Approximating the Value of Energy-Parity Objectives in Simple Stochastic Games

Mohan Dantam
School of Informatics, University of Edinburgh, UK

Richard Mayr
School of Informatics, University of Edinburgh, UK

--- **Abstract** ---

We consider simple stochastic games $\mathcal{G}$ with energy-parity objectives, a combination of quantitative rewards with a qualitative parity condition. The Maximizer tries to avoid running out of energy while simultaneously satisfying a parity condition.

We present an algorithm to approximate the value of a given configuration in 2-NEXPTIME. Moreover, ε-optimal strategies for either player require at most $\mathcal{O}\left(2\text{-EXP}\left(|\mathcal{G}|\right) \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ memory modes.

1 Introduction

Background. *Simple stochastic games* (SSGs) are 2-player turn-based perfect information stochastic games played on finite graphs. They are also called *competitive Markov decision processes* [20], or $2\frac{1}{2}$-*player games* [13, 12]. Introduced by Shapley [36] in 1953, they have since played a central role in the solution of many problems, e.g., synthesis of reactive systems [35, 34] and formal specification and verification [17, 18, 1]. Every state either belongs to one of the players (Maximizer or Minimizer) or is a random state. In each round of the game the player who owns the current state gets to choose the successor state along the game graph. For random states the successor is chosen according to a predefined distribution. Given a start state and strategies of Maximizer and Minimizer, this yields a distribution over induced infinite plays. We consider objectives $\mathtt{O}$ that are measurable subsets of the set of possible plays, and the players try to maximize (resp. minimize) the probability of $\mathtt{O}$.

Many different objectives for SSGs have been studied in the literature. Here we focus on parity, mean-payoff and energy objectives. We assign numeric rewards to transitions and priorities (aka colors), encoded by bounded non-negative numbers, to states. A play satisfies the (min-even) *parity objective* iff the minimal priority that appears infinitely often in a play is even. It subsumes all ω-regular objectives, and in particular safety, liveness, fairness, etc. On finite SSGs, the parity objective can be seen as a special case of the *mean-payoff objective* which requires the limit average reward per transition along a play to be positive (or non-negative). Mean-payoff objectives in SSGs go back to a 1957 paper by Gillette [21] and have been widely studied, due to their relevance for efficient control. The *energy objective* [6] requires that the accumulated reward at any time in a play stays above some finite threshold. The intuition is that a controlled system has some finite initial energy level that must never become depleted. Since the accumulated reward is not bounded a-priori, this essentially turns a finite-state game into an infinite-state one.

Energy-parity. We consider SSGs with *energy-parity* objectives, where plays need to satisfy both an energy and a parity objective. The parity objective specifies functional correctness, while the energy condition can encode efficiency or risk considerations, e.g., the system should not run out of energy since manually recharging would be costly or risky.

Previous work. Much work on combined objectives for stochastic systems is restricted to Markov decision processes (MDPs) [8, 9, 4, 28].

For (stochastic) games, the computational complexity of single objectives is often in NP ∩ coNP, e.g., for parity or mean-payoff objectives [25]. Multi-objective games can be harder, e.g., satisfying *two different* parity objectives leads to coNP completeness [11].

Stochastic mean-payoff parity games can be solved in NP ∩ coNP [10]. However, this does not imply a solution for stochastic energy-parity games, since, unlike in the non-stochastic case [7], there is no known reduction from energy-parity to mean-payoff parity in stochastic games. The reduction in [7] relies on the fact that Maximizer has a winning *finite-memory* strategy for energy-parity, which does not generally hold for stochastic games, or even MDPs [28]. For the same reason, the direct reduction from stochastic energy-parity to ordinary energy games proposed in [8, 9] does not work for general energy-parity but only for energy-Büchi; cf. [28].

Non-stochastic energy-parity games can be solved in NP ∩ coNP (and even in pseudo-quasi-polynomial time [16]) and Maximizer strategies require only finite (but exponential) memory [7].

Stochastic energy-parity games have been studied in [29], where it was shown that the almost-sure problem is decidable and in NP ∩ coNP. That is, given an initial configuration (control-state plus current energy level), does Maximizer have a strategy to ensure that energy-parity is satisfied with probability 1 against any Minimizer strategy? Unlike in many single-objective games, such an almost-surely winning Maximizer strategy (if it exists) requires infinite memory in general. This holds even in MDPs and for energy-coBüchi objectives [28].

However, [29] did not address quantitative questions about energy-parity objectives, such as computing/approximating the value of a given configuration, or the decidability of exact questions like "Is the value of this configuration $\geq k$?" for some constant k (e.g., $k = 1/2$).

The decidability of the latter type of exact question about the energy-parity value is open, but there are strong indications that it is very hard. In fact, even simpler subproblems are already at least as hard as the *positivity problem for linear recurrence sequences*, which in turn is at least as hard as the *Skolem problem* [19]. (The decidability of these problems has been open for decades; see [30] for an overview.) Given an SSG with an energy-parity objective, suppose we remove the parity condition (assume it is always true) and also suppose that Maximizer is passive (does not get to make any decisions). Then we obtain an MDP where the only active player (the Minimizer in the SSG) has a *termination objective*, i.e., to reach a configuration where the energy level is ≤ 0. Exact questions about the value of the termination objective in MDPs are already at least as hard as the positivity problem [31, Section 5.2.3] (see also [32, 33]). Thus exact questions about the energy-parity value in SSGs are also at least as hard as the positivity problem.

Our contributions. Since exact questions about the energy-parity value in SSGs are positivity-hard, we consider the problem of computing approximations of the value. We present an algorithm that, given an SSG $\mathcal{G}$ and error ε, computes ε-close approximations of the energy-parity value of any given configuration in 2-NEXPTIME. Moreover, we show that

ε-optimal Maximizer (resp. Minimizer) strategies can be chosen as deterministic and using only finite memory with $\mathcal{O}\left(2\text{-}\mathsf{EXP}\left(|\mathcal{G}|\right) \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ memory modes. One can understand the idea as a constructive upper bound on the accuracy with which the players need to remember the current energy level in the game. (This is in contrast to the result in [28] that almost-surely winning Maximizer strategies require infinite memory in general.) Once the upper bound on Maximizer's memory for ε-optimal strategies is established, one might attempt a reduction from energy-parity to mean-payoff parity, along similar lines as for non-stochastic games in [7]. However, instead we use a more direct reduction from energy-parity to parity in a derived SSG for our approximation algorithm.

2 Preliminaries

A *probability distribution* over a countable set S is a function $f \colon S \to [0,1]$ with $\sum_{s \in S} f(s) = 1$. $\mathsf{supp}(f) \stackrel{\text{def}}{=} \{s \mid f(s) > 0\}$ denotes the support of f and $\mathcal{D}(S)$ is the set of all probability distributions over S. Given an alphabet Σ, let Σ^ω and Σ^* (Σ^+) denote the set of infinite and finite (non-empty) sequences over Σ, respectively. Elements of Σ^ω or Σ^* are called words.

Games, MDPs and Markov chains. A *Simple Stochastic Game* (SSG) is a finite-state 2-player turn-based perfect-information stochastic game $\mathcal{G} = (S, (S_\square, S_\diamond, S_\bigcirc), E, P)$ where the finite set of states S is partitioned into the states $S_\square$ of the player $\square$ (*Maximizer*), states $S_\diamond$ of player $\diamond$ (*Minimizer*), and chance vertices (aka random states) $S_\bigcirc$. Let $E \subseteq S \times S$ be the transition relation. We write $s \longrightarrow s'$ if $(s, s') \in E$ and assume that $\mathsf{Succ}(s) \stackrel{\text{def}}{=} \{s' \mid sEs'\} \neq \emptyset$ for every state s. The *probability function* P assigns each random state $s \in S_\bigcirc$ a distribution over its successor states, i.e., $P(s) \in \mathcal{D}(\mathsf{Succ}(s))$. For ease of presentation, we extend the domain of P to $S^* S_\bigcirc$ by $P(\rho s) \stackrel{\text{def}}{=} P(s)$ for all $\rho s \in S^+ S_\bigcirc$. An *MDP* is a game where one of the two players does not control any states. An MDP is *maximizing* (resp. *minimizing*) iff $S_\diamond = \emptyset$ (resp. $S_\square = \emptyset$). A *Markov chain* is a game with only random states, i.e., $S_\square = S_\diamond = \emptyset$.

Strategies. A *play* is an infinite sequence $s_0 s_1 \ldots \in S^\omega$ such that $s_i \longrightarrow s_{i+1}$ for all $i \geq 0$. A *path* is a finite prefix of a play. Let $Plays(\mathcal{G}) \stackrel{\text{def}}{=} \left\{\rho = (q_i)_{i \in \mathbb{N}} \mid q_i \longrightarrow q_{i+1}\right\}$ denote the set of all possible plays. A strategy of the player $\square$ ($\diamond$) is a function $\sigma : S^* S_\square \to \mathcal{D}(S)$ ($\pi : S^* S_\diamond \to \mathcal{D}(S)$) that assigns to every path $ws \in S^* S_\square$ ($\in S^* S_\diamond$) a probability distribution over the successors of s. If these distributions are always Dirac then the strategy is called *deterministic* (aka pure), otherwise it is called *randomized* (aka mixed). The set of all strategies of player $\square$ and $\diamond$ in $\mathcal{G}$ is denoted by $\Sigma_\mathcal{G}$ and $\Pi_\mathcal{G}$, respectively. A play/path $s_0 s_1 \ldots$ is compatible with a pair of strategies (σ, π) if $s_{i+1} \in \mathsf{supp}(\sigma(s_0 \ldots s_i))$ whenever $s_i \in S_\square$ and $s_{i+1} \in \mathsf{supp}(\pi(s_0 \ldots s_i))$ whenever $s_i \in S_\diamond$.

