Noname manuscript No.
(will be inserted by the editor)

On Equilibria in Quantitative Games with Reachability/Safety Objectives

Thomas Brihaye · Véronique Bruyère · Julie De Pril

the date of receipt and acceptance should be inserted later

Abstract In this paper, we study turn-based quantitative multiplayer non zero-sum games played on finite graphs with both reachability and safety objectives. In this framework a player with a reachability objective aims at reaching his own goal as soon as possible, whereas a player with a safety objective aims at avoiding his bad set or, if impossible, delaying its visit as long as possible. We prove the existence of Nash equilibria with finite memory in quantitative multiplayer reachability/safety games. Moreover, we prove the existence of finite-memory secure equilibria for quantitative two-player reachability games.

Keywords Nash equilibrium, Turn-based quantitative game, Secure equilibrium, Reachability/Safety objectives

1 Introduction

General framework. The construction of correct and efficient computer systems (hardware or software) is recognized as an extremely difficult task. To support the design and verification of such systems, mathematical logic, automata theory [10] and more recently model-checking [7] have been intensively studied. The model-checking approach, which is now an important part of the design cycle in industries, has proved its efficiency when applied to systems that can be accurately modeled as a finite-state automaton. In contrast, the application of these techniques to computer software, complex systems like embedded systems or distributed systems has been less successful. This could be partly explained by the following reasons: classical automata-based models do not faithfully capture the complex interactive behavior of modern computational systems that are usually composed of several interacting components, also interacting with an environment that is only partially under control. Recent research works show that it is suitable to generalize automata models used in the classical approach to verification, with the more flexible and mathematically deeper game-theoretic framework [14, 15].

University of Mons - UMONS
Place du Parc 20, 7000 Mons, Belgium
`{thomas.brihaye,veronique.bruyere,julie.depril}@umons.ac.be`

Game theory meets automata theory. The basic framework that extends computational models with concepts from game theory is the so-called two-player zero-sum games played on graphs [8]. Many problems in verification and design of reactive systems can be modeled with this approach, like modeling controller-environment interactions. Given a model of a system interacting with a hostile environment, given a control objective (like preventing the system to reach some bad configurations), the controller synthesis problem asks to build a controller ensuring that the control objective is enforced whatever the environment will do. Two-player zero-sum games played on graphs are adequate models to solve this problem [16]. Moves of Player 1 model actions of the controller whereas moves of Player 2 model the uncontrollable actions of the environment, and a winning strategy for Player 1 is an abstract form of a control program that enforces the control objective.

The controller synthesis problem is suitable to model purely antagonist interactions between a controller and a hostile environment. However in order to study more complex systems with more than two components whose objectives are not necessarily antagonist, we need multiplayer and non zero-sum games to model them adequately. Moreover, we do not look for winning strategies, but rather try to find relevant notions of equilibria, for instance the famous notion of Nash equilibria [14]. On the other hand, only qualitative objectives have been considered so far to specify, for example, that a player must be able to reach a target set of states in the underlying game graph. But, in line with the previous point, we also want to express and solve games for quantitative objectives such as forcing the game to reach a particular set of states within a given time bound, or within a given energy consumption limit. In summary, we need to study *equilibria* for *multiplayer non zero-sum* games played on graphs with *quantitative* objectives. This article provides some new results in this research direction.

Related work. Several recent papers have considered two-player zero-sum games played on finite graphs with regular objectives enriched by some *quantitative aspects*. Let us mention some of them: games with *finitary objectives* [6], games with *prioritized requirements* [1], *request-response games* where the waiting times between the requests and the responses are minimized [11,17], and games whose winning conditions are expressed via *quantitative languages* [2].

Other works concern qualitative non zero-sum games. The notion of secure equilibrium, an interesting refinement of Nash equilibrium, has been introduced in [5]. It has been proved that a unique secure equilibrium always exists for two-player non zero-sum games with regular objectives. In [9], general criteria ensuring existence of Nash equilibria, subgame perfect equilibria (resp. secure equilibria) are provided for n-player (resp. 2-player) games, as well as complexity results.

Finally, we mention reference [3] that combines both quantitative and non zero-sum aspects. It is maybe the nearest related work compared to us, however the framework and the objectives are pretty different. In [3], the authors study games played on graphs with terminal vertices where quantitative payoffs are assigned to the players. These games may have cycles but all the infinite plays form a single outcome (like in chess where every infinite play is a draw). That paper gives criteria that ensure the existence of Nash (and subgame perfect) equilibria in pure and memoryless strategies.

Our contribution. We here study turn-based quantitative multiplayer non zero-sum games played on finite graphs with reachability objectives. In this framework each player aims at reaching his own goal as soon as possible. We focus on existence results for two solution concepts: Nash equilibrium and secure equilibrium. We prove the existence of finite-memory Nash (resp. secure) equilibria in n-player (resp. 2-player) games. Moreover, we prove that given a Nash (resp. secure) equilibrium of a n-player (resp. 2-player) game, we can build a finite-memory Nash (resp. secure) equilibrium of the *same type*, i.e. preserving the set of players achieving their objectives. For the case of Nash equilibria, we extend our results in two directions. First we prove that finite-memory Nash equilibria still exist when the model is enriched by allowing n-tuples of non-negative costs on edges (one cost by player). This result provides an answer to a question we posed in [4]. Secondly, we prove the existence of Nash equilibria in quantitative games where both safety and reachability objectives coexist.

Our results are not a direct consequence of the existing results in the qualitative framework, they require some new proof techniques. To the best of our knowledge, this is the first general result about the existence of equilibria in quantitative multiplayer games played on graphs.

Organization of the paper. Section 2 is dedicated to definitions. We present the games and the equilibria we study. In Section 3 we first prove an existence result for Nash equilibria and provide the finite-memory characterization. Similar results concerning secure equilibria in two-player games are established in Section 4. Finally, in Section 5, we discuss the extensions of our results on Nash equilibria.

A part of these results has been published in [4], namely the existence of finite-memory Nash (resp. secure) equilibria in multiplayer (resp. 2-player) games, and the fact that given a Nash equilibrium we can build a finite-memory Nash equilibrium of the same type. Additionally in this paper we give proofs of the previous results and we extend our existence result for Nash equilibria in the two directions mentioned above, namely (i) n-tuples of non-negative costs on edges and (ii) reachability/safety objectives. Moreover, in the two-player case, we prove that given a secure equilibrium, we can build a finite-memory secure equilibrium of the same type.

2 Preliminaries

2.1 Definitions

We consider here *quantitative* games played on a graph where all the players have *reachability*[1] *objectives*. It means that, given a certain set of vertices Goal_i, each player i wants to reach one of these vertices as soon as possible.

This section is mainly inspired by reference [9].

Definition 1 An *infinite turn-based quantitative multiplayer reachability game* is a tuple $\mathcal{G} = (\Pi, V, (V_i)_{i\in\Pi}, v_0, E, (\mathsf{Goal}_i)_{i\in\Pi})$ where

- Π is a finite set of players,
- $G = (V, (V_i)_{i\in\Pi}, v_0, E)$ is a finite directed graph where V is the set of vertices, $(V_i)_{i\in\Pi}$ is a partition of V into the state sets of each player, $v_0 \in V$ is the initial vertex, and $E \subseteq V \times V$ is the set of edges, and
- $\mathsf{Goal}_i \subseteq V$ is the goal set of player i.

[1] The general case of reachability/safety objectives is handled in Subsection 5.1.

We assume that each vertex has at least one outgoing edge. The game is played as follows. A token is first placed on the vertex v_0. Player i, such that $v_0 \in V_i$, has to choose one of the outgoing edges of v_0 and put the token on the vertex v_1 reached when following this edge. Then, it is the turn of the player who owns v_1. And so on.

A *play* $\rho \in V^\omega$ (resp. a *history* $h \in V^+$) of $\mathcal{G}$ is an *infinite* (resp. a *finite*) path through the graph G starting from vertex v_0. Note that a history is always non empty because it starts with v_0. The set $H \subseteq V^+$ is made up of all the histories of $\mathcal{G}$. A *prefix* (resp. *proper prefix*) $\mathfrak{p}$ of a history $h = h_0 \dots h_k$ is a finite sequence $h_0 \dots h_l$, with $l \leq k$ (resp. $l < k$), denoted by $\mathfrak{p} \leq h$ (resp. $\mathfrak{p} < h$). We similarly consider a prefix $\mathfrak{p}$ of a play ρ, denoted by $\mathfrak{p} < \rho$.

We say that a play $\rho = \rho_0\rho_1 \dots$ *visits* a set $S \subseteq V$ (resp. a vertex $v \in V$) if there exists $l \in \mathbb{N}$ such that ρ_l is in S (resp. $\rho_l = v$). The same terminology also stands for a history h. Similarly, we say that ρ *visits* S *after* (resp. *in*) *a prefix* $\rho_0 \dots \rho_k$ if there exists $l > k$ (resp. $l \leq k$) such that ρ_l is in S. For any play ρ we denote by $\mathsf{Visit}(\rho)$ the set of players $i \in \Pi$ such that ρ visits Goal_i. The set $\mathsf{Visit}(h)$ for a history h is defined similarly. The function Last returns, given a history $h = h_0 \dots h_k$, the last vertex h_k of h, and the *length* $|h|$ of h is the number k of its *edges*[2].

For any play $\rho = \rho_0\rho_1 \dots$ of $\mathcal{G}$, we note $\mathsf{Cost}_i(\rho)$ the *cost* of player i, defined by:

$$\mathsf{Cost}_i(\rho) = \begin{cases} l & \text{if } l \text{ is the } \textit{least} \text{ index such that } \rho_l \in \mathsf{Goal}_i, \\ +\infty & \text{otherwise.} \end{cases}$$

We note $\mathsf{Cost}(\rho) = (\mathsf{Cost}_i(\rho))_{i\in\Pi}$ the *cost profile* for the play ρ. The aim of each player i is to *minimize* the cost he has to pay, i.e. reach his goal set Goal_i as soon as possible.

A *strategy* of player i in $\mathcal{G}$ is a function $\sigma : V^*V_i \to V$ assigning to each history hv ending in a vertex v of player i, a next vertex $\sigma(hv)$ such that $(v, \sigma(hv))$ belongs to E. We say that a play $\rho = \rho_0\rho_1 \dots$ of $\mathcal{G}$ is *consistent* with a strategy σ of player i if $\rho_{k+1} = \sigma(\rho_0 \dots \rho_k)$ for all $k \in \mathbb{N}$ such that $\rho_k \in V_i$. The same terminology is used for a history h of $\mathcal{G}$. A *strategy profile* of $\mathcal{G}$ is a tuple $(\sigma_i)_{i\in\Pi}$ where σ_i is a strategy for player i. It determines a unique play of $\mathcal{G}$ consistent with each strategy σ_i, called the *outcome* of $(\sigma_i)_{i\in\Pi}$ and denoted by $\langle(\sigma_i)_{i\in\Pi}\rangle$.

A strategy σ of player i is *memoryless* if σ depends only on the current vertex, i.e. $\sigma(hv) = \sigma(v)$ for all $h \in H$ and $v \in V_i$. More generally, σ is a *finite-memory strategy* if the equivalence relation $\approx_\sigma$ on H defined by $h \approx_\sigma h'$ if $\sigma(h\delta) = \sigma(h'\delta)$ for all $\delta \in V^*V_i$ has finite index. In other words, a finite-memory strategy is a strategy that can be implemented by a finite automaton with output. A strategy profile $(\sigma_i)_{i\in\Pi}$ is called *memoryless* or *finite-memory* if each σ_i is a memoryless or a finite-memory strategy, respectively.

For a strategy profile $(\sigma_i)_{i\in\Pi}$ with outcome ρ and a strategy σ'_j of player j ($j \in \Pi$), we say that *player* j *deviates from* ρ *after a prefix* h *of* ρ if there exists a prefix h' of ρ such that $h \leq h'$, h' is consistent with σ'_j and $\sigma'_j(h') \neq \sigma_j(h')$. We also say that *player* j *deviates from* ρ *just after a prefix* h *of* ρ if h is consistent with σ'_j and $\sigma'_j(h) \neq \sigma_j(h)$.

We now introduce the notion of *Nash equilibrium* and *secure equilibrium*.

Definition 2 A strategy profile $(\sigma_i)_{i\in\Pi}$ of a game $\mathcal{G}$ is a *Nash equilibrium* if for all player $j \in \Pi$ and for all strategy σ'_j of player j, we have:

$$\mathsf{Cost}_j(\rho) \leq \mathsf{Cost}_j(\rho')$$

where $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle$ and $\rho' = \langle\sigma'_j, (\sigma_i)_{i\in\Pi\setminus\{j\}}\rangle$.

[2] Note that the length is not defined as the number of vertices.

This definition means that player j (for all $j \in \Pi$) has no incentive to deviate since he increases his cost when using σ'_j instead of σ_j. Keeping notations of Definition 2 in mind, a strategy σ'_j such that $\mathsf{Cost}_j(\rho) > \mathsf{Cost}_j(\rho')$ is called a *profitable deviation* for player j with respect to $(\sigma_i)_{i\in\Pi}$. In this case either player j pays an infinite cost for ρ and a finite cost for ρ' (ρ' visits Goal_j, but ρ does not), or player j pays a finite cost for ρ and a strictly lower cost for ρ' (ρ' visits Goal_j earlier than ρ does).

As our results on secure equilibria stand for two-player games, we define this notion only in this context. In order to define the concept of secure equilibrium[3] we first need to associate two appropriate binary relations $\prec_1$ and $\prec_2$ on cost profiles with player 1 and 2 respectively. Given two cost profiles (x_1, x_2) and (y_1, y_2):

$$(x_1, x_2) \prec_1 (y_1, y_2) \quad \text{iff} \quad (x_1 > y_1) \vee (x_1 = y_1 \wedge x_2 < y_2)\,.$$

We then say that *player* 1 *prefers* (y_1, y_2) *to* (x_1, x_2). In other words, player 1 prefers a cost profile to another either if he can decrease his own cost, or if he can increase the cost of player 2, while keeping his own cost. We define the relation $\prec_2$ symmetrically.

Definition 3 A strategy profile (σ_1, σ_2) of a two-player game $\mathcal{G}$ is a *secure equilibrium* if there does not exist any strategy σ'_1 of player 1 such that:

$$\mathsf{Cost}(\rho) \prec_1 \mathsf{Cost}(\rho')$$

where $\rho = \langle \sigma_1, \sigma_2 \rangle$ and $\rho' = \langle \sigma'_1, \sigma_2 \rangle$, and there does not exist any strategy σ'_2 of player 2 such that:

$$\mathsf{Cost}(\rho) \prec_2 \mathsf{Cost}(\rho')$$

where $\rho = \langle \sigma_1, \sigma_2 \rangle$ and $\rho' = \langle \sigma_1, \sigma'_2 \rangle$.

In other words, player 1 (resp. 2) has no incentive to deviate, with respect to the relation $\prec_1$ (resp. $\prec_2$). Note that any secure equilibrium is a Nash equilibrium. A strategy σ'_j such that $\mathsf{Cost}(\rho) \prec_j \mathsf{Cost}(\rho')$ is called a $\prec_j$*-profitable deviation* for player j with respect to (σ_1, σ_2) (for $j \in \{1, 2\}$).

Let us go back to the multiplayer framework and define the notion of type of an equilibrium.

