Randomness for Free

Krishnendu Chatterjee[1], Laurent Doyen[2], Hugo Gimbert[3], and Thomas A. Henzinger[1]

[1] IST Austria (Institute of Science and Technology Austria)
[2] LSV, ENS Cachan & CNRS, France
[3] LaBri & CNRS, Bordeaux, France

Abstract. We consider two-player zero-sum games on graphs. These games can be classified on the basis of the information of the players and on the mode of interaction between them. On the basis of information the classification is as follows: (a) partial-observation (both players have partial view of the game); (b) one-sided complete-observation (one player has complete observation); and (c) complete-observation (both players have complete view of the game). On the basis of mode of interaction we have the following classification: (a) concurrent (both players interact simultaneously); and (b) turn-based (both players interact in turn). The two sources of randomness in these games are randomness in transition function and randomness in strategies. In general, randomized strategies are more powerful than deterministic strategies, and randomness in transitions gives more general classes of games. In this work we present a complete characterization for the classes of games where randomness is not helpful in: (a) the transition function (probabilistic transition can be simulated by deterministic transition); and (b) strategies (pure strategies are as powerful as randomized strategies). As consequence of our characterization we obtain new undecidability results for these games.

1 Introduction

Games on graphs. Games played on graphs provide the mathematical framework to analyze several important problems in computer science as well as mathematics. In particular, when the vertices and edges of a graph represent the states and transitions of a reactive system, then the synthesis problem (Church's problem) asks for the construction of a winning strategy in a game played on a graph [5,16,15,13]. Game-theoretic formulations have also proved useful for the verification [1], refinement [10], and compatibility checking [7] of reactive systems. Games played on graphs are dynamic games that proceed for an infinite number of rounds. In each round, the players choose moves; the moves, together with the current state, determine the successor state. An outcome of the game, called a *play*, consists of the infinite sequence of states that are visited.

Strategies and objectives. A strategy for a player is a recipe that describes how the player chooses a move to extend a play. Strategies can be classified as follows: *pure* strategies, which always deterministically choose a move to extend the play, vs. *randomized* strategies, which may choose at a state a probability distribution over the available moves. Objectives are generally Borel measurable functions [12]: the objective for a player is a Borel set B in the Cantor topology on S^ω (where S is the set of states), and the player satisfies the objective iff the outcome of the game is a member of B. In verification, objectives are usually *ω-regular languages*. The ω-regular languages generalize the classical regular languages to infinite strings; they occur in the low levels of the Borel hierarchy (they lie in $\Sigma_3 \cap \Pi_3$) and they form a robust and expressive language for determining payoffs for commonly used specifications.

Classification of games. Games played on graphs can be classified according to the knowledge of the players about the state of the game, and the way of choosing moves. Accordingly, there are (a) *partial-observation* games, where each player only has a partial or incomplete view about the state and the moves of the other player; (b) *one-sided complete-observation* games, where one player has partial knowledge and the other player has complete knowledge about the state and moves of the other player; and (c) *complete-observation* games, where each player has complete knowledge of the game. According to the way of choosing moves, the games on graphs can be classified into *turn-based* and *concurrent* games. In turn-based games, in any given round only one player can choose among multiple moves; effectively, the set of states can be partitioned into the states where it is player 1's turn to play, and the states where it is player 2's turn. In concurrent games, both players may have multiple moves available at each state, and the players choose their moves simultaneously and independently.

Sources of randomness. There are two sources of randomness in these games. First is the randomness in the transition function: given a current state and moves of the players, the transition function defines a probability distribution over the successor states. The second source of randomness is the randomness in strategies (when the players play randomized strategies). In this work we study when randomness can be obtained for *free*; i.e., we study in which classes of games the probabilistic transition function can be simulated by deterministic transition function, and the classes of games where pure strategies are as powerful as randomized strategies.

Motivation. The motivation to study this problem is as follows: (a) if for a class of games it can be shown that randomness is free for transitions, then all future works related to analysis of computational complexity, strategy complexity, and algorithmic solutions can focus on the simpler class with deterministic transitions (the randomness in transition may be essential for modeling appropriate stochastic reactive systems, but the analysis can focus on the deterministic subclass); (b) if for a class of games it can be shown that randomness is free for strategies, then all future works related to correctness results can focus on the simpler class of deterministic strategies, and the results would follow for the more general class of randomized strategies; and (c) the characterization of randomness for free will allow hardness results obtained for the more general class of games (such as games with randomness in transitions) to be carried over to simpler class of games (such as games with deterministic transitions).

Our contribution. Our contributions are as follows:

1. *Randomness for free in transitions.* We show that randomness in the transition function can be obtained for free for complete-observation concurrent games (and any class that subsumes complete-observation concurrent games) and for one-sided complete-observation turn-based games (and any class that subsumes this class). The reduction is polynomial for complete-observation concurrent games, and exponential for one-sided complete-observation turn-based games. It is known that for

complete-observation turn-based games, a probabilistic transition function cannot be simulated by deterministic transition function (see discussion at end of Section 3 for details), and thus we present a complete characterization when randomness can be obtained for free for the transition function.

2. *Randomness for free in strategies.* We show that randomness in strategies is free for complete-observation turn-based games, and for one-player partial-observation games (POMDPs). For all other classes of games randomized strategies are more powerful than pure strategies. It follows from a result of Martin [12] that for one-player complete-observation games with probabilistic transitions (MDPs) pure strategies are as powerful as randomized strategies. We present a generalization of this result to the case of one-player partial-observation games with probabilistic transitions (POMDPs). Our proof is totally different from Martin's proof and based on a new derandomization technique of randomized strategies.
3. *New undecidability results.* As a consequence of our characterization of randomness for free, we obtain new undecidability results. In particular, using our results and results of Baier et al. [2] we show for one-sided complete-observation deterministic games, the problem of almost-sure winning for coBüchi objectives and positive winning for Büchi objectives are undecidable. Thus we obtain the first undecidability result for qualitative analysis (almost-sure and positive winning) of one-sided complete-observation deterministic games with ω-regular objectives.

2 Definitions

In this section we present the definition of concurrent games of partial information and their subclasses, and related notions of strategies and objectives. Our model of game is essentially the same as in [9] and is equivalent to the model of stochastic games with signals [14,3]. A *probability distribution* on a finite set A is a function $\kappa : A \to [0,1]$ such that $\sum_{a \in A} \kappa(a) = 1$. We denote by $\mathcal{D}(A)$ the set of probability distributions on A.

Concurrent games of partial observation. A *concurrent game of partial observation* (or simply a *game*) is a tuple $G = \langle S, A_1, A_2, \delta, \mathcal{O}_1, \mathcal{O}_2 \rangle$ with the following components:

1. *(State space).* S is a finite set of states;
2. *(Actions).* A_i ($i = 1, 2$) is a finite set of actions for Player i;
3. *(Probabilistic transition function).* $\delta : S \times A_1 \times A_2 \to \mathcal{D}(S)$ is a concurrent probabilistic transition function that given a current state s, actions a_1 and a_2 for both players gives the transition probability $\delta(s, a_1, a_2)(s')$ to the next state s';
4. *(Observations).* $\mathcal{O}_i \subseteq 2^S$ ($i = 1, 2$) is a finite set of observations for Player i that partition the state space S. These partitions uniquely define functions $\mathsf{obs}_i : S \to \mathcal{O}_i$ ($i = 1, 2$) that map each state to its observation such that $s \in \mathsf{obs}_i(s)$ for all $s \in S$.

