Permissive Equilibria in Multiplayer Reachability Games

Aline Goeminne, Benjamin Monmege

Permissive Equilibria in Multiplayer Reachability Games

Aline Goeminne
F.R.S.-FNRS & UMONS – Université de Mons, Belgium

Benjamin Monmege
Aix-Marseille Univ, CNRS, LIS, Marseille, France

Abstract

We study multi-strategies in multiplayer reachability games played on finite graphs. A multi-strategy prescribes a set of possible actions, instead of a single action as usual strategies: it represents a set of all strategies that are consistent with it. We aim for profiles of multi-strategies (a multi-strategy per player), where each profile of consistent strategies is a Nash equilibrium, or a subgame perfect equilibrium. The permissiveness of two multi-strategies can be compared with penalties, as already used in the two-player zero-sum setting by Bouyer, Duflot, Markey and Renault [3]. We show that we can decide the existence of a multi-strategy profile that is a Nash equilibrium or a subgame perfect equilibrium, while satisfying some upper-bound constraints on the penalties in PSPACE, if the upper-bound penalties are given in unary. The same holds when we search for multi-strategies where certain players are asked to win in at least one play or in all plays.

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

Keywords and phrases multiplayer reachability games, penalties, permissive equilibria

Digital Object Identifier 10.4230/LIPIcs.CSL.2025.23

Related Version *Full Version*: `https://arxiv.org/abs/2411.13296` [14]

Funding *Aline Goeminne*: Postdoctoral Researcher of the Fonds de la Recherche Scientifique – FNRS.
Benjamin Monmege: This author was partially funded by ANR JCJC Quasy ANR-23-CE48-0008.

1 Introduction

Nowadays, computer systems are ubiquitous and increasingly complex. Errors in such systems can have dramatic consequences. This is why *model checking* provides a formal tool to ensure these systems are *correct* and meet certain *specifications*. *Synthesis*, on the other hand, allows for the construction of a correct-by-construction system model: concepts from *game theory* can be used for this purpose.

Two-player zero-sum games are commonly used to model a system interacting with its environment. In this model, the system aims to achieve a goal while the environment acts antagonistically to prevent it. This situation can be abstracted as a *game played on a graph* involving two players (the system and the environment). The graph represents the different possible configurations of the system, and an infinite path in this graph is a sequence of interactions between the system and the environment. In this model, building a correct system amounts to synthesizing a winning strategy, that is, a way for the system to play that ensures its goal is met regardless of the environment's behavior.

Unlike the purely antagonistic view of two-player zero-sum games, *multiplayer games* allow for modeling situations where the environment may have its own goals, or where the system consists of different interacting components, each with its own specification. In this context, the notion of a winning strategy is no longer appropriate, hence notions of equilibria

are studied: Nash equilibria or subgame perfect equilibria, which more adequately account for the sequential aspect of games played on graphs (avoiding non-credible threats). Intuitively, an equilibrium can be seen as a contract among players such that no player has an incentive to unilaterally change his strategy.

It is well known that different equilibria can coexist in the same game. In particular, a game may include an equilibrium where no player achieves his goal and an equilibrium where all players achieve their goals. The latter equilibrium is more *relevant* than the former. Therefore, it seems appropriate to focus on the existence and synthesis of relevant equilibria (according to certain relevance criteria).

Even if the synthesis process provides an equilibrium, its implementation may fail. This can be due to the occurrence of errors; for example, the action prescribed by the equilibrium may be unavailable. Synthesizing *robust equilibria* against such perturbations is therefore essential. To address these robustness issues, the classic notion of a player's strategy can be replaced by the notion of a *multi-strategy*: unlike a classic strategy that provides a single action at each decision point, a multi-strategy provides a subset of possible actions (see, for example, [2, 3]).

Intuitively, a multi-strategy is more *permissive* than another if the first allows more behaviors than the second. There are different ways to express this permissiveness. A qualitative view of permissiveness is studied in [2], where a multi-strategy is more permissive than another if the set of resulting plays includes those of the second multi-strategy. A quantitative view is addressed in [3] via the notion of *penalty* of multi-strategies, where a cost is associated with each edge not chosen by the multi-strategy. Thus, the penalty of a multi-strategy is the highest sum of blocked edges along a play consistent with the multi-strategy.

Related works. In [2], permissiveness in parity games (a highly expressive winning condition) is studied: considering the qualitative view of permissiveness, there does not necessarily exist a most permissive strategy. However, one exists when restricted to memoryless strategies (which always make the same decision in any given vertex of the game). By reducing to safety games, the authors show that it is possible to compute the most permissive strategy. In [3], the above-mentioned quantitative view of permissiveness is implemented. Several penalty measures and games are used, and the complexity of computing the most permissive strategies in this context is given. More general parity objectives are then studied in [5]. Recently, other methods have explored permissiveness in two-player games using templates to concisely represent multiple strategies in graph games [1]. This approach is also used in multiplayer games for the synthesis of secure equilibria [16].

Independently, different equilibria (Nash or subgame perfect) have been studied in multiplayer games to ensure a strategy profile where no player has an incentive to deviate. Several works have characterized such equilibria and studied the complexity of decision problems related to the existence of relevant equilibria. Notably, these works have focused on the study and characterization of (*i*) Nash equilibria in games where players have classical ω-regular objectives [17, 11], (*ii*) weak subgame perfect equilibria (a variant of subgame perfect equilibria) where players have classical ω-regular objectives [9] (this work also characterizes subgame perfect equilibria when the studied objectives are either qualitative reachability or safety objectives); (*iii*) subgame perfect equilibria for games with quantitative reachability objectives [10]; (*iv*) subgame perfect equilibria for games with parity objectives [6] or (*v*) mean-payoff objectives [8, 7].

Contribution. Our goal is to combine these two research directions by studying permissiveness in strategy profiles that describe equilibria (Nash or subgame perfect). In this first work, we focus on reachability games only. We study permissive strategy profiles such that all the fully described strategy profiles they contain are equilibria. The motivation is to allow greater latitude and robustness of equilibrium profiles without losing quality in the final goal of secure synthesis. With the qualitative view, as in the two-player game framework [2], it is not difficult to show that there does not necessarily exist most permissive profiles that are Nash equilibria (or subgame perfect equilibria). We will thus consider a quantitative view of permissiveness similar to the penalty measures introduced in [3] for two-player games. We obtain a characterization based on trees, and decision algorithms with penalties bounded by a given threshold in polynomial space with respect to the size of the game and the maximal penalty bound (if this is encoded in unary). We also solve the problem of synthesis of robust and relevant equilibria, where the relevance is the constraint that all derived equilibria ensure that all players in a fixed subset satisfy their objective (*strongly winning*), or that at least one derived equilibrium ensures this guarantee (*weakly winning*).

All missing proofs can be found in the long version of this article [14].

2 Multiplayer reachability games

A *(multiplayer) reachability games* is a tuple $(\mathrm{N}, V, (V_i)_{i\in\mathrm{N}}, E, (\mathrm{F}_i)_{i\in\mathrm{N}}, v_0)$, that we denote $(\mathcal{G}, v_0)$ to emphasize the v_0 component, where $\mathrm{N} = \{1, \dots, n\}$ is a finite set of n players, (V, E) is a finite directed graph without deadlocks (for all $v \in V$, there exists $v' \in V$ such that $(v, v') \in E$), $(V_i)_{i\in\mathrm{N}}$ is a partition of V between the players, $\mathrm{F}_i \subseteq V$ is the set of target vertices, called *target set*, of player $i \in \mathrm{N}$, and v_0 is an initial vertex. Given a vertex $v \in V$, we let $\mathrm{Succ}(v) = \{v' \in V \mid (v, v') \in E\}$ be the set of all successors of v.

A *play* in $\mathcal{G}$ is an infinite sequence of vertices consistent with the graph structure, *i.e.,* if $\rho = \rho_0\rho_1\cdots$ is a play, then for all $k \in \mathbb{N}$, $\rho_k \in V$ and $(\rho_k, \rho_{k+1}) \in E$. The set of plays is denoted by Plays, while $\mathrm{Plays}(v)$ denotes the set of plays beginning in v. Given a play $\rho = \rho_0\rho_1\cdots$ and $k \in \mathbb{N}$, $\rho_{\geq k}$ is the suffix $\rho_k\rho_{k+1}\cdots$ of ρ.

For each player $i \in \mathrm{N}$, we let Gain_i be the *gain function* that associates with each play the value 1 if the play is winning for player i, 0 if it is losing. For a reachability game as above, we have $\mathrm{Gain}_i(\rho) = 1$ iff player i reaches his target set in ρ, *i.e.,* $\rho = \rho_0\rho_1\cdots$ and there exists $k \in \mathbb{N}$ with $\rho_k \in \mathrm{F}_i$. **In the rest of this article, $(\mathcal{G}, v_0)$ will always denote a reachability game associated with these gain functions.**

A *history* is a finite sequence of vertices $h = h_0h_1\cdots h_k$ with $k \in \mathbb{N}$ defined similarly. The set of histories is denoted by Hist, while $\mathrm{Hist}(v)$ denotes the set of histories beginning in v. For all $i \in \mathrm{N}$, we write Hist_i to denote the set of histories ending in a vertex owned by player i. If $h = h_0\cdots h_k$ with $k \in \mathbb{N}$ is a history, $\mathrm{Last}(h)$ denotes the last vertex h_k, while $|h|$ denotes its length k. Given a history $h = h_0\cdots h_k$, $\mathrm{Visit}(h) = \{i \in \mathrm{N} \mid \exists 1 \leq \ell \leq k\ h_\ell \in F_i\}$ is the set of players who visit their target set along h.

