On Relevant Equilibria in Reachability Games*

Thomas Brihaye[a], Véronique Bruyère[a], Aline Goeminne[a,b,*], Nathan Thomasset[c]

[a] *Université de Mons (UMONS), Belgium*
[b] *Université libre de Bruxelles (ULB), Belgium*
[c] *ENS Paris-Saclay, Université Paris-Saclay, France*

Abstract

We study multiplayer reachability games played on a finite directed graph equipped with target sets, one for each player. In those reachability games, it is known that there always exists a Nash equilibrium. But sometimes several equilibria may coexist. For instance we can have two equilibria: a first one where no player reaches his target set and an other one where all the players reach their target set. It is thus very natural to identify "relevant" equilibria. In this paper, we consider different notions of relevant Nash equilibria including Pareto optimal equilibria and equilibria with high social welfare. We also study relevant subgame perfect equilibria in reachability games. We provide complexity results for various related decision problems for both Nash equilibria and subgame perfect equilibria.

Keywords:
multiplayer non-zero-sum games played on graphs, reachability objectives, relevant equilibria, social welfare, Pareto optimality

1. Introduction

Two-player zero-sum games played on graphs are commonly used to model *reactive systems* where a system interacts with its environment [1]. In such setting the system wants to achieve a goal - to respect a given property - and the environment acts in an antagonistic way. The system can be described by a game where the two players are the system and the environment, the vertices of the graph are all possible configurations in which the system can be and an infinite path in this graph depicts a possible sequence of interactions between the system and its environment. In such a game, each player chooses a *strategy*: it is the way he plays according to the previous interactions with the other player. Following a strategy for each player results in a *play* in the game. Finding how

*Research partially supported by the PDR project "Subgame perfection in graph games" (T.0088.18 – F.R.S.-FNRS), by COST Action 16228 "GAMENET" (European Cooperation in Science and Technology) and by the ARC project "Non-Zero Sum Game Graphs: Applications to Reactive Synthesis and Beyond" (Fédération Wallonie-Bruxelles).

*Corresponding author
Email addresses: thomas.brihaye@umons.ac.be (Thomas Brihaye), veronique.bruyere@umons.ac.be (Véronique Bruyère), aline.goeminne@umons.ac.be (Aline Goeminne), nathan.thomasset@lsv.fr (Nathan Thomasset)

the system can ensure that a given property is satisfied amounts to finding a *winning strategy* for the system in this game. For some situations, this kind of model is too restrictive and a setting with more than two agents such that each of them has his own not necessarily antagonistic objective is more realistic. These games are called *multiplayer non zero-sum games.* In this setting, the solution concept of winning strategy is not suitable anymore and different notions of *equilibria* can be studied.

In this paper, we focus on *Nash equilibrium* (NE) [2]: given a strategy for each player, no player has an incentive to deviate unilaterally from his strategy. We also consider the notion of *subgame perfect equilibrium* (SPE) [3]: an SPE is a strategy profile that is an NE not only from a given initial configuration but also after any finite number of arbitrary interactions between the players. We study these two notions of equilibria on *reachability games.* In reachability games, we equip each player with a subset of vertices of the graph game that he wants to reach. We are interested in both the *qualitative* and *quantitative* settings. In the qualitative setting, each player only aims at reaching his target set, unlike the quantitative setting where each player wants to reach his target set as soon as possible.

It is well known that both NEs and SPEs exist in both qualitative and quantitative reachability games. But, equilibria such that no player reaches his target set and equilibria such that some players reach it may coexist. This observation has already been made in [4, 5]. In such a situation, one could prefer the second situation to the first one. In this paper, we study different versions of *relevant equilibria.*

Contributions. For quantitative reachability games, we focus on the following three decision problems: *Threshold problem* (Problem 1), *Social welfare problem* (Problem 2), and *Pareto optimal problem* (Problem 3). For the first problem, we ask whether there exists an equilibrium such that the cost of each player, *i.e.*, the number of steps to reach his target set, is upper-bounded by a given threshold. For the second problem, we ask whether there exists an equilibrium such that the *social welfare* is lower-bounded by a given threshold, *i.e.*, such that a certain number of players reach their target set and for all the players reaching their target set, their cost is not too big. For the last problem, we ask whether there exists an equilibrium such that the tuple of the costs obtained by players following this equilibrium is *Pareto optimal* in the set of all the possible costs that players can obtain in the game. We also consider the qualitative adaptations of the three problems.

Our main contributions are the following. *(i)* We study the complexity of the three decision problems. Our results gathered with previous works are summarized in Table 1. *(ii)* In case of a positive answer to any of the three decision problems, we prove that finite-memory strategies are sufficient. Our results and others from previous works are given in Table 2. *(iii)* We identify a subclass of reachability games in which there always exists an SPE where each player reaches his target set. *(iv)* Given a play, we provide a *characterization* which guarantees that this play is the outcome of an NE. This characterization is based on the values in the associated two-player zero-sum games called *coalitional games.*

Table 1: Complexity classes for Problems 1-3.

Complexity	Qualitative Reach.		Quantitative Reach.	
	NE	SPE	NE	SPE
Problem 1	NP-c [6]	PSPACE-c [7]	NP-c	PSPACE-c [8]
Problem 2	NP-c	PSPACE-c	NP-c	PSPACE-c
Problem 3	NP-h/Σ_2^P	PSPACE-c	NP-h/Σ_2^P	PSPACE-c

Table 2: Memory results for Problems 1-3.

Memory	Qualitative Reach.		Quantitative Reach.	
	NE	SPE	NE	SPE
Problem 1	Polynomial [6]	Exponential [7]	Polynomial	Exponential
Problem 2	Polynomial	Exponential	Polynomial	Exponential
Problem 3	Polynomial	Exponential	Polynomial	Exponential

Related Work. There are many results on NEs and SPEs played on graphs for different kinds of qualitative and quantitative objectives. We refer the reader to [9] for a survey and an extended bibliography. Here we focus on the results directly related to our contributions for winning conditions including reachability objectives.

Regarding Problem 1, for NEs, it is shown to be NP-complete in the qualitative setting in [6]; for SPEs it is shown to be PSPACE-complete in both the qualitative and quantitative settings in [7, 8]. Notice that in [4], variants of Problem 1 for games with Streett, parity or co-Büchi winning conditions are shown NP-complete and decidable in polynomial time for Büchi conditions.

Regarding Problem 2, in the setting of games played on matrices, deciding the existence of an NE such that the expected social welfare is at most k is NP-hard [10]. Moreover, in [11] it is shown that deciding the existence of an NE which maximizes the social welfare is undecidable in concurrent games in which a cost profile is associated only with terminal nodes.

Regarding Problem 3, in the setting of zero-sum two-player multidimensional mean-payoff games, the *Pareto-curve* (the set of maximal thresholds that a player can force) is studied in [12] by giving some properties on the geometry of this set. The authors provide a Σ_2^P algorithm to decide if this set intersects a convex set defined by linear inequations.

Regarding the memory, it is shown in [13] that there always exists an NE with polynomial memory in quantitative reachability games, without any constraint on the cost of the NE. It is shown in [5] that, in multiplayer games with ω-regular objectives, there exists an SPE with a given payoff if and only if there exists an SPE with the same payoff but with finite memory. Moreover, in [7] it is claimed that it is sufficient to consider strategies with an exponential memory to solve Problem 1 for SPE in qualitative reachability games.

Finally, we can find several kinds of outcome characterizations for Nash equilibria and variants, *e.g.,* in multiplayer games equipped with prefix-linear cost functions and such that the vertices in coalitional games have a value (summarized in [9]), in multiplayer games with prefix-independent Borel objectives [4], in multiplayer games with classical ω-regular objectives (as reachability) by checking if there exists a play which satisfies an LTL formula [6], in concurrent games [14], etc. Such characterizations are less widespread for subgame perfect equilibria, but one can recover one for quantitative reachability games thanks to a value-iteration procedure [8].

Structure of the Paper. In Section 2, we introduce the needed background and define the different studied problems. In Section 3, we identify families of reachability games for which there always exists a relevant equilibrium, for different notions of relevant equilibria. In Section 4, we state our complexity and memory results in the quantitative setting (see Tables 1 and 2) and provide the material necessary to prove them. In Section 5, we briefly discuss the qualitative setting. The proofs for the qualitative reachability setting are not given because they are in the same spirit as for the quantitative setting.

This article is an extended version of an article that appeared in the Proceedings of RP 2019 [15]. Technical details and proofs about the quantitative setting are here added.

2. Preliminaries and Studied Problems

2.1. Arena, Game and Strategies

An *arena* is a tuple $\mathcal{A} = (\Pi, V, E, (V_i)_{i\in\Pi})$ such that: *(i)* Π is a finite set of players; *(ii)* V is a finite set of vertices; *(iii)* $E \subseteq V \times V$ is a set of edges such that for all $v \in V$ there exists $v' \in V$ such that $(v, v') \in E$ and *(iv)* $(V_i)_{i\in\Pi}$ is a partition of V between the players.

A *play* in $\mathcal{A}$ is an infinite sequence of vertices $\rho = \rho_0\rho_1 \dots$ such that for all $k \in \mathbb{N}$, $(\rho_k, \rho_{k+1}) \in E$. A *history* is a finite sequence $h = h_0h_1 \dots h_k$ with $k \in \mathbb{N}$ defined similarly. Its *length*, denoted by $|h|$, is equal to k which is the number of its vertices minus 1. We denote the set of plays by Plays and the set of histories by Hist. Moreover, the set Hist_i is the set of histories such that their last vertex v is a vertex of Player i, *i.e.*, $v \in V_i$.

Given a play $\rho \in$ Plays and $k \in \mathbb{N}$, the *prefix* $\rho_0\rho_1 \dots \rho_k$ of ρ is denoted by $\rho_{\leq k}$ and its *suffix* $\rho_k\rho_{k+1} \dots$ by $\rho_{\geq k}$. A *lasso* is a play $\rho = h\ell^\omega$ such that $h\ell \in$ Hist. Its *length* is the length of $h\ell$. Notice that ℓ is not necessarily a simple cycle.

A *game* $\mathcal{G} = (\mathcal{A}, (\text{Cost}_i)_{i\in\Pi})$ is an arena equipped with a cost function profile $(\text{Cost}_i)_{i\in\Pi}$ such that for all $i \in \Pi$, $\text{Cost}_i : \text{Plays} \to \mathbb{N} \cup \{+\infty\}$ is a *cost function* which assigns a cost to each play ρ for Player i. We also say that the play ρ has *cost profile* $(\text{Cost}_i(\rho))_{i\in\Pi}$. Given two cost profiles $c, c' \in (\mathbb{N} \cup \{+\infty\})^{|\Pi|}$, we say that $c \leq c'$ if and only if for all $i \in \Pi$, $c_i \leq c'_i$.

An initial vertex $v_0 \in V$ is often fixed, and we call $(\mathcal{G}, v_0)$ an *initialized game*. A play (resp. a history) of $(\mathcal{G}, v_0)$ is then a play (resp. a history) of $\mathcal{G}$ starting in v_0. The set of such plays (resp. histories) is denoted by $\text{Plays}(v_0)$ (resp. $\text{Hist}(v_0)$). The notation $\text{Hist}_i(v_0)$ is used when these histories end in a vertex $v \in V_i$.

Given a game $\mathcal{G}$, a *strategy* for Player i is a function $\sigma_i : \text{Hist}_i \to V$. It assigns to each history hv, with $v \in V_i$, a vertex v' such that $(v, v') \in E$. We denote by Σ_i the set of strategies for Player i. A play $\rho = \rho_0\rho_1 \dots$ is *consistent* with σ_i if for all $\rho_k \in V_i$, $\sigma_i(\rho_0 \dots \rho_k) = \rho_{k+1}$. A strategy σ_i is *positional* if it only depends on the last vertex of the history, *i.e.*, $\sigma_i(hv) = \sigma_i(v)$ for all $hv \in \text{Hist}_i$. It is *finite-memory* if it can be encoded by a finite-state machine [16]. In an initialized game $(\mathcal{G}, v_0)$, a strategy σ_i for Player i needs only to be defined for histories starting in v_0.

A *strategy profile* is a tuple $\sigma = (\sigma_i)_{i\in\Pi}$ of strategies, one for each player. Given an initialized game $(\mathcal{G}, v_0)$ and a strategy profile σ, there exists a unique play from v_0 consistent with each strategy σ_i. We call this play the *outcome* of σ and denote it by $\langle\sigma\rangle_{v_0}$. We say that σ has cost profile $(\mathrm{Cost}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi}$.

2.2. Quantitative Reachability Games

In this article, we are interested in *reachability games*: each player has a target set of vertices that he wants to reach.

Definition 1. A *quantitative reachability game* $\mathcal{G} = (\mathcal{A}, (\mathrm{Cost}_i)_{i\in\Pi}, (F_i)_{i\in\Pi})$ is a game enhanced with a target set $F_i \subseteq V$ for each player $i \in \Pi$ and for all $i \in \Pi$ the cost function Cost_i is defined as follows: for all $\rho = \rho_0\rho_1 \ldots \in \mathrm{Plays}$: $\mathrm{Cost}_i(\rho) = k$ if $k \in \mathbb{N}$ is the least index such that $\rho_k \in F_i$ and $\mathrm{Cost}_i(\rho) = +\infty$ if such index does not exist.

In *quantitative reachability games*, players have to pay a cost equal to the number of edges until visiting their own target set or $+\infty$ if it is not visited. Thus each player aims at *minimizing* his cost.

2.3. Solution Concepts

In the multiplayer game setting, the solution concepts usually studied are *equilibria*. We recall the concepts of Nash equilibrium and subgame perfect equilibrium.

Let $\sigma = (\sigma_i)_{i\in\Pi}$ be a strategy profile in an initialized game $(\mathcal{G}, v_0)$. When we highlight the role of player i, we denote σ by (σ_i, σ_{-i}) where σ_{-i} is the profile $(\sigma_j)_{j\in\Pi\setminus\{i\}}$. A strategy $\sigma'_i \neq \sigma_i$ is a *deviating* strategy of Player i, and it is a *profitable deviation* for him if $\mathrm{Cost}_i(\langle\sigma\rangle_{v_0}) > \mathrm{Cost}_i(\langle\sigma'_i, \sigma_{-i}\rangle_{v_0})$.

