The Complexity of SPEs in Mean-Payoff Games

Léonard Brice ✉
Université libre de Bruxelles, Brussels, Belgium

Jean-François Raskin ✉
Université libre de Bruxelles, Brussels, Belgium

Marie van den Bogaard ✉
Univ Gustave Eiffel, CNRS, LIGM, F-77454 Marne-la-Vallée, France

Abstract

We establish that the subgame perfect equilibrium (SPE) threshold problem for mean-payoff games is **NP**-complete. While the SPE threshold problem was recently shown to be decidable (in doubly exponential time) and **NP**-hard, its exact worst case complexity was left open.

2012 ACM Subject Classification Software and its engineering → Formal methods; Theory of computation → Logic and verification; Theory of computation → Solution concepts in game theory

Keywords and phrases Games on graphs, subgame-perfect equilibria, mean-payoff objectives

Digital Object Identifier 10.4230/LIPIcs.ICALP.2022.116

Category Track B: Automata, Logic, Semantics, and Theory of Programming

Related Version *Full Version*: https://arxiv.org/abs/2202.08499 [4]

Funding This work is partially supported by the ARC project Non-Zero Sum Game Graphs: Applications to Reactive Synthesis and Beyond (Fédération Wallonie-Bruxelles), the EOS project Verifying Learning Artificial Intelligence Systems (F.R.S.-FNRS & FWO), the COST Action 16228 GAMENET (European Cooperation in Science and Technology), and by the PDR project *Subgame perfection in graph games* (F.R.S.- FNRS).

Acknowledgements We wish to thank the anonymous reviewers for their useful comments, in particular for the question that led us to add Lemma 12 to this paper.

1 Introduction

Nash equilibria (NEs), a fundamental solution concept in game theory, are defined as strategy profiles such that no player can improve their payoff by changing unilaterally their strategy. So NEs can be interpreted as self-inforcing contracts from which there is no incentive to deviate unilaterally. Unfortunately, NEs are known to suffer, in sequential games like infinite duration games played on graphs, from the issue of *non-credible threats*: to enforce a NE, some players may threaten other players to play irrationally in order to punish deviations. This is allowed by the definition of NEs, as in case of deviation from one player, the other players are not bound to rational behaviors anymore and they can therefore play irrationally w.r.t. their own objectives in order to sanction the deviating player. This drawback of NEs has triggered the introduction of the notion of *subgame-perfect equilibria* (SPEs) [22], a more complex but more natural solution concept for sequential games. A strategy profile is an SPE if after every history, i.e. in every *subgame*, the strategies of the players still form an NE. Thus, SPEs impose rationality even after a deviation and only rational behaviors can be used to coerce the behavior of other players.

In this paper, we study the complexity of SPE problems in infinite-duration sequential games played on graphs with mean-payoff objectives. While NEs always exist in those games, as proved in [8], SPEs do not always exist as shown in [23, 7]. The *SPE threshold problem*, i.e. the problem of deciding whether a given mean-payoff game admits an SPE satisfying some constraints on the payoffs it grants to the players, has recently been proved to be decidable in [3]. However, its worst-case computational complexity is open: [3] provides only an **NP** lower bound and a 2**ExpTime** upper bound. In this paper, we close this complexity gap and prove that the problem is actually **NP**-complete.

Contributions. The starting point of our algorithm is the characterization of SPEs recently presented in [3], based on the notions of *requirement* and *negotiation function.* A requirement on a game G is a function $\lambda : V \to \mathbb{R} \cup \{\pm\infty\}$, where V is the state space of G. For a given state v, the value $\lambda(v)$ should be understood as the minimal payoff that the player controlling the state v will require in a play traversing v in order to avoid deviating. A requirement captures, therefore, some level of *rationality* of the players. The negotiation function transforms each requirement λ into a (possibly stronger) requirement $\mathsf{nego}(\lambda)$, such that $\mathsf{nego}(\lambda)(v)$ is the best payoff that the player controlling v can ensure, while playing against a coalition of the other players that play rationally with regards to the requirement λ. A play is the outcome of an SPE if and only if it satisfies the requirement λ^*, the least fixed point of the negotiation function – or equivalently, one of its fixed points. We recall that result in Lemma 29. In order to obtain our nondeterministic polynomial time algorithm, the rest of the paper constructs a notion of witness recognizing the positive instances of the SPE threshold problem. Such witnesses admit three pieces. First, we show that the size of λ^* can be bounded by a polynomial function of the size of the game (Theorem 37). This result is obtained by showing that the set of fixed points of the negotiation function can be characterized by a finite union of polyhedra that in turn can be represented by linear inequations. While the number of inequations that are needed for that characterization may be large (it cannot be bounded polynomially), we show that each of those inequations have coefficients and constants whose binary representations can be bounded polynomially. As the least fixed point is the minimal value in this set, it is represented by a vertex of one of those polyhedra. Then this guarantees, using results that bounds the solutions of linear equalities, that the least fixed point has a binary representation that is polynomial and so it can be guessed in polynomial time by a nondeterministic algorithm: it will be the first piece of our notion of witness, in the non-deterministice algorithm we design to solve the SPE threshold problem. Second, we define a witness of polynomial size for the existence of a play, consistent with a given requirement, which generates a payoff vector between the desired thresholds (Theorem 38). This play is not guaranteed to be regular. Third, we define a witness of polynomial size to prove that a requirement is indeed a fixed point of the negotiation function. This notion of certificate relies on a new and more compact game characterization of the negotiation function called the *reduced negotiation game* (Definition 41, Theorem 44). These results are far from trivial as we also show that SPEs may rely on strategy profiles that are not regular and require infinite memory. As both the least fixed point and its two certificates can be guessed and verified in polynomial time, we obtain **NP** membership for the threshold problem, closing the complexity gap left open in [3] (Theorem 48).

Additionally, all the previous results do also apply to ε-SPEs, a quantitative relaxation of SPEs. In particular, Theorem 48 does also apply to the ε-SPE threshold problem.

Related works. Non-zero sum infinite duration games have attracted a large attention in recent years, with applications targeting reactive synthesis problems, see e.g. [1, 9, 10, 14, 19] and their references. We now detail other works more closely related to our contributions.

In [7], Brihaye et al. introduce and study the notion of weak SPE, which is a weakening of the classical notion of SPE. This weakening is equivalent to the original SPE concept on reward functions that are *continuous*. This is the case for example for the quantitative reachability reward function, on which Brihaye et al. solve the SPE threshold problem in [6]. The mean-payoff cost function is not continuous and the techniques used in [7], and generalized in [11], cannot be used to characterize SPEs for the mean-payoff reward function.

In [21], Meunier develops a method based on Prover-Challenger games to solve the problem of the existence of SPEs on games with a finite number of possible payoffs. In mean-payoff games, the number of possible payoffs is uncountably infinite.

In [15], Flesch and Predtetchinski present another characterization of SPEs on games with finitely many possible payoffs, based on a game structure with infinite state space. In [3], Brice et al. define the notions of requirements and negotiation function. They prove that the negotiation function is characterized by a zero-sum two-player game called *abstract negotiation game*, which is similar to the game introduced in the characterization of Flesch and Predtetchinski. As a starting point for algorithms, they also provide an effective representation of this game, called *concrete negotiation game*, which turns out to be a zero-sum finite state multi-mean-payoff games [24]. Finally, they use those tools to prove that the SPE threshold problem is decidable for mean-payoff games. They left open the question of its precise complexity: they provide a **NP** lower bound and a 2**ExpTime** upper bound. In [5], the same authors use those tools to close the complexity gap for the SPE threshold problem in parity games, which had been proved to be **ExpTime**-easy and **NP**-hard by Ummels and Grädel in [16]. They prove that the problem is actually **NP**-complete. The techniques used in that paper heavily rely on the fact that parity objectives are ω-regular, which is not the case of mean-payoff games in general.

In [13], Chatterjee et al. study mean-payoff *automata*, and give a result that can be translated into an expression of all the possible payoff vectors in a mean-payoff game. In [2], Brenguier and Raskin give an algorithm to build the Pareto curve of a multi-dimensional two-player zero-sum mean-payoff game. To do so, they study systems of equations and of inequations, and they prove that they always admit simple solutions (with polynomial size). Those technical results will be used along this paper.

Structure of the paper. In Section 2, we introduce the necessary background. Section 3 recalls the notions of requirement and negotiation function, and link them to NEs and SPEs. Section 4 recalls results about the size of solutions of systems of equations or inequations, and use them to bound the size of the least fixed point of the negotiation function. Section 5 defines a witness for the existence of a λ-consistent play between two given thresholds. Section 6 introduces the reduced negotiation game that is a new compact characterization of the negotiation function. Finally, Section 7 applies those results to prove the **NP**-completeness of the SPE threshold problem on mean-payoff games. The proofs that are not presented here can be found in appendices of [4], the full version of this paper.

2 Background

Games, strategies, equilibria. In all what follows, we study infinite duration turn-based quantitative games on finite graphs with complete information.

▶ **Definition 1** (Game). A *non-initialized game* is a tuple $G = (\Pi, V, (V_i)_{i\in\Pi}, E, \mu)$, where:

- Π is a finite set of *players*;
- (V, E) is a directed graph, called the *underlying graph* of G, whose vertices are sometimes called *states* and whose edges are sometimes called *transitions*, and in which every state has at least one outgoing transition. For the simplicity of writing, a transition $(v, w) \in E$ will often be written vw;

- $(V_i)_{i \in \Pi}$ is a partition of V, in which V_i is the set of states *controlled* by player i;
- $\mu : V^\omega \to \mathbb{R}^\Pi$ is an *payoff function*, that maps each infinite word ρ to the tuple $\mu(\rho) = (\mu_i(\rho))_{i \in \Pi}$ of the players' *payoffs*.

An *initialized game* is a tuple (G, v_0), often written $G_{\upharpoonright v_0}$, where G is a non-initialized game and $v_0 \in V$ is a state called *initial state*. We often use the word *game*, alone, for both initialized and non-initialized games.

▶ **Definition 2** (Play, history). A *play* (resp. history) in the game G is an infinite (resp. finite) path in the graph (V, E). It is also a play (resp. history) in the initialized game $G_{\upharpoonright v_0}$, when v_0 is its first vertex. The set of plays (resp. histories) in the game G (resp. the initialized game $G_{\upharpoonright v_0}$) is denoted by $\mathsf{Plays}G$ (resp. $\mathsf{Plays}G_{\upharpoonright v_0}, \mathsf{Hist}G, \mathsf{Hist}G_{\upharpoonright v_0}$). We write $\mathsf{Hist}_i G$ (resp. $\mathsf{Hist}_i G_{\upharpoonright v_0}$) for the set of histories in G (resp. $G_{\upharpoonright v_0}$) of the form hv, where v is a vertex controlled by player i.

