Strategy Improvement for Stochastic Rabin and Streett Games

Krishnendu Chatterjee
Thomas A. Henzinger

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2006-33
http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-33.html

April 4, 2006

Acknowledgement

This research was supported in part by the AFOSR MURI grant F49620-00-1-0327, and the NSF ITR grant CCR-0225610.

Strategy Improvement for Stochastic Rabin and Streett Games *

Krishnendu Chatterjee[†] Thomas A. Henzinger[†,‡]

[†] EECS, University of California, Berkeley, USA

[‡] EPFL, Switzerland

{c_krish,tah}@eecs.berkeley.edu

April 4, 2006

Abstract

A stochastic graph game is played by two players on a game graph with probabilistic transitions. We consider stochastic graph games with ω-regular winning conditions specified as Rabin or Streett objectives. These games are NP-complete and coNP-complete, respectively. The *value* of the game for a player at a state s given an objective Φ is the maximal probability that the player can guarantee the satisfaction of Φ from s. We present a strategy improvement algorithm to compute values in stochastic Rabin games, where an improvement step involves solving Markov decision processes (MDPs) and non-stochastic Rabin games. The algorithm also computes values for stochastic Streett games but does not directly yield an optimal strategy for Streett objectives. We then show how to obtain an optimal strategy for Streett objectives by solving certain non-stochastic Streett games.

1 Introduction

Graph games. A stochastic graph game [6] is played on a directed graph with three kinds of states: player-1, player-2, and probabilistic states. At player-1 states, player 1 chooses a successor state; at player-2 states, player 2 chooses a successor state; at probabilistic states, a successor state is chosen according to a given probability distribution. The outcome of playing the

*This research was supported in part by the AFOSR MURI grant F49620-00-1-0327, and the NSF ITR grant CCR-0225610.

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 $2^1\!/\!_2$*-player graph game*. If there are only player 1 states and probabilistic states, we refer to the game as a Markov decision process (MDP).

Games with Rabin and Streett objectives. The theory of graph games with ω-regular winning conditions is the foundation for modeling and synthesizing reactive processes with fairness constraints [16, 18]. In the case of $2^1\!/\!_2$-player graph games, the two players represent a reactive system (or plant) and its environment (or controller), and the probabilistic states represent uncertainty. The class of $2^1\!/\!_2$-player graph games with ω-regular objectives provide an adequate model for the problem, because the fairness constraints of reactive processes are ω-regular. Strong fairness conditions are Streett objectives and Rabin objectives are their dual; moreover, every ω-regular objective can be specified as a Rabin and a Streett objective. The quantitative solution problem for a $2^1\!/\!_2$-player game with a Rabin objective Φ asks for each state s, for the maximal probability with which player 1 can ensure the satisfaction of Φ if the game is started from s (this probability is called the *value* of the game at the state s). An *optimal strategy* for player 1 is a strategy that enables player 1 to win with that maximal probability. The existence of *pure memoryless* optimal strategies for $2^1\!/\!_2$-player games with Rabin objectives was established recently [3] (a pure memoryless strategy chooses for each player-1 state a unique successor state; it uses neither randomization nor the history of the game). The existence of pure memoryless optimal strategies implies that the quantitative solution problem for $2^1\!/\!_2$-player games with Rabin objectives can be decided in NP; and the problem is NP-hard even for 2-player games. Hence $2^1\!/\!_2$-player games with Rabin objectives are NP-complete, and dually, coNP-complete for Streett objectives. The optimal strategies for the Streett player requires memory and finite-memory optimal strategies exist for Streett objectives in $2^1\!/\!_2$-player games.

Algorithms. Emerson-Jutla [10] showed that 2-player Rabin and Streett games (*without* probabilistic states) are NP-complete and coNP-completely, respectively. Several algorithms are known to solve 2-player Rabin and Streett games: such as recursive algorithms on game graphs [10, 13], and algorithms obtained by reduction to checking emptiness of weak alternating automata [15]. These algorithms are much better than a brute force enumeration of all possible pure memoryless strategies; especially for Rabin objectives with few Rabin pairs. For example the algorithm of [13] works in time $O(n^d \cdot d!)$ for game graphs with n states and Rabin objectives with

d-pairs. For $2^1\!/\!_2$-player games (*with* probabilistic states), Condon [6] proved containment in NP $\cap$ coNP and gave a strategy improvement algorithm for the restricted case of *reachability* objectives. A strategy improvement scheme iterates local optimizations of a pure memoryless strategy; this works if the iteration can be shown to converge to the global optimum [12]. For $2^1\!/\!_2$-player games with *parity* objectives (parity objectives are a complementation closed sub-class of Rabin and Streett objectives) containment in NP $\cap$ coNP was shown in [5] and a strategy improvement algorithm was given in [4]. However, for $2^1\!/\!_2$-player games with general Rabin objectives, no algorithm has been known which is better than a brute-force enumeration of the set of all possible pure memoryless strategies (choosing the best one as the optimal strategy), or one obtained by reduction of Rabin objectives to parity objectives. However, the reduction of Rabin objectives to parity objectives and then applying the strategy improvement algorithm for $2^1\!/\!_2$-player parity games yields a worst case complexity of double-exponential time[1].

Our results and techniques. We present a direct strategy improvement algorithm for $2^1\!/\!_2$-player Rabin games. The improvement step involves solving MDPs with Streett objectives and solving 2-player Rabin games. Our algorithm combines both techniques for 2-player Rabin games and for $2^1\!/\!_2$-player reachability games, employing a novel reduction from $2^1\!/\!_2$-player Rabin games (with quantitative winning criteria) to 2-player Rabin games (with qualitative winning criteria). A similar idea has been used to obtain a strategy improvement algorithm for $2^1\!/\!_2$-player parity games [4]; however, our present algorithm is more subtle for the following reasons. First, for parity objectives pure memoryless optimal strategies exist for both players, and the analysis of the strategy improvement algorithm for $2^1\!/\!_2$-player parity games can be restricted to pure memoryless strategies. However, the complement of a Rabin objective is a Streett objective: optimal strategies for Streett objectives require memory for $2^1\!/\!_2$-player games and even for MDPs pure optimal strategies require memory. A key insight to our analysis is as follows: once a pure memoryless strategy for a player is fixed we obtain an MDP, and in MDPs with Streett objectives randomized (not necessarily pure) memoryless optimal strategies exist. Since pure memoryless optimal strategies exist for $2^1\!/\!_2$-player games with Rabin objectives, we consider only pure memoryless strategies for the player with the Rabin objective. Then the analysis of the

[1]The reduction of games with n-states and Rabin objectives with d-pairs to parity objectives and applying the strategy improvement algorithm yields a worst case time complexity of $2^{O(n \cdot d!)}$.

counter-optimal strategies for the other player is restricted to randomized memoryless strategies. Second, the algorithm for $2^1\!/\!_2$-player parity games relies on the existence of a strategy improvement algorithm for 2-player parity games. The present algorithm does not depend on any specific algorithm to solve 2-player Rabin games, but uses as a black-box any algorithm to solve 2-player Rabin games for the improvement step. Our strategy improvement algorithm requires exponential many improvement steps in the worst-case and the running time can be bounded by $O\big(2^n \cdot (n \cdot (d+1))^{d+1}\big)$ for game graphs with n states and Rabin objectives with d-pairs. We then present a randomized strategy improvement algorithm with an expected sub-exponential number of iterations, using the techniques of [1] (note that since improvement steps need to solve 2-player Rabin games, the improvement steps may take exponential time). The expected running time of the randomized algorithm can be bounded by $O\big(2^{\sqrt{n \log(n)}} \cdot (n \cdot (d+1))^{d+1}\big)$ for game graphs with n states and Rabin objectives with d-pairs. Since pure memoryless optimal strategies exist for Rabin objectives, we obtain the algorithm for Rabin objectives. While the algorithm obtain the values for Streett objectives also, it does not directly yield an optimal strategy for Streett objectives. We then show that how, once the values are computed, to obtain an optimal strategy for Streett objectives solving certain 2-player games with Streett objectives.

2 Definitions

We consider several classes of turn-based games, namely, two-player turn-based probabilistic games ($2^1\!/\!_2$-player games), two-player turn-based deterministic games (2-player games), and Markov decision processes ($1^1\!/\!_2$-player games).

