On Pure Nash Equilibria in Stochastic Games

Ankush Das[1], Shankara Narayanan Krishna[1], Lakshmi Manasa[1(✉)],
Ashutosh Trivedi[1], and Dominik Wojtczak[2]

[1] Department of Computer Science and Engineering, IIT Bombay, Mumbai, India
lakshmimanasa.g@gmail.com
[2] Department of Computer Science, The University of Liverpool, Liverpool, UK

Abstract. Ummels and Wojtczak initiated the study of finding Nash equilibria in simple stochastic multi-player games satisfying specific bounds. They showed that deciding the existence of pure-strategy Nash equilibria (PURENE) where a fixed player wins almost surely is undecidable for games with 9 players. They also showed that the problem remains undecidable for the finite-strategy Nash equilibrium (FINNE) with 14 players. In this paper we improve their undecidability results by showing that PURENE and FINNE problems remain undecidable for 5 or more players.

Keywords: Stochastic games · Nash equilibrium · Pure strategy · Finite-state strategy

1 Introduction

Stochastic games are well established formalism for analyzing reactive systems under the influence of random events [1]. Such systems are often modeled as games 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, 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, often in the practical settings the system may consist of several components with independent objectives, a situation which is naturally modeled by a multi-player game.

In this paper, we study multi-player *stochastic games* [9] played on finite directed graphs whose vertices are either *stochastic* or controlled by one of the players. A play of such a game evolves by moving a token along edges of the graph in the following manner. The game begins in an initial vertex. 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 probability distribution determines the successor vertex. In the most general case, a measurable function maps plays to payoffs. In this paper we consider so-called *simple stochastic games*, where the possible payoffs of a single play are either 0 or 1 (i.e. each player either wins or loses a given play) and depend only on the *terminal vertex* of the play, i.e. a vertex which only has a self-loop edge.

However, due to the presence of stochastic vertices, a player's *expected payoff* (i.e. her probability of winning) can be an arbitrary probability.

The most common interpretation of rational behavior in multi-player games is captured by the notion of a *Nash equilibrium* [8]. In a Nash equilibrium, no player can improve her payoff by unilaterally switching to a different strategy. Chatterjee et al. in [3] gave an algorithm for computing a Nash equilibrium in a stochastic multi-player game with ω-regular winning conditions. However—as observed by Ummels and Wojtczak [11]—the algorithm proposed by Chatterjee et al. may compute an equilibrium where all players lose almost surely, even when there exist other equilibria where all players win almost surely. The equilibrium where all players win almost surely is more *optimal* than the one where all players lose almost surely.

Ummels and Wojtczak [11] successfully argue that in practice it is desirable to 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. They studied the so-called NE problem as a decision problem where, given a k-player game $\mathcal{G}$ with initial vertex v_0 and two thresholds $\bar{x}, \bar{y} \in [0, 1]^{k\,1}$, the goal is to decide whether $(\mathcal{G}, v_0)$ has a Nash equilibrium with expected payoff at least $\bar{x}$ and at most $\bar{y}$. This problem can be considered as a generalization 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 exceeds a given threshold.

There are several variants of the NE problem depending on the type of strategies permitted. On the one hand, strategies may be *randomized* (allowing randomization over actions) or *pure* (not allowing such randomization). 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 with completely different computational complexity [11].

Contributions. Ummels and Wojtczak [11] showed that deciding the existence of pure-strategy Nash equilibria (PURENE) where a fixed player wins almost surely is undecidable for games with 9 players. They also showed that the problem remains undecidable for the finite-strategy Nash equilibrium (FINNE) with 13 players. In this paper we further refine their undecidability results by showing that PURENE and FINNE problems remain undecidable for 5 or more players.

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 [4,7]. More in the spirit of our

---

[1] The ith element of vector $\bar{x}$ corresponds to the payoff of player i.

work, [6] 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 non-stochastic infinite games, a qualitative version of the NE problem was studied in [10]. 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 in two player case. In the multi-player case, [3] 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 $\bar{x}$ is NP-complete.