Given a play ρ (resp. a history h), we write $\mathsf{Occ}(\rho)$ (resp. $\mathsf{Occ}(h)$) the set of vertices that appear in ρ (resp. h), and $\mathsf{Inf}(\rho)$ the set of vertices that appear infinitely often in ρ. For a given index k, we write $\rho_{\leq k}$ (resp. $h_{\leq k}$), or $\rho_{<k+1}$ (resp. $h_{<k+1}$), the finite prefix $\rho_0 \dots \rho_k$ (resp. $h_0 \dots h_k$), and $\rho_{\geq k}$ (resp. $h_{\geq k}$), or $\rho_{>k-1}$ (resp. $h_{>k-1}$), the infinite (resp. finite) suffix $\rho_k \rho_{k+1} \dots$ (resp. $h_k h_{k+1} \dots h_{|h|-1}$). Finally, we write $\mathsf{first}(\rho)$ (resp. $\mathsf{first}(h)$) the first vertex of ρ (and $\mathsf{last}(h)$ the last vertex of h).

▶ **Definition 3** (Strategy, strategy profile). A *strategy* for player i in the initialized game $G_{\upharpoonright v_0}$ is a function $\sigma_i : \mathsf{Hist}_i G_{\upharpoonright v_0} \to V$, such that $v\sigma_i(hv)$ is an edge of (V, E) for every hv. A history h is *compatible* with a strategy σ_i if and only if $h_{k+1} = \sigma_i(h_0 \dots h_k)$ for all k such that $h_k \in V_i$. A play ρ is compatible with σ_i if all its prefixes are.

A *strategy profile* for $P \subseteq \Pi$ is a tuple $\bar{\sigma}_P = (\sigma_i)_{i \in P}$, where each σ_i is a strategy for player i in $G_{\upharpoonright v_0}$. A play or a history is *compatible* with $\bar{\sigma}_P$ if it is compatible with every σ_i for $i \in P$. A *complete* strategy profile, usually written $\bar{\sigma}$, is a strategy profile for Π. Exactly one play is compatible with a complete strategy profile: we write it $\langle \bar{\sigma} \rangle$, and call it the *outcome* of $\bar{\sigma}$.

When i is a player and when the context is clear, we will often write $-i$ for the set $\Pi \setminus \{i\}$. When $\bar{\tau}_P$ and $\bar{\tau}'_Q$ are two strategy profiles with $P \cap Q = \emptyset$, we write $(\bar{\tau}_P, \bar{\tau}'_Q)$ the strategy profile $\bar{\sigma}_{P \cup Q}$ such that $\sigma_i = \tau_i$ for $i \in P$, and $\sigma_i = \tau'_i$ for $i \in Q$.

Before moving on to SPEs, let us recall that an NE is a strategy profile such that no player has an incentive to deviate unilaterally.

▶ **Definition 4** (Nash equilibrium). Let $G_{\upharpoonright v_0}$ be a game. The strategy profile $\bar{\sigma}$ is a *Nash equilibrium* – or *NE* for short – in $G_{\upharpoonright v_0}$ if and only if for each player i and for every strategy σ'_i, called *deviation of* σ_i, we have the inequality $\mu_i (\langle \sigma'_i, \bar{\sigma}_{-i} \rangle) \leq \mu_i (\langle \bar{\sigma} \rangle)$.

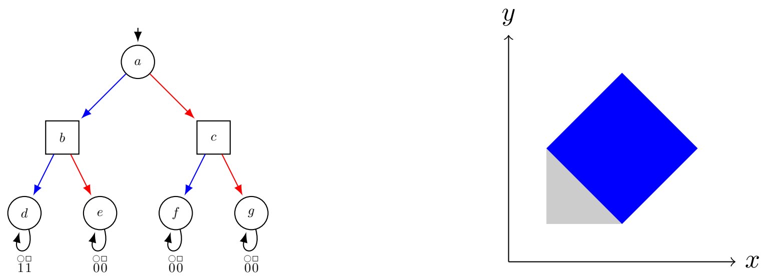
An SPE is a strategy profile whose all substrategy profiles are NEs.

▶ **Definition 5** (Subgame, substrategy). Let hv be a history in the game G. The *subgame* of G after hv is the game $(\Pi, V, (V_i)_i, E, \mu_{\upharpoonright hv})_{\upharpoonright v}$, where $\mu_{\upharpoonright hv}$ maps each play to its payoff in G, assuming that the history hv has already been played: formally, for every $\rho \in \mathsf{Plays}G_{\upharpoonright hv}$, we have $\mu_{\upharpoonright hv}(\rho) = \mu(h\rho)$. If σ_i is a strategy in $G_{\upharpoonright v_0}$, its *substrategy* after hv is the strategy $\sigma_{i \upharpoonright hv}$ in $G_{\upharpoonright hv}$, defined by $\sigma_{i \upharpoonright hv}(h') = \sigma_i(hh')$ for every $h' \in \mathsf{Hist}_i G_{\upharpoonright hv}$.

▶ Remark. The initialized game $G_{\upharpoonright v_0}$ is also the subgame of G after the one-state history v_0.

▶ **Definition 6** (Subgame-perfect equilibrium). Let $G_{\upharpoonright v_0}$ be a game. The strategy profile $\bar{\sigma}$ is a *subgame-perfect equilibrium* – or *SPE* for short – in $G_{\upharpoonright v_0}$ if and only if for every history h in $G_{\upharpoonright v_0}$, the strategy profile $\bar{\sigma}_{\upharpoonright h}$ is a Nash equilibrium in the subgame $G_{\upharpoonright h}$.

The notion of subgame-perfect equilibrium refines the notion of Nash equilibrium and excludes coercion by non-credible threats.



(a) Two Nash equilibria. **(b)** An example for the operator $\llcorner$.

Figure 1 Illustration for preliminary notions.

▶ Example 7. Consider the game pictured in Figure 1a. It is initialized with initial state a, and has two players, player $\bigcirc$ and player $\square$, who own respectively the circle and the square vertices. The payoff function assigns to each player a payoff of 1 for the play abd^ω, and 0 for all the other plays. Two different strategy profiles are represented here, one by the blue colored transitions, which has outcome abd^ω, one by the red colored ones, which has outcome acg^ω. Both are NEs: clearly, no player can increase their payoff by deviating from the blue choices, and in the case of the red profile, a deviation of player $\square$ can only lead to the play acf^ω, and a deviation of player $\bigcirc$ to abe^ω – both plays give to both player the payoff 0. However, for player $\square$, going from b to e is not a rational choice, hence the red profile is not an SPE, while the blue profile is one.

An ε-SPE is a strategy profile which is *almost* an SPE: if a player deviates after some history, they will not be able to improve their payoff by more than a quantity $\varepsilon \geq 0$. Note that a 0-SPE is an SPE, and conversely.

▶ **Definition 8** (ε-SPE)**.** Let $G_{\upharpoonright v_0}$ be a game, and $\varepsilon \geq 0$. A strategy profile $\bar{\sigma}$ from v_0 is an ε-SPE if and only if for every history h, for every player i and every strategy σ'_i, we have $\mu_i(\langle \bar{\sigma}_{-i\upharpoonright h}, \sigma'_{i\upharpoonright h} \rangle) \leq \mu_i(\langle \bar{\sigma}_{\upharpoonright h} \rangle) + \varepsilon$.

Mean-payoff games. We now turn to the definition of mean-payoff objectives.

▶ **Definition 9** (Mean-payoff, mean-payoff game)**.** In a graph (V, E), we associate to each mapping $r : E \to \mathbb{Q}$ the *mean-payoff function*:

$$\mathsf{MP}_r : h_0 \dots h_n \mapsto \frac{1}{n} \sum_{k=0}^{n-1} r_i\left(h_k h_{k+1}\right).$$

A game $G = (\Pi, V, (V_i)_i, E, \mu)$ is a *mean-payoff game* if its underlying graph is finite, and if there exists a tuple $(r_i)_{i \in \Pi}$ of reward functions, such that for each player i and every play ρ:

$$\mu_i(\rho) = \liminf_{n \to \infty} \mathsf{MP}_{r_i}(\rho_{\leq n}).$$

The mapping r_i is called *reward function* of player i: it represents the immediate reward that each action grants to player i. The final payoff of player i is their average payoff along the play, classically defined as the limit inferior[1] over n (since the limit may not be defined) of the average payoff after n steps. When the context is clear, we liberally write $\mathsf{MP}_i(h)$ for $\mathsf{MP}_{r_i}(h)$, and $\mathsf{MP}(h)$ for the tuple $(\mathsf{MP}_i(h))_i$, as well as $r(uv)$ for the tuple $(r_i(uv))_i$.

[1] An alternative definition of mean-payoff games exists, with a limit superior instead of inferior. While in zero-sum one dimensional games, the two definitions lead to the same notion of optimality, this is not the case when considering multiple dimensions, see e.g. [24]. All the results presented in this paper apply only on mean-payoff games defined with a limit inferior.

In the sequel, we develop a worst-case optimal algorithm to solve the *ε-SPE threshold problem*, which is a generalization of the *SPE threshold problem*, defined as follows.

▶ **Definition 10** (ε-SPE threshold problem)**.** Given a rational number $\varepsilon \geq 0$, a mean-payoff game $G_{\restriction v_0}$ and two thresholds $\bar{x}, \bar{y} \in \mathbb{Q}^\Pi$, does there exist an ε-SPE $\bar{\sigma}$ in $G_{\restriction v_0}$ such that $\bar{x} \leq \mu(\langle \bar{\sigma} \rangle) \leq \bar{y}$?

That problem is already known, by [3], to be 2**ExpTime** and **NP**-hard. The proof given in that paper does also show that **NP**-hardness still holds when ε is fixed to 0. Let us also add that the existence of an SPE in a given mean-payoff game, i.e. the same problem with no thresholds and with $\varepsilon = 0$, is itself **NP**-hard.

▶ **Definition 11** (SPE existence problem)**.** Given a mean-payoff game $G_{\restriction v_0}$, does there exist an SPE in $G_{\restriction v_0}$?

▶ **Lemma 12.** *The SPE existence problem is* **NP***-hard.*

Set of possible payoffs. A first important result that we need is the characterization of the set of possible payoffs in a mean-payoff game, which has been introduced in [13]. Given a graph (V, E), we write $\mathsf{SC}(V, E)$ the set of simple cycles it contains. Given a finite set D of dimensions and a set $X \subseteq \mathbb{R}^D$, we write $\mathsf{Conv} X$ the convex hull of X. We will often use the subscript notation $\mathsf{Conv}_{x \in X} f(x)$ for the set $\mathsf{Conv} f(X)$.

▶ **Definition 13** (Downward sealing)**.** Given a set $Y \subseteq \mathbb{R}^D$, the *downward sealing* of Y is the set ${}^\llcorner Y = \left\{ (\min_{\bar{z} \in Z} z_d)_{d \in D} \;\middle|\; Z \text{ is a finite subset of } Y \right\}$.

▶ Example 14. In $\mathbb{R}^2$, if Y is the blue area in Figure 1b, then ${}^\llcorner Y$ is the union of the blue area and the gray area.

▶ **Lemma 15** ([13])**.** *Let G be a mean-payoff game, whose underlying graph is strongly connected. The set of the payoffs $\mu(\rho)$, where ρ is a play in G, is exactly the set:*

$$
{}^\llcorner \left(\underset{c \in \mathsf{SC}(V,E)}{\mathsf{Conv}} \mathsf{MP}(c) \right).
$$

Two-player zero-sum games. We now recall several definitions and two classical results about two-player zero-sum games.

