



HAL
open science

Stochastic Games with Parity Mean-payoff Objective

Krishnendu Chatterjee, Laurent Doyen, Hugo Gimbert, Youssouf Oualhadj

► **To cite this version:**

Krishnendu Chatterjee, Laurent Doyen, Hugo Gimbert, Youssouf Oualhadj. Stochastic Games with Parity Mean-payoff Objective. 2012. hal-00766251

HAL Id: hal-00766251

<https://hal.science/hal-00766251>

Submitted on 17 Dec 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stochastic Games with Parity Mean-payoff Objective

Krichnendu Chatterjee* Laurent Doyen† Hugo Gimbert‡
Youssouf Oualhadj§

Abstract

In this paper, we compute value of two-player games with perfect information equipped with the $\text{Par} \wedge \text{Avg}_{>0}$ objectives. Moreover we show that even though the optimal strategies may require infinite memory, there exists an NP algorithm that computes the almost-sure region.

1 Introduction

Stochastic games with perfect information generalize Markov decision processes in the sense that the model is equipped with an second controller usually called Min whose objective is to minimize the probability that max satisfies her objective. In this model the two-player play in turn and the state space is partitioned into Max's states and Min's states as opposed to concurrent games where the players choose their actions simultaneously.

These games are very useful in modeling problems and providing solutions for verification of open reactive systems even though they are less tractable than Markov decision processes. For instance computing the value of a reachability game is a problem that lies in $\text{NP} \cap \text{CoNP}$ [2] as opposed to the polynomial time algorithm for Markov decision processes.

Our main goal in this paper is to study stochastic games equipped with combination of parity and positive-average objectives. This objective were first studied in the case of non-stochastic games [1].

Contribution and result In this paper we give characterization of the almost-sure region for Max when the objective is $\text{Par} \wedge \text{Avg}_{>0}$, we also give an NP algorithm that computes this region together with an almost-sure strategy even though our almost-sure strategy may require infinite memory. show that for stochastic games the almost-sure region for Max for the objective $\text{Par} \wedge \text{Avg}_{>0}$ can be computed in NP.

*IST Austria (Institute of Science and Technology Austria)

†LSV, ENS Cachan & CNRS, France

‡LaBRI, CNRS, France

§LIF, Aix Marseille université, France

Outline of the document

- In Section 2 we introduce stochastic games and basic tools to study this model.
- In Section 3 we show that deciding whether Max has an almost-sure state for the objective $\text{Par} \wedge \text{Avg}_{>0}$ in a stochastic game is in NP.
- In Section 4 we give an algorithm that computes the value of each state for the objective $\text{Par} \wedge \text{Avg}_{>0}$ in a stochastic game.

2 Two-player Stochastic Games with Perfect Information

Two-player Stochastic Games with Perfect Information are similar to Markov decision processes except there are two kinds of states: states controlled by player Max whose goal is to maximize the probability that some objective is achieved, and states controlled by player Min who has the opposite goal and tries to minimize this probability.

Definition 1 (Stochastic game with perfect information). *A stochastic game with perfect information is a tuple $\mathcal{A} = (S, (S_1, S_2), A, p)$ where:*

- S is a finite set of states,
- (S_1, S_2) is a partition of S ,
- A is a set of actions,
- $p : S \times A \rightarrow \Delta(S)$ is a transition function.

In the sequel we refer to two-player stochastic game with perfect information by stochastic game unless it is not clear by the context.

Definition 2 (Strategies). *A strategy for Max is a function $\sigma : (SA)^*S_1 \rightarrow \Delta(A)$ and a strategy for Min is a function $\tau : (SA)^*S_2 \rightarrow \Delta(A)$.*

Once a couple of strategies chosen (σ, τ) and an initial state s fixed, we associate the probability measure $\mathbb{P}_s^{\sigma, \tau}$ over $s(AS)^\omega$ as the only measure over S^ω such that:

$$\begin{aligned} \mathbb{P}_s^{\sigma, \tau}(S_0 = s) &= 1 \text{ ,} \\ \mathbb{P}_s^{\sigma, \tau}(S_{n+1} = s \mid S_n = s_n \wedge A_{n+1} = a_{n+1}) &= p(s_n, a_{n+1})(s) \text{ ,} \\ \mathbb{P}_s^{\sigma, \tau}(A_{n+1} = a \mid S_0 A_1 S_1 \cdots S_n = s_0 a_1 \cdots s_n) &= \begin{cases} \sigma(s_0 a_1 \cdots s_n) & \text{if } s_n \in S_1 \\ \tau(s_0 a_1 \cdots s_n) & \text{if } s_n \in S_2 \end{cases} \end{aligned}$$

Definition 3. *Let s be a state, (σ, τ) a couple of strategies, and Φ and objective. The value of s with respect to (σ, τ) for Φ is:*

$$\text{Val}(s)_{\sigma, \tau} = \mathbb{P}_s^{\sigma, \tau}(\Phi) \text{ .}$$

Definition 4 (Superior value). *Let s be a state and Φ be an objective, the superior value of s for Φ is:*

$$\overline{\text{Val}}(s) = \inf_{\tau} \sup_{\sigma} \mathbb{P}_s^{\sigma, \tau}(\Phi) .$$

Dually, one defines also the so called *inferior value* of a state with the intuition that now Max chooses her strategy first and let Min defines the best possible answer.

