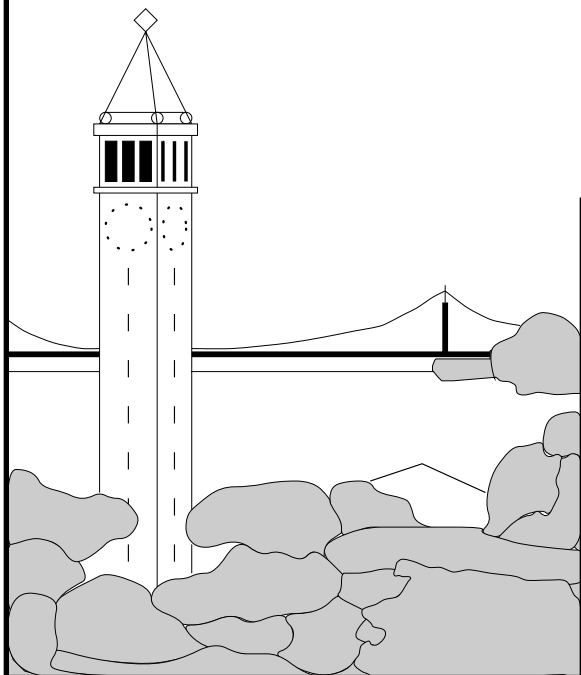


Two-player Nonzero-sum ω -regular Games

Krishnendu Chatterjee



Report No. UCB/CSD-4-1364

December 2004

Computer Science Division (EECS)
University of California
Berkeley, California 94720

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE DEC 2004		2. REPORT TYPE		3. DATES COVERED 00-00-2004 to 00-00-2004	
4. TITLE AND SUBTITLE Two-player Nonzero-sum w-regular Games				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of California at Berkeley, Department of Electrical Engineering and Computer Sciences, Berkeley, CA, 94720				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT We study in finite stochastic games played by two-players on a finite graph with goals specified by sets of finite traces. The games are concurrent (each player simultaneously and independently chooses an action at each round), stochastic (the next state is determined by a probability distribution depending on the current state and the chosen actions), in finite (the game continues for an finite number of rounds) nonzero-sum (the players' goals are not necessarily conflicting), and undiscounted. We show that if each player has an ω-regular objective expressed as a parity objective, then there exists an ω-Nash equilibrium for every $\epsilon > 0$. However, exact Nash equilibria need not exist. We study the complexity of finding values (payoffs) of some ω-Nash equilibrium. We show that the values of some ω-Nash equilibrium in nonzero-sum concurrent parity games can be computed by solving the following two simpler problems: computing the values of zero-sum (the goals of the players are strictly conflicting) concurrent parity games and computing ω-Nash equilibrium values of nonzero-sum concurrent games with reachability objectives. As a consequence we establish that values of some ω-Nash equilibrium can be approximated in FNP (functional NP), and hence in EXPTIME.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 49	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Two-player Nonzero-sum ω -regular Games *

Krishnendu Chatterjee

Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, USA
`c_krish@cs.berkeley.edu`

December 2004

Abstract

We study infinite stochastic games played by two-players on a finite graph with goals specified by sets of infinite traces. The games are *concurrent* (each player simultaneously and independently chooses an action at each round), *stochastic* (the next state is determined by a probability distribution depending on the current state and the chosen actions), *infinite* (the game continues for an infinite number of rounds), *nonzero-sum* (the players' goals are not necessarily conflicting), and undiscounted. We show that if each player has an ω -regular objective expressed as a parity objective, then there exists an ε -Nash equilibrium, for every $\varepsilon > 0$. However, exact Nash equilibria need not exist. We study the complexity of finding values (payoff profile) of some ε -Nash equilibrium. We show that the values of some ε -Nash equilibrium in nonzero-sum concurrent parity games can be computed by solving the following two simpler problems: computing the values of zero-sum (the goals of the players are strictly conflicting) concurrent parity games and computing ε -Nash equilibrium values of nonzero-sum concurrent games with reachability objectives. As a consequence we establish that values of some ε -Nash equilibrium can be approximated in FNP (functional NP), and hence in EXPTIME.

*The work was supported by the AFOSR MURI grant F49620-00-1-0327, by the ONR grant N00014-02-1-0671, and by the NSF grants CCR-0132780, CCR-0234690, CCR-9988172, and CCR-0225610, and by the NSF Career grant CCR-0132780, the NSF grant CCR-0234690, and by the ONR grant N00014-02-1-0671

1 Introduction

Stochastic games. Non-cooperative games provide a natural framework to model interactions between agents [17, 19]. The simplest class of non-cooperative games consists of the “one-step” games — games with single interaction between the agents after which the game ends and the payoffs are decided (e.g., matrix games). However, a wide class of games progress over time and in stateful manner, and the current game depends on the history of interactions. Infinite *stochastic games* [21, 8] are a natural model for such games. A stochastic game is played over a finite *state space* and is played in rounds. In concurrent games, in each round, each player chooses an action from a finite set of available actions, simultaneously and independently of other players. The game proceeds to a new state according to a probabilistic transition relation (stochastic transition matrix) based on the current state and the joint actions of the players. Concurrent games subsumes the simpler class of *turn-based games*, where at every state at most one player can choose between multiple actions. In verification and control of finite state reactive systems such games proceed for infinite rounds, generating an infinite sequence of states, called the *outcome* of the game. The players receive a payoff based on a payoff function that maps every outcome to a real number.

Objectives. Payoffs are generally Borel measurable functions [15]. For example, the payoff set for each player is a Borel set B_i in the Cantor topology on S^ω (where S is the set of states), and player i gets payoff 1 if the outcome of the game is a member of B_i , and 0 otherwise. In verification, payoff functions are usually index sets of ω -regular languages. The ω -regular languages generalizes the classical regular languages to infinite strings, they occur in low levels of the Borel hierarchy (they are in $\Sigma_3 \cap \Pi_3$), and form a robust and expressive language for determining payoffs for commonly used specifications [14]. The simplest ω -regular objectives correspond to safety (“closed sets”) and reachability (“open sets”) objectives.

Zero-sum games. Games may be *zero-sum*, where two players have directly conflicting objectives and the payoff of one player is one minus the payoff of the other, or *nonzero-sum*, where each player has a prescribed payoff function based on the outcome of the game. The fundamental question for games is the existence of equilibrium values. For zero-sum games, this involves showing a *determinacy* theorem that states that the expected optimum value obtained by player 1 is exactly one minus the expected optimum value obtained by player 2. For one-step zero-sum games, this is von

Neumann’s minmax theorem [31]. For infinite games, the existence of such equilibria is not obvious, in fact, by using the axiom of choice, one can construct games for which determinacy does not hold. However, a remarkable result by Martin [15] shows that all stochastic zero-sum games with Borel payoffs are determined.

Nonzero-sum games. For nonzero-sum games, the fundamental equilibrium concept is a *Nash equilibrium* [11], that is, a strategy profile such that no player can gain by deviating from the profile, assuming the other player continue playing the strategy in the profile. Again, for one-step games, the existence of such equilibria is guaranteed by Nash’s theorem [11]. However, the existence of Nash equilibria in infinite games is not immediate: Nash’s theorem holds for finite bimatrix games, but in case of stochastic games, the strategy space is not compact. The existence of Nash equilibria is known only in very special cases of stochastic games. In fact, Nash equilibria may not exist, and the best one can hope for is an ε -Nash equilibrium for all $\varepsilon > 0$, where an ε -Nash equilibrium is a strategy profile where unilateral deviation can only increase the payoff of a player by at most ε . Exact Nash equilibria do exist in discounted stochastic games [9], and other special cases [26, 27]. For concurrent nonzero-sum games with payoffs defined by Borel sets, surprisingly little is known. Secchi and Sudderth [20] showed that exact Nash equilibria do exist when all players have payoffs defined by closed sets (“safety objectives”), where the objective of each player is to stay within a certain set of good states. Formally, each player i has a subset of states F_i as their safe states, and gets a payoff 1 if the play never leaves the set F_i and gets payoff 0 otherwise. This result was generalized to general state and action spaces [20, 13], where only ε -equilibria exist. In the case of open sets (“reachability objectives”), each player i has a subset of states R_i as reachability targets. Player i gets payoff 1 if the outcome visits some state from R_i at some point, and 0 otherwise. The existence of ε -Nash equilibrium in games with payoffs described as open sets, for every $\varepsilon > 0$, has been established in [5]. The above results hold even in the case of n -player games. In one of the most important recent result in stochastic game theory, Vieille shows the existence of ε -Nash equilibrium, for every $\varepsilon > 0$, in two-player concurrent games with limit-average payoff [29, 30]. The existence of ε -Nash equilibrium in two-player concurrent games with objectives in the higher levels of Borel hierarchy has been an intriguing open problem.

Our result and proof techniques. In this paper we show that ε -Nash equilibrium exists, for every $\varepsilon > 0$, for two-player concurrent games with ω -regular objectives. However, exact Nash equilibria need not exist. For

two-player concurrent games our result extends the existence of ε -Nash equilibrium from the lowest level of Borel hierarchy (open and closed sets) to the classical ω -regular objectives that lie in the higher levels of Borel hierarchy; and our result for ω -regular objectives parallels Vieille’s result for limit-average objectives. Our proof technique involves the following key ideas:

1. We first show the existence of ε -Nash equilibrium, for every $\varepsilon > 0$, with ω -regular objectives, for a sub-class of concurrent games, namely single strongly connected component (SSCC) games in Section 3.
2. We extend the above result to all concurrent games in Section 4.

The result for SSCC games involves the following key ideas:

- We identify four sufficient conditions that ensure existence of ε -Nash equilibrium, for every $\varepsilon > 0$, in SSCC games.
- We then show that if the sufficient conditions are not satisfied then the game can be reduced to a nonzero-sum game with reachability objectives, with some desired properties. The result is proved by generalizing a result from [2] and using a fragment of analysis of Vieille [29].
- The existence of ε -Nash equilibrium, for all $\varepsilon > 0$, in the original game is then established by the use of *punishing or spoiling strategies*.

Complexity of ε -Nash equilibrium. Computing the values of a Nash equilibria, when it exists, is another challenging problem [18, 32]. For one-step zero-sum games, equilibrium values and strategies can be computed in polynomial time (by reduction to linear programming) [17]. For one-step nonzero-sum games, no polynomial time algorithm to compute an exact Nash equilibrium in a two-player game is known [18]. In case of concurrent games with limit-average payoff no algorithmic analysis is known even for zero-sum games. However, several algorithms are known for several special cases, e.g., for turn-based games [33, 1, 10]. In case of zero-sum concurrent games with ω -regular objectives several algorithms are known to compute values with in ε -approximation [7, 2]. Since the values can be irrational ε -approximation is the best one can achieve. From the computational aspects, a desirable property of an existence proof of Nash equilibrium is its ease of algorithmic analysis. We show that our proof for existence of ε -Nash equilibrium is completely constructive and algorithmic. Our proof shows that the computation of values of some ε -Nash equilibrium in two-player

concurrent games with parity objectives can be reduced to the following two simpler problems:

1. Computing values of zero-sum concurrent games with parity objectives.
2. Computing values of some special ε -Nash equilibrium of nonzero-sum concurrent games with reachability objectives.

Since solving the more general case of nonzero-sum games must involve solving the special case of zero-sum games, our result reduces the problem of computing ε -Nash equilibrium for ω -regular objectives to solving some special ε -Nash equilibrium of games with reachability objectives. We then prove that the equilibrium values of some ε -Nash equilibrium can be approximated in FNP (functional NP) and hence in EXPTIME. Our result matches the best known complexity bound for the simpler case of turn-based games [5].

Organization. The paper is organized as follows. In section 2 we define the basic notions of games, strategies and objectives. In section 3 we prove existence of ε -Nash equilibrium for a sub-class of concurrent games, and then extend the result for all concurrent games in section 4. We present the complexity result for computing ε -Nash equilibrium values in section 5. We conclude with a few open problems in section 6.

2 Preliminaries

Notation. For a countable set A , a *probability distribution* on A is a function $\delta: A \mapsto [0, 1]$ such that $\sum_{a \in A} \delta(a) = 1$. We denote the set of probability distributions on A by $\mathcal{D}(A)$. Given a distribution $\delta \in \mathcal{D}(A)$, we denote by $\text{Supp}(\delta) = \{x \in A \mid \delta(x) > 0\}$ the *support* of δ .

Definition 1 (Concurrent Games) A (two-player) concurrent game structure $\mathcal{G} = \langle S, \text{Moves}, \Gamma_1, \Gamma_2, \delta \rangle$ consists of the following components:

- A finite state space S .
- A finite set Moves of moves.
- Two move assignments $\Gamma_1, \Gamma_2: S \mapsto 2^{\text{Moves}} \setminus \emptyset$. For $i \in \{1, 2\}$, assignment Γ_i associates with each state $s \in S$ the non-empty set $\Gamma_i(s) \subseteq \text{Moves}$ of moves available to player i at state s .

- A *probabilistic transition function* $\delta : S \times \text{Moves} \times \text{Moves} \rightarrow \mathcal{D}(S)$, that gives the probability $\delta(s, a_1, a_2)(t)$ of a transition from s to t when player 1 plays move a_1 and player 2 plays move a_2 , for all $s, t \in S$ and $a_1 \in \Gamma_1(s)$, $a_2 \in \Gamma_2(s)$. ■

We distinguish the following special classes of concurrent game structures.

- A concurrent game structure $\mathcal{G}$ is *deterministic* if for all $s \in S$ and all $a_1 \in \Gamma_1(s)$, $a_2 \in \Gamma_2(s)$, there is a $t \in S$ such that $\delta(s, a_1, a_2)(t) = 1$.
- A concurrent game structure $\mathcal{G}$ is *turn-based* if at every state at most one player can choose among multiple moves; that is, if for every state $s \in S$ there exists at most one $i \in \{1, 2\}$ with $|\Gamma_i(s)| > 1$.
- A concurrent game structure is a Markov decision process (MDP) if there exists at most one $i \in \{1, 2\}$ such that at every state s , $|\Gamma_i(s)| > 1$. In other words, a MDP is a one-player stochastic game and only one player has a non-trivial choice of moves and for the other player the choice of the moves are fixed.

We define the size of the game structure $\mathcal{G}$ to be equal to the size of the transition function δ ; specifically, $|\mathcal{G}| = \sum_{s \in S} \sum_{a \in \Gamma_1(s)} \sum_{b \in \Gamma_2(s)} \sum_{t \in S} |\delta(s, a, b)(t)|$, where $|\delta(s, a, b)(t)|$ denotes the space to specify the probability distribution. We write n to denote the size of the state space, i.e., $n = |S|$. At every state $s \in S$, player 1 chooses a move $a_1 \in \Gamma_1(s)$, and simultaneously and independently player 2 chooses a move $a_2 \in \Gamma_2(s)$. The game then proceeds to the successor state t with probability $\delta(s, a_1, a_2)(t)$, for all $t \in S$. A state s is called an *absorbing state* if for all $a_1 \in \Gamma_1(s)$ and $a_2 \in \Gamma_2(s)$ we have $\delta(s, a_1, a_2)(s) = 1$. In other words, at s for all choice of moves of the players the next state is always s . A state s is a *turn-based state* if there exists $i \in \{1, 2\}$ such that $|\Gamma_i(s)| = 1$. Moreover, if $|\Gamma_2(s)| = 1$ then the state s is a player-1 turn-based state since the choice of moves for player 2 is trivial; and if $|\Gamma_1(s)| = 1$ then it is a player-2 turn-based state. We assume that the players act *non-cooperatively*, i.e., each player chooses her strategy independently and secretly from the other player, and is only interested in maximizing her own payoff. For all states $s \in S$ and moves $a_1 \in \Gamma_1(s)$ and $a_2 \in \Gamma_2(s)$, we indicate by $\text{Dest}(s, a_1, a_2) = \text{Supp}(\delta(s, a_1, a_2))$ the set of possible successors of s when moves a_1, a_2 are selected.

A *path* or a *play* ω of $\mathcal{G}$ is an infinite sequence $\omega = \langle s_0, s_1, s_2, \dots \rangle$ of states in S such that for all $k \geq 0$, there are moves $a_1^k \in \Gamma_1(s_k)$ and $a_2^k \in \Gamma_2(s_k)$

with $\delta(s_k, a_1^k, a_2^k)(s_{k+1}) > 0$. We denote by Ω the set of all paths and by Ω_s the set of all paths $\omega = \langle s_0, s_1, s_2, \dots \rangle$ such that $s_0 = s$, i.e., the set of plays starting from state s .

2.1 Randomized strategies

A *selector* ξ for player $i \in \{1, 2\}$ is a function $\xi : S \mapsto \mathcal{D}(\text{Moves})$ such that for all $s \in S$ and $a \in \text{Moves}$, if $\xi(s)(a) > 0$ then $a \in \Gamma_i(s)$. We denote by Λ_i the set of all selectors for player $i \in \{1, 2\}$. A selector ξ is *pure* if for every $s \in S$ there is $a \in \text{Moves}$ such that $\xi(s)(a) = 1$; we denote by $\Lambda_i^P \subseteq \Lambda_i$ the set of pure selectors for player i . A *strategy* for player 1 is a function $\sigma : S^+ \rightarrow \Lambda_1$ that associates with every finite non-empty sequence of states, representing the history of the play so far, a selector. Similarly we define strategies π for player 2. A strategy σ for player i is *pure* if it yields only pure selectors, that is, is of type $S^+ \rightarrow \Lambda_i^P$. A strategy with memory can be described as a pair of functions: (a) memory update function $\sigma_u : S \times \mathbb{M} \times \text{Moves} \rightarrow \mathbb{M}$, and (b) next move function $\sigma_m : S \times \mathbb{M} \rightarrow \Lambda_1$. A strategy with memory is finite memory if $\mathbb{M}$ is finite. A memoryless strategy is independent of the history of the play and depends only on the current state. Memoryless strategies coincide with selectors, and we often write σ for the selector corresponding to a memoryless strategy σ . A strategy is pure memoryless if it is pure and memoryless. We denote by $\Sigma^P, \Sigma^F, \Sigma^{PM}$ the family of pure, finite-memory and pure memoryless strategies for player 1 respectively. Analogously we define the families of strategies for player 2. We denote by Σ and Π the set of all strategies for player 1 and player 2, respectively.

