Perfect-Information Stochastic Mean-Payoff Parity Games⋆,⋆⋆

Krishnendu Chatterjee (IST Austria), Laurent Doyen (LSV, ENS Cachan & CNRS, France), Hugo Gimbert (LaBRI, CNRS, France), and Youssouf Oualhadj (LIF, Université Aix Marseille, France)

Abstract The theory of graph games is the foundation for modeling and synthesizing reactive processes. In the synthesis of stochastic processes, we use $2\frac{1}{2}$-player games where some transitions of the game graph are controlled by two adversarial players, the System and the Environment, and the other transitions are determined probabilistically. We consider $2\frac{1}{2}$-player games where the objective of the System is the conjunction of a qualitative objective (specified as a parity condition) and a quantitative objective (specified as a mean-payoff condition). We establish that the problem of deciding whether the System can ensure that the probability to satisfy the mean-payoff parity objective is at least a given threshold is in $\mathsf{NP} \cap \mathsf{coNP}$, matching the best known bound in the special case of 2-player games (where all transitions are deterministic). We present an algorithm running in time $O(d \cdot n^{2d} \cdot \mathsf{MeanGame})$ to compute the set of *almost-sure* winning states from which the objective can be ensured with probability 1, where n is the number of states of the game, d the number of priorities of the parity objective, and $\mathsf{MeanGame}$ is the complexity to compute the set of almost-sure winning states in $2\frac{1}{2}$-player mean-payoff games. Our results are useful in the synthesis of stochastic reactive systems with both functional requirement (given as a qualitative objective) and performance requirement (given as a quantitative objective).

1 Introduction

Perfect-information stochastic games. A perfect-information stochastic graph game [16] is played on a finite directed graph with three kinds of states (or vertices): player-Max, player-Min, and probabilistic states. At player-Max states, player Max chooses a successor state; at player-Min states, player Min (the adversary of player Max) chooses a successor state; and at probabilistic states, a successor state is chosen according to a fixed probability distribution. The result of playing the game forever is an infinite path through the graph. If there are no probabilistic states, we refer to the game as a 2-*player graph game*; otherwise, as a $2\frac{1}{2}$-*player graph game*. There has been a long history of using 2-player graph games for modeling and synthesizing reactive processes [7,23,26]: a reactive system and its environment represent the two players,

⋆ This research was supported by Austrian Science Fund (FWF) Grant No P23499- N23, FWF NFN Grant No S11407-N23 (RiSE), ERC Start grant (279307: Graph Games), Microsoft Faculty Fellowship Award, and European project Cassting (FP7-601148).

⋆⋆ Fuller version: IST Technical Report No IST-2013-128.

whose states and transitions are specified by the states and edges of a game graph. Consequently, $2\frac{1}{2}$-player graph games provide the theoretical foundation for modeling and synthesizing processes that are both reactive and stochastic [17,25]. They subsume both 2-player games which have no probabilistic states, and Markov decision processes (MDPs) which have no player-Min states.

Qualitative and quantitative objectives. In the analysis of reactive systems, the goal is specified as a set of desired paths (such as ω-regular specifications), or as a quantitative optimization objective for a payoff function on the paths. In verification and synthesis of reactive systems all commonly used properties are expressed as ω-regular objectives, and parity objectives are a canonical way to express ω-regular objectives [27]. In a parity objective, an integer priority is assigned to every state, and a path satisfies the objective for player Max if the maximum priority visited infinitely often is even. The most classical example of quantitative objective is the mean-payoff objective [17,24], where a reward is associated with every state and the payoff of a path is the long-run average of the rewards of the path. While traditionally the verification and the synthesis problems were considered with qualitative objectives, recently combinations of qualitative and quantitative objectives have received a lot of attention. Qualitative objectives such as ω-regular objectives specify the functional requirements of reactive systems, whereas the quantitative objectives specify resource consumption requirements (such as for embedded systems or power-limited systems). Combining quantitative and qualitative objectives is crucial in the design of reactive systems with both resource constraints and functional requirements [9,14,5,3]. For example, mean-payoff parity objectives are relevant in synthesis of optimal performance lock-synchronization for concurrent programs [8], where one player is the synchronizer, the opponent is the environment, and the randomization arises due to the randomized scheduler; the performance objective is specified as mean-payoff condition and the functional requirement (e.g., data-race freedom or liveness) as an ω-regular objective. Mean-payoff parity objectives have also been used in other applications such as to define permissivity for parity games [6]. Thus $2\frac{1}{2}$-player mean-payoff parity games provide the theoretical foundation for analysis of stochastic reactive systems with functional as well as performance requirements.

Algorithmic questions in $2\frac{1}{2}$-player games. The study of $2\frac{1}{2}$-player games has a wealth of algorithmic problems. For example, given a $2\frac{1}{2}$-player game with reachability objective (where the goal is to reach a target set of states), whether the player Max can ensure the objective with probability at least $\frac{1}{2}$ (called the value-strategy problem) is in $\mathsf{NP} \cap \mathsf{coNP}$ [16]. This is one of the rare combinatorial problems that belong to $\mathsf{NP} \cap \mathsf{coNP}$, but are not known to be solvable in polynomial time. It is a major and long-standing open question whether the problem can be solved in polynomial time. Moreover, 2-player games with mean-payoff (resp. parity) objectives lies in $\mathsf{NP} \cap \mathsf{coNP}$ (even in $\mathsf{UP} \cap \mathsf{coUP}$) [21,29,20], and again no polynomial time algorithm is known. Both 2-player parity games and 2-player mean-payoff games admit a polynomial reduction to the value-strategy problem of $2\frac{1}{2}$-player reachability games. The value-strategy problem for $2\frac{1}{2}$-player mean-payoff (resp. parity) games also lie in $\mathsf{NP} \cap \mathsf{coNP}$: the key property to show that the problem is in $\mathsf{NP} \cap \mathsf{coNP}$ for mean-payoff (resp. parity) games is to show that it is sufficient to consider positional strategies (that are independent of the past history and depend only on the current state), see [22] for mean-payoff and [15] for parity objectives. In this work we consider $2\frac{1}{2}$-player games with conjunction of mean-payoff and parity objectives for player Max. The study of $2\frac{1}{2}$-player games with conjunction of mean-payoff and parity objectives poses new algorithmic challenges as *infinite-memory* strategies are required. The key challenge is to obtain *succinct* (polynomial) witnesses for the infinite-memory strategies and their characterization to obtain complexity results matching the simpler classes of games where positional strategies suffice. Besides the complexity result, our characterization of strategies will also allow us to obtain algorithms to solve $2\frac{1}{2}$-player mean-payoff parity games.

