Fragility and Robustness in Mean-Payoff Adversarial Stackelberg Games

Mrudula Balachander
Université libre de Bruxelles, Brussels, Belgium

Shibashis Guha
Tata Institute of Fundamental Research, Mumbai, India

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

Abstract

Two-player mean-payoff Stackelberg games are nonzero-sum infinite duration games played on a bi-weighted graph by Leader (Player 0) and Follower (Player 1). Such games are played sequentially: first, Leader announces her strategy, second, Follower chooses his best-response. If we cannot impose which best-response is chosen by Follower, we say that Follower, though strategic, is adversarial towards Leader. The maximal value that Leader can get in this nonzero-sum game is called the adversarial Stackelberg value (ASV) of the game.

We study the robustness of strategies for Leader in these games against two types of deviations: (i) Modeling imprecision - the weights on the edges of the game arena may not be exactly correct, they may be delta-away from the right one. (ii) Sub-optimal response - Follower may play epsilon-optimal best-responses instead of perfect best-responses. First, we show that if the game is zero-sum then robustness is guaranteed while in the nonzero-sum case, optimal strategies for ASV are fragile. Second, we provide a solution concept to obtain strategies for Leader that are robust to both modeling imprecision, and as well as to the epsilon-optimal responses of Follower, and study several properties and algorithmic problems related to this solution concept.

2012 ACM Subject Classification Theory of computation → Solution concepts in game theory; Theory of computation → Mathematical optimization; Theory of computation → Logic and verification

Keywords and phrases mean-payoff, Stackelberg games, synthesis

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2021.9

Related Version *Full Version*: `https://arxiv.org/pdf/2007.07209.pdf`

Funding This work is partially supported by the PDR project Subgame perfection in graph games (F.R.S.-FNRS), 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 and FWO), and the COST Action 16228 GAMENET (European Cooperation in Science and Technology).

1 Introduction

Stackelberg games [16] were first introduced to model strategic interactions among rational agents in markets that consist of Leader and Follower(s). Leader in the market makes her strategy public and Follower(s) respond by playing an optimal response to this strategy. Here, we consider Stackelberg games as a framework for the synthesis of reactive programs [15, 3]. These programs maintain a continuous interaction with the environment in which they operate; they are *deterministic functions* that given a history of interactions so far choose an action. Our work is a contribution to *rational synthesis* [9, 14], a nonzero-sum game setting where both the program and the environment are considered as rational agents that have their own goals. While Boolean ω-regular payoff functions have been studied in [9, 14], here we study the quantitative long-run average (mean-payoff) function.

We illustrate our setting with the example in Figure 1. The set V of vertices is partitioned into V_0 (represented by circles) and V_1 (represented by squares) that are owned by Leader (also called Player 0) and Follower (also called Player 1) respectively. In the tuple on the edges, the first element is the payoff of Leader, while the second one is the payoff of Follower (weights are omitted if they are both equal to 0). Each player's objective is to maximize the long run average of the payoffs that she receives (a.k.a. mean-payoff). In the adversarial Stackelberg setting, Player 0 (Leader) first announces how she will play then Player 1 (Follower) chooses one of his best-responses to this strategy. Here, there are two choices for Player 0: L or R. As Player 1 is assumed to be rational, Player 0 deduces that she must play L. Indeed, the best response of Player 1 is then to play LL and the reward she obtains is 10. This is better than playing R, for which the best-response of Player 1 is RL, and the reward is 8 instead of 10. Note that if there are several possible best responses for Player 1, then we consider the worst-case: Player 0 has no control on the choice of best-responses by Player 1.

Quantitative models and robustness. The study of adversarial Stackelberg games with mean-payoff objectives has been started in [8] with the concept of *adversarial Stackelberg value* (ASV for short). ASV is the best value that Leader can obtain by fixing her strategy and facing any rational response by Follower. As this setting is quantitative, it naturally triggers questions about *robustness* that were left open in this first paper.

Robustness is a *highly desirable* property of quantitative models: small changes in the quantities appearing in a model M (e.g. rewards, probabilities, etc.) should have small impacts on the predictions made from M, see e.g. [2]. Robustness is thus crucial because it accounts for modelling imprecision that are inherent in quantitative modelling and those imprecision may have important consequences. For instance, a reactive program synthesized from a model M should provide acceptable performances if it is executed in a real environment that differ slightly w.r.t. the quantities modeled in M.

Some classes of models are robust. For instance, consider two-player *zero-sum* mean-payoff games where players have fully antagonistic objectives. The *value* of a two-player zero-sum mean-payoff $\mathcal{G}$ is the maximum mean-payoff that Player 0 can ensure against all strategies of Player 1. A strategy σ_0 that enforces the optimal value c in $\mathcal{G}$ is *robust* in the following sense. Let $\mathcal{G}^{\pm\delta}$ be the set of games obtained by increasing or decreasing the weights on the edges of $\mathcal{G}$ by at most δ. Then for all $\delta > 0$, and for all $\mathcal{H} \in \mathcal{G}^{\pm\delta}$, the strategy σ_0 ensures in $\mathcal{H}$ a mean-payoff of at least $c - \delta$ for Player 0 against any strategy of Player 1 (Proposition 1). So slight changes in the quantities appearing in the model have only a small impact on the worst-case value enforced by the strategy.

The situation is more complex and less satisfactory in *nonzero-sum* games. Strategies that enforce the ASV proposed in [8] may be *fragile*: slight differences in the weights of the game, or in the optimality of the response by Player 1, may lead to large differences in the value obtained by the strategy. We illustrate these difficulties on our running example. The strategy of Player 0 that chooses L in v_0 ensures her a payoff of 10 which is the ASV. Indeed, the unique best-response of Player 1 against L is to play LL from v_1. However, if the weights in $\mathcal{G}$ are changed by up to $\pm\delta = \pm 0.6$ then there is a game $\mathcal{H} \in \mathcal{G}^{\pm\delta}$ in which the weight on the self-loop over vertex v_4 changes to e.g. 9.55, and the weight on the self-loop over v_3 changes to e.g. 9.45, and the action LR becomes better for Player 1. So the value of L in $\mathcal{H}$ against a rational adversary is now 0 instead of 10. Thus a slight change in the rewards for Player 1 (due to e.g. modelling imprecision) may have a dramatic effect on the value of the optimal strategy L computed on the model $\mathcal{G}$ when evaluated in $\mathcal{H}$.

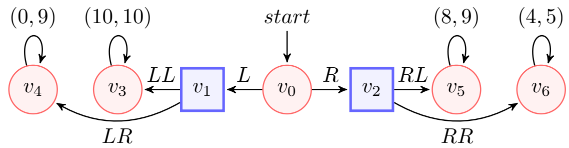

Figure 1 A game in which the strategy of Leader that maximizes the adversarial Stackelberg is fragile while the strategy of Leader that maximizes the $\epsilon = 1$-adversarial Stackelberg value is robust.

Contributions. As a remedy to this situation, we provide an alternative notion of value that is better-suited to synthesize strategies that are robust against perturbations. We consider two types of perturbations. First, the strategies computed for this value are robust against *modeling imprecision*: if a strategy has been synthesized from a weighted game graph with weights that are possibly slightly wrong, the value that this strategy delivers is guaranteed to be close to what the model predicts. Second, strategies computed for this value are robust against *sub-optimal responses*: small deviations from the best-response by the adversary have only limited effect on the value guaranteed by the strategy.

Our solution relies on relaxing the notion of best-responses of Player 1 *in the original model* $\mathcal{G}$: we define the ϵ-adversarial Stackelberg value (ASV^ϵ, for short) as the value that Leader can enforce against all ϵ-best responses of Follower. Obviously, this directly accounts for the second type of perturbations. But we show that, additionally, this accounts for the first type of perturbations: if a strategy σ_0 enforces an ASV^ϵ equal to c then for all games $\mathcal{H} \in G^{\pm\frac{\epsilon}{2}}$, we have that σ_0 enforce a value larger than $c - \epsilon$ in $\mathcal{H}$ (Theorem 5 and Theorem 6).

We illustrate this by considering again the example of Figure 1. Here, if we consider that the adversary can play $2\delta = 1.2$-best responses instead of best responses only, then the optimal strategy of Player 0 is now R and it has a ASV^ϵ equal to 8. This value is guaranteed to be robust for all games $\mathcal{H} \in \mathcal{G}^{\pm\delta}$ as R is guaranteed to enforce a payoff that is larger than $8 - \delta$ in all games in $\mathcal{G}^{\pm\delta}$. Stated otherwise, we use the notion of ASV^ϵ in the original game to find a strategy for Player 0 that she uses in the perturbed model while playing against a rational adversary. Thus we show that in the event of modelling imprecision resulting in a perturbed model, the solution concept to be used is ASV^ϵ instead of ASV since the former provides strategies that are robust to such perturbations.

Table 1 Summary of our results.

	Robustness	Threshold Problem	Computing ASV	Achievability
Adversarial best responses of Follower	**No** **[Proposition 2]**	NP [8] **Finite Memory Strategy [1]** **Memoryless Strategy [1]**	Theory of Reals [8]	No [8]
Adversarial ϵ-best responses of Follower	**Yes** **[Thm 5]**	**NP** **Finite Memory Strategy [Thm 8]** **Memoryless Strategy [Thm 10]**	**Theory of Reals [Thm 12]** **Solving LP in EXPTime [Thm 12]**	**Yes [Thm 16]** **(Requires Infinite Memory [Thm 19])**

In addition to proving the fragility of the original concept introduced in [8] (Proposition 2) and the introduction of the new notion of value ASV^ϵ that is robust against modelling imprecision (Theorem 5), we provide algorithms to handle ASV^ϵ. First, we show how to decide the threshold problem for ASV^{ϵ} in nondeterministic polynomial time and that finite memory strategies suffice (Theorem 8). Second, we provide an algorithm to compute ASV^{ϵ} when ϵ is fixed (Theorem 12). Third, we provide an algorithm that given a threshold value c, computes the largest possible ϵ such that $\mathsf{ASV}^{\epsilon} > c$ (Corollary 15). These three results form the core technical contributions of this paper and they are presented in Section 4 and Section 5. Additionally, in Section 6, we show that ASV^{ϵ} is always achievable (Theorem 16), which is in contrast to the case in [8] where Follower only plays best-responses. Finally, we provide results that concern the memory needed for players to play optimally, and complexity results for subcases (for example when Players are assumed to play memoryless). Our contributions have been summarized in Table 1, where the results obtained in this work are in bold. The new results corresponding to ASV can be found in [1].

