Partial-Observation Stochastic Games: How to Win when Belief Fails

Krishnendu Chatterjee
IST Austria
Klosterneuburg, Austria

Laurent Doyen
LSV, ENS Cachan & CNRS
Cachan, France

***Abstract*—We consider two-player stochastic games played on finite graphs with reachability objectives where the first player tries to ensure a target state to be visited almost-surely (i.e., with probability 1), or positively (i.e., with positive probability), no matter the strategy of the second player.**

We classify such games according to the information and the power of randomization available to the players. On the basis of information, the game can be one-sided with either (a) player 1, or (b) player 2 having partial observation (and the other player has perfect observation), or two-sided with (c) both players having partial observation. On the basis of randomization, the players (a) may not be allowed to use randomization (pure strategies), or (b) may choose a probability distribution over actions but the actual random choice is external and not visible to the player (actions invisible), or (c) may use full randomization.

Our main results for pure strategies are as follows. (1) For one-sided games with player 1 having partial observation we show that (in contrast to full randomized strategies) *belief-based* (subset-construction based) strategies are not sufficient, and we present an exponential upper bound on memory both for almost-sure and positive winning strategies; we show that the problem of deciding the existence of almost-sure and positive winning strategies for player 1 is EXPTIME-complete. (2) For one-sided games with player 2 having partial observation we show that non-elementary memory is both necessary and sufficient for both almost-sure and positive winning strategies. (3) We show that for the general (two-sided) case finite-memory strategies are sufficient for both positive and almost-sure winning, and at least non-elementary memory is required.

We establish the equivalence of the almost-sure winning problems for pure strategies and for randomized strategies with actions invisible. Our equivalence result exhibits serious flaws in previous results of the literature: we show a non-elementary memory lower bound for almost-sure winning whereas an exponential upper bound was previously claimed.

***Keywords*-Partial-observation games, Stochastic games, Reachability and Büchi objectives, Positive and Almost-sure winning, Complexity, Memory bounds.**

I. Introduction

Games on graphs. Two-player games on graphs play a central role in several important problems in computer science, such as controller synthesis [33], [35], verification of open systems [2], realizability and compatibility checking [1], [21], [18], and many others. Most results about two-player games on graphs make the hypothesis of *perfect observation* (i.e., both players have perfect or complete observation about the state of the game). This assumption is often not realistic in practice. For example in the context of hybrid systems, the controller acquires information about the state of a plant using digital sensors with finite precision, which gives imperfect information about the state of the plant [20], [26]. Similarly, in a concurrent system where the players represent individual processes, each process has only access to the public variables of the other processes, not to their private variables [37], [2]. Such problems are better modeled in the more general framework of *partial-observation* games [36], [37], [38], [16], [6] and have been studied in the context of verification and synthesis [30], [2], [20], [45].

Partial-observation stochastic games and subclasses. In two-player partial-observation stochastic games on graphs with a finite state space, in every round, both players independently and simultaneously choose actions which along with the current state give a probability distribution over the successor states in the game. In a general setting, the players may not be able to distinguish certain states which are observationally equivalent for them (e.g., if they differ only by the value of private variables). The state space is partitioned into *observations* defined as equivalence classes and the players do not see the actual state of the game, but only an observation (which is typically different for the two players). The model of partial-observation games we consider is the same as the model of stochastic games with signals [6] and is a standard model in game theory [39], [41]. It subsumes other classical game models such as concurrent games [40], [19], probabilistic automata [34], [32], and partial-observation Markov decision processes (POMDPs) [31] (see also the recent decidability and complexity results for probabilistic automata [3], [4], [5], [9], [10], [11], [24] and for POMDPs [15], [3], [43]).

The special case of *perfect observation* for a player corresponds to every observation for this player being a singleton. Depending on which player has perfect observation, we consider the following *one-sided* subclasses of the general two-sided partial-observation stochastic games: (1) *player*-1 *partial and player*-2 *perfect* where player 2 has perfect observation, and player 1 has partial observation; and (2) *player*-1 *perfect and player*-2 *partial* where player 1 has perfect observation, and player 2 has partial observation. The case where the two players have perfect observation corresponds to the well-known perfect-information (perfect-observation) stochastic games [40], [17], [19].

Note that in a given (two-sided) game G, if player 1 wins in the setting of player-1 partial and player-2 perfect, then player 1 wins in the game G as well. Analogously, if player 1 cannot win in the setting of player 1 perfect and player 2 partial, then player 1 does not win in the game G either. In this sense, the one-sided games are conservative over- and under-approximations of two-sided games. In the context of applications in verification and synthesis, the conservative approximation is that the adversary is all powerful, and hence the games with player 1 partial and player 2 perfect games provide the important worst-case analysis of partial-observation games.