Definition 4 The type of a strategy profile $(\sigma_i)_{i\in\Pi}$ in a reachability game $\mathcal{G}$ is the set of players $j \in \Pi$ such that the outcome ρ of $(\sigma_i)_{i\in\Pi}$ visits Goal_j. It is denoted by $\mathtt{Type}((\sigma_i)_{i\in\Pi})$.

In other words, $\mathtt{Type}((\sigma_i)_{i\in\Pi}) = \mathsf{Visit}(\rho)$.

The previous definitions are illustrated in the following example.

Example 5 Let $\mathcal{G} = (V, V_1, V_2, v_0, E, \mathsf{Goal}_1, \mathsf{Goal}_2)$ be the two-player game depicted in Figure 1. The states of player 1 (resp. 2) are represented by circles (resp. squares)[4]. Thus, according to Figure 1, $V_1 = \{A, C, D\}$ and $V_2 = \{B\}$, the initial vertex v_0 is the vertex A, and we set $\mathsf{Goal}_1 = \{C\}$ and $\mathsf{Goal}_2 = \{D\}$.

An example of play in $\mathcal{G}$ is given by $\rho = (AD)^\omega$, which visits Goal_2 but not Goal_1, leading to the cost profile $\mathsf{Cost}((AD)^\omega) = (+\infty, 1)$. The play ρ is, among others, the

[3] Our definition naturally extends the notion of *secure equilibrium* proposed in [5] to the quantitative reachability framework. A longer discussion comparing the two notions can be found in Section 2.2.

[4] We will keep this convention through the article.

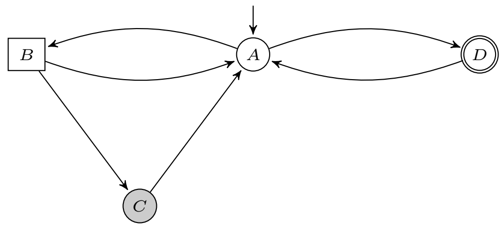

Fig. 1 A two-player game with $\mathsf{Goal}_1 = \{C\}$ and $\mathsf{Goal}_2 = \{D\}$.

outcome of the strategy[5] profile (σ_1, σ_2) where $\sigma_1(hA) = D$ and $\sigma_2(hB) = C$, for all histories h.

Let us show that the strategy profile (σ_1, σ_2) is not a Nash equilibrium, by proving that player 1 has a profitable deviation σ'_1 in which he manages to decrease his own cost. With σ'_1 defined by $\sigma'_1(hA) = B$, we get the play $\langle \sigma'_1, \sigma_2 \rangle = (ABC)^\omega$ such that $\mathsf{Cost}((ABC)^\omega) = (2, +\infty)$, and in particular $\mathsf{Cost}_1((ABC)^\omega) < \mathsf{Cost}_1(\rho)$.

On the opposite side, one can show that (σ'_1, σ_2) is a Nash equilibrium. However (σ'_1, σ_2) is not a secure equilibrium. Indeed, player 2 has a $\prec_2$-profitable deviation in which he can increase player 1's cost without modifying his own cost. With σ'_2 the strategy of player 2 defined by $\sigma'_2(hB) = A$, we get the play $\langle \sigma'_1, \sigma'_2 \rangle = (AB)^\omega$ such that $\mathsf{Cost}((AB)^\omega) = (+\infty, +\infty)$, and $\mathsf{Cost}(\langle \sigma'_1, \sigma_2 \rangle) \prec_2 \mathsf{Cost}(\langle \sigma'_1, \sigma'_2 \rangle)$.

Notice that all strategies discussed so far are memoryless. In order to obtain a Nash equilibrium of type $\{1, 2\}$, finite-memory strategies are necessary. We define the following finite-memory strategy profile (τ_1, τ_2):

$$\tau_1(hA) = \begin{cases} D & \text{if } h = \epsilon \\ B & \text{if } h \neq \epsilon \end{cases} \quad ; \quad \tau_2(hB) = \begin{cases} C & \text{if } h \text{ visits } D \\ A & \text{otherwise.} \end{cases}$$

The outcome $\pi = \langle (\tau_1, \tau_2) \rangle$ is equal to $AD(ABC)^\omega$ and has costs $(4, 1)$. In order to prove that (τ_1, τ_2) is a Nash equilibrium, we prove that no player has a profitable deviation. For player 2 it is clearly impossible to get a cost less than 1. To try to get a cost less than 4, player 1 must use a strategy τ'_1 such that $\tau'_1(A) = B$. But then player 2 chooses $\tau_2(AB) = A$. The prefix ABA of the outcome of (τ'_1, τ_2) shows that player 1 will increase his cost of 4.

However (τ_1, τ_2) is not a secure equilibrium since player 2 has a $\prec_2$-profitable deviation τ'_2 such that $\tau'_2(hB) = A$ for all histories h. One can show that, in this example, there is no secure equilibrium of type $\{1, 2\}$.

The questions studied in this article are the following ones:

Problem 1 *Given $\mathcal{G}$ a quantitative multiplayer (resp. two-player) reachability game, does there exist a Nash equilibrium (resp. a secure equilibrium) in $\mathcal{G}$?*

Problem 2 *Given a Nash equilibrium (resp. a secure equilibrium) in a quantitative multiplayer (resp. two-player) reachability game $\mathcal{G}$, does there exist a finite-memory Nash equilibrium (resp. secure equilibrium) with the same type?*

We provide positive answers in Sections 3 and 4. Notice that these problems have been investigated in the qualitative framework (see [9]).

[5] Note that player 1 has no choice in vertices C and D, that is, $\sigma_1(hv)$ is necessarily equal to A for $v \in \{C, D\}$.

2.2 Qualitative Games vs Quantitative Games

We show in this section that Problems 1 and 2 can not be reduced to problems on qualitative games.

Given a quantitative multiplayer reachability game $\mathcal{G}$, one can naturally define a *qualitative* version of $\mathcal{G}$, denoted by $\overline{\mathcal{G}}$, such that the payoffs[6] are *qualitative*. Given a play ρ of $\mathcal{G}$, the qualitative payoff of player i is defined by:

$$\mathsf{Payoff}_i(\rho) = \begin{cases} \mathsf{Win} & \text{if } \mathsf{Cost}_i(\rho) \text{ is finite} \\ \mathsf{Lose} & \text{otherwise.} \end{cases}$$

We note $\mathsf{Payoff}(\rho) = (\mathsf{Payoff}_i(\rho))_{i\in\Pi}$ the *qualitative* payoff profile for the play ρ. In this framework, player i aims at reaching his own goal set, i.e. at obtaining payoff Win. With this idea in mind, one can naturally adapt the notion of Nash (resp. secure) equilibrium to the qualitative framework.

The existence of Nash (resp. secure) equilibria in n-player (resp. 2-player) qualitative games $\overline{\mathcal{G}}$ has been proved in [9, Corollary 12] (resp. [5, Theorem 2]) for reachability objectives, and more generally for Borel objectives.

The next example illustrates that lifting Nash equilibria in $\overline{\mathcal{G}}$ to Nash equilibria in $\mathcal{G}$ does not work. We developed new ideas in Sections 3 and 4 to solve Problem 1.

Example 6 Let us now consider the two-player game $\mathcal{G}$ depicted in Figure 2, such that $\mathsf{Goal}_1 = \{B, E\}$ and $\mathsf{Goal}_2 = \{C\}$. Notice that only player 1 effectively plays in this game. We are going to exhibit a secure (and thus Nash) equilibrium (σ_1, σ_2) in the qualitative game $\overline{\mathcal{G}}$ that can not be lifted neither to a secure nor to a Nash equilibrium in the quantitative game $\mathcal{G}$. The strategy profile (σ_1, σ_2) is defined such that $\langle(\sigma_1, \sigma_2)\rangle = ADE^\omega$. It is a secure equilibrium in $\overline{\mathcal{G}}$ with the qualitative payoff profile $(\mathsf{Win}, \mathsf{Lose})$. However (σ_1, σ_2) is not a Nash (and thus not a secure) equilibrium in $\mathcal{G}$. Indeed, the play ABC^ω provides a smaller cost to player 1, i.e. $\mathsf{Cost}_1(ABC^\omega) < \mathsf{Cost}_1(ADE^\omega)$. Notice that in this example, there is no equilibrium in $\mathcal{G}$ of type $\{1\}$.

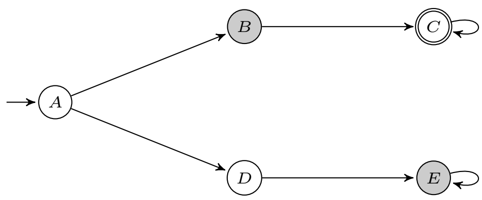

Fig. 2 A game $\mathcal{G}$ with an equilibrium in $\overline{\mathcal{G}}$ that can not be lifted to $\mathcal{G}$.

The next proposition shows that on the opposite side, any Nash equilibrium in a quantitative game $\mathcal{G}$ can be lifted to a Nash equilibrium in the qualitative game $\overline{\mathcal{G}}$.

Proposition 7 *If $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium in a quantitative multiplayer reachability game $\mathcal{G}$, then $(\sigma_i)_{i\in\Pi}$ is also a Nash equilibrium in $\overline{\mathcal{G}}$.*

[6] For qualitative games, we use the notion of *payoff* rather than the notion of *cost* since Win (resp. Lose) can be seen as a payoff of 1 (resp. 0) and the aim of the players is to maximize their payoffs.

Proof For a contradiction, let us assume that in $\overline{\mathcal{G}}$, player j has a profitable deviation σ'_j w.r.t. $(\sigma_i)_{i\in\Pi}$. This is only possible if $\mathsf{Payoff}_j(\langle(\sigma_i)_{i\in\Pi}\rangle) = \mathsf{Lose}$ and $\mathsf{Payoff}_j(\langle\sigma'_j,(\sigma_i)_{i\in\Pi\setminus\{j\}}\rangle) = \mathsf{Win}$. Thus when playing σ'_j against $(\sigma_i)_{i\in\Pi\setminus\{j\}}$, player j manages to visit Goal_j. Clearly enough, σ'_j would also be a profitable deviation w.r.t. $(\sigma_i)_{i\in\Pi}$ in $\mathcal{G}$, contradicting the hypothesis. □

Note that Proposition 7 is false for secure equilibria. To see that, let us come back to the game $\mathcal{G}$ of Figure 2. The strategy profile (σ_1,σ_2) such that $\langle\sigma_1,\sigma_2\rangle = ABC^\omega$ is a secure equilibrium in the quantitative game $\mathcal{G}$ but not in the qualitative game $\overline{\mathcal{G}}$.

2.3 Unraveling

In the proofs of this article we need to unravel the graph $G = (V,(V_i)_{i\in\Pi}, v_0, E)$ from the initial vertex v_0, which ends up in an *infinite tree*, denoted by T. This tree can be seen as a new graph where the set of vertices is the set H of histories of $\mathcal{G}$, the initial vertex is v_0, and a pair $(hv, hvv') \in H\times H$ is an edge of T if $(v,v')\in E$. A history h is a vertex of player i in T if $\mathsf{Last}(h)\in V_i$, and it belongs to the goal set of player i if $\mathsf{Last}(h)\in\mathsf{Goal}_i$.

We denote by $\mathcal{T}$ the related game. This game $\mathcal{T}$ played on the unraveling T of G is equivalent to the game $\mathcal{G}$ that is played on G in the following sense. A play $(\rho_0)(\rho_0\rho_1)(\rho_0\rho_1\rho_2)\dots$ in $\mathcal{T}$ induces a unique play $\rho = \rho_0\rho_1\rho_2\dots$ in $\mathcal{G}$, and conversely. Thus, we denote a play in $\mathcal{T}$ by the respective play in $\mathcal{G}$. The bijection between plays of $\mathcal{G}$ and plays of $\mathcal{T}$ allows us to use the same cost function Cost, and to transform easily strategies in $\mathcal{G}$ to strategies in $\mathcal{T}$ (and conversely).

We also need to study the tree T limited to a certain depth $d\geq 0$: we note $\mathsf{Trunc}_d(T)$ the *truncated tree of* T *of depth* d and $\mathsf{Trunc}_d(\mathcal{T})$ the *finite* game played on $\mathsf{Trunc}_d(T)$. More precisely, the set of vertices of $\mathsf{Trunc}_d(T)$ is the set of histories $h\in H$ of length $\leq d$; the edges of $\mathsf{Trunc}_d(T)$ are defined in the same way as for T except that for the histories h of length d, there exists no edge (h, hv). A play ρ in $\mathsf{Trunc}_d(\mathcal{T})$ corresponds to a history of $\mathcal{G}$ of length *equal to* d. The notions of cost and strategy are defined exactly like in the game $\mathcal{T}$, but limited to the depth d. For instance, a player pays an infinite cost for a play ρ (of length d) if his goal set is not visited by ρ.

2.4 Qualitative Two-player Zero-sum Reachability Games

In this section we recall well-known properties of qualitative two-player zero-sum reachability games [8, Chapter 2]. This will be necessary in our proofs.

Definition 8 A *qualitative two-player zero-sum reachability game* is a tuple $\mathcal{G} = (V, V_1, V_2, E, \mathsf{Goal})$ where

- $G = (V, V_1, V_2, E)$ is a finite directed graph where V is the set of vertices, V_1, V_2 is a partition of V into the state sets of player 1 and player 2, and $E\subseteq V\times V$ is the set of edges,
- $\mathsf{Goal}\subseteq V$ is the goal set of player 1.

Given an initial vertex $v_0\in V$, the notions of *play*, *history* and *strategy* are the same as the ones defined in Section 2.1. Player 1 (resp. player 2) *wins* a play ρ of $\mathcal{G}$

if ρ visits Goal (resp. ρ does not visit Goal). The game is said *zero-sum* because every play is won by exactly one of the two players.

In zero-sum games, it is interesting to know if one of the players can play in such a way that he is sure to win, however the other player plays. We can formalize this by introducing the notion of *winning strategy*. A strategy σ_i for player i is a winning strategy *from an initial vertex* v if all plays of $\mathcal{G}$ starting in v that are consistent with σ_i are won by player i. If player i has a winning strategy in $\mathcal{G}$ from v, we say that player i *wins* the game $\mathcal{G}$ from v. We say that a game $\mathcal{G}$ is *determined* if for all $v \in V$, one of the two players has a winning strategy from v.

Martin showed [13] that every qualitative two-player zero-sum game with a Borel type winning condition is determined. In particular, we have the following proposition:

Proposition 9 ([8]) *Let $\mathcal{G} = (V, V_1, V_2, E, \mathsf{Goal})$ be a qualitative two-player zero-sum reachability game. Then for all $v \in V$, one of the two players has a* memoryless *winning strategy from v (in particular, $\mathcal{G}$ is determined).*
Moreover for all vertices v from which he wins the game, player 1 (resp. player 2) has a memoryless strategy that is independent of v and that forces the play to visit Goal within at most $|V| - 1$ edges (resp. to stay in $V \setminus \mathsf{Goal}$).

3 Nash Equilibria

From now on we will often use the term *game* to denote a quantitative multiplayer reachability game according to Definition 1.

3.1 Existence of a Nash Equilibrium

In this section we positively solve Problem 1 for Nash equilibria.