Definition 5 (Inferior value). *Let s be a state and Φ be an objective, the inferior value of s for Φ is:*

$$\underline{\text{Val}}(s) = \sup_{\sigma} \inf_{\tau} \mathbb{P}_s^{\sigma, \tau}(\Phi) .$$

The following equation always holds.

$$\forall s \in S, \underline{\text{Val}}(s) \leq \overline{\text{Val}}(s) . \quad (1)$$

Equation (1) follows the natural intuition; it is easier to win if one knows the strategy of his opponent. A legitimate question raises. When does these two quantities coincide? The answer follows from Martin's second determinacy theorem [6] extended to stochastic games by Maitra and Sudderth [4], which shows that for any Borel objective both values coincide.

Definition 6 (Determinacy). *Let s be a state and Φ be a objective, then Φ is determined (for finite stochastic games with perfect information) if and only if for every stochastic game with perfect information and finitely many states and actions objective Φ :*

$$\underline{\text{Val}}(s) = \overline{\text{Val}}(s) .$$

In this case we denote the value of a state $\text{Val}(s)$.

Theorem 7 (Borel Determinacy [5, 4]). *Every Borel objective is determined for finite stochastic games with perfect information.*

This determinacy result shows that for Borel objectives, there always exist ε -optimal strategies for both players.

Definition 8 (ε -optimal strategies). *Let $\varepsilon > 0$. A strategy $\sigma^\#$ for player 1 is ε -optimal if*

$$\forall s \in S, \forall \tau, \mathbb{P}_s^{\sigma^\#, \tau}(\Phi) \geq \text{Val}(s) .$$

For player 2 the definition is symmetric. A 0-optimal strategy is simply called optimal.

While ε -strategies are guaranteed to exist in determined games, this is not the case for optimal strategies. However, provided the objective is tail, this existence is guaranteed:

Theorem 9 (Existence of optimal strategies [3]). *In every stochastic game with perfect information equipped with a tail Borel objective, both players have optimal strategies.*

Note that optimal strategies can be characterized as follows:

Definition 10 (Optimal strategies). Let $(\sigma^\sharp, \tau^\sharp)$ be a couple of strategies and Φ be an objective, $(\sigma^\sharp, \tau^\sharp)$ is an pair of optimal strategies if for every pair of strategies (σ, τ)

$$\forall s \in S, \mathbb{P}_s^{\sigma, \tau^\sharp}(\Phi) \leq \mathbb{P}_s^{\sigma^\sharp, \tau^\sharp}(\Phi) \leq \mathbb{P}_s^{\sigma^\sharp, \tau}(\Phi) .$$

If this property holds in a game then the game is determined and:

$$\forall s \in S, \text{Val}(s) = \mathbb{P}_s^{\sigma^\sharp, \tau^\sharp}(\Phi) .$$

Definition 11 (Almost-sure and positive winning strategies). We say that Max wins almost-surely (resp. positively) from a state s if she has a strategy σ such that for every strategy τ $\mathbb{P}_s^{\sigma, \tau}(\Phi) = 1$ (resp. $\mathbb{P}_s^{\sigma, \tau}(\Phi) > 0$).

We will use the following result about qualitative determinacy.

Theorem 12 (Qualitative determinacy [3]). In any stochastic game equipped with a tail objective, each state is either almost-sure for Max or positive for Max and Min or almost-sure for Min.

As a consequence,

Corollary 13. In any stochastic game equipped with a tail objective, the following assertions hold.

1. If there exists an almost-sure strategy with memory M , then there exists an optimal strategy with same memory.
2. The states with value 1 are exactly the almost-sure states.

Remark 14. In the sequel, we say that a game \mathcal{A} is almost-sure (resp positive), if every state in the game is almost-sure (resp positive).

Finally, the notions of positive attractor and subgame will be basic tool notions to build our proofs upon.

Definition 15 (Positive attractor). Let $f : 2^S \rightarrow 2^S$ be the operator such that for any $U \subseteq S$,

$$f(U) = T \cup \{s \in S_1 \mid \exists a \in A, p(s, a)(U) > 0\} \cup \{s \in S_2 \mid \forall a \in A, p(s, a)(U) > 0\} .$$

Then $\overrightarrow{\text{R}}_{\text{Max}}(T, S)$ is the least fixed point of f .

We define also $\overrightarrow{\text{R}}_{\text{Min}}$ as the positive attractor for Min in a dual way.

Definition 16 (Subgame). Let \mathcal{A} be a stochastic game with state space S . $\mathcal{A}[S']$ is a subgame induced by $S' \subseteq S$ if

$$(\forall s \in S'), (\exists a \in A), p(s, a)(S') = 1 .$$

3 A Polynomial Certificate

Parity and Positive-average Stochastic Games In this section we study stochastic games where Max wants to maximize the probability to achieve the objective $\text{Par} \wedge \text{Avg}_{>0}$. Again we focus on the computation of the almost-sure region. We show that deciding whether Max wins almost-surely lies in NP and we give an algorithm to compute the value of each state. The challenging part is to provide a polynomial certificate even though the almost-sure strategies may require infinite memory, hence the usual trick of guessing a strategy for max and checking whether it is almost-sure or no will not work since there are infinitely many possible strategies.

Our goal is to provide a polynomial certificate for the almost-sure winning. We want to solve the following problem

Problem 17. *For a given stochastic game \mathcal{A} with perfect information and a state s , decide whether s is almost-sure for Max for the $\text{Par} \wedge \text{Avg}_{>0}$ objective.*

Our approach consists in providing a polynomial size certificate for a subgame $\mathcal{A}[U]$ of \mathcal{A} that contains s . This notion of certificate is defined by induction on the number of priorities in the arena, and the recursive definition depends on the parity of the highest priority in the subgame $\mathcal{A}[U]$.

