The Non-Cooperative Rational Synthesis Problem for SPEs and ω-Regular Objectives

Véronique Bruyère
Université de Mons, UMONS, Belgium

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

Alexis Reynouard
Université libre de Bruxelles, ULB, Belgium

Marie Van Den Bogaard
Université Gustave Eiffel, CNRS, LIGM, Marne-la-Vallée, France

Abstract

This paper studies the rational synthesis problem for multi-player games played on graphs when rational players are following subgame perfect equilibria. In these games, one player, the system, declares his strategy upfront, and the other players, composing the environment, then rationally respond by playing strategies forming a subgame perfect equilibrium. We study the complexity of the rational synthesis problem when the players have ω-regular objectives encoded as parity objectives. Our algorithm is based on an encoding into a three-player game with imperfect information, showing that the problem is in 2ExpTime. When the number of environment players is fixed, the problem is in ExpTime and is NP- and coNP-hard. Moreover, for a fixed number of players and reachability objectives, we get a polynomial algorithm.

2012 ACM Subject Classification Theory of computation

Keywords and phrases non-zero-sum games, subgame perfect equilibria, rational synthesis

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2025.12

Related Version *Full Version*: `The Non-Cooperative Rational Synthesis Problem for Subgame Perfect Equilibria and` ω`-Regular Objectives` [19]

Funding *Véronique Bruyère*: supported by FNRS under PDR Grant T.0023.22 (Rational).
Jean-François Raskin: supported by Fondation ULB (`https://www.fondationulb.be/en/`), the Thelam fund 2024-F2150080-0021312 (Open Problems on the Decidability and Computational Complexity of Infinite Duration Games), and by FNRS under PDR Grant T.0023.22 (Rational).

1 Introduction

Studying non zero-sum games played on graphs of infinite duration with multiple players [4, 13] poses both theoretical and algorithmic challenges. This paper primarily addresses the (non-cooperative) rational synthesis problem for n-player non zero-sum games featuring ω*-regular* objectives. In this context, the goal is to algorithmically determine the existence of a strategy σ_0 for the system (also called player 0) to enforce his objective against any *rational* response from the environment (players $1, \dots, n$). So, *rational synthesis* supports the automatic synthesis of systems wherein the environment consists of multiple agents having each their *own* objective. These agents are assumed to act rationally towards their own objective rather than being *strictly antagonistic* (to the system). This approach contrasts with the simpler scenario of zero-sum two-player game graphs, the fully antagonistic setting, a framework extensively explored in earlier reactive synthesis research, see [2] and the numerous references therein. While the computational complexity of rational synthesis, where *rationality* is defined by the concept of *Nash equilibrium* (NE), has been explored in [26], this paper revisits the rational synthesis problem using the more encompassing definition of *subgame perfect equilibrium* (SPE) to formalize rationality. Nash equilibria (NEs) have a known limitation in *sequential* games, including the infinite-duration games on graphs that we consider here: they are prone to *non-credible threats*. Such threats involve decisions within subgames (potentially reached after a deviation from the equilibrium) that, while not rational, are intended to coerce other players into specific behaviors. To address this limitation, the concept of subgame-perfect equilibrium (SPE) was introduced, as discussed in [39]. An SPE is a profile of strategies that form Nash equilibria in every subgame, thereby preventing non-rational and thus non-credible threats. Although SPEs align more intuitively with sequential games, their algorithmic treatment in the context of infinite-duration graph games remains underdeveloped. This gap persists primarily because SPEs require more complex algorithmic techniques compared to NEs. Moreover, the standard backward induction method used in *finite* duration sequential games [33] cannot be directly applied to *infinite* duration games due to the non-terminating nature of these interactions.

Kupferman et al. introduced rational synthesis in two distinct forms. The first approach, dubbed *cooperative* rational synthesis [29], considers an environment working collaboratively with the system to determine whether a specific SPE that ensures the specification of the system is met. So, in this model, the agents of the environment engage in an SPE that guarantees a win for player 0, provided such an equilibrium exists. At the opposite, the second approach, termed *non-cooperative* rational synthesis (NCRS) [34], grants the environment greater flexibility. Here, the system first selects a fixed strategy σ_0 and then the environmental agents respond by selecting any strategy that form an SPE with the fixed strategy σ_0 of the system. The central algorithmic challenge is determining whether there exists a strategy σ_0 for the system such that every resulting σ_0-fixed SPE ensures the specification of the system is upheld (SPE-NCRS problem). The computational complexity of decision problems for the cooperative synthesis problem with subgame-perfect equilibria (SPE) is now well understood. While the decidability of this problem was first established in [43], its exact complexity was resolved in [6], where the problem was shown to be NP-complete for parity objectives.

In contrast, the computational complexity of the SPE-NCRS problem remains less thoroughly investigated. Although the decidability of this problem can be determined through an encoding in Strategy Logic [38], such an encoding does not provide clear insights into the effective construction of the strategy σ_0 and is suboptimal from an algorithmic standpoint. For example, for rational environment behaviors modeled by Nash equilibria (NEs) instead of SPEs, the NE-NCRS problem can be solved in 3ExpTime for LTL objectives using a Strategy Logic encoding [34], while it can be solved in 2ExpTime through the use of tree automata [35]. This shows that using a reduction to Strategy Logic does not deliver optimal worst-case complexity. Addressing this gap, our contribution is twofold. First, we introduce an *innovative* algorithm that transforms the SPE-NCRS problem into a three-player imperfect information game. Such games were analyzed in [22], where computationally optimal algorithms for their analysis were presented. Second, our reduction offers a clear advantage over the Strategy Logic encoding, providing improved complexity (double exponential time as opposed to triple exponential time for LTL objectives in [34]). It also enables a precise analysis of the algorithmic complexity when the number of players is *fixed*, a consideration that is practically relevant in cases where the environment consists of a limited number of players.

Technical contributions and sketch of our solution. The summary of our technical contributions is depicted in Figure 1. Our main result shows how to transform the SPE-NCRS problem into a two-player zero-sum parity game with perfect information (Theorem 11), a

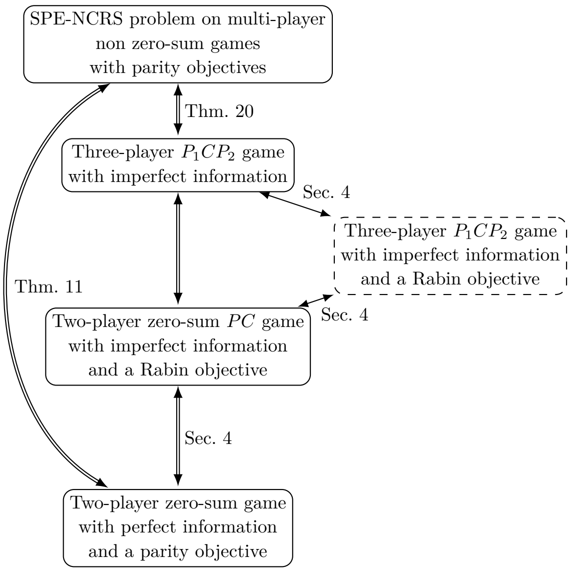

Figure 1 Structure of the article.

well-studied class of games for which algorithms are available. (Thus, our algorithm supports LTL specifications via their translation into deterministic parity automata, yielding parity conditions on the game graph). The transformation is structured in several non-trivial steps.

First, to solve the SPE-NCRS problem, we propose to use the *Prover-Challenger* framework initially introduced for the development of algorithms capable of determining the presence of a simulation relation between transition systems (for more details and extensions of this concept, see e.g. [1, 32, 37]). However, in our context, we need to use *two Provers* P_1 and P_2 with the Challenger C: P_1 aims to demonstrate the existence of a solution, i.e., a strategy σ_0 for the system, to the SPE-NCRS problem, while C seeks to counter this assertion, i.e., with a subgame perfect response $\bar{\sigma}_{-0}$ that results in an unfavorable outcome for player 0. Then, P_2 endeavors to demonstrate that the combined profile $(\sigma_0, \bar{\sigma}_{-0})$ is either not a 0-fixed SPE or its outcome favors player 0. To ensure that σ_0 is fixed and cannot be modified in subgames, we prevent P_1 from adjusting his strategy based on the interactions between C and P_2 by imposing to him *imperfect information* of the game. More intuition and the formal definition of this P_1CP_2 game are given in Section 3 together with a proof of correctness (Theorem 20).

Second, we detail our method for solving the P_1CP_2 game. Given that this game involves three players with imperfect information, specialized techniques are essential for its resolution as multi-player games with imperfect information are undecidable in general (see e.g. [40]). We employ a solution specifically adapted for our context, derived from a transformation introduced in [22]. This transformation was originally proposed for addressing similar types of three-player games with imperfect information. Due to the intricate winning condition present in our Prover-Challenger reduction, we translate it into an explicit Rabin objective. After that transformation of the winning condition, the aforementioned techniques for three-player games with imperfect information can be adapted. As a result of this transformation, we obtain a more conventional two-player zero-sum Rabin game with imperfect information.

Third, techniques to remove imperfect information, see e.g. [41, 42], can then be used to obtain the desired two-player zero-sum parity game with perfect information. Therefore, solving the SPE-NCRS problem reduces to solving this two-player zero-sum parity game for which algorithms are well-known.

In Section 4, we provide a complexity analysis of each step of our construction. This analysis enables us to derive detailed complexity results: solving the SPE-NCRS problem for parity objectives is exponential in the size of the graph and the number of priorities used by the parity objectives, and double-exponential in the number of players (while a PSpace lower bound can be deduced from [26]). Consequently, when the number of players is fixed, we achieve membership in ExpTime. Furthermore, by adapting proofs from the NE-NCRS problem [27], we establish NP- and co-NP-hardness (the NE-NCRS problem is in PSpace and both NP- and co-NP-hard, even for a fixed number of players). Finally, for the specific case of reachability objectives, we obtain a polynomial-time algorithm when the number of players is fixed, as is the case for the NE-NCRS problem [27].

To help the reader navigate the paper, we provide in Figure 1 a visual representation of its structure. Furthermore, for additional details on concepts and proofs, we provide a complete full version in [19] and two Appendices A and B to present a few helpful examples.

