The Complexity of Nash Equilibria in Limit-Average Games*

Michael Ummels[1] and Dominik Wojtczak[2,3]

[1] LSV, CNRS & ENS Cachan, France
ummels@lsv.ens-cachan.fr
[2] University of Liverpool, UK
d.k.wojtczak@liv.ac.uk
[3] Oxford University Computing Laboratory, UK

Abstract. We study the computational complexity of Nash equilibria in concurrent games with limit-average objectives. In particular, we prove that the existence of a Nash equilibrium in randomised strategies is undecidable, while the existence of a Nash equilibrium in pure strategies is decidable, even if we put a constraint on the payoff of the equilibrium. Our undecidability result holds even for a restricted class of concurrent games, where nonzero rewards occur only on terminal states. Moreover, we show that the constrained existence problem is undecidable not only for concurrent games but for turn-based games with the same restriction on rewards. Finally, we prove that the constrained existence problem for Nash equilibria in (pure or randomised) stationary strategies is decidable and analyse its complexity.

1 Introduction

Concurrent games provide a versatile model for the interaction of several components in a distributed system, where the components perform actions in parallel [17]. Classically, such a system is modelled by a family of concurrent two-player games, one for each component, where one component tries to fulfil its specification against the coalition of all other components. In practice, this modelling is often too pessimistic because it ignores the specifications of the other components. We argue that a distributed system is more faithfully modelled by a multiplayer game where each player has her own objective, which is independent of the other players' objectives.

Another objection to the classical theory of verification and synthesis has been that specifications are *qualitative*: either the specification is fulfilled, or it is violated. Examples of such specifications include reachability properties, where a certain set of target states has to be reached, or safety properties, where a certain set of states has to be avoided. In practice, many specifications are of a *quantitative* nature, examples of which include minimising average power

* This work was supported by ESF RNP "Games for Design and Verification" (GAMES), the French project ANR-06-SETI-003 (DOTS) and EPSRC grant EP/G050112/1.

consumption or maximising average throughput. Specifications of the latter kind can be expressed by assigning (positive or negative) rewards to states or transitions and considering the *limit-average* reward gained from an infinite play. In fact, concurrent games where a player's payoff is defined in such a way have been a central topic in game theory (see the related work section below).

The most common solution concept for games with multiple players is that of a Nash equilibrium [20]. In a Nash equilibrium, no player can improve her payoff by changing her strategy unilaterally. Unfortunately, Nash equilibria do not always exist in concurrent games, and if they exist, they may not be unique. In applications, one might look for an equilibrium where some players receive a high payoff while other players receive a low payoff. Formulated as a decision problem, given a game with k players and thresholds $\overline{x}, \overline{y} \in (\mathbb{Q} \cup \{\pm\infty\})^k$, we want to know whether the game has a Nash equilibrium whose payoff lies in-between $\overline{x}$ and $\overline{y}$; we call this decision problem NE.

The problem NE comes in several variants, depending on the type of strategies one considers: On the one hand, strategies may be *randomised* (allowing randomisation over actions) or *pure* (not allowing such randomisation). On the other hand, one can restrict to *stationary* strategies, which only depend on the last state. Indeed, we show that these restrictions give rise to distinct decision problems, which have to be analysed separately.

Our results show that the complexity of NE highly depends on the type of strategies that realise the equilibrium. In particular, we prove the following results, which yield an almost complete picture of the complexity of NE:

1. NE for pure stationary strategies is NP-complete.
2. NE for stationary strategies is decidable in PSPACE, but hard for both NP and SqrtSum.
3. NE for arbitrary pure strategies is NP-complete.
4. NE for arbitrary randomised strategies is undecidable and, in fact, not recursively enumerable.

All of our lower bounds for NE and, in particular, our undecidability result hold already for a subclass of concurrent games where Nash equilibria are guaranteed to exist, namely for *turn-based* games. If this assumption is relaxed and Nash equilibria are not guaranteed to exist, we prove that even the plain existence problem for Nash equilibria is undecidable. Moreover, many of our lower bounds hold already for games where non-zero rewards only occur on terminal states, and thus also for games where each player wants to maximise the *total sum* of the rewards.

As a byproduct of our decidability proof for pure strategies, we give a polynomial-time algorithm for deciding whether in a multi-weighted graph there exists a path whose limit-average weight vector lies between two given thresholds, a result that is of independent interest. For instance, our algorithm can be used for deciding the emptiness of a *multi-threshold mean-payoff language* [2].

Due to space constraints, most proofs are either only sketched or omitted entirely. For the complete proofs, see [27].

Related work. Concurrent and, more generally, stochastic games go back to Shapley [23], who proved the existence of the *value* for *discounted two-player zero-sum* games. This result was later generalised by Fink [13] who proved that every discounted game has a Nash equilibrium. Gillette [16] introduced limit-average objectives, and Mertens and Neyman [19] proved the existence of the value for stochastic two-player zero-sum games with limit-average objectives. Unfortunately, as demonstrated by Everett [12], these games do, in general, not admit a Nash equilibrium (see Example 1). However, Vielle [29, 30] proved that, for all $\varepsilon > 0$, every two-player stochastic limit-average game admits an *ε-equilibrium*, i.e. a pair of strategies where each player can gain at most ε from switching her strategy. Whether such equilibria always exist in games with more than two players is an important open question [21].

Determining the complexity of Nash equilibria has attracted much interest in recent years. In particular, a series of papers culminated in the result that computing a Nash equilibrium of a finite two-player game in *strategic form* is complete for the complexity class PPAD [6, 8]. The constrained existence problem, where one looks for a Nash equilibrium with certain properties, has also been investigated for games in strategic form. In particular, Conitzer and Sandholm [7] showed that deciding whether there exists a Nash equilibrium where player 0's payoff exceeds a given threshold and related decision problems are NP-complete for two-player games in strategic form.

For concurrent games with limit-average objectives, most algorithmic results concern two-player zero-sum games. In the turn-based case, these games are commonly known as *mean-payoff games* [10, 32]. While it is known that the value of such a game can be computed in pseudo-polynomial time, it is still open whether there exists a polynomial-time algorithm for solving mean-payoff games. A related model are *multi-dimensional mean-payoff games* where one player tries to maximise several mean-payoff conditions at the same time [5]. In particular, Velner and Rabinovich [28] showed that the value problem for these games is coNP-complete.