Special cases. We consider the following special cases of partial observation concurrent games, obtained either by restrictions in the observations, the mode of selection of moves, the type of transition function, or the number of players:

– *(Observation restriction).* The games with *one-sided complete-observation* are the special case of games where $\mathcal{O}_1 = \{\{s\} \mid s \in S\}$ (i.e., Player 1 has complete observation) or $\mathcal{O}_2 = \{\{s\} \mid s \in S\}$ (Player 2 has complete observation). The *games of complete-observation* are the special case of games where $\mathcal{O}_1 = \mathcal{O}_2 = \{\{s\} \mid s \in S\}$, i.e., every state is visible to each player and hence both players have complete observation. If a player has complete observation we omit the corresponding observation sets from the description of the game.
– *(Mode of interaction restriction).* A *turn-based state* is a state s such that either (i) $\delta(s, a, b) = \delta(s, a, b')$ for all $a \in A_1$ and all $b, b' \in A_2$ (i.e, the action of Player 1 determines the transition function and hence it can be interpreted as Player 1's turn to play), we refer to s as a Player-1 state, and we use the notation $\delta(s, a, -)$; or (ii) $\delta(s, a, b) = \delta(s, a', b)$ for all $a, a' \in A_1$ and all $b \in A_2$. We refer to s as a Player-2 state, and we use the notation $\delta(s, -, b)$. A state s which is both a Player-1 state and a Player-2 state is called a *probabilistic state* (i.e., the transition function is independent of the actions of the players). We write the $\delta(s, -, -)$ to denote the transition function in s. The *turn-based games* are the special case of games where all states are turn-based.
– *(Transition function restriction).* The *deterministic games* are the special case of games where for all states $s \in S$ and actions $a \in A_1$ and $b \in A_2$, there exists a state $s' \in S$ such that $\delta(s, a, b)(s') = 1$. We refer to such states s as deterministic states. For deterministic games, it is often convenient to assume that $\delta : S \times A_1 \times A_2 \to S$.
– *(Player restriction).* The *1½-player games*, also called *partially observable Markov decision processes* (or POMDP), are the special case of games where A_1 or A_2 is a singleton. Note that 1½-player games are turn-based. Games without player restriction are sometimes called 2½-player games.

The 1½-player games of complete-observation are Markov decision processes (or MDP), and 1½-player deterministic games can be viewed as graphs (and are often called one-player games).

Classes of game graphs. We will use the following abbreviations: we will use Pa for partial observation, Os for one-sided complete-observation, Co for complete-observation, C for concurrent, and T for turn-based. For example, CoC will denote complete-observation concurrent games, and OsT will denote one-sided complete-observation turn-based games. For $\mathcal{C} \in \{\mathsf{Pa}, \mathsf{Os}, \mathsf{Co}\} \times \{\mathsf{C}, \mathsf{T}\}$, we denote by $\mathcal{G}_{\mathcal{C}}$ the set of all $\mathcal{C}$ games. Note that the following strict inclusion: partial observation (Pa) is more general than one-sided complete-observation (Os) and Os is more general than complete-observation (Co), and concurrent (C) is more general than turn-based (T). We will denote by $\mathcal{G}_D$ the set of all games with deterministic transition function.

Plays. In a game structure, in each turn, Player 1 chooses an action $a \in A_1$, Player 2 chooses an action in $b \in A_2$, and the successor of the current state s is chosen according to the probabilistic transition function $\delta(s, a, b)$. A *play* in G is an infinite sequence of states $\rho = s_0 s_1 \dots$ such that for all $i \geq 0$, there exists $a_i \in A_1$ and $b_i \in A_2$ with $\delta(s_i, a_i, b_i, s_{i+1}) > 0$. The *prefix up to* s_n of the play ρ is denoted by $\rho(n)$, its *length* is $|\rho(n)| = n + 1$ and its *last element* is $\mathsf{Last}(\rho(n)) = s_n$. The set of plays in G is denoted $\mathsf{Plays}(G)$, and the set of corresponding finite prefixes is denoted $\mathsf{Prefs}(G)$.

The *observation sequence* of ρ for player i ($i = 1, 2$) is the unique infinite sequence $\mathsf{obs}_i(\rho) = o_0 o_1 \ldots \in O_i^\omega$ such that $s_j \in o_j$ for all $j \geq 0$.

Strategies. A *pure strategy* in G for Player 1 is a function $\sigma : \mathsf{Prefs}(G) \to A_1$. A *randomized strategy* in G for Player 1 is a function $\sigma : \mathsf{Prefs}(G) \to \mathcal{D}(A_1)$. A (pure or randomized) strategy σ for Player 1 is *observation-based* if for all prefixes $\rho, \rho' \in \mathsf{Prefs}(G)$, if $\mathsf{obs}_1(\rho) = \mathsf{obs}_1(\rho')$, then $\sigma(\rho) = \sigma(\rho')$. We omit analogous definitions of strategies for Player 2. We denote by Σ_G, Σ_G^O, Σ_G^P, Π_G, Π_G^O and Π_G^P the set of all Player-1 strategies, the set of all observation-based Player-1 strategies, the set of all pure Player-1 strategies, the set of all Player-2 strategies in G, the set of all observation-based Player-2 strategies, and the set of all pure Player-2 strategies, respectively. Note that if Player 1 has complete observation, then $\Sigma_G^O = \Sigma_G$.

Objectives. An *objective* for Player 1 in G is a set $\phi \subseteq S^\omega$ of infinite sequences of states. A play $\rho \in \mathsf{Plays}(G)$ *satisfies* the objective ϕ, denoted $\rho \models \phi$, if $\rho \in \phi$. Objectives are generally Borel measurable: a Borel objective is a Borel set in the Cantor topology on S^ω [11]. We specifically consider ω-regular objectives specified as parity objectives (a canonical form to express all ω-regular objectives [17]). For a play $\rho = s_0 s_1 \ldots$ we denote by $\mathrm{Inf}(\rho)$ the set of states that occur infinitely often in ρ, that is, $\mathrm{Inf}(\rho) = \{s \mid s_j = s \text{ for infinitely many } j\text{'s}\}$. For $d \in \mathbb{N}$, let $p : S \to \{0, 1, \ldots, d\}$ be a *priority function*, which maps each state to a non-negative integer priority. The *parity* objective $\mathsf{Parity}(p)$ requires that the minimum priority that occurs infinitely often be even. Formally, $\mathsf{Parity}(p) = \{\rho \mid \min\{p(s) \mid s \in \mathrm{Inf}(\rho)\} \text{ is even}\}$. The Büchi and coBüchi objectives are the special cases of parity objectives with two priorities, $p : S \to \{0, 1\}$ and $p : S \to \{1, 2\}$ respectively. We say that an objective ϕ is *visible* for Player i if for all $\rho, \rho' \in S^\omega$, if $\rho \models \phi$ and $\mathsf{obs}_i(\rho) = \mathsf{obs}_i(\rho')$, then $\rho' \models \phi$. For example if the priority function maps observations to priorities (i.e., $p : \mathcal{O}_i \to \{0, 1, \ldots, d\}$), then the parity objective is visible for Player i.

Almost-sure winning, positive winning and value function. An *event* is a measurable set of plays, and given strategies σ and π for the two players, the probabilities of events are uniquely defined [18]. For a Borel objective ϕ, we denote by $\mathrm{Pr}_s^{\sigma,\pi}(\phi)$ the probability that ϕ is satisfied by the play obtained from the starting state s when the strategies σ and π are used. Given a game structure G and a state s, an observation-based strategy σ for Player 1 is *almost-sure winning (almost winning in short)* (resp. *positive winning*) for the objective ϕ from s if for all observation-based randomized strategies π for Player 2, we have $\mathrm{Pr}_s^{\sigma,\pi}(\phi) = 1$ (resp. $\mathrm{Pr}_s^{\sigma,\pi}(\phi) > 0$). The *value function* $\langle\!\langle 1 \rangle\!\rangle_{val}^G : S \to \mathbb{R}$ for Player 1 and objective ϕ assigns to every state the maximal probability with which Player 1 can guarantee the satisfaction of ϕ with an observation-based strategy, against all observation-based strategies for Player 2. Formally we have

$$\langle\!\langle 1 \rangle\!\rangle_{val}^G(\phi)(s) = \sup_{\sigma \in \Sigma_G^O} \inf_{\pi \in \Pi_G^O} \mathrm{Pr}_s^{\sigma,\pi}(\phi).$$

For $\varepsilon \geq 0$, an observation-based strategy is ε*-optimal* for ϕ from s if we have $\inf_{\pi \in \Pi_G^O} \mathrm{Pr}_s^{\sigma,\pi}(\phi) \geq \langle\!\langle 1 \rangle\!\rangle_{val}^G(\phi)(s) - \varepsilon$. An *optimal* strategy is a 0-optimal strategy.