Related work. Recent literature, such as surveys in [4, 13, 14], underscores a growing interest in non zero-sum games for synthesis and design of multi-agents systems. Algorithms have been developed for reasoning about NEs in graph games, both for ω-regular [44] and quantitative objectives [12], or even in a concurrent setting [3]. The concept of secure equilibrium, a refinement of NE, was introduced in [25] and its potential for synthesis was demonstrated in later studies [24]. Similarly, doomsday equilibria, an expansion of secure equilibria for n-player games, is elaborated upon in [23]. All those results are related to the notion of NE. Algorithms to reason on SPEs are more recent and usually more intricate. For that reason, progresses on weaker notions like weak SPE [8, 10, 18] have been needed before establishing tight complexity bounds for SPEs. For SPEs, the exact complexity for the constrained existence for reachability games was established in [9], for parity games in [6], and for mean-payoff games in [5]. All those works introduce new algorithms that are substantially more sophisticated than those needed to study the notion of NE. As mentioned previously, the SPE-NCRS problem with LTL objectives can be addressed by reducing it to the model-checking problem of Strategy Logic, as shown in [34]. However, this encoding results in a triple-exponential complexity, even when the number of players is fixed. Consequently, this approach does not allow for a fine-grained complexity analysis when the number of players is treated as a fixed parameter, and so it cannot yield the ExpTime complexity we have achieved in this case, nor the PTime complexity for a fixed number of players with reachability objectives. Cooperative rational synthesis was first introduced in [29]. The adversarial version was later introduced in [34]. In both cases, as mentioned above, the decidability results were given through a reduction to Strategy Logic [38]. A more detailed analysis of the complexity for a large variety of ω-regular objectives and for rationality expressed as NEs was given in [26] for turn-based games and in [27] for concurrent games, for LTL objectives in [35], and for two players and for mean-payoff and discounted-sum objectives in [28]. Those results do not cover SPEs as we do in this paper. *Rational verification* (instead of rational synthesis) studies the problem of verifying that a given strategy σ_0 for the system is a solution to the NE/SPE-NCRS problem. The complexity of this simpler problem has been studied for several kinds of objectives in [7, 30, 31]. Another notion of rational environment behavior treats the environment as a single agent but with multiple, sometimes conflicting, goals, aiming for behaviors that achieve a *Pareto-optimal* balance among these objectives. Both rational synthesis and verification have been very recently studied for this concept of rationality [11, 15, 16, 17].

2 Preliminaries

In this section, we recall the necessary notions and concepts underpinning this work: First, we specify the model of games on graphs that we study, along with the corresponding definitions of plays, (winning) strategies, and objectives. Second, we present the solution concepts relevant to our non-zero-sum model and specific problem, ranging from Nash equilibria to 0-fixed subgame-perfect equilibria. Finally, we provide a precise statement of the SPE-NCRS problem (see Definition 10) and our main results (see Theorem 11).

▶ **Definition 1** (Game structure). *A* game structure *is a tuple $G = (V, A, \Pi, \delta, v_0)$, where:*

- $\Pi = \{0, \ldots, n\}$ *is a finite set of* players*,*
- $A = \bigcup_{i=0}^{n} A_i$ *is the set of* actions *of the players, such that A_i is the action set of player i, $i \in \Pi$, and $A_0 \cap (\bigcup_{i \neq 0} A_i) = \varnothing$,*
- $V = \bigcup_{i=0}^{n} V_i$ *is the set of* states*, such that V_i is the state set of player i, $i \in \Pi$, and $V_i \cap V_j = \varnothing$ for all $i \neq j$,*
- $v_0 \in V$ *is the* initial state*,*
- $\delta : \bigcup_{i=0}^{n}(V_i \times A_i) \to V$ *is a partial function called the* transition function*, such that:*
  1. *G is* deadlock-free*: for every state $v \in V_i$, there exists $a \in A_i$ an action of player i, such that $\delta(v, a)$ is defined,*
  2. *G is* action-unique*: for every state $v \in V_i$, and for all $a, b \in A_i$ actions of player i, we have $\delta(v, a) = \delta(v, b) \iff a = b$.*

The size *of a game structure is given by the numbers $|V|$, $|A|$, and $|\Pi|$ of its vertices, actions, and players respectively.*

We say that a state $v \in V_i$ is *controlled* or *owned* by player i. Note that the condition $A_0 \cap (\bigcup_{i \neq 0} A_i) = \varnothing$ means that one knows from the actions when player 0 is the one that is playing. This property of the action set of player 0 is not classical nor necessary but will reveal itself useful in the remaining of the paper. Indeed, in the setting we study here, player 0 has a distinguished role compared to the other players, see Section 3. Condition 1 on the transition function ensures that in every state, there is always a possible action to play. Condition 2 requires that a transition between two states can only be achieved via a unique action.

A *play* in a game structure G is an infinite sequence $\rho \in (VA)^\omega$ of states and actions, such that ρ is of the form $v_0 a_0 v_1 a_1 \cdots$ with v_0 the initial state and where for every $k \in \mathbb{N}$, we have $\delta(v_k, a_k) = v_{k+1}$. The set of all plays in G is denoted Plays_G. A *history* is a finite prefix $h \in (VA)^*V$ of a play ending in a state of G, that we describe as $h = v_0 a_0 v_1 a_1 \ldots v_k$. The set of all histories in G is denoted Hist_G, while for $i \in \Pi$, the set of histories ending in a state controlled by player i is denoted Hist_G^i. We write $h \sqsubset \rho$ if the history h is a prefix of the play ρ. We also use both notations $\sqsubset$ and $\sqsubseteq$ for two histories.

A *strategy* for player $i \in \Pi$ is a function $\sigma : \mathsf{Hist}_G^i \to A_i$, that prescribes an action $\sigma(h)$ for player i to choose for every history where it is his turn to play. The set of all strategies σ for player i is denoted by Σ_i. A collection $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ of strategies, one for each player, is called a *profile* of strategies. A play $\rho = v_0 a_0 v_1 a_1 \ldots$ is *compatible* with a strategy σ_i of player i if for every $k \in \mathbb{N}$ such that $v_k \in V_i$, we have $a_k = \sigma_i(v_0 a_0 \ldots v_k)$. Histories compatible with a strategy are defined similarly. A strategy σ is said to be *memoryless* when the prescribed action only depends on the last visited state, that is, for every $hav, h'a'v \in \mathsf{Hist}_G^i$, we have $\sigma(hav) = \sigma(h'a'v)$. Given a profile of strategies $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$, there is a unique play starting in v_0 that is compatible with every strategy of the profile, that we call the *outcome* of $\bar{\sigma}$ and denote it by $\langle \bar{\sigma} \rangle_{v_0}$. Given a strict subset of players $\Pi' \subset \Pi$, we write $\bar{\sigma}_{-\Pi'}$ to refer to a partial profile of strategies that contains a strategy for each player except the ones in Π'. In particular, we will often focus on the strategy of one player, say $\Pi' = \{i\}$, and the profile of strategies of the rest of the players, and use the notation $(\sigma_i, \bar{\sigma}_{-i})$ to denote the complete profile of strategies.

▶ **Definition 2** (Objective and game). *A* winning condition*, or* objective *for player i is a subset $W_i \subseteq \mathsf{Plays}_G$ of plays in the game structure G. We say that a play ρ is* winning *for player i or* satisfies *his objective if $\rho \in W_i$. Otherwise, we say that ρ is* losing *for player i. A* game *is a pair $\mathcal{G} = (G, (W_i)_{i\in\Pi})$ consisting in a game structure G, together with a profile of objectives for the players. When the context is clear, we often only write $\mathcal{G}$ to designate a game. Given a strategy profile $\bar{\sigma}$, its* gain profile *(or simply* gain*) is the Boolean vector $\bar{g}$ such that for all $i \in \Pi$, we have $g_i = 1$ if $\langle\bar{\sigma}\rangle_{v_0} \in W_i$, and $g_i = 0$ if $\langle\bar{\sigma}\rangle_{v_0} \notin W_i$. We also say that $\bar{g}$ is the gain profile of the outcome of $\bar{\sigma}$.*

In this paper, we consider the concept of *parity* objective (which is expressively complete for ω-regular objectives), as well as the particular case of *reachability* objective. Both concepts are defined in the following way:

- For each $i \in \Pi$, let $T_i \subseteq V$ be a *target set*. The reachability objective for player i is then $W_i = \{\rho = v_0a_0v_1a_1 \ldots \in \mathsf{Plays}_G \mid \exists k \in \mathbb{N}, v_k \in T_i\}$ that we also denote by $\mathrm{Reach}(T_i)$. Given such a profile of target sets, we say that $(G, (\mathrm{Reach}(T_i))_{i\in\Pi})$ is a *reachability game*.
- For each $i \in \Pi$, let $C_i = \{0, 1, \ldots, d_i\}$ be a finite set of *priorities* and $\alpha_i : V \to C_i$ be a *priority function*, that is, a function that assigns a priority to each state of the game. The parity objective for player i is then $W_i = \{\rho \in \mathsf{Plays}_G \mid \min_{v\in\mathrm{Inf}(\rho)}(\alpha_i(v)) \text{ is even}\}$[1] that we also denote by $\mathrm{Parity}(\alpha_i)$. Given such a profile of priority functions, we say that $(G, (\mathrm{Parity}(\alpha_i))_{i\in\Pi})$ is a *parity game*. The *size* of each parity objective, denoted by $|\alpha_i|$, is the maximum priority d_i.

▶ Remark 3. Note that the objectives we consider in this work only put constraints on the states of the plays (and not on the actions).

▶ **Definition 4** (Winning strategy). *Let G be a game structure and W_i be an objective for player i. A strategy σ_i of player i is* winning *for W_i, if for every profile $\bar{\sigma}_{-i}$, the outcome ρ of the profile $(\sigma_i, \bar{\sigma}_{-i})$ is winning for player i, that is, $\rho \in W_i$.*

This definition focuses on W_i only. A winning strategy σ_i ensures that player i satisfies his objective W_i against any strategy profile $\bar{\sigma}_{-i}$ of the other players. In particular, it ensures that player i wins even if the other players are *strictly antagonistic*, or *adversarial*, that is, their objective is to make player i lose. This context corresponds to the classical *zero-sum* setting, and we use notation $\mathcal{G} = (G, W_i)$ for the game, since the objective W_j, for all players $j \neq i$, is implicitly defined as $\mathsf{Plays}_G \setminus W_i$. An example of reachability game and a discussion on the winning strategies of both players are given in Appendix A.

Within the classical *zero-sum* context, the *synthesis problem* asks if there exists, for a particular player, a winning strategy (see [2] for an introduction). As recalled in the preliminaries (cf. Definition 4), such a strategy ensures that his player wins against all possible strategies of the other players. However, when we depart from this adversarial hypothesis and consider the richer setting of games that are *non zero-sum*, that is, where each player has his own objective, which may overlap with the objectives of the others, the solution concept of winning strategies shows its limits. Hence the call to solution concepts such as *Nash equilibrium* or *subgame-perfect equilibrium*, on which we focus in this work.

---

[1] For a play $\rho = v_0a_0v_1a_1 \ldots$, notation $\mathrm{Inf}(\rho)$ means $\{v \in V \mid \exists^\infty\ k,\ v_k = v\}$.

▶ **Definition 5** (Nash equilibrium [39]). *Let* $\mathcal{G} = (G, (W_i)_{i\in\Pi})$ *be a game and* $\bar{\sigma} = (\sigma_i)_{i\in\Pi}$ *be a strategy profile. We say that* $\bar{\sigma}$ *is a* Nash equilibrium *(NE for short) if for every player* $i \in \Pi$ *and strategy* $\sigma'_i \in \Sigma_i$*, we have* $\langle \sigma'_i, \bar{\sigma}_{-i}\rangle_{v_0} \in W_i \implies \langle \sigma_i, \bar{\sigma}_{-i}\rangle_{v_0} \in W_i$.