Contributions. The details of our contributions are as follows:

1. We first present polynomial witnesses for infinite-memory strategies required by player Max, and a polynomial-time verification procedure for the witnesses, thereby establishing that the value-strategy problem (of whether player Max can ensure that the probability to satisfy the objective is at least a given threshold) is in NP. The fact that player Max requires infinite-memory strategies follows from the special case of 2-player mean-payoff parity games [14].
2. We show that positional strategies are sufficient for player Min (note that player Max and Min are asymmetric since player Max has a conjunction of parity and mean-payoff objectives to satisfy, whereas player Min has disjunction of parity or mean-payoff objectives to falsify). From the existence of positional strategies for player Min it follows that the value-strategy problem is also in coNP. Our NP $\cap$ coNP bound for the problem matches the special cases of 2-player mean-payoff parity games.
3. We present an algorithm for the computation of the almost-sure winning set (the set of states where the objective can be ensured with probability 1 by player Max) for $2\frac{1}{2}$-player mean-payoff parity games in time $O(d \cdot n^{2d} \cdot \mathsf{MeanGame})$, where n is the number of states of the game graph, d the number of priorities of the parity objective, and MeanGame denotes the complexity to compute the almost-sure winning set in $2\frac{1}{2}$-player mean-payoff games.

In summary, we present results that establish computational, strategy, and algorithmic complexity of solving $2\frac{1}{2}$-player mean-payoff parity games.

Technical difficulty. For 2-player games the NP $\cap$ coNP result for mean-payoff parity objectives was established in [10]: the technique relied on reduction of 2-player mean-payoff parity games to 2-player energy-parity games, and in 2-player energy-parity games finite-memory strategies suffice (for details related to energy objectives see [10,12]). However the technique of reduction of mean-payoff games to energy games (even without the parity condition) for almost-sure winning does not hold in the presence of stochastic transitions because for energy conditions (which are like safety conditions) the precise probabilities do not matter, whereas they matter for mean-payoff conditions. Hence the techniques for 2-player mean-payoff parity games do not extend to $2\frac{1}{2}$-player games, and we need to explicitly construct succinct witness to show the NP $\cap$ coNP result. The succinct witness construction of infinite-memory strategies in the presence of adversary and stochastic transitions is the main technical challenge in $2\frac{1}{2}$-player mean-payoff parity games.

Related works. The problem of 2-player mean-payoff parity games was first studied in [14]. The $\mathsf{NP} \cap \mathsf{coNP}$ complexity bound was established in [10], and an improved algorithm for the problem was given in [6]. The algorithmic analysis of $2\frac{1}{2}$-player mean-payoff games has been studied in [1,4]: a reduction to $2\frac{1}{2}$-player reachability games was presented in [1], and approximation schemes were considered in [4]. The polynomial time complexity for MDPs with mean-payoff parity objectives was established in [11] and the polynomial time complexity for MDPs with positive average parity objectives was shown in [19]. The generalization to $2\frac{1}{2}$-player games with mean-payoff parity objectives gives rise to many delicate issues, such as dealing at the same time with infinite-memory strategies, stochastic transitions, as well as the opponent.

2 Definitions

In this section we present definitions of game graphs, objectives, and the basic decision problems.

Probability distributions. For a finite set S, we denote by $\Delta(S)$ the set of all probability distributions over S, i.e., the set of functions $p : S \to [0,1]$ such that $\sum_{s \in S} p(s) = 1$. For a set $U \subseteq S$ we use the following notation: $p(U) = \sum_{s \in U} p(s)$.

Stochastic games. A *perfect-information stochastic game graph* (for brevity, stochastic game) is a tuple $\mathcal{G} = (S, (S_{\mathsf{Max}}, S_{\mathsf{Min}}), A, \delta)$, where S is a finite set of states, $(S_{\mathsf{Max}}, S_{\mathsf{Min}})$ is a partition of S such that S_{Max} is the set of states controlled by player Max and S_{Min} is the set of states controlled by player Min, A is a finite set of actions, and $\delta : S \times A \to \Delta(S)$ is a probabilistic transition function. Stochastic games are also known as $2\frac{1}{2}$-player games where probabilistic states are explicitly present. In our model, the probabilistic states can be embedded in the probabilistic transition function. A *Markov decision process* (MDP) is the special case of a stochastic game where either $S_{\mathsf{Max}} = \emptyset$, or $S_{\mathsf{Min}} = \emptyset$. Typically in this paper, we obtain MDPs from stochastic games after fixing the action choices of one of the players.

For complexity issues, we assume that the probabilities in stochastic games are rational numbers whose numerator and denominator are encoded in binary. We denote by $|\delta|$ the size of the encoding of the probabilistic transition function δ.

Subgames and traps. Given a stochastic game $\mathcal{G}$, a set $U \subseteq S$ of states induces a subgame if for all $s \in U$, there exists an action $a_s \in A$ such that $\delta(s, a_s)(U) = 1$; the induced subgame is $\mathcal{G}[U] = (U, (U \cap S_{\mathsf{Max}}, U \cap S_{\mathsf{Min}}), A, \delta')$ where, for all states $s \in U$ and action $a \in A$, we have $\delta'(s,a) = \delta(s,a)$ if $\delta(s,a)(U) = 1$, and $\delta'(s,a) = \delta(s, a_s)$ otherwise. We take this definition of subgame to keep the same alphabet of actions in every state. The subgame $\mathcal{G}[U]$ is a *trap* for player Min in the original game $\mathcal{G}$ if for all $s \in U \cap S_{\mathsf{Min}}$ and for all $a \in A$ we have $\delta(s,a)(U) = 1$. A trap for player Max is defined similarly.

Plays and strategies. A *play* $\rho = s_0 s_1 \cdots \in S^\omega$ is an infinite sequence of states such that for all $i \geq 0$ there exists $a \in A$ such that $\delta(s_i, a)(s_{i+1}) > 0$. A *strategy* for Max is a recipe to describe what is the next action to play; formally, it is a function $\sigma : S^* S_{\mathsf{Max}} \to A$. A *positional* strategy is independent of the past and depends only on the current state. We view it as a function $\sigma : S_{\mathsf{Max}} \to A$.

A strategy σ uses *finite memory* if there exists an equivalence relation $\sim$ on S^ω of finite index, such that $\sigma(\rho_1) = \sigma(\rho_2)$ for all plays ρ_1, ρ_2 such that $\rho_1 \sim \rho_2$. We define strategies, positional strategies, and finite-memory strategies analogously for Min. A strategy that is not finite-memory is refered to as an infinite-memory strategies.

