Optimal Strategies in Concurrent Reachability Games

Benjamin Bordais
Université Paris-Saclay, CNRS, ENS Paris-Saclay, LMF, 91190 Gif-sur-Yvette, France

Patricia Bouyer
Université Paris-Saclay, CNRS, ENS Paris-Saclay, LMF, 91190 Gif-sur-Yvette, France

Stéphane Le Roux
Université Paris-Saclay, CNRS, ENS Paris-Saclay, LMF, 91190 Gif-sur-Yvette, France

Abstract

We study two-player reachability games on finite graphs. At each state the interaction between the players is concurrent and there is a stochastic Nature. Players also play stochastically. The literature tells us that 1) Player B, who wants to avoid the target state, has a positional strategy that maximizes the probability to win (uniformly from every state) and 2) from every state, for every $\varepsilon > 0$, Player A has a strategy that maximizes up to ε the probability to win. Our work is two-fold.

First, we present a double-fixed-point procedure that says from which state Player A has a strategy that maximizes (exactly) the probability to win. This is computable if Nature's probability distributions are rational. We call these states *maximizable*. Moreover, we show that for every $\varepsilon > 0$, Player A has a positional strategy that maximizes the probability to win, exactly from maximizable states and up to ε from sub-maximizable states.

Second, we consider three-state games with one main state, one target, and one bin. We characterize the *local interactions* at the main state that guarantee the existence of an optimal Player A strategy. In this case there is a positional one. It turns out that in many-state games, these local interactions also guarantee the existence of a uniform optimal Player A strategy. In a way, these games are well-behaved by design of their elementary bricks, the local interactions. It is decidable whether a local interaction has this desirable property.

2012 ACM Subject Classification Theory of computation → Solution concepts in game theory

Keywords and phrases Concurrent reachability games, Game forms, Optimal strategies

Digital Object Identifier 10.4230/LIPIcs.CSL.2022.7

Related Version *Full Version*: https://arxiv.org/abs/2110.14724

1 Introduction

Stochastic concurrent games. Games on graphs are an intensively studied mathematical tool, with wide applicability in verification and in particular for the controller synthesis problem, see for instance [16, 1]. We consider two-player stochastic concurrent games played on finite graphs. For simplicity (but this is with no restriction), such a game is played over a finite bipartite graph called an arena: some states belong to Nature while others belong to the players. Nature is stochastic, and therefore assigns a probabilistic distribution over the players' states. In each players' state, a local interaction between the two players (called Player A and Player B) happens, specified by a two-dimensional table. Such an interaction is resolved as follows: Player A selects a probability distribution over the rows of the table while Player B selects a probability distribution over the columns of the table; this results into a distribution over the cells of the table, each one pointing to a Nature state of the graph. An example of game arena is given in Figure 1: circle states are players' while square states are Nature's; note that dashed arrows assign only probability 1 to a next state in this example (but in general could give probabilities to several states).

30th EACSL Annual Conference on Computer Science Logic (CSL 2022).
Editors: Florin Manea and Alex Simpson; Article No. 7; pp. 7:1–7:17
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

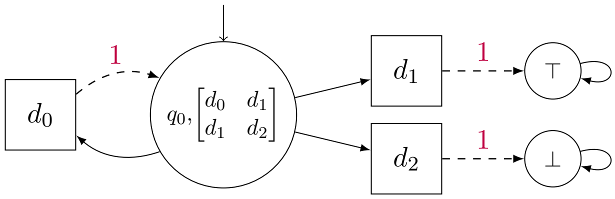

Figure 1 The game starts in q_0 with two actions available for each player. Player A wins if the state $\top$ is reached.

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix}$$

Figure 2 The local interaction at q_0 up to a renaming of the outcomes.

Globally, the game proceeds as follows: starting at an initial state q_0, the two players play the local interaction of the current state, and the joint choice determines (stochastically) the next Nature state of the game, itself moving randomly to players' states; the game then proceeds subsequently from the new players' state. The way players make choices is given by strategies, which, given the sequence of states visited so far (the so-called history), assign local strategies for the local interaction of the state the game is in. For application in controller synthesis, strategies will correspond to controllers, hence it is desirable to have strategies simple to implement. We will be in particular interested in strategies which are *positional*, that is, strategies which only depend on the current state of the game, not on the whole history. When each player has fixed a strategy (say $\mathsf{s_A}$ for Player A and $\mathsf{s_B}$ for Player B), this defines a probability distribution $\mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}$ over infinite sequences of states of the game. The objectives of the two players are opposite (we assume a zero-sum setting): together with the game, a measurable set W of infinite sequences of states is fixed; the objective of Player A is then to maximize the probability of W while the objective of Player B is to minimize this probability.

Back to the example of Figure 1, assume Player A (resp. B) plays the first row (resp. column) with probability p_A (resp. p_B), then the probability to move to $\top$ is $p_\mathsf{A} + p_\mathsf{B} - 2p_\mathsf{A}p_\mathsf{B}$. If Player A repeatedly plays the same strategy at q_0 with $p_\mathsf{A} < 1$, then the probability to reach $\top$ will lie between p_A and 1, depending on Player B; however, if she plays $p_\mathsf{A} = 1$, then by playing $p_\mathsf{B} = 1$, Player B enforces staying in q_0, hence reaching $\top$ with probability 0.

Values and (almost-)optimal strategies. As mentioned above, Player A wants to maximize the probability of W, while Player B wants to minimize this probability. Formally, given a strategy $\mathsf{s_A}$ for Player A, its value is measured by $\inf_{\mathsf{s_B}} \mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}(W)$, and Player A wants to maximize that value. Dually, given a strategy $\mathsf{s_B}$ for Player B, its value is measured by $\sup_{\mathsf{s_A}} \mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}(W)$, and Player B wants to minimize that value. Following Martin's determinacy theorem for Blackwell games [13], it actually holds that when W is Borel, then the game has a *value* given by

$$\chi_{q_0} = \sup_{\mathsf{s_A}} \inf_{\mathsf{s_B}} \mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}(W) = \inf_{\mathsf{s_B}} \sup_{\mathsf{s_A}} \mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}(W)$$

While this ensures the existence of almost-optimal strategies (that is, ε-optimal strategies for every $\varepsilon > 0$) for both players, it says nothing about the existence of optimal strategies, which are strategies achieving χ_{q_0}. In general, as already mentioned in [8], optimal strategies may not exist. Indeed assuming a reachability objective with target $\top$, the game in Figure 1 is such that $\chi_{q_0} = 1$, however Player A can only achieve $1 - \varepsilon$ for every $\varepsilon > 0$ by playing repeatedly at q_0 the first row of the table with probability $1 - \varepsilon$ and the second row with probability ε, but Player A cannot achieve 1.

Our setting. In this paper we focus on reachability games, that is, W is a reachability condition. They are a special case of recursive games (where targets are assigned payoffs), as studied in [8]. As such, they enjoy several nice properties: (i) Player A has positional almost-optimal strategies; (ii) Player B has positional optimal strategies [7]. These properties are specific to reachability games (or slight generalizations thereof), and this is for instance not the case of Büchi games, see [7, Thm. 2].

Our goal is to study *maximizable* and *sub-maximizable* states in (reachability) games: maximizable (resp. sub-maximizable) states are states from which optimal strategies exist (resp. no optimal strategies exist). Our contributions are then mostly twofolds:

1. We characterize via a double-fixed-point procedure maximizable and sub-maximizable states. This characterization cautiously analyzes when and why no optimal strategies will exist. Back to the example of Figure 1, we realize that no optimal strategy exists since at the limit of ε-optimal strategies, i.e. when Player A plays the first row almost-surely, Player B can enforce cycling back to q_0, hence disabling state $\top$.This simple analysis close to the target has to be propagated carefully in the game, in which some strategies which are designated as risky (since they ultimately lead to such a situation) have to be avoided.

   As a byproduct of our construction, we have Theorem 28, which establishes that one can build almost-optimal positional strategies, which are actually optimal where they can be. This refines the result of [8] which did not ensure optimality where it could.

   A consequence of that construction is that maximizable and sub-maximizable states can be computed under slight assumptions, and that witness positional strategies can be computed as well. For these results we rely on Tarski's decidability result of the theory of the reals [15].

   We also show that our result cannot be extended to games with countably many states by exhibiting such a game in which an optimal strategy exists, but there is no optimal positional strategy.

2. Local interactions played by the players are abstracted into game forms, where cells of the matrix are now seen as variables (some of them being equal). For instance, the game form associated with state q_0 in the running example has three outcomes: x, y and z, and it is given in Figure 2. Game forms can be seen as elementary bricks that can be used to build games on graphs. We can embed such a brick into various three-states games with one main state, one target, and one bin (as is done in Figure 1 for the interaction of Figure 2). We characterize the *local interactions* at the main state that guarantee the existence of an optimal Player A strategy. In this case there is a positional one. It turns out that in many-state games, these local interactions also guarantee the existence of a uniform optimal Player A strategy. In a way, these games are well-behaved by design of their elementary bricks, the local interactions. It is decidable whether a local interaction has this desirable property.

   Importantly we exhibit a simple condition on game forms which ensures the above: determined game forms as studied in [2] do satisfy the condition. The latter game forms generalize turn-based local interactions (where each players' state is controlled by a unique player – that is, the matrix defining the local interaction has a single row or a single column). We therefore recover the fact that stochastic turn-based reachability games admit optimal positional strategies, which was shown in [14, 4, 19].

Additional details and complete proofs are available in the arXiv version of this paper [3].

Related work. In [6], the authors characterize using fixed points as well states with value 1: sure-winning states (all generated plays satisfy the reachability condition – as if no probabilities were involved), almost-sure winning states (that is, maximizable states with value 1) and limit-sure winning states (that is, sub-maximizable states with value 1). Our work generalizes this result with states with arbitrary values.

There are many works dedicated to the study of stochastic turn-based games. These games enjoy more properties. Indeed, in parity stochastic turn-based games, Player A always has an optimal pure positional strategy [14, 4, 19]. These results do not extend in general to infinite (turn-based) arenas (even when they are finitely-branching): optimal strategies may not exist, and when they exist, they may require infinite memory [12].