Theorem 10 *In every quantitative multiplayer reachability game, there exists a finite-memory Nash equilibrium.*

The proof of this theorem is based on the following ideas. By Kuhn's theorem (Theorem 11), there exists a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ played on the finite tree $\mathsf{Trunc}_d(T)$, for any depth d. By choosing an adequate depth d, Proposition 13 enables to extend this Nash equilibrium to a Nash equilibrium in the infinite tree T, and thus in $\mathcal{G}$. Let us detail these ideas.

We first recall Kuhn's theorem [12]. A *preference relation* is a total reflexive transitive binary relation.

Theorem 11 (Kuhn's theorem) *Let Γ be a* finite *tree and $\mathcal{G}_\Gamma$ a game played on Γ. For each player $i \in \Pi$, let $\precsim_i$ be a preference relation on cost profiles. Then there exists a strategy profile $(\sigma_i)_{i\in\Pi}$ such that for every player $j \in \Pi$ and every strategy σ'_j of player j in $\mathcal{G}_\Gamma$ we have*

$$\mathsf{Cost}(\rho') \precsim_j \mathsf{Cost}(\rho)$$

where $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle$ and $\rho' = \langle\sigma'_j, (\sigma_i)_{i\in\Pi\setminus\{j\}}\rangle$.

Note that $\mathsf{Cost}(\rho') \precsim_j \mathsf{Cost}(\rho)$ means that player j prefers the cost profile of the play ρ than the one of ρ', or they are equivalent for him.

Corollary 12 *Let $\mathcal{G}$ be a game and T be the unraveling of G. Let $\mathsf{Trunc}_d(\mathcal{T})$ be the game played on the truncated tree of T of depth d, with $d \geq 0$. Then there exists a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$.*

Proof For each player $j \in \Pi$, we define the relation $\precsim_j$ on cost profiles in the following way: let $(x_i)_{i\in\Pi}$ and $(y_i)_{i\in\Pi}$ be two cost profiles, we say that $(x_i)_{i\in\Pi} \precsim_j (y_i)_{i\in\Pi}$ iff $x_j \geq y_j$. It is clearly a preference relation which captures the Nash equilibrium. The strategy profile $(\sigma_i)_{i\in\Pi}$ of Kuhn's theorem is then a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$. □

Proposition 13 states that it is possible to extend a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ to a Nash equilibrium in the game $\mathcal{T}$, if the depth d is equal to $(|\Pi| + 1) \cdot 2 \cdot |V|$. We obtain Theorem 10 as a consequence of Corollary 12 and Proposition 13.

Proposition 13 *Let $\mathcal{G}$ be a game and T be the unraveling of G. Let $\mathsf{Trunc}_d(\mathcal{T})$ be the game played on the truncated tree of T of depth $d = (|\Pi| + 1) \cdot 2 \cdot |V|$. If there exists a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$, then there exists a finite-memory Nash equilibrium in the game $\mathcal{T}$.*

The proof of Proposition 13 roughly works as follows. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$. A well-chosen prefix $\alpha\beta$, with β being a cycle, is first extracted from the outcome ρ of $(\sigma_i)_{i\in\Pi}$. The outcome of the required Nash equilibrium $(\tau_i)_{i\in\Pi}$ in $\mathcal{T}$ will be equal to $\alpha\beta^\omega$. As soon as a player deviates from this play, all the other players form a coalition to punish him in a way that this deviation is not profitable for him. These ideas are detailed in Lemmas 15 and 16. One can see Lemma 15 as a technical result used to prove Lemma 16, which is the main ingredient to show Proposition 13. The proof of Lemma 15 relies on a particular case (stated below) of Proposition 9. More precisely, we consider the qualitative two-player zero-sum game $\mathcal{G}_j$ played on the graph G, where player j plays in order to reach his goal set Goal_j, against the coalition of all other players that wants to prevent him from reaching his goal set. Player j plays on the vertices from V_j and the coalition on $V \setminus V_j$.

Proposition 14 ([8]) *Let $\mathcal{G}_j = (V, V_j, V \setminus V_j, E, \mathsf{Goal}_j)$ be the qualitative two-player zero-sum reachability game associated to player j. Then player j has a memoryless strategy ν_j that enables him to reach Goal_j within $|V| - 1$ edges from each vertex v from which he wins the game $\mathcal{G}_j$. On the contrary, the coalition has a memoryless strategy ν_{-j} that forces the play to stay in $V \setminus \mathsf{Goal}_j$ from each vertex v from which it wins the game $\mathcal{G}_j$.*

Lemma 15 *Suppose $d \geq 0$. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ and ρ the (finite) outcome of $(\sigma_i)_{i\in\Pi}$. Assume that ρ has a prefix $\alpha\beta\gamma$, where β contains at least one vertex, such that*

$$\begin{aligned}
&\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)\\
&\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta)\\
&|\alpha\beta| \leq l \cdot |V|\\
&|\alpha\beta\gamma| = (l + 1) \cdot |V|
\end{aligned}$$

for some $l \geq 1$.
Let $j \in \Pi$ be such that α does not visit Goal_j. Consider the qualitative two-player

zero-sum game $\mathcal{G}_j = (V, V_j, V \setminus V_j, E, \mathsf{Goal}_j)$. *Then for all histories hu of* $\mathcal{G}$ *consistent with* $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ *and such that* $|hu| \leq |\alpha\beta|$, *the coalition of the players* $i \neq j$ *wins the game* $\mathcal{G}_j$ *from* u.

Condition $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$ means that if Goal_i is visited by $\alpha\beta\gamma$, it has already been visited by α. Condition $\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta)$ means that β is a cycle. The play ρ of Lemma 15 is illustrated in Figure 3.

Lemma 15 says in particular that the players $i \neq j$ can play together to prevent player j from reaching his goal set Goal_j, in case he deviates from the play $\alpha\beta$ (as $\alpha\beta$ is consistent with $(\sigma_i)_{i\in\Pi\setminus\{j\}}$). We denote by ν_{-j} the memoryless winning strategy of the coalition. For each player $i \neq j$, let $\nu_{i,j}$ be the memoryless strategy of player i in $\mathcal{G}$ induced by ν_{-j}.

Proof (of Lemma 15) By contradiction suppose that player j wins the game $\mathcal{G}_j$ from u. By Proposition 14 player j has a memoryless winning strategy ν_j which enables him to reach his goal set Goal_j within at most $|V| - 1$ edges from u. We show that ν_j leads to a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in the game $\mathsf{Trunc}_d(\mathcal{T})$, which is impossible by hypothesis.

Let ρ' be a play in $\mathsf{Trunc}_d(\mathcal{T})$ such that hu is a prefix of ρ', and from u, player j plays according to the strategy ν_j and the other players $i \neq j$ continue to play according to σ_i. As the play ρ' is consistent with the memoryless winning strategy ν_j from u, it visits Goal_j and we have

$$\begin{aligned} \mathsf{Cost}_j(\rho') &\leq |hu| + |V| && \text{(by Proposition 14)} \\ &\leq (l+1)\cdot|V| && \text{(by hypothesis)} \\ &\leq d && \text{(as } \alpha\beta\gamma \leq \rho\text{)}. \end{aligned}$$

We consider the following two cases. If $\mathsf{Cost}_j(\rho) = +\infty$ (i.e. ρ does not visit Goal_j), we have

$$\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho) = +\infty.$$

On the contrary, if $\mathsf{Cost}_j(\rho) < +\infty$ (i.e. ρ visits Goal_j, but after the prefix $\alpha\beta\gamma$ by hypothesis), then we have

$$\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho)$$

as $\mathsf{Cost}_j(\rho) > (l+1)\cdot|V|$.

Since ρ' is consistent with $(\sigma_i)_{i\in\Pi\setminus\{j\}}$, the strategy of player j induced by the play ρ' is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in both cases, which is a contradiction. □

Now that we have proved Lemma 15, we use it in order to obtain Lemma 16, which states that one can define a Nash equilibrium $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$, based on the Nash equilibrium $(\sigma_i)_{i\in\Pi}$ in the game $\mathsf{Trunc}_d(\mathcal{T})$.

Lemma 16 *Suppose* $d \geq 0$. *Let* $(\sigma_i)_{i\in\Pi}$ *be a Nash equilibrium in* $\mathsf{Trunc}_d(\mathcal{T})$ *and* $\alpha\beta\gamma$ *be a prefix of* $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle$ *as defined in Lemma 15. Then there exists a Nash equilibrium* $(\tau_i)_{i\in\Pi}$ *in the game* $\mathcal{T}$. *Moreover* $(\tau_i)_{i\in\Pi}$ *is finite-memory, and* $\mathtt{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$.

Proof Let us set $\Pi = \{1, \dots, n\}$. As α and β end in the same vertex, we can consider the infinite play $\alpha\beta^\omega$ in the game $\mathcal{T}$. Without loss of generality we can order the players $i \in \Pi$ so that

$$\begin{array}{lll} \forall i \leq k & \mathsf{Cost}_i(\alpha\beta^\omega) < +\infty & (\alpha \text{ visits } \mathsf{Goal}_i) \\ \forall i > k & \mathsf{Cost}_i(\alpha\beta^\omega) = +\infty & (\alpha \text{ does not visit } \mathsf{Goal}_i) \end{array}$$

where $0 \leq k \leq n$. In the second case, notice that ρ could visit Goal_i (but after the prefix $\alpha\beta\gamma$).

The Nash equilibrium $(\tau_i)_{i\in\Pi}$ required by Lemma 16 is intuitively defined as follows. First the outcome of $(\tau_i)_{i\in\Pi}$ is exactly $\alpha\beta^\omega$. Secondly the first player j who deviates from $\alpha\beta^\omega$ is punished by the coalition of the other players in the following way. If $j \leq k$ and the deviation occurs in the tree $\mathsf{Trunc}_d(T)$, then the coalition plays according to $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ in this tree. It prevents player j from reaching his goal set Goal_j faster than in $\alpha\beta^\omega$. And if $j > k$, the coalition plays according to $(\nu_{i,j})_{i\in\Pi\setminus\{j\}}$ (given by Lemma 15) so that player j does not reach his goal set at all.

We begin by defining a punishment function P on the vertex set H of T such that $P(h)$ indicates the first player j who has deviated from $\alpha\beta^\omega$, with respect to h. We write $P(h) = \bot$ if no deviation has occurred. For v_0, we define $P(v_0) = \bot$ and for $h \in V^+$ such that $\mathsf{Last}(h) \in V_i$ and $v \in V$, we let:

$$P(hv) = \begin{cases} \bot & \text{if } P(h) = \bot \text{ and } hv < \alpha\beta^\omega, \\ i & \text{if } P(h) = \bot \text{ and } hv \not< \alpha\beta^\omega, \\ P(h) & \text{otherwise } (P(h) \neq \bot). \end{cases}$$

The Nash equilibrium $(\tau_i)_{i\in\Pi}$ is then defined as follows: let h be a history ending in a vertex of V_i,

$$\tau_i(h) = \begin{cases} v & \text{if } P(h) = \bot \ (h < \alpha\beta^\omega)\text{; such that } hv < \alpha\beta^\omega, \\ \textit{arbitrary} & \text{if } P(h) = i, \\ \nu_{i,P(h)}(h) & \text{if } P(h) \neq \bot, i \text{ and } P(h) > k, \\ \sigma_i(h) & \text{if } P(h) \neq \bot, i,\ P(h) \leq k \text{ and } |h| < d, \\ \textit{arbitrary} & \text{otherwise } (P(h) \neq \bot, i,\ P(h) \leq k \text{ and } |h| \geq d) \end{cases} \tag{1}$$

where *arbitrary* means that the next vertex is chosen arbitrarily (in a memoryless way). Clearly the outcome of $(\tau_i)_{i\in\Pi}$ is the play $\alpha\beta^\omega$, and $\mathtt{Type}((\tau_i)_{i\in\Pi})$ is equal to $\mathsf{Visit}(\alpha)$ ($= \mathsf{Visit}(\alpha\beta)$).

It remains to prove that $(\tau_i)_{i\in\Pi}$ is a finite-memory Nash equilibrium in the game $\mathcal{T}$. We first show that the strategy profile $(\tau_i)_{i\in\Pi}$ defined in Equation (1) is a Nash equilibrium in the game $\mathcal{T}$. Let τ'_j be a strategy of player j. We show that this is not a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$. We distinguish the following two cases:

(i) $j \leq k$ ($\mathsf{Cost}_j(\alpha\beta^\omega) < +\infty$, α visits Goal_j).

To improve his cost, player j has no incentive to deviate after the prefix α. Thus we assume that the strategy τ'_j causes a deviation from a vertex visited in α. By Equation (1) the other players first play according to $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ in $\mathsf{Trunc}_d(\mathcal{T})$, and then in an arbitrary way.

Suppose that τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$. Let us set $\pi = \langle(\tau_i)_{i\in\Pi}\rangle$ and $\pi' = \langle\tau'_j, (\tau_i)_{i\in\Pi\setminus\{j\}}\rangle$. Then

$$\mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\pi).$$

On the other hand we know that

$$\mathsf{Cost}_j(\pi) = \mathsf{Cost}_j(\rho) \le |\alpha|.$$

So if we limit the play π' in $\mathcal{T}$ to its prefix of length d, we get a play ρ' in $\mathsf{Trunc}_d(\mathcal{T})$ such that

$$\mathsf{Cost}_j(\rho') = \mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\rho).$$

As the play ρ' is consistent with the strategies $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ by Equation (1), the strategy τ'_j restricted to the tree $\mathsf{Trunc}_d(T)$ is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in the game $\mathsf{Trunc}_d(\mathcal{T})$. This contradicts the fact that $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium in this game.

(ii) $j > k$ ($\mathsf{Cost}_j(\alpha\beta^\omega) = +\infty$, $\alpha\beta^\omega$ does not visit Goal_j).

If player j deviates from $\alpha\beta^\omega$ (with the strategy τ'_j), by Equation (1) the other players combine against him and play according to ν_{-j}. By Lemma 15 this coalition wins the game $\mathcal{G}_j$ from any vertex visited by $\alpha\beta^\omega$. So the strategy ν_{-j} of the coalition keeps the play $\langle\tau'_j, (\tau_i)_{i\in\Pi\setminus\{j\}}\rangle$ away from the set Goal_j, whatever player j does. Therefore τ'_j is not a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$.

We now prove that $(\tau_i)_{i\in\Pi}$ is a finite-memory strategy profile. According to the definition of finite-memory strategy (see Section 2) we have to prove that each relation $\approx_{\tau_i}$ on H has finite index (recall that $h \approx_{\tau_i} h'$ if $\tau_i(h\delta) = \tau_i(h'\delta)$ for all $\delta \in V^*V_i$). In this aim we define for each player i an equivalence relation $\sim_{\tau_i}$ with finite index such that

$$\forall h, h' \in H, \quad h \sim_{\tau_i} h' \Rightarrow h \approx_{\tau_i} h'.$$

We first define an equivalence relation $\sim_P$ with finite index related to the punishment function P. For all prefixes h, h' of $\alpha\beta^\omega$, i.e. such that no player is punished, this relation does not distinguish two histories that are identical except for a certain number of cycles β. For the other histories it just has to remember the first player, say i, who has deviated. The definition of $\sim_P$ is as follows:

$$\begin{array}{ll} h \sim_P h' & \text{if } h = \alpha\beta^l\beta',\ h' = \alpha\beta^m\beta',\ \beta' < \beta,\ l, m \ge 0 \\ hv \sim_P h'v' & \text{if } v, v' \in V_i,\ h, h' < \alpha\beta^\omega,\ \text{but } hv, h'v' \not< \alpha\beta^\omega \\ hv \sim_P hv\delta & \text{if } h < \alpha\beta^\omega,\ hv \not< \alpha\beta^\omega,\ \delta \in V^*. \end{array}$$

The relation $\sim_P$ is an equivalence relation on H with finite index.

