THE COMPLEXITY OF NASH EQUILIBRIA IN STOCHASTIC MULTIPLAYER GAMES

MICHAEL UMMELS AND DOMINIK WOJTCZAK

RWTH Aachen University, Germany
e-mail address: ummels@logic.rwth-aachen.de

CWI Amsterdam, The Netherlands
e-mail address: d.k.wojtczak@cwi.nl

Abstract. We analyse the computational complexity of finding Nash equilibria in stochastic multiplayer games with ω-regular objectives. We show that restricting the search space to equilibria whose payoffs fall into a certain interval may lead to undecidability. In particular, we prove that the following problem is undecidable: Given a game $\mathcal{G}$, does there exist a pure-strategy Nash equilibrium of $\mathcal{G}$ where player 0 wins with probability 1. Moreover, this problem remains undecidable if it is restricted to strategies with (unbounded) finite memory. However, if randomised strategies are allowed, decidability remains an open problem; we can only prove NP-hardness in this case. One way to obtain a provably decidable variant of the problem is to restrict the strategies to be positional or stationary. For the complexity of these two problems, we obtain a common lower bound of NP and upper bounds of NP and Pspace respectively. Finally, we single out a special case of the general problem that, in many cases, admits an efficient solution. In particular, we prove that deciding the existence of an equilibrium in which each player either wins or loses with probability 1 can be done in polynomial time for games where, for instance, the objective of each player is given by a parity condition with a bounded number of priorities.

1. Introduction

We study *stochastic games* [50] played by multiple players on a finite, directed graph. Intuitively, a play of such a game evolves by moving a token along edges of the graph: Each vertex of the graph is either controlled by one of the players, or it is *stochastic*. Whenever the token arrives at a non-stochastic vertex, the player who controls this vertex must move the token to a successor vertex; when the token arrives at a stochastic vertex, a fixed

1998 ACM Subject Classification: F.1.2 [Computation by Abstract Devices]: Modes of Computation—Probabilistic computation; G.1.6 [Numerical Analysis]: Optimization—Stochastic Programming; G.3 [Probability and Statistics]: Stochastic processes.

Key words and phrases: Nash equilibria, Stochastic games, Computational complexity.

Preliminary versions of parts of this paper appeared in the *Proceedings of the 36th International Colloquium on Automata, Languages and Programming* (ICALP 2009) and the *Proceedings of the 18th Annual Conference of the European Association for Computer Science Logic* (CSL 2009). This work was supported by the DFG Research Training Group 1298 (AlgoSyn) and the ESF Research Networking Programme GAMES.

LOGICAL METHODS
IN COMPUTER SCIENCE

DOI:10.2168/LMCS-???

© M. Ummels and D. Wojtczak
Creative Commons

probability distribution determines the next vertex. A measurable function maps plays to payoffs. In the simplest case, which we discuss here, the possible payoffs of a single play are binary (i.e. each player either wins or loses a given play). However, due to the presence of stochastic vertices, a player's *expected payoff* (i.e. her probability of winning) can be an arbitrary probability.

Stochastic games with ω-regular objectives have been successfully applied in the verification and synthesis of reactive systems under the influence of random events. Such a system is usually modelled as a game between the system and its environment, where the environment's objective is the complement of the system's objective: the environment is considered hostile. Therefore, the research in this area has traditionally focused on two-player games where each play is won by precisely one of the two players, so-called *two-player zero-sum games*. However, the system may comprise of several components with independent objectives, a situation which is naturally modelled by a multiplayer game.

The most common interpretation of rational behaviour in multiplayer games is captured by the notion of a *Nash equilibrium* [49]. In a Nash equilibrium, no player can improve her payoff by unilaterally switching to a different strategy. Chatterjee et al. [13] gave an algorithm for computing a Nash equilibrium in a stochastic multiplayer games with ω-regular winning conditions. We argue that this is not satisfactory. Indeed, it can be shown that their algorithm may compute an equilibrium where all players lose almost surely (i.e. receive expected payoff 0), while there exist other equilibria where all players win almost surely (i.e. receive expected payoff 1).

In applications, one might look for an equilibrium where as many players as possible win almost surely or where it is guaranteed that the expected payoff of the equilibrium falls into a certain interval. Formulated as a decision problem, we want to know, given a k-player game $\mathcal{G}$ with initial vertex v_0 and two thresholds $\bar{x}, \bar{y} \in [0,1]^k$, whether $(\mathcal{G}, v_0)$ has a Nash equilibrium with expected payoff at least $\bar{x}$ and at most $\bar{y}$. This problem, which we call NE for short, is a generalisation of the *quantitative decision problem* for two-player zero-sum games, which asks whether in such a game player 0 has a strategy that ensures to win the game with a probability that lies above a given threshold.

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 strategies that use (unbounded or bounded) finite memory or even to *stationary* ones (strategies that do not use any memory at all). For the quantitative decision problem, this distinction is often not meaningful since in a two-player, zero-sum simple stochastic game with ω-regular objectives both players have optimal pure strategies with finite memory. Moreover, in many games even positional (i.e. both pure and stationary) strategies suffice for optimality. However, regarding NE this distinction leads to distinct decision problems, which have to be analysed separately.

Our main result is that NE is undecidable if only pure strategies are considered. In fact, even the following, presumably simpler, problem is undecidable: Given a game $\mathcal{G}$, decide whether there exists a pure Nash equilibrium where player 0 wins almost surely. Moreover, the problem remains undecidable if one restricts to pure strategies that use (unbounded) finite memory. However, for the general case of arbitrary randomised strategies, decidability remains an open problem: in this case, we can only prove NP-hardness.

If one restricts to simpler types of strategies like stationary ones, the problem becomes provably decidable. In particular, for positional strategies the problem is typically NP-complete, and for arbitrary stationary strategies it is NP-hard but typically contained in PSPACE. To get a better understanding of the latter problem, we also relate it to the *Square Root Sum Problem* (SqrtSum) by providing a polynomial-time reduction from SqrtSum to NE with the restriction to stationary strategies. It is a long-standing open problem whether SqrtSum falls into the polynomial hierarchy; hence, showing that NE for stationary strategies lies inside the polynomial hierarchy would imply a breakthrough in understanding the complexity of numerical computations.

Although the decidability of NE with respect to arbitrary randomised strategies remains open, we can prove decidability for an important restriction of NE, which we call the strictly qualitative fragment. This fragment arises from NE by restricting the two thresholds to be the same binary payoff. Hence, we are only interested in equilibria where each player either wins or loses with probability 1. Formally, the task is to decide, given a k-player game $\mathcal{G}$ with initial vertex v_0 and a binary payoff $\overline{x} \in \{0,1\}^k$, whether the game has a Nash equilibrium with expected payoff $\overline{x}$. Apart from proving decidability, we show that, depending on the representation of the objective, this problem is typically complete for one of the complexity classes P, NP, coNP and PSPACE, and that the problem is invariant under restricting the search space to equilibria in pure, finite-state strategies.

Outline. In Section 2, we introduce the model that underlies this work and survey earlier work on stochastic two-player zero-sum games. In Section 3, we prove that every stochastic multiplayer game has a Nash equilibrium, thereby addressing an inaccuracy in an earlier proof by Chatterjee et al. [13]. In Section 4, we analyse the complexity of the problem NE with respect to the six modes of strategies we consider in this work: positional strategies, stationary strategies, pure finite-state strategies, randomised finite-state strategies, arbitrary pure strategies, and arbitrary randomised strategies. Finally, in Section 5, we prove that the strictly qualitative fragment of NE is decidable and analyse its complexity.

Related Work. 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 two-player game in strategic form is complete for the complexity class PPAD [19, 15]. More in the spirit of our work, Conitzer and Sandholm [17] showed that deciding whether there exists a Nash equilibrium in a two-player game in strategic form where player 0 receives payoff at least x and related decision problems are all NP-hard. For infinite games (without stochastic vertices), (a qualitative version of) the problem NE was studied in [54]. In particular, it was shown that the problem is NP-complete for games with parity winning conditions but in P for games with Büchi winning conditions.

For stochastic games, most results concern the computation of values and optimal strategies; see Section 2 for a survey of the most important results. We are only aware of two papers that are closely related to our problem: First, Etessami et al. [25] investigated Markov decision processes with, e.g., multiple reachability objectives. Such a system can be viewed as a stochastic multiplayer game where all non-stochastic vertices are controlled by one single player. Under this interpretation, one of their results states that NE is decidable for such games. Second, Chatterjee et al. [13] showed that the problem of deciding whether a (concurrent) stochastic game with reachability objectives has a Nash equilibrium in positional strategies with payoff at least $\overline{x}$ is NP-complete. We sharpen their hardness result by demonstrating that the problem remains NP-hard when it is restricted to games with only three players (as opposed to an unbounded number of players) where, additionally, payoffs are assigned at terminal vertices only (cf. Theorem 4.4).

2. Stochastic games

2.1. **Basic definitions.** Let us start by giving a formal definition of the game model that underlies this paper. The games we are interested in are played by multiple players taken from a finite set Π of players; we usually refer to them as player 0, player 1, player 2, and so on.

The *arena* of the game is basically a directed, coloured graph. Intuitively, the players take turns to form an infinite path through the arena, a *play*. Additionally, there is an element of chance involved: at some vertices, it is not a player who decides how to proceed but *nature* who chooses a successor vertex according to a probability distribution. To model this scenario, we partition the set V of vertices into sets V_i of vertices *controlled by player* $i \in \Pi$ and a set of *stochastic vertices*, and we extend the edge relation to a transition relation that takes probabilities into account. Formally, an arena for a game with players in Π consists of:

– a countable, non-empty set V of *vertices* or *states*,
– for each player i a set $V_i \subseteq V$ of vertices *controlled by player* i ,
– a *transition relation* $\Delta \subseteq V \times ([0,1] \cup \{\bot\}) \times V$, and
– a *colouring function* $\chi : V \to C$ into an arbitrary set C of colours.

We make the assumption that every vertex is controlled by at most one player: $V_i \cap V_j = \emptyset$ if $i \neq j$; vertices that are not controlled by a player are stochastic. For the sake of simplicity, we also assume that for each vertex v the set

$$v\Delta := \{w \in V : \text{exists } p \in (0,1] \cup \{\bot\} \text{ such that } (v,p,w) \in \Delta\}$$

of possible successor vertices is not empty. Moreover, we require that probabilities appear only on transitions originating in stochastic vertices, i.e. if $v \in \bigcup_{i\in\Pi} V_i$ and $(v,p,w) \in \Delta$ then $p = \bot$, and that they are unique: for every pair of a stochastic vertex v and an arbitrary vertex w there exists precisely one $p \in [0,1]$ such that $(v,p,w) \in \Delta$; we denote this probability by p_{vw}. For computational purposes, we assume that these probabilities are rational numbers. Finally, for each stochastic vertex the probabilities on outgoing transitions must sum up to 1: $\sum_{w\in V} p_{vw} = 1$.

The description of a game is completed by a specifying an *objective* for each player. On an abstract level, these are just arbitrary sets of infinite sequences of colours, i.e. subsets of C^ω. Since we want to assign a probability to them, we assume that objectives are Borel sets over the usual topology on infinite sequences, if not stated otherwise. Since objectives specify which plays are winning for a player, they are also called *winning conditions*.

In principal, we will identify an objective $\mathrm{Win} \subseteq C^\omega$ over colours with the corresponding objective $\chi^{-1}(\mathrm{Win}) := \{\pi \in V^\omega : \chi(\pi) \in \mathrm{Win}\} \subseteq V^\omega$ over vertices (which is also Borel since χ, as a mapping $V^\omega \to C^\omega$, is continuous). The reason that we allow objectives to refer to a colouring of the vertices is that the number of colours can be much smaller than the number of vertices, and it is possible that an objective can be represented more succinctly as an objective over colours rather than as an objective over vertices. In particular, this is true for *Muller* objectives, which we are going to introduce in Section 2.2.

If Π is a finite set of players, $(V, (V_i)_{i\in\Pi}, \Delta, \chi)$ is an arena and $(\mathrm{Win}_i)_{i\in\Pi}$ is a collection of objectives, we refer to the tuple $\mathcal{G} = (\Pi, V, (V_i)_{i\in\Pi}, \Delta, \chi, (\mathrm{Win}_i)_{i\in\Pi})$ as a *stochastic multiplayer game (SMG)*. A *play* of $\mathcal{G}$ is an infinite path through the arena of $\mathcal{G}$, i.e. an infinite sequence $\pi = \pi(0)\pi(1)\dots$ of vertices such that for each $k \in \mathbb{N}$ there exists $p \in (0,1] \cup \{\bot\}$ with $(\pi(k), p, \pi(k+1)) \in \Delta$. Finite prefixes of plays are called *histories*. A play or a history of $(\mathcal{G}, v_0)$ is a play respectively a history of $\mathcal{G}$ that starts in v_0. We say that a play π of $\mathcal{G}$ is *won by player i* if the corresponding sequence of colours fulfils player i's objective, i.e. $\chi(\pi) \in \mathrm{Win}_i$; the *payoff* of a play π is the vector $\overline{x} \in \{0,1\}^\Pi$ defined by $x_i = 1$ if and only if $\chi(\pi) \in \mathrm{Win}_i$.

Often, it is convenient to designate an initial vertex $v_0 \in V$; we denote the pair $(\mathcal{G}, v_0)$ an *initialised SMG*. A play or a history of an initialised SMG $(\mathcal{G}, v_0)$ is just a play respectively a history of $\mathcal{G}$ that starts in v_0. In the following, we will refer to both SMGs and initialised SMGs as SMGs; it should always be clear from the context whether the game is initialised or not.

SMGs generalise various stochastic models, each of them the subject of intensive research. First, there are *Markov chains*, the basic model for stochastic processes, in which no control is possible. These are just SMGs where the set Π of players is empty and (as a consequence) there are only stochastic vertices. If we extend Markov chains by a single controller, we arrive at the model of a *Markov decision process (MDP)*, a model introduced by Bellman [4] and heavily used in operations research. Formally, an MDP is an SMG where there is only one player (and only one objective). Finally, in a *(perfect-information) stochastic 2-player zero-sum game (S2G)*, there are only two players, player 0 and player 1, who have opposing objectives: one player wants to fulfil an objective while the other one wants to prevent her from doing so. Hence, one player's objective is the complement of the other player's objective. Due to their competitive nature, these games are also known as *competitive Markov decision processes* (see [26]).

The SMG model also incorporates several non-stochastic models. In particular, we call an SMG *deterministic* if it contains no stochastic vertices. In the 2-player zero-sum setting, the resulting model has found applications in logic and controller synthesis, to name a few.

2.2. **Objectives.** We have introduced objectives as abstract sets of infinite sequences. In order to be amenable for algorithmic solutions, we need to restrict to a class of objectives representable by finite objects. The objectives we consider for this purpose are standard in logic and verification (see [31]). In particular, whether an infinite sequence α fulfils such an objective only depends on the set $\mathrm{Occ}(\alpha)$ of colours occurring in α or on the set $\mathrm{Inf}(\alpha)$ of colours occurring *infinitely often* in α.

- A *reachability objective* is given by a set $F \subseteq C$ of *good* colours, and the objective requires to see a good colour at least once. The corresponding subset of C^ω is $\mathrm{Reach}(F) := \{\alpha \in C^\omega : \mathrm{Occ}(\alpha) \cap F \neq \emptyset\}$.
- A *Büchi objective* is again given by a set $F \subseteq C$ of good colours, but it requires to see a good colour infinitely often. The corresponding subset of C^ω is $\text{Büchi}(F) := \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \cap F \neq \emptyset\}$.
- A *co-Büchi objective* is also given by a set $F \subseteq C$ of good colours; this time, the objective requires to see only good colours *from some point onwards*. The corresponding subset of C^ω is $\text{coBüchi}(F) = \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \subseteq F\}$.
- A *parity objective* is given by a *priority function* $\Omega : C \to \{0, \dots, d\}$, $d \in \mathbb{N}$, which assigns to each colour a certain priority. The objective requires that the least priority

that occurs infinitely often is even. The corresponding subset of C^ω is $\mathrm{Parity}(\Omega) = \{\alpha \in C^\omega : \min(\mathrm{Inf}(\Omega(\alpha))) \text{ is even}\}$.
– A *Streett objective* is given by a set Ω of *Streett pairs* (F, G), $F, G \subseteq C$, and it requires that if a colour on the left-hand side of a pair is seen infinitely often then the same is true about a colour on the right-hand side. The corresponding subset of C^ω is $\mathrm{Streett}(\Omega) = \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \cap F = \emptyset \text{ or } \mathrm{Inf}(\alpha) \cap G \neq \emptyset \text{ for all } (F, G) \in \Omega\}$.
– A *Rabin objective* is given by a set Ω of *Rabin pairs* (F, G), $F, G \subseteq C$; it requires that for some pair a colour on the left-hand side is seen infinitely often while all colours on the right-hand side are seen only finitely often. The corresponding subset of C^ω is $\mathrm{Rabin}(\Omega) = \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \cap F \neq \emptyset \text{ and } \mathrm{Inf}(\alpha) \cap G = \emptyset \text{ for some } (F, G) \in \Omega\}$.
– A *Muller objective* is given by a family $\mathcal{F}$ of *accepting sets* $F \subseteq C$, and it requires that the sets of colours seen infinitely often equals one of these accepting sets. The corresponding subset of C^ω is $\mathrm{Muller}(\mathcal{F}) = \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \in \mathcal{F}\}$.

Parity, Streett, Rabin and Muller objectives are of particular relevance because they provide a standard form for arbitrary ω-regular objectives: any game with arbitrary ω-regular objectives can be reduced to one with parity, Streett, Rabin or Muller objectives (over a larger arena) by taking the product of its original arena with a suitable deterministic word automaton for each player's objective (see [52]).

In this work, for reasons that will become clear later, we are particularly attracted to objectives that are invariant under adding and removing finite prefixes; we call such objectives *prefix-independent*. More formally, an objective is prefix-independent if for each $\alpha \in C^\omega$ and $x \in C^*$ the sequence α satisfies the objective if and only if the sequence $x \cdot \alpha$ does. From the objectives listed above, only reachability objectives are, in general, not prefix-independent. However, many of our results (in particular, many of our lower bounds) apply to games with a prefix-independent form of reachability, which we call *simple reachability*. For these objectives, we assume that each vertex is coloured by itself, i.e. $C = V$, and χ is the identity mapping. The simple reachability objective for a set $F \subseteq V$ coincides with the reachability objective for F, but we require that each $v \in F$ is a *terminal vertex*: $v\Delta = \{v\}$. For any such set F, we have $\mathrm{Occ}(\pi) \cap F \neq \emptyset$ if and only if $\mathrm{Inf}(\pi) \cap F \neq \emptyset$ (or equivalently, $\mathrm{Inf}(\pi) \subseteq F$) for every play π. Hence, simple reachability objectives can be regarded as prefix-independent objectives.

For S2Gs, the distinction between reachability and simple reachability is not important: every S2G with a reachability objective can easily be transformed into an equivalent S2G with a simple reachability objective. For SMGs, we believe that any such transformation requires exponential time: Deciding whether in a deterministic game with simple reachability objectives there exists a play that fulfils each of the objectives can be done in polynomial time whereas the same problem is NP-complete for deterministic games with standard reachability objectives [13, 53].

The resulting hierarchy of objectives is depicted in Figure 1: As explained above, a simple reachability objective can be considered as a (co-)Büchi objective; a Büchi objective can be translated to a parity objective with only two priorities, and any parity objective can be translated to both a Streett and a Rabin objective; in fact, the intersection (union) of any two parity objectives can be represented as a Streett (Rabin) objective. Moreover, any Streett or Rabin objective can be represented as a Muller objective; however, the translation from a set of Streett/Rabin pairs to an equivalent family of accepting sets is, in general, exponential. Finally, note that the complement of a Büchi (Streett) objective is a co-Büchi (Rabin) objective, and vice versa, whereas the class of parity objectives and the

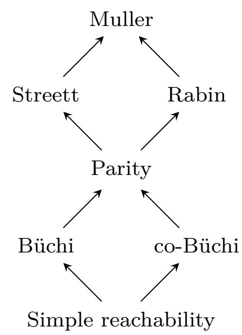


FIGURE 1. A hierarchy of prefix-independent objectives.

class of Muller objectives are closed under complementation. In fact, an objective can be represented as both a Streett and a Rabin objective if and only if it can be represented as a parity objective [59].

To denote the class of SMGs (S2Gs) with a certain type of objectives, we prefix the name *SMG (S2G)* with the name(s) of the objective; for instance, we use the term *Streett-Rabin SMG* to denote SMGs where each player has a Streett or a Rabin objective. For S2Gs, we adopt the convention to name the objective of player 0 first; hence, in a Streett-Rabin S2G player 0 has a Streett objective while player 1 has a Rabin objective. Inspired by Condon [16], we will refer to SMGs with simple reachability objectives and S2Gs with a (simple) reachability objective for player 0 as *simple stochastic multiplayer games (SSMGs)* and *simple stochastic 2-player zero-sum games (SS2Gs)*, respectively.

Drawing an SMG. When drawing an SMG as a graph, we will use the following conventions: The initial vertex is marked by an incoming edge that has no source vertex. Vertices that are controlled by a player are depicted as circles, where the player who controls a vertex is given by the label next to it. Stochastic vertices are depicted as diamonds, where the transition probabilities are given by the labels on its outgoing edges (the default being $\frac{1}{2}$). Finally, terminal vertices are generally represented by their associated payoff vector. In fact, we allow arbitrary vectors of rational probabilities as payoffs. This does not increase the power of the model since such a payoff vector can easily be realised by an SSMG consisting of stochastic and terminal vertices only.

2.3. **Strategies and strategy profiles.**

2.3.1. *Randomised and pure strategies.* The notion of a strategy lies at the heart of game theory. Formally, a *(randomised) strategy of player i in an SMG $\mathcal{G}$* is a mapping $\sigma : V^*V_i \to \mathcal{D}(V)$ assigning to each possible sequence $xv \in V^*V_i$ of vertices ending in a vertex controlled by player i a (discrete) probability distribution over V such that $\sigma(xv)(w) > 0$ only if $(v, \bot, w) \in \Delta$. Instead of $\sigma(xv)(w)$, we usually write $\sigma(w \mid xv)$. We say that a play π of $\mathcal{G}$ is *consistent* with a strategy σ of player i if $\sigma(\pi(k+1) \mid \pi(0)\dots\pi(k)) > 0$ for all $k \in \mathbb{N}$ with $\pi(k) \in V_i$. Similarly, a history $x = v_0 \dots v_n$ is consistent with σ if $\sigma(v_{k+1} \mid v_0 \dots v_k) > 0$ for all $0 \le k < n$.