Ummels and Wojtczak showed in [11] that the NE problem is undecidable if we allow either arbitrary randomized strategies or arbitrary pure strategies. In fact, even the following, presumably simpler, problem was showed undecidable: Given a game $\mathcal{G}$, decide whether there exists a Nash equilibrium (in pure strategies) where player 0 wins almost surely. Moreover, the problem remains undecidable if one restricts to randomized or pure strategies with finite memory. However, it was also shown there that if one restricts to simpler types of strategies like stationary ones, NE becomes decidable [11]. In particular, for positional strategies the problem is NP-complete, and for arbitrary stationary strategies it is NP-hard but contained in Pspace. Also, the strictly qualitative fragment of NE is decidable. This fragment arises from NE by restricting the two thresholds to be the same binary payoff. Hence, they were only interested in equilibria where each player either wins or loses almost surely. Formally, the task is to decide, given a k-player game $\mathcal{G}$ with initial vertex v_0 and a binary payoff $\bar{x} \in \{0,1\}^k$, whether the game has a Nash equilibrium with expected payoff $\bar{x}$. It was shown there that for simple stochastic games, this problem is P-complete [11].

Ummels and Wojtczak studied, in [12], the computational complexity of Nash equilibria in concurrent games with limit-average objectives. They showed that the existence of a Nash equilibrium in randomized strategies is undecidable (for at least 14 players), while the existence of a Nash equilibrium in pure strategies is decidable, even if a constraint is put on the payoff of the equilibrium. Their undecidability result holds even for a restricted class of concurrent games, where nonzero rewards occur only on terminal states. Moreover, they showed that the constrained existence problem is undecidable not only for concurrent games but for turn-based games with the same restriction on rewards. They also showed undecidability of the existence of an (unconstrained) Nash equilibrium in concurrent games with terminal-reward payoffs. Finally, Bouyer et al. [2] showed undecidability of the existence of constrained Nash equilibrium in a very similar model – players do no observe the actions taken but only the state of the game – with only three players and 0/1-rewards (i.e., reachability objectives).

2 Simple Stochastic Multi-player Games

We study multi-player extension of *simple stochastic game* introduced by Condon [5] as studied by Ummels and Wojtczak [11].

Definition 1 (Simple Stochastic Multi-player Games). A simple stochastic multi-player game *(SSMG) is a tuple $(\Pi, V, (V_i)_{i \in \Pi}, \Delta, (F_i)_{i \in \Pi})$ where:*

- $\Pi = \{0, 1, \ldots, k-1\}$ *is a finite set of* players*;*
- V *is a finite set of* vertices*;*
- $V_i \subseteq V$ *is the set of vertices controlled by player i such that $V_i \cap V_j = \emptyset$ for every $i \neq j \in \Pi$;*
- $\Delta \subseteq V \times ([0,1] \cup \{\bot\}) \times V$ *is the* transition relation*, and*
- $F_i \subseteq V$ *for each $i \in \Pi$.*

We say that a vertex $v \in V$ is *controlled by player i* if $v \in V_i$. A vertex $v \in V$ is called a *stochastic vertex* if $v \notin \bigcup_{i \in \Pi} V_i$, that is, v is not contained in any of the sets V_i. We require that a transition is labeled by a probability iff it originates in a stochastic vertex: If $(v, p, w) \in \Delta$ then $p \in [0,1]$ if v is a stochastic vertex and $p = \bot$ if $v \in V_i$ for some $i \in \Pi$. Moreover, for each pair of a stochastic vertex v and an arbitrary vertex w, we require that there exists precisely one $p \in [0,1]$ such that $(v, p, w) \in \Delta$. As usual, for computational purposes we require that all these probabilities are rational.

For a given vertex $v \in V$, the set of all $w \in V$ such that there exists $p \in (0,1] \cup \{\bot\}$ with $(v, p, w) \in \Delta$ is denoted by $v\Delta$. For technical reasons, it is required that $v\Delta \neq \emptyset$ for all $v \in V$. Moreover, for each stochastic vertex v, the outgoing probabilities must sum up to 1: $\sum_{(p,w):(v,p,w) \in \Delta} p = 1$. Finally, it is required that each vertex v that lies in one of the sets F_i is a *terminal (sink) vertex*: $v\Delta = \{v\}$. So if F is the set of all terminal vertices, then $F_i \subseteq F$ for each $i \in \Pi$.

A *(mixed) strategy of player i in $\mathcal{G}$* is a mapping $\sigma : V^* V_i \to \mathcal{D}(V)$ assigning to each possible *history* $xv \in V^* V_i$ 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 $\sigma(w \mid xv)$. A *(mixed) 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 $\mathcal{G}$. Given a strategy profile $\bar{\sigma} = (\sigma_j)_{j \in \Pi}$ and a strategy τ of player i, we denote by $(\bar{\sigma}_{-i}, \tau)$ the strategy profile resulting from $\bar{\sigma}$ by replacing σ_i with τ.