Example 1. Consider the game with one-sided complete observation (Player 2 has complete information) shown in Fig. 1. Consider the Büchi objective defined by the state

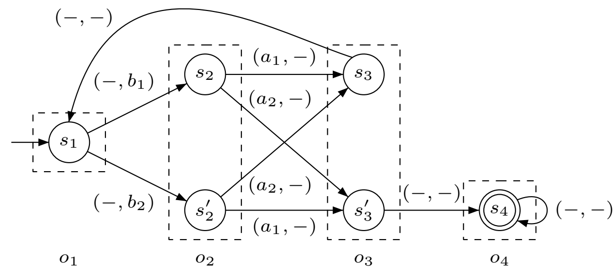

Fig. 1. A game with one-sided complete observation.

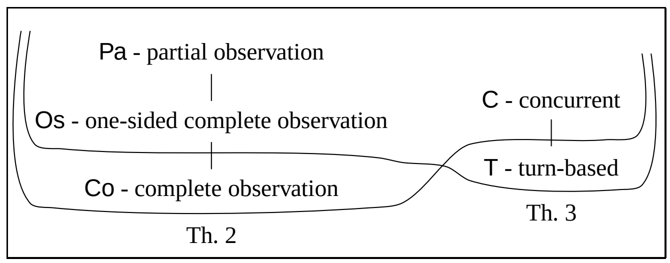

Fig. 2. The various classes of game graphs. The curves materialize the classes for which randomness is for free in transition relation (Theorem 2 and Theorem 3). For $2\,^1\!/_2$-player games, randomness is not free only in complete-observation turn-based games.

s_4 (i.e., state s_4 has priority 0 and other states have priority 1). Because Player 1 has partial observation (given by the partition $\mathcal{O}_i = \{\{s_1\}, \{s_2, s'_2\}, \{s_3, s'_3\}, \{s_4\}\}$), she cannot distinguish between s_2 and s'_2 and therefore has to play the same actions with same probabilities in s_2 and s'_2 (while it would be easy to win by playing a_2 in s_2 and a_1 in s'_2, this is not possible). In fact, Player 1 cannot win using a pure observation-based strategy. However, playing a_1 and a_2 uniformly at random in all states is almost-sure winning. Every time the game visits observation o_2, for any strategy of Player 2, the game visits s_3 and s'_3 with probability $\frac{1}{2}$, and hence also reaches s_4 with probability $\frac{1}{2}$. It follows that against all Player 2 strategies the play eventually reaches s_4 with probability 1, and then stays there.

3 Randomness for Free in Transition Function

In this section we present a precise characterization of the classes of games where the randomness in transition function can be obtained for *free*: in other words, we present the precise characterization of classes of games with probabilistic transition function that can be reduced to the corresponding class with deterministic transition function. We present our results as three reductions: (a) the first reduction allows us to separate probability from the mode of interaction; (b) the second reduction shows how to simulate probability in transition function with CoC (complete-observation concurrent) deterministic transition; and (c) the final reduction shows how to simulate probability in transition with OsT(one-sided complete-observation turn-based) deterministic transition. All our reductions are *local*: they consist of a gadget construction and replacement locally at every state. Our reductions preserve values, existence of ε-optimal strategies for $\varepsilon \geq 0$, and also existence of almost-sure and positive winning strategies. A visual overview is given in Fig. 2.

3.1 Separation of probability and interaction

A concurrent probabilistic game of partial observation G satisfies the *interaction separation* condition if the following restrictions are satisfied (see also Fig. 4): the state space S can be partitioned into (S_A, S_P) such that (1) $\delta : S_A \times A_1 \times A_2 \to S_P$, and (2) $\delta : S_P \times A_1 \times A_2 \to \mathcal{D}(S_A)$ such that for all $s \in S_P$ and all $s' \in S_A$, and for all a_1, a_2, a'_1, a'_2 we have $\delta(s, a_1, a_2)(s') = \delta(s, a'_1, a'_2)(s') = \delta(s, -, -)(s')$. In other words, the choice of actions (or the interaction) of the players takes place at states in S_A and actions determine a unique successor state in S_P, and the transition function at S_P is probabilistic and independent of the choice of the players. In this section, we reduce a class of games to the corresponding class satisfying interaction separation.

Reduction to interaction separation. Let $G = \langle S, A_1, A_2, \delta, \mathcal{O}_1, \mathcal{O}_2 \rangle$ be a concurrent game of partial observation with an objective ϕ. We obtain a concurrent game of partial observation $\overline{G} = \langle S_A \cup S_P, A_1, A_2, \overline{\delta}, \overline{\mathcal{O}}_1, \overline{\mathcal{O}}_2 \rangle$ where $S_A = S$, $S_P = S \times A_1 \times A_2$, and:

– *Observation.* For $i \in \{1, 2\}$, if $\mathcal{O}_i = \{\{s\} \mid s \in S\}$, then $\overline{\mathcal{O}}_i = \{\{s'\} \mid s' \in S_A \cup S_P\}$; otherwise $\overline{\mathcal{O}}_i$ contains the observation $o \cup \{(s, a_1, a_2) \mid s \in o\}$ for each $o \in \mathcal{O}_i$.
– *Transition function.* The transition function is as follows:
  1. We have the following three cases: (a) if s is a Player 1 turn-based state, then pick an action a_2^* and for all a_2 let $\overline{\delta}(s, a_1, a_2) = (s, a_1, a_2^*)$; (b) if s is a Player 2 turn-based state, then pick an action a_1^* and for all a_1 let $\overline{\delta}(s, a_1, a_2) = (s, a_1^*, a_2)$; and (c) otherwise, $\overline{\delta}(s, a_1, a_2) = (s, a_1, a_2)$;
  2. for all $(s, a_1, a_2) \in S_P$ we have $\overline{\delta}((s, a_1, a_2), -, -)(s') = \delta(s, a_1, a_2)(s')$.

  Thus the states in S are S_A where the interaction takes places, and the states in $S \times A_1 \times A_2$ are the purely probabilistic states S_P.
– *Objective mapping.* Given the objective ϕ in G we obtain the objective $\overline{\phi} = \{\langle s_0 s'_0 s_1 s'_1 \ldots \rangle \mid \langle s_0 s_1 \ldots \rangle \in \phi\}$ in $\overline{G}$.

It is easy to map observation-based strategies of the game G to observation-based strategies in $\overline{G}$ and vice-versa that preserves satisfaction of ϕ and $\overline{\phi}$ in G and $\overline{G}$, respectively. Let us refer to the above reduction as $\mathsf{Reduction}$: i.e., $\mathsf{Reduction}(G, \phi) = (\overline{G}, \overline{\phi})$. Then we have the following theorem.

Theorem 1. *Let G be a concurrent game of partial observation with an objective ϕ, and let $(\overline{G}, \overline{\phi}) = \mathsf{Reduction}(G, \phi)$. Then the following assertions hold:*

1. *The reduction* Reduction *is* restriction preserving*: if G is one-sided complete-observation, then so is $\overline{G}$; if G is complete-observation, then so is $\overline{G}$; if G is turn-based, then so is $\overline{G}$.*
2. *For all $s \in S$, there is an observation-based almost-sure (resp. positive) winning strategy for ϕ from s in G iff there is an observation-based almost-sure (resp. positive) winning strategy for $\overline{\phi}$ from s in $\overline{G}$.*
3. *The reduction is objective preserving: if ϕ is a parity objective, then so in $\overline{\phi}$; if ϕ is an objective in the k-the level of the Borel hierarchy, then so is $\overline{\phi}$.*
4. *For all $s \in S$ we have $\langle\langle 1\rangle\rangle_{val}^{G}(\phi)(s) = \langle\langle 1\rangle\rangle_{val}^{\overline{G}}(\overline{\phi})(s)$. For all $s \in S$ there is an observation-based optimal strategy for ϕ from s in G iff there is an observation-based optimal strategy for $\overline{\phi}$ from s in $\overline{G}$.*

Since the reduction is restriction preserving, we have a reduction that separates the interaction and probabilistic transition maintaining the restriction of observation and mode of interaction.