The notion of Nash equilibrium is classical: a strategy profile σ in an initialized game $(\mathcal{G}, v_0)$ is a *Nash equilibrium* (NE) if no player has an incentive to deviate unilaterally from his strategy, *i.e.*, no player has a profitable deviation.

Definition 2 (Nash equilibrium)**.** Let $(\mathcal{G}, v_0)$ be an initialized quantitative reachability game. The strategy profile σ is an NE if for each $i \in \Pi$ and each deviating strategy σ'_i of Player i, we have $\mathrm{Cost}_i(\langle\sigma\rangle_{v_0}) \leq \mathrm{Cost}_i(\langle\sigma'_i, \sigma_{-i}\rangle_{v_0})$.

When considering games played on graphs, a useful refinement of NE is the concept of *subgame perfect equilibrium* (SPE). Whereas an NE is a strategy profile where no player has an incentive to deviate in the initialized game $(\mathcal{G}, v_0)$, an SPE is a strategy profile satisfying this property for the initial vertex v_0 but also after each history $hv \in \mathrm{Hist}(v_0)$. We shortly say that an SPE is a strategy profile that is an NE in each subgame. Formally, given a game $\mathcal{G} = (\mathcal{A}, (\mathrm{Cost}_i)_{i\in\Pi})$, an initial vertex v_0, and a history $hv \in \mathrm{Hist}(v_0)$, the initialized game $(\mathcal{G}_{\restriction h}, v)$ such that $\mathcal{G}_{\restriction h} = (\mathcal{A}, (\mathrm{Cost}_{i\restriction h})_{i\in\Pi})$ where $\mathrm{Cost}_{i\restriction h}(\rho) = \mathrm{Cost}_i(h\rho)$ for all $i \in \Pi$ and $\rho \in V^\omega$ is called a *subgame* of $(\mathcal{G}, v_0)$. Notice that $(\mathcal{G}, v_0)$ is a subgame of itself. Moreover if σ_i is a strategy for player i in $(\mathcal{G}, v_0)$, then $\sigma_{i\restriction h}$ denotes the strategy in $(\mathcal{G}_{\restriction h}, v)$ such that for all histories $h' \in \mathrm{Hist}_i(v)$, $\sigma_{i\restriction h}(h') = \sigma_i(hh')$. Similarly, from a strategy profile σ in $(\mathcal{G}, v_0)$, we derive the strategy profile $\sigma_{\restriction h}$ in $(\mathcal{G}_{\restriction h}, v)$.

Definition 3 (Subgame perfect equilibrium)**.** Let $(\mathcal{G}, v_0)$ be an initialized game. A strategy profile σ is an SPE in $(\mathcal{G}, v_0)$ if for all $hv \in \text{Hist}(v_0)$, $\sigma_{\restriction h}$ is an NE in $(\mathcal{G}_{\restriction h}, v)$.

Clearly, any SPE is an NE and it is stated in Theorem 2.1 in [17] that there always exists an SPE (and thus an NE) in quantitative reachability games.

2.4. Studied Problems

We conclude this section with the problems studied in this article. Let us first recall the concepts of social welfare and Pareto optimality. Let $(\mathcal{G}, v_0)$ be an initialized quantitative reachability game with $\mathcal{G} = (\mathcal{A}, (\text{Cost}_i)_{i\in\Pi}, (F_i)_{i\in\Pi})$. Given $\rho = \rho_0\rho_1 \ldots \in \text{Plays}(v_0)$, we denote by $\text{Visit}(\rho)$ the set of players who visit their target set along ρ, *i.e.*, $\text{Visit}(\rho) = \{i \in \Pi \mid \text{there exists } n \in \mathbb{N} \text{ st. } \rho_n \in F_i\}$.[1] The *social welfare* of ρ, denoted by $\text{SW}(\rho)$, is the pair $(|\text{Visit}(\rho)|, \sum_{i\in\text{Visit}(\rho)} \text{Cost}_i(\rho))$ if $\text{Visit}(\rho) \neq \emptyset$ and the pair $(0, +\infty)$ otherwise. Note that it takes into account both the number of players who visit their target set and their accumulated cost to reach those sets. Finally, let $P = \{(\text{Cost}_i(\rho))_{i\in\Pi} \mid \rho \in \text{Plays}(v_0)\} \subseteq (\mathbb{N} \cup \{+\infty\})^{|\Pi|}$. A cost profile $p \in P$ is *Pareto optimal in* $\text{Plays}(v_0)$ if it is minimal in P with respect to the componentwise ordering $\leq$ on P.[2]

Let us now state the studied decision problems. The first two problems are classical: they ask whether there exists a solution (NE or SPE) σ satisfying certain requirements that impose bounds on either $(\text{Cost}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi}$ or on $\text{SW}(\langle\sigma\rangle_{v_0})$.

Problem 1 (Threshold decision problem)**.** Given an initialized quantitative reachability game $(\mathcal{G}, v_0)$, given a threshold $y \in (\mathbb{N} \cup \{+\infty\})^{|\Pi|}$, decide whether there exists a solution σ such that $(\text{Cost}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi} \leq y$.

The most natural requirements are to impose upper bounds on the costs that the players have to pay and no lower bounds. One might also be interested in imposing an interval $[x_i, y_i]$ in which the cost paid by Player i must lie.

In [8], Problem 1 with upper and lower bounds is already solved for SPEs.

Theorem 4 ([8])**.** *For SPEs, Problem 1 with upper (and lower) bounds is PSPACE-complete.*

In the second problem, constraints are imposed on the social welfare, with the aim to maximize it. We use the lexicographic ordering on $\mathbb{N} \times (\mathbb{N} \cup \{+\infty\})$ such that $(k, c) \succeq (k', c')$ if and only if *(i)* $k > k'$ or *(ii)* $k = k'$ and $c \leq c'$.

Problem 2 (Social welfare decision problem)**.** Given an initialized quantitative reachability game $(\mathcal{G}, v_0)$, given two thresholds $k \in \{0, \ldots, |\Pi|\}$ and $c \in \mathbb{N} \cup \{+\infty\}$, decide whether there exists a solution σ such that $\text{SW}(\langle\sigma\rangle_{v_0}) \succeq (k, c)$.

Notice that with the lexicographic ordering, we want to first maximize the number of players who visit their target set, and then to minimize the accumulated cost to reach those sets. Let us now state the last studied problem.

[1] We can easily adapt this definition to histories.

[2] For convenience, we prefer to say that p is Pareto optimal in $\text{Plays}(v_0)$ rather than in P.

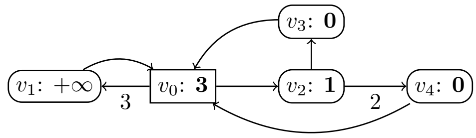

Figure 1: A two-player quantitative reachability game with $F_1 = \{v_3, v_4\}$ and $F_2 = \{v_1, v_4\}$.

Problem 3 (Pareto optimal decision problem)**.** Given an initialized quantitative reachability game $(\mathcal{G}, v_0)$ decide whether there exists a solution σ in $(\mathcal{G}, v_0)$ such that $(\mathrm{Cost}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi}$ is Pareto optimal in $\mathrm{Plays}(v_0)$.

Remark 5. Problems 1 and 2 impose constraints with *non-strict* inequalities. We could also impose strict inequalities or even a mix of strict and non-strict inequalities. The results of this article can be easily adapted to those variants.

We conclude this section with an illustrative example.

Example 6. Consider the quantitative reachability game $(\mathcal{G}, v_0)$ of Figure 1. We have two players such that the vertices of Player 1 (resp. Player 2) are the rounded (resp. rectangular) vertices. For the moment, the reader should not consider the value indicated on the right of the vertices' labeling. Moreover $F_1 = \{v_3, v_4\}$ and $F_2 = \{v_1, v_4\}$. In this figure, an edge (v, v') labeled by x should be understood as a path from v to v' with length x. Observe that F_1 and F_2 are both reachable from the initial vertex v_0. Moreover the two Pareto optimal cost profiles are $(3, 3)$ and $(2, 6)$: take a play with prefix $v_0v_2v_4$ in the first case, and a play with prefix $v_0v_2v_3v_0v_1$ in the second case.

For this example, we claim that there is no NE (and thus no SPE) such that its cost profile is Pareto optimal (see Problem 3). Assume the contrary and suppose that there exists an NE σ such that its outcome ρ has cost profile $(3, 3)$, meaning that ρ begins with $v_0v_2v_4$. Then Player 1 has a profitable deviation such that after history v_0v_2 he goes to v_3 instead of v_4 in a way to pay a cost of 2 instead of 3, which is a contradiction. Similarly assume that there exists an NE σ such that its outcome ρ has cost profile $(2, 6)$, meaning that ρ begins with $v_0v_2v_3v_0v_1$. Then Player 2 has a profitable deviation such that after history v_0 he goes to v_1 instead of v_2, again a contradiction. So there is no NE σ in $(\mathcal{G}, v_0)$ such that $(\mathrm{Cost}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi}$ is Pareto optimal in $\mathrm{Plays}(v_0)$.

The previous discussion shows that there is no NE σ such that $(0, 0) = x \leq (\mathrm{Cost}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi} \leq y = (3, 3)$ (see Problem 1). This is no longer true with $y = (6, 3)$. Indeed, one can construct an NE τ whose outcome has prefix $v_0v_1v_0v_2v_3$ and cost profile $(6, 3)$. This also shows that there exists an NE σ (the same τ as before) that satisfies $\mathrm{SW}(\langle\sigma\rangle_{v_0}) \succeq (k, c) = (2, 9)$ (with τ both players visit their target set and their accumulated cost to reach it equals 9). $\square$

3. Existence Problems

In this section, we show that for particular families of reachability games and requirements, there is no need to solve the related decision problems because they always have a positive answer in this case.

We begin with the family constituted by all reachability games with a *strongly connected* arena. The next theorem then states that there always exists a solution that visits all non-empty target sets.

Theorem 7. *Let $(\mathcal{G}, v_0)$ be an initialized quantitative reachability game such that its arena $\mathcal{A}$ is strongly connected. There exists an SPE σ (and thus an NE) such that its outcome $\langle\sigma\rangle_{v_0}$ visits all target sets F_i, $i \in \Pi$, that are non-empty.*

Let us comment on this result. For this family of games, the answer to Problem 1 is always positive for particular thresholds. In case of quantitative reachability, take strict constraints $< +\infty$ if $F_i \neq \emptyset$ and non-strict constraints $\leq +\infty$ otherwise. We will see later that the strict constraints $< +\infty$ can be replaced by the non-strict constraints $\leq |V| \cdot |\Pi|$ (see Theorem 12). We will also see that, in this setting, the answer to Problem 2 is also always positive for thresholds $k = |\{i \mid F_i \neq \emptyset\}|$ and $c = |\Pi|^2 \cdot |V|$ (see Theorem 12). In order to ease the reading, we relegate the proof of Theorem 7 to Section 3.1.

In the statement of Theorem 7, as the arena is strongly connected, F_i is non-empty if and only if F_i is reachable from v_0. Also notice that the hypothesis that the arena is strongly connected is necessary. Indeed, it is easy to build an example with two players (Player 1 and Player 2) such that from v_0 it is not possible to reach both F_1 and F_2. This is illustrated in Example 8.

Example 8. Consider the initialized quantitative reachability game $(\mathcal{G}, v_0)$ of Figure 2. There are two players, Player 1 who owns round vertices and Player 2 who owns square vertices, and $F_1 = \{v_1\}$, $F_2 = \{v_2\}$. Clearly there is a unique NE $\sigma = (\sigma_1, \sigma_2)$ in $(\mathcal{G}, v_0)$ such that $\sigma_1(v_0) = v_1$ and $\sigma_2(v_1) = v_1$, $\sigma_2(v_2) = v_2$. Its outcome only visits F_1 (and not F_2).

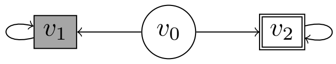

Figure 2: A two-player quantitative reachability game with $F_1 = \{v_1\}$ and $F_2 = \{v_2\}$ where one target set is not reached in equilibrium.

We now turn to the second result of this section. The next theorem states that even with only two players there exists an initialized quantitative reachability game that has no NE with a cost profile which is Pareto optimal. To prove this result, we only have to come back to the quantitative reachability game of Figure 1. We explained in Example 6 that there is no NE in this game such that its cost profile is Pareto optimal.

Theorem 9. *There exists an initialized quantitative reachability game with $|\Pi| = 2$ that has no NE with a cost profile which is Pareto optimal in* $\mathrm{Plays}(v_0)$.

3.1. Technical details and proofs

To prove Theorem 7, we begin with a preliminary lemma and the proof of Theorem 7 follows.

Lemma 10. *Let $\mathcal{G}$ be a quantitative reachability game. Then for all $v_0 \in V$ for which some target set F_j, $j \in \Pi$, is reachable from v_0, there exists an SPE in $(\mathcal{G}, v_0)$ whose outcome ρ visits at least one target set F_i, $i \in \Pi$, that is,* $|\operatorname{Visit}(\rho)| \geq 1$.

Proof. By Theorem 2.1 in [17], there exists an SPE in $(\mathcal{G}, v_0)$ for each initial vertex $v_0 \in V$. Consider the set $U \subseteq V$ of vertices u for which some F_j is reachable from u, and the set $U' \subseteq U$ of those vertices u for which there is an SPE in $(\mathcal{G}, u)$ that visits at least one target set. We have to prove that $U = U'$.

Let us assume that $v_0 \in U \setminus U'$. We claim that there exists an edge (u, u') such that $u \in U \setminus U'$ and $u' \in U'$. Indeed as $v_0 \in U$, there exists a history $h = v_0 v_1 \dots v_k$ with $v_k \in F_j$ for some j. Hence $v_k \in U'$ since the outcome of all SPEs in $(\mathcal{G}, v_k)$ immediately visits F_j. As along h we begin with $v_0 \in U \setminus U'$ and we end with $v_k \in U'$, there must exist an edge $(v_\ell, v_{\ell+1}) = (u, u')$ with $u \in U \setminus U'$ and $u' \in U'$.