▶ **Definition 16** (Two-player zero-sum game)**.** A *two-player zero sum game* is a game G with $\Pi = \{1, 2\}$ and $\mu_2 = -\mu_1$.

▶ **Definition 17** (Borel game)**.** A game G is *Borel* if the function μ, from the set V^ω equipped with the product topology to the Euclidian space $\mathbb{R}^\Pi$, is Borel, i.e. if, for every Borel set $B \subseteq \mathbb{R}^\Pi$, the set $\mu^{-1}(B)$ is Borel.

▶ Remark. Mean-payoff games are Borel (see [12]).

▶ **Lemma 18** (Determinacy of Borel games, [20])**.** *Let $G_{\restriction v_0}$ be a zero-sum Borel game, with $\Pi = \{1, 2\}$. Then, we have the following equality:*

$$
\sup_{\sigma_1} \inf_{\sigma_2} \mu_1(\langle \bar{\sigma} \rangle) = \inf_{\sigma_2} \sup_{\sigma_1} \mu_1(\langle \bar{\sigma} \rangle).
$$

That quantity is called *value* of $G_{\restriction v_0}$, denoted by $\mathsf{val}_1(G_{\restriction v_0})$.

▶ **Definition 19** (Optimal strategy). Let $G_{\restriction v_0}$ be a zero-sum Borel game, with $\Pi = \{1, 2\}$. The strategy σ_1 is *optimal* in $G_{\restriction v_0}$ if $\inf_{\sigma_2} \mu_1(\langle \sigma_1, \sigma_2 \rangle) = \mathsf{val}_1(G_{\restriction v_0})$.

Let us now define memoryless strategies, and a condition under which they can be optimal.

▶ **Definition 20** (Memoryless strategy). A strategy σ_i in a game $G_{\restriction v_0}$ is *memoryless* if for every vertex $v \in V_i$ and for all histories h and h', we have $\sigma_i(hv) = \sigma_i(h'v)$.

We usually write $\sigma_i(\cdot v)$ for the state $\sigma_i(hv)$ for every h. For every game $G_{\restriction v_0}$, we write $\mathsf{ML}\left(G_{\restriction v_0}\right)$ for the set of memoryless strategies in $G_{\restriction v_0}$.

▶ **Definition 21** (Shuffling). Let ρ, η and θ be three plays in a game G. The play θ is a *shuffling* of ρ and η if there exist two sequences of indices $k_0 < k_1 < \dots$ and $\ell_0 < \ell_1 < \dots$ such that $\eta_0 = \rho_{k_0} = \eta_{\ell_0} = \rho_{k_1} = \eta_{\ell_1} = \dots$, and:

$$\theta = \rho_0 \dots \rho_{k_0-1} \eta_0 \dots \eta_{\ell_0-1} \rho_{k_0} \dots \rho_{k_1-1} \eta_{\ell_0} \dots \eta_{\ell_1-1} \dots .$$

▶ **Definition 22** (Convexity, concavity). A function $f : \mathsf{Plays}G \to \mathbb{R}$ is *convex* if every shuffling θ of two plays ρ and η satisfies $f(\theta) \geq \min\{f(\rho), f(\eta)\}$. It is *concave* if $-f$ is convex.

▶ Remark. Mean-payoff functions, defined with a limit inferior, are convex.

▶ **Lemma 23.** *In a two-player zero-sum game played on a finite graph, every player whose payoff function is concave has an optimal strategy that is memoryless.*

Proof. According to [18], this result is true for qualitative objectives, i.e. when μ can only take the values 0 and 1. It follows that for every $\alpha \in \mathbb{R}$, if a player i, whose payoff function is concave, has a strategy that ensures $\mu_i(\rho) \geq \alpha$ (understood as a qualitative objective), then they have a memoryless one. Hence the equality:

$$\mathsf{val}_1(G_{\restriction v_0}) = \sup_{\sigma_1 \in \mathsf{ML}(G_{\restriction v_0})} \inf_{\sigma_2} \mu_1(\langle \bar{\sigma} \rangle).$$

Since the underlying graph (V, E) is finite, memoryless strategies exist in finite number, hence the supremum above is realized by a memoryless strategy σ_1 that is, therefore, optimal. ◀

3 Requirements and negotiation

We now recall some notions and results from [3], which are the starting point of our algorithm.

Requirements. In the sequel, we write $\overline{\mathbb{R}}$ the set $\mathbb{R} \cup \{\pm\infty\}$.

▶ **Definition 24** (Requirement). A *requirement* on the game G is a mapping $\lambda : V \to \overline{\mathbb{R}}$.

For a given state v, the quantity $\lambda(v)$ represents the minimal payoff that the player controlling v will require in a play traversing the state v.

▶ **Definition 25** (λ-consistency). Let λ be a requirement on a game G. A play ρ in G is λ*-consistent* if and only if, for all $i \in \Pi$ and $n \in \mathbb{N}$ with $\rho_n \in V_i$, we have $\mu_i(\rho_{\geq n}) \geq \lambda(\rho_n)$. The set of λ-consistent plays from a state v is denoted by $\lambda\mathsf{Cons}(v)$.

▶ Remark. The set $\lambda\mathsf{Cons}(v)$ can be empty, and is not regular in general.

▶ **Definition 26** (λ-rationality). Let λ be a requirement on a mean-payoff game G. Let $i \in \Pi$. A strategy profile $\bar{\sigma}_{-i}$ is λ*-rational* if and only if there exists a strategy σ_i such that, for every history hv compatible with $\bar{\sigma}_{-i}$, the play $\langle \bar{\sigma}_{\restriction hv} \rangle$ is λ-consistent. We then say that the strategy profile $\bar{\sigma}_{-i}$ is λ-rational *assuming* σ_i. The set of λ-rational strategy profiles in $G_{\restriction v}$ is denoted by $\lambda\mathsf{Rat}(v)$.

Negotiation. In mean-payoff games, as well as in a wider class of games (see [3] and [15]), SPEs are characterized by the fixed points of the *negotiation function*, a function from the set of requirements into itself. We always use the convention $\inf \emptyset = +\infty$.

▶ **Definition 27** (Negotiation function). Let G be a game. The *negotiation function* is the function that transforms each requirement λ on G into a requirement $\mathsf{nego}(\lambda)$ on G defined, for each $i \in \Pi$ and $v \in V_i$, by:

$$\mathsf{nego}(\lambda)(v) = \inf_{\bar{\sigma}_{-i} \in \lambda\mathsf{Rat}(v)} \sup_{\sigma_i} \mu_i(\langle \bar{\sigma} \rangle).$$

The quantity $\mathsf{nego}(\lambda)(v)$ is the best payoff the player controlling the state v can enforce if the other players play rationally with regards to the requirement λ.

▶ Remark. The negotiation function satisfies the following properties.

- It is monotone: if $\lambda \leq \lambda'$ (for the pointwise order), then $\mathsf{nego}(\lambda) \leq \mathsf{nego}(\lambda')$.
- It is also non-decreasing: for every λ, we have $\lambda \leq \mathsf{nego}(\lambda)$.
- There exists a λ-rational strategy profile from v against the player controlling v if and only if $\mathsf{nego}(\lambda)(v) \neq +\infty$.

Link with SPEs. The SPE outcomes in a mean-payoff game are characterized by the fixed points of the negotiation function, or equivalently by its least fixed point. That result can be extended to ε-SPEs. To that end, we recall the notion of ε-fixed points of a function.

▶ **Definition 28** (ε-fixed point). Let $\varepsilon \geq 0$, let D be a finite set and let $f : \overline{\mathbb{R}}^D \to \overline{\mathbb{R}}^D$ be a mapping. A tuple $\bar{x} \in \mathbb{R}^D$ is a *ε-fixed point* of f if for each $d \in D$, for $\bar{y} = f(\bar{x})$, we have $y_d \in [x_d - \varepsilon, x_d + \varepsilon]$.

▶ Remark. A 0-fixed point is a fixed point, and conversely.

▶ **Lemma 29** ([3]). *Let $G_{\restriction v_0}$ be a mean-payoff game, and let $\varepsilon \geq 0$. The negotiation function on G has a least ε-fixed point λ^*, and given a play ρ in $G_{\restriction v_0}$, the three following assertions are equivalent: (1) the play ρ is an ε-SPE outcome; (2) the play ρ is λ-consistent for some ε-fixed point λ of the negotiation function; (3) the play ρ is λ^*-consistent.*

The abstract negotiation game. Given λ and $u \in V$, the quantity $\mathsf{nego}(\lambda)(u)$ can be characterized as the value of a *negotiation game*, a two-player zero-sum game opposing the player *Prover*, who simulates a λ-rational strategy profile and wants to minimize player i's payoff, and the player *Challenger*, who simulates player i's reaction by accepting or refusing Prover's proposals. Two negotiation games were defined in [3]. Conceptually simpler, the *abstract negotiation game* $\mathsf{Abs}_{\lambda i}(G)_{\restriction u}$ unfolds as follows:

- from the state v, Prover chooses a λ-consistent play ρ from the state v and proposes it to Challenger. If Prover has no play to propose, the game is over and Challenger gets the payoff $+\infty$.
- Once a play ρ has been proposed, Challenger can accept it. Or he can *deviate*, and choose a prefix $\rho_{\leq k}$ with $\rho_k \in V_i$ and a new transition $\rho_k w \in E$.
- In the former case, the game is over. In the latter, it starts again from the state w.

If Challenger finally accepts a proposal ρ, then his payoff is $\mu_i(\rho)$. If he deviates infinitely often, then Prover's proposals and his deviations construct a play $\dot{\pi} = \rho^{(0)}_{\leq k_0} \rho^{(1)}_{\leq k_1} \rho^{(2)}_{\leq k_2} \ldots$. Then, Challenger's payoff is $\mu_i(\dot{\pi})$. It has been proved in [3] that the equality $\mathsf{nego}(\lambda)(u) = \mathsf{val}_{\mathbb{C}}(\mathsf{Abs}_{\lambda i}(G)_{\restriction u})$ holds. Thus, the abstract negotiation game captures a first intuition on how the negotiation function can be computed.

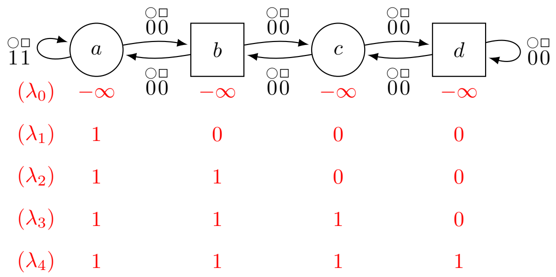


Figure 2 Iterations of the negotiation function.