2 Preliminaries

Consider a non-empty set Q. The *support* $\mathsf{Supp}(\mu)$ of a function $\mu : Q \to [0, 1]$ corresponds to set of non-0s of the function: $\mathsf{Supp}(\mu) = \{q \in Q \mid \mu(q) \in]0, 1]\}$. A *discrete probabilistic distribution* over a non-empty set Q is a function $\mu : Q \to [0, 1]$ such that its support $\mathsf{Supp}(\mu)$ is countable and $\sum_{x \in Q} \mu(x) = 1$. The set of all distributions over the set Q is denoted $\mathcal{D}(Q)$. We also consider the product order on vectors $\preceq : \mathbb{R}^n \times \mathbb{R}^n$ defined for any $n \in \mathbb{N}$ by, for all $v, v' \in \mathbb{R}^n$, we have $v \preceq v' \Leftrightarrow \forall i \in [\![1, n]\!],\ v(i) \leq v'(i)$. For $v \in \mathbb{R}^n$ and $x \in \mathbb{R}$, the notation $v + x$ refers to the vector $v' \in \mathbb{R}^n$ such that, for all $i \in [\![1, n]\!]$, we have $v'(i) = v(i) + x$.

3 Game Forms

We recall the definition of game forms which informally are 2-dim. tables with variables.

▶ **Definition 1** (Game form and game in normal form). *A* game form *is a tuple* $\mathcal{F} = \langle \mathsf{St_A}, \mathsf{St_B}, \mathsf{O}, \varrho \rangle$ *where* $\mathsf{St_A}$ *(resp.* $\mathsf{St_B}$*) is the non-empty set of (pure) strategies available to Player* A *(resp.* B*),* O *is a non-empty set of possible outcomes, and* $\varrho : \mathsf{St_A} \times \mathsf{St_B} \to \mathsf{O}$ *is a function that associates an outcome to each pair of strategies. When the set of outcomes* O *is equal to* $[0, 1]$*, we say that* $\mathcal{F}$ *is a* game in normal form*. For a valuation* $v \in [0, 1]^{\mathsf{O}}$ *of the outcomes, the notation* $\mathcal{F}^v$ *refers to the game in normal form* $\langle \mathsf{St_A}, \mathsf{St_B}, [0, 1], v \circ \varrho \rangle$*. A game form* $\mathcal{F} = \langle \mathsf{St_A}, \mathsf{St_B}, \mathsf{O}, \varrho \rangle$ *is* finite *if the set of pure strategies* $\mathsf{St_A} \cup \mathsf{St_B}$ *is finite.*

In the following, the game form $\mathcal{F}$ will always refer to the tuple $\langle \mathsf{St_A}, \mathsf{St_B}, \mathsf{O}, \varrho \rangle$ unless otherwise stated. Furthermore, we will be interested in valuations of the outcomes in the interval $[0, 1]$. Informally, Player A (the rows) tries to maximize the outcome, whereas Player B (the columns) tries to minimize it.

▶ **Definition 2** (Outcome of a game in normal form). *Consider a game in normal form* $\mathcal{F} = \langle \mathsf{St_A}, \mathsf{St_B}, [0, 1], \varrho \rangle$*. The set* $\mathcal{D}(\mathsf{St_A})$ *corresponds to the set of mixed strategies available to Player* A*, and analogously for Player* B*. For a pair of mixed strategies* $(\sigma_\mathsf{A}, \sigma_\mathsf{B}) \in \mathcal{D}(\mathsf{St_A}) \times \mathcal{D}(\mathsf{St_B})$*, the outcome* $\mathsf{out}_\mathcal{F}(\sigma_\mathsf{A}, \sigma_\mathsf{B})$ *in* $\mathcal{F}$ *of the strategies* $(\sigma_\mathsf{A}, \sigma_\mathsf{B})$ *is defined as:* $\mathsf{out}_\mathcal{F}(\sigma_\mathsf{A}, \sigma_\mathsf{B}) := \sum_{a \in \mathsf{St_A}} \sum_{b \in \mathsf{St_B}} \sigma_\mathsf{A}(a) \cdot \sigma_\mathsf{B}(b) \cdot \varrho(a, b) \in [0, 1]$.

The definition of the value of a game in normal form follows:

▶ **Definition 3** (Value of a game in normal form and optimal strategies). *Consider a game in normal form* $\mathcal{F} = \langle \mathsf{St_A}, \mathsf{St_B}, [0, 1], \varrho \rangle$ *and a strategy* $\sigma_\mathsf{A} \in \mathcal{D}(\mathsf{St_A})$ *for Player* A*. The* value *of strategy* σ_A*, denoted* $\mathsf{val}_\mathcal{F}(\sigma_\mathsf{A})$ *is equal to:* $\mathsf{val}_\mathcal{F}(\sigma_\mathsf{A}) := \inf_{\sigma_\mathsf{B} \in \mathcal{D}(\mathsf{St_B})} \mathsf{out}_\mathcal{F}(\sigma_\mathsf{A}, \sigma_\mathsf{B})$*, and analogously for Player* B*, with a* sup *instead of an* inf*. When* $\sup_{\sigma_\mathsf{A} \in \mathcal{D}(\mathsf{St_A})} \mathsf{val}_\mathcal{F}(\sigma_\mathsf{A}) = \inf_{\sigma_\mathsf{B} \in \mathcal{D}(\mathsf{St_B})} \mathsf{val}_\mathcal{F}(\sigma_\mathsf{B})$*, it defines the* value *of the game* $\mathcal{F}$*, denoted* $\mathsf{val}_\mathcal{F}$*.*

Note that von Neuman's minimax theorem [18] ensures it does as soon as the game $\mathcal{F}$ is finite. A strategy $\sigma_{\mathsf{A}} \in \mathcal{D}(\mathsf{St}_{\mathsf{A}})$ ensuring $\mathsf{val}_{\mathcal{F}} = \mathsf{val}_{\mathcal{F}}(\sigma_{\mathsf{A}})$ is called optimal. *The set of all optimal strategies for Player* A *is denoted $\mathsf{Opt}_{\mathsf{A}}(\mathcal{F}) \subseteq \mathcal{D}(\mathsf{St}_{\mathsf{A}})$, and analogously for Player* B. *Von Neuman's minimax theorem ensures the existence of optimal strategies (for both players).*

As it will be useful in Section 7, we define a least fixed point operator in a game form given a partial valuation of the outcomes.

▶ **Definition 4** (Total valuation induced by a partial valuation). *For a game form $\mathcal{F}$ and a partial valuation $\alpha : \mathsf{O} \setminus E \to [0,1]$ for some $E \subseteq \mathsf{O}$, we define the map $f_\alpha^{\mathcal{F}} : [0,1] \to [0,1]$ by, for all $y \in [0,1]$: $f_\alpha^{\mathcal{F}}(y) := \mathsf{val}_{\mathcal{F}^{\alpha[y]}}$ where $\alpha[y] : \mathsf{O} \to [0,1]$ is such that $\alpha[y][E] = \{y\}$ and $\alpha[y]|_{\mathsf{O}\setminus E} = \alpha$. The map f_α has a least fixed point (by monotonocity), denoted $v_\alpha \in [0,1]$. The valuation $\tilde{\alpha} \in [0,1]^{\mathsf{O}}$ induced by the partial valuation α is then equal to $\tilde{\alpha} = \alpha[v_\alpha]$.*

4 Concurrent stochastic games

In this section, we define the formalism we use throughout this paper for concurrent graph games, strategies and values.

▶ **Definition 5** (Stochastic concurrent games). *A finite* stochastic concurrent arena *$\mathcal{C}$ is a tuple $\langle A, B, Q, \mathsf{D}, \delta, \mathsf{dist}\rangle$ where A (resp. B) is the non-empty finite set of actions of Player* A *(resp.* B*), Q is the non-empty finite set of states,* D *is the non-empty set of Nature states, $\delta : Q \times A \times B \to \mathsf{D}$ is the transition function, $\mathsf{dist} : \mathsf{D} \to \mathcal{D}(Q)$ is the distribution function. A* concurrent reachability game *is a pair $\langle \mathcal{C}, \top\rangle$ where $\top \in Q$ is a target state (for Player* A*). It is supposed to be a self-looping sink: for all $a \in A$ and $b \in B$, we have $\mathsf{Supp}(\delta(\top, a, b)) = \{\top\}$.*

In the following, the arena $\mathcal{C}$ will always refer to the tuple $\langle A, B, Q, \mathsf{D}, \delta, \mathsf{dist}\rangle$ unless otherwise stated, and $\top$ to the target in the game $\langle \mathcal{C}, \top\rangle$, that we assume fixed in the rest of the definitions. Let us now consider a crucial tool in our study: the notion of local interaction. These are game forms induced by the transition function δ in states of the game.

▶ **Definition 6** (Local interaction). *The* local interaction *at state $q \in Q$ is the game form $\mathcal{F}_q := \langle A, B, \mathsf{D}, \delta(q, \cdot, \cdot)\rangle$. That is, the strategies available for Player* A *(resp.* B*) are the actions in A (resp. B) and the outcomes are the Nature states.*

Local interactions also allow us to define the probability transition to go from one state to another, given two local strategies.

▶ **Definition 7** (Probability transition). *Consider a state $q \in Q$ and two local strategies $(\sigma_{\mathsf{A}}, \sigma_{\mathsf{B}}) \in \mathcal{D}(A) \times \mathcal{D}(B)$ in the game form $\mathcal{F}_q$. Let $q' \in Q$. The probability $p^{q,q'}(\sigma_{\mathsf{A}}, \sigma_{\mathsf{B}})$ to go from q to q' if the players opt for strategies σ_{A} and σ_{B} is equal to the outcome of the game form $\mathcal{F}_q$ with the value of a Nature state $d \in \mathsf{D}$ equal to the probability to go from d to q', i.e. it is given by the valuation $\mathsf{dist}(\cdot)(q') \in [0,1]^{\mathsf{D}}$. That is: $p^{q,q'}(\sigma_{\mathsf{A}}, \sigma_{\mathsf{B}}) := \mathsf{out}_{\mathcal{F}_q^{\mathsf{dist}(\cdot)(q')}}(\sigma_{\mathsf{A}}, \sigma_{\mathsf{B}})$.*

Let us now look at the strategies we consider in such concurrent games.

▶ **Definition 8** (Strategies). *A Player* A *strategy is a map $\mathsf{s}_{\mathsf{A}} : Q^+ \to \mathcal{D}(A)$. It is said to be* positional *if, for all $\pi = \rho \cdot q \in Q^+$, we have $\mathsf{s}_{\mathsf{A}}(\pi) = \mathsf{s}_{\mathsf{A}}(q)$: the strategy only depends on the current state. We denote by $\mathsf{S}_{\mathcal{C}}^{\mathsf{A}}$ and $\mathsf{PS}_{\mathcal{C}}^{\mathsf{A}}$ the set of all strategies and positional strategies respectively in arena $\mathcal{C}$ for Player* A. *The definitions are analogous for Player* B.

