gutierrez, Harrenstein, and Wooldridge 2017); the corresponding A-NASH problem asks, instead, whether formula  $\varphi$  holds for all Nash equilibrium profiles in the game.*

Although many models have been studied in the context of rational verification, little research has considered *probabilistic* models, and as such, existing models are limited in the scope of domains they can capture. Our aim in this paper is to rectify this omission: we study rational verification in probabilistic systems in which players have goals represented by LTL formulae. Our basic model is called *concurrent stochastic games* (CSGs), sometimes also referred to as *Markov games*. As in conventional concurrent games, a game is played over an infinite sequence of rounds, and at each round, every player chooses an action to perform. Unlike conventional concurrent games, however, the performance of a profile of actions does not induce a unique successor state, but rather a probability distribution over possible successor states. The main difference between the games we study and standard CSGs is that player preferences in our setting are defined by associating LTL goals  $\gamma_i$  with each player  $i$ . The game is played by each player choosing a strategy (cf. *policy*), which defines how that player will make choices over time when playing the game.

In common with previous work, we model strategies as state machines with output (although we may require them to have infinite memory), though strategies in our setting are not required to choose a unique action at every time step, but instead choose a probability distribution over possible actions. A CSG together with a profile of strategies induces a Markov chain (MC), and given such an MC, we can determine the probability of given temporal formulae being satisfied, and in particular, the probability with which goal formulae  $\gamma_i$  are satisfied. In this paper, we consider rational verification of CSGs in the *qualitative* setting, *i.e.*, where we are interested in checking if the probability of satisfying LTL goals is 1 or greater than 0. These are also known as *almost-sure* (AS) and *non-zero* (NZ) satisfaction respectively, which – together with their negated formulations, probability less than 1 or equal to 0 – form one of the most useful settings

in probabilistic verification (Baier and Katoen 2008). For example, they can be used to capture interesting liveness properties for many randomised protocols (Norman 2004; Lehmann and Rabin 1981; Lengál et al. 2017; Lin and Rümmer 2016).

Against this background, we study the rational verification problem for both cooperative and non-cooperative solution concepts. In the *non-cooperative* case, we focus on *Nash equilibria*, and characterise the complexity of the E-NASH and A-NASH problems, together with related problems. We then investigate *cooperative* solution concepts, adapting the model of the *core* that was introduced by (Gutierrez, Kraus, and Wooldridge 2019), which defines strategy profiles that are stable against beneficial deviations by *groups* of players (*coalitions*), as opposed to individual deviations in the case of Nash equilibria. We consider E-CORE, A-CORE, and related decision problems – which in previous work have been studied only for concurrent multi-agent systems without any probabilistic behaviour.

To the best of our knowledge, this is the first work that considers the rational verification problem for probabilistic systems in which *games* can be cooperative or non-cooperative, players' *preferences* can be expressed using general LTL goals, strategies may have access to *infinite* memory, interactions can take place *concurrently*, and plays may last for an infinite number of rounds (*i.e.*, have an *infinite horizon*). All previous work, in the probabilistic setting, fails to have at least one of these features arising in full interplay with the others, making our framework the most complex so far developed from a theoretical point of view. Indeed, because of the many features we consider together, several new constructions and proof techniques are required to be able to fully account for probabilistic behaviour.

**Structure of the paper** In Section 2, the necessary background is given. Sections 3 and 4 contain the main technical results for non-cooperative games and cooperative games respectively. Then, in Section 5, we present some concluding remarks and a summary of relevant related work.

## 2 Preliminaries

For a finite set  $X$ , a (rational) *probability distribution* over  $X$  is a function  $\text{Pr} : X \rightarrow [0, 1] \cap \mathbb{Q}$  such that  $\sum_{x \in X} \text{Pr}(x) = 1$ . We write  $\text{D}(X)$  for the set of probability distributions on  $X$ , and  $\text{spt}(\text{Pr}) = \{x \in X : \text{Pr}(x) > 0\}$  for the *support* of the distribution  $\text{Pr}$  on  $X$ . For a tuple  $\vec{x} = (x_1, \dots, x_n)$ , we write  $\text{proj}_i(\vec{x}) = x_i$ , *i.e.*, its *i*-th *projection*, and also  $\text{proj}_{x_i}(\vec{x}) = x_i$  when the context is clear.

**Markov chains** A (discrete time) *Markov chain* (MC) is a tuple  $\mathcal{C} = (S, s_i, \text{tr}, \lambda)$ , where  $S$  is a set of states,  $s_i$  is the initial state,  $\text{tr} : S \rightarrow \text{D}(S)$  is a function that assigns a probability distribution (on the set of states  $S$ ) to all states  $s \in S$ , and  $\lambda : S \rightarrow 2^{\text{AP}}$  is a labelling function mapping each state to a set of propositions taken from the set AP.

The set of infinite paths in  $\mathcal{C}$  starting from  $s \in S$  is  $\text{Paths}(\mathcal{C}, s) = \{\pi = s_0 s_1 \dots \in \text{St}^\omega : s_0 = s, \forall k \in \mathbb{N}. \text{tr}(s_k, s_{k+1}) > 0\}$ . The set of all infinite paths in  $\mathcal{C}$  is  $\text{Paths}(\mathcal{C}) = \bigcup_{s \in S} \text{Paths}(\mathcal{C}, s)$ . The set of finite paths

starting from  $s \in S$  is defined as  $\text{Fpaths}(\mathcal{C}, s) = \{\hat{\pi} = s_0 \dots s_n \in S^+ : \exists \pi \in \text{Paths}(\mathcal{C}). \hat{\pi} \pi \in \text{Paths}(\mathcal{C}, s)\}$  and  $\text{Fpaths}(\mathcal{C}) = \bigcup_{s \in S} \text{Fpaths}(\mathcal{C}, s)$ . The *cylinder set* of a finite path  $\hat{\pi} \in \text{Fpaths}(\mathcal{C})$  is defined by  $\text{Cyl}(\hat{\pi}) = \{\pi \in \text{Paths}(\mathcal{C}) : \exists \tilde{\pi} \in \text{Paths}(\mathcal{C}). \pi = \hat{\pi} \tilde{\pi} \in \text{Paths}(\mathcal{C})\}$ . Following (Vardi 1985), we define the probability distribution over the space of infinite paths, as usual, via cylinder sets. We denote this probability distribution over the set of infinite paths beginning from some state  $s$  by  $\text{Pr}_{\mathcal{C}}^s$ . We also write  $\text{Pr}_{\mathcal{C}}$  when  $s$  is clear from the context.

**Concurrent stochastic game arenas** A *concurrent stochastic game arena* (CSGA) is a tuple  $\mathcal{M} = (\text{N}, \text{St}, s^0, (\text{Ac}_i)_{i \in \text{N}}, \text{tr})$ , where  $\text{N}$  is a finite set of players,  $\text{St}$  is a finite set of states,  $s^0$  is the initial state,  $\text{Ac}_i$  is a finite set of actions for each  $i \in \text{N}$ . With each player  $i$  and each state  $s \in \text{St}$ , we associate a non-empty set  $\text{Ac}_i(s)$  of *available* actions that, intuitively,  $i$  can perform when in state  $s$ . When all players have fixed their actions, we have an action profile  $\vec{a} = (a_1, \dots, a_n) \in \vec{\text{Ac}} = \text{Ac}_1 \times \dots \times \text{Ac}_n$  which we refer as a *direction*. A direction  $\vec{a}$  is available in state  $s$  if for all  $i$  we have  $a_i \in \text{Ac}_i(s)$ . We write  $\vec{\text{Ac}}(s)$  for the set of available directions in state  $s$ .

For a given set of players  $A \subseteq \text{N}$  and an action profile  $\vec{a}$ , we let  $\vec{a}_A$  and  $\vec{a}_{-A}$  be two tuples of actions, respectively, one for each player in  $A$  and one for each player in  $\text{N} \setminus A$ . Furthermore, for two directions  $\vec{a}$  and  $\vec{a}'$ , we write  $(\vec{a}_A, \vec{a}'_{-A})$  to denote the direction where the actions for players in  $A$  are taken from  $\vec{a}$  and the actions for players in  $\text{N} \setminus A$  are taken from  $\vec{a}'$ . Finally,  $\text{tr} : \text{St} \times \vec{\text{Ac}} \rightarrow \text{D}(\text{St})$  is a probabilistic transition function. A Markov decision process (MDP), without a reward function, is simply a CSGA with one player only.

**Linear temporal logic** LTL (Pnueli 1977) extends classical propositional logic with two operators, **X** (“next”) and **U** (“until”), which can be used to express properties of paths. The syntax of LTL is defined with respect to a set AP of propositional variables by the following grammar:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathbf{X}\varphi \mid \varphi \mathbf{U} \varphi$$

where  $p \in \text{AP}$ . Other connectives are defined in terms of  $\neg$  and  $\vee$  in the usual way. Two key derived LTL operators are **F** (“eventually”) and **G** (“always”), which are defined in terms of **U** as follows:  $\mathbf{F}\varphi \equiv \top \mathbf{U} \varphi$  and  $\mathbf{G}\varphi \equiv \neg \mathbf{F} \neg \varphi$ .

We interpret formulae of LTL with respect to triples  $(\pi, t, \lambda)$ , where  $\pi \in \text{St}^\omega$  is a path,  $t \in \mathbb{N}$  is a temporal index into  $\pi$ , and  $\lambda : \text{St} \rightarrow 2^{\text{AP}}$  is a labelling function that indicates which propositional variables are true in every state. The semantics of LTL is given by the following rules:

$$\begin{array}{ll} (\pi, t, \lambda) \models \top & \\ (\pi, t, \lambda) \models p & \text{iff } p \in \lambda(\pi_t) \\ (\pi, t, \lambda) \models \neg\varphi & \text{iff it is not the case that } (\pi, t, \lambda) \models \varphi \\ (\pi, t, \lambda) \models \varphi \vee \psi & \text{iff } (\pi, t, \lambda) \models \varphi \text{ or } (\pi, t, \lambda) \models \psi \\ (\pi, t, \lambda) \models \mathbf{X}\varphi & \text{iff } (\pi, t+1, \lambda) \models \varphi \\ (\pi, t, \lambda) \models \varphi \mathbf{U} \psi & \text{iff for some } t' \geq t : ((\pi, t', \lambda) \models \psi \text{ and} \\ & \text{for all } t \leq t'' < t' : (\pi, t'', \lambda) \models \varphi) \end{array}$$

If  $(\pi, 0, \lambda) \models \varphi$ , we write  $\pi \models \varphi$  and say that  $\pi$  *satisfies*  $\varphi$ .