A precise definition of the certificates is given in Definitions 22 and 23 for a start we give a first rough description a d -certificate (where d is the number of priorities) and why they are sufficient to prove that the subgame $\mathcal{A}[U]$ is almost-sure:

- (a) If the highest priority d in the subgame $\mathcal{A}[U]$ is even, then denote S_d the set of vertices with priority d , a d -certificate is a decomposition of $\mathcal{A}[U]$ into $\overrightarrow{\text{R}_{\text{Max}}}(S_d \cap U, U)$ and $U \setminus \overrightarrow{\text{R}_{\text{Max}}}(S_d \cap U, U)$, a $(d-1)$ -certificate for the subgame $\mathcal{A}[U \setminus \overrightarrow{\text{R}_{\text{Max}}}(S_d \cap U, U)]$ and a positional strategy for Max in the subgame $\mathcal{A}[U]$ for the objective $\text{Avg}_{>0}$. This is sufficient to conclude that $\mathcal{A}[U]$ is almost-sure because Max can play as follows. If the play is in $\overrightarrow{\text{R}_{\text{Max}}}(S_d \cap U, U)$, Max applies a strategies to attempt a visit a priority- d state, then switches to an almost-sure strategy for the positive-average objective. Then she either starts these two steps again or in case the play is in $U \setminus \overrightarrow{\text{R}_{\text{Max}}}(S_d \cap U, U)$ Max apply an almost-sure strategy in this subgame.
- (b) If the highest priority d in the subgame $\mathcal{A}[U]$ is odd, then denote S_1 the set of vertices with priority 1, a d -certificate is given by a finite sequence $(R_i)_{0 \leq i \leq |U|-1}$ of disjoint subsets of $U \setminus \overrightarrow{\text{R}_{\text{Min}}}(S_d \cap U, U)$ such that *i*) for every i we have $R_i \subseteq U \setminus \bigcup_{j < i} \overrightarrow{\text{R}_{\text{Max}}}(R_j, U)$, *ii*) a $(d-1)$ -certificate for every R_i , and *iii*) the collection of sets $\overrightarrow{\text{R}_{\text{Max}}}(R_i, U)$ is a partition of U . The intuition beyond this certificate is that Max can apply a positive strategy induced by *ii*) to win the game $\text{Par} \wedge \text{Avg}_{>0}$ if the play starts from some $\overrightarrow{\text{R}_{\text{Max}}}(R_i, U)$, second we show using the qualitative determinacy (c.f. Theorem 12) that *iii*) implies that this strategy is actually almost-sure for the objective $\text{Par} \wedge \text{Avg}_{>0}$. set

In order to provide a polynomial certificate, we proceed in three steps. First we characterize the set of almost-sure states (c.f. Propositions 19 and 21). Second we formally define what is a polynomial certificate (c.f. Definitions 22 and 23) and show that its size is polynomial in the number of states and priorities. Finally we show that the certificate can be checked in time polynomial in the number of states and priorities (c.f. Lemma 25).

3.1 The Almost-sure Region

Lemma 18. *Let \mathcal{A} be a stochastic game and $\mathcal{A}[U]$ be a subgame. Suppose that the highest priority d in $\mathcal{A}[U]$ is even and let S_d be the set of vertices with priority d . Then $\mathcal{A}[U]$ is almost-sure if and only if*

1. $\mathcal{A}[U]$ is almost-sure for the positive-average objective.
2. $\mathcal{A}[U \setminus \overrightarrow{\text{R}}_{\text{Max}}(S_d \cap U, U)]$ is almost-sure for Max

Proof. Let $\mathcal{A}[U]$ be a subgame satisfying items 1. and 2. of Lemma 18. We show that $\mathcal{A}[U]$ is almost-sure for Max for the objective $\text{Par} \wedge \text{Avg}_{>0}$. Let σ_{Sub} , σ_{Attr} , and σ_{Avg} denote the almost-sure strategy in the subgame $\mathcal{A}[U \setminus \overrightarrow{\text{R}}_{\text{Max}}(S_d \cap U, U)]$, the attraction strategy to priority- d states in the subgame $\mathcal{A}[U]$, and the almost-sure strategy for the objective $\text{Avg}_{>0}$ in the subgame $\mathcal{A}[U]$ respectively. We define the application $\text{Mode} : (S \times A)^* \rightarrow \{\text{Sub}, \text{Attr}, \text{Avg}\}$ as follows:

$$\begin{aligned} \text{Mode}(s_0 a_0 \cdots s_n a_n) &= \text{Sub} \text{ if } \begin{cases} s_n \in U \setminus \overrightarrow{\text{R}}_{\text{Max}}(S_0 \cap U, U) \wedge \\ \left[(\text{Mode}(s_0 a_0 \cdots s_{n-1} a_{n-1}) = \text{Sub}) \vee \right. \\ \left. \left(\frac{1}{n} \sum_{i=0}^{n-1} \geq \eta \wedge \text{Mode}(s_0 a_0 \cdots s_{n-1} a_{n-1}) = \text{Avg} \right) \right] , \\ \end{cases} \\ \text{Mode}(s_0 a_0 \cdots s_n a_n) &= \text{Attr} \text{ if } \begin{cases} s_n \in \overrightarrow{\text{R}}_{\text{Max}}(S_0 \cap U, U) \wedge \\ \left[(n - \max \{k \mid \text{Mode}(s_0 a_0 \cdots s_k a_k) \neq \text{Attr}\} < [|U|]) \vee \right. \\ \left. \left(\frac{1}{n} \sum_{i=0}^{n-1} \geq \eta \wedge \text{Mode}(s_0 a_0 \cdots s_{n-1} a_{n-1}) = \text{Avg} \right) \right] , \\ \end{cases} \\ \text{Mode}(s_0 a_0 \cdots s_n a_n) &= \text{Avg} \text{ otherwise.} \end{aligned}$$

We assume also that $\text{Mode}(\epsilon) = \text{Avg}$ where ϵ is the empty word.