In other words, no player has an incentive to *deviate* unilaterally from its fixed strategy σ_i in the profile $\bar{\sigma}$: if he does so, the resulting outcome will not satisfy his objective if the outcome of $\bar{\sigma}$ was not already doing so. Notice that, however, there is no constraint on what players are allowed to do once a deviation has already occurred in the outcome of the NE. In fact, players may change their initial behaviours for completely adversarial choices. That is, NEs allow players to forget about their own objectives once the equilibrium outcome has been left, which can result in *non-credible threats*: promise of irrational behavior with respect to their own objectives. These non-credible threats are one important limitation of NEs as a solution concept that captures rationality of the players in sequential games. In order to enforce rationality in every scenario, even the ones that stem from deviations, one needs to monitor what the strategy profile prescribes after *every* history. This is exactly what *subgame perfect equilibria* (defined below) do [39]. Let $\mathcal{G} = (G, (W_i)_{i\in\Pi})$ be a game. To each history $hv \in \mathsf{Hist}_G$, with $v \in V$, corresponds a *subgame* $\mathcal{G}_{\restriction hv}$ that is the game $\mathcal{G}$ starting *after the history* hv: its plays ρ' start at the initial state v and for all $i \in \Pi$, ρ' is winning for player i if, and only if, $h\rho' \in W_i$. Given a strategy σ_i for player i, this strategy in $\mathcal{G}_{\restriction hv}$ is defined as $\sigma_{i,\restriction hv}(h') = \sigma_i(hh')$ for all histories $h' \in \mathsf{Hist}_G$ starting with the initial vertex v. Given a strategy profile $\bar{\sigma}$, we denote by $\bar{\sigma}_{\restriction hv}$ the profile $(\sigma_{i,\restriction hv})_{i\in\Pi}$ (note that its outcome starts in v).

▶ **Definition 6** (Subgame perfect equilibrium). *Let* $\mathcal{G} = (G, (W_i)_{i\in\Pi})$ *be a game. A* subgame perfect equilibrium *(SPE for short) is a profile* $\bar{\sigma}$ *of strategies such that* $\bar{\sigma}_{\restriction hv}$ *is an NE in each subgame* $\mathcal{G}_{\restriction hv}$ *of* G.

We illustrate both concepts of NE and SPE on an example in Appendix A.

In this work, we focus on *synthesizing* a strategy for a specific player (see Definition 10), referred to as the system (or player 0), which we assume to be trustworthy or whose strategy is fixed by design. For instance, this assumption is justified in scenarios where the system is implemented as a program that is subsequently used by its environment (the other players). Since the program is fixed, the system cannot deviate from its prescribed strategy. As stated at the beginning of this section, we consider the *non zero-sum* setting, where each player has their own objective, which may partially overlap with the objectives of other players. Consequently, we seek a solution concept that addresses two key aspects: (1) the specific player (the system) will not deviate from the prescribed strategy, and (2) the other players may exhibit *adversarial* behavior toward the specific player, as long as such behavior does not compromise their own objectives, given that they are assumed to be rational [34]. We next define 0-*fixed* equilibria, which meet these requirements.

▶ **Definition 7** (0-Fixed NE). *Let* $\mathcal{G}$ *be a game. A strategy profile* $\bar{\sigma} = (\sigma_0, \bar{\sigma}_{-0})$ *is a* 0-fixed NE, *if for every player* i*, where* $i \neq 0$*, and every strategy* $\sigma'_i \in \Sigma_i$*, we have* $\langle \sigma'_i, \bar{\sigma}_{-i}\rangle_{v_0} \in W_i \implies \langle \sigma_i, \bar{\sigma}_{-i}\rangle_{v_0} \in W_i$.

In other words, if the strategy σ_0 of player 0 is fixed, that is, if we assume player 0 *will stick* to his strategy, no other player has an incentive to deviate unilaterally. We also write that a strategy profile is a σ_0-*fixed* NE to insist on the fixed strategy σ_0 of player 0.

In this paper, we focus on 0-fixed SPEs: profiles $\bar{\sigma}$ of strategies that are 0-fixed NE in each subgame of $\mathcal{G}$ compatible with the fixed strategy of player 0. Formally:

▶ **Definition 8** (0-Fixed SPE). *Let $\mathcal{G}$ be a game. A strategy profile $\bar{\sigma} = (\sigma_0, \bar{\sigma}_{-0})$ is a* 0-fixed SPE if for every history hv compatible with σ_0, the profile $\bar{\sigma}_{\restriction hv}$ is a σ_0-fixed NE is the subgame $\mathcal{G}_{\restriction hv}$.

As for NEs, we use the terminology of σ_0-fixed SPE. Furthermore, given a strategy σ_0, we say that the strategy profile $\bar{\sigma}_{-0}$ is a *subgame-perfect response* to σ_0 if the complete profile $(\sigma_0, \bar{\sigma}_{-0})$ is a 0-fixed SPE. The next theorem guarantees the existence of an SPE in every reachability or parity game. This result also holds for 0-fixed SPEs [43].

▶ **Theorem 9** (Existence of (0-fixed) SPEs). *Given a game $\mathcal{G} = (G, (W_i)_{i\in\Pi})$ that is a parity game or a reachability game, there always exists an SPE in $\mathcal{G}$, and, for every strategy σ_0 of player* 0*, there always exists a σ_0-fixed SPE in $\mathcal{G}$.*

In this paper, we study the following synthesis problem. Intuitively, the *non-cooperative rational synthesis problem* asks, for a distinguished player, if there exists a strategy that is *as close as possible* to a winning strategy, taking into account the fact that the others are behaving rationally. The solutions seeked in this problem are indeed strategies that ensure winning for *every* possible way to complete the profile in a rational manner (i.e., other players do not cooperate but take care of their own objectives first).

▶ **Definition 10** (Non-cooperative rational synthesis problem). *The* non-cooperative rational synthesis problem *(SPE-NCRS problem for short), asks, given a game $\mathcal{G} = (G, (W_i)_{i\in\Pi})$, if there exists a strategy σ_0 of player* 0*, such that, for every subgame-perfect response $\bar{\sigma}_{-0}$, the resulting outcome $\langle(\sigma_0, \bar{\sigma}_{-0})\rangle_{v_0}$ is winning for player* 0*.*

The previous definition is illustrated with an example in Appendix A.

We now state our main results.

▶ **Theorem 11** (Complexity of the SPE-NCRS problem). *The SPE-NCRS problem for parity games is in 2ExpTime and PSpace-hard. Given $\mathcal{G} = (G, (\text{Parity}(\alpha_i))_{i\in\Pi})$, this problem is solvable in time exponential in $|V|$ and each $|\alpha_i|$, $i \in \Pi$, and double-exponential in $|\Pi|$.*

Furthermore, if the number of players is fixed, the SPE-NCRS problem is in ExpTime, and NP-hard and co-NP-hard for parity games; and it is solvable in time polynomial in $|V|$ for reachability games.

3 SPE-NCRS Problem and P_1CP_2 Game

3.1 Overview of our solution

One may notice the particular role *player* 0 has in the setting of the SPE-NCRS problem: one needs to find, if it exists, a strategy σ_0 of player 0 that ensures him to win *for all* possible partial profiles $\bar{\sigma}_{-0}$ such that $(\sigma_0, \bar{\sigma}_{-0})$ is a 0-fixed SPE. Our approach is to construct, in the spirit of [36], a zero-sum two-player *Prover-Challenger* game that will, once solved, deliver the answer. The key idea in this approach is that Prover is set to prove that there exists a solution σ_0 to the SPE-NCRS problem, while Challenger wants to disprove this claim. However, the special role of player 0 in the original game $\mathcal{G}$ must be taken into account. To this end, we need to *split* Prover into a coalition of two Provers, *Prover* 1 and *Prover* 2 (we later show, in Section 4, how to come back to the two-player model). This game is called P_1CP_2 *game* and is denoted by $P_1CP_2(\mathcal{G})$.

Let us give some intuition. The three players P_1, C, and P_2, proceed to construct, step by step and with some additional interactions, a play in $P_1CP_2(\mathcal{G})$, that *simulates* a play in $\mathcal{G}$ (see Definition 15 below). Each player has a specific part to play in the construction of this simulating play. As mentioned previously, player P_1 has the particular task to exhibit a candidate solution strategy σ_0 to the SPE-NCRS problem for $\mathcal{G}$. This is done by simulating player 0 playing σ_0 and not deviating from it. Thus, along the play in $P_1CP_2(\mathcal{G})$ being constructed, whenever the corresponding state in the simulated play of $\mathcal{G}$ belongs to player 0, it is then up to P_1 to choose the next state. One of the tasks of player C is to try to complete σ_0 with a subgame perfect response $\bar{\sigma}_{-0}$ from the other players in $\mathcal{G}$. Thus, whenever the corresponding state in the simulated play of $\mathcal{G}$ belongs to a player $i \neq 0$, it is player C's turn to play in $P_1CP_2(\mathcal{G})$: he *proposes* to play an action according to some strategy σ_i of player i in $\mathcal{G}$ and in addition *predicts* the gain profile of what will be the outcome of the profile $(\sigma_0, \bar{\sigma}_{-0})$, that is, of the simulated play being constructed. However, since players other than player 0 in $\mathcal{G}$ can deviate, whenever C has made a proposal for a next state, he has to let player P_2 have a say: P_2 can either accept this proposal or refuse it and *deviate* on behalf of player i by choosing another state as the next one in the simulating play. This phase is called the *decision phase*. If this phase results in a deviation, the game proceeds to the *adjusting* phase: since the current subgame has changed, Challenger has to predict the gain profile of $(\sigma_0, \bar{\sigma}_{-0})$ in this subgame, that shows the deviation of player i was not a profitable one. Then, the play in $P_1CP_2(\mathcal{G})$ resumes to the construction phase again.

How do we ensure that σ_0 is fixed? In other words, what restricts P_1 from adapting his strategy to the way C and P_2 interact, but only according to the so far constructed simulated history in G? The solution we choose is to make use of *imperfect information*: in the P_1CP_2 game, P_1 cannot, in fact, distinguish between all the states. To model the partial view P_1 has on the states of the game, we speak of *observations*. Informally, to each state (and action) is associated an observation, and P_1 has to take action upon sequences of observations only (that we call *observed histories*). In particular, it is assumed that the strategies should *respect* the observations, in the sense that two different histories yielding the same sequence of observations should trigger the same action of P_1. Thus, we only consider these *observation-based strategies* as possible behaviors of P_1. We now proceed to the formal definition of the P_1CP_2 game. In Section 3.4, we show that solving the SPE-NCRS problem is equivalent to solving this game in a sense stated in Theorem 20.

3.2 Observation Functions

We begin by introducing the concepts of observation and of observation-based strategy. Let $G = (V, A, \Pi, \delta, v_0)$ be a game structure and player i be some player in G. Let $\mathcal{O}$ be a partition of V, that defines an *observation function* $\mathcal{O}bs : V \to \mathcal{O}$ for player i, that is, for every state v, we have $\mathcal{O}bs(v) = o \in \mathcal{O}$. Let then $\bar{\mathcal{O}}$ be a partition of A, that extends the previous function $\mathcal{O}bs$ to actions, that is, for every action a, we have $\mathcal{O}bs(a) = \bar{o} \in \bar{\mathcal{O}}$. The function $\mathcal{O}bs$ extends to histories and plays in the straightforward way. We say that player i *observes G through $\mathcal{O}bs$* if he cannot distinguish between states (resp. actions, histories, plays) that yield the same observation via the function $\mathcal{O}bs$. For those states or actions that are the unique element of their observation set, we say that they are *visible*. We say that $\mathcal{O}bs$ is the *identity function* if $\mathcal{O}bs(v) = \{v\}$ for all $v \in V$ and $\mathcal{O}bs(a) = \{a\}$ for all $a \in A$. Whenever player i observes the game structure through the identity function, we say that player i has *perfect information*. If it is through any other function, we say that he has *imperfect information*. We will always assume that a player *sees his own actions perfectly* (that is, his observation function restricted to the domain of his action set corresponds to the identity function). Furthermore, we assume all players have *perfect recall*: they remember the full sequence of observations they witnessed from the start of the play. In the sequel, we only consider *player-stable* observation functions such that the last states of the histories observed similarly all belong to the same player.