We now turn to the definition of $\sim_{\tau_i}$. It is based on the definition of τ_i (given in (1)) and $\sim_P$. To get an equivalence with finite index we proceed as follows. Recall that each strategy $\nu_{i,P(h)}$ is memoryless and when a player plays arbitrarily, his strategy is also memoryless. Furthermore notice that, in the definition of τ_i, the strategy σ_i is only applied to histories h with length $|h| < d$. For histories h such that $\tau_i(h) = v$ with $hv < \alpha\beta^\omega$, it is enough to remember information with respect to $\alpha\beta$ as already done for $\sim_P$. Therefore for $h, h' \in H$ we define $\sim_{\tau_i}$ in the following way:

$$\begin{array}{lll} h \sim_{\tau_i} h' \text{ if } h \sim_P h' & \text{and} & \big(P(h) = \bot \\ & & \text{or } P(h) = i \text{ and } \mathsf{Last}(h) = \mathsf{Last}(h') \\ & & \text{or } P(h) \ne \bot, i,\ P(h) > k \text{ and } \mathsf{Last}(h) = \mathsf{Last}(h') \\ & & \text{or } P(h) \ne \bot, i,\ P(h) \le k,\ |h|, |h'| \ge d \text{ and} \\ & & \quad \mathsf{Last}(h) = \mathsf{Last}(h')\big). \end{array}$$

Notice that this relation satisfies

$$h \sim_{\tau_i} h' \Rightarrow \tau_i(h) = \tau_i(h') \text{ and } \mathsf{Last}(h) = \mathsf{Last}(h')$$

and has finite index. Therefore if $h \sim_{\tau_i} h'$, then $h \approx_{\tau_i} h'$ and the relation $\approx_{\tau_i}$ has finite index. □

We can now proceed to the proof of Proposition 13.

Proof (of Proposition 13) Let us set $\Pi = \{1, \dots, n\}$ and $d = (n+1)\cdot 2\cdot |V|$. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ and ρ its outcome.

To be able to use Lemma 16, we consider the prefix $\mathfrak{pq}$ of ρ of minimal length such that

$$\begin{aligned} \exists\, l \geq 1 \qquad & |\mathfrak{p}| = (l-1)\cdot|V| \\ & |\mathfrak{pq}| = (l+1)\cdot|V| \\ & \mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{pq})\,. \end{aligned} \tag{2}$$

The following statements are true.

- $l \leq 2\cdot n + 1$.
- If $\mathsf{Visit}(\mathfrak{p}) \subsetneq \mathsf{Visit}(\rho)$, then $l < 2\cdot n + 1$.

Indeed the first statement results from the fact that in the worst case, the play ρ visits the goal set of a new player in each prefix of length $i\cdot 2\cdot|V|$, $1 \leq i \leq n$, i.e. $|\mathfrak{p}| = n\cdot 2\cdot|V|$. It follows that $\mathfrak{pq}$ exists as a prefix of ρ, because the length d of ρ is equal to $(n+1)\cdot 2\cdot|V|$ by hypothesis. Thus $\mathsf{Visit}(\mathfrak{p}) \subseteq \mathsf{Visit}(\rho)$. Suppose that there exists $i \in \mathsf{Visit}(\rho) \setminus \mathsf{Visit}(\mathfrak{p})$, then ρ visit Goal_i after the prefix $\mathfrak{pq}$ by Equation (2). The second statement follows easily.

Given the length of $\mathfrak{q}$, one vertex of V is visited at least twice by $\mathfrak{q}$. More precisely, we can write

$$\mathfrak{pq} = \alpha\beta\gamma \quad \text{with} \quad \begin{aligned} & \mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta) \\ & |\alpha| \geq (l-1)\cdot|V| \\ & |\alpha\beta| \leq l\cdot|V|\,. \end{aligned}$$

In particular, $|\mathfrak{p}| \leq |\alpha|$. See Figure 3. We have $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$, and $|\alpha\beta\gamma| = (l+1)\cdot|V|$.

As the hypotheses of Lemma 16 are verified, we can apply it in this context to get a finite-memory Nash equilibrium $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$ with $\mathtt{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$. □

Proposition 13 asserts that given a game $\mathcal{G}$ and the game $\mathsf{Trunc}_d(\mathcal{T})$ played on the truncated tree of T of a well-chosen depth d, one can lift any Nash equilibrium $(\sigma_i)_{i\in\Pi}$ of $\mathsf{Trunc}_d(\mathcal{T})$ to a Nash equilibrium $(\tau_i)_{i\in\Pi}$ of $\mathcal{G}$. The proof of Proposition 13 states that the type of $(\tau_i)_{i\in\Pi}$ is equal to $\mathsf{Visit}(\alpha)$. We give an example that shows that it is impossible to preserve the type of the lifted Nash equilibrium $(\sigma_i)_{i\in\Pi}$.

Example 17 Let us consider the two-player game $\mathcal{G}$ depicted in Figure 4 with $\mathsf{Goal}_1 = \{C\}$, $\mathsf{Goal}_2 = \{E\}$. One can show that $\mathcal{G}$ admits only Nash equilibria of type $\{2\}$ or $\emptyset$. Indeed, on one hand, there is no play of $\mathcal{G}$ where both goals are visited, and on the other hand given a strategy profile $(\sigma_i)_{i\in\Pi}$ such that $\langle(\sigma_i)_{i\in\Pi}\rangle$ visits Goal_1 (i.e. $\langle(\sigma_i)_{i\in\Pi}\rangle$

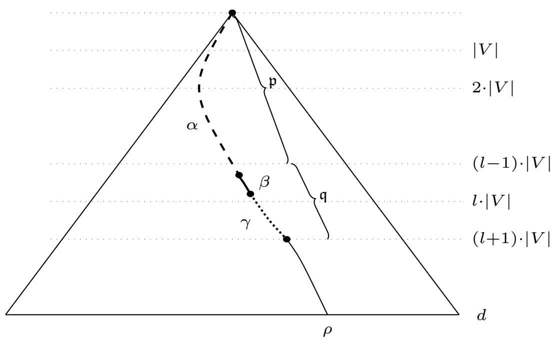

Fig. 3 Slicing of the play ρ in the tree $\mathsf{Trunc}_d(T)$.

is of the form A^+BC^ω), playing D instead of C is clearly a profitable deviation for player 2.

We will now see that for each $d \geq 2$ the game played on $\mathsf{Trunc}_d(T)$ admits a Nash equilibrium of type $\{1\}$. From the above discussion, this equilibrium can not be lifted to a Nash equilibrium of the same type in $\mathcal{G}$. A truncated tree $\mathsf{Trunc}_d(T)$ is depicted in Figure 5. One can show that the strategy profile leading to the outcome $A^{d-1}BC$ (depicted in bold in the figure) is a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ of type $\{1\}$. Following the lines of the proof of Proposition 13, we see that this Nash equilibrium is lifted to a Nash equilibrium of $\mathcal{G}$ with outcome A^ω and type $\emptyset$.

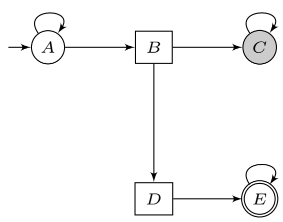

Fig. 4 A game $\mathcal{G}$.

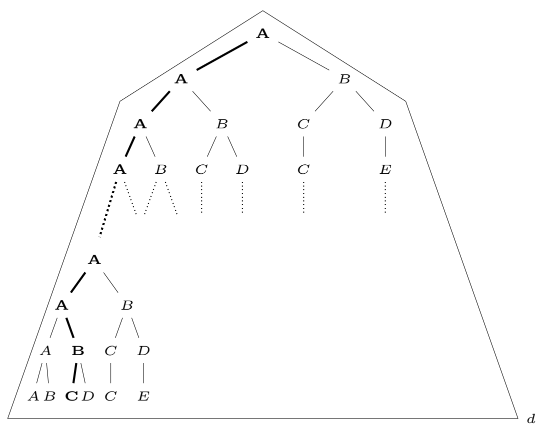

Fig. 5 The truncated tree $\mathsf{Trunc}_d(T)$.

On the other hand, notice that from the proof of Proposition 13, we can construct a Nash equilibrium such that each player pays either an infinite cost, or a cost bounded by $|\Pi| \cdot 2 \cdot |V|$.

3.2 Nash Equilibria with Finite Memory Preserving Types

In this section we show that given a Nash equilibrium, we can construct another Nash equilibrium with the same type such that all its strategies are finite-memory. We then answer to Problem 2 for Nash equilibria.

Theorem 18 *If there exists a Nash equilibrium in a quantitative multiplayer reachability game $\mathcal{G}$, then there exists a finite-memory Nash equilibrium of the same type in $\mathcal{G}$.*

The proof is based on two steps. The first step constructs from $(\sigma_i)_{i\in\Pi}$ another Nash equilibrium $(\tau_i)_{i\in\Pi}$ with the same type such that the play $\langle(\tau_i)_{i\in\Pi}\rangle$ is of the form $\alpha\beta^\omega$ with $\mathsf{Visit}(\alpha) = \mathtt{Type}((\sigma_i)_{i\in\Pi})$. This is possible thanks to Lemmas 19 and 20, by first eliminating unnecessary cycles in the play $\langle(\sigma_i)_{i\in\Pi}\rangle$ and then locating a prefix $\alpha\beta$ such that β is a cycle that can be infinitely repeated.

The second step transforms the Nash equilibrium $(\tau_i)_{i\in\Pi}$ into a finite-memory one thanks to Lemma 16 given in Section 3.1. For that purpose, we consider the strategy profile $(\tau_i)_{i\in\Pi}$ limited to the tree T truncated at a well-chosen depth.

The next lemma indicates how to eliminate a cycle in the outcome of a Nash equilibrium.

Lemma 19 *Let $(\sigma_i)_{i\in\Pi}$ be a strategy profile in a game $\mathcal{G}$ and $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle$ its outcome. Suppose that $\rho = \mathfrak{p}\mathfrak{q}\tilde{\rho}$, where $\mathfrak{q}$ contains at least one vertex, such that*

$$\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{p}\mathfrak{q})$$
$$\mathsf{Last}(\mathfrak{p}) = \mathsf{Last}(\mathfrak{p}\mathfrak{q}).$$

We define a strategy profile $(\tau_i)_{i\in\Pi}$ as follows:

$$\tau_i(h) = \begin{cases} \sigma_i(h) & \text{if } \mathfrak{p} \not\leq h, \\ \sigma_i(\mathfrak{p}\mathfrak{q}\delta) & \text{if } h = \mathfrak{p}\delta \end{cases}$$

where h is a history of $\mathcal{G}$ with $\mathsf{Last}(h) \in V_i$. We get the outcome $\langle(\tau_i)_{i\in\Pi}\rangle = \mathfrak{p}\tilde{\rho}$.
If a strategy τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$, then there exists a profitable deviation σ'_j for player j w.r.t. $(\sigma_i)_{i\in\Pi}$.

Proof Let us set $\Pi = \{1, \ldots, n\}$. We write

$$\rho = \langle(\sigma_i)_{i\in\Pi}\rangle \text{ of cost profile } (x_1, \ldots, x_n),$$
$$\pi = \langle(\tau_i)_{i\in\Pi}\rangle \text{ of cost profile } (y_1, \ldots, y_n).$$

We observe that as $\rho = \mathfrak{p}\mathfrak{q}\tilde{\rho}$, we have $\pi = \mathfrak{p}\tilde{\rho}$ (see Figures 6 and 7). It follows that

$$\forall i \in \Pi, \quad y_i \leq x_i. \tag{3}$$

More precisely,

$$\text{- if } x_i = +\infty, \text{ then } y_i = +\infty; \tag{4}$$
$$\text{- if } x_i < +\infty \text{ and } i \in \mathsf{Visit}(\mathfrak{p}), \text{ then } y_i = x_i;$$
$$\text{- if } x_i < +\infty \text{ and } i \notin \mathsf{Visit}(\mathfrak{p}), \text{ then } y_i = x_i - (|\mathfrak{q}| + 1). \tag{5}$$

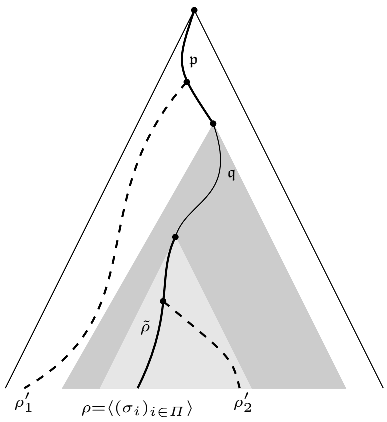

Fig. 6 Play ρ and possible deviations.

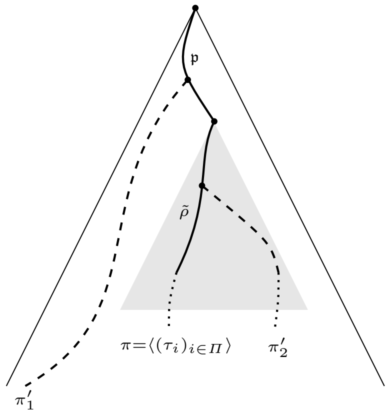

Fig. 7 Play π and possible deviations.

Let τ'_j be a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$, and π' be the outcome of the strategy profile $(\tau'_j, (\tau_i)_{i\in\Pi\setminus\{j\}})$. Then

$$\mathsf{Cost}_j(\pi') < y_j.$$

We show how to construct a profitable deviation σ'_j for player j w.r.t. $(\sigma_i)_{i\in\Pi}$. Two cases occur:

(i) player j deviates from π just after a proper prefix h of $\mathfrak{p}$ (like for the play π'_1 in Figure 7).

We define $\sigma'_j = \tau'_j$ and we denote by ρ' the outcome of $(\sigma'_j, (\sigma_i)_{i\in\Pi\setminus\{j\}})$. Given the definition of the strategy profile $(\tau_i)_{i\in\Pi}$, one can verify that $\rho' = \pi'$ (see the play ρ'_1 in Figure 6). Thus

$$\mathsf{Cost}_j(\rho') = \mathsf{Cost}_j(\pi') < y_j \leq x_j$$

by Equation (3), which implies that σ'_j is a profitable deviation of player j w.r.t. $(\sigma_i)_{i\in\Pi}$.

(ii) player j deviates from π after the prefix $\mathfrak{p}$ (π and π' coincide at least on $\mathfrak{p}$).