A *(randomised) strategy profile of $\mathcal{G}$* is a tuple $\bar{\sigma} = (\sigma_i)_{i\in\Pi}$ where σ_i is a strategy of player i in $\mathcal{G}$. We say that a play or a history of $\mathcal{G}$ is consistent with a strategy profile $\bar{\sigma}$

if it is consistent with each σ_i. Given a strategy profile $\overline{\sigma} = (\sigma_j)_{j\in\Pi}$ and a strategy τ of player i, we denote by $(\overline{\sigma}_{-i}, \tau)$ the strategy profile resulting from $\overline{\sigma}$ by replacing σ_i with τ.

A strategy σ of player i is called *pure* or *deterministic* if for each $xv \in V^*V_i$ there exists $w \in v\Delta$ with $\sigma(w \mid xv) = 1$; note that a pure strategy of player i can be identified with a function $\sigma : V^*V_i \to V$. A strategy profile $\overline{\sigma} = (\sigma_i)_{i\in\Pi}$ is called *pure* if each σ_i is pure.

2.3.2. *The probability space induced by a strategy profile.* Given an initial vertex v_0 and a strategy profile $\overline{\sigma} = (\sigma_i)_{i\in\Pi}$, the *conditional probability of $w \in V$ given $xv \in V^*V$* is the number $\sigma_i(w \mid xv)$ if $v \in V_i$ and the unique $p \in [0,1]$ such that $(v,p,w) \in \Delta$ if v is a stochastic vertex; let us denote this probability by $\overline{\sigma}(w \mid xv)$. The probabilities $\overline{\sigma}(w \mid xv)$ induce a probability measure on the space V^ω in the following way: The probability of a basic open set $v_1 \dots v_k \cdot V^\omega$ is 0 if $v_1 \neq v_0$ and the product of the probabilities $\overline{\sigma}(v_j \mid v_1 \dots v_{j-1})$ for $j = 2, \dots, k$ otherwise. It is a classical result of measure theory that this extends to a unique probability measure assigning a probability to every Borel subset of V^ω, which we denote by $\Pr^{\overline{\sigma}}_{v_0}$.

For a strategy profile $\overline{\sigma}$, we are mainly interested in the probabilities $p_i := \Pr^{\overline{\sigma}}_{v_0}(\mathrm{Win}_i)$ of winning. We call p_i the *(expected) payoff of $\overline{\sigma}$ for player i (from v_0)* and the vector $(p_i)_{i\in\Pi}$ the *(expected) payoff of $\overline{\sigma}$ (from v_0)*. Note that, if $\overline{\sigma}$ is a pure strategy profile of a deterministic game, then its payoff is just the payoff of the the unique play π of $(\mathcal{G}, v_0)$ that is consistent with $\overline{\sigma}$.

In order to apply known results about Markov chains, we can also view the stochastic process induced by a strategy profile $\overline{\sigma}$ as a countable Markov chain $\mathcal{G}^{\overline{\sigma}}$, defined as follows: The set of vertices (states) of $\mathcal{G}^{\overline{\sigma}}$ is the set V^+ of all non-empty sequences of vertices in $\mathcal{G}$. The only transitions from a state xv, $x \in V^*$, $v \in V$, are to states of the form xvw, $w \in V$, and such a transition occurs with probability $p > 0$ if and only if either v is stochastic and $(v,p,w) \in \Delta$ or $v \in V_i$ and $\sigma_i(w \mid xv) = p$. Finally, the colouring χ of vertices is extended to a colouring of states by setting $\chi(xv) = \chi(v)$ for all $x \in V^*$, $v \in V$. With this definition, we can recover the payoff of $\overline{\sigma}$ for player i as the probability of the event $\chi^{-1}(\mathrm{Win}_i)$ in $(\mathcal{G}^{\overline{\sigma}}, v_0)$.

For each player i, the Markov decision process $\mathcal{G}^{\overline{\sigma}_{-i}}$ is defined just as $\mathcal{G}^{\overline{\sigma}}$, but states $xv \in V^*V_i$ are controlled by player i (the unique player in $\mathcal{G}^{\overline{\sigma}_{-i}}$), and there is a transition from such a state to any state of the form xvw, $w \in V$, such that $(v, \bot, w) \in \Delta$; player i's objective is the same as in $\mathcal{G}$.

2.3.3. *Strategies with memory.* A *memory structure* for a game $\mathcal{G}$ with vertices in V is a triple $\mathfrak{M} = (M, \delta, m_0)$ where M is a set of *memory states*, $\delta : M \times V \to M$ is the *update function*, and $m_0 \in M$ is the *initial memory*. A *(randomised) strategy with memory* $\mathfrak{M}$ of player i is a function $\sigma : M \times V_i \to \mathcal{D}(V)$ such that $\sigma(m,v)(w) > 0$ only if $w \in vE$. The strategy σ is a *pure strategy with memory* $\mathfrak{M}$ if additionally the following property holds: for all $m \in M, v \in V$ there exists $w \in V$ such that $\sigma(m,v)(w) = 1$. Hence, a pure strategy with memory $\mathfrak{M}$ can be described by a function $\sigma : M \times V_i \to V$. Finally, a *(pure) strategy profile with memory* $\mathfrak{M}$ is a tuple $\overline{\sigma} = (\sigma_i)_{i\in\Pi}$ such that each σ_i is a (pure) strategy with memory $\mathfrak{M}$ of player i.

A (pure) strategy σ with memory $\mathfrak{M}$ of player i defines a (pure) strategy of player i (in the usual sense) as follows: Let $\delta^*(x)$ be the memory state after $x \in V^*$, defined inductively by $\delta^*(\varepsilon) = m_0$ and $\delta^*(xv) = \delta(\delta^*(x), v)$ for $x \in V^*, v \in V$. If $v \in V_i$, then the distribution (successor vertex) chosen by the strategy σ for the sequence xv is $\sigma(\delta^*(x), v)$. Vice versa, every strategy (profile) of $\mathcal{G}$ can be viewed as a strategy (profile) with memory $\mathfrak{M} := (V^*, \cdot, \varepsilon)$.

A *finite-state strategy (profile)* is a strategy (profile) with memory $\mathfrak{M}$ for a finite memory structure $\mathfrak{M}$. Note that a strategy profile is finite-state if and only if each of its strategies is finite-state because if σ_i is a strategy with memory $\mathfrak{M}_i$ of player i then $\overline{\sigma} := (\sigma_i)_{i\in\Pi}$ can be viewed as a strategy profile with memory $\prod_{i\in\Pi} \mathfrak{M}_i$.

If $|M| = 1$, we call a strategy (profile) with memory $\mathfrak{M}$ *stationary*. Moreover, we call a pure stationary strategy (profile) a *positional* strategy (profile). A stationary strategy of player i can be described by a function $\sigma : V_i \to \mathcal{D}(V)$, and a positional strategy even by a function $\sigma : V_i \to V$.

If $\overline{\sigma} = (\sigma_i)_{i\in\Pi}$ is a strategy profile with memory $\mathfrak{M}$, we refine the Markov chain $\mathcal{G}^{\overline{\sigma}}$ by taking $M \times V$ as its domain. The transition relation is defined as follows: There is a transition from (m, v) to (n, w) with probability $p > 0$ if and only if $\delta(m, v) = n$ and either v is a stochastic vertex of $\mathcal{G}$ and $(v, p, w) \in \Delta$ or $v \in V_i$ and $\sigma_i(m, v)(w) = p$. Finally, a state (m, v) is coloured with the same colour as the vertex v in $\mathcal{G}$. Analogously, we can refine the Markov decision process $\mathcal{G}^{\overline{\sigma}_{-i}}$ by using $M \times V$ as its domain. In $\mathcal{G}^{\overline{\sigma}_{-i}}$, vertices $(m, v) \in M \times V_i$ are controlled by player i, and there is a transition from such a vertex (m, v) to $(n, w) \in M \times V$ if and only if $n = \delta(m, v)$ and $(v, \perp, w) \in \Delta$. Note that the arenas of both $\mathcal{G}^{\overline{\sigma}}$ and $\mathcal{G}^{\overline{\sigma}_{-i}}$ are finite if the memory M and the original arena of $\mathcal{G}$ are finite.

2.3.4. *Residual games and strategies.* Given an SMG $\mathcal{G}$ and a sequence $x \in V^*$ (which will usually be a history), the *residual game* $\mathcal{G}[x]$ has the same arena as $\mathcal{G}$ but different objectives: if $\mathrm{Win}_i \subseteq V^\omega$ is the objective of player i in $\mathcal{G}$, then her objective in $\mathcal{G}[x]$ is $x^{-1}\mathrm{Win}_i := \{\pi \in V^\omega : x\pi \in \mathrm{Win}_i\}$. In particular, if all objectives in $\mathcal{G}$ are prefix-independent, then $\mathcal{G}[x] = \mathcal{G}$.

If player i plays according to a strategy σ in $\mathcal{G}$, then the natural choice for her strategy in $\mathcal{G}[x]$ is the *residual strategy* $\sigma[x]$, defined by $\sigma[x](yv) = \sigma(xyv)$. If $\overline{\sigma} = (\sigma_i)_{i\in\Pi}$ is a strategy profile, then the *residual strategy profile* $\overline{\sigma}[x]$ is just the profile of the residual strategies $\sigma_i[x]$. The following lemma, taken from [58], shows how to compute probabilities with respect to a residual strategy profile.

Lemma 2.1. Let $\overline{\sigma}$ be any strategy profile of $\mathcal{G}$, $xv \in V^*V$ a history of $\mathcal{G}$, and $X \subseteq V^\omega$ a Borel set. Then $\Pr^{\overline{\sigma}}_{v_0}(X \cap xv \cdot V^\omega) = \Pr^{\overline{\sigma}}_{v_0}(xv \cdot V^\omega) \cdot \Pr^{\overline{\sigma}[x]}_v(x^{-1}X)$.

2.4. **Subarenas and end components.** Algorithms for stochastic games often employ a *divide-and-conquer* approach and compute a solution for a complex game from the solution of several smaller games. These smaller games are usually obtained from the original game by restricting it to a *subarena*. Formally, given an SMG $\mathcal{G}$, a set $U \subseteq V$ a subarena if

- $U \neq \emptyset$,
- $v\Delta \cap U \neq \emptyset$ for each $v \in U$, and
- $v\Delta \subseteq U$ for each stochastic vertex $v \in U$.

Clearly, if U is a subarena, then the restriction of $\mathcal{G}$ to vertices in U is again an SMG, which we denote by $\mathcal{G} \upharpoonright U$. Formally,

$$\mathcal{G} \upharpoonright U := (\Pi, U, (V_i \cap U)_{i\in\Pi}, \Delta \cap (U \times ([0,1] \cup \{\perp\}) \times U), (\mathrm{Win}_i \cap U^\omega)_{i\in\Pi}).$$

Of particular interest are the strongly connected subarenas of a game because they are the obvious candidates for the sets $\mathrm{Inf}(\pi)$ of vertices visited infinitely often in a play; we call these sets *end components*. Formally, a set $U \subseteq V$ is an end component if U is a subarena and every vertex $w \in U$ is *reachable* from every other vertex $v \in U$, i.e. there exists a sequence $v = v_1, v_2, \ldots, v_n = w$, $n \geq 1$, such that $v_{i+1} \in v_i\Delta$ for each $0 < i < n$. An end component U is *maximal in a set $S \subseteq V$* if there is no end component U' such that $U \subsetneq U' \subseteq S$. For any finite subset $S \subseteq V$, the set of all end components maximal in S can be computed by standard graph algorithms in quadratic time (see [20]).

The theory of end components has been developed by de Alfaro [20] and Courcoubetis and Yannakakis [18]. The central fact about end components in SMGs with finite arena is that, under *any* strategy profile, the set of vertices visited infinitely often is almost surely an end component.

Lemma 2.2. Let $\mathcal{G}$ be any SMG with finite arena, and let $\overline{\sigma}$ be any strategy profile of $\mathcal{G}$. Then $\mathrm{Pr}_v^{\overline{\sigma}}(\{\alpha \in V^\omega : \mathrm{Inf}(\alpha) \text{ is an end component}\}) = 1$ for every vertex $v \in V$.

Moreover, for any end component U, we can construct a stationary strategy profile, or alternatively a pure finite-state strategy profile, that, when started in U, guarantees to visit all (and only) vertices in U infinitely often. In fact, the stationary profile that ensures this just chooses for each vertex in U a successor in U uniformly at random.

Lemma 2.3. Let $\mathcal{G}$ be any SMG with finite arena, and let U be any end component of $\mathcal{G}$. There exists both a stationary and a pure finite-state strategy profile $\overline{\sigma}$ of $\mathcal{G} \restriction U$ such that $\mathrm{Pr}_v^{\overline{\sigma}}(\{\alpha \in V^\omega : \mathrm{Inf}(\alpha) = U\}) = 1$ for every vertex $v \in U$.

Given an SMG $\mathcal{G}$ with (objectives representable as) Muller objectives $\mathcal{F}_i \subseteq \mathcal{P}(C)$, we say that an end component C is *winning for player i* if $\chi(C) \in \mathcal{F}_i$; the *payoff* of C is the the vector $\overline{x} \in \{0,1\}^\Pi$, defined by $x_i = 1$ if C is winning for player i.

2.5. **Values, determinacy and optimal strategies.** The notions of the value and an *optimal strategy* are central for the theory of 2-player zero-sum games. However, they can also be applied to SMGs.

Given a strategy τ of player i in $\mathcal{G}$ and a vertex $v \in V$, the *value of τ from v* is the number $\mathrm{val}^\tau(v) := \inf_{\overline{\sigma}} \mathrm{Pr}_v^{\overline{\sigma}_{-i},\tau}(\mathrm{Win}_i)$, where $\overline{\sigma}$ ranges over all strategy profiles of $\mathcal{G}$. Moreover, the *value of $\mathcal{G}$ for player i from v* is the supremum of these values: $\mathrm{val}_i^{\mathcal{G}}(v) := \sup_\tau \mathrm{val}^\tau(v)$, where τ ranges over all strategies of player i in $\mathcal{G}$. Intuitively, $\mathrm{val}_i^{\mathcal{G}}(v)$ is the maximal payoff that player i can *ensure* when the game starts from v.

Given an initial vertex $v_0 \in V$, a strategy τ of player i in $\mathcal{G}$ is called *(almost-surely) winning* if $\mathrm{val}^\tau(v_0) = 1$. More generally, τ is called *optimal* if $\mathrm{val}^\tau(v_0) = \mathrm{val}_i^{\mathcal{G}}(v_0)$. For $\varepsilon > 0$, it is called *ε-optimal* if $\mathrm{val}^\tau(v_0) \geq \mathrm{val}_i^{\mathcal{G}}(v_0) - \varepsilon$. A *globally ($\varepsilon$-)optimal* strategy is a strategy that is (ε-)optimal for every possible initial vertex $v_0 \in V$. Note that optimal strategies do not need to exist since the supremum in the definition of $\mathrm{val}_i^{\mathcal{G}}$ is not necessarily attained; in this case, only ε-optimal strategies do exist. However, if for every possible initial vertex there exists an (ε-)optimal strategy, then there also exists a globally (ε-)optimal strategy. Finally, we say that a strategy τ of player i in $(\mathcal{G}, v_0)$ is *residually optimal* if the residual strategy $\tau[x]$ is optimal in the residual game $(\mathcal{G}[x], v)$ for every history xv of $(\mathcal{G}, v_0)$.

Determining values and finding optimal strategies in SMGs actually reduces to performing the same tasks in S2Gs. Formally, given an SMG $\mathcal{G}$, define for each player i the *coalition game* $\mathcal{G}_i$ to be the same game as $\mathcal{G}$ but with only two players, player i and player $\Pi \setminus \{i\}$:

Player $\Pi \setminus \{i\}$ controls all vertices that in $\mathcal{G}$ are controlled by some player $j \neq i$, and her objective is the complement of player i's objective in $\mathcal{G}$. Clearly, $\mathcal{G}_i$ is an S2G, and $\mathrm{val}^{\mathcal{G}_i}(v) = \mathrm{val}_i^{\mathcal{G}}(v)$ for every vertex v. Moreover, any (residually, ε-) optimal strategy for player i in $(\mathcal{G}_i, v_0)$ is (residually, ε-) optimal in $(\mathcal{G}, v_0)$, and vice versa. Hence, when we study values and optimal strategies, we can restrict to 2-player zero-sum games.

A celebrated theorem due to Martin [45] and Maitra and Sudderth [44] states that S2Gs with Borel objectives are *determined*: $\mathrm{val}_0^{\mathcal{G}} = 1 - \mathrm{val}_1^{\mathcal{G}}$ (where the equality holds pointwise).[1] The number $\mathrm{val}^{\mathcal{G}}(v) := \mathrm{val}_0^{\mathcal{G}}(v)$ is consequently called the *value of $\mathcal{G}$ from v*. In fact, an inspection of the proof shows that — for the kind of games we study in this paper — both players do not only have randomised ε-optimal strategies but pure ones.

Theorem 2.4 ([45, 44])**.** Every stochastic 2-player zero-sum game with Borel objectives is determined; for all $\varepsilon > 0$, both players have ε-optimal pure strategies.

For S2Gs with prefix-independent objectives played on finite arenas, Gimbert and Horn [30] showed a stronger result than Theorem 2.4: in these games, both players not only have ε-optimal pure strategies, but optimal ones [30]. In fact, their proof reveals not only the existence of optimal strategies but the existence of residually optimal strategies [37].

Theorem 2.5 ([30])**.** In any S2G with finite arena and prefix-independent objectives, both players have residually optimal pure strategies.

For S2Gs with ω-regular objectives played on finite arenas, even nicer strategies than arbitrary pure strategies suffice for optimality. In particular, in any Rabin-Streett S2G with finite arena there exists a globally optimal positional strategy for player 0, which is, by definition, also residually optimal [42, 14].

Theorem 2.6 ([42, 14])**.** In any Rabin-Streett S2G with finite arena, player 0 has a globally optimal positional strategy.

A consequence of Theorem 2.6 is that the values of any Rabin-Streett S2G with finite arena are rational of bit complexity polynomial in the size of the arena: Given a positional strategy profile $\overline{\sigma}$ of $\mathcal{G}$, the finite MDP $\mathcal{G}^{\overline{\sigma}_{-1}}$ is not larger than the game $\mathcal{G}$. Moreover, if σ_0 is globally optimal, then for every vertex v the value of $\mathcal{G}$ from v and the value of $\mathcal{G}^{\overline{\sigma}_{-1}}$ from v sum up to 1. But the values of any Streett MDP form the optimal solution of a linear programme of polynomial size in the given MDP (see [20]) and are therefore rational of small bit complexity.

Of course, it also follows from Theorem 2.6 that parity S2Gs with finite arena are *positionally determined*: both players have globally optimal positional strategies. This result was first proven for deterministic games (even for games with infinite arena) independently by Emerson and Jutla [23] and Mostowski [48]. For SS2Gs, the existence of optimal positional strategies follows from a result of Liggett and Lippman [43]. Finally, McIver and Morgan [46], Chatterjee et al. [12] and Zielonka [58] extended both results to parity S2Gs.

Corollary 2.7. In any parity S2G with finite arena, both players have globally optimal positional strategies.

Since any S2G with finite arena and ω-regular objectives can be reduced to one with finite arena and parity objectives, we can conclude from Corollary 2.7 that both players

[1] Martin proved the theorem originally for *Blackwell games*; Maitra and Sudderth adapted his proof to stochastic games.

have residually optimal pure finite-state strategies in S2Gs with finite arena and arbitrary ω-regular objectives.

Corollary 2.8. In any finite S2G with finite arena and ω-regular objectives, both players have residually optimal pure finite-state strategies.

2.6. **Algorithmic problems.** For the rest of this section, we only consider 2-player zero-sum games played on finite arenas. The main computational problems for these games are computing the value and optimal strategies for one or both players, if they exist. Instead of computing the value exactly, we can ask whether the value is greater than some given rational probability p, a problem which we call the *quantitative decision problem for S2Gs* :

> Given a S2G $\mathcal{G}$, a vertex v and a rational number $p \in [0, 1]$, decide whether $\mathrm{val}^{\mathcal{G}}(v) \geq p$.

In many cases, it suffices to know whether the value is 1, i.e. whether player 0 has a strategy to win the game almost surely (in the limit, at least). We call the resulting decision problem the *qualitative decision problem for S2Gs* .

Clearly, if we can solve the quantitative decision problem, we can approximate the values $\mathrm{val}^{\mathcal{G}}(v)$ up to any desired precision by using binary search. In fact, for parity S2Gs it is well-known that it suffices to solve the decision problems, since the other problems (computing the values and optimal strategies) are polynomial-time equivalent to the quantitative decision problem (with respect to Turing reductions).

For a Markov decision process whose objective can be represented as a Muller objective, we can compute the values by an analysis of its end components: For a given initial vertex v, the value of the MDP from v is the maximal probability of reaching a winning end component from v. Once the vertices that reside in winning end components have been identified, these probabilities can be computed in polynomial time via linear programming.

For MDPs with Rabin or Muller objectives, it is easy to see that the union of all winning end components can be computed in polynomial time (see [20]); for MDPs with Streett objectives, Chatterjee et al. [14] gave a polynomial-time algorithm for computing this set. Hence, for MDPs with any of these objectives, the quantitative decision problem is solvable in polynomial time.

Theorem 2.9 ([20, 14])**.** The quantitative decision problem for Streett, Rabin or Muller MDPs is in P.

It follows from Theorems 2.6 and 2.9 that the quantitative decision problem for Rabin-Streett S2Gs is in NP: To decide whether $\mathrm{val}^{\mathcal{G}}(v) \geq p$, it suffices to guess a positional strategy for player 0 and to check whether in the resulting Streett MDP the value from v is $\geq p$. By determinacy, this result implies that the quantitative decision problem is in coNP for Streett-Rabin S2Gs, and in NP $\cap$ coNP for parity S2Gs.

Corollary 2.10. The quantitative decision problem for S2Gs is in

- NP for Rabin-Streett S2Gs,
- coNP for Streett-Rabin S2Gs, and
- NP $\cap$ coNP for parity S2Gs.

A corresponding NP-hardness result for Rabin-Streett S2Gs has been established by Emerson and Jutla [22], even for deterministic games. In particular, this hardness result

	Qualitative	Quantitative
SS2Gs	P-complete	NP $\cap$ coNP
Parity[d]	P-complete	NP $\cap$ coNP
Parity	NP $\cap$ coNP	NP $\cap$ coNP
Rabin-Streett	NP-complete	NP-complete
Streett-Rabin	coNP-complete	coNP-complete
Muller	PSPACE-complete	PSPACE-complete

TABLE 1. The complexity of deciding the value in S2Gs.

also holds for the qualitative decision problem. Moreover, by determinacy, this result can be turned into a coNP-hardness result for (deterministic) Streett-Rabin S2Gs.

For S2Gs with Muller objectives, Chatterjee [10] showed that the quantitative decision problem falls into PSPACE; for deterministic games, a polynomial-space algorithm had been given earlier by McNaughton [47]. A matching lower bound was provided by Hunter and Dawar [38], again even for deterministic games.