A *strategy* of player i is a function $\sigma_i \colon \mathrm{Hist}_i(v_0) \to V$ that assigns to each history $hv \in \mathrm{Hist}_i(v_0)$ a vertex v' such that $(v, v') \in E$. A play $\rho = \rho_0\rho_1\cdots$ is consistent with a strategy σ_i if for all $\rho_k \in V_i$, $\sigma_i(\rho_0\cdots\rho_k) = \rho_{k+1}$. A *strategy profile* is a tuple $\sigma = (\sigma_i)_{i\in\mathrm{N}}$ of strategies, one per player: there is a unique play from v_0 which is consistent with each strategy σ_i, and we call this play the *outcome* of σ, denoted by $\langle\sigma\rangle_{v_0}$. To highlight the role of player i, we sometimes write $\sigma = (\sigma_i, \sigma_{-i})$ where σ_{-i} denotes the strategy profile of the players other than player i.

The strategy profile σ is a *Nash equilibrium* (NE) in $(\mathcal{G}, v_0)$ if no player has an incentive to deviate unilaterally from his strategy to increase his gain, *i.e.,* if for all players $i \in \mathrm{N}$ and all strategies σ'_i of player i, $\mathrm{Gain}_i(\langle\sigma\rangle_{v_0}) \geq \mathrm{Gain}_i(\langle\sigma'_i, \sigma_{-i}\rangle_{v_0})$.

The concept of *subgame perfect equilibrium* (SPE) takes more into account the sequential nature of games played on graphs by avoiding non-credible threat, a well-known weakness of NEs in this setting. Informally, a strategy profile is an SPE if it is an NE in all subgames. Given a history $hv \in \mathrm{Hist}(v_0)$, the *subgame* $(\mathcal{G}_{\restriction h}, v)$ is obtained from $\mathcal{G}$ by changing the initial vertex to v, and by considering the gain functions $(\mathrm{Gain}_{i\restriction h})_{i\in\mathrm{N}}$ taking into account the players that have won in history h: we thus write, for each $i \in \mathrm{N}$, $\mathrm{Gain}_{i\restriction h}(\rho) = \mathrm{Gain}_i(h\rho)$ for all $\rho \in \mathrm{Plays}(v)$. Moreover, if σ_i is a strategy of player i in $\mathcal{G}$, then $\sigma_{i\restriction h}$ is the strategy of player i in the subgame $(\mathcal{G}_{\restriction h}, v)$ such that for all $h' \in \mathrm{Hist}_i(v)$, $\sigma_{i\restriction h}(h') = \sigma_i(hh')$. In the same way, from a strategy profile σ in $\mathcal{G}$, we can derive a strategy profile $\sigma_{\restriction h}$ in $(\mathcal{G}_{\restriction h}, v)$. We now define formally the concept of SPEs: a strategy profile σ is an SPE in $\mathcal{G}$ if for all $i \in \mathrm{N}$, for all $hv \in \mathrm{Hist}_i(v_0)$, $\sigma_{\restriction h}$ is an NE in $(\mathcal{G}_{\restriction h}, v)$. Notice that an SPE is an NE and that there always exists an SPE (and thus an NE) in a reachability game [17].

3 Permissiveness in strategies

Our goal is to allow for some permissiveness in strategies of all players, *i.e.,* being able to underspecify the strategies of the players, while maintaining that they describe an NE or an SPE.

A *multi-strategy* of player i is a function $\Theta_i \colon \mathrm{Hist}_i(v_0) \to 2^V \setminus \{\emptyset\}$ that assigns to each history $hv \in \mathrm{Hist}_i(v_0)$ a non-empty set of vertices $\mathrm{A} \subseteq V$ such that for all $v' \in \mathrm{A}$, $(v, v') \in E$. Notice that a strategy σ_i can be seen as a multi-strategy Θ_i where, for all $hv \in \mathrm{Hist}_i(v_0)$, $\Theta_i(hv)$ is the singleton $\{\sigma_i(hv)\}$. A *multi-strategy profile* $\Theta = (\Theta_i)_{i\in\mathrm{N}}$ is a tuple of multi-strategies, one per player.

Unlike strategies, when we fix a game $\mathcal{G}$ and a multi-strategy profile Θ, there exist several plays beginning in v_0 that are consistent with all the multi-strategies Θ_i. To describe them, we say that a strategy σ_i is *consistent* with a multi-strategy Θ_i, written $\sigma_i \lesssim \Theta_i$ if for all $hv \in \mathrm{Hist}_i(v_0)$, $\sigma_i(hv) \in \Theta_i(hv)$. We extend this notation to profiles of strategies, as expected. Then, we let $\langle\Theta\rangle_{v_0}$ be the set of plays $\langle\sigma\rangle_{v_0}$ for all profiles σ of strategies consistent with the multi-strategy Θ. We call this set the *outcomes of* Θ. We also let $\langle\Theta\rangle^{\mathrm{H}}_{v_0}$ be the set of histories consistent with the multi-strategy Θ, *i.e.,* the finite prefixes of plays in $\langle\Theta\rangle_{v_0}$.

Our goal is to compute profiles of multi-strategies such that all profiles of consistent strategies are NEs or SPEs: such profiles of multi-strategies are called permissive NEs or permissive SPEs. By the existence of NEs and SPEs in reachability games, we straightforwardly obtain the existence of permissive NEs and permissive SPEs. We thus want to study *most permissive* NEs or SPEs, *i.e.,* profiles of multi-strategies that are permissive NEs or SPEs, and such that no "more permissive" multi-strategies are still permissive NEs or SPEs.

The natural first attempt would be to look for a notion of "more permissive" that is set-theoretic, with respect to a given solution concept. We would thus say that a profile of multi-strategies Θ is at least as permissive as a profile of multi-strategies Θ' if for all $i \in \mathrm{N}$, for all histories $h \in \mathrm{Hist}_i(v_0)$, $\Theta_i(h) \supseteq \Theta'_i(h)$. Then, Θ would be more permissive than Θ' if it is at least as permissive, while being different (for at most one history). Finally, Θ would be a *most permissive* NE or SPE if it is a permissive NE or SPE, respectively, and no permissive NE or SPE, respectively, is more permissive than Θ.

This natural definition is very problematic in the realm of reachability games (as already noticed in the context of winning strategies in parity games by [2]) where no most permissive NE or SPE could exist, as demonstrated by the game in Figure 1.

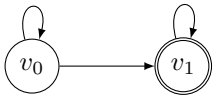

Figure 1 In this game, player 1 owns all vertices and wants to reach v_1. For all $k \in \mathbb{N}$, we define the multi-strategy Θ_1^k such that for all $h \in \mathrm{Hist}(v_0)$, $\Theta_1^k(h) = \{v_0, v_1\}$ if $\mathrm{Last}(h) = v_0$ and $|\{n \in \mathbb{N} \mid h_n = v_0\}| \leq k$, and $\Theta_1^k(h) = \{v_1\}$ otherwise. We have that for all $k \in \mathbb{N}$, for all $\sigma_1 \lesssim \Theta_1^k$, $\mathrm{Gain}_1(\langle\sigma_1\rangle_{v_0}) = 1$ (and thus Θ_1^k is a permissive SPE), but for all $k \in \mathbb{N}$, $\langle\Theta_1^k\rangle_{v_0} \subseteq \langle\Theta_1^{k+1}\rangle_{v_0}$.

We thus propose another way to measure the permissiveness of a multi-strategy, inspired by the definition of penalty used in [3] to describe permissive winning strategies in two-player games. To define the notion of penalty in our context, we equip the game with a function $w\colon E \to \mathbb{N}$ assigning a non-negative weight to each edge: if unspecified, we will consider that every edge has weight 1. The player who owns the vertex at the source of an edge e will pay the penalty $w(e)$ if he decides to not include the edge e in his multi-strategy. All penalties are then counted additively. Formally, for a multi-strategy profile Θ, we first define for each player $i \in \mathrm{N}$ the *penalty of player i w.r.t.* Θ in a play $\rho = \rho_0\rho_1\cdots$ by induction on the length of its prefixes:

- $\mathrm{Penalty}_i^\Theta(\varepsilon) = 0$ where ε denotes the empty prefix;
- for $h = \rho_0\cdots\rho_k$, $\mathrm{Penalty}_i^\Theta(hv) = \begin{cases} \mathrm{Penalty}_i^\Theta(h) + \sum_{v' \in \mathrm{Succ}(v)\setminus\Theta_i(hv)} w(v,v') & \text{if } v \in V_i \\ \mathrm{Penalty}_i^\Theta(h) & \text{otherwise} \end{cases}$;
- $\mathrm{Penalty}_i^\Theta(\rho) = \lim_{k\to+\infty} \mathrm{Penalty}_i^\Theta(\rho_0\cdots\rho_k)$: this limit exists (it may be equal to $+\infty$) since $(\mathrm{Penalty}_i^\Theta(\rho_0\cdots\rho_k))_k$ is a non-decreasing sequence of natural numbers.

There are several ways to associate a penalty with a multi-strategy profile Θ, depending on how we take into account the non-determinism offered in the multi-strategies. A first choice consists in considering a worst-case scenario in the outcomes (without considering the possible deviations). A second choice consists in considering only the deviations of one player, *i.e.,* to consider that the retaliation of other players with respect to the deviation of a player will count in the final penalty. It is then possible to combine both types of penalties, though we will treat them separately in the rest of this article.