**Concurrent stochastic games** A *concurrent stochastic game* (CSG) is a tuple  $\mathcal{G} = (\mathcal{M}, (\gamma_i)_{i \in \mathbb{N}}, \lambda)$ , where  $\mathcal{M}$  is a CSGA,  $\gamma_i$  is a LTL formula that represents the *goal* of player  $i$ , and  $\lambda : \text{St} \rightarrow 2^{\text{AP}}$  a labelling function. A game is played by each player  $i$  selecting a *strategy*  $\sigma_i$  that defines how it makes choices over time. A strategy for player  $i$  can be understood as a function  $\sigma_i : \text{St}^+ \rightarrow \text{D}(\text{Ac}_i)$  that assigns to every non-empty finite sequence of states a probability distribution over player  $i$ 's set of actions. In general, strategies require memory to remember the history of the game. When a strategy remembers a finite amount of information about the past we call it *finite-memory*, and when each distribution  $\sigma_i(s^+)$  is deterministic we call  $\sigma_i$  *pure*.

Formally, a *strategy* in  $\mathcal{G}$  for player  $i$  is a transducer  $\sigma_i = (Q_i, q_i^0, \delta_i, \tau_i)$ , where  $Q_i$  is a (possibly infinite) set of *internal states*,  $q_i^0$  is the *initial state*,  $\delta_i : Q_i \times \text{St} \rightarrow Q_i$  is a deterministic *internal transition function*, and  $\tau_i : Q_i \times \text{St} \rightarrow \text{D}(\text{Ac}_i)$  an *action function* that selects a distribution on  $\text{Ac}_i$  such that for all  $q_i \in Q_i$  and  $s \in \text{St}$ , we have  $\tau_i(q_i, s) \in \text{D}(\text{Ac}_i(s))$ . Let  $\Sigma_i$  be the set of strategies for player  $i$ . A strategy is *memoryless* if there exists a transducer encoding the strategy with  $|Q_i| = 1$ , *i.e.*, the choice of action only depends on the current state of the game, and finite-memory if  $|Q_i| < \infty$ . Moreover, a strategy is said to be *deterministic* if  $\tau_i : Q_i \times \text{St} \rightarrow \text{Ac}_i$ , such that for every  $q_i \in Q_i$  and every  $s \in \text{St}$ , we have that  $\tau_i(q_i, s) \in \text{Ac}_i(s)$ .

Once every player  $i$  has selected a strategy  $\sigma_i$ , we have a *strategy profile*  $\vec{\sigma} = (\sigma_1, \dots, \sigma_n)$ . We write  $\vec{\sigma}_A$  and  $\vec{\sigma}_{-A}$  to denote the strategy profile for players in  $A \subseteq \mathbb{N}$  and  $\mathbb{N} \setminus A$ , respectively. We also write  $(\vec{\sigma}_A, \vec{\sigma}'_{-A})$  to denote the strategy profile where the strategies for players in  $A$  are taken from  $\vec{\sigma}$ , and the strategies for players in  $\mathbb{N} \setminus A$  are taken from  $\vec{\sigma}'$ . Observe that a strategy profile  $\vec{\sigma}$  for a game  $\mathcal{G}$  resolves non-determinism in the underlying  $\mathcal{M}$ . That is, a strategy profile  $\vec{\sigma}$  for a game  $\mathcal{G}$  induces an MC  $\mathcal{C}_{\vec{\sigma}} = (S, s_\iota, \text{tr}', \lambda)$ , where  $S = \text{St} \times \prod_{i \in \mathbb{N}} Q_i$ ,  $s_\iota = (s^0, q_1^0, \dots, q_n^0)$ , and for  $v, v' \in S$ ,  $\text{tr}'(v, v') = \sum_{\vec{a} \in \vec{\text{Ac}}} \prod_{\vec{a}_i \in \vec{a}} \tau_i(\text{proj}_{q_i}(v), \text{proj}_s(v), \vec{a}_i) \cdot \text{tr}(\text{proj}_s(v), \vec{a}, \text{proj}_s(v'))$ , if for each  $i \in \mathbb{N}$ ,  $\text{proj}_{q_i}(v') = \delta_i(\text{proj}_{q_i}(v), \text{proj}_s(v))$ , and is not defined otherwise.

**Automata** A *deterministic automaton on infinite words* is given by a structure  $\mathcal{A} = (\text{AP}, Q, q^0, \rho, \mathcal{F})$ , where  $Q$  is a finite set of states,  $\rho : Q \times \text{AP} \rightarrow Q$  is a transition function,  $q^0$  is an initial state, and  $\mathcal{F}$  is an acceptance condition. A *parity condition*  $\mathcal{F}$  is a partition  $\{F_1, \dots, F_n\}$  of  $Q$ , where  $n$  is the *index* of the parity condition and any  $k \in [1, n]$  is a *priority*. We use a *priority function*  $\alpha : Q \rightarrow \mathbb{N}$  that maps states to priorities such that  $\alpha(q) = k$  if and only if  $q \in F_k$ . For a run  $\pi = q^0 q^1 q^2 \dots$ , let  $\text{inf}(\pi)$  denote the set of states occurring infinitely often in the run  $\text{inf}(\pi) = \{q \in Q : q = q^i \text{ for infinitely many } i\}$ . A run  $\pi$  is accepted by a deterministic parity word (DPW) automaton with condition  $\mathcal{F}$  if the minimum priority that occurs infinitely often is even, *i.e.*, if  $(\min_{k \in [1, n]} (\text{inf}(\pi) \cap F_k) \neq \emptyset) \bmod 2 = 0$ .

For a given game  $\mathcal{G}$  and a strategy profile  $\vec{\sigma}$ , a formula  $\varphi$  is said to be *almost-surely* satisfied, denoted  $\vec{\sigma} \models \text{AS}(\varphi)$ , if and only if,  $\Pr_{\mathcal{C}_{\vec{\sigma}}}(\{\pi \in \text{Paths}(\mathcal{C}_{\vec{\sigma}}, s^0) : \pi \models \varphi\}) = 1$ . Similarly, we say that  $\varphi$  is satisfied with *non-zero* probability, denoted  $\vec{\sigma} \models \text{NZ}(\varphi)$  if  $\Pr_{\mathcal{C}_{\vec{\sigma}}}(\{\pi \in \text{Paths}(\mathcal{C}_{\vec{\sigma}}, s^0) :$

$\pi \models \varphi\}) > 0$ . Observe that NZ can be viewed as the dual of AS, written (with a slight abuse of notation) as  $\text{NZ}(\varphi) \equiv \neg \text{AS}(\neg \varphi)$ . Hence, for ease of exposition, in the remainder of the paper we focus on AS winning conditions, with the understanding that all our results from this case can be also used in the case of NZ winning conditions, and their respective negated formulations.

A *concurrent multiplayer stochastic parity game* (CSPG) is given by a structure  $\mathcal{G}_{\text{PAR}} = (\mathcal{M}, (\alpha_i)_{i \in \mathbb{N}})$  where  $\alpha_i : \text{St} \rightarrow \mathbb{N}$  is the goal of player  $i$ , given as a priority function over the set of states  $\text{St}$ . A path  $\pi$  satisfies a priority function  $\alpha$ , denoted by  $\pi \models \alpha$ , if the minimum number occurring infinitely often in the infinite sequence  $\alpha(\pi_0)\alpha(\pi_1)\alpha(\pi_2)\dots$  is even. Almost-surely satisfaction in CSPGs is then defined in a similar way: we say that  $\vec{\sigma} \models \text{AS}(\alpha)$  if and only if  $\Pr_{\mathcal{C}_{\vec{\sigma}}}(\{\pi \in \text{Paths}(\mathcal{C}_{\vec{\sigma}}, s^0) : \pi \models \alpha\}) = 1$ .

For a CSG  $\mathcal{G}$ , strategy profile  $\vec{\sigma}$ , and state  $s$ , we define the set of winners and losers by  $W_{\mathcal{G}}(\vec{\sigma}, s) = \{i \in \mathbb{N} : (\vec{\sigma}, s) \models \text{AS}(\gamma_i)\}$  and  $L_{\mathcal{G}}(\vec{\sigma}, s) = \{i \in \mathbb{N} : (\vec{\sigma}, s) \models \neg \text{AS}(\gamma_i)\}$ . We also write  $W_{\mathcal{G}}(\vec{\sigma})$  and  $L_{\mathcal{G}}(\vec{\sigma})$ , for  $W_{\mathcal{G}}(\vec{\sigma}, s^0)$  and  $L_{\mathcal{G}}(\vec{\sigma}, s^0)$ . We define the above concepts for CSPGs analogously, with  $\mathcal{G}_{\text{PAR}}$  replacing  $\mathcal{G}$  and  $\alpha_i$  replacing  $\gamma_i$ .

### 3 Non-Cooperative Rational Verification

We now introduce rational verification problems involving non-cooperative solution concepts – and in particular, problems relating to *Nash equilibria* (Osborne and Rubinstein 1994). We begin by defining this concept for our setting: Given a game  $\mathcal{G}$ , a strategy profile  $\vec{\sigma}$  is a *Nash equilibrium* of  $\mathcal{G}$  if, for every player  $i$  and strategy  $\sigma'_i \in \Sigma_i$ , we have

$$(\vec{\sigma}_{-i}, \sigma'_i) \models \text{AS}(\gamma_i) \text{ implies } \vec{\sigma} \models \text{AS}(\gamma_i)$$

where  $(\vec{\sigma}_{-i}, \sigma'_i)$  denotes  $(\sigma_1, \dots, \sigma_{i-1}, \sigma'_i, \sigma_{i+1}, \dots, \sigma_n)$ , the strategy profile where the strategy of player  $i$  in  $\vec{\sigma}$  is replaced by  $\sigma'_i$ . Note that this is equivalent to a more traditional formulation in which the utility function of each player  $i$  is defined as equal to a constant  $a$  if  $\text{AS}(\gamma_i)$  holds and equal to a constant  $b < a$  otherwise. Let  $\text{NE}(\mathcal{G})$  denote the set of Nash equilibria of  $\mathcal{G}$ . We begin by introducing the key rational verification problems for non-cooperative settings; these are the natural adaptation of rational verification for the almost-sure setting.