Let σ^u (resp. $\sigma^{u'}$) be an SPE in $(\mathcal{G}, u)$ (resp. in $(\mathcal{G}, u')$). As $u' \in U'$, we can suppose that the outcome of $\sigma^{u'}$ visits some target set F_j. From σ^u and $\sigma^{u'}$, we are going to construct another SPE τ in $(\mathcal{G}, u)$ whose outcome will now visit this set F_j. This will lead to a contradiction with $u \in U \setminus U'$. We define such a strategy profile τ equal to σ^u except that it is replaced by $\sigma^{u'}$ for all histories with prefix uu'. More precisely,

- for the particular history u, if $u \in V_i$, then $\tau_i(u) = u'$,
- for each history $uu'h \in \mathrm{Hist}_i$, $i \in \Pi$, we define $\tau_i(uu'h) = \sigma_i^{u'}(u'h)$,
- for each history $uv'h \in \mathrm{Hist}_i$, $i \in \Pi$, with $v' \neq u'$, we define $\tau_i(uv'h) = \sigma_i^u(uv'h)$.

Clearly the outcome of τ is equal to $u\langle\sigma^{u'}\rangle_{u'}$ and thus visits F_j. It remains to show that τ is an SPE, *i.e.*, that $\tau_{\restriction h}$ is an NE in the subgame $(\mathcal{G}_{\restriction h}, v)$ for all $hv \in \mathrm{Hist}_i(u)$, $i \in \Pi$.

- For all histories hv that begin with uv' with $v' \neq u'$, clearly $\tau_{\restriction h}$ is an NE in $(\mathcal{G}_{\restriction h}, v)$ because $\tau_{\restriction h} = \sigma^u_{\restriction h}$ and σ^u is an SPE.
- Take any history hv that begin with uu', and let $h = uh'$. Let τ_i' be a deviating strategy for Player i in $(\mathcal{G}_{\restriction h}, v)$. By definition of τ we have

  $$\begin{aligned} \langle\tau_{\restriction h}\rangle_v &= \langle\sigma^{u'}_{\restriction h'}\rangle_v \\ \langle(\tau_i', \tau_{\restriction h, -i})\rangle_v &= \langle(\tau_i', \sigma^{u'}_{\restriction h', -i})\rangle_v \end{aligned}$$

  Moreover, as u belongs to no target set, we have $\mathrm{Cost}_i(u\rho) = 1 + \mathrm{Cost}_i(\rho)$ for all plays $\rho \in \mathrm{Plays}(u')$. It follows that if τ_i' is a profitable deviation for Player i with respect to $\tau_{\restriction h}$, it is also a profitable deviation with respect to $\sigma^{u'}_{\restriction h'}$. The latter case never holds because $\sigma^{u'}$ is an SPE (and in particular $\sigma^{u'}_{\restriction h'}$ is an NE). Therefore $\tau_{\restriction h}$ is an NE in $(\mathcal{G}_{\restriction h}, v)$.
- It remains to consider the history u and to prove that τ is an NE in $(\mathcal{G}, u)$. From what has been gathered so far, only Player i such that $u \in V_i$ might have a profitable deviation by deviating at the initial vertex u with a strategy τ_i' such that $\tau_i'(u) = v' \neq u' = \tau_i(u)$. Notice that since $u \in U \setminus U'$, we have $\mathrm{Cost}_i(\langle\sigma^u\rangle_u) = +\infty$ and since σ^u is an SPE (and in particular an NE), we have $\mathrm{Cost}_i(\langle\tau_i', \sigma^u_{-i}\rangle_u) = +\infty$. Moreover as $\tau_i'(u) = v' \neq u'$ and by definition of τ, we have $\mathrm{Cost}_i(\langle\tau_i', \sigma^u_{-i}\rangle_u) = \mathrm{Cost}_i(\langle\tau_i', \tau_{-i}\rangle_u) = +\infty$. It follows that τ_i' is not a profitable deviation for Player i with respect to τ, and then τ is an NE in $(\mathcal{G}, u)$. □

Proof of Theorem 7. Let $(\mathcal{G}, v_0)$, with $\mathcal{G} = (\mathcal{A}, (\text{Cost}_i)_{i\in\Pi}, (F_i)_{i\in\Pi})$, be an initialized quantitative reachability game such that its arena is strongly connected. Assume by contradiction that there exists no SPE in $(\mathcal{G}, v_0)$ whose outcome visits all target sets F_i, $i \in \Pi$, that are non-empty. By Theorem 2.1 in [17], there exists an SPE σ in $(\mathcal{G}, v_0)$, and we take such an SPE σ whose outcome $\rho = \langle\sigma\rangle_{v_0}$ visits a maximum number of target sets, say $F_{i_1}, F_{i_2}, \dots, F_{i_k}$. Thus by assumption there exists at least one $F_j \neq \emptyset$ with $j \notin \{i_1, \dots, i_k\}$ that is not visited by ρ. Thanks to Lemma 10, we are going to define from σ another SPE τ in $(\mathcal{G}, v_0)$ whose outcome visits all $F_{i_1}, \dots, F_{i_k}$ as well as an additional target set. This will lead to a contradiction.

Consider a prefix $\rho_0\rho_1 \dots \rho_\ell$ of ρ that visits all $F_{i_1}, \dots, F_{i_k}$. We denote it by gu with $u = \rho_\ell$. From $\mathcal{G}$ we define the quantitative reachability game $\mathcal{G}' = (\mathcal{A}, (\text{Cost}'_i)_{i\in\Pi}, (F'_i)_{i\in\Pi})$ with the same arena $\mathcal{A}$ and such that $F'_i = \emptyset$ if $i \in \{i_1, \dots, i_k\}$ and $F'_i = F_i$ otherwise ($(\text{Cost}'_i)_{i\in\Pi}$ is defined with respect to $(F'_i)_{i\in\Pi}$ as in Definition 1). Notice that $F'_j = F_j$ is not empty and it is reachable from u since $\mathcal{A}$ is strongly connected. Therefore by Lemma 10, there exists an SPE σ' in $(\mathcal{G}', u)$ that visits at least one target set $F'_{j'}$. From σ and σ', we define a strategy profile τ in $(\mathcal{G}, v_0)$ as follows: let $h \in \text{Hist}_i(v_0)$,

- if $h = guh'$ for some h', then $\tau_i(h) = \sigma'_i(uh')$,
- otherwise $\tau_i(h) = \sigma_i(h)$.

Thus, τ acts as σ, except that after a history beginning with gu, it acts as σ'. Clearly the outcome of τ is equal to $g\langle\sigma'\rangle_u$ and thus visits $F'_{j'} = F_{j'}$ in addition to $F_{i_1}, \dots, F_{i_k}$. It remains to show that τ is an SPE. Consider $hv \in \text{Hist}_i(v_0)$, $i \in \Pi$, and let us show that $\tau_{\restriction h}$ is an NE in $(\mathcal{G}_{\restriction h}, v)$.

- If neither hv is a prefix of gu nor gu is a prefix of hv, then $\tau_{\restriction h} = \sigma_{\restriction h}$ by definition of τ, and $\tau_{\restriction h}$ is an NE in $(\mathcal{G}_{\restriction h}, v)$ because σ is an SPE in $(\mathcal{G}, v_0)$.
- If gu is a prefix of hv, let h' such that $gh' = h$. Suppose first that hv visits F_i, then Player i has clearly no incentive to deviate in $(\mathcal{G}_{\restriction h}, v)$. Suppose now that hv does not visit F_i, then $i \notin \{i_1, \dots, i_k\}$ and $F'_i = F_i$ by definition of $\mathcal{G}'$. Hence for all plays π in $(\mathcal{G}_{\restriction h}, v)$ that start in v, $h'\pi$ is a play in $(\mathcal{G}', u)$ that starts in u, and we have $\text{Cost}_i(h\pi) = |gu| + \text{Cost}'_i(h'\pi)$. Hence by definition of τ, a profitable deviation for Player i with respect to $\tau_{\restriction h}$ $(\mathcal{G}_{\restriction h}, v)$ would be a profitable deviation with respect to $\sigma'_{\restriction h'}$ in $(\mathcal{G}'_{\restriction h'}, v)$. The latter case cannot happen as σ' is an SPE in $(\mathcal{G}', u)$ and it follows that $\tau_{\restriction h}$ is an NE in $(\mathcal{G}_{\restriction h}, v)$.
- Consider the last case where hv is a prefix of gu with $hv \neq gu$, and let $hh' = g$. Consider τ'_i a deviating strategy for Player i with respect to $\tau_{\restriction h}$ in the subgame $(\mathcal{G}_{\restriction h}, v)$, and let $\rho' = \langle(\tau'_i, \tau_{\restriction h, -i})\rangle_v$. Without loss of generality, we can suppose that $h'u$ is not a prefix of ρ' since this case was treated at the previous item. Notice that if $i \in \{i_1, \dots, i_k\}$, then $\text{Cost}_i(\langle\tau_{\restriction h}\rangle_v) = \text{Cost}_i(\langle\sigma_{\restriction h}\rangle_v)$, otherwise $\text{Cost}_i(\langle\tau_{\restriction h}\rangle_v) \leq +\infty = \text{Cost}_i(\langle\sigma_{\restriction h}\rangle_v)$. In both cases, as $h'u$ is a prefix of both $\langle\tau_{\restriction h}\rangle_v$ and $\langle\sigma_{\restriction h}\rangle_v$, but not a prefix of ρ', if τ'_i was a profitable deviation for Player i with respect to $\tau_{\restriction h}$, it would also be a profitable deviation with respect to $\sigma_{\restriction h}$ which is impossible since σ is an SPE. □

4. Solving Decision Problems

In this section, we present our main results concerning the three decision problems studied in this paper. In Theorem 11 we provide our complexity results and in Theorem 12 the memory requirements for the equilibria.

Theorem 11. *Let $(\mathcal{G}, v_0)$ be a quantitative reachability game.*

- *For NEs: Problem 1 and Problem 2 are NP-complete while Problem 3 is NP-hard and belongs to Σ_2^P.*
- *For SPEs: Problems 1, 2 and 3 are PSPACE-complete.*

Theorem 12. *Let $(\mathcal{G}, v_0)$ be a quantitative reachability game.*

- *For NEs: for each decision problem, if the answer is positive, then there exists a strategy profile σ with memory in $\mathcal{O}(|\Pi| \cdot |V|)$ which satisfies the conditions.*
- *For SPEs: for each decision problem, if the answer is positive, then there exists a strategy profile σ with memory in $\mathcal{O}(2^{|\Pi|} \cdot |\Pi| \cdot |V|^{(|\Pi|+2)\cdot(|V|+3)+1})$ which satisfies the conditions.*

Moreover, for both NEs and SPEs:

- *for Problem 1 and Problem 3, σ is such that: if $i \in \mathrm{Visit}(\langle\sigma\rangle_{v_0})$, then $\mathrm{Cost}_i(\langle\sigma\rangle_{v_0}) \leq |\Pi| \cdot |V|$;*
- *for Problem 2, if $\mathrm{Visit}(\langle\sigma\rangle_{v_0}) \neq \emptyset$, then $\sum_{i\in\mathrm{Visit}(\langle\sigma\rangle_{v_0})} \mathrm{Cost}_i(\langle\sigma\rangle_{v_0}) \leq |\Pi|^2 \cdot |V|$.*

Notice that no assumption is made on the arena of the game. Even if we provide complexity lower bounds in Theorem 11, the main part is to give the upper bounds. Roughly speaking the decision algorithms work as follows: they guess a path and check that it is the outcome of an equilibrium satisfying the relevant property (such as Pareto optimality). In order to verify that a path is an equilibrium outcome, we rely on the outcome characterization of equilibria, presented in Section 4.2. These characterizations rely themselves on the notion of λ-consistent play, introduced in Section 4.1. As the guessed path should be finitely representable, we show that it is sufficient to consider λ-consistent lassoes, in Section 4.3. We then expose in Section 4.4 the philosophy of the algorithms providing the upper bounds on the complexity of the three problems. Finally, all the technical details and proofs are relegated to Section 4.5.

4.1. λ-Consistent Play

Given a *labeling function*, $\lambda : V \to \mathbb{N} \cup \{+\infty\}$, we define in this section the notion of *λ-consistent play.* Intuitively it is a play ρ such that for each vertex v along ρ the value $\lambda(v)$ represents the maximal number of steps within which the player who owns this vertex can use to reach his target set along ρ starting from v.

Definition 13 (λ-consistent play)**.** Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and $\lambda : V \to \mathbb{N} \cup \{+\infty\}$ be a labeling function. Let $\rho \in \mathrm{Plays}$ be a play, we say that $\rho = \rho_0\rho_1 \ldots$ is λ-consistent if for all $i \in \Pi$ and all $k \in \mathbb{N}$ such that $i \notin \mathrm{Visit}(\rho_0 \ldots \rho_k)$ and $\rho_k \in V_i$: $\mathrm{Cost}_i(\rho_{\geq k}) \leq \lambda(\rho_k)$.

Example 14. Let us come back to Example 6 and assume that the values indicated on the right of the vertices' labeling represent the valuation of a labeling function λ. Let us first consider the play $\rho = (v_0 v_2 v_4)^\omega$ with cost profile $(3, 3)$. We have that $\text{Cost}_2(\rho) = 3 \leq \lambda(v_0) = 3$ but $\text{Cost}_1(\rho_{\geq 1}) = \text{Cost}_1(v_2 v_4 (v_0 v_2 v_4)^\omega) = 2 > \lambda(v_2) = 1$. This means that $(v_0 v_2 v_4)^\omega$ is not λ-consistent. Secondly, one can easily see that the play $v_0 v_1 (v_0 v_2 v_3)^\omega$ is λ-consistent.

4.2. Characterizations

4.2.1. Outcome Characterization of Nash Equilibria

In this section we provide an outcome characterization of NEs for quantitative reachability games, thanks to the previous notion of λ-consistent play for a well-chosen labeling function λ. Notice that such a characterization is not new: it is close to the standard folk theorem from repeated games with perfect information (see for instance [18]). In the context of games played on graphs, we also refer to the survey [9] where a similar characterization is recalled for different kinds of qualitative and quantitative objectives, however not including quantitative reachability.

To define this function λ, we need to study the rational behavior of one player playing against the *coalition* of the other players. Let us informally recall this notions and the related useful concepts; formal definitions will be given later (see Definitions 25-27).

With a quantitative reachability game $\mathcal{G} = (\mathcal{A}, (\text{Cost}_i)_{i \in \Pi}, (F_i)_{i \in \Pi})$, we associate $|\Pi|$ *two-player zero-sum quantitative games* $\mathcal{G}_i$ such that $\mathcal{G}_i$ the *coalitional game* associated with Player i [13]. In this game $\mathcal{G}_i$ Player i (which becomes Player *Min*) wants to reach the target set $F = F_i$ within a minimum number of steps, and the coalition of all players except Player i (which forms one player called Player *Max*, aka $-i$) aims to avoid it or, if it is not possible, maximize the number of steps until reaching F.