A pair of strategies then induces a probability measure over paths.

▶ **Definition 9** (Probability measure of paths given two strategies). *For a pair of strategies* $(\mathsf{s_A}, \mathsf{s_B}) \in \mathsf{S}^{\mathsf{A}}_{\mathcal{C}} \times \mathsf{S}^{\mathsf{B}}_{\mathcal{C}}$*, we denote by* $\mathsf{s}^{\pi}_{\mathsf{A}} : Q^+ \to \mathcal{D}(A)$ *the Player* A *residual strategy after* $\pi \in Q^+$ *is seen: for all* $\pi' \in Q^+$, $\mathsf{s}^{\pi}_{\mathsf{A}}(\pi') = \mathsf{s_A}(\pi \cdot \pi')$*. The residual strategy* $\mathsf{s}^{\pi}_{\mathsf{B}}$ *is defined analogously. Then, the probability of occurrence of a finite path* $\pi \in Q^+$ *is defined inductively. For all starting states* $q_0 \in Q$*, for all* $q \cdot \pi \in Q^+$*, if* $q \neq q_0$*, we set* $\mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}(q) := 0$*. Furthermore,* $\mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}(q_0) := 1$ *and for all* $q \cdot \pi \in Q^+$*, we set:*

$$\mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}(q_0 \cdot q \cdot \pi) := p^{q_0,q}(\mathsf{s_A}(q_0), \mathsf{s_B}(q_0)) \cdot \mathbb{P}^{q}_{\mathsf{s}^{q_0}_{\mathsf{A}},\mathsf{s}^{q_0}_{\mathsf{B}}}(q \cdot \pi)$$

A probability measure $\mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}$ *is thus defined over the* σ*-algebra generated by cylinders (which are continuations of finite paths). Standardly (see e.g. [17]), infinite sequences of states visiting some subset* $Q' \subseteq Q$ *is measurable, and we note* $\mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}(Q')$ *(resp.* $\mathbb{P}^{q_0}_{\mathsf{s_A},\mathsf{s_B}}(n, Q')$*) the probability to reach* Q' *(resp. in at most* n *steps) from state* q_0*.*

Finally, we can define what is the value of strategies (for both players) and of the game.

▶ **Definition 10** (Value of strategies and of the game). *The value* $\chi^{\mathcal{C}}_{\mathsf{s_A}}(q)$ *of a Player* A *strategy* $\mathsf{s_A}$ *from a state* $q \in Q$ *is equal to* $\chi^{\mathcal{C}}_{\mathsf{s_A}}(q) := \inf_{\mathsf{s_B} \in \mathsf{S}^{\mathsf{B}}_{\mathcal{C}}} \mathbb{P}^{q}_{\mathsf{s_A},\mathsf{s_B}}(\top)$*. The value* $\chi^{\mathcal{C}}_{\mathsf{A}}(q)$ *of the game for Player* A *from* q *is:* $\chi^{\mathcal{C}}_{\mathsf{A}}(q) := \sup_{\mathsf{s_A} \in \mathsf{S}^{\mathsf{A}}_{\mathcal{C}}} \chi^{\mathcal{C}}_{\mathsf{s_A}}(q)$*. It is analogous for Player* B*, by inverting the* inf *and* sup*. When equality of these two values holds, it defines the* value *at state* q*, denoted* $\chi^{\mathcal{C}}(q)$*:* $\chi^{\mathcal{C}}(q) := \chi^{\mathcal{C}}_{\mathsf{A}}(q) = \chi^{\mathcal{C}}_{\mathsf{B}}(q) \in [0,1]$*. The value of the game is then given by the valuation* $\chi^{\mathcal{C}} \in [0,1]^Q$*. Since the game is finite, [13] gives that this equality is always ensured. A strategy* $\mathsf{s_A} \in \mathsf{S}^{\mathsf{A}}_{\mathcal{C}}$ *such that* $\chi^{\mathcal{C}}_{\mathsf{s_A}}(q) = \chi^{\mathcal{C}}_{\mathsf{A}}(q)$ *(resp.* $\chi^{\mathcal{C}}_{\mathsf{s_A}}(q) \geq \chi^{\mathcal{C}}_{\mathsf{A}}(q) - \varepsilon$ *for some* $\varepsilon > 0$*) is called a Player* A optimal *strategy (resp.* ε-optimal*) from state* q*. If* $\chi^{\mathcal{C}}_{\mathsf{s_A}} = \chi^{\mathcal{C}}_{\mathsf{A}}$*, the strategy* $\mathsf{s_A}$ *is* uniformly *optimal. This is defined analogously for Player* B*. For a valuation* $v \in [0,1]^Q$ *of the states, a Player* A *strategy* $\mathsf{s_A} \in \mathsf{S}^{\mathsf{A}}_{\mathcal{C}}$ *such that* $v \preceq \chi^{\mathcal{C}}_{\mathsf{s_A}}$ *is said to* guarantee *the valuation* v*.*

Value of the game and least fixed point. In the context of a reachability game, the value of the game is the least fixed point (lfp) of an operator on valuations on states. We define this operator here.

▶ **Definition 11** (Valuation of the Nature states and operator on values). *For* $v \in [0,1]^Q$*, we define the valuation* $\mu_v \in [0,1]^{\mathsf{D}}$ *of the Nature states by* $\mu_v(d) := \sum_{q \in Q} \mathsf{dist}(d)(q) \cdot v(q)$ *for all* $d \in \mathsf{D}$*. For the operator* $\Delta : [0,1]^Q \to [0,1]^Q$*, for all valuations* $v \in [0,1]^Q$*, we set* $\Delta(v)(\top) := 1$ *and, for all* $q \neq \top \in Q$*, we set* $\Delta(v)(q) := \mathsf{val}_{\mathcal{F}^{\mu_v}_q}$*.*

As the operator Δ is monotonous, it has an lfp for the product order $\preceq$. This lfp gives the value of the game. Furthermore, Player B has an optimal positional strategy:

▶ **Theorem 12** ([8, 9]). *Let* m *denote the lfp of the operator* Δ*. Then:* $\chi^{\mathcal{C}} = \mathsf{m}$*. Furthermore, there exists a positional strategy* $\mathsf{s_B} \in \mathsf{PS}^{\mathcal{C}}_{\mathsf{B}}$ *for Player* B *ensuring* $\chi^{\mathcal{C}}_{\mathsf{s_B}} = \chi^{\mathcal{C}} = \mathsf{m}$*.*

Markov decision process induced by a positional strategy. Once a Player A positional strategy is fixed, we obtain a Markov decision process, which, informally, is a game where only one player (here, Player B) plays (against probabilistic transitions).

▶ **Definition 13** (Induced Markov decision process). *Consider a Player* A *positional strategy* $\mathsf{s_A} \in \mathsf{PS}^{\mathsf{A}}_{\mathcal{C}}$*. The* Markov decision process Γ *(MDP for short) induced by the strategy* $\mathsf{s_A}$ *is the triplet* $\Gamma := \langle Q, B, \iota \rangle$ *where* Q *is the set of states,* B *is the set of actions and* $\iota : Q \times B \to \mathcal{D}(Q)$ *is a map associating to a state and an action a distribution over the states. For all* $q \in Q$*,* $b \in B$ *and* $q' \in Q$*, we set* $\iota(q,b)(q') := p^{q,q'}(\mathsf{s_A}(q), b)$*.*

Note that the set of Player B strategies in an induced MDP Γ is the same as in the concurrent game $\mathcal{C}$. Furthermore, the useful objects in MDPs are the end components [5]: informally, sub-MDPs that are strongly connected.

▶ **Definition 14** (End component). *Consider a Player* A *positional strategy* $\mathsf{s_A} \in \mathsf{PS}^{\mathsf{A}}_{\mathcal{C}}$ *and consider the MDP* Γ *induced by that strategy. An* end component *(EC for short)* H *in* Γ *is a pair* (Q_H, β) *such that* $Q_H \subseteq Q$ *is a subset of states and* $\beta : Q_H \to \mathcal{P}(B) \setminus \emptyset$ *associates to each state a non-empty set of actions compatible with the EC* H *such that:*

- *for all* $q \in Q_H$ *and* $b \in \beta(q)$*, we have* $\mathsf{Supp}(\iota(q, b)) \subseteq Q_H$*;*
- *the underlying graph* (Q_H, E) *is strongly connected where* $(q, q') \in E$ *iff* $q' \in \mathsf{Supp}(\iota(q, \beta(q)))$*.*

We denote by $\mathsf{D}_H \subseteq \mathsf{D}$ *the set of Nature states compatible with the EC* H*:* $\mathsf{D}_H = \{d \in \mathsf{D} \mid \mathsf{Supp}(d) \subseteq Q_H\}$*. Note that, for all* $q \in Q_H$ *and* $b \in \beta(q)$*, we have* $\delta(q, \mathsf{Supp}(\mathsf{s_A}(q)), b) \subseteq \mathsf{D}_H$*.*

The interest of ECs lies in the proposition below: in the MDP induced by a Player A strategy, for all Player B (positional) strategies (thus inducing a Markov chain), from all states, there is a non-zero probability to reach an EC from which it is impossible to exit.

▶ **Proposition 15.** *Consider a Player* A *positional strategy* $\mathsf{s_A} \in \mathsf{PS}^{\mathsf{A}}_{\mathcal{C}}$*. Let* $\mathcal{H}$ *denote the set of all ECs in the MDP induced by the strategy* $\mathsf{s_A}$*. For all Player* B *strategies* $\mathsf{s_B} \in \mathsf{PS}^{\mathsf{B}}_{\mathcal{C}}$*, there exists a subset of end components* $\mathcal{H}_{\mathsf{s_B}} \subseteq \mathcal{H}$ *called* bottom strongly conneted components *(BSCC for short): for all* $H = (Q_H, \beta) \in \mathcal{H}_{\mathsf{s_B}}$ *and* $q \in Q_H$*, we have* $\mathbb{P}^q_{\mathsf{s_A},\mathsf{s_B}}(Q \setminus Q_H) = 0$*. Furthermore, if* $q \in Q$*, we have:* $\mathbb{P}^q_{\mathsf{s_A},\mathsf{s_B}}(n, \cup_{H \in \mathcal{H}_{\mathsf{s_B}}} H) > 0$ *where* $n = |Q|$*.*

5 Crucial proposition

We fix a concurrent reachability game $\langle \mathcal{C}, T \rangle$ and a valuation $v \in [0,1]^Q$ of the states that Player A wants to guarantee. That is, she seeks a strategy $\mathsf{s_A}$ ensuring that for all $q \in Q$, it holds $\chi^{\mathcal{C}}_{\mathsf{s_A}}(q) \geq v(q)$. In particular, when $v = \mathsf{m}$, such a strategy $\mathsf{s_A}$ would be optimal. We state a sufficient condition for Player A positional strategies to ensure such a property.