Related Works. Stackelberg games on graphs have been first considered in [9], where the authors study rational synthesis for ω-regular objectives with co-operative Follower(s). In [8], Stackelberg mean-payoff games in adversarial setting, and Stackelberg discounted sum games in both adversarial and co-operative setting have been considered. However, as pointed out earlier, the model of [8] is not robust to perturbations. In [10], mean-payoff Stackelberg games in the co-operative setting have been studied. In [13], the authors study the effects of limited memory on both Nash and Stackelberg (or leader) strategies in multi-player discounted sum games. Incentive equilibrium over bi-matrix games and over mean-payoff games in a co-operative setting have been studied in [11] and [12] respectively. In [14], adversarial rational synthesis for ω-regular objectives have been studied. In [7], precise complexity results for various ω-regular objectives have been established for both adversarial and co-operative settings. In [6, 4], secure Nash equilibrium has been studied, where each player first maximises her own payoff, and then minimises the payoff of the other player; Player 0 and Player 1 are symmetric there unlike in Stackelberg games. For discounted sum objectives, in [8], the gap problem has been studied. Given rationals c and δ, a solution to the gap problem can decide if $\mathsf{ASV} > c + \delta$ or $\mathsf{ASV} < c - \delta$. The threshold problem was left open in [8], and is technically challenging. We leave the case of analysing robustness for discounted sum objective for future work.

A full version of this work with detailed proofs appears in [1].

2 Preliminaries

We denote by $\mathbb{N}$, $\mathbb{N}^{+}$, $\mathbb{Q}$, and $\mathbb{R}$ the set of naturals, the set of naturals excluding 0, the set of rationals, and the set of reals respectively.

Arenas. An (bi-weighted) arena $\mathcal{A} = (V, E, \langle V_0, V_1\rangle, w_0, w_1)$ consists of a finite set V of vertices, a set $E \subseteq V \times V$ of edges such that for all $v \in V$ there exists $v' \in V$ and $(v, v') \in E$, a partition $\langle V_0, V_1\rangle$ of V, where V_0 (resp. V_1) is the set of vertices for Player 0 (resp. Player 1), and two edge weight functions $w_0 : E \to \mathbb{Z}$, $w_1 : E \to \mathbb{Z}$. In the sequel, we denote the maximum absolute value of a weight in $\mathcal{A}$ by W. A strongly connected component of a directed graph is a subgraph that is strongly connected. Unless otherwise mentioned, *SCC* denotes a subgraph that is strongly connected, and which may or may not be maximal.

Plays and histories. A play in $\mathcal{A}$ is an infinite sequence of vertices $\pi = \pi_0\pi_1 \cdots \in V^{\omega}$ such that for all $k \in \mathbb{N}$, we have $(\pi_k, \pi_{k+1}) \in E$. A *history* in $\mathcal{A}$ is a (non-empty) prefix of a play in $\mathcal{A}$. Given $\pi = \pi_0\pi_1 \cdots \in \mathsf{Plays}_{\mathcal{A}}$ and $k \in \mathbb{N}$, the prefix $\pi_0\pi_1 \dots \pi_k$ of π is denoted by $\pi_{\leqslant k}$.

We denote by $\inf(\pi)$ the set of vertices v that appear infinitely many times along π, i.e., $\inf(\pi) = \{v \in V \mid \forall i \in \mathbb{N} \cdot \exists j \in \mathbb{N}, j \geqslant i : \pi(j) = v\}$. It is easy to see that $\inf(\pi)$ forms an SCC in the underlying graph of the arena $\mathcal{A}$. We denote by $\mathsf{Plays}_{\mathcal{A}}$ and $\mathsf{Hist}_{\mathcal{A}}$ the set of plays and the set of histories in $\mathcal{A}$ respectively; the symbol $\mathcal{A}$ is omitted when clear from the context. Given $i \in \{0,1\}$, the set $\mathsf{Hist}^i_{\mathcal{A}}$ denotes the set of histories such that their last vertex belongs to V_i. We denote the first vertex and the last vertex of a history h by $\mathsf{first}(h)$ and $\mathsf{last}(h)$ respectively.

Games. A *mean-payoff game* $\mathcal{G} = (\mathcal{A}, \langle \underline{\mathsf{MP}}_0, \underline{\mathsf{MP}}_1 \rangle)$ consists of a bi-weighted arena $\mathcal{A}$, payoff functions $\underline{\mathsf{MP}}_0 : \mathsf{Plays}_{\mathcal{A}} \to \mathbb{R}$ and $\underline{\mathsf{MP}}_1 : \mathsf{Plays}_{\mathcal{A}} \to \mathbb{R}$ for for Player 0 and Player 1 respectively which are defined as follows. Given a play $\pi \in \mathsf{Plays}_{\mathcal{A}}$ and $i \in \{0,1\}$, the payoff $\underline{\mathsf{MP}}_i(\pi)$ is given by $\underline{\mathsf{MP}}_i(\pi) = \liminf_{k \to \infty} \frac{1}{k} w_i(\pi_{\leqslant k})$, where the weight $w_i(h)$ of a history $h \in \mathsf{Hist}$ is the sum of the weights assigned by w_i to its edges. In our definition of the mean-payoff, we have used $\liminf$ as the limit of the successive average may not exist.

Strategies and payoffs. A strategy for Player $i \in \{0,1\}$ in the game $\mathcal{G}$ is a function $\sigma : \mathsf{Hist}^i_{\mathcal{A}} \to V$ that maps histories ending in a vertex $v \in V_i$ to a successor of v. The set of all strategies of Player $i \in \{0,1\}$ in the game $\mathcal{G}$ is denoted by $\Sigma_i(\mathcal{G})$, or Σ_i when $\mathcal{G}$ is clear from the context. A strategy has memory M if it can be realized as the output of a state machine with M states. A memoryless strategy is a function that only depends on the last element of the history $h \in \mathsf{Hist}$. We denote by Σ_i^{ML} the set of memoryless strategies of Player i, and by Σ_i^{FM} her set of finite memory strategies. A *profile* is a pair of strategies $\overline{\sigma} = (\sigma_0, \sigma_1)$, where $\sigma_0 \in \Sigma_0(\mathcal{G})$ and $\sigma_1 \in \Sigma_1(\mathcal{G})$. As we consider games with perfect information and deterministic transitions, any profile $\overline{\sigma}$ yields, from any history h, a unique play or *outcome*, denoted $\mathsf{Out}_h(\mathcal{G}, \overline{\sigma})$. Formally, $\mathsf{Out}_h(\mathcal{G}, \overline{\sigma})$ is the play π such that $\pi_{\leqslant |h|-1} = h$ and $\forall k \geqslant |h| - 1$ it holds that $\pi_{k+1} = \sigma_i(\pi_{\leqslant k})$ if $\pi_k \in V_i$. We write $h \leqslant \pi$ whenever h is a prefix of π. The set of outcomes compatible with a strategy $\sigma \in \Sigma_{i \in \{0,1\}}(\mathcal{G})$ after a history h is $\mathsf{Out}_h(\mathcal{G}, \sigma) = \{\pi | \exists \sigma' \in \Sigma_{1-i}(\mathcal{G})$ such that $\pi = \mathsf{Out}_h(\mathcal{G}, (\sigma, \sigma'))\}$. Each outcome $\pi \in \mathcal{G} = (\mathcal{A}, \langle \underline{\mathsf{MP}}_0, \underline{\mathsf{MP}}_1 \rangle)$ yields a payoff $\underline{\mathsf{MP}}(\pi) = (\underline{\mathsf{MP}}_0(\pi), \underline{\mathsf{MP}}_1(\pi))$.

Usually, we consider instances of games such that the players start playing at a fixed vertex v_0. Thus, we call an initialized game a pair $(\mathcal{G}, v_0)$, where $\mathcal{G}$ is a game and $v_0 \in V$ is the initial vertex. When v_0 is clear from context, we use $\mathcal{G}$, $\mathsf{Out}(\mathcal{G}, \overline{\sigma})$, $\mathsf{Out}(\mathcal{G}, \sigma)$, $\underline{\mathsf{MP}}(\overline{\sigma})$ instead of $\mathcal{G}_{v_0}$, $\mathsf{Out}_{v_0}(\mathcal{G}, \overline{\sigma})$, $\mathsf{Out}_{v_0}(\mathcal{G}, \sigma)$, $\underline{\mathsf{MP}}_{v_0}(\overline{\sigma})$. We sometimes omit $\mathcal{G}$ when it is clear from the context.

Best-responses, ϵ-best-responses. Let $\mathcal{G} = (\mathcal{A}, \langle \underline{\mathsf{MP}}_0, \underline{\mathsf{MP}}_1 \rangle)$ be a two-dimensional mean-payoff game on the bi-weighted arena $\mathcal{A}$. Given a strategy σ_0 for Player 0, we define

1. Player 1's best responses to σ_0, denoted by $\mathsf{BR}_1(\sigma_0)$, as:

$$\{\sigma_1 \in \Sigma_1 \mid \forall v \in V . \forall \sigma_1' \in \Sigma_1 : \mathsf{MP}_1(\mathsf{Out}_v(\sigma_0, \sigma_1)) \geqslant \mathsf{MP}_1(\mathsf{Out}_v(\sigma_0, \sigma_1'))\}$$

2. Player 1's ϵ-best-responses to σ_0, for $\epsilon > 0$[1], denoted by $\mathsf{BR}_1^{\epsilon}(\sigma_0)$, as:

$$\{\sigma_1 \in \Sigma_1 \mid \forall v \in V \cdot \forall \sigma_1' \in \Sigma_1 : \underline{\mathsf{MP}}_1(\mathsf{Out}_v(\sigma_0, \sigma_1)) > \underline{\mathsf{MP}}_1(\mathsf{Out}_v(\sigma_0, \sigma_1')) - \epsilon\}$$

[1] Since we will use ϵ in ASV^{ϵ} to add robustness, we only consider the cases in which ϵ is strictly greater than 0.