One subclass of concurrent games with limit-average objectives that has been studied in the multiplayer setting are concurrent games with reachability objectives. In particular, Bouyer et al. [3] showed that the constrained existence problem for Nash equilibria is NP-complete for these games (see also [25, 14]). We extend their result to limit-average objectives. However, we assume that strategies can observe actions (a common assumption in game theory), which they do not. Hence, while our result is more general w.r.t. the type of objectives we consider, their result is more general w.r.t. the type of strategies they allow.

In a recent paper [26], we studied the complexity of Nash equilibria in *stochastic* games with reachability objectives. In particular, we proved that NE for pure strategies is undecidable in this setting. Since we prove here that this problem is decidable in the non-stochastic setting, this undecidability result can be explained by the presence of probabilistic transitions in stochastic games. On the other hand, we prove in this paper that randomisation in strategies also leads to undecidability, a question that was left open in [26].

2 Concurrent Games

Concurrent games are played by finitely many players on a finite state space. Formally, a concurrent game is given by

- a finite nonempty set Π of *players*, e.g. $\Pi = \{0, 1, \ldots, k-1\}$,
- a finite nonempty set S of *states*,
- for each player i and each state s a nonempty set $\Gamma_i(s)$ of *actions* taken from a finite set Γ,
- a *transition function* $\delta\colon S \times \Gamma^\Pi \to S$,
- for each player $i \in \Pi$ a *reward function* $r_i\colon S \to \mathbb{R}$.

For computational purposes, we assume that all rewards are rational numbers with numerator and denominator given in binary. We say that an action profile $\bar{a} = (a_i)_{i\in\Pi}$ is *legal* at state s if $a_i \in \Gamma_i(s)$ for each $i \in \Pi$. Finally, we call a state s *controlled* by player i if $|\Gamma_j(s)| = 1$ for all $j \neq i$, and we say that a game is *turn-based* if each state is controlled by (at least) one player. For turn-based games, an action of the controlling player prescribes to go to a certain state. Hence, we will usually omit actions in turn-based games.

For a tuple $\bar{x} = (x_i)_{i\in\Pi}$, where the elements x_i belong to an arbitrary set X, and an element $x \in X$, we denote by $\bar{x}_{-i}$ the restriction of $\bar{x}$ to $\Pi \setminus \{i\}$ and by $(\bar{x}_{-i}, x)$ the unique tuple $\bar{y} \in X^\Pi$ with $y_i = x$ and $\bar{y}_{-i} = \bar{x}_{-i}$.

A play of a game $\mathcal{G}$ is an infinite sequence $s_0\bar{a}_0s_1\bar{a}_1\ldots \in (S\cdot\Gamma^\Pi)^\omega$ such that $\delta(s_j, \bar{a}_j) = s_{j+1}$ for all $j \in \mathbb{N}$. For each player, a play $\pi = s_0\bar{a}_0s_1\bar{a}_1\ldots$ gives rise to an infinite sequence of rewards. There are different criteria to evaluate this sequence and map it to a *payoff*. In this paper, we consider the *limit-average* (or *mean-payoff*) criterion, where the payoff of π for player i is defined by

$$\phi_i(\pi) \coloneqq \liminf_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} r_i(s_j).$$

Note that this payoff mapping is *prefix-independent*, i.e. $\phi_i(\pi) = \phi_i(\pi')$ if π' is a suffix of π. An important special case are games where non-zero rewards occur only on *terminal* states, i.e. states s with $\delta(s, \bar{a}) = s$ for all (legal) $\bar{a} \in \Gamma^\Pi$. These games were introduced by Everett [12] under the name *recursive games*, but we prefer to call them *terminal-reward games*. Hence, in a terminal-reward game, $\phi_i(\pi) = r_i(s)$ if π enters a terminal state s and $\phi_i(\pi) = 0$ otherwise.

Often, it is convenient to designate an *initial* state. An *initialised* game is thus a tuple $(\mathcal{G}, s_0)$ where $\mathcal{G}$ is a concurrent game and s_0 is one of its states.

Strategies and strategy profiles. For a finite set X, we denote by $\mathcal{D}(X)$ the set of probability distributions over X. A *(randomised) strategy* for player i in $\mathcal{G}$ is a mapping $\sigma\colon (S\cdot\Gamma^\Pi)^*\cdot S \to \mathcal{D}(\Gamma)$ assigning to each possible *history* $xs \in (S\cdot\Gamma^\Pi)^*\cdot S$ a probability distribution $\sigma(xs)$ over actions such that $\sigma(xs)(a) > 0$ only if $a \in \Gamma_i(s)$. We write $\sigma(a \mid xs)$ for the probability assigned to $a \in \Gamma$ by the distribution $\sigma(xs)$. A *(randomised) strategy profile of* $\mathcal{G}$ is a tuple $\bar{\sigma} = (\sigma_i)_{i\in\Pi}$ of strategies in $\mathcal{G}$, one for each player.

A strategy σ for player i is called *pure* if for each $xs \in (S \cdot \Gamma^{\Pi})^* \cdot S$ the distribution $\sigma(xs)$ is *degenerate*, i.e. there exists $a \in \Gamma_i(s)$ with $\sigma(a \mid xs) = 1$. Note that a pure strategy can be identified with a function $\sigma\colon (S \cdot \Gamma^{\Pi})^* \cdot S \to \Gamma$. A strategy profile $\overline{\sigma} = (\sigma_i)_{i \in \Pi}$ is called *pure* if each σ_i is pure, in which case we can identify $\overline{\sigma}$ with a mapping $(S \cdot \Gamma^{\Pi})^* \cdot S \to \Gamma^{\Pi}$. Note that, given an initial state s_0 and a pure strategy profile $\overline{\sigma}$, there exists a unique play $\pi = s_0 \bar{a}_0 s_1 \bar{a}_1 \ldots$ such that $\overline{\sigma}(s_0 \bar{a}_0 \ldots \bar{a}_{j-1} s_j) = \bar{a}_j$ for all $j \in \mathbb{N}$; we call π the play *induced* by $\overline{\sigma}$ from s_0.