Theorem 2.11 ([10, 38])**.** The quantitative and the qualitative decision problem for Muller S2Gs are PSPACE-complete.

Theorem 2.11 does not hold if the Muller objectives are given by a family of subsets of *vertices*: Horn [36, 35] showed that the qualitative decision problem for *explicit* Muller S2Gs is in P, and that the quantitative problem is in NP $\cap$ coNP.

Another class of S2Gs for which the qualitative decision problem is in P is, for each $d \in \mathbb{N}$, the class Parity[d] of all parity S2Gs whose priority function refers to at most d priorities [21]. In particular, the qualitative decision problem for SS2Gs as well as (co-)Büchi S2Gs is in P.

Theorem 2.12 ([21])**.** For each $d \in \mathbb{N}$, the qualitative decision problem for parity S2Gs with at most d priorities is in P.

Table 1 summarises the results about the complexity of the quantitative and the qualitative decision problem for S2Gs: P-hardness (via LOGSPACE-reductions) for all these problems follows from the fact that *and-or graph reachability* is P-complete [39].

The results summarised in Table 1 leave open the possibility that at least one of the following problems is decidable in polynomial time:

1. the qualitative decision problem for parity S2Gs,
2. the quantitative decision problem for SS2Gs,
3. the quantitative decision problem for parity S2Gs.

Note that, given that all of them are contained in both NP and coNP, it is unlikely that one of them is NP- or coNP-hard (because if it were, then NP would equal coNP and the polynomial hierarchy would collapse to its first level).

For the first problem, Chatterjee et al. [11] gave a polynomial-time reduction from the qualitative decision problem for stochastic 2-player zero-sum parity games. Hence, solving the qualitative decision problem for parity S2Gs is not harder than deciding which of the two players has a winning strategy in a deterministic (2-player zero-sum) parity game. Whether the latter problem is decidable in polynomial time is a long-standing open problem: Several years after Emerson and Jutla [23] put the problem into NP$\cap$coNP, Jurdziński [40] improved this bound slightly to UP $\cap$ coUP. Together with Paterson and Zwick [41], he also gave an algorithm that solves these games in subexponential time; a *randomised* subexponential algorithm had been given earlier by Björklund et al. [6]. On the other hand, Friedmann [27] recently showed that the most promising candidate for a polynomial-time algorithm so far, the *discrete strategy improvement* algorithm due to Vöge and Jurdziński [56], requires exponential time in the worst case.

Regarding the second problem, only some progress towards a polynomial-time algorithm has been made since Condon [16] showed that the problem falls into NP$\cap$coNP: For example, Björklund and Vorobyov [5] gave a randomised subexponential algorithm for solving SS2Gs, and Gimbert and Horn [29] showed that the quantitative decision problem for SS2Gs is fixed-parameter tractable with respect to the number of stochastic vertices as parameter.

For the third problem, Andersson and Miltersen [2] recently established a polynomial-time Turing reduction to the second. Hence, there exists a polynomial-time algorithm for 2. if and only if there exists one for 3. In particular, a polynomial-time algorithm for 2. would also give a polynomial-time algorithm for 1. However, to the best of our knowledge, it is plausible that the qualitative decision problem for parity S2Gs is in P while the quantitative decision problem for SS2Gs is not.

3. Existence of Nash equilibria

To capture rational behaviour of (selfish) players, John Nash [49] introduced the notion of, what is now called, a Nash equilibrium.

Definition 3.1. Given a strategy profile $\overline{\sigma}$ of a game $(\mathcal{G}, v_0)$, a strategy τ of player i in $\mathcal{G}$ is called a *best response to* $\overline{\sigma}$ if τ maximises the expected payoff of player i: $\Pr_{v_0}^{\overline{\sigma}_{-i},\tau'}(\mathrm{Win}_i) \leq \Pr_{v_0}^{\overline{\sigma}_{-i},\tau}(\mathrm{Win}_i)$ for all strategies τ' of player i. A strategy profile $\overline{\sigma} = (\sigma_i)_{i\in\Pi}$ is a *Nash equilibrium* if each σ_i is a best response to $\overline{\sigma}$.

In a Nash equilibrium, no player can improve her payoff by (unilaterally) switching to a different strategy. In fact, it suffices if no player can gain from switching to a pure strategy.

Proposition 3.2. A strategy profile $\overline{\sigma}$ of a game $(\mathcal{G}, v_0)$ is a Nash equilibrium if and only if, for each player i and for each *pure* strategy τ of player i in $\mathcal{G}$, $\Pr_{v_0}^{\overline{\sigma}_{-i},\tau}(\mathrm{Win}_i) \leq \Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i)$.

Proof. Clearly, if $\overline{\sigma}$ is a Nash equilibrium, then $\Pr_{v_0}^{\overline{\sigma}_{-i},\tau}(\mathrm{Win}_i) \leq \Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i)$ for each *pure* strategy τ of player i in $\mathcal{G}$. Now, assume that $\overline{\sigma}$ is not a Nash equilibrium. Hence, $p := \sup_\tau \Pr_{v_0}^{\overline{\sigma}_{-i},\tau}(\mathrm{Win}_i) = \Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i) + \varepsilon$ for some player i and some $\varepsilon > 0$. Consider the Markov decision process $\mathcal{G}^{\overline{\sigma}_{-i}}$: Clearly, the value of $\mathcal{G}^{\overline{\sigma}_{-i}}$ from v_0 equals p. By Theorem 2.4, there exists an $\frac{\varepsilon}{2}$-optimal pure strategy τ in $\mathcal{G}^{\overline{\sigma}_{-i}}$. Since the arena of $\mathcal{G}^{\overline{\sigma}_{-i}}$ is a forest, we can assume that τ is a positional strategy, which can be viewed as a pure strategy in $\mathcal{G}$. We have: $\Pr_{v_0}^{\overline{\sigma}_{-i},\tau}(\mathrm{Win}_i) \geq p - \frac{\varepsilon}{2} > p - \varepsilon = \Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i)$. ∎

For 2-player zero-sum games, a Nash equilibrium is nothing else than a pair of optimal strategies.

Proposition 3.3. Let $(\mathcal{G}, v_0)$ be a 2-player zero-sum game. A strategy profile (σ, τ) of $(\mathcal{G}, v_0)$ is a Nash equilibrium if and only if both σ and τ are optimal. In particular, every Nash equilibrium of $(\mathcal{G}, v_0)$ has payoff $(\mathrm{val}^{\mathcal{G}}(v_0), 1 - \mathrm{val}^{\mathcal{G}}(v_0))$.

Proof. ($\Rightarrow$) Assume that both σ and τ are optimal, but that (σ, τ) is not a Nash equilibrium. Hence, one of the players, say player 1, can improve her payoff by playing some strategy τ'. Hence, $\mathrm{val}^{\mathcal{G}}(v_0) = \mathrm{Pr}_{v_0}^{\sigma,\tau}(\mathrm{Win}_0) > \mathrm{Pr}_{v_0}^{\sigma,\tau'}(\mathrm{Win}_0)$. However, since σ is optimal, it must also be the case that $\mathrm{val}^{\mathcal{G}}(v_0) \leq \mathrm{Pr}_{v_0}^{\sigma,\tau'}(\mathrm{Win}_0)$, a contradiction. The reasoning in the case that player 0 can improve is analogous.

($\Leftarrow$) Let (σ, τ) be a Nash equilibrium of $(\mathcal{G}, v_0)$, and let us first assume that σ is not optimal, i.e. $\mathrm{val}^{\sigma}(v_0) < \mathrm{val}^{\mathcal{G}}(v_0)$. By the definition of $\mathrm{val}^{\mathcal{G}}$, there exists another strategy σ' of player 0 such that $\mathrm{val}^{\sigma}(v_0) < \mathrm{val}^{\sigma'}(v_0) \leq \mathrm{val}^{\mathcal{G}}(v_0)$. Moreover, since (σ, τ) is a Nash equilibrium:

$$\mathrm{Pr}_{v_0}^{\sigma,\tau}(\mathrm{Win}_0) \leq \mathrm{val}^{\sigma}(v_0) < \mathrm{val}^{\sigma'}(v_0) = \inf_{\tau'} \mathrm{Pr}_{v_0}^{\sigma',\tau'}(\mathrm{Win}_0) \leq \mathrm{Pr}_{v_0}^{\sigma',\tau}(\mathrm{Win}_0).$$

Thus, player 0 can improve her payoff by playing σ' instead of σ, a contradiction to the fact that (σ, τ) is a Nash equilibrium. The argumentation in the case that τ is not optimal is analogous. □

It follows from Proposition 3.3 that every 2-player zero-sum stochastic game with prefix-independent objectives has a Nash equilibrium in pure strategies. The question arises whether this is still true if the 2-player zero-sum assumption is relaxed.

Clearly, a strategy profile $\overline{\sigma}$ can only be a Nash equilibrium if for no player i there exists a history xv, consistent with $\overline{\sigma}$, such that $\mathrm{Pr}_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \mid xv \cdot V^*) < \mathrm{val}_i^{\mathcal{G}}(v)$ because otherwise player i could improve her payoff by switching to an optimal strategy after history xv. The next lemma shows that the we can turn every strategy profile that fulfils this property into a Nash equilibrium. The proof uses so-called *threat* (or *trigger*) strategies, which are added on top of the given strategy profile: any player threatens to change her behaviour when one of the other players deviates from the prescribed strategy. Before having been applied to infinite-duration games, this concept has proven fruitful in the (related) area of *repeated games* (see [51, Chapter 8] and [3]).

Lemma 3.4. Let $(\mathcal{G}, v_0)$ be any SMG with finite arena and prefix-independent objectives. If $\overline{\sigma}$ is a pure strategy profile such that $\mathrm{Pr}_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \mid xv \cdot V^{\omega}) \geq \mathrm{val}_i^{\mathcal{G}}(v)$ for each player i and for each history xv of $(\mathcal{G}, v_0)$ that is consistent with $\overline{\sigma}$, then $(\mathcal{G}, v_0)$ has a pure Nash equilibrium $\overline{\sigma}^*$ with $\mathrm{Pr}_{v_0}^{\overline{\sigma}} = \mathrm{Pr}_{v_0}^{\overline{\sigma}^*}$.

Proof. Let $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, \Delta, \chi, (\mathrm{Win}_i)_{i \in \Pi})$. By Theorem 2.5, for each player i we can fix a globally optimal pure strategy τ_i of the coalition $\Pi \setminus \{i\}$ in the coalition game $\mathcal{G}_i$; denote by $\tau_{j,i}$ the corresponding pure strategy of player j in $\mathcal{G}$. To simplify notation, we also define $\tau_{i,i}$ to be an arbitrary pure strategy of player i in $\mathcal{G}$. Player i's equilibrium strategy σ_i^* is defined as follows:

$$\sigma_i^*(xv) = \begin{cases} \sigma_i(xv) & \text{if } xv \text{ is consistent with } \overline{\sigma}, \\ \tau_{i,j}(x_2 v) & \text{otherwise,} \end{cases}$$

where, in the latter case, $x = x_1 x_2$ with x_1 being the longest prefix of xv such that $\mathrm{Pr}_{v_0}^{\overline{\sigma}}(x_1 \cdot V^{\omega}) > 0$ and $j \in \Pi$ being the player who has deviated from $\overline{\sigma}$, i.e. x_1 ends in V_j; if x_1 is empty or ends in a stochastic vertex, we set $j = i$. Intuitively, σ_i^* behaves like σ_i as long as no other player j deviates from playing σ_j, in which case σ_i^* starts to behave like $\tau_{i,j}$.

Note that $\mathrm{Pr}_{v_0}^{\overline{\sigma}^*} = \mathrm{Pr}_{v_0}^{\overline{\sigma}}$. We claim that, additionally, $\overline{\sigma}^*$ is a Nash equilibrium of $(\mathcal{G}, v_0)$. Let ρ be any pure strategy of player i in $\mathcal{G}$; by Proposition 3.2, it suffices to show that $\mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i) \leq \mathrm{Pr}_{v_0}^{\overline{\sigma}^*}(\mathrm{Win}_i)$.

Let us call a history $xv \in V^* \cdot V_i$ a *deviation history* if xv is consistent with both $\overline{\sigma}$ and $(\overline{\sigma}_{-i}, \rho)$, but $\sigma_i(xv) \neq \rho(xv)$; we denote the set of all deviation histories by D. Clearly, $\mathrm{Pr}_{v_0}^{\overline{\sigma}}(xv \cdot V^\omega) = \mathrm{Pr}_{v_0}^{\overline{\sigma}^*}(xv \cdot V^\omega) = \mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(xv \cdot V^\omega)$ for every $xv \in D$.

Claim. $\mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(X \setminus D \cdot V^\omega) = \mathrm{Pr}_{v_0}^{\overline{\sigma}}(X \setminus D \cdot V^\omega)$ for every Borel set $X \subseteq V^\omega$.

Proof. The claim is obviously true for the basic open sets $X = x \cdot V^\omega$, $x \in V^*$, and thus also for finite, disjoint unions of such sets, which are precisely the *clopen sets* (i.e. sets of the form $W \cdot V^\omega$ for finite $W \subseteq V^*$). Since the class of clopen sets is closed under complementation and taking finite unions, by the *monotone class theorem* [32], the closure of the class of all clopen sets under taking limits of chains contains the smallest σ-algebra containing all clopen sets, which is just the Borel σ-algebra. Hence, it suffices to show that whenever we are given measurable sets $X_1, X_2, \ldots \subseteq V^\omega$ with $X_1 \subseteq X_2 \subseteq \ldots$ or $X_1 \supseteq X_2 \supseteq \ldots$ such that the claim holds for each X_n, then the claim also holds for $\lim_n X_n$, where $\lim_n X_n = \bigcup_{n\in\mathbb{N}} X_n$ or $\lim_n X_n = \bigcap_{n\in\mathbb{N}} X_n$, respectively. So assume that $X_1, X_2, \ldots \subseteq V^\omega$ is a chain such that $\mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(X_n \setminus D \cdot V^\omega) = \mathrm{Pr}_{v_0}^{\overline{\sigma}}(X_n \setminus D \cdot V^\omega)$ for each $n \in \mathbb{N}$. Clearly, $(\lim_n X_n) \setminus D \cdot V^\omega = \lim_n (X_n \setminus D \cdot V^\omega)$. Moreover, since measures are continuous from above and below:

$$
\begin{aligned}
&\mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\lim_n (X_n \setminus D \cdot V^\omega)) \\
=\ &\lim_n \mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(X_n \setminus D \cdot V^\omega) \\
=\ &\lim_n \mathrm{Pr}_{v_0}^{\overline{\sigma}}(X_n \setminus D \cdot V^\omega) \\
=\ &\mathrm{Pr}_{v_0}^{\overline{\sigma}}(\lim_n (X_n \setminus D \cdot V^\omega)).
\end{aligned}
$$

$\square$

Claim. $\mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i \mid xv \cdot V^\omega) \leq \mathrm{val}_i^{\mathcal{G}}(v)$ for every $xv \in X$.

Proof. By the definition of the strategies $\tau_{j,i}$, we have that $\mathrm{Pr}_v^{(\tau_{j,i})_{j\neq i},\rho}(\mathrm{Win}_i) \leq \mathrm{val}_i^{\mathcal{G}}(v)$ for *every* vertex $v \in V$ and *every* strategy ρ of player i. On the other hand, if xv is a deviation history, then for each player j the residual strategy $\sigma_j^*[xv]$ is equal to $\tau_{j,i}$ on histories that start in $w := \rho(xv)$. Hence, by Lemma 2.1, and since Win_i is prefix-independent, we get:

$$
\begin{aligned}
&\mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i \mid xv \cdot V^\omega) \\
=\ &\mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i \mid xvw \cdot V^\omega) \\
=\ &\mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i \cap xvw \cdot V^\omega) \;/\; \mathrm{Pr}_{v_0}^{\overline{\sigma}^*_{-i},\rho}(xvw \cdot V^\omega) \\
=\ &\mathrm{Pr}_{w}^{\overline{\sigma}^*_{-i}[xv],\rho[xv]}(\mathrm{Win}_i) \\
=\ &\mathrm{Pr}_{w}^{(\tau_{j,i})_{j\neq i},\rho[xv]}(\mathrm{Win}_i) \\
\leq\ &\mathrm{val}_i^{\mathcal{G}}(w) \\
\leq\ &\mathrm{val}_i^{\mathcal{G}}(v).
\end{aligned}
$$

$\square$

Using the previous two claims, we can prove that $\Pr_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i) \leq \Pr_{v_0}^{\overline{\sigma}^*}(\mathrm{Win}_i)$ as follows:

$$\begin{aligned}
&\Pr_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i)\\
=\ &\Pr_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i \setminus X \cdot V^\omega) + \sum_{xv\in X} \Pr_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i \cap xv \cdot V^\omega)\\
=\ &\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \setminus X \cdot V^\omega) + \sum_{xv\in X} \Pr_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i \cap xv \cdot V^\omega)\\
=\ &\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \setminus X \cdot V^\omega) + \sum_{xv\in X} \Pr_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i \mid xv \cdot V^\omega) \cdot \Pr_{v_0}^{\overline{\sigma}^*_{-i},\rho}(xv \cdot V^\omega)\\
=\ &\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \setminus X \cdot V^\omega) + \sum_{xv\in X} \Pr_{v_0}^{\overline{\sigma}^*_{-i},\rho}(\mathrm{Win}_i \mid xv \cdot V^\omega) \cdot \Pr_{v_0}^{\overline{\sigma}}(xv \cdot V^\omega)\\
\leq\ &\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \setminus X \cdot V^\omega) + \sum_{xv\in X} \mathrm{val}_i^{\mathcal{G}}(v) \cdot \Pr_{v_0}^{\overline{\sigma}}(xv \cdot V^\omega)\\
\leq\ &\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \setminus X \cdot V^\omega) + \sum_{xv\in X} \Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \mid xv \cdot V^\omega) \cdot \Pr_{v_0}^{\overline{\sigma}}(xv \cdot V^\omega)\\
=\ &\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \setminus X \cdot V^\omega) + \sum_{xv\in X} \Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \cap xv \cdot V^\omega)\\
=\ &\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i)\\
=\ &\Pr_{v_0}^{\overline{\sigma}^*}(\mathrm{Win}_i).
\end{aligned}$$

□

A variant of Lemma 3.4 deals with games with ω-regular objectives and finite-state strategies.

Lemma 3.5. Let $(\mathcal{G}, v_0)$ be any SMG with finite arena and prefix-independent ω-regular objectives. If $\overline{\sigma}$ is a pure finite-state strategy profile such that $\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \mid xv \cdot V^\omega) \geq \mathrm{val}_i^{\mathcal{G}}(v)$ for each player i and for each history xv consistent with $\overline{\sigma}$, then there exists a pure finite-state Nash equilibrium $\overline{\sigma}^*$ with $\Pr_{v_0}^{\overline{\sigma}} = \Pr_{v_0}^{\overline{\sigma}^*}$.

Proof. The proof is analogous to the proof of Lemma 3.4: Since, by Corollary 2.8, there exist optimal pure finite-state strategies in every SMG with finite arena and ω-regular objectives, the strategies $\tau_{j,i}$ as defined there can be assumed to be pure finite-state strategies. Consequently, the equilibrium profile $\overline{\sigma}^*$ can be implemented using finite-state strategies as well. □

Using Lemma 3.4, we can now prove the existence of a pure Nash equilibrium in any SMG with finite arena and prefix-independent objectives.

Theorem 3.6. There exists a pure Nash equilibrium in any SMG with finite arena and prefix-independent objectives.

Proof. Let $(\mathcal{G}, v_0)$ be any SMG with prefix-independent objectives whose arena is finite. By Theorem 2.5, for each player i there exists a residually optimal strategy σ_i in $\mathcal{G}$. Let $\overline{\sigma} = (\sigma_i)_{i\in\Pi}$. For every history xv of $(\mathcal{G}, v_0)$ that is consistent with $\overline{\sigma}$ and each player i, we have $\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Win}_i \mid xv \cdot V^\omega) = \Pr_v^{\overline{\sigma}[x]}(x^{-1}\mathrm{Win}_i) \geq \mathrm{val}_i^{\mathcal{G}[x]}(v) = \mathrm{val}_i^{\mathcal{G}}(v)$. By Lemma 3.4, this implies that there exists a pure Nash equilibrium of $(\mathcal{G}, v_0)$. □

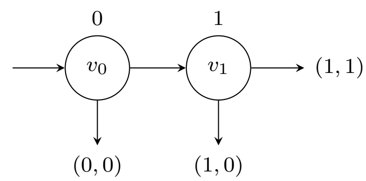


FIGURE 2. A 2-player game with a pair of optimal strategies that cannot be extended to a Nash equilibrium.

For SMGs with ω-regular objectives played on a finite arena, we can even show the existence of a pure finite-state equilibrium

Theorem 3.7. There exists a pure finite-state Nash equilibrium in any SMG with finite arena and ω-regular objectives.

Proof. Since any SMG with ω-regular objectives can be reduced to one with parity objectives, it suffices to consider parity SMGs. For these games, the claim follows from Corollary 2.7 and Lemma 3.5 using the same argumentation as in the proof of Theorem 3.6. $\square$

Theorem 3.7 and a variant of Theorem 3.6 appeared originally in [13]. However, their proof contains an inaccuracy: Essentially, they claim that any profile of optimal strategies can be extended to a Nash equilibrium with the same payoff (by adding threat strategies on top). This is, in general, not true, as the following example demonstrates.

Example 3.8. Consider the deterministic 2-player game $(\mathcal{G}, v_0)$ depicted in Figure 2. Clearly, the value $\mathrm{val}_0^{\mathcal{G}}(v_0)$ for player 0 from v_0 is 1, and player 0's optimal strategy σ is to play from v_0 to v_1. For player 1, the value from v_0 is 0, and both of her positional strategies are optimal (albeit not necessarily globally optimal). In particular, her strategy τ of playing from v_1 to the terminal vertex with payoff $(1, 0)$ is optimal. The payoff of the strategy profile (σ, τ) is $(1, 0)$. However, there is no Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $(1, 0)$: In any Nash equilibrium of $(\mathcal{G}, v_0)$, player 0 will move from v_0 to v_1 with probability 1. To have a Nash equilibrium, player 1 must play from v_1 to the terminal vertex with payoff $(1, 1)$ with probability 1; hence, every Nash equilibrium of this game has payoff $(1, 1)$.

4. COMPLEXITY OF NASH EQUILIBRIA

For the rest of this paper, we will only deal with SMGs whose arena is *finite*. Previous research on algorithms for finding Nash equilibria in such games has focused on computing *some* Nash equilibrium [13]. However, a game may have several Nash equilibria with different payoffs, and one might not be interested in *any* Nash equilibrium but in one whose payoff fulfils certain requirements. For example, one might look for a Nash equilibrium where certain players win almost surely while certain others lose almost surely. This idea leads us to the following decision problem, which we call NE:[2]

> Given an SMG $(\mathcal{G}, v_0)$ and thresholds $\overline{x}, \overline{y} \in [0, 1]^{\Pi}$, decide whether there exists a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\geq \overline{x}$ and $\leq \overline{y}$.