We also introduce the following notation for zero-sum games (that are needed as intermediary steps in our algorithms). Let $\mathcal{A}$ be an arena, $v \in V$ one of its states, and $\mathcal{O} \subseteq \mathsf{Plays}_{\mathcal{A}}$ be a set of plays (called objective), then we write $\mathcal{A}, v \models \ll i \gg \mathcal{O}$, if:

$$\exists \sigma_i \in \Sigma_i \cdot \forall \sigma_{1-i} \in \Sigma_{1-i} : \mathsf{Out}_v(\mathcal{A}, (\sigma_i, \sigma_{1-i})) \in \mathcal{O}, \text{for } i \in \{0,1\}$$

All the zero-sum games we consider in this paper are *determined* meaning that for all $\mathcal{A}$, for all objectives $\mathcal{O} \subseteq \mathsf{Plays}_{\mathcal{A}}$ we have that $\mathcal{A}, v \models \ll i \gg \mathcal{O} \iff \mathcal{A}, v \not\models \ll 1-i \gg \mathsf{Plays}_{\mathcal{A}} \setminus \mathcal{O}$. We sometimes omit $\mathcal{A}$ when the arena being referenced is clear from the context.

Convex hull and $\mathsf{F}_{\min}$. Given a finite dimension d, a finite set $X \subset \mathbb{Q}^d$ of rational vectors, we define the convex hull $\mathsf{CH}(X) = \{v \mid v = \sum_{x \in X} \alpha_x \cdot x \wedge \forall x \in X : \alpha_x \in [0,1] \wedge \sum_{x \in X} \alpha_x = 1\}$ as the set of all their convex combinations. Let $f_{\min}(X)$ be the vector $v = (v_1, v_2, \ldots, v_d)$ where $v_i = \min\{c \mid \exists x \in X : x_i = c\}$ i.e. the vector v is the pointwise minimum of the vectors in X. For $S \subseteq \mathbb{Q}^d$, we define $\mathsf{F}_{\min}(S) = \{f_{\min}(P) \mid P \text{ is a finite subset of } S\}$.

Mean-payoffs induced by simple cycles. A *cycle* c is a sequence of edges that starts and stops in a given vertex v, it is simple if it does not contain repetition of any other vertex. Given an SCC S, we write $\mathbb{C}(S)$ for the set of simple cycles inside S. Given a simple cycle c, for $i \in \{0,1\}$, let $\mathsf{MP}_i(c) = \frac{w_i(c)}{|c|}$ be the mean of the weights[2] in each dimension along the edges in the simple cycle c, and we call the pair $(\mathsf{MP}_0(c), \mathsf{MP}_1(c))$ the mean-payoff coordinate of the cycle c. We write $\mathsf{CH}(\mathbb{C}(S))$ for the convex-hull of the set of mean-payoff coordinates of simple cycles of S.

Adversarial Stackelberg Value for MP. Since the set of best-responses in mean-payoff games can be empty (See Lemma 3 of [8]), we use the notion of ϵ-best-responses for the definition of ASV which are guaranteed to always exist[3]. We define

$$\mathsf{ASV}(v) = \sup_{\sigma_0 \in \Sigma_0, \epsilon > 0} \inf_{\sigma_1 \in \mathsf{BR}_1^{\epsilon}(\sigma_0)} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)).^4$$

We also associate a (adversarial) value to a strategy $\sigma_0 \in \Sigma_0$ of Player 0, denoted

$$\mathsf{ASV}(\sigma_0)(v) = \sup_{\epsilon > 0} \inf_{\sigma_1 \in \mathsf{BR}_1^{\epsilon}(\sigma_0)} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)).$$

Clearly, we have that $\mathsf{ASV}(v) = \sup_{\sigma_0 \in \Sigma_0} \mathsf{ASV}(\sigma_0)(v)$.

We define the adversarial Stackelberg values, where strategies of Player 0 are restricted to finite memory strategies, as

$$\mathsf{ASV}_{\mathsf{FM}}(v) = \sup_{\sigma_0 \in \Sigma_0^{\mathsf{FM}}} \inf_{\sigma_1 \in \mathsf{BR}_1(\sigma_0)} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1))$$

where Σ_0^{FM} refers to the set of all finite memory strategies of Player 0. We note that for every finite memory strategy σ_0 of Player 0, a best-response of Player 1 to σ_0 always exists as noted in [8].

We also define the adversarial Stackelberg values, where Player 0 is restricted to using memoryless strategies, as

$$\mathsf{ASV}_{\mathsf{ML}}(v) = \sup_{\sigma_0 \in \Sigma_0^{\mathsf{ML}}} \inf_{\sigma_1 \in \mathsf{BR}_1(\sigma_0)} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1))$$

where Σ_0^{ML} is the set of all memoryless strategies of Player 0.

---

2 We do not use $\underline{\mathsf{MP}}_i$ since lim inf and lim sup are the same for a finite sequence of edges.
3 For a game $\mathcal{G}$, we also use $\mathsf{ASV}_{\mathcal{G}}$ and $\mathsf{ASV}^{\epsilon}_{\mathcal{G}}$, and drop the subscript $\mathcal{G}$ when it is clear from the context.
4 The definition of ASV, as it appears in [8], is syntactically different but the two definitions are equivalent, and the one presented here is simpler.

In the sequel, unless otherwise mentioned, we refer to a two-dimensional nonzero-sum two-player mean-payoff game simply as a mean-payoff game.

Zero-sum case. Zero-sum games are special cases of nonzero-sum games, where for all edges $e \in E$, we have that $w_0(e) = -w_1(e)$, i.e. the gain of one player is always equal to the opposite (the loss) of the other player. For zero-sum games, the classical concept is the notion of (worst-case) value. It is defined as

$$\mathsf{Val}_{\mathcal{G}}(v) = \sup_{\sigma_0 \in \Sigma_0} \inf_{\sigma_1 \in \Sigma_1} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)).$$

Additionally, we define the *value* of a Player 0 strategy σ_0 from a vertex v in a zero-sum mean-payoff game $\mathcal{G}$ as $\mathsf{Val}_{\mathcal{G}}(\sigma_0)(v) = \inf_{\sigma_1 \in \Sigma_1} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1))$.

3 Fragility and robustness in games

In this section, we study fragility and robustness properties in *zero-sum* and *nonzero-sum* games. Additionally, we provide a notion of value, for the nonzero-sum case, that is well-suited to synthesize strategies that are robust against two types of perturbations:

- Modeling imprecision: We want guarantees about the value that is obtained by a strategy in the Stackelberg game even if this strategy has been synthesized from a weighted game graph with weights that are possibly slightly wrong: small perturbations of the weight should have only limited effect on the value guaranteed by the strategy.
- Sub-optimal responses: We want guarantees about the value that is obtained by a strategy in the Stackelberg game even if the adversary responds with an ϵ-best response instead of a perfectly optimal response (for some $\epsilon > 0$): small deviations from the best-response by the adversary should have only limited effect on the value guaranteed by the strategy.

Formalizing deviations. To formalize modeling imprecision, we introduce the notion of a *perturbed game graph*. Given a game $\mathcal{G}$ with arena $\mathcal{A}_{\mathcal{G}} = (V, E, \langle V_0, V_1\rangle, w_0, w_1)$, and a value $\delta > 0$, we write $\mathcal{G}^{\pm\delta}$ for the set $\mathcal{H}$ of games with arena $\mathcal{A}_{\mathcal{H}} = (V, E, \langle V_0, V_1\rangle, w'_0, w'_1)$ where edge weight functions respect the following constraints:

$$\forall (v_1, v_2) \in E, \forall i \in \{0, 1\}, \quad w'_i(v_1, v_2) \in (w_i(v_1, v_2) + \delta, w_i(v_1, v_2) - \delta).$$

We note that as the underlying game graph (V, E) is not altered, for both players, the set of strategies in $\mathcal{G}$ is identical to the set of strategies in $\mathcal{H}$. Finally, to formalize *sub-optimal responses*, we naturally use the notion of ϵ-best response introduced in the previous section.

Robustness in zero-sum games. In zero-sum games, the worst-case value $\mathsf{Val}_{\mathcal{G}}(\sigma_0)$ is robust against both modeling imprecision and sub-optimal responses of Player 1.

▶ **Proposition 1** (Robustness in zero-sum games). *For all zero-sum mean-payoff games $\mathcal{G}$ with a set V of vertices, for all Player 0 strategies σ_0, and for all vertices $v \in V$ we have that:*

$$\forall \delta, \epsilon > 0 : \forall \mathcal{H} \in \mathcal{G}^{\pm\delta} : \inf_{\sigma_1 \in \mathsf{BR}^{\epsilon}_{1,\mathcal{H}}(\sigma_0)} \underline{\mathsf{MP}}^{\mathcal{H}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)) > \mathsf{Val}_{\mathcal{G}}(\sigma_0)(v) - \delta.$$

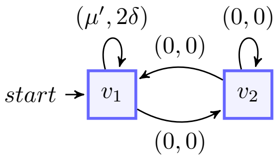

Figure 2 In this game $\mathcal{G}$, all vertices are controlled by Player 1. Here, $\mathsf{ASV}_{\mathcal{G}}(v_1) = \mu'$.

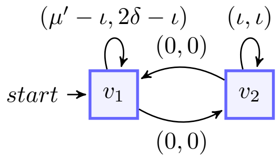

Figure 3 An δ-perturbed game $\mathcal{H}$ of $\mathcal{G}$ in Figure 2. Here, we consider $0 < \iota < \delta$. Here, $\mathsf{ASV}_{\mathcal{H}}(v_1) = \mu' - \iota$.

Fragility in non-zero sum games. On the contrary, the adversarial Stackelberg value $\mathsf{ASV}(\sigma_0)$ is fragile against both modeling imprecision and sub-optimal responses.

▶ **Proposition 2** (Fragility - modeling imprecision). *For all $\mu > 0$, we can construct a nonzero-sum mean-payoff game $\mathcal{G}$ and a Player 0 strategy σ_0, such that there exist $\delta > 0$, a perturbed game $\mathcal{H} \in \mathcal{G}^{\pm\delta}$, and a vertex v in $\mathcal{G}$ with $\mathsf{ASV}_{\mathcal{H}}(\sigma_0)(v) < \mathsf{ASV}_{\mathcal{G}}(\sigma_0)(v) - \mu$.*

▶ **Proposition 3** (Fragility - sub-optimal responses). *For all $\mu > 0$, we can construct a nonzero-sum mean-payoff game $\mathcal{G}$ and a Player 0 strategy σ_0, such that there exist $\epsilon > 0$ and a vertex v in $\mathcal{G}$ with $\inf_{\sigma_1 \in \mathsf{BR}_1^{\epsilon}(\sigma_0)} \underline{\mathsf{MP}}_0^{\mathcal{G}}(\mathsf{Out}_v(\sigma_0, \sigma_1)) < \mathsf{ASV}_{\mathcal{G}}(\sigma_0)(v) - \mu$.*

Note that μ can be arbitrarily large and thus the adversarial Stackelberg value in the model under deviations can be arbitrarily worse than in the original model.