A strategy σ is called *stationary* if σ depends only on the last state: $\sigma(xs) = \sigma(s)$ for all $xs \in (S \cdot \Gamma^{\Pi})^* \cdot S$. A strategy profile $\overline{\sigma} = (\sigma_i)_{i \in \Pi}$ is called *stationary* if each σ_i is stationary. Finally, we call a strategy (profile) *positional* if it is both pure and stationary.

The probability measure induced by a strategy profile. Given an initial state $s_0 \in S$ and a strategy profile $\overline{\sigma} = (\sigma_i)_{i \in \Pi}$, the *conditional probability* of $\bar{a} \in \Gamma^{\Pi}$ given the history $xs \in (S \cdot \Gamma^{\Pi})^* \cdot S$ is $\overline{\sigma}(\bar{a} \mid xs) := \prod_{i \in \Pi} \sigma_i(a_i \mid xs)$. The probabilities $\overline{\sigma}(\bar{a} \mid xs)$ induce a probability measure on the Borel σ-algebra over $(S \cdot \Gamma^{\Pi})^{\omega}$ as follows: The probability of a basic open set $s_1 \bar{a}_1 \ldots s_n \bar{a}_n \cdot (S \cdot \Gamma^{\Pi})^{\omega}$ equals the product $\prod_{j=1}^{n} \overline{\sigma}(\bar{a}_j \mid s_1 \bar{a}_1 \ldots \bar{a}_{j-1} s_j)$ if $s_1 = s_0$ and $\delta(s_j, \bar{a}_j) = s_{j+1}$ for all $1 \leq j < n$; in all other cases, this probability is 0. By *Carathéodory's extension theorem*, this extends to a unique probability measure assigning a probability to every Borel subset of $(S \cdot \Gamma^{\Pi})^{\omega}$, which we denote by $\Pr_{s_0}^{\overline{\sigma}}$. Via the natural projection $(S \cdot \Gamma^{\Pi})^{\omega} \to S^{\omega}$, we obtain a probability measure on the Borel σ-algebra over S^{ω}. We abuse notation and denote this measure also by $\Pr_{s_0}^{\overline{\sigma}}$; it should always be clear from the context to which measure we are referring to. Finally, we denote by $\mathrm{E}_{s_0}^{\overline{\sigma}}$ the expectation operator that corresponds to $\Pr_{s_0}^{\overline{\sigma}}$, i.e. $\mathrm{E}_{s_0}^{\overline{\sigma}}(f) = \int f \, \mathrm{d}\Pr_{s_0}^{\overline{\sigma}}$ for all Borel measurable functions $f\colon (S \cdot \Gamma^{\Pi})^{\omega} \to \mathbb{R} \cup \{\pm\infty\}$ or $f\colon S^{\omega} \to \mathbb{R} \cup \{\pm\infty\}$. In particular, we are interested in the quantities $p_i := \mathrm{E}_{s_0}^{\overline{\sigma}}(\phi_i)$. We call p_i the *(expected) payoff* of $\overline{\sigma}$ for player i and the vector $(p_i)_{i \in \Pi}$ the (expected) payoff of $\overline{\sigma}$.

Drawing concurrent games. When drawing a concurrent game as a graph, we will adhere to the following conventions: States are usually depicted as circles, but terminal states are depicted as squares. The initial state is marked by a dangling incoming edge. An edge from s to t with label $\bar{a}$ means that $\delta(s, \bar{a}) = t$ and that $\bar{a}$ is legal at s. However, the label $\bar{a}$ might be omitted if it is not essential. In turn-based games, the player who controls a state is indicated by the label next to it. Finally, a label of the form $i\colon x$ next to state s indicates that $r_i(s) = x$; if this reward is 0, the label will usually be omitted.

3 Nash Equilibria

To capture rational behaviour of selfish players, Nash [20] introduced the notion of — what is now called — a *Nash equilibrium*. Formally, given a game $\mathcal{G}$ and an

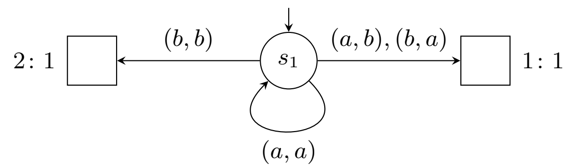

Fig. 1. A terminal-reward game that has no Nash equilibrium

initial state s_0, a strategy τ for player i is a *best response* to a strategy profile $\overline{\sigma}$ if τ maximises the expected payoff for player i, i.e.

$$\mathrm{E}_{s_0}^{\overline{\sigma}_{-i},\tau'}(\phi_i) \leq \mathrm{E}_{s_0}^{\overline{\sigma}_{-i},\tau}(\phi_i)$$

for all strategies τ' for player i. A strategy profile $\overline{\sigma} = (\sigma_i)_{i \in \Pi}$ is a *Nash equilibrium* of $(\mathcal{G}, s_0)$ if for each player i the strategy σ_i is a best response to $\overline{\sigma}$. Hence, in a Nash equilibrium no player can improve her payoff by (unilaterally) switching to a different strategy. As the following example demonstrates, Nash equilibria are not guaranteed to exist in concurrent games.

Example 1. Consider the terminal-reward game $\mathcal{G}_1$ depicted in Fig. 1, which is a variant of the game *hide-or-run* as presented by de Alfaro et al. [9]. We claim that $(\mathcal{G}_1, s_1)$ does not have a Nash equilibrium. First note that, for each $\varepsilon > 0$, player 1 can ensure a payoff of $1 - \varepsilon$ by the stationary strategy that selects action b with probability ε. Hence, every Nash equilibrium (σ, τ) of $(\mathcal{G}_1, s_1)$ must have payoff $(1, 0)$. Now consider the least k such that $p := \sigma(b \mid (s_1(a, a))^k s_1) > 0$. By choosing action b with probability 1 for the history $(s_1(a, a))^k s_1$ and choosing action a with probability 1 for all other histories, player 2 can ensure payoff p, a contradiction to (σ, τ) being a Nash equilibrium.

It follows from Nash's theorem [20] that every game whose arena is a tree (or a DAG) has a Nash equilibrium. Another important special case of concurrent limit-average games where Nash equilibria always exist are turn-based games. For these games, Thuijsman and Raghavan [24] proved not only the existence of arbitrary Nash equilibria but of pure finite-state ones.