Objectives and qualitative problems. In this work we consider partial-observation stochastic games with *reachability* objectives where the goal of player 1 is to reach a set of target states, and games with *Büchi* objectives where the goal of player 1 is to visit some target state infinitely often. The study of partial-observation games is considerably more complicated than games of perfect observation. For example, in contrast to perfect-observation games, strategies in partial-observation games require both randomization and memory for reachability objectives; and the *quantitative* problem of deciding whether there exists a strategy for player 1 to ensure that the target is reached with probability at least $\frac{1}{2}$ can be decided in NP $\cap$ coNP for perfect-observation stochastic games [17], whereas the problem is undecidable even for partial-observation stochastic games with only one player [32]. Since the quantitative problem is undecidable, we consider the following *qualitative* problems: the *almost-sure* problem for reachability (resp. Büchi) objectives asks whether there exists a strategy for player 1 to ensure that the target set is reached (resp. visited infinitely often) with probability 1; the *positive* problem asks the same question, but requires positive probability instead of probability 1. For Büchi objectives, the positive problem is undecidable [3], and the almost-sure problem is polynomially equivalent to the almost-sure problem for reachability objectives [3]. Therefore, we discuss reachability objectives, and the results for Büchi objectives follow.

Classes of strategies. In general, randomized strategies are necessary to win with probability 1 in a partial-observation game with reachability objective [16]. However, there exist two types of randomized strategies where either (i) actions are visible, the player can observe the action he played [16], [6], or (ii) actions are invisible, the player may choose a probability distribution over actions, but the source of randomization is external and the actual (random) choice of the action is invisible to the player [25]. The second model is more general since the qualitative problems of randomized strategies with actions visible can be reduced in polynomial time to randomized strategies with actions invisible, by modeling the visibility of actions using the observations on states.

With actions visible, the almost-sure (resp. positive) problem was shown to be EXPTIME-complete (resp. PTIME-complete) for one-sided games with player 1 partial and player 2 perfect [16], and 2EXPTIME-complete (resp. EXPTIME-complete) in the two-sided case [6]. For the positive problem memoryless randomized strategies exist, and for the almost-sure problem *belief-based* strategies exist (strategies based on subset construction that consider the possible current states of the game). It was remarked (without any proof) in [16, p.4] that these results easily extend to randomized strategies with actions invisible for one-sided games with player 1 partial and player 2 perfect. It was claimed in [25] (Theorems 1 & 2) that the almost-sure problem is 2EXPTIME-complete for randomized strategies with actions invisible for two-sided games, and that belief-based strategies are sufficient for player 1. Thus it is believed that the two qualitative problems with actions visible or actions invisible are essentially equivalent.

Pure strategies and motivation. In this paper, we consider the class of *pure* strategies, which do not use randomization at all. Pure strategies arise naturally in the synthesis of controllers and processes that do not have access to any source of randomization, such as synchronizers for lock placement in concurrent programs [8], and controllers for robot planning [29]. Moreover we will establish deep connections between the qualitative problems for pure strategies and for randomized strategies with actions invisible, which on one hand exhibit major flaws in previous results of the literature (the remark without proof of [16] and the main results of [25]), and on the other hand show that the solution for almost-sure winning randomized strategies with actions invisible (which is the most general case) can be surprisingly obtained by solving the problem for pure strategies.

Contributions. The contributions of the paper are as follows.

1) *Player* 1 *partial and player* 2 *perfect.* We show that both for almost-sure and positive winning, belief-based pure strategies are not sufficient. This implies that the classical approaches relying on the belief-based subset construction cannot work for solving the qualitative problems for pure strategies. However, we present an optimal exponential upper bound on the memory needed by pure strategies (the exponential lower bound follows from the special case of non-stochastic games [7]). By a reduction to perfect-observation games of exponential size, we show that both the almost-sure and positive problems are EXPTIME-complete for one-sided games with perfect-observation for player 2. In contrast to the previous proofs of EXPTIME upper bound that rely either on subset constructions or enumeration of belief-based strategies, our correctness proof relies on a novel rank-based argument that works uniformly both for positive and almost-sure winning. The structure of this construction also provides symbolic antichain-based algorithms (see [22] for a survey of the antichain approach) for solving the qualitative problems that avoids the explicit exponential construction. Thus for the important special case of player 1 partial and player 2 perfect we establish optimal memory bound, complexity bound, and obtain symbolic algorithmic solutions for the qualitative problems.

	one-sided player 2 perfect		one-sided player 1 perfect		two-sided	
	Positive	Almost-sure	Positive	Almost-sure	Positive	Almost-sure
Randomized (actions visible)	Memoryless	Exponential (belief-based)	Memoryless	Memoryless	Memoryless	Exponential (belief-based)
Randomized (actions invisible)	Memoryless	**Exponential (belief is not sufficient)**	Memoryless	Memoryless	Memoryless	**Non-elem. low. bound** **Finite upp. bound**
Pure	**Exponential (belief is not sufficient)**	**Exponential (belief is not sufficient)**	**Non-elem. complete**	**Non-elem. complete**	**Non-elem. low. bound** **Finite upp. bound**	**Non-elem. low. bound** **Finite upp. bound**

TABLE I
MEMORY REQUIREMENT FOR PLAYER 1 AND REACHABILITY OBJECTIVE.