Game graphs. A *turn-based probabilistic game graph* (*$2^1\!/\!_2$-player game graph*) $G = ((S,E),(S_1,S_2,S_\bigcirc),\delta)$ consists of a directed graph (S,E), a partition (S_1, S_2, $S_\bigcirc$) of the finite set S of states, and a probabilistic transition function δ: $S_\bigcirc \to \mathcal{D}(S)$, where $\mathcal{D}(S)$ denotes the set of probability distributions over the state space S. The states in S_1 are the *player*-1 states, where player 1 decides the successor state; the states in S_2 are the *player*-2 states, where player 2 decides the successor state; and the states in $S_\bigcirc$ are the *probabilistic* states, where the successor state is chosen according to the probabilistic transition function δ. We assume that for $s \in S_\bigcirc$ and $t \in S$, we have $(s,t) \in E$ iff $\delta(s)(t) > 0$, and we often write $\delta(s,t)$ for $\delta(s)(t)$. For technical convenience we assume that every state in the graph (S,E) has at least one outgoing edge. For a state $s \in S$, we write $E(s)$ to denote the set $\{ t \in S \mid (s,t) \in E \}$ of possible successors.

A set $U \subseteq S$ of states is called *δ-closed* if for every probabilistic state $u \in U \cap S_{\bigcirc}$, if $(u,t) \in E$, then $t \in U$. The set U is called *δ-live* if for every nonprobabilistic state $s \in U \cap (S_1 \cup S_2)$, there is a state $t \in U$ such that $(s,t) \in E$. A δ-closed and δ-live subset U of S induces a *subgame graph* of G, indicated by $G \upharpoonright U$.

The *turn-based deterministic game graphs* (*2-player game graphs*) are the special case of the $2^1\!/\!_2$-player game graphs with $S_{\bigcirc} = \emptyset$. The *Markov decision processes* ($1^1\!/\!_2$*-player game graphs*) are the special case of the $2^1\!/\!_2$-player game graphs with $S_1 = \emptyset$ or $S_2 = \emptyset$. We refer to the MDPs with $S_2 = \emptyset$ as *player*-1 *MDPs*, and to the MDPs with $S_1 = \emptyset$ as *player*-2 *MDPs*.

Plays and strategies. An infinite path, or *play*, of the game graph G is an infinite sequence $\omega = \langle s_0, s_1, s_2, \ldots \rangle$ of states such that $(s_k, s_{k+1}) \in E$ for all $k \in \mathbb{N}$. We write Ω for the set of all plays, and for a state $s \in S$, we write $\Omega_s \subseteq \Omega$ for the set of plays that start from the state s.

A *strategy* for player 1 is a function σ: $S^* \cdot S_1 \to \mathcal{D}(S)$ that assigns a probability distribution to all finite sequences $\vec{w} \in S^* \cdot S_1$ of states ending in a player-1 state (the sequence represents a prefix of a play). Player 1 follows the strategy σ if in each player-1 move, given that the current history of the game is $\vec{w} \in S^* \cdot S_1$, she chooses the next state according to the probability distribution $\sigma(\vec{w})$. A strategy must prescribe only available moves, i.e., for all $\vec{w} \in S^*$, $s \in S_1$, and $t \in S$, if $\sigma(\vec{w} \cdot s)(t) > 0$, then $(s,t) \in E$. The strategies for player 2 are defined analogously. We denote by Σ and Π the set of all strategies for player 1 and player 2, respectively.

Once a starting state $s \in S$ and strategies $\sigma \in \Sigma$ and $\pi \in \Pi$ for the two players are fixed, the outcome of the game is a random walk $\omega_s^{\sigma,\pi}$ for which the probabilities of events are uniquely defined, where an *event* $\mathcal{A} \subseteq \Omega$ is a measurable set of paths. Given strategies σ for player 1 and π for player 2, a play $\omega = \langle s_0, s_1, s_2, \ldots \rangle$ is *feasible* if for every $k \in \mathbb{N}$ the following three conditions hold: (1) if $s_k \in S_{\bigcirc}$, then $(s_k, s_{k+1}) \in E$; (2) if $s_k \in S_1$, then $\sigma(s_0, s_1, \ldots, s_k)(s_{k+1}) > 0$; and (3) if $s_k \in S_2$ then $\pi(s_0, s_1, \ldots, s_k)(s_{k+1}) > 0$. Given two strategies $\sigma \in \Sigma$ and $\pi \in \Pi$, and a state $s \in S$, we denote by $\mathrm{Outcome}(s, \sigma, \pi) \subseteq \Omega_s$ the set of feasible plays that start from s given strategies σ and π. For a state $s \in S$ and an event $\mathcal{A} \subseteq \Omega$, we write $\mathrm{Pr}_s^{\sigma,\pi}(\mathcal{A})$ for the probability that a path belongs to $\mathcal{A}$ if the game starts from the state s and the players follow the strategies σ and π, respectively. In the context of player-1 MDPs we often omit the argument π, because Π is a singleton set.

We classify strategies according to their use of randomization and memory. The strategies that do not use randomization are called pure. A player-1 strategy σ is *pure* if for all $\vec{w} \in S^*$ and $s \in S_1$, there is a state $t \in S$ such that $\sigma(\vec{w} \cdot s)(t) = 1$. The pure strategies for player 2 are defined analogously. We denote by $\Sigma^P \subseteq \Sigma$ the set of pure strategies for player 1. A strategy that is not necessarily pure is called *randomized*. Let $\mathtt{M}$ be a set called *memory*. A player-1 strategy can be described as a pair of functions: a *memory-update* function σ_u: $S \times \mathtt{M} \to \mathtt{M}$ and a *next-move* function σ_m: $S_1 \times \mathtt{M} \to \mathcal{D}(S)$. The strategy (σ_u, σ_m) is *finite-memory* if the memory $\mathtt{M}$ is finite. We denote by Σ^F the set of finite-memory strategies for player 1, and by Σ^{PF} the set of *pure finite-memory* strategies; that is, $\Sigma^{PF} = \Sigma^P \cap \Sigma^F$. The strategy (σ_u, σ_m) is *memoryless* if $|\mathtt{M}| = 1$. A memoryless player-1 strategy does not depend on the history of the play but only on the current state and hence can be represented as a function σ: $S_1 \to \mathcal{D}(S)$. A *pure memoryless strategy* is a pure strategy that is memoryless. A pure memoryless strategy for player 1 can be represented as a function σ: $S_1 \to S$. We denote by Σ^M the set of memoryless strategies for player 1,and by Σ^{PM} the set of pure memoryless strategies; that is, $\Sigma^{PM} = \Sigma^P \cap \Sigma^M$. Analogously we define the family Π^M and Π^{PM} of memoryless and pure memoryless strategies for player 2.

Given a memoryless strategy $\sigma \in \Sigma^M$, let G_σ be the game graph obtained from G under the constraint that player 1 follows the strategy σ. The corresponding definition G_π for a player-2 strategy $\pi \in \Pi^M$ is analogous, and we write $G_{\sigma,\pi}$ for the game graph obtained from G if both players follow the memoryless strategies σ and π, respectively. Observe that given a $2^1\!/\!_2$-player game graph G and a memoryless player-1 strategy σ, the result G_σ is a player-2 MDP. Similarly, for a player-1 MDP G and a memoryless player-1 strategy σ, the result G_σ is a Markov chain. Hence, if G is a $2^1\!/\!_2$-player game graph and the two players follow memoryless strategies σ and π, the result $G_{\sigma,\pi}$ is a Markov chain. These observations will be useful in the analysis of $2^1\!/\!_2$-player games.

Objectives. We specify objectives for the players by providing the set of *winning plays* $\Phi \subseteq \Omega$ for each player. In this paper we study only zero-sum games [17, 11], where the objectives of the two players are strictly competitive. In other words, it is implicit that if the objective of one player is Φ, then the objective of the other player is $\Omega \setminus \Phi$. A general class of objectives are the Borel objectives [14]. A *Borel objective* $\Phi \subseteq S^\omega$ is a Borel set in the Cantor topology on S^ω. In this paper we consider *ω-regular objectives* [18] specified as Rabin and Streett objectives. which lie in the first $2^1\!/\!_2$ levels of the Borel hierarchy (i.e., in the intersection of Σ_3 and Π_3). For a play $\omega = \langle s_0, s_1, s_2, \ldots \rangle$, let $\mathrm{Inf}(\omega)$ be the set $\{ s \in S \mid s = s_k \text{ for infinitely many } k \geq 0 \}$ of states that occur infinitely often in ω. We use colors to define objectives independent of game graphs. For a set C of colors, we write $[\![\cdot]\!]\colon C \to 2^S$ for a function that maps each color to a set of states. Inversely, given a set $U \subseteq S$ of states, we write $[U] = \{ c \in C \mid [\![c]\!] \cap U \neq \emptyset \}$ for the set of colors that occur in U. Note that a state can have multiple colors.