To measure the complexity of Nash equilibria in concurrent games, we introduce the following decision problem, which we call NE:

> Given a game $\mathcal{G}$, a state s_0 and thresholds $\overline{x}, \overline{y} \in (\mathbb{Q} \cup \{\pm\infty\})^{\Pi}$, decide whether $(\mathcal{G}, s_0)$ has a Nash equilibrium with payoff $\geq \overline{x}$ and $\leq \overline{y}$.

Note that we have not put any restriction on the type of strategies that realise the equilibrium. It is natural to restrict the search space to profiles of pure, stationary or positional strategies. These restrictions give rise to different decision problems, which we call PureNE, StatNE and PosNE, respectively.

Before we analyse the complexity of these problems, let us convince ourselves that these problems are not just different faces of the same coin. We first show that the decision problems where we look for equilibria in randomised strategies are distinct from the ones where we look for equilibria in pure strategies.

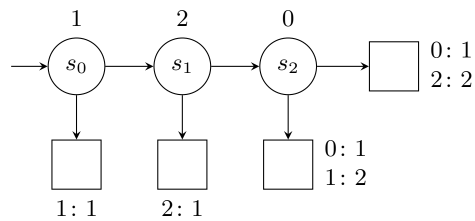

Fig. 2. A game with no pure Nash equilibrium where player 0 wins with positive probability

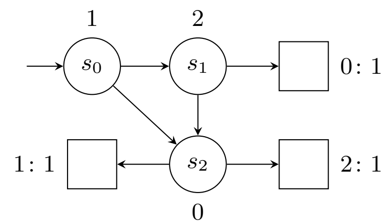

Fig. 3. A game with no stationary Nash equilibrium where player 0 wins with positive probability

Proposition 2. *There exists a turn-based terminal-reward game that has a stationary Nash equilibrium where player* 0 *receives payoff* 1 *but that has no pure Nash equilibrium where player* 0 *receives payoff* > 0.

Proof. Consider the game depicted in Fig. 2 played by three players 0, 1 and 2. Clearly, the stationary strategy profile where from state s_2 player 0 selects both outgoing transitions with probability $\frac{1}{2}$ each, player 1 plays from s_0 to s_1 and player 2 plays from s_1 to s_2 is a Nash equilibrium where player 0 receives payoff 1. However, in any pure strategy profile where player 0 receives payoff > 0, either player 1 or player 2 receives payoff 0 and could improve her payoff by switching her strategy at s_0 or s_1, respectively. ☐

Now we show that it makes a difference whether we look for an equilibrium in stationary strategies or not.

Proposition 3. *There exists a turn-based terminal-reward game that has a pure Nash equilibrium where player* 0 *receives payoff* 1 *but that has no stationary Nash equilibrium where player* 0 *receives payoff* > 0.

Proof. Consider the game $\mathcal{G}$ depicted in Fig. 3 and played by three players 0, 1 and 2. Clearly, the pure strategy profile that leads to the terminal state with payoff 1 for player 0 and where player 0 plays "right" if player 1 has deviated and "left" if player 2 has deviated is a Nash equilibrium of $(\mathcal{G}, s_0)$ with payoff 1 for player 0. Now consider any stationary equilibrium of $(\mathcal{G}, s_0)$ where player 0 receives payoff > 0. If the stationary strategy of player 0 prescribes to play "right" with positive probability, then player 2 can improve her payoff by playing to s_2 with probability 1, and otherwise player 1 can improve her payoff by playing to s_2 with probability 1, a contradiction. ☐

It follows from Proposition 2 that NE and StatNE are different from PureNE and PosNE, and it follows from Proposition 3 that NE and PureNE are different from StatNE and PosNE. Hence, all of these decision problems are pairwise distinct, and their decidability and complexity has to be studied separately.

4 Positional Strategies

In this section, we show that the problem PosNE is NP-complete. Since we can check in polynomial time whether a positional strategy profile is a Nash equilibrium (using a result of Karp [18]), membership in NP is straightforward.

Theorem 4. *PosNE is in NP.*

A result by Chatterjee et al. [5, Lemma 15] implies that PosNE is NP-hard, even for turn-based games with rewards taken from $\{-1, 0, 1\}$ (but the number of players is unbounded). We strengthen their result by showing that the problem remains NP-hard if there are only three players and rewards are taken from $\{0, 1\}$.

Theorem 5. *PosNE is NP-hard, even for turn-based three-player games with rewards* 0 *and* 1*.*

Proof. We reduce from the Hamiltonian cycle problem. Given a graph $G = (V, E)$, we define a turn-based three-player game $\mathcal{G}$ as follows: the set of states is V, all states are controlled by player 0, and the transition function corresponds to E (i.e. $\Gamma_0(v) = vE$ and $\delta(v, \bar{a}) = w$ if and only if $a_0 = w$). Let $n = |V|$ and $v_0 \in V$. The reward of state v_0 to player 1 equals 1. All other rewards for player 0 and player 1 equal 0. Finally, player 2 receives reward 0 at v_0 and reward 1 at all other states. We claim that there is a Hamiltonian cycle in G if and only if $(\mathcal{G}, v_0)$ has a positional Nash equilibrium with payoff $\geq (0, 1/n, (n-1)/n)$. □

By combining our reduction with a game that has no positional Nash equilibrium, we can prove the following stronger result for non-turn-based games.

Corollary 6. *Deciding the existence of a positional Nash equilibrium in a concurrent limit-average game is NP-complete, even for three-player games.*

5 Stationary Strategies

To prove the decidability of StatNE, we appeal to results established for the *existential theory of the reals*, the set of all existential first-order sentences that hold in the ordered field $\mathfrak{R} := (\mathbb{R}, +, \cdot, 0, 1, \leq)$. The best known upper bound for the complexity of the associated decision problem is PSPACE [4], which leads to the following theorem.

Theorem 7. *StatNE is in* PSPACE*.*

The next theorem shows that StatNE is NP-hard, even for turn-based games with rewards 0 and 1. Note that this does not follow from the NP-hardness of PosNE, but requires a different proof.