2) *Player* 1 *perfect and player* 2 *partial.*
   a) We show a very surprising result that both for positive and almost-sure winning, pure strategies for player 1 require memory of non-elementary size (i.e., a tower of exponentials). This is in sharp contrast with (i) the case of randomized strategies (with or without actions visible) where memoryless strategies are sufficient for positive winning, and with (ii) the previous case where player 1 has partial observation and player 2 has perfect observation, where pure strategies for positive winning require only exponential memory. Surprisingly and perhaps counter-intuitively when player 1 has more information and player 2 has less information, the positive winning strategies for player 1 require much more memory (non-elementary as compared to exponential). With more information player 1 can win from more states, but the winning strategy is much harder to implement.
   b) We present a non-elementary upper bound for the memory needed by pure strategies for positive winning. We then show with an example that for almost-sure winning more memory may be required as compared to positive winning. Finally, we show how to combine pure strategies for positive winning in a recharging scheme to obtain a non-elementary upper bound for the memory required by pure strategies for almost-sure winning. Thus we establish non-elementary complete bounds for pure strategies both for positive and almost-sure winning.
3) *General (two-sided) case.* We show that in the general case finite memory strategies are sufficient both for positive and almost-sure winning. The result is obtained essentially by a simple generalization of König's Lemma [28]. A non-elementary lower bound for memory follows from the special case when player 1 has perfect observation and player 2 has partial observation.
4) *Randomized strategies with actions invisible.* For randomized strategies with actions invisible we give two reductions to establish connections with pure strategies. First, we show that the almost-sure problem for randomized strategies with actions invisible reduces in polynomial time to the almost-sure problem for pure strategies. The reduction requires to first establish that finite-memory randomized strategies are sufficient in two-sided games. Second, we show that the problem of almost-sure winning with pure strategies reduces in polynomial time to the problem of randomized strategies with actions invisible. For this reduction it is crucial that the actions are not visible.
   Our reductions have deep consequences. They unexpectedly imply that the problems of almost-sure winning with *pure* strategies or *randomized* strategies with actions invisible are polynomial-time *equivalent*. Moreover, it follows that even in one-sided games with player 1 partial and player 2 perfect, belief-based randomized strategies with actions invisible are not sufficient for almost-sure winning. This shows that the remark (without proof) of [16] that the results (such as existence of belief-based strategies) of randomized strategies with actions visible carry over to actions invisible is an oversight. However from our first reduction and our results for pure strategies it follows that there is an exponential upper bound on memory and the problem is EXPTIME-complete for one-sided games with player 1 partial and player 2 perfect. More importantly, our results exhibit a serious flaw[1] in the main result of [25] which showed that belief-based randomized strategies with actions invisible are sufficient for almost-sure winning in two-sided games, and concluded that enumerating over such strategies yields a 2EXPTIME algorithm for the problem. Our second reduction and lower bound for pure strategies show that the result is incorrect, and that the exponential (belief-based) upper bound is far off. Instead, the lower bound on memory for almost-sure winning with randomized strategies and actions invisible is non-elementary. Thus, contrary to the general belief, there is a sharp contrast for randomized strategies with or without actions visible: if actions are visible, then exponential memory is sufficient for almost-sure winning while if actions are not visible, then memory of non-elementary size is necessary in general.

The memory requirements are summarized in Table I and the results of this paper are shown in bold font. We explain

[1]This flaw was presented in [13] to the authors of [25] and acknowledged in August 2011.

how the other results of the table follow from results of the literature. For randomized strategies (with or without actions visible), if a positive winning strategy exists, then a memoryless strategy that plays all actions uniformly at random is also positive winning. Thus the memoryless result for positive winning strategies follows for all cases of randomized strategies. The belief-based bound for memory of almost-sure winning randomized strategies with actions visible follows from [16], [6]. The memoryless strategies results for almost-sure winning for one-sided games with player 1 perfect and player 2 partial are obtained as follows: when actions are visible, then belief-based strategies coincide with memoryless strategies as player 1 has perfect observation. If player 1 has perfect observation, then for memoryless strategies whether actions are visible or not is irrelevant and thus the memoryless result also follows for randomized strategies with actions invisible. Thus we obtain Table I. Proofs omitted due to lack of space are available in a technical report released in July 2011 [13].

II. Definitions

A *probability distribution* on a finite set S is a function $\kappa : S \to [0,1]$ such that $\sum_{s\in S} \kappa(s) = 1$. The *support* of κ is the set $\mathsf{Supp}(\kappa) = \{s \in S \mid \kappa(s) > 0\}$. We denote by $\mathcal{D}(S)$ the set of probability distributions on S. Given $s \in S$, the *Dirac distribution* on s assigns probability 1 to s.