- *Reachability objectives.* Given a set $T \subseteq S$ of "target" states, the reachability objective requires that some state of T be visited. The set of winning plays is thus $\mathrm{Reach}(T) = \{ \omega = \langle s_0, s_1, s_2, \ldots \rangle \in \Omega \mid s_k \in T \text{ for some } k \geq 0 \}$.

- *Rabin, parity, and Streett objectives.* A *Rabin objective* is specified as a set $P = \{(e_1, f_1), \ldots, (e_d, f_d)\}$ of pairs of colors $e_i, f_i \in C$. Intuitively, the Rabin condition P requires that for some $1 \leq i \leq d$, all states of color e_i be visited finitely often and some state of color f_i be visited infinitely often. Let $[\![P]\!] = \{(E_1, F_1), \ldots, (E_d, F_d)\}$ be the corresponding set of so-called *Rabin pairs*, where $E_i = [\![e_i]\!]$ and $F_i = [\![f_i]\!]$ for all $1 \leq i \leq d$. Formally, the set of winning plays is $\mathrm{Rabin}(P) = \{\omega \in \Omega \mid \exists\, 1 \leq i \leq d.\ (\mathrm{Inf}(\omega) \cap E_i = \emptyset \wedge \mathrm{Inf}(\omega) \cap F_i \neq \emptyset)\}$. Without loss of generality, we require that $\left(\bigcup_{i \in \{1,2,\ldots,d\}} (E_i \cup F_i)\right) = S$. The *parity* (or *Rabin-chain*) objectives are the special case of Rabin objectives such that $E_1 \subset F_1 \subset E_2 \subset F_2 \ldots \subset E_d \subset F_d$. A *Streett objective* is again specified as a set $P = \{(e_1, f_1), \ldots, (e_d, f_d)\}$ of pairs of colors. The Streett condition P requires that for each $1 \leq i \leq d$, if some state of color f_i is visited infinitely often, then some state of color e_i be visited infinitely often. Formally, the set of winning plays is $\mathrm{Streett}(P) = \{ \omega \in \Omega \mid \forall\, 1 \leq i \leq d.\ (\mathrm{Inf}(\omega) \cap E_i \neq \emptyset \ \vee\ \mathrm{Inf}(\omega) \cap F_i = \emptyset) \}$, for the set $[\![P]\!] = \{(E_1, F_1), \ldots, (E_d, F_d)\}$ of so-called *Streett pairs*. Note that the Rabin and Streett objectives are dual; i.e., the complement of a Rabin objective is a Streett objective, and vice versa. Moreover, every parity objective is both a Rabin objective and a Streett objective.

Sure winning, almost-sure winning, and optimality. Given a player-1 objective Φ, a strategy $\sigma \in \Sigma$ is *sure winning* for player 1 from a state $s \in S$ if for every strategy $\pi \in \Pi$ for player 2, we have $\mathrm{Outcome}(s, \sigma, \pi) \subseteq \Phi$. The strategy σ is *almost-sure winning* for player 1 from the state s for the objective Φ if for every player-2 strategy π, we have $\Pr_s^{\sigma,\pi}(\Phi) = 1$. The sure and almost-sure winning strategies for player 2 are defined analogously. Given an objective Φ, the *sure winning set* $\langle\langle 1 \rangle\rangle_{sure}(\Phi)$ for player 1 is the set of states from which player 1 has a sure winning strategy. The *almost-sure winning set* $\langle\langle 1 \rangle\rangle_{almost}(\Phi)$ for player 1 is the set of states from which player 1 has an almost-sure winning strategy. The sure winning set $\langle\langle 2 \rangle\rangle_{sure}(\Omega \setminus \Phi)$ and the almost-sure winning set $\langle\langle 2 \rangle\rangle_{almost}(\Omega \setminus \Phi)$ for player 2 are defined analogously. It follows from the definitions that for all $2^1\!/\!_2$-player game graphs and all objectives Φ, we have $\langle\langle 1 \rangle\rangle_{sure}(\Phi) \subseteq \langle\langle 1 \rangle\rangle_{almost}(\Phi)$. A game is sure (resp. almost-sure) winning for player i, if player i wins surely (resp. almost-surely) from every state in the game. Computing sure and almost-sure winning sets and strategies is referred to as the *qualitative* analysis of $2^1\!/\!_2$-player games [9].

Given ω-regular objectives $\Phi \subseteq \Omega$ for player 1 and $\Omega \setminus \Phi$ for player 2, we define the *value* functions $\langle\langle 1 \rangle\rangle_{val}$ and $\langle\langle 2 \rangle\rangle_{val}$ for the players 1 and 2, respectively, as the following functions from the state space S to the interval $[0, 1]$ of reals: for all states $s \in S$, let $\langle\langle 1 \rangle\rangle_{val}(\Phi)(s) = \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \Pr_s^{\sigma,\pi}(\Phi)$ and $\langle\langle 2 \rangle\rangle_{val}(\Omega \setminus \Phi)(s) = \sup_{\pi\in\Pi} \inf_{\sigma\in\Sigma} \Pr_s^{\sigma,\pi}(\Omega \setminus \Phi)$. In other words, the value $\langle\langle 1 \rangle\rangle_{val}(\Phi)(s)$ gives the maximal probability with which player 1 can achieve her objective Φ from state s, and analogously for player 2. The strategies that achieve the value are called optimal: a strategy σ for player 1 is *optimal* from the state s for the objective Φ if $\langle\langle 1 \rangle\rangle_{val}(\Phi)(s) = \inf_{\pi\in\Pi} \Pr_s^{\sigma,\pi}(\Phi)$. The optimal strategies for player 2 are defined analogously. Computing values is referred to as the *quantitative* analysis of $2^1\!/\!_2$-player games. The set of states with value 1 is called the *limit-sure winning set* [9]. For $2^1\!/\!_2$-player game graphs with ω-regular objectives the almost-sure and limit-sure winning sets coincide [5].

Let $\mathcal{C} \in \{P, M, F, PM, PF\}$ and consider a family $\Sigma^{\mathcal{C}} \subseteq \Sigma$ of special strategies for player 1. We say that the family $\Sigma^{\mathcal{C}}$ *suffices* with respect to a player-1 objective Φ on a class $\mathcal{G}$ of game graphs for *sure winning* if for every game graph $G \in \mathcal{G}$ and state $s \in \langle\langle 1 \rangle\rangle_{sure}(\Phi)$, there is a player-1 strategy $\sigma \in \Sigma^{\mathcal{C}}$ such that for every player-2 strategy $\pi \in \Pi$, we have $\mathrm{Outcome}(s, \sigma, \pi) \subseteq \Phi$. Similarly, the family $\Sigma^{\mathcal{C}}$ *suffices* with respect to the objective Φ on the class $\mathcal{G}$ of game graphs for *almost-sure winning* if for every game graph $G \in \mathcal{G}$ and state $s \in \langle\langle 1 \rangle\rangle_{almost}(\Phi)$, there is a player-1 strategy $\sigma \in \Sigma^{\mathcal{C}}$ such that for every player-2 strategy $\pi \in \Pi$, we have $\Pr_s^{\sigma,\pi}(\Phi) = 1$; and for *optimality*, if for every game graph $G \in \mathcal{G}$ and state $s \in S$, there is a player-1 strategy $\sigma \in \Sigma^{\mathcal{C}}$ such that $\langle\langle 1 \rangle\rangle_{val}(\Phi)(s) = \inf_{\pi\in\Pi} \Pr_s^{\sigma,\pi}(\Phi)$.

For sure winning, the $1^1\!/\!_2$-player and $2^1\!/\!_2$-player games coincide with 2-player (deterministic) games where the random player (who chooses the successor at the probabilistic states) is interpreted as an adversary, i.e., as player 2. Theorem 1 and Theorem 2 state the classical determinacy results for 2-player and $2^1\!/_2$-player game graphs with Rabin and Streett objectives.

Theorem 1 (Qualitative determinacy [10]**)** *For all 2-player game graphs and Rabin objectives* Φ *and Streett objective* $\Omega \setminus \Phi$*, we have* $\langle\langle 1\rangle\rangle_{sure}(\Phi) = S \setminus \langle\langle 2\rangle\rangle_{sure}(\Omega \setminus \Phi)$*. Moreover, on 2-player game graphs, the family of pure memoryless strategies suffices for sure winning with respect to Rabin objectives and the family of pure finite-memory strategies suffices for sure winning with respect to Streett objectives.*

Theorem 2 (Quantitative determinacy [3]**)** *For all* $2^1\!/_2$*-player game graphs, all Rabin objectives* Φ*, all Streett objectives* $\Omega \setminus \Phi$*, and all states* s*, we have* $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) + \langle\langle 2\rangle\rangle_{val}(\Omega \setminus \Phi)(s) = 1$*. The family of pure memoryless strategies suffices for optimality with respect to Rabin objectives and the family of pure finite-memory strategies suffices for sure winning with respect to Streett objectives on* $2^1\!/_2$*-player game graphs.*

Since in $2^1\!/_2$-player games with Rabin objectives, pure memoryless strategies suffices for optimality, in sequel we consider only pure memoryless strategies for the player with Rabin objective. Moreover, since Rabin and Streett objectives are infinitary objectives the following proposition is immediate.