#### MEMBERSHIP

*Given:* Game  $\mathcal{G}$ , strategy profile  $\vec{\sigma}$ .

*Question:* Is it the case that  $\vec{\sigma} \in \text{NE}(\mathcal{G})$ ?

#### E-NASH

*Given:* Game  $\mathcal{G}$ , LTL formula  $\varphi$ .

*Question:* Is it the case that  $\exists \vec{\sigma} \in \text{NE}(\mathcal{G}). \vec{\sigma} \models \text{AS}(\varphi)$ ?

We can also ask the obvious counterpart of E-NASH:

#### A-NASH

*Given:* Game  $\mathcal{G}$ , LTL formula  $\varphi$ .

*Question:* Is it the case that  $\forall \vec{\sigma} \in \text{NE}(\mathcal{G}). \vec{\sigma} \models \text{AS}(\varphi)$ ?

The intuitively simpler question of asking whether a game  $\mathcal{G}$  has any Nash equilibria, typically known as NON-EMPTYNESS in the rational verification literature, can be solved simply by checking if  $(\mathcal{G}, \top) \in \text{E-NASH}$ . Note that

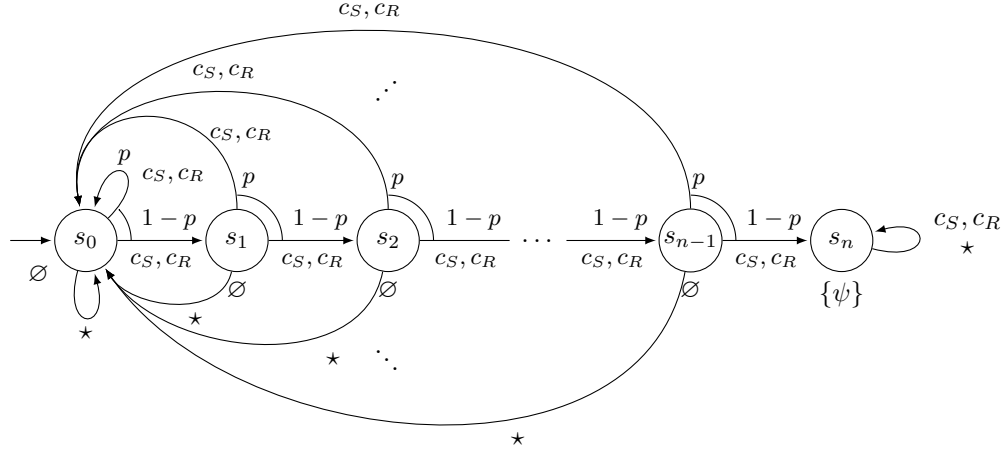


Figure 1: The CSGA for Examples 1 and 2. Edges without probability labels mean they have probability 1, and edges labelled  $\star$  are executed if the action profile is not equal to  $c_S, c_R$ .

the question of NON-EMPTYNESS may be non-trivial, as the fact that in our setting strategies can have infinite memory (and thus there are infinitely many of them) means we cannot straightforwardly apply Nash’s theorem.

To illustrate some of the concepts introduced above, we will make use of the following example.

**Example 1.** Suppose we have a sender and a receiver who want to transmit some data. The sender and the receiver can be in either of two modes: idle or communicating. The data is sent sequentially in  $n$  blocks and successful transmission is assumed to be continuous, i.e., no gaps (missing blocks) are allowed between blocks; if there are gaps, the transmission fails and has to be restarted from the beginning. In order to be able to send the data, both sender and receiver have to be in the communicating mode. Furthermore, suppose that the network is noisy, thus for each block being transmitted, it may be lost with probability  $p$ .

To capture this, consider a game with  $\mathbb{N} = \{S, R\}$  representing the sender  $S$  and receiver  $R$ . The set of actions for player  $j \in \mathbb{N}$  is  $\text{Ac}_j = \{c_j, i_j\}$ , where  $c_j$  and  $i_j$  mean that player  $j$  is communicating or idle, respectively. The arena of the game is shown in Figure 1. Being in state  $s_i$  indicates that  $i$  blocks have been successfully transmitted, with  $s_n$  forming a sink state. The goal of each player  $j$  is to almost-surely satisfy  $\gamma_j = \mathbf{F}\psi$ .

There are infinitely many Nash equilibria in this game. However they can be classified into two categories: (a) those that satisfy the goals of each player; and (b) those that do not. For category (a), observe that any strategy profile that prescribes the action  $c_j$  for each player  $j$  with probability strictly greater than zero in every state  $s \in \{s_0, \dots, s_{n-1}\}$  is a Nash equilibrium. For (b), any strategy profile that prescribes  $i_j$  with probability one for each player  $j$  in any state  $s \in \{s_0, \dots, s_{n-1}\}$  is also a Nash equilibrium.

Thus, the answer E-NASH with  $\varphi \equiv \mathbf{F}\psi$  (i.e., the data is eventually sent) is “yes”, since there exist Nash equilibria

that satisfy  $\varphi$  with probability one, namely, the strategy profiles that belong to category (a). On the other hand, the answer to the A-NASH query with the same property  $\varphi$  is “no”, since we have equilibria that belong to category (b).

Similar, but more realistic, versions of Example 1 can be constructed. For example, consider a distributed system with  $n$  servers  $\{S_i\}_{0 \leq i \leq n}$ , each of which has a message inbox (a FIFO queue channel) that can contain up to  $k$  messages. Each server  $S_i$  can send a message  $m$  to another server  $S_j$ , where  $m \in \Sigma$  for some finite  $\Sigma$ . This happens instantaneously in that  $m$  is placed on the FIFO channel of the receiver, although with a probability  $p \in (0, 1)$  that this fails. We denote this action by  $\text{snd}(i, j, m)$ . Each server can also opt to perform two other actions: pop the first message in the inbox (denoted by  $\text{pop}(j, i, m)$ , meaning that  $S_j$  pops message  $m$  sent by  $S_i$ ), or remain idle. In addition, we require that each server must pop a message when the inbox is full. The system operates fully concurrently: each server acts completely independently of each other.

The goal of each server  $S_i$  is

$$\gamma_i = \bigwedge_{j,m} \mathbf{G}(\text{snd}(i, j, m) \rightarrow \mathbf{F}\text{pop}(j, i, m));$$

namely, each message that is sent has to be eventually received. It is not difficult to show that given the property  $\varphi = \bigwedge_{i=1}^n \gamma_i$ , both E-NASH and A-NASH queries return “yes” answers. This is because a server  $S_i$  that has sent a message  $m$  to server  $S_j$  can resend the message, until it knows that the message has been popped by  $S_j$ , thus forcing the LTL goal to be satisfied almost surely.

The above system is an example of a so-called stochastic lossy channel system (Abdulla et al. 2008; Baier, Bertrand, and Schnoebelen 2007), but restricted to bounded channels, which is reasonable in practice, wherein memories are bounded. Of course, communicating systems that employ channels are quite realistic in practice (as in, e.g., the Erlang

programming language), and handling the possibility of a message loss is important in the study of large communicating and distributed computer systems in general.

In the remainder of the paper, it will be useful to sometimes consider a two-player zero-sum variant of an existing game in which the set of players is partitioned into two coalitions,  $A \subseteq N$  and  $N \setminus A$ . In the case of LTL objectives then  $A$  has goal  $\psi$  and  $N \setminus A$  has goal  $\neg\psi$ , and for parity objectives  $A$  has even parity for priority function  $\alpha : St \rightarrow \mathbb{N}$  and  $N \setminus A$  odd parity. We define this formally as follows.

**Definition 1.** Let  $\mathcal{G} = (\mathcal{M}, (\gamma_i)_{i \in N}, \lambda)$  be a CSG whose underlying arena is  $\mathcal{M} = (N, (Ac_i)_{i \in N}, St, s^0, tr)$  and let  $A \subseteq N$ . Then the *two-player coalition game arena* is defined as  $\mathcal{M}^A = ((1, 2), (\times_{i \in A} Ac_i, \times_{i \in N \setminus A} Ac_i), St, s^0, tr_A)$  where  $tr_A(s, (a_1, a_2)) = tr(s, (\bar{a}_A, \bar{a}_{-A}))$ . The *two-player LTL coalition game* with respect to  $\mathcal{G}$ ,  $A$ , and LTL formula  $\psi$ , is thus defined as  $\mathcal{G}^{A, \psi} = (\mathcal{M}^A, (\psi, \neg\psi), \lambda)$ , and the *two-player parity coalition game* with respect to  $\mathcal{G}$ ,  $A$ , and priority function  $\alpha$ , is defined as  $\mathcal{G}^{A, \alpha} = (\mathcal{M}^A, (\alpha, \bar{\alpha}))$  where  $\bar{\alpha}(s) = \alpha(s) + 1$  for any state  $s \in St$ .

In the remaining subsections we address the three main decision problems considered in the non-cooperative setting.

### 3.1 MEMBERSHIP

Recall that the MEMBERSHIP problem requires two inputs: a game  $\mathcal{G}$  and a strategy profile  $\bar{\sigma}$ . We then ask if  $\bar{\sigma}$  forms a Nash equilibrium. Note that, in general, infinite memory strategies are needed to play concurrent  $\omega$ -regular games with almost-sure winning conditions (Chatterjee and Henzinger 2012), however for this problem we assume that the input  $\bar{\sigma}$  is represented by some finite state transducer. An optimal procedure for solving MEMBERSHIP is given by Algorithm 1, as shown by the following theorem.