▶ Example 30. Let G be the game of Figure 2, where each edge is labelled by the rewards $r_{\bigcirc}$ and $r_{\Box}$. Below the states, we present the requirements $\lambda_0 : v \mapsto -\infty$, $\lambda_1 = \mathsf{nego}(\lambda_0)$, $\lambda_2 = \mathsf{nego}(\lambda_1)$, $\lambda_3 = \mathsf{nego}(\lambda_2)$, and $\lambda_4 = \mathsf{nego}(\lambda_3)$. Let us explicate those computations, using the abstract negotiation game. From λ_0 to λ_1: since every play is λ_0-consistent, Prover can always propose whatever she wants. From the state a, whatever she (trying to minimize player $\bigcirc$'s payoff) proposes, Challenger can always make player $\bigcirc$ deviate in order to loop on the state a. Then, in the game G, player $\bigcirc$ gets the payoff 1, hence $\lambda_1(a) = 1$. From the state b, Prover (trying to minimize player $\Box$'s payoff) can propose the play $(bc)^\omega$. If Challenger makes player $\Box$ deviate to go to the state a, then Prover can propose the play $a(bc)^\omega$. Even if Challenger makes player $\Box$ deviate infinitely often, he cannot give him more than the payoff 0, hence $\lambda_1(b) = 0$. Similar situations happen from the states c and d, hence $\lambda_1(c) = \lambda_1(d) = 0$. From λ_1 to λ_2: now, from the state b, whatever Prover proposes at first, Challenger can make player $\Box$ deviate and go to the state a. From there, since we have $\lambda_1(a) = 1$, Prover has to propose a play in which player $\bigcirc$ gets the payoff 1. The only such plays do also give the payoff 1 to player $\Box$, hence $\lambda_2(b) = 1$. Similar situations explain $\lambda_3(c) = 1$, and $\lambda_4(c) = 1$. Finally, plays ending with the loop a^ω are all λ_4-consistent, hence Prover can always propose them, hence the requirement λ_4 is a fixed point of the negotiation function – and therefore the least. By Lemma 29, the SPE plays in G are exactly the plays in which both player $\bigcirc$ and player $\Box$ get the payoff 1.

The concrete negotiation game. The abstract negotiation game cannot be directly used for an algorithmic purpose, since it has an infinite state space. However, it can be turned into a game on a finite graph if Prover does not propose plays as a whole, but edge by edge. In the *concrete negotiation game* $\mathsf{Conc}_{\lambda i}(G)_{\restriction(u,\{u\})}$, the states controlled by Prover have the form (v, M), where $M \subseteq V$ memorizes the states seen since the last time Challenger did deviate, in order to control that the play Prover is constructing since that moment is effectively λ-consistent: for each $u \in M$, Prover has to give to the player controlling u at least the payoff $\lambda(u)$. Similarly, the states controlled by Challenger are of the form (vv', M), where $vv' \in E$ is an edge proposed by Prover. The game unfolds as follows:

- from the state (v, M), Prover chooses a transition vv' and proposes it to Challenger.
- Once a transition vv' has been proposed, Challenger can accept it. Or, if $v \in V_i$, he can *deviate*, and choose a new transition vw.
- If the former case, the game starts again from the state $(v', M \cup \{v'\})$. In the latter, it starts from the state $(w, \{w\})$.

For every play $\pi = (\rho_0, M_0)(\rho_0\rho_0', M_0)(\rho_1, M_1)(\rho_1\rho_1', M_1)\dots$ in $\mathsf{Conc}_{\lambda i}(G)$, we write $\dot\pi = \rho_0\rho_1\dots$ the play in G constructed by Prover's proposals and Challenger's deviations. Then, Challenger's payoff in the play π is either $+\infty$ if there exists an index k such that the suffix $\pi_{\geq 2k}$ contains no deviation and $\dot\pi_{\geq k}$ is not λ-consistent, and $\mu_i(\dot\pi)$ otherwise.

In [3], a first algorithm was proposed to solve the ε-SPE threshold problem, using the fact that in the concrete negotiation game, Challenger has a memoryless optimal strategy, to design a complete representation of the negotiation function, and to compute its least ε-fixed point. However, that algorithm requires doubly exponential time, because it needs to enumerate all the memoryless strategies available for Challenger, whose number is exponential in the size of the concrete game, itself exponential in the size of G. Here, we make use of the concrete negotiation game only to bound the size of the least ε-fixed point: our algorithm will use a third negotiation game, the *reduced negotiation game*.

4 Size of the least ε-fixed point

In this section, after having recalled some results about the sizes of solutions to linear equations and inequations, we prove that the least ε-fixed point of the negotiation function in a game G has a size that is polynomial in the size of G and ε. The first piece of the witnesses identifying positive instances of the ε-SPE threshold problem will then be an ε-fixed point of the negotiation function of polynomial size.

About size, equations and inequations. We define here the notion of size that we use.

▶ **Definition 31** (Size). The *size* of a rational number $r = \frac{p}{q}$, where $p, q \in \mathbb{Z}$ are co-prime, is the quantity $\|r\| = 1 + \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(|q| + 1) \rceil$. The size of an irrational number is $+\infty$. The size of the infinite numbers is $\| +\infty \| = \| -\infty \| = 1$. The size of a tuple $\bar{x} \in O^D$, where D is a set and O is a set of objects for which the notion of size has been defined, is the quantity $\mathsf{card} D + \sum_{d \in D} \|x_d\|$. Similarly, the size of a function $f : D \to X$ is the quantity $\mathsf{card} D + \sum_{d \in D} \|f(d)\|$, and the size of a set $X \subseteq O$ is the quantity $\mathsf{card} X + \sum_{x \in X} \|x\|$.

The proof of Theorem 37 below requires the manipulation of *polytopes*, e.g. downward sealings of convex hulls (from Lemma 15), expressed as solution sets of *systems of linear inequations*.

▶ **Definition 32** (Linear equations, inequations, systems). Let D be a finite set. A *linear equation* in $\mathbb{R}^D$ is a pair $(\bar{a}, b) \in \left(\mathbb{R}^D \setminus \left\{\bar{0}\right\}\right) \times \mathbb{R}$. The *solution set* of the equation $(\bar{a}, b)$ is the set $\mathsf{Sol}_=(\bar{a}, b) = \{\bar{x} \in \mathbb{R}^D \mid \bar{a} \cdot \bar{x} = b\}$, where $\cdot$ denotes the canonical scalar product on the euclidian space $\mathbb{R}^D$. A set $X \subseteq \mathbb{R}^D$ is a *hyperplane* of $\mathbb{R}^D$ if it is the solution set of some linear equation. A *system of linear equations* is a finite set Σ of linear equations. The *solution set* of the system Σ is the set $\mathsf{Sol}_= \Sigma = \bigcap_{(\bar{a}, b) \in \Sigma} \mathsf{Sol}_=(\bar{a}, b)$. A set $X \subseteq \mathbb{R}^D$ is a *linear subspace* of $\mathbb{R}^D$ if it is the solution set of some system of linear equations.

A *linear inequation* in $\mathbb{R}^D$ is a pair $(\bar{a}, b) \in \left(\mathbb{R}^D \setminus \left\{\bar{0}\right\}\right) \times \mathbb{R}$. The *solution set* of the inequation $(\bar{a}, b)$ is the set $\mathsf{Sol}_\geq(\bar{a}, b) = \{\bar{x} \in \mathbb{R}^D \mid \bar{a} \cdot \bar{x} \geq b\}$. A set $X \subseteq \mathbb{R}^D$ is a *half-space* of $\mathbb{R}^D$ if it is the solution set of some linear inequation. A *system of linear inequations* is a finite set Σ of linear inequations. The *solution set* of the system Σ is the set $\mathsf{Sol}_\geq \Sigma = \bigcap_{(\bar{a}, b) \in \Sigma} \mathsf{Sol}_\geq(\bar{a}, b)$. A set $X \subseteq \mathbb{R}^D$ is a *polyhedron* of $\mathbb{R}^D$ if it is the solution set of some system of linear inequations Σ. A *vertex* of X is a point $\bar{x} \in \mathbb{R}^D$ such that $\{\bar{x}\} = \mathsf{Sol}_=(\Sigma')$ for some subset $\Sigma' \subseteq \Sigma$. A *polytope* is a bounded polyhedron.

▶ Remark. Polyhedra are closed sets. The polytopes of $\mathbb{R}^D$ are exactly the sets of the form $\mathsf{Conv}(S)$, where S is a finite subset of $\mathbb{R}^D$.

▶ **Lemma 33** ([13]). *Let Σ be a system of inequations, and let $X = \mathsf{Sol}_\geq(\Sigma)$. The set ${}^\llcorner X$ is itself a polyhedron, and there exists a system of inequations Σ' such that ${}^\llcorner X = \mathsf{Sol}_\geq(\Sigma')$ and that for every $(\bar{a}', b') \in \Sigma'$, there exists $(\bar{a}, b) \in \Sigma$ with $\|(\bar{a}', b')\| \leq \|(\bar{a}, b)\|$.*

▶ **Lemma 34** ([2], Theorem 1). *There exists a polynomial P_1 such that, for every system of equations Σ, there exists a point $\bar{x} \in \mathsf{Sol}_=\Sigma$, such that $\|\bar{x}\| \leq P_1\left(\max_{(\bar{a},b)\in\Sigma}\|(\bar{a},b)\|\right)$.*

▶ **Corollary 35.** *For every system of inequations Σ, each vertex $\bar{x}$ of the polyhedron $\mathsf{Sol}_\geq(\Sigma)$ has size $\|x\| \leq P_1\left(\max_{(\bar{a},b)\in\Sigma}\|(\bar{a},b)\|\right)$.*

Note that in Lemma 33, in Lemma 34 and in Corollary 35, the number of equations or inequations has no influence. A consequence of Lemma 34 is the following result.

▶ **Lemma 36.** *There exists a polynomial P_2 such that, for each finite set D and every finite subset $X \subseteq \mathbb{R}^D$, there exists a system of linear inequations Σ, such that $\mathsf{Sol}_\geq(\Sigma) = \mathsf{Conv}(X)$ and $\|(\bar{a},b)\| \leq P_2(\|X\|)$ for every $(\bar{a},b) \in \Sigma$.*

Size of the least ε-fixed point. The following theorem bounds the size of the least fixed point of the negotiation function. As we used the notation $\bar{x}$ for tuples so far, we use the notation $\bar{\bar{x}}$ for tuples of tuples.

▶ **Theorem 37.** *There exists a polynomial P_3 such that for every mean-payoff game G, the least ε-fixed point λ^* of the negotiation function has size $\|\lambda^*\| \leq P_3(\|G\| + \|\varepsilon\|)$.*

Proof sketch. It has been proved in [3] that Challenger has a memoryless optimal strategy in every concrete negotiation game. Given a requirement λ, a player i, a state $v \in V_i$ and a memoryless strategy $\tau_{\mathbb{C}}$, we can construct the set of payoff vectors $\mu(\dot{\pi})$, where π is a play in $\mathsf{Conc}_{\lambda i}(G)_{\restriction(v,\{v\})}$ compatible with $\tau_{\mathbb{C}}$, as a union of polytopes defined using Lemma 15. If we intersect the upward closures of those sets, then $\mathsf{nego}(\lambda)(v)$ is equal to the least value x_i, where $\bar{x}$ belongs to that intersection. Therefore, if $X_\lambda \subseteq \mathbb{R}^{V\times\Pi}$ is the product of those intersections, then for each i and $v \in V_i$, we have $\mathsf{nego}(\lambda) = \inf\{x_{vi} \mid \bar{\bar{x}} \in X_\lambda\}$.