Once the starting state s and the strategies σ and π for the two players have been chosen, the game is reduced to an ordinary stochastic process. Hence, the probabilities of events are uniquely defined, where an *event* $\mathcal{A} \subseteq \Omega_s$ is a measurable set of paths. For an event $\mathcal{A} \subseteq \Omega_s$, we denote by $\text{Pr}_s^{\sigma, \pi}(\mathcal{A})$ the probability that a path belongs to $\mathcal{A}$ when the game starts from s and the players follow the strategies σ and π . For $i \geq 0$, we also denote by $\Theta_i : \Omega_s \rightarrow S$ the random variable denoting the i -th state along a path.

2.2 Objectives.

An *objective* for a player in a game $\mathcal{G}$ is a set $\mathcal{W} \subseteq \Omega$ of infinite paths. We consider the following objectives.

- *Reachability objective.* For a set $R \subseteq S$ of *target* states, the Reachability objective is defined as $\text{Reach}(R) = \{ \langle s_0, s_1, s_2, \dots \rangle \in \Omega \mid \exists k \in \mathbb{N}. s_k \in R \}$.

- *Safety objective.* For a set $F \subseteq S$ of *safe* states, the Safety objective is defined as $\text{Safe}(F) = \{ \langle s_0, s_1, s_2, \dots \rangle \in \Omega \mid \forall k \in \mathbb{N}. s_k \in F \}$. Note that $\Omega \setminus \text{Reach}(R) = \text{Safe}(S \setminus R)$. Hence the reachability objective with target set R is complementary to the safety objective with safe set $S \setminus R$.
- *Parity objective.* Given $d \in \mathbb{N}$, we write $[d]$ for the set $\{0, 1, 2, \dots, d\}$ and $[d]_+$ for the set $\{1, 2, \dots, d\}$. Let $p : S \mapsto [d]$ be a function that assigns a *priority* $p(s)$ to every state $s \in S$, where $d \in \mathbb{N}$. For an infinite path $\omega = \langle s_0, s_1, s_2, \dots \rangle \in \Omega$, we define $\text{Inf}(\omega) = \{ i \in [d] \mid p(s_k) = i \text{ for infinitely many } k \geq 0 \}$. The *parity objective* is defined as $\text{Parity}(p) = \{ \omega \in \Omega \mid \min(\text{Inf}(\omega)) \text{ is even} \}$. Informally we say that a path ω satisfy the parity objective, $\text{Parity}(p)$, if $\omega \in \text{Parity}(p)$.

The ability to solve games with Rabin-chain (parity) objectives suffices for solving games with arbitrary LTL (or ω -regular) objectives: in fact, it suffices to encode the ω -regular objective as a deterministic Rabin-chain automaton, solving then the game consisting of the synchronous product of the original game with the Rabin-chain automaton [16, 24].

A concurrent nonzero-sum parity game consists of a game structure $\mathcal{G}$ and two priority function p_1 and p_2 for player 1 and player 2, respectively. The objective of player 1 and player 2 are $\text{Parity}(p_1)$ and $\text{Parity}(p_2)$, respectively. In general we write Ψ for a arbitrary parity objective. We write the objective of player 1 and player 2 as Ψ_1 and Ψ_2 , respectively, where Ψ_1 and Ψ_2 are arbitrary ω -regular objective formalized as a parity objective. We also use Ψ_1 to denote the set of paths $\omega \in \Omega$ such that $\omega \in \text{Parity}(p_1)$. Similarly we write Ψ_2 to denote the set of paths $\text{Parity}(p_2)$. Given a parity objective Ψ , the set of paths Ψ is measurable for any choice of strategies for the two players [28]. Given a state s we write Ψ_{1s} to denote $\Omega_s \cap \Psi_1$ and similarly we write Ψ_{2s} to denote $\Omega_s \cap \Psi_2$. We also write Ψ_s to denote $\Omega_s \cap \Psi$. Hence, the probability that a path satisfies objective Ψ starting from state $s \in S$ under strategies σ, π for the two players is $\text{Pr}_s^{\sigma, \pi}(\Psi_s)$.

Concurrent zero-sum games. A concurrent game is zero-sum if the objectives of the players are complementary, i.e., $\Psi_1 = \Omega \setminus \Psi_2$. The zero-sum values for the players in a zero-sum concurrent game is defined as follows.

Definition 2 (Zero-sum values) *Given a state $s \in S$ we call the maximal probability with which player 1 can ensure that Ψ_1 holds from s against any strategy of player 2 is the zero-sum value of player 1 at s . The zero-sum value for player 2 is defined symmetrically. Formally, the zero-sum value*

for player 1 and player 2 are given by functions $\langle\langle 1 \rangle\rangle_{val}(\Psi_1) : S \mapsto [0, 1]$ and $\langle\langle 2 \rangle\rangle_{val}(\Psi_2) : S \mapsto [0, 1]$, defined for all $s \in S$ by

$$\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s) = \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \Pr_s^{\sigma, \pi}(\Psi_{1s})$$

$$\langle\langle 2 \rangle\rangle_{val}(\Psi_2)(s) = \sup_{\pi \in \Pi} \inf_{\sigma \in \Sigma} \Pr_s^{\sigma, \pi}(\Psi_{2s}).$$

■

Concurrent zero-sum games satisfy a *quantitative* version of determinacy [15], stating that for all parity objective, Ψ_1 and Ψ_2 such that $\Psi_1 = \Omega \setminus \Psi_2$, and all $s \in S$, we have

$$\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s) + \langle\langle 2 \rangle\rangle_{val}(\Psi_2)(s) = 1.$$

A strategy σ for player 1 is *optimal* if for all $s \in S$ we have

$$\inf_{\pi \in \Pi} \Pr_s^{\sigma, \pi}(\Psi_{1s}) = \langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s).$$

For $\varepsilon > 0$, a strategy σ for player 1 is ε -*optimal* if for all $s \in S$ we have

$$\inf_{\pi \in \Pi} \Pr_s^{\sigma, \pi}(\Psi_{1s}) \geq \langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s) - \varepsilon.$$

We define optimal and ε -optimal strategies for player 2 symmetrically. Note that the quantitative determinacy of concurrent zero-sum games is equivalent to the existence of ε -optimal strategies for both players for all $\varepsilon > 0$, at all states $s \in S$.

Definition 3 (Cooperative value) *Given an objective Ψ we define the cooperative value of the game as the maximal probability with which player 1 and player 2 can cooperate to satisfy the objective Ψ . Formally, the cooperative value is given by the function $\langle\langle 1, 2 \rangle\rangle_{val}(\Psi) : S \mapsto [0, 1]$ defined for all $s \in S$ by*

$$\langle\langle 1, 2 \rangle\rangle_{val}(\Psi)(s) = \sup_{(\sigma, \pi) \in \Sigma \times \Pi} \Pr_s^{\sigma, \pi}(\Psi_s).$$

■

Note that the computation of the cooperative value function $\langle\langle 1, 2 \rangle\rangle_{val}(\Psi)$ can be interpreted as the computation of a value function in a MDP with objective Ψ , where player 1 and player 2 cooperatively choose strategies.

Definition 4 (ε -Nash equilibrium) *Let $\mathcal{G}$ be a game and let the objectives for player 1 and player 2 be Ψ_1 and Ψ_2 , respectively. For $\varepsilon \geq 0$, a strategy profile $(\sigma^*, \pi^*) \in \Sigma \times \Pi$ is an ε -Nash equilibrium for a state $s \in S$ iff the following two conditions hold:*

$$\forall \sigma \in \Sigma. \Pr_s^{\sigma, \pi^*}(\Psi_{1s}) \leq \Pr_s^{\sigma^*, \pi^*}(\Psi_{1s}) + \varepsilon$$

$$\forall \pi \in \Pi. \Pr_s^{\sigma^*, \pi}(\Psi_{2s}) \leq \Pr_s^{\sigma^*, \pi^*}(\Psi_{2s}) + \varepsilon.$$

An exact Nash equilibrium is an 0-Nash equilibrium. ■

It may be noted that in case of zero-sum concurrent games with parity objectives optimal strategies need not exist, and only existence of ε -optimal strategies can be guaranteed, for all $\varepsilon > 0$. Hence in the general case of nonzero-sum concurrent games with parity objectives Nash equilibrium need not exist, and existence of ε -Nash equilibrium, for all $\varepsilon > 0$, is the best one can achieve.

2.3 The branching structure of plays

Many of the arguments developed in this paper rely on a detailed analysis of the branching process resulting from the strategies chosen by the players, and from the probabilistic transition relation of the game. In order to make our arguments precise, we need some definitions. A play is *feasible* if each of its transitions could have arisen according to the transition relation of the game.

Definition 5 (Feasible plays and outcomes) *Given strategies σ for player 1 and π for player 2, a play $\omega = \langle s_0, s_1, s_2, \dots \rangle$ is feasible in a concurrent game graph $\mathcal{G}$, if for every $k \in \mathbb{N}$ the following conditions hold:*

1. $s_{k+1} \in \text{Dest}(s_k, a_1, a_2)$;
2. $\sigma(s_0, s_1, \dots, s_k)(a_1) > 0$; and
3. $\pi(s_0, s_1, \dots, s_k)(a_2) > 0$.

Given strategies $\sigma \in \Sigma$ and $\pi \in \Pi$, and a state s , we denote by $\text{Outcome}(s, \sigma, \pi) \subseteq \Omega_s$ the set of feasible plays that start from s , given strategies σ and π . ■

In order to make precise statements about the branching process arising from a game play, we define below trees labeled by game states.

Definition 6 (Infinite trees, S -labeled trees and trees for events)

An infinite tree is a set $\text{Tr} \subseteq \mathbb{N}^*$ such that

- if $x \cdot i \in \text{Tr}$ where $x \in \mathbb{N}^*$ and $i \in \mathbb{N}$ then $x \in \text{Tr}$;
- for all $x \in \text{Tr}$ there exists $i \in \mathbb{N}$ such that $x \cdot i \in \text{Tr}$. We refer to $x \cdot i$ as a successor of x .

We call the elements in Tr as nodes and the empty word ϵ is the root of the tree. An infinite path τ of Tr is a set $\tau \subseteq \text{Tr}$ such that

- $\epsilon \in \tau$;
- for every x in τ there is a unique $i \in \mathbb{N}$ such that $x \cdot i \in \tau$. Note that for every $i \in \mathbb{N}$, there is a unique element $x \in \tau$ such that $|x| = i$. We denote by τ_i the element $x \in \tau$ such that $|x| = i$.

Given an infinite tree Tr and a node $x \in \text{Tr}$, we denote by $\text{Tr}(x)$ the sub-tree rooted at node x . Formally, $\text{Tr}(x)$ denotes the set $\{x' \in \text{Tr} \mid x \text{ is a prefix of } x'\}$.

A S -labeled tree $\mathcal{T}$ is a pair $(\text{Tr}, \langle \cdot \rangle)$, where Tr is a tree and $\langle \cdot \rangle : \text{Tr} \rightarrow S$ maps each node of Tr to a state $s \in S$. Given a S -labelled tree $\mathcal{T}$, and a infinite path $\tau \subseteq \text{Tr}$, we denote by $\langle \tau \rangle$ the play $\langle s_0, s_1, s_2, \dots \rangle$, such that $s_0 = \langle \epsilon \rangle$ and for all $i > 0$ we have $s_i = \langle \tau_i \rangle$. A S -labeled tree $\mathcal{T}_s = (\text{Tr}_s, \langle \cdot \rangle)$, where $\langle \epsilon \rangle = s$, represents a set of infinite paths, denoted as $\mathcal{L}(\mathcal{T}_s) \subseteq \Omega_s$, such that

$$\mathcal{L}(\mathcal{T}_s) = \{ \omega = \langle s_0 = s, s_1, s_2, \dots \rangle \in \Omega_s \mid \exists \tau \subseteq \text{Tr}_s. \langle \tau \rangle = \omega \}.$$

A S -labeled tree $\mathcal{T}_s$ represents an event $\mathcal{A} \subseteq \Omega_s$ if and only if $\mathcal{L}(\mathcal{T}_s) = \mathcal{A}$. We denote by $\mathcal{T}_{\mathcal{A},s}$ a S -labeled tree that represents an event $\mathcal{A} \subseteq \Omega_s$, and denote by $\text{Tr}_{\mathcal{A},s}$ the tree of $\mathcal{T}_{\mathcal{A},s}$. ■

Several of the following results will be phrased in terms of the S -labeled tree $\mathcal{T}_{\mathcal{A},s}^{\sigma,\pi}$, which represents the outcomes from $s \in S$ that result from player 1 using strategy σ and player 2 using strategy π , and that belong to a specified event $\mathcal{A}$.

Definition 7 (Trees for strategies) Given a measurable event $\mathcal{A}$, strategies σ, π , a state s , such that $\Pr_s^{\sigma,\pi}(\mathcal{A}) > 0$, we denote by $\mathcal{T}_{\mathcal{A},s}^{\sigma,\pi}$ a S -labeled tree to represent $\mathcal{A} \cap \text{Outcome}(s, \sigma, \pi)$, and we also denote by $\text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$ the tree of $\mathcal{T}_{\mathcal{A},s}^{\sigma,\pi}$. Given strategy σ, π , we denote by $\mathcal{T}_s^{\sigma,\pi}$ the S -labeled tree $\mathcal{T}_{\text{Outcome}(s, \sigma, \pi), s}^{\sigma,\pi}$, and we also denote by $\text{Tr}_s^{\sigma,\pi}$ the tree of $\mathcal{T}_s^{\sigma,\pi}$. ■

Notations. Let $\mathcal{T} = (\text{Tr}, \langle \cdot \rangle)$ be a S -labeled tree and $x \in \text{Tr}$ such that $|x| = n$. We denote by x_i the prefix of x of length i . We denote by $\text{hist}(x) = (\langle \epsilon \rangle, \langle x_1 \rangle, \dots, \langle x_n \rangle)$, the history represented by the path from root to the node x . We denote by $\text{Cone}(x) = \{ \omega = (s_0, s_1, s_2, \dots) \mid \langle x_i \rangle = s_i \text{ for all } 0 \leq i \leq n \}$ the set of paths with the prefix $\text{hist}(x)$. Given a measurable event $\mathcal{A} \subseteq \Omega_s$, strategies σ and π such that $\Pr_s^{\sigma, \pi}(\mathcal{A}) > 0$, consider the S -labeled tree $\mathcal{T}_{\mathcal{A}, s}^{\sigma, \pi}$ to represent $\mathcal{A} \cap \text{Outcome}(s, \sigma, \pi)$. Consider the event $\mathcal{A}_{\text{nil}} = \{ \text{Cone}(x) \mid x \in \text{Tr}_{\mathcal{A}, s}^{\sigma, \pi}, \Pr_s^{\sigma, \pi}(\text{Cone}(x) \cap \mathcal{A}) = 0 \}$. Since $\mathcal{A}_{\text{nil}}$ is the countable union of measurable sets each with measure 0 we have $\Pr_s^{\sigma, \pi}(\mathcal{A}_{\text{nil}} \cap \mathcal{A}) = 0$. Hence, in sequel without loss of generality given any event $\mathcal{A}$, we only consider the event $\mathcal{A} \setminus \mathcal{A}_{\text{nil}}$ and by a little abuse of notation use $\mathcal{T}_{\mathcal{A}, s}^{\sigma, \pi}$ to represent the stochastic tree $\mathcal{T}_{(\mathcal{A} \setminus \mathcal{A}_{\text{nil}}), s}^{\sigma, \pi}$. Hence, without loss of generality we assume for any $x \in \text{Tr}_{\mathcal{A}, s}^{\sigma, \pi}$ we have $\Pr_s^{\sigma, \pi}(\text{Cone}(x) \cap \mathcal{A}) > 0$. Henceforth, for any $x \in \text{Tr}_{\mathcal{A}, s}^{\sigma, \pi}$ we write $\Pr_x^{\sigma, \pi}(\mathcal{B} \mid \mathcal{A})$ to denote $\Pr_s^{\sigma, \pi}(\mathcal{B} \mid \text{Cone}(x), \mathcal{A})$.

Definition 8 (Perennial ε -optimal and perennial ε -spoiling strategies)

For a parity objective Ψ , for $\varepsilon > 0$, a strategy σ is a perennial ε -optimal strategy for player 1, from state s , with respect to objective Ψ if for all strategy π , for all node x in the stochastic tree $\text{Tr}_s^{\sigma, \pi}$, we have $\Pr_x^{\sigma, \pi}(\Omega_{es}) \geq \langle 1 \rangle_{\text{val}}(\Psi)(\langle x \rangle) - \varepsilon$, i.e., in the stochastic sub-tree rooted at x player 1 is ensured the zero-sum value of the game at $\langle x \rangle$ within ε -precision. Perennial ε -optimal strategies for player 2 are defined analogously. Given a nonzero-sum concurrent game with objective Ψ_1 for player 1 and Ψ_2 for player 2, a strategy σ_ε is perennial ε -optimal if it is perennial ε -optimal with respect to objective Ψ_1 , and a strategy $\bar{\sigma}_\varepsilon$ is perennial ε -spoiling if it is perennial ε -optimal with respect to objective $\bar{\Psi}_2 = \Omega \setminus \Psi_2$. Perennial ε -optimal and perennial ε -spoiling strategies for player 2 are defined similarly. We denote by Σ_ε and Π_ε the set of perennial ε -optimal strategies for player 1 and player 2, respectively. Similarly, we denote by $\bar{\Sigma}_\varepsilon$ and $\bar{\Pi}_\varepsilon$ the set of perennial ε -spoiling strategies for player 1 and player 2, respectively. ■

The ε -optimal strategies constructed for parity objectives in [7] are perennial ε -optimal strategies. This gives us the following Proposition.

Proposition 1 *The following assertions hold:*

1. For all $\varepsilon > 0$, we have $\Sigma_\varepsilon \neq \emptyset$ and $\Pi_\varepsilon \neq \emptyset$.
2. For all $\varepsilon > 0$, we have $\bar{\Sigma}_\varepsilon \neq \emptyset$ and $\bar{\Pi}_\varepsilon \neq \emptyset$.

3 Single strongly connected component games

In this section we prove the existence of ε -Nash equilibrium for every $\varepsilon > 0$, in a subclass of concurrent games, namely, single strongly connected component games. In the next section we generalize the existence of ε -Nash equilibrium, for every $\varepsilon > 0$, to all concurrent games using the result for single strongly connected component games. Given a game structure $\mathcal{G}$ we define a underlying graph $G_{\mathcal{G}}$ associated with $\mathcal{G}$.