Probability measures. Given a finite prefix $\rho \in S^*$ of a play, denote by $|\rho|$ the length of ρ and by $\mathsf{Cone}(\rho)$ the set of plays with prefix ρ. If $\rho \in S^+$ is nonempty, we denote by $\mathsf{Last}(\rho)$ the last state of ρ. Given a pair of strategies (σ, τ) for Max and Min, and an initial state s, we first define the probability measure on cones inductively as follows: for all $s' \in S$, let

$$\mathbb{P}_s^{\sigma,\tau}(\mathsf{Cone}(s')) = \begin{cases} 1 & \text{if } s' = s \\ 0 & \text{if } s' \neq s \end{cases}$$

and for all $\rho \in S^+$ (where $S^+ = S^* \setminus \{\epsilon\}$ and ϵ is the empty string), let

$$\mathbb{P}_s^{\sigma,\tau}(\mathsf{Cone}(\rho \cdot s')) = \begin{cases} \mathbb{P}_s^{\sigma,\tau}(\mathsf{Cone}(\rho)) \cdot \delta(\mathsf{Last}(\rho), \sigma(\rho))(s') & \text{if } \mathsf{Last}(\rho) \in S_{\mathsf{Max}} \\ \mathbb{P}_s^{\sigma,\tau}(\mathsf{Cone}(\rho)) \cdot \delta(\mathsf{Last}(\rho), \tau(\rho))(s') & \text{if } \mathsf{Last}(\rho) \in S_{\mathsf{Min}} \end{cases}$$

By Caratheodary's extension theorem, there is a unique extension of this probability measure to S^ω which is also denoted as $\mathbb{P}_s^{\sigma,\tau}(\cdot)$ [2].

Mean-payoff parity objectives. An *objective* is a measurable set $\varphi \subseteq S^\omega$ of plays. Let $\mathsf{rwd} : S \times S \to \mathbb{Q}$ be a *reward function* defined on edges and $\chi : S \to \mathbb{N}$ be a *priority function* defined on states. Given a set of states $U \subseteq S$ and a priority $d \in \mathbb{N}$, we denote by $U(d)$ the set $\{s \in U \mid \chi(s) = d\}$ of states with priority d. The *mean-payoff* objective $\mathsf{Mean} = \left\{ s_0 s_1 \cdots \in S^\omega \mid \limsup_{n\to\infty} \frac{1}{n} \cdot \sum_{i=0}^{n-1} \mathsf{rwd}(s_i, s_{i+1}) \geq 0 \right\}$ requires that the long-run average of rewards be non-negative. The *parity objective* $\mathsf{Par} = \{s_0 s_1 \cdots \in S^\omega \mid \limsup_{n\to\infty} \chi(s_n) \text{ is even}\}$ requires that the maximal priority visited infinitely often be even. The *mean-payoff parity objective* $\mathsf{Mean} \cap \mathsf{Par}$ is the conjunction of a mean-payoff objective Mean and a parity objective Par.

Almost-sure and positive winning. We say that player Max wins almost-surely (resp., positively) from an initial state s for an objective φ if there exists a strategy σ for Max such that for every strategy τ of player Min we have $\mathbb{P}_s^{\sigma,\tau}(\varphi) = 1$ (resp., $\mathbb{P}_s^{\sigma,\tau}(\varphi) > 0$). The state s is called *almost-sure* (resp., *positive*) winning for Max. In the sequel, we say that a game $\mathcal{G}$ is almost-sure (resp., positive) winning, if every state in $\mathcal{G}$ is almost-sure (resp., positive) winning for Max. We use analogous definitions for player Min. Note that almost-sure winning for Max is the dual of positive winning for Min.

Value-strategy problem and reduction to almost-sure winning. Given a threshold λ, the *value-strategy* problem for an objective asks whether there exists a strategy for player Max to ensure against all strategies of player Min that the objective is satisfied with probability at least λ. A strategy for player Max is *optimal* if it ensures the maximal value λ (for stochastic mean-payoff parity games, optimal strategies are guaranteed to exist [18]). In this paper we focus on the *almost-sure winning problem*, which is to decide whether there exists an almost-sure winning strategy for player Max for a mean-payoff parity objective, that is the value-strategy problem for $\lambda = 1$. While for player Max infinite-memory strategies are necessary [14], we will show that for player Min positional strategies are sufficient, and that the almost-sure winning problem is in $\mathsf{NP} \cap \mathsf{coNP}$.

Remark 1. It follows from the results of [13, Lemma 7] and [18, Theorem 4.1] that since mean-payoff parity objectives are *tail* objectives (independent of finite prefixes), the memory requirement for optimal strategies of both players is the same as for almost-sure winning strategies, and if the almost-sure winning problem is in $\mathsf{NP} \cap \mathsf{coNP}$, then the value-strategy problem is also in $\mathsf{NP} \cap \mathsf{coNP}$. The details are as follows: The results of [13, Lemma 7] and [18, Theorem 4.1] show that for the quantitative analysis of tail objectives it suffices to guess the *value classes* (where a value class for r, with $0 \leq r \leq 1$, is the set of states with value r), almost-sure winning witness in a modified game for each value class, and then the verification problem requires the almost-sure witness verification in each value class, and verification of MDPs which is polynomial time. Since $\mathsf{NP} \cap \mathsf{coNP}$ bound for the almost-sure problem imply polynomial witness and polynomial-time verification for the witness, it follows (using the results of [13,18]) that the $\mathsf{NP} \cap \mathsf{coNP}$ bound for almost-sure winning imply that there exists polynomial witness and polynomial-time verification for quantitative analysis, and thereby establish the $\mathsf{NP} \cap \mathsf{coNP}$ bound. Thus from our results it will follow that the value-strategy problem is in $\mathsf{NP} \cap \mathsf{coNP}$ for $2\frac{1}{2}$-player mean-payoff parity games.

Positive attractors. Given a stochastic game $\mathcal{G}$, let $U \subseteq S$ induce a subgame $\mathcal{G}[U]$ with probabilistic transition function $\delta : U \times A \to \Delta(U)$. For $T \subseteq U$, let $f_T : 2^U \to 2^U$ be the operator such that for all $Z \subseteq U$,

$$\begin{aligned} f_T(Z) = T &\cup \{s \in S_{\mathsf{Max}} \cap U \mid \exists a \in A : \delta(s,a)(Z) > 0\} \\ &\cup \{s \in S_{\mathsf{Min}} \cap U \mid \forall a \in A : \delta(s,a)(Z) > 0\} \,. \end{aligned}$$

Then $\mathsf{Attr}_{\mathsf{Max}}(T, \mathcal{G}[U])$ is the least fixed point of f_T, called the *positive attractor* for Max to T in $\mathcal{G}[U]$. It can be computed as the limit of the iteration $(f_T^i(\emptyset))_{i \in \mathbb{N}}$. There exists a positional strategy for Max (referred to as *positive-attractor* strategy) to ensure that from all states in $\mathsf{Attr}_{\mathsf{Max}}(T, \mathcal{G}[U])$, the set T is reached within $|U|$ steps with positive probability. We define $\mathsf{Attr}_{\mathsf{Min}}(T, \mathcal{G}[U])$ as the positive attractor for Min in an analogous way. An important property of positive attractors is that if X is a positive attractor for Max in $\mathcal{G}[U]$, then $\mathcal{G}[U \setminus X]$ is a subgame and it is a trap for Max. Analogous statement holds for Min.