This case is illustrated by the play π'_2 in Figure 7. We define for all histories h ending in a vertex of V_j:

$$\sigma'_j(h) = \begin{cases} \sigma_j(h) & \text{if } \mathfrak{pq} \not\leq h, \\ \tau'_j(\mathfrak{p}\delta) & \text{if } h = \mathfrak{pq}\delta. \end{cases}$$

Let us set $\rho' = \langle \sigma'_j, (\sigma_i)_{i\in\Pi\setminus\{j\}}\rangle$. As player j deviates after $\mathfrak{p}$ with the strategy τ'_j, one can prove that

$$\pi' = \mathfrak{p}\tilde{\pi}' \quad \text{and} \quad \rho' = \mathfrak{pq}\tilde{\pi}'$$

by definition of $(\tau_i)_{i\in\Pi}$ (see the play ρ'_2 in Figure 6). As $\mathsf{Cost}_j(\pi') < y_j$, it means that $j \notin \mathsf{Visit}(\mathfrak{p})$ (otherwise the deviation would not be profitable for player j). Since $\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{pq})$, we also have

$$\mathsf{Cost}_j(\pi') + (|\mathfrak{q}| + 1) = \mathsf{Cost}_j(\rho').$$

By Equations (4) and (5), we get

– either $x_j = y_j = +\infty$ and $\mathsf{Cost}_j(\rho') < x_j$,
– or $x_j = y_j + (|\mathfrak{q}| + 1)$ and $\mathsf{Cost}_j(\rho') < x_j$,

which proves that σ'_j is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$. □

While Lemma 19 deals with elimination of unnecessary cycles, Lemma 20 deals with repetition of a useful cycle.

Lemma 20 *Let $(\sigma_i)_{i\in\Pi}$ be a strategy profile in a game $\mathcal{G}$ and $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle$ its outcome. We assume that $\rho = \mathfrak{p}\mathfrak{q}\tilde{\rho}$, where $\mathfrak{q}$ contains at least one vertex, such that*

$$\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\rho)$$
$$\mathsf{Last}(\mathfrak{p}) = \mathsf{Last}(\mathfrak{p}\mathfrak{q}).$$

We define a strategy profile $(\tau_i)_{i\in\Pi}$ as follows:

$$\tau_i(h) = \begin{cases} \sigma_i(h) & \text{if } \mathfrak{p} \not\leq h, \\ \sigma_i(\mathfrak{p}\delta) & \text{if } h = \mathfrak{p}\mathfrak{q}^k\delta,\ k \in \mathbb{N}, \text{ and } \mathfrak{q} \not\leq \delta \end{cases}$$

where h is a history of $\mathcal{G}$ with $\mathsf{Last}(h) \in V_i$. We get the outcome $\langle(\tau_i)_{i\in\Pi}\rangle = \mathfrak{p}\mathfrak{q}^\omega$.
If a strategy τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$, then there exists a profitable deviation σ'_j for player j w.r.t. $(\sigma_i)_{i\in\Pi}$.

Proof We use the same notations as in the proof of Lemma 19. Here we have $x_i = y_i$ for all $i \in \Pi$ since $\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\rho)$. One can prove that $\pi = \mathfrak{p}\mathfrak{q}^\omega$ (see Figures 8 and 9).

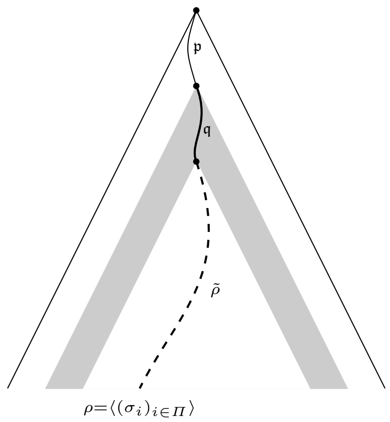

Fig. 8 Play ρ and its prefix $\mathfrak{p}\mathfrak{q}$.

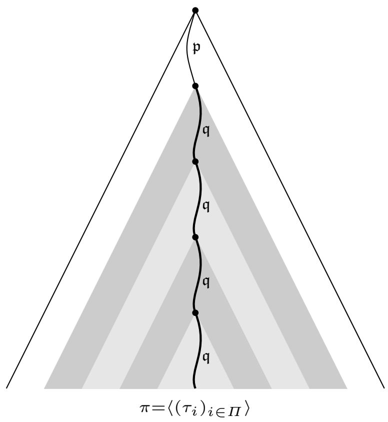

Fig. 9 Play $\pi = \mathfrak{p}\mathfrak{q}^\omega$.

We show how to define a profitable deviation σ'_j from the deviation τ'_j. We distinguish the following two cases:

(i) player j deviates from π just after a proper prefix h of $\mathfrak{p}\mathfrak{q}$.
We define $\sigma'_j = \tau'_j$. As in the first case of the proof of Lemma 19, we have $\mathsf{Cost}_j(\rho') < x_j$, which implies that σ'_j is a profitable deviation of player j w.r.t. $(\sigma_i)_{i\in\Pi}$.

(*ii*) player j deviates from π after the prefix $\mathtt{pq}$, i.e. after a prefix $\mathtt{pq}^k$ and strictly before the prefix $\mathtt{pq}^{k+1}$ ($k \geq 1$).

We define for all histories h ending in a vertex of V_j:

$$\sigma'_j(h) = \begin{cases} \sigma_j(h) & \text{if } \mathtt{p} \not\leq h, \\ \tau'_j(\mathtt{pq}^k\delta) & \text{if } h = \mathtt{p}\delta. \end{cases}$$

One can prove that

$$\pi' = \mathtt{pq}^k\tilde{\pi}' \quad \text{and} \quad \rho' = \mathtt{p}\tilde{\pi}'.$$

And then, in the point of view of costs we have

$$\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\pi') < y_j = x_j,$$

which proves that σ'_j is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$. □

The next proposition achieves the first step of the proof of Theorem 18 as mentioned in Section 3.2. It shows that one can construct from a Nash equilibrium another Nash equilibrium with the same type and with an outcome of the form $\alpha\beta^\omega$. Its proof uses Lemmas 19 and 20.

Proposition 21 *Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in a game $\mathcal{G}$. Then there exists a Nash equilibrium $(\tau_i)_{i\in\Pi}$ with the same type and such that $\langle(\tau_i)_{i\in\Pi}\rangle = \alpha\beta^\omega$, where $\mathsf{Visit}(\alpha) = \mathtt{Type}((\sigma_i)_{i\in\Pi})$ and $|\alpha\beta| < (|\Pi| + 1) \cdot |V|$.*

Proof Let us set $\Pi = \{1, \ldots, n\}$ and $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle$. Without loss of generality suppose that

$$\mathsf{Cost}(\rho) = (x_1, \ldots, x_n) \quad \text{such that } x_1 \leq \ldots \leq x_k < +\infty$$
$$\text{and } x_{k+1} = \ldots = x_n = +\infty$$

where $0 \leq k \leq n$. We consider two cases:

(*i*) $x_1 \geq |V|$.

Then, there exists a prefix $\mathtt{pq}$ of ρ, with $\mathtt{q}$ containing at least one vertex, such that

$$|\mathtt{pq}| < x_1$$
$$\mathsf{Visit}(\mathtt{p}) = \mathsf{Visit}(\mathtt{pq}) = \emptyset$$
$$\mathsf{Last}(\mathtt{p}) = \mathsf{Last}(\mathtt{pq}).$$

We define the strategy profile $(\tau_i)_{i\in\Pi}$ as proposed in Lemma 19. By this lemma it is actually a Nash equilibrium in $\mathcal{G}$. With $\pi = \langle(\tau_i)_{i\in\Pi}\rangle$, we have

$$\rho = \mathtt{pq}\tilde{\rho} \quad \text{and} \quad \pi = \mathtt{p}\tilde{\rho}.$$

Thus if the cost profile for the play π is $(y_1, \ldots, y_n)$, we have

$$y_1 < x_1, \ldots, y_k < x_k$$
$$y_{k+1} = x_{k+1} = +\infty, \ldots, y_n = x_n = +\infty.$$

(ii) $(x_{l+1} - x_l) \geq |V|$ for $1 \leq l \leq k-1$.

Then, there exists a prefix $\mathfrak{pq}$ of ρ, with $\mathfrak{q}$ containing at least one vertex, such that

$$\begin{aligned}&x_l < |\mathfrak{pq}| < x_{l+1}\\&\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{pq}) = \{1,\dots,l\}\\&\mathsf{Last}(\mathfrak{p}) = \mathsf{Last}(\mathfrak{pq}).\end{aligned}$$

We define the strategy profile $(\tau_i)_{i\in\Pi}$ given in Lemma 19. It is then a Nash equilibrium in $\mathcal{G}$, and for $\pi = \langle(\tau_i)_{i\in\Pi}\rangle$, we have

$$\rho = \mathfrak{pq}\tilde{\rho} \quad \text{and} \quad \pi = \mathfrak{p}\tilde{\rho}.$$

Hence if the cost profile for the play π is $(y_1,\dots,y_n)$, we have

$$\begin{aligned}&y_1 = x_1,\dots,y_l = x_l;\\&y_{l+1} < x_{l+1},\dots,y_k < x_k;\\&y_{k+1} = x_{k+1} = +\infty,\dots,y_n = x_n = +\infty.\end{aligned}$$

By applying finitely many times the two previous cases, we can assume without loss of generality that $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium with a cost profile $(x_1,\dots,x_n)$ such that

$$\begin{aligned}&x_i < i\cdot|V| \quad \text{for } i \leq k;\\&x_i = +\infty \quad\;\; \text{for } i > k.\end{aligned}$$

Let us go further. We can write $\rho = \alpha\beta\tilde{\rho}$ such that

$$\begin{aligned}&\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)\\&\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta)\\&|\alpha\beta| < (k+1)\cdot|V| \leq (n+1)\cdot|V|.\end{aligned}$$

Indeed, the prefix h of ρ of length $(k+1)\cdot|V|$ visits each goal set Goal_i, with $i \leq k$, and after the last visited Goal_k, there remains enough vertices to observe a cycle. Notice that $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta) = \mathsf{Visit}(\rho)$ $(= \mathtt{Type}((\sigma_i)_{i\in\Pi}))$.

If we define the strategy profile $(\tau_i)_{i\in\Pi}$ like in Lemma 20, we get a Nash equilibrium in $\mathcal{G}$ with outcome $\alpha\beta^\omega$ and the same type as $(\sigma_i)_{i\in\Pi}$. □

We are now ready to prove Theorem 18.

Proof (of Theorem 18) Let us set $\Pi = \{1,\dots,n\}$. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in the game $\mathcal{G}$. The first step consists in constructing a Nash equilibrium as in Proposition 21. Let us denote it again by $(\sigma_i)_{i\in\Pi}$. Let us set $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle = \alpha\beta^\omega$ such that $\mathsf{Visit}(\alpha) = \mathtt{Type}((\sigma_i)_{i\in\Pi})$ and $|\alpha\beta| < (n+1)\cdot|V|$. The strategy profile $(\sigma_i)_{i\in\Pi}$ is also a Nash equilibrium in the game $\mathcal{T}$ played on the unraveling T of G.

For the second step we consider $\mathsf{Trunc}_d(T)$ the truncated tree of T of depth $d = (n+2)\cdot|V|$. It is clear that the strategy profile $(\sigma_i)_{i\in\Pi}$ limited to this tree is also a Nash equilibrium of $\mathsf{Trunc}_d(\mathcal{T})$.

We know that $|\alpha\beta| < (n+1)\cdot|V|$ and we set γ such that $\alpha\beta\gamma$ is a prefix of ρ and $|\alpha\beta\gamma| = (n+2)\cdot|V|$. Furthermore we have $\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta)$ and $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$ (since $\mathsf{Visit}(\alpha) = \mathtt{Type}(\rho)$). Then this prefix $\alpha\beta\gamma$ satisfies the properties described in Lemma 15 (by setting $l = n+1$). By Lemma 16 we conclude that there exists a Nash equilibrium $(\tau_i)_{i\in\Pi}$ with finite memory such that $\mathtt{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$, that is, with the same type as the initial Nash equilibrium $(\sigma_i)_{i\in\Pi}$. □

4 Secure Equilibria

In the previous section, we positively solved Problem 1 and Problem 2 for Nash equilibria. We here solve these two problems for secure equilibria, but in two-player games only. The main results are stated in Theorems 22 and 28 below. In this section, we exclusively consider two-player games.

Theorem 22 *In every quantitative two-player reachability game, there exists a finite-memory secure equilibrium.*

The proof of Theorem 22 is based on the same ideas as for the proof of Theorem 10 (existence of a Nash equilibrium). By Kuhn's theorem (Theorem 11), there exists a secure equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ played on the finite tree $\mathsf{Trunc}_d(T)$, for any depth d. By choosing an adequate depth d, Proposition 25 enables to extend this secure equilibrium to a secure equilibrium in the infinite tree T, and thus in $\mathcal{G}$.

The notion of secure equilibrium is based on the binary relations $\prec_j$ of Definition 3. One can easily see that $\prec_j$ is not reflexive. To be able to apply Kuhn's theorem, it is more convenient to define secure equilibria via a preference relation. Given two cost profiles (x_1, x_2) and (y_1, y_2):

$$(x_1, x_2) \precsim_j (y_1, y_2) \quad \text{iff} \quad (x_1, x_2) \prec_j (y_1, y_2) \quad \vee \quad (x_1 = y_1 \wedge x_2 = y_2)\,.$$

The relation $\precsim_j$ is clearly a preference relation[7]. We can now provide an equivalent definition of secure equilibrium.

Proposition 23 *A strategy profile (σ_1, σ_2) of a game $\mathcal{G}$ is a secure equilibrium iff for all strategies σ_1' of player 1 in $\mathcal{G}$, we have:*

$$\mathsf{Cost}(\rho') \precsim_1 \mathsf{Cost}(\rho)$$

where $\rho = \langle \sigma_1, \sigma_2 \rangle$ and $\rho' = \langle \sigma_1', \sigma_2 \rangle$, and symmetrically for all strategies σ_2' of player 2.

Since $\precsim_1$ and $\precsim_2$ are preference relations, we get the next corollary by Kuhn's theorem.

Corollary 24 *Let $\mathcal{G}$ be a quantitative two-player reachability game and T be the unraveling of G. Let $\mathsf{Trunc}_d(\mathcal{T})$ be the game played on the truncated tree of T of depth d, with $d \geq 0$. Then there exists a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$.*

Now that we can guarantee the existence of secure equilibrium in finite trees, it remains to show how to lift them to infinite trees. The next proposition states that it is possible to extend a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ to a secure equilibrium in the game $\mathcal{T}$ with the same type, if the depth d is greater or equal to $(|\Pi| + 1) \cdot 2 \cdot |V|$ and there are only two players. It also says that we can construct a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ from a secure equilibrium in $\mathcal{T}$, while keeping the same type.

Proposition 25 *Let $\mathcal{G}$ be a two-player game and T be the unraveling of G.*

(i) *If there exists a secure equilibrium of a certain type in the game $\mathcal{T}$, then there exists a secure equilibrium of the same type in the game $\mathsf{Trunc}_d(\mathcal{T})$, for some depth $d \geq (|\Pi| + 1) \cdot 2 \cdot |V|$.*

[7] Remark that $\precsim_j$ is a kind of lexicographic order on $(\mathbb{N} \cup \{+\infty\}) \times (\mathbb{N} \cup \{+\infty\})$.

(*ii*) *If there exists a secure equilibrium of a certain type in the game* $\mathsf{Trunc}_d(\mathcal{T})$*, where* $d \geq (|\Pi| + 1) \cdot 2 \cdot |V|$*, then there exists a finite-memory secure equilibrium of the same type in the game* $\mathcal{T}$*.*

To prove Proposition 25, we need the following technical lemma whose hypotheses are the same as in Lemma 15. Recall that Lemma 15 states that for all $j \in \Pi$ such that α does not visit Goal_j, the players $i \neq j$ can play together to prevent player j from reaching his goal set Goal_j from any history hu consistent with $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ and such that $|hu| \leq |\alpha\beta|$. We denote by ν_{-j} the memoryless winning strategy of the coalition, and for each player $i \neq j$, $\nu_{i,j}$ the memoryless strategy of player i in $\mathcal{G}$ induced by ν_{-j}.