A strategy σ of player i is called *pure* 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 $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ is called *pure* if each σ_i is pure. More generally, a pure strategy σ is called *finite-state* if it can be implemented by a finite automaton with output or, equivalently, if the equivalence relation $\sim \subseteq V^* \times V^*$ defined by $x \sim y$ if $\sigma(xz) = \sigma(yz)$ for all $z \in V^* V_i$ has only finitely many equivalence classes. In general, this definition is applicable to mixed strategies as well, but here, we identify finite-state strategies with pure finite-state strategies. Finally, a *finite-state strategy profile* is a profile consisting of finite-state strategies only.

It is sometimes convenient to designate an initial vertex $v_0 \in V$ of the game. We call the tuple $(\mathcal{G}, v_0)$ an *initialized SSMG*. A strategy (strategy profile) of $(\mathcal{G}, v_0)$ is just a strategy (strategy profile) of $\mathcal{G}$. In the following, we will use the abbreviation SSMG also for initialized SSMGs. It should always be clear from the context if the game is initialized or not.

When drawing an SSMG as a graph, we continue to use the conventions of [11]. 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. 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 realized by an SSMG consisting of stochastic and terminal vertices only.

Given an SSMG $(\mathcal{G}, v_0)$ and a strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$, the *conditional probability of $w \in V$ given the history $xv \in V^*V$* 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. We abuse notation and denote this probability by $\bar{\sigma}(w \mid xv)$. The probabilities $\bar{\sigma}(w \mid xv)$ induce a probability measure on the space V^ω in the following way: The probability of a basic open set $v_1 \ldots v_k \cdot V^\omega$ is 0 if $v_1 \neq v_0$ and the product of the probabilities $\bar{\sigma}(v_j \mid v_1 \ldots v_{j-1})$ for $j = 2, \ldots, k$ 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 $\Pr_{v_0}^{\bar{\sigma}}$. For a set $U \subseteq V$, let $\mathrm{Reach}(U) := V^* \cdot U \cdot V^\omega$.

Given a strategy profile $\bar{\sigma}$, a strategy τ of player i is called a *best response to $\bar{\sigma}$* if τ maximizes the expected payoff of player i, i.e. for all strategies τ' of player i we have that $\Pr_{v_0}^{(\bar{\sigma}_{-i}, \tau')}(\mathrm{Reach}(F_i)) \leq \Pr_{v_0}^{(\bar{\sigma}_{-i}, \tau)}(\mathrm{Reach}(F_i))$. A *Nash equilibrium* is a strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ such that each σ_i is a best response to $\bar{\sigma}$. Hence, in a Nash equilibrium no player can improve her payoff by (unilaterally) switching to a different strategy. In this paper we study the following decision problem.

Definition 2 (Decision Problem NE). *Given an initialized simple stochastic multi-player game $(\mathcal{G}, v_0)$ and two thresholds $\bar{x}, \bar{y} \in [0, 1]^\Pi$, decide whether there exists a Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$.*

As usual, for computational purposes we assume that the thresholds $\bar{x}$ and $\bar{y}$ are vectors of rational numbers. The threshold-free variant of the above problem which omits the thresholds just asks about a Nash equilibrium where some distinguished player, say player 0, wins almost surely.

The following is the key result of this paper.

Theorem 1. *The existence of a pure-strategy-Nash equilibrium SSMG where player 0 wins almost surely is undecidable for games with 5 or more players.*

3   Improved Undecidability Result

In this section we construct an SSMG $\mathcal{G}$ for which we show the undecidability of the existence of pure-strategy Nash equilibria of $(\mathcal{G}, v_0)$ where player 0 wins almost surely, whenever $\mathcal{G}$ has 5 or more players. We then explain how this proof can be adapted to show undecidability of

– finite-strategy Nash equilibrium where player 0 wins almost surely whenever $\mathcal{G}$ has 5 or more players.

3.1   Pure-Strategy 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 in [11]. 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 k);
– "zero(j) ? goto k: dec(j); goto l"   (if the value of counter j is zero, go to instruction k; otherwise, decrement counter j by one and go to instruction 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 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.