To each tuple of tuples $\bar{\bar{x}} \in \mathbb{R}^{V\times\Pi}$, we associate the requirement $\lambda_{\bar{\bar{x}}}$ defined by $\lambda_{\bar{\bar{x}}}(v) = x_{vi} - \varepsilon$ for each $i \in \Pi$ and $v \in V_i$. Then, we define $X = \left\{\bar{\bar{x}} \;\middle|\; \bar{\bar{x}} \in X_{\lambda_{\bar{\bar{x}}}}\right\}$, and we show that any requirement λ is an ε-fixed point of the negotiation function if and only $\lambda = \lambda_{\bar{\bar{x}}}$ for some $\bar{\bar{x}} \in X$. Then, the set X is itself a union of polyhedra, hence the linear mapping $\bar{\bar{x}} \mapsto \sum_v \lambda_{\bar{\bar{x}}}(v)$ has its minimum over X on some vertex $\bar{\bar{x}}$ of one of those polyhedra. The requirement λ^* is equal to $\lambda_{\bar{\bar{x}}}$, hence its size can be bounded using Corollary 35. ◀

5 Constrained existence of a λ-consistent play

We claim that a non-deterministic algorithm can recognize the positive instances of the ε-SPE threshold problem by guessing an ε-fixed point λ of the negotiation function. Once λ has been guessed, according to Lemma 29, two assertions must be proved: on the one hand, that there exists a λ-consistent play between the two desired thresholds, and on the other hand, that λ is actually an ε-fixed point of the negotiation function. The latter will be handled later through the concept of *reduced negotiation game*. Now, we tackle the former, and provide the second piece of our notion of witness: to prove the existence of a λ-consistent play ρ with $\bar{x} \leq \mu(\rho) \leq \bar{y}$, we need to guess the sets $W = \mathsf{Inf}(\rho)$ and $W' = \mathsf{Occ}(\rho)$, and a tuple of tuples $\bar{\bar{\alpha}} \in [0,1]^{\Pi\times\mathsf{SC}(W)}$ indicating how ρ combines the cycles of W, i.e. such that:

$$\mu(\rho) = \left(\min_{j\in\Pi} \sum_{c\in\mathsf{SC}(W)} \alpha_{jc}\mathsf{MP}_i(c)\right)_i.$$

▶ **Theorem 38.** *There exists a polynomial P_4 such that for every mean-payoff game $G_{\restriction v_0}$, for every $\bar{x},\bar{y} \in \mathbb{R}^V$, and for every requirement λ on G, there exists a λ-consistent play ρ in $G_{\restriction v_0}$ satisfying $\bar{x} \leq \mu(\rho) \leq \bar{y}$ if and only if there exist two sets $W \subseteq W' \subseteq V$ and a tuple of tuples $\bar{\bar{\alpha}} \in [0,1]^{\Pi\times\mathsf{SC}(W)}$ such that:*

- *the set W is strongly connected in (V, E), and accessible from the state v_0 using only and all the states of W';*
- *for each player i, we have $\sum_c \alpha_{ic} = 1$, and:*

$$x_i \leq \min_{j \in \Pi} \sum_{c \in \mathsf{SC}(W)} \alpha_{jc} \mathsf{MP}_i(c) \leq y_i;$$

- *for each player i and $v \in W \cap V_i$, we have:*

$$\min_{j \in \Pi} \sum_{c \in \mathsf{SC}(W)} \alpha_{jc} \mathsf{MP}_i(c) \geq \lambda(v);$$

- $\|\bar{\bar{\alpha}}\| \leq P_4(\|G, \bar{x}, \bar{y}, \lambda\|)$.

Proof. Let us first notice that given a set $X \subseteq \mathbb{R}^\Pi$, the elements of the set ${}^{\llcorner}(\mathsf{Conv}X)$ are exactly the tuples of the form:

$$\left(\min_{j \in \Pi} \sum_{x \in X} \alpha_{jx} x \right)_{i \in \Pi}$$

for some tuple $\bar{\bar{\alpha}} \in \mathbb{R}^{\Pi \times X}$ satisfying $\sum_x \alpha_{ix} = 1$ for each x.

Now, let us assume that W, W' and $\bar{\bar{\alpha}}$ exist. Then, there exists a play η with $\mathsf{Occ}(\eta) = \mathsf{Inf}(\eta) = W$ with payoff vector:

$$\mu(\eta) = \left(\min_{j \in \Pi} \sum_{x \in X} \alpha_{jx} x \right)_{i \in \Pi}.$$

Moreover, since W is accessible from v_0 using all and only the vertices of W', there exists a history $h\eta_0$ from v_0 to η_0 with $\mathsf{Occ}(h) = W'$. Then, the play $\rho = h\eta$ is λ-consistent and satisfies $\bar{x} \leq \mu(\rho) \leq \bar{y}$.

Conversely, if the play ρ exists: let $W = \mathsf{Inf}(\rho)$ and $W' = \mathsf{Occ}(\rho)$. The polytope:

$$Z = \left\{ \mu(\eta) \;\middle|\; \begin{array}{c} \eta \in \lambda\mathsf{Cons}(G_{\restriction v_0}), \\ \mathsf{Inf}(\eta) = W, \\ \mathsf{Occ}(\eta) = W', \\ \text{and } \bar{x} \leq \mu(\eta) \leq \bar{y} \end{array} \right\} = \left\{ \bar{z} \in {}^{\llcorner}\left(\underset{c \in \mathsf{SC}(W)}{\mathsf{Conv}} \mathsf{MP}(c) \right) \;\middle|\; \begin{array}{c} \bar{x} \leq \bar{z} \leq \bar{y}, \text{ and} \\ \forall i, \forall v \in W' \cap V_i, z_i \geq \lambda(v) \end{array} \right\}$$

(the equality holds by Lemma 15) is nonempty (it contains at least $\mu(\rho)$). By Lemma 36, the set $\mathsf{Conv}_{c \in SC(W)} \mathsf{MP}(c)$ is defined by a system of inequations which all have size $\|(\bar{a}, b)\| \leq P_2\left(\max_c \|\mathsf{MP}(c)\|\right)$. Since by Lemma 33, the inequations defining ${}^{\llcorner}\left(\mathsf{Conv}_{c \in SC(W)} \mathsf{MP}(c)\right)$ are not larger, there exists a polynomial P_6, independent of $G, \bar{x}, \bar{y}$ and λ, such that Z is defined by a system of inequations Σ such that for every $(\bar{a}, b) \in \Sigma$, we have $\|(\bar{a}, b)\| \leq P_6(\|(G, \bar{x}, \bar{y}, \lambda)\|)$. Therefore, by Corollary 35, the polytope Z admits a vertex $\bar{z}$ of size $\|\bar{z}\| \leq P_1(P_6(\|(G, \bar{x}, \bar{y}, \lambda)\|))$.

Then, since we have $\bar{z} \in {}^{\llcorner}\left(\underset{c \in \mathsf{SC}(W)}{\mathsf{Conv}} \mathsf{MP}(c) \right)$, that vertex is, according to Definition 13, of the form:

$$\bar{z} = \left(\min_j \sum_c \alpha_{jc} \mathsf{MP}_j(c) \right)_i$$

for some tuple of tuples $\bar{\bar{\alpha}} \in [0,1]^{\Pi \times \mathsf{SC}(W)}$ with $\sum_c \alpha_{ic} = 1$ and having, by Corollary 35 again, size $\|\bar{\bar{\alpha}}\| \leq P_1\left(\max_{i \in \Pi} \sum_{c \in \mathsf{SC}(W)} \|\mathsf{MP}_i(c)\| + \|z_i\|\right)$, i.e. $\|\bar{\bar{\alpha}}\| \leq P_4(\|(G, \bar{x}, \bar{y}, \lambda)\|)$ for some polynomial P_4 independent of $G, \bar{x}, \bar{y}$ and λ. ◀

Now, we need the third piece of our witness, which will be evidence of the fact that the requirement λ is an ε-fixed point of the negotiation function.

6 The reduced negotiation game

The abstract negotiation game has an infinite (and uncountable) state space in general, and the concrete negotiation game has an exponential one. In [5], the infiniteness of the abstract negotiation game has been handled in the case of parity games, by proving that Prover has an optimal strategy that is memoryless, and that proposes only simple plays with a finite representation. Unfortunately, this result does not apply to mean-payoff games, where Prover needs infinite memory in general. That fact is illustrated in the next example.

▶ Example 39. In the game of Figure 3a, the requirement λ defined by $\lambda(a) = \lambda(b) = 1$ is a fixed point of the negotiation function (it is actually the least fixed point). Indeed, from the state a (the situation is symmetrical from the state b), consider the strategy for Prover that proposes always, from the state v, the play $vb^{|h|^2}(a^3b^3)^\omega$, where h is the history that has already been constructed by her proposals and Challenger's deviations. If Challenger accepts such a play, then he gets the payoff 1. If he deviates infinitely often, then Prover loops longer and longer on the state b, and he also gets the payoff 1. The loop on b corresponds to what we will call later a *punishing cycle*. Now, if Prover uses only finite memory, Challenger can get a payoff better than 1 by always deviating and go to b as soon as he can: then, edges giving to player $\bigcirc$ the reward 2 will occur with a nonzero frequency.

However, the plays proposed by Prover in the previous example are very similar: only the number of repetitions of the loop b does increase. More generally, one observes that Prover can play optimally while always proposing a play of the form $hc^n\rho$, where h, c and ρ are constant, and only the number n increases, quadratically with the time – so that Challenger's payoff is dominated by the mean-payoff $\mathsf{MP}_i(c)$ if he deviates infinitely often.

▶ **Definition 40** (Punishment family)**.** A *punishment family* is a set of plays of the form:

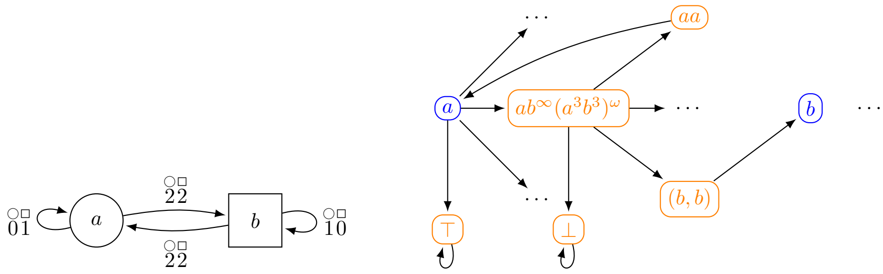
$$\{hc^n\rho \mid n > 0, \mu(\rho) = \bar{x}, \mathsf{Occ}(\rho) = W\}$$

where h is a simple history, c is a nonempty simple cycle, and where $W \subseteq V$ and $\bar{x} \in \mathbb{R}^\Pi$. The cycle c is called its *punishing cycle*. For every $\beta \in \mathbb{N}$, a *β-punishment family* is a punishment family with $\|\bar{x}\| \leq \beta$. A β-punishment family is represented by the data h, c, $\mu(\rho)$ and $\mathsf{Occ}(\rho)$, and that representation has a size smaller than or equal to the quantity $3\mathsf{card}V\lceil\log_2(\mathsf{card}V + 1)\rceil + \beta$.