Given a coalitional game $\mathcal{G}_i$ and a vertex $v \in V$, the *value*[3] of $\mathcal{G}_i$ from v, depicted by $\text{Val}_i(v)$, allows us to know what is the lowest (resp. greatest) cost (resp. gain) that Player *Min* (resp. Player *Max*) can ensure to obtain from v. Moreover, as quantitative coalitional games are determined these values always exist and can be computed in polynomial time [13, 19, 20].

An *optimal strategy* for Player *Min* (resp. Player *Max*) in a coalitional game $\mathcal{G}_i$ is a strategy which ensures that, from all vertices $v \in V$, Player *Min* (resp. Player *Max*) will pay at most (resp. obtain at least) $\text{Val}_i(v)$ by following this strategy whatever the strategy of the other player. For each $i \in \Pi$, we know that there always exist optimal strategies for both players in $\mathcal{G}_i$. Moreover, we can always find optimal strategies which are positional [13].

The next theorem states that the outcomes of NEs are exactly the plays that are Val-consistent, with the labeling function Val defined in this way: for all $v \in V$, if $v \in V_i$, then $\text{Val}(v) = \text{Val}_i(v)$.

Theorem 15 (Characterization of NEs)**.** *Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and let $\rho \in \text{Plays}(v_0)$ be a play, the following assertions are equivalent:*

1. *there exists an NE σ such that $\langle \sigma \rangle_{v_0} = \rho$;*

[3]also known as minmax value [3].

2. *the play ρ is* Val*-consistent.*

The main idea is that if the second assertion is false, then there exists a player i who has an incentive to deviate along ρ. Indeed, if there exists $k \in \mathbb{N}$ such that $\mathrm{Cost}_i(\rho_{\geq k}) > \mathrm{Val}_i(\rho_k)$ ($\rho_k \in V_i$) it means that Player i can ensure a better cost for him even if the other players play in coalition and in an antagonistic way. Thus, Player i has a profitable deviation. For the second implication, the Nash equilibrium σ is defined as follows: all players follow the outcome ρ but if one player, assume it is Player i, deviates from ρ the other players form a coalition $-i$ and punish the deviator by playing the optimal strategy of Player $-i$ in the coalitional game $\mathcal{G}_i$.

Example 16. Let us go back to Example 14, in this example the used labeling function λ is in fact the labeling function Val. We proved in Example 14 that the play $(v_0v_2v_4)^\omega$ is not Val-consistent and so not the outcome of an NE by Theorem 15. On the contrary, we have seen that the play $v_0v_1(v_0v_2v_3)^\omega$ is Val-consistent and it means that it is the outcome of an NE (again by Theorem 15). Notice that we have already proved these two facts in Example 6.

Up to our knowledge, there is no formal proof of Theorem 15 in the literature. For the sake of completeness we provide such a proof in Section 4.5.1.

4.2.2. Outcome Characterization of Subgame Perfect Equilibria

In the previous section, we proved that the set of plays which are Val-consistent is equal to the set of outcomes of NEs. We now want to have the same kind of characterization for SPEs. We may not use the notion of Val-consistent plays because there exist plays which are Val-consistent but which are not the outcome of an SPE. But, we can recover the characterization of SPEs thanks to a different labeling function defined in [8] that we depict by λ^*.

Let us present this function λ^*; notice that λ^* is not defined on the vertices of the game $\mathcal{G}$ but on the vertices of the *extended game* $\mathcal{X}$ associated with $\mathcal{G}$. The reader is referred to [8] for further details about λ^*.

The set of players in the extended game is the same as in the game $\mathcal{G}$ (*i.e.*, Π) and its vertices (v, I) store a vertex $v \in V$ as well as a subset $I \subseteq \Pi$ of players that have already visited their target sets. The extended game is also a reachability game such that (v, I) is in the target set of Player i as soon as $i \in I$ (*i.e.*, Player i has visited F_i). Therefore all concepts and definitions introduced in Section 2 hold.

Definition 17 (Extended game)**.** Let $\mathcal{G} = (\mathcal{A}, (\mathrm{Cost}_i)_{i\in\Pi}, (F_i)_{i\in\Pi})$ be a quantitative reachability game with an arena $\mathcal{A} = (\Pi, V, E, (V_i)_{i\in\Pi})$. Let v_0 be an initial vertex. The *extended game* of $\mathcal{G}$ is equal to $\mathcal{X} = (\mathcal{A}^X, (\mathrm{Cost}_i)_{i\in\Pi}, (F_i^X)_{i\in\Pi})$. Its arena $\mathcal{A}^X = (\Pi, V^X, E^X, (V_i^X)_{i\in\Pi})$ is such that $V^X = V \times 2^\Pi$ with $V_i^X = \{(v, I) \mid v \in V_i\}$, and $((v, I), (v', I')) \in E^X$ if and only if $(v, v') \in E$ and $I' = I \cup \{i \in \Pi \mid v' \in F_i\}$. Each target set F_i^X is defined such that $(v, I) \in F_i^X$ if and only if $i \in I$. The extended game $\mathcal{X}$ is a quantitative reachability such that for each play ρ in $\mathcal{A}^X$, its cost $\mathrm{Cost}_i(\rho)$ is equal (as in $\mathcal{G}$) to the least index k such that $\rho_k \in F_i^X$, and to $+\infty$ if no such index exists.

The initialized extended game $(\mathcal{X}, x_0)$ associated with the initialized game $(\mathcal{G}, v_0)$ is such that $x_0 = (v_0, I_0)$ with $I_0 = \{i \in \Pi \mid v_0 \in F_i\}$.

There is a one-to-one correspondence between SPEs in $(\mathcal{G}, v_0)$ and its associated initialized extended game $(\mathcal{X}, x_0)$. This is the reason why we solve the different decision problems on the extended game instead of the initial game. However, it is very important to notice that some of our results depend on $|V|$ (resp. $|\Pi|$) that are the number of vertices (resp. players) in $\mathcal{G}$ and not in $\mathcal{X}$.

The labeling function λ^* is obtained thanks to a fixpoint algorithm [8]. We start from an initial labeling function λ^0 that imposes no constraints on the plays in the extended game $\mathcal{X}$. We then iterate an operator that reinforces the constraints step after step, up to obtaining a fixpoint which leads to the required function λ^*. In other words, suppose that λ^k denotes the labeling function computed at step k and each $\Lambda^k(v)$, with $v \in V^X$, denotes the related set of λ^k-consistent plays beginning in v. Initially we have $\Lambda^0(v) = \text{Plays}(v)$, and step by step, the constraints imposed by λ^k become stronger and the sets $\Lambda^k(v)$ become smaller, until a fixpoint is reached.

The initial function λ^0 is defined in a way that all plays in $\mathcal{X}$ are λ^0-consistent. For each $v \in V^X$, let i be such that $v \in V_i^X$, then $\lambda^0(v) = 0$ for all $v \in F_i^X$ and $\lambda^0(v) = +\infty$ otherwise.

Suppose that λ^k, with $k \geq 0$, has been computed. The labeling update to obtain λ^{k+1} is defined as follows.[4] For each $v \in V^X$, let i be such that $v \in V_i^X$, then

$$\lambda^{k+1}(v) = \begin{cases} 0 & \text{if } v \in F_i^X, \\ 1 + \min\limits_{(v,v') \in E^X} \sup\{\text{Cost}_i(\rho) \mid \rho \in \Lambda^k(v')\} & \text{otherwise.} \end{cases}$$

Intuitively, when it is updated, the value $\lambda^{k+1}(v)$ represents what is the best cost that Player i can ensure for himself from v with a "one-shot" choice by only taking into account plays of $\Lambda^k(v')$ with $(v, v') \in E^X$.

The next theorem is the counterpart of Theorem 15 for SPEs.

Theorem 18 ([8] Characterization of SPEs)**.** *Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and $(\mathcal{X}, x_0)$ be its extended game and let $\rho = \rho_0\rho_1 \ldots \in \text{Plays}(x_0)$ be a play in the extended game, the following assertions are equivalent:*

1. *there exists a subgame perfect equilibrium σ such that $\langle\sigma\rangle_{x_0} = \rho$;*
2. *the play ρ is λ^*-consistent.*

4.3. Sufficiency of Lassoes

In this section, we provide technical results which given a λ-consistent play produce an associated λ-consistent lasso. In the rest of this document, we show that working with these lassoes is sufficient for the algorithms.

The associated lassoes are built by eliminating some *unnecessary cycles* and then identifying a prefix $h\ell$ such that ℓ can be repeated infinitely often. An unnecessary cycle is a cycle inside of which no new player visits his target set. More formally, let $\rho = \rho_0\rho_1 \ldots \rho_k \ldots \rho_{k+\ell} \ldots$ be a play in $\mathcal{G}$, if $\rho_k = \rho_{k+\ell}$

[4] To ease the reading, the definition presented here is simplified compared to the one given in [8]. To be correct, the labeling update must be limited to some *region*, i.e, to some set of vertices $v \in V^X$ with the same second component $I \subseteq \Pi$. Moreover, those regions must be treated according to a certain ordering (see [8]).

and $\text{Visit}(\rho_0 \dots \rho_k) = \text{Visit}(\rho_0 \dots \rho_{k+\ell})$ then the cycle $\rho_k \dots \rho_{k+\ell}$ is called an unnecessary cycle.

We call:

- (P1) the procedure which eliminates an unnecessary cycle, *i.e.*, let $\rho = \rho_0\rho_1 \dots \rho_k \dots \rho_{k+\ell} \dots$ such that $\rho_k \dots \rho_{k+\ell}$ is an unnecessary cycle, ρ becomes $\rho' = \rho_0 \dots \rho_k \rho_{k+\ell+1} \dots$
- (P2) the procedure which turns ρ into a lasso $\rho' = h\ell^\omega$ by copying ρ long enough for all players to visit their target set and then to form a cycle after the last player has visited his target set. If no player visits his target set along ρ, then (P2) only copies ρ long enough to form a cycle.

Notice that, given $\rho \in \text{Plays}$, applying (P1) or (P2) may involve a decreasing of the costs but for (P1) and (P2) $\text{Visit}(\rho) = \text{Visit}(\rho')$. Additionally, after applying (P2) we have that $\text{Visit}(h) = \text{Visit}(\rho')$. Moreover, applying (P1) until it is no longer possible and then (P2) leads to a lasso with length at most $(|\Pi|+1) \cdot |V|$ and cost less than or equal to $|\Pi| \cdot |V|$ for players who have visited their target set.

Lemma 19. *Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and $\rho \in \text{Plays}$ be a play.*

- *If ρ' is obtained by applying (P1) on ρ, then $(\text{Cost}_i(\rho'))_{i\in\Pi} \leq (\text{Cost}_i(\rho))_{i\in\Pi}$*
- *If ρ' is obtained by applying (P2) on ρ, then $(\text{Cost}_i(\rho'))_{i\in\Pi} = (\text{Cost}_i(\rho))_{i\in\Pi}$.*
- *Applying (P1) until it is no longer possible and then (P2), leads to a lasso ρ' with length at most $(|\Pi| + 1) \cdot |V|$ and $\text{Cost}_i(\rho') \leq |V| \cdot |\Pi|$ for each $i \in \text{Visit}(\rho')$.*

□

Remark 20 (about Lemma 19). Notice that, given a quantitative reachability game $(\mathcal{G}, v_0)$, as its extended game $(\mathcal{X}, x_0)$ is also a quantitative reachability game, all statements of Lemma 19 also hold for the latter game. Notice that the third assertion applied to $(\mathcal{X}, x_0)$ leads to upper bounds where $|V|$ must be replaced by $|V^X|$ which is exponential in $|\Pi|$ (see Definition 17).

In fact, even in the extended game $(\mathcal{X}, x_0)$ we can obtain the same result: applying (P1) until it is no longer possible and then (P2), leads to a lasso ρ' with size at most $(|\Pi| + 1) \cdot |V|$ and $\text{Cost}_i(\rho') \leq |V| \cdot |\Pi|$ for each $i \in \text{Visit}(\rho')$. This is because along a play ρ in the extended game, the second components of the vertices of ρ form a non-decreasing sequence.

Additionally, applying (P1) or (P2) on λ-consistent plays preserves this property. This is stated in Lemma 21 which is in particular true for extended games.

Lemma 21. *Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and $\rho \in \text{Plays}$ be a λ-consistent play for a given labeling function λ. If ρ' is the play obtained by applying (P1) or (P2) on ρ, then ρ' is λ-consistent.* □

Lemmas 19 and 21 allow us to claim that it is sufficient to deal with lassoes with polynomial length to solve Problems 1, 2 and 3. Moreover, it yields some bounds on the needed memory and the costs for each problem as stated in the next two propositions.

The first proposition is used to solve Problems 1 and 2.

Proposition 22. *Let σ be an NE (resp. SPE) in a quantitative reachability game $(\mathcal{G}, v_0)$ (resp. $(\mathcal{X}, x_0)$ its extended game). Let $w_0 = v_0$ (resp. $w_0 = x_0$). Then there exists τ an NE (resp. SPE) in $(\mathcal{G}, v_0)$ (resp. $(\mathcal{X}, x_0)$) such that:*

- *$\langle\tau\rangle_{w_0}$ is a lasso $h\ell^\omega$ such that $|h\ell| \leq (|\Pi| + 1) \cdot |V|$;*
- *for each $i \in \mathrm{Visit}(\langle\tau\rangle_{w_0})$, $\mathrm{Cost}_i(\langle\tau\rangle_{w_0}) \leq |\Pi| \cdot |V|$;*
- *τ has memory in $\mathcal{O}(|\Pi| \cdot |V|)$ (resp. $\mathcal{O}(2^{|\Pi|} \cdot |\Pi| \cdot |V|^{(|\Pi|+2)\cdot(|V|+3)+1})$).*

Moreover, given $y \in (\mathbb{N} \cup \{+\infty\})^{|\Pi|}$, $k \in \{0, \dots, |\Pi|\}$ and $c \in \mathbb{N} \cup \{+\infty\}$:

- *If $(\mathrm{Cost}_i(\langle\sigma\rangle_{w_0}))_{i\in\Pi} \leq y$, then $(\mathrm{Cost}_i(\langle\tau\rangle_{w_0}))_{i\in\Pi} \leq y$;*
- *If $\mathrm{SW}(\langle\sigma\rangle_{w_0}) \succeq (k, c)$, then $SW(\langle\tau\rangle_{w_0}) \succeq (k, c)$.*

The following proposition is used to solve Problem 3.