Consider a Player A positional strategy $\mathsf{s_A} \in \mathsf{PS}^{\mathcal{C}}_{\mathsf{A}}$. The probability distribution chosen by this strategy only depends on the current state. In fact, this strategy is built with one (local) strategy per local interaction: for all state $q \in Q$, $\mathsf{s_A}(q) \in \mathcal{D}(A)$ is a strategy in the game form $\mathcal{F}_q$. As Player A wants to guarantee the valuation v, the valuation of interest of the outcomes of the game form $\mathcal{F}_q = \langle A, B, \mathsf{D}, \delta(q, \cdot, \cdot) \rangle$ is $\mu_v \in [0,1]^{\mathsf{D}}$ – lifting the valuation v to the Nature states. To ensure that $\chi^{\mathcal{C}}_{\mathsf{s_A}}(q) \geq v(q)$, one may think that it suffices to choose $\mathsf{s_A}(q)$ so that its value in the game in normal form $\mathcal{F}^{\mu_v}_q$ is at least $v(q)$, that is: $\mathsf{val}_{\mathcal{F}^{\mu_v}_q}(\mathsf{s_A}(q)) \geq v(q)$. In that case, the strategy $\mathsf{s_A}$ is said to locally dominate the valuation v:

▶ **Definition 16** (Strategy locally dominating a valuation). *A Player* A *positional strategy* $\mathsf{s_A} \in \mathsf{PS}^{\mathsf{A}}_{\mathcal{C}}$ locally dominates *the valuation* v *if, for all* $q \in Q$*, we have:* $\mathsf{val}_{\mathcal{F}^{\mu_v}_q}(\mathsf{s_A}(q)) \geq v(q)$*.*

However, this is not sufficient in the general case, as examplified in Figure 1. For the valuation $v = \chi^{\mathcal{C}}$ such that $v(q_0) = v(\top) = 1$ and $v(\bot) = 0$, a Player A positional strategy $\mathsf{s_A}$ that plays the first row in $\mathcal{F}_{q_0}$ with probability 1 ensures that $\mathsf{val}_{\mathcal{F}^{\mu_v}_{q_0}}(\mathsf{s_A}(q_0)) = 1 \geq v(q_0)$. However, we have seen that it does not ensure that $\chi^{\mathcal{C}}_{\mathsf{s_A}}(q_0) = 1$ since, if Player B always plays the first column, the game indefinitely loops in q_0. The issue is that, in the MDP induced by the strategy $\mathsf{s_A}$, the trivial end component $\{q_0\}$ is a trap, as it does not intersect the target set $\top$ – and therefore, the probability to reach $\top$ from q_0 is equal to 0 – whereas $\chi^{\mathcal{C}}(q_0) > 0$. In fact, as soon as this issue is avoided, if the strategy $\mathsf{s_A}$ locally dominates the valuation v, the desired property on $\mathsf{s_A}$ holds. Indeed:

▶ **Proposition 17.** *Consider a Player* A *positional strategy* $\mathsf{s_A} \in \mathsf{PS}_{\mathcal{C}}^{\mathsf{A}}$ *locally dominating* v*, and assume that* $v \preceq \mathsf{m}$*. Assume that for all end components* $H = (Q_H, \beta)$ *in the MDP induced by the strategy* $\mathsf{s_A}$*, if* $Q_H \neq \{\top\}$*, for all* $q_H \in Q_H$*, we have* $\chi^{\mathcal{C}}(q_H) = 0$ *(in other words, for all* $q \in Q$*, if* $\chi^{\mathcal{C}}_{\mathsf{s_A}}(q) = 0$ *then* $\chi^{\mathcal{C}}(q) = 0$*). In that case, for all* $q \in Q$*, we have* $\chi^{\mathcal{C}}_{\mathsf{s_A}}(q) \geq v(q)$ *(i.e. the strategy* $\mathsf{s_A}$ *guarantees the valuation* v*).*

Proof Sketch. Consider some $\varepsilon > 0$ and, for $x \in \{\varepsilon, \varepsilon/2\}$, the valuations $v_x = v - x \in [0,1]^Q$. We show that $\mathsf{s_A}$ guarantees v_ε. As this holds for all $\varepsilon > 0$, it follows that $\mathsf{s_A}$ guarantees v. Consider an arbitrary positional strategy $\mathsf{s_B}$ for Player B. Let κ_{A} be a Player A strategy guaranteeing $v_{\varepsilon/2}$ in $n \geq 0$ steps from every state (which exists since $v_{\varepsilon/2} \prec \mathsf{m}$) and a strategy κ_{B} for Player B optimal against κ_{A}. So $\mathbb{P}^q_{\kappa_{\mathsf{A}},\kappa_{\mathsf{B}}}(n, \top) \geq v_{\varepsilon/2}(q)$ for all $q \in Q$. Now, for all $l \geq 0$, we consider the strategy $\mathsf{s}^l_{\mathsf{A}}$ that plays $\mathsf{s_A}$ l times and then plays κ_{A} (and similarly for a strategy $\mathsf{s}^l_{\mathsf{B}}$ for Player B). As $\mathsf{s_A}$ locally dominates v, it also locally dominates $v_{\varepsilon/2}$ which is obtained from v by translation. Therefore, for any state $q \in Q$, if the local strategy $\mathsf{s_A}(q)$ is played in q, then the convex combination of the values of the successors of q w.r.t. the valuation $v_{\varepsilon/2}$ is at least $v_{\varepsilon/2}(q)$. In other words, the probability to reach $\top$ from q in $1+n$ steps if the strategy $\mathsf{s}^1_{\mathsf{A}}$ is played is at least $v_{\varepsilon/2}(q)$: $\mathbb{P}^q_{\mathsf{s}^1_{\mathsf{A}},\mathsf{s}^1_{\mathsf{B}}}(1+n, \top) \geq v_{\varepsilon/2}(q)$. In fact, by induction, this holds for all $l \geq 0$: $\mathbb{P}^q_{\mathsf{s}^l_{\mathsf{A}},\mathsf{s}^l_{\mathsf{B}}}(l+n, \top) \geq v_{\varepsilon/2}(q)$. Now, with strategies $\mathsf{s}^l_{\mathsf{A}}$ and $\mathsf{s}^l_{\mathsf{B}}$, consider the state of the game after l steps: either it is in a BSCC (w.r.t. $\mathsf{s_A}$ and $\mathsf{s_B}$) or it is not. For a sufficiently large l, the probability not to have reached a BSCC is as close to 0 as we want. Furthermore, for a state q_H in a BSCC H that is not $\{\top\}$, by assumption, we have that $\chi^{\mathcal{C}}(q_H) = 0$, hence $\mathbb{P}^{q_H}_{\kappa_{\mathsf{A}},\kappa_{\mathsf{B}}}(\top) = 0$. In addition, if the state is in the trivial BSCC $\{\top\}$, then $\top$ is reached. Hence, for l large enough, the two probabilities $\mathbb{P}^q_{\mathsf{s}^l_{\mathsf{A}},\mathsf{s}^l_{\mathsf{B}}}(l+n, \top)$ and $\mathbb{P}^q_{\mathsf{s}^l_{\mathsf{A}},\mathsf{s}^l_{\mathsf{B}}}(l, \top)$ are as close to one another as we want. Finally, note that the strategies $\mathsf{s}^l_{\mathsf{A}}, \mathsf{s}^l_{\mathsf{B}}$ behave exactly like the strategies $\mathsf{s_A}, \mathsf{s_B}$ in the first l steps. That is, for l large enough, and $q \in Q$, we have $\mathbb{P}^q_{\mathsf{s_A},\mathsf{s_B}}(\top) \geq \mathbb{P}^q_{\mathsf{s_A},\mathsf{s_B}}(l, \top) = \mathbb{P}^q_{\mathsf{s}^l_{\mathsf{A}},\mathsf{s}^l_{\mathsf{B}}}(l, \top) \geq \mathbb{P}^q_{\mathsf{s}^l_{\mathsf{A}},\mathsf{s}^l_{\mathsf{B}}}(l+n, \top) - \varepsilon/2 \geq v_{\varepsilon/2}(q) - \varepsilon/2 = v_\varepsilon(q)$. ◀

Fix a Player A positional strategy $\mathsf{s_A}$ locally dominating the valuation v and let Γ be the MDP induced by $\mathsf{s_A}$. For $\mathsf{s_A}$ to guarantee the valuation v, it suffices to ensure that any EC in Γ that is not the trivial EC $\{\top\}$ has all its states of value 0. It does not necessarily hold for $\mathsf{s_A}$ (recall the explanations before Proposition 17). However, we do have the following: fix an EC H in Γ. Then, all the states H have the same value w.r.t. the valuation v. It is stated in the proposition below.

▶ **Proposition 18.** *Consider a Player* A *positional strategy* $\mathsf{s_A} \in \mathsf{PS}_{\mathcal{C}}^{\mathsf{A}}$ *locally dominating a valuation* $v \in [0,1]^Q$*. For all EC* $H = (Q_H, \beta)$ *in the MDP induced by the strategy* $\mathsf{s_A}$*, there exists* $v_H \in [0,1]$ *such that, for all* $q \in Q_H$*, we have* $v(q) = v_H$*. Furthermore, for all* $q \in Q_H$*, we have* $\mathsf{val}_{\mathcal{F}^{\mu_v}_q}(\mathsf{s_A}(q)) = v(q)$*.*

6 Positional optimal and ε-optimal strategies

The aim of this section is, given a concurrent reachability game, to determine exactly from which states Player A has an optimal strategy. This, in turn, will give that whenever she has an optimal strategy, she has one that is positional which therefore extends Everett [8] (the existence of positional ε-optimal strategies). We fix a concurrent reachability game $\langle \mathcal{C}, \top \rangle$ for the rest of this section. Let us first introduce some terminology.