We write $hc^\infty\rho$ for the punishment family $\{hc^n\rho' \mid n > 0, \mu(\rho') = \mu(\rho), \mathsf{Occ}(\rho') = \mathsf{Occ}(\rho)\}$. Beware that the play ρ matters only for its payoff vector and the vertices it traverses: if $\mathsf{Occ}(\rho) = \mathsf{Occ}(\rho')$ and $\mu(\rho) = \mu(\rho')$, then $hc^\infty\rho = hc^\infty\rho'$. We write $\mu(hc^\infty\rho)$ for the common payoff vector of all elements of $hc^\infty\rho$, and we will say that $hc^\infty\rho$ is λ-consistent if all its elements are (or equivalently, if one of its elements is). Let us clarify that a punishment family is not an equivalence class: for example, in the game of Figure 3a, the play ab^ω belongs to both $a^\infty b^\omega$ and $ab^\infty b^\omega$, which are distinct. We can now define the *reduced negotiation game*, where Prover proposes β-punishment families instead of plays.

▶ **Definition 41** (Reduced negotiation game)**.** Let G be a mean-payoff game, let λ be a requirement, let i be a player, let $v_0 \in V_i$ and let β be a natural integer. The corresponding *reduced negotiation game* is the game $\mathsf{Red}^\beta_{\lambda i}(G)_{\restriction v_0} = (\{\mathbb{P}, \mathbb{C}\}, S, (S_\mathbb{P}, S_\mathbb{C}), \Delta, \nu)_{\restriction v_0}$, where:

- the player $\mathbb{P}$ is called *Prover*, and the player $\mathbb{C}$ *Challenger*;
- the states controlled by Prover are the states of G, i.e. $S_\mathbb{P} = V$;
- the states controlled by Challenger are the states of the form $[hc^\infty\rho]$, (c, u) or $[h'v]$, where:



(a) The game G.   **(b)** The reduced negotiation game.

Figure 3 A game on which Prover needs infinite memory.

- $hc^\infty\rho$ is a λ-consistent β-punishment family,
- there exists a state $\rho_k \in V_i$ along the play ρ such that $\rho_k u \in E$,
- and $h'v$ is a history such that h' is a prefix of the history hc, and $\mathsf{last}(h') \in V_i$;

plus two additional states, written $\top$ and $\bot$;

- with the same notations, the set Δ contains the transitions of the form:
  - $v[hc^\infty\rho]$ (Prover proposes a punishment family);
  - $v\bot$ (Prover gives up);
  - $[hc^\infty\rho]\top$ (Challenger accepts Prover's proposal);
  - $[hc^\infty\rho][h'v]$ (Challenger deviates before the punishing cycle – *pre-cycle deviation*);
  - $[hc^\infty\rho](c, u)$ (Challenger deviates after the punishing cycle – *post-cycle deviation*);
  - $[h'v]v$ and $(c, u)u$ (Prover has now to propose a new play);
  - $\top\top$ and $\bot\bot$ (the play is over);
- given a history $H = H_0 \dots H_n \in \mathsf{HistRed}^{\beta}_{\lambda i}(G)$ that does not reach the state $\bot$, we write $\dot{H} = h^{(1)} \dots h^{(n)}$ the history or play in G defined by, for each k:
  - if $H_{k-1}H_k = v[hc^\infty\rho]$, then $h^{(k)}$ is empty;
  - if $H_{k-1}H_k = [hc^\infty\rho]\top$, then $h^{(k)} \dots h^{(n)} = hc^{|h^{(1)}\dots h^{(k-1)}h|^2}\rho$ (the number of times the cycle c is repeated depends quadratically on the time);
  - if $H_{k-1}H_k = [hc^\infty\rho][h'v]$, then $h^{(k)} = h'$;
  - if $H_{k-1}H_k = [hc^\infty\rho](c, v)$, then $h^{(k)} = hc^{|h^{(1)}\dots h^{(k-1)}h|^2}h'$, where h' is among the shortest histories such that $\mathsf{Occ}(h') \subseteq \mathsf{Occ}(\rho)$, $\mathsf{last}(h') \in V_i$ and $\mathsf{last}(h')v \in E$;
  - if $H_{k-1}H_k = [h'v]v$ or $(c, v)v$, then $h^{(k)}$ is empty;

  and that definition is naturally extended to plays: for example, if G is the game of Figure 3a and if $\pi = a[ab^\infty a^\omega][aa]a[ab^\infty a^\omega](b, b)b[b^\infty a^\omega]\top^\omega$, then $\dot{\pi} = a \cdot ab^{2^2} \cdot ab^{7^2}a^\omega = a^2b^4ab^{49}a^\omega$;
- the payoff function ν is defined, for each play π, by $\nu_{\mathbb{C}}(\pi) = -\nu_{\mathbb{P}}(\pi) = +\infty$ if π reaches the state $\bot$, and $\nu_{\mathbb{C}}(\pi) = -\nu_{\mathbb{P}}(\pi) = \mu_i(\dot{\pi})$ otherwise.

▶ Example 42. Figure 3b illustrates a (small) part of the game $\mathsf{Red}^2_{\lambda\bigcirc}(G)_{\restriction a}$, where G is the game of Figure 3a, and $\lambda(a) = \lambda(b) = 1$. Blue states are owned by Prover, orange ones by Challenger. When Prover proposes the punishment family $ab^\infty(a^3b^3)^\omega$, the function $\nu_{\mathbb{C}}$ interprets it as the play $ab^{|h|^2}(a^3b^3)^\omega$, where h is the history that has already been constructed so far.

▶ Remark. Reduced negotiation games are Borel, and are played on a finite graph.

Link with the negotiation function. We will now prove that the reduced negotiation game captures the negotiation function, as do the abstract and concrete ones. For that purpose, we first need the following key result.

▶ **Lemma 43.** *In a reduced negotiation game, Prover has a memoryless optimal strategy.*

Proof. This lemma is a consequence of Lemma 23: the payoff function $\nu_\mathbb{C}$ is concave. Indeed, let ξ be a shuffling of two plays π and χ. If either π or χ reaches the state $\bot$ (in which case both do), then we immediately have $\nu_\mathbb{C}(\dot{\xi}) \leq \max\{\nu_\mathbb{C}(\dot{\pi}), \nu_\mathbb{C}(\dot{\chi})\} = +\infty$. Otherwise, the play $\dot{\xi}$ is a shuffling of $\dot{\pi}$ and $\dot{\chi}$, and since mean-payoff objectives defined with a limit inferior are convex, we have $\nu_\mathbb{C}(\dot{\xi}) \leq \max\{\nu_\mathbb{C}(\dot{\pi}), \nu_\mathbb{C}(\dot{\chi})\}$. ◀

This lemma enables us to prove that the reduced negotiation game is equivalent to the other negotiation games.

▶ **Theorem 44.** *There exists a polynomial P_4 such that for every mean-payoff game G, every requirement λ with rational values, each player i and each $v_0 \in V_i$, for every $\beta \geq P_4(\|G\| + \|\lambda\|)$, we have* $\mathsf{nego}(\lambda)(v_0) = \mathsf{val}_\mathbb{C}\left(\mathsf{Red}^\beta_{\lambda i}(G)_{\restriction v_0}\right)$.

Proof. For every mean-payoff game G and every requirement λ, we assume:

$$\beta \geq P_1\left(P_2\left(\|\{\mathsf{MP}(c) \mid c \in \mathsf{SC}(G)\}\|\right)\right)$$

and for each $v \in V$:

$$\beta \geq \|\lambda(v)\| + 3,$$

which are indeed quantities that are bounded by a polynomial of $\|G\| + \|\lambda\|$.

- *First direction:* $\mathsf{nego}(\lambda)(v_0) \geq \mathsf{val}_\mathbb{C}\left(\mathsf{Red}^\beta_{\lambda i}(G)_{\restriction v_0}\right)$.
  Let $\bar{\sigma}_{-i}$ be a strategy profile in G that is λ-rational assuming a strategy σ_i, and let $x = \sup_{\sigma'_i} \mu_i(\langle \bar{\sigma}_{-i}, \sigma'_i \rangle)$. We wish to prove that there exists a strategy $\tau_\mathbb{P}$ in the reduced negotiation game such that $\sup_{\tau_\mathbb{C}} \nu_\mathbb{C}(\langle \bar{\tau} \rangle) \leq x$. Thus, we will have proved that the quantity $\mathsf{val}_\mathbb{C}\left(\mathsf{Red}^\beta_{\lambda i}(G)_{\restriction v_0}\right)$ is smaller than or equal to every such x, and therefore smaller than or equal to $\mathsf{nego}(\lambda)(v_0)$.
  Let us define simultaneously the strategy $\tau_\mathbb{P}$ and a mapping $\varphi : \mathsf{Hist}_\mathbb{P}\mathsf{Red}^\beta_{\lambda i}(G)_{\restriction v_0} \to \mathsf{Hist}G_{\restriction v_0}$, such that for each history H, the punishment family $\tau_\mathbb{P}(H)$ will be defined from the play $\langle \bar{\sigma}_{\restriction \varphi(H)} \rangle$. We guarantee inductively that if $H \in \mathsf{Hist}_\mathbb{P}\mathsf{Red}^\beta_{\lambda i}(G)_{\restriction v_0}$ is compatible with $\tau_\mathbb{P}$, then $\varphi(H) \in \mathsf{Hist}G_{\restriction v_0}$ is compatible with $\bar{\sigma}_{-i}$. First, let us define $\varphi(v_0) = v_0$.
  Let $H \in \mathsf{Hist}_\mathbb{P}\mathsf{Red}^\beta_{\lambda i}(G)_{\restriction v_0}$ be a history compatible with $\tau_\mathbb{P}$ as it has been defined so far, and such that $\varphi(H)$ has already been defined. Let $\eta^0 = \langle \bar{\sigma}_{\restriction \varphi(H)} \rangle$. By induction hypothesis, the history $\varphi(H)$ is compatible with $\bar{\sigma}_{-i}$, hence the play η^0 is λ-consistent, and satisfies $\mu_i(\eta^0) \leq x$.
  Let $\eta^0_{\leq \ell}$ be the shortest prefix of η^0 that is not simple, i.e. such that there exists $k < \ell$ with $\eta^0_k = \eta^0_\ell$. If $\mathsf{MP}_i\left(\eta^0_{k+1} \dots \eta^0_\ell\right) \leq x$, then we define $\tau_\mathbb{P}(H) = \left[\eta^0_{\leq k}\left(\eta^0_{k+1} \dots \eta^0_\ell\right)^\omega \rho\right]$, where ρ is a play such that $\mathsf{Occ}(\rho) = \mathsf{Occ}\left(\eta^0_{>\ell}\right)$, that $\mu_i(\rho) \leq x$, and that $\|\mu(\rho)\| \leq \beta$. Such a play exists, because the polytope:

$$Z = \left\{ \mu(\rho) \;\middle|\; \begin{array}{c} \forall j, \forall v \in V_j \cap \mathsf{Occ}(\eta^0), \mu_j(\rho) \geq \lambda(v), \\ \text{and } \mathsf{Occ}(\rho) = \mathsf{Occ}\left(\eta^0_{>\ell}\right) \end{array} \right\}$$

  is nonempty (it contains $\eta^0_{>\ell}$), and has at least one vertex $\bar{z}$ with $z_i \leq x$ (because $\mu_i(\eta^0_{>\ell}) \leq x$), which by Lemma 36 and Corollary 35 has size $\|\bar{z}\| \leq \beta$.