Games. Given finite alphabets A_i of actions for player i ($i = 1, 2$), a *stochastic game* on A_1, A_2 is a tuple $G = \langle Q, q_0, \delta\rangle$ where Q is a finite set of states, $q_0 \in Q$ is the initial state, and $\delta : Q \times A_1 \times A_2 \to \mathcal{D}(Q)$ is a probabilistic transition function that, given a current state q and actions a, b for the players gives the transition probability $\delta(q,a,b)(q')$ to the next state q'. The game is called *deterministic* if $\delta(q, a, b)$ is a Dirac distribution for all $(q, a, b) \in Q \times A_1 \times A_2$. A state q is *absorbing* if $\delta(q, a, b)$ is the Dirac distribution on q for all $(a, b) \in A_1 \times A_2$. In some examples, we allow an initial distribution of states. This can be encoded in our game model by a probabilistic transition from the initial state.

A *player*-1 *state* is a state q where $\delta(q,a,b) = \delta(q,a,b')$ for all $a \in A_1$ and all $b, b' \in A_2$. We use the notation $\delta(q, a, -)$. *Player*-2 *states* are defined analogously. In figures, we use boxes to emphasize that a state is a player-2 state, and we represent probabilistic branches using diamonds (which are not real 'states', e.g., as in Fig. 1).

In a (two-sided) *partial-observation* game, the players have a partial or incomplete view of the states visited and of the actions played in the game. This view may be different for the two players and it is defined by equivalence relations $\approx_i$ on the states and on the actions ($i = 1, 2$). For player i, equivalent states (or actions) are indistinguishable. We denote by $\mathcal{O}_i \subseteq 2^Q$ ($i = 1, 2$) the $\approx_i$-equivalence classes of states which define two partitions of the state space Q, and we call them *observations* (for player i). These partitions uniquely define functions $\mathsf{obs}_i : Q \to \mathcal{O}_i$ such that $q \in \mathsf{obs}_i(q)$ for all $q \in Q$, that map each state q to its observation for player i.

In the case where all states and actions are equivalent (i.e., the relation $\approx_i$ is the set $(Q \times Q) \cup (A_1 \times A_1) \cup (A_2 \times A_2)$), we say that player i is *blind* and the actions are *invisible*. In this case, we have $\mathcal{O}_i = \{Q\}$ because all states have the same observation. Note that the case of perfect observation for player i corresponds to the case $\mathcal{O}_i = \{\{q_0\}, \{q_1\}, \dots, \{q_n\}\}$ (given $Q = \{q_0, q_1, \dots, q_n\}$), and $a \approx_i b$ iff $a = b$, for all actions a, b.

For $s \subseteq Q$, $a \in A_1$, and $b \in A_2$, let $\mathsf{Post}_{a,b}(s) = \bigcup_{q\in s} \mathsf{Supp}(\delta(q, a, b))$ denote the set of possible successors of q given action a and b, and let $\mathsf{Post}_{a,-}(s) = \bigcup_{b\in A_2} \mathsf{Post}_{a,b}(s)$.

Plays and observations. Initially, the game starts in the initial state q_0. In each round, player 1 chooses an action $a \in A_1$, player 2 (simultaneously and independently) chooses an action $b \in A_2$, and the successor of the current state q is chosen according to the probabilistic transition function $\delta(q, a, b)$. A *play* in G is an infinite sequence $\rho = q_0\, a_0 b_0\, q_1\, a_1 b_1\, q_2 \dots$ such that q_0 is the initial state and $\delta(q_j, a_j, b_j)(q_{j+1}) > 0$ for all $j \geq 0$ (the actions a_j's and b_j's are the actions *associated* to the play). Its *length* is $|\rho| = \infty$. The length of a play prefix $\rho = q_0\, a_0 b_0\, q_1 \dots q_k$ is $|\rho| = k$, and its last element is $\mathsf{Last}(\rho) = q_k$. A state $q \in Q$ is *reachable* if it occurs in some play. We denote by $\mathsf{Plays}(G)$ the set of plays in G, and by $\mathsf{Prefs}(G)$ the set of corresponding finite prefixes. For $i = 1, 2$, the *observation sequence* for player i of a play (prefix) ρ is the unique (in)finite sequence $\mathsf{obs}_i(\rho) = \gamma_0\gamma_1 \dots$ such that $\gamma_j = \mathsf{obs}_i(q_j)$ for all $0 \leq j \leq |\rho|$.

The games with *one-sided partial-observation* are the special case where either $\approx_1$ is equality and hence $\mathcal{O}_1 = \{\{q\} \mid q \in Q\}$ (player 1 has complete observation) or $\approx_2$ is equality and hence $\mathcal{O}_2 = \{\{q\} \mid q \in Q\}$ (player 2 has complete observation). The games with *perfect observation* are the special cases where $\approx_1$ and $\approx_2$ are equality, i.e., every state and action is visible to both players.