Finite-memory deterministic (FD) strategies are a subclass of strategies described by deterministic transducers $\mathsf{T} = (\mathsf{M}, \mathsf{m}_0, \mathsf{upd}, \mathsf{nxt})$ where M is a finite set of memory modes with initial mode m_0, $\mathsf{upd} : \mathsf{M} \times E \mapsto \mathsf{M}$ updates the memory mode upon observing a transition and $\mathsf{nxt} : \mathsf{M} \times S_\bigcirc \mapsto S$ chooses the successor state based on the current memory mode and state. FD strategies without memory ($|\mathsf{M}| = 1$) are called memoryless deterministic (MD). For deterministic strategies, there is no difference between public memory (observable by the other player) and private memory.

Measure. A game $\mathcal{G}$ with initial state s_0 and strategies (σ, π) yields a probability space $(s_0 S^\omega, \mathcal{F}_{s_0}, \mathcal{P}^\mathcal{G}_{\sigma, \pi, s_0})$ where $\mathcal{F}_{s_0}$ is the σ-algebra generated by the cylinder sets $s_0 s_1 \ldots s_n S^\omega$ for $n \geq 0$. The probability measure $\mathcal{P}^\mathcal{G}_{\sigma, \pi, s_0}$ is first defined on the cylinder sets. For

$\rho = s_0 \ldots s_n$, let $\mathcal{P}^{\mathcal{G}}_{\sigma, \pi, s_0}(\rho) \stackrel{\text{def}}{=} 0$ if ρ is not compatible with σ, π and otherwise $\mathcal{P}^{\mathcal{G}}_{\sigma, \pi, s_0}(\rho S^\omega) \stackrel{\text{def}}{=} \prod_{i=0}^{n-1} \tau(s_0 \ldots s_i)(s_{i+1})$ where τ is σ or π or P depending on whether $s_i \in S_\square$ or $S_\diamond$ or $S_\bigcirc$, respectively. By Carathéodory's extension theorem [2], this defines a unique probability measure on the σ-algebra.

Objectives and Payoff functions. General objectives are defined by real-valued measurable functions. However, we only consider indicator functions of measurable sets. Hence our objectives can be described by measurable subsets $\mathsf{O} \subseteq S^\omega$ of plays. The payoff, under strategies (σ, π), is the probability that plays belong to O.

We use the syntax and semantics of the LTL operators [14] F (eventually), G (always) and X (next) to specify some conditions on plays.

Reachability & Safety. A reachability objective is defined by a set of target states $T \subseteq S$. A play $\rho = s_0 s_1 \ldots$ belongs to $\mathsf{F}\,T$ iff $\exists i \in \mathbb{N}\, s_i \in T$. Similarly, ρ belongs to $\mathsf{F}^{\leq n} T$ (resp. $\mathsf{F}^{\geq n} T$) iff $\exists i \leq n$ (resp. $i \geq n$) such that $s_i \in T$. Dually, the safety objective $\mathsf{G}\,T$ consists of all plays which never leave T. We have $\mathsf{G}\,T = \neg\mathsf{F}\neg T$.

Parity. A parity objective is defined via bounded function $Col : S \to \mathbb{N}$ that assigns non-negative priorities (aka colors) to states. Given an infinite play $\rho = s_0 s_1 \ldots$, let $\mathrm{Inf}(\rho)$ denote the set of numbers that occur infinitely often in the sequence $Col(s_0) Col(s_1) \ldots$. A play ρ satisfies *even parity* w.r.t. Col iff the minimum of $\mathrm{Inf}(\rho)$ is even. Otherwise, ρ satisfies *odd parity*. The objective even parity is denoted by $\mathtt{EPAR}(Col)$ and odd parity is denoted by $\mathtt{OPAR}(Col)$. Most of the time, we implicitly assume that the coloring function is known and just write $\mathtt{EPAR}$ and $\mathtt{OPAR}$. Observe that, given any coloring Col, we have $\overline{\mathtt{EPAR}} = \mathtt{OPAR}$ and $\mathtt{OPAR}(Col) = \mathtt{EPAR}(Col + 1)$ where $Col + 1$ is the function which adds 1 to the color of every state. This justifies to consider only one of the even/odd parity objectives, but, for the sake of clarity, we distinguish these objectives wherever necessary.

Energy/Reward/Counter based objectives. Let $r : E \to \{-R, \ldots, 0, \ldots, R\}$ be a bounded function that assigns weights to transitions. Depending on context, the sum of these weights in a path can be viewed as energy, cost/reward or a counter. If $s \longrightarrow s'$ and $r((s, s')) = c$, we write $s \stackrel{c}{\longrightarrow} s'$. Let $\rho = s_0 \stackrel{c_0}{\longrightarrow} s_1 \stackrel{c_1}{\longrightarrow} \ldots$ be a play. We say that ρ satisfies

1. the *k-energy* objective $\mathsf{EN}(k)$ iff $\left(k + \sum_{i=0}^{n-1} c_i\right) > 0$ for all $n \geq 0$.

2. the *l-storage condition* if $l + \sum_{i=m}^{n-1} c_i \geq 0$ holds for every infix $s_m \stackrel{c_m}{\longrightarrow} s_{m+1} \ldots s_n$ of the play. Let $\mathsf{ST}(k, l)$ denote the set of plays that satisfy both the k-energy and the l-storage condition. Let $\mathsf{ST}(k) \stackrel{\text{def}}{=} \bigcup_l \mathsf{ST}(k, l)$. Clearly, $\mathsf{ST}(k) \subseteq \mathsf{EN}(k)$.

3. *k-Termination* $\mathtt{Term}(k)$ iff there exists $n \geq 0$ such that $\left(k + \sum_{i=0}^{n-1} c_i\right) \leq 0$.

4. *Limit* objective $\mathtt{LimInf}(\triangleright z)$ iff $\left(\liminf_{n \to \infty} \sum_{i=0}^{n-1} c_i\right) \triangleright z$ for $\triangleright \in \{<, \leq, =, \geq, >\}$ and $z \in \mathbb{R} \cup \{\infty, -\infty\}$ and similarly for $\mathtt{LimSup}(\triangleright z)$.

5. *Mean payoff* $\mathtt{MP}(\triangleright c)$ for some constant $c \in \mathbb{R}$ iff $\left(\liminf_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} c_i\right) \triangleright c$.

Observe that the objectives k-energy and k-termination are mutually exclusive and cover all of the plays. A different way to consider these objectives is to encode the energy level (the sum of the transition weights so far) into the state space and then consider the obtained infinite-state game with safety/reachability objective, respectively.

An objective O is called *shift-invariant* iff for all finite paths ρ and plays $\rho' \in S^\omega$, we have $\rho\rho' \in \mathsf{O} \iff \rho' \in \mathsf{O}$. Parity and mean payoff objectives are shift-invariant, but energy and termination objectives are not. Objective O is called *submixing* iff for all sequences of finite non-empty words $u_0, v_0, u_1, v_1 \ldots$ we have $u_0 v_0 u_1 v_1 \ldots \in \mathsf{O} \implies ((u_0 u_1 \ldots \in \mathsf{O}) \vee (v_0 v_1 \ldots \in \mathsf{O}))$.

Determinacy. Given an objective $\mathtt{O}$ and a game $\mathcal{G}$, state s has value (w.r.t to $\mathtt{O}$) iff

$$\sup_{\sigma \in \Sigma_{\mathcal{G}}} \inf_{\pi \in \Pi_{\mathcal{G}}} \mathcal{P}^{\mathcal{G}}_{\sigma,\pi,s}(\mathtt{O}) = \inf_{\pi \in \Pi_{\mathcal{G}}} \sup_{\sigma \in \Sigma_{\mathcal{G}}} \mathcal{P}^{\mathcal{G}}_{\sigma,\pi,s}(\mathtt{O}).$$

If s has value then $\mathtt{val}^{\mathcal{G}}_{\mathtt{O}}(s)$ denotes the value of s defined by the above equality. A game with an objective is called *weakly determined* if every state has value. Stochastic games with Borel objectives are weakly determined [26, 27]. Our objectives above are Borel, hence any boolean combination of them is also weakly determined. For $\varepsilon > 0$ and state s, a strategy
1. $\sigma \in \Sigma_{\mathcal{G}}$ is ε-optimal (maximizing) iff $\mathcal{P}^{\mathcal{G}}_{\sigma,\pi,s}(\mathtt{O}) \geq \mathtt{val}^{\mathcal{G}}_{\mathtt{O}}(s) - \varepsilon$ for all $\pi \in \Pi_{\mathcal{G}}$.
2. $\pi \in \Pi_{\mathcal{G}}$ is ε-optimal (minimizing) iff $\mathcal{P}^{\mathcal{G}}_{\sigma,\pi,s}(\mathtt{O}) \leq \mathtt{val}^{\mathcal{G}}_{\mathtt{O}}(s) + \varepsilon$ for all $\sigma \in \Sigma_{\mathcal{G}}$.
A 0-optimal strategy is called *optimal*. An MD strategy is called *uniformly ε-optimal* (resp. uniformly optimal) if it is so from every start state. An optimal strategy for player $\square$ from state s is *almost surely* winning if $\mathtt{val}^{\mathcal{G}}_{\mathtt{O}}(s) = 1$. By $\mathsf{AS}(\mathtt{O})$ we denote the set of states that have an almost surely winning strategy for objective $\mathtt{O}$. For ease of presentation, we drop subscripts and superscripts wherever possible if they are clear from the context.