Uniform-n-ary concurrent probabilistic games. The class of *uniform-n-ary probabilistic games* are the special class of probabilistic games such that every state $s \in S_P$ has n successors and the transition probability to each successor is $\frac{1}{n}$. It follows from the results of [19] that every CoC probabilistic game with rational transition probabilities can be reduced in polynomial time to an equivalent polynomial size uniform-binary (i.e., $n = 2$) CoC probabilistic game for all parity objectives. The reduction is achieved by adding dummy states to simulate the probability, and the reduction extends to all objectives (in the reduced game we need to consider the objective whose projection in the original game gives the original objective).

In the case of partial information, the reduction to uniform-binary probabilistic games of [19] is not valid. To see this, consider Fig. 3 where two probabilistic states s_1, s_2 have the same observation (i.e., $\mathsf{obs}_1(s_1) = \mathsf{obs}_1(s_2)$) and the outgoing probabilities are $\langle\frac{1}{4}, \frac{3}{4}\rangle$ from s_1 and $\langle\frac{1}{3}, \frac{2}{3}\rangle$ from s_2. The corresponding uniform-binary game (given in Fig. 3) is not equivalent to the original game because the number of steps needed to simulate the probabilities is not always the same from s_1 and from s_2. From s_1 two steps are always sufficient, while from s_2 more than two steps may be necessary (with probability $\frac{1}{4}$). Therefore with probability $\frac{1}{4}$, Player 1 observing more than 2 steps would infer that the game was for sure in s_2, thus artificially improving his knowledge and increasing his value function.

Therefore in the case of partial observation, we can only reduce a probabilistic game G to a uniform-n-ary probabilistic game with $n = 1/r$ where r is the greatest common divisor of all probabilities in the original game G (a rational r is a divisor of a rational p if $p = q \cdot r$ for some integer q). Note that the number $n = 1/r$ is an integer. We denote by $[n]$ the set $\{0, 1, \ldots, n-1\}$. For a probabilistic state $s \in S_P$, we define the n-tuple $\mathsf{Succ}(s) = \langle s_0', \ldots, s_{n-1}'\rangle$ in which each state $s' \in S$ occurs $n \cdot \delta(s, -, -)(s')$ times. Then, we can view the transition relation $\delta(s, -, -)$ as a function assigning the same probability $r = 1/n$ to each element of $\mathsf{Succ}(s)$ (and then adding up the probabilities of identical elements).

Note that the above reduction is worst-case exponential (because so can be the least common multiple of all probability denominators). This is necessary to have the property that all probabilistic states in the game have the same number of successors. We

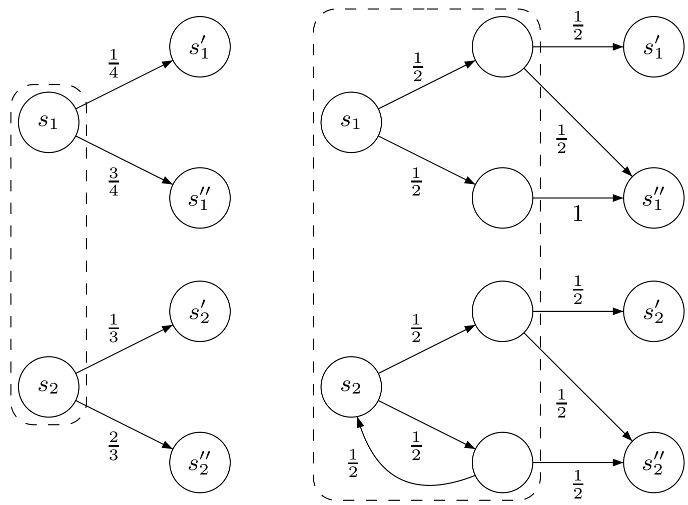

Fig. 3. An example showing why the uniform-binary reduction cannot be used with partial observation.

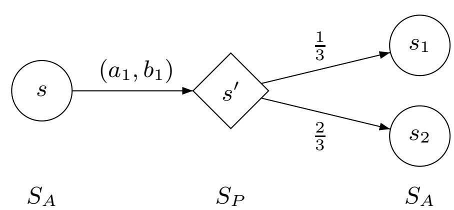

Fig. 4. Example of interaction separation for $\delta(s, a_1, b_1)(s_1) = \frac{1}{3}$ and $\delta(s, a_1, b_1)(s_2) = \frac{2}{3}$.

will see that this property is crucial because it determines the number of actions available to Player 1 in the reductions presented in Section 3.2 and 3.3, and the number of available actions should not differ in states that have the same observation.

3.2 Simulation of probability with complete-observation concurrent determinism

In this section, we show that probabilistic states can be simulated by CoC deterministic gadgets (and hence also by OsC and PaC deterministic gadgets). By Theorem 1, we focus on games that satisfy interaction separation.

Theorem 2. *Let $a \in \{\mathsf{Pa}, \mathsf{Os}, \mathsf{Co}\}$ and $b \in \{\mathsf{C}, \mathsf{T}\}$, and let $\mathcal{C} = ab$ and $\mathcal{C}' = a\mathsf{C}$. Let G be a game in $\mathcal{G}_\mathcal{C}$ with probabilistic transition function with rational probabilities and an objective ϕ. A game $\overline{G} \in \mathcal{G}_{\mathcal{C}'} \cap \mathcal{G}_D$ (in the class that subsumes $\mathcal{G}_\mathcal{C}$ with concurrent* *interaction) with deterministic transition function can be constructed in (a) polynomial time if $a =$ Co, and (b) in exponential time if $a =$ Pa or Os, with an objective $\overline{\phi}$ such that the state space of G is a subset of the state space of $\overline{G}$ and the following assertions hold.*

1. *For all $s \in S$ there is an observation-based almost-sure (resp. positive) winning strategy from s for ϕ in G iff there is an observation-based almost-sure (resp. positive) winning strategy for $\overline{\phi}$ from s in $\overline{G}$.*
2. *For all $s \in S$ we have $\langle\langle 1 \rangle\rangle_{val}^{G}(\phi)(s) = \langle\langle 1 \rangle\rangle_{val}^{\overline{G}}(\overline{\phi})(s)$. For all $s \in S$ there is an observation-based optimal strategy for ϕ from s in G iff there is an observation-based optimal strategy for $\overline{\phi}$ from s in $\overline{G}$.*

Proof. To prove the desired result we show how an uniform-n-ary probabilistic state can be simulated by a CoC deterministic gadget. For simplicity we present the details for the case when $n = 2$, and the gadget for the general case is given in the Appendix. Our reduction will be as follows: we consider a uniform-binary CoC probabilistic game such that there is only one probabilistic state, and reduce it to a CoC deterministic game. For uniform-binary CoC probabilistic games with multiple probabilistic states the reduction can be applied to each state one at a time and we would obtain the desired reduction from uniform-binary CoC probabilistic games to CoC deterministic games. Hence we prove the following claim.

Claim. Consider a uniform-binary CoC probabilistic game G with a single probabilistic state s^* with two successors s_1 and s_2. Consider the CoC deterministic game G' obtained from G by transforming the state s^* to a concurrent deterministic state as follows: the actions available for player 1 at s^* are a_1 and a_2 and the actions available for player 2 at s^* are b_1 and b_2; and the transition function is as follows: $\delta(s^*, a_1, b_1) = \delta(s^*, a_2, b_2) = s_1$ and $\delta(s^*, a_1, b_2) = \delta(s^*, a_2, b_1) = s_2$. Then for all objectives ϕ, the following assertions hold.

1. For all $s \in S$ there is an observation-based almost-sure (resp. positive) winning strategy from s for ϕ in G iff there is an observation-based almost-sure (resp. positive) winning strategy for ϕ from s in G'.
2. For all $s \in S$ we have $\langle\langle 1 \rangle\rangle_{val}^{G}(\phi)(s) = \langle\langle 1 \rangle\rangle_{val}^{G'}(\phi)(s)$. For all $s \in S$ there is an observation-based optimal strategy for ϕ from s in G iff there is an observation-based optimal strategy for ϕ from s in G'.

The reduction is illustrated in Figure 5. We prove the claim as follows. Let the value for the objective ϕ player 1 at a state s be $v(s)$ and $v'(s)$ in G and G', respectively, and let the value for player 2 be $w(s)$ and $w'(s)$ in G and G', respectively. By determinacy of CoC games [12] we have $w(s) = 1 - v(s)$ and $w'(s) = 1 - v'(s)$. We present two inequalities to complete the proof.