Lemma 26 *Suppose* $d \geq 0$*. Let* (σ_1, σ_2) *be a secure equilibrium in* $\mathsf{Trunc}_d(\mathcal{T})$ *and* $\rho = \langle\sigma_1, \sigma_2\rangle$ *its outcome. Assume that* ρ *has a prefix* $\alpha\beta\gamma$*, where* β *contains at least one vertex, such that*

$$\begin{aligned}
&\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)\\
&\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta)\\
&|\alpha\beta| \leq l \cdot |V|\\
&|\alpha\beta\gamma| = (l+1) \cdot |V|
\end{aligned}$$

for some $l \geq 1$*. Then we have*

$$(\mathsf{Visit}(\alpha) \neq \emptyset \vee \mathsf{Visit}(\rho) \neq \{1,2\}) \Rightarrow \mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho).$$

In particular, Lemma 26 implies that if α visits none of the goal sets, then ρ visits either both goal sets or none. Notice that in the case of Nash equilibria, we can have situations contradicting Lemma 26, and in particular the previous situation, as it can be seen in Example 17.

Proof By contradiction, assume that $2 \in \mathsf{Visit}(\rho)\setminus\mathsf{Visit}(\alpha)$ (the case where $1 \in \mathsf{Visit}(\rho)\setminus\mathsf{Visit}(\alpha)$ is symmetric). The hypothesis implies that $1 \in \mathsf{Visit}(\alpha)$ or $1 \notin \mathsf{Visit}(\rho)$.

By Lemma 15, player 1 wins the game $\mathcal{G}_2$ from $\mathsf{Last}(\alpha)$, that is, has a memoryless winning strategy $\nu_{1,2}$ from this vertex. Then if player 1 plays according to σ_1 until depth $|\alpha|$, and then switches to $\nu_{1,2}$ from $\mathsf{Last}(\alpha)$, this strategy is a $\prec_1$-profitable deviation for player 1 w.r.t. (σ_1, σ_2). Indeed, if $1 \in \mathsf{Visit}(\alpha)$, player 1 manages to increase player 2's cost while keeping his own cost. On the other hand, if $1 \notin \mathsf{Visit}(\rho)$, either player 1 succeeds in reaching his goal set (i.e. strictly decreasing his cost), or he does not reach it (then gets the same cost as in ρ) but succeeds in increasing player 2's cost. Thus we get a contradiction. □

We can now give the proof of Proposition 25. The idea for showing case (*i*) is to look at the play π of the secure equilibrium in $\mathcal{T}$ and consider the depth d needed to visit all the goal sets of the players in $\mathsf{Visit}(\pi)$. Then, the secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ is defined exactly as the secure equilibrium of $\mathcal{T}$.

The proof of case (*ii*) works pretty much as the one of Proposition 13 (whereas the latter proposition does not preserve the type of the Nash equilibrium). Thanks to Lemma 26, the proof reduces into only two cases depending on when the goal sets are visited. In the most interesting case, a well-chosen prefix $\alpha\beta$, with β being a cycle, is first extracted from the outcome ρ of the secure equilibrium (σ_1, σ_2) of $\mathsf{Trunc}_d(\mathcal{T})$. The outcome of the required secure equilibrium of $\mathcal{T}$ will be equal to $\alpha\beta^\omega$. As soon as a player deviates from this play, the other player punishes him, but the way to define the punishment is here more involved than in the proof of Proposition 13. In the other case, the proof is simpler, but the ideas are quite the same.

Before entering the details, let us introduce a notation. For any play $\rho = \rho_0\rho_1\ldots$ of $\mathcal{G}$ and any player $i \in \Pi$, we define $\mathsf{Index}_i(\rho)$ as the least index l such that $\rho_l \in \mathsf{Goal}_i$ if it exists, or -1 if not[8].

Proof (of Proposition 25) First let us begin with the proof of (i). Suppose that there exists a secure equilibrium (τ_1, τ_2) in $\mathcal{T}$ and that the play π is the outcome of this strategy profile. Let us set $d := \max\{(|\Pi|+1)\cdot 2\cdot |V|, \mathsf{Index}_1(\pi), \mathsf{Index}_2(\pi)\}$ and define (σ_1, σ_2) as the strategy profile in $\mathsf{Trunc}_d(\mathcal{T})$ corresponding to the strategies (τ_1, τ_2) restricted to the finite tree. Clearly the outcome ρ of (σ_1, σ_2) is a prefix of π and $\mathsf{Visit}(\rho) = \mathsf{Visit}(\pi)$, so (σ_1, σ_2) and (τ_1, τ_2) are of the same type. It remains to show that (σ_1, σ_2) is a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$.

Assume by contradiction that player 1 has a $\prec_1$-profitable deviation σ_1' w.r.t. (σ_1, σ_2) (the case of player 2 is symmetric). We write ρ' for the outcome of (σ_1', σ_2) in $\mathsf{Trunc}_d(\mathcal{T})$. There are two cases to consider: either player 1 manages to decrease his cost in ρ' w.r.t. ρ, or he pays the same cost as in ρ but he is able to increase the cost of player 2 in ρ' w.r.t. ρ. In both cases, if player 1 plays according to σ_1' in $\mathcal{T}$ until depth d and then arbitrarily, one can easily be convinced that we get a $\prec_1$-profitable deviation[9] w.r.t. (τ_1, τ_2) in $\mathcal{T}$.This leads to a contradiction.

Now let us proceed to the proof of (ii). Let (σ_1, σ_2) be a secure equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$, where $d \geq (|\Pi| + 1) \cdot 2 \cdot |V|$, and ρ its outcome. We define the prefixes $\mathfrak{pq}$ and $\alpha\beta\gamma$ as in the proof of Proposition 13 (see Figure 3).

By Lemma 26 there are only two cases to consider:

(a) $\mathsf{Visit}(\alpha) = \emptyset$ and $\mathsf{Visit}(\rho) = \{1, 2\}$;
(b) $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$.

We define a different secure equilibrium according to the case.

Let us start with case (a): $\mathsf{Visit}(\alpha) = \emptyset$ and $\mathsf{Visit}(\rho) = \{1, 2\}$. We define the following strategy profile:

$$\tau_i(h) = \begin{cases} \sigma_i(h) & \text{if } |h| < \max\{\mathsf{Index}_1(\rho), \mathsf{Index}_2(\rho)\}, \\ \textit{arbitrary} & \text{otherwise.} \end{cases}$$

where $i = 1, 2$, and *arbitrary* means that the next vertex is chosen arbitrarily, but in a memoryless way. Note that the outcome of (τ_1, τ_2) is of the form $\alpha'(\beta')^\omega$ where $\mathsf{Visit}(\alpha') = \mathsf{Visit}(\rho) = \{1, 2\}$ and β' is a cycle. So, (τ_1, τ_2) has the same type as (σ_1, σ_2). It remains to prove that (τ_1, τ_2) is a finite-memory secure equilibrium in $\mathcal{T}$.

Assume by contradiction that player 1 has a $\prec_1$-profitable deviation τ_1' w.r.t. (τ_1, τ_2) in $\mathcal{T}$ (the case for player 2 is symmetric). The strategy σ_1' equal to τ_1' in $\mathsf{Trunc}_d(\mathcal{T})$ is clearly a $\prec_1$-profitable deviation w.r.t. (σ_1, σ_2), which is a contradiction with the fact that (σ_1, σ_2) is a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$. Moreover, as done in the proof of Lemma 16, (τ_1, τ_2) is a finite-memory strategy profile.

[8] We are conscious that it is counterintuitive to use the particular value -1, but it is helpful in the proofs.

[9] Notice that in the second case, when ρ does not visit Goal_1 in $\mathsf{Trunc}_d(\mathcal{T})$, player 1 may reach his goal set in $\mathcal{T}$ when deviating in this way, and this would be profitable for him in this game.

Now we consider case (b): $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$. Like in the proof of Lemma 16 we consider the infinite play $\alpha\beta^\omega$ in the game $\mathcal{T}$. The basic idea of the strategy profile (τ_1, τ_2) is the same as for the Nash equilibrium case: player 2 (resp. 1) plays according to $\alpha\beta^\omega$ and punishes player 1 (resp. 2) if he deviates from $\alpha\beta^\omega$, in the following way. Suppose that player 1 deviates (the case for player 2 is similar). Then player 2 plays according to σ_2 until depth $|\alpha|$, and after that, he plays arbitrarily if α visits Goal_1, otherwise he plays according to $\nu_{2,1}$.

We define the same punishment function P as in the proof of Lemma 16: for v_0, we define $P(v_0) = \perp$ and for $h \in V^+$ such that $\mathsf{Last}(h) \in V_i$ and $v \in V$, we let:

$$P(hv) = \begin{cases} \perp & \text{if } P(h) = \perp \text{ and } hv < \alpha\beta^\omega, \\ i & \text{if } P(h) = \perp \text{ and } hv \not< \alpha\beta^\omega, \\ P(h) & \text{otherwise } (P(h) \neq \perp). \end{cases}$$

The definition of the secure equilibrium (τ_1, τ_2) is as follows: for $h \in H$ such that $\mathsf{Last}(h) \in V_i$:

$$\tau_i(h) = \begin{cases} v & \text{if } P(h) = \perp \ (h < \alpha\beta^\omega)\text{; such that } hv < \alpha\beta^\omega, \\ \textit{arbitrary} & \text{if } P(h) = i, \\ \sigma_i(h) & \text{if } P(h) \neq \perp, i \text{ and } |h| \leq |\alpha|, \\ \nu_{i,P(h)}(h) & \text{if } P(h) \neq \perp, i,\ |h| > |\alpha| \text{ and } \alpha \text{ does not visit } \mathsf{Goal}_{P(h)}, \\ \textit{arbitrary} & \text{otherwise } (P(h) \neq \perp, i,\ |h| > |\alpha| \text{ and } \alpha \text{ visits } \mathsf{Goal}_{P(h)}) \end{cases}$$

where $i = 1, 2$, and *arbitrary* means that the next vertex is chosen arbitrarily (in a memoryless way). Clearly the outcome of (τ_1, τ_2) is the play $\alpha\beta^\omega$, and the type of (τ_1, τ_2) is equal to $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$, the type of (σ_1, σ_2). Moreover, as done in the proof of Lemma 16, (τ_1, τ_2) is a finite-memory strategy profile.

Remark that the definition of the strategy profile (τ_1, τ_2) is a little different from the one in the proof of Lemma 16 because here, if player 1 deviates (for example), then player 2 has to prevent him from reaching his goal set Goal_1 (faster), or having the same cost but succeeding in increasing player 2's cost.

It remains to show that (τ_1, τ_2) is a secure equilibrium in the game $\mathcal{T}$. Assume by contradiction that there exists a $\prec_1$-profitable deviation τ_1' for player 1 w.r.t. (τ_1, τ_2). The case of a $\prec_2$-profitable deviation τ_2' for player 2 is similar. We construct a play ρ' in $\mathsf{Trunc}_d(\mathcal{T})$ as follows: player 1 plays according to the strategy τ_1' restricted to $\mathsf{Trunc}_d(\mathcal{T})$ (denoted by σ_1') and player 2 plays according to σ_2. Thus the play ρ' coincide with the play $\pi' = \langle \tau_1', \tau_2 \rangle$ at least until depth $|\alpha|$ (by definition of τ_2); it can differ afterwards. We have:

$$\begin{aligned} \rho\ &= \langle \sigma_1, \sigma_2 \rangle \text{ of cost profile } (x_1, x_2) \\ \rho' &= \langle \sigma_1', \sigma_2 \rangle \text{ of cost profile } (x_1', x_2') \\ \pi\ &= \langle \tau_1, \tau_2 \rangle \text{ of cost profile } (y_1, y_2) \\ \pi' &= \langle \tau_1', \tau_2 \rangle \text{ of cost profile } (y_1', y_2'). \end{aligned}$$

The situation is depicted in Figure 10.

By contradiction, we assumed that τ_1' is a $\prec_1$-profitable deviation for player 1 w.r.t. (τ_1, τ_2), i.e. $(y_1, y_2) \prec_1 (y_1', y_2')$. Now we are going to show that $(x_1, x_2) \prec_1 (x_1', x_2')$, meaning that σ_1' is a $\prec_1$-profitable deviation for player 1 w.r.t. (σ_1, σ_2) in $\mathsf{Trunc}_d(\mathcal{T})$. This will lead to the contradiction. As τ_1' is a $\prec_1$-profitable deviation w.r.t. (τ_1, τ_2), one of the following three cases stands.

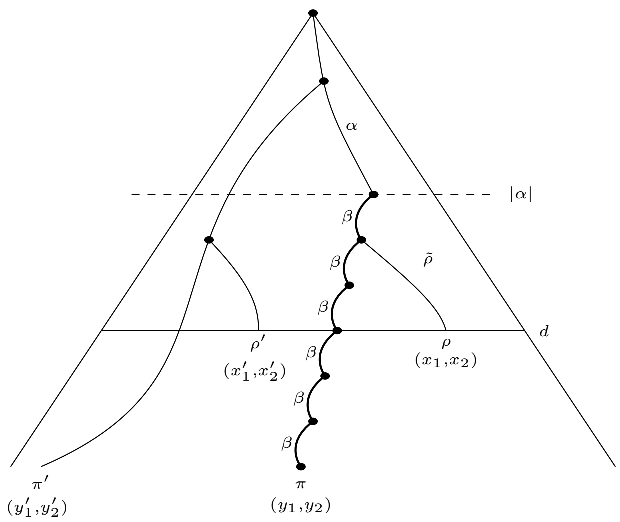

Fig. 10 Plays ρ and π, and their respective deviations ρ' and π'.

(1) $y_1' < y_1 < +\infty$.
As $\pi = \alpha\beta^\omega$, it means that α visits F_1, and then:

$$y_1' < y_1 = x_1 \leq |\alpha|.$$

As $y_1' < |\alpha|$, we have $x_1' = y_1'$ (as ρ' and π' coincide until depth $|\alpha|$). Therefore $x_1' < x_1$, and $(x_1, x_2) \prec_1 (x_1', x_2')$.

(2) $y_1' < y_1 = +\infty$.
If $y_1' \leq |\alpha|$, we have $x_1' = y_1'$ (by the same argument as before). As $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$, we have $x_1 = y_1 = +\infty$ and $x_1' < x_1$ (and so $(x_1, x_2) \prec_1 (x_1', x_2')$).
We show that the case $y_1' > |\alpha|$ is impossible. By definition of τ_2 the play π' is consistent with σ_2 until depth $|\alpha|$, and then with $\nu_{2,1}$ (as $y_1 = +\infty$). By Lemma 15 the play π' can not visit Goal_1 after a depth $> |\alpha|$.