In order to prove Theorem 1, we show that one can compute from a two-counter machine $\mathcal{M}$ an SSMG $(\mathcal{G}, v_0)$ with five players such that the computation of $\mathcal{M}$ is infinite iff $(\mathcal{G}, v_0)$ has a pure Nash equilibrium where player 0 wins almost surely. This establishes a reduction from the non-halting problem to PURENE.

The game $\mathcal{G}$ is played by player 0 and four other players A^t and B^t, indexed by $t \in \{0, 1\}$. Let $\Gamma = \{\text{init}, \text{inc}(j), \text{dec}(j), \text{zero}(j) : j = 1, 2\}$, and let $q_1 = 2$, $q_2 = 3$ be two primes. 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 $S_{i,\gamma}^t$, $I_{i,\gamma}^t$ and $C_{j,\gamma}^t$, which are depicted in Fig. 1. In the figure, squares represent terminal vertices (the edge leading from a terminal vertex to itself being implicit), and the labeling indicates which players win at the respective vertex. Moreover, the dashed edge inside $C_{j,\gamma}^t$ is present iff $\gamma \notin \{\text{init}, \text{zero}(j)\}$. The initial vertex v_0 of $\mathcal{G}$ is the black vertex inside the gadget $S_{1,\text{init}}^0$.

For any pure strategy profile $\bar{\sigma}$ of $\mathcal{G}$ where player 0 wins almost surely, let $x_0 v_0 \prec x_1 v_1 \prec x_2 v_2 \prec \ldots$ ($x_i \in V^*, v \in V, x_0 = \epsilon$) be the (unique) sequence of all consecutive histories such that, for each $n \in \mathbb{N}$, v_n is a black vertex and $\Pr_{v_0}^{\bar{\sigma}}(x_n v_n \cdot V^\omega) > 0$. Additionally, let $\gamma_0, \gamma_1, \ldots$ be the corresponding sequence of instructions, i.e. $\gamma_n = \gamma$ for the unique instruction γ such that v_n lies in one of the gadgets $S_{i,\gamma}^t$ (where $t = n \mod 2$). For each $j \in \{1, 2\}$ and $n \in \mathbb{N}$, we define two conditional probabilities a_n and p_n as follows:

$$a_n := \Pr\nolimits_{v_0}^{\bar{\sigma}}\left(\text{Reach}(F_{A^n \mod 2}) \mid x_n v_n \cdot V^\omega\right) \text{ and}$$
$$p_n := \Pr\nolimits_{v_0}^{\bar{\sigma}}\left(\text{Reach}(F_{A^n \mod 2}) \mid x_n v_n \cdot V^\omega \setminus x_{n+2} v_{n+2} \cdot V^\omega\right).$$

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 $t = n \mod 2$). 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
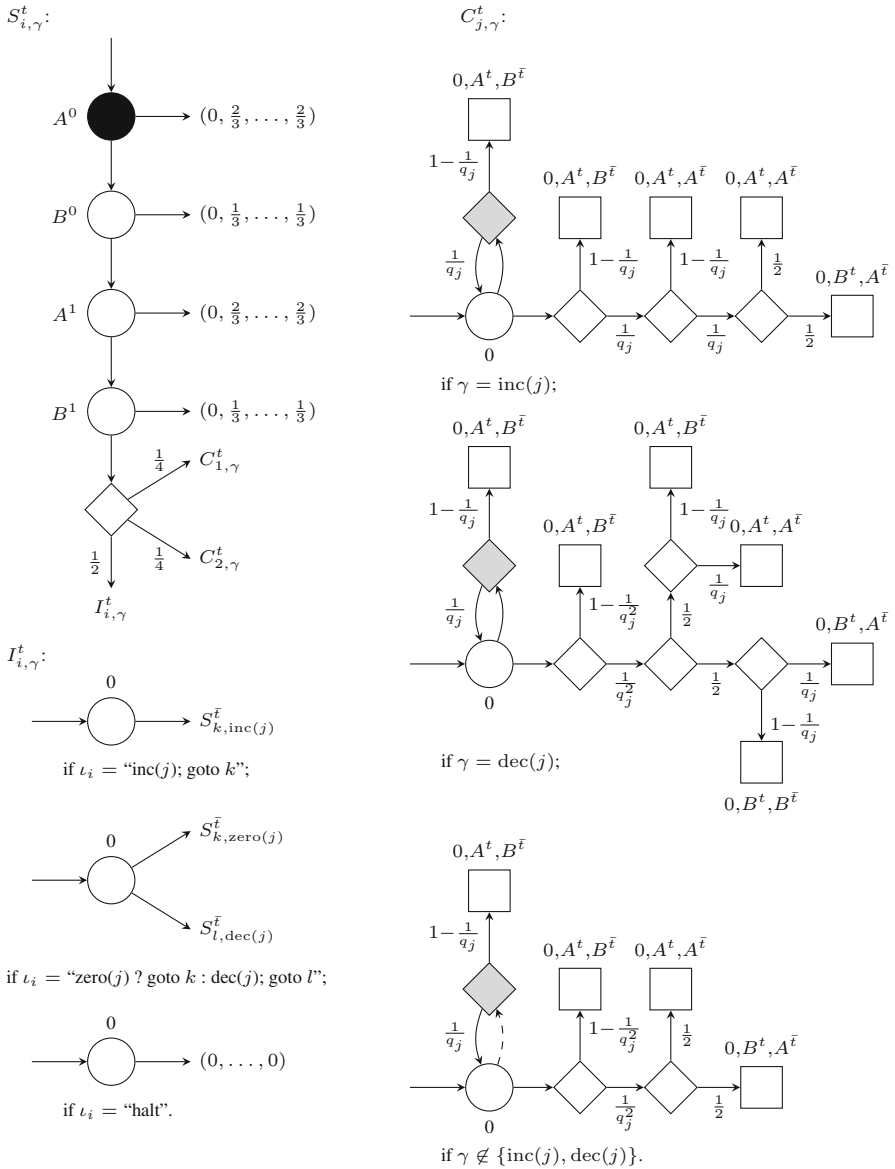