The strategy σ that Max applies is as follows.

- For every $w \in (S \times A)^*$ if the $\text{Mode}(w) = x$, then apply the strategy σ_x for $x \in \{\text{Sub}, \text{Attr}, \text{Avg}\}$.

We show that σ is almost-sure. Let $s \in U$, then if

$$\forall \tau, \mathbb{P}_s^{\sigma, \tau} (\exists N \geq 0, \forall n \geq N, \text{Mode}(S_0 A_0 \cdots S_n A_n) = \text{Sub}) = 1 .$$

Then Max plays consistent with strategy σ_{Sub} , and by definition of σ_{Sub} Max wins almost-surely.

If we have:

$$\forall \tau, \mathbb{P}_s^{\sigma, \tau} (\exists^\infty n \geq 0, \text{Mode}(S_0 A_0 \cdots S_n A_n) = \text{Attr}) = 1 .$$

First, we show that the parity objective is satisfied. Let A_n be the following sequence of events:

$$A_0 = \left\{ S^\omega \mid \left(S_0 \in \overrightarrow{\text{R}_{\text{Max}}}(S_0 \cap U, U) \right) \wedge (\forall 0 \leq i \leq |U|, \chi(S_i) \neq 0) \right\} ,$$

$$A_n = \left\{ S^\omega \mid (\exists i_0, \dots, i_n), \left(\bigcap_{j=0}^n A_{i_j} \right) \wedge (\forall j \leq i_n, (S_j \notin \overrightarrow{\text{R}_{\text{Max}}}(S_d \cap U, U)) \vee (\chi(S_j) \neq 0)) \right\} .$$

Intuitively, a play of \mathcal{M} belongs to A_n if it reaches the positive attractor to S_d n consecutive times and misses a state with priority- d . We show that that can happen only for finite number of time. Let m be the least transition probability of the \mathcal{M} , we have

$$\begin{aligned} \forall \tau, (\forall s \in S), \mathbb{P}_s^{\sigma, \tau}(A_n) &\leq (1 - m^{|U|})^{n+1} \\ &\leq (1 - m^{|S|})^{n+1} . \end{aligned}$$

Since

$$(1 - m^{|S|}) < 1 ,$$

we get

$$\forall \tau, \sum_{n>0} \mathbb{P}_s^{\sigma, \tau}(A_n) < \infty .$$

According to Borel-Cantelli Lemma we get:

$$\forall \tau, \forall s \in \overrightarrow{\text{R}_{\text{Max}}}(S_d \cap U, U), \mathbb{P}_s^{\sigma, \tau} (\text{i.o. } A_k) = 0 .$$

Hence a state with priority d is eventually visited, and the parity objective is satisfied with probability 1 when the play stays in $\overrightarrow{\text{R}_{\text{Max}}}(S_d \cap U, U)$.

Second, we prove that the positive-average objective is satisfied. By definition of σ_{Avg} there exists an integer $\eta > 0$ such that:

$$\forall \tau, \forall s \in W_{=1}[\text{Avg}_{>0}], \mathbb{P}_s^{\sigma_{\text{Avg}}, \tau} \left(\exists^\infty n \geq 0, \frac{1}{n+1} \sum_{i=0}^n r(S_i) \geq \eta \right) = 1 .$$

To show that σ satisfies the objective $\text{Avg}_{>0}$ with probability 1, we still need to show that Max can make the average reward go above η , but this is always possible since the play is happening in the almost-sure region for the positive-average condition it follows that

$$\forall \tau, \forall s \in W_{=1}[\text{Avg}_{>0}], \mathbb{P}_s^{\sigma, \tau} \left(\exists n \geq 0, \frac{1}{n+1} \sum_{i=0}^n r(S_i) \geq \eta \right) = 1 .$$

Thus the $\text{Avg}_{>0}$ objective is achieved almost-surely. The above facts show that σ is almost-sure. This show that $\mathcal{A}[U]$ is almost-sure.

Let us show that any winning region satisfies items 1 and 2. Denote W the almost-sure region for $\text{Par} \wedge \text{Avg}_{>0}$ objective played in \mathcal{A} . We prove that W satisfies items 1 and 2. That W satisfies item 1 follows from the fact that Max can win almost-surely $\text{Par} \wedge \text{Avg}_{>0}$ in $\mathcal{A}[W]$. To see that item 2 holds, note that $\mathcal{A}[W \setminus \overrightarrow{\text{R}}_{\text{Max}}(S_0 \cap W, W)]$ is a trap for Max. So if she plays her almost-sure strategy σ defined on W , she wins almost-surely the $\text{Par} \wedge \text{Avg}_{>0}$ objective. This terminates the proof. \square

Proposition 19. *Let \mathcal{A} be a stochastic game with a tail winning condition. Then the almost-sure region is given by the largest subset $W \subseteq S$ that induces a trap for Min and such that $\mathcal{A}[W]$ is almost-sure for Max.*

Proof. We show that the collection of subsets inducing a subgame and satisfying Lemma 18 is closed under union.