Otherwise, if $\mathsf{MP}_i\left(\eta^0_{k+1}\dots\eta^0_\ell\right) > x$, we define $\eta^1 = \eta^0_{\leq k}\eta^0_{>\ell}$, and we iterate the process, which does necessarily terminate – because $\mu_i(\eta^0) \leq x$. As a consequence, it effectively defines the proposal $\tau_{\mathbb{P}}(H) = \left[\eta^n_{\leq k}\left(\eta^n_{k+1}\dots\eta^n_\ell\right)^\infty\rho\right]$, for some n. Then, for each prefix hv, we define $\varphi\left(H\left[\eta^n_{\leq k}\left(\eta^n_{k+1}\dots\eta^n_\ell\right)^\infty\rho\right][hv]v\right) = \varphi(H)\eta^0_{\leq m}$, where $\eta^0_{\leq m}$ is the prefix of η^0 of which n simple cycles have been pulled out to obtain the prefix h of η^n; and similarly, for each pair (c,v), we define $\varphi\left(H\left[\eta^n_{\leq k}\left(\eta^n_{k+1}\dots\eta^n_\ell\right)^\infty\rho\right](c,v)v\right) = \varphi(H)\eta^0_{\leq m}$, where $\eta^0_{\leq m}$ is the prefix of η^0 from which n simple cycles have been pulled out to obtain the shortest prefix $\eta^n_{\leq p}$ of η^n such that $\eta^n_p \in V_i$ and $\eta^n_p v \in E$.

Thus, the mapping φ is defined on every history compatible with $\tau_{\mathbb{P}}$, and the image of such a history is always a history compatible with $\bar{\sigma}_{-i}$. We define it arbitrarily on other histories. Note that for each history H, the history $\dot{H}$ can be obtained from $\varphi(H)$ by pulling out cycles c satisfying $\mathsf{MP}_i(c) > x$, and adding cycles d with $\mathsf{MP}_i(d) \leq x$. As a consequence, if $\mathsf{MP}_i(\varphi(H)) \leq x$, then $\mathsf{MP}_i(\dot{H}) \leq x$ – and the same result is true when we naturally extend the mapping φ to plays.

Let us now prove that $\sup_{\tau_{\mathbb{C}}} \nu_{\mathbb{C}}(\langle\bar{\tau}\rangle) \leq \sup_{\sigma'_i} \mu_i(\langle\bar{\sigma}_{-i}, \sigma'_i\rangle)$. Let π be a play compatible with $\tau_{\mathbb{P}}$:

- the state $\perp$ does not appear in π, because Prover's strategy does never use a transition to it.
- If π has the form $\pi = H[hc^\infty\rho]\top^\omega$: then, we have $\nu_{\mathbb{C}}(\pi) = \mu_i(\rho) \leq x$.
- If π is made of infinitely many deviations: the play $\varphi(\pi)$ is compatible with $\bar{\sigma}_{-i}$, hence $\mu_i(\varphi(\pi)) \leq x$; which implies $\mu_i(\dot{\pi}) \leq x$, i.e. $\nu_{\mathbb{C}}(\pi) \leq x$.

- *Second direction:* $\mathsf{nego}(\lambda)(v_0) \leq \mathsf{val}_{\mathbb{C}}\left(\mathsf{Red}^\beta_{\lambda i}(G)_{\restriction v_0}\right)$.

Let $\tau_{\mathbb{P}}$ be a memoryless strategy for Prover in the reduced negotiation game, and let $y = \sup_{\tau_{\mathbb{C}}} \nu_{\mathbb{C}}(\langle\bar{\tau}\rangle)$. We want to show that $\mathsf{nego}(\lambda)(v_0) \leq y$: by Lemma 43, it will be enough to conclude. If $y = +\infty$, it is clear. Let us assume that $y \neq +\infty$. Then, we will define a strategy profile $\bar{\sigma}$, where $\bar{\sigma}_{-i}$ is λ-rational assuming σ_i, such that $\sup_{\sigma'_i} \mu_i(\langle\bar{\sigma}_{-i}, \sigma'_i\rangle) \leq y$: we proceed inductively by defining the play $\langle\bar{\sigma}_{\restriction hv}\rangle$ for each history hv compatible with $\bar{\sigma}_{-i}$ such that h is empty, or $\mathsf{last}(h) \in V_i$ and $v \neq \sigma_i(h)$. Such a history is called a *bud history*. After other histories, the strategy profile can be defined arbitrarily. To that end, we construct a mapping ψ which maps each bud history to a history $\psi(hv) \in \mathsf{Hist}_{\mathbb{P}}\mathsf{Red}^\beta_{\lambda i}(G)_{\restriction v_0}$ that is compatible with $\tau_{\mathbb{P}}$. This mapping will induce a definition of $\bar{\sigma}$: since $y \neq +\infty$, we have $\tau_{\mathbb{P}}(\psi(hv)) \neq \perp$: let then $[h'c^\infty\rho] = \tau_{\mathbb{P}}(\psi(hv))$. We then define $\langle\bar{\sigma}_{\restriction hv}\rangle = h'c^{|hh'|^2}\rho$, which is a λ-consistent play since $h'c^\infty\rho$ is a λ-consistent punishment family, by definition of the reduced negotiation game.

Let now h_0v be a bud history: we assume that $\bar{\sigma}$ has been defined on every prefix of h_0, but not on h_0v itself. If h_0 is empty, that is if $h_0v = v_0$, then we define $\psi(h_0v) = v_0$. Otherwise, let us write $h_0 = h_1wh_2$, where h_1w is the longest prefix of h_0 that is a bud history – that is, its longest prefix such that $\psi(h_1w)$ has been defined, or its shortest prefix such that wh_2 is compatible with $\bar{\sigma}_{\restriction h_1w}$. Let $H = \psi(h_1w)$, and let $[hc^\infty\rho] = \tau_{\mathbb{P}}(H)$. We have defined $\langle\bar{\sigma}_{\restriction h_1w}\rangle = hc^{|h_1h|^2}\rho$, and consequently, the history wh_2 is a prefix of that play. If it is a prefix of the history hc, then we define $\psi(h_0v) = H[hc^\infty\rho][wh_2v]v$. Otherwise, we define $\psi(h_0v) = H[hc^\infty\rho](c,v)v$.

Now, the strategy profile $\bar{\sigma}$ has been defined, and since all the punishment families proposed by Prover are λ-consistent, the strategy profile $\bar{\sigma}_{-i}$ is λ-rational assuming σ_i. Let η be a play compatible with $\bar{\sigma}_{-i}$, and let us prove that $\mu_i(\eta) \leq y$. If η has finitely many prefixes that are bud histories, then let $\eta_{\leq n}$ be the longest one: we have $\eta_{\geq n} = hc^{n+|h|}\rho$, where $[hc^\infty\rho] = \tau_{\mathbb{P}}(\psi(\eta_{\leq n}))$. Then, we have $\mu_i(\eta) = \mu_i(\rho) \leq y$.

Now, if η has infinitely many such prefixes, then there exists a unique play π in the reduced negotiation game such that for any prefix $\eta_{\leq n}$ of η that is a bud history, the history $\psi(\eta_{\leq n})$ is a prefix of π. Then, if π contains finitely many post-cycle deviations, then there exist two indices m and n such that $\eta_{\geq m} = \dot{\pi}_{\geq n}$, hence $\mu_i(\eta) = \mu_i(\dot{\pi}) \leq y$. Finally, if π contains infinitely many post-cycle deviations, i.e. infinitely many occurrences of a state of the form (c, v), then let us choose such state that minimizes the quantity $\mathsf{MP}_i(c)$. The play η has the form:

$$\eta = h_0 c^{k_0^2} h_1 c^{k_1^2} h_2 \dots,$$

where for each n, we have $k_n = \left| h_0 c^{k_0^2} \dots c^{k_{n-1}^2} h_n \right|$. Then, if we write $M = \max r_i$, we have:

$$\mathsf{MP}_i\left(h_0 c^{k_0^2} \dots h_n c^{k_n^2}\right) \leq \frac{1}{k_n + k_n^2|c| - 1}\left(k_n M + \left(k_n^2|c| - 1\right)\mathsf{MP}_i(c)\right),$$

which converges to $\mathsf{MP}_i(c)$ when n tends to $+\infty$, hence $\mu_i(\eta) \leq \mathsf{MP}_i(c)$. Now, since $\tau_{\mathbb{P}}$ is memoryless, there exists a play of the form HC^ω that is compatible with it, and such that $(c, v) \in \mathsf{Occ}(C) \subseteq \mathsf{Inf}(\pi)$; and by definition of y, we have $\mathsf{MP}_i(\dot{C}) = \nu_{\mathbb{C}}(HC^\omega) \leq y$. By minimality of $\mathsf{MP}_i(c)$, we have $\mathsf{MP}_i(\dot{C}) = \mathsf{MP}_i(c)$, hence $\mathsf{MP}_i(c) \leq y$, and therefore $\mu_i(\eta) \leq y$. ◀

Thus, a given requirement λ is an ε-fixed point of nego if and only if for each i and $v \in V_i$, there exists a memoryless strategy $\tau_{\mathbb{P}}$ in the game $\mathsf{Red}^{\beta}_{\lambda i}$, with $\beta = P_4(\|G\| + \|\lambda\|)$, such that $\sup_{\tau_{\mathbb{C}}} \nu_{\mathbb{C}}(\langle \bar{\tau} \rangle) \leq \lambda(v) + \varepsilon$. The reduced negotiation game has an exponential size, but it contains only $\mathsf{card} V$ states that are controlled by Prover: memoryless strategies for Prover are therefore objects of polynomial size. Thus, such memoryless strategies constitute the third and last piece of our notion of witness.

7 Algorithm and complexity

We are now in a position to define formally our notion of witness.

▶ **Definition 45** (Witness). Let $I = (G_{\upharpoonright v_0}, \bar{x}, \bar{y}, \varepsilon)$ be an instance of the ε-SPE threshold problem. A *witness* for I is a tuple $\left(W, W', \bar{\bar{\alpha}}, \lambda, (\tau_{\mathbb{P}}^v)_v\right)$, where $W \subseteq W' \subseteq V$; $\bar{\bar{\alpha}} \in [0,1]^{\Pi \times \mathsf{SC}(W)}$; λ is a requirement; and each $\tau_{\mathbb{P}}^v$ is a memoryless strategy in the game $\mathsf{Red}^{\beta}_{\lambda i}(G)_{\upharpoonright v}$, where $\beta = P_4(\|G\| + \|\lambda\|)$. A witness is *valid* if:

- each strategy $\tau_{\mathbb{P}}^v$ satisfies the inequality $\sup_{\tau_{\mathbb{C}}} \nu_{\mathbb{C}}(\langle \tau_{\mathbb{P}}^v, \tau_{\mathbb{C}} \rangle) \leq \lambda(v) + \varepsilon$;
- the sets W and W' and the tuple of tuples $\bar{\bar{\alpha}}$ satisfy the hypotheses of Theorem 38.

