ε-Stationary Nash Equilibria in Multi-Player Stochastic Graph Games

Ali Asadi
Institute of Science and Technology Austria, Klosterneuburg, Austria

Léonard Brice
Université Libre de Bruxelles, Belgium

Krishnendu Chatterjee
Institute of Science and Technology Austria, Klosterneuburg, Austria

K. S. Thejaswini
Institute of Science and Technology Austria, Klosterneuburg, Austria

Abstract

A strategy profile in a multi-player game is a Nash equilibrium if no player can unilaterally deviate to achieve a strictly better payoff. A profile is an ε-Nash equilibrium if no player can gain more than ε by unilaterally deviating from their strategy. In this work, we use ε-Nash equilibria to approximate the computation of Nash equilibria. Specifically, we focus on turn-based, multiplayer stochastic games played on graphs, where players are restricted to stationary strategies – strategies that use randomness but not memory.

The problem of deciding the constrained existence of stationary Nash equilibria – where each player's payoff must lie within a given interval – is known to be $\exists\mathbb{R}$-complete in such a setting (Hansen and Sølvsten, 2020). We extend this line of work to stationary ε-Nash equilibria and present an algorithm that solves the following promise problem: given a game with a Nash equilibrium satisfying the constraints, compute an ε-Nash equilibrium that ε-*satisfies* those same constraints – satisfies the constraints up to an ε additive error. Our algorithm runs in $\mathsf{FNP}^{\mathsf{NP}}$ time.

To achieve this, we first show that if a constrained Nash equilibrium exists, then one exists where the non-zero probabilities are at least an inverse of a double-exponential in the input. We further prove that such a strategy can be encoded using floating-point representations, as in the work of Frederiksen and Miltersen (2013), which finally gives us our $\mathsf{FNP}^{\mathsf{NP}}$ algorithm.

We further show that the decision version of the promise problem is NP-hard. Finally, we show a partial tightness result by proving a lower bound for such techniques: if a constrained Nash equilibrium exists, then there must be one where the probabilities in the strategies are double-exponentially small.

2012 ACM Subject Classification Theory of computation → Automata over infinite objects

Keywords and phrases Nash Equilibria, ε-Nash equilibria, Approximation, Existential Theory of Reals

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2025.9

Related Version *Full Version*: arXiv:2508.15356 [2]

Funding This work is a part of project VAMOS that has received funding from the European Research Council (ERC), grant agreement No 101020093.

1 Introduction

Modelling decentralised systems involving multiple agents requires understanding the interactions between different players, each with their own objectives. Stochastic games – graphs in which some nodes are controlled by agents and others are controlled by a random

45th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2025).
Editors: C. Aiswarya, Ruta Mehta, and Subhajit Roy; Article No. 9; pp. 9:1–9:17
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

environment – provide a modelling framework for a wide range of domains, including epidemic processes [17], formal verification [13], learning theory [1], cyber-physical systems [20], distributed and probabilistic programs [10], and probabilistic planning [22].

These stochastic multi-player turn-based games are a specific class of models that capture interactions between multiple agents on such graph arenas. These games have been extensively studied, with several results characterizing their computational complexity [8, 24] and others studying restrictions and different payoff functions that yield tractable fragments [24, 7]. While it is known that Nash equilibria (NEs) always exist in these settings, the equilibria that arise in proofs are often adversarial in nature, offering poor outcomes for all players [8, 23]. A more compelling question is whether one can decide, and compute if possible, a Nash Equilibrium that satisfies certain payoff constraints. However, this problem becomes undecidable when players are allowed unbounded memory. Indeed, it is already known that even checking whether a player can achieve a payoff within a given interval is undecidable with just 10 agents [24, Theorem 4.9]. Restricting players to deterministic (pure) strategies does not help, as the problem remains undecidable. Thus, for any hope of tractability, we must restrict each player to use bounded memory. In practice, this is also desirable: strategies requiring large memory lead to complex controllers and are harder to implement. Fortunately, finite-memory strategies can be encoded directly in the state space of the game, and in many cases, it suffices to consider stationary strategies, which require no memory at all.

In this work, we only focus on Nash equilibria where each player is restricted to stationary strategies. Indeed, henceforth, whenever we mention strategies, we only mean stationary strategies, unless specified otherwise. However, it is known that the constrained existence problem even when restricted to such stationary strategies is $\exists\mathbb{R}$-complete [16]. Moreover, for the functional version of the problem, the probabilities used to indicate the distribution of these stationary strategies can have irrational values [24, Theorem 4.6] even for games with just four players. Indeed, such strategies cannot be executed in real-world scenarios since it is hard to simulate choosing an edge with such irrational probabilities.

The standard definition of Nash equilibrium requires that no player can improve their payoff by any non-zero amount by unilaterally deviating from their strategy even if the improvement is arbitrarily small [19]. This strict notion of optimality contributes significantly to the computational complexity of deciding whether a constrained Nash equilibrium exists. From a practical perspective, however, it is reasonable to assume that players are unlikely to deviate from their current strategies for only negligible gains.

Our setting. Motivated by this discussion, we consider the relaxed variant of the equilibrium concept that is also well-studied [18, 9]. Specifically, we study the problem of computing ε-approximate Nash equilibria – in which the expected utility of each player is required to be within ε of the optimum response to the other players' strategies – while still satisfying given constraints on the players' payoffs. Formally, we focus on multi-player turn-based games with terminal rewards and for a (stationary) strategy profile that is an ε-NE that also approximately satisfies the constraints (the constraints are also satisfied up to an ε-additive factor), assuming there is an exact stationary NE.

The decision version of this problem is defined as follows: given such a game, determine whether either (i) there exists an exact NE that satisfies the constraints, or (ii) no ε-approximate NE satisfies the constraints ε-approximately. Instances that fall outside these two cases – that is, where an ε-approximate equilibrium exists but no exact one does – are considered indeterminate, and the output may be arbitrary in those cases. Note that whenever there is an exact NE, by definition, there exists an ε-NE.

Our results. We show that this approximate variant is computationally more tractable than the general problem of deciding the existence of a stationary constrained equilibrium. In particular, we present an algorithm that solves the problem within the class FNP with access to an NP oracle. We also show that the probabilities required for the exact version of the problem can be double-exponentially small, proving some evidence of hardness of the problem. Finally, we show that the decision version of the problem is NP-hard.

Technical overview

Upper bound. Recall our result that the problem lies in $\mathsf{FNP}^{\mathsf{NP}}$. Our approach relies on showing that if a stationary equilibrium exists, then there is one in which the probability values used in the strategies are not too small – specifically, they are lower-bounded by the inverse of a doubly exponential function of the input size. This structural property is established using techniques from the work of Hansen, Koucký, and Miltersen on zero-sum concurrent reachability games [15]. However, this bound alone is insufficient, because representing such tiny probabilities explicitly requires at least exponentially many bits to represent using fixed-point presentation.

To represent the probabilities used in strategies more succinctly than with fixed-point representation, we use floating-point values, a standard tool in numerical analysis. While floating-point representations cannot express the exact probabilities required in an exact Nash equilibrium, they are sufficient for representing approximate equilibria. We prove (see Lemma 9) that if an exact Nash equilibrium exists, then there is also an ε-NE that can be encoded using floating-point numbers.

The proof relies on two ideas from the work of Frederickson and Milterson [14]. First, any probability distribution can be approximated by one using floating-point values, such that the difference between them is small in a precise sense. Second, when all the probabilities in a Markov chain are approximated in this way, the expected value of the chain changes by at most a polynomial function of the approximation error. To formalise this notion of approximation, a distance metric introduced by Solan [21] is used to show how values of Markov chains behave under small perturbations.