Relation between the two types of deviations. In nonzero-sum mean-payoff games, robustness against modeling imprecision does not imply robustness against sub-optimal responses.

▶ **Lemma 4.** *For all $\mu, \delta, \epsilon > 0$, we can construct a nonzero-sum mean-payoff game $\mathcal{G}$ such that for all Player 0 strategies σ_0 and vertex v in $\mathcal{G}$, we have that:*

$$\forall \mathcal{H} \in \mathcal{G}^{\pm\delta} : \mathsf{ASV}_{\mathcal{H}}(\sigma_0)(v) > \inf_{\sigma_1 \in \mathsf{BR}_{1,\mathcal{G}}^{\epsilon}} \underline{\mathsf{MP}}_0^{\mathcal{G}}(\mathsf{Out}_v(\sigma_0, \sigma_1)) + \mu.$$

Proof. Consider the game $\mathcal{G}$ shown in Figure 2. Here, since all the vertices are controlled by Player 1, the strategy of Player 0 is inconsequential. For every $\delta > 0$, we claim that the best strategy for Player 1 across all perturbed games $\mathcal{H} \in \mathcal{G}^{\pm\delta}$ is to play $v_1 \to v_1$ forever. One such example of a perturbed game is shown in Figure 3. Here, for every $0 < \iota < \delta$, we have that $v_1 \to v_1$ is the only best-response for Player 1. Therefore, we have that $\inf_{\mathcal{H} \in \mathcal{G}^{\pm\delta}} \mathsf{ASV}_{\mathcal{H}}(\sigma_0)(v_1) = \mu' - \delta$, for all $\delta > 0$.

However, if we relax the assumption that Player 1 plays optimally and assume that he plays an ϵ-best response in the game $\mathcal{G}$, we note that Player 1 can play a strategy $(v_1^{k_1+1} v_2^{k_2+1})^{\omega}$, for some $k_1, k_2 \in \mathbb{N}$, such that $\frac{2\delta \cdot k_1}{k_1+k_2+2} > 2\delta - \epsilon$, and Player 0 gets a payoff of $\frac{k_1 \cdot \mu'}{k_1+k_2+2} > \mu'(1 - \frac{\epsilon}{2\delta})$. Thus, we have that $\inf_{\sigma_1 \in \mathsf{BR}_{1,\mathcal{G}}^{\epsilon}} \underline{\mathsf{MP}}_0^{\mathcal{G}}(\mathsf{Out}_v(\sigma_0, \sigma_1)) = \mu'(1 - \frac{\epsilon}{2\delta})$. We note that the choice of μ' is arbitrary, and we can have a μ' such that $\mu' - \delta > \mu'(1 - \frac{\epsilon}{2\delta}) + \mu$, i.e, we choose μ' to be large enough so that $\mu < \mu' \cdot \frac{\epsilon}{2\delta} - \delta$. ◀

On the contrary, robustness against sub-optimal responses implies robustness against modeling imprecision.

▶ **Theorem 5** (Robust strategy in non-zero sum games). *For all non-zero sum mean-payoff games $\mathcal{G}$ with a set V of vertices, for all $\epsilon > 0$, for all vertices $v \in V$, for all strategies σ_0 of Player 0, we have that $\forall \mathcal{H} \in \mathcal{G}^{\pm\epsilon} : \mathsf{ASV}_{\mathcal{H}}(\sigma_0)(v) > \inf_{\sigma_1 \in \mathsf{BR}^{2\epsilon}_{1,\mathcal{G}}} \underline{\mathsf{MP}}^{\mathcal{G}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)) - \epsilon$.*

Proof. Consider a nonzero-sum mean-payoff game $\mathcal{G}$ and a vertex v in $\mathcal{G}$ and a strategy σ_0 of Player 0. We let $\inf_{\sigma_1 \in \mathsf{BR}^{2\epsilon}_{1,\mathcal{G}}} \underline{\mathsf{MP}}^{\mathcal{G}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)) = c$, for some $c \in \mathbb{Q}$. Let the supremum of the payoffs that Player 1 gets when Player 0 plays σ_0 be y, where $y \in \mathbb{Q}$, i.e., $\sup\{\underline{\mathsf{MP}}_1(\rho) \mid \rho \in \mathsf{Out}_v(\mathcal{G}, \sigma_0))\} = y$. For all outcomes ρ which are in Player 1's 2ϵ-best response of σ_0, we have that $\underline{\mathsf{MP}}_1(\rho) > y - 2\epsilon$ and $\underline{\mathsf{MP}}_0(\rho) \geqslant c$.

Now, consider a game $\mathcal{H} \in \mathcal{G}^{\pm\epsilon}$ and a Player 0 strategy σ_0 played in $\mathcal{H}$. We can see that the maximum payoff that Player 1 gets when Player 0 plays σ_0 is bounded by $y + \epsilon$ and $y - \epsilon$, i.e., $y - \epsilon < \sup\{\underline{\mathsf{MP}}_1(\rho) \mid \rho \in \mathsf{Out}_v(\mathcal{H}, \sigma_0))\} < y + \epsilon$. We let this value be denoted by $y_{\mathcal{H}}$. We note that if $\sup_{\rho \in \mathsf{Out}_v(\mathcal{H},\sigma_0)}(\underline{\mathsf{MP}}_1(\rho)) = y_{\mathcal{H}}$, then for the corresponding play $\rho_{\mathcal{H}}$ in the game $\mathcal{G}$, the mean-payoff of Player 1 in $\rho_{\mathcal{H}}$ is $\underline{\mathsf{MP}}_1(\rho_{\mathcal{H}}) > y - 2\epsilon$. Thus, in the game $\mathcal{G}$, we note that $\underline{\mathsf{MP}}_0(\rho_{\mathcal{H}}) \geqslant c$ and for the corresponding play in $\mathcal{H}$, we have $\underline{\mathsf{MP}}_0(\rho_{\mathcal{H}}) > c - \epsilon$. Thus, we have $\mathsf{ASV}_{\mathcal{H}}(\sigma_0)(v) > c - \epsilon = \inf_{\sigma_1 \in \mathsf{BR}^{2\epsilon}_{1,\mathcal{G}}} \underline{\mathsf{MP}}^{\mathcal{G}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)) - \epsilon$. ◀

We note that in the above theorem, we need to consider a strategy that is robust against 2ϵ-best-responses to ensure robustness against ϵ weight perturbations.

ϵ-Adversarial Stackelberg Value. The results above suggest that, in order to obtain some robustness guarantees in nonzero-sum mean-payoff games, we must consider a solution concept that accounts for ϵ-best responses of the adversary. This leads to the following definition: Given an $\epsilon > 0$, we define the adversarial value of Player 0 strategy σ_0 when Player 1 plays ϵ-best-responses as

$$\mathsf{ASV}^{\epsilon}(\sigma_0)(v) = \inf_{\sigma_1 \in \mathsf{BR}^{\epsilon}_1(\sigma_0)} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)) \tag{1}$$

and the ϵ-Adversarial Stackelberg value at vertex v is: $\mathsf{ASV}^{\epsilon}(v) = \sup_{\sigma_0 \in \Sigma_0} \mathsf{ASV}^{\epsilon}(\sigma_0)(v)$, and we note that $\mathsf{ASV}(v) = \sup_{\epsilon > 0} \mathsf{ASV}^{\epsilon}(v)$. We can now state a theorem about combined robustness of ASV^{ϵ}.

▶ **Theorem 6** (Combined robustness of ASV^{ϵ}). *For all nonzero-sum mean-payoff games $\mathcal{G}$ with a set V of vertices, for all $\epsilon > 0$, for all $\delta > 0$, for all $\mathcal{H} \in \mathcal{G}^{\pm\delta}$, for all vertices $v \in V$, and for all strategies σ_0, we have that if $\mathsf{ASV}^{2\delta+\epsilon}_{\mathcal{G}}(\sigma_0)(v) > c$, then for all $\mathcal{H} \in \mathcal{G}^{\pm\delta}$, we have that $\inf_{\sigma_1 \in \mathsf{BR}^{\epsilon}_{\mathcal{H}}(\sigma_0)} \underline{\mathsf{MP}}^{\mathcal{H}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)) > c - \delta$.*

Proof. The proof for Theorem 6 is very similar to the proof of Theorem 5 and involves looking at the set of ϵ-best-responses in the game $\mathcal{H}$ and showing that the corresponding plays lie in the set of $(2\delta + \epsilon)$-best-responses in the game $\mathcal{G}$. This would imply that the corresponding Player 0 mean-payoffs for the ϵ-best-responses of Player 1 in every perturbed game $\mathcal{H} \in \mathcal{G}^{\pm\delta}$ would always be greater than $c - \delta$. Therefore, we can extrapolate that $\mathsf{ASV}^{\epsilon}_{\mathcal{H}}(\sigma_0)(v) > c - \delta$. ◀

In the rest of the paper we study properties of ASV^{ϵ} and solve the following two problems:

- Threshold Problem of ASV^{ϵ}: Given $\mathcal{G}$, $c \in \mathbb{Q}$, an $\epsilon > 0$, and a vertex v, we provide a nondeterministic polynomial time algorithm to decide if $\mathsf{ASV}^{\epsilon}(v) > c$ (see Theorem 8).

- Computation of ASV^ϵ and largest ϵ: Given $\mathcal{G}$, an $\epsilon > 0$, and a vertex v, we provide an exponential time algorithm to compute $\mathsf{ASV}^\epsilon(v)$ (see Theorem 12). We also establish that ASV^ϵ is achievable (see Theorem 16). Then we show, given a fixed threshold c, how to computation of largest ϵ such that $\mathsf{ASV}^\epsilon(v) > c$. Formally, we compute $\sup\{\epsilon > 0 \mid \mathsf{ASV}^\epsilon(v) > c\}$ (See Corollary 15).

4 Threshold problem for the ASV^ϵ

In this section, given $c \in \mathbb{Q}$, and a vertex v in game $\mathcal{G}$, we study the threshold problem which is to determine if $\mathsf{ASV}^\epsilon(v) > c$.