**Theorem 1.** MEMBERSHIP is 2EXPTIME-complete.

*Proof.* Observe that checking line 2 of Algorithm 1 amounts to (qualitative) model checking of the LTL formula  $\gamma_i$  on the resulting MC  $\mathcal{C}_{\bar{\sigma}}$ , that is, after non-determinism in  $\mathcal{M}$  is resolved by  $\bar{\sigma}$ . This step can be done in PSPACE (Courcoubetis and Yannakakis 1995). Checking line 3 amounts to LTL model checking on the CSGA  $\mathcal{M}_{\bar{\sigma}_i}$ , i.e., model checking LTL over MDP, which is 2EXPTIME-complete (Courcoubetis and Yannakakis 1995). Therefore, we have a 2EXPTIME procedure for solving MEMBERSHIP.

For hardness, we reduce from qualitative LTL model checking on MDPs. Given an MDP  $\mathcal{K}$  with labelling function  $\lambda$  and an LTL formula  $\varphi$ , we build a corresponding MEMBERSHIP instance  $(\mathcal{G}, \bar{\sigma})$  as follows.  $\mathcal{G} = (\mathcal{M}, \gamma_1, \lambda')$  is a CSG, where  $\gamma_1 = \mathbf{X}(\varphi \wedge p)$  and  $p$  is a fresh variable. The CSGA  $\mathcal{M} = (N, St, s^0, (Ac_i)_{i \in N}, tr)$  is built from  $\mathcal{K}$  with only a single player, two additional states, and two fresh additional actions. Formally,  $N = \{1\}$ ,  $St = S \cup \{s_\infty, s^0\}$ ,  $Ac_1 = Ac \cup \{a, \bar{a}\}$ , and

$$tr(s, a) = \begin{cases} s_\infty, & \text{if } (s, a_1) = (s^0, a) \text{ or } s = s_\infty \\ s_l, & \text{if } (s, a_1) = (s^0, \bar{a}) \\ Pr(s, a), & \text{otherwise.} \end{cases}$$

### Algorithm 1 MEMBERSHIP

---

**input:**  $\mathcal{G}, \bar{\sigma}$   
1: **for**  $i \in N$  **do**  
2:     **if**  $\bar{\sigma} \not\models AS(\gamma_i)$  **then**  
3:         **if**  $\exists \sigma'_i \in \Sigma_i$  s.t.  $(\bar{\sigma}_{-i}, \sigma'_i) \models AS(\gamma_i)$  **then**  
4:             **return** “no”  
5: **return** “yes”

---

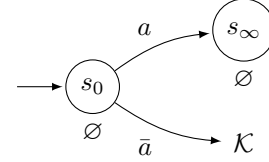


Figure 2: The CSGA for our reduction from MEMBERSHIP to qualitative LTL model checking in MDPs. Edges without probability labels mean they have probability 1.

The labelling function  $\lambda'$  is the same as  $\lambda$  except that  $\lambda(s_\infty) = \emptyset$  and  $\lambda'(s_l) = \lambda(s_l) \cup \{p\}$ . An illustration of the construction of  $\mathcal{M}$  is shown in Figure 2.

The strategy profile is defined as  $\bar{\sigma} = (\sigma_1)$ , where  $\sigma_1(s^0) = a$ , i.e., action  $a$  is chosen with probability 1. Observe that a “yes” answer to the MEMBERSHIP query means that  $\bar{\sigma} \in NE(\mathcal{G})$ , which implies that  $\varphi$  is not satisfied in  $\mathcal{K}$  with probability one. On the other hand, a “no” answer (i.e.,  $\bar{\sigma} \notin NE(\mathcal{G})$ ) implies that  $\varphi$  is satisfied in  $\mathcal{K}$  with probability one. Furthermore, the construction can be done in polynomial time, concluding the proof.  $\square$

### 3.2 E-NASH and A-NASH

For a given  $\mathcal{G}$  and formula  $\varphi$ , E-NASH asks whether *some* Nash equilibrium almost-surely satisfies  $\varphi$ . On the other hand, A-NASH asks whether *all* Nash equilibria almost-surely satisfy  $\varphi$ . Observe that A-NASH is closely related to E-NASH, i.e., it can be framed as an instance of E-NASH, with a small modification: rather than check whether the formula  $AS(\varphi)$  is satisfied in some Nash equilibrium, we check if  $NZ(\neg\varphi)$  is satisfied. A Nash equilibrium that satisfies  $NZ(\neg\varphi)$  is a negative witness to A-NASH. Thus, we first provide a decision procedure for solving E-NASH, and later adapt the procedure to handle the A-NASH problem.

To solve the problem, we adapt the technique presented in (Gutierrez et al. 2020). At this point, it is important to note that while our approach is inspired by the one proposed in (Gutierrez et al. 2020), the setting considered in this paper differs in multiple ways. Firstly, here we consider stochastic games and randomised strategies, while (Gutierrez et al. 2020) only considers deterministic games and pure strategies. Secondly, we allow strategies to have infinite number of states, instead of finite (albeit unbounded) states. Finally, we use almost-sure winning conditions, which does not apply to deterministic games played with pure strategies.

To describe our approach, we begin with some definitions. Let  $\mathcal{G} = (\mathcal{M}, \lambda, (\gamma_i)_{i \in N})$  be a CSG whose

underlying arena is  $\mathcal{M} = (\mathbb{N}, (Ac_i)_{i \in \mathbb{N}}, St, s^0, tr)$ , and let  $\mathcal{A}_{\gamma_i} = \langle 2^{AP}, Q_i, q_i^0, \rho_i, \alpha_i \rangle$  be the DPW corresponding to player  $i$ 's goal  $\gamma_i$  in  $\mathcal{G}$ , and  $\mathcal{A}_\varphi = \langle 2^{AP}, Q_\varphi, q_\varphi^0, \rho_\varphi, \alpha_\varphi \rangle$  to the formula  $\varphi$ . The CSPG  $\mathcal{G}_{PAR}$  associated to  $\mathcal{G}$  is  $\mathcal{G}_{PAR} = (\mathcal{M}', (\alpha'_i)_{i \in \mathbb{N}}, \alpha'_\varphi)$ , where  $\mathcal{M}' = (\mathbb{N}, (Ac_i)_{i \in \mathbb{N}}, St', s^{0'}, tr')$  and  $(\alpha'_i)_{i \in \mathbb{N}}$  are as follows:

- $St' = St \times \times_{i \in \mathbb{N}} Q_i \times Q_\varphi$  and  $s^{0'} = (s^0, q_1^0, \dots, q_n^0, q_\varphi^0)$ ;
- for each state  $(s, q_1, \dots, q_n, q_\varphi) \in St'$  and action profile  $\vec{a}$ , we define  $tr'((s, q_1, \dots, q_n, q_\varphi), \vec{a}) = (tr(s, \vec{a}), \rho_1(q_1, \lambda(s)), \dots, \rho_n(q_n, \lambda(s)), \rho_\varphi(q_\varphi, \lambda(s)))$ ;
- $\alpha'_i(s, q_1, \dots, q_n, q_\varphi) = \alpha_i(q_i)$ .
- $\alpha'_\varphi(s, q_1, \dots, q_n, q_\varphi) = \alpha_\varphi(q_\varphi)$

Observe that in the translation of  $\mathcal{G}$  to its associated  $\mathcal{G}_{PAR}$ , the set of actions for each player is unchanged. Therefore, the set of strategies in both  $\mathcal{G}$  and  $\mathcal{G}_{PAR}$  is the same, since for every state  $s \in St$  and action profile  $\vec{a}$ , it follows that  $\vec{a}$  is available in  $s$  if and only if it is available in  $(s, q_1, \dots, q_n, q_\varphi) \in St'$ , for all  $(q_1, \dots, q_n, q_\varphi) \in \times_{i \in \mathbb{N}} Q_i \times Q_\varphi$ . This, in turn, means that, for a given strategy profile  $\vec{\sigma}$ , we obtain MCs  $\mathcal{C}_{\vec{\sigma}}$  and  $\mathcal{C}_{\vec{\sigma}}^{PAR}$  that correspond to  $\mathcal{G}$  and  $\mathcal{G}_{PAR}$ , respectively. Furthermore, since the construction of  $tr'$  preserves the probability distribution assignments induced by  $tr$ , we have the following lemma.

**Lemma 1.** *For a CSG  $\mathcal{G}$  and its associated CSPG  $\mathcal{G}_{PAR}$ , it holds that  $\pi' \in Paths(\mathcal{C}_{\vec{\sigma}}^{PAR}, s^{0'})$  if and only if  $proj_s(\pi') \in Paths(\mathcal{C}_{\vec{\sigma}}, s^0)$ , where  $proj_s(\pi')$  is the  $s$  component of  $\pi'$ .*

Now, suppose that  $\vec{\sigma} \models AS(\gamma_i)$  in  $\mathcal{G}$ ; thus we have that  $Pr_{\mathcal{C}_{\vec{\sigma}}}(\{\pi \in Paths(\mathcal{C}_{\vec{\sigma}}, s^0) : \pi \models \gamma_i\}) = 1$ . Moreover, consider the component  $proj_{q_i}(\pi')$ ,  $i \in \mathbb{N}$ . By the construction of  $\mathcal{G}_{PAR}$ , it holds that  $proj_{q_i}(\pi')$  is the run executed by the DPW  $\mathcal{A}_{\gamma_i}$  when  $\lambda(\pi)$  is read, and the parity of  $\pi'$  with respect to  $\alpha'_i$  corresponds to the one recognised by  $\mathcal{A}_{\gamma_i}$ . Thus, by Lemma 1, it holds that  $Pr_{\mathcal{C}_{\vec{\sigma}}^{PAR}}(\{\pi' \in Paths(\mathcal{C}_{\vec{\sigma}}^{PAR}, s^{0'}) : proj_{q_i}(\pi') \models \alpha'_i\}) = 1$ , which implies that  $\vec{\sigma} \models AS(\alpha_i)$  in  $\mathcal{G}_{PAR}$ . Therefore, we obtain the following lemma.