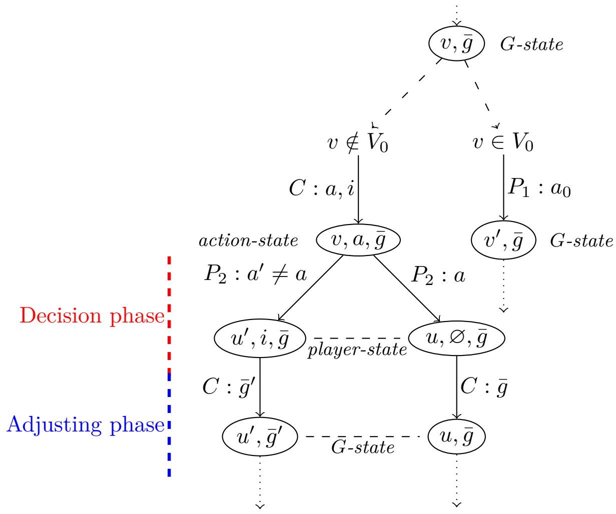

Figure 2 General structure of a play in $P_1CP_2(\mathcal{G})$.

▶ **Definition 12** (Player-stable observation function). *An observation function $\mathcal{O}bs$ for player i is* player-stable *if for every two histories $h = v_0a_0v_1a_1 \ldots v_k$ and $g = u_0b_0u_1b_1 \ldots u_k$ such that $\mathcal{O}bs(h) = \mathcal{O}bs(g)$, then the states v_ℓ, u_ℓ are controlled by the same player, for all $\ell \in \{0, \ldots, k\}$.*

The identity observation function is trivially player-stable. A game $\mathcal{G}$ with objectives W_i for all its players $i \in \Pi$ and a player-stable observation function $\mathcal{O}bs$ for one of its players is denoted $\mathcal{G} = (G, (W_i)_{i\in\Pi}, \mathcal{O}bs)$. When $\mathcal{O}bs$ is the identity, we do not mention it in this notation. Notice that with a player-stable observation function for player i, we have that, given a history $h \in \mathsf{Hist}^i_G$, all histories h' with the same observation as h also belong to Hist^i_G. It is therefore natural to ask that a strategy σ_i for player i is *observation-based*, that is, the same action is played by player i after all histories observed similarly.

▶ **Definition 13** (Observation-based strategy). *Given a player-stable observation function $\mathcal{O}bs$, a strategy σ_i of player i is* observation-based *if for every pair of histories h, h', if $h \in \mathsf{Hist}^i_G$ and $\mathcal{O}bs(h) = \mathcal{O}bs(h')$, then $\sigma_i(h) = \sigma_i(h')$.*

Both concepts of observation and observation-based strategy are illustrated in an example in Appendix B.1.

3.3 Definition of the P_1CP_2 Game

In this section, let us fix a game $\mathcal{G} = (G, (W_i)_{i\in\Pi})$ with a game structure $G = (V, A, \Pi, \delta, v_0)$ and objectives $(W_i)_{i\in\Pi}$. Later, this game will be a reachability game or a parity game. We here formally define the corresponding P_1CP_2 game, that was informally presented in Section 3.1. Recall that the role of player P_1 is to simulate a strategy σ_0 of player 0 in $\mathcal{G}$, while player C has to propose a subgame perfect response to σ_0 such that $(\sigma_0, \bar{\sigma}_{-0})$ is losing for player 0 in $\mathcal{G}$, response to which player P_2 can react by choosing some deviations (decision phase). The interactions between the three players result in a play in $P_1CP_2(\mathcal{G})$ that simulates a play in $\mathcal{G}$. Challenger has to predict a gain profile for this simulated play, which he can update upon deviations from P_2 (adjusting phase). We begin by defining the game structure of the new game $P_1CP_2(\mathcal{G})$.[2] We then define the observation function of P_1 in Definition 19; and we finally define and describe the objectives of the three players. Figure 2 should help the reader for understanding those definitions. The states (except its initial state v_0') of the P_1CP_2 game structure have several components: All of them have a state of G as their first component, that we call *G-component*. Similarly, all of them have a *gain-component*, which consists of a gain profile $\bar{g} \in \{0,1\}^{|\Pi|}$ for the players of G, where in each component g_i of $\bar{g}$, 0 symbolizes a loss and 1 a win with respect to objective W_i. The states that have only these two components are called *G-states* and belong either to P_1 or C. A projection on the G-components of these states determines the simulated play in G (see Definition 15). As already stated in Section 3.1, whenever the current state of the simulated play belongs to player 0 in G, the next state of the simulated play is chosen directly by P_1. However, when it is not the case, C has to make a proposition, which should then be validated or changed by P_2 (decision phase). Thus, to reflect this proposal while remaining hidden to P_1, the successor states of G-states belonging to C have an extra component, called *action-component*, which records the action of G proposed by C. From such states called *action-states*, P_2 has to react by confirming or changing the action that was proposed by C. This is modeled by a third kind of states, called *player-states*. Such states have a *player-component*, which consists of a player from G different from player 0 to signal to C this player has indeed deviated (or the empty set, to signal acceptance by Prover 2 of Challenger's proposal).

▶ **Definition 14** (States of the P_1CP_2 game)**.** *Given a game $\mathcal{G}$, the game structure of $P_1CP_2(\mathcal{G})$ is a game structure $G' = (V', A', \{P_1, C, P_2\}, \delta', v_0')$, where P_1, C, and P_2 are the three players and the set of states $V' = V'_{P_1} \cup V'_C \cup V'_{P_2}$ is as follows:*

- $V'_{P_1} = \{(v,\bar{g}) \mid v \in V_0,\ \bar{g} \in \{0,1\}^{|\Pi|}\}$,
- $V'_C = \{v_0'\} \cup \{(v,\bar{g}) \mid v \in V \setminus V_0,\ \bar{g} \in \{0,1\}^{|\Pi|}\} \cup \{(v,i,\bar{g}) \mid v \in V,\ i \in \Pi \setminus \{0\} \cup \{\varnothing\},\ \bar{g} \in \{0,1\}^{|\Pi|}\}$,
- $V'_{P_2} = \{(v,a,\bar{g}) \mid v \in V \setminus V_0,\ a \in A \setminus A_0$ *and* $\delta(v,a)$ *is defined*, $\bar{g} \in \{0,1\}^{|\Pi|}\}$,
- v_0' *is the initial state.*

Among those states, $(v,\bar{g})$ are G-states, $(v,a,\bar{g})$ are action-states, and $(v,i,\bar{g})$ are player-states. Moreover, v is a G-component, $\bar{g}$ is a gain-component, a is an action-component, and i is a player-component. Notice that the sets V'_{P_1}, V'_C, V'_{P_2} are pairwise disjoint.

Let us now describe the transitions of the P_1CP_2 game and explain further the actions of the players, see also Figure 2 . The goal of Challenger is to prove the existence of a σ_0-fixed SPE losing for player 0 in $\mathcal{G}$. To be able to verify this claim in $P_1CP_2(\mathcal{G})$, Challenger has to predict the gain of the simulated play that is being constructed. To do so, the gain-components of the states of $P_1CP_2(\mathcal{G})$ are used as follows: First, Challenger owns the initial state v_0', and has to choose a gain profile $\bar{g}$, losing for player 0,[3] to start the construction of the simulated play in the G-state $(v_0,\bar{g})$. Then, whenever P_2 chooses to make some player i deviate, before reaching a G-state, Challenger has to respond by choosing an adjusted gain profile with a lower or equal gain for this player i (adjusting phase). This new gain shows to P_2 the absence of any profitable deviation for player i. Note that for the P_1CP_2 game structure, two modeling decisions have been made to help handle the subsequent

[2] Note that to avoid confusion with the players of G, and without loss of generality, the three players of the P_1CP_2 game have been given explicit names rather than numbers.

[3] As an outcome of a subgame perfect response to σ_0 losing for player 0 is exactly what Challenger wants to exhibit, for any σ_0 simulated by P_1.

developments. First, to ensure a certain regularity of the plays' shape in the P_1CP_2 game and that P_1 cannot infer whether P_2 actually made a deviation, the adjusting phase is played regardless of a deviation occurring or not: if there was no deviation, Challenger has no choice but to play the same gain profile. Second, to avoid confusion on who, among Challenger and Prover 2, is currently playing some action $a \in A \setminus A_0$, this action is of the form (a, i) for C and a for P_2, that is, we additionally recall the player i from G performing the action a when Challenger is playing.

As already mentioned in Section 3.1, given a play in $P_1CP_2(\mathcal{G})$, there exists a unique corresponding play in G being simulated by the interactions of the Provers and Challenger.

▶ **Definition 15** (Simulated play and gain in G). *Let $\rho = v'_0a'_0v'_1a'_1v'_2a'_2\dots$ be a play in the game structure of $P_1CP_2(\mathcal{G})$. The* simulated play *of ρ is the play $\mathsf{sim}(\rho) \in \mathsf{Plays}_G$ being the projection of ρ on the G-component of G-states (which belong either to P_1 or C) and on actions of P_1 and P_2.*[4] *The definition extends naturally to histories. The* simulated gain *of ρ is a Boolean vector $\mathsf{simGain}(\rho) \in \{0,1\}^{|\Pi|}$, such that $\mathsf{simGain}(\rho)_i = 0$ if $\mathsf{sim}(\rho)$ is losing for player i, and $\mathsf{simGain}(\rho)_i = 1$ if $\mathsf{sim}(\rho)$ is winning for player i.*

Let us give an example:

▶ **Example 16.** Consider the fictional history $h = v'_0\ \bar{g}\ (v_0, \bar{g})\ a_0\ (v_1, \bar{g})\ (a_1, i)\ (v_1, a_1, \bar{g})\ a'_1\ (v_2, i, \bar{g})\ \bar{g}'\ (v_2, \bar{g}')$ in the P_1CP_2 game of some game $\mathcal{G}$. It starts in the initial state v'_0, where C chooses a gain profile $\bar{g}$ where $g_0 = 0$. Thus the second state of the history is the G-state $(v_0, \bar{g})$. Looking at the form of the third state $(v_1, \bar{g})$ in h, we can deduce that $(v_0, \bar{g})$ belongs to Prover 1, that is, v_0 belongs to player 0 in G. Moreover $\delta(v_0, a_0) = v_1$. Then, the fourth state $(v_1, a_1, \bar{g})$ in h is an action-state, which means that the previous state $(v_1, \bar{g})$ belongs to Challenger, such that v_1 belongs to player $i \neq 0$ in G, and a_1 is the action from G proposed by C in this scenario. One can see that P_2 then chooses the action a'_1, which is different from a_1 since the fifth state $(v_2, i, \bar{g})$ of h is a state whose player-component i is not equal to $\varnothing$. This indicates that P_2 makes player i deviate in G and that $\delta(v_1, a'_1) = v_2$. By now, Challenger has to choose a new gain profile $\bar{g}'$, yielding the last state $(v_2, \bar{g}')$ of h, which is a G-state. By projecting h on the G-components of its G-states and on the actions of the Provers, one can check that the simulated history of h in G is $\mathsf{sim}(h) = v_0a_0v_1a'_1v_2$. Consider now the fictional play $\rho = h\ \big((a_2, j)\ (v_2, a_2, \bar{g}')\ a_2\ (v_2, \varnothing, \bar{g}')\ \bar{g}'\ (v_2, \bar{g}')\big)^\omega$ in the P_1CP_2 game. Looking at ρ, one can deduce that from $(v_2, \bar{g}')$, the last state of h, Challenger proposes action a_2 from G for player $j \neq 0$ and that P_2 accepts this proposal. Furthermore, Challenger cannot adjust the gain profile, leading to state $(v_2, \bar{g}')$. One can see that this behavior repeats indefinitely, thus the corresponding simulated play is $\mathsf{sim}(\rho) = \mathsf{sim}(h)(a_2v_2)^\omega$. The simulated gain $\mathsf{simGain}(\rho)$ of ρ is a Boolean vector deduced from $\mathsf{sim}(\rho)$, such that its i-th component is equal to 1 if, and only if, $\mathsf{sim}(\rho)$ is winning for player i.