(3) $y_1 = y_1'$ and $y_2 < y_2'$.
Note that this implies $y_2 < +\infty$ and $x_2 = y_2$ (as $\pi = \alpha\beta^\omega$). Since ρ' and π' coincide until depth $|\alpha|$, $y_2 < y_2'$ and $x_2 = y_2 \leq |\alpha|$, we have

$$x_2 = y_2 < x_2'$$

showing that the cost of player 2 is increased. In order to ensure that σ_1' is a $\prec_1$-profitable deviation, it remains to show that either player 1 keeps the same cost, or he decreases his cost.
If $y_1' = y_1 < +\infty$, it follows as in the first case that:

$$y_1 = x_1 \leq |\alpha| \quad \text{and} \quad x_1' = y_1'.$$

Therefore $x_1 = x_1'$, i.e. player 1 has the same cost in ρ and ρ'. And so, $(x_1, x_2) \prec_1 (x_1', x_2')$.
On the contrary, if $y_1' = y_1 = +\infty$, it follows that $x_1 = +\infty$ (as $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$). And so, we have that $x_1' < +\infty = x_1$, or $x_1' = x_1$. But in both cases, it holds that $(x_1, x_2) \prec_1 (x_1', x_2')$.

In conclusion, we constructed a $\prec_1$-profitable deviation σ_1' w.r.t. (σ_1, σ_2) in $\mathsf{Trunc}_d(\mathcal{T})$, and then we get a contradiction. $\square$

Remark 27 Let us notice that in case (i) of Proposition 25, the proof remains valid if we take $d = \max\{0, \mathsf{Index}_1(\pi), \mathsf{Index}_2(\pi)\}$. Thus, in the statement of case (i), the constraint $d \geq (|\Pi| + 1) \cdot 2 \cdot |V|$ can be replaced by $d \geq 0$.

We can now proceed to the proof of Theorem 22.

Proof (of Theorem 22) Let us set $d := (\Pi + 1) \cdot 2 \cdot |V|$ and apply Corollary 24 on the game $\mathsf{Trunc}_d(\mathcal{T})$. Then we get a secure equilibrium in this game. By Proposition 25 there exists in $\mathcal{G}$ a finite-memory secure equilibrium with the same type. $\square$

Theorem 22 positively answers to Problem 1 for secure equilibria in *two-player* games. The next theorem solves Problem 2 for the same kind of games.

Theorem 28 *If there exists a secure equilibrium in a quantitative two-player reachability game $\mathcal{G}$, then there exists a finite-memory secure equilibrium of the same type in $\mathcal{G}$.*

Proof Let (σ_1, σ_2) be a secure equilibrium in $\mathcal{G}$. By the first part of Proposition 25, there exists a secure equilibrium of the same type in the game $\mathsf{Trunc}_d(\mathcal{T})$, for a certain depth $d \geq (\Pi + 1) \cdot 2 \cdot |V|$. If we apply the second part of Proposition 25, we get a finite-memory secure equilibrium of the same type as (σ_1, σ_2) in $\mathcal{G}$. $\square$

The proof of Theorem 28 is based on Proposition 25 which, roughly speaking, ensures that every secure equilibrium of $\mathsf{Trunc}_d(\mathcal{T})$ can be lifted to a secure equilibrium *of the same type* in $\mathcal{T}$, and conversely. Notice that Proposition 25 has no counterpart for Nash equilibria, since we can not guarantee that the type can be preserved, as it can be seen from Example 17. This approach makes the proof of Theorem 28 rather different than the proof of Theorem 18.

Notice that Proposition 25 stands for two-player games because its proof uses Lemma 26 that has been proved only for two players.

5 Extensions of the Model

5.1 Safety Objectives

Let us now consider quantitative games played on a graph where some players have *reachability objectives*, whereas others have *safety objectives.* As previously, the players with reachability objectives want to reach their goal set as soon as possible. The players with safety objectives want to avoid their *bad set* or, if impossible, delay its visit as long as possible. Let us make that precise through the following definition.

Definition 29 An *infinite turn-based quantitative multiplayer reachability/safety game* is a tuple $\mathcal{G} = (\Pi, \Pi_r, \Pi_s, V, (V_i)_{i \in \Pi}, v_0, E, (\mathsf{Goal}_i)_{i \in \Pi_r}, (\mathsf{Bad}_i)_{i \in \Pi_s})$ where

- Π is a finite set of players partitioned into Π_r and Π_s which are the players with reachability and safety objectives respectively,

- $G = (V, (V_i)_{i\in\Pi}, v_0, E)$ is a finite directed graph where V is the set of vertices, $(V_i)_{i\in\Pi}$ is a partition of V into the state sets of each player, $v_0 \in V$ is the initial vertex, and $E \subseteq V \times V$ is the set of edges, and
- $\mathsf{Goal}_i \subseteq V$ is the goal set of player i, for $i \in \Pi_r$; $\mathsf{Bad}_i \subseteq V$ is the bad set of player i, for $i \in \Pi_s$.

For any play $\rho = \rho_0\rho_1 \ldots$ of $\mathcal{G}$, we note $\mathsf{Cost}_i(\rho)$ the *cost* of player i. For $i \in \Pi_r$ the cost is defined as before and for $i \in \Pi_s$ the cost is defined by:

$$\mathsf{Cost}_i(\rho) = \begin{cases} -l & \text{if } l \text{ is the } \textit{least} \text{ index such that } \rho_l \in \mathsf{Bad}_i, \\ -\infty & \text{otherwise.} \end{cases}$$

As before, the aim of each player i is to *minimize* his cost, i.e. reach his goal set Goal_i as soon as possible for $i \in \Pi_r$, or delay the visit of Bad_i as long as possible for $i \in \Pi_s$. The notions of play, strategy, outcome and Nash equilibrium extend in a natural way. The main result of this subsection is the following theorem which solves Problem 1 in this framework.

Theorem 30 *In every quantitative multiplayer reachability/safety game, there exists a finite-memory Nash equilibrium.*

In order to prove Theorem 30, we have to revisit the results of Section 3. Let us first notice that Lemma 15 remains true in this framework when player j belongs to Π_r. Lemma 16 remains true, however we have to slightly adapt its proof.

Proof (of Lemma 16 in the case of reachability/safety objectives)
Let us first introduce some notations. In the rest of the proof, we denote by Π_r^f (resp. Π_s^f) the subset of players $i \in \Pi_r$ (resp. $i \in \Pi_s$) such that α visits Goal_i (resp. Bad_i) and by Π_r^∞ (resp. Π_s^∞) the set $\Pi_r \setminus \Pi_r^f$ (resp. $\Pi_s \setminus \Pi_s^f$).

The punishment function P is defined exactly as in the proof of Lemma 16. For v_0, we define $P(v_0) = \perp$ and for $h \in V^+$ such that $\mathsf{Last}(h) \in V_i$ and $v \in V$, we let:

$$P(hv) = \begin{cases} \perp & \text{if } P(h) = \perp \text{ and } hv < \alpha\beta^\omega, \\ i & \text{if } P(h) = \perp \text{ and } hv \not< \alpha\beta^\omega, \\ P(h) & \text{otherwise } (P(h) \neq \perp)\,. \end{cases}$$

The difference with the proof of Lemma 16 arises in the definition of the Nash equilibrium $(\tau_i)_{i\in\Pi}$. The new equilibrium needs to incorporate an adequate punishment for the players with safety objectives. More precisely, in order to dissuade a player $j \in \Pi_s^f$ from deviating, the other players punish him by playing the strategies $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ in $\mathsf{Trunc}_d(T)$. Notice that a player $j \in \Pi_s^\infty$ has no incentive to deviate. Formally we define the Nash equilibrium $(\tau_i)_{i\in\Pi}$ as follows. For $h \in H$ such that $\mathsf{Last}(h) \in V_i$,

$$\tau_i(h) = \begin{cases} v & \text{if } P(h) = \perp \ (h < \alpha\beta^\omega); \text{ such that } hv < \alpha\beta^\omega, \\ \textit{arbitrary} & \text{if } P(h) = i, \\ \nu_{i,P(h)}(h) & \text{if } P(h) \neq \perp, i \text{ and } P(h) \in \Pi_r^\infty, \\ \sigma_i(h) & \text{if } P(h) \neq \perp, i,\ P(h) \in \Pi_r^f \cup \Pi_s^f \text{ and } |h| < d, \\ \textit{arbitrary} & \text{otherwise,} \end{cases} \tag{6}$$

where *arbitrary* means that the next vertex is chosen arbitrarily (in a memoryless way). Clearly the outcome of $(\tau_i)_{i\in\Pi}$ is the play $\alpha\beta^\omega$, and $\mathtt{Type}((\tau_i)_{i\in\Pi})$ is equal to $\mathsf{Visit}(\alpha)$ $(= \mathsf{Visit}(\alpha\beta))$.

It remains to prove that $(\tau_i)_{i\in\Pi}$ is a finite-memory Nash equilibrium in the game $\mathcal{T}$. In order to do so, we prove that none of the players has a profitable deviation. For players with reachability objectives, the arguments are exactly the same as the ones provided in the proof of Lemma 16. Let us now consider players with safety objectives. In the case where $j \in \Pi_s^\infty$, player j has clearly no incentive to deviate. In the case where $j \in \Pi_s^f$, to decrease his cost, player j has no incentive to deviate after the prefix α. Thus we assume that the strategy τ'_j causes a deviation from a vertex visited in α. By Equation (6) the other players first play according to $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ in $\mathsf{Trunc}_d(\mathcal{T})$, and then in an arbitrary way.

Suppose that τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$. Let us set $\pi = \langle(\tau_i)_{i\in\Pi}\rangle$ and $\pi' = \langle\tau'_j, (\tau_i)_{i\in\Pi\setminus\{j\}}\rangle$. Then

$$\mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\pi).$$

On the other hand we know that

$$\mathsf{Cost}_j(\pi) = \mathsf{Cost}_j(\rho) \leq |\alpha|.$$

So if we limit the play π' in $\mathcal{T}$ to its prefix of length d, we get a play ρ' in $\mathsf{Trunc}_d(\mathcal{T})$ such that

$$\mathsf{Cost}_j(\rho') \leq \mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\rho).$$

Notice that we do not necessarily have that $\mathsf{Cost}_j(\rho') = \mathsf{Cost}_j(\pi')$ (as in the proof of Lemma 16) since the bad set Bad_j can be visited by π' and not by ρ'. As the play ρ' is consistent with the strategies $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ by Equation (6), the strategy τ'_j restricted to the tree $\mathsf{Trunc}_d(T)$ is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in the game $\mathsf{Trunc}_d(\mathcal{T})$. This is impossible. Moreover, as done in the proof of Lemma 16, (τ_1, τ_2) is a finite-memory strategy profile.

□

Since Lemma 16 holds in the context of reachability/safety objectives, Proposition 13 ensures that the equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ provided by Kuhn's theorem (Corollary 12) can be lifted to $\mathcal{T}$. This proves Theorem 30.

5.2 Tuples of Costs on Edges

In this subsection, we come back to a pure reachability framework and we extend our model in the following way: we assume that edges are labelled with tuples of positive costs (one cost for each player). Here we do not only count the number of edges to reach the goal of a player, but we sum up his costs along the path until his goal is reached. His aim is still to minimize his global cost for a play. We generalize Definition 1.

Definition 31 An *infinite turn-based quantitative multiplayer reachability game with tuples of costs on edges* is a tuple $\mathcal{G} = (\Pi, V, (V_i)_{i\in\Pi}, v_0, E, (\mathsf{Cost}_i)_{i\in\Pi}, (\mathsf{Goal}_i)_{i\in\Pi})$ where

- Π is a finite set of players,
- $G = (V, (V_i)_{i\in\Pi}, v_0, E)$ is a finite directed graph where V is the set of vertices, $(V_i)_{i\in\Pi}$ is a partition of V into the state sets of each player, $v_0 \in V$ is the initial vertex, and $E \subseteq V \times V$ is the set of edges,
- $\mathsf{Cost}_i : E \to \mathbb{R}^{>0}$ is the cost function of player i defined on the edges of the graph,

- $\mathsf{Goal}_i \subseteq V$ is the goal set of player i.

We also positively solve Problem 1 for Nash equilibria in this context.

Theorem 32 *In every quantitative multiplayer reachability game with tuples of costs on edges, there exists a finite-memory Nash equilibrium.*

To prove Theorem 32, we follow the same scheme as in Section 3. In particular, we rely on Kuhn's theorem (Corollary 12) and need to prove a counterpart of Lemma 15, Lemma 16 and Proposition 13 in this framework.

Let us first introduce some notations that will be useful in this context. We define $\mathsf{c_{min}} := \min_{i\in\Pi} \min_{e\in E} \mathsf{Cost}_i(e)$, $\mathsf{c_{max}} := \max_{i\in\Pi} \max_{e\in E} \mathsf{Cost}_i(e)$ and $\mathsf{K} := \left\lceil \frac{\mathsf{c_{max}}}{\mathsf{c_{min}}} \right\rceil$. It is clear that $\mathsf{c_{min}}, \mathsf{c_{max}} > 0$ and $\mathsf{K} \geq 1$.

We also adapt the definition of $\mathsf{Cost}_i(\rho)$, the *cost* of player i for a play $\rho = \rho_0\rho_1\ldots$,

$$\mathsf{Cost}_i(\rho) = \begin{cases} \displaystyle\sum_{k=1}^{k=l} \mathsf{Cost}_i((\rho_{k-1},\rho_k)) & \text{if } l \text{ is the \textit{least} index such that } \rho_l \in \mathsf{Goal}_i, \\ +\infty & \text{otherwise.} \end{cases}$$

The counterpart of Lemma 15 is the following one, taking into account the constant K defined before.

Lemma 33 *Suppose $d \geq 0$. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ and ρ the outcome of $(\sigma_i)_{i\in\Pi}$. Assume that ρ has a prefix $\alpha\beta\gamma$, where β contains at least one vertex, such that*

$$\begin{aligned} &\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma) \\ &\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta) \\ &|\alpha\beta| \leq l\cdot|V| \\ &|\alpha\beta\gamma| = (l+\mathsf{K})\cdot|V| \end{aligned}$$

for some $l \geq 1$.
Let $j \in \Pi$ be such that α does not visit Goal_j. Consider the qualitative two-player zero-sum game $\mathcal{G}_j = (V, V_j, V\setminus V_j, E, \mathsf{Goal}_j)$. Then for all histories hu of $\mathcal{G}$ consistent with $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ and such that $hu \leq \alpha\beta$, the coalition of the players $i \neq j$ wins the game $\mathcal{G}_j$ from u.

Proof (Sketch) As for the proof of Lemma 15 we proceed by contradiction and define a play ρ' in the very same way. We can deduce that

$$\begin{aligned} \mathsf{Index}_j(\rho') &\leq |hu| + |V| && \text{(by Proposition 14)} \\ &\leq (l+1)\cdot|V| && \text{(by hypothesis)} \\ &\leq (l+\mathsf{K})\cdot|V| && (\text{as } \mathsf{K}\geq 1) \\ &\leq d && (\text{as } \alpha\beta\gamma \leq \rho). \end{aligned}$$

The case where $\mathsf{Cost}_j(\rho) = +\infty$ is solved in the same way. For the other case $\mathsf{Cost}_j(\rho) < +\infty$, we note $\mathsf{c}_j(hu)$ the sum of the costs of player j along the prefix hu.