Definition 9 (Graph of a game $\mathcal{G}$) *Given a concurrent game structure $\mathcal{G} = \langle S, \text{Moves}, \Gamma_1, \Gamma_2, \delta \rangle$ the graph of game $\mathcal{G}$ is a directed graph $G_{\mathcal{G}} = (S_{\mathcal{G}}, E_{\mathcal{G}})$ that is defined as follows:*

- $S_{\mathcal{G}} = S$, i.e., the set of states of $G_{\mathcal{G}}$ is same as the state space of $\mathcal{G}$.
- $E_{\mathcal{G}} = \{ (s, t) \mid \exists a_1 \in \Gamma_1(s), \exists a_2 \in \Gamma_2(s). t \in \text{Dest}(s, a_1, a_2) \}$.

■

Definition 10 (Single strongly connected component (SSCC) games)

Let $\mathcal{G}$ be a concurrent game and $G_{\mathcal{G}}$ be the graph of $\mathcal{G}$. We call $\mathcal{G}$ a single strongly connected component (SSCC) game if the graph $G_{\mathcal{G}}$ satisfy the following conditions:

- The state space $S_{\mathcal{G}}$ can be partitioned into three sets: $C, U, T = \{ t_{00}, t_{10}, t_{01}, t_{11} \}$.
- C is a strongly connected component in the graph $G_{\mathcal{G}}$.
- The states $t_{ij} \in T$ are absorbing states, for $i, j \in \{ 0, 1 \}$. The priority function for the states in T are as follows: $p_1(t_{ij}) = i$ and $p_2(t_{ij}) = j$, for $i, j \in \{ 0, 1 \}$. Note that at state t_{00} objective of both the players are satisfied; at state t_{01} only player 1's objective is satisfied; at state t_{10} only player 2's objective is satisfied and at state t_{11} none of the players objective is satisfied.
- For every state $s \in U$ we have $|\Gamma_i(s)| = 1$ for $i \in \{ 1, 2 \}$ and $(\{ s \} \times S_{\mathcal{G}}) \cap E_{\mathcal{G}} \subseteq \{ s \} \times T$. In other words, at states in U there is no non-trivial choice of moves for the players and thus for any state s in U the game proceeds to the set T according to the probability distribution of the transition relation at s .
- $C \times (S_{\mathcal{G}} \setminus C) \cap E_{\mathcal{G}} \subseteq C \times U$, i.e., the edges going out of C ends at a state in U .

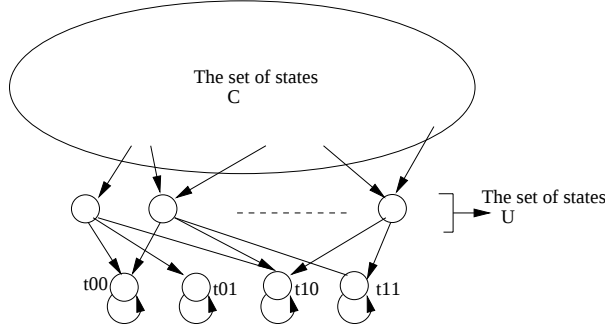


Figure 1: A SSCC game

Figure 1 illustrates a SSCC game. ■

The following Proposition states that if existence of ε -Nash equilibrium is established at a state s , then state s can be replaced by some gadget and to prove existence of ε -Nash equilibrium in original game it suffices to prove existence of ε -Nash equilibrium in the transformed game with the gadget replaced for state s .

Proposition 2 *Let $\mathcal{G}$ be a SSCC game with objective Ψ_1 and Ψ_2 for player 1 and player 2, respectively. Suppose (σ^*, π^*) is an ε -Nash equilibrium profile at s , with $\varepsilon \rightarrow 0$, such that $x_1(s) = \Pr_s^{\sigma^*, \pi^*}(\Psi_{1s})$ and $x_2(s) = \Pr_s^{\sigma^*, \pi^*}(\Psi_{2s})$. The game graph $\mathcal{G}$ can be transformed to a game graph $\mathcal{G}'$ by replacing the state s with the following gadget (Figure 2) such that if there is an ε -Nash equilibrium in the transformed game $\mathcal{G}'$, for every $\varepsilon > 0$, then there is an ε -Nash equilibrium in the original game $\mathcal{G}$, for every $\varepsilon > 0$. The gadget is as follows:*

- *Without loss of generality let $x_1(s) \leq x_2(s)$ (when $x_2(s) \leq x_1(s)$ the gadget is symmetric). Then gadget to replace s is as follows: $\Gamma_1(s) = \{a\}$, $\Gamma_2(s) = \{b\}$, and*

$$\delta(s, a, b)(t_{00}) = x_1(s), \quad \delta(s, a, b)(t_{10}) = x_2(s) - x_1(s),$$

$$\delta(s, a, b)(t_{11}) = 1 - x_2(s), \quad \delta(s, a, b)(t_{01}) = 0$$

where t_{ij} are as defined in Definition 10.

The gadget is illustrated in figure Fig 2. The construction ensures that at state s the set $\{t_{00}, t_{01}\}$ of states is reached with probability $x_1(s)$, i.e., player 1's objective is satisfied with probability $x_1(s)$, and the set $\{t_{00}, t_{10}\}$ of

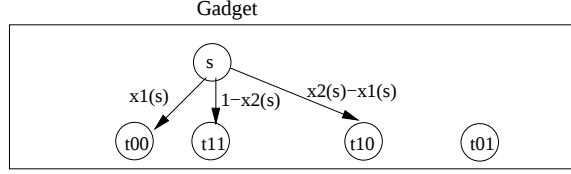


Figure 2: The gadget

states is reached with probability $x_2(s)$, i.e., player 2's objective is satisfied with probability $x_2(s)$. ■

The result follows from the observation that player 1 and player 2 can switch to strategies (σ^*, π^*) when the game reaches s .

Lemma 1 *Let $\mathcal{G}$ be SSCC game with objective Ψ_1 and Ψ_2 for player 1 and player 2, respectively. If any of the following four properties (P1-P4) hold, then for every $\varepsilon > 0$, there is an ε -Nash equilibrium (σ^*, π^*) for every state $s \in C$. The properties (P1-P4) are as follows:*

- (P1) *There is a state $s \in C$ such that $\langle\langle 1, 2 \rangle\rangle_{val}(\Psi_1 \cap \Psi_2)(s) = 1$.*
- (P2) *There is a state $s \in C$ such that $\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s) = 1$.*
- (P3) *There is a state $s \in C$ such that $\langle\langle 2 \rangle\rangle_{val}(\Psi_2)(s) = 1$.*
- (P4) *There is a state $s \in C$ such that $\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s) = 0$ and $\langle\langle 2 \rangle\rangle_{val}(\Psi_2)(s) = 0$.*

Proof. The proof is by induction on the size of C , i.e., induction on $|C|$. It is trivial for the base case when $|C| = 0$. We now prove the inductive case:

1. Suppose there is a state $s \in C$ such that $\langle\langle 1, 2 \rangle\rangle_{val}(\Psi_1 \cap \Psi_2)(s) = 1$, then there is a strategy profile (σ, π) such that $\Pr_s^{\sigma, \pi}(\Psi_{1s}) = 1$ and $\Pr_s^{\sigma, \pi}(\Psi_{2s}) = 1$. Since 1 is the maximum payoff a player can achieve, clearly (σ, π) is a Nash equilibrium. By Proposition 2 we can replace s by the gadget as described in Proposition 2. This breaks C into smaller strongly connected components. We can then apply the induction hypothesis on the smaller strongly connected components in a bottom-up order. The idea is as follows: consider a strongly connected component $C' \subset C$ in the game where s is replaced by the gadget of Proposition 2. By inductive hypothesis it follows that for every strongly connected component $C_1 \subseteq C$ such that C_1 is lower than C' (i.e., there is a path

from some state in C' to a state in C_1 in the graph of the transformed game), ε -Nash equilibrium exists for every state $s_1 \in C_1$. This follows by induction hypothesis since $|C_1| < |C|$. By Proposition 2 every state $s_1 \in C_1$ can be replaced by the gadget of Proposition 2. Hence the strongly connected component C' and the set of strongly connected components lower than C' replaced by the gadget, form a SSSC game. Since $|C'| < |C|$, by induction hypothesis on C' there exists ε -Nash equilibrium (hence also ε -Nash equilibrium values) from every state $s' \in C'$. Then by applying Proposition 2 we can replace each state $s' \in C'$ by the gadget as described in Proposition 2 and proceed.

2. Suppose there is a state $s \in C$ such that $\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s) = 1$. Then for every $\varepsilon > 0$, there is an ε -optimal strategy σ_ε for player 1 such that $\inf_{\pi \in \Pi} \Pr_s^{\sigma_\varepsilon, \pi}(\Psi_{1s}) \geq 1 - \varepsilon$. Consider a strategy π^* such that $\Pr_s^{\sigma_\varepsilon, \pi^*}(\Psi_{2s}) \geq \sup_{\pi \in \Pi} \Pr_s^{\sigma_\varepsilon, \pi}(\Psi_{2s}) - \varepsilon$. In other words we fix an ε -optimal strategy σ_ε for player 1 and a strategy π^* for player 2 that ensures player 2 the maximal probability to satisfy Ψ_2 against the strategy σ_ε , within ε -precision. Thus we have

$$\sup_{\sigma \in \Sigma} \Pr_s^{\sigma, \pi^*}(\Psi_{1s}) \leq 1 \leq \Pr_s^{\sigma_\varepsilon, \pi^*}(\Psi_{1s}) + \varepsilon;$$

$$\sup_{\pi \in \Pi} \Pr_s^{\sigma_\varepsilon, \pi}(\Psi_{2s}) \leq \Pr_s^{\sigma_\varepsilon, \pi^*}(\Psi_{2s}) + \varepsilon.$$

Hence $(\sigma_\varepsilon, \pi^*)$ is an ε -Nash equilibrium. Hence we can fix the value of ε -Nash equilibrium at state s and then the argument to prove that ε -Nash equilibrium exists for every state in C follows from the induction hypothesis and Proposition 2 as described earlier. The proof for the case when we have a state s such that $\langle\langle 2 \rangle\rangle_{val}(\Psi_2)(s) = 1$ is symmetric.

3. Suppose there is a state $s \in C$ such that $\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s) = 0$ and $\langle\langle 2 \rangle\rangle_{val}(\Psi_2)(s) = 0$. Then consider ε -spoiling strategy pair $(\sigma_\varepsilon, \pi_\varepsilon) \in \overline{\Sigma_\varepsilon} \times \overline{\Pi_\varepsilon}$. Since σ_ε is an ε -spoiling strategy it follows that

$$\sup_{\pi \in \Pi} \Pr_s^{\sigma_\varepsilon, \pi}(\Psi_{2s}) \leq \varepsilon.$$

Similarly, since π_ε is an ε -spoiling strategy we have

$$\sup_{\sigma \in \Sigma} \Pr_s^{\sigma, \pi_\varepsilon}(\Psi_{1s}) \leq \varepsilon.$$

Hence $(\sigma_\varepsilon, \pi_\varepsilon)$ is an ε -Nash equilibrium at s . The argument to prove that there is an ε -Nash equilibrium at every state in C follows from argument similar to the previous cases.

The desired result follows. ■

In the next sub-section we show that if the four properties (P1-P4) of Lemma 1 are not satisfied then the nonzero-sum SSCC game with parity objectives can be reduced to a nonzero-sum game with reachability objectives with some desired properties. The reachability objectives are $\text{Reach}(\{t_{00}, t_{01}\})$ for player 1 and $\text{Reach}(\{t_{00}, t_{10}\})$ for player 2. We then establish the existence of ε -Nash equilibrium, for every $\varepsilon > 0$, in the original game by the use of punishing or spoiling strategies.

3.1 Non-zero sum reachability game

Let $W_1 = \{t_{00}, t_{01}\}$ and $W_2 = \{t_{00}, t_{10}\}$. We consider the nonzero-sum reachability game $\mathcal{G}_R$ such that the objective for player 1 is to reach W_1 , i.e., $\text{Reach}(W_1)$, and objective for player 2 is to reach W_2 , i.e., $\text{Reach}(W_2)$. Lemma 2 to Lemma 4 were proved in [2]; we present the proofs for sake of completeness.

In sequel, we consider stochastic trees $\mathcal{T}_{\mathcal{A},s}^{\sigma,\pi}$ such that $\Pr_s^{\sigma,\pi}(\mathcal{A}) > 0$. Given a stochastic tree $\mathcal{T}_{\mathcal{A},s}^{\sigma,\pi}$, let κ be a subset of nodes, i.e., $\kappa \subseteq \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$. Analogous to the definition of reachability and safety we define the following notions of reachability and safety in the stochastic tree:

1. *Reachability in tree.* For a set $\kappa \subseteq \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, let

$$\text{ReachTree}(\kappa) = \{\langle \tau \rangle \mid \tau \text{ is an infinite path in } \text{Tr}_{\mathcal{A},s}^{\sigma,\pi} \text{ such that } \exists i \in \mathbb{N}. \tau_i \in \kappa\},$$

denote the set of paths that reach the subset κ of nodes.

2. *Safety in tree.* For a set $\kappa \subseteq \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, let

$$\text{SafeTree}(\kappa) = \{\langle \tau \rangle \mid \tau \text{ is an infinite path in } \text{Tr}_{\mathcal{A},s}^{\sigma,\pi} \text{ such that } \forall i \in \mathbb{N}. \tau_i \in \kappa\},$$

denote the set of paths that stay safe in the subset κ of nodes.

Given a positive integer k and a set $\kappa \subseteq \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, we define by $\text{ReachTree}^k(\kappa) = \{\langle \tau \rangle \mid \exists x \in \tau. \exists i \leq k. x_i \in \kappa\}$, i.e., the set of paths that reaches κ within k steps.

Lemma 2 (Reachability Lemma) *Let $\mathcal{T}_{\mathcal{A},s}^{\sigma,\pi}$ be a stochastic tree.*

1. *For a set $\kappa \subseteq \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, if $\inf_{x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}} \Pr_x^{\sigma,\pi}(\text{ReachTree}(\kappa) \mid \mathcal{A}) > 0$, then $\Pr_x^{\sigma,\pi}(\text{ReachTree}(\kappa) \mid \mathcal{A}) = 1$, for all nodes $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$.*

2. For a set $B \subseteq S$, if $\inf_{x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}} \Pr_x^{\sigma,\pi}(\text{Reach}(B) \mid \mathcal{A}) > 0$, then $\Pr_x^{\sigma,\pi}(\text{Reach}(B) \mid \mathcal{A}) = 1$, for all nodes $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$.

Proof. We prove the first case and show that the second case is an immediate consequence.

1. Let $0 < c \leq \inf_{x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}} \Pr_x^{\sigma,\pi}(\text{ReachTree}(\kappa) \mid \mathcal{A})$. Chose $0 < c' < c$. For every node $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, there exists k_x such that $\Pr_x^{\sigma,\pi}(\text{ReachTree}^{k_x}(\kappa) \mid \mathcal{A}) \geq c'$. Consider $k_1 = k_\epsilon$ (recall that ϵ is the root of the tree) and consider the frontier F_1 of $\text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$ at depth k_1 . Given a frontier F at depth k , let $\overline{F}$ be the set of nodes x in F such that the path from the root to x has not visited a node in κ , i.e., none of $\epsilon, x_1, x_2, \dots, x_{|x|}$ is in κ . For a frontier F_i , define $k_{i+1} = \max\{k_x \mid x \in \overline{F}_i\}$. Inductively, define the frontier F_{i+1} at depth $\sum_{j=1}^{i+1} k_j$. It follows that for $k = \sum_{i=1}^n k_i$ we have $\Pr_s^{\sigma,\pi}(\Omega \setminus \text{ReachTree}^k(\kappa) \mid \mathcal{A}) \leq (1 - c')^n$. Since $\lim_{n \rightarrow \infty} (1 - c')^n = 0$, the desired result follows for the root of the tree. Since $\inf_{x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}} \Pr_x^{\sigma,\pi}(\text{ReachTree}(\kappa) \mid \mathcal{A}) > 0$, it follows that for all node $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$ we have $\inf_{x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)} \Pr_{x_1}^{\sigma,\pi}(\text{ReachTree}(\kappa) \mid \mathcal{A}) > 0$. Arguing similarly for the subtree rooted at the node x the desired result follows.

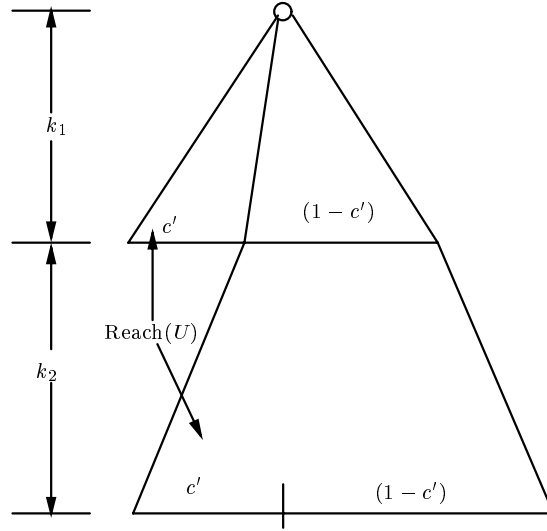


Figure 3: The Stochastic Tree for Reachability

2. Observe that with $\kappa = \{ x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi} \mid \langle x \rangle \in B \}$, we have $\text{Reach}(B) = \text{ReachTree}(\kappa)$. The result is immediate from part 1. ■

Notations. Let $\mathcal{A} \subseteq \Omega_s$ be a measurable event such that $\Pr_s^{\sigma,\pi}(\mathcal{A}) > 0$. For a set $B \subseteq S$, let $\text{InfSet}(B) = \{\omega \mid \text{Inf}(\omega) \subseteq B\}$. For a set $B \subseteq S$, let $\text{InfSetEq}(B) = \{\omega \mid \text{Inf}(\omega) = B\}$. Given a node x in $\text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, and $\varepsilon > 0$, we define $\mathcal{C}_{\mathcal{A},\varepsilon}^{\sigma,\pi}(x)$ as follows:

$$\mathcal{C}_{\mathcal{A},\varepsilon}^{\sigma,\pi}(x) = \{ B \subseteq S \mid \Pr_x^{\sigma,\pi}(\text{InfSet}(B) \mid \mathcal{A}) \geq 1 - \varepsilon \}.$$

Note that for $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$, such that $\varepsilon_1 \leq \varepsilon_2$, for all node $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, if $B \in \mathcal{C}_{\mathcal{A},\varepsilon_1}^{\sigma,\pi}(x)$ then $B \in \mathcal{C}_{\mathcal{A},\varepsilon_2}^{\sigma,\pi}(x)$. We define by $\mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x) = \lim_{\varepsilon \rightarrow 0} \mathcal{C}_{\mathcal{A},\varepsilon}^{\sigma,\pi}(x)$. The monotonicity property of $\mathcal{C}_{\mathcal{A},\varepsilon}^{\sigma,\pi}$ with respect to ε ensures that $\mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x)$ exists for all $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$.