▶ **Definition 19** (Maximizable and sub-maximizable states)**.** *A state* $q \in Q$ *from which Player* A *has (resp. does not have) an optimal strategy is called* maximizable *(resp.* sub-maximizable*). The set of such states is denoted* $\mathsf{MaxQ_A}$ *(resp.* $\mathsf{SubMaxQ_A}$*).*

The value of that game is given by the vector $\mathsf{m} \in [0,1]^Q$ (from Definition 11). We want to build an optimal (and positional) strategy for Player A when possible. To be optimal, a Player A positional strategy $\mathsf{s_A}$ has to play optimally at each local interaction $\mathcal{F}_q$ (for $q \in Q$) with regard to the valuation $\mu_{\mathsf{m}} \in [0,1]^{\mathsf{D}}$ (lifting the valuation m to Nature states). However, it is not sufficient in general: in the snow-ball game of Figure 1, when Player A plays optimally in $\mathcal{F}_{q_0}$ w.r.t. the valuation μ_{m} (that is, plays the first line with probability 1), Player B can enforce the play never to leave the state $q_0 \neq \top$. Hence, locally, we want to have strategies that not only play optimally but, regardless of the choice of Player B, have a non-zero probability to get closer to the target $\top$. Such strategies will be called *progressive strategies*. To properly define them, we introduce the following notation.

▶ **Definition 20** (Optimal action)**.** *Let $q \in Q$ be a state of the game. Consider the game in normal form $\mathcal{F}_q^{\mu_{\mathsf{m}}}$. For all strategies $\sigma_{\mathsf{A}} \in \mathcal{D}(\mathsf{St_A})$, we define the set $B_{\sigma_{\mathsf{A}}}$ of* optimal actions *w.r.t. the strategy σ_{A} by $B_{\sigma_{\mathsf{A}}} := \{b \in B \mid \mathsf{out}_{\mathcal{F}_q^{\mu_{\mathsf{m}}}}(\sigma_{\mathsf{A}}, b) = \mathsf{val}_{\mathcal{F}_q^{\mu_{\mathsf{m}}}}(\sigma_{\mathsf{A}})\}$.*

In Figure 3, the set $B_{\sigma_{\mathsf{A}}}$ of optimal actions w.r.t. the strategy σ_{A} are represented in bold purple: the weighted values of these actions is the value of the strategy: $1/2$.

We can now define the set of *progressive* strategies.

▶ **Definition 21** (Progressive strategies)**.** *Consider a state $q \in Q$ and a set of states $\mathsf{Gd} \subseteq Q$ that Player A wants to reach. The set of Nature states $\mathsf{Gd_D} \subseteq \mathsf{D}$ corresponds to the Nature states with a non-zero probability to reach the set Gd: $\mathsf{Gd_D} := \{d \in \mathsf{D} \mid \mathsf{Supp}(\mathsf{dist}(d)) \cap \mathsf{Gd} \neq \emptyset\}$. Then, the set of* progressive strategies *$\mathsf{Prog}_q(\mathsf{Gd})$ at state q w.r.t. Gd is defined by $\mathsf{Prog}_q(\mathsf{Gd}) := \{\sigma_{\mathsf{A}} \in \mathsf{Opt_A}(\mathcal{F}_q^{\mu_{\mathsf{m}}}) \mid \forall b \in B_{\sigma_{\mathsf{A}}},\ \delta(q, \mathsf{Supp}(\sigma_{\mathsf{A}}), b) \cap \mathsf{Gd_D} \neq \emptyset\}$.*

In Figure 3, the Nature states in $\mathsf{Gd_D}$ are arbitrarily chosen for the example and circled in green. The depicted strategy is progressive as, for all bold purple actions, there is a green-circled state in the support of the strategy (the circled $3/4$).

However, in an arbitrary game, some states may be sub-maximizable. In that case, playing optimally implies avoiding these states. Given a set $\mathsf{Bd} \subseteq Q$ of states to avoid, an optimal strategy that has a non-zero probability to reach that set of states Bd is called *risky*.

▶ **Definition 22** (Risky strategies)**.** *Let $q \in Q$ be a state of the game and $\mathsf{Bd} \subseteq Q$ be a set of sub-maximizable states. The corresponding set of Nature states $\mathsf{Bd_D} \subseteq \mathsf{D}$ is defined similarly to $\mathsf{Gd_D}$ in Definition 21: $\mathsf{Bd_D} := \{d \in \mathsf{D} \mid \mathsf{Supp}(\mathsf{dist}(d)) \cap \mathsf{Bd} \neq \emptyset\}$. Then, the set of* risky strategies *$\mathsf{Risk}_q(\mathsf{Bd})$ at state q w.r.t. Bd is defined by $\mathsf{Risk}_q(\mathsf{Bd}) := \{\sigma_{\mathsf{A}} \in \mathsf{Opt_A}(\mathcal{F}_q^{\mu_{\mathsf{m}}}) \mid \exists b \in B_{\sigma_{\mathsf{A}}},\ \delta(q, \mathsf{Supp}(\sigma_{\mathsf{A}}), b) \cap \mathsf{Bd_D} \neq \emptyset\}$.*

In Figure 3, the set of Nature states $\mathsf{Bd_D}$ are also arbitrarily chosen for the example and circled in red. The strategy σ_{A} is not risky since no red-squared state appears in the intersection of the support of σ_{A} and the purple actions in B_{σ_A}.

In fact, we want for local strategies to be *efficient*, that is both progressive and not risky.

▶ **Definition 23** (Efficient strategies)**.** *Let $q \in Q$ be a state of the game and $\mathsf{Gd}, \mathsf{Bd} \subseteq Q$ be sets of states. The set of* efficient strategies *$\mathsf{Eff}_q(\mathsf{Gd}, \mathsf{Bd})$ at state q w.r.t. Gd and Bd is defined by $\mathsf{Eff}_q(\mathsf{Gd}, \mathsf{Bd}) := \mathsf{Prog}_q(\mathsf{Gd}) \setminus \mathsf{Risk}_q(\mathsf{Bd})$.*

In Figure 3, the strategy σ_{A} is efficient as it is both progressive and not risky.

We can now compute inductively the set of maximizable and sub-maximizable states. First, given a set of sub-maximizable states Bd, we define iteratively below a set of *secure* states w.r.t. Bd, there are the states with a non-zero probability to get closer to the target $\top$ while avoiding the set Bd. The construction is illustrated in Figure 4.

σ_A :				
0	1/2	**1/2**	**0**	0
0.5	1	**3/4**	**1/4**	1
0.5	3/4	**1/4**	**3/4**	1/2
0	0	**0**	**1/2**	1

Figure 3 A game in normal form with an optimal strategy depicted in brown on the left. Its value is $1/2 = 1/2 \cdot 3/4 + 1/2 \cdot 1/4$.

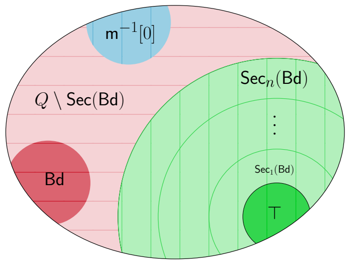

Figure 4 The construction of Definition 24 of the set of states $\mathsf{Sec}(\mathsf{Bd})$: it is the reunion of the blue and green vertical stripe areas.

▶ **Definition 24** (Secure states). *Consider a set of states* $\mathsf{Bd} \subseteq Q$. *We set* $\mathsf{Sec}_0(\mathsf{Bd}) := \{\top\}$ *and, for all* $i \geq 0$, $\mathsf{Sec}_{i+1}(\mathsf{Bd}) := \mathsf{Sec}_i(\mathsf{Bd}) \cup \{q \in Q \setminus \mathsf{Bd} \mid \mathsf{Eff}_q(\mathsf{Sec}_i(\mathsf{Bd}), \mathsf{Bd}) \neq \emptyset\}$. *The set* $\mathsf{Sec}(\mathsf{Bd})$ *of states* secure *w.r.t.* Bd *is:* $\mathsf{Sec}(\mathsf{Bd}) := \cup_{n \in \mathbb{N}} \mathsf{Sec}_n(\mathsf{Bd}) \cup \mathsf{m}^{-1}[0]$.

Note that, as the game $\mathcal{C}$ is finite, this procedure ends in at most $n = |Q|$ steps. Furthermore, the states of value 0 are added since any state of value 0 is maximizable. The interest of this construction lies in the lemma below: if all states in Bd are sub-maximizable, then all states in $Q \setminus \mathsf{Sec}(\mathsf{Bd})$ also are.

▶ **Lemma 25.** *Assume that a set of states* Bd *is such that* $\mathsf{Bd} \subseteq \mathsf{SubMaxQ_A}$. *Then, the set of states* $Q \setminus \mathsf{Sec}(\mathsf{Bd})$ *is such that* $Q \setminus \mathsf{Sec}(\mathsf{Bd}) \subseteq \mathsf{SubMaxQ_A}$ *(these correspond to the red horizontal stripe areas in Figure 4).*

Proof Sketch. For an arbitrary Player A strategy $\mathsf{s_A} \in \mathsf{S}_{\mathcal{C}}^{\mathsf{A}}$ to be optimal, it roughly needs, on all relevant paths, to be optimal. More precisely, on any finite path $\pi = \pi' \cdot q \in Q^+$ with a non-zero probability to occur if Player B plays (locally) optimal actions against the strategy $\mathsf{s_A}$ (called a relevant path), the strategy $\mathsf{s_A}$ needs to play an optimal (local) strategy in the local interaction $\mathcal{F}_q$ and it[1] has to be optimal from q in the reachability game. Therefore, on all relevant paths, the strategy $\mathsf{s_A}$, locally, has to play optimal strategies that are not risky. However, in any local interaction of a state $q \in Q \setminus \mathsf{Sec}(\mathsf{Bd})$, there is no efficient strategies available to Player A. Therefore, if the game starts from a state $q \in Q \setminus \mathsf{Sec}(\mathsf{Bd})$ an optimal strategy $\mathsf{s_A}$ for Player A (which therefore is locally optimal but not progressive) would allow Player B to ensure staying in the set $Q \setminus \mathsf{Sec}(\mathsf{Bd})$ while playing optimal actions. In that case, the game never leaves the set $Q \setminus \mathsf{Sec}(\mathsf{Bd})$, which induces a value of 0, whereas $\chi^{\mathcal{C}}(q) > 0$ since $q \notin \mathsf{Sec}(\mathsf{Bd})$. Thus, there is no optimal strategy for Player A from a state in $Q \setminus \mathsf{Sec}(\mathsf{Bd})$. ◀

We define inductively the set of bad states (which, in turn, will correspond to the set of sub-maximizable states) below.

▶ **Definition 26** (Set of sub-maximizable states). *Let* $\mathsf{Bad}_0 := \emptyset$ *and, for all* $i \geq 0$, $\mathsf{Bad}_{i+1} := Q \setminus \mathsf{Sec}(\mathsf{Bad}_i)$. *Then, the set* Bad *of bad states is equal to* $\mathsf{Bad} := \cup_{n \in \mathbb{N}} \mathsf{Bad}_n$ *for* $n = |Q|$.