1. Consider a strategy π for player 2 in G and we construct a strategy π' for player 2 in G as follows: the strategy π' follows the strategy π for all histories other than when the current state is s^*; and if the current state is s^*, then strategy π' plays the actions b_1 and b_2 uniformly with probability $\frac{1}{2}$. Given the strategy π', if the current state is s^*, then for any probability distribution over a_1 and a_2, the successor states

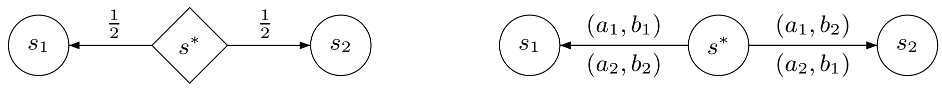

Fig. 5. The reduction of uniform-binary CoC probabilistic games.

are s_1 and s_2 with probability $\frac{1}{2}$ (i.e., it plays exactly the role of state s^* in G). It follows that the value for player 1 in G' is no more than the value in G, i.e., for all s we have $v'(s) \leq v(s)$.

2. Consider a strategy σ for player 1 in G and we construct a strategy σ' for player 1 in G' as follows: the strategy σ' follows the strategy σ for all histories other than when the current state is s^*, and if the current state is s^*, then the strategy σ' plays the actions a_1 and a_2 uniformly with probability $\frac{1}{2}$. Given the strategy σ', if the current state is s^*, then for any probability distribution over b_1 and b_2, the successor states are s_1 and s_2 with probability $\frac{1}{2}$ (i.e., it plays exactly the role of state s^* in G). It follows that the value for player 2 in G' is no more than the value in G, i.e., for all s we have $w'(s) \leq w(s)$.

It follows from above that $v(s) = v'(s)$ for all states s, and the desired result follows. Observe that the reduction also ensures that from an optimal strategy in G we can construct an optimal strategy in G' and vice-versa. Our proof shows how probabilistic states can be simulated by CoC deterministic states, and it follows that probabilistic states can be simulated by OsC deterministic states and PaC deterministic states. The result follows. ∎

3.3 Simulation of probability with one-sided complete-observation turn-based determinism

We show that probabilistic states can be simulated by OsT (one-sided complete-observation turn-based) states, and by Theorem 1 we consider games that satisfy interaction separation. The reduction is illustrated in Fig. 6: each probabilistic state s is transformed into a Player-2 state with n successor Player-1 states (where n is chosen such that the probabilities in s are integer multiples of $1/n$, here $n = 3$). Because all successors of s have the same observation, Player 1 has no advantage in playing after Player 2, and because by playing all actions uniformly at random each player can unilaterally decide to simulate the probabilistic state, the value and properties of strategies of the game are preserved.

Theorem 3. *Let $a \in \{\mathsf{Pa}, \mathsf{Os}, \mathsf{Co}\}$ and $b \in \{\mathsf{C}, \mathsf{T}\}$, and let $a' = a$ if $a \neq \mathsf{Co}$, and $a' = \mathsf{Os}$ otherwise. Let $\mathcal{C} = ab$ and $\mathcal{C}' = a'b$. Let G be a game in $\mathcal{G}_{\mathcal{C}}$ with probabilistic transition function with rational transition probabilities and an objective ϕ. A game $G' \in \mathcal{G}_{\mathcal{C}'} \cap \mathcal{G}_D$ (in the class that subsumes one-sided complete-observation turn-based games and the class $\mathcal{G}_{\mathcal{C}}$) with deterministic transition function can be constructed in exponential time with an objective ϕ' such that the state space of G is a subset of the state space of G' and the following assertions hold.*

1. *For all $s \in S$ there is an observation-based almost-sure (resp. positive) winning strategy from s for ϕ in G iff there is an observation-based almost-sure (resp. positive) winning strategy for ϕ' from s in G'.*
2. *For all $s \in S$ we have $\langle\langle 1 \rangle\rangle_{val}^{G}(\phi)(s) = \langle\langle 1 \rangle\rangle_{val}^{G'}(\phi')(s)$. For all $s \in S$ there is an observation-based optimal strategy for ϕ from s in G iff there is an observation-based optimal strategy for ϕ' from s in G'.*

Proof. First, we present the proof for $a \neq \mathsf{Co}$, assuming that Player 2 has complete observation. Let $G = \langle S_A \cup S_P, A_1, A_2, \delta, \mathcal{O}_1 \rangle$ and assume w.l.o.g. (according to Theorem 1) that G satisfies interaction separation (i.e., states in S_A are deterministic states, and S_P are probabilistic states) and G is uniform-n-ary, i.e. all probabilities are equal to $\frac{1}{n}$. For each probabilistic state $s \in S_P$, let $\mathsf{Succ}(s) = \langle s'_0, \dots, s'_{n-1} \rangle$ be the n-tuple of states such that $\delta(s, -, -)(s'_i) = \frac{1}{n}$ for each $1 \leq i \leq n$.

We present a reduction that replaces the probabilistic states in G by a gadget with Player-1 and Player-2 turn-based states. From G, we construct the one-sided complete-observation game G' where Player-2 has complete observation. A similar construction where Player-1 instead of Player-2 has complete observation is obtained symmetrically. The game $G' = \langle S', A'_1, A'_2, \delta', \mathcal{O}'_1 \rangle$ is defined as follows: $S' = S \cup (S \times [n]) \cup \{\mathsf{sink}\}$, $A'_1 = A_1 \cup [n]$, $A'_2 = A_2 \cup [n]$, $\mathcal{O}'_1 = \{o \cup \{(s,i) \mid s \in o\} \mid o \in \mathcal{O}_1\}$, and δ' is obtained from δ by applying the following transformation for each state $s \in S$:

1. if s is a deterministic state in G, then $\delta'(s,a,b) = \delta(s,a,b)$ for all $a \in A_1, b \in A_2$, and $\delta'(s,-,j) = \delta'(s,i,-) = \mathsf{sink}$ for all $i,j \in [n]$;
2. if s is a probabilistic state in G, then s is a Player-2 state in G' and for all $i,j \in [n]$ we define $\delta'(s,-,i) = (s,i)$ and $\delta'((s,i),j,-) = s'_k$ such that s'_k is the element in position k in $\mathsf{Succ}(s)$ with $k = i + j \mod n$ (and let $\delta'(s,-,b) = \delta'((s,i),a,-) = \delta'(\mathsf{sink},-,-) = \mathsf{sink}$ for all $a \in A_1, b \in A_2$).

Note that turn-based states in G remain turn-based in G' and the states $(s, \cdot)$ are Player-1 states with the same observation as s. A sequence of observation $o_1, \dots, o_m$ in G corresponds to the sequence $o_1, o_2, o_2, o_3, o_4, o_4, \dots, o_m$ in G' because deterministic and probabilistic states alternate in G, and in G', transitions from probabilistic states have intermediate states with duplicated observation. The objective ϕ' is defined as the set $\{o_1, o_2, o_2, o_3, o_4, o_4, \dots \mid o_1, o_2, \dots \in \phi\}$. Intuitively, each player in G' has the possibility to force faithful simulation of the probabilistic states of G by playing actions in $[n]$ uniformly at random. For instance, if Player 1 does so, then irrespective of the (possibly randomized) choice of Player 2 among the states $(s,1), \dots, (s,n)$, the states in $\mathsf{Succ}(s)$ are reached with probability $1/n$, as in G. And the same holds if Player 2 plays in $[n]$ uniformly at random, no matter what Player 1 does. Therefore, Player 1 can achieve the objective ϕ' in G' with the same probability as for ϕ in G, but not more.