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 $\rho = q_0\, a_0 b_0\, q_1 \dots$ and $\rho' = q_0'\, a_0' b_0'\, q_1' \dots$, if $a_j \approx_1 a_j'$ and $b_j \approx_1 b_j'$ for all $j \geq 0$, and $\mathsf{obs}_1(\rho) = \mathsf{obs}_1(\rho')$, then $\sigma(\rho) = \sigma(\rho')$. In the sequel, strategies are meant to be observation-based in partial-observation games. If for all actions a and b we have $a \approx_1 b$ iff $a = b$, and $a \approx_2 b$ iff $a = b$ (all actions are distinguishable), then the strategy is *action visible*, and if for all actions a and b we have $a \approx_1 b$ and $a \approx_2 b$ (all actions are indistinguishable), then the strategy is *action invisible*. We say that a play (prefix) $\rho = q_0\, a_0 b_0\, q_1 \dots$ is *compatible* with a pure (resp., randomized) strategy σ if the associated action of player 1 in step j is $a_j = \sigma(q_0\, a_0 b_0 \dots q_{j-1})$ (resp., $a_j \in \mathsf{Supp}(\sigma(q_0\, a_0 b_0 \dots q_{j-1}))$) for all $0 \leq j \leq |\rho|$.

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.

Remark 1. *The model of games with partial observation on both actions and states can be encoded in a model of games with actions invisible and observations on states only: when actions are invisible, we can use the state space to keep track of the last action played, and reveal information about the last action played using observations on the states [25]. Therefore, in the sequel we assume that the actions are invisible to the players with partial observation. A play is then viewed as a sequence of states only, and the definition of strategies is updated accordingly. Note that a player with perfect observation has actions and states visible (and the equivalence relation $\approx_i$ is equality).*

Remark 2. *The important special case of partial-observation Markov decision processes (POMDP) corresponds to the case where either all states in the game are player-1 states (player-1 POMDP) or all states are player-2 states (player-2 POMDP). For POMDP it is known that randomization is not necessary, and pure strategies are as powerful as randomized strategies [14].*

Finite-memory strategies. A player-1 strategy uses *finite-memory* if it can be encoded by a deterministic transducer $\langle \mathsf{Mem}, m_0, \alpha_u, \alpha_n \rangle$ where Mem is a finite set (the memory of the strategy), $m_0 \in \mathsf{Mem}$ is the initial memory value, $\alpha_u : \mathsf{Mem} \times \mathcal{O}_1 \to \mathsf{Mem}$ is an update function, and $\alpha_n : \mathsf{Mem} \times \mathcal{O}_1 \to \mathcal{D}(A_1)$ is a next-move function. The *size* of the strategy is the number $|\mathsf{Mem}|$ of memory values. If the current observation is o, and the current memory value is m, then the strategy chooses the next action according to the probability distribution $\alpha_n(m, o)$, and the memory is updated to $\alpha_u(m, o)$. Formally, $\langle \mathsf{Mem}, m_0, \alpha_u, \alpha_n \rangle$ defines the strategy σ such that $\sigma(\rho \cdot q) = \alpha_n(\hat{\alpha}_u(m_0, \mathsf{obs}_1(\rho)), \mathsf{obs}_1(q))$ for all $\rho \in Q^*$ and $q \in Q$, where $\hat{\alpha}_u$ extends α_u to sequences of observations as expected. This definition extends to infinite-memory strategies by dropping the assumption that the set Mem is finite. A strategy is *memoryless* if $|\mathsf{Mem}| = 1$.

Objectives and winning modes. An *objective* (for player 1) in G is a set $\varphi \subseteq \mathsf{Plays}(G)$ of plays. A play $\rho \in \mathsf{Plays}(G)$ *satisfies* the objective φ, denoted $\rho \models \varphi$, if $\rho \in \varphi$. Objectives are generally Borel measurable: a Borel objective is a Borel set in the Cantor topology [27]. Given strategies σ and π for the two players, the probabilities of a measurable objective φ is uniquely defined [44]. We denote by $\mathrm{Pr}_{q_0}^{\sigma,\pi}(\varphi)$ the probability that φ is satisfied by the play obtained from the starting state q_0 when the strategies σ and π are used.

We specifically consider the following well-known objectives. Given a set $\mathcal{T} \subseteq Q$ of target states, the *reachability objective* requires that the play visit the set $\mathcal{T}$: $\mathsf{Reach}(\mathcal{T}) = \{q_0 a_0 b_0 q_1 \ldots \in \mathsf{Plays}(G) \mid \exists i \geq 0 : q_i \in \mathcal{T}\}$, and the *Büchi objective* requires that the play visit the set $\mathcal{T}$ infinitely often, $\mathsf{Büchi}(\mathcal{T}) = \{q_0 a_0 b_0 q_1 \ldots \in \mathsf{Plays}(G) \mid \forall i \geq 0 \cdot \exists j \geq i : q_j \in \mathcal{T}\}$. Our solution for reachability objectives will also use the dual notion of *safety objectives* that require the play to stay within the set $\mathcal{T}$: $\mathsf{Safe}(\mathcal{T}) = \{q_0 a_0 b_0 q_1 \ldots \in \mathsf{Plays}(G) \mid \forall i \geq 0 : q_i \in \mathcal{T}\}$. In figures, the target states in $\mathcal{T}$ are double-lined and labeled by ☺.