3 Characterization of the Almost-sure Winning Set

In this section we present the key lemmas that enable an inductive characterization of certificates and a polynomial-time verification procedure for the existence of almost-sure winning strategies, showing that the almost-sure winning problem is in NP for stochastic games with mean-payoff parity objectives.

It follows from the results of [14] that finite-memory strategies are not sufficient for Max and infinite-memory strategies are required for almost-sure winning. We present polynomial witnesses and polynomial-time verification procedure for the infinite-memory almost-sure winning strategies. The polynomial witnesses consists of a trap U for player Min that defines a subgame where all states are almost-sure winning for player Max, together with a certificate defined as an inductive decomposition of the subgame induced by U constructed according to the parity of the largest priority d in

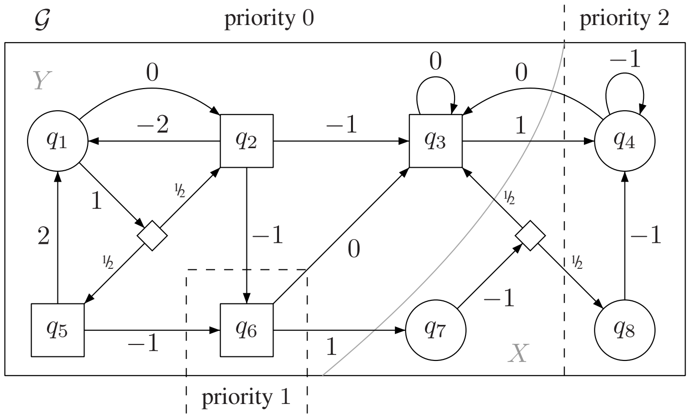

Figure 1. Stochastic game $\mathcal{G}$ with largest priority even.

U. If d is even we refer to the certificate as an *even certificate*, if d is odd as an odd certificate.

Intuitive description. To present the intuition of the (inductive) certificates, we informally explain some key properties in establishing that all states in a (sub)game are almost-sure winning for Max. In figures, we denote states of player Max by circles, and states of player Min by square boxes. Probability distributions over states are shown by a diamond. We omit actions and assume that every outgoing edge from player-Max and player-Min states corresponds to a different action. Let $\mathcal{G}$ be a (sub)game with state space S where all states are almost-sure winning. Then, we describe a certificate according to the parity of the largest priority d in $\mathcal{G}$ as follows.

1. If d is even (see Example 1 and Figure 1), let $X = \mathsf{Attr}_{\mathsf{Max}}(S(d), \mathcal{G})$ and $Y = S \setminus X$. An even certificate for $\mathcal{G}$ ensures that (1) in $\mathcal{G}$ all states are almost-sure winning for the objective Mean; and (2) in $\mathcal{G}[Y]$ all states are almost-sure winning for Max for the objective Mean $\cap$ Par (using a certificate defined recursively in the subgame $\mathcal{G}[Y]$, which has at least one less priority as there is no priority-d state in Y). In other words, the even certificate consists of (i) a positional positive attractor strategy in X for the target $S(d)$; (ii) a positional almost-sure winning strategy in $\mathcal{G}$ for the mean-payoff objective; and (iii) a certificate for $\mathcal{G}[Y]$. We establish that the above two conditions ensure that in $\mathcal{G}$ all states are almost-sure winning for Max for the objective Mean $\cap$ Par. An almost-sure winning strategy for Max is as follows: if the current state is in the subgame $\mathcal{G}[Y]$, then player Max ignores the history of the play up to the last state that was not in Y, and uses an almost-sure winning strategy in $\mathcal{G}[Y]$ (such a strategy exists in $\mathcal{G}[Y]$ by the certificate). If the opponent decides to visit the positive attractor X, then player Max switches to a (positional) positive-attractor strategy for at most $|S|$ steps. Then, either after $|S|$ steps or before (e.g., if a state with priority d is reached), player Max switches to an almost-sure winning strategy for Mean and plays it for a long finite time (that increases over the play). After that, the play might be in Y or in X, and player Max restarts from scratch

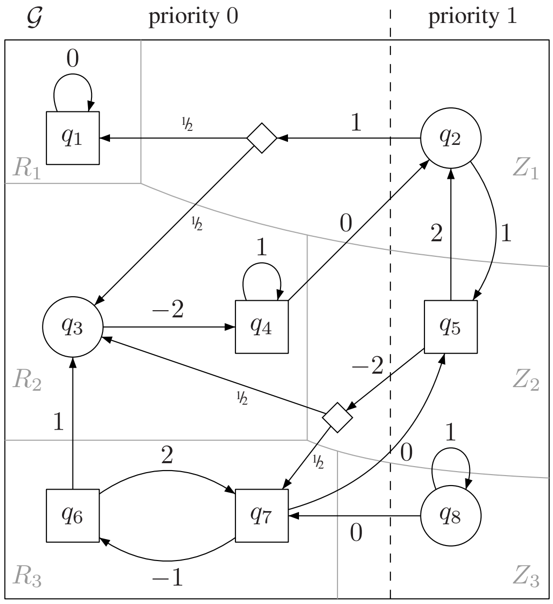

Figure 2. Stochastic game $\mathcal{G}$ with largest priority odd.

the same process of playing. Intuitively, if the play keeps visiting X, then with probability 1 the positive-attractor strategy ensures infinitely many visits to a state with priority d (thus the parity condition is satisfied), and the almost-sure winning strategy for Mean played for increasing number of steps ensures that the mean-payoff objective is satisfied. On the other hand, if the play eventually stays in $\mathcal{G}[Y]$ forever, then the almost-sure winning strategy in $\mathcal{G}[Y]$ ensures the mean-payoff parity objective is satisfied with probability 1 (since the objective is independent of finite prefixes).

Example 1. Consider the stochastic game $\mathcal{G}$ in Figure 1 where the largest priority is 2. All states are almost-sure winning for the Mean objective, and a positional strategy for player Max is as follows: for state q_1 choose the edge labeled reward 1; and for state q_4 choose the edge to q_3. The positive attractor for Max to the largest priority is $X = \{q_4, q_7, q_8\}$. In the subgame induced by $Y = \{q_1, q_2, q_3, q_5, q_6\}$ there is one less priority, and player Min can decide to leave the subgame in states q_3 and q_6. An (odd) certificate defined in the subgame $\mathcal{G}[Y]$ witnesses that all states in $\mathcal{G}[Y]$ are almost-sure winning for the mean-payoff parity objective. Thus the even certificate consists of the positional strategy for Mean, the positive-attractor strategy, and a certificate for $\mathcal{G}[Y]$.