The above reduction can be easily adapted to the case $a = \mathsf{Pa}$ of games with partial information for both players. ∎

Role of probabilistic transition in CoT **games and** POMDPs**.** We have already shown that for CoC games and OsT games, randomness in transition can be obtained for free. We complete the picture by showing that for CoT (complete-observation turn-based) games randomness in transition cannot be obtained for free. It follows from the result

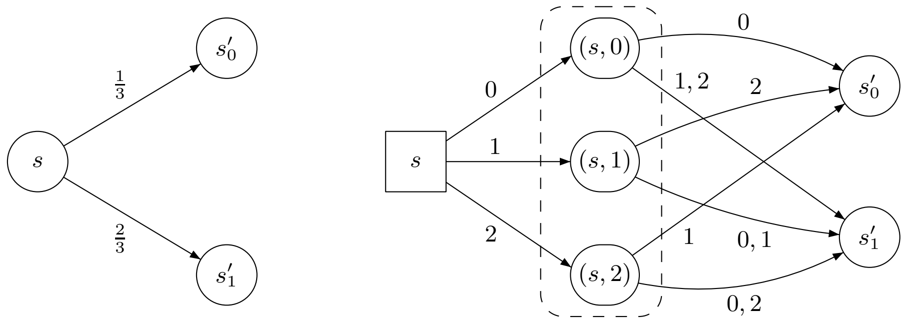

Fig. 6. For the probabilistic state s (on the left), we have $\mathsf{Succ}(s) = \langle s'_0, s'_1, s'_1 \rangle$ and $n = 3$ is the gcd of the probabilities denominators. Therefore, we apply the reduction of Theorem 3 to obtain the turn-based game on the right, where s is a Player-2 states.

	2 ½-player			1 ½-player	
	complete	one-sided	partial	MDP	POMDP
turn-based	not	free	free	not	not
concurrent	free	free	free	(NA)	(NA)

Table 1. When randomness is for free in the transition function. In particular, probabilities can be eliminated in all classes of 2-player games except complete-observation turn-based games.

of Martin [12] that for all CoT deterministic games and all objectives, the values are either 1 or 0; however, MDPs with reachability objectives can have values in the interval $[0, 1]$ (not value 0 and 1 only). Thus the result follows for CoT games. It also follows that "randomness in transitions" can be replaced by "randomness in strategies" is not true: in CoT deterministic games even with randomized strategies the values are either 1 or 0 [12]; whereas MDPs can have values in the interval $[0, 1]$. For POMDPs, we show in Theorem 5 that pure strategies are sufficient, and it follows that for POMDPs with deterministic transition function the values are 0 or 1, and since MDPs with reachability objectives can have values other than 0 and 1 it follows that randomness in transition cannot be obtained for free for POMDPs. The probabilistic transition also plays an important role in the complexity of solving games in case of CoT games: for example, CoT deterministic games with reachability objectives can be solved in linear time, but for probabilistic transition the problem lies in NP $\cap$ coNP and no polynomial time algorithm is known. In contrast, for CoC games we present a polynomial time reduction from probabilistic transition to deterministic transition. Table 1 summarizes our results characterizing the classes of games where randomness in transition can be obtained for free.

4 Randomness for Free in Strategies

It is known from the results of [8] that in CoC games randomized strategies are more powerful than pure strategies; for example, values achieved by pure strategies are lower than values achieved by randomized strategies and randomized almost-sure winning strategies may exist whereas no pure almost-sure winning strategy exists. Similar results also hold in the case of OsT games (see [6] for an example). By contrast we show that in one-player games, restricting the set of strategies to pure strategies does not decrease the value nor affect the existence of almost-sure and positive winning strategies. We first start with a lemma, then present a result that can be derived from Martin's theorem for Blackwell games [12], and finally present our results precisely in a theorem.

Lemma 1. *Let G be a* POMDP *with initial state s_* and an objective $\phi \subseteq S^\omega$. Then for every randomized observation-based strategy $\sigma \in \Sigma_O$ there exists a* pure *observation-based strategy $\sigma_P \in \Sigma_P \cap \Sigma_O$ such that:*

$$\mathrm{Pr}_{s_*}^{\sigma}(\phi) \leq \mathrm{Pr}_{s_*}^{\sigma_P}(\phi) \ . \tag{1}$$

Proof. Let $G = \langle S, A, \delta, \mathcal{O} \rangle$ a POMDP. Let $\sigma : \mathcal{O}^* \to \mathcal{D}(A)$ be a randomized observation-based strategy and fix $s_* \in S$ an initial state.

To simplify notations, we suppose that $A = \{0, 1\}$ contains only two actions, and that given a state $s \in S$ and an action $a \in \{0, 1\}$ there are only two possible successors $L(s, a) \in S$ and $R(s, a) \in S$ chosen with respective probabilities $\delta(s, a, L(s, a))$ and $\delta(s, a, R(s, a)) = 1 - \delta(s, a, L(s, a))$. The proof is for an arbitrary finite set of actions and more than two successors is essentially the same, with more complicated notations.

There is a natural way to "derandomize" the randomized strategy σ. Fix an infinite sequence $x = (x_n)_{n\in\mathbb{N}} \in [0, 1]^\omega$ and define the deterministic strategy σ_x as follows. For every $o_0, o_1, \ldots, o_n \in \mathcal{O}^*$,

$$\sigma_x(o_0, o_1, \ldots, o_n) = \begin{cases} 0 & \text{if } x_n \leq \sigma(o_0, o_1, \ldots, o_n)(0) \\ 1 & \text{otherwise.} \end{cases}$$

Intuitively, the sequence x fixes in advance the sequence of results of coin tosses used for playing with σ.

To prove the lemma, we show that $[0, 1]^\omega$ can be equipped with a probability measure ν such that the mapping $x \mapsto \mathrm{Pr}_{s_*}^{\sigma_x}(\phi)$ from $[0, 1]^\omega$ to $[0, 1]$ is measurable and:

$$\mathrm{Pr}_{s_*}^{\sigma}(\phi) = \int_{x\in[0,1]^\omega} \mathrm{Pr}_{s_*}^{\sigma_x}(\phi)\, d\nu(x) \ . \tag{2}$$

Suppose that (2) holds. Then there exists $x \in [0, 1]^\omega$ (actually many x's) such that $\mathrm{Pr}_{s_*}^{\sigma}(\phi) \leq \mathrm{Pr}_{s_*}^{\sigma_x}(\phi)$ and since strategy σ_x is deterministic, this proves the lemma.

To complete the proof of Lemma 1, it is thus enough to construct a probability measure ν on $[0, 1]^\omega$ such that (2) holds.

We start with the definition of the probability measure ν. The set $[0, 1]^\omega$ is equipped with the σ-field generated by *sequence-cylinders* which are defined as follows. For every finite sequence $x = x_0, x_1, \ldots, x_n \in [0, 1]^*$ the sequence-cylinder $\mathcal{O}(x)$ is the subset $[0, x_0] \times [0, x_1] \times \ldots \times [0, x_n] \times [0, 1]^\omega \subseteq [0, 1]^\omega$. According to Tulcea's theorem [4], there is a unique product probability measure ν on $[0,1]^\omega$ such that $\nu(\mathcal{O}(\epsilon)) = 1$ and for every sequence $x_0, \dots, x_n, x_{n+1}$ in $[0,1]$,

$$\nu(\mathcal{O}(x_0, \dots, x_n, x_{n+1})) = x_{n+1} \cdot \nu(\mathcal{O}(x_0, \dots, x_n)) \ .$$

Now that ν is defined, it remains to prove that the mapping $x \mapsto \Pr_{s_*}^{\sigma_x}(\phi)$ from $[0,1]^\omega$ to $[0,1]$ is measurable and that (2) holds. For that, we introduce the following mapping:

$$f_{s_*,\sigma} : [0,1]^\omega \times [0,1]^\omega \to (SA)^\omega,$$

that associates with every pair of sequences $((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}})$ the infinite history $h = s_0 a_1 s_1 a_2 \ldots \in (SA)^\omega$ defined recursively as follows. First $s_0 = s_*$, and for every $n \in \mathbb{N}$,

$$a_{n+1} = \begin{cases} 0 & \text{if } x_n \leq \sigma(\mathsf{obs}(s_0 s_1 \cdots s_n))(0), \\ 1 & \text{otherwise.} \end{cases}$$

$$s_{n+1} = \begin{cases} L(s_n, a_{n+1}) & \text{if } y_n \leq \delta(s_n, a_{n+1}, L(s_n, a_{n+1})), \\ R(s_n, a_{n+1}) & \text{otherwise.} \end{cases}$$

Intuitively, $(x_n)_{n\in\mathbb{N}}$ fixes in advance the coin tosses used by the strategy, while $(y_n)_{n\in\mathbb{N}}$ takes care of coin tosses used by the probabilistic transitions, and $f_{s_*,\sigma}$ produces the resulting description of the play. Thanks to the mapping $f_{s_*,\sigma}$, randomness related to the use of the randomized strategy σ is separated from randomness due to transitions of the game, which allows to represent the randomized strategy σ by mean of a probability measure over the set of deterministic strategies $\{\sigma_x \mid x \in [0,1]^\omega\}$.