Proposition 1 (Optimality conditions) *For a Rabin objective* Φ*, for every* $s \in S$ *the following conditions hold.*

1. *If* $s \in S_1$*, then for all* $t \in E(s)$ *we have* $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) \geq \langle\langle 1\rangle\rangle_{val}(\Phi)(t)$*, and for some* $t \in E(s)$ *we have* $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) = \langle\langle 1\rangle\rangle_{val}(\Phi)(t)$*.*
2. *If* $s \in S_2$*, then for all* $t \in E(s)$ *we have* $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) \leq \langle\langle 1\rangle\rangle_{val}(\Phi)(t)$*, and for some* $t \in E(s)$ *we have* $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) = \langle\langle 1\rangle\rangle_{val}(\Phi)(t)$*.*
3. *If* $s \in S_{\bigcirc}$*, then* $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) = \big(\sum_{t \in E(s)} \langle\langle 1\rangle\rangle_{val}(\Phi)(t)\ \delta(s,t)\big)$*.*

Similar conditions hold for the value function $\langle\langle 2\rangle\rangle_{val}(\Omega \setminus \Phi)$ *of player 2.*

3 Strategy Improvement for $2^1\!/_2$-player Rabin and Streett Games

In section 3.1 we first recall a few key properties of $2^1\!/_2$-player games with Rabin objectives that were proved in [3]. We use the properties in section 3.2 to develop a strategy improvement algorithm for $2^1\!/_2$-player games with Rabin objectives.

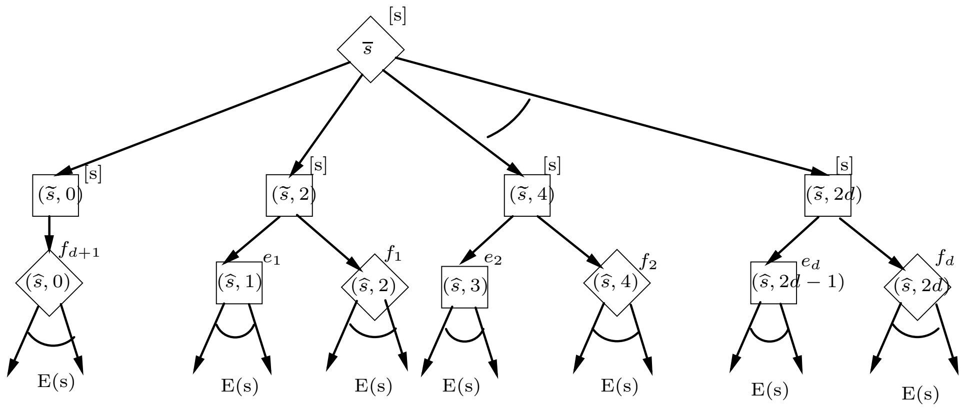

Figure 1: Gadget for the reduction of 2½-player Rabin games to 2-player Rabin games.

3.1 Key properties

We present a reduction of 2½-player parity games to 2-player parity games preserving the ability of player 1 to win almost-surely.

Reduction. Given a 2½-player game graph $G = ((S, E), (S_1, S_2, S_{\bigcirc}), \delta)$, a set $C = \{e_1, f_1, \ldots, e_d, f_d\}$ of colors, and a color map $[\cdot]\colon S \to 2^C \setminus \emptyset$, we construct a 2-player game graph $\overline{G} = ((\overline{S}, \overline{E}), (\overline{S}_1, \overline{S}_2), \overline{\delta})$ together with a color map $[\cdot]\colon \overline{S} \to 2^{\overline{C}} \setminus \emptyset$ for the extended color set $\overline{C} = C \cup \{e_{d+1}, f_{d+1}\}$. The construction is specified as follows. For every nonprobabilistic state $s \in S_1 \cup S_2$, there is a corresponding state $\overline{s} \in \overline{S}$ such that (1) $\overline{s} \in \overline{S}_1$ iff $s \in S_1$, and (2) $[\overline{s}] = [s]$, and (3) $(\overline{s}, \overline{t}) \in \overline{E}$ iff $(s, t) \in E$. Every probabilistic state $s \in S_{\bigcirc}$ is replaced by the gadget shown in Figure 1. In the figure, diamond-shaped states are player-2 states (in $\overline{S}_2$), and square-shaped states are player-1 states (in $\overline{S}_1$). From the state $\overline{s}$ with $[\overline{s}] = [s]$, the players play the following 3-step game in $\overline{G}$. First, in state $\overline{s}$ player 2 chooses a successor $(\widetilde{s}, 2k)$, for $k \in \{0, 1, \ldots, d\}$. For every state $(\widetilde{s}, 2k)$, we have $[(\widetilde{s}, 2k)] = [s]$. For $k > 1$, in state $(\widetilde{s}, 2k)$ player 1 chooses from two successors: state $(\widehat{s}, 2k-1)$ with $[(\widehat{s}, 2k-1)] = e_k$, or state $(\widehat{s}, 2k)$ with $[(\widehat{s}, 2k)] = f_k$. The state $(\widetilde{s}, 0)$ has only one successor $(\widehat{s}, 0)$, with $[(\widehat{s}, 0)] = f_{d+1}$. Note that no state in $\overline{S}$ is labeled by the new color e_{d+1}, that is, $[\![e_{d+1}]\!] = \emptyset$. Finally, in each state $(\widehat{s}, j)$ the choice is between all states $\overline{t}$ such that $(s, t) \in E$, and it belongs to player 1 if k is odd, and to player 2 if k is even.

We consider $2^1\!/_2$-player games played on the graph G with $P = \{(e_1, f_1), \ldots, (e_d, f_d)\}$ and the Rabin objective $\mathrm{Rabin}(P)$ for player 1. We denote by $\overline{G} = \mathrm{Tr}^1_{\mathrm{as}}(G)$ the 2-player game, with Rabin objective $\mathrm{Rabin}(\overline{P})$, where $\overline{P} = \{(e_1, f_1), \ldots, (e_{d+1}, f_{d+1})\}$, as defined by the reduction above. Also given a strategy (pure memoryless) $\overline{\sigma}$ in the 2-player game $\overline{G}$, a strategy $\sigma = \mathrm{Tr}^1_{\mathrm{as}}(\overline{\sigma})$ in the $2^1\!/_2$-player game G is defined as follows:

$$\sigma(s) = t, \text{ if and only if } \overline{\sigma}(\overline{s}) = \overline{t}; \text{ for all } s \in S_1.$$

Similar definitions hold for player 2.

Lemma 1 ([3]) *Given a $2^1\!/_2$-player game graph G with the Rabin objective Rabin(P) for player 1, let $\overline{U}_1$ and $\overline{U}_2$ be the sure winning sets for players 1 and 2, respectively, in the 2-player game graph $\overline{G} = \mathrm{Tr}^1_{\mathrm{as}}(G)$ with the modified parity objective Rabin($\overline{P}$). Define the sets U_1 and U_2 in the original $2^1\!/_2$-player game graph G by $U_1 = \{ s \in S \mid \overline{s} \in \overline{U}_1 \}$ and $U_2 = \{ s \in S \mid \overline{s} \in \overline{U}_2 \}$. Then the following assertions hold:*

1. *(a) $U_1 = \langle\langle 1 \rangle\rangle_{almost}(Rabin(P)) = (S \setminus U_2)$; and*
2. *(b) if $\overline{\sigma}$ is a pure memoryless sure winning strategy for player 1 from $\overline{U}_1$ in $\overline{G}$, then $\sigma = \mathrm{Tr}^1_{\mathrm{as}}(\overline{\sigma})$ is an almost-sure winning strategy for player 1 from U_1 in G.*

A similar reduction exists that preserves almost-sure winning for player 2 (i.e., the player with Streett objective) and we refer to the reduction for player 2 as $\mathrm{Tr}^2_{\mathrm{as}}$. Also there is a simple mapping of finite-memory sure winning strategies $\overline{\pi}$ in $\mathrm{Tr}^2_{\mathrm{as}}(G)$ to finite-memory almost-sure winning strategy $\pi = \mathrm{Tr}^2_{\mathrm{as}}(\overline{\pi})$ in G.