**Lemma 2.** *For a CSG  $\mathcal{G}$  and its associated CSPG  $\mathcal{G}_{PAR}$ , it is the case that for every strategy profile  $\vec{\sigma}$  and player  $i$ ,  $\vec{\sigma} \models AS(\gamma_i)$  if and only if  $\vec{\sigma} \models AS(\alpha_i)$ .*

With Lemma 2 in hand, we can show that the set of Nash equilibria for any CSG  $\mathcal{G}$  exactly corresponds to the set of Nash equilibria of its associated CSPG  $\mathcal{G}_{PAR}$ . Formally, we have the following proposition.

**Proposition 1.** *Given a CSG  $\mathcal{G}$  and its associated CSPG  $\mathcal{G}_{PAR}$ , we have  $NE(\mathcal{G}) = NE(\mathcal{G}_{PAR})$ .*

*Proof.* We prove the proposition by double inclusion. Assume  $\vec{\sigma} \in NE(\mathcal{G})$ , and, by contradiction,  $\vec{\sigma} \notin NE(\mathcal{G}_{PAR})$ . Due to Lemma 2, it holds that  $W_{\mathcal{G}}(\vec{\sigma}) = W_{\mathcal{G}_{PAR}}(\vec{\sigma})$ . Then, there is a player  $j \in L_{\mathcal{G}}(\vec{\sigma})$  and a strategy  $\sigma'_j$  such that  $(\vec{\sigma}_{-j}, \sigma'_j) \models AS(\alpha_j)$  in  $\mathcal{G}_{PAR}$ . This implies that  $\sigma'_j$  is also a beneficial deviation for  $j$  in  $\mathcal{G}$  – a contradiction. On the other hand, for every  $\vec{\sigma} \in NE(\mathcal{G}_{PAR})$ , we can also reason in a symmetric way to conclude that  $\vec{\sigma} \in NE(\mathcal{G})$ .  $\square$

Proposition 1 allows us to compute Nash equilibria in CSG  $\mathcal{G}$  via its associated CSPG  $\mathcal{G}_{PAR}$ . To do this, we use Nash equilibrium characterisation presented in (Gutierrez, Harrenstein, and Wooldridge 2015a) which employs two concepts: *punishment* and *attributability*. For the former, we introduce the notion of *punishing strategy*.

**Definition 2.** For a game  $\mathcal{G}_{PAR}$ , player  $j$ , and state  $s$ , the strategy profile  $\vec{\sigma}_{-j}$  is a *punishing strategy* for player  $j$  in  $s$  if  $((\vec{\sigma}_{-j}, \sigma'_j), s) \models \neg AS(\alpha_j)$ , for every possible  $\sigma'_j$ .

We say that a state  $s$  is *punishing* for  $j$  if there exists a punishing strategy profile for  $j$  on  $s$ . Moreover, we denote by  $Pun_j(\mathcal{G}_{PAR})$  the set of punishing states for player  $j$  in  $\mathcal{G}_{PAR}$ . To compute  $Pun_j(\mathcal{G}_{PAR})$ , we solve the two-player parity coalition game  $\mathcal{G}_{PAR}^{A, \alpha_j}$  where  $A = \mathbb{N} \setminus \{j\}$ . Let  $Win_A(\mathcal{G}_{PAR}^{A, \alpha_j})$  be the set of winning states of coalition player  $A$  in  $\mathcal{G}_{PAR}^{A, \alpha_j}$ . Then the set  $Win_A(\mathcal{G}_{PAR}^{A, \alpha_j})$ , corresponds exactly to  $Pun_j(\mathcal{G}_{PAR})$ . A pair  $(s, \vec{a}) \in St \times Ac$  is *punishing-secure* for player  $j$ , if  $spt(tr(s, (\vec{a}_{-j}, a'_j))) \subseteq Pun_j(\mathcal{G}_{PAR})$  for every action  $a'_j$ . We can then extend the notion of punishing-secure pairs just defined to MCs as follows.

**Definition 3.** Given a CSPG  $\mathcal{G}_{PAR}$  and strategy profile  $\vec{\sigma}$ , the associated MC  $\mathcal{C}_{\vec{\sigma}} = (S, s_L, tr', \lambda)$  is *punishing-secure* for  $j \in L_{\mathcal{G}_{PAR}}(\vec{\sigma})$  if for every  $s, s' \in S$  and every associated  $\vec{a} \in Ac$  such that  $tr'(s, s') > 0$ ,  $(s, \vec{a})$  is punishing-secure.

Now, with those definitions in place, we can characterise Nash equilibria in CSPGs as follows.

**Proposition 2.** *Let  $\mathcal{C}_{\vec{\sigma}}$  be the associated MC of a given CSPG  $\mathcal{G}_{PAR}$  and strategy profile  $\vec{\sigma}$ . It holds that  $\vec{\sigma} \in NE(\mathcal{G}_{PAR})$  if and only if for every player  $j \in L_{\mathcal{G}_{PAR}}(\vec{\sigma})$ ,  $\mathcal{C}_{\vec{\sigma}}$  is punishing-secure for  $j$ .*

*Proof.* From left to right, suppose  $\vec{\sigma} \in NE(\mathcal{G}_{PAR})$ , and assume for a contradiction that  $\mathcal{C}_{\vec{\sigma}} = (S, s_L, tr', \lambda)$  is not punishing-secure for some  $j \in L_{\mathcal{G}_{PAR}}(\vec{\sigma})$ . This means that there is a state  $s \in S$  and action  $a'_j \in Ac_j$  such that there exists  $s' \in spt(tr(s, (\vec{a}_{-j}, a'_j)))$  where  $s' \notin Pun_j(\mathcal{G}_{PAR})$ . This, in turn, means that there exists a (deviating) strategy of player  $j$  such that there is non-zero probability of player  $j$  escaping the punishing area. By the determinacy of two-player concurrent parity games with almost-sure winning conditions (de Alfaro and Henzinger 2000), player  $j$  can thus achieve its goal with probability 1. Since this is a beneficial deviation for player  $j$ , then  $\vec{\sigma}$  is not in the set of Nash equilibria of  $\mathcal{G}_{PAR}$  – which is a contradiction.

From right to left, we first assume the existence of some  $\mathcal{C}_{\vec{\sigma}}$  that is punishing-secure for every losing player  $j$ . Such an MC can be generated by a (possibly infinite state) transducer  $T$ . Moreover, for every losing player  $j$  and every state  $s \in S$  of  $\mathcal{C}_{\vec{\sigma}}$ , there is a punishing strategy for  $j$ . Combining  $T$  with such punishing strategies, we obtain a strategy profile  $\vec{\sigma}$  that follows  $T$ , until a losing player  $j$  deviates. At this point, the concept of attributability is required, since to be able to punish, the coalition  $\mathbb{N} \setminus \{j\}$  needs to know who should be punished once a deviation happens. In order to do this, the players must be able to remember the history of play from the beginning of the game up until the point

at which a deviation happens. Notice that in general this requires infinite memory. In such a case,  $\vec{\sigma}$  would start punishing player  $j$ . Therefore, there is no beneficial deviation for player  $j$ , and strategy profile  $\vec{\sigma}$  is a Nash equilibrium of the game.  $\square$

Proposition 2 characterises Nash equilibria through the concept of a *punishing region*, i.e., a region where, for a given set of losing players  $L$ , each player  $j \in L$  can be punished. This region, denoted as  $\mathcal{G}_{\text{PAR}}^{-L}$ , is the game resulting from  $\mathcal{G}_{\text{PAR}}$  after the removal of the states that are not punishing for some  $j \in L$ , and the edges  $(s, \vec{a})$  that are not punishing-secure for some  $j \in L$ . We next observe that a positive answer for Algorithm 2 corresponds to the existence of Nash equilibrium in  $\mathcal{G}$  that satisfies  $\varphi$  with probability one, and prove that it is optimal.

**Theorem 2.** E-NASH and A-NASH are 2EXPTIME-complete.

*Proof.* Algorithm 2 runs in doubly exponential time. The underlying structure  $\mathcal{M}'$  of  $\mathcal{G}_{\text{PAR}}$  is doubly exponential in the size of the LTL goals of  $\mathcal{G}$  and formula  $\varphi$ , but the priority functions sets  $(\alpha_i)_{i \in N}$  and  $\alpha_\varphi$  are only singly exponential (Piterman 2006). Computing  $\text{Pun}_j(\mathcal{G}_{\text{PAR}})$  is polynomial in the size of  $\mathcal{M}'$  and exponential time in the size of priority functions set (de Alfaro and Henzinger 2000). Line 6 in the algorithm corresponds to checking the realisability problem for a *qualitative parity logic* formula containing conjunctions of almost-sure atoms over the MDP  $\mathcal{K}^{-L}$  resulting from  $\mathcal{G}_{\text{PAR}}^{-L}$  when all players are as one, which can be solved in polynomial time (Berthon, Guha, and Raskin 2020). The formula expresses that the objective  $\alpha_\varphi$  (representing  $\varphi$ ) is satisfied with probability 1, and each winning player  $i$  (that cannot be punished in  $\mathcal{M}^{-L}$ ) achieves its goal with probability 1. The overall complexity of the algorithm is thus in 2EXPTIME.

Now that we have a procedure to solve E-NASH, we can adapt it to solve A-NASH. The adaptation is straightforward, and goes as follows. First, when building  $\mathcal{G}_{\text{PAR}}$  from  $\mathcal{G}$ , instead of DPW  $\mathcal{A}_\varphi$ , we use  $\mathcal{A}_{\neg\varphi}$ , i.e., a DPW built from  $\neg\varphi$ . Thus, in Algorithm 2, the formula in line 6 is replaced by  $\text{NZ}(\alpha_{\neg\varphi}) \wedge \bigwedge_{i \in W} \text{AS}(\alpha_i)$  and the positive and negative answers in lines 7 and 8 respectively are swapped. Thus we also have a 2EXPTIME algorithm for A-NASH.