Witnesses for ASV^ϵ. For a game $\mathcal{G}$ and $\epsilon > 0$, we associate with each vertex v in $\mathcal{G}$, a set $\Lambda^\epsilon(v)$ of pairs or real numbers (c, d) such that Player 1 has a strategy that ensures a mean-payoff greater than d - ϵ for himself while restricting the payoff of Player 0 to at most c. Formally, we have:

$$\Lambda^\epsilon(v) = \{(c, d) \in \mathbb{R}^2 \mid v \vDash \ll 1 \gg \underline{\mathsf{MP}}_0 \leqslant c \wedge \underline{\mathsf{MP}}_1 > d - \epsilon\}.$$

A vertex v is $(c, d)^\epsilon$-bad if $(c, d) \in \Lambda^\epsilon(v)$. Let $c' \in \mathbb{R}$. A play π of $\mathcal{G}$ is called a $(c', d)^\epsilon$-witness of $\mathsf{ASV}^\epsilon(v) > c$ if $(\underline{\mathsf{MP}}_0(\pi), \underline{\mathsf{MP}}_1(\pi)) = (c', d)$ where $c' > c$, and π does not contain any $(c, d)^\epsilon$-bad vertex. A play π is called a witness for $\mathsf{ASV}^\epsilon(v) > c$ if it is a $(c', d)^\epsilon$-witness for $\mathsf{ASV}^\epsilon(v) > c$ for some c', d. We now show that polynomial-size witnesses for $\mathsf{ASV}^\epsilon > c$ exist:

▶ **Theorem 7.** *For all mean-payoff games $\mathcal{G}$, for all vertices v in $\mathcal{G}$, for all $\epsilon > 0$, and $c \in \mathbb{Q}$, we have that $\mathsf{ASV}^\epsilon(v) > c$ if and only if there exists a $(c', d)^\epsilon$-witness of $\mathsf{ASV}^\epsilon(v) > c$, where $d \in \mathbb{Q}$. Furthermore, the $(c', d)^\epsilon$-witness can be chosen as a regular witness $\pi = u \cdot v^\omega$, where u and v are finite paths of polynomial size.*

Proof sketch. We consider only the left to right direction here since the other direction of the proof is similar to showing that existence of a witness for $\mathsf{ASV}(v) > c$ implies $\mathsf{ASV}(v) > c$ [1]. We are given that $\mathsf{ASV}^\epsilon(v) > c$. First we show that $\mathsf{ASV}^\epsilon(v) > c$ iff there exists a strategy σ_0 of Player 0 such that $\mathsf{ASV}^\epsilon(\sigma_0)(v) > c$. Thus, there exists a $\delta > 0$, such that $\inf_{\sigma_1 \in \mathsf{BR}_1^\epsilon(\sigma_0)} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)) = c' = c + \delta$ Let $d = \sup_{\sigma_1 \in \mathsf{BR}_1^\epsilon(\sigma_0)} \underline{\mathsf{MP}}_1(\mathsf{Out}_v(\sigma_0, \sigma_1))$. We show that for all $\sigma_1 \in \mathsf{BR}_1^\epsilon(\sigma_0)$, we have that $\mathsf{Out}_v(\sigma_0, \sigma_1)$ does not cross a $(c, d)^\epsilon$-bad vertex. We then consider a sequence $(\sigma_i)_{i \in \mathbb{N}}$ of Player 1 strategies such that $\sigma_i \in \mathsf{BR}_1^\epsilon(\sigma_0)$ for all $i \in \mathbb{N}$, and $\lim_{i \to \infty} \underline{\mathsf{MP}}_1(\mathsf{Out}_v(\sigma_0, \sigma_i)) = d$. Let $\pi_i = \mathsf{Out}_v(\sigma_0, \sigma_i)$. W.l.o.g., we can have that all the plays $\mathsf{Out}_v(\sigma_0, \sigma_i)$ end up in the same SCC, say S.

Now using the fact that $\mathsf{F}_{\min}(\mathsf{CH}(\mathbb{C}(S)))$ is a closed set, and using a result from [8, 5] which states that for every pair of points (x, y) in $\mathsf{F}_{\min}(\mathsf{CH}(\mathbb{C}(S)))$, there exists a play π in the SCC S such that $(\underline{\mathsf{MP}}_0(\pi), \underline{\mathsf{MP}}_1(\pi)) = (x, y)$, we can show that there exists a play $\pi^* \in \mathsf{Out}_v(\sigma_0)$ with $(\underline{\mathsf{MP}}_0(\pi^*), \underline{\mathsf{MP}}_1(\pi^*)) = (c^*, d)$ and $c^* \geqslant c'$. That π^* is a $(c^*, d)^\epsilon$-witness now follows since the vertices appearing in π^* are not $(c, d)^\epsilon$-bad. We have thus shown that if $\mathsf{ASV}^\epsilon(v) > c$, then there exists a $(c^*, d)^\epsilon$-witness. Finally, by using the Carathéodory baricenter theorem, we show that two simple cycles, and three acyclic finite plays suffice to construct a regular witness. ◀

The following statement can be obtained by exploiting the existence of finite regular witnesses of polynomial size proved above.

▶ **Theorem 8.** *For all mean-payoff games $\mathcal{G}$, for all vertices v in $\mathcal{G}$, for all $\epsilon > 0$, and for all $c \in \mathbb{Q}$, it can be decided in nondeterministic polynomial time if $\mathsf{ASV}^\epsilon(v) > c$, and a pseudopolynomial memory strategy of Player 0 suffices for this threshold. Furthermore, this decision problem is at least as hard as solving zero-sum mean-payoff games.*

As a corollary of Theorem 8, we can deduce that the ϵ-adversarial Stackelberg value achievable using finite memory strategies which defined as :

$$\mathsf{ASV}^\epsilon_{\mathsf{FM}}(v) = \sup_{\sigma_0 \in \Sigma_0^{\mathsf{FM}}} \inf_{\sigma_1 \in \mathsf{BR}_1^\epsilon(\sigma_0)} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1))$$

where Σ_0^{FM} refers to the set of all finite memory strategies of Player 0, is equal to ASV^ϵ:

▶ **Corollary 9.** *For all games $\mathcal{G}$, for all vertices v in $\mathcal{G}$, and for all $\epsilon > 0$, we have that $\mathsf{ASV}^\epsilon_{\mathsf{FM}}(v) = \mathsf{ASV}^\epsilon(v)$.*

This corollary is important from a practical point of view as it implies that the ASV^ϵ value can be approached to any precision with a finite memory strategy. Nevertheless, we show in Theorem 16 that infinite memory is necessary to achieve the exact ASV^ϵ.

Memoryless strategies of Player 0. We now establish that the threshold problem is NP-complete when Player 0 is restricted to play *memoryless* strategies. First we define

$$\mathsf{ASV}^\epsilon_{\mathsf{ML}}(v) = \sup_{\sigma_0 \in \Sigma_0^{\mathsf{ML}}} \inf_{\sigma_1 \in \mathsf{BR}_1^\epsilon(\sigma_0)} \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1))$$

where Σ_0^{ML} is the set of all memoryless strategies of Player 0.

▶ **Theorem 10.** *For all mean-payoff games $\mathcal{G}$, for all vertices v in $\mathcal{G}$, for all $\epsilon > 0$, and for all rationals c, the problem of deciding if $\mathsf{ASV}^\epsilon_{\mathsf{ML}}(v) > c$ is* NP-Complete.

The proof of hardness is a reduction from the partition problem while easiness is straightforwardly obtained by techniques used in the proof of Theorem 8.

5 Computation of the ASV^ϵ and the largest ϵ possible

Here, we express the ASV^ϵ as a formula in the theory of reals by adapting a method provided in [8] for ASV. We then provide a new EXPTime algorithm to compute the ASV^ϵ based on construction of linear programs (LPs) which in turn is applicable to ASV as well.

Extended mean-payoff game. Given a mean-payoff game $\mathcal{G}$ with a set V of vertices in its arena, we construct an extended mean-payoff game $\mathcal{G}^{\mathsf{ext}}$, whose arena consists of vertices $V^{\mathsf{ext}} = V \times 2^V$. With a history h in $\mathcal{G}$, we associate a vertex in $\mathcal{G}^{\mathsf{ext}}$ which is a pair (v, P), where $v = \mathsf{last}(h)$ and P is the set of the vertices traversed along h. The set E^{ext} of edges, and the weight functions w_i^{ext} for $i \in \{0, 1\}$ are defined as $E^{\mathsf{ext}} = \{((v, P), (v', P')) \mid (v, v') \in E, P' = P \cup \{v'\}\}$ and $w_i^{\mathsf{ext}}((v, P), (v', P')) = w_i(v, v')$ respectively. There exists a bijection between the plays π in $\mathcal{G}$ and the plays π^{ext} in $\mathcal{G}^{\mathsf{ext}}$. Note that the second component of the vertices of the play π^{ext} stabilises into a set of vertices of $\mathcal{G}$ which we denote by $V^*(\pi^{\mathsf{ext}})$.

We characterize $\mathsf{ASV}^\epsilon(v)$ with the notion of witness introduced earlier and the decomposition of $\mathcal{G}^{\mathsf{ext}}$ into SCCs. For a vertex v in V, let $\mathsf{SCC}^{\mathsf{ext}}(v)$ be the set of strongly-connected components in $\mathcal{G}^{\mathsf{ext}}$ which are reachable from $(v, \{v\})$.

▶ **Lemma 11.** *For all mean-payoff games $\mathcal{G}$ and for all vertices v in $\mathcal{G}$, we have*

$$\mathsf{ASV}^{\epsilon}(v) = \max_{S \in \mathsf{SCC}^{\mathsf{ext}}(v)} \sup\{c \in \mathbb{R} \mid \exists \pi^{\mathsf{ext}} : \pi^{\mathsf{ext}} \text{ is a witness for } \mathsf{ASV}^{\epsilon}(v, \{v\}) > c \text{ and } V^*(\pi^{\mathsf{ext}}) = S\}.$$

By definition of $\mathcal{G}^{\mathsf{ext}}$, for every SCC S of $\mathcal{G}^{\mathsf{ext}}$, there exists a set $V^*(S)$ of vertices of $\mathcal{G}$ such that every vertex of S is of the form $(v', V^*(S))$, where v' is a vertex in $\mathcal{G}$. Now, we define $\Lambda^{\mathsf{ext}}_S = \bigcup_{v \in V^*(S)} \Lambda^{\epsilon}(v)$ as the set of (c, d) such that Player 1 can ensure $v \models \ll 1 \gg \underline{\mathsf{MP}}_0 \leqslant c \wedge \underline{\mathsf{MP}}_1 > d - \epsilon$ from some vertex $v \in S$. The set Λ^{ext}_S can be represented by a formula $\Psi^{\epsilon}_S(x, y)$ in the first order theory of reals with addition, $\langle \mathbb{R}, +, < \rangle$, with two free variables. We refer the reader to [1] for a formal statement and a proof of this. We can now state the following theorem about the computability of $\mathsf{ASV}^{\epsilon}(v)$:

▶ **Theorem 12.** *For all mean-payoff games $\mathcal{G}$, for all vertices v in $\mathcal{G}$ and for all $\epsilon > 0$, the $\mathsf{ASV}^{\epsilon}(v)$ can be effectively expressed by a formula in $\langle \mathbb{R}, +, < \rangle$, and can be computed from this formula. Furthermore, the formula can be effectively transformed into exponentially many linear programs which establish membership in* $\mathsf{EXPTime}$.