Note that, as in the case of the set of secure states, since the game $\mathcal{C}$ is finite, this procedure ends in at most $n = |Q|$ steps. Lemma 25 ensures that the set of states Bad is included in $\mathsf{SubMaxQ_A}$. In addition, we have that there exists a Player A positional strategy optimal from all states q in its complement $\mathsf{Sec}(\mathsf{Bad}) = Q \setminus \mathsf{Bad}$, as stated in the lemma below.

[1] In fact, the residual strategy $\mathsf{s}_{\mathsf{A}}^{\pi'}$.

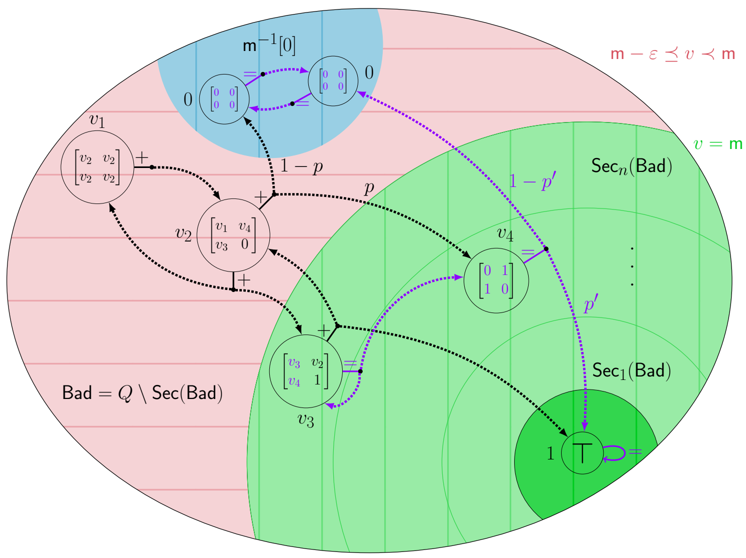

Figure 5 An illustration of the proof of Lemma 27 on the MDP induced by the strategy $\mathsf{s_A}$. Labels $v_1, \ldots, v_4$ is the value of the corresponding states given by the valuation v.

▶ **Lemma 27.** *For all $\varepsilon > 0$, there exists a positional strategy $\mathsf{s_A} \in \mathsf{PS}_\mathsf{A}^\mathcal{C}$ s.t.:*

- *for all $q \in \mathsf{Sec}(\mathsf{Bad})$, we have $\chi^\mathcal{C}_{\mathsf{s_A}}(q) = \mathsf{m}(q)$;*
- *for all $q \in \mathsf{Bad}$, we have $\chi^\mathcal{C}_{\mathsf{s_A}}(q) \geq \mathsf{m}(q) - \varepsilon$.*

In particular, it follows that $\mathsf{Sec}(\mathsf{Bad}) \subseteq \mathsf{MaxQ_A}$.

Proof Sketch. To prove this lemma, we define a Player A positional strategy $\mathsf{s_A} \in \mathsf{PS}^\mathsf{A}_\mathcal{C}$, a valuation $v \in [0,1]^Q$ of the states, prove that the strategy $\mathsf{s_A}$ locally dominates that valuation and prove that the only EC compatible with $\mathsf{s_A}$ that is not the target has value 0. This will show that is guarantees the valuation v by applying Proposition 17. As we want the strategy $\mathsf{s_A}$ to be optimal from all secure states, we consider a partial valuation v such that $v|_{\mathsf{Sec}(\mathsf{Bad})} := \mathsf{m}|_{\mathsf{Sec}(\mathsf{Bad})}$ (we will define it later on Bad). Then, on all secure states $q \in \mathsf{Sec}_i(\mathsf{Bad})$, we set $\mathsf{s_A}(q)$ to be an efficient strategy w.r.t. $\mathsf{Sec}_{i-1}(\mathsf{Bad})$ and Bad, i.e. $\mathsf{s_A}(q) \in \mathsf{Eff}_q(\mathsf{Sec}_{i-1}(\mathsf{Bad}), \mathsf{Bad})$. In particular, $\mathsf{s_A}(q)$ is optimal in the game form $\mathcal{F}_q$ w.r.t. the valuation μ_m. However, we know that no strategy can be optimal from states in Bad. Hence, we consider a valuation v that is ε-close to the valuation m on states in Bad for a well-chosen $\varepsilon > 0$. This ε is chosen so that the value of the local strategy $\mathsf{s_A}(q)$ for $q \in \mathsf{Sec}(\mathsf{Bad})$ is at least $v(q)$ w.r.t. the valuation μ_v[2]. We can now define the valuation v and the strategy $\mathsf{s_A}$ on Bad such that the value of $\mathsf{s_A}(q)$ in $\mathcal{F}_q$ w.r.t. μ_v is greater than $v(q)$: $\mathsf{val}_{\mathcal{F}_q^{\mu_v}}(\mathsf{s_A}(q)) > v(q)$ (this requires a careful use the fact that the operator Δ from Section 4 is 1-Lipschitz). The valuation v and the strategy $\mathsf{s_A}$ are now completely defined on Q. By definition, the strategy $\mathsf{s_A}$ locally dominates the valuation v.

The MDP induced by the strategy $\mathsf{s_A}$ is schematically depicted in Figure 5. The different split arrows appearing in the figure correspond to the actions (or columns in the local interactions) available to Player B. Black +-labeled-split arrows correspond to the actions of Player B that increase the value of v (i.e. in a state q, such that the convex combination –

[2] Specifically, ε has to be chosen smaller than the smallest difference between the values of an optimal actions $b \in B_{\mathsf{s_A}(q)}$ and a non-optimal action $b \in B_{\mathsf{s_A}(q)}$.

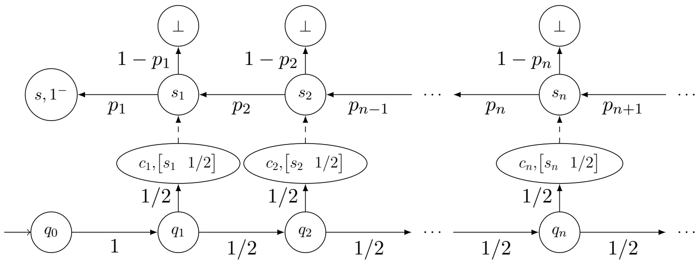

Figure 6 An infinite concurrent reachability game $\mathcal{C}$ (the Nature states are omitted). The probabilities p_k are such that, for all $i \geq 1$, the value of the state s_i is $\chi^{\mathcal{C}}(s_i) = \Pi_{k=1}^{i} p_k = (1/2{+}1/2^i)$.

w.r.t. to the probabilities chosen by the strategy $\mathsf{s_A}$ – of the values w.r.t. v of the successor states of q is greater than $v(q)$). For instance, we have $v_2 < p \cdot v_4 + (1-p) \cdot 0$, where the probability $p \in [0,1]$ is set by the strategy $\mathsf{s_A}$. On the other hand, purple =-labeled-split arrows correspond to the actions whose values do not increase the value of the state. For instance $v_4 = (1-p') \cdot 0 + p' \cdot 1$. We can see that the only split arrows exiting states in Bad (the red horizontal stripe area) are black (since $\mathsf{val}_{\mathcal{F}_q^{\mu_v}}(\mathsf{s_A}(q)) > v(q)$ for all $q \in \mathsf{Bad}$). However, from a secure state $q \in \mathsf{Sec}(\mathsf{Bad})$ (the green and blue vertical stripe areas) there are also purple split arrows. Note that, in these secure states $q \in \mathsf{Sec}(\mathsf{Bad})$, purple split arrows correspond to the optimal actions $B_{\mathsf{s_A}(q)}$ at the local interaction $\mathcal{F}_q$. Furthermore, these split arrows cannot exit the set of secure states $\mathsf{Sec}(\mathsf{Bad})$ since the local strategy $\mathsf{s_A}(q)$ is not risky.

We can then prove that the strategy $\mathsf{s_A}$ guarantees the valuation v by applying Proposition 17: since $\mathsf{s_A}$ locally dominates the valuation v, it remains to show that all the ECs different from $\{\top\}$ have only states of value 0. In the figure, this corresponds to having ECs only in the blue upper circle and dark green bottom right inner circle areas. In fact, Proposition 18 gives that any state q in an EC ensures $\mathsf{val}_{\mathcal{F}_q^{\mu_v}}(\mathsf{s_A}(q)) = v(q)$, which implies that no state in Bad can be in an EC. This can be seen in the figure between the states of value v_1 and v_2: because of the black arrow from v_1 to v_2, we necessarily have $v_1 < v_2$. Then, v_2 cannot loop (with probability one) to v_1 since this would imply $v_2 < v_1$. As all the split arrows are black for states in Bad, no EC can appear in this region. Furthermore, the optimal actions in the secure states always have a non-zero probability to get closer to the target $\top$. In the figure, this corresponds to the fact that there is always one tip of a purple split arrow that goes down in the $(\mathsf{Sec}_i(\mathsf{Bad}))_{i \in \mathbb{N}}$ hierarchy (since the strategy $\mathsf{s_A}(q)$ is progressive): in the example, from v_3 to v_4 and from v_4 to the target $\top$. Therefore, the only loop (with probability one) that can occur in the set $(\mathsf{Sec}_i(\mathsf{Bad}))_{i \in \mathbb{N}}$ is at the target $\top$. We conclude by applying Proposition 17. ◀

Overall, we obtain the theorem below summarizing the results proved in this section.

▶ **Theorem 28.** *In a concurrent reachability game* $\langle \mathcal{C}, \top \rangle$, *we have* $\mathsf{Bad} = \mathsf{SubMaxQ_A}$ *and* $\mathsf{Sec}(\mathsf{Bad}) = \mathsf{MaxQ_A}$. *Furthermore, for all* $\varepsilon > 0$, *there is a Player* A *positional strategy* $\mathsf{s_A}$ *optimal from all states in* $\mathsf{MaxQ_A}$ *and* ε*-optimal from all states in* $\mathsf{SubMaxQ_A}$.