[2]In the definition of NE, the ordering $\leq$ is applied componentwise.

For computational purposes, we assume that the thresholds $\bar{x}$ and $\bar{y}$ are vectors of rational numbers and that all objectives are ω-regular. A variant of the problem which omits the thresholds just asks about a Nash equilibrium where some distinguished player, say player 0, wins with probability 1:

> Given an SMG $(\mathcal{G}, v_0)$, decide whether there exists a Nash equilibrium of $(\mathcal{G}, v_0)$ where player 0 wins almost surely.

Clearly, every instance of the threshold-free variant can easily be turned into an instance of NE (by adding the thresholds $\bar{x} = (1, 0, \ldots, 0)$ and $\bar{y} = (1, \ldots, 1)$). Hence, NE is, a priori, more general than its threshold-free variant.

Our main concern in this paper are variants of NE where we restrict the type of strategies that are allowed in the definition of the problem: Let PureNE, FinNE, StatNE and PosNE be the problems that arise from NE by restricting the desired Nash equilibrium to consist of pure strategies, finite-state strategies, stationary strategies, and positional strategies, respectively. In the rest of this paper, we are going to prove upper and lower bounds on the complexity of these problems, where all lower bounds hold for the threshold-free variants as well.

Our first observation is that neither stationary nor pure strategies are sufficient to implement any Nash equilibrium, even for SSMGs and even if we are only interested in whether a player wins or loses almost surely in the equilibrium. Together with a later result from this section (namely Proposition 4.10), this demonstrates that the problems NE, PureNE, FinNE, StatNE, and PosNE are pairwise distinct problems, which have to be analysed separately. This in sharp contract to the situation for SS2Gs where all these problems coincide because these games admit globally optimal positional strategies.

Proposition 4.1. There exists an SSMG that has a finite-state Nash equilibrium where player 0 wins almost surely but that has no stationary Nash equilibrium where player 0 wins with positive probability.

Proof. Consider the game $\mathcal{G}$ depicted in Figure 3 played by three players 0, 1 and 2 (with payoffs in this order). Obviously, the following finite-state strategy profile is a Nash equilibrium where player 0 wins almost surely: Player 1 plays from vertex v_2 to vertex v_3 at the first visit of v_2 but leaves the game immediately (by playing to the neighbouring terminal vertex) at all subsequent visits to v_2; from vertex v_0 player 1 plays to v_1; player 2 plays from vertex v_3 to vertex v_4 at the first visit of v_3 but leaves the game immediately at all subsequent visits to v_3; from vertex v_1 player 2 plays to v_2.

It remains to show that there is no stationary Nash equilibrium of $(\mathcal{G}, v_0)$ where player 0 wins with positive probability. Any stationary Nash equilibrium of $(\mathcal{G}, v_0)$ where player 0 wins with positive probability induces a stationary Nash equilibrium of $(\mathcal{G}, v_2)$ where both players 1 and 2 receive payoff at least $\frac{1}{2}$ since otherwise one of these players could improve her payoff by changing her strategy at v_0 or v_1. Hence, it suffices to show that $(\mathcal{G}, v_2)$ has no stationary Nash equilibrium where both players 1 and 2 receive payoff at least $\frac{1}{2}$. Assume there exists such an equilibrium and denote by p the probability that player 2 plays from v_3 to v_4. Since player 1 wins with probability > 0, it must be the case that $p > 0$. But then, to have a Nash equilibrium, player 1 must play from v_2 to v_3 with probability 1, giving player 2 a payoff of 0, a contradiction. □

Proposition 4.2. There exists an SSMG that has a stationary Nash equilibrium where player 0 wins almost surely but that has no pure Nash equilibrium where player 0 wins with positive probability.

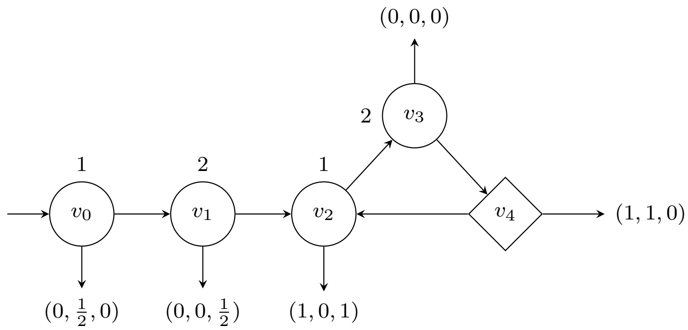


FIGURE 3. An SSMG that has a pure finite-state Nash equilibrium where player 0 wins almost surely but no stationary Nash equilibrium where player 0 wins with positive probability.

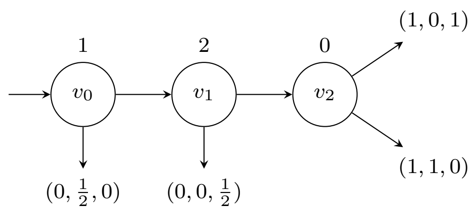


FIGURE 4. An SSMG that has a stationary Nash equilibrium where player 0 wins almost surely but no pure Nash equilibrium where player 0 wins with positive probability.

Proof. Consider the game depicted in Figure 4 played by three players 0, 1 and 2 (with payoffs given in this order). Clearly, the stationary strategy profile where from vertex v_2 player 0 selects both outgoing edges with probability $\frac{1}{2}$ each, player 1 plays from v_0 to v_1 and player 2 plays from v_1 to v_2 is a Nash equilibrium where player 0 wins almost surely. However, for any pure strategy profile where player 0 wins with positive probability (i.e. with probability 1), either player 1 or player 1 receives payoff 0 and could improve her payoff by switching her strategy at v_0 or v_1, respectively. □

4.1. **Positional equilibria.** In this subsection, we analyse the complexity of the (presumably) simplest of the decision problems introduced so far: PosNE. Not surprisingly, this problem is decidable; in fact, it is NP-complete for all types of objectives we consider in this paper. Let us start by proving membership to NP. Since simple reachability, (co-)Büchi and parity objectives can easily be translated to Rabin or Streett objectives, it suffices to consider Streett-Rabin and Muller SMGs.

Theorem 4.3. PosNE is in NP for SMGs with Streett-Rabin or Muller objectives.

Proof. To decide PosNE, on input $\mathcal{G}, v_0, \overline{x}, \overline{y}$ we can guess a positional strategy profile $\overline{\sigma}$, i.e. a mapping $\bigcup_{i \in \Pi} V_i \to V$; then, we verify whether $\overline{\sigma}$ is a Nash equilibrium with the desired payoff. To do this, we first compute the payoff z_i of $\overline{\sigma}$ for each player i by computing the probability of the event Win_i in the (finite) Markov chain $(\mathcal{G}^{\overline{\sigma}}, v_0)$. Once each z_i is computed, we can easily check whether $x_i \le z_i \le y_i$. To verify that $\overline{\sigma}$ is a Nash equilibrium, we additionally compute, for each player i, the value r_i of the (finite) MDP $(\mathcal{G}^{\overline{\sigma}_{-i}}, v_0)$. Clearly, $\overline{\sigma}$ is a Nash equilibrium if and only if $r_i \leq z_i$ for each player i. Since we can compute the value of any MDP (and thus any Markov chain) with one of the above objectives in polynomial time (Theorem 2.9), all these checks can be carried out in polynomial time. □

To establish NP-completeness, it remains to prove NP-hardness. In fact, the reduction we are going to present does not only work for PosNE, but also for StatNE, where we allow arbitrary stationary equilibria.

Theorem 4.4. PosNE and StatNE are NP-hard, even for SSMGs with only two players (three players for the threshold-free variant).

Proof. The proof is by reduction from SAT. Let $\varphi = C_1 \wedge \ldots \wedge C_m$, $m \geq 1$, be a formula in conjunctive normal form over propositional variables $X_1, \ldots, X_n$ (where, without loss of generality, each clause is non-empty). Our aim is to construct a 2-player SSMG $(\mathcal{G}, v_0)$ such that the following statements are equivalent:

1. φ is satisfiable.
2. $(\mathcal{G}, v_0)$ has a positional Nash equilibrium with payoff $(1, \frac{1}{2})$.
3. $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium with payoff $(1, \frac{1}{2})$.

Provided that the game can be constructed in polynomial time, these equivalences establish both reductions. The game $\mathcal{G}$ is depicted in Figure 5. The game proceeds from the initial vertex v_0 to X_i or $\overline{X_i}$ with probability $\frac{1}{2^{i+1}}$ each, and to vertex φ with probability $\frac{1}{2^{n+1}}$; with the remaining probability of $\frac{1}{2^{n+1}}$ the game proceeds to a terminal vertex with payoff $(1, 0)$. From φ, the game proceeds to vertex C_j with probability $\frac{1}{m+1}$ each; with the remaining probability of $\frac{1}{m+1}$, the game proceeds to a terminal vertex with payoff $(1, 1)$. From vertex C_j (controlled by player 1) there is an edge to vertex X_i or $\overline{X_i}$ if and only if X_i respectively $\neg X_i$ occurs inside the clause C_j. Obviously, the game $\mathcal{G}$ can be constructed from φ in polynomial time. It remains to show that 1.–3. are equivalent.

$(1. \Rightarrow 2.)$ Assume that $\alpha : \{X_1, \ldots, X_n\} \to \{\text{true}, \text{false}\}$ is a satisfying assignment of φ. In the positional Nash equilibrium of $(\mathcal{G}, v_0)$, player 0 moves from a literal L, $L = X_i$ or $L = \overline{X_i}$, to the neighbouring $\top$-labelled vertex if and only if L is mapped to true by α, and player 1 moves from vertex C_j to a (fixed) literal L that is contained in C_j and mapped to true by α (which is possible since α is a satisfying assignment). At $\top$-labelled vertices, player 1 never plays to a terminal vertex. Obviously, player 0 wins almost surely with this strategy profile. For player 1, the payoff is

$$\frac{1}{2^{n+1}} + \sum_{i=1}^{n} \frac{1}{2^{i+1}} = \frac{1}{2^{n+1}} + \frac{1}{2}\Big(\sum_{i=1}^{n} \frac{1}{2^i}\Big) = \frac{1}{2^{n+1}} + \frac{1}{2}\Big(1 - \frac{1}{2^n}\Big) = \frac{1}{2},$$

where the first summand is the probability of going from the initial vertex to φ, from where player 1 wins almost surely since from every clause vertex she plays to a "true" literal. Obviously, changing her strategy cannot give her a better payoff. Therefore, we have found a Nash equilibrium.

$(2. \Rightarrow 3.)$ Trivial.

$(3. \Rightarrow 1.)$ Let $\overline{\sigma} = (\sigma_0, \sigma_1)$ be a stationary Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $(1, \frac{1}{2})$. Our first aim is to show that σ_0 is actually a positional strategy. Towards a contradiction, assume that there exists a literal L such that $\sigma_0(L)$ assigns probability $0 < q < 1$ to the neighbouring $\top$-labelled vertex. Since player 0 wins almost surely, player 1 never plays to a terminal vertex with payoff $(0, 1)$. Hence, the expected payoff for player 1 from vertex L

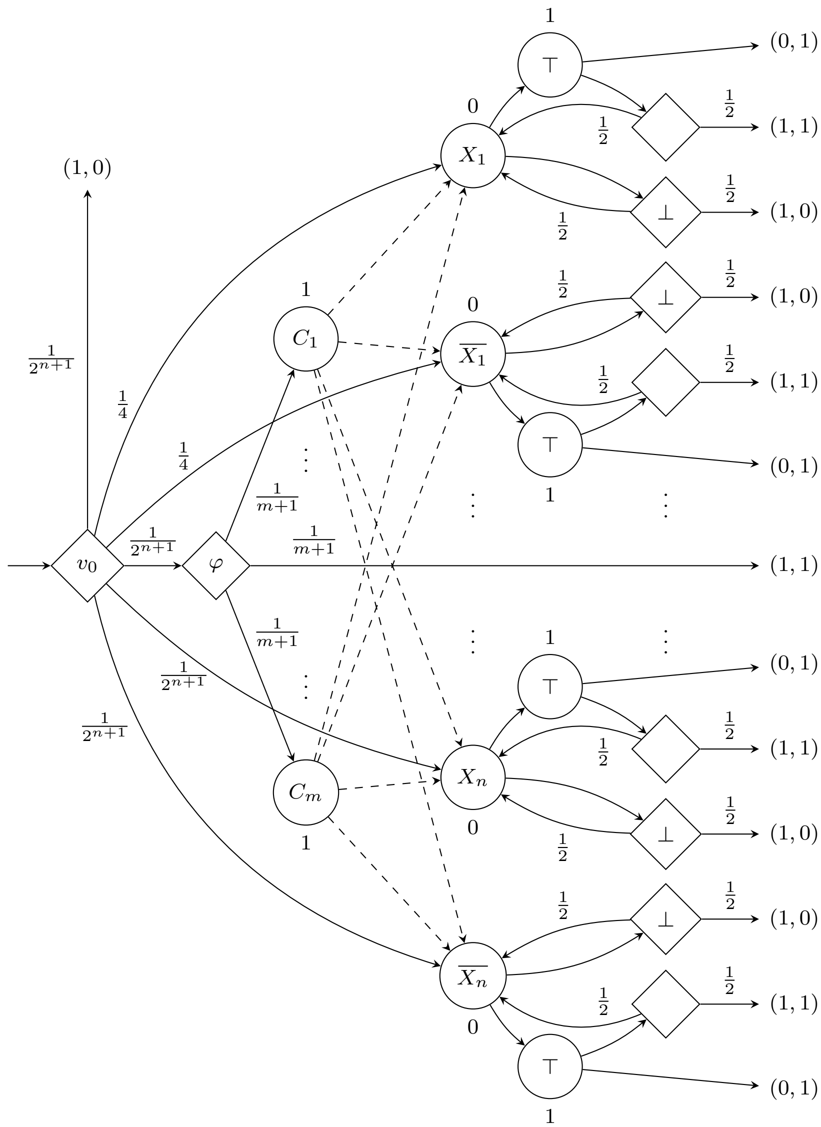


FIGURE 5. Reducing SAT to PosNE and StatNE.

(i.e. in the game $(\mathcal{G}, L)$) is precisely q. However, by playing to a terminal vertex with payoff $(0, 1)$, she could receive payoff $\frac{2q}{1+q} > q$. Therefore, $\bar{\sigma}$ is not a Nash equilibrium, a contradiction.

Knowing that σ_0 is a positional strategy, we can define a *pseudo assignment* $\alpha : \{X_1, \neg X_1, \dots, X_n, \neg X_n\} \to \{\text{true}, \text{false}\}$ by setting $\alpha(L) = \text{true}$ if σ_1 prescribes to go from vertex L to the neighbouring $\top$-labelled vertex. Our next aim is to show that α is actually an assignment: $\alpha(X_i) = \text{true}$ if and only if $\alpha(\neg X_i) = \text{false}$. To see this, note that we can compute player 1's expected payoff as follows:

$$\frac{1}{2} = \frac{p}{2^{n+1}} + \sum_{i=1}^{n} \frac{a_i}{2^{i+1}}, \quad a_i = \begin{cases} 0 & \text{if } \alpha(X_i) = \alpha(\neg X_i) = \text{false}, \\ 1 & \text{if } \alpha(X_i) \neq \alpha(\neg X_i), \\ 2 & \text{if } \alpha(X_i) = \alpha(\neg X_i) = \text{true}, \end{cases}$$

where p is the expected payoff for player 1 from vertex φ. By the construction of $\mathcal{G}$, we have $p > 0$, and the equality only holds if $p = 1$ and $a_i = 1$ for all $i = 1, \ldots, n$, which proves that α is an assignment.

Finally, we claim that α satisfies φ. If this were not the case, there would exist a clause C such that player 1's expected payoff from vertex C is 0 and therefore $p < 1$. This is a contradiction to the fact that $p = 1$, as we have shown above.

To show that the threshold-free variants of PosNE and StatNE are also NP-hard, it suffices to modify the game $\mathcal{G}$ as follows: First, we add one new player, player 2, who wins at precisely those terminal vertices where player 1 loses. Second, we add two new vertices v_1 and v_2. At v_1, player 1 has the choice to leave the game; if she decides to stay inside the game, the play proceeds to v_2, where player 2 has the choice to leave the game; if she also decides to stay inside the game, the play proceeds to vertex v_0 from where the game continues as described above; if player 1 or player 2 decide to leave the game, then both receive payoff $\frac{1}{2}$, but player 0 receives payoff 0. Let us denote the modified game by $\mathcal{G}'$. It is straightforward to see that the following statements are equivalent:

1. $(\mathcal{G}', v_1)$ has a stationary Nash equilibrium where player 0 wins almost surely.
2. $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium with payoff $(1, \frac{1}{2})$.
3. φ is satisfiable.
4. $(\mathcal{G}, v_0)$ has a positional Nash equilibrium with payoff $(1, \frac{1}{2})$.
5. $(\mathcal{G}', v_1)$ has a positional Nash equilibrium where player 0 wins almost surely.

□

4.2. **Stationary equilibria.** To prove the decidability of StatNE, we appeal to results established for the *Existential Theory of the Reals*, $\exists\mathrm{Th}(\mathfrak{R})$, the set of all existential first-order sentences (over the appropriate signature) 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 [9], which leads to the following theorem.

Theorem 4.5. StatNE is in PSPACE for SMGs with Streett-Rabin or Muller objectives.

Proof. Since PSPACE = NPSPACE, it suffices to provide a nondeterministic algorithm with polynomial space requirements for deciding StatNE. On input $\mathcal{G}, v_0, \overline{x}, \overline{y}$, where without loss of generality $\mathcal{G}$ is an SMG with Muller objectives $\mathcal{F}_i \subseteq \mathcal{P}(C)$, the algorithm starts by guessing the *support* $S \subseteq V \times V$ of a stationary strategy profile $\overline{\sigma}$ of $\mathcal{G}$, i.e.

$$S = \{(v, w) \in V \times V : \overline{\sigma}(w \mid v) > 0\}.$$

From the set S alone, by standard graph algorithms, one can compute (in polynomial time) for each player i the following sets:

1. the union F_i of all end components (i.e. bottom SCCs) of the Markov chain $\mathcal{G}^{\overline{\sigma}}$ that are winning for player i;
2. the set R_i of vertices v such that $\Pr_v^{\overline{\sigma}}(\mathrm{Reach}(F_i)) > 0$;
3. the union T_i of all end components of the MDP $\mathcal{G}^{\overline{\sigma}_{-i}}$ that are winning for player i.

After computing all these sets, the algorithm evaluates an existential first-order sentence ψ, which can be computed in polynomial time from $\mathcal{G}$, v_0, $\bar{x}$, $\bar{y}$, $(R_i)_{i\in\Pi}$, $(F_i)_{i\in\Pi}$ and $(T_i)_{i\in\Pi}$ over $\mathfrak{R}$ and returns the answer to this query.

It remains to describe a suitable sentence ψ. Let $\bar{\alpha} = (\alpha_{vw})_{v,w\in V}$, $\bar{r} = (r_v^i)_{i\in\Pi,v\in V}$ and $\bar{z} = (z_v^i)_{i\in\Pi,v\in V}$ be three sets of variables. The formula

$$\varphi(\bar{\alpha}) := \bigwedge_{v\in\bigcup V_i}\Bigg(\bigwedge_{w\in v\Delta}\alpha_{vw}\geq 0\ \wedge \bigwedge_{w\in V\setminus v\Delta}\alpha_{vw}=0\ \wedge\sum_{w\in v\Delta}\alpha_{vw}=1\Bigg)\wedge$$
$$\bigwedge_{\substack{v\in V\setminus\bigcup V_i\\ w\in V}}\alpha_{vw}=p_{vw}\ \wedge\bigwedge_{(v,w)\in S}\alpha_{vw}>0\ \wedge\bigwedge_{(v,w)\notin S}\alpha_{vw}=0,$$

where p_{vw} is the unique number such that $(v, p_{vw}, w)\in\Delta$, states that the mapping $\bar{\sigma}: V\to\mathcal{D}(V)$ defined by $\bar{\sigma}(v)(w)=\alpha_{vw}$ constitutes a valid stationary strategy profile of $\mathcal{G}$ whose support is S. Provided that $\varphi(\bar{\alpha})$ holds in $\mathfrak{R}$, the formula

$$\eta_i(\bar{\alpha},\bar{z}) := \bigwedge_{v\in F_i} z_v^i = 1\ \wedge\bigwedge_{v\in V\setminus R_i} z_v^i=0\ \wedge\bigwedge_{v\in V\setminus F_i} z_v^i=\sum_{w\in v\Delta}\alpha_{vw}z_w^i$$

states that $z_v^i = \Pr_v^{\bar{\sigma}}(\mathrm{Win}_i)$ for each $v\in V$, where $\bar{\sigma}$ is defined as above. This follows from a well-known results about Markov chains, namely that the vector of the aforementioned probabilities is the unique solution of the given system of equations. Finally, the formula

$$\vartheta_i(\bar{\alpha},\bar{r}) := \bigwedge_{v\in V} r_v^i\geq 0\ \wedge\bigwedge_{v\in T_i} r_v^i=1\ \wedge\bigwedge_{\substack{v\in V_i\\ w\in v\Delta}} r_v^i\geq r_w^i\ \wedge\bigwedge_{v\in V\setminus V_i} r_v^i=\sum_{w\in v\Delta}\alpha_{vw}r_w^i$$

states that $\bar{r}$ is a solution of the linear programme for computing the maximal payoff that player i can achieve when playing against the strategy profile $\bar{\sigma}_{-i}$. In particular, the formula is fulfilled if $r_v^i = \sup_\tau \Pr_v^{(\bar{\sigma}_{-i},\tau)}(\mathrm{Reach}(T_i)) = \sup_\tau \Pr_v^{(\bar{\sigma}_{-i},\tau)}(\mathrm{Win}_i)$ (the latter equality follows from Lemmas 2.2 and 2.3), and every other solution is greater than this one (in each component).