▶ Remark 17. Note that while there exists a unique simulated play $\mathsf{sim}(\rho)$ in $\mathcal{G}$ for every play ρ in $P_1CP_2(\mathcal{G})$, the converse does not hold. Indeed, several different sequences of interactions between P_2 and C yield the same simulated play in $\mathcal{G}$. For instance, Challenger can propose different actions that P_2 can refuse.

We have seen that each state and action of G appear in several contexts in the game structure G' of $P_1CP_2(\mathcal{G})$. Let us now state how the size of G' depends on the size of G. The proof of this lemma directly follows from the definition of the P_1CP_2 game.

[4] This projection does not take into account the action- and player-states and the actions of Challenger.

▶ **Lemma 18.** *Given the size $|V|$, $|A|$, and $|\Pi|$ of the game structure of $\mathcal{G}$, the game structure of $P_1CP_2(\mathcal{G})$ has 3 players and a size $|V'|$ linear in $|V|$ and $|A|$, and exponential in $|\Pi|$, and $|A'|$ linear in $|A|$ and exponential in $|\Pi|$.*

In the P_1CP_2 game, recall that C and P_2 have perfect information whereas P_1 has imperfect information to ensure that P_1 cannot adapt his strategy from observing the interactions between C and P_2. The observation function $\mathcal{O}bs$ for player P_1 is defined below on V' and A'. For each state v' of V', he is only able to observe the G-component of v'. Concerning the actions of A', those of $A'_{P_1} \cup A'_{P_2}$ are all visible whereas those of A'_C are not visible at all. In the next definition, there is an abuse of notation in $\mathcal{O}$ and $\bar{\mathcal{O}}$.

▶ **Definition 19** (Information in the P_1CP_2 game)**.** *Given the game structure G' of $P_1CP_2(\mathcal{G})$, players C and P_2 have perfect information and the observation function $\mathcal{O}bs$ of player P_1 is defined as follows: We have $\mathcal{O}bs : V' \to \mathcal{O} = \{v'_0\} \cup \{v \mid v \in V\}$ such that $\mathcal{O}bs(v, \bar{g}) = v$ for all $(v, \bar{g}) \in V'_{P_1}$, $\mathcal{O}bs(v'_0) = v'_0$, $\mathcal{O}bs(v, \bar{g}) = v$ for all $(v, \bar{g}) \in V'_C$, $\mathcal{O}bs(v, i, \bar{g}) = v$ for all $(v, i, \bar{g}) \in V'_C$, and $\mathcal{O}bs(v, a, \bar{g}) = v$ for all $(v, a, \bar{g}) \in V'_{P_2}$. Furthermore, we have $\mathcal{O}bs : A' \to \bar{\mathcal{O}} = \{\sharp\} \cup \{a \mid a \in A\}$ such that $\mathcal{O}bs(a') = a'$ for all $a' \in A'_{P_1}$, $\mathcal{O}bs(a') = \sharp$ for all $a' \in A'_C$, and $\mathcal{O}bs(a') = a'$ for all $a' \in A'_{P_2}$.*

To finalize the definition of the P_1CP_2 game, it remains to define the objectives of the three players. Recall that the objective of Challenger is to show that for each strategy σ_0 of player 0 in $\mathcal{G}$, there exists a subgame perfect response $\bar{\sigma}_{-0}$ such that the outcome of $(\sigma_0, \bar{\sigma}_{-0})$ is losing for player 0. Let us give some intuition on what is a play ρ in the P_1CP_2 game that is winning for C. Three winning situations may occur:

- **(iC)** Eventually, that is, after reaching some subgame, P_2 always accepts the action proposals of C and the gain predicted by C in the subgame is correct (it is equal to the simulated gain in this subgame), or
- **(iiC)** eventually, P_2 keeps making one unique player i deviate in the decision phase, but C is able to adjust the gain to show that this deviation is not profitable for player i, or finally
- **$(iiiC)$** P_2 keeps making at least two different players deviate, essentially conceding the play, as the only deviations that can be considered within the scope of (0-fixed) SPEs are *unilateral* deviations.

In the P_1CP_2 game, the two Provers have the same objective W_P that is opposed to the objective W_C of Challenger. Indeed, recall that their objective is to exhibit a strategy σ_0 for player 0 in G, such that for every subgame-perfect response $\bar{\sigma}_{-0}$ to σ_0, the outcome of the resulting profile $(\sigma_0, \bar{\sigma}_{-0})$ is winning for player 0. Let us give some intuition on what is a winning play for the Provers. Two winning situations may occur for the Provers along a play ρ in the P_1CP_2 game:

- **(iP)** Eventually, after reaching some subgame, P_2 always accepts the action proposals of C and the gain predicted by C is incorrect, or
- **(iiP)** eventually, P_2 keeps making one unique player i deviate in the decision phase, and C is not able to adjust the gain to show that this deviation is not profitable for player i.

A full formal definition of the players' objectives is provided in Appendix B.2.

Given a game $\mathcal{G}$, we have $P_1CP_2(\mathcal{G}) = (G', W_P, \mathcal{O}bs)$ its P_1CP_2 game, where W_P is the objective of the Provers.

3.4 Equivalence Between SPE-NCRS Problem and P_1CP_2 Game

In this section, we show that solving the SPE-NCRS problem for a game $\mathcal{G}$ is equivalent to solving its P_1CP_2 game, as stated in the following theorem. To keep things short, we give here a sketch of the proof, and refer the reader to [19] for a full and detailed account.

▶ **Theorem 20** (Equivalence theorem between $\mathcal{G}$ and $P_1CP_2(\mathcal{G})$). *Let $\mathcal{G} = (G, (W_i)_{i\in\Pi})$ be a game, and $P_1CP_2(\mathcal{G}) = (G', W_P, \mathcal{O}bs)$ be its P_1CP_2 game. There exists in $\mathcal{G}$ a strategy σ_0 of player 0, such that, for every subgame-perfect response $\bar{\sigma}_{-0}$, the play $\langle(\sigma_0, \bar{\sigma}_{-0})\rangle_{v_0}$ is winning for player 0 if, and only if, there exists in $P_1CP_2(\mathcal{G})$ an observation-based strategy τ_{P_1} of P_1 such that for all strategies τ_C of C, there exists a strategy τ_{P_2} of P_2 such that the play $\langle\tau_{P_1}, \tau_C, \tau_{P_2}\rangle_{v_0'}$ belongs to W_P.*

Proof sketch. Let σ_0 be a solution to the SPE-NCRS problem of $\mathcal{G}$. Let τ_{P_1} be the observation-based strategy of P_1 that simulates σ_0: it is naturally defined as the strategy that makes the same choice of action as σ_0 at h in $\mathcal{G}$ at all the histories that are observed as h for P_1 in $P_1CP_2(\mathcal{G})$. Let τ_C be a strategy of C. Let $\bar{\sigma}_{-0}$ be the strategy profile obtained from τ_C by extracting the actions chosen by C at histories that indeed simulate histories in $\mathcal{G}$. There are two possibilities, either $(\sigma_0, \bar{\sigma}_{-0})$ is a 0-fixed SPE in $\mathcal{G}$, or it is not a 0-fixed SPE.

1. Suppose $(\sigma_0, \bar{\sigma}_{-0})$ is a 0-fixed SPE in $\mathcal{G}$. Let τ_{P_2} be τ_{acc}, the *accepting strategy* of P_2, where he accepts every action proposal of C and never deviates. By definition of τ_{P_2}, the simulated play of $\langle\tau_{P_1}, \tau_C, \tau_{P_2}\rangle_{v_0'}$ is $\langle\sigma_0, \bar{\sigma}_{-0}\rangle_{v_0}$. In that case, since σ_0 is a solution to the SPE-NCRS problem of $\mathcal{G}$, we know that $\langle\sigma_0, \bar{\sigma}_{-0}\rangle_{v_0} \in W_0$. However, by construction of $P_1CP_2(\mathcal{G})$ and the fact that τ_{P_2} is the accepting strategy of P_2, we know that C has predicted a loss for player 0 from the start, and has never had to adjust his prediction. Thus, we are in the case (iP) of the winning condition for the Provers.
2. Suppose now $(\sigma_0, \bar{\sigma}_{-0})$ is not a 0-fixed SPE in $\mathcal{G}$. This means that it is not a 0-fixed NE in some subgame of $\mathcal{G}$. Let h be a history compatible with σ_0 such that $(\sigma_0, \bar{\sigma}_{-0})_{\restriction h}$ is not an NE. Let player $i \neq 0$ be a player in $\mathcal{G}$ that has a profitable deviation σ_i' in the subgame $\mathcal{G}_{\restriction h}$ starting after h. In $P_1CP_2(\mathcal{G})$, we can show that there exists a unique history h' ending in a G-state, compatible with τ_C, such that $\mathsf{sim}(h') = h$ and that is also compatible with τ_{P_1}. Depending on Challenger's prediction for player i at h', the strategy τ_{P_2} has to be different: if Challenger predicted a loss, we let τ_{P_2} bring the play in h', then simulate the profitable deviation of player i. If Challenger predicted a win, we let τ_{P_2} bring the play in h', then switch to τ_{acc}: since player i has a profitable deviation from h', it is necessary the case that the resulting play in $\mathcal{G}$ is losing for player i.