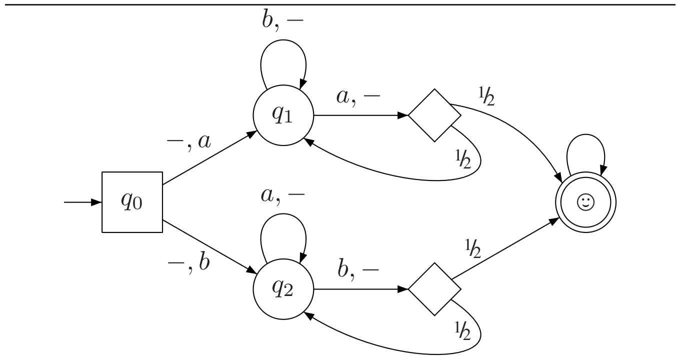

Fig. 1. Belief-based pure strategies are not sufficient for positive and almost-sure reachability.

Given a game structure G and a state q, an observation-based strategy σ for player 1 is *almost-sure winning* (resp. *positive winning*) for the objective φ from q if for all observation-based randomized strategies π for player 2, we have $\mathrm{Pr}_q^{\sigma,\pi}(\varphi) = 1$ (resp. $\mathrm{Pr}_q^{\sigma,\pi}(\varphi) > 0$). The strategy σ is *sure winning* if all plays compatible with σ satisfy φ. We also say that the state q is almost-sure (or positive, or sure) winning for player 1.

Positive and almost-sure winning problems. We are interested in the problems of deciding, given a game structure G, a state q, and an objective φ, whether there exists a {pure, randomized} strategy which is {almost-sure, positive} winning from q for the objective φ. For safety objectives almost-sure winning coincides with sure winning, however for reachability objectives they are different. The sure winning problem for the objectives we consider has been studied in [36], [16], [12]. The almost-sure winning problem for Büchi objectives can be easily reduced to the almost-sure winning problem for reachability objectives [3]. The positive winning problem for Büchi objectives is undecidable even for POMDPs [3]. Hence in this paper we mostly focus on reachability objectives.

Remark 3. *(Almost-sure Büchi to almost-sure reachability [3]). The reduction of almost-sure Büchi to almost-sure reachability is as follows: given a two-sided stochastic game with Büchi objective* $\mathsf{Büchi}(\mathcal{T})$*, we add a new absorbing state q_T, make q_T the target state for the reachability objective, and from every state $q \in \mathcal{T}$ we add positive probability transitions to q_T (details and correctness proof follow from [3, Lemma 13]).*

III. One-sided Games: Player 1 Partial and Player 2 Perfect

In Sections III and IV, we consider one-sided games with partial observation: one player has perfect observation, and the other player has partial observation. The player with perfect observation sees the states visited and the actions played in the game. We present the results for positive and almost-sure

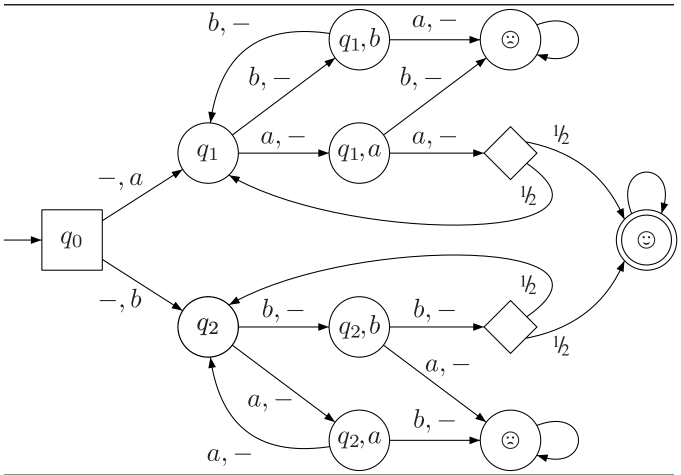

Fig. 2. Belief-based randomized action-invisible strategies are not sufficient for almost-sure reachability.

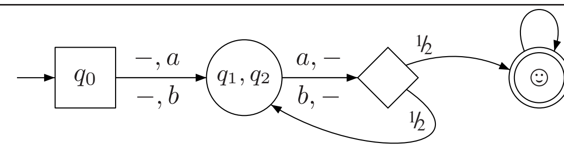

Fig. 3. A belief-based subset construction for the reachability game of Fig. 1.

winning for reachability objectives along with examples that illustrate key elements of the problem such as the memory required for winning strategies.

Note that the case of player 1 partial and player 2 perfect is important in the context of controller synthesis as it is a conservative approximation of two-sided games for player 1 (if player 1 wins in the one-sided game, then he also wins in the two-sided game). In the following example we show that for pure strategies *belief-based* strategies are not sufficient for positive as well as almost-sure winning. A strategy is belief-based if its memory relies only on the subset construction, i.e., the strategy plays only depending on the set of possible current states of the game which is called *belief*.