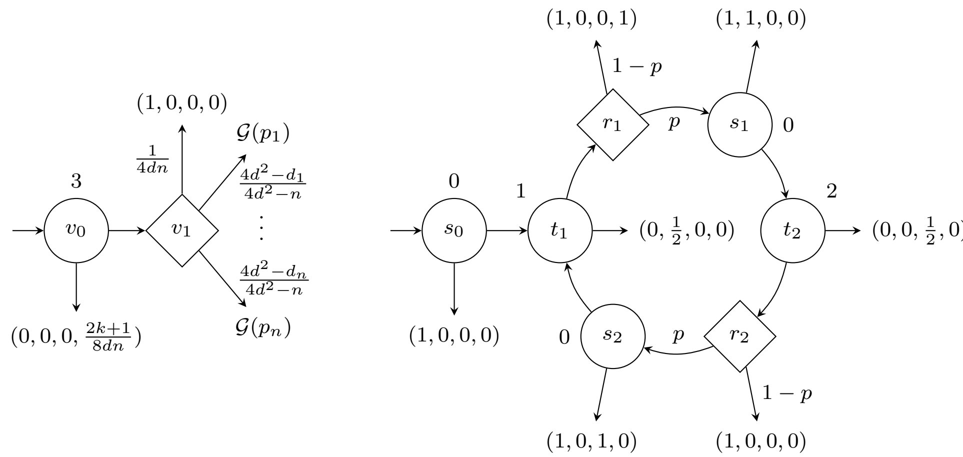
The desired sentence ψ is the existential closure of the conjunction of φ and, for each player i, the formulae η_i and ϑ_i combined with formulae stating that player i cannot improve her payoff and that the expected payoff for player i lies in between the given thresholds:

$$\psi := \exists\bar{\alpha}\,\exists\bar{r}\,\exists\bar{z}\left(\varphi(\bar{\alpha})\wedge\bigwedge_{i\in\Pi}(\eta_i(\bar{\alpha},\bar{z})\wedge\vartheta_i(\bar{\alpha},\bar{r})\wedge r_{v_0}^i\leq z_{v_0}^i\wedge x_i\leq z_{v_0}^i\leq y_i)\right).$$

Clearly, ψ holds in $\mathfrak{R}$ if and only if $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium with payoff at least $\bar{x}$ and at most $\bar{y}$ whose support is S. Consequently, the algorithm is correct. □

In the previous subsection, we showed that StatNE is NP-hard, leaving a considerable gap to our upper bound of PSPACE. Towards gaining a better understanding of the problem, we relate StatNE to the *square root sum problem* (SqrtSum) of deciding, given numbers $d_1,\dots,d_n,k\in\mathbb{N}$, 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*, which is a slight improvement over the previously known PSPACE upper bound. However, it is an open question since the 1970s whether SqrtSum falls into the polynomial hierarchy [28, 24]. We identify a polynomial-time reduction from SqrtSum to StatNE



(a) The game $\mathcal{G}$

(b) The game $\mathcal{G}(p)$

FIGURE 6. Reducing SqrtSum to StatNE.

for SSMGs.[3] 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 computation.

Theorem 4.6. SqrtSum is polynomial-time reducible to StatNE, even for 4-player SSMGs.

Before we start with the proof of the theorem, let us first examine the game $\mathcal{G}(p)$, $p \in [\frac{1}{2}, 1)$, which is depicted in Figure 6 (b).

Claim. The maximal payoff player 3 receives in a stationary Nash equilibrium of $(\mathcal{G}(p), s_0)$ is $\frac{\sqrt{2-2p}-p+1}{2p+2}$.

Proof. Note that a stationary strategy profile $\overline{\sigma}$ can only be a Nash equilibrium where player 3 receives payoff > 0 if player 1 plays from t_1 to r_1 with probability 1 and player 2 plays from t_2 to r_2 with probability 1 (or if t_2 is not reachable with $\overline{\sigma}$ in which case player 3 receives payoff $\leq 1-p$) because otherwise player 0 receives payoff < 1 and would prefer to leave the game a v_0 where she could get payoff 1. Moreover, the maximum payoff for player 3 can only be attained when player 0 plays with probability 1 from s_0 to t_1 because, if player 0 plays from s_0 to t_1 with probability $0 < x < 1$, then setting x to 1 yields a Nash equilibrium with a better payoff for player 3. Hence, the only variable quantities are the probabilities x_1 and x_2 that player 0 plays from s_1 to t_2 respectively from s_2 to t_1. Given x_1 and x_2, we can compute the probabilities $p_1(x_1, x_2) := \mathrm{Pr}^{\overline{\sigma}}_{t_1}(\mathrm{Win}_1)$ and $p_2(x_1, x_2) := \mathrm{Pr}^{\overline{\sigma}}_{t_2}(\mathrm{Win}_2)$ as follows: $p_1(x_1, x_2) = \frac{p(1-x_1)}{1-x_1x_2p^2}$, and $p_2(x_1, x_2) = \frac{p(1-x_2)}{1-x_1x_2p^2}$. To have a Nash equilibrium, it must be the case that $p_1(x_1, x_2), p_2(x_1, x_2) \geq \frac{1}{2}$ since otherwise player 1 or player 2 would prefer to leave the game at t_1 or t_2, respectively, where they could obtain payoff $\frac{1}{2}$ immediately. Vice versa, if $p_1(x_1, x_2), p_2(x_1, x_2) \geq \frac{1}{2}$ then $\overline{\sigma}$ is a Nash equilibrium with expected payoff $\frac{1-p}{1-x_1x_2p^2} \geq 1-p$ for player 3.

[3]Some authors define SqrtSum with $\leq$ instead of $\geq$. With this definition, we would reduce from the complement of SqrtSum instead.

Hence, to determine the maximum payoff for player 3 in a stationary Nash equilibrium, we have to maximise $\frac{1-p}{1-x_1x_2p^2}$, the expected payoff for player 3, under the constraints $p_1(x_1,x_2), p_2(x_1,x_2) \geq \frac{1}{2}$ and $0 \leq x_1, x_2 \leq 1$. We claim that the maximum is reached only if $x_1 = x_2$; if, for example, $x_1 > x_2$ then we can achieve a higher payoff for player 3 by setting $x_2' := x_1$, and the constraints are still satisfied:

$$\frac{p(1-x_2')}{1-x_1x_2'p^2} = \frac{p(1-x_1)}{1-x_1^2p^2} \geq \frac{p(1-x_1)}{1-x_1x_2p^2} \geq \frac{1}{2}.$$

Thus, in fact, we have to maximise $\frac{1-p}{1-x^2p^2}$ under the constraints $\frac{p(1-x)}{1-x^2p^2} \geq \frac{1}{2}$ and $0 \leq x \leq 1$. Since $p \in [\frac{1}{2}, 1)$, this is equivalent to maximising $\frac{1-p}{1-x^2p^2}$ subject to the constraints

$$\begin{aligned} &p^2x^2 - 2px + 2p - 1 \geq 0, \\ &0 \leq x \leq 1. \end{aligned}$$

The roots of the former polynomial are $(1 \pm \sqrt{2-2p}) \,/\, p$, but $(1 + \sqrt{2-2p}) \,/\, p > 1$ for $p \in [\frac{1}{2}, 1)$. Therefore, any solution must be less than (or be equal to) $x_0 := (1 - \sqrt{2-2p}) / p$. In fact, we always have $0 \leq x_0 < 1$ for $p \in (\frac{1}{2}, 1)$. Therefore, x_0 is the optimal solution, and the maximal payoff for player 3 is indeed

$$\frac{1-p}{1-x_0^2p^2} = \frac{1-p}{1-(1-\sqrt{2-2p})^2} = \frac{\sqrt{2-2p}-p+1}{2p+2}.$$

□

Proof of Theorem 4.6. To prove the theorem, we show how to construct from an instance $(d_1, \dots, d_n, k)$ of SqrtSum a 4-player SSMG $(\mathcal{G}, v_0)$ such that $\sum_{i=1}^n \sqrt{d_i} \geq k$ if and only if $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium where player 0 wins almost surely.

Let $(d_1, \dots, d_n, k)$ be an instance of SqrtSum where, without loss of generality, $n > 0$, $d_i > 0$ for each $i = 1, \dots, n$, and $k \leq d := \sum_{i=1}^n d_i$. Define $p_i := 1 - \frac{d_i}{2d^2}$ for $i = 1, \dots, n$. Note that $p_i \in [\frac{1}{2}, 1)$ since $0 < d_i \leq d \leq d^2$. For the reduction, we use n copies of the game $\mathcal{G}(p)$, where in the ith copy we set p to p_i. The complete game $\mathcal{G}$ is depicted in Figure 6 (a). Clearly, $\mathcal{G}$ can be constructed in polynomial time from $(d_1, \dots, d_n, k)$.

By the above claim, the maximal payoff player 3 receives in a stationary Nash equilibrium of $(\mathcal{G}(p_i), s_0)$ is

$$\frac{\sqrt{2-2p_i}-p_i+1}{2p_i+2} = \frac{\frac{1}{d}\sqrt{d_i} - (1-\frac{d_i}{2d^2}) + 1}{4 - \frac{d_i}{d^2}} = \frac{d\sqrt{d_i} + \frac{d_i}{2}}{4d^2 - d_i}.$$

Consequently, the maximal payoff player 3 receives in a stationary Nash equilibrium of $(\mathcal{G}, v_1)$ is

$$\sum_{i=1}^n \frac{4d^2-d_i}{4d^2n} \cdot \frac{d\sqrt{d_i}+\frac{d_i}{2}}{4d^2-d_i} = \sum_{i=1}^n \frac{\sqrt{d_i}}{4dn} + \sum_{i=1}^n \frac{d_i}{8d^2n} = \sum_{i=1}^n \frac{\sqrt{d_i}}{4dn} + \frac{1}{8dn}.$$

To prove the theorem, it remains to be shown that $\sum_{i=1}^n \sqrt{d_i} \geq k$ if and only if $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium where player 0 wins almost surely.

($\Rightarrow$) Assume that $\sum_{i=1}^n \sqrt{d_i} \geq k$. Then also $\sum_{i=1}^n \frac{\sqrt{d_i}}{4dn} + \frac{1}{8dn} \geq \frac{2k+1}{8dn}$, and any stationary Nash equilibrium $\overline{\sigma}$ of $(\mathcal{G}, v_1)$ with this payoff for player 3 can be extended to a Nash equilibrium of $(\mathcal{G}, v_0)$ where player 0 wins almost surely by setting $\overline{\sigma}(v_1 \mid v_0) = 1$.

($\Leftarrow$) Assume that $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium where player 0 wins almost surely, but $\sum_{i=1}^{n} \sqrt{d_i} < k$. Then also $\sum_{i=1}^{n} \frac{\sqrt{d_i}}{4dn} + \frac{1}{8dn} < \frac{2k+1}{8dn}$, and in every stationary Nash equilibrium of $(\mathcal{G}, v_0)$ player 3 leaves the game at v_0, which gives payoff 0 to player 0, a contradiction. □

4.3. **Pure equilibria.** In this section, we show that the problem PureNE is undecidable by exhibiting a reduction from an undecidable problem about *two-counter machines*. Our construction is inspired by a construction used by Brázdil et al. [7] to prove the undecidability of stochastic games with branching-time objectives; see Remark 4.9 below.

A two-counter machine $\mathcal{M}$ is given by a list of instructions $\iota_1, \ldots, \iota_m$ where each instruction is one of the following:

– "inc(j); goto k" (increment counter j by 1 and go to instruction number k);
– "zero(j) ? goto k : dec(j); goto l" (if the value of counter j is 0, go to instruction number k; otherwise, decrement counter j by 1 and go to instruction number l);
– "halt" (stop the computation).

Here j ranges over $1, 2$ (the two counters), and $k \neq l$ range over $1, \ldots, m$. A configuration of $\mathcal{M}$ is a triple $C = (i, c_1, c_2) \in \{1, \ldots, m\} \times \mathbb{N} \times \mathbb{N}$, where i denotes the number of the current instruction and c_j denotes the current value of counter j. A configuration C' is the *successor* of configuration C, denoted by $C \vdash C'$, if it results from C by executing instruction ι_i; a configuration $C = (i, c_1, c_2)$ with $\iota_i =$ "halt" has no successor configuration. Finally, the *computation of* $\mathcal{M}$ is the unique maximal sequence $\rho = \rho(0)\rho(1)\ldots$ such that $\rho(0) \vdash \rho(1) \vdash \ldots$ and $\rho(0) = (1, 0, 0)$ (the *initial configuration*). Note that ρ is either infinite or finite, in which case it ends in a configuration $C = (i, c_1, c_2)$ such that $\iota_i =$ "halt".

The *halting problem* is to decide, given a machine $\mathcal{M}$, whether the computation of $\mathcal{M}$ is finite. It is well-known that two-counter machines are Turing powerful, which makes the halting problem and its dual, the *non-halting problem*, undecidable.

Theorem 4.7. PureNE is undecidable, even for 9-player SSMGs.

In order to prove Theorem 4.7, we show that one can compute from a two-counter machine $\mathcal{M}$ an SSMG $(\mathcal{G}, v_0)$ with 9 players such that the computation of $\mathcal{M}$ is infinite if and only if $(\mathcal{G}, v_0)$ has a pure Nash equilibrium where player 0 wins almost surely, which establishes a reduction from the non-halting problem to PureNE.

The game $\mathcal{G}$ is played by player 0 and eight other players A_j^t and B_j^t, indexed by $j \in \{1, 2\}$ and $t \in \{0, 1\}$. Intuitively, player 0 builds up the computation of $\mathcal{M}$, and the other players make sure that player 0 updates the counters correctly: If player 0 cheats or the computation stops, one of them will prefer to play a strategy that gives a bad payoff to player 0. More precisely, in every step of the computation, the players A_j^0 and A_j^1 make sure that the value of counter j is not too high, and the players B_j^0 and B_j^1 make sure that the value of counter j is not too low. Hereby, they alternate: The first step of the computation is monitored by the players A_j^0 and B_j^0, the second step by the players A_j^1 and B_j^1, the third step again by the players A_j^0 and B_j^0, and so on.

Let $\Gamma = \{\text{init}, \text{inc}(j), \text{dec}(j), \text{zero}(j) : j = 1, 2\}$. If $\mathcal{M}$ has instructions $\iota_1, \ldots, \iota_m$, then for each $i \in \{1, \ldots, m\}$, each $\gamma \in \Gamma$, each $j \in \{1, 2\}$ and each $t \in \{0, 1\}$, the game $\mathcal{G}$ contains the gadgets $S_{i,\gamma}^t$, I_i^t and $C_{j,\gamma}^t$, which are depicted in Figure 7. For better readability, terminal vertices are depicted as squares, where the label indicates which players win. Note that the

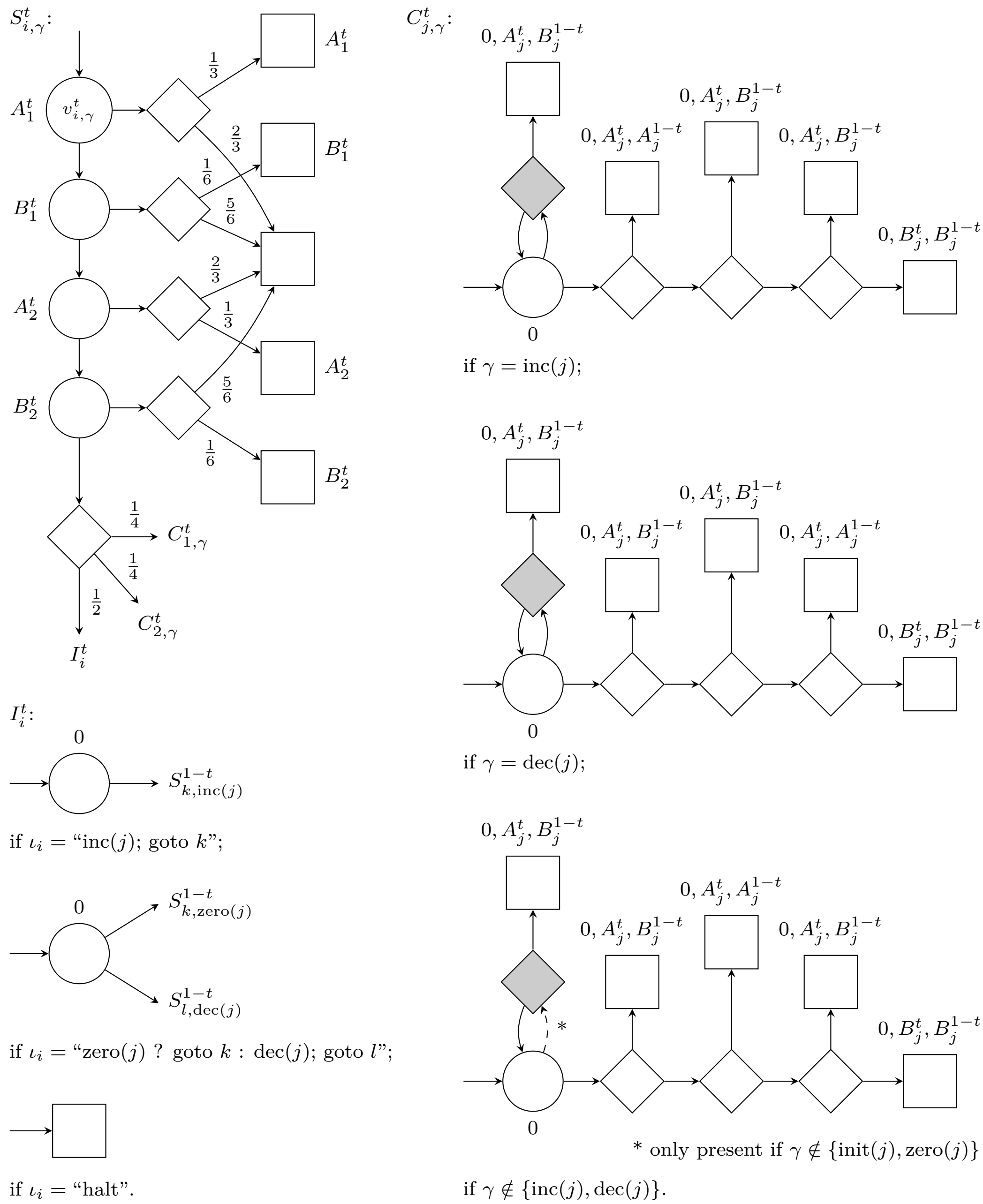


FIGURE 7. Simulating a two-counter machine.

dashed edge inside $C^t_{j,\gamma}$ is present if and only if $\gamma \notin \{\mathrm{init}, \mathrm{zero}(j)\}$. The initial vertex of $\mathcal{G}$ is the initial vertex of the gadget $S^0_{1,\mathrm{init}}$.

For any pure strategy profile $\overline{\sigma}$ of $\mathcal{G}$ where player 0 wins almost surely, let $x_0 v_0 \prec x_1 v_1 \prec x_2 v_2 \prec \ldots$ ($x_i \in V^*, v \in V$, $x_0 = \varepsilon$) be the (unique) sequence of all consecutive histories, consistent with $\overline{\sigma}$, such that, for each $n \in \mathbb{N}$, $v_n = v^t_{i,\gamma}$ for suitable i, γ, t; this sequence is infinite because player 0 wins almost surely. Additionally, let $\gamma_0, \gamma_1, \ldots$ be the corresponding sequence of instructions, i.e. $v_n = v^t_{i,\gamma_n}$. For each $j \in \{1, 2\}$ and $n \in \mathbb{N}$, we define two conditional probabilities a_j^n and p_j^n as follows:

$$a_j^n := \Pr_{v_0}^{\bar{\sigma}}(\text{player } A_j^{n \bmod 2} \text{ wins} \mid x_n v_n \cdot V^\omega),$$

and

$$p_j^n := \Pr_{v_0}^{\bar{\sigma}}(\text{player } A_j^{n \bmod 2} \text{ wins} \mid x_n v_n \cdot V^\omega \setminus x_{n+2} v_{n+2} \cdot V^\omega).$$

Finally, for each $j \in \{1, 2\}$ and $n \in \mathbb{N}$, we define an ordinal number $c_j^n \leq \omega$ as follows: After the history $x_n v_n$, with probability $\frac{1}{4}$ the play proceeds to the vertex controlled by player 0 in the counter gadget C_{j,γ_n}^t (where $t = n \bmod 2$). The number c_j^n is defined to be the maximal number of subsequent visits to the grey vertex inside this gadget (where $c_j^n = \omega$ if, on one path, the grey vertex is visited infinitely often). Note that, by the construction of $C_{j,\gamma}^t$, it holds that $c_j^n = 0$ if $\gamma_n = \text{zero}(j)$ or $\gamma_n = \text{init}$.

Lemma 4.8. Let $\bar{\sigma}$ be a pure strategy profile of $(\mathcal{G}, v_0)$ where player 0 wins almost surely. Then $\bar{\sigma}$ is a Nash equilibrium if and only if

$$c_j^{n+1} = \begin{cases} 1 + c_j^n & \text{if } \gamma_{n+1} = \text{inc}(j), \\ c_j^n - 1 & \text{if } \gamma_{n+1} = \text{dec}(j), \\ c_j^n = 0 & \text{if } \gamma_{n+1} = \text{zero}(j), \\ c_j^n & \text{otherwise,} \end{cases} \tag{4.1}$$

for each $j \in \{1, 2\}$ and all $n \in \mathbb{N}$.

Here, $+$ and $-$ denote the usual addition and subtraction of ordinal numbers, respectively (satisfying $1 + \omega = \omega - 1 = \omega$). The proof of Lemma 4.8 goes through several claims. In the following, let $\bar{\sigma}$ be a pure strategy profile of $(\mathcal{G}, v_0)$ where player 0 wins almost surely. The first claim gives a necessary and sufficient condition on the probabilities a_j^n for $\bar{\sigma}$ to be a Nash equilibrium.

Claim. The profile $\bar{\sigma}$ is a Nash equilibrium if and only if $a_j^n = \frac{1}{3}$ for each $j \in \{1, 2\}$ and all $n \in \mathbb{N}$.

Proof. ($\Rightarrow$) Assume that $\bar{\sigma}$ is a Nash equilibrium. Clearly, this implies that $a_j^n \geq \frac{1}{3}$ for all $n \in \mathbb{N}$ since otherwise some player A_j^t could improve her payoff by leaving one of the gadgets $S_{i,\gamma}^t$. Let

$$b_j^n := \Pr_{v_0}^{\bar{\sigma}}(\text{player } B_j^{n \bmod 2} \text{ wins} \mid x_n v_n \cdot V^\omega).$$

We have $b_j^n \geq \frac{1}{6}$ for all $n \in \mathbb{N}$ since otherwise some player B_j^t could improve her payoff by leaving one of the gadgets $S_{i,\gamma}^t$. Note that at every terminal vertex of the counter gadgets $C_{j,\gamma}^t$ and $C_{j,\gamma}^{1-t}$ either player A_j^t or player B_j^t wins. The conditional probability that, given the history $x_n v_n$, we reach one of those gadgets is $\sum_{k \in \mathbb{N}} \frac{1}{2^k} \cdot \frac{1}{4} = \frac{1}{2}$ for all $n \in \mathbb{N}$; so, $a_j^n = \frac{1}{2} - b_j^n$ for all $n \in \mathbb{N}$. Since $b_j^n \geq \frac{1}{6}$, we arrive at $a_j^n \leq \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$, which proves the claim.