Using these results, we design an algorithm for the problem. The algorithm first guesses a stationary approximate equilibrium that satisfies the payoff constraints up to the given error. Our earlier results ensure that such a guess can be represented using floating-point numbers with only polynomially many bits. The algorithm then verifies that this strategy profile is indeed an approximate equilibrium by checking, for each player, whether any alternative strategy would improve their payoff by more than the allowed margin of error. This check reduces to solving a Markov decision process for each player, where the strategies of the other players are fixed. We show that solving this single-player game can be done with a single call to an NP oracle for each player (see Lemma 13). Therefore, the search problem lies in the class $\mathsf{FNP}^{\mathsf{NP}}$, and the corresponding decision problem is in $\mathsf{NP}^{\mathsf{NP}}$.

Small probabilities. We first consider the limits of the above approach. Prior work by Deligkas, Fearnley, Melissourgos, and Spirakis [11] observed that the class ε-$\exists\mathbb{R}$, which captures the complexity of deciding whether approximate solutions exist to systems of real-valued constraints, is polynomial time reducible to the class $\exists\mathbb{R}$ itself. This equivalence in the complexity of $\exists\mathbb{R}$ and its approximation version hinges on the fact that solutions to such existential theory of the reals (ETR) problems can involve values that are double-exponentially large. To address this, they identified a restricted fragment of ETR for which ε-approximate solutions can be computed using a quasi-polynomial time approximation scheme (QPTAS).

In our setting, we use a key structural result (see Lemma 16) showing that stationary Nash equilibria, when they exist, can involve probabilities that are double-exponentially small, but not smaller. One might suspect that such small probabilities are merely theoretical and not actually be needed to represent probabilities occurring in such equilibria. However, in Section 4, we present a concrete example with just five players where double-exponentially small probabilities are indeed necessary for an exact equilibrium to exist. We note that previous results establishing $\exists\mathbb{R}$-completeness for related problems required seven players, underscoring the tightness of our example.

Hardness results. Finally, in Section 5, we also show that even the decision version of the promise problem is NP-hard, via a reduction from the classical 3SAT problem. While it was already known that the constrained existence of stationary Nash equilibria is NP-hard [24, Theorem 4.4] and later shown to be even $\exists\mathbb{R}$-complete, we prove that even the relaxed version of the problem – deciding whether an approximate (i.e., ε-close) stationary Nash equilibrium exists that satisfies the given constraints – is also NP-hard.

Related work. The approximation techniques we use to establish our $\mathsf{FNP}^{\mathsf{NP}}$ upper bound have appeared in various forms in the literature, particularly in the context of approximating values in concurrent games. For example, Frederiksen and Miltersen [14] introduced similar floating-point approximation methods to compute values of concurrent reachability games up to arbitrary precision. Their work placed the problem in the complexity class $\mathsf{TFNP}^{\mathsf{NP}}$. However, for this class of games, no corresponding hardness results – such as the NP-hardness we establish for our setting – were previously known.

These techniques have also been extended to other objectives in the same setting of concurrent games. In particular, Asadi, Chatterjee, Saona, and Svoboda [3] applied related methods to show $\mathsf{TFNP}^{\mathsf{NP}}$ upper bounds for games with stateful-discounted objectives and parity objectives.

More recently, Bose, Ibsen-Jensen, and Totzke [6] studied ε-approximate Nash equilibria in concurrent and partial-observation settings, also using approximation-based techniques. However, their work differs from ours in two key ways. First, they do not address the *constrained existence problem* for ε-Nash equilibria. We note that the unconstrained variant of the problem, which they consider, is significantly easier in terms of computational complexity. Second, they assume the number of players is fixed, whereas in our setting, the number of players is part of the input. To our knowledge, our work is the first to analyse constrained ε-Nash equilibria in multi-player games where the number of players is not fixed.

2 Preliminaries

We assume that the reader is familiar with the basics of probability and graph theory. However, we define some concepts for establishing notation.

Probabilities. Given a (finite or infinite) set of outcomes Ω and a probability measure $\mathbb{P}$ over Ω, let X be a random variable over Ω, i.e., a mapping $X : \Omega \to \mathbb{R}$. We write $\mathbb{E}^{\mathbb{P}}[X]$, or simply $\mathbb{E}[X]$, for the expected value of X, when defined. Given a finite set S, a *probability distribution* over S is a mapping $d : S \to [0, 1]$ that satisfies the equality $\sum_{x \in S} d(x) = 1$.

Size of numbers. Integers are always implicitly assumed to be represented with their binary encoding. For a given integer $n \in \mathbb{N}$, we therefore write $\mathsf{bit}(n) = \lceil \log_2(n+1) \rceil$ for the number of bits required to write n. Similarly, unless stated otherwise, a rational number $\frac{p}{q}$, where p and q are co-prime, is represented by the pair of the encodings of p and q, and we write $\mathsf{bit}\left(\frac{p}{q}\right) = \mathsf{bit}(p) + \mathsf{bit}(q) + 1$.

Graphs and games. A directed graph (V, E) consists of a set V of *vertices* and a set E of ordered pairs of vertices, called *edges*. For simplicity, we often write uv for an edge $(u, v) \in E$. A *path* in the directed graph (V, E) is a (finite or infinite) ordered sequence of vertices from V such that every pair of two consecutive elements is an edge. A *cycle* is a path with distinct vertices such that the last element of the sequence and the first one also forms an edge.

Throughout this paper, we use the word *game* for multiplayer simple quantitative turn-based games played on graphs.

▶ **Definition 1** (Game)**.** *A* game *is a tuple $\mathcal{G}$ that consists of:*

- *a directed graph (V, E), called the* underlying *graph of $\mathcal{G}$;*
- *a finite set Π of* players*;*
- *a partition $(V_i)_{i \in \Pi \cup \{?\}}$ of the set V, where V_i denotes the set of vertices* controlled *by player i, and the vertices in $V_?$ are called* stochastic vertices*;*
- *a* probability function $\mathsf{p} : E(V_?) \to [0, 1]$*, such that for each stochastic vertex s, the restriction of p to $E(s)$ is a probability distribution;*
- *a mapping $\mu : T \to [0, 1]^\Pi$ called* payoff function*, where T is the set of* terminal vertices*, that is, vertices of the graph (V, E) that have no outgoing edges. We also write μ_i, for each player i, for the function that maps a terminal vertex t to the i^{th} coordinate of the tuple $\mu(t)$.*
- *a vertex $v_0 \in V$, which is the* initial *vertex of the game.*

When referring to a game $\mathcal{G}$, we will often use the notations V for vertices, E for edges, Π for players, and so on without necessarily recalling them.

▶ **Definition 2** (Markov decision process, Markov chain)**.** *A* Markov decision process *is a game with one player. A* Markov chain *is a game with zero players.*

Plays. We call *play* a path in the underlying graph that is infinite, or whose last vertex is a terminal vertex. The payoff functions μ and μ_i, defined only for terminal vertices, are extended naturally to plays as follows: for a play of the form ht, with $t \in T$, we define $\mu(ht) = \mu(t)$, and for an infinite play π, we define $\mu(\pi) = (0)_{i \in \Pi}$ (all players receive the payoff 0 in an infinite path where no terminal vertex is reached).

Strategies, and strategy profiles. In this paper, by *strategy*, we mean a stationary strategy, that is, a strategy that only depends on the current vertex.

Thus, in a game $\mathcal{G}$, a *strategy* for player i is a mapping σ_i that maps each vertex $v \in V_i$ to a probability distribution over the neighbours of v. A *strategy profile* for a subset $P \subseteq \Pi$ is a tuple $(\sigma_i)_{i \in P}$, that we usually write $\bar{\sigma}$ if $P = \Pi$, and $\bar{\sigma}_{-i}$ if $P = \Pi \backslash \{i\}$ for some player i. We then write (σ_{-i}, σ'_i) to denote the strategy profile $\bar{\tau}$ defined by $\tau_i = \sigma'_i$ and $\tau_j = \sigma_j$ for $j \neq i$.