Energy-parity. We are concerned with approximating the value for the combined energy-parity objective $\mathtt{EN}(k) \cap \mathtt{EPAR}$ and building ε-optimal strategies.

In our constructions we use some auxiliary objectives. Following [29], these are defined as $\mathtt{Gain} \overset{\mathrm{def}}{=} \mathtt{LimInf}(> -\infty) \cap \mathtt{EPAR}$ and $\mathtt{Loss} \overset{\mathrm{def}}{=} \overline{\mathtt{Gain}} = \mathtt{LimInf}(= -\infty) \cup \mathtt{OPAR}$.

▶ **Remark 1.** For finite-state SSGs and the following objectives there exist optimal MD strategies for both players. Moreover, if the SSG is just a maximizing MDP then the set of states that are almost surely winning for Maximizer can be computed in polynomial time.
1. $\mathsf{F}\,T$ [15]
2. $\mathtt{LimInf}(\rhd - \infty)$, $\mathtt{LimInf}(\rhd\infty)$, $\mathtt{LimSup}(\rhd - \infty)$, $\mathtt{LimSup}(\rhd\infty)$, $\mathtt{MP}(> 0)$ [5, Prop. 1]
3. $\mathtt{EPAR}$ [37]

3 The Main Result

The following theorem states our main result.

▶ **Theorem 2.** *Let $\mathcal{G} = (S, (S_{\square}, S_{\diamond}, S_{\bigcirc}), E, P)$ be an SSG with transition rewards in unary assigned by function r and colors assigned to states by function Col. For every state $s \in S$, initial energy level $i \geq 0$ and error margin $\varepsilon > 0$, one can compute*
1. *a rational number v' such that $0 \leq v' - \mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}(s) \leq \varepsilon$ in 2-NEXPTIME.* [1]
2. *ε-optimal FD strategies σ_{ε} and π_{ε} for Maximizer and Minimizer, resp., in 2-NEXPTIME. These strategies use $\mathcal{O}\left(2\text{-}\mathsf{EXP}\left(|\mathcal{G}|\right) \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ memory modes.*
For rewards in binary, the bounds above increase by one exponential.

We outline the main steps of the proof; details in the following sections. We begin with the observation that $\mathtt{EN}(i) \subseteq \mathtt{EN}(j)$ for $i \leq j$, and thus for all states s we have $\mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}(s) \leq \mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(j) \cap \mathtt{EPAR}}(s) \leq 1$. So $\lim_{n\to\infty} \mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(n) \cap \mathtt{EPAR}}(s)$ exists. We define

$$\mathtt{Lval}^{\mathcal{G}}(s) \overset{\mathrm{def}}{=} \lim_{n\to\infty} \mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(n) \cap \mathtt{EPAR}}(s). \tag{1}$$

---

[1] We write "computing a number v' in 2-NEXPTIME" as a shorthand for the property that questions like $v' \leq c$ for constants c are decidable in 2-NEXPTIME.

We will see that $\text{Lval}^{\mathcal{G}}(s)$ and $\text{val}^{\mathcal{G}}_{\text{Gain}}(s)$ are in fact equal (a consequence of Lemma 9) and $\text{val}^{\mathcal{G}}_{\text{Gain}}(s)$ can be computed in nondeterministic polynomial time (Theorem 5). Intuitively, for high energy levels, the precise energy level does not matter much for the value.

The main steps of the approximation algorithm are as follows.

1. Compute FD strategies $\sigma^*(s)$ that are optimal maximizing for the objective Gain starting from state s in $\mathcal{G}$. Compute an MD strategy π^* that is uniformly optimal minimizing for the objective Gain. Compute the value $\text{val}^{\mathcal{G}}_{\text{Gain}}(s)$ for every $s \in S$. See Section 4.

2. Compute a natural number N such that for all $s \in S$ and all $i \geq N$ we have

$$0 \leq \text{val}^{\mathcal{G}}_{\text{Gain}}(s) - \text{val}^{\mathcal{G}}_{\text{EN}(i) \cap \text{EPAR}}(s) \leq \varepsilon.$$

   N will be doubly exponential. See Section 5.

3. Consider the finite-state parity game $\mathcal{G}'$ derived from $\mathcal{G}$ by encoding the energy level up-to N into the states, i.e., the states of $\mathcal{G}'$ are of the form (s, k) for $s \in S$ and $0 \leq k \leq N$, and colors are inherited from s. Moreover, we add gadgets that ensure that states (s, N) at the upper end win with probability $\text{val}^{\mathcal{G}}_{\text{Gain}}(s)$ and states $(s, 0)$ at the lower end lose. By the previous item, $\text{val}^{\mathcal{G}}_{\text{Gain}}(s)$ is ε-close to $\text{val}^{\mathcal{G}}_{\text{EN}(N) \cap \text{EPAR}}(s)$. Thus, for $k < N$ we can ε-approximate the value $v = \text{val}^{\mathcal{G}}_{\text{EN}(k) \cap \text{EPAR}}(s)$ by $v' \stackrel{\text{def}}{=} \text{val}^{\mathcal{G}'}_{\text{EPAR}}((s, k))$. If $k \geq N$ we can ε-approximate v by $v' \stackrel{\text{def}}{=} \text{val}^{\mathcal{G}}_{\text{Gain}}(s)$.

   Moreover, we obtain ε-optimal FD strategies σ_ε for Maximizer (resp. π_ε for Minimizer) for $\text{EN}(k) \cap \text{EPAR}$ in $\mathcal{G}$. Let $\hat{\sigma}$ (resp. $\hat{\pi}$) be optimal MD strategies for Maximizer (resp. Minimizer) for the objective EPAR in $\mathcal{G}'$. Then σ_ε plays as follows. While the current energy level j (k plus the sum of the rewards so far) stays $< N$, then, at any state s', play like $\hat{\sigma}$ at state (s', j) in $\mathcal{G}'$. Once the energy level reaches a value $\geq N$ at some state s' for the first time, then play like $\sigma^*(s')$ forever. Similarly, π_ε plays as follows. While the current energy level j (k plus the sum of the rewards so far) stays $< N$, then, at any state s', play like $\hat{\pi}$ at state (s', j) in $\mathcal{G}'$. Once the energy level reaches a value $\geq N$ (at any state) for the first time, then play like π^* forever. See Section 6.

As a technical tool, we sometimes consider the dual of a game $\mathcal{G}$ (resp. the dual maximizing MDP of some minimizing MDP). Consider $\mathcal{G}^d \stackrel{\text{def}}{=} (S', (S'_\square, S'_\diamond, S'_\bigcirc), E', P')$ with the complement objective $\overline{\text{EN}(k) \cap \text{EPAR}} = \text{Term}(k) \cup \text{OPAR}$, where $\mathcal{G}^d$ is simply the game with the roles of Maximizer and Minimizer reversed, i.e.,

$$S' = S \qquad S'_\square = S_\diamond \qquad S'_\diamond = S_\square \qquad S'_\bigcirc = S_\bigcirc \qquad E' = E \qquad P' = P$$

Hence $\Sigma_{\mathcal{G}^d} = \Pi_{\mathcal{G}}$ and $\Pi_{\mathcal{G}^d} = \Sigma_{\mathcal{G}}$. It is easy to see that for any objective $\mathtt{O}$ and start state s

1. $\text{val}^{\mathcal{G}}_{\mathtt{O}}(s) + \text{val}^{\mathcal{G}^d}_{\overline{\mathtt{O}}}(s) = 1$.
2. σ is ε-optimal maximizing for $\mathtt{O}$ in $\mathcal{G}$ iff it is ε-optimal minimizing for $\overline{\mathtt{O}}$ in $\mathcal{G}^d$.
3. π is ε-optimal minimizing for $\mathtt{O}$ in $\mathcal{G}$ iff it is ε-optimal maximizing for $\overline{\mathtt{O}}$ in $\mathcal{G}^d$.

So approximating the value of $\text{EN}(k) \cap \text{EPAR}$ in $\mathcal{G}$ can be reduced in linear time to approximating the value of $\text{Term}(k) \cup \text{OPAR}$ in $\mathcal{G}^d$.

4   Computing $\text{val}^{\mathcal{G}}_{\text{Gain}}(s)$

Given an SSG $\mathcal{G} = (S, (S_\square, S_\diamond, S_\bigcirc), E, P)$ and a start state s, we will show how to compute $\text{val}^{\mathcal{G}}_{\text{Gain}}(s)$ and the optimal strategies for both players.