Example 1. Belief is not sufficient for positive (as well as almost-sure) reachability. *Consider the game in Fig. 1 where player 1 is blind (all states have the same observation except the target state, and actions are invisible) and player 2 has perfect observation. Initially, player 2 chooses the state q_1 or q_2 (which player 1 does not see). The belief of player 1 is thus the set $\{q_1, q_2\}$ (see Fig. 3). We claim that the belief is not a sufficient information to win with a pure strategy for player 1 because the belief-based subset construction in Fig. 3 suggests that playing always the same action (say a) when the belief is $\{q_1, q_2\}$ is an almost-sure winning strategy. However, in the original game this is not even a positive winning strategy (the counter strategy of player 2 is to choose q_2 initially). A winning strategy for player 1 is to alternate between a and b when the belief is $\{q_1, q_2\}$, showing that remembering the belief is not sufficient.* ■

We present reductions of the almost-sure and positive winning problem for reachability objective to the problem of sure-winning in a game of perfect observation with Büchi objective, and reachability objective respectively. The two reductions are based on the same construction of a game where the state space $L = \{(s, o) \mid o \subseteq s \subseteq Q\}$ contains the subset construction s enriched with *obligation sets* $o \subseteq s$ which ensure that from all states in s, the target set $\mathcal{T}$ is reached with positive probability. The Büchi (resp. reachability) objective is to visit the empty obligation set infinitely often (resp. at least once). Instead of a naive solution which would keep track of all successors of probabilistic choices, we use a rank-based argument on the obligation set to show correctness of the construction. The key argument considers arbitrary (possibly infinite-memory) almost-sure winning strategy σ, and proves the existence of a finite ranking in the infinite tree obtained from σ such that target states have rank 0, the rank is strictly decreasing for non-target states, and the root gets a finite rank.

The construction is as follows. Given $G = \langle Q, q_0, \delta\rangle$ over alphabets A_1, A_2 and observation set $\mathcal{O}_1$ for player 1, with reachability objective $\mathsf{Reach}(\mathcal{T})$, we construct the following (deterministic) game of perfect observation $H = \langle L, \ell_0, \delta_H\rangle$ over alphabets A'_1, A'_2 such that player 1 has a pure observation-based almost-sure (resp., positive) winning strategy in G from q_0 if and only if player 1 has a sure winning strategy in H from ℓ_0 for the objective $\mathsf{Büchi}(\alpha)$ (resp., $\mathsf{Reach}(\alpha)$) defined by $\alpha \subseteq L$ where:

- $L = \{(s, o) \mid o \subseteq s \subseteq Q\}$. Intuitively, s is the belief of player 1 and o is a set of obligation states that "owe" a visit to $\mathcal{T}$ with positive probability.
- $\ell_0 = (\{q_0\}, \{q_0\})$ if $q_0 \notin \mathcal{T}$, and $\ell_0 = (\emptyset, \emptyset)$ if $q_0 \in \mathcal{T}$;
- $A'_1 = A_1 \times 2^Q$. In a pair $(a, u) \in A'_1$, we call a the action, and u the witness set;
- $A'_2 = \mathcal{O}_1$. In the game H, player 2 simulate player 2's choice in game G, as well as resolves the probabilistic choices. This amounts to choosing a possible successor state, and revealing its observation;
- $\alpha = \{(s, \emptyset) \in L\}$;
- δ_H is defined as follows. First, the state $(\emptyset, \emptyset)$ is absorbing. Second, in every other state $(s, o) \in L$ the function δ_H ensures that (i) player 1 chooses a pair (a, u) such that $\mathsf{Supp}(\delta(q, a, b)) \cap u \neq \emptyset$ for all $q \in o$ and $b \in A_2$, and (ii) player 2 chooses an observation $\gamma \in \mathcal{O}_1$ such that $\mathsf{Post}_{a,-}(s) \cap \gamma \neq \emptyset$. If a player violates this, then a losing absorbing state is reached with probability 1. Assuming the above condition on (a, u) and γ is satisfied, define $\delta_H((s, o), (a, u), \gamma)$ as the Dirac distribution on the state (s', o') such that:
  - $s' = (\mathsf{Post}_{a,-}(s) \cap \gamma) \setminus \mathcal{T}$;
  - $o' = s'$ if $o = \emptyset$; and $o' = (\mathsf{Post}_{a,-}(o) \cap \gamma \cap u) \setminus \mathcal{T}$ if $o \neq \emptyset$.