Theorem 8. *StatNE is NP-hard, even for turn-based games with rewards* 0 *and* 1*.*

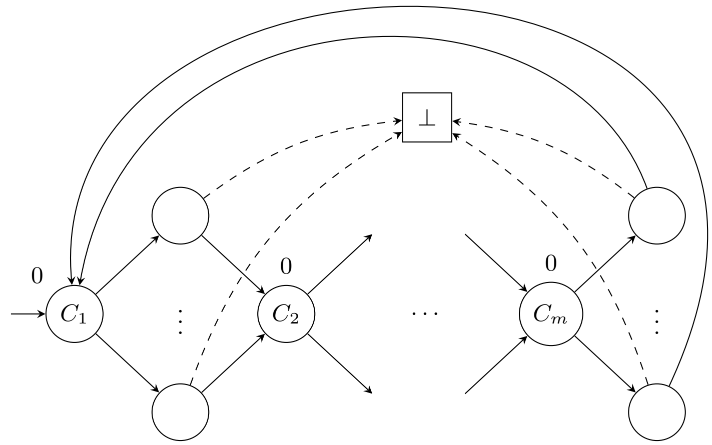

Fig. 4. Reducing SAT to StatNE

Proof. We employ a reduction from SAT, which resembles a reduction in [25]. Given a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_m$ in conjunctive normal form over propositional variables $X_1, \ldots, X_n$, where w.l.o.g. $m \geq 1$ and each clause is nonempty, we build a turn-based game $\mathcal{G}$ played by players $0, 1, \ldots, n$ as follows: The game $\mathcal{G}$ has states $C_1, \ldots, C_m$ controlled by player 0 and for each clause C and each literal L that occurs in C a state (C, L), controlled by player i if $L = X_i$ or $L = \neg X_i$; additionally, the game contains a terminal state $\bot$. There are transitions from a clause C_j to each state (C_j, L) such that L occurs in C_j and from there to $C_{(j \bmod m)+1}$, and there is a transition from each state of the form $(C, \neg X)$ to $\bot$. Each state except $\bot$ has reward 1 for player 0, whereas $\bot$ has reward 0 for player 0. For player i, each state except states of the form (C, X_i) has reward 1; states of the form (C, X_i) have reward 0. The structure of $\mathcal{G}$ is depicted in Fig. 4. Clearly, $\mathcal{G}$ can be constructed from φ in polynomial time. We claim that φ is satisfiable if and only if $(\mathcal{G}, C_1)$ has a stationary Nash equilibrium with payoff ≥ 1 for player 0. $\square$

By combining our reduction with the game from Example 1, which has no Nash equilibrium, we can prove the following stronger result for concurrent games.

Corollary 9. *Deciding the existence of a stationary Nash equilibrium in a concurrent limit-average game with rewards* 0 *and* 1 *is NP-hard.*

So far we have shown that StatNE is contained in PSPACE and hard for NP, leaving a considerable gap between the two bounds. In order to gain a better understanding of StatNE, we relate this problem to the *square root sum problem* (SqrtSum), an important problem about numerical computations. Formally, SqrtSum is the following decision problem: Given numbers $d_1, \ldots, d_n, k \in \mathbb{N}$, decide whether $\sum_{i=1}^{n} \sqrt{d_i} \geq k$. Recently, Allender et al. [1] showed that SqrtSum belongs to the fourth level of the *counting hierarchy*, a slight improvement over the previously known PSPACE upper bound. However, it has been an open question since the 1970s as to whether SqrtSum falls into the polynomial hierarchy

[15, 11]. We give a polynomial-time reduction from SqrtSum to StatNE for turn-based terminal-reward games. Hence, StatNE is at least as hard as SqrtSum, and showing that StatNE resides inside the polynomial hierarchy would imply a major breakthrough in understanding the complexity of numerical computations. While our reduction is similar to the one in [26], it requires new techniques to simulate stochastic states.

Theorem 10. *SqrtSum is polynomial-time reducible to StatNE for turn-based 8-player terminal-reward games.*

Again, we can combine our reduction with the game from Example 1 to prove a stronger result for games that are not turn-based.

Corollary 11. *Deciding whether a concurrent 8-player terminal reward game has a stationary Nash equilibrium is hard for SqrtSum.*

Remark 12. By appealing to results on Markov decision processes with limit-average objectives (see e.g. [22]), the positive results of Sects. 4 and 5 can be extended to stochastic games (with the same complexity bounds).

6 Pure Strategies

In this section, we show that PureNE is decidable and, in fact, NP-complete. Let $\mathcal{G}$ be a concurrent game, $s \in S$ and $i \in \Pi$. We define

$$\mathrm{pval}_i^{\mathcal{G}}(s) = \inf_{\overline{\sigma}} \sup_{\tau} \mathrm{E}_s^{\overline{\sigma}_{-i},\tau}(\phi_i),$$

where $\overline{\sigma}$ ranges over all *pure* strategy profiles of $\mathcal{G}$ and τ ranges over all strategies of player i. Intuitively, $\mathrm{pval}_i^{\mathcal{G}}(s)$ is the lowest payoff that the coalition $\Pi \setminus \{i\}$ can inflict on player i by playing a pure strategy.

By a reduction to a turn-based two-player zero-sum game, we can show that there is a positional strategy profile that attains this value.

Proposition 13. *Let $\mathcal{G}$ be a concurrent game, and $i \in \Pi$. There exists a positional strategy profile $\overline{\sigma}^*$ such that $\mathrm{E}_s^{\overline{\sigma}^*_{-i},\tau}(\phi_i) \leq \mathrm{pval}_i^{\mathcal{G}}(s)$ for all states s and all strategies τ of player i.*

Given a payoff vector $\overline{z} \in (\mathbb{R} \cup \{\pm\infty\})^{\Pi}$, we define a directed graph $G(\overline{z}) = (V, E)$ (with self-loops) as follows: $V = S$, and there is an edge from s to t if and only if there is an action profile $\overline{a}$ with $\delta(s, \overline{a}) = t$ such that (1) $\overline{a}$ is legal at s and (2) $\mathrm{pval}_i^{\mathcal{G}}(\delta(s, (\overline{a}_{-i}, b))) \leq z_i$ for each player i and each action $b \in \Gamma_i(s)$. Following [3], we call any $\overline{a}$ that fulfils (1) and (2) $\overline{z}$-*secure* at s.