Fig. 1. Simulating a two-counter machine.

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 1. *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 the following equation holds.*

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

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

Here $+$ and $-$ denote the usual addition and subtraction of ordinal numbers respectively (satisfying $1 + \omega = \omega - 1 = \omega$). The proof of Lemma 1 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_n for $\bar{\sigma}$ to be a Nash equilibrium.

Proposition 1. *The profile $\bar{\sigma}$ is a Nash equilibrium iff $a_n = \frac{2}{3}$ for all $n \in \mathbb{N}$.*

Proof. ($\Rightarrow$) Assume that $\bar{\sigma}$ is a Nash equilibrium. Clearly, this implies that $a_n \geq \frac{2}{3}$ for all $n \in \mathbb{N}$ since otherwise some player A^t could improve her payoff by leaving one of the gadgets $S_{i,\gamma}^t$. Let $b_n := \text{Pr}_{v_0}^{\bar{\sigma}}(\text{Reach}(F_{B^n \mod 2}) \mid x_n v_n \cdot V^\omega)$. We have $b_n \geq \frac{1}{3}$ for all $n \in \mathbb{N}$ since otherwise some player B^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 $C_{j,\gamma}^{\bar{t}}$ either player A^t or player B^t wins. The conditional probability that, given the history $x_n v_n$, we reach either of those gadgets is $\sum_{k \in \mathbb{Z}} (\frac{1}{2})^k \cdot \frac{1}{2} = 1$ for all $n \in \mathbb{N}$, so we have $a_n = 1 - b_n$ for all $n \in \mathbb{N}$. Since $b_n \geq \frac{1}{3}$, we arrive at $a_n \leq 1 - \frac{1}{3} = \frac{2}{3}$, which proves the claim.

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

The second claim relates the probabilities a_n and p_n.

Proposition 2. *$a_n = \frac{2}{3}$ for all $n \in \mathbb{N}$ if and only if $p_n = \frac{1}{2}$ for all $n \in \mathbb{N}$.*

Proof. ($\Rightarrow$) Assume that $a_n = \frac{2}{3}$ for all $n \in \mathbb{N}$. We have $a_n = p_n + \frac{1}{4} \cdot a_{n+2}$ and therefore $\frac{2}{3} = p_n + \frac{1}{6}$ for all $n \in \mathbb{N}$. Hence, $p_n = \frac{1}{2}$ for all $n \in \mathbb{N}$.

($\Leftarrow$) Assume that $p_n = \frac{1}{2}$ for all $n \in \mathbb{N}$. Since $a_n = p_n + \frac{1}{4} \cdot a_{n+2}$ for all $n \in \mathbb{N}$, the numbers a_n must satisfy the following recurrence: $a_{n+2} = 4a_n - 2$. Since all the numbers a_n are probabilities, we have $0 \leq a_n \leq 1$ for all $n \in \mathbb{N}$. It is easy to see that the only values for a_0 and a_1 such that $0 \leq a_n \leq 1$ for all $n \in \mathbb{N}$ are $a_0 = a_1 = \frac{2}{3}$. But this implies that $a_n = \frac{2}{3}$ for all $n \in \mathbb{N}$. $\square$

Finally, the last claim relates the numbers p_n to Eq. (1).