Proposition 23. *Let σ be an NE (resp. SPE) in a quantitative reachability game $(\mathcal{G}, v_0)$ (resp. $(\mathcal{X}, x_0)$ its extended game). Let $w_0 = v_0$ (resp. $w_0 = x_0$). If we have that $(\mathrm{Cost}_i(\langle\sigma\rangle_{w_0}))_{i\in\Pi}$ is Pareto optimal in $\mathrm{Plays}(w_0)$, then:*

- *for all $i \in \mathrm{Visit}(\langle\sigma\rangle_{w_0})$, $\mathrm{Cost}_i(\langle\sigma\rangle_{w_0}) \leq |V| \cdot |\Pi|$;*
- *there exists τ an NE (resp. SPE) such that $\langle\tau\rangle_{w_0} = h\ell^\omega$, $|h\ell| \leq (|\Pi|+1)\cdot|V|$ and $(\mathrm{Cost}_i(\langle\sigma\rangle_{w_0}))_{i\in\Pi} = (\mathrm{Cost}_i(\langle\tau\rangle_{w_0}))_{i\in\Pi}$;*
- *τ has memory in $\mathcal{O}(|\Pi| \cdot |V|)$ (resp. $\mathcal{O}(2^{|\Pi|} \cdot |\Pi| \cdot |V|^{(|\Pi|+2)\cdot(|V|+3)+1})$.*

For the sake of clarity we relegate the proofs of Propositions 22 and 23 to Section 4.5.2.

4.4. Algorithms and memory requirements

In this section, we provide the main ideas behind the results stated in Theorems 11 and 12.

4.4.1. Algorithm for NEs

We first focus on Theorem 11 for NEs, *i.e.*, Problem 1 and Problem 2 are NP-complete while Problem 3 is NP-hard and belongs to Σ_2^P. We only provide algorithms to solve these problems and their related complexity, since the proof for the NP-hardness is very similar to the one given in [6]. Recall that Σ_2^P is by definition the class $\mathrm{NP}^{\mathrm{NP}}$, it is also equal to $\mathrm{NP}^{\text{co-NP}}$. The algorithm for each problem works as follows:

1. it guesses a lasso of polynomial length;
2. it verifies that the cost profile of this lasso satisfies the conditions[5] given by the problem;
3. it verifies that the lasso is the outcome of an NE.

Let us comment on the different steps of these algorithms.

[5]Satisfying the conditions is either satisfying the constraints (Problem 1 and Problem 2) or having a cost profile which is Pareto optimal (Problem 3).

- Step 1: For Problem 1 and Problem 2 (resp. Problem 3), it is sufficient to consider plays which are lassoes with polynomial length thanks to Proposition 22 (resp. Proposition 23).
- Step 3: This property is verified thanks to Theorem 15. This is done in polynomial time as the lasso has a polynomial length and the values of the coalitional games are computed in polynomial time.
- Step 2: For Problem 1 and Problem 2, this verification can be obviously done in polynomial time. For Problem 3, we need to have an oracle allowing us to know if the cost profile of the lasso is Pareto optimal. As a consequence, we study Problem 4 which lies in co-NP.

Problem 4. Given a quantitative reachability game $(\mathcal{G}, v_0)$ (resp. its extended game $(\mathcal{X}, x_0)$) and a lasso $\rho \in \text{Plays}(v_0)$ (resp. $\rho \in \text{Plays}(x_0)$), we want to decide if $(\text{Cost}_i(\rho))_{i\in\Pi}$ is Pareto optimal in $\text{Plays}(v_0)$ (resp. $\text{Plays}(x_0)$).

Proposition 24. *Problem 4 lies in co-NP.*

Proof. Let $(\mathcal{G}, v_0)$ be a quantitative reachability game (resp. $(\mathcal{X}, x_0)$ be its extended game) and let $\rho \in \text{Plays}(v_0)$ (resp. $\rho \in \text{Plays}(x_0)$) be a lasso. If ρ is not Pareto optimal, there exists a play ρ' such that $(\text{Cost}_i(\rho))_{i\in\Pi} \geq (\text{Cost}_i(\rho'))_{i\in\Pi}$ and $(\text{Cost}_i(\rho))_{i\in\Pi} \neq (\text{Cost}_i(\rho'))_{i\in\Pi}$. Moreover, thanks to Lemma 19, one may assume that ρ' is a lasso with size at most $(|\Pi| + 1) \cdot |V|$. So, we only have to guess such a lasso ρ' and to verify that $(\text{Cost}_i(\rho))_{i\in\Pi} \geq (\text{Cost}_i(\rho'))_{i\in\Pi}$ and $(\text{Cost}_i(\rho))_{i\in\Pi} \neq (\text{Cost}_i(\rho'))_{i\in\Pi}$. This can be done in polynomial time. □

4.4.2. Algorithm for SPEs

We now focus on Theorem 11 for SPEs, *i.e.*, Problems 1, 2 and 3 are PSPACE-complete. The PSPACE-completeness of Problem 1 is already solved (see Theorem 4). We thus provide algorithms to solve Problems 2 and 3 and their related complexity. We do not provide the proof for the PSPACE-hardness as it is very similar to the one given in [8].

The algorithm for Problem 2 and 3 works as follows:

1. it guesses a lasso of polynomial length;
2. it verifies that the cost profile c of this lasso satisfies the conditions given by the problem;
3. it checks, whether there exists an SPE with cost profile equal to c.

The explanations for the first and the second steps are the same as for the algorithms for NEs. Finally, we know that the third step can be done in PSPACE by Theorem 4.

4.4.3. Memory requirements

We now turn to Theorem 12 that provides memory requirements for the equilibria in case of positive answer to the studied decision problems. Its proof directly follows from Proposition 22 (resp. Proposition 23) for Problems 1 and 2 (resp. Problem 3).

4.5. Technical details and proofs

In order to better highlight the results and to ease the reading, we choose to present the technical proofs in this section.

4.5.1. Proof of Theorem 15 (Characterization of NEs)

We first provide the formal definitions of coalitional game, value and optimal strategy.

Definition 25 (Coalitional game)**.** Let $\mathcal{A} = (\Pi, V, E, (V_i)_{i\in\Pi})$ be an arena and $\mathcal{G} = (\mathcal{A}, (\text{Cost}_i)_{i\in\Pi}, (F_i)_{i\in\Pi})$ be a quantitative reachability game with $|\Pi| \geq 2$. With each player $i \in \Pi$, we can associate a two-player zero-sum quantitative reachability game depicted by $\mathcal{G}_i = (\mathcal{A}_i, (\text{Cost}_{\text{Min}}, \text{Gain}_{\text{Max}}), F)$ and defined as follows: *i)* $\mathcal{A}_i = (\{i, -i\}, V, E, (V_i, V\backslash V_i))$ where Player i (resp. $-i$) can be called Player Min (resp. Player Max); *ii)* $\text{Cost}_{\text{Min}} = \text{Cost}_i$ and $\text{Gain}_{\text{Max}} = \text{Cost}_{\text{Min}}$ and *iii)* $F = F_i$.

Definition 26 (Value)**.** Let $\mathcal{G}_i$ be a coalitional game and $v \in V$ be a vertex, we define the value of $\mathcal{G}_i$ from v as :

$$\text{Val}_i(v) = \inf_{\sigma_1\in\Sigma_{Min}} \sup_{\sigma_2\in\Sigma_{Max}} \text{Cost}_{\text{Min}}(\langle\sigma_1, \sigma_2\rangle_v). \tag{1}$$

Remark that, as for each $i \in \Pi$ the coalitional game $\mathcal{G}_i$ is determined ([13]) and $\text{Cost}_{\text{Min}} = \text{Gain}_{\text{Max}}$, the equality (1) could be defined as $\text{Val}_i(v) = \sup_{\sigma_2\in\Sigma_{Max}} \inf_{\sigma_1\in\Sigma_{Min}} \text{Gain}_{\text{Max}}(\langle\sigma_1, \sigma_2\rangle_v)$.

Definition 27 (Optimal strategy)**.** Let $\mathcal{G}_i$ be a coalitional game, we say that $\sigma_1^* \in \Sigma_{Min}$ is an *optimal strategy* for player Min if, for all $v \in V$, we have that: $\sup_{\sigma_2\in\Sigma_{Max}} \text{Cost}_{\text{Min}}(\langle\sigma_1^*, \sigma_2\rangle_v) \leq \text{Val}_i(v)$. Similarly, we say that $\sigma_2^* \in \Sigma_{Max}$ is an optimal strategy for player Max if, for all $v \in V$, we have that:
$\inf_{\sigma_1\in\Sigma_{Min}} \text{Gain}_{\text{Max}}(\langle\sigma_1, \sigma_2^*\rangle_v) \geq \text{Val}_i(v)$.

Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and $\rho \in \text{Plays}(v_0)$ be a play. Let us prove Theorem 15, *i.e.*, there exists an NE σ such that $\langle\sigma\rangle_{v_0} = \rho$ if and only if the play ρ is Val-consistent.

Proof of Theorem 15. Let us first recall that, for all $i \in \Pi$, the coalitional game $\mathcal{G}_i$ is determined and there are optimal positional strategies for both players $(\sigma_i^*, \sigma_{-i}^*)$. Moreover since $\text{Cost}_{\text{Min}} = \text{Gain}_{\text{Max}}$, for all $v \in V$, we have:

$$\inf_{\sigma_i\in\Sigma_{Min}} \text{Cost}_{\text{Min}}(\langle\sigma_i, \sigma_{-i}^*\rangle_v) = \text{Val}_i(v) = \sup_{\sigma_{-i}\in\Sigma_{Max}} \text{Cost}_{\text{Min}}(\langle\sigma_i^*, \sigma_{-i}\rangle_v).$$

From the optimal strategy σ_{-i}^* in $\mathcal{G}_i$ we can extract a strategy $\sigma_{j,i}^*$ of Player j in $\mathcal{G}$. Notice also that even if σ_i^* is a strategy in $\mathcal{G}_i$, we can use it as a strategy of Player i in $\mathcal{G}$.

Let us prove the equivalence between the two assertions of Theorem 15.

1 $\Rightarrow$ **2**: Let σ be a Nash equilibrium in $(\mathcal{G}, v_0)$ such that $\langle\sigma\rangle_{v_0} = \rho$. Let us assume by contradiction that there exist $i \in \Pi$ and $k \in \mathbb{N}$ such that $i \notin \text{Visit}(\rho_0 \dots \rho_k)$ and $\rho_k \in V_i$ such that:

$$\text{Cost}_i(\rho_{\geq k}) > \text{Val}_i(\rho_k). \tag{2}$$

Let $h = \rho_0 \dots \rho_{k-1}$, we can write:

$$\text{Cost}_i(\rho_{\geq k}) = \text{Cost}_i(\langle\sigma_{\restriction h}\rangle_{\rho_k}). \tag{3}$$

Additionally, by definition of value in a coalitional game and thanks to the fact that the optimal strategies are positional:

$$\begin{aligned}\mathrm{Val}_i(\rho_k) &= \sup_{\tau_{-i}\in\Sigma_{Max}} \mathrm{Cost}_{\mathrm{Min}}(\langle\sigma_i^*, \tau_{-i}\rangle_{\rho_k}) \\ &\geq \mathrm{Cost}_{\mathrm{Min}}(\langle\sigma_i^*, \sigma_{-i\restriction h}\rangle_{\rho_k}) \\ &= \mathrm{Cost}_i(\langle\sigma_i^*, \sigma_{-i\restriction h}\rangle_{\rho_k}) \qquad (4)\end{aligned}$$

where σ_i^* is the optimal strategy of Player i in $\mathcal{G}_i$ and σ_{-i} is an abuse of notation to depict the strategy of the coalition $-i = \Pi\backslash\{i\}$ which follows strategies σ_j for all $j \neq i$.

By (2), (3) and (4), it follows that:

$$\mathrm{Cost}_i(\langle\sigma_i^*, \sigma_{-i\restriction h}\rangle_{\rho_k}) < \mathrm{Cost}_i(\langle\sigma_{\restriction h}\rangle_{\rho_k}).$$

As $i \notin \mathrm{Visit}(h)$ by hypothesis, we can conclude that:

$$\mathrm{Cost}_i(h\langle\sigma_i^*, \sigma_{-i\restriction h}\rangle_{\rho_k}) < \mathrm{Cost}_i(h\langle\sigma_{\restriction h}\rangle_{\rho_k}) = \mathrm{Cost}_i(\rho).$$

This means that following σ_i along h and then σ_i^* once he reaches ρ_k is a profitable deviation for Player i. This concludes the proof.

2 ⇒ 1: Let τ be a strategy profile such that $\langle\tau\rangle_{v_0} = \rho$. From τ we aim to construct a Nash equilibrium with the same outcome. The main idea is the following one: first, all players play according to τ. But if a player, let us call him Player i deviates from τ_i, the other players form a coalition and each of them plays their strategy obtained thanks to the strategy σ_{-i}^* in $\mathcal{G}_i$.