Let U_1 and U_2 be two subsets inducing subgames and satisfying Lemma 18, we show that $\mathcal{A}[U_1 \cup U_2]$ is almost-sure for Max for the objective $\text{Par} \wedge \text{Avg}_{>0}$ i.e. we show that $\mathcal{A}[U_1 \cup U_2]$ satisfies Lemma 18.

First we show that if $\mathcal{A}[U_1]$ and $\mathcal{A}[U_2]$ are almost-sure for $\text{Avg}_{>0}$ then $\mathcal{A}[U_1 \cup U_2]$ is almost-sure for the objective $\text{Avg}_{>0}$ as well. Let σ_i be the almost-sure strategy for the objective $\text{Avg}_{>0}$ played in the subgame $\mathcal{A}[U_i]$ for $i \in \{1, 2\}$, then in the subgame $\mathcal{A}[U_1 \cup U_2]$ Max plays as follows:

- If the play is in $\mathcal{A}[U_i]$, apply the strategy σ_i for $i \in \{1, 2\}$.

This strategy is clearly almost-sure since each $\mathcal{A}[U_i]$ is a trap for Min.

Second, since the condition is tail the almost-sure winning region W is a trap for Min and obviously $\mathcal{A}[W]$ is almost-sure. \square

Lemma 20. *Let \mathcal{A} be a stochastic game and $\mathcal{A}[U]$ a subgame. Suppose that the highest priority d in $\mathcal{A}[U]$ is odd, then $\mathcal{A}[U]$ is almost-sure if and only if there exists a sequence of disjoint subsets $(R_i)_{0 \leq i \leq |U|-1}$ such that*

1. Every R_i is a trap for Min in $\mathcal{A} \left[U \setminus \left(\overrightarrow{\text{R}}_{\text{Min}}(S_d \cap U, U) \cup \bigcup_{j=0}^i \overrightarrow{\text{R}}_{\text{Max}}(R_j, U) \right) \right]$,
2. every $\mathcal{A}[R_i]$ is almost-sure for the objective $\text{Par} \wedge \text{Avg}_{>0}$,
3. $U = \bigcup_{i=0}^{|U|-1} \overrightarrow{\text{R}}_{\text{Max}}(R_i, U)$,

Proof. Let $\mathcal{A}[U]$ be a subgame induced by a subset $U \subseteq S$ and let $(R_i)_{0 \leq i \leq |U|-1}$ be a sequence of disjoint subsets of $S \setminus \overrightarrow{\text{R}}_{\text{Min}}(S_d \cap U, U)$ such that 1,2 and 3 hold. We show that $\mathcal{A}[U]$ is almost-sure for the objective $\text{Par} \wedge \text{Avg}_{>0}$. Max applies the following strategy σ . For any state $s \in \bigcup_{i=0}^{|U|-1} \overrightarrow{\text{R}}_{\text{Max}}(R_i, U)$ we say that:

- s is *locked* if $s \in \bigcup_{i=0}^{|U|-1} R_i$ and denote $\text{ind}(s)$ the least i such that $s \in R_i$,

- s is *unlocked* if $s \in \bigcup_{i=0}^{|U|-1} \overrightarrow{\text{R}_{\text{Max}}}(R_i, U) \bigcup_{i=0}^{|U|-1} R_i$ and denote $\text{ind}(s)$ the least i such that $s \in \overrightarrow{\text{R}_{\text{Max}}}(R_i, U)$.

As long as the current state is unlocked, Max plays the attractor strategy to reach $R_{\text{ind}(s)}$ with positive probability. When the current state is locked, Max switches to her almost-sure strategy for the objective $\text{Par} \wedge \text{Avg}_{>0}$ in the subgame $\mathcal{A}[R_{\text{ind}(s)}]$ which exists according to condition 2. We show that using this strategy guarantees almost-surely that ultimately the current state S_n is locked forever and that $\text{ind}(S_n)$ remains ultimately constant. Precisely:

$$\mathbb{P}_s^\sigma (\exists 0 \leq i \leq |U| - 1, \exists N \geq 0, \forall n \geq N, S_n \in R_i) = 1 . \quad (2)$$

Since the arena is finite, there exists $x > 0$ such that for every i playing the attractor strategy to R_i in $\overrightarrow{\text{R}_{\text{Max}}}(R_i, S)$ ensures to reach R_i in at most $|U|$ steps with probability at least x . As a consequence, according to condition 1, for every $0 \leq m \leq |U| - 1$

$$\left(\exists^\infty k, S_k \in \overrightarrow{\text{R}_{\text{Max}}}(R_m, S) \right) \implies \left(\exists^\infty k, S_k \in R_{m-1} \vee \exists N \geq 0, \forall n \geq N, S_n \in R_m \right) , \quad (3)$$

\mathbb{P}_s^σ almost-surely. Let M be the random variable with values in $\{0 \dots |U| - 1\}$ defined as follows:

$$M = \liminf_n \text{ind}(S_n) ,$$

then according to (3)

$$\mathbb{P}_s^\sigma (\exists N \geq 0, \forall n \geq N, S_n \in R_M) = 1 , \quad (4)$$

which shows (2) and terminates the proof of the direct implication.