We start with the case of maximizing MDPs. The following lemma summarizes some previous results ([29, Lemmas 30,16], [28, Lemma 26], [24, Proposition 4]).

▶ **Lemma 3.** *Let $\mathcal{M}$ be a maximizing MDP.*

1. $\mathtt{Lval}^{\mathcal{M}}(s) = \mathtt{val}^{\mathcal{M}}_{\mathtt{Gain}}(s)$ *for all states $s \in S$.*

2. *Optimal strategies for $\mathtt{Gain}$ in $\mathcal{M}$ exist and can be chosen FD, with $\mathcal{O}(\exp(|\mathcal{M}|^{\mathcal{O}(1)}))$ memory modes, and exponential memory is also necessary.*

3. *For any state $s \in S$, $\mathtt{Lval}^{\mathcal{M}}(s)$ is rational and can be computed in $\mathcal{O}(|\mathcal{M}|^8)$ deterministic polynomial time if rewards are in unary, and in NP and coNP if rewards are in binary.*

Proof. Item 1 holds by [29, Lemma 30].

Towards Item 2, we follow the proof of [29, Lemma 16]. Since $\mathtt{Gain} = \mathtt{LimInf}(> -\infty) \cap \mathtt{EPAR}$ is shift-invariant, there exist optimal strategies by [22]. By [28, Theorem 18] and Item 1, an optimal strategy for $\mathtt{Gain}$ can be constructed as follows. Let $A \stackrel{\text{def}}{=} \bigcup_{k \in \mathbb{N}} \mathsf{AS}\,(\mathsf{ST}(k) \cap \mathtt{EPAR})$ and $B \stackrel{\text{def}}{=} \mathsf{AS}\,(\mathtt{LimInf}(= \infty) \cap \mathtt{EPAR})$ be the subsets of states from which there exist almost surely winning strategies for the objectives $\mathsf{ST}(k) \cap \mathtt{EPAR}$ and $\mathtt{LimInf}(= \infty) \cap \mathtt{EPAR}$, respectively. By [28, Theorem 8], we can restrict the values k in the definition of A by some $k' = \mathcal{O}(|S| \cdot R)$, i.e., $A = \bigcup_{k \le k'} \mathsf{AS}\,(\mathsf{ST}(k) \cap \mathtt{EPAR})$. An optimal strategy σ for $\mathtt{Gain}$ works in two phases. First it plays an optimal strategy σ_R towards reaching the set $A \cup B$, where σ_R can be chosen MD by Remark 1. Then, upon reaching A (resp. B), it plays an almost surely winning strategy σ_A for the objective $\mathsf{ST}(k) \cap \mathtt{EPAR}$ (resp. σ_B for the objective $\mathtt{LimInf}(= \infty) \cap \mathtt{EPAR}$). By [28, Theorem 8], the strategy σ_A requires $\mathcal{O}(k \cdot |S|)$ memory modes for a given k and thus at most $\mathcal{O}(|S|^2 \cdot R)$, since we can assume that $k \le k'$. Towards the strategy σ_B, we first observe that in finite MDPs a strategy is almost-surely winning for $\mathtt{LimInf}(= \infty) \cap \mathtt{EPAR}$ iff it is almost-surely winning for $\mathtt{MP}(> 0) \cap \mathtt{EPAR}$. By [24, Proposition 4], there exist optimal deterministic strategies for $\mathtt{MP}(> 0) \cap \mathtt{EPAR}$ that use exponential memory, i.e., $\mathcal{O}(\exp(|\mathcal{M}|^{\mathcal{O}(1)}))$ memory modes. The memory required for σ_B exceeds that of σ_R and σ_A (even when R is given in binary), and the one extra memory mode to record the switch from σ_R to σ_A (resp. σ_B) is negligible in comparison. Thus we can conclude that σ uses $\mathcal{O}(\exp(|\mathcal{M}|^{\mathcal{O}(1)}))$ memory modes. [24, Fig. 1 and Prop. 4] shows that exponential memory is necessary.

Towards Item 3, let $d \stackrel{\text{def}}{=} |Col(S)|$ be the number of priorities in the parity condition. By [28, Lemma 26], for each $s \in S$, $\mathtt{Lval}^{\mathcal{M}}(s)$ is rational and can be computed in deterministic time $\tilde{\mathcal{O}}(|E| \cdot d \cdot |S|^4 \cdot R + d \cdot |S|^{3.5} \cdot (|P| + |r|)^2)$ (and still in NP and coNP if R is given in binary). So $\mathtt{Lval}^{\mathcal{M}}(s)$ can be computed in $\mathcal{O}(|\mathcal{M}|^8)$ deterministic polynomial time if weights are given in unary, and in NP and coNP if weights are given in binary. ◀

In order to extend Lemma 3 from MDPs to games, we need the notion of derived MDPs, obtained by fixing the choices of one player according to some FD strategy. Given an SSG $\mathcal{G} = (S, (S_\square, S_\diamond, S_\bigcirc), E, P)$ and a finite memory deterministic (FD) strategy π for Minimizer (resp. σ for Maximizer) from a state s, described by $(\mathsf{M}, \mathsf{m}_0, \mathsf{upd}, \mathsf{nxt})$, let $\mathcal{G}_\pi$ (resp. $\mathcal{G}^\sigma$) be the maximizing (resp. minimizing) MDP with state space $\mathsf{M} \times S$ obtained by fixing Minimizer's (resp. Maximizer's) choices according to π (resp. σ).

▶ **Lemma 4.** *For every SSG $\mathcal{G}$, objective $\mathtt{O}$ and Minimizer (resp. Maximizer) FD strategy $\pi = (\mathsf{M}, \mathsf{m}_0, \mathsf{upd}, \mathsf{nxt})$ (resp. σ), from state s we get $\mathtt{val}^{\mathcal{G}^\sigma}_{\mathtt{O}}((\mathsf{m}_0, s)) \le \mathtt{val}^{\mathcal{G}}_{\mathtt{O}}(s) \le \mathtt{val}^{\mathcal{G}_\pi}_{\mathtt{O}}((\mathsf{m}_0, s))$ and equality holds if π (resp. σ) is optimal from state s.*

▶ **Theorem 5.** *Consider an SSG $\mathcal{G} = (S, (S_\square, S_\diamond, S_\bigcirc), E, P)$ with the $\mathtt{Gain}$ objective.*

1. *Optimal Minimizer strategies exist and can be chosen uniform MD.*

2. $\mathtt{val}^{\mathcal{G}}_{\mathtt{Gain}}(s)$ *is rational and questions about it, i.e., $\mathtt{val}^{\mathcal{G}}_{\mathtt{Gain}}(s) \le c$ for constants c, are decidable in NP.*

3. *Optimal Maximizer strategies exist and can be chosen FD, with $\mathcal{O}(\exp(|\mathcal{G}|^{\mathcal{O}(1)}))$ memory modes. Moreover, exponential memory is also necessary.*

Proof. Towards Item 1, observe that since both the objectives $\texttt{LimInf}(=-\infty)$ and $\texttt{OPAR}$ are shift-invariant and submixing, so is their union, i.e., $\overline{\texttt{Gain}}$ is shift-invariant and submixing. Hence, by [23, Theorem 1.1], an optimal MD strategy π_s^* for Minimizer exists from any state $s \in S$. Since S is finite and $\texttt{Gain}$ is shift-invariant, we can also obtain a uniformly optimal MD strategy π^*, i.e., π^* is optimal from every state.

Towards Item 2, consider the maximizing MDP $\mathcal{G}_{\pi^*}$ obtained from $\mathcal{G}$ by fixing π^*. Since π^* is MD, the states of $\mathcal{G}_{\pi^*}$ are the same as the states as $\mathcal{G}$. Since π^* is optimal for Minimizer from every state s, we obtain that $\texttt{val}_{\texttt{Gain}}^{\mathcal{G}}(s) = \texttt{val}_{\texttt{Gain}}^{\mathcal{G}_{\pi^*}}(s)$ for every state s by Lemma 4. By Lemma 3, the latter is rational and can be computed in polynomial time for weights in unary (resp. in NP and coNP for weights in binary). Thus, by guessing π^*, we can decide questions $\texttt{val}_{\texttt{Gain}}^{\mathcal{G}}(s) \leq c$ in NP.

Towards Item 3, we again use the property that $\overline{\texttt{Gain}}$ is shift-invariant and submixing (see above). By [29, Theorem 6, Def. 24], optimal FD Maximizer strategies for $\texttt{Gain}$ in an SSG require only $|S_\diamond| \cdot \lceil \log(|E|) \rceil$ many extra bits of memory above the memory required for optimal Maximizer strategies in any derived MDP $\mathcal{M}$ where Minimizer's choices are fixed. Hence, by Lemma 3, one can obtain optimal FD Maximizer strategies in $\mathcal{G}$ that use at most $2^{|S_\diamond| \cdot \lceil \log(|E|) \rceil} \cdot \mathcal{O}(\exp(|\mathcal{M}|^{\mathcal{O}(1)})) = \mathcal{O}(\exp(|\mathcal{G}|^{\mathcal{O}(1)}))$ memory modes. The corresponding exponential lower bound on the memory holds already for MDPs by Lemma 3. ◀

5   Computing the Upper Bound N

We show how to compute the upper bound N, up-to which Maximizer needs to remember the energy level, for any given error margin $\varepsilon > 0$. Similarly as in Section 4, we first solve the problem for maximizing MDPs and then extend the solution to SSGs.