In order to define properly the Nash equilibrium that we are looking for, we have to define a punishment function $P : \mathrm{Hist}(v_0) \to \Pi \cup \{\perp\}$ which allows us to know who is the player who has deviated for the first time from τ. So for all $h \in \mathrm{Hist}(v_0)$, $P(h) =\perp$ if no player has yet deviated and $P(h) = i$ for some $i \in \Pi$ if Player i is the first player who has deviated along h. We can define P as follows: for the initial vertex $P(v_0) =\perp$ and then for all history $hv \in \mathrm{Hist}(v_0)$ with $v \in V$:

$$P(hv) = \begin{cases} \perp & \text{if } P(h) =\perp \text{ and } hv \text{ is a prefix of } \rho, \\ i & \text{if } P(h) =\perp \text{, } hv \text{ is not a prefix of } \rho \text{ and } h \in \mathrm{Hist}_i, \\ P(h) & \text{otherwise.} \end{cases}$$

We now define σ. For all $i \in \Pi$ and for all $h \in \mathrm{Hist}_i(v_0)$:

$$\sigma_i(h) = \begin{cases} \tau_i(h) & \text{if } P(h) =\perp, \\ \sigma_i^*(h) & \text{if } P(h) = i, \\ \sigma_{i,P(h)}^*(h) & \text{otherwise.} \end{cases}$$

It is clear that $\langle\sigma\rangle_{v_0} = \rho$. It remains to prove that σ is a Nash equilibrium in $(\mathcal{G}, v_0)$. Let us assume that σ is not an NE. It means that there exists a profitable deviation depicted by $\tilde{\sigma}_i$ for some player $i \in \Pi$. We choose i such that Player i is the first player who has a profitable deviation from σ along ρ. Let $\tilde{\rho} = \langle\tilde{\sigma}_i, \sigma_{-i}\rangle_{v_0}$ the outcome such that Player i plays his profitable deviation. As $\tilde{\sigma}_i$ is a profitable deviation we have:

$$\mathrm{Cost}_i(\tilde{\rho}) < \mathrm{Cost}_i(\rho). \tag{5}$$

Moreover as ρ and $\tilde{\rho}$ both begin in v_0, they have a common prefix. Let $hv \in \mathrm{Hist}_i$ this longest common prefix. We have that: $\rho = h\langle \sigma_{\restriction h}\rangle_v$ and $\tilde{\rho} = h\langle \tilde{\sigma}_{i\restriction h}, \sigma_{-i\restriction h}\rangle_v$. Notice that $i \notin \mathrm{Visit}(hv)$. But, by definition of σ and as the optimal strategies in $\mathcal{G}_i$ are positional, we can rewrite these two equalities as follows: $\rho = h\langle \tau_{\restriction h}\rangle_v$ and $\tilde{\rho} = h\langle \tilde{\sigma}_{i\restriction h}, (\sigma^*_{j,i})_{j\in\Pi\setminus\{i\}}\rangle_v$. Additionally, thanks to the definition of the value in the coalitional game $\mathcal{G}_i$:

$$\begin{aligned}\mathrm{Val}_i(v) &= \inf_{\mu_i \in \Sigma_{Min}} \mathrm{Cost}_{\mathrm{Min}}(\langle \mu_i, \sigma^*_{-i}\rangle_v) \\ &\leq \mathrm{Cost}_{\mathrm{Min}}(\langle \tilde{\sigma}_{i\restriction h}, \sigma^*_{-i}\rangle_v) \\ &= \mathrm{Cost}_i(\langle \tilde{\sigma}_{i\restriction h}, (\sigma^*_{j,i})_{j\in\Pi\setminus\{i\}}\rangle_v). \end{aligned} \tag{6}$$

By hypothesis, as hv is a prefix of ρ and $i \notin \mathrm{Visit}(hv)$, we have that $\mathrm{Val}_i(v) \geq \mathrm{Cost}_i(\langle \tau_{\restriction h}\rangle_v)$. Thus by (6), it follows that:

$$\mathrm{Cost}_i(\langle \tilde{\sigma}_{i\restriction h}, (\sigma^*_{j,i})_{j\in\Pi\setminus\{i\}}\rangle_v) \geq \mathrm{Cost}_i(\langle \tau_{\restriction h}\rangle_v).$$

And thanks to the definition of the cost function associated with quantitative reachability games, we have that:

$$\mathrm{Cost}_i(h\langle \tilde{\sigma}_{i\restriction h}, (\sigma^*_{j,i})_{j\in\Pi\setminus\{i\}}\rangle_v) \geq \mathrm{Cost}_i(h\langle \tau_{\restriction h}\rangle_v).$$

Thus, we can conclude that $\mathrm{Cost}_i(\tilde{\rho}) \geq \mathrm{Cost}_i(\rho)$ which leads to a contradiction with (5). This concludes the proof. □

Theorem 15 and its previous proof can be easily adapted to lassoes as follows.

Corollary 28 (of Theorem 15)**.** *Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and let $\rho = h\ell^\omega \in \mathrm{Plays}(v_0)$ be a lasso, the following assertions are equivalent:*

1. *there exists an NE σ with memory in $\mathcal{O}(|h\ell|+|\Pi|)$ and such that $\langle\sigma\rangle_{v_0} = \rho$.*
2. *the play ρ is* Val-*consistent.*

Proof. Let us now assume that $\rho = h\ell^\omega$ is a lasso. The implication **1** $\Rightarrow$ **2** is the same as in the previous proof. Thus we only have to prove that, in the implication **2** $\Rightarrow$ **1**, the previously built strategy σ has memory in $\mathcal{O}(|h\ell| + |\Pi|)$. The intuition is the following. If $\rho = h\ell^\omega$, a player has to remember: *(i)* $h\ell$ to know both what he has to play and if someone has deviated and *(ii)* who is the deviator. Once a deviation has occured, both players play memoryless strategies. □

4.5.2. Proof of Propositions 22 and 23

In this section, we prove Propositions 22 and 23. To this end, we need to discuss memory requirements for SPEs as done for NEs in Corollary 28. This will done in Proposition 31 (item 4). Let us introduce some additional technical notions and intermediate results.

By adapting the concept of *(good) symbolic witness* (a set of lassoes with some good properties) used in [7], we can show that if there exists an SPE with a cost profile c then, there exists one with the same cost profile but with a finite-memory. This leads to Proposition 31 which allows us to prove Proposition 22 for SPEs.

Before the statement of the proposition, we have to introduce the notion of *very weak subgame perfect equilibrium* (very weak SPE) and to formally introduce what is a (good) symbolic witness. This latter notion was introduced in [7] for games with prefix-independent gain functions. We adapt it for quantitative reachability games for which the cost function is not prefix-independent.

Very weak SPE. We begin by recalling the concept of very weak SPE introduced in [21, 22]. Let $(\mathcal{G}, v_0)$ be an initialized game and $\sigma = (\sigma_i)_{i\in\Pi}$ be a strategy profile. Given $i \in \Pi$, we say that a strategy σ'_i is *one-shot deviating* from σ_i if σ'_i and σ_i only differ on the initial vertex v_0. A strategy profile σ is a *very weak NE* in $(\mathcal{G}, v_0)$ if, for each player $i \in \Pi$, for each strategy σ'_i of Player i that is one-shot deviating from σ_i, we have $\text{Cost}_i(\langle\sigma\rangle_{v_0}) \leq \text{Cost}_i(\langle\sigma'_i, \sigma_{-i}\rangle_{v_0})$. A strategy profile σ is a *very weak SPE* in $(\mathcal{G}, v_0)$ if, for all $hv \in \text{Hist}(v_0)$, $\sigma_{\restriction h}$ is a very weak NE in $(\mathcal{G}_{\restriction h}, v)$.

Every SPE is a very weak SPE. Additionally, for quantitative reachability games these two concepts are equivalent.

Proposition 29 ([21, 22])**.** *Let $(\mathcal{G}, v_0)$ be an initialized quantitative reachability game and σ be a strategy profile in $(\mathcal{G}, v_0)$. Then σ is an SPE if and only if σ is a very weak SPE.*

(Good) symbolic witness. A *symbolic witness* is a finite set of plays and a *good symbolic witness* is a symbolic witness such that the plays respect some good property. The intuition behind this property is the following: each time you consider a play ρ in the good symbolic witness no player has an incentive to deviate and to follow another play of the good symbolic witness. We will see that a good symbolic witness is all we need to build a very weak SPE. Moreover, if each play of the good symbolic witness is a lasso, the very weak SPE requires finite memory and this good symbolic witness provides a finite representation of this equilibrium.

Given an initialized quantitative reachability game $(\mathcal{G}, v_0)$ and $(\mathcal{X}, (v_0, I_0))$ its extended game (see Definition 17), we define the set $\mathcal{I}$, a subset of $(\Pi \cup \{0\}) \times V \times 2^{\Pi}$ such that:

$$\begin{aligned}\mathcal{I} = \{(0, v_0, I_0)\} \;\cup\; &\{(i, v', I') \mid \text{there exists } ((v, I), (v', I')) \in E^X \\ &\text{with } (v, I), (v', I') \in \text{Succ}^*(v_0, I_0) \text{ and } v \in V_i\}.\end{aligned}$$

Notation $(v, I) \in \text{Succ}^*(v_0, I_0)$ means that vertex (v, I) is reachable from the initial vertex (v_0, I_0) in the extended game. The set $\mathcal{I}$ is thus composed of some such reachable vertices (with additional information depending on players).

Definition 30 ((Good) Symbolic witness)**.** Let $(\mathcal{G}, v_0)$ be an initialized quantitative reachability game and $(\mathcal{X}, (v_0, I_0))$ its extended game.

- A *symbolic witness* is a set $\mathcal{P} = \{\rho_{i,v,I} \mid (i, v, I) \in \mathcal{I}\}$ such that each $\rho_{i,v,I}$ is a lasso in $\mathcal{X}$ with first vertex equal to (v, I).

- A symbolic witness $\mathcal{P}$ is *good* if for all $\rho_{j,u,J},\ \rho_{i,v',I'} \in \mathcal{P}$, for all suffixes $\rho \in \text{Plays}(v, I)$ of $\rho_{j,u,J}$ such that $((v, I), (v', I')) \in E^X$ and $(v, I) \in V_i^X$, if $i \notin I$, then we have:

$$\text{Cost}_i(\rho) \leq 1 + \text{Cost}_i(\rho_{i,v'I'}).$$

In Figure 3, we provide an illustration of the condition which should be satisfied to be a good symbolic witness (see Definition 30).

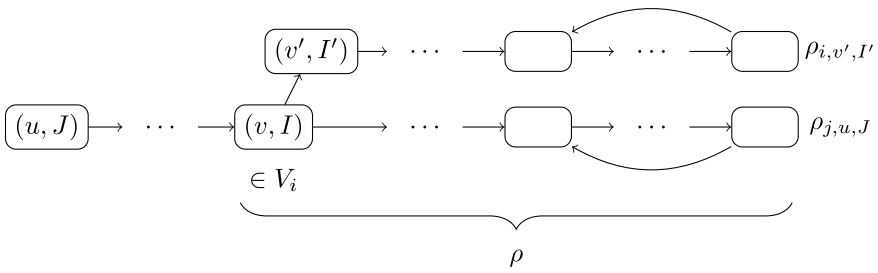

Figure 3: The condition of Definition 30.

From Theorem 18 we know that if ρ is the outcome of an SPE then ρ is λ^*-consistent. In fact, it is possible to build from the sets of λ^*-consistent plays a good symbolic witness and then to construct a very weak SPE that is composed of finite-memory strategies. The following proposition formalizes these results. We use the notation $\Lambda^*(v, I) = \{\rho \in \text{Plays}(v, I) \mid \rho \text{ is } \lambda^*\text{-consistent}\}$.

Proposition 31. *Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and $(\mathcal{X}, (v_0, I_0))$ be its extended game, let $\rho \in \text{Plays}(v_0, I_0)$ and $c \in (\mathbb{N} \cup \{+\infty\})^{|\Pi|}$ such that $(\text{Cost}_i(\rho))_{i\in\Pi} = c$. Let $M = \max_{i\in\Pi}\{c_i \mid c_i < +\infty\}$ if this max exists, $M = 0$ otherwise. The following assertions are equivalent:*

1. *There exists an SPE with outcome ρ in $(\mathcal{X}, (v_0, I_0))$;*
2. *$\Lambda^*(v, I) \neq \emptyset$ for all $(v, I) \in \text{Succ}^*(v_0, I_0)$ and $\rho \in \Lambda^*(v_0, I_0)$;*
3. *There exists a good symbolic witness $\mathcal{P}$ that contains a lasso ρ_{0,v_0,I_0} with cost profile c and $|\rho_{0,v_0,I_0}| \leq M + |V|$. Moreover, for each $\rho_{i,v,I} \in \mathcal{P}$, $|\rho_{i,v,I}| \leq \mathcal{O}(|V|^{(|\Pi|+2)\cdot(|V|+3)}) + |\Pi| \cdot |V|$;*
4. *There exists a finite-memory SPE σ with cost profile c in $(\mathcal{X}, (v_0, I_0))$ such that its memory is in $\mathcal{O}(M + 2^{|\Pi|} \cdot |\Pi| \cdot |V|^{(|\Pi|+2)\cdot(|V|+3)+1})$.*

Remark 32. The details of Proposition 2.13 in [8] shows that if there exists a very weak SPE in $(\mathcal{X}, (v_0, I_0))$ then for each $(v, I) \in V^X$, $\Lambda^*(v, I) \neq \emptyset$. Since it is always true in the setting of quantitative reachability game [17, Theorem 2.1], the second item may be replaced by $\rho \in \Lambda^*(v_0, I_0)$ and we get back exactly to Theorem 18.

The proofs of ($2 \Rightarrow 3$) and ($3 \Rightarrow 4$) are nearly identical to the one of the one of Proposition 2.14 in [8]. Notice that in the proof given in [8] there is no consideration about the memory of the built SPE. In order to do so, in the proof below, we choose more adequatly the plays of the form $\rho_{i,v',I'}$ by picking lassoes in the sets of λ^*-consistent plays beginning in (v', I'). To obtain a bound on the length of these lassoes, we use the following lemma.

Lemma 33 ([8]). *Let v be a vertex in the extended game, let* $\text{MaxCost}_i(v) = \max\{\text{Cost}_i(\rho) \mid \rho \in \text{Plays}(v)$ *and ρ is λ^*-consistent*$\}$. *If* $\text{MaxCost}_i(v) < +\infty$, *then* $\text{MaxCost}_i(v) \leq \mathcal{O}(|V|^{(|\Pi|+2)\cdot(|V|+3)})$.

Proof sketch of Proposition 31. **1** $\Rightarrow$ **2**: Proposition 2.13 in [8].

2 $\Rightarrow$ **3**: We build a symbolic witness $\mathcal{P}$ step by step and then we prove that it is good. At the initialization, $\mathcal{P} = \emptyset$.

Let $\rho \in \Lambda^*(v_0, I_0)$ such that $(\text{Cost}_i(\rho))_{i\in\Pi} = c$. We apply (P2) on ρ to obtain a lasso ρ_{0,v_0,I_0} such that $|\rho_{0,v_0,I_0}| \leq M + |V|$ and $(\text{Cost}_i(\rho_{0,v_0,I_0}))_{i\in\Pi} = c$. Moreover, as ρ is λ^*-consistent, ρ_{0,v_0,I_0} is also λ^*-consistent (by Lemma 21). We add ρ_{0,v_0,I_0} to $\mathcal{P}$.