($\Leftarrow$) Assume that $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$. Clearly, this implies that none of the players A_j^t can improve her payoff. To show that none of the players B_j^t can improve her payoff, it suffices to show that $b_j^n \geq \frac{1}{6}$ for all $n \in \mathbb{N}$. But with the same argumentation as above, we have $b_j^n = \frac{1}{2} - a_j^n$ and thus $b_j^n = \frac{1}{6}$ for all $n \in \mathbb{N}$, which proves the claim. $\square$

The second claim relates the probabilities a_j^n and p_j^n.

Claim. Let $j \in \{1,2\}$. Then $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$ if and only if $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$.

Proof. ($\Rightarrow$) Assume that $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$. We have $a_j^n = p_j^n + \frac{1}{4} \cdot a_j^{n+2}$ and therefore $\frac{1}{3} = p_j^n + \frac{1}{12}$ for all $n \in \mathbb{N}$. Hence, $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$.

($\Leftarrow$) Assume that $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$. Since $a_j^n = p_j^n + \frac{1}{4} \cdot a_j^{n+2}$ for all $n \in \mathbb{N}$, the numbers a_j^n must satisfy the following recurrence: $a_j^{n+2} = 4a_j^n - 1$. Since all the numbers a_j^n are probabilities, we have $0 \leq a_j^n \leq 1$ for all $n \in \mathbb{N}$. It is easy to see that the only values for a_j^0 and a_j^1 such that $0 \leq a_j^n \leq 1$ for all $n \in \mathbb{N}$ are $a_j^0 = a_j^1 = \frac{1}{3}$. But this implies that $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$. □

Finally, the last claim relates the numbers p_j^n to (4.1).

Claim. Let $j \in \{1,2\}$. Then $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$ if and only if (4.1) holds for all $n \in \mathbb{N}$.

Proof. Let $n \in \mathbb{N}$, and let $t = n \bmod 2$. The probability p_j^n can be expressed as the sum of the probability that the play reaches a terminal vertex that is winning for player A_j^t inside C_{j,γ_n}^t and the probability that the play reaches such a vertex inside $C_{j,\gamma_{n+1}}^{1-t}$. The first probability does not depend on γ_n, but the second depends on γ_{n+1}. Let us consider the case that $\gamma_{n+1} = \mathrm{inc}(j)$. In this case, the aforementioned sum is equal to

$$\frac{1}{4} \cdot \left(1 - \left(\frac{1}{2}\right)^{c_j^n+3}\right) + \frac{1}{8} \cdot \left(\frac{1}{2}\right)^{c_j^{n+1}+1} = \frac{1}{4} - \frac{1}{16} \cdot \left(\frac{1}{2}\right)^{c_j^n+1} + \frac{1}{16} \cdot \left(\frac{1}{2}\right)^{c_j^{n+1}}.$$

Obviously, this sum is equal to $\frac{1}{4}$ if and only if $c_j^{n+1} = 1 + c_j^n$. For any other value of γ_{n+1}, the argumentation is similar, and we omit it here. □

Proof of Lemma 4.8. By the first claim, the profile $\bar{\sigma}$ is a Nash equilibrium if and only if $a_j^n = \frac{1}{3}$ for all $j \in \{1,2\}$ and $n \in \mathbb{N}$. By the second claim, the latter is true if $p_j^n = \frac{1}{4}$ for all $j \in \{1,2\}$ and $n \in \mathbb{N}$. Finally, by the last claim, this is the case if and only if (4.1) holds for all $j \in \{1,2\}$ and $n \in \mathbb{N}$. □

To establish our reduction, it remains to show that the computation of $\mathcal{M}$ is infinite if and only if $(\mathcal{G}, v_0)$ has a pure Nash equilibrium where player 0 wins almost surely.

($\Rightarrow$) Assume that the computation $\rho = \rho(0)\rho(1)\ldots$ of $\mathcal{M}$ is infinite. We define a pure strategy σ_0 for player 0 as follows: For a history that ends in one of the instruction gadgets $I_{i,\gamma}^t$ after visiting a black vertex exactly n times, player 0 tries to move to the neighbouring gadget $S_{k,\gamma'}^{1-t}$ such that $\rho(n)$ refers to instruction number k (which is always possible if $\rho(n-1)$ refers to instruction number i; in any other case, σ_0 might be defined arbitrarily). In particular, if $\rho(n-1)$ refers to instruction $\iota_i =$ "zero(j) ? goto k : dec(j); goto l", then player 0 will move to the gadget $S_{k,\mathrm{zero}(j)}^{1-t}$ if the value of the counter in configuration $\rho(n-1)$ is 0 and to the gadget $S_{l,\mathrm{dec}(j)}^{1-t}$ otherwise. For a history that ends in one of the gadgets $C_{j,\gamma}^t$ after visiting a black vertex exactly n times and a grey vertex exactly m times, player 0 will move to the grey vertex again if and only if m is strictly less than the value of counter j in configuration $\rho(n-1)$. So after entering $C_{j,\gamma}^t$, player 0's strategy is to loop through the grey vertex exactly as many times as given by the value of counter j in configuration $\rho(n-1)$. Any other player's pure strategy is to "move down" at any time. We claim that the resulting strategy profile $\bar{\sigma}$ is a Nash equilibrium of $(\mathcal{G}, v_0)$ where player 0 wins almost surely.

Since, according to her strategy, player 0 follows the computation of $\mathcal{M}$, no vertex inside an instruction gadget $I^t_{i,\gamma}$ where ι_i is the halt instruction is ever reached in the Markov chain $(\mathcal{G}^{\overline{\sigma}}, v_0)$. Hence, and because all other players never leave the game, a terminal vertex in one of the counter gadgets is reached with probability 1. Since player 0 wins at any such vertex, we can conclude that she wins almost surely with $\overline{\sigma}$.

It remains to show that $\overline{\sigma}$ is a Nash equilibrium. By the definition of $\overline{\sigma}$, we have the following for all $n \in \mathbb{N}$: 1. c^n_j is the value of counter j in configuration $\rho(n)$; 2. c^{n+1}_j is the value of counter j in configuration $\rho(n+1)$; 3. γ_{n+1} is the instruction corresponding to the counter update from configuration $\rho(n)$ to $\rho(n+1)$. Hence, (4.1) holds, and $\overline{\sigma}$ is a Nash equilibrium by Lemma 4.8.

($\Leftarrow$) Assume that $\overline{\sigma}$ is a pure Nash equilibrium of $(\mathcal{G}, v_0)$ where player 0 wins almost surely. We define an infinite sequence $\rho = \rho(0)\rho(1)\ldots$ of *pseudo configurations* (where the counters may take the value ω) of $\mathcal{M}$ as follows. Let $n \in \mathbb{N}$, and assume that v_n lies inside the gadget S^t_{i,γ_n} (where $t = n \bmod 2$); then $\rho(n) := (i, c^n_1, c^n_2)$.

We claim that ρ is, in fact, the (infinite) computation of $\mathcal{M}$. It suffices to verify the following two properties:

– $\rho(0) = (1, 0, 0)$;
– $\rho(n) \vdash \rho(n+1)$ for all $n \in \mathbb{N}$.

Note that we do not have to show explicitly that each $\rho(n)$ is a configuration of $\mathcal{M}$ since this follows easily by induction. Verifying the first property is easy: v_0 lies inside $S^0_{1,\mathrm{init}}$ (and we are at instruction 1), which is linked to the counter gadgets $C^0_{1,\mathrm{init}}$ and $C^0_{2,\mathrm{init}}$. The edge leading to the grey vertex is missing in these gadgets. Hence, both c^0_1 and c^0_2 equal 0.

For the second property, let $\rho(n) = (i, c_1, c_2)$ and $\rho(n+1) = (i', c'_1, c'_2)$. Hence, v_n lies inside $S^t_{i,\gamma}$, and v_{n+1} lies inside $S^{1-t}_{i',\gamma'}$ for suitable γ, γ' and $t = n \bmod 2$. We only prove the claim for ι_i = "zero(2) ? goto k : dec(2); goto l"; the other cases are straightforward. Note that, by the construction of the gadget $I^t_{i,\gamma}$, it must be the case that either $i' = k$ and $\gamma' = \mathrm{zero}(2)$, or $i' = l$ and $\gamma' = \mathrm{dec}(2)$. By Lemma 4.8, if $\gamma' = \mathrm{zero}(2)$, then $c'_2 = c_2 = 0$ and $c'_1 = c_1$, and if $\gamma' = \mathrm{dec}(2)$, then $c'_2 = c_2 - 1$ and $c'_1 = c_1$. This implies $\rho(n) \vdash \rho(n+1)$: On the one hand, if $c_2 = 0$, then $c'_2 \neq c_2 - 1$, which implies $\gamma' \neq \mathrm{dec}(2)$ and thus $\gamma' = \mathrm{zero}(2)$, $i' = k$ and $c'_2 = c_2 = 0$. On the other hand, if $c_2 > 0$, then $\gamma' \neq \mathrm{zero}(2)$ and thus $\gamma' = \mathrm{dec}(2)$, $i' = l$ and $c'_2 = c_2 - 1$. □

Remark 4.9. The proof of Theorem 4.7 can also be viewed as a proof for the undecidability of the following problem: Given a labelled Markov decision process $(\mathcal{G}, v_0)$ and a formula φ of the logic PCTL [33], decide whether the controller has a strategy σ such that the Markov chain $(\mathcal{G}^\sigma, v_0)$ is a model of φ. Undecidability for this problem was first proven by Brázdil et al. [7] with a similar reduction than ours. However, we can derive a stronger result from our reduction, namely that there exists a *fixed* formula φ for which the problem is undecidable. To prove this, it suffices to modify our construction as follows:

1. For each player $A \in \{A^0_1, A^1_1, A^0_2, A^1_2\}$, we add one proposition A that holds at precisely those terminal vertices that are winning for player A.
2. For each $t = 0, 1$, we add one proposition Z^t that holds at each vertex of the form $v^t_{i,\gamma}$.
3. We modify the gadget $S^t_{i,\gamma}$ in such a way that all non-stochastic vertices are controlled by player 0.

4. We add two more propositions P and Q where P holds at all vertices controlled by player 0 and Q holds at precisely one successor of any such vertex.

Finally, the PCTL formula for which undecidability is proven is given by[4]

$$\varphi := \bigwedge_{t=0,1} \mathrm{G}^{=1}(Z^t \to \mathrm{F}^{=1/3} A_1^t \wedge \mathrm{F}^{=1/3} A_2^t) \wedge \mathrm{G}^{=1}(P \to \mathrm{X}^{=1} Q \vee \mathrm{X}^{=1} \neg Q).$$

4.4. **Finite-state equilibria.** We can use the construction in the proof of Theorem 4.7 to show that Nash equilibria may require infinite memory (even for SSMGs and if we are only interested in whether a player wins with probability 0 or 1).

Proposition 4.10. There exists an SSMG that has a pure Nash equilibrium where player 0 wins almost surely but that has no finite-state Nash equilibrium where player 0 wins with positive probability.

Proof. Consider the game $(\mathcal{G}, v_0)$ constructed in the proof of Theorem 4.7 for the machine $\mathcal{M}$ consisting of the single instruction "inc(1); goto 1". We modify this game by adding a new initial vertex v_1 which is controlled by a new player, player 1, and from where she can either move to v_0 or to a new terminal vertex where she receives payoff 1 and every other player receives payoff 0. Additionally, player 1 wins at every terminal vertex of the game $\mathcal{G}$ that is winning for player 0. Let us denote the modified game by $\mathcal{G}'$.

Since the computation of $\mathcal{M}$ is infinite, the game $(\mathcal{G}, v_0)$ has a pure Nash equilibrium where player 0 wins almost surely. This equilibrium induces a pure Nash equilibrium of $(\mathcal{G}', v_1)$ where both player 0 and player 1 win almost surely. Now assume that there exists a finite-state Nash equilibrium of $(\mathcal{G}', v_1)$ where player 0 wins with positive probability. Such an equilibrium induces a finite-state Nash equilibrium $\overline{\sigma}$ of $(\mathcal{G}, v_0)$ where player 1, and thus also player 0, wins almost surely: otherwise, player 1 would prefer to play to v_0 with probability 1. Using the same notation as in the proof of Theorem 4.7, it is easy to see that

$$\mathrm{Pr}_{v_0}^{\overline{\sigma}}(\text{player } B_1^{n \bmod 2} \text{ wins} \mid x_n v_n \cdot V^\omega \setminus x_{n+1} v_{n+1} \cdot V^\omega) = \frac{1}{2^{n+5}}$$

for each $n \in \mathbb{N}$. But this is impossible if $\overline{\sigma}$ uses only finite memory. □

It follows from Proposition 4.10 (together with Proposition 4.2) that the decision problems NE, FinNE, PureNE and PureFinNE are pairwise distinct. Another way to see that PureNE and PureFinNE are distinct is to observe that that PureFinNE is recursively enumerable: To decide whether an SSMG $(\mathcal{G}, v_0)$ has a pure finite-state Nash equilibrium with payoff $\geq \overline{x}$ and $\leq \overline{y}$, one can just enumerate all possible pure finite-state profiles $\overline{\sigma}$ and check for each of them whether it constitutes a Nash equilibrium with the desired properties by analysing the finite Markov chain $\mathcal{G}^{\overline{\sigma}}$. Hence, to prove the undecidability of PureFinNE, we cannot reduce from the non-halting problem; instead, we give a reduction from the halting problem (which is recursively enumerable itself).

Theorem 4.11. PureFinNE is undecidable, even for 13-player SSMGs.

Proof. The construction is similar to the one for proving undecidability of PureNE. Given a two-counter machine $\mathcal{M}$, we modify the SSMG $\mathcal{G}$ constructed in the proof of Theorem 4.7 by adding another "counter" (together with four more players for checking whether the counter is updated correctly) that has to be incremented in each step. Moreover, additionally to

[4]The last conjunct forces player 0 to use a pure strategy.

the terminal vertices in the gadgets $C^t_{j,\gamma}$, we let player 0 win at the terminal vertex in each of the gadgets $I_{i,\gamma}$ where ι_i = "halt". Let us denote the new game by $\mathcal{G}'$. Now, if $\mathcal{M}$ does not halt, any pure Nash equilibrium of $(\mathcal{G}', v_0)$ where player 0 wins almost surely needs infinite memory: to win with probability 1, player 0 must follow the computation of $\mathcal{M}$ and increment the new counter at each step. On the other hand, if $\mathcal{M}$ halts, we can easily construct a pure finite-state Nash equilibrium of $(\mathcal{G}', v_0)$ where player 0 wins almost surely (see the proof of Theorem 4.7; since $\mathcal{M}$ halts, the equilibrium described there can be implemented with finite memory). Hence, $(\mathcal{G}', v_0)$ has a pure finite-state Nash equilibrium where player 0 wins almost surely if and only if the machine $\mathcal{M}$ halts. □

4.5. **Randomised equilibria.** The decidability of NE and FinNE (with respect to arbitrary randomised strategies) remains an open question: We do not even know whether the SqrtSum-hardness result for StatNE can be adapted to NE. What we can show, however, is that NE and FinNE are NP-hard, even for SSMGs. Arguably, this indicates that these problems are intractable.

Theorem 4.12. NE and FinNE are NP-hard for SSMGs.

Proof. The proof is by a reduction from SAT: Given a Boolean formula φ in conjunctive normal form with clauses $C_1, \dots, C_m$, $m \geq 1$, over variables $X_1, \dots, X_n$ (where, without loss of generality, each clause is non-empty), we build an SSMG $\mathcal{G}$ played by players $0, 1, \dots, n$ as follows: $\mathcal{G}$ has vertices $C_1, \dots, C_m$ controlled by player 0, and for each clause C and each literal L, $L = X_i$ or $L = \neg X_i$, that occurs in C, a vertex (C, L), controlled by player i, and a stochastic vertex $(C, L, 1)$. Additionally, there are terminal vertices $\bot, \top, X_1, \dots, X_n$. There are edges from a clause C_j to each vertex (C_j, L) such that L occurs as a literal in C_j and from there to $(C_j, L, 1)$. From $(C_j, L, 1)$ there are two outgoing edges, each taken with probability $\frac{1}{2}$, one to L or $\top$ if L is a positive or a negative literal, respectively, and one to $C_{(j \bmod m)+1}$. Payoffs are assigned to terminal vertices as follows:

- Player 0 wins at every terminal vertex except $\bot$;
- Player $i \neq 0$ wins at every terminal vertex except X_i.

For the (satisfiable) formula $\varphi = (X_1 \vee X_2 \vee X_3) \wedge (\neg X_2 \vee X_3) \wedge \neg X_3$ the game $\mathcal{G}$ is schematically depicted in Figure 8; terminal vertices are again shown as squares.

Clearly, $\mathcal{G}$ can be constructed from φ in polynomial time. To establish both reductions, it suffices to show that the following three statements are equivalent:

1. φ is satisfiable.
2. $(\mathcal{G}, C_1)$ has a finite-state Nash equilibrium where player 0 wins almost surely.
3. $(\mathcal{G}, C_1)$ has a Nash equilibrium where player 0 wins almost surely.

(1. ⇒ 2.) Assume that $\alpha : \{X_1, \dots, X_k\} \to \{\text{true}, \text{false}\}$ is a satisfying assignment of φ. We show that the positional strategy profile $\overline{\sigma}$ where at any time player 0 plays from a clause C to a (fixed) vertex (C, L) such that $\alpha(L) = \text{true}$ and each player $i \neq 0$ never plays to $\bot$ is a Nash equilibrium of $(\mathcal{G}, C_1)$ where player 0 wins almost surely. First note that $\bot$ is reached with probability 0 when playing $\overline{\sigma}$; hence player 0 wins almost surely. Now consider any player $i \neq 0$ who wins with probability less than 1. This can only happen if the probability that the play ends at vertex X_i is > 0. But, as player 0 plays according to the satisfying assignment, this means that each vertex of the form $(C, \neg X_i)$ is visited with probability 0; hence, player i has no chance to improve her payoff by playing to $\bot$.

(2. ⇒ 3.) Trivial.

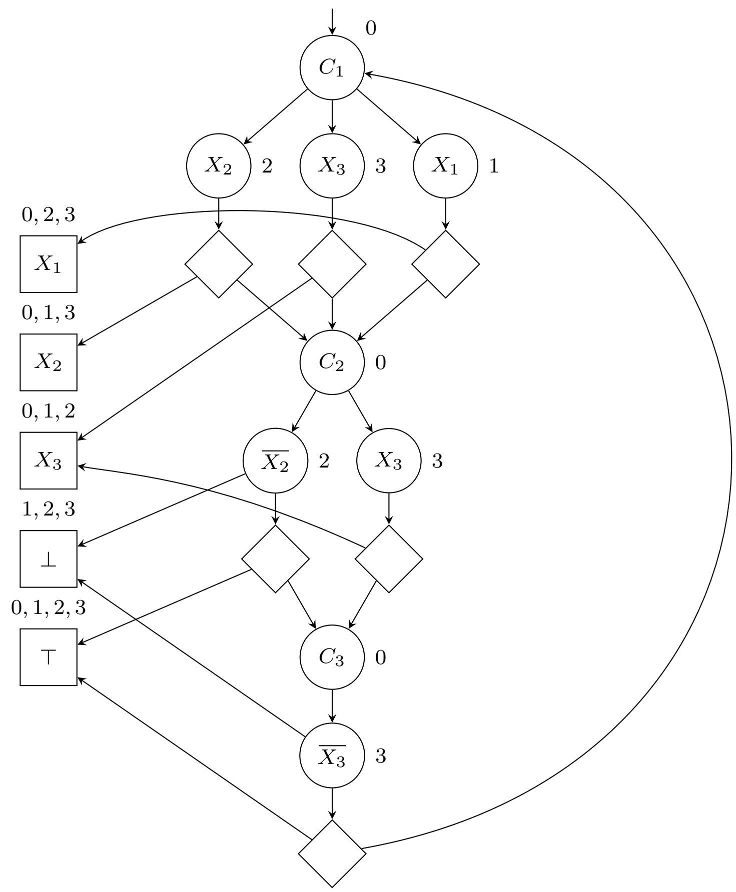


FIGURE 8. Reducing SAT to NE.

(3. $\Rightarrow$ 1.) Assume that $(\mathcal{G}, C_1)$ has a Nash equilibrium $\bar{\sigma}$ where player 0 wins almost surely. Thus, $\bot$ is visited with probability 0 when playing $\bar{\sigma}$. Hence, by the construction of $\mathcal{G}$, there exists an infinite play π of $(\mathcal{G}, C_1)$ that is consistent with $\bar{\sigma}$. We claim that it is not possible that both a vertex (C, X_i) and a vertex $(C', \neg X_i)$ are visited infinitely often in π. Towards a contradiction, assume that both (C, X_i) and $(C', \neg X_i)$ are visited infinitely often. Let $\pi|_n := \pi(0) \ldots \pi(n-1)$ be any prefix of π ending in $(C', \neg X_i)$. By the construction of π, the history $\pi|_n$ appears with positive probability when $\bar{\sigma}$ is played, and the conditional probability that player i wins given the history $\pi|_n$ is less than 1, since (by the construction of π) with positive probability the vertex $(C, X_i, 1)$ is visited later on, from where we move to vertex X_i with probability $\frac{1}{2}$. Consider the strategy σ' of player i that behaves like σ_i but moves to $\bot$ with probability 1 after history $\pi|_n$. With this strategy, the conditional probability that player 1 wins given the history $\pi|_n$ equals 1 while the conditional probability that player i wins given that $\pi|_n$ is not a history remains the same. Hence, player i can improve her payoff by switching to σ', a contradiction to the fact that $\bar{\sigma}$ is a Nash equilibrium.

Now consider the variable assignment that maps X_i to true if some vertex (C, X_i) is visited infinitely often in π. We claim that this assignment satisfies the formula. To see this, consider any clause C. By the construction of $\mathcal{G}$, there exists a literal L in C such that the vertex (C, L) is visited infinitely often in π. In both cases, the defined assignment maps the literal to true and satisfies C. $\square$

5. The strictly qualitative fragment

In this section, we prove that the fragment of NE that arises from restricting the thresholds to be the same binary payoff (i.e., each entry is either 0 or 1) is decidable for games with ω-regular objectives; we denote this problem by QualNE. Formally, QualNE is defined as follows:

> Given an SMG $(\mathcal{G}, v_0)$ and $\overline{x} \in \{0,1\}^\Pi$, decide whether there exists a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\overline{x}$.

To prove the decidability of QualNE, we first characterise the existence of a Nash equilibrium with a binary payoff in any game with prefix-independent objectives.

5.1. **Characterisation of existence.** Given an SMG $\mathcal{G}$ and a player i, we denote by $W_i^{>0}$ the set of all vertices $v \in V$ such that $\mathrm{val}_i^{\mathcal{G}}(v) > 0$.