Infinite arenas. In this paper, we only consider finite arenas and the constructions we have exhibited and results we have shown hold in that setting. Note that Theorem 28 does not hold on infinite arenas (i.e. with an infinite number of states): Figure 6 depicts an infinite concurrent reachability game where the state q_0 is maximizable but, from q_0, Player A does not have any positional optimal strategy. Indeed, in state s is plugged the game of Figure 1, whose value is 1 but Player A does not have an optimal strategy. Then, for all $i \geq 0$, the probability to reach s from s_i is equal to $v_i = (1/2 + 1/2^i) > 1/2$. Hence, if Player A plays an $0 < \varepsilon_i$-optimal strategy in s such that $(1 - \varepsilon_i) \cdot q_i > 1/2$, then the value of the state s_i is greater than $1/2$. In that case, in the states c_i, Player B plays the second columns obtaining the value $1/2$. This induces that the value in all states q_i is $1/2$. However, this is only possible if Player A has (infinite) memory, since the greater the index i considered, the smaller the value of ε_i needs to be to ensure $(1 - \varepsilon_i) \cdot q_i \geq 1/2$ while still ensuring $\varepsilon_i > 0$ (since Player A does not have an optimal strategy from s). In particular, for any Player A positional strategy s_A from q_0 that is $0 < \varepsilon$-optimal in s, the value – w.r.t. the strategy s_A – of all states s_i for indexes i such that $(1 - \varepsilon) \cdot q_i < 1/2$ is smaller than $1/2$. In which case, Player B plays the first column in c_i, thus obtaining a value smaller than $1/2$. It follows that the value of all states $(q_n)_{n \geq 0}$ – w.r.t. the strategy s_A – is smaller than $1/2$. Hence, any Player A positional strategy is not optimal from q_0. Note that, when considering MDPs instead of two-player games, optimal strategies need not exist but when they do there necessarily are positional ones (see for instance [10]).

Computing the set of maximizable states. Finally, consider the problem, given a finite concurrent reachability game, to effectively compute the set of maximizable and sub-maximizable states (assuming the probability distribution of the Nature states are rational). In fact, this can be done by using the theory of the reals.

▶ **Definition 29** (First-order theory of the reals)**.** *The* first-order theory of the reals *(denoted* FO-$\mathbb{R}$*) corresponds to the well-formed sentences of first-order logic (i.e. with universal and existential quandtificators), also involving logical combinations of equalities and inequalities of real polynomials, with integer coefficients.*

The first-order theory of the reals is decidable [15], i.e. determining if a given formula belonging to that theory is true is decidable. Now, let us consider a finite concurrent reachability game $\mathcal{C}$ and a state $q \in Q$. It is possible to encode, with an FO-$\mathbb{R}$ formula, that the state q is maximizable, i.e. $q \in \mathsf{MaxQ}_\mathsf{A}$. First, note that, given two positional strategies s_A and s_B for both players, it is possible to compute the value of the game with the theory of reals: it amounts to finding the least fixed point of the operator Δ with the strategies of both players fixed. Then, q being maximizable, denoting $u := \chi^{\mathcal{C}}(q) \in [0, 1]$ its value, is equivalent to having a Player A positional strategy ensuring at least u (against all Player B positional strategies) and no Player A positional strategy ensures more than u (as ε-optimal positional strategies always exists for Player A [8]). This can be expressed in FO-$\mathbb{R}$. The theorem below follows.

▶ **Theorem 30.** *In a finite concurrent reachability game with rational distributions, the set of maximizable states is computable.*

7 Maximizable states and game forms

In the previous section, we were given a concurrent reachability game and we considered a construction to compute exactly the sets of maximizable and sub-maximizable states. It is rather cumbersome as it requires two nested fixed point procedures. Now, we would like to have a structural condition ensuring that if a game is built correctly (i.e. built from reach-maximizable local interactions), then all states are maximizable. More specifically, in this section, we characterize exactly the *reach-maximizable* game forms, that is the game forms such that every reachability game built with these game forms as local interactions have only maximizable states.

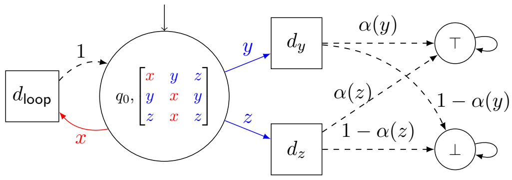

Figure 7 The three-state reachability game $\langle \mathcal{C}_{(\mathcal{F},\alpha)}, \top \rangle$ built from the game form $\mathcal{F}$ for some partial valuation $\alpha : \mathsf{O} \setminus E \to [0,1]$ with $E = \{x\}$.

$$\mathcal{F} = \begin{bmatrix} x & y & z \\ y & x & y \\ z & x & z \end{bmatrix}$$

Figure 8 The game form that constitutes the local interaction in the state q_0.

First, let us characterize a necessary condition for game forms to be reach-maximizable. We want for reach-maximizable game forms to behave properly when used individually. That is, from a game form $\mathcal{F}$ and a partial valuation $\alpha : \mathsf{O} \setminus E \to [0,1]$ of the outcomes, we define a three-state reachability game $\langle \mathcal{C}_{(\mathcal{F},\alpha)}, \top \rangle$. Note that such games were previously studied in [11]. We illustrate this construction on an example.

▶ **Example 31.** In Figure 7, a three-state reachability game $\langle \mathcal{C}_{(\mathcal{F},\alpha)}, \top \rangle$ is built from a game form $\mathcal{F} = \langle \mathsf{St}_\mathsf{A}, \mathsf{St}_\mathsf{B}, \{x, y, z\}, \varrho \rangle$ – with ϱ depicted in Figure 8 – and a partial valuation $\alpha : \{y, z\} \to [0, 1]$. We have a one-to-one correspondence between the outcomes of the game form $\mathcal{F}$ and the Nature states of the reachability game $\langle \mathcal{C}_{(\mathcal{F},\alpha)}, \top \rangle$ via the bijection $g : \{x, y, z\} \to \mathsf{D}$ such that $g(x) = d_{\mathsf{loop}}$ and for $u \in \{y, z\}$, $g(u) = d_u$. Furthermore, in the reachability game $\langle \mathcal{C}_{(\mathcal{F},\alpha)}, \top \rangle$, we have $\mathsf{m}(\top) = 1$ and $\mathsf{m}(\bot) = 0$. Therefore, for $u \in \{y, z\}$, we have $\mu_\mathsf{m} \circ g(u) = \alpha(u)$. In fact, this game is built so that $v_\alpha = \mathsf{m}(q_0)$ and $\mu_\mathsf{m} = \tilde{\alpha} \circ g^{-1}$ (recall that $\tilde{\alpha}$ is the (total) valuation induced by the partial valuation α from Definition 4).

Let us now determine at which condition on the pair $(\mathcal{F}, \alpha)$ is the starting state q_0 maximizable in $\mathcal{C}_{(\mathcal{F},\alpha)}$. If we have $v_\alpha = \mathsf{m}(q_0) = 0$, the state q_0 is maximizable in any case. Now, assume that $v_\alpha = \mathsf{m}(q_0) > 0$. Recall the construction of the previous section, specifically the set of secure states w.r.t. a set of bad states (Definition 24). Initially, $\mathsf{Bad}_0 = \emptyset$, so we want for the state q_0 to be in $\mathsf{Sec}(\emptyset)$, i.e. we want (and need) an efficient strategy in the state q_0 where the set of good states Gd is the target $\mathsf{Gd} = \{\top\}$ and the set of bad states is empty. In that case, the set of efficient strategies coincide with the set of progressive strategies. Thus, q_0 is maximizable if and only if $\mathsf{Prog}_{q_0}(\{\top\}) \neq \emptyset$. We assume for simplicity that $\alpha(y), \alpha(z) > 0$, hence the set Nature states Gd_D with a non-zero probability to reach $\top$ is $\{g(y), g(z)\} \subseteq \mathsf{D}$. By definition of Prog (Definition 21), $\mathsf{Prog}_{q_0}(\{\top\}) \neq \emptyset$ amounts to have an optimal strategy σ_A in $\mathcal{F}^{\mu_\mathsf{m}}_{q_0}$ such that, for all $b \in B_{\sigma_\mathsf{A}} : \delta(q_0, \mathsf{Supp}(\sigma_\mathsf{A}), b) \cap \{g(y), g(z)\} \neq \emptyset$ or, equivalently, $\delta(q_0, \mathsf{Supp}(\sigma_\mathsf{A}), b) \not\subseteq \{g(x)\}$. In terms of $\mathcal{F}$ and α, the state q_0 is maximizable if and only if there is an optimal strategy σ_A in $\mathcal{F}^{\tilde{\alpha}}$ such that, for all $b \in B_{\sigma_\mathsf{A}} : \varrho(\mathsf{Supp}(\sigma_\mathsf{A}), b) \not\subseteq \{x\} = E$ if the partial valuation α is defined as $\alpha : \mathsf{O} \setminus E \to [0, 1]$ for $\mathsf{O} = \{x, y, z\}$ and $E = \{x\}$.

This suggests the definition below of *reach-maximizable game form* w.r.t. a partial valuation.

▶ **Definition 32** (Reach-maximizable game forms w.r.t. a partial valuation)**.** *Consider a game form $\mathcal{F}$ and a partial valuation of the outcomes $\alpha : \mathsf{O} \setminus E \to [0, 1]$. The game form $\mathcal{F}$ is* reach-maximizable *w.r.t. the partial valuation α if $v_\alpha = 0$ or there exists an optimal strategy $\sigma_\mathsf{A} \in \mathsf{Opt}_\mathsf{A}(\mathcal{F}^{\tilde{\alpha}})$ such that for all $b \in B_{\sigma_\mathsf{A}}$, we have $\varrho(\mathsf{Supp}(\sigma_\mathsf{A}), b) \not\subseteq E$. Such strategies are said to be* reach-maximizing *w.r.t. α.*

This definition was chosen to ensure the lemma below.

▶ **Lemma 33.** *Consider a game form $\mathcal{F}$ and a partial valuation of the outcomes $\alpha : \mathsf{O} \setminus E \to [0,1]$. The initial state (and thus all states) in the three-state reachability game $\mathcal{C}_{(\mathcal{F},\alpha)}$ is maximizable if and only if the game form $\mathcal{F}$ is reach-maximizable w.r.t. the partial valuation α.*

The definition of reach-maximizable game form is then obtained via a universal quantification over the partial valuations considered.

▶ **Definition 34** (Reach-maximizable game form)**.** *Consider a game form $\mathcal{F} = \langle \mathsf{St_A}, \mathsf{St_B}, \mathsf{O}, \varrho \rangle$. It is a* reach-maximizable *(RM for short) game form if it is reach-maximizable w.r.t. all partial valuations $\alpha : \mathsf{O} \setminus E \to [0,1]$.*

Lemma 33 gives that RM game forms behave properly when used individually, such as in three-state reachability games. Let us now look at how such game forms behave collectively, that is we consider concurrent reachability games where all local interactions are RM. In fact, in such a setting, all states are maximizable. This is stated in the lemma below.