Proof sketch. Using Lemma 11, we have that for every S in $\mathsf{SCC}^{\mathsf{ext}}(v)$, a value of c such that $\mathsf{ASV}^{\epsilon}(v) > c$ can be encoded by the formula $\rho^S_v(c) \equiv \exists x, y \cdot x > c \wedge \Phi_S(x, y) \wedge \neg \Psi^{\epsilon}_S(c, y)$ where the formula $\Phi_S(x, y)$ expresses the symbolic encoding of the pair of values (x, y) which represents the mean-payoff values of some play in S, and the formula $\neg \Psi^{\epsilon}_S(c, y)$ expresses that the play does not cross a $(c, y)^{\epsilon}$-bad vertex. We then construct a formula $\rho^S_{\max,v}(z)$ which is satisfied by a value that is the supremum over the set of values c such that c satisfies the formula ρ^S_v. From the formula $\rho^S_{\max,v}$, we can compute the $\mathsf{ASV}^{\epsilon}(v)$ by quantifier elimination, and by finding the maximum across all the SCCs S in $\mathsf{SCC}^{\mathsf{ext}}(v)$.

For the $\mathsf{EXPTime}$ algorithm, first note that for each SCC S in $\mathcal{G}^{\mathsf{ext}}$, the set satisfying the formula $\Phi_S(x, y)$, which is the symbolic encoding of $\mathsf{F}_{\min}(\mathsf{CH}(\mathbb{C}(S)))$, can be expressed as a set of *exponentially* many inequalities [5]. Also the formula $\Psi^{\epsilon}_S(x, y)$ can be expressed using exponentially many LPs. We refer the interested reader to [1] for more details. It follows that the formula $\rho^S_v(c)$ can be expressed with exponentially many LPs. In each LP, the objective is to maximize c. The algorithm runs in $\mathsf{EXPTime}$ since there can be exponentially many SCCs. ◀

▶ **Example 13.** We illustrate the computation of ASV^{ϵ} with an example. Consider the mean-payoff game $\mathcal{G}$ depicted in Figure 4 and its extension $\mathcal{G}^{\mathsf{ext}}$ as shown in Figure 5.

Note that in $\mathcal{G}^{\mathsf{ext}}$ there exist three SCCs which are $S_1 = \{v'_{0^2}, v'_1\}, S_2 = \{v'_{2^1}\}$, and $S_3 = \{v'_{2^2}\}$. The SCCs S_2 and S_3 are similar, and thus $\rho^{S_2}_{\max,v_0}(z)$ and $\rho^{S_3}_{\max,v_0}(z)$ would be equivalent. We start with SCC S_1 that contains two cycles $v'_1 \to v'_1$ and $v'_{0^2} \to v'_1, v'_1 \to v'_{0^2}$, and SCC S_2 contains one cycle $v'_{2^1} \to v'_{2^1}$. Since S_3 is similar to S_2, we consider only S_2 in our example. Thus, the set $\mathsf{F}_{\min}(\mathsf{CH}(\mathbb{C}(S_1)))$ is represented by the Cartesian points within the triangle represented by $(0, 2), (1, 1)$ and $(0, 1)$ [5] and $\mathsf{F}_{\min}(\mathsf{CH}(\mathbb{C}(S_2))) = \{(0, 1)\}$. Thus, we get that $\Phi_{S_1}(x, y) \equiv (x \geqslant 0 \wedge x \leqslant 1) \wedge (y \geqslant 1 \wedge y \leqslant 2) \wedge (x + y) \leqslant 2$ and $\Phi_{S_2}(x, y) \equiv x = 0 \wedge y = 1$. Now, we calculate $\Lambda^{\epsilon}(v'_{0^2})$, $\Lambda^{\epsilon}(v'_{2^1})$ and $\Lambda^{\epsilon}(v'_1)$ for some value of ϵ less tan 1. We note that the vertex v'_1 is not $(0, 2 + \epsilon - \delta)^{\epsilon}$-bad, for all $0 < \delta < 1$, as Player 0 can always choose the edge (v'_1, v'_{0^2}) from v'_1, thus giving Player 1 a mean-payoff of 1. Additionally, the vertex v'_{0^2} is both $(0, 1 + \epsilon - \delta)^{\epsilon}$-bad, for all $\delta > 0$, since Player 1 can choose the edge (v'_{0^2}, v'_{2^2}) from v'_{0^2}, and

[5] Note that the coordinate $(0, 1)$ is obtained as the pointwise minimum over the two coordinates separately.

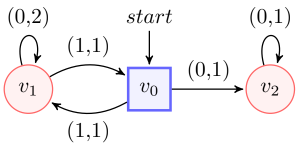

Figure 4 Example to calculate $\mathsf{ASV}^{\epsilon}(v)$.

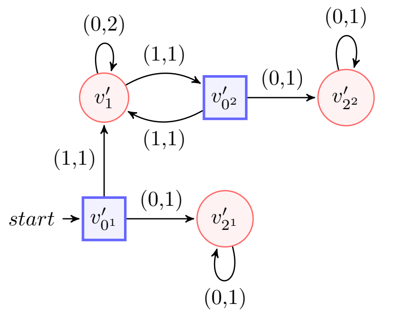

Figure 5 Extended Mean-Payoff Game where $v'_{0^1} = (v_0, \{v_0\})$, $v'_{0^2} = (v_0, \{v_0, v_1\})$, $v'_1 = (v_1, \{v_0, v_1\})$, $v'_{2^1} = (v_1, \{v_0, v_2\})$, and $v'_{2^2} = (v_2, \{v_0, v_1, v_2\})$.

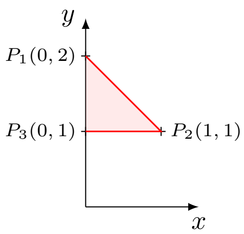

Figure 6 The red triangle represents the set of points in Φ_{S_1}.

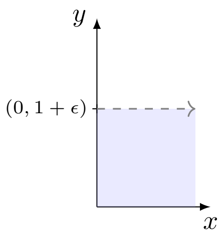

Figure 7 The blue region under and excluding the line $y = (1-\epsilon)$ represents the set of points in $\Psi^{\epsilon}_{S_1}$ and $\Psi^{\epsilon}_{S_2}$.

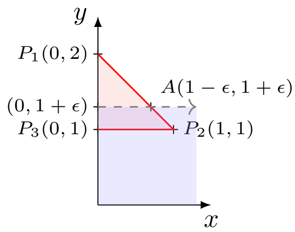

Figure 8 The formula $\rho^{S_1}(c)$ is represented by the points in Φ_{S_1} and not in $\Psi^{\epsilon}_{S_1}$, i.e., the points in the triangle which are not strictly below the line $y = (1-\epsilon)$. Here, the max c value is represented by point A.

$(1, 1+\epsilon-\delta)^{\epsilon}$-bad, for all $\delta > 0$, since Player 1 can choose the edge (v'_{0^2}, v'_1) from v'_{0^2}. Thus, we get that $\Lambda^{\epsilon}(v'_1) = \Lambda^{\epsilon}(v'_{0^2}) = \{(c,d) \mid (c \geqslant 1 \wedge d < 1+\epsilon)\} \bigcup \{(c,d) \mid (c \geqslant 0 \wedge d < 1+\epsilon)\}$ which is the same as $\{(c,d) \mid (c \geqslant 0 \wedge d < 1+\epsilon)\}$, and $\Lambda^{\epsilon}(v'_{2^1}) = \{(c,d) \mid (c \geqslant 0 \wedge d < 1+\epsilon)\}$. Therefore, we have that $\Lambda^{\mathsf{ext}}_{S_1} = \Lambda^{\mathsf{ext}}_{S_2} = \{(c,d) \mid (c \geqslant 0 \wedge d < 1+\epsilon)\}$. Hence, we get that $\Psi^{\epsilon}_{S_1}(x,y) = \Psi^{\epsilon}_{S_2}(x,y) \equiv (x \geqslant 0 \wedge y < 1+\epsilon)$ as shown in Figure 7. From Figure 8, the formula $\rho^{S_1}(c)$ holds true for values of c less than $(1-\epsilon)$ and the formula $\rho^{S_2}(c)$ holds for values of c less than 0. Hence, by assigning $(1-\epsilon)$ to x, and $(1+\epsilon)$ to y, we get that $\rho^{S_1}_{\max,v_0}(z)$ holds true for $z = (1-\epsilon)$. Additionally, by assigning 0 to x, and 1 to y, we get that $\rho^{S_2}_{\max,v_0}(z)$ holds true for $z = 0$. It follows that $\mathsf{ASV}^{\epsilon}(v_0) = 1-\epsilon$ for $\epsilon < 1$ as it is the maximum of the values over all the SCCs. ◀

We now illustrate the LP formulation for $\rho^S_v(c)$ for each SCC S with the following example, and provide details for computing $\mathsf{ASV}^{\epsilon}(v_0)$.