For hardness, we reduce from qualitative LTL model checking over MDPs. Given an MDP  $\mathcal{K}$  with a labelling function  $\lambda$  and formula  $\varphi$ , then solving the E-NASH problem with input given by the one-player game  $\mathcal{G} = (\mathcal{K}, (\top), \lambda)$  and formula is equivalent to LTL model checking over the MDP. This fact, and the duality between A-NASH and E-NASH concludes the proof.  $\square$

We now turn our attention to the cooperative setting, in which equilibria are instead characterised by the *core*.

## 4 Cooperative Rational Verification

Nash equilibrium is a non-cooperative solution concept: it assumes that players must act in isolation, without the possibility of forming binding agreements to cooperate. In many

---

### Algorithm 2 E-NASH

---

```

input:  $\mathcal{G}, \varphi$ 
1: build  $\mathcal{G}_{\text{PAR}}$  from  $\mathcal{G}$ 
2: for  $W \subseteq N$  do
3:   for  $j \in L = N \setminus W$  do
4:     compute  $\text{Pun}_j(\mathcal{G}_{\text{PAR}})$ 
5:   build  $\mathcal{G}_{\text{PAR}}^{-L}$  and obtain  $\mathcal{K}^{-L}$ 
6:   if  $\mathcal{K}^{-L} \models \text{AS}(\alpha_\varphi) \wedge \bigwedge_{i \in W} \text{AS}(\alpha_i)$  then
7:     return “yes”
8: return “no”

```

---

settings, however, binding agreements are possible, and for these it is appropriate to consider *cooperative* solution concepts, of which the *core* is the most prominent. While Nash equilibrium considers strategy profiles that are stable against individual deviations, the core considers possible beneficial deviations by groups of players (*coalitions*).

We follow the definition of *core* from (Gutierrez, Kraus, and Wooldridge 2019). We first define the notion of a *deviation* and a *beneficial deviation*. A deviation is a joint strategy  $\vec{\sigma}'_A$  for the coalition  $A \subseteq N$ , with  $A \neq \emptyset$ . For a strategy profile  $\vec{\sigma}$ , we say  $\vec{\sigma}'_A$  is a beneficial deviation from  $\vec{\sigma}$  if  $A \subseteq L(\vec{\sigma})$  and for all  $\vec{\sigma}'_{-A}$ , we have  $A \subseteq W((\vec{\sigma}'_A, \vec{\sigma}'_{-A}))$ . The core of a game  $\mathcal{G}$ , denoted  $\text{core}(\mathcal{G})$ , is then defined to be the set of strategy profiles that admit no beneficial deviation.

Given the above definitions, we can introduce the key decision problems relating to rational verification and the core.

#### E-CORE:

*Given:* Game  $\mathcal{G}$ , LTL formula  $\varphi$ .

*Question:* Is it the case that

$$\exists \vec{\sigma} \in \text{core}(\mathcal{G}). \vec{\sigma} \models \text{AS}(\varphi)?$$

#### A-CORE:

*Given:* Game  $\mathcal{G}$ , LTL formula  $\varphi$ .

*Question:* Is it the case that

$$\forall \vec{\sigma} \in \text{core}(\mathcal{G}). \vec{\sigma} \models \text{AS}(\varphi)?$$

**Example 2.** Recall the game in Figure 1. As we saw in Example 1, the set of strategy profiles  $\vec{\sigma}$  in which player S/R chooses action  $i_S/i_R$  with probability one in any  $s \in \{s_0, \dots, s_{n-1}\}$  is a Nash equilibrium, because unilateral deviations cannot improve the situation. However, this strategy profile is not in the core, because there is a cooperative beneficial deviation to the strategy  $\vec{\sigma}$  in which player S/R chooses action  $c_S/c_R$  with probability strictly greater than zero in every state  $s \in \{s_0, \dots, s_{n-1}\}$ . This means that, while the A-NASH query with the property  $\varphi \equiv \mathbf{F}\psi$  returns “no”, the A-CORE query with the same property returns “yes”, since every strategy profile in the core satisfies  $\varphi$  with probability one – in general they are always Pareto-optimal.

Alongside E-CORE and A-CORE, we will also often be interested in the question of whether a particular alternative strategy  $\vec{\sigma}'_A$  represents a beneficial deviation from  $\vec{\sigma}$  for a coalition  $A$  of players. This question, along with a version of MEMBERSHIP for cooperative games, forms the final two decision problems we investigate in this work.

**CORE MEMBERSHIP:**

Given: Game  $\mathcal{G}$ , strategy profile  $\vec{\sigma}$ .

Question: Is it the case that  $\vec{\sigma} \in \text{core}(\mathcal{G})$ ?

**BENEFICIAL DEVIATION:**

Given: Game  $\mathcal{G}$ , strategy profile  $\vec{\sigma}$ , deviation  $\vec{\sigma}'_A$ .

Question: Is  $\vec{\sigma}'_A$  a beneficial deviation from  $\vec{\sigma}$  in  $\mathcal{G}$ ?

As noted above, the core can be viewed as the set of strategy profiles from which no *coalition* (as opposed to *player* in the case of NE) has a beneficial deviation. Our complexity results for this cooperative solution concept follow a similar high-level line of reasoning as taken in previous work in the non-stochastic setting (Gutierrez, Kraus, and Wooldridge 2019). However, as it will be seen next, the “inner workings” to obtain the main complexity results rely on very different techniques needed to be able to account for the various probabilistic features in the game.

**4.1 E-CORE and A-CORE**

We begin by noting that the ability of a coalition  $A$  to achieve an LTL goal  $\psi$  can be interpreted as its possession of a winning strategy in the two-player coalition game  $\mathcal{G}^{A,\psi}$ , as defined in the previous section. We say that such a game is *winnable* if player 1 has a strategy for achieving  $\psi$ . Using these concepts, we restate the following result, with some adaptations for our stochastic setting.

**Lemma 3** ((Gutierrez, Kraus, and Wooldridge 2019)). *Let  $\mathcal{G} = (\mathcal{M}, \lambda, (\gamma_i)_{i \in N})$  be a CSG whose underlying arena is  $\mathcal{M} = (N, (Ac_i)_{i \in N}, St, s^0, tr)$  and let  $\varphi$  be an LTL formula. Then  $\mathcal{G}$  and  $\varphi$  satisfy E-CORE if and only if there exists  $W \subseteq N$  such that*

- *There exists some  $\vec{\sigma}$  such that  $\vec{\sigma} \models \chi_W$*
- *For all  $L \subseteq N \setminus W$ ,  $\mathcal{G}^{L,\psi_L}$  is not winnable*

where  $\chi_W = \text{AS}(\varphi) \wedge \bigwedge_{i \in W} \text{AS}(\gamma_i) \wedge \bigwedge_{i \in N \setminus W} \neg \text{AS}(\gamma_i)$  and  $\psi_L = \bigwedge_{i \in L} \text{AS}(\gamma_i)$ .

This result leads us to the procedure shown in Algorithm 3 for determining whether some game  $\mathcal{G}$  and LTL formula  $\varphi$  satisfy E-CORE. Moreover, due to the duality between AS and NZ winning conditions we may express the A-CORE problem for a game  $\mathcal{G}$  and formula  $\varphi$  as the negation of the E-CORE problem for  $\mathcal{G}$  where the conjunct  $\text{AS}(\varphi)$  is replaced by  $\neg \text{AS}(\varphi) \equiv \text{NZ}(\neg \varphi)$ . We thus have the following complexity results.

**Theorem 3.** *E-CORE and A-CORE are 2EXPTIME-complete.*

*Proof.* The loop in line 1 and the check in line 4 (for each  $A \subseteq N$ ) in Algorithm 3 are executed  $2^{|N|}$  times each. Next, observe that we may also write  $\chi_W$  as  $\text{AS}(\varphi) \wedge \bigwedge_{i \in W} \text{AS}(\gamma_i) \wedge \bigwedge_{i \in N \setminus W} \text{NZ}(\neg \gamma_i)$ , a conjunction of AS and NZ conditions. Thus, by expressing  $\varphi, \gamma_i$  for  $i \in W$ , and  $\neg \gamma_i$  for  $i \in N \setminus W$  as DPWs and constructing the game  $\mathcal{G}_{\text{PAR}}$ , we may perform this model checking problem in time polynomial in the size of  $\mathcal{G}_{\text{PAR}}$  using qualitative parity logic (Berthon, Guha, and Raskin 2020). As a DPW formed from an LTL formula  $\psi$  has states and labels doubly and singly exponential in the size of  $\psi$  respectively (Piterman 2006),

**Algorithm 3** E-CORE

---

**input:**  $\mathcal{G}, \varphi$   
1: **for**  $A \subseteq N$  **do**  
2:     check if  $\exists \vec{\sigma}$  s.t.  $\vec{\sigma} \models \chi_A$ , written  $W_A$   
3:     check that  $\mathcal{G}^{A,\psi_A}$  is not winnable, written  $L_A$   
4:     **if**  $\exists A \subseteq N$  s.t.  $W_A$  and  $L_B \forall B \subseteq N \setminus A$  **then**  
5:         **return** “yes”  
6:     **return** “no”

---

then we see that  $\mathcal{G}_{\text{PAR}}$  has size doubly exponential in the original input and hence that the model checking step in the algorithm can be solved in 2EXPTIME.

For line 3 we begin by noting that  $\vec{\sigma} \models \bigwedge_{i \in A} \text{AS}(\gamma_i)$  if and only if  $\vec{\sigma} \models \text{AS}(\bigwedge_{i \in A} \gamma_i)$ . Let us define  $\psi_A^\wedge = \text{AS}(\bigwedge_{i \in A} \gamma_i)$ . Thus, we instead form the two-player parity coalition game  $\mathcal{G}_{\text{PAR}}^{A,\psi_A^\wedge}$ . As  $|\psi_A^\wedge|$  is linear in  $|\gamma_1|, \dots, |\gamma_n|$ , then the number of states in  $\mathcal{G}_{\text{PAR}}^{A,\psi_A^\wedge}$  is doubly exponential in the size of the original input, and the number of pairs in the parity accepting condition of  $\mathcal{G}_{\text{PAR}}^{A,\psi_A^\wedge}$  is singly exponential in the size of the original input. Whether player 1 has a winning strategy in this product game can be checked in time polynomial in the former and singly exponential in the latter for AS-winning conditions (de Alfaro and Henzinger 2000), meaning this step remains in 2EXPTIME as well.