Lemma 14. *Let $\overline{z} \in (\mathbb{R} \cup \{\pm\infty\})^{\Pi}$. If there exists an infinite path π in $G(\overline{z})$ from s_0 with $z_i \leq \phi_i(\pi)$ for each player i, then $(\mathcal{G}, s_0)$ has a pure Nash equilibrium with payoff $\phi_i(\pi)$ for player i.*

Proof. Let $\pi = s_0 s_1 \ldots$ be an infinite path in $G(\bar{z})$ from s_0 with $z_i \le \phi_i(\pi)$ for each player i. We define a pure strategy profile $\bar{\sigma}$ as follows: For histories of the form $x = s_0 \bar{a}_0 s_1 \ldots s_{k-1} \bar{a}_{k-1} s_k$, we set $\bar{\sigma}(x)$ to an action profile $\bar{a}$ with $\delta(s_k, \bar{a}) = s_{k+1}$ that is $\bar{z}$-secure at s_k. For all other histories $x = t_0 \bar{a}_0 t_1 \ldots t_{k-1} \bar{a}_{k-1} t_k$, consider the least j such that $s_{j+1} \ne t_{j+1}$. If $\bar{a}_j$ differs from a $\bar{z}$-secure action profile $\bar{a}$ at s_j in precisely one entry i, we set $\bar{\sigma}(x) = \bar{\sigma}^*(t_k)$, where $\bar{\sigma}^*$ is a (fixed) positional strategy profile such that $\mathrm{E}_s^{\bar{\sigma}^*_{-i},\tau}(\phi_i) \le \mathrm{pval}_i^{\mathcal{G}}(s)$ for all $s \in S$ (which is guaranteed to exist by Proposition 13); otherwise, $\bar{\sigma}(x)$ can be chosen arbitrarily. It is easy to see that $\bar{\sigma}$ is a Nash equilibrium with induced play π. ◻

Lemma 15. *Let $\bar{\sigma}$ be a pure Nash equilibrium of $(\mathcal{G}, s_0)$ with payoff $\bar{z}$. Then there exists an infinite path π in $G(\bar{z})$ from s_0 with $\phi_i(\pi) = z_i$ for each player i.*

Proof. Let $s_0 \bar{a}_0 s_1 \bar{a}_1 \ldots$ be the play induced by $\bar{\sigma}$. We claim that $\pi \coloneqq s_0 s_1 \ldots$ is a path in $G(\bar{z})$. Otherwise, consider the least k such that (s_k, s_{k+1}) is not an edge in $G(\bar{z})$. Hence, there exists no $\bar{z}$-secure action profile at $s \coloneqq s_k$. Since $\bar{a}_k$ is certainly legal at s, there exists a player i and an action $b \in \Gamma_i(s)$ such that $\mathrm{pval}_i^{\mathcal{G}}(\delta(s, (\bar{a}_{-i}, b))) > z_i$. But then player i can improve her payoff by switching to a strategy that mimics σ_i until s is reached, then plays action b, and after that mimics a strategy τ with $\mathrm{E}_{\delta(s,(\bar{a}_{-i},b))}^{\bar{\sigma}_{-i},\tau}(\phi_i) > z_i$. This contradicts the assumption that $\bar{\sigma}$ is a Nash equilibrium. ◻

Using Lemmas 14 and 15, we can reduce the task of finding a pure Nash equilibrium to the task of finding a path in a multi-weighted graph whose limit-average weight vector falls between two thresholds. The latter problem can be solved in polynomial time by solving a linear programme with one variable for each pair of a weight function and an edge in the graph.

Theorem 16. *Given a finite directed graph $G = (V, E)$ with weight functions $r_0, \ldots, r_{k-1} \colon V \to \mathbb{Q}$, $v_0 \in V$, and $\bar{x}, \bar{y} \in (\mathbb{Q} \cup \{\pm\infty\})^k$, we can decide in polynomial time whether there exists an infinite path $\pi = v_0 v_1 \ldots$ in G with $x_i \le \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} r_i(v_j) \le y_i$ for all $i = 0, \ldots, k-1$.*

We can now describe a nondeterministic algorithm to decide the existence of a pure Nash equilibrium with payoff $\ge \bar{x}$ and $\le \bar{y}$ in polynomial time. The algorithm starts by guessing, for each player i, a positional strategy profile $\bar{\sigma}^i$ of $\mathcal{G}$ and computes $p_i(s) \coloneqq \sup_\tau \mathrm{E}_s^{\bar{\sigma}^i_{-i},\tau}(\phi_i)$ for each $s \in S$; these numbers can be computed in polynomial time using the algorithm given by Karp [18]. The algorithm then guesses a vector $\bar{z} \in (\mathbb{R} \cup \{\pm\infty\})^{\Pi}$ by setting z_i either to x_i or to $p_i(s)$ for some $s \in S$ with $x_i \le p_i(s)$, and constructs the graph $G'(\bar{z})$, which is defined as $G(\bar{z})$ but with $p_i(s)$ substituted for $\mathrm{pval}_i^{\mathcal{G}}(s)$. Finally, the algorithm determines (in polynomial time) whether there exists an infinite path π in $G(\bar{z})$ from s_0 with $z_i \le \phi_i(\pi) \le y_i$ for all $i \in \Pi$. If such a path exists, the algorithm accepts; otherwise it rejects.

Theorem 17. *PureNE is in NP.*

Proof. We claim that the algorithm described above is correct, i.e. sound and complete. To prove soundness, assume that the algorithm accepts its input. Hence, there exists an infinite path π in $G'(\bar{z})$ from s_0 with $z_i \leq \phi_i(\pi) \leq y_i$. Since $\mathrm{pval}_i^{\mathcal{G}}(s) \leq p_i(s)$ for all $i \in \Pi$ and $s \in S$, the graph $G'(\bar{z})$ is a subgraph of $G(\bar{z})$. Hence, π is also an infinite path in $G(\bar{z})$. By Lemma 14, we can conclude that $(\mathcal{G}, s_0)$ has a pure Nash equilibrium with payoff $\geq \bar{z} \geq \bar{x}$ and $\leq \bar{y}$.