A strategy profile $\bar{\sigma}$ for all players in the game $\mathcal{G}$ defines a probability measure $\mathbb{P}_{\bar{\sigma}}$ over plays – which turns the payoff functions μ_i into random variables. We then write $\mathbb{E}(\bar{\sigma})$ for the expectation operator $\mathbb{E}^{\mathbb{P}_{\bar{\sigma}}}$. If the game is a Markov chain, then we just write $\mathbb{E}$ to represent this value. We say the maximum payoff of an MDP is the maximum of all expectations over all stationary strategies and write $\mathsf{val} = \max_\sigma \mathbb{E}(\bar{\sigma})$ to denote it.

Equilibria. In such games, we study *equilibria*, that is, strategy profiles that offer some stability guarantees. The most famous notion of equilibrium is the *Nash equilibrium* [19].

▶ **Definition 3** (Nash equilibrium). *A* Nash equilibrium*, or* NE *for short, in the game $\mathcal{G}$ is a strategy profile $\bar{\sigma}$ such that for each player i and every strategy σ'_i, we have*

$$\mathbb{E}(\bar{\sigma}_{-i}, \sigma'_i)[\mu_i] \leqslant \mathbb{E}(\bar{\sigma})[\mu_i].$$

A classical quantitative relaxation of Nash equilibria is that of ε-approximate Nash equilibria, or just ε-Nash equilibria.

▶ **Definition 4** (ε-Nash equilibrium). *Let $\varepsilon > 0$. An* ε-Nash equilibrium*, or* ε-NE *for short, in the game $\mathcal{G}$ is a strategy profile $\bar{\sigma}$ such that for each player i and every strategy σ'_i, we have $\mathbb{E}(\bar{\sigma}_{-i}, \sigma'_i)[\mu_i] - \varepsilon \leqslant \mathbb{E}(\bar{\sigma})[\mu_i]$.*

Problem. A classical decision problem about Nash equilibria is the following one.

▶ **Problem 1** (Constrained existence problem of Nash equilibria). *Given a game $\mathcal{G}$ and two vectors $\bar{x}, \bar{y} \in \mathbb{Q}^\Pi$, called* threshold vectors*, does there exist an NE $\bar{\sigma}$ in $\mathcal{G}$ that satisfies the inequality $\bar{x} \leqslant \mathbb{E}(\bar{\sigma})[\mu] \leqslant \bar{y}$?*

Problem 1 is known to be undecidable if the players are allowed to use memory [24], which is excluded by our formalism here. Where players are restricted to *stationary* strategy – as in our definition – this problem is known to be $\exists\mathbb{R}$-complete [16], where $\exists\mathbb{R}$ is the complexity class of problems that can be reduced to the satisfiability of a formula in the existential theory of the reals. In order to obtain a more tractable problem, we consider here an approximated version defined using the notion of ε-NEs: we wish to obtain, for sure, the answer *yes* when an NE satisfying the constraint exists, and to obtain, for sure, the answer *no* when we are far from having such an NE because there is not even an *ε-NE* that satisfies the constraint *up to ε* – but on the limit case, we can accept any answer. Formally, we want an algorithm that solves the following problem.

▶ **Problem 2** (Approximate constrained existence problem of Nash equilibria). *Given a game $\mathcal{G}$, two vectors $\bar{x}, \bar{y} \in \mathbb{Q}^\Pi$, and a rational number $\varepsilon > 0$ input using fixed point representation, such that either:*

1. *there is an NE $\bar{\sigma}$ in $\mathcal{G}$ with $x_i \leqslant \mathbb{E}(\bar{\sigma})[\mu_i] \leqslant y_i$ for each player i; or*
2. *there is no ε-NE in $\mathcal{G}$ satisfying $x_i - \varepsilon \leqslant \mathbb{E}(\bar{\sigma})[\mu_i] \leqslant y_i + \varepsilon$ for each player i;*

holds, then are we in case 1?

We call *functional* approximate constrained NE problem the functional version of this problem, in which the answer *yes* is replaced by a succinct representation of an ε-NE that satisfies the constraint up to ε. In the sequel, we show a $\mathsf{FNP}^{\mathsf{NP}}$ upper bound on this functional problem, which implies an $\mathsf{NP}^{\mathsf{NP}}$ upper bound on the non-functional one; and an NP lower bound on the non-functional approximated problem, which implies an FNP lower bound on the functional one.

3 Finding ε-Nash equilibria

In this section, we present the $\mathsf{FNP}^{\mathsf{NP}}$ upper bound for the functional approximate constrained NE problem. We always use n to represent the number of vertices in the game and m to represent the number of edges in the game, and τ which corresponds to the bit size used to represent the probabilities.

Throughout this section, we assume that we are given an instance $(\mathcal{G}, \bar{x}, \bar{y}, \varepsilon)$ of the functional approximate constrained problem. We call *(ε-)constrained NE* an (ε-)NE in $\mathcal{G}$ such that each player i's payoff lies in the interval $[x_i, y_i]$ (resp. $[x_i - \varepsilon, y_i + \varepsilon]$). The rest of this section is dedicated to the proof of the following theorem.

▶ **Theorem 5.** *There exists an $\mathsf{FNP}^{\mathsf{NP}}$ procedure to solve the functional approximated constrained problem of NEs.*

We begin by defining the existential theory of the reals and showing how the problem of computing stationary Nash equilibria can be encoded using sentences in it.

By applying a result of Basu, Pollack, and Roy [5] (restated in Lemma 8), it can be shown that any solution to the resulting system of polynomial inequalities lies within a ball of at most double-exponential radius. This implies that, in the worst case, the smallest non-zero probabilities in a stationary equilibrium may be as small as inverse double-exponential in the size of the input.

To represent such small probabilities succinctly, we introduce a floating-point representation that allows encoding of some double-exponentially small values using only polynomially many bits. Next, using a result of Solan [21] (see Lemma 11), we observe that small perturbations in transition probabilities lead to only small changes in the value of the associated Markov chain. This continuity result ensures that approximate equilibria obtained where probabilities are slightly perturbed still yield approximately correct values. This shows that it is enough to represent the probabilities using floating-point numbers with polynomial precision.

Finally, we provide a coNP procedure (see Lemma 13) to decide whether the maximum payoff in a Markov Decision Process (MDP), where probabilities are given in floating-point representation, lies above or below a given threshold up to an additive ε.

Together, these components give us our desired $\mathsf{FNP}^{\mathsf{NP}}$ procedure: we guess a stationary strategy profile that approximately satisfies the constraints and verify it using the above results, and use the fact that such strategies admit succinct polynomial-size representations.

Existential theory of the reals. A sentence in the *existential theory of the reals* is of the form

$$\varphi \;=\; \exists x_1\, \exists x_2\; \cdots\; \exists x_k\; F(x_1, \ldots, x_k)\,,$$

where $F(x_1, \ldots, x_k)$ is a *quantifier-free* formula in the language of ordered fields over the reals $\mathbb{R}$. More concretely, atomic sub-formulas are

$$p(x_1, \ldots, x_k) = 0 \quad \text{or} \quad p(x_1, \ldots, x_k) < 0,$$

where $p \in \mathbb{R}[x_1, \ldots, x_k]$ is a polynomial over $\mathbb{R}$, and arbitrary formulas are built by Boolean connectives.

▶ **Lemma 6.** *If there exists a constrained Nash equilibrium $\bar{\sigma}$, then there exists a constrained Nash equilibrium such that the non-zero probabilities are at least $2^{-\tau 2^{O(m|\Pi|)}}$.*

Proof. This statement is similar to a result in the work of Hansen, Koucký, and Miltersen [15, Thm. 4]. We write the ETR sentence for solving constrained stationary Nash equilibrium in the following proposition.