▶ **Definition 1** (Penalties). *Let Θ be a multi-strategy profile in $(\mathcal{G}, v_0)$. The* main penalty *and* retaliation penalty *of player i with respect to Θ are defined respectively as*

$$\mathrm{MPenalty}_i(\Theta, v_0) = \sup_{\rho \in \langle\Theta\rangle_{v_0}} \mathrm{Penalty}_i^\Theta(\rho)$$

$$\mathrm{RPenalty}_i(\Theta, v_0) = \sup_{hv \in \mathrm{Hist}_i(v_0)\setminus\langle\Theta\rangle_{v_0}^{\mathrm{H}}} \sup\{\mathrm{Penalty}_i^\Theta(\rho) \mid \rho \in \langle\Theta_{\restriction hv}\rangle_v\}$$

If there are no histories hv in $\mathrm{Hist}_i(v_0) \setminus \langle\Theta\rangle_{v_0}^{\mathrm{H}}$, we let $\mathrm{RPenalty}_i(\Theta, v_0) = 0$.

The existence of a multi-strategy profile which satisfies some upper-bounds on penalties does not provide any certainty about the satisfaction of the reachability objectives of the players. For this reason, we also consider multi-strategy profiles that satisfy some properties on the set of players who satisfy their objective. Let Win be a subset of players and Θ be a multi-strategy profile. Then, Θ is said *weakly winning* if **there exists** a strategy profile σ which is consistent with Θ and such its outcome is winning for all players in Win. Similarly, Θ is said *strongly winning* if **for each** strategy profile σ which is consistent with Θ, its outcome is winning for all players in Win.

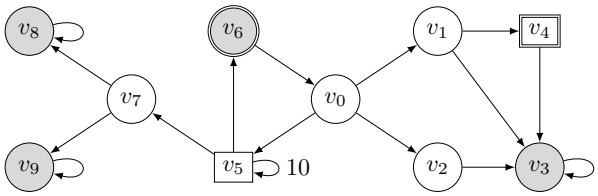

Figure 2 An example of a reachability game where player 1 (resp. player 2) owns circle (resp. rectangle) vertices. The initial vertex is v_0. Target vertices $F_1 = \{v_3, v_6, v_8, v_9\}$ of player 1 and $F_2 = \{v_4, v_6\}$ of player 2 are drawn with gray vertices and double-bordered vertices respectively.

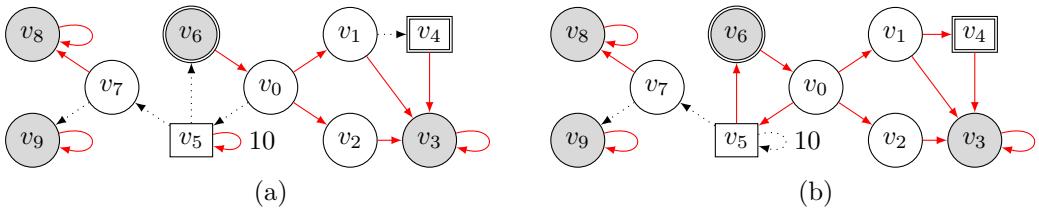

Figure 3 Examples of permissive equilibria: (a) a permissive NE and (b) a permissive SPE.

▶ **Definition 2** (Weakly and strongly winning). *Given a subset of player* Win $\subseteq$ N *and a multi-strategy profile* Θ*,*

- Θ *is said* weakly winning with respect to Win *if there exists a strategy profile* σ *such that* $\sigma \lesssim \Theta$ *and for all* $i \in$ Win, $\mathrm{Gain}_i(\langle\sigma\rangle_{v_0}) = 1$.
- Θ *is said* strongly winning with respect to Win *if for all strategy profiles* σ *such that* $\sigma \lesssim \Theta$*, we have that for all* $i \in$ Win, $\mathrm{Gain}_i(\langle\sigma\rangle_{v_0}) = 1$.

▶ **Example 3.** An example of a reachability game with two players is depicted in Figure 2. The edge labelled with 10 corresponds to the penalty if player 2 decides not to allow this edge: all other penalties are set to 1 by default. A multi-strategy is represented with red edges (black dotted edges are thus the ones that are not selected in the multi-strategy) in Figure 3(a).[1] All strategy profiles that are consistent with this multi-strategy depend on the choice of successor for v_0 among $\{v_1, v_2\}$. It is indeed a permissive NE since the consistent strategies are NEs: player 1 has no interest in deviating from either v_1 or v_2 in v_0, since all strategies lead to plays where he visits his target set, while going to v_5 make him lose. It has a main penalty of 2 for player 1 and 0 for player 2. Player 1 can do slightly better by allowing the edge (v_1, v_4) in the multi-strategy: this remains a permissive NE (now player 2 wins in certain plays, but he is left with no real choices to make), and player 1 now gets a main penalty of 1. This modified permissive NE is strongly winning w.r.t. $\{1\}$, and weakly winning w.r.t. $\{1, 2\}$. It is not a permissive SPE since player 2 has a profitable deviation from v_5 by going to v_6 where he wins. A permissive SPE is depicted in Figure 3(b), that is strongly winning w.r.t. $\{1\}$, but only weakly winning w.r.t $\{1, 2\}$. Player 2 has a main penalty of 11 (because he cuts edges (v_5, v_5) and (v_5, v_7)), while player 1 has a retaliation penalty of 1 (because he cuts edge (v_7, v_9)). If we want a permissive SPE that is strongly winning w.r.t. $\{1, 2\}$, we need to increase the main penalty of player 1 to 2 by removing edges (v_1, v_3) and (v_0, v_2). However, we may decrease to 0 the retaliation penalty of player 1 by adding the edge (v_7, v_9) (since it is equally good to him anyway).

[1] Notice that, in this example, the set of successors prescribed by multi-strategies only depends on the current vertex and not on the past history.

We now define the problems we study in the rest of the article, where we use the word "equilibrium" to either mean NE or SPE, depending on the solution concept we want to check. In all these problems, we give different penalty bounds for the main penalty and the retaliation penalty. Notice though that the bounds can be set to $+\infty$, relaxing the constraints in this case.

▶ **Problem 1** (Constrained penalty problem). *Given a reachability game* $(\mathcal{G}, v_0)$*,* $m \in (\mathbb{N} \cup \{\infty\})^n$ *and* $r \in (\mathbb{N} \cup \{\infty\})^n$*, does there exist a permissive equilibrium* Θ *in* $(\mathcal{G}, v_0)$ *such that for all* $i \in \mathrm{N}$*,* $\mathrm{MPenalty}_i(\Theta, v_0) \leq m_i$ *and* $\mathrm{RPenalty}_i(\Theta, v_0) \leq r_i$*?*

▶ **Problem 2** (Weakly winning with constrained penalty problem). *Given a reachability game* $(\mathcal{G}, v_0)$*,* $m \in (\mathbb{N} \cup \{\infty\})^n$*,* $r \in (\mathbb{N} \cup \{\infty\})^n$ *and* $\mathrm{Win} \subseteq \mathrm{N}$*, does there exist a permissive equilibrium* Θ *in* $(\mathcal{G}, v_0)$ *such that* (i) *for all* $i \in \mathrm{N}$*,* $\mathrm{MPenalty}_i(\Theta, v_0) \leq m_i$ *and* $\mathrm{RPenalty}_i(\Theta, v_0) \leq r_i$ *and* (ii) Θ *is weakly winning w.r.t.* Win*?*

▶ **Problem 3** (Strongly winning with constrained penalty problem). *Given a reachability game* $(\mathcal{G}, v_0)$*,* $m \in (\mathbb{N} \cup \{\infty\})^n$*,* $r \in (\mathbb{N} \cup \{\infty\})^n$ *and* $\mathrm{Win} \subseteq \mathrm{N}$*, does there exist a permissive equilibrium* Θ *in* $(\mathcal{G}, v_0)$ *such that* (i) *for all* $i \in \mathrm{N}$*,* $\mathrm{MPenalty}_i(\Theta, v_0) \leq m_i$ *and* $\mathrm{RPenalty}_i(\Theta, v_0) \leq r_i$ *and* (ii) Θ *is strongly winning w.r.t.* Win*?*

We show in the rest of this article that all these problems, for NEs and SPEs, are decidable in PSPACE, if the upper-bound penalties are encoded in unary. To do so, we characterize the permissive equilibria in the various problems in Section 4. In Section 5, we then show that tree-like witnesses can be found if the according permissive equilibria exist. These witnesses have a height bounded by a polynomial depending on the size of the game and the largest upper-bound on penalties. We use these witnesses to obtain the PSPACE decision procedures.

4 Characterizations of permissive equilibria

We now characterize permissive equilibria of the reachability game $(\mathcal{G}, v_0)$. This is a first step towards their computation in the next section. We provide a characterization for permissive NEs in Section 4.1 and one for permissive SPEs in Section 4.2. These characterizations are inspired by existing ones for classical NEs (resp. SPEs) [11, 9]. The latter rely on properties that a play (resp. a set of plays) must satisfy in order to be the outcome of an NE (resp. the set of subgame outcomes of an SPE). However, the outcomes of permissive equilibria are a set of plays and not a simple play. For that reason, the characterizations of permissive equilibria employ trees that we first formally define.