5.1   Computing N for maximizing MDPs

Given a maximizing MDP $\mathcal{M} = (S, S_\square, S_\bigcirc, E, P)$ and $\varepsilon > 0$, we will compute an $N \in \mathbb{N}$ such that for all $s \in S$ and all $j \geq N$

$$0 \leq \texttt{val}_{\texttt{Term}(j) \cup \texttt{OPAR}}^{\mathcal{M}}(s) - \texttt{val}_{\texttt{Loss}}^{\mathcal{M}}(s) \leq \varepsilon.$$

Recall that $\texttt{Loss} = \texttt{LimInf}(=-\infty) \cup \texttt{OPAR}$. We now define the sets of states $W_0 \overset{\text{def}}{=} \texttt{AS}(\texttt{Loss})$, $W_1 \overset{\text{def}}{=} \texttt{AS}(\texttt{LimInf}(=-\infty))$ and $W_2 \overset{\text{def}}{=} \texttt{AS}(\texttt{OPAR})$. By Remark 1, there exist optimal MD strategies for $\texttt{LimInf}(=-\infty)$ and $\texttt{OPAR}$. Since $\texttt{Loss}$ is shift-invariant and submixing, there exists an optimal MD strategy for it by [23, Theorem 1.1].

▶ **Lemma 6.** *For every state s in the MDP $\mathcal{M}$ we have*
1. $W_1 \cup W_2 \subseteq W_0$
2. $\texttt{val}_{\mathsf{F}\,W_0}(s) \leq \texttt{val}_{\texttt{Loss}}(s)$
3. $\texttt{val}_{\texttt{OPAR} \cap \overline{\mathsf{F}\,W_2}}(s) = 0$
4. *for every initial energy level $j \geq 0$*

$$\texttt{val}_{(\texttt{Term}(j) \cup \texttt{OPAR}) \cap \mathsf{F}\,W_0}(s) = \texttt{val}_{\mathsf{F}\,W_0}(s) \tag{2}$$

$$\texttt{val}_{\texttt{Loss}}(s) \leq \texttt{val}_{\texttt{Term}(j) \cup \texttt{OPAR}}(s) \leq \texttt{val}_{\texttt{Loss}}(s) + \sup_\sigma \mathcal{P}_{\sigma,s}\left(\texttt{Term}(j) \cap \overline{\mathsf{F}\,W_1}\right) \tag{3}$$

Proof.
1. This follows directly from the definitions of W_0, W_1, W_2.
2. Let σ' be an optimal MD strategy for $\mathsf{F}\,W_0$ from s and σ'' be an almost surely winning MD strategy for $\texttt{Loss}$ from any state in W_0. Let σ be the strategy that plays σ' until reaching W_0 and then switches to σ''. We have $\texttt{val}_{\texttt{Loss}}(s) \geq \mathcal{P}_{\sigma,s}(\texttt{Loss}) \geq \mathcal{P}_{\sigma',s}(\mathsf{F}\,W_0) = \texttt{val}_{\mathsf{F}\,W_0}(s)$.

3. For $s \in W_2$ the statement is obvious. So let $s \notin W_2$ and consider the modified MDP $\mathcal{M}' = \left(S', S'_\square, S'_\bigcirc, E', P'\right)$ where all states in W_2 are collapsed into a losing sink. I.e., $S' \stackrel{\text{def}}{=} (S \setminus W_2) \uplus \{trap\}$, with $trap$ a new random sink state having color 0 (thus losing for objective $\mathtt{OPAR}$), E' contains all of $(E \cap \{(S \setminus W_2) \times (S \setminus W_2)\} \cup (trap, trap))$ and all transitions to W_2 are redirected to $trap$ and P' is derived accordingly from P. Then $\mathtt{val}^{\mathcal{M}'}_{\mathtt{OPAR}}(\hat{s}) = \mathtt{val}^{\mathcal{M}}_{\mathtt{OPAR} \cap \overline{\mathsf{F}W_2}}(\hat{s})$ for all states $\hat{s} \in S \setminus W_2$. Towards a contradiction, assume that $\mathtt{val}^{\mathcal{M}}_{\mathtt{OPAR} \cap \overline{\mathsf{F}W_2}}(s) > 0$. Hence $\mathtt{val}^{\mathcal{M}'}_{\mathtt{OPAR}}(s) > 0$. Then, by [22, Theorem 3.2], there exists a state $s' \in S'$ such that $\mathtt{val}^{\mathcal{M}'}_{\mathtt{OPAR}}(s') = 1$, and it is easy to see that $s' \neq trap$ and thus $s' \in S \setminus W_2$. But this implies that $\mathtt{val}^{\mathcal{M}}_{\mathtt{OPAR}}(s') = 1$ and thus $s' \in W_2$, a contradiction.

4. Let $\mathtt{O} \stackrel{\text{def}}{=} \mathtt{Term}(j) \cup \mathtt{OPAR}$. For Equation (2), the first inequality $\mathtt{val}_{\mathtt{O} \cap \mathsf{F}W_0}(s) \leq \mathtt{val}_{\mathsf{F}W_0}(s)$ is trivial, since $\mathtt{O} \cap \mathsf{F}W_0 \subseteq \mathsf{F}W_0$. To show the reverse inequality, consider the strategy σ that first plays like an optimal MD strategy σ' for the objective $\mathsf{F}W_0$ and after reaching W_0 switches to an almost surely winning MD strategy σ'' for the objective $\mathtt{Loss}$. Then $\mathtt{val}_{\mathtt{O} \cap \mathsf{F}W_0}(s) \geq \mathcal{P}_{\sigma,s}(\mathtt{O} \cap \mathsf{F}W_0) \geq \mathcal{P}_{\sigma,s}(\mathtt{Loss} \cap \mathsf{F}W_0) = \mathcal{P}_{\sigma',s}(\mathsf{F}W_0) = \mathtt{val}_{\mathsf{F}W_0}(s)$, where the second inequality is due to $\mathtt{LimInf}(= -\infty) \subseteq \mathtt{Term}(j)$.

   For Equation (3), the first inequality is again due to the fact that $\mathtt{LimInf}(= -\infty) \subseteq \mathtt{Term}(j)$ for all $j \geq 0$. Towards the second inequality of Equation (3) we have

$$\mathtt{val}_\mathtt{O}(s)$$
$$= \sup_\sigma \mathcal{P}_{\sigma,s}(\mathtt{O})$$
$$= \sup_\sigma \left(\mathcal{P}_{\sigma,s}\left(\mathtt{O} \cap \mathsf{F}W_0\right) + \mathcal{P}_{\sigma,s}\left(\mathtt{O} \cap \overline{\mathsf{F}W_0}\right)\right) \qquad \text{(Law of total probability)}$$
$$\leq \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{O} \cap \mathsf{F}W_0\right) + \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{O} \cap \overline{\mathsf{F}W_0}\right) \qquad (\sup(f+g) \leq \sup f + \sup g)$$
$$= \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathsf{F}W_0\right) + \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{O} \cap \overline{\mathsf{F}W_0}\right) \qquad (\textit{Equation (2)})$$
$$\leq \mathtt{val}_\mathtt{Loss}(s) + \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{O} \cap \overline{\mathsf{F}W_0}\right) \qquad (\textit{Item 2})$$

We can upper-bound the second summand above as follows.

$$\sup_\sigma \mathcal{P}_{\sigma,s}(\mathtt{O} \cap \overline{\mathsf{F}W_0})$$
$$= \sup_\sigma \mathcal{P}_{\sigma,s}\left((\mathtt{Term}(j) \cup \mathtt{OPAR}) \cap \overline{\mathsf{F}W_0}\right)$$
$$\leq \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{Term}(j) \cap \overline{\mathsf{F}W_0}\right) + \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{OPAR} \cap \overline{\mathsf{F}W_0}\right) \qquad \text{(Union bound)}$$
$$\leq \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{Term}(j) \cap \overline{\mathsf{F}W_1}\right) + \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{OPAR} \cap \overline{\mathsf{F}W_2}\right) \qquad (\textit{Item 1})$$
$$= \sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{Term}(j) \cap \overline{\mathsf{F}W_1}\right) \qquad (\textit{Item 3}) \qquad \blacktriangleleft$$

We show that the term $\sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{Term}(j) \cap \overline{\mathsf{F}W_1}\right)$ in Equation (3) can be made arbitrarily small for large j. To this end, we use [3, Lemma 3.9] (adapted to our notation).

▶ **Lemma 7** ([3, Lemma 3.9 and Claim 6]). *Let* $\mathcal{M} = (S, S_\square, S_\bigcirc, E, P)$ *be a maximizing finite MDP with rewards in unary and* $W_1 \stackrel{\text{def}}{=} \mathsf{AS}(\mathtt{LimInf}(= -\infty))$. *One can compute, in polynomial time, a rational constant* $c < 1$, *and an integer* $h \geq 0$ *such that for all* $j \geq h$ *and* $s \in S$

$$\sup_\sigma \mathcal{P}_{\sigma,s}\left(\mathtt{Term}(j) \cap \overline{\mathsf{F}W_1}\right) \leq \frac{c^j}{1-c}.$$

Moreover, $1/(1-c) \in \mathcal{O}\left(\exp(|\mathcal{M}|^{\mathcal{O}(1)})\right)$ *and* $h \in \mathcal{O}(\exp\left(|\mathcal{M}|^{\mathcal{O}(1)}\right))$.