We have the following inequalities (see Figure 11):

$$\begin{aligned}
\mathsf{Cost}_j(\rho') &\leq \mathsf{c}_j(hu) + \mathsf{c}_{\mathsf{max}} \cdot |V| \\
\mathsf{Cost}_j(\rho) &> \mathsf{c}_j(hu) + \mathsf{c}_{\mathsf{min}} \cdot \mathsf{K} \cdot |V| && (\text{as } \mathsf{Index}_j(\rho) > (l+\mathsf{K}) \cdot |V|) \\
&\geq \mathsf{c}_j(hu) + \mathsf{c}_{\mathsf{min}} \cdot \frac{\mathsf{c}_{\mathsf{max}}}{\mathsf{c}_{\mathsf{min}}} \cdot |V| && (\text{by definition of } \mathsf{K}) \\
&= \mathsf{c}_j(hu) + \mathsf{c}_{\mathsf{max}} \cdot |V| .
\end{aligned}$$

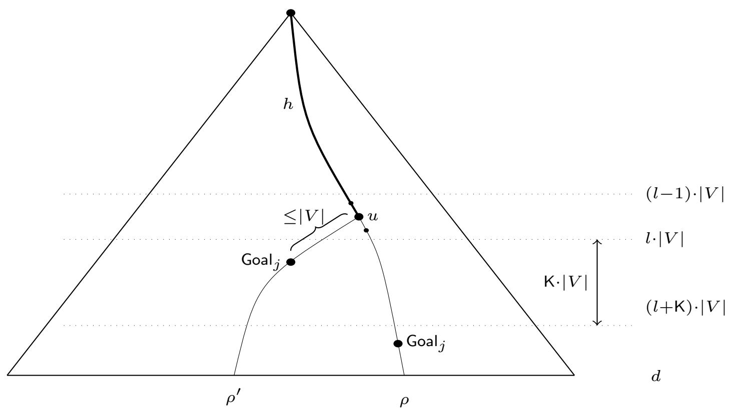

Fig. 11 Plays ρ and ρ' with their common prefix hu.

Then we have $\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho)$, and since ρ' is consistent with $(\sigma_i)_{i\in\Pi\setminus\{j\}}$, the strategy of player j induced by the play ρ' is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$. This contradicts the fact that $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$. □

The following lemma is the counterpart of Lemma 16.

Lemma 34 *Suppose $d \geq 0$. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ and $\alpha\beta\gamma$ be a prefix of $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle$ as defined in Lemma 33 where $|\alpha\beta\gamma| = (l+\mathsf{K}) \cdot |V|$ for some $l \geq 1$ such that $l \leq \frac{d}{|V|\cdot\mathsf{K}} + 1$.*
Then there exists a Nash equilibrium $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$. Moreover $(\tau_i)_{i\in\Pi}$ is finite-memory, and $\mathtt{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$.

Proof We prove this result in the very same way as Lemma 16. The only difference lies in the case[10] $j \leq k$ when we show that $(\tau_i)_{i\in\Pi}$ is a Nash equilibrium. We suppose that τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$. So we have $\mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\pi)$, where $\pi = \langle(\tau_i)_{i\in\Pi}\rangle$ and $\pi' = \langle\tau'_j, (\tau_i)_{i\in\Pi\setminus\{j\}}\rangle$. As $\mathsf{Index}_j(\pi) \leq |\alpha|$, we know that $\mathsf{Cost}_j(\pi) \leq |\alpha| \cdot \mathsf{c}_{\mathsf{max}}$. It follows that $\mathsf{Cost}_j(\pi') < |\alpha| \cdot \mathsf{c}_{\mathsf{max}}$ and

$$\begin{aligned}
\mathsf{Index}_j(\pi') &< |\alpha| \cdot \frac{\mathsf{c}_{\mathsf{max}}}{\mathsf{c}_{\mathsf{min}}} \\
&\leq (l-1) \cdot |V| \cdot \mathsf{K} \\
&\leq d && (\text{by hypothesis}).
\end{aligned}$$

[10] Indeed when $j > k$, i.e. when player j has not reached his goal set, the coalition punishes him in the exact same way as Lemma 16 by preventing him from visiting his goal set.

The first inequality can be justified as follows. For a contradiction, let us assume that $\mathsf{Index}_j(\pi') \geq |\alpha| \cdot \frac{\mathsf{c_{max}}}{\mathsf{c_{min}}}$. It follows that $\mathsf{Cost}_j(\pi') \geq \mathsf{c_{min}} \cdot |\alpha| \cdot \frac{\mathsf{c_{max}}}{\mathsf{c_{min}}}$, this contradicts the fact that $\mathsf{Cost}_j(\pi') < |\alpha| \cdot \mathsf{c_{max}}$.

As in the proof of Lemma 16, we limit the play π' in $\mathcal{T}$ to its prefix of length d and get a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in the game $\mathsf{Trunc}_d(\mathcal{T})$, contradicting the fact that $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$.

Moreover, as done in the proof of Lemma 16, (τ_1, τ_2) is a finite-memory strategy profile. □

As a consequence of the two previous lemmas, Proposition 13 remains true in this context, we only have to adjust the depth d of the finite tree.

Proposition 35 *Let $\mathcal{G}$ be a game and T be the unraveling of G. Let $\mathsf{Trunc}_d(\mathcal{T})$ be the game played on the truncated tree of T of depth $d = \max\{(|\Pi| + 1) \cdot (\mathsf{K} + 1) \cdot |V|, (|\Pi| \cdot (\mathsf{K} + 1) + 1) \cdot |V| \cdot \mathsf{K}\}$.*
If there exists a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$, then there exists a finite-memory Nash equilibrium in the game $\mathcal{T}$.

Proof The proof is similar to the proof of Proposition 13. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ and ρ its outcome. We consider the prefix $\mathfrak{pq}$ of ρ of minimal length such that

$$\exists\, l \geq 1 \qquad \begin{aligned} &|\mathfrak{p}| = (l-1) \cdot |V| \\ &|\mathfrak{pq}| = (l + \mathsf{K}) \cdot |V| \\ &\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{pq}). \end{aligned}$$

In the worst case, the play ρ visits the goal set of a new player in each prefix of length $i \cdot (\mathsf{K} + 1) \cdot |V|$, $1 \leq i \leq |\Pi|$, i.e. $|\mathfrak{p}| = |\Pi| \cdot (\mathsf{K} + 1) \cdot |V|$. So we know that $l \leq |\Pi| \cdot (\mathsf{K} + 1) + 1$ and $\mathfrak{pq}$ exists as a prefix of ρ, because the length d of ρ is greater or equal to $(|\Pi| + 1) \cdot (\mathsf{K} + 1) \cdot |V|$ by hypothesis.

Given the length of $\mathfrak{q}$ ($\mathsf{K} \geq 1$), one vertex of V is visited at least twice by $\mathfrak{q}$. More precisely, we can write

$$\mathfrak{pq} = \alpha\beta\gamma \quad \text{with} \quad \begin{aligned} &\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta) \\ &|\alpha| \geq (l-1) \cdot |V| \\ &|\alpha\beta| \leq l \cdot |V|. \end{aligned}$$

We have $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$, and $|\alpha\beta\gamma| = (l + \mathsf{K}) \cdot |V|$.

Moreover, the following inequality holds:

$$d \geq (|\Pi| \cdot (\mathsf{K} + 1) + 1) \cdot |V| \cdot \mathsf{K} \geq l \cdot |V| \cdot \mathsf{K} \quad \text{and so,} \quad l \leq \frac{d}{|V| \cdot \mathsf{K}}.$$

Then, we can apply Lemma 34 and get a finite-memory Nash equilibrium $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$ such that $\mathtt{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$. □

Thanks to Corollary 12 and Proposition 35, one can easily deduce Theorem 32.

Let us comment on the depth d chosen in Proposition 35. It is defined as the maximum between $d_1 := (|\Pi| + 1) \cdot (\mathsf{K} + 1) \cdot |V|$ and $d_2 := (|\Pi| \cdot (\mathsf{K} + 1) + 1) \cdot |V| \cdot \mathsf{K}$. One can easily prove that $d_1 < d_2$ if and only if $\mathsf{K}^2 > \frac{|\Pi|+1}{|\Pi|}$.

We now investigate an alternative method to handle *simple cost functions*. More precisely we only consider cost functions $(\mathsf{Cost}_i)_{i\in\Pi}$ such that for all i, $j \in \Pi$ we have that $\mathsf{Cost}_i = \mathsf{Cost}_j$ and $\mathsf{Cost}_i : E \rightarrow \mathbb{N}_0$. In other words, it means that there is a unique non-zero natural cost on every edge. Later on we are going to compare the depths of the finite trees obtained by the two methods.

In the case of these simple cost functions, we can directly deduce Theorem 32 by replacing any edge of cost c by a path of length c composed of c new edges (of cost 1) and then applying the results of Section 3 on this new game. If we write $\mathcal{G}' = (\Pi, V', (V'_i)_{i\in\Pi}, v_0, E', (\mathsf{Goal}_i)_{i\in\Pi})$ the new game obtained by adding new vertices and edges when necessary, it holds that:

$$\begin{aligned}|V'| &\leq |V| + (\mathsf{c_{max}} - 1)\cdot|E| \\ &\leq |V| + (\mathsf{c_{max}} - 1)\cdot|V|^2, \text{ and} \\ |E'| &\leq \mathsf{c_{max}}\cdot|E|\,.\end{aligned}$$

If we apply Proposition 13, the depth d' of the finite tree that is considered satisfies:

$$\begin{aligned}d' &= (|\Pi|+1)\cdot 2\cdot|V'| \\ &\leq (|\Pi|+1)\cdot 2\cdot(|V| + (\mathsf{c_{max}} - 1)\cdot|E|) \\ &\leq (|\Pi|+1)\cdot 2\cdot\left(|V| + (\mathsf{c_{max}} - 1)\cdot|V|^2\right)\,.\end{aligned}$$

Whereas if we apply Proposition 35 directly on the initial game $\mathcal{G}$, we have the following equality:

$$d = \max\{(|\Pi|+1)\cdot(\mathsf{K}+1)\cdot|V|, (|\Pi|\cdot(\mathsf{K}+1)+1)\cdot|V|\cdot\mathsf{K}\}\,.$$

Let us first notice that if all the edges of $\mathcal{G}$ are labelled with the same cost (i.e., $\mathsf{c_{max}} = \mathsf{c_{min}}$ and $\mathsf{K} = 1$), then

$$\begin{aligned}d' &= (|\Pi|+1)\cdot 2\cdot(|V| + (\mathsf{c_{max}} - 1)\cdot|E|), \text{ and} \\ d &= (|\Pi|+1)\cdot 2\cdot|V|\,.\end{aligned}$$

And so,

$$\begin{aligned}&\text{if } \mathsf{c_{max}} = \mathsf{c_{min}} = 1, \quad \text{then} \quad d' = d = (|\Pi|+1)\cdot 2\cdot|V|, \text{ and} \\ &\text{if } \mathsf{c_{max}} = \mathsf{c_{min}} > 1, \quad \text{then} \quad d' > d\,.\end{aligned}$$

When $\mathsf{K} > 1$, the comparison between d and d' depends on the values of many parameters of the game. For example, if the graph of the game has five vertices, three edges of cost 1 and one edge of cost 100, then it is more interesting to use the game $\mathcal{G}'$ and techniques from Section 3 to construct the Nash equilibrium, because in this case, $d' = (\Pi + 1)\cdot 2\cdot 104$ and $d = (|\Pi|\cdot 101 + 1)\cdot 5\cdot 101$, and so $d >> d'$.

6 Conclusion and Perspectives

In this paper, we first prove the existence of finite-memory Nash equilibria for quantitative multiplayer reachability games played on finite graphs. We also prove that this result remains true when the model is enriched by allowing n-tuples of non-negative costs on edges (one cost by player), answering a question we posed in [4]. Moreover we extend our existence result to quantitative games where both safety and reachability objectives coexist. Secondly, we prove the existence of finite-memory secure equilibria for quantitative two-player reachability games played on finite graphs.

There are several interesting directions for further research. First, we intend to investigate the existence of secure equilibria in the n-player framework. Notice that the proof techniques related to our results on secure equilibria rely on the two-player assumption. Furthermore, we also want to investigate deeper the size of the memory needed in the equilibria. This could be a first step towards a study of the complexity of computing equilibria with certain requirements, in the spirit of [9]. We also intend to look for existence results for *subgame perfect equilibria*. Finally we would like to address these questions for other objectives such as Büchi or request-response.

Acknowledgements This work has been partly supported by the ESF project GASICS and a grant from the National Bank of Belgium. The third author is supported by a grant from L'Oreal-UNESCO/F.R.S.-FNRS. The authors are grateful to Jean-François Raskin and Hugo Gimbert for useful discussions.

References

1. R. Alur, A. Kanade, and G. Weiss. Ranking automata and games for prioritized requirements. In *Computer Aided Verification, 20th International Conference, CAV 2008*, volume 5123 of *Lecture Notes in Computer Science*, pages 240–253. Springer, 2008.
2. R. Bloem, K. Chatterjee, T. Henzinger, and B. Jobstmann. Better quality in synthesis through quantitative objectives. In *CAV: Computer-Aided Verification*, volume 5643 of *Lecture Notes in Computer Science*, pages 140–156. Springer, 2009.
3. E. Boros and V. Gurvich. Why chess and back gammon can be solved in pure positional uniformly optimal strategies. Rutcor Research Report 21-2009, Rutgers University, 2009.
4. T. Brihaye, V. Bruyère, and J. De Pril. Equilibria in quantitative reachability games. In F. M. Ablayev and E. W. Mayr, editors, *CSR*, volume 6072 of *Lecture Notes in Computer Science*, pages 72–83. Springer, 2010.
5. K. Chatterjee, T. Henzinger, and M. Jurdziński. Games with secure equilibria. *Theoretical Computer Science*, 365(1-2):67–82, 2006.
6. K. Chatterjee and T. A. Henzinger. Finitary winning in omega-regular games. In *TACAS*, volume 3920 of *Lecture Notes in Computer Science*, pages 257–271. Springer, 2006.
7. E. Clarke, O. Grumberg, and D. Peled. *Model Checking.* MIT Press, Cambridge, MA, 2000.
8. E. Grädel, W. Thomas, and T. Wilke. Automata, logics, and infinite games. volume 2500 of *Lecture Notes in Computer Science*. Springer, 2002.
9. E. Grädel and M. Ummels. Solution concepts and algorithms for infinite multiplayer games. In K. Apt and R. van Rooij, editors, *New Perspectives on Games and Interaction*, volume 4 of *Texts in Logic and Games*, pages 151–178. Amsterdam University Press, 2008.
10. J. E. Hopcroft and J. D. Ullman. *Introduction to automata theory, languages, and computation.* Addison-Wesley Publishing Co., Reading, Mass., 1979. Addison-Wesley Series in Computer Science.
11. F. Horn, W. Thomas, and N. Wallmeier. Optimal strategy synthesis in request-response games. In *Automated Technology for Verification and Analysis, 6th International Symposium, ATVA 2008*, volume 5311 of *Lecture Notes in Computer Science*, pages 361–373. Springer, 2008.
12. H. Kuhn. Extensive games and the problem of information. *Classics in Game Theory*, pages 46–68, 1953.
13. D. Martin. Borel determinacy. *Ann. of Math. (2)*, 102(2):363–371, 1975.
14. J. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences of the United States of America*, 36(1):48–49, 1950.
15. M. Osborne and A. Rubinstein. *A course in game theory.* MIT Press, Cambridge, MA, 1994.
16. W. Thomas. On the synthesis of strategies in infinite games. In *STACS 95 (Munich, 1995)*, volume 900 of *Lecture Notes in Computer Science*, pages 1–13. Springer, Berlin, 1995.

17. M. Zimmermann. Time-optimal winning strategies for poset games. In *CIAA*, volume 5642 of *Lecture Notes in Computer Science*, pages 217–226. Springer, 2009.