Let us prove the converse implication, we proceed by induction on the size of $|U|$. First we show that if $\mathcal{A}[U]$ is almost-sure then the subgame $\mathcal{A}[U \setminus \overrightarrow{\text{R}_{\text{Min}}}(S_d \cap U, U)]$ contains a non-empty set R such that $\mathcal{A}[R]$ is almost-sure for Max. Assume towards a contradiction the contrary, it follows that the arena $\mathcal{A}[U \setminus \overrightarrow{\text{R}_{\text{Min}}}(S_d \cap U, U)]$ is almost-sure for Min which in turn shows that $\mathcal{A}[U]$ is almost-sure for Min since Min would have a strategy to either win in the subgame $\mathcal{A}[U \setminus \overrightarrow{\text{R}_{\text{Min}}}(R_d \cap U, U)]$ or visit a state with priority 1 infinitely often (using similar argument as in the proof of Lemma 18). Hence there exists a non-empty set R_0 in $U \setminus \overrightarrow{\text{R}_{\text{Min}}}(S_d \cap U, U)$ such that R_0 is almost-sure for Max. If $S_1 = S$ we are over. Otherwise we can now use the same argument to build a subset $R_1 \subseteq U \setminus \overrightarrow{\text{R}_{\text{Max}}}(R_0, U)$ such that $\mathcal{A}[R_1]$ is almost-sure for Max. Since at each step we obtain a subgame which contains at least one state less the result follows. \square

Proposition 21. *Let \mathcal{A} be a stochastic game such that the highest priority d is odd, then the almost-sure region is given by the largest trap satisfying Lemma 20.*

Proof. This is a direct corollary of Proposition 19 and Lemma 20. \square

Now we are ready to give a formal definition of a certificate for the Problem 17.

3.2 Polynomial Size Certificate

Definition 22 (Even Certificate). *Let \mathcal{A} be a stochastic equipped with the objective $\text{Par} \wedge \text{Avg}_{>0}$ with d priorities such that the highest priority d is even, then a d -certificate for the almost-sure winning for a subgame $\mathcal{A}[U]$ is given by:*

- A positional strategy σ for Max in $\mathcal{A}[U]$,
- a $(d - 1)$ -certificate C_{d-1} for the almost-sure winning for the subgame $\mathcal{A}[U \setminus \overrightarrow{\text{R}}_{\text{Max}}(S_d \cap U, U)]$.

Definition 23 (Odd Witness). *Let \mathcal{A} be a stochastic equipped with the objective $\text{Par} \wedge \text{Avg}_{>0}$ with d -priorities and such that the highest priority d is odd, then a d -certificate for the almost-sure winning for a subgame $\mathcal{A}[U]$ is given by:*

- A sequence of disjoint subsets $(R_i)_{0 \leq i \leq |U|-1} \subseteq S \setminus \overrightarrow{\text{R}}_{\text{Min}}(S_d \cap U, U)$ such that conditions 1 and 3 of Lemma 20 hold,
- a $(d - 1)$ -certificate C_{d-1} for the almost-sure winning for the subgame $\mathcal{A}[R_i]$ for every $0 \leq i \leq |U| - 1$.

Lemma 24. *Let \mathcal{A} a stochastic game and $\mathcal{A}[U]$ a subgame of \mathcal{A} . There exists a certificate of size $O(nd)$ where n is the size of U and d the number of priorities in U which shows that $\mathcal{A}[U]$ is almost-sure for Max.*

Proof. Let $\mathcal{A}[U]$ be a subgame of a stochastic game \mathcal{A} . denote $C(n, d)$ the maximal size of a certificate for a subgame U with n vertices and d priorities. In each inductive step of the recursive definition of a certificate the size of U is reduced by at least one priority and one state. If the highest priority d is even then $|\overrightarrow{\text{R}}_{\text{Max}}(S_d, S)| \leq n$ thus,

$$C(n, d) \leq n + C(n - 1, d - 1) .$$

If the highest priority d is odd, the subsets R_i are disjoint hence $\sum_{i=0}^{|U|-1} |R_i| \leq n$ and

$$C(n, d) \leq n + \max_{\substack{n_1, \dots, n_{|U|-1} \\ n_1 + \dots + n_{|U|-1} \leq n}} \sum_i C(n_i, d - 1) ,$$

Since $C(n, 1) \leq O(n)$, it follows that

$$C(n, d) \leq O(nd) .$$

□

3.3 Checking the Certificate in Polynomial Time

Lemma 25. *Let \mathcal{A} be a stochastic game equipped with the objective $\text{Par} \wedge \text{Avg}_{>0}$, let $\mathcal{A}[U]$ be a subgame and let C be a certificate for $\mathcal{A}[U]$, then one can verify in time $O(dn^3)$ where d is the number of priorities of \mathcal{A} and n the number of states in \mathcal{A} that C is a valid certificate.*

Proof. Let C be a certificate for the almost-sure winning in the subgame $\mathcal{A}[U]$, first notice that if there is only one priority available in $\mathcal{A}[U]$, then either it is odd and $\mathcal{A}[U]$ surely losing or it is even and checking W amounts to checking if the strategy σ is almost-sure for the objective $\text{Avg}_{>0}$. For that consider the Markov decision process $\mathcal{A}[\sigma]$ induced by σ , in $\mathcal{A}[\sigma]$ then one can compute the value of every state for the mean payoff objective in $O(n^3)$ [7] and check these values are strictly positive. According to the proof of Theorem ?? this guarantees that σ is almost-sure for the $\text{Avg}_{>0}$ objective.