▶ **Lemma 8.** *Consider a maximizing MDP $\mathcal{M} = (S, S_\square, S_\bigcirc, E, P)$, $\varepsilon > 0$ and the constants c, h from Lemma 7. For rewards in unary and $i \geq N$ we have $\mathtt{val}^{\mathcal{M}}_{\mathtt{Term}(i) \cup \mathtt{OPAR}}(s) - \mathtt{val}^{\mathcal{M}}_{\mathtt{Loss}}(s) \leq \varepsilon$ where $N \stackrel{def}{=} \max(h, \lceil \log_c(\varepsilon \cdot (1-c)) \rceil) \in \mathcal{O}\left(\exp(|\mathcal{M}|^{\mathcal{O}(1)}) \cdot \log(1/\varepsilon)\right)$.*

For rewards in binary we have $N \in \mathcal{O}\left(\exp(\exp(|\mathcal{M}|^{\mathcal{O}(1)})) \cdot \log(1/\varepsilon)\right)$, i.e., the size of N increases by one exponential.

Proof sketch. For rewards in unary, the result follows from Lemma 6(Equation (3)) and Lemma 7. For rewards in binary, the constants increase by one exponential via encoding binary rewards into unary rewards in a modified MDP. ◀

5.2 Computing N for SSGs

In order to compute the bound N for an SSG $\mathcal{G}$, we first consider bounds $N(s)$ for individual states s and then take their maximum. Given a state s, we can use Theorem 5(Item 3) to obtain an optimal FD strategy (with $\mathcal{O}(\exp(|\mathcal{G}|^{\mathcal{O}(1)}))$ memory modes) $\sigma^*(s) = (\mathsf{M}, \mathsf{m}_0, \mathsf{upd}, \mathsf{nxt})$ for Maximizer from state s w.r.t. the $\mathtt{Gain}$ objective. Theorem 5(Item 1) yields a uniform MD strategy π^* that is optimal for Minimizer from all states s w.r.t. the $\mathtt{Gain}$ objective.

▶ **Lemma 9.** *Given an SSG $\mathcal{G} = (S, (S_\square, S_\diamond, S_\bigcirc), E, P)$ and $\varepsilon > 0$, we can compute a number $N \in \mathbb{N}$ such that for all $i \geq N$ and states $s \in S$ we have*

$$\mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}(s) - \varepsilon \leq \mathtt{val}^{\mathcal{G}}_{\mathtt{Gain}}(s) - \varepsilon \leq \inf_\pi \mathcal{P}^{\mathcal{G}}_{\sigma^*(s), \pi, s}(\mathtt{EN}(i) \cap \mathtt{EPAR}) \leq \mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}(s) \quad (4)$$

i.e., $\sigma^*(s)$ is ε-optimal for Maximizer for $\mathtt{EN}(i) \cap \mathtt{EPAR}$ for all $i \geq N$. In particular, $0 \leq \mathtt{val}^{\mathcal{G}}_{\mathtt{Gain}}(s) - \mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}(s) \leq \varepsilon$.

Moreover, π^* is ε-optimal for Minimizer from any state s for $i \geq N$.

$$\sup_\sigma \mathcal{P}^{\mathcal{G}}_{\sigma, \pi^*, s}(\mathtt{EN}(i) \cap \mathtt{EPAR}) \leq \sup_\sigma \mathcal{P}^{\mathcal{G}}_{\sigma, \pi^*, s}(\mathtt{Gain}) = \mathtt{val}^{\mathcal{G}}_{\mathtt{Gain}}(s) \leq \mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}(s) + \varepsilon \quad (5)$$

For rewards in unary, N is doubly exponential, i.e., $N \in \mathcal{O}\left(\exp(\exp(|\mathcal{G}|^{\mathcal{O}(1)})) \cdot \log(1/\varepsilon)\right)$ and it can be computed in exponential time. For rewards in binary, the size of N and its computation time increase by one exponential, respectively.

Proof. Assume that rewards are in unary. The first inequality of (4) holds because $\mathtt{EN}(i) \cap \mathtt{EPAR} \subseteq \mathtt{Gain}$ for any i. The third inequality of (4) follows from the definition of the value. Towards the second inequality of (4), we consider the minimizing MDP $\mathcal{M}(s) \stackrel{def}{=} \mathcal{G}^{\sigma^*(s)}$ obtained by fixing the Maximizer strategy $\sigma^*(s)$. Since $\sigma^*(s)$ is optimal for Maximizer from state s wrt. the objective $\mathtt{Gain}$, Lemma 4 yields that

$$\mathtt{val}^{\mathcal{G}}_{\mathtt{Gain}}(s) = \mathtt{val}^{\mathcal{M}(s)}_{\mathtt{Gain}}((\mathsf{m}_0, s)). \tag{6}$$

Since $\sigma^*(s)$ has $\mathcal{O}(\exp(|\mathcal{G}|^{\mathcal{O}(1)}))$ memory modes, the size of $\mathcal{M}(s)$ is exponential in $|\mathcal{G}|$ and $\mathcal{M}(s)$ can be computed in exponential time.

Now we consider the dual maximizing MDP $\mathcal{M}(s)^d$ and the objectives $\mathtt{Term}(i) \cup \mathtt{OPAR}$ and $\mathtt{Loss}$. (Note that $\mathcal{M}(s)^d$ has the same size as $\mathcal{M}(s)$.) From Lemma 8, we obtain a bound $N(s) \in \mathbb{N}$ such that for all $i \geq N(s)$

$$0 \leq \mathtt{val}^{\mathcal{M}(s)^d}_{\mathtt{Term}(i) \cup \mathtt{OPAR}}((\mathsf{m}_0, s)) - \mathtt{val}^{\mathcal{M}(s)^d}_{\mathtt{Loss}}((\mathsf{m}_0, s)) \leq \varepsilon. \tag{7}$$

By Lemma 8 and Lemma 7, $N(s)$ is exponential in $|\mathcal{M}(s)^d|$ and thus doubly exponential in $|\mathcal{G}|$, i.e., $N(s) \in \mathcal{O}\left(\exp(\exp(|\mathcal{G}|^{\mathcal{O}(1)})) \cdot \log(1/\varepsilon)\right)$. Moreover, $N(s)$ can be computed in time polynomial in $|\mathcal{M}(s)^d|$ and thus in time exponential in $|\mathcal{G}|$. By duality, we can rewrite Equation (7) for $\mathcal{M}(s)$ as follows. For all $i \geq N(s)$

$$0 \leq \mathtt{val}_{\mathtt{Gain}}^{\mathcal{M}(s)}\left((\mathsf{m}_0, s)\right) - \mathtt{val}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}^{\mathcal{M}(s)}\left((\mathsf{m}_0, s)\right) \leq \varepsilon. \tag{8}$$

In order to get a uniform upper bound that holds for all states, let $N \stackrel{\text{def}}{=} \max_{s \in S} N(s)$. Since $|S|$ is linear, we still have $N \in \mathcal{O}\left(\exp(\exp(|\mathcal{G}|^{\mathcal{O}(1)})) \cdot \log\left(1/\varepsilon\right)\right)$ and it can be computed in exponential time in $|\mathcal{G}|$. Finally, we can show the second inequality of (4).

$$\begin{aligned}
&\inf_{\pi} \mathcal{P}_{\sigma^*(s), \pi, s}^{\mathcal{G}}\left(\mathtt{EN}(i) \cap \mathtt{EPAR}\right) \\
&= \inf_{\pi} \mathcal{P}_{\pi, (\mathsf{m}_0, s)}^{\mathcal{M}(s)}\left(\mathtt{EN}(i) \cap \mathtt{EPAR}\right) \\
&= \mathtt{val}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}^{\mathcal{M}(s)}\left((\mathsf{m}_0, s)\right) \\
&\geq \mathtt{val}_{\mathtt{Gain}}^{\mathcal{M}(s)}\left((\mathsf{m}_0, s)\right) - \varepsilon && \text{by } i \geq N \geq N(s) \text{ and Equation (8)} \\
&= \mathtt{val}_{\mathtt{Gain}}^{\mathcal{G}}(s) - \varepsilon && \text{by (6)}
\end{aligned}$$

The first inequality of (5) holds because $\mathtt{EN}(i) \cap \mathtt{EPAR} \subseteq \mathtt{Gain}$ for any i. The equality in (5) holds by the optimality of π^*. The second inequality of (5) follows from the previously stated consequence of (4).