▶ **Lemma 35.** *Consider a concurrent reachability game $\langle \mathcal{C}, \top \rangle$ and assume that all local interactions are RM game forms. Then, all states are maximizable: $Q = \mathsf{MaxQ_A}$.*

Proof Sketch. We show that $Q = \mathsf{MaxQ_A}$ by showing that $\mathsf{Bad} = \emptyset$, which is equivalent since, by Theorem 28, we have $\mathsf{Bad} = \mathsf{SubMaxQ_A} = Q \setminus \mathsf{MaxQ_A}$. That is, we consider the iterative construction of the set of sub-maximizable states of the previous section and we show that $\mathsf{Bad}_1 = Q \setminus (\mathsf{Sec}(\mathsf{Bad}_0)) = \emptyset = \mathsf{Bad}_0$ (see Definition 26), which induces that $\mathsf{Bad} = \emptyset$. Let us assume towards a contradiction that $Q \setminus (\mathsf{Sec}_n(\emptyset) \cup \mathsf{m}^{-1}[0]) \neq \emptyset$ for $n = |Q|$. Since $\mathsf{Risk}_q(\emptyset) = \emptyset$ for all $q \in q$, any efficient strategy in a state q w.r.t. to the sets $\mathsf{Sec}_n(\emptyset)$ and $\emptyset$ is in fact a progressive strategy w.r.t. the set $\mathsf{Sec}_n(\emptyset)$. Hence, the goal is to exhibit such a progressive strategy in a state $q \in Q \setminus \mathsf{Sec}(\emptyset)$, thus showing a contradiction with the fact that $q \notin \mathsf{Sec}(\emptyset)$. We consider the states with the greatest value – w.r.t. m – as we can hope that they are the more likely to have progressive strategies. That is, for $x := \max_{q \in Q \setminus \mathsf{Sec}_n(\emptyset)} \mathsf{m}(q) > 0$ the maximum of m, we set $Q_x := \mathsf{m}^{-1}[x] \setminus \mathsf{Sec}_n(\emptyset) \neq \emptyset$ the set of states realizing that maximum. We want to use the assumption that all local interactions are RM. That is, we need to define a partial valuation on the outcomes of the local interactions, i.e. on Nature states. First, let us define its domain. We can find intuition in the example of the three-state reachability game in Figure 7: the outcome that is not valued by the partial valuation considered is the Nature state looping on the state q_0. Note that its value w.r.t. μ_{m} is the same as the value of the state q_0 w.r.t. m. In our case, we consider the set of Nature states D_x realizing this value x that cannot reach the set $\mathsf{Sec}_n(\emptyset)$, that is $\mathsf{D}_x := \mu_{\mathsf{m}}^{-1}[x] \setminus \mathsf{Sec}_n(\emptyset)_{\mathsf{D}}$. Then, we define the partial valuation of the Nature states $\alpha : \mathsf{D} \setminus \mathsf{D}_x \to [0,1]$ by $\alpha := \mu_{\mathsf{m}}|_{\mathsf{D} \setminus \mathsf{D}_x}$. Now, we can show that there exists a state $q \in Q_x$ such that $\tilde{\alpha} = \mu_{\mathsf{m}}$ in the game form $\mathcal{F}_q$. By maximality of x, we can prove that any local strategy σ_A in $\mathcal{F}_q$ that is reach-maximizing w.r.t. the partial valuation α of the outcomes of $\mathcal{F}_q$ is a progressive strategy w.r.t. $\mathsf{Sec}_n(\emptyset)$ in $\mathcal{F}_q$. Equivalently, σ_{A} is efficient w.r.t. $\mathsf{Sec}_n(\emptyset)$ and $\emptyset$. Hence the contradiction with the fact that $q \notin \mathsf{Sec}(\emptyset)$. ◀

Overall, we obtain the theorem below.

▶ **Theorem 36.** *For a set of game forms $\mathcal{G}$, all states in all concurrent reachability games with local interactions in $\mathcal{G}$ are maximizable if and only if all game forms in $\mathcal{G}$ are RM.*

Deciding if game forms are RM. Consider the following decision problem RMGF: given a game form, decide if it is a RM game form. We proved Theorem 30 by showing that the fact that a state is maximizable in a concurrent reachability game can be encoded in the theory of the reals (FO-$\mathbb{R}$). Since Lemma 33 ensures that a game form $\mathcal{F}$ is RM w.r.t. a partial valuation α if and only if the initial state in the three-state reachability game $\mathcal{C}_{(\mathcal{F},\alpha)}$ is maximizable, it follows that, via a universal quantification over partial valuations, the fact that a game form is RM can be encoded in the theory of the reals. Note that it can also be encoded directly from the definition of RM game form. We obtain the theorem below.

▶ **Proposition 37.** *The problem* RMGF *is decidable.*

Determined game forms and RM game forms. In [2], the authors have studied a problem similar to the one we considered in this section: determining the game forms ensuring that, when used as local interaction in a concurrent game (with an arbitrary Borel winning condition), the game is determined (i.e. either of the players has a winning strategy). The authors have shown that these game forms exactly correspond to *determined* game forms. These roughly correspond to game forms where, for all subsets of outcomes $E \subseteq \mathsf{O}$, there is either of line of outcomes in E or a column of outcomes in $\mathsf{O} \setminus E$, as formally defined below.

▶ **Definition 38** (Determined game forms)**.** *Consider a game form* $\mathcal{F} = \langle \mathsf{St_A}, \mathsf{St_B}, \mathsf{O}, \varrho \rangle$*. It is* determined *if, for all subsets of outcomes* $E \subseteq \mathsf{O}$*, either there exists some* $a \in \mathsf{St_A}$ *such that* $\varrho(a, \mathsf{St_B}) \subseteq E$ *or there exists some* $b \in \mathsf{St_B}$ *such that* $\varrho(\mathsf{St_A}, b) \subseteq \mathsf{O} \setminus E$*.*

In fact, they proved an equivalence between turn-based games and concurrent games using determined game forms as local interactions, which holds also when the game is stochastic. In fact, positional optimal strategies exists for both players in turn-based reachability games [4], it is also the case in concurrent reachability games with determined local interactions. This result, combined with Theorem 36 gives immediately that determined game forms are RM. Interestingly, determined game forms can also be characterized with the least fixed point operator as in the proposition below.

▶ **Proposition 39.** *A game form* $\mathcal{F}$ *is* determined *if and only if, for all partial valuations* $\alpha : \mathsf{O} \setminus E \to [0, 1]$ *of the outcomes, we have* $v_\alpha = f_\alpha^{\mathcal{F}}(0)$*. In particular, this implies that all determined game forms are RM.*

8 Future Work

In this paper we give a double-fixed-point procedure to compute maximizable and sub-maximizable states in a stochastic concurrent reachability (finite) game. Our procedure yields *de facto* positional witnesses for the strategies. As further natural work, we seek studying more general objectives. It is however interesting to notice that, as mentioned in the introduction, it will not be so easy since even Büchi games do not enjoy positional almost optimal strategies [7, Theorem 2].

We also plan to better grasp RM game forms, and understand what are RM game forms for the two players, or analyze the complexity of the RMGF problem.

References

1 Roderick Bloem, Krishnendu Chatterjee, and Barbara Jobstmann. *Handbook of Model Checking*, chapter Graph games and reactive synthesis, pages 921–962. Springer, 2018.
2 Benjamin Bordais, Patricia Bouyer, and Stéphane Le Roux. From local to global determinacy in concurrent graph games. Technical Report abs/2107.04081, CoRR, 2021. arXiv:2107.04081.
3 Benjamin Bordais, Patricia Bouyer, and Stéphane Le Roux. Optimal strategies in concurrent reachability games. *CoRR*, abs/2110.14724, 2021. arXiv:2110.14724.

4 Krishnendu Chatterjee, Marcin Jurdziński, and Thomas A. Henzinger. Quantitative stochastic parity games. In *Proc. of 15th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'04)*, pages 121–130. SIAM, 2004.

5 Luca de Alfaro. *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University, 1997.

6 Luca de Alfaro, Thomas Henzinger, and Orna Kupferman. Concurrent reachability games. *Theoretical Computer Science*, 386(3):188–217, 2007.

7 Luca de Alfaro and Rupak Majumdar. Quantitative solution of omega-regular games. *Journal of Computer and System Sciences*, 68:374–397, 2004.

8 Hugh Everett. Recursive games. *Annals of Mathematics Studies – Contributions to the Theory of Games*, 3:67–78, 1957.

9 Jerzy Filar and Koos Vrieze. *Competitive Markov decision processes*. Springer Science & Business Media, 2012.

10 Stefan Kiefer, Richard Mayr, Mahsa Shirmohammadi, and Patrick Totzke. Strategy complexity of parity objectives in countable mdps. In *Proc. 31st International Conference on Concurrency Theory (CONCUR'20)*, volume 171 of *LIPIcs*, pages 39:1–39:17. Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CONCUR.2020.39.

11 Elon Kohlberg. Repeated games with absorbing states. *The Annals of Statistics*, pages 724–738, 1974.

12 Antonín Kučera. *Lectures in Game Theory for Computer Scientists*, chapter Turn-Based Stochastic Games, pages 146–184. Cambridge University Press, 2011.

13 Donald A. Martin. The determinacy of blackwell games. *The Journal of Symbolic Logic*, 63(4):1565–1581, 1998.

14 Annabelle McIver and Carroll Morgan. Games, probability and the quantitative μ-calculus $qm\mu$. In *Proc. 9th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR'02)*, volume 2514 of *Lecture Notes in Computer Science*, pages 292–310. Springer, 2002.

15 James Renegar. On the computational complexity and geometry of the first-order theory of the reals. part iii: Quantifier elimination. *Journal of Symbolic Computation*, 13(3):329–352, 1992. doi:10.1016/S0747-7171(10)80005-7.

16 Wolfgang Thomas. Infinite games and verification. In *Proc. 14th International Conference on Computer Aided Verification (CAV'02)*, volume 2404 of *Lecture Notes in Computer Science*, pages 58–64. Springer, 2002. Invited Tutorial.

17 Moshe Y. Vardi. Automatic verification of probabilistic concurrent finite-state programs. In *Proc. 26th Annual Symposium on Foundations of Computer Science (FOCS'85)*, pages 327–338. IEEE Computer Society Press, 1985.

18 John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton Univ. Press, Princeton, 1944.

19 Wiesław Zielonka. Perfect-information stochastic parity games. In *Proc. 7th International Conference on Foundations of Software Science and Computation Structures (FoSSaCS'04)*, volume 2987 of *Lecture Notes in Computer Science*, pages 499–513. Springer, 2004.