Assume by induction that the result holds for any subgame with less than d priorities and let $\mathcal{A}[U]$ be a subgame with d priorities.

If the highest priority d in $\mathcal{A}[U]$ is even then to check that C_d is a valid certificate, we perform the following steps:

- check that the positional strategy σ for Max is almost-sure for the objectives $\text{Avg}_{>0}$ in the subgame $\mathcal{A}[U]$.
- compute the set $\overrightarrow{\text{R}}_{\text{Max}}(S_d \cap U, U)$,
- check that C_{d-1} is a valid $(d-1)$ -certificate for the subgame $\mathcal{A}[U \setminus \overrightarrow{\text{R}}_{\text{Max}}(S_d \cap U, U)]$.

Let us show that these three steps can be performed in polynomial time. In order to verify that positional strategy σ is almost-sure in polynomial time consider the Markov decision process $\mathcal{A}[\sigma]$ induced by σ , in $\mathcal{A}[\sigma]$ one can compute the value of every state for the mean payoff objective in $O(n^3)$ [7], the computation of the set $U \setminus \overrightarrow{\text{R}}_{\text{Max}}(S_d \cap U, U)$ can be done in time $O(n^2)$ and verifying the certificate C_{d-1} can be done in polynomial time by induction hypothesis. Let $T(n, d)$ be the time complexity of the verification parametrized by n the number of states of U and d the number of priorities in U , thus:

$$T(n, d) \leq n^3 + T(n - 1, d - 1) . \quad (5)$$

If the highest priority d in $\mathcal{A}[U]$ is odd then we proceed as follows. For each $0 \leq i \leq |U| - 1$,

- check that C_i is a valid $(d-1)$ -certificate in the subgame $\mathcal{A}[R_i]$,
- compute the attractor $\overrightarrow{\text{R}}_{\text{Max}}(R_i, U)$,
- remove from U the set $\overrightarrow{\text{R}}_{\text{Max}}(R_i, U)$,
- repeat with $i \leftarrow i + 1$.

Computing the attractor set can be done in time $O(n^2)$.

Let $T(n, d)$ be the complexity of the verification parametrized by n the number of states of U and d the number of priorities in U , then:

$$T(n, d) \leq n^3 + \max_{\substack{n_1, \dots, n_{|U|-1} \\ n_1 + \dots + n_{|U|-1} \leq n}} \sum_{i=1}^n T(n_i, d-1) . \quad (6)$$

From Equations (5) and (6) and the concavity of $x \rightarrow x^3$ we obtain

$$T(n, d) = O(dn^3) .$$

□

Theorem 26. *Given a stochastic game equipped with parity and positive-average objective, whether Max has an almost-sure winning strategy from a state s can be decided in NP.*

Proof. An NP algorithm that solves this problem starts first by guessing a subset U containing state s . It first checks whether U induces a subgame, then according to Lemma 25 one can check in polynomial time whether $\mathcal{A}[U]$ is almost-sure. Hence the result. □

4 Computing the Values

In this section we give a deterministic version of the the NP algorithm presented in Section 3 and show that computation of the almost-sure region can be done in time $O(nmd + n^d)$.

The algorithm considers two cases: (a) when the highest priority d is even, and (b) when the highest priority d is odd. The details of the two cases are as follows:

- (a) If the highest priority d in the game is even, then we compute the almost-sure states of Max as the fixed point of the procedure where in each iteration removes from \mathcal{A}' some states that are positive for Min. The subgame $R \subseteq \mathcal{A}'$ contains the almost-sure states for the objective $\text{Avg}_{>0}$ (Line 5) which is a necessary condition to win according to Lemma 18. We decompose R into $\overrightarrow{R_{\text{Max}}}(S_d \cap R, R)$ and $R \setminus \overrightarrow{R_{\text{Max}}}(S_d \cap R, R)$. $R \setminus \overrightarrow{R_{\text{Max}}}(S_d \cap R, R)$ has strictly less priorities than R . The states in $R \setminus R'$ are positive for Min in the original game since $R \setminus \overrightarrow{R_{\text{Max}}}(S_d \cap R, R)$ is a trap for Max, we remove $\overrightarrow{R_{\text{Min}}}(R \setminus R')$. The correctness argument is similar to the proof of Lemma 18, namely that when $R' = R \setminus \overrightarrow{R_{\text{Max}}}(S_d \cap R, R)$, Max wins almost-surely by applying an almost-sure strategy in $R \setminus \overrightarrow{R_{\text{Max}}}(S_d \cap R, R)$, that exists by an inductive argument, and by alternating between the attraction strategy and the positive-average strategy in $\overrightarrow{R_{\text{Max}}}(S_d \cap R, R)$ as shown in the proof of Lemma 18.

Algorithm 1

Input: Stochastic game \mathcal{A} with state space S .