2. If d is odd (see Example 2 and Figure 2), an odd certificate is a *layer-decomposition* of the state space of $\mathcal{G}$ into non-empty sets $R_1, \ldots, R_k$ and $Z_1, \ldots, Z_k$ defined recursively as follows: (1) $R_1 \subseteq S \setminus S(d)$ is a trap for player Min in $\mathcal{G}$ that contains no priority-d state, and such that all states in R_1 are almost-sure winning for Max

for the objective $\mathsf{Mean} \cap \mathsf{Par}$ (using a certificate defined recursively in the subgame $\mathcal{G}[R_1]$, which has at least one less priority since priority d does not occur in R_1), (2) $Z_1 = \mathsf{Attr}_{\mathsf{Max}}(R_1, \mathcal{G})$ is the positive attractor for player Max to R_1 in $\mathcal{G}$, and (3) the sets R_2 and Z_2 are defined analogously in the subgame $\mathcal{G}[S \setminus Z_1]$, and the sets R_3 and Z_3 in the subgame $\mathcal{G}[S \setminus Z_2]$ where $Z_2 = \mathsf{Attr}_{\mathsf{Max}}(R_2, \mathcal{G}[S \setminus Z_1])$, and so on to obtain the layer-decomposition of $\mathcal{G}$. Such a decomposition must cover the state space, and thus the sets $Z_1, \dots, Z_k$ form a partition of S (and $k \leq |S|$). An almost-sure winning strategy for player Max is as follows: if the current state is in a subgame R_i, then player Max ignores the history of the play up to the last state that was not in R_i, and uses an almost-sure winning strategy (that exists in R_i by the certificate). If the current state is in $Z_i \setminus R_i$, then player Max uses the positive-attractor strategy defined in Z_i. We show that almost-surely, one of the sets R_i is never left from some point on, and then the almost-sure winning strategy in $\mathcal{G}[R_i]$ ensures that the mean-payoff parity objective is satisfied with probability 1 (since the objective is independent of finite prefixes).

Example 2. Consider the stochastic game $\mathcal{G}$ in Figure 2 where the largest priority is 1. A layer-decomposition is shown where $R_1 = \{q_1\}$ is a trap of almost-sure winning states for Max, and $Z_1 = \{q_1, q_2\}$ is the positive attractor to R_1. In the subgame $\mathcal{G}[S \setminus Z_1]$, there is no edge from q_4 to q_2, and it follows that the states in $R_2 = \{q_3, q_4\}$ form a trap of almost-sure winning states in this subgame, and the positive attractor to R_2 is $Z_2 = R_2 \cup \{q_5\}$. The last layer consists of $R_3 = \{q_6, q_7\}$ and $Z_3 = R_3 \cup \{q_8\}$. As this layer-decomposition covers the state space of $\mathcal{G}$, it gives an odd certificate for player Max.

Given the basic intuitions, we now present the formal proofs. We start with a basic lemma, and then consider the two cases when the largest priority is even or odd.

Lemma 1. *Let $\mathcal{G}$ be a stochastic mean-payoff game with state space S where all states are almost-sure winning for the mean-payoff objective* Mean. *Then there exists a positional strategy σ for player* Max *such that against all strategies τ for* Min*, for all $s \in S$ and for all $\epsilon > 0$, there exists k_ϵ such that for all $k \geq k_\epsilon$ we have* $\mathbb{P}_s^{\sigma,\tau}\left(\left\{s_0 s_1 \ldots \in S^\omega \mid \sum_{i=0}^{k-1} \frac{1}{k} \cdot \mathsf{rwd}(s_i, s_{i+1}) \geq -\epsilon\right\}\right) \geq 1 - \epsilon$.

Lemma 2. *Let $\mathcal{G}$ be a stochastic mean-payoff parity game with state space S and such that the largest priority d in $\mathcal{G}$ is even. Let $X = \mathsf{Attr}_{\mathsf{Max}}(S(d), \mathcal{G})$ and $Y = S \setminus X$. All states in $\mathcal{G}$ are almost-sure winning for player* Max *with the mean-payoff parity objective* $\mathsf{Mean} \cap \mathsf{Par}$ *if and only if:*

1. *all states in $\mathcal{G}$ are almost-sure winning for the mean-payoff objective* Mean *for* Max*, and*
2. *all states in $\mathcal{G}[Y]$ are almost-sure winning for the mean-payoff parity objective* $\mathsf{Mean} \cap \mathsf{Par}$ *for* Max.

Proof. Let $\mathcal{G}$ satisfy the conditions of the lemma. We first show that all states in $\mathcal{G}$ are almost-sure winning for Max for the objective $\mathsf{Mean} \cap \mathsf{Par}$. Let σ_{Sub} be an almost-sure winning strategy for $\mathsf{Mean} \cap \mathsf{Par}$ in the subgame $\mathcal{G}[Y]$ induced by Y, let σ_{Attr} be a positional positive-attractor strategy to $S(d)$ in $\mathcal{G}$, and let σ_{Mean} be an almost-sure winning strategy for Mean in $\mathcal{G}$. Let $W = \max_{s,s' \in S} |\mathsf{rwd}(s, s')|$ be the largest absolute reward and for every $j > 0$, let $\epsilon_j = \frac{1}{j}$ and let $K_j = \max \left\{ k_{\epsilon_j}, j^2 \cdot W \right\}$ where k_{ϵ_j} is defined in Lemma 1.

The strategy σ that Max uses is played in *rounds* numbered $1, 2, \cdots$, and at round i, the strategy σ is defined as follows:

Phase 1: (Mean-payoff phase). Let j be the length of the current play prefix until the end of phase 3 of round $i-1$; then play according to the positional strategy σ_{Mean} for K_j steps. Switch to Phase 2.

Phase 2: (Subgame phase). While the current play ρ is in Y, let ρ' be the suffix of ρ obtained by ignoring the prefix of ρ up to the end of Phase 1 of the current round. Play $\sigma_{\mathsf{Sub}}(\rho')$. If the play leaves Y (and thus reaches X), then switch to Phase 3.

Phase 3: (Attractor phase). Play σ_{Attr} for at most $|S|$ steps, or until a state with priority d is reached, or the positive attractor X is left. Switch to Phase 1 in round $i+1$.

We show that σ is almost-sure winning for the $\mathsf{Mean} \cap \mathsf{Par}$ objective. Consider the following events:

$$A = \{s_0 s_1 \cdots \mid \exists J \geq 0 \cdot \forall j \geq J : s_j \in Y\},$$
$$B = \{s_0 s_1 \cdots \mid \forall J \geq 0 \cdot \exists j \geq J : s_j \in X\}.$$

Intuitively, A denotes that from some point on the play remains only in the subgame Y (and thus the strategy σ remains forever in the subgame phase), and B denotes that the set X (the positive attractor to priority d) is visited infinitely often. Let τ be a strategy for Min, then any play consistent with (σ, τ) belongs to $A \cup B$ and since $A \cap B = \emptyset$ we have $\mathbb{P}_s^{\sigma,\tau}(A \cup B) = \mathbb{P}_s^{\sigma,\tau}(A) + \mathbb{P}_s^{\sigma,\tau}(B) = 1$. We now consider two cases to establish that σ is almost-sure winning.