Trees. We call *tree over* $\mathcal{G}$ *rooted at* v (for some $v \in V$) any subset $\mathcal{T}$ of non-empty histories of $\mathcal{G}$ that contains v and such that if $hu \in \mathcal{T}$ then $h \in \mathcal{T}$. All $h \in \mathcal{T}$ are called *nodes* of the tree, the particular node v is called the *root* of the tree, and for all $hu \in \mathcal{T}$, h is called the *parent* of hu, and hu a *child* of h.

As for histories in an arena, for all $hu \in \mathcal{T}$, we let $\mathrm{Last}(hu) = u$. The *depth* of a node $h \in \mathcal{T}$, written $\mathrm{depth}(h)$, is equal to $|h|$ and its *height*, denoted by $\mathrm{height}(h)$, is given by $\sup\{|\,\mathrm{Last}(h)h'|\,\mid h' \in \mathrm{Hist} \text{ and } hh' \in \mathcal{T}\}$. The height of the tree corresponds to the height of its root. A node $h \in \mathcal{T}$ is called a *leaf* if $\mathrm{height}(h) = 0$.

We denote by $\mathcal{T}_{\restriction hu}$, the subtree of $\mathcal{T}$ rooted at u for some $hu \in \mathcal{T}$, that is the set of non-empty histories $h' \in \mathrm{Hist}(u)$ such that $hh' \in \mathcal{T}$.

A (finite or infinite) *branch* of the tree is a maximal (finite or infinite) sequence of nodes $h_0h_1\cdots$ such that for all $k \in \mathbb{N}$, h_k is the parent of h_{k+1}. Finally, we denote by $\mathcal{T}^\infty$ the set of plays in $\mathcal{G}$ represented by infinite branches in $\mathcal{T}$, *i.e.,*

$$\mathcal{T}^\infty = \{\rho_0\rho_1\cdots \in \text{Plays} \mid \text{there exists a branch } h_0h_1\cdots \in \mathcal{T} \text{ st. } \forall k \in \mathbb{N},\ \rho_k = \text{Last}(h_k)\}$$

In what follows, we consider outcomes of multi-strategies as trees. Indeed, given a multi-strategy Θ, $\langle\Theta\rangle^{\text{H}}_{v_0}$ can be seen as a tree $\mathcal{T}$ over $\mathcal{G}$ rooted at v_0 and $\langle\Theta\rangle_{v_0}$ corresponds to $\mathcal{T}^\infty$. In particular, penalties can also be defined on trees, mimicking the definition for profiles of multi-strategies. The penalty of a tree $\mathcal{T}$ for a player i, denoted by $\text{Penalty}_i(\mathcal{T})$, is the maximal penalty of a branch of $\mathcal{T}$, the penalty of a branch being equal to the penalty of the associated play ρ w.r.t. any profile of multi-strategies that is consistent with the choices appearing in $\mathcal{T}$ along the play ρ. Formally, let $\mathcal{T}$ be a tree and $i \in \text{N}$ be a player. For each $hv = v_1\cdots v_kv \in \mathcal{T}$, we define $\text{Blocked}(h) = \{u \in \text{Succ}(v_k) \mid hu \notin \mathcal{T}\}$ as the set of blocked successors of h in $\mathcal{T}$ and

$$\text{Penalty}_i(hv) = \begin{cases} 0 & \text{if } h = \varepsilon \\ \text{Penalty}_i(h) + \sum_{u\in\text{Blocked}(h)} w(v_k,u) & \text{if } v_k \in V_i \\ \text{Penalty}_i(h) & \text{otherwise.} \end{cases}$$

Moreover, for all plays $\rho = \rho_0\rho_1\cdots \in \mathcal{T}^\infty$, we let $\text{Penalty}_i(\rho) = \lim_{k\to+\infty} \text{Penalty}_i(\rho_0\cdots\rho_k)$. Thus, the penalty of a tree $\mathcal{T}$ for a player $i \in \text{N}$ is naturally defined as:

$$\text{Penalty}_i(\mathcal{T}) = \sup\{\text{Penalty}_i(\rho) \mid \rho \in \mathcal{T}^\infty\}.$$

4.1 Characterization of permissive Nash equilibria

In order to characterize permissive Nash equilibria, we start by defining *good* trees, by checking two conditions. The first one, called *resistance to internal deviations*, means that at any node h of the tree such that $\text{Last}(h)$ belongs to player i, if h has at least two children, the plays starting with h are either all losing, or all winning, for player i. The second one, called *resistance to external deviations*, means that at any node hu of the tree with u belonging to player i, if player i has the possibility to play to a successor u' not in the tree from which he has a winning strategy, then all plays in the subtree from hu must be winning for player i.

▶ **Definition 4.** *Let $\mathcal{T}$ be a tree over $(\mathcal{G}, v_0)$.*

1. *Given a subset of players $\text{D} \subseteq \text{N}$, the tree $\mathcal{T}$ is* D-resistant to internal deviations *if for all $i \in \text{D}$ and for all $hv \in \mathcal{T}$ such that $v \in V_i$ and $|\{hvv' \in \mathcal{T} \mid v' \in V\}| \geq 2$, we have that for all $\rho, \rho' \in \mathcal{T}^\infty_{\restriction hv}$, $\text{Gain}_i(h\rho) = \text{Gain}_i(h\rho')$. If $\text{D} = \text{N}$, we simply say that $\mathcal{T}$ is* resistant to internal deviations.
2. *The tree $\mathcal{T}$ is* resistant to external deviations *if for all $hu \in \mathcal{T}$ with $u \in V_i$ and $i \notin \text{Visit}(hu)$, if there exists $u' \in \text{Succ}(u)$ such that $huu' \notin \mathcal{T}$ and player i has a winning strategy from u' (against the coalition of the other players), then for all plays $\rho \in \mathcal{T}^\infty_{\restriction hu}$, $\text{Gain}_i(\rho) = 1$.*
3. *The tree $\mathcal{T}$ is* good *if it is resistant to internal and external deviations.*

The resistance to internal and external deviations leads to the characterization of outcomes of permissive NEs (Theorem 5): given a good tree $\mathcal{T}$, there exists a permissive NE such that its outcomes are the plays corresponding to the infinite branches of $\mathcal{T}$ iff $\mathcal{T}$ is good.

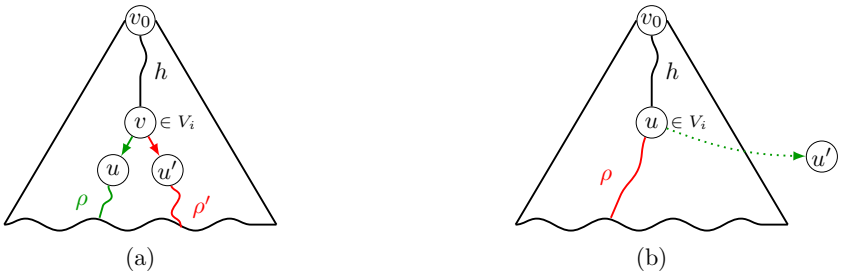

Figure 4 Examples of trees that do not respect: (a) the resistance to internal deviations since $\mathrm{Gain}_i(\rho') = 0$ but $\mathrm{Gain}_i(\rho) = 1$; (b) the resistance to external deviations since $\mathrm{Gain}_i(\rho) = 0$ but Player i can win from u'.

▶ **Theorem 5.** *Let $\mathcal{T}$ be a tree over $(\mathcal{G}, v_0)$ rooted at v_0. The following assertions are equivalent:*

1. *There exists a permissive NE Θ in $(\mathcal{G}, v_0)$ such that $\langle\Theta\rangle^{\mathrm{H}}_{v_0} = \mathcal{T}$;*
2. *The tree $\mathcal{T}$ is good.*

▶ Remark 6. For all multi-strategies Θ, and all players $i \in \mathrm{N}$, the penalty $\mathrm{MPenalty}_i(\Theta, v_0)$ is equal to the penalty of player i in the good tree $\langle\Theta\rangle^{\mathrm{H}}_{v_0}$, *i.e.,* $\mathrm{Penalty}_i(\langle\Theta\rangle^{\mathrm{H}}_{v_0})$. The construction of Theorem 5 thus also preserves the main penalties.

Proof sketch. For $(1 \Rightarrow 2)$ let us assume that Θ is a permissive NE and that $\langle\Theta\rangle^{\mathrm{H}}_{v_0} = \mathcal{T}$. We have to prove that $\mathcal{T}$ is good. If $\mathcal{T}$ is not resistant to internal deviations that means that from some vertex v there exists two plays ρ, crossing u, and ρ', crossing $u' \neq u$, such that: ρ is winning for player i and ρ' is losing for player i, see Figure 4(a). In particular, we can build a strategy profile σ consistent with Θ such that $\langle\sigma_{\restriction h}\rangle_v = \rho'$ and $\langle\sigma_{\restriction hv}\rangle_u = \rho_{\geq 1}$. Meaning that player i should deviate by choosing u instead of u' from v, meaning that σ is not an NE and Θ is not a permissive NE. If $\mathcal{T}$ is not resistant to external deviations, that means that from some vertex u of player i there exists a play ρ such that $\mathrm{Gain}_i(\rho) = 0$ and u' a successor of u outside $\mathcal{T}$ from which player i can win, see Figure 4(b). Thus we can build a strategy profile σ consistent with Θ such that $\langle\sigma\rangle_{v_0} = h\rho$. In this way, player i should choose to go in u' and then follow a winning strategy meaning that σ is not an NE and Θ not a permissive NE.