▶ **Proposition 7.** *There exists an ETR sentence for computing constrained stationary Nash equilibrium with $O(m|\Pi|)$ variables, where all polynomials are of degree at most 2.*

Proof. The following part is similar to the ETR sentence for constrained stationary Nash equilibria found in the work of Ummels and Wojtczak [24, Theorem 4.5]. However, we require the explicit bounds on the number of variables, the degree polynomials, and the length of the sentence. Thus, we rewrite the explicit ETR sentence. First, by the assumption, there exists a strategy profile that is a stationary Nash equilibrium $\bar{\sigma}$. Let S be the set of all non-zero edges in the strategy profile $\bar{\sigma}$. Furthermore, we define the set of vertices V_S from which terminals can be reached with non-zero probability. Second, we have a formula that states that the variables p_{vw} indeed describe a strategy. We enforce that the non-zero variables belong to the set S. We further ensure that for stochastic vertices, the variable p_{vw} encodes exactly the value dictated by the probability function p by the stochastic vertex. This involves satisfying the following constraints:

1. $p_{vw} > 0$; for each $vw \in S$
2. $p_{vw} \leqslant 1$; for each $vw \in S$
3. $p_{vw} = 0$; for each $vw \notin S$
4. $\sum_{w \in E(v)} p_{vw} = 1$; for each non-stochastic vertex v
5. $p_{vw} = \mathsf{p}(vw)$; for all stochastic vertices v.

The variables r_v^i correspond to the payoff player i receives by starting from the vertex v. The payoffs r_v^i are equal to 0 for all vertices in V_S because no terminal is reachable from such vertices. The payoffs for other vertices satisfy the following constraints.

6. $r_t^i = \mu_i(t)$, for $t \in T$;
7. $r_v^i = 0$, for $v \notin V_S$;
8. $r_v^i = \sum_{w \in E(v)} p_{vw}\, r_v^i$, for $v \in V_S \backslash T$;

We now check that there is no unilateral deviation for the players by adding the following constraints. This following check is enough because, when fixing the strategies of all players except player i, the resulting structure is an MDP. For MDPs, Bellman equations characterise the value of each state and the equation below enforces that the value at a state v is at least as large as the maximum value of any successor, thereby ensuring consistency with the Bellman equations of the induced MDP [4, Theorem 10.109].

9. $r_v^i \geqslant r_w^i$, for $v \in V_i, w \in E(v)$;

Finally, we add the auxiliary variables g_{vw} to keep track of the size of p_{vw}, which satisfy the following constraint:

10. $g_{vw}\, p_{vw} = 1$, for each $vw \in S$;

Thus, our ETR statement is: $\exists p\, \exists r\, \exists g$ that satisfies constraints 1-10, completing the proof. ◁

The following result from the work of Basu, Pollack, and Roy [5] guarantees that if a system of polynomial inequalities is satisfiable, then there is a valuation satisfiable within a ball with radius double-exponentially large with respect to the input size. Before we state the result, we recall that a partition of semi-algebraic subsets of $\mathbb{R}^k$ generated by a set of polynomials $\mathcal{P}$ is the collection of all nonempty sets of the form $\{x \in \mathbb{R}^k \mid \text{sign}(p(x)) = \sigma(p) \text{ for all } p \in \mathcal{P}\}$, where $\sigma : \mathcal{P} \to \{-1, 0, 1\}$ ranges over all possible sign assignments. Intuitively, the partition is obtained by slicing $\mathbb{R}^k$ according to which side of each polynomial in $\mathcal{P}$ a point lies on (positive, zero, or negative).

▶ **Lemma 8** ([5, Theorem 1.3.5]). *Given a set $\mathcal{P}$ of polynomials of degree d in k variables with coefficients in $\mathbb{Z}$ of bit-sizes τ, then there exists a ball of radius $2^{\tau d^{O(k)}}$ that intersects every part of the partition of semi-algebraic subsets of $\mathbb{R}^k$ generated by $\mathcal{P}$ (cells of $\mathcal{P}$).*

Note that each of the constraints in the ETR sentence is a polynomial of degree at most 2. The total number of variables is $O(m|\Pi|)$. Note that we had introduced auxiliary variables g_{vw} and required that $g_{vw}\, p_{vw} = 1$, for each $vw \in S$. Since this is also a polynomial in $\mathcal{P}$, using Lemma 8, for all $vw \in S$, we have $g_{vw} \leqslant 2^{\tau 2^{O(m|\Pi|)}}$. Consequently, we have $p_{vw} \geqslant 2^{-\tau 2^{O(m|\Pi|)}}$, which completes the proof. ◀

We recall some results used in the work of Frederiksen and Milterson [14] where they adapted arguments that are standard in the context of numerical analysis to provide better algorithms for approximating the values of concurrent reachability games.

Floating-point number representation. We define the set of floating-point numbers with precision ℓ as

$$\mathcal{F}(\ell) := \left\{ m \cdot 2^e \quad \middle| \quad m \in \{0, \cdots, 2^\ell - 1\}, \quad e \in \mathbb{Z} \right\} .$$

The floating-point representation of an element $x = m \cdot 2^e \in \mathcal{F}(\ell)$ uses $\mathsf{bit}(m) + \mathsf{bit}(|e|)$ bits. We define the relative distance of two positive real numbers $x, \widetilde{x}$ as

$$\mathsf{rel}(x, \widetilde{x}) := \max \left\{ \frac{x}{\widetilde{x}}, \frac{\widetilde{x}}{x} \right\} - 1 .$$

Intuitively, $\mathsf{rel}(x, \widetilde{x})$ describes the multiplicative distance. Observe that if the values are closer to 0, but ε apart for some small ε, their relative distance is larger than if these values are much larger and only ε apart. We say x is (ℓ, i)-close to $\widetilde{x}$ if $\mathsf{rel}(x, \widetilde{x}) \leqslant (1 - 2^{1-\ell})^{-i} - 1$, where ℓ is a positive integer and i is a non-negative integer. (ℓ, i)-closeness measures how far two numbers can differ after i steps of rounding in an ℓ-bit floating-point system. Here, ℓ controls precision (smaller gaps between representable numbers), and i allows for cumulative rounding error. So x is (ℓ, i)-close to $\widetilde{x}$ if their relative difference is within about $i \cdot 2^{1-\ell}$, that is, within i units of machine precision.

Arithmetic operations. We define $\oplus^\ell, \ominus^\ell, \otimes^\ell, \oslash^\ell$ as finite precision arithmetic operations $+, -, *, /$ respectively by truncating the result of the exact arithmetic operation to ℓ bits. We drop the superscript ℓ if context is clear.

Floating-point probability distribution representation. We denote by $\mathcal{D}(\ell)$ the set of all floating-point probability distributions with precision ℓ. A probability distribution $\mu \in \Delta([t])$ belongs to $\mathcal{D}(\ell)$ if there exists $w_1, w_2, \cdots, w_t \in \mathcal{F}(\ell)$ such that

- For all $i \in [t]$, we have $\mu(i) = \frac{w_i}{\sum_{j \in [t]} w_j}$; and
- $\sum_{j \in [t]} w_j$ and 1 are (ℓ, t)-close.

We define the relative distance rel for probability distributions as $\mathsf{rel}(\mu, \widetilde{\mu}) := \max\{\mathsf{rel}(\mu(i), \widetilde{\mu}(i)) : i \in [t]\}$. We say μ is (ℓ, i)-close to $\widetilde{\mu}$ if $\mathsf{rel}(\mu, \widetilde{\mu}) \leqslant (1 - 2^{1-\ell})^{-i} - 1$, where ℓ is a positive integer and i is a non-negative integer.