1. *(Under event A).* Observe that both parity and mean-payoff objectives are independent of finite prefixes, and if a play belongs to A, then the finite prefix of the play after which the play only visits states in Y does not change the mean-payoff nor the parity objective. Since σ_{Sub} is almost-sure winning in the subgame induced by Y, it follows that for all $s \in S$ and all strategies τ of player Min in $\mathcal{G}$ we have $\mathbb{P}_s^{\sigma,\tau}(\mathsf{Mean} \cap \mathsf{Par} \mid A) = 1$ (if $\mathbb{P}_s^{\sigma,\tau}(A) \neq 0$).
2. *(Under event B).* We now reason under the event B and show that both the parity and the mean-payoff objectives are satisfied almost-surely. We first show that the parity objective is satisfied almost-surely. Consider an arbitrary strategy τ for player Min in $\mathcal{G}$ and a state $s \in S$.

   Parity objective almost-surely. Given the event B, the strategy is in attractor mode infinitely often. Given the strategy is in the attractor phase, the probability to reach a priority-d state within the next $|S|$ steps after the attractor mode starts is at least $x = (p_{\min})^{|S|} > 0$, where $p_{\min}$ is the minimum positive transition probability (i.e., $p_{\min} = \min\{\delta(s, a)(t) > 0 \mid s, t \in S, a \in A\}$). It follows that if the strategy is switching k times to the attractor phase, then the probability not to visit the priority-d set is at most $(1-x)^k$. The event B ensures that the strategy is in the attractor

phase infinitely often, and thus the probability that given the event B after some point a priority d state is not visited at all is $\lim_{k\to\infty}(1-x)^k = 0$. Hence given event B, the best even priority d is visited infinitely often almost-surely, ensuring that the parity objective is satisfied, that is for all $s \in S$ and all strategies τ of player Min in $\mathcal{G}$ we have $\mathbb{P}_s^{\sigma,\tau}(\mathsf{Par} \mid B) = 1$ (if $\mathbb{P}_s^{\sigma,\tau}(B) \neq 0$).
In other words, given that the positive attractor to a set T is visited infinitely often, it follows that the set T is visited infinitely often with probability 1, and we refer to this property as the *almost-sure positive attractor property*.

Mean-payoff objective almost-surely. We now prove that the mean-payoff objective is almost-surely satisfied. Given the event B, the strategy σ is in the mean-payoff phase infinitely often. Consider the finite prefixes of play $\rho = s_0 \cdots s_{j+1}$ consistent with (σ,τ) that are in the mean-payoff phase for the first time in the current round. Then by the definition of the strategy σ, every play prefix $\rho' = \rho \cdot s_{j+1} \cdots s_{j+i}$ consistent with (σ,τ) that extends ρ, for all $0 < i \leq K_j$, is in the mean-payoff phase. The sum of the rewards for all prefixes of length j is at least $-j \cdot W$ and then applying Lemma 1 we have

$$\mathbb{P}_s^{\sigma,\tau}\left(\left\{ s_0 s_1 \cdots \mid \frac{1}{j+K_j} \cdot \sum_{i=0}^{j+K_j} \mathsf{rwd}(s_i, s_{i+1}) \geq -\frac{\epsilon_j \cdot K_j + j \cdot W}{j+K_j} \right\} \mid \mathsf{Cone}(\rho)\right) \geq 1-\epsilon_j$$

By the choice of K_j (that $K_j \geq j^2 \cdot W$) and $\epsilon_j = \frac{1}{j}$, we have $-\frac{\epsilon_j \cdot K_j + j \cdot W}{j+K_j} \geq -\frac{\epsilon_j \cdot K_j}{K_j} - \frac{j \cdot W}{j^2 \cdot W} \geq -\frac{2}{j}$. Consider the function f that given a number ℓ returns the maximum number j such that $j + K_j \leq \ell$. Note that f is a non-decreasing function and as ℓ tends to ∞, also $f(\ell)$ tends to ∞. Given the event B, there are infinitely many prefixes ρ consistent with (σ,τ) that are in the mean-payoff phase for the first time in the current round. Hence we have

$$\limsup_{\ell\to\infty} \mathbb{P}_s^{\sigma,\tau}\left(\left\{ s_0 s_1 \cdots \mid \frac{1}{\ell} \cdot \sum_{i=0}^{\ell} \mathsf{rwd}(s_i, s_{i+1}) \geq -\frac{2}{f(\ell)} \right\} \mid B\right) \geq \limsup_{\ell\to\infty} 1 - \frac{1}{f(\ell)} = 1.$$

By Fatou's lemma [2] we know that for an event sequence $\mathcal{E}_\ell$ we have that $\limsup_{\ell\to\infty} \mathbb{P}(\mathcal{E}_\ell) \leq \mathbb{P}(\limsup_{\ell\to\infty} \mathcal{E}_\ell)$. Hence an application of the Fatou's lemma gives us that

$$\mathbb{P}_s^{\sigma,\tau}\left(\limsup_{\ell\to\infty} \left\{ s_0 s_1 \cdots \mid \frac{1}{\ell} \cdot \sum_{i=0}^{\ell} \mathsf{rwd}(s_i, s_{i+1}) \geq -\frac{2}{f(\ell)} \right\} \mid B\right) = 1.$$

Let $\varphi_\ell = \left\{ s_0 s_1 \cdots \mid \frac{1}{\ell} \cdot \sum_{i=0}^{\ell} \mathsf{rwd}(s_i, s_{i+1}) \geq -\frac{2}{f(\ell)} \right\}$ and $\varphi = \limsup_{\ell\to\infty} \varphi_\ell$. Consider a play $\rho = s_0 s_1 \cdots \in \varphi$. Fix $\epsilon > 0$, and consider ℓ_0 such that $\frac{2}{f(\ell_0)} \leq \epsilon$. Since $\rho \in \varphi$, there exists infinitely many $\ell \geq \ell_0$ such that $\rho \in \varphi_\ell$, and hence for infinitely many ℓ we have $\frac{1}{\ell} \cdot \sum_{i=1}^{\ell-1} \mathsf{rwd}(s_i, s_{i+1}) \geq -\epsilon$. Hence $\limsup_{\ell\to\infty} \frac{1}{\ell} \cdot \sum_{i=1}^{\ell-1} \mathsf{rwd}(s_i, s_{i+1}) \geq -\epsilon$. Since this holds for all $\epsilon > 0$, it follows that $\limsup_{\ell\to\infty} \frac{1}{\ell} \cdot \sum_{i=1}^{\ell-1} \mathsf{rwd}(s_i, s_{i+1}) \geq 0$. In other words, we have $\varphi \subseteq \mathsf{Mean}$ and hence for all $s \in S$ and all strategies τ of player Min in $\mathcal{G}$ we have $\mathbb{P}_s^{\sigma,\tau}(\mathsf{Mean} \mid B) = 1$ (if $\mathbb{P}_s^{\sigma,\tau}(B) \neq 0$).