Boundary probabilistic states. Given a set U of states, let $Bou(U) = \{ s \in U \cap S_{\bigcirc} \mid \exists t \in E(s), t \not\in U \}$, be the set of *boundary* probabilistic states that have an edge out of U. Given a set U of states and a Rabin objective $\mathrm{Rabin}(P)$ for player 1, we define two transformations $\mathrm{Tr}_{\mathrm{win}_1}(U)$ and $\mathrm{Tr}_{\mathrm{win}_2}(U)$ of U as follows: every state s in $Bou(U)$ is converted to an *absorbing* state (state with only a self-loop) and (a) in $\mathrm{Tr}_{\mathrm{win}_1}(U)$ it is assigned the color f_1 and (b) in $\mathrm{Tr}_{\mathrm{win}_2}(U)$ it is assigned the color e_1; i.e., every state in $Bou(U)$ is converted to a sure winning state for player 1 in $\mathrm{Tr}_{\mathrm{win}_1}(U)$ and every state in $Bou(U)$ is converted to a sure winning state for player 2 in $\mathrm{Tr}_{\mathrm{win}_2}(U)$. Observe that if U is δ-live, then $\mathrm{Tr}_{\mathrm{win}_1}(G \upharpoonright U)$ and $\mathrm{Tr}_{\mathrm{win}_2}(G \upharpoonright U)$ is a game graph.

Value classes. Given a Rabin objective Φ, for every real $r \in \mathbb{R}$ the *value class* with value r, $\mathrm{VC}(r) = \{ s \in S \mid \langle\langle 1 \rangle\rangle_{val}(\Phi)(s) = r \}$, is the set of states with value r for player 1. It follows from Proposition 1 that for every $r > 0$, the value class $\mathrm{VC}(r)$ is δ-live. The following lemma establishes a connection between value classes, the transformations $\mathrm{Tr}_{\mathrm{win}_1}$ and $\mathrm{Tr}_{\mathrm{win}_2}$ and the almost-sure winning states.

Lemma 2 (Almost-sure winning reduction[3]) *The following assertions hold.*

1. *For every value class* $\mathrm{VC}(r)$*, for* $r > 0$*, the game* $\mathrm{Tr}_{\mathrm{win}_1}(G \restriction \mathrm{VC}(r))$ *is almost-sure winning for player 1.*

2. *For every value class* $\mathrm{VC}(r)$*, for* $r < 1$*, the game* $\mathrm{Tr}_{\mathrm{win}_2}(G \restriction \mathrm{VC}(r))$ *is almost-sure winning for player 2.*

Lemma 3 (Optimal strategies[3]) *The following assertions hold.*

1. *If a strategy* σ *is an almost-winning strategy in the game* $\mathrm{Tr}_{\mathrm{win}_1}(G \restriction \mathrm{VC}(r))$*, for every value class* $\mathrm{VC}(r)$*, then* σ *is an optimal strategy.*

2. *If a strategy* π *is an almost-winning strategy in the game* $\mathrm{Tr}_{\mathrm{win}_2}(G \restriction \mathrm{VC}(r))$*, for every value class* $\mathrm{VC}(r)$*, then* π *is an optimal strategy.*

It follows from Lemma 1 and Lemma 2, that for every value class $\mathrm{VC}(r)$, with $r > 0$, the game $\mathrm{Tr}^1_{\mathrm{as}}(\mathrm{Tr}_{\mathrm{win}_1}(G \restriction \mathrm{VC}(r)))$ is sure winning for player 1.

Properties of almost-sure winning states. The results of [9] shows that for ω-regular objectives specified as parity objectives if the set of limit-winning states for a player is empty, then the other player wins almost-surely from all states in the game. Since Rabin and Streett objectives can be reduced to parity objectives [18] and in $2^1\!/\!_2$-player games limit-sure and almost-sure winning sets coincide we have the following result.

Lemma 4 *Given a* $2^1\!/\!_2$*-player game* G *and a Rabin objective* $\mathit{Rabin}(P)$ *if* $\langle\!\langle 1 \rangle\!\rangle_{almost}(\mathit{Rabin}(P)) = \emptyset$*, then* $\langle\!\langle 2 \rangle\!\rangle_{almost}(\Omega \setminus \mathit{Rabin}(P)) = S$*.*

Property of MDPs with Streett objectives. The following lemma is a result from [2].

Lemma 5 ([2]) *The family of randomized memoryless strategies suffices for optimality with respect to Streett objectives on MDPs.*

3.2 Strategy Improvement Algorithm

We now present an algorithm to compute values for 2½-player games with Rabin objective $\text{Rabin}(P)$ for player 1. By quantitative determinacy (Theorem 2) the algorithm also computes values for Streett objective $\text{Streett}(P)$ for player 2. Recall that since pure memoryless strategies exist for Rabin objectives we will only consider pure memoryless strategies σ for player 1.

Notation. Given a strategy σ and a set U of states, we denote by $(\sigma \upharpoonright U)$ a strategy that for every state in U follows the strategy σ.

Values and value class given strategies. Given a player-1 strategy σ and a Rabin objective Φ, we denote the value of player 1 given the strategy σ as follows: $\langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(s) = \inf_{\pi\in\Pi} \text{Pr}_s^{\sigma,\pi}(\Phi)$. Similarly we define the value classes given strategy σ as $\text{VC}^{\sigma}(r) = \{ s \in S \mid \langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(s) = r \}$.

Ordering of strategies. We define an ordering relation $\prec$ on strategies as follows: given two strategies σ and σ', we have $\sigma \prec \sigma'$ if and only if

- for all states s we have $\langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(s) \leq \langle\langle 1\rangle\rangle_{val}^{\sigma'}(\Phi)(s)$ and for some state s we have $\langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(s) < \langle\langle 1\rangle\rangle_{val}^{\sigma'}(\Phi)(s)$.

Improve strategy. Given a strategy σ for player 1, we describe a procedure `Improve` to "improve" the strategy for player 1. The procedure is described in Algorithm 1. An informal description of the procedure is as follows: given a strategy σ, the algorithm computes the values $\langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(s)$ for all states. Since σ is a pure memoryless strategy, $\langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(s)$ can be computed by solving the MDP G_{σ} with the Streett objective $\Omega \setminus \Phi$. If there is a state $s \in S_1$, such that the strategy can be "value improved", i.e., there is a state $t \in E(s)$, with $\langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(t) > \langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(s)$, then the strategy σ is modified by setting $\sigma(s)$ to t. This is achieved in Step 2.1 of `Improve`. Else in every value class $\text{VC}^{\sigma}(r)$, the strategy σ is "improved" for the game $\text{Tr}_{\text{as}}^{1}(\text{Tr}_{\text{win}_2}(G \upharpoonright \text{VC}^{\sigma}(r)))$ by solving the 2-player game $\text{Tr}_{\text{as}}^{1}(\text{Tr}_{\text{win}_2}(G \upharpoonright \text{VC}^{\sigma}(r)))$ by an algorithm to solve 2-player Rabin games. The computation of `Improve` is discussed in Lemma 11. In the algorithm the strategy σ for player 1 is always a pure memoryless strategy (this is sufficient since pure memoryless strategies suffices for optimality in 2½-player games with Rabin objectives (Theorem 2)). Moreover, given a pure memoryless strategy σ the game G_{σ} is a player-2 MDP and by Lemma 5 there is a randomized memoryless counter-optimal strategy for player 2. Hence fixing a pure memoryless strategy for player 1 we only consider randomized memoryless strategies for player 2.

Algorithm 1 `Improve`

Input : A 2½-player game G with Rabin objective Φ for player 1 and a strategy σ for player 1.
Output: A strategy σ' for player 1 such that either $\sigma' = \sigma$ or $\sigma \prec \sigma'$.
1. (Step 1.) Compute $\langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(s)$ for all states s.
2. (Step 2.) Consider the set $I = \{ s \in S_1 \mid \exists t \in E(s).\ \langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(t) > \langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(s) \}$.
   2.1 (Value improvement.) **if** $I \neq \emptyset$, then set σ' as follows:
      $\sigma'(s) = \sigma(s)$ for $s \in S_1 \setminus I$; and
      $\sigma'(s) = t$ for $s \in I$, and $t \in E(s)$, such that $\langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(t) > \langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(s)$.
   2.2 (Qualitative improvement.) **else** for every value class $\mathrm{VC}^{\sigma}(r)$,
      let $\overline{G}_r$ be the 2-player game $\mathrm{Tr}_{\mathrm{as}}^1(\mathrm{Tr}_{\mathrm{win}_2}(G \upharpoonright \mathrm{VC}^{\sigma}(r)))$
      For every r, solve the Rabin game $\overline{G}_r$ by `TwoPlRabinGame`$(\overline{G}_r)$
         (by a two-player game solving algorithm for Rabin games).
      If for some r, we have a non-empty sure winning set $\overline{U}_r$ for player 1 in $\overline{G}_r$,
      let $\overline{\sigma}$ be the sure winning strategy for player 1 in $\overline{U}_r$ and
      U_r be the corresponding set in G,
      set $(\sigma' \upharpoonright U_r) = \mathrm{Tr}_{\mathrm{as}}^1(\overline{\sigma} \upharpoonright \overline{U}_r)$.
3. **return** σ'.