▶ **Example 14.** We previously showed that the $\mathsf{ASV}^{\epsilon}(v_0)$ can be computed by quantifier elimination of a formula in the theory of reals with addition. Now, we compute the $\mathsf{ASV}^{\epsilon}(v_0)$ by solving a set of linear programs for every SCC in $\mathcal{G}^{\mathsf{ext}}$. We recall that there are three SCCs S_1, S_2 and S_3 in $\mathcal{G}^{\mathsf{ext}}$. From a result in [5], we have that $\mathsf{F}_{\min}(\mathsf{CH}(\mathbb{C}(S_i)))$ for $i \in \{1,2,3\}$ can be defined using a set of linear inequalities. Now recall that $\mathsf{F}_{\min}(\mathsf{CH}(\mathbb{C}(S_2)) = \mathsf{F}_{\min}(\mathsf{CH}(\mathbb{C}(S_3))) = \{(0,1)\}$, and $\mathsf{F}_{\min}(\mathsf{CH}(\mathbb{C}(S_1)))$ is represented by the set of points enclosed by the triangle formed by connecting the points $(0,1), (1,1)$ and $(0,2)$ as shown in Figure 6, and $\Lambda^{\epsilon}(v'_{0^2}) = \Lambda^{\epsilon}(v'_{2^1}) = \Lambda^{\epsilon}(v'_{2^2}) = \Lambda^{\epsilon}(v'_1) = \{(c,y) \mid c \geqslant 0 \wedge y < 1+\epsilon\}$. Now, we consider the

SCC S_1, and the formula $\neg\Psi^{\epsilon}_{S_1}$. We start this by finding the complement of $\Lambda^{\epsilon}(v'_{0^2})$ and $\Lambda^{\epsilon}(v'_1)$, that is, $\overline{\Lambda}^{\epsilon}(v'_{0^2}) = \overline{\Lambda}^{\epsilon}(v'_1) = \mathbb{R}\times\mathbb{R} - \Lambda^{\epsilon}(v'_{0^2}) = \mathbb{R}\times\mathbb{R} - \Lambda^{\epsilon}(v'_1) = \{(c,y) \mid c < 0 \vee y \geqslant 1+\epsilon\}$. Now, we get that $\neg\Psi^{\epsilon}_{S_1} = \overline{\Lambda}^{\epsilon}(v'_{0^2}) \bigcap \overline{\Lambda}^{\epsilon}(v'_1) = \{(c,y) \mid c < 0 \vee y \geqslant 1+\epsilon\}$.

Similarly for the SCC S_2 and SCC S_3, we calculate the complement of $\Lambda^{\epsilon}(v'_{2^1})$ and $\Lambda^{\epsilon}(v'_{2^2})$, that is, $\overline{\Lambda}^{\epsilon}(v'_{2^2}) = \overline{\Lambda}^{\epsilon}(v'_{2^1}) = \mathbb{R}\times\mathbb{R} - \Lambda^{\epsilon}(v'_{2^1}) = \mathbb{R}\times\mathbb{R} - \Lambda^{\epsilon}(v'_{2^2}) = \{(c,y) \mid c < 0 \vee y \geqslant 1+\epsilon\}$ and obtain $\neg\Psi^{\epsilon}_{S_2} = \overline{\Lambda}^{\epsilon}(v'_{2^1}) = \{(c,y) \mid c < 0 \vee y \geqslant 1+\epsilon\}$ and $\neg\Psi^{\epsilon}_{S_3} = \overline{\Lambda}^{\epsilon}(v'_{2^2}) = \{(c,y) \mid c < 0 \vee y \geqslant 1+\epsilon\}$. Note that the formulaes $\Phi_{S_2}(x,y)$ and $\Phi_{S_3}(x,y)$ are represented by the set of linear inequations $x = 0 \wedge y = 1$ and the formula $\Phi_{S_1}(x,y)$ is represented by the set of linear inequations $y \geqslant 1 \wedge y \leqslant 2 \wedge x \leqslant 1 \wedge (x+y) \leqslant 2$. Now the formula $\rho^{S_1}_{v_0}(c)$ can be expressed using a set of linear equations and inequalities as follows: $x > c \wedge y \geqslant 1 \wedge y \leqslant 2 \wedge x \leqslant 1 \wedge (x+y) \leqslant 2 \wedge (c < 0 \vee y \geqslant 1+\epsilon)$ and the formula $\rho^{S_2}_{v_0}(c)$ can be expressed using a set of linear equations and inequalities as follows: $x > c \wedge x = 0 \wedge y = 1 \wedge (c < 0 \vee y \geqslant 1+\epsilon)$. We maximise the value of c in the formula $\rho^{S_1}_{v_0}(c)$ to get the following two linear programs: *maximise* c *in* $(x > c \wedge y \geqslant 1 \wedge y \leqslant 2 \wedge x \leqslant 1 \wedge (x+y) \leqslant 2 \wedge c < 0)$ which gives a solution $\{0\}$ and *maximise* c *in* $(x > c \wedge y \geqslant 1 \wedge y \leqslant 2 \wedge x \leqslant 1 \wedge (x+y) \leqslant 2 \wedge y \geqslant (1+\epsilon))$ which gives us a solution $\{(1-\epsilon)\}$. Similarly, maximising c in the formulaes $\rho^{S_2}_{v_0}(c)$ and $\rho^{S_3}_{v_0}(c)$ would give us the following two linear programs: *maximise* c *in* $(x > c \wedge x = 0 \wedge y = 1 \wedge c < 0)$ which gives a solution $\{0\}$ and *maximise* c *in* $(x > c \wedge x = 0 \wedge y = 1 \wedge y \geqslant (1+\epsilon))$ which gives us a solution $\{0\}$. Thus, we conclude that $\mathsf{ASV}^{\epsilon}(v_0) = 1-\epsilon$ which is the maximum value amongst all the SCCs. Note that in an LP, the strict inequalities are replaced with non-strict inequalities, and computing the supremum in the objective function is replaced by maximizing the objective function.

Again, for every SCC S and for every LP corresponding to that of S, we fix a value of c and change the objective function to *maximise* ϵ from *maximise* c in order to obtain the maximum value of ϵ that allows $\mathsf{ASV}^{\epsilon}(v_0) > c$. For example, consider the LP $(x > c \wedge y \geqslant 1 \wedge y \leqslant 2 \wedge x \leqslant 1 \wedge (x+y) \leqslant 2 \wedge y \geqslant (1+\epsilon))$ in SCC S_1 and fix a value of c, and then maximize the value of ϵ. Doing this over all linear programs in an SCC, and over all SCCs, reachable from v_0 for a fixed c gives us the supremum value of ϵ such that we have $\mathsf{ASV}^{\epsilon}(v_0) > c$. ◀

On the other hand, we note that in every SCC S, the value c is a function of ϵ, for illustration, in the example above, $\rho^{S_1}(c)$ holds true for values of c less than $1-\epsilon$. Thus if we fix a value of c, we can find the supremum over ϵ which allows $\mathsf{ASV}^{\epsilon}(v) > c$ in S. Again, taking the maximum over all SCCs reachable from $(v, \{v\})$ gives us the largest ϵ possible so that we have $\mathsf{ASV}^{\epsilon}(v) > c$. We state the following corollary.

▶ **Corollary 15.** *For all mean-payoff games $\mathcal{G}$, for all vertices v in $\mathcal{G}$, and for all $c \in Q$, we can compute in* EXPTime *the maximum possible value of ϵ such that $\mathsf{ASV}^{\epsilon}(v) > c$.*

6 Additional Properties of ASV^{ϵ}

In this section, we first show that the ASV^{ϵ} is *achievable*, i.e., there exists a Player 0 strategy that achieves the ASV^{ϵ}. Then we study the memory requirement in strategies of Player 0 for achieving the ASV^{ϵ}, as well as the memory requirement by Player 1 for playing the ϵ-best-responses.

Achievability of the ASV^{ϵ}. We formally define achievability as follows. Given $\epsilon > 0$, we say that $\mathsf{ASV}^{\epsilon}(v) = c$ is achievable from a vertex v, if there exists a strategy σ_0 for Player 0 such that $\forall \sigma_1 \in \mathsf{BR}^{\epsilon}_1(\sigma_0) : \underline{\mathsf{MP}}_0(\mathsf{Out}_v(\sigma_0, \sigma_1)) \geqslant c$. We note that this result is in contrast to the case for ASV as shown in [8].

▶ **Theorem 16.** *For all mean-payoff games $\mathcal{G}$, for all vertices v in $\mathcal{G}$, and for all $\epsilon > 0$, we have that the $\mathsf{ASV}^{\epsilon}(v)$ is achievable.*

The rest of this section is devoted to proving Theorem 16. We start by defining the notion of a witness for $\mathsf{ASV}^{\epsilon}(\sigma_0)(v)$ for a strategy σ_0 of Player 0.

Witness for $\mathsf{ASV}^{\epsilon}(\sigma_0)(v)$. Given a mean-payoff game $\mathcal{G}$, a vertex v in $\mathcal{G}$, and an $\epsilon > 0$, we say that a play π is a witness for $\mathsf{ASV}^{\epsilon}(\sigma_0)(v) > c$ for a strategy σ_0 of Player 0 if (i) $\pi \in \mathsf{Out}_v(\sigma_0)$, and (ii) π is a witness for $\mathsf{ASV}^{\epsilon}(v) > c$ when Player 0 uses strategy σ_0 where the strategy σ_0 is defined as follows:

1. σ_0 follows π if Player 1 does not deviate from π.
2. If Player 1 deviates π, then for each vertex $v \in \pi$, we have that σ_0 consists of a memoryless strategy that establishes $v \nvDash \ll 1 \gg \underline{\mathsf{MP}}_0 \leqslant c \wedge \underline{\mathsf{MP}}_1 > d - \epsilon$, where $d = \underline{\mathsf{MP}}_1(\pi)$. The existence of such a memoryless strategy of Player 0 has been established in Section 4.

Assume that the $\mathsf{ASV}^{\epsilon}(v)$ cannot be achieved by a finite memory strategy. We show that for such cases, it can indeed be achieved by an infinite memory strategy.

Let $\mathsf{ASV}^{\epsilon}(v) = c$. For every $c' < c$, from Theorem 8, there exists a finite memory strategy σ_0 such that $\mathsf{ASV}^{\epsilon}(\sigma_0)(v) > c'$, and recall from Theorem 7 that there exists a corresponding regular witness. First we state the following proposition.

▶ **Proposition 17.** *There exists a sequence of increasing real numbers, $c_1 < c_2 < c_3 < \ldots < c$, such that the sequence converges to c, and a set of finite memory strategies $\sigma_0^1, \sigma_0^2, \sigma_0^3, \ldots$ of Player 0 such that for each c_i, we have $\mathsf{ASV}^{\epsilon}(\sigma_0^i)(v) > c_i$, and there exists a play π^i that is a witness for $\mathsf{ASV}^{\epsilon}(\sigma_0^i)(v) > c_i$, where $\pi^i = \pi_1(l_1^{\alpha \cdot k_i} \cdot \pi_2 \cdot l_2^{\beta \cdot k_i} \cdot \pi_3)^{\omega}$, and π_1, π_2 and π_3 are simple finite plays, and l_1, l_2 are simple cycles in the arena of the game $\mathcal{G}$.*

These witnesses or plays in the sequence are regular, and they differ from each other only in the value of k_i that they use.