For $(2 \Rightarrow 1)$, let us assume that $\mathcal{T}$ is a good tree. We build a permissive NE Θ such that its outcomes are the plays corresponding to the infinite branches of $\mathcal{T}$. Additionally, if a player i deviates from $\mathcal{T}$, the coalition of the other players plays its retaliation[2] strategy to prevent player i from deviating. ◀

4.2 Characterization of permissive subgame perfect equilibria

Permissive subgame perfect equilibria are intrinsically more complex than permissive Nash equilibria. Thus their characterization cannot only rely on the outcomes from the initial vertex, it should also take into account the outcomes in all subgames. This is the reason why, in order to deal with a compact representation of outcomes of a permissive SPE and its subgames, we introduce the notion of *forest*. Then, we generalize the definition of good trees to define *good forests* needed to characterize SPEs instead of NEs.

[2] This retaliation strategy corresponds to the winning strategy of player 2 in a two-player zero-sum reachability game in which player 1 is player i and wants to reach F_i and player 2 is the coalition of the other players and wants to avoid visiting F_i [15, Chapter 2].

Forests and penalties of forests. Trees of the forest are indexed by tuples $(i, v, \mathrm{I}) \in \mathrm{N} \times V \times 2^{\mathrm{N}}$. More precisely, we let

$$\mathcal{I} = \{(0, v_0, \mathrm{I}_0)\} \cup \{(i, v, \mathrm{I}) \in \mathrm{N} \times V \times 2^{\mathrm{N}} \mid \exists hv' \in \mathrm{Hist}_i(v_0) \text{ st. } v \in \mathrm{Succ}(v') \wedge \mathrm{I} = \mathrm{Visit}(hv'v)\}$$

where $\mathrm{I}_0 = \{i \in \mathrm{N} \mid v_0 \in F_i\}$. Apart from the special tuple $(0, v_0, \mathrm{I}_0)$, a tuple (i, v, I) represents the fact that v is a vertex played by player i and reachable from v_0, and that all players in I have already seen their target when v is reached. A forest in $(\mathcal{G}, v_0)$ is thus a set of trees $\mathcal{F} = \{\mathcal{T}_{i,v,I} \mid (i, v, \mathrm{I}) \in \mathcal{I}\}$ such that $\mathcal{T}_{i,v,\mathrm{I}}$ is a tree without leaves over $\mathcal{G}$ rooted at v. The intuition behind this object is that the tree $\mathcal{T}_{0,v_0,\mathrm{I}_0}$ represents the outcomes of a multi-strategy Θ and the other trees $\mathcal{T}_{i,v,\mathrm{I}}$ represent the outcomes of $\Theta_{\restriction hv'}$ in the subgames $(\mathcal{G}_{\restriction hv'}, v)$ for all $hv' \in \mathrm{Hist}_i(v_0)$ such that $\mathrm{Visit}(hv'v) = \mathrm{I}$.

Moreover the main (resp. retaliation) penalty of a forest $\mathcal{F}$ for a player $i \in \mathrm{N}$ are respectively given by

$$\mathrm{MPenalty}_i(\mathcal{F}) = \mathrm{Penalty}_i(\mathcal{T}_{0,v_0,\mathrm{I}_0}) \quad \text{and} \quad \mathrm{RPenalty}_i(\mathcal{F}) = \sup_{\substack{\mathcal{T}_{i,v,I} \in \mathcal{F} \\ (i,v,\mathrm{I}) \in \mathrm{Out}}} \mathrm{Penalty}_i(\mathcal{T}_{i,v,I})$$

where $\mathrm{Out} = \{(i, v, I) \in \mathcal{I} \setminus \{(0, v_0, \mathrm{I}_0)\} \mid \exists hv \in \mathrm{Hist}(v_0) \text{ st. } hv \notin \mathcal{T}_{(0,v_0,\mathrm{I}_0)} \wedge \mathrm{Last}(h) \in V_i \wedge \mathrm{I} = \mathrm{Visit}(hv)\}$ described the indices of trees in the forest that are deviations from the main tree $\mathcal{T}_{0,v_0,\mathrm{I}_0}$. If Out is empty, we let $\mathrm{RPenalty}_i(\mathcal{F}) = 0$.

Characterization. Following the same philosophy as for permissive NEs, a forest is *good* if each tree $\mathcal{T}_{i,v,\mathrm{I}}$ of the forest satisfies two properties. The first one is that $\mathcal{T}_{i,v,\mathrm{I}}$ has to be $(\mathrm{N} \setminus \mathrm{I})$-resistant to internal deviations, exactly as for permissive NEs except that we take into account players who have already visited their target set, *i.e.,* players in I. The second one, called *resistance to constrained external deviations*, means that at any node hu of the tree such that u belongs to player j, if player j has the possibility to jump to another tree $\mathcal{T}_{j,u',\mathrm{I}'}$ by playing to a successor u' not in the tree and if there exists a play in this latter tree which is winning for player j, then all plays after hu in $\mathcal{T}_{i,v,\mathrm{I}}$ have to be winning for player j.

▶ **Definition 7** (Good forest). *Let $\mathcal{F}$ be a forest in $(\mathcal{G}, v_0)$.*

1. *A tree $\mathcal{T}_{i,v,\mathrm{I}} \in \mathcal{F}$ is* resistant to constrained external deviations *if it satisfies the following property: for all $hu \in \mathcal{T}_{i,v,\mathrm{I}}$ and $j \in \mathrm{N}$ such that we have that* (i) *$u \in V_j$ and $j \notin I \cup \mathrm{Visit}(hu)$ and* (ii) *there exists $u' \in \mathrm{Succ}(u)$ such that $huu' \notin \mathcal{T}_{i,v,\mathrm{I}}$, if there exists $\rho' \in \mathcal{T}^{\infty}_{j,u',\mathrm{I}'}$, where $\mathrm{I}' = \mathrm{I} \cup \mathrm{Visit}(huu')$, such that $\mathrm{Gain}_j(\rho') = 1$, then for all $\rho \in \mathcal{T}^{\infty}_{i,v,\mathrm{I}\restriction hu}$, $\mathrm{Gain}_j(\rho) = 1$.*
2. *The forest $\mathcal{F}$ is* good *if each tree $\mathcal{T}_{i,v,\mathrm{I}} \in \mathcal{F}$ is $(\mathrm{N} \setminus \mathrm{I})$-resistant to internal deviations (see* (1) *in Definition 4) and resistant to constrained external deviations.*

Thanks to good forests, we are able to characterize the outcomes of permissive SPEs: given a good tree $\mathcal{T}^*$, there exists a permissive SPE such that its outcomes correspond to $\mathcal{T}^*$ iff there exists a good forest whose "main" tree is $\mathcal{T}^*$, *i.e.,* $\mathcal{T}_{0,v_0,\mathrm{I}_0} = \mathcal{T}^*$. With some other constraints, this also preserves strongly (resp. weakly) winning and penalty properties.

▶ **Theorem 8.** *Let $m \in (\mathbb{N} \cup \{\infty\})^n$ and $r \in (\mathbb{N} \cup \{\infty\})^n$ be upper thresholds. Let $\mathcal{T}^*$ be a tree rooted at v_0 and $\mathrm{Win} \subseteq \mathrm{N}$ be a set of players. The following assertions are equivalent:*

1. *There exists a permissive SPE Θ in $(\mathcal{G}, v_0)$ such that:*
   - **a.** *$\langle \Theta \rangle^{\mathrm{H}}_{v_0} = \mathcal{T}^*$;*
   - **b.** *Θ is strongly winning w.r.t.* Win*;*
   - **c.** *for all $i \in \mathrm{N}$, $\mathrm{MPenalty}_i(\Theta, v_0) \leq m_i$ and $\mathrm{RPenalty}_i(\Theta, v_0) \leq r_i$.*

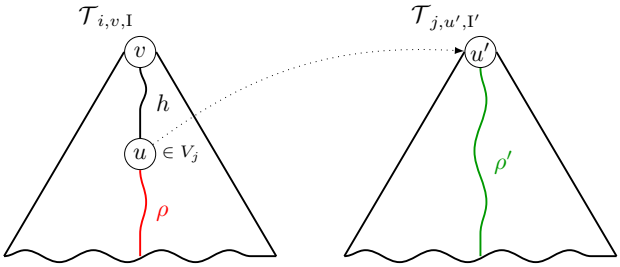

Figure 5 Example of forest that does not respect the resistance to constrained external deviations since $\mathrm{Gain}_i(\rho) = 0$ but $\mathrm{Gain}_i(\rho') = 1$.

2. *There exists a good forest $\mathcal{F}$ in $(\mathcal{G}, v_0)$ such that:*
   a. $\mathcal{T}_{0,v_0,\mathrm{I}_0} = \mathcal{T}^*$;
   b. *for all* $\rho \in \mathcal{T}^\infty_{0,v_0,\mathrm{I}_0}$, *for all* $i \in \mathrm{Win}$, $\mathrm{Gain}_i(\rho) = 1$;
   c. *for all* $i \in \mathrm{N}$, $\mathrm{MPenalty}_i(\mathcal{F}) \leq m_i$ *and* $\mathrm{RPenalty}_i(\mathcal{F}) \leq r_i$.