For rewards in binary, the sizes of the numbers $N(s)$ (and hence N) and the time to compute it increase by one exponential by Lemma 8.   ◀

6 Unfolding the Game to Energy Level N

Given an SSG $\mathcal{G} = (S, (S_\square, S_\diamond, S_\bigcirc), E, P)$ and error tolerance $\varepsilon > 0$, for each state $s \in S$ and energy level $i \geq 0$, we want to compute a rational number v' which satisfies $0 \leq v' - \mathtt{val}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}^{\mathcal{G}}(s) \leq \varepsilon$, and ε-optimal FD strategies σ_ε and π_ε for Maximizer and Minimizer, resp. We achieve this by constructing a finite-state parity game $\mathcal{G}'$ that closely approximates the original game $\mathcal{G}$, as described in Section 3(Item 3).

For clarity, we explain the construction in two steps. In the first step, we consider a finite-state parity game $\mathcal{G}[N]$. (Unlike $\mathcal{G}'$, the game $\mathcal{G}[N]$ is not actually constructed. It just serves as a part of the correctness proof.) $\mathcal{G}[N]$ encodes the energy level up-to $N + R$ (where R is the maximal transition reward) into the states, i.e., it has states of the form (s, k) with $k \leq N + R$. It imitates the original game $\mathcal{G}$ till energy level $N + R$, but at any state (s, i) with energy level $i \geq N$ it jumps to a winning state with probability $\mathtt{val}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}^{\mathcal{G}}(s)$ and to a losing state with probability $1 - \mathtt{val}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}^{\mathcal{G}}(s)$. (We need the margin up-to $N + R$, because transitions can have rewards > 1, so the level N might not be hit exactly.) Similarly, at states $(s, 0)$ with energy level 0, we jump to a losing state. The coloring function in the new game $\mathcal{G}[N]$ derives its colors from the colors in the original game $\mathcal{G}$, i.e., all states (s, i) have the same color as s in $\mathcal{G}$.

By construction of $\mathcal{G}[N]$, for $i \leq N$, the $\mathtt{EPAR}$ value of (s, i) in $\mathcal{G}[N]$ coincides with $\mathtt{val}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}^{\mathcal{G}}(s)$.

In the second step, since we do not know the exact values $\mathtt{val}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}^{\mathcal{G}}(s)$ for $N + R \geq i > N$, we approximate these by the slightly larger $\mathtt{val}_{\mathtt{Gain}}^{\mathcal{G}}(s)$. I.e., we modify $\mathcal{G}[N]$ by replacing the probability values $\mathtt{val}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}^{\mathcal{G}}(s)$ for the jumps to the winning state by $\mathtt{val}_{\mathtt{Gain}}^{\mathcal{G}}(s)$. Let $\mathcal{G}'$ be the resulting finite-state parity game. It follows from Lemma 9 that $0 \leq \mathtt{val}_{\mathtt{Gain}}^{\mathcal{G}}(s) - \mathtt{val}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}^{\mathcal{G}}(s) \leq \varepsilon$ for $i \geq N$ and $\mathtt{Lval}_{\mathtt{EN} \cap \mathtt{EPAR}}^{\mathcal{G}}(s) = \mathtt{val}_{\mathtt{Gain}}^{\mathcal{G}}(s)$. Thus $\mathcal{G}'$ ε-over-approximates $\mathcal{G}[N]$ and $\mathcal{G}$, and we obtain the following lemma.

▶ **Lemma 10.** *For all states s and all $0 \leq i \leq N$*

$$\mathtt{val}^{\mathcal{G}[N]}_{\mathtt{EPAR}}\left((s,i)\right) = \mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}(s), \text{ and}$$

$$0 \leq \mathtt{val}^{\mathcal{G}'}_{\mathtt{EPAR}}\left((s,i)\right) - \mathtt{val}^{\mathcal{G}[N]}_{\mathtt{EPAR}}\left((s,i)\right) \leq \varepsilon.$$

Now we are ready to prove the main theorem.

▶ **Theorem 2.** *Let $\mathcal{G} = (S, (S_\square, S_\diamond, S_\bigcirc), E, P)$ be an SSG with transition rewards in unary assigned by function r and colors assigned to states by function Col. For every state $s \in S$, initial energy level $i \geq 0$ and error margin $\varepsilon > 0$, one can compute*
1. *a rational number v' such that $0 \leq v' - \mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(i) \cap \mathtt{EPAR}}(s) \leq \varepsilon$ in 2-NEXPTIME.* [2]
2. *ε-optimal FD strategies σ_ε and π_ε for Maximizer and Minimizer, resp., in 2-NEXPTIME. These strategies use $\mathcal{O}\left(2\text{-}\mathsf{EXP}\left(|\mathcal{G}|\right) \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ memory modes.*
For rewards in binary, the bounds above increase by one exponential.

Proof. For $i > N$ we output $v' = \mathtt{val}^{\mathcal{G}}_{\mathtt{Gain}}(s)$, which satisfies the condition by Lemma 9. For $i \leq N$ we output $v' = \mathtt{val}^{\mathcal{G}'}_{\mathtt{EPAR}}\left((s,i)\right)$, which satisfies the condition by Lemma 10. By Theorem 5, the values $\mathtt{val}^{\mathcal{G}}_{\mathtt{Gain}}(s)$ are rational for all states s. Therefore all probability values in $\mathcal{G}'$ are rational and thus the EPAR values of all states in $\mathcal{G}'$ are rational. Hence our numbers v' are always rational.

By Theorem 5, the values $\mathtt{val}^{\mathcal{G}}_{\mathtt{Gain}}(s)$ for all states $s \in S$ can be computed in exponential time. By Lemma 9, $N \in \mathcal{O}\left(\exp(\exp(|\mathcal{G}|^{\mathcal{O}(1)})) \cdot \log(1/\varepsilon)\right)$ is doubly exponential. Therefore, we can construct $\mathcal{G}'$ in $\mathcal{O}\left(\exp(\exp(|\mathcal{G}|^{\mathcal{O}(1)})) \cdot \log(1/\varepsilon)\right)$ time and space. Questions about the parity values of states in $\mathcal{G}'$ can be decided in nondeterministic time polynomial in $|\mathcal{G}'|$. Thus the numbers v' are computed in 2-NEXPTIME.

Towards Item 2, we construct ε-optimal FD strategies σ_ε for Maximizer (resp. π_ε for Minimizer) for $\mathtt{EN}(i) \cap \mathtt{EPAR}$ in $\mathcal{G}$. Let $\hat{\sigma}$ (resp. $\hat{\pi}$) be optimal MD strategies for Maximizer (resp. Minimizer) for the objective EPAR in $\mathcal{G}'$, which exist by Remark 1. Since these strategies are MD, they can be guessed in nondeterministic time polynomial in the size $|\mathcal{G}'|$, and thus in $\mathcal{O}\left(\exp(\exp(|\mathcal{G}|^{\mathcal{O}(1)})) \cdot \log(1/\varepsilon)\right)$ nondeterministic time.

Then σ_ε plays as follows. While the current energy level j (i plus the sum of the rewards so far) stays $< N$, then, at any state s', play like $\hat{\sigma}$ at state (s', j) in $\mathcal{G}'$. Once the energy level reaches a value $\geq N$ at some state s' for the first time, then play like $\sigma^*(s')$ forever. (Recall that $\sigma^*(s')$ is the optimal FD Maximizer strategy for Gain from state s' from Section 5.2.) σ_ε is ε-optimal by Lemma 10 and Lemma 9. It needs to remember the energy level up-to N while simulating $\hat{\sigma}$. Moreover, $\sigma^*(s')$ needs $\mathcal{O}(\exp(|\mathcal{G}|^{\mathcal{O}(1)}))$ memory modes by Theorem 5. Finally, it needs to remember the switch from $\hat{\sigma}$ to $\sigma^*(s')$. Since $N \in \mathcal{O}\left(\exp(\exp(|\mathcal{G}|^{\mathcal{O}(1)})) \cdot \log(1/\varepsilon)\right)$ dominates the rest, σ_ε uses $\mathcal{O}\left(\exp(\exp(|\mathcal{G}|^{\mathcal{O}(1)})) \cdot \log(1/\varepsilon)\right)$ memory modes.

Similarly, π_ε plays as follows. While the current energy level j stays $< N$, at any state s', play like $\hat{\pi}$ at state (s', j) in $\mathcal{G}'$. Once the energy level reaches a value $\geq N$ (at any state) for the first time, then play like π^* forever (where π^* is the uniform optimal MD Minimizer strategy for Gain from Section 5.2.) π_ε is ε-optimal by Lemma 10 and Lemma 9. While π^* is MD and does not use any memory, π_ε still needs to remember the energy level up-to N while simulating $\hat{\pi}$, and thus it uses $\mathcal{O}\left(\exp(\exp(|\mathcal{G}|^{\mathcal{O}(1)})) \cdot \log(1/\varepsilon)\right)$ memory modes.

For rewards in binary, all bounds increase by one exponential via an encoding of $\mathcal{G}$ into an exponentially larger but equivalent game with rewards in unary. ◀

---

[2] We write "computing a number v' in 2-NEXPTIME" as a shorthand for the property that questions like $v' \leq c$ for constants c are decidable in 2-NEXPTIME.

No nontrivial lower bounds are known on the computational complexity of approximating $\mathtt{val}^{\mathcal{G}}_{\mathtt{EN}(i)\,\cap\,\mathtt{EPAR}}(s)$. However, even without the parity part, the problem appears to be hard. The best known algorithm for approximating the value of the energy objective (resp. the dual termination objective) runs in NEXPTIME for SSGs with rewards in unary [3].