Proposition 2 *Given a strategy σ for player 1, for every state $s \in \mathrm{VC}^{\sigma}(r) \cap S_2$, if $t \in E(s)$, then we have $\langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(t) \geq r$, i.e., $E(s) \subseteq \bigcup_{q \geq r} \mathrm{VC}^{\sigma}(q)$.*

Proof. The result is proved by contradiction. Suppose the assertion of the proposition fails, i.e., there exists s and $t \in E(s)$, such that $s \in \mathrm{VC}^{\sigma}(r)$ and $\langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(t) < r$, then consider the strategy $\pi \in \Pi$ for player 2 that at s chooses successor t, and from t ensures Φ is satisfied with probability at most $\langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(t)$ against strategy σ. Hence we have $\langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(s) \leq \langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(t) < r$. This contradicts that $s \in \mathrm{VC}^{\sigma}(r)$. Hence player 2 can only choose edges with the target of the edge in equal or higher value classes. ∎

Rabin winning set. A set $C \subseteq S$ is Rabin winning for a Rabin objective $\mathrm{Rabin}(P)$, if for all plays ω with $\mathrm{Inf}(\omega) = C$ we have $\omega \in \mathrm{Rabin}(P)$.

Proposition 3 *Given a strategy σ for player 1, for all strategies $\pi \in \Pi^M$ for player 2, if there is a closed recurrent class C in the Markov chain $G_{\sigma,\pi}$, with $C \subseteq \mathrm{VC}^{\sigma}(r)$, for $r > 0$, then C is Rabin winning.*

Proof. The result is again proved by contradiction. Suppose the assertion of the proposition fails, i.e., for some strategy $\pi \in \Pi^M$ for player 2, for some $r > 0$, C is a closed recurrent class in the Markov chain $G_{\sigma,\pi}$, with $C \subseteq \mathrm{VC}^\sigma(r)$ and C is not Rabin winning. Then player 2 by playing strategy π ensures that for all states $s \in C$ we have $\mathrm{Pr}_s^{\sigma,\pi}(\Phi) = 0$ (since C is not Rabin winning and given C is a closed recurrent class, all states in C are visited infinitely often). This contradicts that $C \subseteq \mathrm{VC}^\sigma(r)$ and $r > 0$. ∎

Lemma 6 *Consider a strategy σ to be an input to Algorithm 1, and let σ' be an output, i.e., $\sigma' =$ `Improve`(G, σ). If the set I in Step 2 of Algorithm 1 is non-empty, then we have*

$$\langle\langle 1\rangle\rangle_{val}^{\sigma'}(\Phi)(s) \geq \langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(s) \ \ \forall s \in S; \qquad \langle\langle 1\rangle\rangle_{val}^{\sigma'}(\Phi)(s) > \langle\langle 1\rangle\rangle_{val}^{\sigma}(\Phi)(s) \ \ \forall s \in I.$$

Proof. Consider a switch of the strategy of player 1 from σ to σ', as constructed when Step 2.1 of Algorithm 1. Consider a strategy $\pi \in \Pi^M$ for player 2 and a closed recurrent class C in $G_{\sigma',\pi}$ such that $C \subseteq \bigcup_{r>0} \mathrm{VC}^\sigma(r)$. Let $z = \max\{\, r > 0 \mid C \cap \mathrm{VC}^\sigma(r) \neq \emptyset \,\}$, i.e., $\mathrm{VC}^\sigma(z)$ is the greatest value class with non-empty intersection with C. A state $s \in \mathrm{VC}^\sigma(z) \cap C$ satisfy the following conditions.

1. If $s \in S_2$, then we have $\mathrm{Supp}(\pi(s)) \subseteq \mathrm{VC}^\sigma(z)$. This follows since by Proposition 2 we have $E(s) \subseteq \bigcup_{q \geq z} \mathrm{VC}^\sigma(q)$ and $C \cap \mathrm{VC}^\sigma(q) = \emptyset$ for $q > z$.

2. If $s \in S_1$, then $\sigma'(s) \in \mathrm{VC}^\sigma(z)$. This follows since by construction $\sigma'(s) \in \bigcup_{q \geq z} \mathrm{VC}^\sigma(q)$ and $C \cap \mathrm{VC}^\sigma(q) = \emptyset$ for $q > z$. Also since $s \in \mathrm{VC}^\sigma(z)$ and $\sigma'(s) \in \mathrm{VC}^\sigma(z)$, it follows that $\sigma'(s) = \sigma(s)$.

3. If $s \in S_\bigcirc$, then $E(s) \subseteq \mathrm{VC}^\sigma(z)$. This follows because for $s \in S_\bigcirc$ if $E(s) \subsetneq \mathrm{VC}^\sigma(z)$, then $E(s) \cap (\bigcup_{q>z} \mathrm{VC}^\sigma(q)) \neq \emptyset$. Since C is closed, and $C \cap \mathrm{VC}^\pi(q) = \emptyset$ for $q > z$, the claim follows.

It follows that $C \subseteq \mathrm{VC}^\sigma(z)$ and for all states $s \in C \cap S_1$, we have $\sigma'(s) = \sigma(s)$. Hence by Proposition 3 we have C is Rabin winning.

It follows that if player 1 switches to the strategy σ', as constructed when Step 2.1 of Algorithm 1 is executed, then for all strategies $\pi \in \Pi^M$ for player 2 the following assertion hold: if there is a closed recurrent class $C \subseteq (S \setminus \mathrm{VC}^\sigma(0))$ in the Markov chain $G_{\sigma',\pi}$, then C is Rabin winning for player 1. Hence given strategy σ', a counter-optimal strategy for player 2 maximizes the probability to reach $\mathrm{VC}^\sigma(0)$. The desired result follows from arguments similar to 2½-player games with reachability objectives [7], with $\mathrm{VC}^\sigma(0)$ as the target for player 2, and the value improvement step (Step 2.1 of Algorithm 1). ∎

Lemma 7 *Consider a strategy σ to be an input to Algorithm 1, and let σ' be an output, i.e., $\sigma' = \texttt{Improve}(G, \sigma)$, such that $\sigma' \neq \sigma$. If the set I in Step 2 of Algorithm 1 is empty, then*

1. *for all states s we have $\langle\langle 1 \rangle\rangle_{val}^{\sigma'}(\Phi)(s) \geq \langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(s)$; and*
2. *for some state s we have $\langle\langle 1 \rangle\rangle_{val}^{\sigma'}(\Phi)(s) > \langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(s)$.*

Proof. It follows from Proposition 3 that for all strategies $\pi \in \Pi^M$ for player 2, if C is a closed recurrent class in $G_{\sigma,\pi}$ and $C \subseteq \mathrm{VC}^\sigma(r)$, for $r > 0$, then C is Rabin winning. Let σ' be the strategy constructed from σ in Step 2.2 of Algorithm 1. The set U_r where σ is modified to obtain σ', the strategy $\sigma' \upharpoonright U_r$ is an almost-winning strategy in U_r in the sub-game $\mathrm{Tr}_{\mathrm{win}_2}(G \upharpoonright \mathrm{VC}^\sigma(r))$ This follows from Lemma 1 since $\sigma' \upharpoonright U_r = \mathrm{Tr}_{\mathrm{as}}^1(\overline{\sigma} \upharpoonright \overline{U}_r)$ and $\overline{\sigma} \upharpoonright \overline{U}_r$ is a sure-winning strategy for player 1 in $\overline{U}_r$ in the sub-game $\mathrm{Tr}_{\mathrm{as}}^1(\mathrm{Tr}_{\mathrm{win}_2}(G \upharpoonright \mathrm{VC}^\sigma(r)))$. It follows that if C is a closed recurrent class in $G_{\sigma',\pi}$ and $C \subseteq \mathrm{VC}^\sigma(r)$, then C is Rabin winning. Arguments similar to Lemma 6 shows that the following assertion hold: for all strategies $\pi \in \Pi^M$ for player 2, if there is a closed recurrent class $C \subseteq (S \setminus \mathrm{VC}^\sigma(0))$ in the Markov chain $G_{\sigma',\pi}$, then C is Rabin winning. Since in strategy σ' player 1 chooses every edge in the same value class as σ, it can be shown that for all states s we have $\langle\langle 1 \rangle\rangle_{val}^{\sigma'}(\Phi)(s) \geq \langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(s)$. If $\sigma \neq \sigma'$, then the set U_r where the strategy σ is modified is non-empty. Since $\sigma' \upharpoonright U_r$ is an almost-winning strategy in U_r in $\mathrm{Tr}_{\mathrm{win}_2}(G \upharpoonright \mathrm{VC}^\sigma(r))$, it follows that given σ', any counter-optimal strategy $\pi \in \Pi^M$ of player 2 either moves to a higher value class or player 1 wins almost-surely in U_r. In either case for a state $s \in U_r$ we have $\langle\langle 1 \rangle\rangle_{val}^{\sigma'}(\Phi)(s) > \langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(s)$. ∎

Lemma 6 and Lemma 7 yields Lemma 8.