Lemma 3 *For every node $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, there is a unique minimal element of $\mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x)$ under the $\subset$ ordering.*

Proof. Consider a node $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$. Let B_1 and B_2 be two distinct minimal elements in $\mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x)$. Consider any arbitrary $\varepsilon > 0$. It follows from the definition that we have $\Pr_x^{\sigma,\pi}(\text{InfSet}(B_i) \mid \mathcal{A}) \geq 1 - \frac{\varepsilon}{2}$, for $i \in \{1, 2\}$. By definition we must have $\Pr_x^{\sigma,\pi}(\text{InfSet}(B_1 \cup B_2) \mid \mathcal{A}) \leq 1$. Hence we have the following equation:

$$\Pr_x^{\sigma,\pi}(\text{InfSet}(B_1) \mid \mathcal{A}) + \Pr_x^{\sigma,\pi}(\text{InfSet}(B_2) \mid \mathcal{A}) - \Pr_x^{\sigma,\pi}(\text{InfSet}(B_1 \cap B_2) \mid \mathcal{A}) \leq 1$$

Hence it follows that $\Pr_x^{\sigma,\pi}(\text{InfSet}(B_1 \cap B_2) \mid \mathcal{A}) \geq 1 - \varepsilon$. Hence for every $\varepsilon > 0$, we have $\Pr_x^{\sigma,\pi}(\text{InfSet}(B_1 \cap B_2) \mid \mathcal{A}) \geq 1 - \varepsilon$. Hence, $B_1 \cap B_2 \in \mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x)$. However, this is a contradiction to the assumption that B_1 and B_2 are distinct minimal elements of $\mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x)$. ■

We define the function $\mathcal{M}_{\mathcal{A}}^{\sigma,\pi} : \text{Tr}_{\mathcal{A},s}^{\sigma,\pi} \rightarrow 2^S$ that assigns to every node $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$ the minimum element of $\mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x)$. Formally, we have

$$\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x) = \bigcap_{B \in \mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x)} B = \lim_{\varepsilon \rightarrow 0} \bigcap_{B \in \mathcal{C}_{\mathcal{A},\varepsilon}^{\sigma,\pi}(x)} B.$$

Proposition 3 *For every $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, for every successor x_1 of x we have $\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x_1) \subseteq \mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x)$.*

Proof. By definition for all nodes $x, x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, such that x_1 is a successor of x we have $\mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x_1) \subseteq \mathcal{C}_{\mathcal{A}}^{\sigma,\pi}(x)$. The result is an easy consequence of the above fact. ■

Lemma 4 Given a S -labeled tree $\mathcal{T}_{\mathcal{A},s}^{\sigma,\pi}$, for all node $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, for all $\varepsilon > 0$, there is a set $B \subseteq S$ and $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$, such that for all node $x_2 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x_1)$ we have

$$\Pr_{x_2}^{\sigma,\pi}(\text{InfSetEq}(B) \mid \mathcal{A}) \geq 1 - \varepsilon.$$

Proof. The proof is by induction on $|\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x)|$.

Base Case. If $|\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x)| = 1$, let $\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x) = \{s\}$. Then for all nodes $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$ we have $\Pr_{x_1}^{\sigma,\pi}(\text{InfSet}(\{s\}) \mid \mathcal{A}) \geq 1 - \varepsilon$, for all $\varepsilon > 0$. Thus for all nodes $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$, for all $\varepsilon > 0$, we have $\Pr_{x_1}^{\sigma,\pi}(\text{InfSetEq}(\{s\}) \mid \mathcal{A}) \geq 1 - \varepsilon$.

Inductive Case. Suppose there exist a node $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$ such that $\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x_1) \subsetneq \mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x)$, then $|\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x_1)| < |\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x)|$ and the result follows by inductive hypothesis at x_1 . Otherwise for every node $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$ we have $\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x_1) = \mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x)$. Let the set $\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x)$ be B . We have $\lim_{\varepsilon \rightarrow 0} \bigcap_{x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)} \left(\bigcap_{D \in \mathcal{C}_{\mathcal{A},\varepsilon}^{\sigma,\pi}(x_1)} D \right) = B$.

- Suppose we have $\inf_{x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)} \Pr_{x_1}^{\sigma,\pi}(\text{Reach}(\{s\}) \mid \mathcal{A}) > 0$, for all states $s \in B$. Then it follows from Lemma 2 that for all nodes $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$ we have $\Pr_{x_1}^{\sigma,\pi}(\text{Reach}(\{s\}) \mid \mathcal{A}) = 1$. Hence for all nodes $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$ we have $\Pr_{x_1}^{\sigma,\pi}(\text{InfSetEq}(B) \mid \mathcal{A}) = 1$.
- Otherwise, consider a state $s \in B$ such that $\inf_{x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)} \Pr_{x_1}^{\sigma,\pi}(\text{Reach}(\{s\}) \mid \mathcal{A}) = 0$. Hence it follows, for every $\varepsilon > 0$, there is a node $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$ such that $\Pr_{x_1}^{\sigma,\pi}(\text{InfSet}(B \setminus \{s\}) \mid \mathcal{A}) \geq 1 - \varepsilon$. Formally, we have $\lim_{\varepsilon \rightarrow 0} \bigcap_{x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)} \left(\bigcap_{D \in \mathcal{C}_{\mathcal{A},\varepsilon}^{\sigma,\pi}(x_1)} D \right) \subseteq B \setminus \{s\}$. This is a contradiction to the fact that for all nodes $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$ we have $\mathcal{M}_{\mathcal{A}}^{\sigma,\pi}(x_1) = B$ (i.e., $\lim_{\varepsilon \rightarrow 0} \bigcap_{x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)} \left(\bigcap_{D \in \mathcal{C}_{\mathcal{A},\varepsilon}^{\sigma,\pi}(x_1)} D \right) = B$).

The desired result follows. ■

Lemma 5 Given a stochastic tree $\mathcal{T}_{\mathcal{A},s}^{\sigma,\pi}$, for all node $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, for every $\varepsilon > 0$, there is a node $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$ such that for all node $x_2 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x_1)$ one of the following conditions (C1-C4) hold:

1. (C1) $\Pr_{x_2}^{\sigma,\pi}(\Psi_{1s} \mid \mathcal{A}) \geq 1 - \varepsilon$ and $\Pr_{x_2}^{\sigma,\pi}(\Psi_{2s} \mid \mathcal{A}) \geq 1 - \varepsilon$;
2. (C2) $\Pr_{x_2}^{\sigma,\pi}(\Psi_{1s} \mid \mathcal{A}) \geq 1 - \varepsilon$ and $\Pr_{x_2}^{\sigma,\pi}(\Psi_{2s} \mid \mathcal{A}) \leq \varepsilon$;
3. (C3) $\Pr_{x_2}^{\sigma,\pi}(\Psi_{1s} \mid \mathcal{A}) \leq \varepsilon$ and $\Pr_{x_2}^{\sigma,\pi}(\Psi_{2s} \mid \mathcal{A}) \geq 1 - \varepsilon$;
4. (C4) $\Pr_{x_2}^{\sigma,\pi}(\Psi_{1s} \mid \mathcal{A}) \leq \varepsilon$ and $\Pr_{x_2}^{\sigma,\pi}(\Psi_{2s} \mid \mathcal{A}) \leq \varepsilon$.

Proof. It follows from Lemma 4 that for every $\varepsilon > 0$, there is a node $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$, and a set B such that for all node $x_2 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x_1)$ we have $\Pr_{x_2}^{\sigma,\pi}(\text{InfSetEq}(B) \mid \mathcal{A}) \geq 1 - \varepsilon$. The following case analysis proves the result:

1. If $\min(p_1(B))$ is even and $\min(p_2(B))$ is even then condition 1 (C1) is satisfied.
2. If $\min(p_1(B))$ is even and $\min(p_2(B))$ is odd then condition 2 (C2) is satisfied.
3. If $\min(p_1(B))$ is odd and $\min(p_2(B))$ is even then condition 3 (C3) is satisfied.
4. If $\min(p_1(B))$ is odd and $\min(p_2(B))$ is odd then condition 4 (C4) is satisfied.

Hence, it also follows that for every stochastic tree $\text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$ for all node $x \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}$, for every $\varepsilon > 0$, there is a node $x_1 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x)$ such that for all node $x_2 \in \text{Tr}_{\mathcal{A},s}^{\sigma,\pi}(x_1)$ either $\max\{\Pr_{x_2}^{\sigma,\pi}(\Psi_{1s} \mid \mathcal{A}), \Pr_{x_2}^{\sigma,\pi}(\Psi_{2s} \mid \mathcal{A})\} \geq 1 - \varepsilon$; or $\min\{\Pr_{x_2}^{\sigma,\pi}(\Psi_{1s} \mid \mathcal{A}), \Pr_{x_2}^{\sigma,\pi}(\Psi_{2s} \mid \mathcal{A})\} \leq \varepsilon$. ■

Punishing perennial ε -optimal strategy construction. We consider punishing perennial ε -optimal strategy profile $(\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon)$ that are defined as follows:

$$\hat{\sigma}_\varepsilon(s_0, s_1, \dots, s_k) = \begin{cases} \sigma_\varepsilon(s_0, s_1, \dots, s_k) & \text{if } \langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s_k) > 0 \\ \bar{\sigma}_\varepsilon(s_0, s_1, \dots, s_k) & \text{if } \langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s_k) = 0 \end{cases}$$

where $\sigma_\varepsilon \in \Sigma_\varepsilon$ and $\bar{\sigma}_\varepsilon \in \overline{\Sigma_\varepsilon}$. That is player 1 follows a perennial ε -optimal strategy σ_ε when the play is in a state with positive zero-sum value for player 1; else it follows a perennial ε -spoiling strategy $\bar{\sigma}_\varepsilon$. It is easy to observe that since $\sigma_\varepsilon \in \Sigma_\varepsilon$ we have $\hat{\sigma}_\varepsilon \in \Sigma_\varepsilon$. Similarly we define the strategy $\hat{\pi}_\varepsilon$ as follows:

$$\hat{\pi}_\varepsilon(s_0, s_1, \dots, s_k) = \begin{cases} \pi_\varepsilon(s_0, s_1, \dots, s_k) & \text{if } \langle\langle 2 \rangle\rangle_{\text{val}}(\Psi_2)(s_k) > 0 \\ \bar{\pi}_\varepsilon(s_0, s_1, \dots, s_k) & \text{if } \langle\langle 2 \rangle\rangle_{\text{val}}(\Psi_2)(s_k) = 0 \end{cases}$$

where $\pi_\varepsilon \in \Pi_\varepsilon$ and $\bar{\pi}_\varepsilon \in \overline{\Pi_\varepsilon}$.

Lemma 6 *Let $(\sigma, \pi) \in \Sigma \times \Pi$ be an arbitrary strategy profile, and let $\kappa = \{x \in \text{Tr}_s^{\sigma,\pi} \mid \Pr_x^{\sigma,\pi}(\text{Safe}(C)) > 0\}$. For all node $x \in \text{Tr}_s^{\sigma,\pi}$ the following assertions hold:*

1. $\Pr_x^{\sigma, \pi}(\text{Safe}(C)) = \Pr_x^{\sigma, \pi}(\text{SafeTree}(\kappa))$.
2. If $\Pr_x^{\sigma, \pi}(\text{Safe}(C)) > 0$, then for every $\eta > 0$, there exists $x_1 \in \text{Tr}_s^{\sigma, \pi}(x)$ such that $\Pr_{x_1}^{\sigma, \pi}(\text{SafeTree}(\kappa)) \geq 1 - \eta$.

Proof.

1. For every node $x_1 \in \bar{\kappa} = (\text{Tr}_s^{\sigma, \pi} \setminus \kappa)$ we have $\Pr_{x_1}^{\sigma, \pi}(\text{Reach}(U)) = 1$. Hence for all node $x_1 \in \text{Tr}_s^{\sigma, \pi}(x)$ we have $\Pr_{x_1}^{\sigma, \pi}(\text{Reach}(U) \mid \text{ReachTree}(\bar{\kappa})) = 1$, i.e., $\Pr_x^{\sigma, \pi}(\text{Safe}(C) \mid \text{ReachTree}(\bar{\kappa})) = 0$. For every node $x_1 \in \kappa$, since $\Pr_{x_1}^{\sigma, \pi}(\text{Safe}(C)) > 0$, we have $\langle x_1 \rangle \in C$. Thus for every node x we have $\Pr_x^{\sigma, \pi}(\text{Safe}(C) \mid \text{SafeTree}(\kappa)) = 1$. Hence we have

$$\begin{aligned}
\Pr_x^{\sigma, \pi}(\text{Safe}(C)) &= \Pr_x^{\sigma, \pi}(\text{Safe}(C) \mid \text{SafeTree}(\kappa)) \cdot \Pr_x^{\sigma, \pi}(\text{SafeTree}(\kappa)) \\
&+ \Pr_x^{\sigma, \pi}(\text{Safe}(C) \mid \text{ReachTree}(\bar{\kappa})) \cdot \Pr_x^{\sigma, \pi}(\text{ReachTree}(\bar{\kappa})) \\
&= \Pr_x^{\sigma, \pi}(\text{Safe}(C) \mid \text{SafeTree}(\kappa)) \cdot \Pr_x^{\sigma, \pi}(\text{SafeTree}(\kappa)) \\
&= \Pr_x^{\sigma, \pi}(\text{SafeTree}(\kappa))
\end{aligned}$$

The desired result follows.

2. It follows from above that for all node x if $\Pr_x^{\sigma, \pi}(\text{Safe}(C)) > 0$, then $\Pr_x^{\sigma, \pi}(\text{SafeTree}(\kappa)) > 0$. Hence we must have $\inf_{x_1 \in \text{Tr}_s^{\sigma, \pi}(x)} \Pr_{x_1}^{\sigma, \pi}(\text{ReachTree}(\bar{\kappa})) = 0$; otherwise, if $\inf_{x_1 \in \text{Tr}_s^{\sigma, \pi}(x)} \Pr_{x_1}^{\sigma, \pi}(\text{ReachTree}(\bar{\kappa})) > 0$, then it follows from Lemma 2 that $\Pr_x^{\sigma, \pi}(\text{ReachTree}(\bar{\kappa})) = 1$, i.e., $\Pr_x^{\sigma, \pi}(\text{SafeTree}(\kappa)) = 0$. Since $\inf_{x_1 \in \text{Tr}_s^{\sigma, \pi}(x)} \Pr_{x_1}^{\sigma, \pi}(\text{ReachTree}(\bar{\kappa})) = 0$ we have $\sup_{x_1 \in \text{Tr}_s^{\sigma, \pi}(x)} \Pr_{x_1}^{\sigma, \pi}(\text{SafeTree}(\kappa)) = 1$. Hence for all $\eta > 0$, there exists $x_1 \in \text{Tr}_s^{\sigma, \pi}(x)$ such that $\Pr_{x_1}^{\sigma, \pi}(\text{SafeTree}(\kappa)) \geq 1 - \eta$.

■

Lemma 7 *Let x be a node in the stochastic tree $\text{Tr}_s^{\sigma, \pi}$, and $\eta > 0$, and $\mathcal{A}$ be an event such that $\Pr_x^{\sigma, \pi}(\mathcal{A}) \geq 1 - \eta$. For all objective Ψ the following assertions hold:*

1. If $\Pr_x^{\sigma, \pi}(\Psi \mid \mathcal{A}) \geq 1 - \varepsilon$, then $\Pr_x^{\sigma, \pi}(\Psi) \geq 1 - \varepsilon - \eta$.
2. If $\Pr_x^{\sigma, \pi}(\Psi \mid \mathcal{A}) \leq \varepsilon$, then $\Pr_x^{\sigma, \pi}(\Psi) \leq \varepsilon + \eta$.

Proof.

1. If $\Pr_x^{\sigma, \pi}(\Psi \mid \mathcal{A}) \geq 1 - \varepsilon$, then

$$\begin{aligned} \Pr_x^{\sigma, \pi}(\Psi) &\geq \Pr_x^{\sigma, \pi}(\Psi \cap \mathcal{A}) \\ &= \Pr_x^{\sigma, \pi}(\Psi \mid \mathcal{A}) \cdot \Pr_x^{\sigma, \pi}(\mathcal{A}) \\ &\geq (1 - \varepsilon) \cdot (1 - \eta) = 1 - \varepsilon - \eta + \eta \cdot \varepsilon \geq 1 - \varepsilon - \eta. \end{aligned}$$

2. If $\Pr_x^{\sigma, \pi}(\Psi \mid \mathcal{A}) \leq \varepsilon$, then

$$\begin{aligned} \Pr_x^{\sigma, \pi}(\Psi) &= \Pr_x^{\sigma, \pi}(\Psi \cap \mathcal{A}) + \Pr_x^{\sigma, \pi}(\Psi \cap \overline{\mathcal{A}}) \\ &\leq \Pr_x^{\sigma, \pi}(\Psi \mid \mathcal{A}) \cdot \Pr_x^{\sigma, \pi}(\mathcal{A}) + \Pr_x^{\sigma, \pi}(\overline{\mathcal{A}}) \\ &\leq \varepsilon + \eta. \end{aligned}$$

Hence the Lemma follows. ■

Lemma 8 *Suppose properties (P1-P4) of Lemma 1 are not satisfied. For all state $s \in S$, we have $\Pr_s^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Reach}(U)) = 1$, where $\hat{\sigma}_\varepsilon$ and $\hat{\pi}_\varepsilon$ are punishing perennial ε -optimal strategies and $\varepsilon \rightarrow 0$.*

Proof. Let

$$\alpha_1^{\min} = \min\{ \langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s) \mid s \in S, \langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s) > 0 \}$$

denote the least positive zero-sum value for player 1 for a state $s \in S$; and

$$\alpha_1^{\max} = \max\{ \langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s) \mid s \in S, \langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s) < 1 \}$$

denote the greatest zero-sum value for player 1 that is less than 1 for a state $s \in S$. Similarly, let

$$\alpha_2^{\min} = \min\{ \langle\langle 2 \rangle\rangle_{\text{val}}(\Psi_2)(s) \mid s \in S, \langle\langle 2 \rangle\rangle_{\text{val}}(\Psi_2)(s) > 0 \}$$

denote the least positive zero-sum value for player 2 for a state $s \in S$; and

$$\alpha_2^{\max} = \max\{ \langle\langle 2 \rangle\rangle_{\text{val}}(\Psi_2)(s) \mid s \in S, \langle\langle 2 \rangle\rangle_{\text{val}}(\Psi_2)(s) < 1 \}$$

denote the greatest zero-sum value for player 2 that is less than 1 for a state $s \in S$. Let

$$\alpha_{\langle 1, 2 \rangle}^{\max} = \max\{ \langle\langle 1, 2 \rangle\rangle_{\text{val}}(\Psi_1 \cap \Psi_2)(s) \mid s \in C, \langle\langle 1, 2 \rangle\rangle_{\text{val}}(\Psi_1 \cap \Psi_2)(s) < 1 \}$$

denote the greatest cooperative value with objective $\Psi_1 \cap \Psi_2$ that is less than 1, for a state $s \in C$. Let $\alpha = \min\{ \alpha_1^{\min}, \alpha_2^{\min}, 1 - \alpha_1^{\max}, 1 - \alpha_2^{\max}, 1 - \alpha_{\langle 1, 2 \rangle}^{\max} \}$

$\alpha_{\langle 1,2 \rangle}^{\max}$ }. Note that $0 < \alpha \leq \frac{1}{2}$. Fix β such that $0 < 6\beta < \alpha$, and fix η and ε such that $0 < \varepsilon < \eta < \beta^2$.