To prove that the algorithm is complete, let $\bar{\sigma}$ be a pure Nash equilibrium of $(\mathcal{G}, s_0)$ with payoff $\bar{z}$, where $\bar{x} \leq \bar{z} \leq \bar{y}$. By Proposition 13, the algorithm can guess positional strategy profiles $\bar{\sigma}^i$ such that $p_i(s) = \mathrm{pval}_i^{\mathcal{G}}(s)$ for all $s \in S$. If the algorithm additionally guesses the payoff vector $\bar{z}'$ defined by $z'_i = \max\{x_i, \mathrm{pval}_i^{\mathcal{G}}(s) : s \in S, \mathrm{pval}_i^{\mathcal{G}}(s) \leq z_i\}$ for all $i \in \Pi$, then the graph $G(\bar{z})$ coincides with the graph $G(\bar{z}')$ (and thus with $G'(\bar{z}')$). By Lemma 15, there exists an infinite path π in $G(\bar{z})$ from s_0 such that $z'_i \leq z_i = \phi_i(\pi) \leq y_i$ for all $i \in \Pi$. Hence, the algorithm accepts. □

The following theorem shows that PureNE is NP-hard. In fact, NP-hardness holds even for turn-based games with rewards 0 and 1.

Theorem 18. *PureNE is NP-hard, even for turn-based games with rewards* 0 *and* 1*.*

Proof. Again, we reduce from SAT. Given a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_m$ in conjunctive normal form over propositional variables $X_1, \ldots, X_n$, where w.l.o.g. $m \geq 1$ and each clause is nonempty, let $\mathcal{G}$ be the turn-based game described in the proof of Theorem 8 and depicted in Fig. 4. We claim that φ is satisfiable if and only if $(\mathcal{G}, C_1)$ has a pure Nash equilibrium with payoff ≥ 1 for player 0. □

It follows from Theorems 17 and 18 that PureNE is NP-complete. By combining our reduction with a game that has no pure Nash equilibrium, we can prove the following stronger result for non-turn-based games.

Corollary 19. *Deciding the existence of a pure Nash equilibrium in a concurrent limit-average game is NP-complete, even for games with rewards* 0 *and* 1*.*

Note that Theorem 18 and Corollary 19 do not apply to terminal-reward games. In fact, PureNE is decidable in P for these games, which follows from two facts about terminal-reward games: (1) the numbers $\mathrm{pval}_i^{\mathcal{G}}(s)$ can be computed in polynomial time (using a reduction to a turn-based two-player zero-sum game and applying a result of Washburn [31]), and (2) the only possible vectors that can emerge as the payoff of a pure strategy profile are the zero vector and the reward vectors at terminal states.

Theorem 20. *PureNE is in P for terminal-reward games.*

7 Randomised Strategies

In this section, we show that the problem NE is undecidable and, in fact, not recursively enumerable for turn-based terminal-reward games.

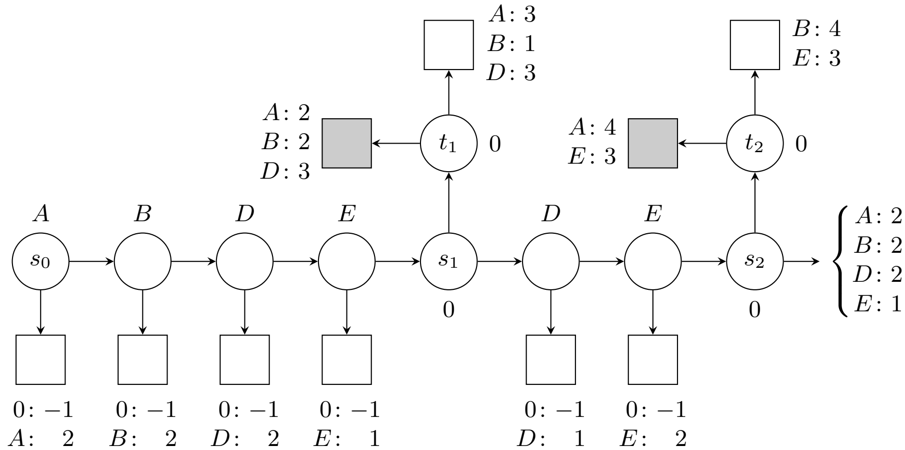

Fig. 5. Incrementing a counter

Theorem 21. *NE is not recursively enumerable, even for turn-based 14-player terminal-reward games.*

Proof (Sketch). The proof is by a reduction from the non-halting problem for two-counter machines: we show that one can compute from a deterministic two-counter machine $\mathcal{M}$ a turn-based 14-player terminal-reward game $(\mathcal{G}, s_0)$ such that the computation of $\mathcal{M}$ is infinite if and only if $(\mathcal{G}, s_0)$ has a Nash equilibrium where player 0 receives payoff ≥ 0.

To get a flavour of the full proof, let us consider a one-counter machine $\mathcal{M}$ that contains an increment instruction. A (simplified) part of the game $\mathcal{G}$ is depicted in Fig. 5. The counter values before and after the increment operation are encoded by the probabilities $c_1 = 2^{-i_1}$ and $c_2 = 2^{-i_2}$ that player 0 plays from t_1, respectively t_2, to the neighbouring grey state. We claim that $c_2 = \frac{1}{2}c_1$, i.e. $i_2 = i_1 + 1$, in any Nash equilibrium $\overline{\sigma}$ of $(\mathcal{G}, s_0)$ where player 0 receives payoff ≥ 0. First note that player 0 has to choose both outgoing transitions with probability $\frac{1}{2}$ each at s_1 and s_2 because otherwise player D or player E would have an incentive to play to a state where player 0 receives payoff < 0. Now consider the payoffs $a = \mathrm{E}^{\overline{\sigma}}_{s_0}(\phi_A)$ and $b = \mathrm{E}^{\overline{\sigma}}_{s_0}(\phi_B)$ for players A and B. We have $a, b \geq 2$ because otherwise one of these two players would have an incentive to play to a state where player 0 receives payoff < 0. On the other hand, the payoffs of players A and B sum up to at most 4 in every terminal state. Hence, $a + b \leq 4$ and therefore $a = b = 2$. Finally, the expected payoff for player A equals

$$a = \tfrac{1}{2}\big(c_1 \cdot 2 + (1 - c_1) \cdot 3\big) + \tfrac{1}{4} \cdot c_2 \cdot 4 + \tfrac{1}{4} \cdot 2 = 2 - \tfrac{1}{2}c_1 + c_2 \,.$$

Obviously, $a = 2$ if and only if $c_2 = \frac{1}{2}c_1$. □

For games that are not turn-based, by combining our reduction with the game from Example 1, we can show the stronger theorem that deciding the existence of *any* Nash equilibrium is not recursively enumerable.