Output: Outputs the almost-sure winning region for Max for the objective Max .

```
1 Let  $d$  be the highest priority of  $\mathcal{A}$ .
2  $S' \leftarrow S$ 
3 if  $d$  is even then
4   repeat
5     Let  $R$  be the almost-sure winning region for Max in the subgame  $\mathcal{A}[S']$  for the
      objective  $\text{Avg}_{\geq 0}$ .
6     Compute  $\overrightarrow{R_{\text{Max}}}(S_d \cap R, R)$ , the positive attractor of Max to priority- $d$  states in
       $R$ 
7     Let  $R'$  be the almost-sure winning region for Max in the subgame  $\mathcal{A}[R \setminus$ 
       $\overrightarrow{R_{\text{Max}}}(S_d \cap R, R)]$  for the objective  $\text{Par} \wedge \text{Avg}_{> 0}$ 
8     Compute  $\overrightarrow{R_{\text{Min}}}(R \setminus R', R)$ , the positive attractor of Min to  $R \setminus R'$  in the subgame
       $\mathcal{A}[R]$ 
9      $S' \leftarrow R \setminus \overrightarrow{R_{\text{Min}}}(R \setminus R', R)$ 
10    until  $R' = R \setminus \overrightarrow{R_{\text{Max}}}(S_d \cap R, R)$ 
11    return  $S'$ 
12  else if  $d$  is odd then
13     $R' \leftarrow \emptyset$ 
14    repeat
15      Compute  $\overrightarrow{R_{\text{Min}}}(S'_d, S')$ , the positive attractor of Min to priority- $d$  states in  $\mathcal{A}[S']$ 
16      Let  $R$  be the almost-sure region for Max in the subgame  $\mathcal{A}' \setminus \overrightarrow{R_{\text{Min}}}(S'_s, S')$  for
      the objective  $\text{Par} \wedge \text{Avg}_{> 0}$ 
17      Compute  $\overrightarrow{R_{\text{Max}}}(R, S')$ , the positive attractor of Max to  $R$  in  $\mathcal{A}[S']$ 
18       $R' \leftarrow R' \cup \overrightarrow{R_{\text{Max}}}(R, S')$ 
19       $S' \leftarrow S' \setminus \overrightarrow{R_{\text{Max}}}(R, S')$ 
20    until  $R = \emptyset$ 
21    return  $R'$ 
22 end if
```

- (b) The second part of the algorithm is when the highest priority d in the game is odd, the set of almost-sure states is computed in rounds as the union of the almost-sure region for the objective $\text{Par} \wedge \text{Avg}_{> 0}$ in the subgame $\mathcal{A}' \setminus \overrightarrow{R_{\text{Min}}}(S_d \cap R, R)$. The correctness argument follows from two facts: First, according to Lemmas 20, Max wins almost-surely in the subgame induced by the union of $\overrightarrow{R_{\text{Max}}R'}$. Second, since Max cannot win in $\mathcal{A}' \setminus \overrightarrow{R_{\text{Min}}}(S_d \cap R, R)$ we are ensured that the computed set is the largest set of almost-sure winning states.

Theorem 27 (Algorithmic Complexity). *In stochastic games, one can compute the almost-sure region for the objective $\text{Par} \wedge \text{Avg}_{> 0}$ in time $O(nmd + n^d)$, where m is the*

time one needs to solve positive-average objectives.

Proof. This problem is solved by Alg 1, the correctness follows from the arguments above. Let $O(m)$ be the time complexity one needs to solve positive-average objectives. Denote $T(d)$ the complexity of Alg 1 parametrized by the number of priorities in the input game. The computation of the attractors in lines 6,8,15,17 is subsumed by the computation of the almost-sure region for the objective positive-average since solving these games lie in $\text{NP} \cap \text{CoNP}$. In each recursive call the set of states reduces by at least one state and one priority and since there are at most n recursive calls we get

$$T(d) \leq n(m + T(d - 1)) ,$$

where m is the time one needs to solve positive-average objectives. It follows that

$$T(d) \leq nmd + n^d ,$$

hence the result. □

5 Conclusion

In this paper we studied the problem of almost-sure winning for stochastic games equipped with the objective $\text{Par} \wedge \text{Avg}_{>0}$ and the main result we obtain is: despite the fact that almost-sure strategies may require infinite memory, there exists an NP algorithm that computes the almost-sure region and an almost-sure strategy. Unfortunately this procedure does work only for the lim sup semantics. Indeed the correctness proof for the almost-sure strategy described does not provide any lower bound on the accumulated average reward and hence the main argument used breaks in the case of $\text{Par} \wedge \underline{\text{Avg}}_{>0}$. However we believe that the almost-sure region for the objective $\text{Par} \wedge \text{Avg}_{>0}$ and $\text{Par} \wedge \underline{\text{Avg}}_{>0}$ are equivalent, we finish this paper by the following conjecture:

Conjecture 28. *Let \mathcal{A} be a stochastic game and let s be a state, then:*

$$s \in W_{=1}[\text{Par} \wedge \text{Avg}_{>0}] \iff s \in W_{=1}[\text{Par} \wedge \underline{\text{Avg}}_{>0}] .$$

a similar procedure works for the lim inf semantics if c but we do not have any proof of it.

References

- [1] Krishnendu Chatterjee, Tom Henzinger, and Marcin Jurdzinski. Mean-payoff parity games. In *LICS 05*, June 2005.
- [2] Anne Condon. The complexity of stochastic games. *Inf. Comput.*, 96(2):203–224, 1992.

- [3] Hugo Gimbert and Florian Horn. Solving Simple Stochastic Tail Games. page 1000, 01 2010.
- [4] Maitra and Sudderth. Stochastic games with borel payoffs. *A Maitra, W Sudderth - Stochastic Games and Applications, NATO . . . , 2003 - ratio.huji.ac.il*.
- [5] D. A. Martin. Borel determinacy. *Annals of Mathematics*, 102:363–371, 1975.
- [6] Donald A. Martin. The determinacy of blackwell games. *J. Symb. Log.*, 63(4):1565–1581, 1998.
- [7] Martin L. Putterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, New York, NY, 1994.