These assertions are still equivalent by replacing 1b by "Θ is weakly winning w.r.t. Win*" and 2b by "there exists $\rho \in \mathcal{T}^\infty_{0,v_0,\mathrm{I}_0}$ such that for all $i \in \mathrm{Win}$, $\mathrm{Gain}_i(\rho) = 1$".*

Proof sketch. For $(1 \Rightarrow 2)$, let us assume that Θ is a permissive SPE. We build a good forest $\mathcal{F}$ such that $\mathcal{T}_{0,v_0,\mathrm{I}_0}$ is the outcomes of Θ, *i.e.*, $\mathcal{T}_{0,v_0,\mathrm{I}_0} = \langle\Theta\rangle^{\mathrm{H}}_{v_0}$, and a tree $\mathcal{T}_{i,v,\mathrm{I}}$ is a representative of the outcomes of $\Theta_{\restriction hv'}$ in some subgame $(\mathcal{G}_{\restriction hv'}, v)$ such that $v' \in V_i$ and $\mathrm{Visit}(hv'v) = \mathrm{I}$. In order to obtain a good forest and since several $hv'v$ could satisfy those properties, each representative $\mathcal{T}_{i,v,\mathrm{I}}$ has to be chosen in a proper way: it has to minimize the maximal gain of player i for plays in $\mathcal{T}_{i,v,\mathrm{I}}$. More formally, for each $(i, v, \mathrm{I}) \in \mathcal{I}$, we let $\mathcal{O}(i, v, \mathrm{I}) = \{\langle\Theta_{\restriction hv'}\rangle^{\mathrm{H}}_v \mid hv'v \in \mathrm{Hist}(v_0) \wedge v' \in V_i \wedge \mathrm{I} = \mathrm{Visit}(hv'v)\}$ and we choose $\mathcal{T}_{i,v,\mathrm{I}} \in \mathcal{O}(i, v, \mathrm{I})$ such that $\max\{\mathrm{Gain}_i(\rho) \mid \rho \in \mathcal{T}^\infty_{i,v,\mathrm{I}}\} = \min_{\mathcal{T} \in \mathcal{O}(i,v,\mathrm{I})} \max\{\mathrm{Gain}_i(\rho) \mid \rho \in \mathcal{T}^\infty\}$.

Thanks to this latter property, $\mathcal{F}$ is good. Indeed, let $\mathcal{T}_{i,v,\mathrm{I}}$ be a tree of $\mathcal{F}$. Exactly as for permissive NEs, if $\mathcal{T}_{i,v,\mathrm{I}}$ is not $(\mathrm{N} \setminus \mathrm{I})$-resistant to internal deviations, we can build a strategy profile σ consistent with Θ such that the restriction of σ is not an NE in a subgame corresponding to $\mathcal{T}_{i,v,\mathrm{I}}$. If $\mathcal{T}_{i,v,\mathrm{I}}$ is not resistant to constrained external deviations that means that from some node u, owned by player j, there exists a play ρ losing for player j and player j could choose to play outside $\mathcal{T}_{i,v,\mathrm{I}}$ by jumping to a tree $\mathcal{T}_{j,u',\mathrm{I}'}$ in which there exists a play ρ' winning for him, see Figure 5. Let g be the history such that $\mathcal{T}_{i,v,\mathrm{I}}$ represents the outcomes of $\Theta_{\restriction g}$ in $(\mathcal{G}_{\restriction g}, v)$. Notice that $\langle\Theta_{\restriction gh}\rangle_{u'}$ may be different from $\mathcal{T}_{j,u',I'}$. However thanks to the way in which this representative is chosen, we have that there exists a play ρ'' in $\langle\Theta_{\restriction gh}\rangle_{u'}$ with $\mathrm{Gain}_j(\rho'') = 1$. Thus, we can build a strategy σ consistent with Θ such that $\langle\sigma_{\restriction gh}\rangle_u = \rho$ and $\langle\sigma_{\restriction ghu}\rangle_{u'} = \rho''$. This means that player j could deviate by choosing u' instead of u from v in the subgame $(\mathcal{G}_{\restriction g}, v)$, thus σ would not be an SPE and Θ not a permissive SPE.

For $(2 \Rightarrow 1)$, from a good forest $\mathcal{F}$ a multi-strategy is build such that its subgame outcomes are the trees of $\mathcal{F}$. This forms a permissive SPE because $\mathcal{F}$ is good. ◀

For now, good trees and trees in good forests are infinite, but Section 5 will show that we can represent some trees using a finite representation (intuitively, by supposing that every branch ends with a lasso in the game). It is this finite representation of good trees and good forests that will be used to decide the constrained penalty problems for permissive NEs and permissive SPEs, thanks to the characterizations of Theorems 5 and 8.

5 Computation of permissive equilibria

Theorems 5 and 8 characterize permissive NEs and SPEs with respect to infinite tree-shaped objects. In this section, we use these characterizations in order to decide the various penalty problems defined in Section 3: we check the existence of the good infinite tree-shaped objects by checking the existence of finite symbolic representations of such objects. We start by describing for a single tree this symbolic representation, and show that there exists a *polynomial-size* such representation (when the penalty upper-bounds are encoded in unary).

5.1 Symbolic trees and forests

▶ **Definition 9.** *A* symbolic tree *is a pair* $\mathcal{U} = (\mathcal{T}, f)$ *with* $\mathcal{T}$ *a finite tree (*i.e., *a finite subset of non-empty histories of* $\mathcal{G}$*), and* f *a function mapping each leaf* h *of* $\mathcal{U}$ *to a non-empty set of successor nodes* h' *that are ancestors of* h *in* $\mathcal{U}$ *such that* $(\mathrm{Last}(h), \mathrm{Last}(h')) \in E$.

A symbolic tree can be *unfolded* into an infinite tree by repeatedly expanding the leaves of $\mathcal{U}$ using as successors the choice prescribed by f. We denote by $\widetilde{\mathcal{U}}$ the infinite tree obtained by unfolding the symbolic tree $\mathcal{U}$. Similarly, the notions of symbolic forest $\mathcal{F}$, where every tree in it is a symbolic tree, and unfolding of symbolic forest $\widetilde{\mathcal{F}}$ can be defined.

In order to treat simultaneously NEs and SPEs, we introduce a new definition generalizing the resistance to external deviations and constrained external deviations. For a vector $\gamma \in \{0,1\}^{\mathrm{N} \times V \times 2^{\mathrm{N}}}$ of gains, and a subset $D \subseteq \mathrm{N}$ of players (that represent players that did not already win at the beginning of the tree), we say that a tree $\mathcal{T}$ is (γ, D)-resistant if for all $hu \in \mathcal{T}$ with $u \in V_i$ and $u' \in \mathrm{Succ}(u)$ with $huu' \notin \mathcal{T}$, if $\gamma_{i,u',(\mathrm{N}\setminus D)\cup \mathrm{Visit}(huu')} = 1$, if $i \notin (\mathrm{N} \setminus D) \cup \mathrm{Visit}(hu)$, then for all plays $\rho \in \mathcal{T}^{\infty}_{\restriction hu}$, $\mathrm{Gain}_i(\rho) = 1$.

▶ Remark 10. The notion of (γ, D)-resistance is close to the resistance to external deviations and constrained external deviations, so that we directly obtain from Theorems 5 and 8:

- Let $\gamma^{\mathcal{G}}$ be defined as follows: for all (i, u, I), we let $\gamma^{\mathcal{G}}_{i,u,\mathrm{I}}$ equals 1 iff player i belongs to I or can win from u against the coalition of the other players in $\mathcal{G}$. Let $\mathcal{T}$ be a tree. Then, $\mathcal{T}$ is a good tree iff $\mathcal{T}$ is resistant to internal deviations and $(\gamma^{\mathcal{G}}, \mathrm{N})$-resistant.
- Let $\mathcal{F}$ be a forest and let $\gamma^{\mathcal{F}}$ defined as follows: for all (j, u, J), we let $\gamma^{\mathcal{F}}_{j,u,\mathrm{J}}$ equals 1 iff player j belongs to J or the tree $\mathcal{T}_{j,u,\mathrm{J}}$ contains at least one branch with a vertex of F_j. Then, $\mathcal{F}$ is a good forest iff each each tree $\mathcal{T}_{i,v,\mathrm{I}}$ of $\mathcal{F}$ is $(\mathrm{N} \setminus \mathrm{I})$-resistant to internal deviations and $(\gamma^{\mathcal{F}}, \mathrm{N} \setminus \mathrm{I})$-resistant.

The challenge to make this remark a decision procedure is to make the tree and forest finitely representable. We treat each tree independently of each other, thus explaining how to symbolically represent one single tree in the following proposition:

▶ **Proposition 11.** *Let* $\mathcal{T}$ *be a tree that is* D*-resistant to internal deviations, with* $D \subseteq \mathrm{N}$. *We let* $\gamma \in \{0,1\}^{\mathrm{N} \times V \times 2^{\mathrm{N}}}$ *be a vector of gains such that* $\mathcal{T}$ *is* (γ, D)*-resistant, and* $(P_i)_{i \in \mathrm{N}'}$ *be finite constraints on penalties for a subset* $\mathrm{N}' \subseteq \mathrm{N}$ *of players. There exists a symbolic tree* $\mathcal{U}$*, that is a subtree of* $\mathcal{T}$*, of height polynomial in the number of players and vertices of* $\mathcal{G}$*, and in the largest bound on penalty* P_i*, such that the infinite tree* $\widetilde{\mathcal{U}}$ *satisfies the following properties:*

1. $\widetilde{\mathcal{U}}$ *is* D*-resistant to internal deviations;*
2. *in* $\widetilde{\mathcal{U}}$*, every player* $i \in \mathrm{N}'$ *has a penalty at most* P_i*;*
3. $\widetilde{\mathcal{U}}$ *is* (γ, D)*-resistant.*

Moreover, for a subset Win *of players, if we start with* $\mathcal{T}$ *that is strongly (respectively, weakly) winning w.r.t.* Win*, then we can make the above construction so that moreover* $\widetilde{\mathcal{U}}$ *is strongly (respectively, weakly) winning w.r.t.* Win.