Turning to the other direction, assume that there is no solution to the SPE-NCRS problem in $\mathcal{G}$. Let τ_{P_1} be a strategy of P_1 in $P_1CP_2(\mathcal{G})$. Let σ_0 be the strategy of player 0 in $\mathcal{G}$ such that τ_{P_1} is its simulation. By Theorem 9, there exists a σ_0-fixed SPE in $\mathcal{G}$. Among all those σ_0-fixed SPEs, one must have an outcome losing for player 0 as, by assumption, σ_0 is not a solution of the SPE-NCRS problem in $\mathcal{G}$. Let $\bar{\sigma} = (\sigma_0, \bar{\sigma}_{-0})$ be such a σ_0-fixed SPE. From these, we can define a strategy τ_C for C that will precisely enact $\bar{\sigma}_{-0}$ on relevant histories, and predict correctly the gain profile as the one σ in every subgame. In particular, at the initial state, it predicts the gain profile $\bar{g}$ of $\bar{\sigma}$, such that $g_0 = 0$ by choice of the 0-fixed SPE.

Let now τ_{P_2} be a strategy of P_2. Consider the play $\rho = \langle\tau_{P_1}, \tau_C, \tau_{P_2}\rangle_{v_0'}$. If there are infinitely many deviations by two different players of $\mathcal{G}$ in the play ρ, then ρ satisfies the winning condition $(iiiC)$ of Challenger. Assume now that there is at most one player that P_2 makes deviate. There are two cases:

1. Either τ_{P_2} prescribes a finite (possibly null) number of deviations along ρ, and then switches to the accepting strategy τ_{acc}. Eventually, after reaching some subgame, P_2 always accepts the action proposals of C and the gain predicted by C in the subgame is correct. Hence ρ satisfies the winning condition (iC) of Challenger.
2. Or τ_{P_2} prescribes an infinite number of deviations for the same player i along ρ. As there are only two values (0 and 1) for the gain-components and there is no opportunity for the i-component of the predicted gain to increase, we eventually get a stable gain-component

g_i for player i. If $g_i = 1$, we clearly have $g_i = 1 \geq \mathsf{simGain}(\rho)_i$, showing that (iiC) is satisfied. If $g_i = 0$, recall that $\bar{\sigma}$ is a 0-fixed SPE, thus it is an NE in the subgame from which g_i was stabilized. The outcome of $\bar{\sigma}$ in this subgame is losing for player i as Challenger correctly predicts a gain of 0 with τ_C. Any deviation from this outcome is thus also losing for player i, including $\mathsf{sim}(\rho)$. Thus, (iiC) is again satisfied. ◀

4 Solving the P_1CP_2 Game

This section is the last puzzle piece to complete the proof of Theorem 11. We present here the key ideas behind the manyfold process of solving the P_1CP_2 game, while the full technical constructions and details can be found in [19]. Thanks to Theorem 20, we have an equivalence between the existence of a solution of the SPE-NCRS problem for a game and the fact that the Provers are able to win the associated P_1CP_2 game. More precisely, the situation where P_1 has *a* strategy, such that for *every* strategy of Challenger, P_2 has *a* strategy to make the Provers win. The remaining question is how to determine whether this is the case? In other words, how can one *solve*, in this particular sense, this three-player game with imperfect information? As shown on Figure 1 , this involves several steps. We start by a first technical step to obtain a P_1CP_2 game with a *Rabin objective* for the Provers. Then, we get rid of the three-player setting, by eliminating one Prover to obtain a two-player PC game with imperfect information. Once we have such a PC game, we focus on getting rid of the imperfect information and work with a parity objective instead of a Rabin one. This allows us to obtain an equivalent *two-player zero-sum parity game with perfect information*, which is effectively solvable. The techniques used are similar to what can be found in the literature ([21, 22, 41, 42]). However, their settings each differ slightly from the others and ours. Thus, to obtain finer complexity measures, we adapt their techniques to our setting.

The P_1CP_2 Game as a Rabin Game. Given a parity game $\mathcal{G}$, we show that its corresponding P_1CP_2 game can be seen as a three-player game with the objective W_P for the two Provers translated into a *Rabin* objective. The approach is to use a deterministic automaton $\mathcal{O}$ that *observes* the states of $P_1CP_2(\mathcal{G})$. Then the *synchronized product* of the game structure of $P_1CP_2(\mathcal{G})$ with this observer automaton $\mathcal{O}$ is equipped with a Rabin objective translating W_P and thus leads to the announced three-player Rabin game.

From two Provers to one Prover. The next step is to merge P_1 and P_2 into a unique Prover by using a technique inspired from [22]. The main idea is to use imperfect information to ensure merging the two Provers does not grant too much knowledge to the new single Prover. Indeed, if the new Prover had perfect information, he could not simulate Prover 1 truthfully. Thus we let the new Prover have the same level of information as Prover 1 in the P_1CP_2 game. However, in order to let the new Prover have as much actions available as Prover 2 and stay observation-based, we modify the action set to include all *functions* from states of P_2 to actions of P_2. This way, the "merged" Prover preselects actions for each possible Challenger move, effectively encoding both Provers' strategies via a state-to-action function. This yields an exponential blowup in the number of actions. The Rabin objective remains the same. We obtain the corresponding PC game, denoted $PC(\mathcal{G})$.

Eliminating Imperfect Information. To solve the two-player game $PC(\mathcal{G})$, we get rid of the imperfect information and then apply standard game-theoretic techniques. In the literature [21, 41, 42], this is usually done in two steps: first, make the objective *visible*, that is, such that any two similarly observed plays agree on the winning condition, second, apply the *subset construction* to recall the set of possible visited states, and letting them be observed. This can be done in only one step, by simultaneously modifying the game structure to both entail the subset construction on the states of the PC game, and the product with an automaton that monitors the winning condition along the plays. This allows us to limit the exponential blowups to only one in the size of the state space and the Rabin condition. Essentially, the monitoring automaton is a deterministic parity automaton that is obtained by complementing a non-deterministic Streett-automaton that detects sets of similarly observable plays that contain a losing play for the now unique Prover.

Complexity. Recall that we set out from a parity game $\mathcal{G} = (G, (W_i)_{i\in\Pi})$ with $|V|$, $|A|$, and $|\Pi|$ be the size of the game structure G, and $|\alpha_i|$ be the size of each parity condition W_i. By the construction mentioned above, we have an equivalence with a Rabin P_1CP_2 game with a state set of size linear in $|V|$ and $|A|$, and exponential in $|\Pi|$, an action set linear in $|A|$ and exponential in $|\Pi|$, while its Rabin objective has a size linear in $|\Pi|$ and the $|\alpha_i|$. Reducing to only one Prover, we obtain an exponential number of actions compared to the two-Provers version, thus exponential in $|V|$, $|A|$ and double-exponential in $|\Pi|$. As mentioned above, with the removal of imperfect information, we obtain a game with a size of the state set that is now exponential in $|V|$, $|A|$, the $|\alpha_i|$ and double exponential in $|\Pi|$, while the new parity condition is polynomial in $|V|, |A|, |\alpha_i|, \forall i$ and exponential in $|\Pi|$. Finally, by [20], solving the ultimate parity game, or equivalently the SPE-NCRS problem, is in time exponential in $|V|, |A|, |\alpha_i|, \forall i$, and double-exponential in $|\Pi|$. As the game structure of $\mathcal{G}_0$ is action-unique (see Definition 1), it follows that $|A| \leq |V|^2$, thus leading to an algorithm in time exponential in $|V|$ and each $|\alpha_i|$, and double-exponential in $|\Pi|$. This completes the proof of the complexity upper bound of Theorem 11. For the PSpace lower bound, the QBF-reduction of [26] (Theorem 7) applies here; all NE responses to σ_0 in that proof are also SPEs. For the complexity lower bound of Theorem 11 when the number of players is fixed, we adapt the reduction proposed for the NE-NCRS problem in [26], in a way to deal with SPEs instead of NEs.

The Case of Reachability. A careful analysis of the simpler case of reachability games, when *the number of players is fixed*, leads to a more fine-tuned solution and a better complexity result: Theorem 11 states a *polynomial complexity* instead of the exponential complexity of the parity case. This approach follows the same steps as presented above for parity games, but with a few adjustments. The key idea is that monitoring reachability objectives is simpler than parity ones, which means that the synchronized product with an observer automaton $\mathcal{O}$ can omit some information from the original game structure. Indeed, when monitoring a play in the P_1CP_2 game that stems from an original reachability game, one still needs to keep track of the actions of both Provers and Challenger, but the checking of the gain component is simplified compared to parity: only the information of whether each player has already visited his target set is to be remembered. This can be exploited to avoid the exponential blowup of the state space in the subset construction phase (with respect to the original state set V) to get rid of imperfect information and thus obtain a better complexity.

▶ Remark 21. Suppose that our algorithm establishes the existence of a solution σ_0 to the SPE-NCRS problem for parity games. As it is obtained from a memoryless winning strategy in the *final* parity zero-sum game, we get a finite-memory solution σ_0 whose memory size is exponential when the number of players is fixed, doubly exponential otherwise. For a lower bound on the memory required, it is fairly straightforward to show that it requires exponential memory by reducing a Streett zero-sum game to our problem.

5 Conclusion

In this work, we introduce a novel algorithm to solve the SPE-NCRS problem for parity objectives. Unlike previous methods that converted the problem into a model-checking problem for Strategy Logic, our algorithm reduces the SPE-NCRS problem to a three-player zero-sum game with imperfect information, framed as a Prover-Challenger game. This new angle yields improved complexity upper bounds: exponential time in the number of vertices of the game structure and the number of priorities of the parity objectives, doubly exponential time in the number of players. In particular, our algorithm runs within exponential time for a fixed number of players, which is particularly relevant since the number of players is typically small in practical scenarios. Moreover, we establish a lower bound that indicates the impossibility of solving the SPE-NCRS problem in polynomial time unless $\mathsf{P} = \mathsf{NP}$ even for a fixed number of players. For the particular case of reachability objectives, when the number of players is fixed, we prove polynomial complexity like for the NE-NCRS problem [26]. We believe that the Prover-Challenger framework, based on a three-player model with imperfect information, may be applicable for other synthesis challenges beyond our current application.

References

1 Rajeev Alur, Thomas A. Henzinger, Orna Kupferman, and Moshe Y. Vardi. Alternating refinement relations. In Davide Sangiorgi and Robert de Simone, editors, *CONCUR '98: Concurrency Theory, 9th International Conference, Nice, France, September 8-11, 1998, Proceedings*, volume 1466 of *Lecture Notes in Computer Science*, pages 163–178. Springer, 1998. doi:10.1007/BFB0055622.

2 Roderick Bloem, Krishnendu Chatterjee, and Barbara Jobstmann. Graph games and reactive synthesis. In Edmund M. Clarke, Thomas A. Henzinger, Helmut Veith, and Roderick Bloem, editors, *Handbook of Model Checking*, pages 921–962. Springer, 2018. doi:10.1007/978-3-319-10575-8_27.

3 Patricia Bouyer, Romain Brenguier, Nicolas Markey, and Michael Ummels. Pure Nash equilibria in concurrent deterministic games. *Log. Methods Comput. Sci.*, 11(2), 2015. doi:10.2168/LMCS-11(2:9)2015.