We consider the perennial ε -optimal strategy profile $(\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon)$ as described by punishing perennial ε -optimal strategy construction. Let $B_s^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon} = \{x \in \text{Tr}_s^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon} \mid \Pr_x^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Safe}(C)) > 0\}$. If $B_s^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}$ is empty the Lemma follows. Assume for the sake of contradiction that $B_s^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}$ is non-empty.

Let $x \in B_s^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}$, i.e., $\Pr_x^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Safe}(C)) > 0$. Let $\kappa = \{x_1 \in \text{Tr}_s^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon} \mid \Pr_{x_1}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Safe}(C)) > 0\}$. Since $\Pr_x^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Safe}(C)) > 0$, it follows from Lemma 6 that $\Pr_x^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{SafeTree}(\kappa)) > 0$. Consider the event $\mathcal{A} = \text{SafeTree}(\kappa)$. Let $x_1 \in \text{Tr}_{\mathcal{A}, s}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(x)$ such that one of the conditions (C1-C4) of Lemma 5 are satisfied for ε , i.e., for every $x_2 \in \text{Tr}_{\mathcal{A}, s}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(x_1)$ one of the conditions (C1-C4) of Lemma 5 hold for ε . Since $\mathcal{A} = \text{SafeTree}(\kappa)$, and $x_1 \in \text{Tr}_{\mathcal{A}, s}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(x)$ we have $x_1 \in \kappa$. Hence $\Pr_{x_1}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Safe}(C)) > 0$, and it follows from Lemma 6 that $\Pr_{x_1}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{SafeTree}(\kappa)) > 0$. Again it follows from Lemma 6 that there is a node $x_3 \in \text{Tr}_{\mathcal{A}, s}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(x_1)$ such that $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{SafeTree}(\kappa)) > 1 - \eta$ and also x_3 satisfies one of the conditions (C1-C4) for ε . We analyze the following four cases:

1. If condition (C1) of Lemma 5 holds, then we have $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1 \mid \text{SafeTree}(\kappa)) \geq 1 - \varepsilon$ and $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_2 \mid \text{SafeTree}(\kappa)) \geq 1 - \varepsilon$. Since $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{SafeTree}(\kappa)) \geq 1 - \eta$, from Lemma 7 we have $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1) \geq 1 - \varepsilon - \eta \geq 1 - 2\eta$ and $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_2) \geq 1 - \varepsilon - \eta \geq 1 - 2\eta$. It follows that we have $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1 \cap \Psi_2) \geq 1 - 4\eta \geq 1 - 4\beta^2 \geq 1 - 4\beta$. Let $\langle x_3 \rangle = s_1 \in C$, and consider the strategy pair $(\tilde{\sigma}, \tilde{\pi})$ defined as follows:

$$\tilde{\sigma}(s_0, s_1, \dots, s_k) = \hat{\sigma}_\varepsilon(\text{hist}(x_3), s_0, s_1, \dots, s_k)$$

and

$$\tilde{\pi}(s_0, s_1, \dots, s_k) = \hat{\pi}_\varepsilon(\text{hist}(x_3), s_0, s_1, \dots, s_k)$$

i.e., the strategies follows $\hat{\sigma}_\varepsilon$ and $\hat{\pi}_\varepsilon$ from x_3 . Hence $\Pr_{s_1}^{\tilde{\sigma}, \tilde{\pi}}(\Psi_1 \cap \Psi_2) = \Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1 \cap \Psi_2) \geq 1 - 4\beta > 1 - \alpha \geq \alpha_{\langle 1,2 \rangle}^{\max}$. Hence we must have $\sup_{(\sigma, \pi) \in \Sigma \times \Pi} \Pr_{s_1}^{\sigma, \pi}(\Psi_1 \cap \Psi_2) = 1$ and thus the property (P1) of Lemma 1 is satisfied.

2. If condition (C2) of Lemma 5 holds, then we have $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1 \mid \text{SafeTree}(\kappa)) \geq 1 - \varepsilon$ and $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_2 \mid \text{SafeTree}(\kappa)) \leq \varepsilon$. Since $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{SafeTree}(\kappa)) \geq 1 - \eta$, from Lemma 7 we have $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1) \geq 1 - \varepsilon - \eta \geq 1 - 2\eta$ and $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_2) \leq \varepsilon + \eta \leq 2\eta$. Let $W_0 = \{s \in$

$S \mid \langle\langle 2 \rangle\rangle_{val}(\Psi_2)(s) = 0 \}$ denote the set of states where the zero-sum value for player 2 is 0; and let $\overline{W_0} = S \setminus W_0$. Note that for every state $s \in \overline{W_0}$ we have $\langle\langle 2 \rangle\rangle_{val}(\Psi_2)(s) \geq \alpha_2^{\min} \geq \alpha$. Then we have

$$\begin{aligned} 2\eta \geq \Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_2) &\geq \Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_2 \cap \text{Reach}(\overline{W_0})) \\ &= \Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_2 \mid \text{Reach}(\overline{W_0})) \cdot \Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Reach}(\overline{W_0})) \\ &\geq (\alpha - \varepsilon) \cdot \Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Reach}(\overline{W_0})) \quad (\text{since } \hat{\sigma}_\varepsilon \in \Sigma_\varepsilon) \end{aligned}$$

Since $6\beta < \alpha$ and $\varepsilon < \eta < \beta^2 < \beta$ we have

$$\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Reach}(\overline{W_0})) \leq \frac{2\eta}{(\alpha - \varepsilon)} \leq \frac{2\eta}{5\beta} \leq \frac{2\beta^2}{5\beta} \leq \frac{2\beta}{5} < \beta.$$

Hence $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{Safe}(W_0)) \geq 1 - \beta$. By construction of $\hat{\pi}_\varepsilon$ if the current state s_k of the play is in W_0 then player 2 follows an ε -spoiling strategy $\bar{\pi}_\varepsilon \in \bar{\Pi}_\varepsilon$. Let $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \bar{\pi}_\varepsilon}(\Psi_1) = \rho$. Then we have

$$1 - 2\eta \leq \Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1) \leq (1 - \beta)\rho + \beta.$$

Since $\eta < \beta^2 < \beta < \alpha \leq \frac{1}{2}$ we have

$$\rho \geq \frac{1 - 2\eta - \beta}{1 - \beta} \geq \frac{1 - 3\beta}{1 - \beta} \geq 1 - \frac{2\beta}{1 - \beta} \geq 1 - 4\beta.$$

Hence we have $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \bar{\pi}_\varepsilon}(\Psi_1) \geq 1 - 4\beta > 1 - 6\beta + \varepsilon \geq 1 - \alpha + \varepsilon \geq \alpha_1^{\max} + \varepsilon$. Since $\bar{\pi}_\varepsilon$ is an ε -spoiling strategy it follows that $\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(\langle x_3 \rangle) = 1$. Since $\langle x_3 \rangle \in C$, the property (P2) of Lemma 1 holds.

3. Argument similar to previous case shows that if condition (C3) of Lemma 5 holds, then the property (P3) of Lemma 1 holds.
4. If condition (C4) of Lemma 5 holds, then we have $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1 \mid \text{SafeTree}(\kappa)) \leq \varepsilon$ and $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_2 \mid \text{SafeTree}(\kappa)) \leq \varepsilon$. Since $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\text{SafeTree}(\kappa)) \geq 1 - \eta$, from Lemma 7 we have $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1) \leq \varepsilon + \eta \leq 2\eta$ and $\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_2) \leq \varepsilon + \eta \leq 2\eta$. Since $\hat{\sigma}_\varepsilon$ is perennial ε -optimal strategy and

$$\Pr_{x_3}^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon}(\Psi_1) \leq 2\eta \leq 3\eta - \varepsilon < 3\beta - \varepsilon < \alpha_1^{\min} - \varepsilon,$$

it follows that $\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(\langle x_3 \rangle) = 0$. Similarly we also have that $\langle\langle 2 \rangle\rangle_{val}(\Psi_2)(\langle x_3 \rangle) = 0$. Since $\langle x_3 \rangle \in C$, the property (P4) of Lemma 1 holds.

Since by assumption of the Lemma properties (P1-P4) of Lemma 1 are not satisfied, we have a contradiction. Hence $B_s^{\hat{\sigma}_\varepsilon, \hat{\pi}_\varepsilon} = \emptyset$. The Lemma follows. ■

Definition 11 (Locally optimal strategy) *A selector function ξ_1^l is locally optimal if it is optimal in the “one-step” matrix game where each state is assigned a reward value $\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s)$. Formally, for all state s , for all move $a_2 \in \Gamma_2(s)$ we have*

$$\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s) \leq E[\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(\Theta_1) \mid s, \xi_1^l(s), a_2].$$

Locally optimal selector ξ_2^l for player 2 is defined symmetrically. We denote by Λ_1^l and Λ_2^l be the set of locally optimal selectors for player 1 and player 2, respectively. ■

The following Lemma is an easy consequence of Lemma 8 and Theorem 3 of [2].

Lemma 9 *For every $\varepsilon > 0$, there exists perennial ε -optimal strategy profile $(\sigma_\varepsilon, \pi_\varepsilon) \in \Sigma_\varepsilon \times \Pi_\varepsilon$ and there exists locally optimal selector $(\bar{\sigma}, \bar{\pi}) \in \Lambda_1^l \times \Lambda_2^l$ such that the following conditions hold:*

1. $\lim_{\varepsilon \rightarrow 0} \sigma_\varepsilon = \bar{\sigma}; \lim_{\varepsilon \rightarrow 0} \pi_\varepsilon = \bar{\pi}$.
2. For all state s , $\Pr_s^{\sigma_\varepsilon, \pi_\varepsilon}(\text{Reach}(U)) = 1$.

In sequel by a little abuse of notation we also denote by $\bar{\sigma}$ and $\bar{\pi}$ the memoryless strategies constructed from the locally optimal selectors $\bar{\sigma}$ and $\bar{\pi}$ of Lemma 9.

Notation. For notational simplicity we denote by $v_1(s)$ and $v_2(s)$ the zero-sum values of the games, i.e., $v_1(s) = \langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s)$ and $v_2(s) = \langle\langle 2 \rangle\rangle_{val}(\Psi_2)(s)$.

1. For a play $\omega = \langle s_0, s_1, s_2, \dots \rangle$ we write $e_{\hat{C}}(\omega) = \inf\{n \geq 1 \mid s_n \notin \hat{C}\}$, to denote the first time the play leaves $\hat{C}$.
2. For a memoryless strategy x of player 1 we denote by x_s the distribution described by the strategy x at state s . Similar notations are used for memoryless strategies y of player 2.
3. For a memoryless strategy y for player 2 we define $H_1(y, \hat{C}) = \max_{a \in \Gamma_1(s)} \max_{s \in \hat{C}} E[v_1(\Theta_1) \mid s, a, y_s]$.

4. For a memoryless strategy x for player 1 we define $H_2(x, \hat{C}) = \max_{b \in \Gamma_2(s)} \max_{s \in \hat{C}} \mathbb{E}[v_2(\Theta_1) \mid s, b, x_s]$.

Definition 12 (Perturbation) Let μ and $\tilde{\mu}$ be two distribution over S , then $\tilde{\mu}$ is a perturbation of μ if $\text{Supp}(\mu) \subseteq \text{Supp}(\tilde{\mu})$. ■

Definition 13 (Perturbed graph [29]) Given a pair of memoryless strategies (x, y) and a subset $\hat{C} \subseteq S$, the perturbed graph $\tilde{G}_{\hat{C}}(x, y)$ is a directed graph defined as follows:

- the set of states is $\hat{C}$;
- for $s, s' \in \hat{C}$, there is an edge (s, s') if there exists perturbation $(\tilde{x}_s, \tilde{y}_s)$ of (x_s, y_s) such that $\delta(s' \mid s, \tilde{x}_s, \tilde{y}_s) > 0$ and $\delta(\hat{C} \mid s, \tilde{x}_s, \tilde{y}_s) = 1$.

Intuitively, the definition captures the idea that the players can play perturbation of (x, y) to reach from s to s' without leaving the set $\hat{C}$. ■

Definition 14 (Weak-communicating sets [23, 29]) Let (x, y) be a pair of memoryless strategies. A set $\hat{C} \subseteq S$ is weak-communicating under (x, y) if the graph $\tilde{G}_{\hat{C}}(x, y)$ is strongly connected. ■

Intuitively, weak-communicating set $\hat{C}$ under (x, y) means that playing perturbation $(\tilde{x}, \tilde{y})$ of (x, y) every state $s \in \hat{C}$ is reached almost-surely without leaving $\hat{C}$. This is formalized in the Lemma below.

Lemma 10 Let $\hat{C}$ be a weak-communicating set under a memoryless strategy pair (x, y) . There exist memoryless strategies $(\tilde{x}, \tilde{y})$ such that

1. for each $s \in \hat{C}$, $(\tilde{x}_s, \tilde{y}_s)$ is a perturbation of (x_s, y_s) ;
2. $\hat{C}$ is closed under $(\tilde{x}, \tilde{y})$, i.e., for all state $s \in \hat{C}$ we have $\delta(\hat{C} \mid s, \tilde{x}_s, \tilde{y}_s) = 1$;
3. for all $s \in \hat{C}$, s is reached almost-surely (with probability 1) in finite time from every state $\hat{s} \in \hat{C}$, under $(\tilde{x}, \tilde{y})$.

Proof. For every edge $e = (s, s')$ in $\tilde{G}_{\hat{C}}(x, y)$ there is a perturbation (x_e, y_e) of (x_s, y_s) such that $\delta(s' \mid s, x_e, y_e) > 0$ and $\delta(\hat{C} \mid s, x_e, y_e) = 1$. For any state s with out-going edges $e_1, e_2, \dots, e_k$, let $(x_{e_1}, y_{e_1}), (x_{e_2}, y_{e_2}), \dots, (x_{e_k}, y_{e_k})$,

be the corresponding perturbation with the property described. Define the perturbation $(\tilde{x}, \tilde{y})$ as follows:

$$\tilde{x}_s = \alpha_1 x_{e_1} + \alpha_2 x_{e_2} + \dots + \alpha_k x_{e_k} \quad \text{such that } \alpha_i > 0, \sum_{i=1}^k \alpha_i = 1.$$

$$\tilde{y}_s = \beta_1 y_{e_1} + \beta_2 y_{e_2} + \dots + \beta_k y_{e_k} \quad \text{such that } \beta_i > 0, \sum_{i=1}^k \beta_i = 1.$$

Since $(\tilde{x}, \tilde{y})$ is a convex combination of the perturbations it follows that $\delta(\hat{C} \mid s, \tilde{x}_s, \tilde{y}_s) = 1$, for all $s \in \hat{C}$. Observe that $\hat{C}$ is a closed recurrent class under the strategy $(\tilde{x}, \tilde{y})$. Hence the desired result follows. ■

Lemma 11 *The following assertions hold.*

1. *Let $s \in S$, and a memoryless strategy y for player 2 be given. There exists $a \in \Gamma_1(s)$ such that*

$$\mathbb{E}[v_1(\Theta_1) \mid s, a, y_s] \geq v_1(s)$$

2. *Let $\hat{C} \subseteq S$, and a memoryless strategy y for player 2 be given.*

$$H_1(y, \hat{C}) \geq \max_{s \in \hat{C}} v_1(s)$$

Proof.

1. It follows from the results of [7] that the zero-sum values $v_1(\cdot)$ are characterized by fixed points of a matrix game. The result then follows from the fact that in any matrix game, given a distribution y for player 2 there is an optimal move a that maximizes the expected one-step payoff for player 1.
2. Follows from part 1 and definition of $H_1(y, \hat{C})$. ■

Definition 15 (Exit distributions [22, 23]) *Given $\hat{C} \subseteq C$, an exit distribution from $\hat{C}$ is a distribution $q \in \mathcal{D}(S)$ such that $q(\hat{C}) < 1$, i.e., $\sum_{s \in \hat{C}} q(s) < 1$. Let (x, y) be a pair of memoryless strategies and $\hat{C} \subseteq C$ be given. We define unilateral and joint exit as follows:*

1. *Player 1 unilateral exit:*

$$Q_1^{\hat{C}}(x, y) = \{ \delta(\cdot \mid s, a, y_s) \text{ where } s \in \hat{C}, a \in \Gamma_1(s), \delta(\hat{C} \mid s, a, y_s) < 1 \}$$

i.e., player 1 force the play out of $\hat{C}$ with positive probability playing move a against the memoryless strategy y .

2. *Player 2 unilateral exit:*

$$Q_2^{\widehat{C}}(x, y) = \{ \delta(\cdot \mid s, x_s, b) \text{ where } s \in \widehat{C}, b \in \Gamma_2(s), \delta(\widehat{C} \mid s, x_s, b) < 1 \}$$

i.e., player 2 force the play out of $\widehat{C}$ with positive probability playing move b against the memoryless strategy x .