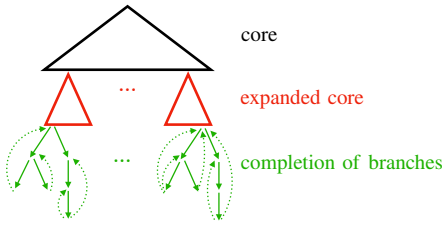

Figure 6 Construction of the symbolic tree.

The proof of this result goes by several steps, that we briefly sketch here only in the case where $\mathcal{T}$ is strongly winning w.r.t. Win. Figure 6 depicts the notions used in the construction of the symbolic tree. First, we consider the smallest subtree of $\mathcal{T}$ where leaves are such that all players of Win have visited their target set: this subtree is finite by König's lemma, since all branches of $\mathcal{T}$ have such a node where all players of Win have won, and the tree is finitely branching. This subtree is called the *core*. We then continue considering the parts of $\mathcal{T}$ outside the core, in order to complete the branches so that: the D-resistance to internal deviations is fulfilled (if a player has won in a certain branch of a subtree, he must win in all of them), the (γ, D)-resistance is fulfilled (if γ gives a constraint in the current node for a player i, all the branches of this subtree should visit a target vertex of i). This extension of the core is cut into two parts: the *expanded core* that ends in places where all the new players that must visit their target because of D-resistance to internal deviations and (γ, D)-resistance have indeed won; the *completion of branches* in order to then find leaves of the symbolic tree where all successors can be replaced (with function f) by similar nodes in the same branch, and the lassos thus formed are such that the penalty of players that have a finite penalty threshold does not increase along them. We show that these completions of branches can be chosen of polynomial length. We then compress the core and expanded core so that they also have polynomial height.

The symbolic tree $\mathcal{U}$ thus built is a subtree of $\mathcal{T}$ (even if its unfolding $\widetilde{\mathcal{U}}$ is not): in particular, as a corollary, if a player j has no winning play in $\mathcal{T}$, he does not have a winning play in $\mathcal{U}$ neither. In particular, when we apply independently this proposition to all the trees of a forest $\mathcal{F}$, to obtain a symbolic forest $\mathcal{H}$, this remark allows us to check that the new vector $\gamma^{\widetilde{\mathcal{H}}}$ has all its components not above the corresponding ones in $\gamma^{\mathcal{F}}$ (if $\gamma^{\mathcal{F}}_{i,v,\mathrm{I}} = 0$ then $\gamma^{\widetilde{\mathcal{H}}}_{i,v,\mathrm{I}} = 0$). In particular, if the tree $\mathcal{T}_{i,v,\mathrm{I}}$ of the forest $\mathcal{F}$ is $(\gamma^{\mathcal{F}}, \mathrm{N} \setminus \mathrm{I})$-resistant, then the tree $\widetilde{\mathcal{U}}_{i,v,\mathrm{I}}$ of the symbolic forest $\mathcal{H}$ is $(\gamma^{\widetilde{\mathcal{H}}}, \mathrm{N} \setminus \mathrm{I})$-resistant.

Finally, by combining this result with Remark 10, we obtain the following corollaries that allow us to obtain the PSPACE decision procedures:

▶ **Corollary 12.** *Let $m \in (\mathbb{N} \cup \{\infty\})^n$ be upper thresholds, and M be the largest such upper threshold. The following assertions are equivalent:*

1. *There exists a permissive NE Θ in $(\mathcal{G}, v_0)$ such that:*
   (a) Θ is strongly winning w.r.t. Win*; (b) for all $i \in \mathrm{N}$,* $\mathrm{MPenalty}_i(\Theta, v_0) \leq m_i$.
2. *There exists a symbolic tree $\widetilde{\mathcal{T}}$ in $(\mathcal{G}, v_0)$ of height polynomial in the number of players and vertices of $\mathcal{G}$ and in M, such that*
   (a) $\widetilde{\mathcal{T}}$ is resistant to internal deviations, and $(\gamma^{\mathcal{G}}, \mathrm{N})$-resistant, with $\gamma^{\mathcal{G}}$ defined in Remark 10; and (b) for all $\rho \in \widetilde{\mathcal{T}}^{\infty}$ and $i \in$ Win*,* $\mathrm{Gain}_i(\rho) = 1$*; and (c) for all $i \in \mathrm{N}$,* $\mathrm{Penalty}_i(\widetilde{\mathcal{T}}) \leq m_i$.

These assertions are still equivalent by replacing 1(a) by "Θ is weakly winning w.r.t. Win*"* *and 2(b) by "there exists $\rho \in \widetilde{\mathcal{T}}^{\infty}$ such that for all $i \in$* Win*,* $\mathrm{Gain}_i(\rho) = 1$*".*

▶ **Corollary 13.** *Let $m \in (\mathbb{N} \cup \{\infty\})^n$ and $r \in (\mathbb{N} \cup \{\infty\})^n$ be upper thresholds, and M be the largest such upper threshold. The following assertions are equivalent:*

1. *There exists a permissive SPE Θ in $(\mathcal{G}, v_0)$ such that:*
   (a) Θ is strongly winning w.r.t. Win*; and (b) for all $i \in \mathrm{N}$, $\mathrm{MPenalty}_i(\Theta, v_0) \leq m_i$ and $\mathrm{RPenalty}_i(\Theta, v_0) \leq r_i$.*
2. *There exists a symbolic forest $\mathcal{F}$ in $(\mathcal{G}, v_0)$, where each symbolic tree has a height polynomial in the number of players and vertices of $\mathcal{G}$ and in M, such that (a) each tree $\widetilde{\mathcal{T}}_{i,v,\mathrm{I}}$ is $(\mathrm{N} \setminus \mathrm{I})$-resistant to internal deviations, and $(\gamma^{\mathcal{F}}, \mathrm{N})$-resistant, with $\gamma^{\mathcal{F}}$ defined in Remark 10; (b) for all $\rho \in \widetilde{\mathcal{T}}^{\infty}_{0,v_0,\mathrm{I}_0}$ and $i \in \mathrm{Win}$, $\mathrm{Gain}_i(\rho) = 1$; and (c) for all $i \in \mathrm{N}$, $\mathrm{MPenalty}_i(\widetilde{\mathcal{F}}) \leq m_i$ and $\mathrm{RPenalty}_i(\widetilde{\mathcal{F}}) \leq r_i$.*

These assertions are still equivalent by replacing 1(a) by "Θ is weakly winning w.r.t. Win*" and 2(b) by "there exists $\rho \in \widetilde{\mathcal{T}}^{\infty}_{0,v_0,\mathrm{I}_0}$ such that for all $i \in \mathrm{Win}$, $\mathrm{Gain}_i(\rho) = 1$".*

5.2 Decision problems over permissive Nash equilibria

For permissive NEs, it makes little sense to take into consideration the retaliation penalties, since the punishment after a deviation should definitely make the deviator lose whatever the penalty from now on. We thus obtain the following decision result:

▶ **Theorem 14.** *The constrained penalty problem, the weakly winning with constrained penalty problem and the strongly winning with constrained penalty problem, all with infinite (and thus no) constraints on retaliation penalties and for NEs are decidable in* PSPACE *(when the penalty bounds are encoded in unary).*

Proof. We build upon Corollary 12, looking for a finite symbolic tree with the corresponding properties. We first explain how to solve the constrained penalty problem, and explain afterwards the adaptation for the two other problems. The idea is to use an alternating polynomial time Turing machine (since AP = PSPACE [12]) to guess a symbolic tree, checking the various constraints over it by using branch per branch. We describe the construction by supposing that the states of the Turing machine are split between existential states (where the machine accepts if at least one execution accepts) and universal states (where the machine accepts if all the executions accept). Existential states thus allow us to non-deterministically guess the finite symbolic tree node after node. We use a polynomial counter to keep track of the polynomially bounded height of the tree: if the counter goes over the polynomial bound, the execution of the alternating machine fails. At each node, existential states guess non-deterministically the set of successors on the working tape.

Universal states allow us to check several pieces of information on the guessed symbolic tree: the resistance to internal deviations, the constraint on the penalty for each player, and the $(\gamma^{\mathcal{G}}, \mathrm{N})$-resistance, with $\gamma^{\mathcal{G}}$ as in Remark 10. Notice that this vector has exponential size, but the index I in a triple (i, v, I) is useless (apart from knowing if $i \in \mathrm{I}$), and can thus be ignored: moreover, this set I will be maintained along the execution of the algorithm. This vector can thus be precomputed in (deterministic) polynomial time by determining, for each player, their set of winning vertices (against the coalition of the other players) [15].

The various checks can be performed branch per branch by keeping some pieces of information in memory, not only for the current node of the symbolic tree, but also for the whole current branch (this remains in polynomial space). Universal states are thus used to perform the checks on all the branches of the guessed tree.

- *Checking the penalty for player i.* If we have to check that the main penalty of player i is bounded by a threshold m_i (*i.e.,* that the penalty of player i over each branch is bounded by m_i), we keep in memory the current penalty, forbidding for it to go above m_i.