Lemma 8 *For a strategy σ, if $\sigma \neq \texttt{Improve}(G, \sigma)$, then $\sigma \prec \texttt{Improve}(G, \sigma)$.*

The key argument to establish that if a strategy σ satisfy that $\sigma = \texttt{Improve}(G, \sigma)$, then σ is an optimal strategy is as follows: let σ be a strategy such that $\sigma = \texttt{Improve}(G, \sigma)$. It follows that the strategy σ cannot be "value-improved". Moreover, for all value-classes $\mathrm{VC}^\sigma(r)$ we have the set of almost-winning set in $\mathrm{Tr}_{\mathrm{win}_2}(G \upharpoonright \mathrm{VC}^\sigma(r))$ for player 1 is empty. Hence by Lemma 4 we have all states in $\mathrm{Tr}_{\mathrm{win}_2}(G \upharpoonright \mathrm{VC}^\sigma(r))$ is almost-winning for player 2. Consider a strategy π for player 2 such that for all value class $\mathrm{VC}^\sigma(r)$, the strategy π is almost-winning in $\mathrm{Tr}_{\mathrm{win}_2}(G \upharpoonright \mathrm{VC}^\sigma(r))$. Given π, for all strategies σ of player 1 and for all states $s \in (S \setminus \mathrm{VC}^\sigma(1))$ we have $\mathrm{Pr}_s^{\sigma,\pi}(\Phi \mid \mathrm{Safe}((S \setminus \mathrm{VC}^\sigma(1))) = 0$. Hence given the strategy π, any counter-optimal strategy for player 1 maximizes the probability to reach $\mathrm{VC}^\sigma(1)$.

Algorithm 2 StrategyImprovementAlgorithm

Input : A $2^1\!/\!_2$-player game G with Rabin objective Φ for player 1.
Output: An optimal strategy σ^* for player 1.
1. Pick an arbitrary strategy σ for player 1.
2. **while** $\sigma \neq \texttt{Improve}(G,\sigma)$
   do $\sigma = \texttt{Improve}(G,\sigma)$.
3. **return** $\sigma^* = \sigma$.

Since the strategy σ cannot be "value improved" it follows from arguments similar to [7] for $2^1\!/\!_2$-player reachability games that for all strategies σ', for all states $s \in \mathrm{VC}^\sigma(r)$, we have $\mathrm{Pr}_s^{\sigma',\pi}(\Phi) \leq r$. Hence we have $\langle\langle 1 \rangle\rangle_{val}(\Phi)(s) \leq r$. For all states $s \in \mathrm{VC}^\sigma(r)$, we have $r = \langle\langle 1 \rangle\rangle_{val}^{\sigma}(\Phi)(s) \leq \langle\langle 1 \rangle\rangle_{val}(\Phi)(s)$. This establishes optimality of σ, and yields the following lemma.

Lemma 9 *For a strategy σ if $\sigma = \texttt{Improve}(G,\sigma)$, then σ is an optimal strategy for player 1.*

A strategy improvement algorithm using the `Improve` procedure is described in Algorithm 2. Observe that it follows from Lemma 8 that if Algorithm 2 outputs a strategy σ^*, then $\sigma^* = \texttt{Improve}(G,\sigma^*)$. The correctness of the algorithm follows from Lemma 9 and yields Theorem 2. Given an optimal strategy σ for player 1, the values for both the players can be computed in polytime by computing the values of the MDP G_σ [3, 8]. Since there are at most 2^n possible pure memoryless strategies, it easily follows that Algorithm 2 requires at most 2^n-iterations. Each iteration can be computed in time $O\big((n \cdot (d+1))^{d+1}\big)$ for game graphs with n states and Rabin objectives with d-pairs. (see Lemma 11 and the following discussion). This gives us the following theorem.

Theorem 3 (Correctness of Algorithm 2) *Let σ^* be an output of Algorithm 2. Then the strategy σ^* is an optimal strategy for player 1. The running time of Algorithm 2 can be bounded by $O(2^n \cdot (n \cdot (d+1))^{d+1})$ on game graphs with n states and Rabin objectives with d-pairs.*

4 Randomized Algorithm

We now present a randomized algorithm for $2^1\!/\!_2$-player Rabin games, by combining an algorithm of Björklund et al. [1] and the procedure `Improve`.

Games and improving subgames. Given $l, m \in \mathbb{N}$, let $\mathcal{G}(l,m)$ be the class of $2^1\!/\!_2$-player game graphs with the set S_1 of player 1 states partitioned into two sets as follows: (a) $O_1 = \{s \in S_1 \mid |E(s)| = 1\}$, i.e., the set of states with out-degree 1; and (b) $O_2 = S_2 \setminus O_1$, with $O_2 \leq l$ and $\sum_{s \in O_2} |E(s)| \leq m$. There is no restriction for player 2. Given a game $G \in \mathcal{G}(l,m)$, a state $s \in O_2$, and an edge $e = (s,t)$, we define the subgame $\widetilde{G}_e$ by deleting all edges from s other than the edge e. Observe that $\widetilde{G}_e \in \mathcal{G}(l-1, m-|E(s)|)$, and hence also $\widetilde{G}_e \in \mathcal{G}(l,m)$. If σ is a strategy for player 1 in $G \in \mathcal{G}(l,m)$, then a subgame $\widetilde{G}$ is *σ-improving* if some strategy σ' in $\widetilde{G}$ satisfies that $\sigma \prec \sigma'$.

Informal description of Algorithm 3. The algorithm takes a $2^1\!/\!_2$-player Rabin game and an initial strategy σ^0, and proceeds in three steps. In Step 1, it constructs r pairs of σ^0-improving subgames $\widetilde{G}$ and corresponding improved strategy σ in $\widetilde{G}$. This is achieved by the procedure `ImprovingSubgames`. The parameter r will be chosen to obtain a suitable complexity analysis. In Step 2, the algorithm selects uniformly at random one of the improving subgames $\widetilde{G}$ with corresponding strategy σ, and recursively computes an optimal strategy σ^* in $\widetilde{G}$ from σ as the initial strategy. If the strategy σ^* is optimal in the original game G, then the algorithm terminates and returns σ^*. Otherwise it improves σ^* by a call to `Improve`, and continues at Step 1 with the improved strategy `Improve`(G, σ^*) as the initial strategy.

The procedure `ImprovingSubgames` constructs a sequence of game graphs $G^0, G^1, \ldots, G^{r-l}$ with $G^i \in \mathcal{G}(l, l+i)$ such that all $(l+i)$-subgames $\widetilde{G}^i_e$ of G^i are σ^0-improving. The subgame G^{i+1} is constructed from G^i as follows: we compute an optimal strategy σ^i in G^i, and if σ^i is optimal in G, then we have discovered an optimal strategy; otherwise we construct G^{i+1} by adding any *target* edge e of `Improve`(G, σ^i) in G^i, i.e., e is an edge required in the strategy `Improve`(G, σ^i) that is not in the strategy σ^i.

The correctness of the algorithm can be seen as follows. Observe that every time Step 1 is executed, the initial strategy is improved with respect to the ordering $\prec$ on strategies. Since the number of strategies is bounded, the termination of the algorithm is guaranteed. Step 3 of Algorithm 3 and Step 1.2.1 of procedure `ImprovingSubgames` ensure that on termination of the algorithm, the returned strategy is optimal. Lemma 10 bounds the expected number of iterations of Algorithm 3. The analysis is similar to the results of [1].