3. *Joint exit of the players:*

$$Q_3^{\widehat{C}}(x, y) = \{ \delta(\cdot \mid s, a, b) \text{ where } s \in \widehat{C}, a \in \Gamma_1(s), b \in \Gamma_2(s), \\ \delta(\widehat{C} \mid s, a, y_s) = \delta(\widehat{C} \mid s, x_s, b) = 1, \text{ and } \delta(\widehat{C} \mid s, a, b) < 1 \}$$

i.e., playing a against y , and b against x keeps the play in $\widehat{C}$ with probability 1, but playing a and b jointly the players can ensure the play to leave $\widehat{C}$ with positive probability.

Let $Q^{\widehat{C}}(x, y) = \text{convex-hull}(Q_1^{\widehat{C}}(x, y) \cup Q_2^{\widehat{C}}(x, y) \cup Q_3^{\widehat{C}}(x, y))$ denote the convex combination of the distributions of unilateral and joint exit distribution. For all distribution $Q \in Q^{\widehat{C}}(x, y)$, the distribution can be represented as

$$Q = \sum_{l_1 \in L_1} \eta_{l_1} P_{l_1} + \sum_{l_2 \in L_2} \eta_{l_2} P_{l_2} + \sum_{l_3 \in L_3} \eta_{l_3} P_{l_3}$$

where $P_{l_j} \in Q_j^{\widehat{C}}(x, y)$ for $l_j \in L_j$. ■

For a distribution $Q \in \mathcal{D}(S)$ and a payoff γ for all states, by $E_Q[\gamma(\Theta_1)]$ we denote $\sum_{s \in S} \gamma(s) \cdot Q(s)$.

Definition 16 (Controllable exit distributions and controllable sets [22, 23])

A distribution $Q \in \mathcal{D}(S)$ is a controllable exit distribution from $\widehat{C}$ w.r.t. to a payoff vector $\gamma = (\gamma_1, \gamma_2)$, if for every $\varepsilon > 0$, there exists strategy pair $(\sigma_\varepsilon, \pi_\varepsilon)$ and two bounded stopping times τ_1 and τ_2 such that for all $s \in \widehat{C}$ the following conditions hold:

1. $\Pr_s^{\sigma_\varepsilon, \pi_\varepsilon}(e_{\widehat{C}} < \infty) = 1$, *i.e., the play leaves $\widehat{C}$ with probability 1.*
2. $\Pr_s^{\sigma_\varepsilon, \pi_\varepsilon}(\Theta_{e_{\widehat{C}}} = s') = Q(s')$, *i.e., the exit distribution from $\widehat{C}$ is the distribution Q .*
3. $\Pr_s^{\sigma_\varepsilon, \pi_\varepsilon}(\min\{\tau_1, \tau_2\} \leq e_{\widehat{C}}) \leq \varepsilon$, *i.e., the stopping times are smaller than the exit time with small probability ε .*

4. For all strategy σ ,

$$\mathbb{E}^{\sigma, \pi_\varepsilon}[\gamma_1(\Theta_{e_{\hat{C}}})\mathbf{1}_{(e_{\hat{C}} < \tau_1)}] + \mathbb{E}^{\sigma, \pi_\varepsilon}[v_1(\Theta_{\tau_1})\mathbf{1}_{(e_{\hat{C}} \geq \tau_1)}] \leq \mathbb{E}_Q[\gamma_1(\Theta_1)] + \varepsilon,$$

where $\mathbf{1}_{(e_{\hat{C}} < \tau_1)}$ is the indicator function of the event $\{e_{\hat{C}} < \tau_1\}$, and $\mathbf{1}_{(e_{\hat{C}} \geq \tau_1)}$ is the indicator function of the event $\{e_{\hat{C}} \geq \tau_1\}$. Intuitively, if play leaves $\hat{C}$ within time τ_1 then the payoff is defined by the exit distribution $\Theta_{e_{\hat{C}}}$ and payoff γ_1 , and if the game stays in $\hat{C}$ for more than τ_1 steps, then the payoff is defined by the distribution at time τ_1 and the payoff v_1 . Then the expected payoff is at most $\mathbb{E}_Q[\gamma_1(\Theta_1)] + \varepsilon$. Similarly, for all strategy π ,

$$\mathbb{E}^{\sigma_\varepsilon, \pi}[\gamma_2(\Theta_{e_{\hat{C}}})\mathbf{1}_{(e_{\hat{C}} < \tau_2)}] + \mathbb{E}^{\sigma_\varepsilon, \pi}[v_2(\Theta_{\tau_2})\mathbf{1}_{(e_{\hat{C}} \geq \tau_2)}] \leq \mathbb{E}_Q[\gamma_2(\Theta_1)] + \varepsilon,$$

where $\mathbf{1}_{(e_{\hat{C}} < \tau_2)}$ is the indicator function of the event $\{e_{\hat{C}} < \tau_2\}$, and $\mathbf{1}_{(e_{\hat{C}} \geq \tau_2)}$ is the indicator function of the event $\{e_{\hat{C}} \geq \tau_2\}$.

A controlled set $(\hat{C}, Q)$ is a set $\hat{C} \subseteq S$, and Q is a controllable exit distribution for any payoff vector $\gamma \geq v = (v_1, v_2)$. ■

A notion that complements the notion of controlled set is a *blocking pair*. We will establish the relation between a blocking pair and a controlled set in Lemma 13 and Lemma 15.

Definition 17 (Blocking pairs) Let $D \subseteq S$, and y be a memoryless strategy for player 2. The pair (y, D) is a blocking pair for player 1 (i.e., player 2 blocks) if for all $s \in D$, and for all $a \in \Gamma_1(s)$ we have

$$\delta(D \mid s, a, y_s) < 1 \Rightarrow \mathbb{E}[v_1(\Theta_1) \mid s, a, y_s] < \max_{s \in D} v_1(s)$$

Informally, by playing the strategy y player 2 ensures that if player 1 leaves the set D , then the expected payoff for player 1, assuming all state s has reward $v_1(s)$, is less than the maximum value $v_1(\cdot)$ of player 1 in D . Blocking pair (x, D) for player 2 (i.e., player 1 blocks) is defined by exchanging the roles of the players. ■

Reduced game. Let $\hat{C}$ be any controlled set. Then the game $\mathcal{G}_{\hat{C}}$ is obtained from $\mathcal{G}$ by collapsing the set $\hat{C}$ to a single dummy state $\{\hat{C}\}$, and the transition function at $\{\hat{C}\}$ defined by the controllable exit distribution Q , i.e., at state $\{\hat{C}\}$ players have only a single move $*$ and the transition function at state $\{\hat{C}\}$ given the moves $(*, *)$ is given by the distribution Q . Hence the state space of $\mathcal{G}_{\hat{C}}$ is $(S \setminus \hat{C}) \cup \{\hat{C}\}$ and we denote $\delta_{\hat{C}}$ to denote the transition function of $\mathcal{G}_{\hat{C}}$. We refer to this process as *collapsing* of a controlled set.

Definition 18 (Reduced blocking pair) A reduced blocking pair is a blocking pair in a reduced game. A pair (y, D) is a reduced blocking pair in a game $\mathcal{G}_{\hat{C}}$ if for all state $s \in D$, for all $a \in \Gamma_1(s)$ we have

$$\delta_{\hat{C}}(D \mid s, a, y_s) < 1 \Rightarrow \mathbb{E}_{\delta_{\hat{C}}}[v_1(\Theta_1) \mid s, a, y_s] < \max_{s \in D} v_1(s)$$

Note that $v_1(s)$ is the value of the original game and not the reduced game. ■

The following proof is similar to Vieille's proof [29].

Lemma 12 ([29]) Let (x, D) be a blocking pair for player 2. Then there exists $\overline{D} \subseteq D$ such that

1. (C1) $v_2(\cdot)$ is constant in $\overline{D}$, i.e., for all $s, s' \in \overline{D}$ we have $v_2(s) = v_2(s')$.
2. (C2) $\overline{D}$ is weak-communicating w.r.t. $(x, \overline{\pi})$.
3. (C3) $(x, \overline{D})$ is a blocking pair for player 2.

Proof. Let $\tilde{D} = \{s \in D \mid v_2(s) = \max_{s' \in D} v_2(s')\}$ be the set of states where v_2 is maximum. Since (x, D) is a blocking pair for player 2 it follows immediately that $(x, \tilde{D})$ is also blocking pair for player 2. Consider the perturbed graph $\tilde{G}_{\tilde{D}}(x, \overline{\pi})$ and let $\overline{D}$ be a terminal strongly connected end-component in the graph. There is no edge out of $\overline{D}$, $\overline{D}$ is closed and weak-communicating. Since $\overline{D} \subseteq \tilde{D}$ it follows that $v_2(\cdot)$ is constant in $\overline{D}$. ■

The following Lemmas will be the basic principle of a reduction mechanism of the original game.

Lemma 13 ([29]) Let (x, D) be a blocking pair for player 2. Then there exists $\tilde{D} \subseteq D$ such that one of the following two conditions hold:

1. $\tilde{D}$ is a controlled set.
2. $(\overline{\sigma}, \tilde{D})$ and $(\overline{\pi}, \tilde{D})$ are blocking pairs, $(\overline{\sigma}, \tilde{D})$ and $(\overline{\pi}, \tilde{D})$ are weak-communicating and v_1 and v_2 are constant in $\tilde{D}$.

Proof. Given (x, D) is a blocking pair consider $(x, \overline{D})$ that satisfies condition (C1 – C3) of Lemma 12.

1. If $\delta(\overline{D} \mid s, a, \overline{\pi}_s) < 1$ and $E[v_1(\Theta_1) \mid s, a, \overline{\pi}_s] \geq \max_{s' \in \overline{D}} v_1(s')$, for some state s and $a \in \Gamma_1(s)$, then we verify that $\overline{D}$ is a controlled set. Chose s^*, a^* such that it maximizes $E[v_1(\Theta_1) \mid s, a, \overline{\pi}_s]$. Then by construction we have

$$E[v_1(\Theta_1) \mid s^*, a^*, \overline{\pi}_s] \geq H_1(\overline{\pi}, \overline{D}).$$

For player 2 notice that

$$E[v_2(\Theta_1) \mid s^*, a^*, \overline{\pi}_s] \geq v_2(s^*) = \max_{s \in \overline{D}} v_2(s) \quad (\text{since } v_2 \text{ constant in } \overline{D}).$$

Hence it follows, that playing x with perturbation to a^* is a controllable exit distribution.

2. Otherwise, note that $(\overline{\pi}, \overline{D})$ is blocking pair for player 1. Apply Lemma 12 to $(\overline{\pi}, \overline{D})$ with roles of the players exchanged and let $\tilde{D} \subseteq \overline{D}$ be the corresponding subset. Then as above,
 - either there exists exit distribution using unilateral exits of player 2; or
 - $(\overline{\sigma}, \tilde{D})$ is a blocking pair for player 2. Moreover, $(\overline{\sigma}, \tilde{D})$ and $(\overline{\pi}, \tilde{D})$ satisfy all the desired conditions. ■

We present a sketch of the following Lemma as described in [22, 23]. For details see [22, 23].

Lemma 14 ([22, 23]) *Let $\hat{C} \subseteq C$ and let $(x, \hat{C})$ and $(y, \hat{C})$ be blocking pairs and $Q \in \text{convex-hull}(Q_{\hat{C}}^3(x, y))$ be an exit distribution. Let $\hat{C}$ be weak-communicating under (x, y) . Assume that the following conditions hold:*

1. *Let γ be a payoff-vector such that $\gamma_i(s) \geq v_i(s)$ for all state $s \in \hat{C}$, for $i = 1, 2$.*
2. *For all $s \in \hat{C}$, for any $a \in \Gamma_1(s)$ we have $E[v_1(\Theta_1) \mid s, a, y_s] \leq E_Q[\gamma_1(\Theta_1)]$.*
3. *For all $s \in \hat{C}$, for any $b \in \Gamma_2(s)$ we have $E[v_2(\Theta_1) \mid s, x_s, b] \leq E_Q[\gamma_2(\Theta_1)]$.*

Then Q is a controllable exit distribution.

Proof. (Sketch). Fix $\beta > 0$ and $\varepsilon > 0$ to be sufficiently small. By definition of weak-communication we have that for all $s \in \hat{C}$, exists $(\tilde{x}, \tilde{y})$ such that

- $\|(\tilde{x}, \tilde{y}) - (x, y)\| \leq \alpha$.

- if both players play $(\tilde{x}, \tilde{y})$ then the game leaves $\hat{C}$ with probability 0, and s is reached with probability 1 in finite time for any state $s' \in \hat{C}$.

The strategy (σ, π) is defined as follows. In a cyclic manner do the following for exit distributions P_{l_3} for $l_3 \in L_3$.

- Step 1. Denote by z the state where the joint exit distribution occurs. Play $(\tilde{x}, \tilde{y})$ till the game reaches z .
- Step 2. Let $\alpha = \beta \cdot \eta_{l_3}$. At z play the following strategy

$$((1 - \sqrt{\alpha})\tilde{x} + \sqrt{\alpha}a, (1 - \sqrt{\alpha})\tilde{y} + \sqrt{\alpha}b)$$

- Step 3. Continue cyclically.

Stopping times τ_1 and τ_2 .

- If player 1 (resp. player 2) plays an move that is not compatible with σ (resp. π) then τ_1 (resp. τ_2) is stopped.
- For every $l_3 \in L_3$, consider all stages that the play has been in Step 2 and check if the opponent has perturbed to a (or to b) approximately in the specified frequency, i.e., the ratio $\sqrt{\alpha}$ and the number of times the move played by player 1 (resp. player 2) was a (resp. b) is in $(1 - \varepsilon, 1 + \varepsilon)$, for small ε .

The statistical test is done only if the number of rounds the players is in Step 2 is sufficiently large, so that the probability of false detection of deviation is small.

If α and ε are sufficiently small, the test can be employed effectively, since exiting $\hat{C}$ occurs after $O(\frac{1}{\alpha})$ stages whereas the players should perturb with probability $\sqrt{\alpha}$. Hence until the exit occurs, each player should perturb $O(\frac{1}{\sqrt{\alpha}})$ times, which is enough for the statistical test. Once the statistical test fails, or the stopping time is reached the players play the spoiling strategies of the zero-sum games, ensuring that the other player's payoff is no more than her value in the zero-sum game.

Since the players switches to their spoiling strategies ultimately no player has an unilateral incentive to stay in $\hat{C}$. Thus there is no profitable deviation for the players and hence Q is a controllable exit distribution. ■

The following Lemma is an proof of Vieille using Solan's result.

Lemma 15 ([29]) *Let $(\bar{\sigma}, \tilde{D})$ and $(\bar{\pi}, \tilde{D})$ be blocking pairs such that $\bar{D}$ is weak-communicating under $(\bar{\sigma}, \bar{\pi})$. Then there is a controllable exit distribution Q such that $Q \in \text{convex-hull}(Q_D^3(\bar{\sigma}, \bar{\pi}))$ and $(\tilde{D}, Q)$ is a controlled set.*

Proof. Consider the perennial ε -optimal strategies σ_ε and π_ε such that $\Pr_s^{\sigma_\varepsilon, \pi_\varepsilon}(\text{Reach}(U)) = 1$ (recall Lemma 8). Hence we have $\Pr_s^{\sigma_\varepsilon, \pi_\varepsilon}(e_{\tilde{D}} < \infty) = 1$. Let us denote by

$$Q_\varepsilon = \Pr_s^{\sigma_\varepsilon, \pi_\varepsilon}(\Theta_{e_{\tilde{D}}} = \cdot)$$

the law of exit distribution from $\tilde{D}$ under strategy σ_ε and π_ε . Since $e_{\tilde{D}}$ is finite almost-surely (with probability 1) and the strategies σ_ε and π_ε are perennial ε -optimal strategies we have

$$E_{Q_\varepsilon}[v_1(\Theta_{e_{\tilde{D}}})] \geq v_1(s) - \varepsilon; \quad E_{Q_\varepsilon}[v_2(\Theta_{e_{\tilde{D}}})] \geq v_2(s) - \varepsilon.$$

Since $(\bar{\sigma}, \tilde{D})$ and $(\bar{\pi}, \tilde{D})$ are blocking pairs it follows that for any history $\omega_n = \langle s_0, s_1, \dots, s_n \rangle$, we have $\delta(\tilde{D} \mid s_n, \sigma_\varepsilon(\omega_n), \bar{\pi}_{s_n}) = 1$ and $\delta(\tilde{D} \mid s_n, \bar{\sigma}_{s_n}, \pi_\varepsilon(\omega_n)) = 1$. Hence $Q_\varepsilon \in \text{convex-hull}(Q_D^3)$. It follows from Solan's result (Lemma 14) that there is a exit distribution $Q \in \text{convex-hull}(Q_D^3)$ such that

$$E_Q[v_1(\Theta_{e_{\tilde{D}}})] \geq v_1(s); \quad E_Q[v_2(\Theta_{e_{\tilde{D}}})] \geq v_2(s).$$

Since, Q involves no unilateral exit it follows from Solan's result (Lemma 14) that $(\tilde{D}, Q)$ is controllable for all $\gamma \geq v = (v_1, v_2)$. ■

Reduction sequence. It follows from Lemma 13 and Lemma 15 that if there is a blocking pair in a game there is a controlled set. The analysis of Vieille (Lemma 41-45 of [29]) presents an mechanism to collapse the controlled sets of the game to get a game $\mathcal{G}^*$ such that there is no reduced blocking pair in the game $\mathcal{G}^*$. The key idea is as follows: (a) let the original game be $\mathcal{G}_0 = \mathcal{G}_R$; if $\mathcal{G}_0$ has no blocking pair then $\mathcal{G}^* = \mathcal{G}_0$, (b) else there is a sequence of controlled sets $\hat{C}_0, \hat{C}_1, \dots, \hat{C}_k$, such that each $\hat{C}_i$ is a maximal controlled set, and the game can be reduced by collapsing the controlled sets in sequence, i.e., $\mathcal{G}_{i+1}$ is obtained from $\mathcal{G}_i$ by collapsing the controlled set $\hat{C}_i$. Since $\mathcal{G}_{i+1}$ has fewer moves than $\mathcal{G}_i$ and the move set is finite the process stop after finite number of steps (say k steps). The analysis of Vieille shows that the game $\mathcal{G}^* = \mathcal{G}_{k+1}$ has no reduced blocking pair.

Lemma 16 *Let $\mathcal{G}^*$ be the game with no reduced blocking pair and the transition function of the game be δ^* . The following assertions hold in the game $\mathcal{G}^*$:*

1. For every memoryless strategy x for player 1 there is a memoryless strategy y for player 2 such that $\Pr_s^{x,y}(\text{Reach}(U)) = 1$, and $\Pr_s^{x,y}(\text{Reach}(W_2)) \geq v_2(s)$.
2. For every memoryless strategy y for player 2 there is a memoryless strategy x for player 1 such that $\Pr_s^{x,y}(\text{Reach}(U)) = 1$, and $\Pr_s^{x,y}(\text{Reach}(W_1)) \geq v_1(s)$.