Corollary 22. *The set of all initialised concurrent 14-player terminal-reward games that have a Nash equilibrium is not recursively enumerable.*

8 Conclusion

We have analysed the complexity of Nash equilibria in concurrent games with limit-average objectives. In particular, we showed that randomisation in strategies leads to undecidability, while restricting to pure strategies retains decidability. This is in contrast to stochastic games, where pure strategies lead to undecidability [26]. While we provided matching and lower bounds in most cases, there remain some problems where we do not know the exact complexity. Apart from StatNE, these include the problem PureNE when restricted to a bounded number of players.

References

1. E. Allender, P. Bürgisser, J. Kjeldgaard-Pedersen, P. B. Miltersen. On the complexity of numerical analysis. *SIAM Journal on Computing*, 38(5):1987–2006, 2009.
2. R. Alur, A. Degorre, O. Maler, G. Weiss. On omega-languages defined by mean-payoff conditions. In *FOSSACS 2009*, vol. 5504 of *LNCS*, pp. 333–347. Springer-Verlag, 2009.
3. P. Bouyer, R. Brenguier, N. Markey. Nash equilibria for reachability objectives in multi-player timed games. In *CONCUR 2010*, vol. 6269 of *LNCS*, pp. 192–206. Springer-Verlag, 2010.
4. J. Canny. Some algebraic and geometric computations in PSPACE. In *STOC '88*, pp. 460–469. ACM Press, 1988.
5. K. Chatterjee, L. Doyen, T. A. Henzinger, J.-F. Raskin. Generalized mean-payoff and energy games. In *FSTTCS 2010*, vol. 8 of *LIPICS*. Schloss Dagstuhl, 2010.
6. X. Chen, X. Deng, S.-H. Teng. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM*, 56(3), 2009.
7. V. Conitzer, T. Sandholm. Complexity results about Nash equilibria. In *IJCAI 2003*, pp. 765–771. Morgan Kaufmann, 2003.
8. C. Daskalakis, P. W. Goldberg, C. H. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.
9. L. de Alfaro, T. A. Henzinger, O. Kupferman. Concurrent reachability games. *Theoretical Computer Science*, 386(3):188–217, 2007.
10. A. Ehrenfeucht, J. Mycielski. Positional strategies for mean payoff games. *International Journal of Game Theory*, 8:109–113, 1979.
11. K. Etessami, M. Yannakakis. On the complexity of Nash equilibria and other fixed points. *SIAM Journal on Computing*, 39(6):2531–2597, 2010.
12. H. Everett. Recursive games. In M. Dresher, A. W. Tucker, P. Wolfe, eds., *Contributions to the Theory of Games III*, vol. 39 of *Annals of Mathematical Studies*, pp. 47–78. Princeton University Press, 1957.

13. A. M. Fink. Equilibrium in a stochastic n-person game. *Journal of Science in Hiroshima University*, 28(1):89–93, 1964.
14. D. Fisman, O. Kupferman, Y. Lustig. Rational synthesis. In *TACAS 2010*, vol. 6015 of *LNCS*, pp. 190–204. Springer-Verlag, 2010.
15. M. R. Garey, R. L. Graham, D. S. Johnson. Some NP-complete geometric problems. In *STOC '76*, pp. 10–22. ACM Press, 1976.
16. D. Gillette. Stochastic games with zero stop probabilities. In M. Dresher, A. W. Tucker, P. Wolfe, eds., *Contributions to the Theory of Games III*, vol. 39 of *Annals of Mathematical Studies*, pp. 179–187. Princeton University Press, 1957.
17. T. A. Henzinger. Games in system design and verification. In *TARK 2005*, pp. 1–4. National University of Singapore, 2005.
18. R. M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete Mathematics*, 23(3):309–311, 1978.
19. J.-F. Mertens, A. Neyman. Stochastic games. *International Journal of Game Theory*, 10(2):53–66, 1981.
20. J. F. Nash, Jr. Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences of the USA*, 36:48–49, 1950.
21. A. Neyman, S. Sorin, eds. *Stochastic Games and Applications*, vol. 570 of *NATO Science Series C*. Springer-Verlag, 2003.
22. M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.
23. L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences of the USA*, 39:1095–1100, 1953.
24. F. Thuijsman, T. E. S. Raghavan. Perfect-information stochastic games and related classes. *International Journal of Game Theory*, 26:403–408, 1997.
25. M. Ummels. The complexity of Nash equilibria in infinite multiplayer games. In *FOSSACS 2008*, vol. 4962 of *LNCS*, pp. 20–34. Springer-Verlag, 2008.
26. M. Ummels, D. Wojtczak. The complexity of Nash equilibria in simple stochastic multiplayer games. In *ICALP 2009*, vol. 5556 of *LNCS*, pp. 297–308. Springer-Verlag, 2009.
27. M. Ummels, D. Wojtczak. The complexity of Nash equilibria in limit-average games. Tech. Rep. LSV-11-15, ENS Cachan, 2011.
28. Y. Velner, A. Rabinovich. Church synthesis problem for noisy input. In *FOSSACS 2011*, vol. 6604 of *LNCS*, pp. 275–289. Springer-Verlag, 2011.
29. N. Vielle. Two-player stochastic games I: A reduction. *Israel Journal of Mathematics*, 119(1):55–91, 2000.
30. N. Vielle. Two-player stochastic games II: The case of recursive games. *Israel Journal of Mathematics*, 119(1):93–126, 2000.
31. A. R. Washburn. Deterministic graphical games. *Journal of Mathematical Analysis and Applications*, 153:84–96, 1990.
32. U. Zwick, M. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158(1–2):343–359, 1996.