Proposition 5.1. Let $(\mathcal{G}, v_0)$ be any SMG with prefix-independent objectives, and let $\overline{x} \in \{0,1\}^\Pi$. Then the following statements are equivalent:

1. There exists a Nash equilibrium with payoff $\overline{x}$.
2. There exists a strategy profile $\overline{\sigma}$ with payoff $\overline{x}$ such that $\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Reach}(W_i^{>0})) = 0$ for each player i with $x_i = 0$.
3. There exists a pure strategy profile $\overline{\sigma}$ with payoff $\overline{x}$ such that $\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Reach}(W_i^{>0})) = 0$ for each player i with $x_i = 0$.
4. There exists a pure Nash equilibrium with payoff $\overline{x}$.

If additionally all objectives are ω-regular, then each of the above statements is equivalent to each of the following statements:

5. There exists a pure finite-state strategy profile $\overline{\sigma}$ with payoff $\overline{x}$ such that $\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Reach}(W_i^{>0})) = 0$ for each player i with $x_i = 0$.
6. There exists a pure finite-state Nash equilibrium with payoff $\overline{x}$.

Proof. (1. $\Rightarrow$ 2.) Let $\overline{\sigma}$ be a Nash equilibrium with payoff $\overline{x}$. We claim that $\overline{\sigma}$ is already the strategy profile we are looking for: $\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Reach}(W_i^{>0})) = 0$ for each player i with $x_i = 0$. Towards a contradiction, assume that $\Pr_{v_0}^{\overline{\sigma}}(\mathrm{Reach}(W_i^{>0})) > 0$ for some player i with $x_i = 0$. Since V is finite, there exists a history xv, $v \in W_i^{>0}$, such that $\Pr_{v_0}^{\overline{\sigma}}(xv \cdot V^\omega) > 0$. Let τ be an optimal strategy for player i in the game $(\mathcal{G}, v)$, and consider the strategy σ' of player i defined by

$$\sigma'(yw) = \begin{cases} \sigma(yw) & \text{if } xv \not\preceq yw, \\ \tau(y_1 w) & \text{otherwise,} \end{cases}$$

where, in the latter case, $y = xy_1$. Clearly, $\Pr^{\overline{\sigma}}_{v_0}(xv \cdot V^\omega) = \Pr^{(\overline{\sigma}_{-i},\sigma')}_{v_0}(xv \cdot V^\omega)$. Using Lemma 2.1, we can infer that $\Pr^{\overline{\sigma}_{-i},\sigma'}_{v_0}(\mathrm{Win}_i) > 0$ as follows:

$$\begin{aligned}
&\Pr^{\overline{\sigma}_{-i},\sigma'}_{v_0}(\mathrm{Win}_i)\\
\geq\ &\Pr^{\overline{\sigma}_{-i},\sigma'}_{v_0}(\mathrm{Win}_i \cap xv \cdot V^\omega)\\
=\ &\Pr^{\overline{\sigma}_{-i},\sigma'}_{v_0}(xv \cdot V^\omega) \cdot \Pr^{(\overline{\sigma}_{-i},\sigma')[x]}_{v}(\mathrm{Win}_i)\\
=\ &\Pr^{\overline{\sigma}}_{v_0}(xv \cdot V^\omega) \cdot \Pr^{\overline{\sigma}_{-i}[x],\tau}_{v}(\mathrm{Win}_i)\\
\geq\ &\Pr^{\overline{\sigma}}_{v_0}(xv \cdot V^\omega) \cdot \mathrm{val}^{\mathcal{G}}_i(v)\\
>\ &0.
\end{aligned}$$

Hence, player i can improve her payoff by playing σ' instead of σ_i, a contradiction to the fact that $\overline{\sigma}$ is a Nash equilibrium.

(2. ⇒ 3.) Let $\overline{\sigma}$ be a strategy profile of $(\mathcal{G}, v_0)$ with payoff $\overline{x}$ such that $\Pr^{\overline{\sigma}}_{v_0}(\mathrm{Reach}(W^{>0}_i)) = 0$ for each player i with $x_i = 0$. Consider the MDP $\mathcal{M}$ that is obtained from $\mathcal{G}$ by removing all vertices $v \in V$ such that $v \in W^{>0}_i$ for some player i with $x_i = 0$, merging all players into one, and imposing the objective

$$\mathrm{Win} = \bigwedge_{\substack{i\in\Pi\\ x_i=1}} \mathrm{Win}_i \wedge \bigwedge_{\substack{i\in\Pi\\ x_i=0}} \neg\mathrm{Win}_i.$$

The MDP $\mathcal{M}$ is well-defined since its domain is a subarena of $\mathcal{G}$. Moreover, the value $\mathrm{val}^{\mathcal{M}}(v_0)$ of $\mathcal{M}$ is equal to 1 because the strategy profile $\overline{\sigma}$ induces a strategy σ in $\mathcal{M}$ satisfying $\Pr^{\sigma}_{v_0}(\mathrm{Win}) = 1$. Since each Win_i is prefix-independent, so is the set Win. Hence, by Theorem 2.5, $(\mathcal{M}, v_0)$ admits an optimal pure strategy τ. Since $\mathrm{val}^{\mathcal{M}}(v_0) = 1$, we have $\Pr^{\tau}_{v_0}(\mathrm{Win}) = 1$, and τ induces a pure strategy profile of $\mathcal{G}$ with the desired properties.

(3. ⇒ 4.) Let $\overline{\sigma}$ be a pure strategy profile of $(\mathcal{G}, v_0)$ with payoff $\overline{x}$ such that $\Pr^{\overline{\sigma}}_{v_0}(\mathrm{Reach}(W^{>0}_i)) = 0$ for each player i with $x_i = 0$. We show that the requirements of Lemma 3.4 are fulfilled: $\Pr^{\overline{\sigma}}_{v_0}(\mathrm{Win}_i \mid xv \cdot V^\omega) \geq \mathrm{val}^{\mathcal{G}}_i(v)$ for each player i and each history xv of $(\mathcal{G}, v_0)$ that is consistent with $\overline{\sigma}$. Let xv be such a history, and let $i \in \Pi$. By the assumption on $\overline{\sigma}$, we have $\mathrm{val}^{\mathcal{G}}_i(v) = 0$ or $x_i = 1$. In the first case, there is nothing to show. In the second case, we claim that $p := \Pr^{\overline{\sigma}}_{v_0}(\mathrm{Win}_i \mid xv \cdot V^\omega) = 1$ (and therefore $p \geq \mathrm{val}^{\mathcal{G}}_i(v)$). Towards a contradiction, assume that $p < 1$. Then

$$\begin{aligned}
x_i &= \Pr^{\overline{\sigma}}_{v_0}(\mathrm{Win}_i)\\
&= \Pr^{\overline{\sigma}}_{v_0}(xv \cdot V^\omega) \cdot p + \Pr^{\overline{\sigma}}_{v_0}(\mathrm{Win}_i \setminus xv \cdot V^\omega)\\
&< \Pr^{\overline{\sigma}}_{v_0}(xv \cdot V^\omega) + \Pr^{\overline{\sigma}}_{v_0}(V^\omega \setminus xv \cdot V^\omega)\\
&= 1,
\end{aligned}$$

a contradiction. Hence, Lemma 3.4 is applicable, and there exists a pure Nash equilibrium $\overline{\sigma}^*$ of $(\mathcal{G}, v_0)$ with $\Pr^{\overline{\sigma}}_{v_0} = \Pr^{\overline{\sigma}^*}_{v_0}$. In particular, $\overline{\sigma}^*$ has payoff $\overline{x}$.

(4. ⇒ 1.) Trivial.

Under the additional assumption that all objectives are ω-regular, the implications (2. ⇒ 5.) and (5. ⇒ 6.) are proven analogously (using Lemma 3.5 instead of Lemma 3.4); the implication (6. ⇒ 1.) is trivial. □

As an immediate consequence of Proposition 5.1, we can conclude that pure finite-state strategies are as powerful as arbitrary randomised strategies as far as the existence of a Nash equilibrium with a binary payoff in SMGs with ω-regular objectives is concerned.

Corollary 5.2. Let $(\mathcal{G}, v_0)$ be any SMG with ω-regular objectives, and let $x \in \{0,1\}^\Pi$. There exists a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ if and only if there exists a pure finite-state Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$.

Proof. The claim follows from Proposition 5.1 and the fact that every SMG with ω-regular objectives can be reduced to one with parity objectives. □

5.2. **Computational Complexity.** We can now give an algorithm that decides QualNE for SMGs with Muller objectives. The algorithm relies on the characterisation we gave in Proposition 5.1, which allows to reduce the problem to a problem about a certain MDP.

Formally, given a Muller SMG $\mathcal{G} = (\Pi, V, (V_i)_{i\in\Pi}, \Delta, \chi, (\mathcal{F}_i)_{i\in\Pi})$ and a binary payoff $\bar{x} \in \{0,1\}^\Pi$, we define the Markov decision process $\mathcal{G}(\bar{x})$ as follows: Let $Z \subseteq V$ be the set of all v such that $\mathrm{val}_i^{\mathcal{G}}(v) = 0$ for each player i with $x_i = 0$; the set of vertices of $\mathcal{G}(\bar{x})$ is precisely the set Z, with the set of vertices controlled by player 0 being $Z_0 := \bigcup_{i\in\Pi}(V_i \cap Z)$; if $Z = \emptyset$, we define $\mathcal{G}(\bar{x})$ to be a trivial MDP with the empty set as its objective. The transition relation of $\mathcal{G}(\bar{x})$ is the restriction of Δ to transitions between Z-states. Note that the transition relation of $\mathcal{G}(\bar{x})$ is well-defined since Z is a subarena of $\mathcal{G}$. Finally, the single objective in $\mathcal{G}(\bar{x})$ is $\mathrm{Reach}(T)$ where $T \subseteq Z$ is the union of all end components $U \subseteq Z$ with payoff $\bar{x}$.

Lemma 5.3. Let $(\mathcal{G}, v_0)$ be any Muller SMG, and let $\bar{x} \in \{0,1\}^\Pi$. Then $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\bar{x}$ if and only if $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$.

Proof. ($\Rightarrow$) Assume that $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\bar{x}$. By Proposition 5.1, this implies that there exists a strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(V \setminus Z)) = 0$. We claim that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(T)) = 1$. Otherwise, by Lemma 2.2, there would exist an end component $U \subseteq Z$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\{\alpha \in V^\omega : \mathrm{Inf}(\alpha) = U\}) > 0$, and U is either not winning for some player i with $x_i = 1$ or it is winning for some player i with $x_i = 0$. But then, $\bar{\sigma}$ cannot have payoff $\bar{x}$, a contradiction. Now, since $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(V \setminus Z)) = 0$, $\bar{\sigma}$ induces a strategy σ in $\mathcal{G}(\bar{x})$ such that $\mathrm{Pr}_{v_0}^{\sigma}(X) = \mathrm{Pr}_{v_0}^{\bar{\sigma}}(X)$ for every Borel set $X \subseteq Z^\omega$. In particular, $\mathrm{Pr}_{v_0}^{\sigma}(\mathrm{Reach}(T)) = 1$ and hence $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$.

($\Leftarrow$) Assume that $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$ (in particular, $v_0 \in Z$), and let σ be an optimal strategy in $(\mathcal{G}(\bar{x}), v_0)$. From σ, using Lemma 2.3, we can devise a strategy σ' such that $\mathrm{Pr}_{v_0}^{\sigma'}(\{\alpha \in V^\omega : \mathrm{Inf}(\alpha) \text{ has payoff } \bar{x}\}) = 1$. Finally, σ' can can be extended to a strategy profile $\bar{\sigma}$ of $\mathcal{G}$ with payoff $\bar{x}$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(V \setminus Z)) = 0$. By Proposition 5.1, this implies that $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\bar{x}$. □

Since the value of an MDP with reachability objectives can be computed in polynomial time, the difficult part lies in computing the MDP $\mathcal{G}(\bar{x})$ from $\mathcal{G}$ and $\bar{x}$ (i.e. its domain Z and the target set T).

Theorem 5.4. QualNE is in PSPACE for Muller SMGs.

Proof. We describe a polynomial-space algorithm for solving QualNE on Muller SMGs: On input $\mathcal{G}, v_0, \bar{x}$, the algorithm starts by computing for each player i with $x_i = 0$ the set of vertices v such that $\mathrm{val}_i^{\mathcal{G}}(v) = 0$, which can be done in polynomial space by Theorem 2.11. The intersection of these sets is the domain Z of the Markov decision process $\mathcal{G}(\bar{x})$. If v_0 is not contained in this intersection, the algorithm immediately rejects. Otherwise, the algorithm determines the union T of all end components with payoff $\bar{x}$ contained in Z by enumerating all subsets of Z one after another and checking which ones are end components with payoff $\bar{x}$. Finally, the algorithm computes (in polynomial time) the value $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0)$ of the MDP $\mathcal{G}(\bar{x})$ and accepts if the value is 1. In all other cases, the algorithm rejects. The correctness of the algorithm follows immediately from Lemma 5.3. □

Since the qualitative decision problem for Muller S2Gs is already PSPACE-hard, it follows from Theorem 5.4 that QualNE is, in fact, PSPACE-complete for Muller games. By contrast, for SMGs with Streett objectives, the problem becomes NP-complete. First, we prove the upper bound.

Theorem 5.5. QualNE is in NP for Streett SMGs.

Proof. We describe a nondeterministic polynomial-time algorithm for solving QualNE: On input $\mathcal{G}, v_0, \bar{x}$, the algorithm starts by guessing a subarena $Z' \subseteq V$ and, for each player i with $x_i = 0$, a positional strategy τ_i of the coalition $\Pi \setminus \{i\}$ in the S2G $\mathcal{G}_i$. In the next step, the algorithm checks (in polynomial time) whether $\mathrm{val}^{\tau_i}_{\Pi\setminus\{i\}}(v) = 1$ for each vertex $v \in Z'$ and each player i with $x_i = 0$. If not, the algorithm rejects immediately. Otherwise, the algorithm proceeds by guessing (at most) $n := |V|$ subsets $U_1, \dots, U_n \subseteq Z'$ and checks whether they are end components with payoff $\bar{x}$ (which can be done in polynomial time). If yes, the algorithm sets $T' := \bigcup_{j=1}^{n} U_j$ and computes (in polynomial time) the value $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0)$ of the MDP $\mathcal{G}(\bar{x})$ with Z' substituted for Z and T' substituted for T. If this value is 1, the algorithm accepts; otherwise, it rejects.

It remains to show that the algorithm is correct: On the one hand, if $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\bar{x}$, then the run of the algorithm where it guesses $Z' = Z$, globally optimal positional strategies τ_i (which exist by Theorem 2.6) and end components U_i such that $T' = T$ will be accepting since then, by Lemma 5.3, $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$. On the other hand, in any accepting run of the algorithm we have $Z' \subseteq Z$ and $T' \subseteq T$, and the value that the algorithm computes cannot be higher than $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0)$; hence, $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$, and Lemma 5.3 guarantees the existence of a Nash equilibrium with payoff $\bar{x}$. □

The matching lower bound does even hold for deterministic 2-player Streett games and was established in [54].

Theorem 5.6. QualNE is NP-hard for deterministic 2-player Streett games.

Proof. The proof is a variant of the proof for NP-hardness of the qualitative decision problem for deterministic 2-player zero-sum Rabin–Streett games [22] and by a reduction from SAT. Given a Boolean formula $\varphi = C_1 \wedge \ldots \wedge C_m$, $m \geq 1$, in conjunctive normal form (where, without loss of generality, each clause is non-empty), we construct a deterministic 2-player Streett game $\mathcal{G}$ as follows: For each clause C the game $\mathcal{G}$ has a vertex C, which is controlled by player 0, and for each literal L occurring in φ there is a vertex L, which is controlled by player 1. There are edges from a clause to each literal that occurs in this clause, and from a literal to each clause occurring in φ. (Without loss of generality, we assume that there is at least one clause and that all clauses are non-empty.) The structure of the game is depicted in Figure 9. Player 0's objective is given by the single Streett pair $(\emptyset, V)$, i.e. she wins every play of the game, whereas player 1's objective consists of all Streett pairs of the

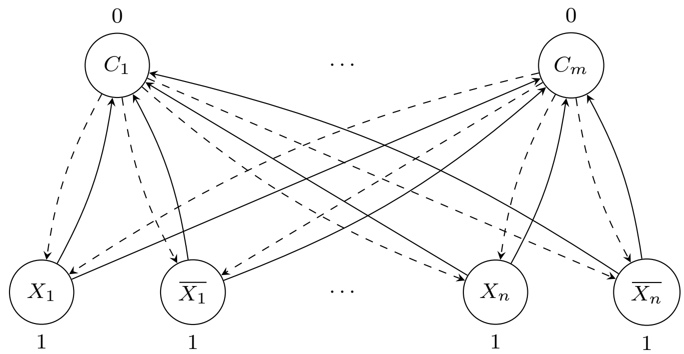


FIGURE 9. Reducing SAT to QualNE for games with Streett objectives.

form $(\{X\},\{\neg X\})$ or $(\{\neg X\},\{X\})$, i.e. she wins if, for each variable X, either X and $\neg X$ are both visited infinitely often or neither of them is.

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 Nash equilibrium with payoff $(1, 0)$.

($\Rightarrow$) Assume that φ is satisfiable, and consider the following positional strategy σ_0 of player 0: Whenever the play reaches a clause, then σ_0 plays to a literal that is mapped to true by the satisfying assignment. This strategy ensures that, for each variable X, at most one of the literals X or $\neg X$ is visited infinitely often. Hence, (σ_0, σ_1) is a Nash equilibrium of $(\mathcal{G}, C_1)$ with payoff $(1, 0)$ for every (positional) strategy σ_1 of player 1.

($\Leftarrow$) Let (σ_0, σ_1) be a Nash equilibrium of $(\mathcal{G}, C_1)$ with payoff $(1, 0)$, and assume that φ is not satisfiable. Consider the coalition game $(\mathcal{G}_1, C_1)$, a Rabin-Streett game. We claim that player 1 does have a winning strategy in this game, which she could use to improve her payoff in $(\mathcal{G}, C_1)$, a contradiction to the fact that (σ_0, σ_1) is a Nash equilibrium. By determinacy, it suffices to show that player 0 does not have a winning strategy. Let τ be an optimal positional strategy of player 0 (which exists by Theorem 2.6). If player 0 has a winning strategy, then τ must be winning as well. Since φ is unsatisfiable, there must exist a variable X and clauses C and C' such that $\tau(C) = X$ and $\tau(C') = \neg X$. But player 1 can counter this strategy by playing from X to C' and from any other literal to C. Hence, τ is not a winning strategy. □

For games with Rabin objective, the situation is more delicate. One might think that, because of the duality of Rabin and Streett objectives, QualNE would be in coNP for SMGs with Rabin objectives.[5] However, as we will see later, this is rather unlikely, and we can only show that the problem lies in the class $\mathrm{P}^{\mathrm{NP}[\log]}$ of problems solvable by a deterministic polynomial-time algorithm that may perform at most $\mathrm{O}(\log n)$ queries to an oracle for any problem in NP. In fact, the same upper bound holds for games with a Streett *or* a Rabin objective for each player.

Theorem 5.7. QualNE is in $\mathrm{P}^{\mathrm{NP}[\log]}$ for Streett-Rabin SMGs.

Proof. Let us describe a polynomial-time algorithm performing a logarithmic number of queries to an NP oracle for the problem. On input $\mathcal{G}, v_0, \overline{x}$, the algorithm starts by determining for each vertex v and each player i with $x_i = 0$ who has a Rabin objective whether $\mathrm{val}_i^{\mathcal{G}}(v) = 0$. Naively implemented, this requires a super-logarithmic number of queries to

[5]In fact, in [55] we claimed that the problem *is* in coNP.

the oracle. To reduce the number of queries, we use a neat trick, due to Hemachandra [34]. Let us denote by R and S the set of players $i \in \Pi$ with $x_i = 0$ that have a Rabin and a Streett objective, respectively. Instead of looping through all pairs of a vertex and a player, we begin by determining the number p of all pairs (v, i) such that $i \in R$ and $\mathrm{val}_i^{\mathcal{G}}(v) = 0$. It is not difficult to see that this number can be computed using binary search by performing only $\mathrm{O}(\log n)$ queries to an NP oracle, which we can use for deciding whether $\mathrm{val}_i^{\mathcal{G}}(v) > 0$ (by Corollary 2.10). Then we perform one more query: We ask whether there exist sets $(Z_i)_{i \in R \cup S}$, $Z_i \subseteq V$, $U_1, \ldots, U_n \subseteq V$, $n = |V|$, and positional strategies $(\sigma_i)_{i \in R}$ and $(\tau_i)_{i \in S}$, where σ_i is a strategy of player i and τ_i is a strategy of the coalition $\Pi \setminus \{i\}$, such that

1. $\sum_{i \in R} |Z_i| = p$,
2. $\mathrm{val}^{\sigma_i}(v) > 0$ for each player $i \in R$ and each $v \in V \setminus Z_i$,
3. $\mathrm{val}^{\tau_i}(v) = 1$ for each player $i \in S$ and each $v \in Z_i$,
4. each U_j is an end component of $\mathcal{G}$ with payoff $\bar{x}$, and
5. the value of the MDP that results from $\mathcal{G}$ by restricting to vertices inside $\bigcap_{i \in R \cup S} Z_i$ and imposing the objective $\mathrm{Reach}(\bigcup_{j=1}^{n} U_j)$ is 1.

This query can be decided by an NP oracle by guessing suitable sets and strategies and verifying 1.–5. in polynomial time. If the answer to the query is *yes*, the algorithm accepts, otherwise it rejects.

Obviously, the algorithm runs in polynomial time. To see that the algorithm is correct, first note that for each player $i \in R$ the set Z_i is precisely the set of all $v \in V$ such that $\mathrm{val}_i^{\mathcal{G}}(v) = 0$. Otherwise, there would exist a vertex $v \in Z_i$ such that $\mathrm{val}_i^{\mathcal{G}}(v) > 0$. But then the number of pairs (v, i) with $i \in R$ and $\mathrm{val}_i^{\mathcal{G}}(v) = 0$ would be strictly less than p, a contradiction. Now, the correctness of the algorithm follows with the same reasoning as in the proof of Theorem 5.5. □

Remark 5.8. For a bounded number of players, QualNE is in coNP for SMGs with Rabin objectives.

Regarding lower bounds for QualNE in SMGs with Rabin objectives, we start by proving that the problem is coNP-hard, even for deterministic 2-player games. In particular, unless NP = coNP, QualNE cannot lie in NP for SMGs with Rabin objectives.

Theorem 5.9. QualNE is coNP-hard for deterministic 2-player Rabin games.

Proof. The proof is similar to the proof of Theorem 5.6 and by a reduction from the *unsatisfiability* problem for Boolean formulae in conjunctive normal form.