Proof. We prove the result for case 1, and argument for case 2 is symmetric. Let a memoryless strategy x for player 1 be given. For $s \in S$ define

$$B(s) = \{ b \in \Gamma_2^*(s) \mid E_{\delta^*}[v_2(\Theta_1) \mid s, x_s, b] \geq v_2(s) \}.$$

Note that for any controlled set $\hat{C}$ that is reduced we have $E_{\delta^*}[v_2(\Theta_1) \mid s, *, *] = E_{Q_{\hat{C}}}[v_2(\Theta_{e_{\hat{C}}})] \geq v_2(s)$, where $Q_{\hat{C}}$ is the controllable exit distribution from $\hat{C}$. Hence it follows that $B(s) \neq \emptyset$ for all states s . Choose a y such that $\text{Supp}(y_s) = B(s)$. Consider the Markov chain under the memoryless strategy pair (x, y) . Consider any arbitrary $F \subseteq C$. Let $\overline{F} = \{ s \in F \mid v_2(s) = \max_{s' \in F} v_2(s') \}$. Since there is no reduced blocking pair in the game $\mathcal{G}^*$, we must have for some state $s \in \overline{F}$ that $\delta^*(\overline{F} \mid s, x_s, b) < 1$ and $E_{\delta^*}[v_2(\Theta_1) \mid s, x_s, b] \geq \max_{s \in \overline{F}} v_2(s)$. Since $\overline{F}$ consists of the set of states of F with maximum value for player 2, we have that if $\delta^*(\overline{F} \mid s, x_s, b) < 1$ and $E_{\delta^*}[v_2(\Theta_1) \mid s, x_s, b] \geq \max_{s \in \overline{F}} v_2(s)$, then $\delta^*(F \mid s, x_s, b) < 1$. Hence no subset $F \subseteq C$ is closed. Since under (x, y) we have a Markov chain it follows that $\Pr_s^{x,y}(\text{Reach}(U)) = 1$.

Finally observe that $\Pr_s^{x,y}(\text{Reach}(W_2)) \geq v_2(s)$, since $(v_2(\Theta_n))_n$ is a sub-martingale under (x, y) . ■

Recall that $W_1 = \{ t_{00}, t_{01} \}$ and $W_2 = \{ t_{00}, t_{10} \}$ and the game $\mathcal{G}_R$ is the nonzero-sum reachability game with objective $\text{Reach}(W_1)$ for player 1 and $\text{Reach}(W_2)$ for player 2.

Lemma 17 *The following assertions hold:*

1. For every $\varepsilon > 0$, there is an ε -Nash equilibrium (σ_k^*, π_k^*) in the nonzero-sum reachability game $\mathcal{G}^*$ such that $\Pr_s^{\sigma_k^*, \pi_k^*}(\text{Reach}(U)) = 1$; and $\Pr_s^{\sigma_k^*, \pi_k^*}(\text{Reach}(W_1)) \geq v_1(s) - \varepsilon$ and $\Pr_s^{\sigma_k^*, \pi_k^*}(\text{Reach}(W_2)) \geq v_2(s) - \varepsilon$.
2. For every $\varepsilon > 0$, there is an ε -Nash equilibrium (σ^*, π^*) in the nonzero-sum reachability game $\mathcal{G}_R$ such that $\Pr_s^{\sigma^*, \pi^*}(\text{Reach}(U)) = 1$; and $\Pr_s^{\sigma^*, \pi^*}(\text{Reach}(W_1)) \geq v_1(s) - \varepsilon$ and $\Pr_s^{\sigma^*, \pi^*}(\text{Reach}(W_2)) \geq v_2(s) - \varepsilon$.

Proof.

1. For $\varepsilon > 0$, let (σ_k^*, π_k^*) be a memoryless ε -Nash equilibrium in the game $\mathcal{G}^*$. The existence of memoryless ε -Nash equilibrium in two-player nonzero-sum games with reachability objectives follows from [5]. The result then follows from Lemma 16.
2. It follows from the definition of controlled sets that if $(\sigma_{i+1}^*, \pi_{i+1}^*)$ is an ε -Nash equilibrium, with $\varepsilon \rightarrow 0$, in the game $\mathcal{G}_{i+1}$, satisfying the assumptions of part 1, then the following strategy profile (σ_i^*, π_i^*) is an ε -Nash equilibrium in $\mathcal{G}_i$
 - (a) Let $\hat{C}_i$ be the controlled set collapsed in game $\mathcal{G}_i$ to obtain game $\mathcal{G}_{i+1}$. Then the strategy profile are as follows: play $(\sigma_{i+1}^*, \pi_{i+1}^*)$ for all states in $S_i \setminus \hat{C}_i$ and play the controllable exit distribution (σ, π) at every state in $\hat{C}_i$. Then the strategy (σ_i^*, π_i^*) is an ε -Nash equilibrium satisfying the required assumptions. By induction the result follows for $\mathcal{G} = \mathcal{G}_0$.

The desired result follows. ■

Lemma 18 *For every $\varepsilon > 0$, there is an ε -Nash equilibrium (σ^*, π^*) in the nonzero-sum reachability game $\mathcal{G}_R$, and there exists $k \in \mathbb{N}$ such that*

1. $\Pr_s^{\sigma^*, \pi^*}(\text{Reach}^k(U)) \geq 1 - \varepsilon$;
2. *for every history $\omega \in \text{Outcome}(s, \sigma^*, \pi^*)$, if $\omega_k = s_k$, then*
 - (a) $\Pr_{s_k}^{\sigma^*, \pi^*}(\text{Reach}(W_1)) \geq v_1(s_k) - \varepsilon$;
 - (b) $\Pr_{s_k}^{\sigma^*, \pi^*}(\text{Reach}(W_2)) \geq v_2(s_k) - \varepsilon$.

Proof. Let us denote the ε -Nash equilibrium profile satisfying the conditions of Lemma 17 as $(\tilde{\sigma}, \tilde{\pi})$, i.e., $\Pr_s^{\tilde{\sigma}, \tilde{\pi}}(\text{Reach}(U)) = 1$. Hence for $\varepsilon > 0$, there exists k such that $\Pr_s^{\tilde{\sigma}, \tilde{\pi}}(\text{Reach}^k(U)) = 1 - \varepsilon$. The strategy (σ^*, π^*) is defined as follows:

- $(\sigma^*, \pi^*) = (\tilde{\sigma}^k + \tilde{\sigma}, \tilde{\pi}^k + \tilde{\pi})$, i.e., the players play $(\tilde{\sigma}, \tilde{\pi})$ for k steps and then again switches to $(\tilde{\sigma}, \tilde{\pi})$. Formally, for a history $\omega = \langle s_0, s_1, \dots \rangle$ we have

$$\sigma^*(s_0 s_1 \dots s_n) = \begin{cases} \tilde{\sigma}(s_0 s_1 \dots s_n) & \text{if } n < k \\ \tilde{\sigma}(s_k \dots s_n) & \text{if } n \geq k \end{cases}$$

$$\pi^*(s_0 s_1 \dots s_n) = \begin{cases} \tilde{\pi}(s_0 s_1 \dots s_n) & \text{if } n < k \\ \tilde{\pi}(s_k \dots s_n) & \text{if } n \geq k \end{cases}$$

Since for all state s the following conditions hold:

1. $\Pr_s^{\tilde{\sigma}, \tilde{\pi}}(\text{Reach}(U)) = 1$;
2. $\Pr_s^{\tilde{\sigma}, \tilde{\pi}}(\text{Reach}(W_1)) \geq v_1(s) - \varepsilon$; and
3. $\Pr_s^{\tilde{\sigma}, \tilde{\pi}}(\text{Reach}(W_2)) \geq v_2(s) - \varepsilon$;

it follows that (σ^*, π^*) is an ε -Nash equilibrium with the desired property. ■

Theorem 1 (ε -Nash equilibrium in SSGC game) *Let $\mathcal{G}$ be a SSGC game with parity objective Ψ_1 for player 1 and Ψ_2 for player 2. For every $\varepsilon > 0$, there is an ε -Nash equilibrium for every state $s \in C$.*

Proof. The case when properties P1-P4 of Lemma 1 hold the result follows from Lemma 1. The case when properties P1-P4 are not satisfied we consider the nonzero-sum reachability game $\mathcal{G}_R$. We obtain a ε -Nash equilibrium in the original game considering the ε -Nash equilibrium of the reachability game $\mathcal{G}_R$ and then using spoiling strategies.

Fix arbitrary $\varepsilon > 0$, and we show that there is an 3ε -Nash equilibrium for every state $s \in C$. Since ε is arbitrary the result follows. Let (σ^*, π^*) be an ε -Nash equilibrium of the reachability game $\mathcal{G}_R$ as constructed in Lemma 18. Consider the strategy σ_ε^* for player 1 defined as follows:

$$\sigma_\varepsilon^*(s_0, s_1, \dots, s_l) = \begin{cases} \sigma^*(s_0, s_1, \dots, s_l) & \text{if } l < k \text{ (} k \text{ of Lemma 18)} \\ \bar{\sigma}_\varepsilon(s_0, s_1, \dots, s_l) & \text{if } l \geq k \text{ (} k \text{ of Lemma 18 and } \bar{\sigma}_\varepsilon \in \overline{\Sigma_\varepsilon}) \end{cases}$$

i.e., player 1 plays σ^* for k steps and then switches to an ε -spoiling strategy $\bar{\sigma}_\varepsilon$. Similarly, the strategy for player 2 is defined as follows:

$$\pi_\varepsilon^*(s_0, s_1, \dots, s_l) = \begin{cases} \pi^*(s_0, s_1, \dots, s_l) & \text{if } l < k \text{ (} k \text{ of Lemma 18)} \\ \bar{\pi}_\varepsilon(s_0, s_1, \dots, s_l) & \text{if } l \geq k \text{ (} k \text{ of Lemma 18 and } \bar{\pi}_\varepsilon \in \overline{\Pi_\varepsilon}) \end{cases}$$

Since $\Pr_s^{\sigma^*, \pi^*}(\text{Reach}^k(U)) \geq 1 - \varepsilon$, we have that

$$\Pr_s^{\sigma_\varepsilon^*, \pi_\varepsilon^*}(\Psi_{1s}) \geq \Pr_s^{\sigma^*, \pi^*}(\text{Reach}(W_1)) - \varepsilon$$

and

$$\Pr_s^{\sigma_\varepsilon^*, \pi_\varepsilon^*}(\Psi_{2s}) \geq \Pr_s^{\sigma^*, \pi^*}(\text{Reach}(W_2)) - \varepsilon.$$

Recall that (σ^*, π^*) is an ε -Nash equilibrium of the reachability game such that for every history $\omega \in \text{Outcome}(s, \sigma^*, \pi^*)$ we have if $\omega_k = s_k$ then

$\Pr_{s_k}^{\sigma^*, \pi^*}(\text{Reach}(W_1)) \geq v_1(s_k) - \varepsilon$ and $\Pr_{s_k}^{\sigma^*, \pi^*}(\text{Reach}(W_2)) \geq v_2(s_k) - \varepsilon$. Since the players play an ε -spoiling strategy after k -steps it follows that

$$\forall \sigma \in \Sigma. \Pr_s^{\sigma, \pi_\varepsilon^*}(\Psi_{1s}) \leq \Pr_s^{\sigma^*, \pi^*}(\text{Reach}(W_1)) + 2\varepsilon \leq \Pr_s^{\sigma_\varepsilon^*, \pi_\varepsilon^*}(\Psi_{1s}) + 3\varepsilon$$

and

$$\forall \pi \in \Pi. \Pr_s^{\sigma_\varepsilon^*, \pi}(\Psi_{1s}) \leq \Pr_s^{\sigma_\varepsilon^*, \pi^*}(\text{Reach}(W_2)) + 2\varepsilon \leq \Pr_s^{\sigma_\varepsilon^*, \pi_\varepsilon^*}(\Psi_{2s}) + 3\varepsilon.$$

Hence it follows that $(\sigma_\varepsilon^*, \pi_\varepsilon^*)$ is an 3ε -Nash equilibrium. ■

4 Existence of ε -Nash equilibrium

In this section we show that for all nonzero-sum concurrent game $\mathcal{G}$, with ω -regular objectives specified as parity objectives Ψ_1 and Ψ_2 for player 1 and player 2, respectively, for every $\varepsilon > 0$, there exists an ε -Nash equilibrium for every state s of game $\mathcal{G}$. The proof follows from an inductive argument: by induction on the size of the state space of the $\mathcal{G}$ and by application of Theorem 1. We assume without loss of generality that there are four special states $\{t_{00}, t_{01}, t_{10}, t_{11}\}$ in $\mathcal{G}$, as defined in Definition 10.

Lemma 19 *Let $\mathcal{G}$ be a concurrent game with parity objectives Ψ_1 and Ψ_2 for player 1 and player 2, respectively. Let $G_{\mathcal{G}}$ be the graph of $\mathcal{G}$ and TC be a terminal strongly connected component in $G_{\mathcal{G}}$. Then for every $\varepsilon > 0$, there is an ε -Nash equilibrium for every state $s \in \text{TC}$.*

Proof. The proof is by induction on the size of TC. It is easy to argue when $|\text{TC}| = 1$, i.e., TC consists of an absorbing state. Consider the sub-game induced by the set of states TC and call the sub-game $\mathcal{G}_{\text{TC}}$.

- Suppose there is a state $s \in \text{TC}$ such that $\langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s) = 1$. Then fix an ε -optimal strategy σ for player 1 and let π be an ε -optimal strategy for player 2 against σ . Then (σ, π) is an ε -Nash equilibrium. We can replace s by the gadget described in Proposition 2. This will break TC into (possibly many) smaller strongly connected components. By induction hypothesis, Theorem 1 and the bottom-up evaluation procedure described in Lemma 1 it follows that ε -Nash equilibrium exists at every state in TC. Similar arguments hold if there is a state $s \in \text{TC}$ such that $\langle\langle 2 \rangle\rangle_{\text{val}}(\Psi_2)(s) = 1$.

- Suppose for every state $s \in \text{TC}$ we have $\langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s) < 1$ and $\langle\langle 2 \rangle\rangle_{\text{val}}(\Psi_2)(s) < 1$. It follows from Corollary 1 of [6] that in a zero-sum concurrent game with ω -regular objectives if for every state s we have $\langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s) < 1$, then for every state s in the game we have $\langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s) = 0$, i.e., if the zero-sum value is positive for player 1 at some state, then there exists a state s where the zero-sum value is 1. Hence it follows from the above condition that for all state $s \in \text{TC}$ we have $\langle\langle 1 \rangle\rangle_{\text{val}}(\Psi_1)(s) = 0$ and $\langle\langle 2 \rangle\rangle_{\text{val}}(\Psi_2)(s) = 0$. Let $\bar{\pi}_\varepsilon$ be an ε -spoiling strategy for player 2 and $\bar{\sigma}_\varepsilon$ be an ε -spoiling strategy for player 1. Hence we have the following inequalities:

$$\forall \sigma \in \Sigma. \text{Pr}_s^{\sigma, \bar{\pi}_\varepsilon}(\Psi_{1s}) \leq \varepsilon \text{ and } \forall \pi \in \Pi. \text{Pr}_s^{\bar{\sigma}_\varepsilon, \pi}(\Psi_{2s}) \leq \varepsilon.$$

Hence we have $(\bar{\sigma}_\varepsilon, \bar{\pi}_\varepsilon)$ is an ε -Nash equilibrium for all state $s \in \text{TC}$. ■

Theorem 2 (ε -Nash equilibrium) *Let $\mathcal{G}$ be a concurrent game with parity objectives Ψ_1 and Ψ_2 for player 1 and player 2, respectively. For every $\varepsilon > 0$, there is an ε -Nash equilibrium for every state $s \in S$.*

Proof. Let $G_{\mathcal{G}}$ be the graph of $\mathcal{G}$. It follows from Lemma 19 that for any state s in a terminal strongly connected component of $G_{\mathcal{G}}$ there is an ε -Nash equilibrium. By Proposition 2 we can replace every state s of a terminal strongly connected component by the gadget described in Proposition 2. For the rest of the strongly connected components we proceed in a bottom-up order as follows: consider a strongly connected component C when all the strongly connected component below it are replaced by the gadgets of Proposition 2. The sub-game induced by C and the gadgets of the strongly connected components below C form a SSCC game. By Theorem 1 we have there is an ε -Nash equilibrium for every state $s \in C$. ■

5 Computational Complexity

In this section we show how to compute the values of some ε -Nash equilibrium of SSCC games within ε -precision. We prove that every case of the existence proof of ε -Nash equilibrium is constructive and computable. It may be noted that even in the case of zero-sum concurrent games with parity objectives the values can be irrational (for an example see [7]). Hence, one can only achieve ε -approximation of the values in the general case of nonzero-sum concurrent parity games. It follows from the inductive argument of Theorem 2 that the values of ε -Nash equilibrium for concurrent

games can be computed by n -iterations of a procedure to compute ε -Nash equilibrium values for SSCC games.

Complexity of ε -Nash equilibrium in SSCC games. To analyze the complexity of computing values of some ε -Nash equilibrium in SSCC games we consider the following cases:

1. Case 1. Compute the values of ε -Nash equilibrium when the property P1 of Lemma 1 is satisfied.
2. Case 2. Compute the values of ε -Nash equilibrium when the property P4 of Lemma 1 is satisfied.
3. Case 3. Compute the values of ε -Nash equilibrium when the property P2 or P3 of Lemma 1 is satisfied.
4. Case 4. Compute the values of some special ε -Nash equilibrium of SSCC games with reachability objectives.

We analyze the above cases below.