▶ **Lemma 9.** *If there exists a constrained stationary Nash equilibrium $\bar{\sigma}$, then there exists an $(32n^2 2^{-\ell})$-constrained stationary Nash equilibrium $\bar{\sigma}'$ such that for all players i and vertices $v \in V_i$, we have $\bar{\sigma}'(v) \in \mathcal{D}(\ell)$, where $\ell \geqslant 1000n^2$.*

Proof. We first recall the result in the work of Frederiksen and Milterson [14] related to approximation of probability distributions by floating-point distributions.

▶ **Lemma 10** ([14, Lemma 5]). *Consider $x_1, \cdots, x_t \in \mathcal{F}(\ell)$. Let $\mu(i) := x_i \oslash \left(\bigoplus_{j=1}^{t} x_j\right)$. Then, there exists $\widetilde{\mu} \in \mathcal{D}(\ell)$ such that for all i, we have $\widetilde{\mu}(i) = \mu(i)/\left(\sum_{j=1}^{t} \mu(j)\right)$, and μ and $\widetilde{\mu}$ are $(\ell, 2t)$-close.*

A stationary strategy comprises of probability distributions over actions for each state. We can truncate a strategy by truncating each of these probability distributions. Let the strategy profile $\bar{\sigma}'$ be the truncation of the strategy profile $\bar{\sigma}$ defined in the above result. Therefore, for all players i and vertices $v \in V_i$, we have that $\bar{\sigma}(v)$ and $\bar{\sigma}'(v)$ are $(\ell, 2n)$-close. Consequently, for all vertices $v, w \in V$,

$$\begin{aligned}
\mathsf{rel}(\bar{\sigma}(v)(w), \bar{\sigma}'(v)(w)) &\leqslant \frac{1}{(1-2^{1-\ell})^{2n}} - 1 \\
&\leqslant \frac{1}{1-(2n)2^{1-\ell}} - 1 && \text{(Bernoulli inequality)} \\
&\leqslant \frac{(2n)2^{1-\ell}}{1-(2n)2^{1-\ell}} && \text{(rearrange)} \\
&\leqslant 4n2^{-\ell} && \left(\ell \geqslant 1000n^2\right) \qquad (1)
\end{aligned}$$

We then recall the result in the work of Solan [21]. This result provides an upper bound on the difference between the reachability values of two Markov chains based on the relative distance of their transition functions.

▶ **Lemma 11** ([21, Thm. 6]). *Consider two MCs $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ with identical vertex sets and a target set T. Let μ be a reward function, where $\mu(v) = 1$ for all vertices $v \in T$ and $\mu(v) = 0$ for all vertices $v \in V \backslash T$. We denote by val and $\widetilde{\mathsf{val}}$ the expected payoff of $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ respectively. Fix $\varepsilon := \max_{v,w \in V} \mathsf{rel}(\mathsf{p}(vw), \widetilde{\mathsf{p}}(vw))$. Then, we have*

$$|\mathsf{val} - \widetilde{\mathsf{val}}| \leqslant 4n\varepsilon\,.$$

By Equation (1) and Lemma 11, we bound the the difference of payoffs for $\bar{\sigma}$ and $\bar{\sigma}'$. For all players i, we have $|\mathbb{E}(\bar{\sigma})[\mu_i] - \mathbb{E}(\bar{\sigma}')[\mu_i]| \leqslant 16n^2 2^{-\ell}$. Since the strategy profile $\bar{\sigma}$ is a constrained Nash equilibrium, $\bar{\sigma}'$ is a $(32n^2 2^{-\ell})$-constrained Nash equilibrium, which completes the proof. ◀

In the above result, we show that it is sufficient to consider floating-point representable strategies to solve the problem. We now present the $\mathsf{NP}^{\mathsf{NP}}$ procedure. We first recall the result in the work of Frederiksen and Milterson [14] related to computing the approximate value of MCs with floating-point distributions.

▶ **Lemma 12** ([14, Thm. 4]). *Consider an MC $\mathcal{M}$ and a target set T. Let μ be a reward function, where $\mu(v) = 1$ for all vertices $v \in T$ and $\mu(v) = 0$ for all vertices $v \in V \backslash T$. For all vertices $v \in V$, we have $\mathsf{p}(v, .) \in \mathcal{D}(\ell)$ where $\ell \geqslant 1000n^2$. Then, there exists a polynomial-time algorithm that for all vertices $v \in V$, computes an approximation $r \in \mathcal{F}(\ell)$ for $\mathbb{E}[\mu]$ such that*

$$|r - \mathbb{E}[\mu]| \leqslant 80n^4 2^{-\ell}\,,$$

where $\mathsf{val}_T(v)$ is the reachability value for the vertex v.

▶ **Lemma 13.** *The problem of deciding if the payoff for MDPs is below a threshold up to an additive error is in* coNP *where the input is a MDP $\mathcal{P}$, a reward function μ, a vertex v, a threshold $0 \leqslant \alpha \leqslant 1$, an additive error $\varepsilon = 2^{-\kappa}$ and a positive integer ℓ such that, for all stochastic vertices $w \in V$, we have*

$$\mathsf{p}(w, .) \in D(\ell), \quad \ell \geqslant 1000n^2 + \kappa .$$

Define $\mathsf{val} := \sup_\sigma \mathbb{E}(\sigma)[\mu]$. *Note that the numbers α and ε are represented in fixed-point binary and the NP procedure is such that*

- *If $\alpha \leqslant \mathsf{val} - \varepsilon$, then it outputs* no*; and*
- *If $\alpha \geqslant \mathsf{val} + \varepsilon$, then it outputs* yes*.*

Proof. We first present the coNP procedure and then prove its soundness and completeness.

Procedure. The procedure guesses a pure stationary strategy σ for the player. Note that the size of the representation of a pure stationary strategy is polynomial with respect to the size of the representation of $\mathcal{P}$. By fixing σ, we obtain an MC $\mathcal{M}$. We denote by val_σ the value $\mathbb{E}(\sigma)[\mu]$. By Lemma 12, there exists a polynomial time algorithm that computes an ε-approximation $\widehat{\mathsf{val}}_\sigma$ of val_σ. Our procedure outputs *no* if $\alpha \leqslant \widehat{\mathsf{val}}_\sigma$. If there exists no such pure stationary strategy, the procedure outputs *yes*.

Completeness. If $\alpha \leqslant \mathsf{val} - \varepsilon$, then, by [12], there exists a pure stationary strategy σ such that $\mathsf{val} = \mathsf{val}_\sigma$. The procedure non-deterministically guesses σ. By Lemma 12, we have $\widehat{\mathsf{val}}_\sigma + \varepsilon \geqslant \mathsf{val}_\sigma$. Therefore, we have $\alpha \leqslant \widehat{\mathsf{val}}_\sigma$, and the procedure outputs *no*.

Soundness. If $\alpha \geqslant \mathsf{val} + \varepsilon$, then for all pure stationary strategies σ, we have $\alpha \geqslant \mathsf{val}_\sigma + \varepsilon$. By Lemma 12, we have $\widehat{\mathsf{val}}_\sigma - \varepsilon \leqslant \mathsf{val}_\sigma$. Therefore, $\alpha \geqslant \widehat{\mathsf{val}}_\sigma$ which implies that the procedure outputs *yes* and yields the result. ◀

Proof of Theorem 5. We first present the procedure and then prove its soundness and completeness.

Procedure. Let ℓ be sufficiently large. The procedure guesses a strategy profile and verifies that it is an $\varepsilon/8$-constrained stationary Nash equilibrium $\bar{\sigma}$. By Lemma 6 and Lemma 9, the size of the representation of $\bar{\sigma}$ is polynomial with respect to the size of game and $\mathsf{bit}(\varepsilon)$. By fixing the strategy profile $\bar{\sigma}$, we obtain an MC $\mathcal{M}$. We denote by val_i the payoff of $\mathcal{M}$ where the reward function is μ_i. By Lemma 11, the procedure computes the $\varepsilon/8$-approximate payoff $\widehat{\mathsf{val}}_i$ in polynomial time. For each player i, by fixing the strategy profile $\bar{\sigma}_{-i}$, we obtain an MDP $\mathcal{P}_i$. We denote by $\widetilde{val}_i$ the value of $\mathcal{P}_i$ where the reward function is μ_i. Then, the procedure checks if there exists an ε-unilateral deviation for player i by deciding if $\widetilde{\mathsf{val}}_i$ for is at most $\widehat{\mathsf{val}}_i(v) + 3/4\varepsilon$ up to additive error $\varepsilon/8$. The procedure finally checks if the payoff constraints are satisfied.