4 Romain Brenguier, Lorenzo Clemente, Paul Hunter, Guillermo A. Pérez, Mickael Randour, Jean-François Raskin, Ocan Sankur, and Mathieu Sassolas. Non-zero sum games for reactive synthesis. In Adrian-Horia Dediu, Jan Janousek, Carlos Martín-Vide, and Bianca Truthe, editors, *Language and Automata Theory and Applications - 10th International Conference, LATA 2016, Prague, Czech Republic, March 14-18, 2016, Proceedings*, volume 9618 of *Lecture Notes in Computer Science*, pages 3–23. Springer, 2016. doi:10.1007/978-3-319-30000-9_1.

5 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. The complexity of SPEs in mean-payoff games. In Mikolaj Bojanczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022, Paris, France*, volume 229 of *LIPIcs*, pages 116:1–116:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPICS.ICALP.2022.116.

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

7 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. Rational Verification for Nash and Subgame-Perfect Equilibria in Graph Games. In Jérôme Leroux, Sylvain Lombardy, and David Peleg, editors, *48th International Symposium on Mathematical Foundations of*

Computer Science, MFCS 2023, August 28 to September 1, 2023, Bordeaux, France, volume 272 of *LIPIcs*, pages 26:1–26:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPICS.MFCS.2023.26.

8 Thomas Brihaye, Véronique Bruyère, Aline Goeminne, and Jean-François Raskin. Constrained existence problem for weak subgame perfect equilibria with ω-regular Boolean objectives. *Inf. Comput.*, 278:104594, 2021. doi:10.1016/J.IC.2020.104594.

9 Thomas Brihaye, Véronique Bruyère, Aline Goeminne, Jean-François Raskin, and Marie van den Bogaard. The complexity of subgame perfect equilibria in quantitative reachability games. *Log. Methods Comput. Sci.*, 16(4), 2020. URL: https://lmcs.episciences.org/6883.

10 Thomas Brihaye, Véronique Bruyère, Noémie Meunier, and Jean-François Raskin. Weak subgame perfect equilibria and their application to quantitative reachability. In Stephan Kreutzer, editor, *24th EACSL Annual Conference on Computer Science Logic, CSL 2015, September 7-10, 2015, Berlin, Germany*, volume 41 of *LIPIcs*, pages 504–518. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015. doi:10.4230/LIPICS.CSL.2015.504.

11 Thomas Brihaye, Véronique Bruyère, and Gaspard Reghem. Quantitative Reachability Stackelberg-Pareto Synthesis is NEXPTIME-Complete. In Olivier Bournez, Enrico Formenti, and Igor Potapov, editors, *Reachability Problems - 17th International Conference, RP 2023, Nice, France, October 11-13, 2023, Proceedings*, volume 14235 of *Lecture Notes in Computer Science*, pages 70–84. Springer, 2023. doi:10.1007/978-3-031-45286-4_6.

12 Thomas Brihaye, Julie De Pril, and Sven Schewe. Multiplayer cost games with simple Nash equilibria. In Sergei N. Artëmov and Anil Nerode, editors, *Logical Foundations of Computer Science, International Symposium, LFCS 2013, San Diego, CA, USA, January 6-8, 2013. Proceedings*, volume 7734 of *Lecture Notes in Computer Science*, pages 59–73. Springer, 2013. doi:10.1007/978-3-642-35722-0_5.

13 Véronique Bruyère. Computer aided synthesis: A game-theoretic approach. In Émilie Charlier, Julien Leroy, and Michel Rigo, editors, *Developments in Language Theory - 21st International Conference, DLT 2017, Liège, Belgium, August 7-11, 2017, Proceedings*, volume 10396 of *Lecture Notes in Computer Science*, pages 3–35. Springer, 2017. doi:10.1007/978-3-319-62809-7_1.

14 Véronique Bruyère. Synthesis of equilibria in infinite-duration games on graphs. *ACM SIGLOG News*, 8(2):4–29, 2021. doi:10.1145/3467001.3467003.

15 Véronique Bruyère, Baptiste Fievet, Jean-François Raskin, and Clément Tamines. Stackelberg-Pareto synthesis. *ACM Trans. Comput. Log.*, 25(2):14:1–14:49, 2024. doi:10.1145/3651162.

16 Véronique Bruyère, Christophe Grandmont, and Jean-François Raskin. As soon as possible but rationally. In Rupak Majumdar and Alexandra Silva, editors, *35th International Conference on Concurrency Theory, CONCUR 2024, September 9-13, 2024, Calgary, Canada*, volume 311 of *LIPIcs*, pages 14:1–14:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024. doi:10.4230/LIPICS.CONCUR.2024.14.

17 Véronique Bruyère, Jean-François Raskin, and Clément Tamines. Pareto-Rational Verification. In Bartek Klin, Slawomir Lasota, and Anca Muscholl, editors, *33rd International Conference on Concurrency Theory, CONCUR 2022, September 12-16, 2022, Warsaw, Poland*, volume 243 of *LIPIcs*, pages 33:1–33:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.CONCUR.2022.33.

18 Véronique Bruyère, Stéphane Le Roux, Arno Pauly, and Jean-François Raskin. On the existence of weak subgame perfect equilibria. *Inf. Comput.*, 276:104553, 2021. doi:10.1016/J.IC.2020.104553.

19 Véronique Bruyère, Jean-François Raskin, Alexis Reynouard, and Marie Van Den Bogaard. The non-cooperative rational synthesis problem for subgame perfect equilibria and omega-regular objectives, 2024. doi:10.48550/arXiv.2412.08547.

20 Cristian S. Calude, Sanjay Jain, Bakhadyr Khoussainov, Wei Li, and Frank Stephan. Deciding parity games in quasi-polynomial time. *SIAM J. Comput.*, 51(2):17–152, 2022. doi:10.1137/17M1145288.

21 Krishnendu Chatterjee and Laurent Doyen. The complexity of partial-observation parity games. In Christian G. Fermüller and Andrei Voronkov, editors, *Logic for Programming, Artificial Intelligence, and Reasoning - 17th International Conference, LPAR-17, Yogyakarta, Indonesia, October 10-15, 2010. Proceedings*, volume 6397 of *Lecture Notes in Computer Science*, pages 1–14. Springer, 2010. doi:10.1007/978-3-642-16242-8_1.

22 Krishnendu Chatterjee and Laurent Doyen. Games with a weak adversary. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II*, volume 8573 of *Lecture Notes in Computer Science*, pages 110–121. Springer, 2014. doi:10.1007/978-3-662-43951-7_10.

23 Krishnendu Chatterjee, Laurent Doyen, Emmanuel Filiot, and Jean-François Raskin. Doomsday equilibria for omega-regular games. *Inf. Comput.*, 254:296–315, 2017. doi:10.1016/J.IC.2016.10.012.

24 Krishnendu Chatterjee and Thomas A. Henzinger. Assume-guarantee synthesis. In Orna Grumberg and Michael Huth, editors, *Tools and Algorithms for the Construction and Analysis of Systems, 13th International Conference, TACAS 2007, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2007 Braga, Portugal, March 24 - April 1, 2007, Proceedings*, volume 4424 of *Lecture Notes in Computer Science*, pages 261–275. Springer, 2007. doi:10.1007/978-3-540-71209-1_21.

25 Krishnendu Chatterjee, Thomas A. Henzinger, and Marcin Jurdzinski. Games with secure equilibria. *Theor. Comput. Sci.*, 365(1-2):67–82, 2006. doi:10.1016/J.TCS.2006.07.032.

26 Rodica Condurache, Emmanuel Filiot, Raffaella Gentilini, and Jean-François Raskin. The complexity of rational synthesis. In Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi, editors, *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, volume 55 of *LIPIcs*, pages 121:1–121:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPICS.ICALP.2016.121.

27 Rodica Condurache, Youssouf Oualhadj, and Nicolas Troquard. The complexity of rational synthesis for concurrent games. In Sven Schewe and Lijun Zhang, editors, *29th International Conference on Concurrency Theory, CONCUR 2018, September 4-7, 2018, Beijing, China*, volume 118 of *LIPIcs*, pages 38:1–38:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPICS.CONCUR.2018.38.

28 Emmanuel Filiot, Raffaella Gentilini, and Jean-François Raskin. The Adversarial Stackelberg Value in Quantitative Games. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 127:1–127:18, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2020.127.

29 Dana Fisman, Orna Kupferman, and Yoad Lustig. Rational synthesis. In Javier Esparza and Rupak Majumdar, editors, *Tools and Algorithms for the Construction and Analysis of Systems, 16th International Conference, TACAS 2010, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2010, Paphos, Cyprus, March 20-28, 2010. Proceedings*, volume 6015 of *Lecture Notes in Computer Science*, pages 190–204. Springer, 2010. doi:10.1007/978-3-642-12002-2_16.

30 Julian Gutierrez, Muhammad Najib, Giuseppe Perelli, and Michael J. Wooldridge. Automated temporal equilibrium analysis: Verification and synthesis of multi-player games. *Artif. Intell.*, 287:103353, 2020. doi:10.1016/J.ARTINT.2020.103353.

31 Julian Gutierrez, Muhammad Najib, Giuseppe Perelli, and Michael J. Wooldridge. On the complexity of rational verification. *Ann. Math. Artif. Intell.*, 91(4):409–430, 2023. doi:10.1007/S10472-022-09804-3.

32 Thomas A. Henzinger, Orna Kupferman, and Sriram K. Rajamani. Fair simulation. In Antoni W. Mazurkiewicz and Józef Winkowski, editors, *CONCUR '97: Concurrency Theory, 8th*

International Conference, Warsaw, Poland, July 1-4, 1997, Proceedings, volume 1243 of *Lecture Notes in Computer Science*, pages 273–287. Springer, 1997. doi:10.1007/3-540-63141-0_19.

33 Harold W. Kuhn. Extensive games and the problem of information. *Classics in Game Theory*, pages 46–68, 1953.

34 Orna Kupferman, Giuseppe Perelli, and Moshe Y. Vardi. Synthesis with rational environments. *Ann. Math. Artif. Intell.*, 78(1):3–20, 2016. doi:10.1007/S10472-016-9508-8.

35 Orna Kupferman and Noam Shenwald. The complexity of LTL rational synthesis. In Dana Fisman and Grigore Rosu, editors, *Tools and Algorithms for the Construction and Analysis of Systems - 28th International Conference, TACAS 2022, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2022, Munich, Germany, April 2-7, 2022, Proceedings, Part I*, volume 13243 of *Lecture Notes in Computer Science*, pages 25–45. Springer, 2022. doi:10.1007/978-3-030-99524-9_2.

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

37 Robin Milner. An algebraic definition of simulation between programs. In D. C. Cooper, editor, *Proceedings of the 2nd International Joint Conference on Artificial Intelligence. London, UK, September 1-3, 1971*, pages 481–489. William Kaufmann, 1971. URL: http://ijcai.org/Proceedings/71/Papers/044.pdf.

38 Fabio Mogavero, Aniello Murano, Giuseppe Perelli, and Moshe Y. Vardi. What makes Atl* decidable? A decidable fragment of strategy logic. In Maciej Koutny and Irek Ulidowski, editors, *CONCUR 2012 - Concurrency Theory - 23rd International Conference, CONCUR 2012, Newcastle upon Tyne, UK, September 4-7, 2012. Proceedings*, volume 7454 of *Lecture Notes in Computer Science*, pages 193–208. Springer, 2012. doi:10.1007/978-3-642-32940-1_15.