▶ Remark. The sets W and W', as well as the tuple of strategies $(\tau_{\mathbb{P}}^v)_v$, have polynomial size. In order to bound the size of witnesses by a polynomial, we only have to bound $\|\lambda\|$ and $\|\bar{\bar{\alpha}}\|$.

The ε-SPE threshold problem will be **NP**-easy if we show, first, that there exists a valid witness of polynomial size if and only if the instance is positive, and second, that the validity of a witness can be decided in polynomial time. The former is a consequence of Lemma 29, Theorem 37, Theorem 38, Theorem 44, and Lemma 43:

▶ **Lemma 46.** *There exists a polynomial P_5 such that an instance I of the ε-SPE threshold problem admits a valid witness of size $P_5(\|I\|)$ if and only if it is a positive instance.*

Let us now tackle the latter.

▶ **Lemma 47.** *Given an instance of the ε-SPE threshold problem and a witness for it, deciding whether that witness is valid is* $\mathbf{P}$*-easy.*

Proof. The validity of a witness is defined by two conditions. As regards the second one, all the hypotheses of Theorem 38 can be checked in polynomial time with classical algorithms. Let us now show how the first condition can also be checked in polynomial time.

Let $n = \mathsf{card} V$. Given a memoryless strategy $\tau_{\mathbb{P}}^v$ of Prover in a reduced negotiation game, one can construct in a time polynomial in $\|\tau_{\mathbb{P}}^v\|$ the graph $\mathsf{Red}_{\lambda i}^{\beta}(G)[\tau_{\mathbb{P}}^v]$, defined as the underlying graph of $\mathsf{Red}_{\lambda i}^{\beta}(G)$ where all the transitions that are not compatible with $\tau_{\mathbb{P}}^v$ have been omitted, as well as all the states that are, then, no longer accessible from the state v. That graph has indeed a polynomial size, because it is composed only of:

- at most n vertices of the form $w \in V$;
- at most n vertices of the form $\tau_{\mathbb{P}}(\cdot w)$ (either equal to $\perp$ or of the form $[hc^\infty\rho]$);
- at most n^2 vertices of the form (c, w);
- at most $2n^2$ vertices of the form $[h'w']$, where h' is a prefix of the history hc for some punishment family $[hc^\infty\rho] = \tau_{\mathbb{P}}(\cdot w)$;
- possibly the state $\top$.

We call this connected graph the *deviation graph.* Note that if among those vertices, there is the vertex $\perp$, then since the vertices that are not accessible have been removed, we have $\sup_{\tau_{\mathbb{C}}} \nu_{\mathbb{C}}(\langle\bar{\tau}\rangle) = +\infty$ and the problem can be solved immediately. In what follows, we assume that it is not the case, i.e. that for each w, the state $\tau_{\mathbb{P}}(\cdot w)$ has the form $[hc^\infty\rho]$. Deciding whether $\sup_{\tau_{\mathbb{C}}} \nu_{\mathbb{C}}(\langle\bar{\tau}\rangle) \leq \alpha$ is then equivalent to deciding whether there exists a path π, in that graph, such that $\nu_{\mathbb{C}}(\pi) > \alpha$. Such a play can have three forms.

- It can end in the state $\top$, i.e. with Challenger accepting Prover's proposal. The existence of such a play can be decided immediately, by checking whether in the deviation graph, there exists a vertex of the form $[hc^\infty\rho]$ with $\mu_i(hc^\infty\rho) > \alpha$.
- It can avoid the state $\top$, and comprise finitely many post-cycle deviations. This is the case if and only if there exists a cycle C in the deviation graph, without post-cycle deviations, such that $\mathsf{MP}_i(\dot{C}) > \alpha$. The existence of such a cycle can be decided in polynomial time with Karp's algorithm (see [17]).
- It can avoid the state $\top$, and comprise infinitely many post-cycle deviations. In that case, we have $\nu_{\mathbb{C}}(\pi) \leq \mathsf{MP}_i(c)$ for each state of the form (c, w) appearing infinitely often along π; then, there exists a cycle C in the deviation graph, such that every state of the form (c, w) along C satisfies $\mathsf{MP}_i(c) > \alpha$. Conversely, if such a cycle exists, then π exists. The existence of such a cycle can be decided in polynomial time with Karp's algorithm.

Therefore, the existence of such a play is decidable in polynomial time. ◀

Thus, given an instance of the ε-SPE threshold problem, a valid witness can be guessed and checked in polynomial time. Since the ε-SPE threshold problem has been proved to be $\mathbf{NP}$-hard in [3], we finally obtain the following theorem:

▶ **Theorem 48.** *The ε-SPE threshold problem in mean-payoff games is* $\mathbf{NP}$*-complete.*

References

1 Romain Brenguier, Lorenzo Clemente, Paul Hunter, Guillermo A. Pérez, Mickael Randour, Jean-François Raskin, Ocan Sankur, and Mathieu Sassolas. Non-zero sum games for reactive synthesis. In *Language and Automata Theory and Applications - 10th International Conference, LATA 2016, Prague, Czech Republic, March 14-18, 2016, Proceedings*, volume 9618 of *Lecture Notes in Computer Science*, pages 3–23. Springer, 2016.

2 Romain Brenguier and Jean-François Raskin. Pareto curves of multidimensional mean-payoff games. In Daniel Kroening and Corina S. Pasareanu, editors, *Computer Aided Verification - 27th International Conference, CAV 2015, San Francisco, CA, USA, July 18-24, 2015, Proceedings, Part II*, volume 9207 of *Lecture Notes in Computer Science*, pages 251–267. Springer, 2015.

3 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. Subgame-perfect equilibria in mean-payoff games. In Serge Haddad and Daniele Varacca, editors, *32nd International Conference on Concurrency Theory, CONCUR 2021, August 24-27, 2021, Virtual Conference*, volume 203 of *LIPIcs*, pages 8:1–8:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. `doi:10.4230/LIPIcs.CONCUR.2021.8`.

4 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. The complexity of spes in mean-payoff games. *CoRR*, abs/2202.08499, 2022. `arXiv:2202.08499`.

5 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. On the complexity of spes in parity games. In Florin Manea and Alex Simpson, editors, *30th EACSL Annual Conference on Computer Science Logic, CSL 2022, February 14-19, 2022, Göttingen, Germany (Virtual Conference)*, volume 216 of *LIPIcs*, pages 10:1–10:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. `doi:10.4230/LIPIcs.CSL.2022.10`.

6 Thomas Brihaye, Véronique Bruyère, Aline Goeminne, Jean-François Raskin, and Marie van den Bogaard. The complexity of subgame perfect equilibria in quantitative reachability games. In *CONCUR*, volume 140 of *LIPIcs*, pages 13:1–13:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

7 Thomas Brihaye, Véronique Bruyère, Noémie Meunier, and Jean-François Raskin. Weak subgame perfect equilibria and their application to quantitative reachability. In *24th EACSL Annual Conference on Computer Science Logic, CSL 2015, September 7-10, 2015, Berlin, Germany*, volume 41 of *LIPIcs*, pages 504–518. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015.

8 Thomas Brihaye, Julie De Pril, and Sven Schewe. Multiplayer cost games with simple nash equilibria. In *Logical Foundations of Computer Science, International Symposium, LFCS 2013, San Diego, CA, USA, January 6-8, 2013. Proceedings*, volume 7734 of *Lecture Notes in Computer Science*, pages 59–73. Springer, 2013.

9 Véronique Bruyère. Computer aided synthesis: A game-theoretic approach. In *Developments in Language Theory - 21st International Conference, DLT 2017, Liège, Belgium, August 7-11, 2017, Proceedings*, volume 10396 of *Lecture Notes in Computer Science*, pages 3–35. Springer, 2017.

10 Véronique Bruyère. Synthesis of equilibria in infinite-duration games on graphs. *ACM SIGLOG News*, 8(2):4–29, 2021.

11 Véronique Bruyère, Stéphane Le Roux, Arno Pauly, and Jean-François Raskin. On the existence of weak subgame perfect equilibria. In *Foundations of Software Science and Computation Structures - 20th International Conference, FOSSACS 2017, Uppsala, Sweden, April 22-29, 2017, Proceedings*, volume 10203 of *Lecture Notes in Computer Science*, pages 145–161, 2017.

12 Krishnendu Chatterjee. Concurrent games with tail objectives. *Theor. Comput. Sci.*, 388(1-3):181–198, 2007.

13 Krishnendu Chatterjee, Laurent Doyen, Herbert Edelsbrunner, Thomas A. Henzinger, and Philippe Rannou. Mean-payoff automaton expressions. In Paul Gastin and François Laroussinie, editors, *CONCUR 2010 - Concurrency Theory, 21th International Conference, CONCUR 2010, Paris, France, August 31-September 3, 2010. Proceedings*, volume 6269 of *Lecture Notes in Computer Science*, pages 269–283. Springer, 2010.

14 Dana Fisman, Orna Kupferman, and Yoad Lustig. Rational synthesis. In Javier Esparza and Rupak Majumdar, editors, *Tools and Algorithms for the Construction and Analysis of Systems, 16th International Conference, TACAS 2010, Paphos, Cyprus, March 20-28, 2010. Proceedings*, volume 6015 of *Lecture Notes in Computer Science*, pages 190–204. Springer, 2010.

15 János Flesch and Arkadi Predtetchinski. A characterization of subgame-perfect equilibrium plays in borel games of perfect information. *Math. Oper. Res.*, 42(4):1162–1179, 2017.

16 Erich Grädel and Michael Ummels. Solution Concepts and Algorithms for Infinite Multiplayer Games. In Krzysztof Apt and Robert van Rooij, editors, *New Perspectives on Games and Interaction*, volume 4 of *Texts in Logic and Games*, pages 151–178. Amsterdam University Press, 2008.

17 Richard M. Karp. A characterization of the minimum cycle mean in a digraph. *Discret. Math.*, 23(3):309–311, 1978.

18 Eryk Kopczynski. Half-positional determinacy of infinite games. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming, 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proceedings, Part II*, volume 4052 of *Lecture Notes in Computer Science*, pages 336–347. Springer, 2006.

19 Orna Kupferman, Giuseppe Perelli, and Moshe Y. Vardi. Synthesis with rational environments. *Ann. Math. Artif. Intell.*, 78(1):3–20, 2016.

20 Donald A. Martin. Borel determinacy. *Annals of Mathematics*, pages 363–371, 1975.

21 Noémie Meunier. *Multi-Player Quantitative Games: Equilibria and Algorithms.* PhD thesis, Université de Mons, 2016.

22 Martin J. Osborne. *An introduction to game theory.* Oxford Univ. Press, 2004.

23 Eilon Solan and Nicolas Vieille. Deterministic multi-player dynkin games. *Journal of Mathematical Economics*, 39(8):911–929, 2003.

24 Yaron Velner, Krishnendu Chatterjee, Laurent Doyen, Thomas A. Henzinger, Alexander Moshe Rabinovich, and Jean-François Raskin. The complexity of multi-mean-payoff and multi-energy games. *Inf. Comput.*, 241:177–196, 2015.