To show that $\lim_{i \to \infty} \mathsf{ASV}^{\epsilon}(\sigma_0^i)(v) = c$, we construct a play π^* that starts from v, follows π^1 until the mean-payoff of Player 0 over the prefix becomes greater than c_1. Then for $i \in \{2, 3, \ldots\}$, starting from $\mathsf{first}(l_1)$, it follows π^i, excluding the initial simple finite play π_1, until the mean-payoff of the prefix of π^i becomes greater than c_i. Then the play π^* follows the prefix of the play π^{i+1}, excluding the initial finite play π_1, and so on. Clearly, we have that $\underline{\mathsf{MP}}_1(\pi^*) = c$. We let $\underline{\mathsf{MP}}_1(\pi^*) = d = \alpha \cdot \mathsf{MP}_1(l_1) + \beta \cdot \mathsf{MP}_1(l_2)$.

For the sequence of plays $(\pi^i)_{i \in \mathbb{N}^+}$ which are witnesses for $(\mathsf{ASV}^{\epsilon}(\sigma_0^i)(v) > c_i)_{i \in \mathbb{N}^+}$ for the strategies $(\sigma_0^i)_{i \in \mathbb{N}^+}$, we let $\underline{\mathsf{MP}}_1(\pi_i) = d_i$. We state the following proposition.

▶ **Proposition 18.** *The sequence $(d_i)_{i \in \mathbb{N}^+}$ is monotonic, and it converges to d in the limit.*

The above two propositions establish the existence of an infinite sequence of regular witnesses $\mathsf{ASV}^{\epsilon}(\sigma_0^i)(v) > c_i$ for a sequence of increasing numbers $c_1 < c_2 < \ldots < c$, such that the mean-payoffs of the witnesses are monotonic and at the limit, the mean-payoffs of the witnesses converge to c and d for Player 0 and Player 1 respectively. These observations show the existence of a witness π^* which gives Player 0 a mean-payoff value at least c and Player 1 a mean-payoff value equal to d. Assuming that Player 0 has a corresponding strategy σ_0, we show that Player 1 does not have an ϵ-best response to σ_0 that gives Player 0 a payoff less than c. Now, we have the ingredients to prove Theorem 16.

Proof sketch of Theorem 16. We consider a sequence of increasing numbers $c_1 < c_2 < c_3 < \ldots < c$ such that for every $i \in \mathbb{N}^+$, by Theorem 8, we consider a finite memory strategy σ_0^i of Player 0 that ensures $\mathsf{ASV}^{\epsilon}(\sigma_0^i)(v) > c_i$.

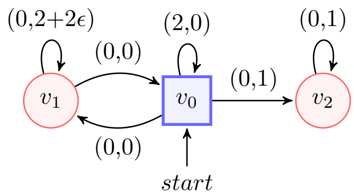

Figure 9 Finite memory strategy of Player 0 may not achieve $\mathsf{ASV}^\epsilon(v_0)$. Also, no finite memory ϵ-best response exists for Player 1 for the strategy σ_0 of Player 0.

If the ASV^ϵ is not achievable, then there exists a strategy of Player 1 to enforce some play π' such that $\underline{\mathsf{MP}}_0(\pi') = c' < c$ and $\underline{\mathsf{MP}}_1(\pi') = d' > d - \epsilon$. Now, we use the monotonicity of the sequence $(d_i)_{i \in \mathbb{N}^+}$ established in Proposition 18 to show a contradiction. Since the sequence $(d_i)_{i \in \mathbb{N}^+}$ is monotonic, there can be two cases:

1. The sequence $(d_i)_{i \in \mathbb{N}^+}$ is monotonically non-decreasing.
2. The sequence $(d_i)_{i \in \mathbb{N}^+}$ is monotonically decreasing.

For each of these cases, we reach a contradiction if we assume that $\mathsf{ASV}^\epsilon(v)$ is not achievable, i.e. Player 1 deviates from π^* to enforce the play π' where $\underline{\mathsf{MP}}_0(\pi') = c' < c$ and $\underline{\mathsf{MP}}_1(\pi') = d'$. ◀

Memory requirements of the players' strategies. First we show that there exists a mean-payoff game $\mathcal{G}$ in which Player 0 needs an infinite memory strategy to achieve the ASV^ϵ.

▶ **Theorem 19.** *There exist a mean-payoff game $\mathcal{G}$, a vertex v in $\mathcal{G}$, and an $\epsilon > 0$ such that Player 0 needs an infinite memory strategy to achieve the $\mathsf{ASV}^\epsilon(v)$.*

Proof sketch. Consider the example in Figure 9. We show that in this example the $\mathsf{ASV}^\epsilon(v_0) = 1$, and that this value can only be achieved using an infinite memory strategy. Assume a strategy σ_0 for Player 0 such that the game is played in rounds. In round k: (i) if Player 1 plays $v_0 \to v_0$ repeatedly at least k times before playing $v_0 \to v_1$, then from v_1, play $v_1 \to v_1$ repeatedly k times and then play $v_1 \to v_0$ and move to round $k+1$; (ii) else, if Player 1 plays $v_0 \to v_0$ less than k times before playing $v_0 \to v_1$, then from v_1 , play $v_1 \to v_0$. Note that σ_0 is an infinite memory strategy. The best-response for Player 1 to strategy σ_0 would be to choose k sequentially as $k = 1, 2, 3, \ldots$, to get a play $\pi = ((v_0)^i (v_1)^i)_{i \in \mathbb{N}}$. We have that $\underline{\mathsf{MP}}_1(\pi) = 1 + \epsilon$ and $\underline{\mathsf{MP}}_0(\pi) = 1$. Player 1 can only sacrifice an amount that is less than ϵ to minimize the mean-payoff of Player 0, and thus he would not play $v_0 \to v_2$. We can show that $\mathsf{ASV}^\epsilon(\sigma_0)(v_0) = \mathsf{ASV}^\epsilon(v_0)$, and that no finite memory strategy of Player 0 can achieve an $\mathsf{ASV}^\epsilon(v_0)$ of 1. ◀

There also exist mean-payoff games in which a finite memory (but not memoryless) strategy for Player 0 can achieve the ASV^ϵ.

Further, we show that there exist games such that for a strategy σ_0 of Player 0, and an $\epsilon > 0$, there does not exist any finite memory best-response of Player 1 to the strategy σ_0.

▶ **Theorem 20.** *There exist a mean-payoff game $\mathcal{G}$, an $\epsilon > 0$, and a Player 0 strategy σ_0 in $\mathcal{G}$ such that every Player 1 strategy $\sigma_1 \in \mathsf{BR}_1^\epsilon(\sigma_0)$ is an infinite memory strategy.*

References

1 Mrudula Balachander, Shibashis Guha, and Jean-François Raskin. Fragility and robustness in mean-payoff adversarial stackelberg games. *CoRR*, abs/2007.07209, 2020. `arXiv:2007.07209`.

2 Roderick Bloem, Krishnendu Chatterjee, Karin Greimel, Thomas A. Henzinger, Georg Hofferek, Barbara Jobstmann, Bettina Könighofer, and Robert Könighofer. Synthesizing robust systems. *Acta Informatica*, 51(3-4):193–220, 2014.

3 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*, pages 3–23, 2016.

4 Véronique Bruyère, Noémie Meunier, and Jean-François Raskin. Secure equilibria in weighted games. In *Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), CSL-LICS '14, Vienna, Austria, July 14–18, 2014*, pages 26:1–26:26. ACM, 2014.

5 Krishnendu Chatterjee, Laurent Doyen, Herbert Edelsbrunner, Thomas A. Henzinger, and Philippe Rannou. Mean-payoff automaton expressions. In *CONCUR 2010 – Concurrency Theory, 21th International Conference, Paris, France, August 31-September 3, 2010. Proceedings*, pages 269–283, 2010.

6 Krishnendu Chatterjee, Thomas A. Henzinger, and Marcin Jurdzinski. Games with secure equilibria. *Theor. Comput. Sci.*, 365(1-2):67–82, 2006.

7 Rodica Condurache, Emmanuel Filiot, Raffaella Gentilini, and Jean-François Raskin. The complexity of rational synthesis. In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, pages 121:1–121:15, 2016.

8 Emmanuel Filiot, Raffaella Gentilini, and Jean-François Raskin. The adversarial stackelberg value in quantitative games. In *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8-11, 2020, Saarbrücken, Germany (Virtual Conference)*, pages 127:1–127:18, 2020.

9 Dana Fisman, Orna Kupferman, and Yoad Lustig. Rational synthesis. In *Tools and Algorithms for the Construction and Analysis of Systems, 16th International Conference, TACAS 2010, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2010, Paphos, Cyprus, March 20-28, 2010. Proceedings*, pages 190–204, 2010.

10 Anshul Gupta and Sven Schewe. Quantitative verification in rational environments. In *21st International Symposium on Temporal Representation and Reasoning, TIME 2014, Verona, Italy, September 8-10, 2014*, pages 123–131, 2014.

11 Anshul Gupta and Sven Schewe. Buying optimal payoffs in bi-matrix games. *Games*, 9(3):40, 2018.

12 Anshul Gupta, Sven Schewe, Ashutosh Trivedi, Maram Sai Krishna Deepak, and Bharath Kumar Padarthi. Incentive stackelberg mean-payoff games. In *Software Engineering and Formal Methods – 14th International Conference, SEFM 2016, Held as Part of STAF 2016, Vienna, Austria, July 4-8, 2016, Proceedings*, pages 304–320, 2016.

13 Anshul Gupta, Sven Schewe, and Dominik Wojtczak. Making the best of limited memory in multi-player discounted sum games. In *Proceedings Sixth International Symposium on Games, Automata, Logics and Formal Verification, GandALF 2015, Genoa, Italy, 21-22nd September 2015*, pages 16–30, 2015.

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

15 Amir Pnueli and Roni Rosner. On the synthesis of a reactive module. In *Conference Record of the Sixteenth Annual ACM Symposium on Principles of Programming Languages, Austin, Texas, USA, January 11-13, 1989*, pages 179–190, 1989.

16 Heinrich Freiherr von Stackelberg. *Marktform und Gleichgewicht.* Wien und Berlin, J. Springer, 1934.