Proposition 3. $p_n = \frac{1}{2}$ *for all* $n \in \mathbb{N}$ *if and only if Eq. (1) holds for all* $n \in \mathbb{N}$.

Proof. Let $n \in \mathbb{N}$, and let $t = n \bmod 2$. The probability p_n can be expressed as the sum of the probability that the play reaches a terminal vertex that is winning for player A^t inside C^t_{j,γ_n} (this probability is denoted as α^j_n) and the probability that the play reaches a terminal vertex winning for player $A^{\bar{t}}$ inside $C^{\bar{t}}_{j,\gamma_{n+1}}$ (denoted as α^j_{n+1}). For counter 1 gadgets, the probability α^1_n of A^t winning in counter gadget C^t_{1,γ_n} is

$$
\begin{aligned}
\alpha^1_n &= \Sigma_{0 \leq i \leq c^n_1 - 1} \left(1 - \frac{1}{q_1}\right) \frac{1}{q^i_1} + \frac{1}{q^{c^n_1}_1} \left\{ \left(1 - \frac{1}{q^2_1}\right) + \frac{1}{2q^2_1} \right\} \\
&= 1 - \frac{1}{q^{c^n_1}_1} + \frac{1}{q^{c^n_1}_1} \left\{ \left(1 - \frac{1}{q^2_1}\right) + \frac{1}{2q^2_1} \right\} \\
&= 1 - \frac{1}{q^{c^n_1}_1} + \frac{1}{q^{c^n_1}_1} \left\{ 1 - \frac{1}{2q^2_1} \right\} \\
&= 1 - \frac{1}{2q^{c^n_1+2}_1}
\end{aligned}
$$

Suppose $\gamma_{n+1} = \mathrm{inc}(1)$.

Then the probability α^1_{n+1} of $A^{\bar{t}}$ winning in counter gadget $C^{\bar{t}}_{1,\gamma_{n+1}}$ is $\dfrac{1}{q^{c^{n+1}_1}_1} \cdot \dfrac{1}{q_1}$

Similarly, the probabilities α^2_n and α^2_{n+1} corresponding to counter 2 gadgets are as follows:

$$
\alpha^2_n = 1 - \frac{1}{2q^{c^n_1+2}_1} \text{ and } \alpha^2_{n+1} = \frac{1}{q^{c^{n+1}_2}_2} \cdot \frac{1}{q^2_2}
$$

Given, these probabilities, p_n is as follows.

$$
\begin{aligned}
p_n &= \frac{1}{4}\left[\alpha^1_n + \frac{1}{2}\alpha^1_{n+1}\right] + \frac{1}{4}\left[\alpha^2_n + \frac{1}{2}\alpha^2_{n+1}\right] \\
&= \frac{1}{4}\left[1 - \frac{1}{2q^{c^n_1+2}_1} + \frac{1}{2q^{c^{n+1}_1+1}_1}\right] + \frac{1}{4}\left[1 - \frac{1}{2q^{c^n_2+2}_2} + \frac{1}{2q^{c^{n+1}_2+2}_2}\right] \\
&= \frac{1}{2} - \frac{1}{8}\left[\frac{1}{q^{c^n_1+2}_1} - \frac{1}{q^{c^{n+1}_1+1}_1}\right] - \frac{1}{8}\left[\frac{1}{q^{c^n_2+2}_2} - \frac{1}{q^{c^{n+1}_2+2}_2}\right]
\end{aligned}
$$

As q_1 and q_2 are primes, this sum is equal to $\frac{1}{2}$ iff $c^{n+1}_1 = 1 + c^n_1$ and $c^{n+1}_2 = c^n_2$. For γ_{n+1} being any other instruction like decrement, other instructions, the argument is similar. $\square$

Proof (Proof of Lemma 1). By Proposition 1, the profile $\bar{\sigma}$ is a Nash equilibrium iff $a_n = \frac{2}{3}$ for all $n \in \mathbb{N}$. By Proposition 2, the latter is true if $p_n = \frac{1}{2}$ for all $n \in \mathbb{N}$. Finally, by Proposition 3, this is the case iff Eq. (1) holds for all $j \in \{1, 2\}$ and $n \in \mathbb{N}$.   □