For each $(i, v, I) \in \mathcal{I}$, let ρ be such that $\rho \in \underset{\rho' \in \Lambda^*(v,I)}{\text{argmax}} \{\text{Cost}_i(\rho')\}$. We obtain $\rho_{i,v,I}$ by copying ρ until Player i has visited his target set, then by removing the unnecessary cycles and apply (P2). If Player i does not visit his target set along ρ, we remove all the unnecessary cycles (by applying iteratively (P1)) and then we apply (P2). By the same kind of arguments used in Lemma 19 and Lemma 21, we obtain that: *i)* $\rho_{i,v,I}$ is λ^*-consistent, *ii)* $\text{Cost}_i(\rho_{i,v,I}) = \text{Cost}_i(\rho)$ and *iii)* $|\rho_{i,v,I}| \leq \mathcal{O}(|V|^{(|\Pi|+2)\cdot(|V|+3)}) + |\Pi| \cdot |V|$ (by Lemma 33). We add $\rho_{i,v,I}$ to $\mathcal{P}$.

By construction $\mathcal{P}$ is a symbolic witness. If we look at the proof of Proposition 2.14 in [8], we can be convinced that $\mathcal{P}$ is a good symbolic witness due to the same kind of arguments.

3 $\Rightarrow$ **4**: Given a good symbolic witness $\mathcal{P}$ with properties given in statement 3, an SPE with cost profile c and finite memory is built from the lassoes in $\mathcal{P}$ in the same way as in the proof of Proposition 2.14. in [8]. The inequality (2.10) in this proof, which allows to conclude that the built strategy profile is a very weak SPE, is similar to the condition to be a good symbolic witness. Thus with the same kind of arguments, we are able to prove that the strategy profile is a very weak SPE and so an SPE.

It remains to prove that σ is finite-memory with memory in $\mathcal{O}(M + 2^{|\Pi|} \cdot |\Pi| \cdot |V|^{(|\Pi|+1)\cdot(|\Pi|+|V|)+1})$. Having (j, u, J) in memory (the last deviating player j and the vertex (u, J) where he moved), the machine $\mathcal{M}_i$, $i \in \Pi$, which represents the strategy σ_i, has to produce the lasso $\rho_{j,u,J}$ of length bounded by $M + |V|$ for ρ_{0,v_0,I_0} and by $\mathcal{O}(|V|^{(|\Pi|+2)\cdot(|V|+3)} + |\Pi| \cdot |V|)$ for the others (at most $|\Pi| \cdot |V| \cdot 2^{|\Pi|}$ such lassoes). It leads to a memory in $\mathcal{O}(M + |V| + |\Pi| \cdot |V| \cdot 2^{|\Pi|}(|V|^{(|\Pi|+2)\cdot(|V|+3)} + |\Pi| \cdot |V|)) = \mathcal{O}(M + |\Pi| \cdot 2^{|\Pi|} \cdot |V|^{(|\Pi|+2)\cdot(|V|+3)+1})$ (we assume without loss of generality that $|\Pi| \leq |V|$).

4 $\Rightarrow$ **1**: Obvious.

□

We are now able to prove Propositions 22 and 23. We begin by the first one. Recall that it states that if there exists an NE σ (resp. SPE) in a reachability game, then one can construct another one, τ, such that its outcome is a lasso of polynomial length and τ is composed of finite-memory strategies with polynomial (resp. exponential) size. Moreover, if σ is a solution to Problem 1 (resp. Problem 2), it is also the case for τ.

Proof of Proposition 22.

- **For NEs**: Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and σ be an NE in $(\mathcal{G}, v_0)$. Let $\rho = \langle\sigma\rangle_{v_0}$. We apply procedure (P1) on ρ until there is no longer any unnecessary cycle and then we apply (P2). In this way, we obtain a lasso $\rho' = h\ell^\omega \in \text{Plays}(v_0)$. By Lemma 19, $|h\ell| \leq (|\Pi| + 1) \cdot |V|$ and $\text{Cost}_i(h\ell^\omega) \leq \min\{\text{Cost}_i(\langle\sigma\rangle_{v_0}), |\Pi| \cdot |V|\}$ if $i \in \text{Visit}(\langle\sigma\rangle_{v_0})$ and $\text{Cost}_i(h\ell^\omega) = +\infty$ otherwise.

  By hypothesis and thanks to Theorem 15, we know that ρ is Val-consistent. Thus, by Lemma 21, ρ' is Val-consistent. Thanks to Corollary 28, there exists an NE τ such that $\langle\tau\rangle_{v_0} = \rho' = h\ell^\omega$ with memory $\mathcal{O}(|\Pi| \cdot |V|)$ (we can assume without loss of generally that $|V|, |\Pi| \geq 1$).

  Let $y \in (\mathbb{N} \cup \{+\infty\})^{|\Pi|}$, let us assume that $(\text{Cost}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi} \leq y$, by Lemma 19 we have that $(\text{Cost}_i(\langle\tau\rangle_{v_0})) \leq y$.

  Let $k \in \{0, \dots, |\Pi|\}$ and $c \in \mathbb{N} \cup \{+\infty\}$, let us assume that $\text{SW}(\langle\sigma\rangle_{v_0}) \succeq (k, c)$. If $\text{SW}(\langle\sigma\rangle_{v_0}) = (k_1, c_1)$ and $\text{SW}(\langle\tau\rangle_{v_0}) = (k_2, c_2)$, we have $k_1 = k_2$ and $c_2 \leq c_1 \leq c$ thanks to Lemma 19. Thus, we have that $\text{SW}(\langle\tau\rangle_{v_0}) \succeq (k, c)$.

- **For SPEs:**

  Let $(\mathcal{G}, v_0)$ be a quantitative reachability game and $(\mathcal{X}, x_0)$ be its extended game. Let σ be an SPE in $(\mathcal{X}, x_0)$. Let $\rho = \langle\sigma\rangle_{x_0}$. We apply procedure (P1) on ρ until there is no longer any unnecessary cycle and then we apply (P2). In this way, we obtain a lasso $\rho' = h\ell^\omega \in \text{Plays}(x_0)$. By Lemma 19, $|h\ell| \leq (|\Pi| + 1) \cdot |V|$ and $\text{Cost}_i(h\ell^\omega) \leq \min\{\text{Cost}_i(\langle\sigma\rangle_{x_0}), |\Pi| \cdot |V|\}$ if $i \in \text{Visit}(\langle\sigma\rangle_{x_0})$ and $\text{Cost}_i(h\ell^\omega) = +\infty$ otherwise.

  By hypothesis and by Theorem 18, we know that ρ is λ^*-consistent. Thus, by Lemma 21, ρ' is λ^*-consistent. Thanks to Proposition 31, there exists an SPE τ such that $(\text{Cost}_i(\langle\tau\rangle_{x_0}))_{i\in\Pi} = (\text{Cost}_i(\rho'))_{i\in\Pi}$ with memory $\mathcal{O}(2^{|\Pi|} \cdot |\Pi| \cdot |V|^{(|\Pi|+2)\cdot(|V|+3)+1})$.

  Let $y \in (\mathbb{N} \cup \{+\infty\})^{|\Pi|}$, let us assume that $(\text{Cost}_i(\langle\sigma\rangle_{x_0}))_{i\in\Pi} \leq y$, by Lemma 19 we have that $(\text{Cost}_i(\langle\tau\rangle_{x_0})) = (\text{Cost}_i(\rho'))_{i\in\Pi} \leq y$.

  Let $k \in \{0, \dots, |\Pi|\}$ and $c \in \mathbb{N} \cup \{+\infty\}$, let us assume that $\text{SW}(\langle\sigma\rangle_{x_0}) \succeq (k, c)$. If $\text{SW}(\langle\sigma\rangle_{x_0}) = (k_1, c_1)$ and $\text{SW}(\langle\tau\rangle_{x_0}) = (k_2, c_2)$, we have $k_1 = k_2$ and $c_2 \leq c_1 \leq c$ thanks to Lemma 19. Thus, we have that $\text{SW}(\langle\tau\rangle_{x_0}) \succeq (k, c)$.

□

Finally, we prove Proposition 23. Recall that this proposition states that if there exists an NE σ (resp. SPE) whose outcome is Pareto optimal, then one can construct another one, τ, such that its outcome is a lasso of polynomial length, has the same cost as σ (thus is also Pareto optimal), and τ uses finite-memory strategies with polynomial (resp. exponential) size.

Proof of Proposition 23. The second and third items are a direct consequence of the first one. Thus, let us prove the first item.

Let σ be an NE such that its cost profile is Pareto optimal in $\text{Plays}(v_0)$. To get a contradiction, assume that there exists $i \in \text{Visit}(\langle\sigma\rangle_{v_0})$ such that $\text{Cost}_i(\langle\sigma\rangle_{v_0}) > |V| \cdot |\Pi|$. It means that there exists an unnecessary cycle before Player i reaches his target set. By removing this cycle (applying (P1)), we obtain a new play ρ' such that $\mathrm{Cost}_i(\rho') < \mathrm{Cost}_i(\langle\sigma\rangle_{v_0})$ and for Player j ($j \neq i$), $\mathrm{Cost}_j(\rho') \leq \mathrm{Cost}_j(\langle\sigma\rangle_{v_0})$ (by Lemma 19). It leads to a contradiction with the fact that $(\mathrm{Cost}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi}$ is Pareto optimal in $\mathrm{Plays}(v_0)$.

The same proof holds for SPE. □

5. The qualitative setting

In the previous sections, we have considered quantitative reachability problems. In this section we consider the qualitative variant and investigate the difference with the previously obtained results.

5.1. Qualitative reachability games

All along this section we focus on *qualitative reachability games*. Unlike quantitative reachability games, the arena is equipped with a gain function profile $(\mathrm{Gain}_i)_{i\in\Pi}$ such that for all $i \in \Pi$, $\mathrm{Gain}_i : \mathrm{Plays} \to \{0, 1\}$ is a *gain function* which assigns a gain 0 or 1 to each play ρ for Player i. We also say that the play ρ has gain profile $(\mathrm{Gain}_i(\rho))_{i\in\Pi}$ and similarly if we consider the outcome of the strategy profile σ from v_0, we say that σ has gain profile $(\mathrm{Gain}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi}$.

Definition 34. A *qualitative reachability game* $\mathcal{G} = (\mathcal{A}, (\mathrm{Gain}_i)_{i\in\Pi}, (F_i)_{i\in\Pi})$ is a game enhanced with a target set $F_i \subseteq V$. For all $i \in \Pi$, the gain function Gain_i is defined as follows: for all $\rho = \rho_0\rho_1 \ldots \in \mathrm{Plays}$: $\mathrm{Gain}_i(\rho) = 1$ if there exists $k \in \mathbb{N}$ such that $\rho_k \in F_i$ and $\mathrm{Gain}_i(\rho) = 0$ otherwise.

In this particular setting, players only aim at reaching their target set but do not take into account the number of steps it takes. Player i receives a gain of 1 if ρ visits his target set F_i, and a gain of 0 otherwise. Thus each player i wants to maximize his gain.

For qualitative reachability games, it is easy to adapt the definitions of NE and SPE defined in Section 2.3 by reversing the inequality and replacing cost functions by gain functions, as players want to maximize their gain instead of minimizing their cost. This leads to the following Lemma.

Lemma 35. *Let $(\mathcal{G}, v_0)$ be an initialized quantitative reachability game and σ be a strategy profile. Consider the related qualitative reachability game $\mathcal{G}'$ with the same arena $\mathcal{A}$ and target sets $(F_i)_{i\in\Pi}$, but the gain functions $(\mathrm{Gain}_i)_{i\in\Pi}$. Then if σ is an NE (resp. SPE) in $(\mathcal{G}, v_0)$, then σ is also an NE (resp. SPE) in $(\mathcal{G}', v_0)$.*

Thus, as it is proved that there always exists an SPE (and thus an NE) in a quantitative reachability game, there always exists one in a qualitative reachability game.

Theorem 36. *In every initialized qualitative reachability game, there always exists an SPE, and thus also an NE.*

In case of qualitative reachability, as for quantitative reachability games, we are interested in a solution that fulfills certain requirements. For example, we would like to know whether there exists a solution such that a maximum number of players visit their target sets.

Let $(\mathcal{G}, v_0)$ be an initialized qualitative reachability game with $\mathcal{G} = (\mathcal{A}, (\text{Gain}_i)_{i\in\Pi}, (F_i)_{i\in\Pi})$. Given $\rho \in \text{Plays}(v_0)$, we denote by $\text{Visit}(\rho)$ the set of players i such that ρ visits F_i, that is, $\text{Visit}(\rho) = \{i \in \Pi \mid \text{Gain}_i(\rho) = 1\}$. The *social welfare* $\text{SW}(\rho)$ of ρ is the size of $\text{Visit}(\rho)$. Let $P \subseteq \{0,1\}^{|\Pi|}$ be the set of all gain profiles $p = (\text{Gain}_i(\rho))_{i\in\Pi}$, with $\rho \in \text{Plays}(v_0)$. A cost profile $p \in P$ is called *Pareto optimal in* $\text{Plays}(v_0)$ if it is maximal in P with respect to the componentwise ordering $\leq$ on P. Notice that if there exists ρ with $\text{Visit}(\rho) = \Pi$, then its social welfare is the largest possible and there exists a unique Pareto optimal gain profile equal to $(1, 1, \ldots, 1)$. Notice also that certain target sets F_i might be empty or not reachable from the initial vertex v_0. Hence in this case, the best that we can hope is a (unique) Pareto optimal gain profile p such that $p_i = 1$ if and only if F_i is reachable[6] from v_0.

Qualitative variant of Problem 1. Given an initialized qualitative reachability game $(\mathcal{G}, v_0)$, given two thresholds $x, y \in \{0,1\}^{|\Pi|}$, decide whether there exists a solution σ such that $x \leq (\text{Gain}_i(\rho))_{i\in\Pi} \leq y$.

Imposing a lower bound $x_i = 1$ means that player i has to visit his target set whereas imposing an upper bound $y_i = 0$ means that player i cannot visit his target set.

Unlike quantitative reachability, social welfare in qualitative reachability games only aims to maximize the number of players who visit their target set.

Qualitative variant of Problem 2. Given an initialized qualitative reachability game $(\mathcal{G}, v_0)$, given a threshold $k \in \{0, \ldots, |\Pi|\}$, decide whether there exists a solution σ such that $\text{SW}(\langle\sigma\rangle_{v_0}) \geq k$.