We equip both sets $(SA)^\omega$ and $[0,1]^\omega \times [0,1]^\omega$ with σ-fields that make $f_{s_*,\sigma}$ measurable. First, $(SA)^\omega$ is equipped with the σ-field generated by cylinders, defined as follows. An action-cylinder is any subset $\mathcal{O}(h) \subseteq (SA)^\omega$ such that $\mathcal{O}(h) = h(SA)^\omega$ for some $h \in (SA)^*$. A state-cylinder is any subset $\mathcal{O}(h) \subseteq (SA)^\omega$ such that $\mathcal{O}(h) = h(AS)^\omega$ for some $h \in (SA)^* S$. The set of cylinders is the union of the sets of action-cylinders and state-cylinders. Second, $[0,1]^\omega \times [0,1]^\omega$ is equipped with the σ-field generated by products of sequence-cylinders. Checking that $f_{s_*,\sigma}$ is measurable is an elementary exercise.

Now we define two probability measures μ and μ' on $(SA)^\omega$ and prove that they coincide.

On one hand, the measurable mapping $f_{s_*,\sigma} : [0,1]^\omega \times [0,1]^\omega \to (SA)^\omega$ defines naturally a probability measure μ' on $(SA)^\omega$. Equip the set $[0,1]^\omega \times [0,1]^\omega$ with the product measure $\nu \times \nu$. Then for every measurable subset $B \subseteq (SA)^\omega$,

$$\mu'(B) = (\nu \times \nu)(f_{s_*,\sigma}^{-1}(B)) \ .$$

On the other hand, the strategy σ and the initial state s_* naturally define another probability measure μ on $(SA)^\omega$. According to Tulcea's theorem [4], there exists a unique product probability measure μ on $(SA)^\omega$ such that $\mu(\mathcal{O}(s_*)) = 1$, $\mu(\mathcal{O}(s)) = 0$ for $s \in S \setminus \{s_*\}$, and for $h = s_0 a_1 s_1 a_2 \cdots s_n \in (SA)^* S$ and $(a,t) \in A \times S$,

$$\begin{aligned} \mu(\mathcal{O}(ha)) &= \sigma(\mathsf{obs}(s_0 \cdots s_n))(a) \cdot \mu(\mathcal{O}(h)) \\ \mu(\mathcal{O}(hat)) &= \delta(s_n, a, t) \cdot \mu(\mathcal{O}(ha)) \ . \end{aligned}$$

We have defined $f_{s_*,\sigma}$ in such a way that μ and μ' coincide. To prove that μ and μ' coincide, it is enough to prove that μ and μ' coincide on the set of cylinders, that is for every cylinder $\mathcal{O}(h) \subseteq (SA)^\omega$,

$$\mu(\mathcal{O}(h)) = (\nu \times \nu)(f_{s_*,\sigma}^{-1}(\mathcal{O}(h))) \ . \tag{3}$$

For $h = s_*$ or $h = s \in S \setminus \{s_*\}$ then (3) is obvious. The general case goes by induction. Let $h = s_0 a_1 s_1 a_2 \cdots s_n \in (SA)^* S$ and $(a, t) \in A \times S$. Let $I = [0, 1]$. Let $I_a = [0, \sigma(h)(a)]$ if $a = 0$ and $I_a = [\sigma(h)(a), 1]$ if $a = 1$. Let $I_t = [0, \delta(s_n, a, t)]$ if $t = L(s_n, a)$ and $I_t = [\delta(s_n, a, t), 1]$ if $t = R(s_n, a)$. Then:

$$\begin{aligned}
\mu(\mathcal{O}(ha) \mid \mathcal{O}(h)) &= \sigma(h)(a) \\
&= (\nu \times \nu)((I \times I)^n (I_a \times I)(I \times I)^\omega) \\
&= (\nu \times \nu)(f_{s_*,\sigma}^{-1}(\mathcal{O}(ha)) \mid f_{s_*,\sigma}^{-1}(\mathcal{O}(h))) \\
\mu(\mathcal{O}(hat) \mid \mathcal{O}(ha)) &= \delta(s_n, a, t) \\
&= (\nu \times \nu)((I \times I)^n (I \times I_t)(I \times I)^\omega) \\
&= (\nu \times \nu)(f_{s_*,\sigma}^{-1}(\mathcal{O}(hat)) \mid f_{s_*,\sigma}^{-1}(\mathcal{O}(ha))) \ ,
\end{aligned}$$

which proves that (3) holds for every cylinder h.

Now all the tools needed to prove (2) have been introduced, and we can state the main relation between $f_{s_*,\sigma}$ and $\mathrm{Pr}_{s_*}^\sigma(\phi)$. Let $\phi' \subseteq (SA)^\omega$ be the set of histories $s_0 a_1 s_1 \cdots$ such that $s_0 s_1 \cdots \in \phi$, and let $\mathbf{1}_\phi$ and $\mathbf{1}_{\phi'}$ be the indicator functions of ϕ and ϕ'. Then:

$$\begin{aligned}
\mathrm{Pr}_{s_*}^\sigma(\phi) &= \int_{p \in S^\omega} \mathbf{1}_\phi(p) \, d\mathrm{Pr}_{s_*}^\sigma(p) = \int_{p \in (SA)^\omega} \mathbf{1}_{\phi'}(p) \, d\mu(p) = \int_{p \in (SA)^\omega} \mathbf{1}_{\phi'}(p) \, d\mu'(p) \\
&= \int_{(x,y) \in [0,1]^\omega \times [0,1]^\omega} \mathbf{1}_{\phi'}(f_{s_*,\sigma}(x, y)) \, d(\nu \times \nu)(x, y) \\
&= \int_{x \in [0,1]^\omega} \left(\int_{y \in [0,1]^\omega} \mathbf{1}_{\phi'}(f_{s_*,\sigma}(x, y)) \, d\nu(y) \right) \, d\nu(x) \ ,
\end{aligned} \tag{4}$$

where the first and second equalities are by definition of $\mathrm{Pr}_{s_*}^\sigma(\phi)$, the third equality holds because $\mu = \mu'$, the fourth equality is a basic property of image measures, and the fifth equality holds by Fubini's theorem [4] that we can use since $\mathbf{1}_{\phi'} \circ f_{s_*,\sigma}$ is positive.

To complete the proof, we prove that for every $x \in [0, 1]^\omega$,

$$\mathrm{Pr}_s^{\sigma_x}(\phi) = \int_{y \in [0,1]^\omega} \mathbf{1}_{\phi'}(f_{s_*,\sigma}(x, y)) \, d\nu(y) \ , \tag{5}$$

Equation (4) holds for every observation-based strategy σ, hence in particular for strategy σ_x. But strategy σ_x has the following property: for every $x' \in]0, 1[^\omega$ and every $y \in [0, 1]^\omega$, $f_{s_*,\sigma_x}(x', y) = f_{s_*,\sigma}(x, y)$. Together with (4), this gives (5). This completes the proof, since (4) and (5) immediately give (2). ∎

	2½-player			1½-player	
	complete	one-sided	partial	MDP	POMDP
turn-based	$\epsilon > 0$	not	not	$\epsilon \geq 0$	$\epsilon \geq 0$
concurrent	not	not	not	(NA)	(NA)

Table 2. When deterministic (ϵ-optimal) strategies are as powerful as randomized strategies. The case $\epsilon = 0$ in complete-observation turn-based games is open.

Theorem 4 ([12]). *Let G be a* CoT *stochastic game with initial state s_* and an objective $\phi \subseteq S^\omega$. Then the following equalities hold:* $\inf_{\pi \in \Pi_O} \sup_{\sigma \in \Sigma_O} \Pr_{s_*}^{\sigma,\pi}(\phi) = \sup_{\sigma \in \Sigma_O} \inf_{\pi \in \Pi_O} \Pr_{s_*}^{\sigma,\pi}(\phi) = \sup_{\sigma \in \Sigma_O \cap \Sigma_P} \inf_{\pi \in \Pi_O} \Pr_{s_*}^{\sigma,\pi}(\phi)$.