To establish the reduction, it remains to show that the computation of $\mathcal{M}$ is infinite iff the game $(\mathcal{G}, v_0)$ 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 neighboring gadget $S_{k,\gamma'}^{\bar{t}}$ 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 $S_{k,\text{zero}(j)}^{\bar{t}}$ if the value of the counter in configuration $\rho(n-1)$ is 0 and to the gadget $S_{l,\text{dec}(j)}^{\bar{t}}$ 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 iff m is strictly less than the value of the 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 the counter j in configuration $\rho(n-1)$.

Any other player's pure strategy is "moving 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 $I_{i,\gamma}^t$ where ι_i is the halt instruction is ever reached. Hence, with probability 1 a terminal vertex in one of the counter gadgets is reached. Since player 0 wins at any such vertex, we can conclude that she wins almost surely.

It remains to show that $\bar{\sigma}$ is a Nash equilibrium. By the definition of player 0's strategy σ_0, we have the following for all $n \in \mathbb{N}$: 1. c_j^n is the value of counter j in configuration $\rho(n)$; 2. c_j^{n+1} 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, Eq. (1) holds, and $\bar{\sigma}$ is a Nash equilibrium by Lemma 1.

($\Leftarrow$) Assume that $\bar{\sigma}$ 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_{i,γ_n}^t (where $t = n \bmod 2$); then $\rho(n) := (i, c_1^n, c_2^n)$.

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

1. $\rho(0) = (1, 0, 0)$;
2. $\rho(n) \vdash \rho(n+1)$ 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 from 1. and 2. Verifying the first property is easy: v_0 lies inside $S_{1,\text{init}}^0$ (and we are at instruction 1), which is linked to the counter gadgets $C_{1,\text{init}}^0$ and $C_{2,\text{init}}^0$. The edge leading to the grey vertex is missing in these gadgets. Hence, c_1^0 and c_2^0 are both equal to 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 $S_{i,\gamma}^t$ and v_{n+1} inside $S_{i',\gamma'}^{\bar{t}}$ for suitable γ, γ' and $t = n \mod 2$. We only prove the claim for the case that $\iota_i = $ "zero(2) ? goto k : dec(2); goto l"; the other cases are straightforward. Note that, by the construction of the gadget $I_{i,\gamma}^t$, it must be the case that either $i' = k$ and $\gamma' = \text{zero}(2)$, or $i' = l$ and $\gamma' = \text{dec}(2)$. By Lemma 1, if $\gamma' = \text{zero}(2)$, then $c_2' = c_2 = 0$ and $c_1' = c_1$, and if $\gamma' = \text{dec}(2)$, 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 $\gamma' \neq \text{dec}(2)$ and thus $\gamma' = \text{zero}(2)$, $i' = k$ and $c_2' = c_2 = 0$. On the other hand, if $c_2 > 0$, then $\gamma' \neq \text{zero}(2)$ and thus $\gamma' = \text{dec}(2)$, $i' = l$ and $c_2' = c_2 - 1$. □

3.2 Finite-State Equilibria

Theorem 2. *The existence of a finite-strategy-Nash equilibrium SSMG where player 0 wins almost surely is undecidable for games with 5 or more players.*

We now move on to prove Theorem 2. Before showing the undecidability of the existence of FINNE, we first note that FINNE is recursively enumerable: To decide whether an SSMG $(\mathcal{G}, v_0)$ has a finite-state Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$, one can just enumerate all possible finite-state profiles and check for each of them whether the profile is a Nash equilibrium with the desired properties by analyzing the finite Markov chain that is generated by this profile (where one identifies states that correspond to the same vertex and memory state). Hence, to show the undecidability of FINNE, we cannot reduce from the non-halting problem but from the halting problem for two-counter machines (which is recursively enumerable itself).

We now explain how to adapt the proof of Theorem 1 to show the undecidability of FINNE. 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 1 by adding another "counter" (sharing the four players from the other two gadgets, but using an additional new prime, say $q_3 = 5$ for checking whether the counter is updated correctly) that has to be incremented in each step. Moreover, additionally to the terminal vertices in the gadgets $C_{j,\gamma}^t$, we let player 0 win at the terminal vertex in each of the gadgets $I_{i,\gamma}$ where $\iota_i = $ "halt". The gadget $\gamma = \text{inc}(j)$ in Fig. 1 is a generic one and when we put $q_j = 5$, it becomes the increment gadget for this new counter. Correctly incrementing this counter comes from Proposition 3 that $p_n = \frac{1}{2}$ iff Eq. (1) is correct. With the extra counter, p_n is the sum of A^t winning in the gadgets of all the three counters. Hence, this will ensure correct updates of all counters.