Lemma 10 *Algorithm 3 computes an optimal strategy. The expected number of iterations $T(\cdot,\cdot)$ of Algorithm 3 for a game $G \in \mathcal{G}(l,m)$ is bounded by the following recurrence: $T(l,m) \leq \sum_{i=l}^{r} T(l,i) + T(l-1, m-2) + \frac{1}{r} \cdot$*

Algorithm 3 RandomizedAlgorithm ($2^1\!/_2$-player Rabin games)

Input: a $2^1\!/_2$-player game graph $G \in \mathcal{G}(l,m)$, a Rabin objective Rabin(p) for pl. 1 and an initial strategy σ^0 for pl. 1.
Output : an optimal strategy σ^* for player 1.
1. (Step 1) Collect a set I of r pairs $(\widetilde{G},\sigma)$ of subgames $\widetilde{G}$ of G, and corresponding strategies σ in $\widetilde{G}$ such that $\sigma^0 \prec \sigma$.
 (This is achieved by the procedure `ImprovingSubgames` below).
2. (Step 2) Select a pair $(\widetilde{G},\sigma)$ from I uniformly at random.
 2.1 Find an optimal strategy in $\sigma^* \in \widetilde{G}$ by applying the algorithm recursively, with σ as the initial strategy.
3. (Step 3) **if** σ^* is an optimal strategy in the original game G **then return** σ^*.
 else let $\sigma =$ `Improve`(G,σ^*), and
 goto Step 1 with G and σ as the initial strategy.

procedure `ImprovingSubgames`
1. Construct sequence $G^0, G^1, \ldots, G^{r-l}$ of subgames with $G^i \in \mathcal{G}(l, l+i)$ as follows:
 1.1 G^0 is the game where each edge is fixed according to σ^0.
 1.2 Let σ^i be an optimal strategy in G^i;
  1.2.1 **if** σ^i is an optimal strategy in the original game G
   then return σ^i.
  1.2.2 **else** let e be any target of `Improve`(G,σ^i);
   the subgame G^{i+1} is G^i with the edge e added.
2. **return** r subgames (fixing one of the r edges in G^{r-l}) and associated strategies.

$\sum_{i=1}^{r} T(l, m-i) + 1$.

For a game graph G with $|S| = n$, we obtain a bound of n^2 for m. Using this fact and an analysis of Kalai for linear programming, Björklund et al. [1] showed that $m^{O\left(\sqrt{n/\log(n)}\right)} = 2^{O\left(\sqrt{n \cdot \log(n)}\right)}$ is a solution to the recurrence of Lemma 10, by choosing $r = \max\{\, n, \frac{m}{2} \,\}$.

Lemma 11 *Procedure* `Improve` *can be computed in time* $O(poly(n)) \cdot O($`TwoPlRabinGame`$(n \cdot d, d+1))$, *where poly represents a polynomial function.*

In Lemma 11 we denote by $O($`TwoPlRabinGame`$(n \cdot d, d+1))$ the time complexity of a 2-player Rabin game solving algorithm with $n \cdot d$ states and $d+1$ Rabin pairs. Recall the reduction $\mathrm{Tr}^1_{\mathrm{as}}$ blows up states in $S_{\bigcirc}$ by a factor of d and adds a new Rabin pair. A call to `Improve` requires solving an MDP with Streett objectives quantitatively (Step 1 of `Improve`; for a polynomial-time procedure, see [3, 8]) and computing Step 2.2 requires to solve at most n two-player Rabin games (since there can be at most n value-classes). Hence the lemma follows. Also recall that by the results of [13] we have $O(\texttt{TwoPlRabinGame}(n\cdot d, d+1)) = O\big((n\cdot d)^{d+1}\cdot(d+1)!\big) = O\big((n\cdot(d+1))^{d+1}\big)$. This analysis yields Theorem 4.

Theorem 4 *Given a $2^1\!/\!_2$-player game graph G and a Rabin objective $Rabin(P)$ with d-pairs the value $\langle\langle 1\rangle\rangle_{val}(Rabin(P))(s)$ can be computed for all states $s \in S$ in expected time $2^{O\left(\sqrt{n\cdot\log(n)}\right)} \cdot O\big(poly(n)\big) \cdot O(\texttt{TwoPlRabinGame}(n\cdot d, d+1)) = 2^{O\left(\sqrt{n\cdot\log(n)}\right)} \cdot O\big(poly(n)\big) \cdot O\big((n\cdot(d+1))^{d+1}\big)$. where poly represents a polynomial function.*

5 Optimal Strategy Construction for Streett Objectives

The algorithms, Algorithm 2 and Algorithm 3, computes values for both player 1 and player 2 (i.e., both for Rabin and Streett objectives), but only constructs an optimal strategy for player 1 (i.e., the player with Rabin objective). Since pure memoryless optimal strategies exist for player 1 (with Rabin objective), it is much simpler to analyze and obtain the values and optimal strategy for player 1. We now show that once the values are computed how to obtain an optimal strategy for player 2 (with Streett objective) by algorithms to obtain sure-winning strategies in 2-player games with Streett objectives.

Optimal strategy construction. Given a $2^1\!/\!_2$-player game G with Rabin objective $\text{Rabin}(P)$ for player 1 and the complementary objective $\text{Streett}(P)$ for player 2, first we compute $\langle\langle 1\rangle\rangle_{val}(\text{Rabin}(P))(s)$ for all states $s \in S$. An optimal strategy π^* for player 2 is constructed as follows: for a value-class $\text{VC}(r)$, obtain a sure-winning strategy $\overline{\pi}_r$ for player 2 in $\text{Tr}^2_{\text{as}}(\text{Tr}_{\text{win}_2}(G \restriction \text{VC}(r)))$, and let $\pi^* \restriction \text{VC}(r) = \text{Tr}^2_{\text{as}}(\overline{\pi}_r \restriction \text{VC}(r))$. By Lemma 3 it follows that π^* is an optimal strategy, and given the values the construction of π^* requires n-calls to a procedure to solve 2-player games with Streett objectives.

Theorem 5 *Let G be a $2^1\!/\!_2$-player game with Rabin objective $Rabin(P)$ for player 1 and Streett objective $Streett(P)$ for player 2, where P has d-pairs. Given the values $\langle\langle 1\rangle\rangle_{val}(Rabin(P))(s) = 1 - \langle\langle 2\rangle\rangle_{val}(Streett(P))(s)$, for all* *states $s \in S$, an optimal strategy π^* for player 2 can be constructed in time $n \cdot O(\texttt{TwoPlStreettGame}(n \cdot d, d+1))$, where* `TwoPlStreettGame` *is an algorithm to produce sure-winning strategies in 2-player games with Streett objectives.*

References

[1] H. Bjorklund, S. Sandberg, and S. Vorobyov. A discrete subexponential algorithms for parity games. In *STACS'03*, pages 663–674. LNCS 2607, Springer, 2003.

[2] K. Chatterjee, L. de Alfaro, and T.A. Henzinger. Trading memory for randomness. In *QEST'04*, pages 206–217. IEEE Computer Society Press, 2004.

[3] K. Chatterjee, L. de Alfaro, and T.A. Henzinger. The complexity of stochastic Rabin and Streett games. In *ICALP'05*, pages 878–890. LNCS 3580, Springer, 2005.

[4] K. Chatterjee and T.A. Henzinger. Strategy improvement and randomized subexponential algorithms for stochastic parity games. In *STACS'06*, LNCS 3884, pages 512–523. Springer, 2006.

[5] K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Simple stochastic parity games. In *CSL'03*, LNCS 2803, pages 100–113. Springer, 2003.

[6] A. Condon. The complexity of stochastic games. *Information and Computation*, 96:203–224, 1992.

[7] A. Condon. On algorithms for simple stochastic games. In Jin-Yi Cai, editor, *Advances in Computational Complexity Theory*, volume 13 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 51–73. American Mathematical Society, 1993.

[8] L. de Alfaro. *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University, 1997.

[9] L. de Alfaro and T. A. Henzinger. Concurrent ω-regular games. In *LICS'00*, pages 141–154. IEEE Computer Society Press, 2000.

[10] E.A. Emerson and C. Jutla. The complexity of tree automata and logics of programs. In *FOCS'88*, pages 328–337. IEEE Computer Society Press, 1988.

[11] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, 1997.

[12] A. Hoffman and R. Karp. On nonterminating stochastic games. *Management Science*, 12:359–370, 1966.

[13] F. Horn. Street games on finite graphs. In *GDV'05*, 2005.

[14] A. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.

[15] O. Kupferman and M.Y. Vardi. Weak alternating automata and tree automata emptiness. In *STOC'98*, pages 224–233. ACM Press, 1998.

[16] A. Pnueli and R. Rosner. On the synthesis of a reactive module. In *POPL'89*, pages 179–190. ACM Press, 1989.

[17] 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.

[18] W. Thomas. Languages, automata, and logic. In *Handbook of Formal Languages*, volume 3, Beyond Words, chapter 7, pages 389–455. Springer, 1997.