Thus given either event A or B, the mean-payoff parity objective is satisfied almost-surely. Note that if one of the event has probability 0, then the other has probability 1. It follows that the mean-payoff parity objective is satisfied almost-surely. This concludes one direction of the proof that if the conditions of the lemma are satisfied, then almost-sure winning for $\mathsf{Mean} \cap \mathsf{Par}$ is ensured with probability 1.

We now prove the converse. Consider a game $\mathcal{G}$ such that all states in its state space S are almost-sure winning for the objective $\mathsf{Mean} \cap \mathsf{Par}$ for player Max. First, observe that since $\mathsf{Mean} \cap \mathsf{Par} \subseteq \mathsf{Mean}$, almost-sure winning for $\mathsf{Mean} \cap \mathsf{Par}$ implies almost-sure winning for Mean. This implies the first condition. Second, observe that Y is a trap for player Max. If player Max does not have an almost-sure winning strategy for a non-empty set $Z \subseteq Y$ in the subgame $\mathcal{G}[Y]$, then player Max does not have an almost-sure winning strategy from Z in $\mathcal{G}$, which contradicts that all states in $\mathcal{G}$ are almost-sure winning. This proves the second condition of the lemma and completes the proof. $\square$

Lemma 3. *Let $\mathcal{G}$ be a stochastic mean-payoff parity game with state space S, and such that the largest priority d in $\mathcal{G}$ is odd. All states in $\mathcal{G}$ are almost-sure winning for the objective $\mathsf{Mean} \cap \mathsf{Par}$ if and only if there exists a partition $\{Z_i\}_{1 \le i \le k}$ of S and non-empty sets R_i, U_i for $i = 1, \ldots, k$, and U_{k+1} such that $U_1 = S$ and for all $1 \le i \le k$: (1) $R_i \subseteq U_i \setminus U_i(d)$ is a trap for Min in $\mathcal{G}[U_i]$, and all states in R_i are almost-sure winning for the objective $\mathsf{Mean} \cap \mathsf{Par}$ in $\mathcal{G}[U_i]$; (2) $Z_i = \mathsf{Attr}_{\mathsf{Max}}(R_i, \mathcal{G}[U_i])$; and (3) $U_{i+1} = U_i \setminus Z_i$.*

Lemma 3 presents a characterization of the certificate for almost-sure winning when the largest priority is odd. The key correctness argument uses the almost-sure positive attractor property to show that the event that from some point on only states in R_i are visited for some i has probability 1. From the above fact and the almost-sure winning strategies in R_i we obtain an almost-sure winning strategy in $\mathcal{G}$.

We remark that it follows from our proofs that the infinite-memory required by the strategies can be captured in terms of counter-based strategies that keep track of the number of steps that certain positional strategies need to be played.

4 Algorithm

In this section we present an algorithm for the almost-sure winning problem. Let $\mathcal{G}$ be a stochastic mean-payoff parity game with largest priority d. Our algorithm computes the set R of almost-sure winning states for Max, by iterations that, from the state space S of $\mathcal{G}$ remove positive winning states of player Min. When a fixpoint is obtained, we show that it satisfies the characterization of Lemma 2 and Lemma 3, hence it is the almost-sure winning set. Starting with $R = S$, the algorithm considers two cases:

(a) If d is even: First, compute the almost-sure wining region U for the Mean objective in $\mathcal{G}[R]$. Compute the positive attractor X for player Max to the set of states with priority d in U, and let Y be the complement. Recursively compute the almost-sure winning region R' in $\mathcal{G}[Y]$ for the mean-payoff parity objective, and iterate (until $R' = Y$) in the subgame induced by the complement $U \setminus Z$ of the player-Min positive attractor $Z = \mathsf{Attr}_{\mathsf{Min}}(Y \setminus R', \mathcal{G}[U])$ (i.e., removing some positive winning states for player Min).

(b) If d is odd: In each iteration of the main loop, the algorithm computes a set of positive winning states for player Min as the positive attractor (for Min) to the set U computed in the inner loop. The inner loop computes in R' the almost-sure winning states of player Max in the subgame induced by the complement Y of player-Min positive attractor to priority d, using a recursive call. The positive attractor for Max to R' is removed, and the next iteration starts (if $R' \neq \emptyset$) with a strictly smaller state space U. The main loop terminates when there is nothing to remove ($U = \emptyset$).

Correctness and termination. The correctness and termination of our algorithm (which we refer to as AlgStMPP, algorithm for stochastic mean-payoff parity games) is established using an argument by induction on the depth of the recursive calls, which are always invoked with games that have at least one less priority than the current game, and using Lemma 2 and Lemma 3.

The complexity of AlgStMPP is exponential in the number of priorities in the game, like the basic algorithm for parity games [28]. The key differences to the basic algorithm for parity games are as follows: (i) in our algorithm there is an extra nested loop when the maximum priority is odd; and (ii) in addition to the basic attractor computation for parity games we also need to compute the almost-sure winning set for stochastic mean-payoff games.

Theorem 1. *Given a stochastic mean-payoff parity game $\mathcal{G}$ with n states, probabilistic transition function δ, priorities in $\{0, 1, \ldots, d-1\}$, and largest absolute reward W,* AlgStMPP *computes the almost-sure winning region of $\mathcal{G}$ in time $O(d \cdot n^{2d} \cdot \mathsf{MeanGame}(n, |\delta|, W))$ where $\mathsf{MeanGame}(n, |\delta|, W)$ is the time complexity of solving the almost-sure winning problem for stochastic games with only a mean-payoff objective.*

Note that $\mathsf{MeanGame}(n, |\delta|, W) \in |A|^n \cdot Poly(n, |\delta|, W)$ by simply enumerating over all positional strategies and then solving in polynomial time the MDP obtained by fixing the positional strategy.

5 Computational Complexity

In this section we establish the NP ∩ coNP complexity bound for the almost-sure winning problem.

The NP **Membership.** Although infinite-memory strategies are necessary for player Max to win mean-payoff parity games almost surely [14], we show that the almost-sure winning problem can be solved in NP by guessing a polynomial-size decomposition of the state space along with positional strategies that allow to construct an almost-sure winning strategy, possibly with infinite memory. The polynomial certificate is obtained from the characterization of Lemma 2 and Lemma 3; and the verification procedure requires solving MDPs with mean-payoff parity objectives, which can be done in polynomial time [11].

Lemma 4. *The almost-sure winning problem for stochastic mean-payoff parity games is in* NP.