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 almost surely, 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, then we can easily construct a finite-state Nash equilibrium of $(\mathcal{G}', v_0)$ where player 0 wins almost surely. Hence, $(\mathcal{G}', v_0)$ has a finite-state Nash equilibrium where player 0 wins almost surely iff the machine $\mathcal{M}$ halts.

We shall now compare the above described improved results with their counterparts in [11]. The PURENE undecidability proof in [11] reduced the non-halting problem to a game with 9 players. The game has 4 dedicated players to ensure correctness of each counter - thus using 8 additional players. While we follow their idea of reduction, with the help of primes q_1, q_2 we re-use the 4 players A^t and B^t, $t \in \{0, 1\}$ across the gadgets of both counters. Addtionally, FINNE undecidability proof is achieved by incrementing a third additional counter. While the proof for FINNE in [11] uses 4 new players for the third counter, we use another prime q_3 and re-use the 4 players (A^t and B^t, $t \in \{0, 1\}$) for the third counter.

4   Conclusion

We have showed that PURENE where player 0 wins almost surely is undecidable when the game has 5 or more players. A closely related open problem is PURENE where player 0 wins with probability $p \in [0, 1)$. The decidability of the existence of mixed-strategy NE is an interesting open problem. A further line of work is to explore concurrent moves by all the non-stochastic players, and study the decidability of the existence of various kinds of Nash equilibrium. This concurrent extension of SSMGs is inspired by [12], where the authors consider concurrent moves of all players on finite graphs, with reward vectors attached to the terminal vertices.

References

1. Baier, C., Größer, M., Leucker, M., Bollig, B., Ciesinski, F.: Probabilistic controller synthesis. In: International Conference on Theoretical Computer Science, IFIP TCS 2004, pp. 493–506. Kluwer Academic Publishers (2004)
2. Bouyer, P., Markey, N., Stan, D.: Mixed Nash equilibria in concurrent games. In: FSTTCS (to appear, 2014)
3. Chatterjee, K., Majumdar, R., Jurdziński, M.: On nash equilibria in stochastic games. In: Marcinkowski, J., Tarlecki, A. (eds.) CSL 2004. LNCS, vol. 3210, pp. 26–40. Springer, Heidelberg (2004)
4. Chen, X., Deng, X., Teng, S.-H.: Settling the complexity of computing two-player Nash equilibria. J. ACM **56**(3), 1–57 (2009)
5. Condon, A.: On algorithms for simple stochastic games. Adv. Comput. Complex. Theor. **13**, 51–73 (1993)
6. Conitzer, V., Sandholm, T.: Complexity results about Nash equilibria. In: Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI 2003, pp. 765–771. Morgan Kaufmann (2003)
7. Daskalakis, C., Goldberg, P.W., Papadimitriou, C.H.: The complexity of computing a Nash equilibrium. SIAM J. Comput. **39**(1), 195–259 (2009)

8. Nash, J.F.: Equilibrium points in *N*-person games. Proc. National Acad. Sci. USA **36**, 48–49 (1950)
9. Neyman, A., Sorin, S. (eds.): Stochastic Games and Applications. NATO Science Series C, vol. 570. Springer, Berlin (2003)
10. Ummels, M.: The complexity of Nash equilibria in infinite multiplayer games. In: Amadio, R.M. (ed.) FOSSACS 2008. LNCS, vol. 4962, pp. 20–34. Springer, Heidelberg (2008)
11. Ummels, M., Wojtczak, D.: The complexity of Nash equilibria in simple stochastic multiplayer games. In: Albers, S., Marchetti-Spaccamela, A., Matias, Y., Nikoletseas, S., Thomas, W. (eds.) ICALP 2009, Part II. LNCS, vol. 5556, pp. 297–308. Springer, Heidelberg (2009)
12. Ummels, M., Wojtczak, D.: The complexity of Nash equilibria in limit-average games. In: Katoen, J.-P., König, B. (eds.) CONCUR 2011. LNCS, vol. 6901, pp. 482–496. Springer, Heidelberg (2011)