Thus, Algorithm 3 above can be seen to run in 2EXPTIME, providing an upper bound for E-CORE. To see that this bound is tight, note that we may reduce (qualitative) model checking of LTL over MDPs to E-CORE. Given an MDP  $\mathcal{K}$  with a labelling function  $\lambda$  and formula  $\varphi$ , we input the one-player game  $\mathcal{G} = (\mathcal{K}, (\top), \lambda)$  and  $\varphi$  to Algorithm 3, which returns “yes” if and only if there is a strategy in  $\mathcal{K}$  satisfying  $\varphi$  almost-surely. As qualitative LTL model checking over MDPs is in 2EXPTIME (Courcoubetis and Yannakakis 1995), and our construction is polynomial in the size of the original input, this concludes the proof.  $\square$

**4.2 CORE MEMBERSHIP**

For the problem of CORE MEMBERSHIP and BENEFICIAL DEVIATION we assume, as in the non-cooperative setting, that the given strategies are all finite memory. By taking a product of the finite state transducers representing  $\vec{\sigma}$  with the game  $\mathcal{G}$  then we may check CORE MEMBERSHIP by first checking for every player  $i$  whether  $\vec{\sigma} \models \text{AS}(\gamma_i)$ , and then checking whether any subset  $L \subseteq N \setminus W$  can deviate to achieve the formula  $\psi_L = \bigwedge_{i \in L} \text{AS}(\gamma_i)$ . This procedure is shown in full in Algorithm 4 and gives rise to the following complexity result.

**Theorem 4.** *CORE MEMBERSHIP is 2EXPTIME-complete.*

*Proof.* The first step (line 4) of Algorithm 4 is equivalent to performing qualitative LTL model checking on the MC  $\mathcal{C}_{\vec{\sigma}}$ , which can be done in PSPACE (Courcoubetis and Yannakakis 1995). The second step (line 4) can be done in 2EXPTIME, as noted above, and thus we have that CORE MEMBERSHIP is also in 2EXPTIME. As a lower bound we



---

**Algorithm 4** CORE MEMBERSHIP
 

---

**input:**  $\mathcal{G}, \vec{\sigma}$   
 1: **for**  $i \in N$  **do**  
 2:     check if  $\vec{\sigma} \models \text{AS}(\gamma_i)$ , written  $W_i$   
 3: **for**  $L \subseteq N \setminus \{i : W_i = \top\}$  **do**  
 4:     **if**  $\vec{\mathcal{G}}^{L, \psi_L}$  is winnable **then**  
 5:         **return** “no”  
 6: **return** “yes”

---

note that the problem of CORE MEMBERSHIP is the same as the problem of MEMBERSHIP when  $|N| = 1$ , and so we may use exactly the same reduction given in the proof of Theorem 1 with the construction illustrated in Figure 2.  $\square$

### 4.3 BENEFICIAL DEVIATION

We now study BENEFICIAL DEVIATION, solved using Algorithm 5, and note that although the complexities for the other problems in cooperative rational verification are the same as in the deterministic setting, BENEFICIAL DEVIATION is only in PSPACE for deterministic games, and hence the stochastic setting is significantly more difficult.

**Theorem 5.** BENEFICIAL DEVIATION is 2EXPTIME-complete.

*Proof.* We begin, in Algorithm 5, by checking whether  $\vec{\sigma} \models \bigwedge_{i \in A} \neg \text{AS}(\gamma_i)$ , or equivalently whether  $\vec{\sigma} \models \text{NZ}(\neg \gamma_i)$  for each player  $i \in A$ , as otherwise there is no beneficial deviation for the coalition  $A$  and we are done (line 3). This is again a simple qualitative LTL model checking problem on  $\mathcal{C}_{\vec{\sigma}}$ , which is in PSPACE, as remarked in the previous proof. The second condition we check is whether, when the coalition  $A$  is instead playing  $\vec{\sigma}'_A$ , there exists a way for the remaining players  $N \setminus A$  to achieve  $\bigvee_{i \in A} \neg \text{AS}(\gamma_i)$ . This can be done by taking the product of the original game  $\mathcal{G}$  and the finite state transducers representing  $\vec{\sigma}'_A$ , and then model checking the resulting MDP  $\mathcal{K}_{-A}$ , in which  $N \setminus A$  is viewed as a single agent, with respect to  $\text{NZ}(\neg \gamma_i)$  (line 5), which is in 2EXPTIME and hence so is BENEFICIAL DEVIATION.

To see that this bound is tight note that we can reduce qualitative LTL model checking on MDPs to BENEFICIAL DEVIATION. Given an MDP  $\mathcal{K}$  and LTL formula  $\varphi$ , let  $\mathcal{C}^\varphi$  be a minimal (deterministic) MC satisfying  $\varphi$  and similarly for  $\mathcal{C}^{\neg\varphi}$ . We then form the CSGA  $\mathcal{M}$  shown in Figure 3, where player 1 has control over all the actions in  $\mathcal{K}$ , and the resulting game  $\mathcal{G} = (\mathcal{M}, (\mathbf{X}\varphi, \neg\mathbf{X}\varphi), \lambda)$  where  $\lambda$  is the combination of the labelling functions for  $\mathcal{K}$ ,  $\mathcal{C}^\varphi$ , and  $\mathcal{C}^{\neg\varphi}$  with  $\lambda(s_0) = \emptyset$ . Let  $\vec{\sigma}$  be such that  $\sigma_1(s_0) = a$  and  $\sigma_2(s_0) = b$ , and for  $A = \{2\}$  let  $\vec{\sigma}'_A$  be the same as  $\vec{\sigma}$  except for having  $\sigma_2(s_0) = \bar{b}$ .

It can be seen immediately both that  $\mathcal{G}$ ,  $\vec{\sigma}$ , and  $\vec{\sigma}'_A$  are of size polynomial in the original input  $\mathcal{K}$  and  $\varphi$ , and that  $\vec{\sigma}'_A$  is a beneficial deviation from  $\vec{\sigma}$  if and only if it is not the case that there exists a strategy in  $\mathcal{K}$  satisfying  $\text{AS}(\varphi)$ ; if there was then player 1 could switch to such a strategy and play  $\bar{a}$  in  $s_0$  in order to ensure that the coalition  $\{2\}$  does not achieve their goal.  $\square$

---

**Algorithm 5** BENEFICIAL DEVIATION
 

---

**input:**  $\mathcal{G}, \vec{\sigma}, \vec{\sigma}'_A$   
 1: build  $\mathcal{K}_{-A}$  from  $\mathcal{G}$  and  $\vec{\sigma}'_A$   
 2: **for**  $i \in A$  **do**  
 3:     **if**  $\vec{\sigma} \models \text{AS}(\gamma_i)$  **then**  
 4:         **return** “no”  
 5:     **if**  $\mathcal{K}_{-A} \models \text{NZ}(\neg \gamma_i)$  **then**  
 6:         **return** “no”  
 7: **return** “yes”

---

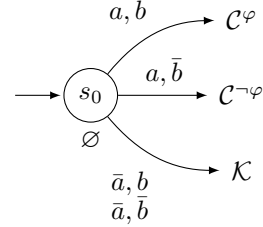


Figure 3: The CSGA for our reduction from BENEFICIAL DEVIATION to qualitative LTL model checking in MDPs. Edges without probability labels mean they have probability 1.

## 5 Discussion and Related Work

Our results account for cooperative and non-cooperative settings, and required the development of new techniques with respect to previous work. In particular, most verification techniques for deterministic systems, many of which are used in the context of rational verification, no longer apply when probabilistic behaviour is allowed. We then conclude with a more detailed comparison and analysis against previous work on rational verification and probabilistic systems.

**Special Cases** Our results pertain to general-sum multiplayer games with perfect information and memoryful strategies. A few observations can be made as a result. The proofs for hardness of most of our results show that the problems studied here remain in 2EXPTIME even in the single-player case. The same is true if we consider zero-sum games, which are frequently used in formal verification in the two-player setting. This shows, in turn, that the 2EXPTIME results hold regardless of the number of players (as long as it is more than one) or how antagonistic they are.

A less obvious situation is when strategies are restricted or when the game has multiple players, but each control their own set of states – a multiplayer turn-based game. Concrete results in these cases are yet to be obtained (for instance, for memoryless or finite-memory strategies), and therefore have been left as directions for future work.

Another special case is where players possess simpler temporal logic goals. It is known that two-player games with goals expressed using various kinds of LTL fragments are 2EXPTIME-complete even for deterministic systems (Alur, Torre, and Madhusudan 2003). However, if only one player is allowed in the game, the problem can have a significantly lower complexity (NP or EXPTIME) in case of qualitative probabilistic model checking (Kini and Viswanathan 2017).

A much different complexity landscape is found when the quantitative probabilistic setting is considered, with similar decision problems having much higher complexity in the quantitative setting, and requiring, yet again, a different toolset of verification techniques.

**On Rational Verification** Most relevant related work on rational verification has focused on deterministic systems, leaving all reasoning about stochastic systems largely overlooked. One important question, not discussed in the present paper, is the problem of whether a game has at least one equilibrium: the NON-EMPTINESS problem in rational verification. This problem always has a positive answer in the cooperative case – see (Gutierrez, Kraus, and Wooldridge 2019), the argument therein also extends to the stochastic setting considered in the current work – and in the deterministic, non-cooperative case is solvable in 2EXPTIME when considering arbitrary LTL goals and strategies. In the stochastic, non-cooperative setting, however, it is an open problem and known results on game theory and games with probabilistic behaviour do not seem to provide an answer.

Two important types of games in the rational verification literature for multi-agent systems are the case where players control Boolean variables, as in an iterated version of a Boolean game (Gutierrez, Harrenstein, and Wooldridge 2015b), and the case where systems are succinctly represented using a guarded command-like language for multi-agent system specifications (Alur and Henzinger 1999). While all of these problems are also in 2EXPTIME in the non-stochastic setting, they require considerably different proof techniques for their solutions, typically resorting to the use of logics for strategic reasoning (Mogavero et al. 2014), automata (Fisman, Kupferman, and Lustig 2010), or reductions to other game representations (Gutierrez et al. 2020; Gutierrez et al. 2018; Bouyer et al. 2015). In the probabilistic setting considered here, several new constructions and proof techniques were required, including, *e.g.*, the need for infinite-memory strategies to achieve optimal behaviour.