The coNP **Membership.** We show that positional strategies are sufficient for player Min to win positively in stochastic mean-payoff parity games. Using the fact that AlgStMPP maintains in variable R an over-approximation of the almost-sure winning set for player Max, we construct a positional strategy for player Min from all states that are removed from R by the algorithm.

Lemma 5. *To win positively in stochastic mean-payoff parity games, positional strategies are sufficient for player* Min.

We then show how to use the positional strategy for positive winning to obtain a positional strategy for almost-sure winning for player Min. By Remark 1 it follows that positional optimal strategies exist for player Min. Lemma 4, the existence of positional optimal strategies for player Min, and the fact that MDPs with mean-payoff parity objectives can be solved in polynomial time [11], gives us the following result.

Theorem 2. *The following assertions hold: (1) Positional optimal strategies exist for player* Min *in stochastic mean-payoff parity games (2) The almost-sure winning and the value-strategy problem for stochastic mean-payoff parity games can be decided in* NP $\cap$ coNP.

Remark 2. The complexity result of Theorem 2 matches the best known complexity for stochastic mean-payoff games [22], stochastic parity games [15] (also see [1] for relationship of stochastic mean-payoff and stochastic parity games), and (non-stochastic) mean-payoff parity games [12].

Concluding remarks. In this work we studied the computational and strategy complexity of the value-strategy problem for $2\frac{1}{2}$-player mean-payoff parity games. In addition we presented an algorithm for computing the almost-sure winning states which requires the computation of the almost-sure winning states for $2\frac{1}{2}$-player mean-payoff games. Improved algorithmic solutions for the computation of the almost-sure winning states in $2\frac{1}{2}$-player mean-payoff games is an interesting question. Our algorithm for almost-sure winning and the general technique mentioned in Remark 1 for $2\frac{1}{2}$-player games with tail objectives provide an exponential-time algorithm for the value-strategy problem. Whether more specialized algorithms (such as strategy-iteration algorithms) can be developed for the value-strategy problem in $2\frac{1}{2}$-player mean-payoff parity games is another interesting algorithmic question.

References

1. D. Andersson and P. B. Miltersen. The complexity of solving stochastic games on graphs. In *Proc. of ISAAC*, LNCS 5878, pages 112–121. Springer, 2009.
2. P. Billingsley, editor. *Probability and Measure*. Wiley-Interscience, 1995.
3. R. Bloem, K. Chatterjee, T. A. Henzinger, and B. Jobstmann. Better quality in synthesis through quantitative objectives. In *Proc. of CAV*, LNCS 5643, pages 140–156. Springer, 2009.
4. E. Boros, K. M. Elbassioni, M. Fouz, V. Gurvich, K. Makino, and B. Manthey. Stochastic mean payoff games: Smoothed analysis and approximation schemes. In *ICALP (track A)*, pages 147–158, 2011.

5. P. Bouyer, U. Fahrenberg, K. G. Larsen, N. Markey, and J. Srba. Infinite runs in weighted timed automata with energy constraints. In *Proc. of FORMATS*, LNCS 5215, pages 33–47. Springer, 2008.
6. P. Bouyer, N. Markey, J. Olschewski, and M. Ummels. Measuring permissiveness in parity games: Mean-payoff parity games revisited. In *Proc. of ATVA*, LNCS 6996, pages 135–149. Springer, 2011.
7. J. R. Büchi and L. H. Landweber. Solving sequential conditions by finite-state strategies. *Transactions of the AMS*, 138:295–311, 1969.
8. P. Cerný, K. Chatterjee, T. A. Henzinger, A. Radhakrishna, and R. Singh. Quantitative synthesis for concurrent programs. In *Proc. of CAV*, LNCS 6806, pages 243–259. Springer, 2011.
9. A. Chakrabarti, L. de Alfaro, T. A. Henzinger, and M. Stoelinga. Resource interfaces. In *Proc. of EMSOFT*, LNCS 2855, pages 117–133. Springer, 2003.
10. K. Chatterjee and L. Doyen. Energy parity games. In *Proc. of ICALP: Automata, Languages and Programming (B)*, LNCS 6199, pages 599–610. Springer, 2010.
11. K. Chatterjee and L. Doyen. Energy and mean-payoff parity Markov decision processes. In *Proc. of MFCS*, LNCS 6907, pages 206–218. Springer, 2011.
12. K. Chatterjee and L. Doyen. Energy parity games. *TCS*, 458(2):49–60, 2012.
13. K. Chatterjee, T. A. Henzinger, and F. Horn. Stochastic games with finitary objectives. In *Proc. of MFCS*, LNCS 5734, pages 34–54. Springer, 2009.
14. K. Chatterjee, T. A. Henzinger, and M. Jurdziński. Mean-payoff parity games. In *Proc. of LICS*, pages 178–187. IEEE Computer Society, 2005.
15. K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Quantitative stochastic parity games. In *Proc. of SODA*, pages 121–130. SIAM, 2004.
16. A. Condon. The complexity of stochastic games. *Inf. Comput.*, 96(2):203–224, 1992.
17. J. A. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, 1997.
18. H. Gimbert and F. Horn. Solving simple stochastic tail games. In *Proc. of SODA*, pages 847–862. SIAM, 2010.
19. H. Gimbert, Y. Oualhadj, and S. Paul. Computing optimal strategies for Markov decision processes with parity and positive-average conditions. Technical report, LaBRI, Université de Bordeaux II, 2011.
20. V. Gurvich, A. Karzanov, and L. Khachivan. Cyclic games and an algorithm to find minimax cycle means in directed graphs. *USSR Computational Mathematics and Mathematical Physics*, 28(5):85–91, 1988.
21. M. Jurdziński. Deciding the winner in parity games is in UP $\cap$ co-UP. *Information Processing Letters*, 68(3):119–124, 1998.
22. T. A. Liggett and S. A. Lippman. Stochastic games with perfect information and time average payoff. *Siam Review*, 11:604–607, 1969.
23. A. Pnueli and R. Rosner. On the synthesis of a reactive module. In *Proc. of POPL*, pages 179–190. ACM Press, 1989.
24. M. L. Puterman. *Markov Decision Processes*. John Wiley and Sons, 1994.
25. T. E. S. Raghavan and J. A. Filar. Algorithms for stochastic games — a survey. *ZOR — Methods and Models of Operations Research*, 35:437–472, 1991.
26. P. J. Ramadge and W. M. Wonham. Supervisory control of a class of discrete-event processes. *SIAM Journal of Control and Optimization*, 25(1):206–230, 1987.
27. W. Thomas. Languages, automata, and logic. In *Handbook of Formal Languages*, volume 3, Beyond Words, chapter 7, pages 389–455. Springer, 1997.
28. W. Zielonka. Infinite games on finitely coloured graphs with applications to automata on infinite trees. In *Theoretical Computer Science*, volume 200(1-2), pages 135–183, 1998.
29. U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *TCS*, 158(1&2):343–359, 1996.