39 Martin J. Osborne and Ariel Rubinstein. *A course in Game Theory*. MIT Press, Cambridge, MA, 1994.

40 Gary L. Peterson and John H. Reif. Multiple-person alternation. In *20th Annual Symposium on Foundations of Computer Science, San Juan, Puerto Rico, 29-31 October 1979*, pages 348–363. IEEE Computer Society, 1979. doi:10.1109/SFCS.1979.25.

41 Jean-François Raskin, Krishnendu Chatterjee, Laurent Doyen, and Thomas A. Henzinger. Algorithms for omega-regular games with imperfect information. *Log. Methods Comput. Sci.*, 3(3), 2007. doi:10.2168/LMCS-3(3:4)2007.

42 John H. Reif. The complexity of two-player games of incomplete information. *J. Comput. Syst. Sci.*, 29(2):274–301, 1984. doi:10.1016/0022-0000(84)90034-5.

43 Michael Ummels. Rational behaviour and strategy construction in infinite multiplayer games. In S. Arun-Kumar and Naveen Garg, editors, *FSTTCS 2006: Foundations of Software Technology and Theoretical Computer Science, 26th International Conference, Kolkata, India, December 13-15, 2006, Proceedings*, volume 4337 of *Lecture Notes in Computer Science*, pages 212–223. Springer, 2006. doi:10.1007/11944836_21.

44 Michael Ummels. The complexity of Nash equilibria in infinite multiplayer games. In Roberto M. Amadio, editor, *Foundations of Software Science and Computational Structures, 11th International Conference, FOSSACS 2008, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2008, Budapest, Hungary, March 29 - April 6, 2008. Proceedings*, volume 4962 of *Lecture Notes in Computer Science*, pages 20–34. Springer, 2008. doi:10.1007/978-3-540-78499-9_3.

A Helpful Examples for Games on Graphs

In this appendix, we present additional examples and explanations for notions referred to in Section 2.

▶ **Example 22** (A simple reachability game). Consider the game structure G pictured in Figure 3. Its initial state is v_0 and there are two players, player 0, who owns the circle states, and player 1, who owns the square state. Transitions are represented by the arrows

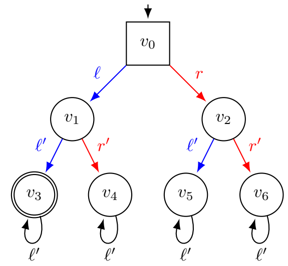

Figure 3 Two NEs and one SPE.

between states, such that there exists an arrow from a state v to another state v' if there exists an action a (either left or right here) such that $\delta(v, a) = v'$. The set of actions A is thus partitioned into $A_0 = \{\ell', r'\}$ and $A_1 = \{\ell, r\}$.

Both players have the same reachability objective, with $T_0 = T_1 = \{v_3\}$, denoted in the figure by the fact that state v_3 is double-circled. None of the two players have a winning strategy: player 0 can prevent player 1 to ever reach state v_3 by choosing to go to v_4, v_5 or v_6, depending on the first action of player 1, while player 1 can prevent player 0 from winning by going to state v_2.

▶ **Example 23** (A game with two NEs)**.** Let us come back to the reachability game of Figure 3. Two strategy profiles are represented in Figure 3: the first one, $\bar{\sigma}$, by the red transitions between states and the second one, $\bar{\sigma}'$, by the blue ones. Both profiles are NEs. Indeed, for the red profile with outcome $v_0 r v_2 r' (v_6 \ell')^\omega$, if player 0 deviates from σ_0, the resulting play $v_0 r v_2 \ell' (v_5 \ell')^\omega$ is still losing for him, since player 1 chooses to go to state v_2. On the other hand, if player 1 deviates from σ_1, the resulting play is $v_0 \ell v_1 r' (v_4 \ell')^\omega$, since player 0 chooses to go to state v_4, and this play is losing for player 1. Thus, none of the players have a profitable deviation from $\bar{\sigma}$. Notice that the gain profile of the red profile is equal to $(0, 0)$. Similarly, one can easily check that the blue profile is also an NE with a gain profile equal to $(1, 1)$.

In the red profile $\bar{\sigma}$ of the previous example, the possible choice of going from v_1 to v_4 for player 0 is irrational: after all, his objective is to reach state v_3. Notice, however, that this behavior is part of the NE $\bar{\sigma}$.

▶ **Example 24** (A game with an SPE)**.** Consider again the game in Figure 3. In the subgame $\mathcal{G}_{\restriction v_0 \ell v_1}$, starting after the history $v_0 \ell v_1$, the red profile is not an NE. Indeed, choosing to go to state v_4 is a non-credible threat from player 0, as his target set $\{v_3\}$ is accessible from state v_1, which he owns. The player 0 strategy that goes to v_3 from v_1 is thus a profitable deviation in the subgame $\mathcal{G}_{\restriction v_0 \ell v_1}$. Therefore, the red profile is not an SPE in $\mathcal{G}$. On the other hand, one can check that the blue profile is an SPE in $\mathcal{G}$, as it is an NE in every of its subgames.

▶ **Example 25** (A solution to the SPE-NCRS problem)**.** The answer to the SPE-NCRS problem for the game in Figure 3 is positive. Indeed, consider the strategy σ_0 of player 0 that chooses action ℓ' in both states v_1 and v_2. The unique subgame-perfect response $\bar{\sigma}_{-0}$ for player 1 is to choose action ℓ in v_0. The resulting strategy profile is thus the blue one whose outcome is winning for player 0.

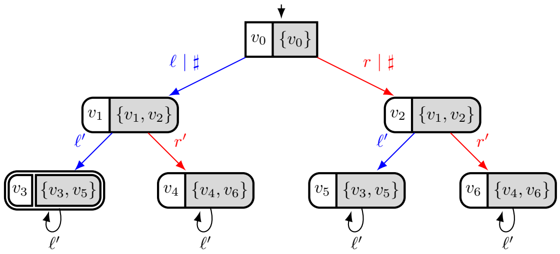

Figure 4 Imperfect information in a reachability game.

B SPE-NCRS Problem and P_1CP_2 Game

In this appendix, we provide additional material for Section 3.

B.1 Imperfect Information

▶ **Example 26** (Observation). In Figure 4, a two-player game structure is pictured, together with the observations of player 0, who owns the circle states (we suppose that player 1 has perfect information). Each state is divided in two sections: on the left, with a white background, the name of the state is displayed, while on the right side, with a gray background, the observation of player 0 is displayed. Similarly, the actions ℓ and r of player 1, who owns the square state, are accompanied by a # on their right side, which is an abuse of notation to mean $\{\ell, r\}$, that is, player 0 cannot distinguish between the two possible actions of player 1. Notice that from states owned by player 0, the actions do not have an observation attached to them, as player 0 knows his own actions, by hypothesis.

Let us now look at which histories player 0 can or cannot distinguish. From v_0, player 0 cannot distinguish between the actions of player 1, and both states v_1 and v_2 have the same observation set $\{v_1, v_2\}$. Therefore, the two histories $v_0\ell v_1$ and $v_0 r v_2$ are indistinguishable for player 0, i.e., $\mathcal{O}bs(v_0\ell v_1) = \mathcal{O}bs(v_0 r v_2)$. Similarly, we have $\mathcal{O}bs(v_0\ell v_1\ell' v_3) = \mathcal{O}bs(v_0 r v_2 \ell' v_5)$.

▶ **Example 27** (Observation-based strategy). Let us come back to the game structure of Figure 4. Suppose that the objective of player 0 is to reach the target set $\{v_3\}$. One can verify that player 0 does not have an observation-based winning strategy. Indeed, as noted in Example 26, player 0 cannot distinguish between histories $h_1 = v_0\ell v_1$ and $h_2 = v_0 r v_2$. By definition, any observation-based strategy must prescribe the same action from both h_1 and h_2. If this action is r', then both resulting plays are losing for player 0. If it is ℓ', then one of the resulting play is winning, while the other is losing. Thus, no observation-based strategy ensures player 0 to reach $\{v_3\}$.

B.2 Details on the P_1CP_2 Game Structure

▶ **Definition 28** (Objectives of the P_1CP_2 game). *Let $\rho = v'_0 a'_0 v'_1 a'_1 v'_2 a'_2 \dots$ be a play in the game structure of $P_1CP_2(\mathcal{G})$. Let* $\mathsf{simGain}(\rho)$ *be its simulated gain.*

- *The play ρ is* winning for Challenger *if one of the following conditions is satisfied:*
  (iC) *there exist $n \in \mathbb{N}$ and $g \in \{0,1\}^{|\Pi|}$, such that for every state v'_k with $k > n$, if v'_k is a player-state, then it is of the form $(v_k, i_k, \bar{g}_k)$ such that its player-component i_k equals $\varnothing$ and its gain-component satisfies*

$$\bar{g}_k = \mathsf{simGain}(\rho),$$

(iiC) *there exists a player* $i \neq 0$ *such that*

- *for every* $n \in \mathbb{N}$, *there exists* $k > n$ *for which* v'_k *is a player-state with its player-component being equal to* i,
- *there exists* $n \in \mathbb{N}$ *such that for every state* v'_k *with* $k > n$, *if* $v'_k = (v_k, i_k, \bar{g}_k)$ *is a player-state with* $i_k \neq \varnothing$, *then its player-component* i_k *equals* i *and the* i*-th component* $\bar{g}_{k,i}$ *of its gain-component* $\bar{g}_k$ *satisfies*

$$\bar{g}_{k,i} \geq \mathsf{simGain}(\rho)_i,$$

$(iiiC)$ *there exist two distinct players* $i, j \neq 0$ *such that for every* $n \in \mathbb{N}$, *there exist* $k, \ell > n$ *for which* v'_k *and* v'_ℓ *are player-states, with their player-component being respectively* i *and* j.

The set of plays satisfying one of these conditions is denoted W_C.

- *The play* ρ *is* winning for Prover 1 and Prover 2 *if one of the following conditions is satisfied:*

(iP) *there exist* $n \in \mathbb{N}$ *and* $g \in \{0,1\}^{|\Pi|}$, *such that for every state* v'_k *with* $k > n$, *if* v'_k *is a player-state, then it is of the form* $(v_k, i_k, \bar{g}_k)$ *such that its player-component* i_k *equals* $\varnothing$ *and its gain-component satisfies*

$$\bar{g}_k \neq \mathsf{simGain}(\rho),$$

(iiP) *there exists a player* $i \neq 0$ *such that*

- *for every* $n \in \mathbb{N}$, *there exists* $k > n$ *for which* v'_k *is a player-state with its player-component being equal to* i,
- *there exists* $n \in \mathbb{N}$ *such that for every state* v'_k *with* $k > n$, *if* $v'_k = (v_k, i_k, \bar{g}_k)$ *is a player-state with* $i_k \neq \varnothing$, *then its player-component* i_k *equals* i *and the* i*-th component* $\bar{g}_{k,i}$ *of its gain-component* $\bar{g}_k$ *satisfies*

$$\bar{g}_{k,i} < \mathsf{simGain}(\rho)_i,$$

The set of plays satisfying one of these conditions is denoted W_P.

▶ Remark 29. We have $W_{P_1} = W_{P_2} = W_P$ and $W_P = \mathsf{Plays}_{G'} \setminus W_C$.