- *Checking the resistance to internal deviations and* $(\gamma^{\mathcal{G}}, \mathrm{N})$*-resistance.* At each node of the guessed tree, if the existential states guessed at least two successors, or depending on the vector $\gamma^{\mathcal{G}}$ (for a vertex v where $\gamma^{\mathcal{G}}$ has value 1, and that has not been chosen among the set of successors), we must remember constraints on the successors: either (a) all plays in their subtrees must be winning for a certain player i, or (b) none. We could add neither constraint (a) nor (b) for a certain player (if only one successor has been chosen, and the $\gamma^{\mathcal{G}}$ value of all the other successors is 0). In the case where only the resistance to internal deviation applies (if at least two successors have been chosen, and the $\gamma^{\mathcal{G}}$ value of all the other successors is 0), the choice of constraint (a) or (b) is guessed non-deterministically. These constraints are kept all along the guessed branch except if a vertex of the target set of player i is visited; in this case the constraint (a) is released. Moreover, the constraint (b) for a player i forbids to select a successor in the future where player i visits one of his target vertices.
- *The end of the branches.* The existential states decide when to stop the branch of the symbolic tree (before the counter runs out of the polynomial bound). Notice that the branch cannot stop if one of the type (a) constraints is not released. Then existential states provide the set of successors taken in the ancestors so that for players that have a finite upper threshold on their penalty, ancestors must have the same current penalty as the leaf (to ensure that their penalty does not raise to $+\infty$ in the long run).

For the strongly winning variants, universal states also check the constraint that every player of Win must win at the end of each branch. For the weakly winning variant, the existential states are also used to propose a branch where all players of Win will win. The universal states moreover check whether this condition is fulfilled for this particular branch. ◀

5.3 Decision problems over permissive subgame perfect equilibria

▶ **Theorem 15.** *The constrained penalty problem, the weakly winning with constrained penalty problem and the strongly winning with constrained penalty problem for SPEs are decidable in* PSPACE *(when the penalty bounds are encoded in unary).*

Proof. The proof is the same as for NEs, instead of the fact that we use Corollary 13, with a vector $\gamma^{\mathcal{F}}$ that is partially guessed non-deterministically when it is needed. When the existential states extend a branch of the tree $\mathcal{T}$ from a vertex of player i, the universal states does not only explore the chosen successors (with constraints (a) or (b) as in the previous proof), but now also explores the other vertices u by starting a fresh exploration of another tree $\mathcal{T}_{i,u,I}$ of the forest. Existential states also non-deterministically guess if player i is weakly winning in $\mathcal{T}_{i,u,I}$. If so, this gives new constraints (a) in the tree $\mathcal{T}$. The guessed weakly winning constraints are then checked in the fresh exploration: if player i must be weakly winning, this is a constraint of the same type as a weakly winning constraint in the "main" tree; if player i must not be weakly winning, this is a constraint of type (b) (none of the play must be winning for player i) that we deal as before.

Main penalties are checked as before. For the retaliation penalties, for each player, we check that the total penalty of all new symbolic trees $\mathcal{T}_{i,u,\mathrm{I}}$ is below the given upper threshold. To ensure polynomial time termination, we maintain a polynomial counter, and the set of trees (more precisely, the set of triples (i, u, I) used to index the trees of the forest) we jumped in so far. The polynomial counter again takes care of the depth of the branch we explore in the current tree (we reset this counter when we jump from a tree to another one). The set of trees we jumped in so far is maintained to forbid several explorations of the same tree of the forest. As for NEs, the exploration is losing if the depth of the current branch is longer than

the polynomial bound. The cardinal of the set of triples (i, u, I) we must maintain is also polynomial (bounded by $|\mathrm{N}| \times |V| \times |\mathrm{N}|$, even though there are exponentially many trees in a forest), since the subset I of winning players does not decrease along the jumps from a tree to the next one. This also implies that the total length of the executions of the Turing machine is indeed polynomial.

Notice that weakly and strongly winning conditions have only to be checked on the "main" tree as for permissive NEs. ◀

6 Conclusion

We studied the permissiveness in Nash, and subgame perfect equilibria over multiplayer reachability games. We showed that several associated problems are decidable in PSPACE: they ask for the existence of such equilibria with various constraints, both on the set of players who reach their target set, and on the penalties that allow us to compare the permissiveness of two equilibria. The polynomial space depends on the size of the game, and the largest upper threshold on the penalties. We were not able to decrease the space dependency to be only polynomial in the logarithm of the penalty thresholds: we leave for future work to investigate if this is possible, or if there is a matching lower bound on complexity.

As other ideas for future works, we would like to extend our study to other objectives than reachability, like more general ω-regular objectives (*e.g.,* parity games), but also weighted games like mean-payoff games, discounted-payoff games, or shortest-path games (where the reachability objective is combined with an objective to reach the target with the smallest possible total weight). An even more challenging problem is to extend this study to the setting of timed games, where the permissiveness is not only on the choice of edges, but also on the choice of delays spent in a given vertex. Work along these lines has been carried out on timed automata and two-player timed games [4, 13].

References

1 Ashwani Anand, Satya Prakash Nayak, and Anne-Kathrin Schmuck. Synthesizing permissive winning strategy templates for parity games. In *CAV 2023*, volume 13964 of *LNCS*, pages 436–458. Springer, 2023. doi:10.1007/978-3-031-37706-8_22.

2 Julien Bernet, David Janin, and Igor Walukiewicz. Permissive strategies: from parity games to safety games. *RAIRO Theor. Informatics Appl.*, 36(3):261–275, 2002. doi:10.1051/ITA:2002013.

3 Patricia Bouyer, Marie Duflot, Nicolas Markey, and Gabriel Renault. Measuring permissivity in finite games. In *CONCUR 2009*, volume 5710 of *LNCS*, pages 196–210. Springer, 2009. doi:10.1007/978-3-642-04081-8_14.

4 Patricia Bouyer, Erwin Fang, and Nicolas Markey. Permissive strategies in timed automata and games. *Electron. Commun. Eur. Assoc. Softw. Sci. Technol.*, 72, 2015. doi:10.14279/TUJ.ECEASST.72.1015.

5 Patricia Bouyer, Nicolas Markey, Jörg Olschewski, and Michael Ummels. Measuring permissiveness in parity games: Mean-payoff parity games revisited. In *ATVA 2011*, volume 6996 of *LNCS*, pages 135–149. Springer, 2011. doi:10.1007/978-3-642-24372-1_11.

6 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. On the Complexity of SPEs in Parity Games. In *CSL 2022*, volume 216 of *LIPIcs*, pages 10:1–10:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.CSL.2022.10.

7 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. The Complexity of SPEs in Mean-Payoff Games. In *ICALP 2022*, volume 229 of *LIPIcs*, pages 116:1–116:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.ICALP.2022.116.

8 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. Subgame-perfect equilibria in mean-payoff games. *Logical Methods in Computer Science*, 19, 2023. doi:10.46298/LMCS-19(4:6)2023.

9 Thomas Brihaye, Véronique Bruyère, Aline Goeminne, and Jean-François Raskin. Constrained existence problem for weak subgame perfect equilibria with ω-regular boolean objectives. In *GandALF 2018*, volume 277 of *EPTCS*, pages 16–29, 2018. doi:10.4204/EPTCS.277.2.

10 Thomas Brihaye, Véronique Bruyère, Aline Goeminne, Jean-François Raskin, and Marie van den Bogaard. The complexity of subgame perfect equilibria in quantitative reachability games. In *CONCUR 2019*, volume 140 of *LIPIcs*, pages 13:1–13:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPICS.CONCUR.2019.13.

11 Thomas Brihaye, Véronique Bruyère, Aline Goeminne, and Nathan Thomasset. On relevant equilibria in reachability games. In *RP 2019*, volume 11674 of *LNCS*, pages 48–62. Springer, 2019. doi:10.1007/978-3-030-30806-3_5.

12 Ashok K. Chandra, Dexter C. Kozen, and Larry J. Stockmeyer. Alternation. *Journal of the ACM*, 28(1):114–133, 1981. doi:10.1145/322234.322243.

13 Emily Clement, Thierry Jéron, Nicolas Markey, and David Mentré. Computing maximally-permissive strategies in acyclic timed automata. In *FORMATS 2020*, volume 12288 of *LNCS*, pages 111–126. Springer, 2020. doi:10.1007/978-3-030-57628-8_7.

14 Aline Goeminne and Benjamin Monmege. Permissive equilibria in multiplayer reachability games. Technical Report 2411.13296, arXiv, 2024. doi:10.48550/arXiv.2411.13296.

15 Erich Grädel, Wolfgang Thomas, and Thomas Wilke, editors. *Automata, Logics, and Infinite Games: A Guide to Current Research [outcome of a Dagstuhl seminar, February 2001]*, volume 2500 of *LNCS*. Springer, 2002. doi:10.1007/3-540-36387-4.

16 Satya Prakash Nayak and Anne-Kathrin Schmuck. Most general winning secure equilibria synthesis in graph games. In *TACAS 2024*, volume 14572 of *LNCS*, pages 173–193. Springer, 2024. doi:10.1007/978-3-031-57256-2_9.

17 Michael Ummels. Rational behaviour and strategy construction in infinite multiplayer games. In *FSTTCS 2006*, volume 4337 of *LNCS*, pages 212–223. Springer, 2006. doi:10.1007/11944836_21.