As for lower bounds on the strategy complexity, ε-optimal Maximizer strategies need at least an exponential number of memory modes (for any $0 < \varepsilon < 1$) even in maximizing MDPs. This can easily be shown by extending the example in Lemma 3(Item 2) and [24, Fig. 1 and Prop. 4] that shows the lower bound for the Gain objective. First loop in a state with an unfavorable color to accumulate a sufficiently large reward (depending on ε) and then switch to the MDP in [24, Fig. 1 and Prop. 4] to play for Gain (since $\mathtt{EN}(i) \cap \mathtt{EPAR}$ will be very close to Gain then). Even the latter part requires exponentially many memory modes.

7 Conclusion & Extensions

We gave a procedure to compute ε-approximations of the value of combined energy-parity objectives in SSGs. The decidability of questions about the exact values is open, but the problem is at least as hard as the positivity problem for linear recurrence sequences [31, Section 5.2.3]. Unlike almost surely winning Maximizer strategies which require infinite memory in general [28, 29], ε-optimal strategies for either player require only finite memory with at most doubly exponentially many memory modes.

An interesting topic for further study is whether these results can be extended to other combined objectives where the parity part is replaced by something else, i.e., energy-X for some objective X (e.g., some other color-based condition like Rabin/Streett, or a quantitative objective about multi-dimensional transition rewards). While our proofs are not completely specific to parity, they do use many strong properties that parity satisfies.

- Shift-invariance of EPAR is used in several places, e.g. in Lemma 6 (and thus its consequences) and for the correctness of the constructions in Section 6.
- We use the fact that EPAR goes well together with $\mathtt{LimInf}(> -\infty)$, i.e., the objective $\mathtt{Gain} = \mathtt{LimInf}(> -\infty) \cap \mathtt{EPAR}$ allows optimal FD strategies for Maximizer in MDPs; cf. Lemma 3.
- The submixing property of $\mathtt{OPAR} = \overline{\mathtt{EPAR}}$ is used in Theorem 5 to lift Lemma 3 from MDPs to SSGs.

─── **References** ───

1 Rajeev Alur, Thomas A Henzinger, and Orna Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49(5):672–713, 2002.

2 Patrick Billingsley. *Probability and measure.* John Wiley & Sons, 2008.

3 T. Brázdil, V. Brožek, K. Etessami, and A. Kučera. Approximating the Termination Value of One-Counter MDPs and Stochastic Games. *Information and Computation*, 222:121–138, 2013. `arXiv:1104.4978`.

4 T. Brázdil, A. Kučera, and P. Novotný. Optimizing the expected mean payoff in energy Markov decision processes. In *International Symposium on Automated Technology for Verification and Analysis (ATVA)*, volume 9938 of *LNCS*, pages 32–49, 2016.

5 Tomás Brázdil, Václav Brozek, and Kousha Etessami. One-Counter Stochastic Games. In Kamal Lodaya and Meena Mahajan, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2010)*, volume 8 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 108–119, Dagstuhl, Germany, 2010. Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik. Full version at `arXiv:1009.5636`. `doi:10.4230/LIPIcs.FSTTCS.2010.108`.

6   Arindam Chakrabarti, Luca De Alfaro, Thomas A Henzinger, and Mariëlle Stoelinga. Resource interfaces. In *International Workshop on Embedded Software*, pages 117–133, 2003.

7   Krishnendu Chatterjee and Laurent Doyen. Energy parity games. In *International Colloquium on Automata, Languages and Programming (ICALP)*, volume 6199 of *LNCS*, pages 599–610, 2010.

8   Krishnendu Chatterjee and Laurent Doyen. Energy and mean-payoff parity Markov decision processes. In *International Symposium on Mathematical Foundations of Computer Science (MFCS)*, volume 6907, pages 206–218, 2011.

9   Krishnendu Chatterjee and Laurent Doyen. Games and Markov decision processes with mean-payoff parity and energy parity objectives. In *Mathematical and Engineering Methods in Computer Science (MEMICS)*, volume 7119 of *LNCS*, pages 37–46. Springer, 2011.

10   Krishnendu Chatterjee, Laurent Doyen, Hugo Gimbert, and Youssouf Oualhadj. Perfect-information stochastic mean-payoff parity games. In *International Conference on Foundations of Software Science and Computational Structures (FoSSaCS)*, volume 8412 of *LNCS*, 2014.

11   Krishnendu Chatterjee, Thomas A. Henzinger, and Nir Piterman. Generalized parity games. In Helmut Seidl, editor, *International Conference on Foundations of Software Science and Computational Structures (FoSSaCS)*, volume 4423 of *LNCS*, pages 153–167. Springer, 2007. `doi:10.1007/978-3-540-71389-0_12`.

12   Krishnendu Chatterjee, Marcin Jurdziński, and Thomas A. Henzinger. Simple stochastic parity games. In *Computer Science Logic (CSL)*, volume 2803 of *LNCS*, pages 100–113. Springer, 2003.

13   Krishnendu Chatterjee, Marcin Jurdziński, and Thomas A. Henzinger. Quantitative stochastic parity games. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 121–130. SIAM, 2004.

14   E.M. Clarke, O. Grumberg, and D. Peled. *Model Checking*. MIT Press, December 1999.

15   Anne Condon. The complexity of stochastic games. *Information and Computation*, 96(2):203–224, 1992. `doi:10.1016/0890-5401(92)90048-K`.

16   Laure Daviaud, Martin Jurdziński, and Ranko Lazić. A pseudo-quasi-polynomial algorithm for mean-payoff parity games. In *Logic in Computer Science (LICS)*, pages 325–334, 2018.

17   Luca De Alfaro and Thomas A Henzinger. Interface automata. *ACM SIGSOFT Software Engineering Notes*, 26(5):109–120, 2001.

18   David L. Dill. *Trace theory for automatic hierarchical verification of speed-independent circuits*, volume 24. MIT press Cambridge, 1989.

19   Graham Everest, Alfred Jacobus van der Poorten, Igor Shparlinski, and Thomas Ward. *Recurrence sequences*. ACM, 2003.

20   J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, 1997.

21   Dean Gillette. Stochastic games with zero stop probabilities. *Contributions to the Theory of Games*, 3:179–187, 1957.

22   Hugo Gimbert and Florian Horn. Solving simple stochastic tail games. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 847–862, 2010. URL: `http://epubs.siam.org/doi/abs/10.1137/1.9781611973075.69`.

23   Hugo Gimbert and Edon Kelmendi. Submixing and shift-invariant stochastic games. working paper or preprint, 2022. `arXiv:1401.6575`.

24   Hugo Gimbert, Youssouf Oualhadj, and Soumya Paul. Computing optimal strategies for Markov decision processes with parity and positive-average conditions. working paper or preprint, 2011. URL: `https://hal.science/hal-00559173/en/`.

25   Marcin Jurdziński. Deciding the winner in parity games is in UP ∩ co-UP. *Information Processing Letters*, 68(3):119–124, 1998.

26   A. Maitra and W. Sudderth. Stochastic games with Borel payoffs. In *Stochastic Games and Applications*, pages 367–373. Kluwer, Dordrecht, 2003.

27   Donald A. Martin. The determinacy of Blackwell games. *Journal of Symbolic Logic*, 63(4):1565–1581, 1998.

28   Richard Mayr, Sven Schewe, Patrick Totzke, and Dominik Wojtczak. MDPs with Energy-Parity Objectives. In *Logic in Computer Science (LICS)*. IEEE, 2017. `arXiv:1701.02546`.

29   Richard Mayr, Sven Schewe, Patrick Totzke, and Dominik Wojtczak. Simple stochastic games with almost-sure energy-parity objectives are in NP and coNP. In *Proc. of Fossacs*, volume 12650 of *LNCS*, 2021. Extended version on arXiv. `arXiv:2101.06989`.

30   Joël Ouaknine and James Worrell. On linear recurrence sequences and loop termination. *ACM SIGLOG News*, 2(2):4–13, 2015.

31   Jakob Piribauer. *On non-classical stochastic shortest path problems*. PhD thesis, Technische Universität Dresden, Germany, 2021. URL: `https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-762812`.

32   Jakob Piribauer and Christel Baier. On Skolem-Hardness and Saturation Points in Markov Decision Processes. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *Proc. of ICALP*, volume 168 of *LIPIcs*, pages 138:1–138:17, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/LIPIcs.ICALP.2020.138`.

33   Jakob Piribauer and Christel Baier. Positivity-hardness results on Markov decision processes. working paper or preprint, 2023. `arXiv:2302.13675`.

34   Amir Pnueli and Roni Rosner. On the synthesis of a reactive module. In *Annual Symposium on Principles of Programming Languages (POPL)*, pages 179–190, 1989.

35   Peter J. Ramadge and W. Murray Wonham. Supervisory control of a class of discrete event processes. *SIAM journal on control and optimization*, 25(1):206–230, 1987.

36   Lloyd S. Shapley. Stochastic games. *Proceedings of the national academy of sciences*, 39(10):1095–1100, 1953.

37   W. Zielonka. Infinite games on finitely coloured graphs with applications to automata on infinite trees. *Theoretical Computer Science*, 200(1-2):135–183, 1998.