Soundness. We assume that $\bar{\sigma}$ is not an ε-constrained Nash equilibrium, i.e., there exists a player i such that ε-unilateral deviation is possible. Equivalently, we have $\mathsf{val}_i \leqslant \widetilde{\mathsf{val}}_i - \varepsilon$. Therefore, we have

$$\begin{aligned}\widehat{\mathsf{val}}_i + 3/4\varepsilon &\leqslant \mathsf{val}_i + 7/8\varepsilon && \left(\widehat{\mathsf{val}}_i \text{ is an } \varepsilon/8\text{-approximation of } \mathsf{val}_i\right)\\ &\leqslant \widetilde{\mathsf{val}}_i - \varepsilon/8 . && (\varepsilon\text{-unilateral deviation})\end{aligned}$$

However, for this case, the coNP procedure defined in Lemma 13, successfully outputs *no*. Therefore, our procedure does not output $\bar{\sigma}$ as ε-constrained stationary Nash equilibrium, which yields the soundness of our procedure.

Completeness. By Lemma 6 and Lemma 9, there exists an $\varepsilon/8$-constrained stationary Nash equilibrium $\bar{\sigma}$ which is polynomial-size representable. The procedure non-deterministically guesses $\bar{\sigma}$. Since the strategy profile is $\varepsilon/8$-Nash equilibrium, for all players i, we have $\widetilde{\mathsf{val}}_i \leqslant \mathsf{val}_i + \varepsilon/8$. Therefore, we get

$$\begin{aligned}
\widehat{\mathsf{val}}_i + 3/4\varepsilon &\geqslant \mathsf{val}_i + 5/8\varepsilon && \left(\widehat{\mathsf{val}}_i \text{ is an } \varepsilon/8\text{-approximation of } \mathsf{val}_i\right)\\
&\geqslant \widetilde{\mathsf{val}}_i + 4/8\varepsilon && (\widetilde{\mathsf{val}}_i \leqslant \mathsf{val}_i + \varepsilon/8)\\
&\geqslant \widetilde{\mathsf{val}}_i + \varepsilon/8\,.
\end{aligned}$$

Thus, the coNP procedure outputs *yes*. Therefore, our procedure successfully decides that $\bar{\sigma}$ is a ε-constrained stationary Nash equilibrium. This yields the completeness of the procedure and completes the proof. ◀

4 About (very) small probabilities

We now provide a tight lower bound for Lemma 6, by showing that in a constrained NE, players might need to use double-exponentially small probabilities. This result itself is not surprising due to the $\exists\mathbb{R}$-completeness of the problem [16] even in games with 7 players or more. But we tighten this result by producing an explicit game with only 5 players where such double-exponentially small probabilities are required.

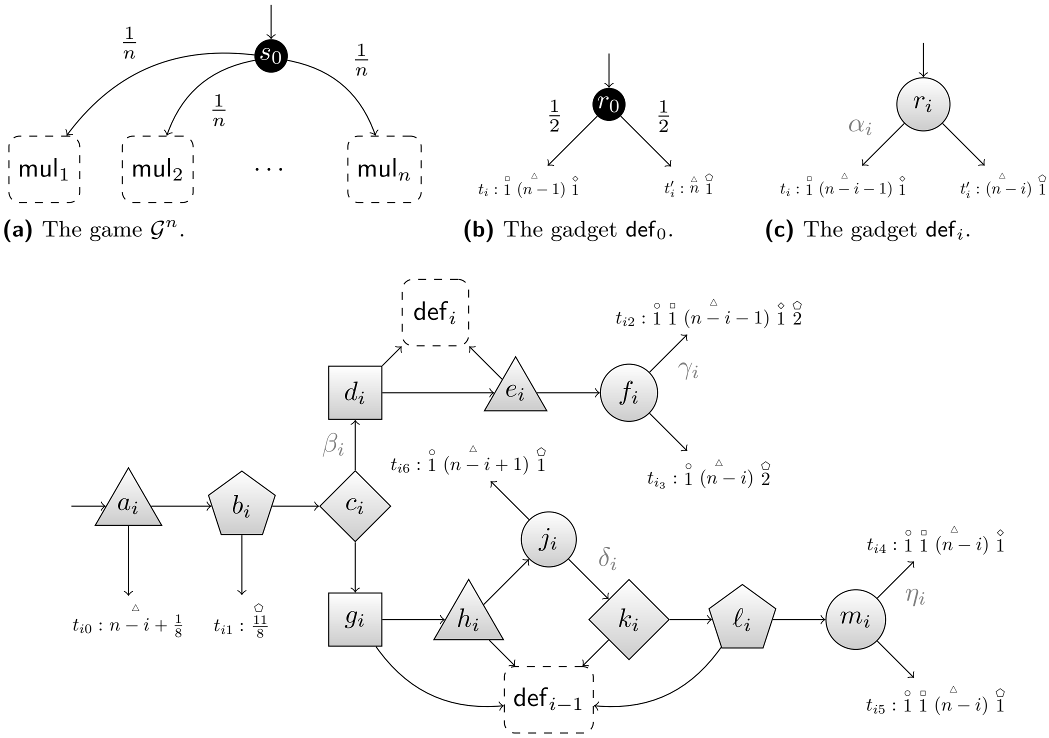
▶ **Theorem 14.** *Let $n \in \mathbb{N}$. There exist two threshold vectors $\bar{x}$ and $\bar{y}$ and a game $\mathcal{G}$ where there is a stationary Nash equilibrium $\bar{\sigma}$ satisfying $\bar{x} \leqslant \mathbb{E}(\bar{\sigma}) \leqslant \bar{y}$, but where all such stationary Nash equilibria contain probabilities that are double-exponentially small in n.*

Let $n \in \mathbb{N}$. The game $\mathcal{G}^n$ is depicted by Figure 1. It contains five players, named ○, □, △, ◇, and ⬠: the shape of each vertex indicates which player controls it, and black vertices are stochastic ones – the probability distribution is then indicated on the outgoing edges. All omitted rewards are 0. For convenience, we allow rewards greater than 1 here, accounting for the fact that all rewards can easily be normalised by dividing them by the greatest of them (namely n). For now, the reader should ignore the notations β_i, γ_i, δ_i, η_i.

Intuitively, the construction is the following: in each gadget def_i, from the vertex r_i, player ○ has to randomise between the terminal vertices t_i and t'_i (in the gadget def_0 the randomisation is imposed, because player ○ does not control the vertex r_0). The probability of going to the vertex t_i will be written α_i. We will be interested in NEs such that player ○ gets expected payoff 1: that means that the players should never go to those gadgets, nor to the terminal vertices t_{i0} or t_{i1}. But then, deviating and going to those gadgets or vertices must never be a profitable deviation. This imposes strict restrictions on which expected payoffs the players should get in each gadget mul_i, and therefore on how they can randomise their own strategies. In particular, from the vertex c_i, we will see that player ◇ must proceed to a strict randomisation, which is possible in a Nash equilibrium only if both sides are equivalent from her perspective: with this idea, each gadget mul_i binds the values of α_{i-1} and α_i, imposing $\alpha_i = \alpha_{i-1}^2$. Then, by induction, we find $\alpha_n = \frac{1}{2^{2^n}}$.