Given a Boolean formula $\varphi = C_1 \wedge \ldots \wedge C_m$, $m \geq 1$, in conjunctive normal form (where, without loss of generality, each clause is non-empty), we construct a deterministic 2-player Rabin game $\mathcal{G}$ as follows: The arena of $\mathcal{G}$ is the same as in the proof of Theorem 5.6, depicted in Figure 9: For each clause C there is a vertex C, which is controlled by player 0, and for each literal L occurring in φ there is a vertex L, which is controlled by player 1. There are edges from a clause to each literal that occurs in this clause, and from a literal to each clause occurring in φ. Player 1 wins every play of the game, whereas player 0's objective consists of all Rabin pairs of the form $(\{X\}, \{\neg X\})$ or $(\{\neg X\}, \{X\})$.

Clearly, $\mathcal{G}$ can be constructed from φ in polynomial time. We claim that φ is unsatisfiable if and only if $(\mathcal{G}, C_1)$ has a Nash equilibrium with payoff $(0, 1)$.

($\Rightarrow$) Assume that φ is unsatisfiable, and consider the 2-player zero-sum game $\mathcal{G}_0$ where player 1's objective is the complement of player 0's objective. Let σ_1 be a globally optimal strategy for player 1 in this game. We claim that σ_1 is winning in $(\mathcal{G}_0, C_1)$. Consequently,

(σ_0, σ_1) is a Nash equilibrium of $(\mathcal{G}, C_1)$ for every strategy σ_0 of player 0. Otherwise, let τ be a globally optimal positional strategy for player 0 in $\mathcal{G}_0$ (which exists by Theorem 2.6). By determinacy, τ would be winning in $(\mathcal{G}_0, C_1)$. But a positional strategy τ of player 0 picks for each clause a literal contained in this clause. Since φ is unsatisfiable, there must exist a variable X and clauses C and C' such that $\tau(C) = X$ and $\tau(C') = \neg X$. Player 1 could counter this strategy by playing from X to C' and from any other literal to C, a contradiction.

($\Leftarrow$) Let (σ_0, σ_1) be a Nash equilibrium of $(\mathcal{G}, C_1)$ with payoff $(0, 1)$, and assume that φ is satisfiable. Consider the following positional strategy τ of player 0: Whenever the play reaches a clause, then τ plays to a literal that is mapped to true by the satisfying assignment. This strategy ensures that for each variable X at most one of the literals X or $\neg X$ is visited infinitely often. Since the construction of $\mathcal{G}$ ensures that, under any strategy profile, at least one literal is visited infinitely often, τ ensures a winning play for player 0. Hence, player 0 can improve her payoff by playing τ instead of σ_0, a contradiction to the fact that (σ_0, σ_1) is a Nash equilibrium. □

Theorem 5.9 leaves open the possibility that QualNE is not only coNP-hard for SMGs with Rabin objectives, but also coNP-complete, even if the numbers of players is unbounded. However, unless NP = coNP, this is not the case because QualNE is also NP-hard for SMGs with Rabin objectives. In fact, it is even NP-hard to decide whether in a deterministic Rabin game there exists a play that fulfils the objective of each player.

Proposition 5.10. The problem of deciding, given a deterministic Rabin game, whether there exists a play that is won by each player is NP-hard.

Proof. We reduce from SAT: Given a Boolean formula $\varphi = C_1 \wedge \ldots \wedge C_m$, $m \geq 1$, in conjunctive normal form over propositional variables $X_1, \ldots, X_n$ (where, without loss of generality, every clause is non-empty), we show how to construct (in polynomial time) a deterministic $(n+1)$-player Rabin game $\mathcal{G}$ such that φ is satisfiable if and only if there exists a play of $\mathcal{G}$ that is won by each player. The game has vertices $C_1, \ldots, C_m$ and, for each clause C and each literal X or $\neg X$ that occurs in C, a vertex (C, X) or $(C, \neg X)$, respectively. All vertices are controlled by player 0. There are edges from a clause C_j to each vertex (C_j, L) such that L occurs in C_j and from there to $C_{(j \bmod m)+1}$. The arena of $\mathcal{G}$ is schematically depicted in Figure 10. The Rabin objectives are defined as follows:

- Player 0 wins every play of $\mathcal{G}$.
- Player $i \neq 0$ wins if each vertex of the form (C, X_i) is visited only finitely often *or* each vertex of the from $(C, \neg X_i)$ is visited only finitely often (two Rabin pairs).

Clearly, $\mathcal{G}$ can be constructed from φ in polynomial time. To establish the reduction, it remains to show that φ is satisfiable if and only if there exists a play of $\mathcal{G}$ that is won by each player.

($\Rightarrow$) Assume that $\alpha : \{X_1, \ldots, X_n\} \to \{\text{true}, \text{false}\}$ is a satisfying assignment of φ. Clearly, the positional strategy of player 0 where, from each clause C, she plays to a fixed vertex (C, L) such that L is mapped to true by α induces a play that is won by each player

($\Leftarrow$) Assume that there exists a play π of $\mathcal{G}$ that is won by each player. Obviously, it is not possible that both a vertex (C, X_i) and a vertex $(C', \neg X_i)$ are visited infinitely often in π since this would mean that player i loses π. Now consider the variable assignment that maps X to true if some vertex (C, X) is visited infinitely often in π. This assignment satisfies the formula because, by the construction of $\mathcal{G}$, for each clause C there exists a literal L in C such that the vertex (C, L) is visited infinitely often in π. □

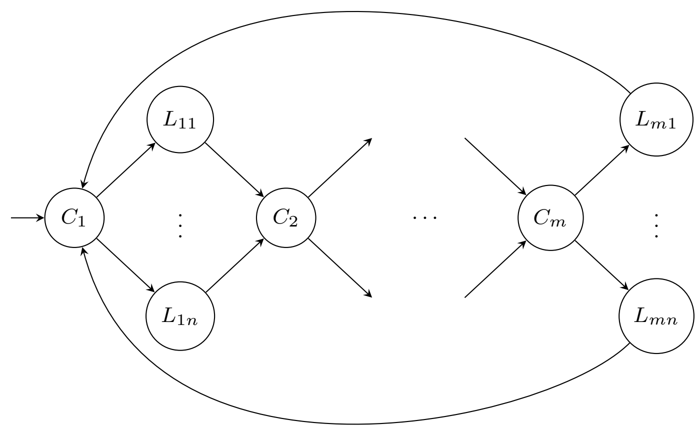


FIGURE 10. Reducing SAT to deciding the existence of a play winning for all players in a deterministic Rabin game.

It follows from Theorem 5.9 and Proposition 5.10 that, unless NP = coNP, QualNE is not contained in NP $\cup$ coNP, even for deterministic Rabin games. For stochastic Rabin games, we can show a completeness result: for these games, QualNE is also hard for $\mathrm{P}^{\mathrm{NP}[\log]}$.

Theorem 5.11. QualNE is $\mathrm{P}^{\mathrm{NP}[\log]}$-hard for Rabin SMGs.

Proof. Wagner [57] and, independently, Buss and Hay [8] showed that $\mathrm{P}^{\mathrm{NP}[\log]}$ equals the closure of NP with respect to *polynomial-time Boolean formula reducibility*. The canonical complete problem for this class is to decide, given a Boolean combination α of statements of the form "φ is satisfiable" (where φ ranges over all Boolean formulae) whether α evaluates to true. We claim that for every such statement α we can construct (in polynomial time) a Rabin SMG $(\mathcal{G}, v_0)$ such that α evaluates to true if and only if $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $(0, 1, \ldots, 1)$. The game $\mathcal{G}$ is constructed by induction on the complexity of α, where we assume without loss of generality that negations are only applied to atoms. If α is of the form "φ is satisfiable" or "φ is *not* satisfiable", then the existence of a suitable game $\mathcal{G}$ follows from Proposition 5.10 or Theorem 5.9, respectively.

Now, let $\alpha = \alpha_1 \wedge \alpha_2$, and assume that we already have constructed suitable games $(\mathcal{G}_1, v_1)$ and $(\mathcal{G}_2, v_2)$ (without loss of generality played by the same players $0, 1, \ldots, n$). The game $\mathcal{G}$ is the disjoint union of $\mathcal{G}_1$ and $\mathcal{G}_2$ combined with one new stochastic vertex v_0. From v_0, the game moves with probability $\frac{1}{2}$ each to v_1 and v_2. Obviously, $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $(0, 1, \ldots, 1)$ if and only if both $(\mathcal{G}_1, v_1)$ and $(\mathcal{G}_2, v_2)$ have such an equilibrium.

Finally, let $\alpha = \alpha_1 \vee \alpha_2$, and assume that we already have constructed suitable games $(\mathcal{G}_1, v_1)$ and $(\mathcal{G}_2, v_2)$ (again without loss of generality played by the same players $0, 1, \ldots, k$, $k \geq 1$). As in the previous case, the game $\mathcal{G}$ is the disjoint union of $\mathcal{G}_1$ and $\mathcal{G}_2$ combined with one new vertex v_0, which has transitions to both v_1 and v_2. However, this time v_0 is controlled by player 1. Obviously, $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $(0, 1, \ldots, 1)$ if and only if at least one of the games $(\mathcal{G}_1, v_1)$ and $(\mathcal{G}_2, v_2)$ has such an equilibrium. $\square$

To solve QualNE for parity SMGs, we will employ Algorithm 5.1, which computes for a game $\mathcal{G}$ with parity objectives Ω_i, $i \in \Pi$, $S \subseteq V$, and $\bar{x} \in \{0,1\}^{\Pi}$ the union of all end components with payoff $\bar{x}$ that are contained in S. The algorithm is a straightforward

Algorithm 5.1. Computing the union of all end components with payoff $\bar{x}$ contained in U.

input parity SMG $\mathcal{G} = (\Pi, V, (V_i)_{i\in\Pi}, \Delta, \chi, (\Omega_i)_{i\in\Pi})$, $S \subseteq V$, $\bar{x} \in \{0,1\}^{\Pi}$
return FindEC(S)

procedure FindEC(X)
  $Z := \emptyset$
  Compute all end components of $\mathcal{G}$ maximal in X
  for each such end component U **do**
    $P := \{i \in \Pi : \min \Omega_i(\chi(U)) \equiv x_i \bmod 2\}$
    if $P = \emptyset$ **then**
      (* U is an end component with payoff $\bar{x}$ *)
      $Z := Z \cup U$
    else
      (* U has the wrong payoff *)
      $Y := \bigcap_{i\in P}\{v \in U : \Omega_i(\chi(v)) > \min \Omega_i(\chi(U))\}$
      $Z := Z \cup \text{FindEC}(Y)$
    end if
  end for
  return Z
end procedure

adaptation of the algorithm for computing the union of all winning end components in a Streett MDP [14].

Note that on input X, FindEC calls itself at most $|X|$ times. Since, additionally, the set of all end components maximal in a set X can be computed in polynomial time, this proves that Algorithm 5.1 runs in polynomial time (all other operations are simple).

Theorem 5.12. QualNE is in NP $\cap$ coNP for parity SMGs.

Proof. Since any parity objective can be translated into an equivalent Streett objective in polynomial time, membership in NP follows from Theorem 5.5. To prove membership in coNP, we describe a nondeterministic polynomial-time algorithm for the complement of QualNE. On input $\mathcal{G}, v_0, \bar{x}$, the algorithm starts by guessing a subarena $Z' \subseteq V$ and, for each player i with $x_i = 0$, a positional strategy σ_i of player i in $\mathcal{G}$. In the next step, the algorithm checks whether for each vertex $v \in Z'$ there exists some player i with $x_i = 0$ and $\text{val}_i^{\sigma_i}(v) > 0$. If not, the algorithm rejects immediately. Otherwise, the algorithm uses Algorithm 5.1 to determine the union T' of all end components with payoff $\bar{x}$ that are contained in $V \setminus Z'$. Finally, the algorithm computes (in polynomial time) the value $\text{val}^{\mathcal{G}(\bar{x})}(v_0)$ of the MDP $\mathcal{G}(\bar{x})$ with $V \setminus Z'$ substituted for Z and T' substituted for T. If this value is not 1, the algorithm accepts; otherwise, it rejects. The correctness of the algorithm is proven in a similar fashion as in the proof of Theorem 5.5. □

Recall from Section 2.6 that it is a major open problem whether the qualitative decision problem for S2Gs with parity objectives is in P. This would imply that QualNE is decidable in polynomial time for games with parity objectives since this would allow us to compute the domain of the MDP $\mathcal{G}(\bar{x})$ in polynomial time. For each $d \in \mathbb{N}$, a class of games where the qualitative decision problem is provably in P is the class of all parity S2Gs with a bounded number of priorities (Theorem 2.12). Hence, we can solve QualNE for parity SMGs with a bounded number of priorities (and thus also for (co-)Büchi SMGs) in polynomial time as well.

Theorem 5.13. For each $d \in \mathbb{N}$, QualNE is in P for parity SMGs whose objectives refer to at most d priorities. □

References

[1] E. Allender, P. Bürgisser, J. Kjeldgaard-Pedersen, and P. B. Miltersen. On the complexity of numerical analysis. In *Proceedings of the 21st Annual IEEE Conference on Computational Complexity, CCC '06*, pages 331–339. IEEE Computer Society Press, 2006.

[2] D. Andersson and P. B. Miltersen. The complexity of solving stochastic games on graphs. In *Proceedings of the 20th International Symposium on Algorithms and Computation, ISAAC 2009*, 2009. To appear.

[3] R. J. Aumann. Survey of repeated games. In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, pages 11–42. Bibliographisches Institut Mannheim/Wien/Zürich, 1981.

[4] R. E. Bellman. *Dynamic Programming.* Princeton University Press, 1957.

[5] H. Björklund and S. Vorobyov. Combinatorial structure and randomized subexponential algorithms for infinite games. *Theoretical Computer Science*, 349:347–360, 2005.

[6] H. Björklund, S. Sandberg, and S. Vorobyov. A discrete subexponential algorithm for parity games. In *Proceedings of the 20th Annual Symposium on Theoretical Aspects of Computer Science, STACS 2003*, volume 2607 of *LNCS*, pages 663–674. Springer-Verlag, 2003.

[7] T. Brázdil, V. Brožek, V. Forejt, and A. Kučera. Stochastic games with branching-time winning objectives. In *Proceedings of the 21st IEEE Symposium on Logic in Computer Science, LICS 2006*, pages 349–358. IEEE Computer Society Press, 2006.

[8] S. R. Buss and L. Hay. On truth-table reducibility to SAT. *Information and Computation*, 91(1):86–102, 1991.

[9] J. Canny. Some algebraic and geometric computations in PSPACE. In *Proceedings of the 20th annual ACM Symposium on Theory of Computing, STOC '88*, pages 460–469. ACM Press, 1988.

[10] K. Chatterjee. Stochastic Müller games are PSPACE-complete. In *Proceedings of the 27th International Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2007*, volume 4855 of *LNCS*, pages 436–448. Springer-Verlag, 2007.

[11] K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Simple stochastic parity games. In *Proceedings of the 12th Annual Conference of the European Association for Computer Science Logic, CSL 2003*, volume 2803 of *LNCS*, pages 100–113. Springer-Verlag, 2003.

[12] K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Quantitative stochastic parity games. In *Proceedings of the 15th ACM-SIAM Symposium on Discrete Algorithms, SODA 2004*, pages 121–130. ACM Press, 2004.

[13] K. Chatterjee, R. Majumdar, and M. Jurdziński. On Nash equilibria in stochastic games. In *Proceedings of the 13th Annual Conference of the European Association for*

Computer Science Logic, CSL 2004, volume 3210 of *LNCS*, pages 26–40. Springer-Verlag, 2004.

[14] K. Chatterjee, L. de Alfaro, and T. A. Henzinger. The complexity of stochastic Rabin and Streett games. In *Proceedings of the 32nd International Colloquium on Automata, Languages and Programming, ICALP 2005*, volume 3580 of *LNCS*, pages 878–890. Springer-Verlag, 2005.

[15] X. Chen and X. Deng. Settling the complexity of two-player Nash equilibrium. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2006*, pages 261–272. IEEE Computer Society Press, 2006.

[16] A. Condon. The complexity of stochastic games. *Information and Computation*, 96(2): 203–224, 1992.

[17] V. Conitzer and T. Sandholm. Complexity results about Nash equilibria. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI 2003*, pages 765–771. Morgan Kaufmann, 2003.

[18] C. A. Courcoubetis and M. Yannakakis. Markov decision processes and regular events. *IEEE Transactions on Automatic Control*, 43(10):1399–1418, 1998.

[19] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a Nash equilibrium. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, STOC 2006*, pages 71–78. ACM Press, 2006.

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

[21] L. de Alfaro and T. A. Henzinger. Concurrent omega-regular games. In *Proceedings of the 15th IEEE Symposium on Logic in Computer Science, LICS 2000*, pages 141–154. IEEE Computer Society Press, 2000.

[22] E. A. Emerson and C. S. Jutla. The complexity of tree automata and logics of programs (extended abstract). In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science, FoCS '88*, pages 328–337. IEEE Computer Society Press, 1988.

[23] E. A. Emerson and C. S. Jutla. Tree automata, mu-calculus and determinacy (extended abstract). In *Proceedings of the 32nd Annual Symposium on Foundations of Computer Science, FoCS '91*, pages 368–377. IEEE Computer Society Press, 1991.

[24] K. Etessami and M. Yannakakis. On the complexity of Nash equilibria and other fixed points. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2007*, pages 113–123. IEEE Computer Society Press, 2007.

[25] K. Etessami, M. Z. Kwiatkowska, M. Y. Vardi, and M. Yannakakis. Multi-objective model checking of Markov decision processes. *Logical Methods in Computer Science*, 4 (4), 2008.

[26] J. Filar and K. Vrieze. *Competitive Markov decision processes*. Springer-Verlag, 1997.

[27] O. Friedmann. An exponential lower bound for the parity game strategy improvement algorithm as we know it. In *Proceedings of the 24th IEEE Symposium on Logic in Computer Science, LICS 2009*, pages 145–156. IEEE Computer Society Press, 2009.

[28] M. R. Garey, R. L. Graham, and D. S. Johnson. Some NP-complete geometric problems. In *Proceedings of the 8th Annual ACM Symposium on Theory of Computing, STOC '76*, pages 10–22. ACM Press, 1976.

[29] H. Gimbert and F. Horn. Solving simple stochastic games with few random vertices. *Logical Methods in Computer Science*, 5(2), 2009.

[30] H. Gimbert and F. Horn. Solving simple stochastic tail games. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms, SODA 2010*, 2010. To appear.

[31] E. Grädel, W. Thomas, and T. Wilke, editors. *Automata, Logics, and Infinite Games*, volume 2500 of *LNCS*. Springer-Verlag, 2002.

[32] P. R. Halmos. *Measure Theory*, volume 18 of *Graduate Texts in Mathematics*. Springer-Verlag, 1974.

[33] H. Hansson and B. Jonsson. A logic for reasoning about time and reliability. *Formal Aspects of Computing*, 6(5):512–535, 1994.

[34] L. A. Hemachandra. The strong exponential hierarchy collapses. *Journal of Computer and System Sciences*, 39(3):299–322, 1989.

[35] F. Horn. *Random Games*. PhD thesis, Université Paris 7, 2008.

[36] F. Horn. Explicit Muller games are PTIME. In *Proceedings of the 28th International Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2008*, volume 2 of *Leibniz International Proceedings in Informatics*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany, 2008.

[37] F. Horn. Personal Communication, 2009.

[38] P. Hunter and A. Dawar. Complexity bounds for regular games. In *Proceedings of the 30th International Symposium on Mathematical Foundations of Computer Science, MFCS 2005*, volume 3618 of *LNCS*, pages 495–506. Springer-Verlag, 2005.

[39] N. Immerman. Number of quantifiers is better than number of tape cells. *Journal of Computer and System Sciences*, 22(3):384–406, 1981.

[40] M. Jurdziński. Deciding the winner in parity games is in UP $\cap$ co-UP. *Information Processing Letters*, 68(3):119–124, 1998.

[41] M. Jurdziński, M. Paterson, and U. Zwick. A deterministic subexponential algorithm for solving parity games. *SIAM Journal on Computing*, 38(5):1519–1532, 2008.

[42] N. Klarlund. Progress measures, immediate determinacy, and a subset construction for tree automata. *Annals of Pure and Applied Logic*, 69(2–3):243–268, 1994.

[43] T. M. Liggett and S. A. Lippman. Stochastic games with perfect information and time average payoff. *SIAM Review*, 11(4):604–607, 1969.

[44] A. P. Maitra and W. D. Sudderth. Finitely additive stochastic games with Borel measurable payoffs. *International Journal of Game Theory*, 27(2):257–267, 1998.

[45] D. A. Martin. The determinacy of Blackwell games. *Journal of Symbolic Logic*, 63(4): 1565–1581, 1998.

[46] A. McIver and C. Morgan. Games, probability and the quantitative μ-calculus. In *Proceedings of the 9th International Conference on Logic for Programming, Artificial Intelligence and Reasoning, LPAR 2002*, volume 2514 of *LNCS*, pages 292–310. Springer-Verlag, 2002.

[47] R. McNaughton. Infinite games played on finite graphs. *Annals of Pure and Applies Logic*, 65(2):149–184, 1993.

[48] A. W. Mostowski. Games with forbidden positions. Technical Report 78, Instytut Matematyki, Uniwersytet Gdański, Poland, 1991.

[49] J. F. Nash, Jr. Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences of the USA*, 36:48–49, 1950.

[50] A. Neyman and S. Sorin, editors. *Stochastic Games and Applications*, volume 570 of *NATO Science Series C*. Springer-Verlag, 2003.

[51] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[52] W. Thomas. Automata on infinite objects. In *Handbook of Theoretical Computer Science*, volume B: Formal Models and Semantics, pages 133–192. Elsevier, 1990.

[53] M. Ummels. Rational behaviour and strategy construction in infinite multiplayer games. Diploma Thesis, RWTH Aachen, Germany, 2005.

[54] M. Ummels. The complexity of Nash equilibria in infinite multiplayer games. In *Proceedings of the 11th International Conference on Foundations of Software Science and Computation Structures, FOSSACS 2008*, volume 4962 of *LNCS*, pages 20–34. Springer-Verlag, 2008.

[55] M. Ummels and D. Wojtczak. Decision problems for Nash equilibria in stochastic games. In *Proceedings of the 18th Annual Conference of the European Association for Computer Science Logic, CSL '09*, volume 5771 of *LNCS*, pages 515–529. Springer-Verlag, 2009.

[56] J. Vöge and M. Jurdziński. A discrete strategy improvement algorithm for solving parity games. In *Proceedings of the 12th International Conference on Computer Aided Verification, CAV 2000*, volume 1855 of *LNCS*, pages 202–215. Springer-Verlag, 2000.

[57] K. W. Wagner. Bounded query classes. *SIAM Journal on Computing*, 19(5):833–846, 1990.

[58] W. Zielonka. Perfect-information stochastic parity games. In *Proceedings of the 7th International Conference on Foundations of Software Science and Computation Structures, FOSSACS 2004*, volume 2987 of *LNCS*, pages 499–513. Springer-Verlag, 2004.

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