Let us now state the last studied problem for qualitative reachability games.

Qualitative variant of Problem 3. Given an initialized qualitative reachability game $(\mathcal{G}, v_0)$ decide whether there exists a solution σ in $(\mathcal{G}, v_0)$ such that $(\text{Gain}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi}$ is Pareto optimal in $\text{Plays}(v_0)$.

The latter problem has some connections with the two previous ones. For instance in case of qualitative reachability, suppose there exists a play in $\text{Plays}(v_0)$ that visits all target sets. As already explained, there is only one Pareto optimal gain $(1, \ldots, 1)$. Asking for the existence of a solution σ such that $(\text{Gain}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi}$ is Pareto optimal is equivalent to asking for the existence of a solution σ such that $\text{Gain}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi} \geq (1, \ldots, 1)$ (see Qualitative variant of Problem 1), or such that $\text{SW}(\langle\sigma\rangle_{v_0}) \geq |\Pi|$ (see Qualitative variant of Problem 2).

We can now state the qualitative variant of Theorem 11.

Theorem 37. *Let $(\mathcal{G}, v_0)$ be a qualitative reachability game.*

[6] Notice that if F_i is reachable from v_0, then it is necessarily not empty.

- *For NEs: the Qualitative variants of Problem 1 and Problem 2 are NP-complete while the Qualitative variant of Problem 3 is NP-hard and belongs to* Σ_2^P*.*
- *For SPEs: the Qualitative variants of Problems 1, 2 and 3 are PSPACE-complete.*

5.3. Existence problem

The following Theorem is a direct consequence of Theorem 7 and Lemma 35.

Theorem 38. *Let* $(\mathcal{G}, v_0)$ *be an initialized qualitative reachability game such that its arena* $\mathcal{A}$ *is strongly connected. Then there exists an SPE* σ *(and thus an NE) such that its outcome* $\langle\sigma\rangle_{v_0}$ *visits all target sets* F_i*,* $i \in \Pi$*, that are non-empty.*

Let us comment on this result. For this family of games, the answer to the Qualitative variant of Problem 1 is always positive for particular thresholds. Take thresholds x, y such that $x_i = 1$ (and thus $y_i = 1$) if and only if $F_i \neq \emptyset$. The answer to the Qualitative variant of Problem 2 is also always positive for threshold $k = |\{i \mid F_i \neq \emptyset\}|$. Finally, the answer to the Qualitative variant of Problem 3 is also always positive since there exists an unique Pareto optimal gain profile p such that $p_i = 1$ if and only if $F_i \neq \emptyset$.

Recall that we explained before why it was enough to prove Theorem 7 for SPEs and for quantitative reachability games only. Notice that in case of qualitative reachability games, there exists a simpler construction of the required NE or SPE. Indeed, as the arena is strongly connected, there exists a play $\rho \in \text{Plays}(v_0)$ that visits all non-empty target sets. *(i)* Hence to get an NE, construct a strategy profile σ in $(\mathcal{G}, v_0)$ such that $\langle\sigma\rangle_{v_0} = \rho$. As the gain profile of σ is the best that each player can hope, no player has an incentive to deviate and σ is then an NE. *(ii)* The construction is a little more complex to get an SPE. We again construct a strategy profile σ in $(\mathcal{G}, v_0)$ such that $\langle\sigma\rangle_{v_0} = \rho$, and inductively extend its construction to all subgames $(\mathcal{G}_{\restriction h}, v)$ as follows. Assume that $\sigma_{\restriction h}$ is not yet constructed, then extend the construction of σ such that $\sigma_{\restriction h} = g\rho$ for some gv_0 starting in v and ending in v_0 (such a history gv_0 exists because the arena is strongly connected). In this way, the outcome of $\sigma_{\restriction h}$ in each subgame $(\mathcal{G}, v)$ has gain profile $(1, \ldots, 1)$ and no player has an incentive to deviate. It follows that σ is an SPE.

The next theorem states that the Qualitative variant of Problem 3 has a positive answer for all *qualitative* reachability games with a number of players *limited to two*, and that this existence result cannot be extended to three players.

Theorem 39. *Let* $(\mathcal{G}, v_0)$ *be an initialized qualitative reachability game,*

- *Let* $(\mathcal{G}, v_0)$ *be an initialized qualitative reachability game such that* $|\Pi| = 2$*, there exists an SPE* σ *(and thus an NE) with a gain profile that is Pareto optimal in* $\text{Plays}(v_0)$*.*
- *There exists an initialized qualitative reachability games with* $|\Pi| = 3$ *that has no NE with a gain profile that is Pareto optimal in* $\text{Plays}(v_0)$*.*

Let us focus on the proof of Theorem 39 which is based on the next lemma, which is interesting in its own right.

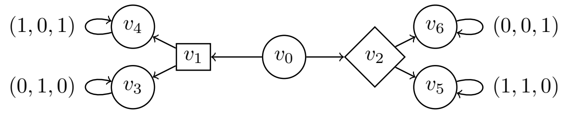

Figure 4: A qualitative reachability game that has no NE with a gain profile that is Pareto optimal

Lemma 40. *Let $(\mathcal{G}, v_0)$ be an initialized qualitative reachability game. Let p be a gain profile equal to $(0, 0, \ldots, 0)$ or $(1, 1, \ldots, 1)$. If p is Pareto optimal*[7] *in* $\text{Plays}(v_0)$*, then there exists an SPE σ with gain profile p.*

Proof. The case $p = (0, 0, \ldots, 0)$ is easy to solve. By Pareto optimality, all plays in $\text{Plays}(v_0)$ have gain profile p. Hence every strategy profile σ is trivially an SPE with gain profile p. Let us turn to case $p = (1, 1, \ldots, 1)$ and let $\rho = \rho_0\rho_1 \ldots \in \text{Plays}(v_0)$ with gain profile p. By Theorem 2.1 in [17][8], there exists an SPE σ in $(\mathcal{G}, v_0)$. If $(\text{Gain}_i(\langle\sigma\rangle_{v_0}))_{i\in\Pi} = p$, we are done. Otherwise let us show how to modify σ into another SPE τ with outcome ρ and thus with gain profile p. Let $h \in \text{Hist}_i(v_0)$, $i \in \Pi$,

- if h is a prefix of ρ, then $\tau_i(h) = \rho_{|h|+1}$,
- otherwise, $\tau_i(h) = \sigma_i(h)$.

Let us prove that τ is an SPE. Clearly for each history hv that is not a prefix of ρ, $\tau_{\restriction h} = \sigma_{\restriction h}$ is an NE in the subgame $(\mathcal{G}_{\restriction h}, v)$. So let $hv = \rho_0 \ldots \rho_k$. As $\langle\tau_{\restriction h}\rangle_v$ has gain profile $(1, 1, \ldots, 1)$ in $(\mathcal{G}_{\restriction h}, v)$, player i such that $v \in V_i$ has no incentive to deviate, and then $\tau_{\restriction h}$ is also an NE in $(\mathcal{G}_{\restriction h}, v)$. □

Proof of Theorem 39. We begin with the first item. There are three cases to study: either the unique Pareto optimal gain profile of $\text{Plays}(v_0)$ is equal to $(0, 0)$, or it is equal to $(1, 1)$, or there are one or two Pareto optimal gain profiles that belong to $\{(0, 1), (1, 0)\}$. In the first two cases, we get the required SPE by Lemma 40. Hence it remains to treat the last case. From Lemma 10, we know that there exists an SPE in $(\mathcal{G}, v_0)$ whose outcome ρ visits a least one target set F_i, $i \in \{1, 2\}$. Therefore the gain profile of ρ is either equal to $(0, 1)$ or $(1, 0)$ as required.

For the second item, consider the initialized qualitative reachability game $(\mathcal{G}, v_0)$ of Figure 4. We have three players such that player 3 owns diamond vertices. Moreover, $F_1 = \{v_4, v_5\}$, $F_2 = \{v_3, v_5\}$, and $F_3 = \{v_4, v_6\}$. There are four plays in $\text{Plays}(v_0)$ whose gain profile is indicated below each of them. The set of Pareto optimal gain profiles in $\text{Plays}(v_0)$ is equal to $\{(1, 0, 1), (1, 1, 0)\}$. Consider a strategy profile σ with outcome $v_0v_1v_4^\omega$ and gain profile $(1, 0, 1)$. Then it is not an NE because player 2 has a profitable deviation by going from v_1 to v_3 (instead of v_4). Similarly the strategy profile σ with outcome $v_0v_2v_5^\omega$ and gain profile $(1, 1, 0)$ is not an NE. Therefore there is no NE in $(\mathcal{G}, v_0)$ with a gain profile that is Pareto optimal. □

[7] $(1, 1, \ldots, 1)$ is trivially Pareto optimal.

[8] Notice that we cannot apply Theorem 7 since the arena is not necessarily strongly connected.

6. Conclusion

In the present paper, we have considered multiplayer (both qualitative and quantitative) reachability games, with a focus on two concepts of equilibrium: NE and SPE. It is well-known that NEs and SPEs are guaranteed to exist in reachability games in both qualitative and quantitative settings. Here we have investigated three decision problems about the existence of what we have called relevant equilibria. More precisely we have considered the Threshold problem (Problem 1), the Social welfare problem (Problem 2), and the Pareto optimal problem (Problem 3). For the three problems we provided their complexity class (summarized in Table 1). Let us notice that our results for NEs heavily rely on a characterization of plays which are NE outcomes. Such a characterization is not new but was missing for quantitative reachability games. The results concerning SPEs rely on a characterisation of SPE outcomes that has been recently obtained in [8]. In the case where a relevant equilibrium exists (for the three variants of relevant equilibrium), we have described the size of the finite-memory strategies needed in the equilibrium (summarized in Table 2). In this quest of finding relevant equilibria, we finally identified a subclass of reachability games in which there always exists an SPE where each player reaches his target set. Future work could include an extension of the present results to richer objectives and to other concepts of equilibrium.

References

[1] A. Pnueli, R. Rosner, On the synthesis of a reactive module, in: POPL Proceedings, ACM Press, 1989, pp. 179–190.

[2] J. F. Nash, Equilibrium points in n-person games, in: PNAS, Vol. 36, National Academy of Sciences, 1950, pp. 48–49.

[3] M. Osborne, An introduction to game theory, Oxford Univ. Press, 2004.

[4] M. Ummels, The complexity of Nash equilibria in infinite multiplayer games, in: FOSSACS Proceedings, Vol. 4962 of Lecture Notes in Computer Science, Springer, 2008, pp. 20–34.

[5] M. Ummels, Rational behaviour and strategy construction in infinite multiplayer games, in: FSTTCS Proceedings, Vol. 4337 of Lecture Notes in Computer Science, Springer, 2006, pp. 212–223.

[6] R. Condurache, E. Filiot, R. Gentilini, J.-F. Raskin, The Complexity of Rational Synthesis, in: ICALP Proceedings, Vol. 55 of Leibniz International Proceedings in Informatics, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016, pp. 121:1–121:15.

[7] T. Brihaye, V. Bruyère, A. Goeminne, J. Raskin, Constrained existence problem for weak subgame perfect equilibria with ω-regular boolean objectives, in: GandALF, Vol. 277 of EPTCS, 2018, pp. 16–29.

[8] T. Brihaye, V. Bruyère, A. Goeminne, J. Raskin, M. van den Bogaard, The complexity of subgame perfect equilibria in quantitative reachability games, Log. Methods Comput. Sci. 16 (4) (2020).

[9] V. Bruyère, Computer aided synthesis: A game-theoretic approach, in: DLT Proceedings, Vol. 10396 of Lecture Notes in Computer Science, Springer, 2017, pp. 3–35.

[10] V. Conitzer, T. Sandholm, Complexity results about Nash equilibria, in: IJCAI Proceedings, Morgan Kaufmann Publishers Inc., 2003, pp. 765–771.

[11] P. Bouyer, N. Markey, D. Stan, Mixed Nash equilibria in concurrent terminal-reward games, in: FSTTCS Proceedings, Vol. 29 of Leibniz International Proceedings in Informatics, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2014, pp. 351–363.

[12] R. Brenguier, J. Raskin, Pareto curves of multidimensional mean-payoff games, in: CAV Proceedings, Vol. 9207 of Lecture Notes in Computer Science, Springer, 2015, pp. 251–267.

[13] T. Brihaye, J. De Pril, S. Schewe, Multiplayer cost games with simple Nash equilibria, in: LFCS Proceedings, Vol. 7734 of Lecture Notes in Computer Science, Springer, 2013, pp. 59–73.

[14] A. Haddad, Characterising Nash equilibria outcomes in fully informed concurrent games, available at http://web1.ulb.ac.be/di/verif/haddad/H16.pdf.

[15] T. Brihaye, V. Bruyère, A. Goeminne, N. Thomasset, On relevant equilibria in reachability games, in: RP Proceedings, Vol. 11674 of Lecture Notes in Computer Science, Springer, 2019, pp. 48–62.

[16] E. Grädel, W. Thomas, T. Wilke (Eds.), Automata, Logics, and Infinite Games: A Guide to Current Research, Vol. 2500 of Lecture Notes in Computer Science, Springer, 2002.

[17] T. Brihaye, V. Bruyère, J. De Pril, H. Gimbert, On subgame perfection in quantitative reachability games, Logical Methods in Computer Science 9 (1) (2012).

[18] M. J. Osborne, A. Rubinstein, A course in game theory, MIT Press, Cambridge, MA, 1994.

[19] T. Brihaye, G. Geeraerts, A. Haddad, B. Monmege, Pseudopolynomial iterative algorithm to solve total-payoff games and min-cost reachability games, Acta Inf. 54 (1) (2017) 85–125.

[20] L. Khachiyan, E. Boros, K. Borys, K. Elbassioni, V. Gurvich, G. Rudolf, J. Zhao, On short paths interdiction problems: Total and node-wise limited interdiction, Theory of Computing Systems 43 (2) (2008) 204–233.

[21] T. Brihaye, V. Bruyère, N. Meunier, J. Raskin, Weak subgame perfect equilibria and their application to quantitative reachability, in: CSL Proceedings, Vol. 41 of Leibniz International Proceedings in Informatics, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2015, pp. 504–518.

[22] V. Bruyère, S. L. Roux, A. Pauly, J. Raskin, On the existence of weak subgame perfect equilibria, in: FOSSACS Proceedings, Vol. 10203 of Lecture Notes in Computer Science, 2017, pp. 145–161.