**Verification of Probabilistic Systems** Probabilistic systems have been studied extensively, but mostly without a focus on rational behaviour as considered in this work. Instead, the traditional verification approaches to probabilistic systems have considered zero-sum games,  $\omega$ -regular goals, and many types of winning, of which almost-sure is just one of many possibilities. A comprehensive survey of main results in this area can be found in (Chatterjee and Henzinger 2012). Of the many studies on the analysis and verification of probabilistic systems, the work in (Kwiatkowska et al. 2020a; Kwiatkowska et al. 2019) is the closest to that presented here. In common with (Kwiatkowska et al. 2020a; Kwiatkowska et al. 2019), we use CSGs as the underlying model and focus on properties satisfied in equilibrium. However, several aspects of our work are different. On one hand, we model players’ preferences using LTL formulae, allow infinite-horizon plays unrestrictedly, provide optimal complexity results for key decision problems, and look at different game-theoretic solution concepts; in fact, no other paper investigates the core for CSGs. On the other hand, we do not consider probabilistic reasoning in the quantitative

setting, and do not have a practical implementation.

**Practical Implementations** So far, only PRISM-games can be used to verify the satisfaction of properties in equilibrium in CSGs, and until recently, only games without concurrency were supported (Kwiatkowska et al. 2020b). The current PRISM-games implementation supports concurrency, but verification is restricted to games with a finite horizon, while the procedures we have developed consider plays with an infinite horizon and strategies having access to infinite memory. In addition, PRISM-games considers a non-cooperative solution concept different from Nash equilibrium and does not support cooperative solution concepts.

The closest implementation to the work in this paper, but in a non-probabilistic setting, is EVE (Gutierrez et al. 2018; Gutierrez et al. 2020), one of the most efficient software verification tools for the analysis and verification of temporal logic properties of multi-agent systems. An avenue for future work on the practical side is to extend the functionalities of EVE to account for the more complex probabilistic framework we have studied here, both for cooperative and for non-cooperative games. MCMAS (Lomuscio, Qu, and Raimondi 2017), a verification tool for multi-agent systems, also provides some support to model check logics for strategic reasoning, some of which can express both Nash equilibrium and the core; however, at the time of writing, an implementation that can account for the kind of probabilistic systems we have considered here is not available.

**Future Work** As pointed out before, a few problems seem to lead to interesting avenues for future research. On the practical side, just discussed, the immediate work to do would be to implement the algorithms herein proposed – an powerful way to extend the current capabilities of, say, EVE’s verification engine which at present does not support any kind of probabilistic reasoning. On the other hand, on the theory side, we would like to understand better two specific problems: firstly, NON-EMPTINESS in the stochastic, non-cooperative setting, and secondly, whether our results can be extended to the probabilistic quantitative setting.

## Acknowledgments

Lewis Hammond acknowledges the support of an EP-SRC Doctoral Training Partnership studentship (Reference: 2218880). Anthony Lin and Muhammad Najib acknowledge the support of ERC Starting Grant 759969 (AV-SMP) and Max-Planck Fellowship. Michael Wooldridge was supported by JP Morgan and the Alan Turing Institute.

## References

- Abdulla, P. A.; Henda, N. B.; de Alfaro, L.; Mayr, R.; and Sandberg, S. 2008. Stochastic games with lossy channels. In Amadio, R. M., ed., *FoSSaCS*, volume 4962 of *Lecture Notes in Computer Science*, 35–49. Springer.
- Alur, R., and Henzinger, T. 1999. Reactive modules. *Formal Methods in System Design* 15(1):7–48.
- Alur, R.; Torre, S. L.; and Madhusudan, P. 2003. Playing games with boxes and diamonds. In Amadio, R. M., and

- Lugiez, D., eds., *CONCUR*, volume 2761 of *LNCS*, 127–141. Springer.
- Baier, C., and Katoen, J.-P. 2008. *Principles of Model Checking (Representation and Mind Series)*. The MIT Press.
- Baier, C.; Bertrand, N.; and Schnoebelen, P. 2007. Verifying nondeterministic probabilistic channel systems against  $\omega$ -regular linear-time properties. *ACM Trans. Comput. Log.* 9(1):5.
- Berthon, R.; Guha, S.; and Raskin, J.-F. 2020. Mixing probabilistic and non-probabilistic objectives in markov decision processes. In *Proceedings of the 35th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '20*, 195–208. New York, NY, USA: Association for Computing Machinery.
- Bouyer, P.; Brenguier, R.; Markey, N.; and Ummels, M. 2015. Pure nash equilibria in concurrent deterministic games. *Log. Methods Comput. Sci.* 11(2).
- Chatterjee, K., and Henzinger, T. A. 2012. A survey of stochastic  $\omega$ -regular games. *Journal of Computer and System Sciences* 78(2):394 – 413. Games in Verification.
- Courcoubetis, C., and Yannakakis, M. 1995. The complexity of probabilistic verification. *J. ACM* 42(4):857–907.
- de Alfaro, L., and Henzinger, T. A. 2000. Concurrent omega-regular games. In *Proceedings of the 15th Annual IEEE Symposium on Logic in Computer Science, LICS '00*, 141. USA: IEEE Computer Society.
- Fisman, D.; Kupferman, O.; and Lustig, Y. 2010. Rational synthesis. In Esparza, J., and Majumdar, R., eds., *TACAS*, volume 6015 of *LNCS*, 190–204. Springer.
- Gutierrez, J.; Najib, M.; Perelli, G.; and Wooldridge, M. J. 2018. EVE: A tool for temporal equilibrium analysis. In Lahiri, S. K., and Wang, C., eds., *ATVA*, volume 11138 of *LNCS*, 551–557. Springer.
- Gutierrez, J.; Najib, M.; Perelli, G.; and Wooldridge, M. J. 2020. Automated temporal equilibrium analysis: Verification and synthesis of multi-player games. *Artif. Intell.* 287:1–70.
- Gutierrez, J.; Harrenstein, P.; and Wooldridge, M. 2015a. Expressiveness and Complexity Results for Strategic Reasoning. In *26th International Conference on Concurrency Theory (CONCUR 2015)*.
- Gutierrez, J.; Harrenstein, P.; and Wooldridge, M. 2015b. Iterated boolean games. *Inf. Comput.* 242(C):53–79.
- Gutierrez, J.; Harrenstein, P.; and Wooldridge, M. J. 2017. From model checking to equilibrium checking: Reactive modules for rational verification. *Artif. Intell.* 248:123–157.
- Gutierrez, J.; Kraus, S.; and Wooldridge, M. 2019. Cooperative concurrent games. *AAMAS '19*, 1198–1206. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.
- Kini, D., and Viswanathan, M. 2017. Complexity of model checking mdps against LTL specifications. In Lokam, S. V., and Ramanujam, R., eds., *FSTTCS*, volume 93 of *LIPICs*, 35:1–35:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.
- Kwiatkowska, M.; Norman, G.; Parker, D.; and Santos, G. 2019. Equilibria-based probabilistic model checking for concurrent stochastic games. In ter Beek, M. H.; McIver, A.; and Oliveira, J. N., eds., *FM*, volume 11800 of *LNCS*, 298–315. Springer.
- Kwiatkowska, M.; Norman, G.; Parker, D.; and Santos, G. 2020a. Multi-player equilibria verification for concurrent stochastic games. In Gribaudo, M.; Jansen, D. N.; and Remke, A., eds., *QEST*, volume 12289 of *LNCS*, 74–95. Springer.
- Kwiatkowska, M.; Norman, G.; Parker, D.; and Santos, G. 2020b. Prism-games 3.0: Stochastic game verification with concurrency, equilibria and time. In Lahiri, S. K., and Wang, C., eds., *CAV*, volume 12225 of *LNCS*, 475–487. Springer.
- Lehmann, D., and Rabin, M. 1981. On the advantage of free choice: A symmetric and fully distributed solution to the dining philosophers problem (extended abstract). In *POPL*, 133–138.
- Lengál, O.; Lin, A. W.; Majumdar, R.; and Rümmer, P. 2017. Fair termination for parameterized probabilistic concurrent systems. In *TACAS*, 499–517.
- Lin, A. W., and Rümmer, P. 2016. Liveness of randomised parameterised systems under arbitrary schedulers. In *CAV*, 112–133.
- Lomuscio, A.; Qu, H.; and Raimondi, F. 2017. MCMAS: an open-source model checker for the verification of multi-agent systems. *Int. J. Softw. Tools Technol. Transf.* 19(1):9–30.
- Mogavero, F.; Murano, A.; Perelli, G.; and Vardi, M. Y. 2014. Reasoning about strategies: On the model-checking problem. *ACM Trans. Comput. Log.* 15(4):34:1–34:47.
- Norman, G. 2004. Analysing randomized distributed algorithms. In *Validation of Stochastic Systems - A Guide to Current Research*, 384–418.
- Osborne, M. J., and Rubinstein, A. 1994. *A Course in Game Theory*. MIT Press.
- Piterman, N. 2006. From nondeterministic buchi and streett automata to deterministic parity automata. In *21st Annual IEEE Symposium on Logic in Computer Science (LICS'06)*, 255–264.
- Pnueli, A. 1977. The temporal logic of programs. In *FOCS*, 46–57. IEEE.
- Vardi, M. Y. 1985. Automatic verification of probabilistic concurrent finite state programs. In *26th Annual Symposium on Foundations of Computer Science (sfcs 1985)*, 327–338.
- Wooldridge, M. J.; Gutierrez, J.; Harrenstein, P.; Marchioni, E.; Perelli, G.; and Toumi, A. 2016. Rational verification: From model checking to equilibrium checking. In Schuurmans, D., and Wellman, M. P., eds., *AAAI*, 4184–4191. AAAI Press.