1. Case 1. Given Ψ_1 and Ψ_2 are parity objectives, the objective $\Psi_1 \cap \Psi_2$ is a Streett objective [25]. To analyze the computation of $\sup_{(\sigma, \pi) \in \Sigma \times \Pi} \Pr_s^{\sigma, \pi}(\Psi_1 \cap \Psi_2)$, observe that this is equivalent to the computation of values of one-player games (MDPs) where player 1 and player 2 cooperates to achieve the objective $\Psi_1 \cap \Psi_2$. Hence the computation reduces to computing values in a MDP with Streett objective. This can be achieved in polynomial time [3].
2. Case 2. After the computation of the zero-sum values $v_1(\cdot)$ and $v_2(\cdot)$, it is easy to determine if there is a state s such that $v_1(s) = 0$ and $v_2(s) = 0$. Hence Case 2 can be solved by computing the zero-sum values for player 1 and player 2.
3. Case 3. Given the zero-sum values for player 1 and player 2 are computed, we describe a polynomial time procedure to determine the values of some ε -Nash equilibrium when property P2 or P3 of Lemma 1 is satisfied. We prove the result for the case when property P2 is satisfied and the result for the case when property P3 is satisfied is symmetric. Consider the set $W = \{ s \mid v_1(s) = 1 \}$ of states that have zero-sum value 1 for player 1. Since property P2 is satisfied, we have $W \cap C \neq \emptyset$. Given a state $s \in W$, consider the set $\text{SafeAct}(s) = \{ a \in \Gamma_1(s) \mid \forall b \in \Gamma_2(s). \text{Dest}(s, a, b) \subseteq W \}$ of moves

for player 1 that ensures that the set W is never left. Consider a reduced sub-game $\mathcal{G}'$ induced by W such that at every state $s \in W$ the available moves for player 1 is $\text{SafeAct}(s)$. Let Σ' be the set of strategies such that player 1 plays only moves in $\text{SafeAct}(s)$ for every state $s \in W$, i.e., the set of strategies in $\mathcal{G}'$. We compute the values $val(s) = \sup_{(\sigma, \pi) \in \Sigma' \times \Pi} \Pr_s^{\sigma, \pi}(\Psi_1 \cap \Psi_2)$.

Observe that there exists ε -optimal strategy σ_ε of the original game such that for every strategy $\pi \in \Pi$ we have $\Pr_s^{\sigma_\varepsilon, \pi}(\Psi_1 \cap \text{Safe}(W)) \geq 1 - \varepsilon$, for all state $s \in W$. Hence it follows that

$$\begin{aligned} \Pr_s^{\sigma_\varepsilon, \pi}(\Psi_2) &\leq \Pr_s^{\sigma_\varepsilon, \pi}((\Psi_1 \cap \text{Safe}(W)) \cap \Psi_2) + \varepsilon \\ &\leq \sup_{(\sigma, \pi) \in \Sigma' \times \Pi} \Pr_s^{\sigma, \pi}(\Psi_1 \cap \Psi_2) + \varepsilon. \end{aligned} \quad (1)$$

- If for some state $s \in W \cap C$ we have $val(s) = 1$, then property P1 of Lemma 1 is satisfied and then Case 1 is followed.
- Else for every state $s \in W \cap C$ we have $val(s) < 1$. It follows from property of MDPs that for any ω -regular objective Ψ , the maximum probability to satisfy Ψ is equal to the maximum probability of reaching the set of states where the value is 1. Hence we have

$$\sup_{(\sigma, \pi) \in \Sigma' \times \Pi} \Pr_s^{\sigma, \pi}(\Psi_1 \cap \Psi_2) = \sup_{(\sigma, \pi) \in \Sigma' \times \Pi} \Pr_s^{\sigma, \pi}(\text{Reach}(t_{00})) \quad (2)$$

We show that for every state $s \in W \cap C$, the profile $(1, val(s))$ is the value of some ε -Nash equilibrium profile. Let $(\hat{\sigma}, \hat{\pi})$ be a memoryless strategy such that $\Pr_s^{\hat{\sigma}, \hat{\pi}}(\text{Reach}(t_{00})) = val(s)$, for all $s \in W \cap C$. The existence of such a memoryless strategy follows from [4]. For any $\varepsilon > 0$, let $k \in \mathbb{N}$ be such that $\Pr_s^{\hat{\sigma}, \hat{\pi}}(\text{Reach}^k(t_{00})) \geq val(s) - \varepsilon$. The strategy profile (σ^*, π^*) is described as follows:

$$\sigma^*(s_0, s_1, \dots, s_l) = \begin{cases} \hat{\sigma}(s_0, s_1, \dots, s_k) & l < k \\ \sigma_\varepsilon(s_0, s_1, \dots, s_k) & l \geq k, \sigma_\varepsilon \in \Sigma_\varepsilon \end{cases}$$

and $\pi^* = \hat{\pi}$. Given strategy σ^* , for any strategy π the play never leaves W within k steps, since $\hat{\sigma} \in \Sigma'$. Since $\sigma_\varepsilon \in \Sigma_\varepsilon$ and for every state $s \in W$ we have $\langle\langle 1 \rangle\rangle_{val}(\Psi_1)(s) = 1$ it follows that $\Pr_s^{\sigma^*, \pi^*}(\Psi_1) \geq 1 - \varepsilon$. Since σ^* follows $\hat{\sigma}$ for k steps, it follows that $\Pr_s^{\sigma^*, \pi^*}(\Psi_2) \geq \Pr_s^{\sigma^*, \pi^*}(\text{Reach}(t_{00})) - \varepsilon$. It follows from equation 1

and 2 that $\sup_{\pi \in \Pi} \Pr_s^{\sigma^*, \pi}(\Psi_2) \leq \Pr_s^{\sigma^*, \pi^*}(\text{Reach}(t_{00})) + \varepsilon$. Hence it follows that $(1, \text{val}(s))$ is an ε -Nash equilibrium value profile for all state $s \in W \cap C$, for all $\varepsilon > 0$.

It follows from above that the values of some ε -Nash equilibrium of states $s \in C$ can be computed by a polynomial procedure and solving the zero-sum values for player 1 and player 2 when Case 1, Case 2 or Case 3 is satisfied. The analysis of Case 4 involves solving some special ε -Nash equilibrium values of a game with reachability objectives. We argue below the existence of polynomial witness and polynomial time verification procedure for Case 4.

4. Case 4. The polynomial witness and the polynomial time verification procedure for the witness consists of the analysis of the following two cases:
 - (a) Witness and verification procedure for the reduction sequence that is described after Lemma 15.
 - (b) Witness and verification procedure for ε -Nash equilibrium in the reachability game $\mathcal{G}^*$, when the reduction sequence terminates.

Observe that in the reduction sequence defined after Lemma 15, the length of the of the sequence is linear in the size of the game-graph. The fact follows since every reduction step decreases the number of moves by at least 1. We show that there are polynomial witness for every reduction step and thereby establish existence of polynomial witness for the entire reduction sequence. The polynomial witness for a reduction step consists of a controlled set. It follows from the results of [5, 12] that any memoryless strategy (or a memoryless distribution) can be suitably ε -approximated by k -uniform memoryless strategies, where a k -uniform memoryless strategy is a strategy that assigns probabilities to every move as multiples of $\frac{1}{\ell}$, where $\ell \leq k$. Moreover, k is polynomial in $|\mathcal{G}|$ and $\frac{1}{\varepsilon}$. We now analyze the following cases to provide the polynomial witness and verification procedure for controlled set.

- (a) If Case 1 of Lemma 13 holds then the witness of a controlled set consists of a set $\overline{D}$, memoryless strategy pair x and $\overline{\pi}$ such that
 - i. $(x, \overline{D})$ is a blocking pair for player 2;
 - ii. $v_2(\cdot)$ is constant in $\overline{D}$;
 - iii. $\overline{D}$ is weak-communicating under $(x, \overline{\pi})$.

Since x and $\bar{\pi}$ are memoryless strategies the witnesses can be ε -approximated by k -uniform memoryless strategies and the witnesses are polynomial. It is easy to verify that $(x, \bar{D})$ is a blocking pair by verifying that for every state $s \in \bar{D}$ we have if $\delta(\bar{D} \mid s, x_s, b) < 1$ then $\mathbb{E}[v_2(\Theta_1) \mid s, x_s, b] < \max_{s' \in \bar{D}} v_2(\cdot)$. Since the zero-sum values $v_2(\cdot)$ is computed the verification procedure is achieved in polynomial time. Again since the zero-sum values $v_2(\cdot)$ is computed it is easy to verify that $v_2(\cdot)$ is constant over $\bar{D}$. To conclude $\bar{D}$ is weak-communicating under $(x, \bar{\pi})$ it is sufficient to construct the perturbed graph $\tilde{G}_{\bar{D}}(x, \bar{\pi})$ and verify that $\bar{D}$ is strongly connected. The last step of the verification procedure checks that for some state $s \in \bar{D}$ and $a \in \Gamma_1(s)$ we have $\delta(\bar{D} \mid s, a, \bar{\pi}_s) < 1$ and $\mathbb{E}[v_1(\Theta_1) \mid s, a, \bar{\pi}_s] \geq \max_{s' \in \bar{D}} v_1(s')$. The procedure then chooses (s^*, a^*) that maximizes $\mathbb{E}[v_1(\Theta_1) \mid s, a, \bar{\pi}_s]$. The controlled exit distribution, as described in Lemma 13, then consists of playing the distribution x with perturbation to a^* . This establishes the existence of polynomial witness and polynomial time verification procedure for the case when part 1 of Lemma 13 holds. Similar arguments hold for the symmetric case when the condition holds for player 2. Otherwise, part 2 of Lemma 13 holds. We analyze the case below.

- (b) If part 2 of Lemma 13 holds then there are blocking pairs $(\bar{\sigma}, \tilde{D})$ and $(\bar{\pi}, \tilde{D})$ such that $\bar{D}$ is weak-communicating under $(\bar{\sigma}, \bar{\pi})$. Since $\bar{\sigma}$ and $\bar{\pi}$ are memoryless strategies arguments analogous to the previous case proves the existence of polynomial witness and polynomial time verification procedure for the above condition. It follows from Lemma 15 that if the above condition holds then there is a controlled exit distribution $Q \in \text{convex-hull}(Q_{\bar{D}}^3(\bar{\sigma}, \bar{\pi}))$. Since Q is memoryless distribution it can be ε -approximated by a k -uniform memoryless strategy such that k is polynomial in $|\mathcal{G}|$ and $\frac{1}{\varepsilon}$. To verify that Q is a controlled exit distribution the conditions of Lemma 14 needs to be verified. The verification resembles the analysis of a Markov chain under the distribution Q . Since Q can be approximated by a polynomial k -uniform memoryless strategy the verification is achieved in polynomial time.

It follows from above that at every reduction step the witness of a controlled set is polynomial and can be verified in polynomial time. We now consider the case when the reduction sequence terminates and ε -Nash equilibrium of the reachability game $\mathcal{G}^*$ needs to be computed

(recall Lemma 16). The existence of polynomial witness and polynomial time verification procedure to compute values of such ε -Nash equilibrium follows from [5].

Let $\mathbf{ZeroSum}(\mathcal{G}, \varepsilon)$ denote the time complexity of an algorithm to compute the zero-sum values of a concurrent parity game within ε -precision. Let $\mathbf{NonzeroSumReachability}(\mathcal{G}, \varepsilon)$ denote the complexity of an algorithm to compute the values of some ε -Nash equilibrium, greater than some specified value, of a concurrent game with reachability objectives. It follows from [2] and [5] that there exists $\mathbf{ZeroSum}(\mathcal{G}, \varepsilon)$ and $\mathbf{NonzeroSumReachability}(\mathcal{G}, \varepsilon)$ that are in the complexity class FNP, for constant $\varepsilon > 0$. The above analysis gives us the following Theorem on complexity of computing the values of some ε -Nash equilibrium in concurrent games with parity objectives.

Theorem 3 (Complexity of ε -Nash equilibrium) *Let $\mathcal{G}$ be a two-player concurrent game structure with n states. Then the following assertions hold:*

1. *The value of some ε -Nash equilibrium of a nonzero-sum concurrent game with parity objectives can be computed in time*

$$O(n(\mathbf{ZeroSum}(\mathcal{G}, \varepsilon) + \mathbf{NonzeroSumReachability}(\mathcal{G}, \varepsilon)) + O(p(|\mathcal{G}|)))$$

where p is a polynomial function.

2. *For every constant $\varepsilon > 0$, the values of some ε -Nash equilibrium of a nonzero-sum concurrent game with parity objectives can be ε -approximated in FNP; and hence in EXPTIME.*

6 Conclusion

In case of two-player concurrent games we extend the existence of ε -Nash equilibrium, for every $\varepsilon > 0$, from safety and reachability objectives to the class of ω -regular objectives. Our analysis also shows that computation of values of some ε -Nash equilibrium can be reduced to two simpler problems: (a) computing values of zero-sum games; and (b) computing values of ε -Nash equilibrium of nonzero-sum reachability games. The possible extension of the result can be made in two directions:

1. **More players.** The existence of ε -Nash equilibrium, for all $\varepsilon > 0$, in n -player games with ω -regular objectives remains an open problem.

The problem is likely to require more involved analysis. In case of n -player games with safety objectives the existence of Nash equilibrium proof critically relies on the existence of finite counterexamples for safety objectives. In case of n -player games with reachability objectives the existence of ε -Nash equilibrium, for all $\varepsilon > 0$, is achieved by analyzing discounted version of the original game. Unfortunately, both the above ideas fails for infinitary objectives like ω -regular objectives. In case of n -player games application of punishing strategies is complicated and to the best of the authors' knowledge, no general result is known for existence of ε -Nash equilibrium in case of n -player games that is achieved by applying punishing strategies.

2. **More objectives.** The existence of ε -Nash equilibrium, for all $\varepsilon > 0$, in case of two-player games with objectives in the higher levels of Borel hierarchy than ω -regular objectives remains another open problem.

Acknowledgment

I am grateful to Tom Henzinger for all his support and guidance, and several helpful discussions. I thank Luca de Alfaro for several key insights on concurrent ω -regular games that I have received from him. The techniques developed with Tom and Luca for the zero-sum games was the starting point for this work. I thank Rupak Majumdar for pointing me to this problem as an interesting and challenging problem. I am deeply indebted to Nicollas Vieille for several key insights of his results that he gave me. His insights helped me to identify the relevant fragment of his analysis in this work.

References

- [1] H. Bjorklund, S. Sandberg, and S. Vorobyov. A combinatorial strongly subexponential strategy improvement algorithm for mean payoff games. In *MFCS'04*, pages 673–685, 2004.
- [2] K. Chatterjee, L. de Alfaro, and T.A. Henzinger. The complexity of quantitative concurrent parity games. 2004. Technical Report: UCB/CSD-3-1354 (November 2004).
- [3] K. Chatterjee, L. de Alfaro, and T.A. Henzinger. The complexity of stochastic rabin and streett games. 2004. Technical Report: UCB/CSD-3-1355 (November 2004).

- [4] K. Chatterjee, L. de Alfaro, and T.A. Henzinger. Trading memory for randomness. In *Proceedings of the First Annual Conference on Quantitative Evaluation of Systems*. IEEE Computer Society Press, 2004.
- [5] K. Chatterjee, R. Majumdar, and M. Jurdziński. On nash equilibria in stochastic games. In *CSL 04: Computer Science Logic*, pages 26–40. LNCS, 2004.
- [6] L. de Alfaro and T.A. Henzinger. Concurrent omega-regular games. In *Proceedings of the 15th Annual Symposium on Logic in Computer Science*, pages 141–154. IEEE Computer Society Press, 2000.
- [7] L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. In *STOC 01: Symposium on Theory of Computing*, pages 675–683. ACM Press, 2001. Full version to appear in JCSS.
- [8] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer-Verlag, 1997.
- [9] A.M. Fink. Equilibrium in a stochastic n-person game. *Journal of Science of Hiroshima University*, 28:89–93, 1964.
- [10] A.J. Hoffman and R.M. Karp. On nonterminating stochastic games. *Management Sciences*, 12(5):359–370, 1966.
- [11] J.F. Nash Jr. Equilibrium points in n -person games. *Proceedings of the National Academy of Sciences USA*, 36:48–49, 1950.
- [12] R.J. Lipton, E. Markakis, and A. Mehta. Playing large games using simple strategies. In *EC 03: Electronic Commerce*, pages 36–41. ACM Press, 2003.
- [13] A. Maitra and W.D. Sudderth. Borel stay-in-a-set games. *International Journal of Game Theory*, 32:97–108, 2003.
- [14] Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems: Specification*. Springer-Verlag, 1992.
- [15] D.A. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63(4):1565–1581, 1998.
- [16] A.W. Mostowski. Regular expressions for infinite trees and a standard form of automata. In *5th Symposium on Computation Theory*, LNCS 208, pages 157–168. Springer-Verlag, 1984.

- [17] G. Owen. *Game Theory*. Academic Press, 1995.
- [18] C.H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and Systems Sciences*, 48(3):498–532, 1994.
- [19] C.H. Papadimitriou. Algorithms, games, and the internet. In *STOC 01: Symposium on Theory of Computing*, pages 749–753. ACM Press, 2001.
- [20] P. Secchi and W.D. Sudderth. Stay-in-a-set games. *International Journal of Game Theory*, 30:479–490, 2001.
- [21] L.S. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. USA*, 39:1095–1100, 1953.
- [22] E. Solan. Three-player absorbing games. *Mathematics of Operations Research*, 24:669–698, 1999.
- [23] E. Solan. Perturbations of markov chains with applications to stochastic games. *Chapter 17, Stochastic Games and Applications, NATO Science Series C, Mathematical and Physical Sciences*, 570:265–280, 2000. (editors A. Neyman and S. Sorin).
- [24] W. Thomas. On the synthesis of strategies in infinite games. In *STACS 95: Theoretical Aspects of Computer Science*, volume 900 of *Lecture Notes in Computer Science*, pages 1–13. Springer-Verlag, 1995.
- [25] W. Thomas. Languages, automata, and logic. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3, Beyond Words, chapter 7, pages 389–455. Springer, 1997.
- [26] F. Thuijsman. *Optimality and Equilibria in Stochastic Games*. CWI-Tract 82, CWI, Amsterdam, 1992.
- [27] F. Thuijsman and T.E.S. Raghavan. Perfect information stochastic games and related classes. *International Journal of Game Theory*, 26:403–408, 1997.
- [28] M.Y. Vardi. Automatic verification of probabilistic concurrent finite-state systems. In *Proceedings of the 26th Annual Symposium on Foundations of Computer Science*, pages 327–338. IEEE Computer Society Press, 1985.

- [29] N. Vieille. Two player stochastic games I: a reduction. *Israel Journal of Mathematics*, 119:55–91, 2000.
- [30] N. Vieille. Two player stochastic games II: the case of recursive games. *Israel Journal of Mathematics*, 119:93–126, 2000.
- [31] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, 1947.
- [32] B. von Stengel. Computing equilibria for two-person games. *Chapter 45, Handbook of Game Theory*, 3:1723–1759, 2002. (editors R.J. Aumann and S. Hart).
- [33] U. Zwick and M.S. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158:343–359, 1996.