We obtain the following result as a consequence of Lemma 1.

Theorem 5. *Let G be a* POMDP *with initial state s_* and an objective $\phi \subseteq S^\omega$. Then the following assertions hold:*

1. $\sup_{\sigma \in \Sigma_O} \Pr_{s_*}^{\sigma}(\phi) = \sup_{\sigma \in \Sigma_O \cap \Sigma_P} \Pr_{s_*}^{\sigma}(\phi)$.
2. *If there is a randomized optimal (resp. almost-sure winning, positive winning) strategy for ϕ from s_*, then there is a pure optimal (resp. almost-sure winning, positive winning) strategy for ϕ from s_*.*

Theorem 4 can be derived as a consequence of Martin's proof of determinacy of Blackwell games [12]: the result states that for CoT stochastic games pure strategies can achieve the same value as randomized strategies, and as a special case the result also holds for MDPs. Theorem 5 shows that the result can be generalized to POMDPs, and a stronger result (item (2) of Theorem 5) can be proved for POMDPs (and MDPs as a special case). It remains open whether result similar to item (2) of Theorem 5 can be proved for CoT stochastic games. The results summarizing when randomness can be obtained for free for strategies is shown in Table 2.

Undecidability result for POMDPs. The results of [2] shows that the emptiness problem for probabilistic coBüchi (resp. Büchi) automata under the almost-sure (resp. positive) semantics [2] is undecidable. As a consequence it follows that for POMDPs the problem of deciding if there is a pure observation-based almost-sure (resp. positive) winning strategy for coBüchi (resp. Büchi) objectives is undecidable, and as a consequence of Theorem 5 we obtain the same undecidability result for randomized strategies. This result closes an open question discussed in [9]. The undecidability result holds even if the coBüchi (resp. Büchi) objectives are visible.

Corollary 1. *Let G be a* POMDP *with initial state s_* and let $\mathcal{T} \subseteq S$ be a subset of states (or subset of observations). Whether there exists a pure or randomized almost-sure winning strategy for Player 1 from s in G for the objective* $\mathsf{coBuchi}(\mathcal{T})$ *is undecidable; and whether there exists a pure or randomized positive winning strategy for Player 1 from s in G for the objective* $\mathsf{Buchi}(\mathcal{T})$ *is undecidable.*

Undecidability result for one-sided complete-observation turn-based games. The undecidability results of Corollary 1 also holds for OsT stochastic games (as they subsume POMDPs as a special case). It follows from Theorem 3 that OsT stochastic games can be reduced to OsT deterministic games. Thus we obtain the first undecidability result for OsT deterministic games (the following corollary), solving the open question of [6].

Corollary 2. *Let G be an* OsT *deterministic game with initial state s_* and let $\mathcal{T} \subseteq S$ be a subset of states (or subset of observations). Whether there exists a pure or randomized almost-sure winning strategy for Player 1 from s in G for the objective* $\mathsf{coBuchi}(\mathcal{T})$ *is undecidable; and whether there exists a pure or randomized positive winning strategy for Player 1 from s in G for the objective* $\mathsf{Buchi}(\mathcal{T})$ *is undecidable.*

5 Conclusion

In this work we have presented a precise characterization for classes of games where randomization can be obtained for free in transitions and in strategies. As a consequence of our characterization we obtain new undecidability results. The other impact of our characterization is as follows: for the class of games where randomization is free in transition, future algorithmic and complexity analysis can focus on the simpler class of deterministic games; and for the class of games where randomization is free in strategies, future analysis of such games can focus on the simpler class of deterministic strategies. Thus our results will be useful tools for simpler analysis techniques in the study of games.

References

1. R. Alur, T.A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49:672–713, 2002.
2. C. Baier, N. Bertrand, and M. Größer. On decision problems for probabilistic Büchi automata. In *FoSSaCS: Foundations of Software Science and Computational Structures*, LNCS 4962, pages 287–301. Springer, 2008.
3. Nathalie Bertrand, Blaise Genest, and Hugo Gimbert. Qualitative determinacy and decidability of stochastic games with signals. In *Proc. of LICS*, pages 319–328. IEEE Computer Society, 2009.
4. P. Billingsley, editor. *Probability and Measure*. Wiley-Interscience, 1995.
5. J.R. Büchi and L.H. Landweber. Solving sequential conditions by finite-state strategies. *Transactions of the AMS*, 138:295–311, 1969.
6. K. Chatterjee, L. Doyen, T. A. Henzinger, and J.-F. Raskin. Algorithms for omega-regular games of incomplete information. *Logical Methods in Computer Science*, 3(3:4), 2007.
7. L. de Alfaro and T.A. Henzinger. Interface theories for component-based design. In *EMSOFT'01*, LNCS 2211, pages 148–165. Springer, 2001.
8. H. Everett. Recursive games. In *Contributions to the Theory of Games III*, volume 39 of *Annals of Mathematical Studies*, pages 47–78, 1957.
9. Vincent Gripon and Olivier Serre. Qualitative concurrent stochastic games with imperfect information. In *Proc. of ICALP (2): Automata, Languages and Programming*, LNCS 5556, pages 200–211. Springer, 2009.

10. T.A. Henzinger, O. Kupferman, and S. Rajamani. Fair simulation. *Information and Computation*, 173:64–81, 2002.
11. A. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.
12. D.A. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63(4):1565–1581, 1998.
13. R. McNaughton. Infinite games played on finite graphs. *Annals of Pure and Applied Logic*, 65:149–184, 1993.
14. Jean-François Mertens, Sylvain Sorin, and Shmuel Zamir. Repeated games. *Core Discussion Papers*, 9422, 1994.
15. A. Pnueli and R. Rosner. On the synthesis of a reactive module. In *POPL'89*, pages 179–190. ACM Press, 1989.
16. P.J. Ramadge and W.M. Wonham. Supervisory control of a class of discrete-event processes. *SIAM Journal of Control and Optimization*, 25(1):206–230, 1987.
17. W. Thomas. Languages, automata, and logic. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3, Beyond Words, chapter 7, pages 389–455. Springer, 1997.
18. M.Y. Vardi. Automatic verification of probabilistic concurrent finite-state systems. In *Proceedings of the 26th Annual Symposium on Foundations of Computer Science*, pages 327–338. IEEE Computer Society Press, 1985.
19. U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158:343–359, 1996.

A Appendix

Gadget for uniform-n-ary probability reduction for Theorem 2. We now show how to simulate a probabilistic state s^*, with n successors $s_0, s_1, \ldots, s_{n-1}$ such that the transition probability is $1/n$ to each of the successor, by a concurrent deterministic state. In the concurrent deterministic state s^* there are n actions $a_0, a_1, \ldots, a_{n-1}$ available for player 1 and n actions $b_0, b_1, \ldots, b_{n-1}$ available for player 2. The transition function is as follows: for $0 \leq i < n$ and $0 \leq j < n$ we have $\delta(s^*, a_i, b_j) = s_{(i+j) \mod n}$. Intuitively, the transition function matrix is obtained as follows: the first row is filled with states $s_0, s_1, \ldots, s_{n-1}$, and from a row i, the row $i+1$ is obtained by moving the state of the first column of row i to the last column in row $i+1$ and left-shifting by one position all the other states; the construction is illustrated on an example with $n = 4$ successors in (6). The construction ensures that in every row and every column each state $s_0, s_1, \ldots, s_{n-1}$ appears exactly once. It follows that if player 1 plays all actions uniformly at random, then against any probability distribution of player 2 the successor states are $s_0, s_1, \ldots, s_{n-1}$ with probability $1/n$ each; and a similar result holds if player 2 plays all actions uniformly at random. The correctness of the reduction for uniform-n-ary probabilistic state is then exactly as the proof of Theorem 2.

$$\begin{bmatrix} s_0 & s_1 & s_2 & s_3 \\ s_1 & s_2 & s_3 & s_0 \\ s_2 & s_3 & s_0 & s_1 \\ s_3 & s_0 & s_1 & s_2 \end{bmatrix} \tag{6}$$