▶ **Lemma 15.** *In the game $\mathcal{G}^n$, if there exists a stationary Nash equilibrium $\bar{\sigma}$ in which player ○ gets expected payoff 1, then we have $\sigma_\circ(r_n) = \frac{1}{2^{2^n}}$.*

We now know that if a stationary NE satisfying this constraint exists, it necessarily includes a double-exponentially small probability. We now prove that such an NE exists.



(a) The game $\mathcal{G}^n$. **(b)** The gadget def$_0$. **(c)** The gadget def$_i$.

(d) The gadget mul$_i$.

Figure 1 A game where very small probabilities are necessary. The rewards corresponding to all the players not explicitly mentioned are zero.

▶ **Lemma 16.** *For every $n \in \mathbb{N}$, the game $\mathcal{G}^n$ has a stationary Nash equilibrium where player ○'s expected payoff is* 1*.*

Together, Lemmas 15 and 16 prove Theorem 14. This theorem should be understood as a lower bound on Lemma 6, that shows that one cannot expect a better complexity than $\mathsf{NP}^{\mathsf{NP}}$ using the techniques we use here.

However, let us now note that the above result is true only for exact Nash equilibria. For ε-NEs, our construction does not lead to the same result, as shown by the following lemma.

▶ **Lemma 17.** *Let $n \in \mathbb{N}$ and $\varepsilon > 0$. In the game $\mathcal{G}^n$, there exists a stationary ε-NE $\bar{\tau}$ where player ○'s expected payoff is* 1*, and where for each player i and every edge uv with $u \in V_i$, the size $\mathsf{bit}\left(\tau_i(u)(v)\right)$ is bounded by a polynomial of $\mathsf{bit}(n)$ and $\mathsf{bit}(\varepsilon)$.*

We conjecture that this can be generalised, and that as soon as we consider ε-NEs instead of NEs, double-exponentially small probabilities are no longer required.

▶ **Conjecture 18.** *Let $\mathcal{G}$ be a game, let $\bar{x}, \bar{y}$ be two vectors denoting the constraints, and let $\varepsilon > 0$. If there exists a stationary ε-NE in $\mathcal{G}$ where the expected payoffs are between $\bar{x}$ and $\bar{y}$, then there exists one that has polynomial representation.*

Proving such a result would lead to an NP upperbound on our approximate problem, since such an ε-NE can then be guessed and checked in polynomial time.

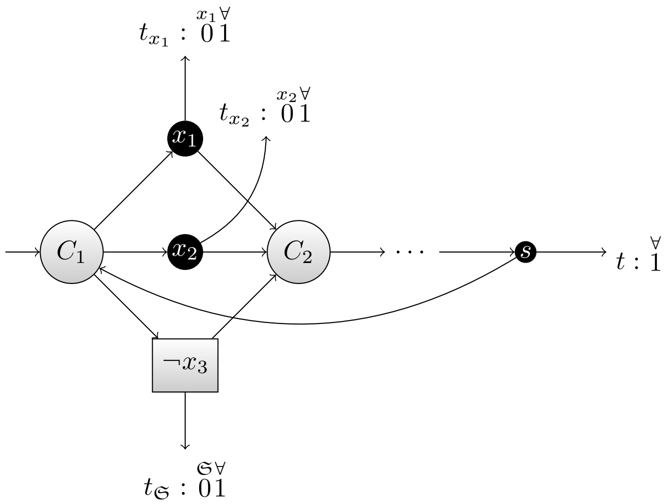


Figure 2 A reduction from 3SAT.

5 A lower bound

We dedicate this section to give a lower bound on our problem, and we show NP-hardness of the approximate version of the constrained existence problem.

▶ **Theorem 19.** *The approximate constrained existence problem of NEs is* NP*-hard.*

Proof. We proceed by reduction from the 3SAT problem. Let $\varphi = \bigvee_{i=1}^{m} (L_{i1} \wedge L_{i2} \wedge L_{i3})$ be a CNF formula, over the variable set X.

Construction of the game $\mathcal{G}^\varphi$

We define the game $\mathcal{G}^\varphi$ as follows.

- There is a player called *Solver* and written $\mathfrak{S}$, and for each variable $x \in X$, there is a player x.
- For each clause C_i, there is a vertex C_i controlled by Solver. The initial vertex is C_1.
- From each clause vertex C_i, Solver can move to the vertices (C_i, L_{i1}), (C_i, L_{i2}), and (C_i, L_{i3}).
- Each vertex of the form (C_i, x) (positive literals) is a stochastic vertex. From there, there are two edges, each taken with probability $\frac{1}{2}$: one to the terminal vertex t_x, and one to the next clause vertex C_{i+1} (or the vertex s if $i = m$).
- Each vertex of the form $(C_i, \neg x)$ is controlled by player x. From there, player x can move either to the terminal vertex $t_\mathfrak{S}$, or to the next clause vertex C_{i+1} (or$s$ if $i = m$).
- The vertex s is a stochastic vertex with two outgoing edges, each with probability $\frac{1}{2}$: one to the vertex C_1, one to the terminal vertex t.
- In the terminal vertex t_i, player i receives the reward 0, and every other player receives the reward 1. In the terminal vertex t, every player receives the reward 1.

That game is depicted by Figure 2, in a case where $C_1 = x_1 \vee x_2 \vee \neg x_3$. Solver controls the circle vertices, and the square vertex is controlled by player x_3. We wrote simply L for every vertex of the form (C, L). The symbol $\forall$ should be interpreted as *"every (other) player"*. Outgoing edges of stochastic vertices all have probability $\frac{1}{2}$.

Nash equilibria from a satisfiable formula

We now state the following result, whose proof is found in the full version of the paper.

▶ **Proposition 20.** *If the formula φ is satisfiable, then there exists a Nash equilibrium in the game $\mathcal{G}^\varphi$ such that Solver's expected payoff is* 1*.*

Non-existence of ε-Nash equilibria from unsatisfiable formula

For the converse, we show this stronger result, whose proof is also in full version of the paper.

▶ **Proposition 21.** *For m large enough, if the formula φ is not satisfiable, then there is no ε-Nash equilibrium in the game $\mathcal{G}^\varphi$ such that Solver's expected payoff is greater than or equal to $1-\varepsilon$, where $\varepsilon = 2^{-3m}$.*

To conclude, given the formula φ, we define the game $\mathcal{G}^\varphi$ as above, and $\varepsilon = 2^{-3m}$ (which can be written with polynomially many bits since m is given in unary by the formula). We show that, for m large enough, these games constructed are such that the unsatisfiable instances are mapped to the negative instance of the approximate existence problem (no approximate Nash equilibria satisfies the constraints), while the satisfiable formulas, and only them, are mapped to positive instances. Therefore, our problem is NP-hard. ◀

We conclude by saying that in the full version of this paper [2] we have a proof of NP hardness even when the number of players is at most two.

6 Conclusion and discussion

We studied the (functional) approximate constrained problem of NEs. Specifically, we asked the following: assuming the existence of a stationary Nash equilibrium that satisfies given payoff constraints, can we efficiently compute a stationary strategy profile that is an ε-Nash equilibrium and satisfies the constraints up to an additive error ε?

Although our results focus on turn-based stochastic games, we note that the lower bound also holds in more general settings, including concurrent games (games on graphs where the players select actions simultaneously) and broader objectives such as limit-average (mean-payoff) objectives. For the upper bound, we remark that the existential theory of reals formulations we use for turn-based games with reachability objectives can be naturally adapted to concurrent games with mean-payoff objectives, with only a constant increase in the degree of the polynomials involved. By following similar proof techniques, one can still obtain an $\mathsf{FNP}^{\mathsf{NP}}$ upper bound for the corresponding constrained ε-NE problem in these more general settings. However, we restrict our presentation and main results to turn-based games with reachability objectives in this paper to maintain clarity.

We showed that the decision version of this problem lies in the class $\mathsf{NP}^{\mathsf{NP}}$, and the corresponding search problem belongs to $\mathsf{FNP}^{\mathsf{NP}}$. On the other hand, we established only an NP-hardness lower bound. This leaves a gap in the known complexity of the problem, and closing this gap remains an open question.

One possible direction is to prove that if an ε-Nash equilibrium exists, then there also exists one in which all probabilities are at least inverse-exponential in the size of the input, assuming the existence of a Nash equilibrium. Such a result would not contradict our example in Section 4, which shows that double-exponentially small probabilities are required in the exact version of the problem. If exponentially small probabilities suffice in the approximate setting, then a strategy profile could be represented using only polynomially many bits in the fixed-point representation. Since such a strategy yields a polynomial certificate, and verification requires only finding the values of MDPs with fixed-point representation, this would yield an NP-completeness result for the decision version.

Alternatively, if such a bound does not hold, and there exists a counterexample where all ε-Nash equilibria satisfying the constraints require double-exponentially small probabilities, this would suggest an inherent complexity. It would also point to intrinsic limitations in our techniques, analogous to those demonstrated by Hansen, Koucký, and Miltersen [15], who showed that double-exponentially small probabilities are necessary for ε-optimal strategies in concurrent stochastic games. Their result implies that any algorithm that explicitly manipulates these probabilities must use at least exponential space in the worst case, and our results would also imply the same for our problem.

We end by recalling our conjecture in Section 4.

▶ **Conjecture 18.** *Let $\mathcal{G}$ be a game, let $\bar{x}, \bar{y}$ be two vectors denoting the constraints, and let $\varepsilon > 0$. If there exists a stationary ε-NE in $\mathcal{G}$ where the expected payoffs are between $\bar{x}$ and $\bar{y}$, then there exists one that has polynomial representation.*

References

1 Agarwal Alekh, Nan Jiang, Sham M. Kakade, and Sun Wen. *Reinforcement Learning: Theory and Algorithms.* https://rltheorybook.github.io/, 2021.

2 Ali Asadi, Léonard Brice, Krishnendu Chatterjee, and K. S. Thejaswini. ε-stationary nash equilibria in multi-player stochastic graph games, 2025. doi:10.48550/arXiv.2508.15356.

3 Ali Asadi, Krishnendu Chatterjee, Raimundo Saona, and Jakub Svoboda. Concurrent Stochastic Games with Stateful-Discounted and Parity Objectives: Complexity and Algorithms. In *FSTTCS 2024*, volume 323 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:17, 2024. doi:10.4230/LIPIcs.FSTTCS.2024.5.

4 Christel Baier and Joost-Pieter Katoen. *Principles of Model Checking.* The MIT Press, 2008.

5 Saugata Basu, Richard Pollack, and Marie-Françoise Roy. On the combinatorial and algebraic complexity of quantifier elimination. *Journal of the ACM (JACM)*, 43(6):1002–1045, 1996. doi:10.1145/235809.235813.

6 Sougata Bose, Rasmus Ibsen-Jensen, and Patrick Totzke. Bounded-memory strategies in partial-information games. In *Proceedings of the 39th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '24, New York, NY, USA, 2024. Association for Computing Machinery. doi:10.1145/3661814.3662096.

7 Léonard Brice, Thomas A. Henzinger, and K. S. Thejaswini. Finding Equilibria: Simpler for Pessimists, Simplest for Optimists. In *50th International Symposium on Mathematical Foundations of Computer Science (MFCS 2025)*, volume 345, pages 30:1–30:18, 2025. doi:10.4230/LIPIcs.MFCS.2025.30.

8 Krishnendu Chatterjee, Rupak Majumdar, and Marcin Jurdzinski. On Nash equilibria in stochastic games. In *Computer Science Logic, 18th International Workshop, CSL 2004*, volume 3210 of *Lecture Notes in Computer Science*, pages 26–40. Springer, 2004. doi:10.1007/978-3-540-30124-0_6.

9 Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009. doi:10.1137/070699652.

10 Luca de Alfaro, Thomas A. Henzinger, and Ranjit Jhala. Compositional methods for probabilistic systems. In *CONCUR 2001 — Concurrency Theory*, pages 351–365, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg. doi:10.1007/3-540-44685-0_24.

11 Argyrios Deligkas, John Fearnley, Themistoklis Melissourgos, and Paul G. Spirakis. Approximating the existential theory of the reals. In George Christodoulou and Tobias Harks, editors, *Web and Internet Economics - 14th International Conference, WINE 2018, Oxford,*

UK, December 15-17, 2018, Proceedings, volume 11316 of *Lecture Notes in Computer Science*, pages 126–139. Springer, 2018. doi:10.1007/978-3-030-04612-5_9.

12 Jerzy Filar and Koos Vrieze. *Competitive Markov Decision Processes*. Springer-Verlag, 1997.

13 Vojtěch Forejt, Marta Kwiatkowska, Gethin Norman, and David Parker. *Automated Verification Techniques for Probabilistic Systems*, pages 53–113. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. doi:10.1007/978-3-642-21455-4_3.

14 Søren Kristoffer Stiil Frederiksen and Peter Bro Miltersen. Approximating the value of a concurrent reachability game in the polynomial time hierarchy. In *ISAAC*, volume 8283 of *Lecture Notes in Computer Science*, pages 457–467. Springer, 2013. doi:10.1007/978-3-642-45030-3_43.

15 Kristoffer Arnsfelt Hansen, Michal Koucký, and Peter Bro Miltersen. Winning concurrent reachability games requires doubly-exponential patience. In *LICS*, pages 332–341. IEEE Computer Society, 2009. doi:10.1109/LICS.2009.44.

16 Kristoffer Arnsfelt Hansen and Steffan Christ Sølvsten. $\exists\mathbb{R}$-completeness of stationary nash equilibria in perfect information stochastic games. In *45th International Symposium on Mathematical Foundations of Computer Science, MFCS 2020*, volume 170 of *LIPIcs*, pages 45:1–45:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.MFCS.2020.45.

17 Claude Lefèvre. Optimal control of a birth and death epidemic process. *Operations Research*, 29(5):971–982, 1981. doi:10.1287/OPRE.29.5.971.

18 Richard J. Lipton, Evangelos Markakis, and Aranyak Mehta. Playing large games using simple strategies. In Daniel A. Menascé and Noam Nisan, editors, *Proceedings 4th ACM Conference on Electronic Commerce (EC-2003), San Diego, California, USA, June 9-12, 2003*, pages 36–41. ACM, 2003. doi:10.1145/779928.779933.

19 John F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950. doi:10.1073/pnas.36.1.48.

20 Dawei Shi, Robert J Elliott, and Tongwen Chen. On finite-state stochastic modeling and secure estimation of cyber-physical systems. *IEEE Transactions on Automatic Control*, 62(1):65–80, 2016. doi:10.1109/TAC.2016.2541919.

21 Eilon Solan. Continuity of the value of competitive Markov decision processes. *Journal of Theoretical Probability*, 16:831–845, 2003.

22 Florent Teichteil-Königsbuch, Ugur Kuter, and Guillaume Infantes. Incremental plan aggregation for generating policies in MDPs. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 - Volume 1*, AAMAS '10, pages 1231–1238. International Foundation for Autonomous Agents and Multiagent Systems, 2010. URL: https://dl.acm.org/citation.cfm?id=1838366.

23 Michael Ummels. *Stochastic multiplayer games: theory and algorithms*. PhD thesis, RWTH Aachen, Germany, January 2010.

24 Michael Ummels and Dominik Wojtczak. The Complexity of Nash Equilibria in Stochastic Multiplayer Games. *Logical Methods in Computer Science*, Volume 7, Issue 3, September 2011. doi:10.2168/LMCS-7(3:20)2011.