Lemma 1. *Given a one-sided partial-observation stochastic game G with player 1 partial and player 2 perfect with a reachability objective for player 1, we can construct in time* *exponential in the size of the game and polynomial in the size of action sets a perfect-information deterministic game H with a Büchi objective (resp. reachability objective) such that player 1 has a pure almost-sure (resp. positive) winning strategy in G iff player 1 has a sure-winning strategy in H.*

It follows from the construction in the proof of Lemma 1 that pure strategies with exponential memory are sufficient for positive (as well as almost-sure) winning, and the exponential lower bound follows from the special case of non-stochastic games [7]. Lemma 1 also gives EXPTIME upper bound for the problem since perfect-observation Büchi games can be solved in polynomial time [42]. The EXPTIME-hardness follows from the sure winning problem for non-stochastic games [37], where pure almost-sure (positive) winning strategies coincide with sure winning strategies. Theorem 1 summarizes the results, and note that by Remark 3 all the results of the theorem for almost-sure winning also hold for Büchi objectives.

Theorem 1. *Given one-sided partial-observation stochastic games with player 1 partial and player 2 perfect, the following assertions hold for reachability objectives for player 1:*

1) (Memory bound). *Belief-based pure strategies are not sufficient both for positive and almost-sure winning; exponential memory is necessary and sufficient both for positive (memory of size $\sum_{\gamma \in \mathcal{O}_1} 2^{|\gamma|}$ is sufficient) and almost-sure winning (memory of size $\sum_{\gamma \in \mathcal{O}_1} 3^{|\gamma|}$ is sufficient) for pure strategies, where $|\gamma|$ is the cardinality of γ.*
2) (Algorithm). *The problems of deciding the existence of a pure almost-sure and a pure positive winning strategy can be solved in time exponential in the state space of the game and polynomial in the size of the action sets.*
3) (Complexity). *The problems of deciding the existence of a pure almost-sure and a pure positive winning strategy are EXPTIME-complete.*

Symbolic algorithms. The exponential Büchi (or reachability) game constructed in the proof of Lemma 1 can be solved by computing classical fixpoint formulas [23]. However, it is not necessary to construct the exponential game structure explicitly. Instead, we can exploit the structure induced by the pre-order $\preceq$ defined by $(s, o) \preceq (s', o')$ if (i) $s \subseteq s'$, (ii) $o \subseteq o'$, and (iii) $o = \emptyset$ iff $o' = \emptyset$. Intuitively, if a state (s', o') is winning for player 1, then all states $(s, o) \preceq (s', o')$ are also winning because they correspond to a better belief and a looser obligation. Hence all sets computed by the fixpoint algorithm are downward-closed and thus they can be represented symbolically by the antichain of their maximal elements (see [16] for details related to antichain algorithms). This technique provides a symbolic algorithm without explicitly constructing the exponential game.

IV. One-sided Games: Player 1 Perfect and Player 2 Partial

Recall that we are interested in finding a pure winning strategy for player 1. We present the key ideas of the main three results for one-sided games with player 1 perfect and player 2 partial.

Lower bound on memory. We present a family of games where player 1 needs memory of non-elementary size to satisfy both almost-sure and positive reachability. The key idea is that player 1 needs to remember not only the possible current states of the game (belief of player 2), but also how many paths that player 2 cannot distinguish end up in each state. Then we show that player 1 needs to simulate a counter system where the operations on counters are increment and division by 2 (with round down) which requires to store non-elementary values of the counters in the worst case. The key challenge is to construct a polynomial-size game to simulate non-elementary counter values. We show how to use the partial observation of player 2 to achieve this. This establishes the surprising non-elementary lower bound. See [13, Theorem 2] for details.

Upper bound for positive reachability with almost-sure safety. We show a matching non-elementary upper bound for pure strategies to ensure positive reachability along with almost-sure safety. We obtain the solution for positive reachability as a special case and on the other hand it will be required for solving almost-sure reachability. The result is achieved in the following steps. First, we compute the set of states from which player 1 can satisfy the objective with a *randomized action-visible* strategy. Second, we show how pure strategies can simulate randomized strategies by using the stochasticity of the transition relation and the fact that player 2 cannot distinguish observationally-equivalent paths. This is the main novel idea behind this proof. Finally, we show that if the number of indistinguishable paths is non-elementary, then player 1 achieves the full power of randomized action-visible strategies and is winning using the computation of the first step. The crux of the final step is to analyze a new class of counter systems (with division by a constant and increment) and show that counters with non-elementary value suffice. See [13, Theorem 3] for details.

Upper bound for almost-sure reachability. We show an example of a game where memoryless positive winning strategies exist, but almost-sure winning strategies require memory [13, Example 4]. We then present a construction of a pure almost-sure winning strategy (when such a strategy exists) by repeatedly playing a strategy for positive reachability along with almost-sure safety in a *recharging* scheme. As a consequence we obtain a non-elementary upper bound on the memory size of almost-sure winning strategies. Let Q_B be the set of states such that if the belief of player 2 is a state in Q_B, then against all strategies of player 1, player 2 can ensure that with positive probability the target is not reached. Hence an almost-sure winning strategy must ensure almost-sure safety for the set $Q_G = Q \setminus Q_B$. From Q_G player 1 can ensure both positive reachability to the target as well as safety for the set Q_G. We show that repeatedly playing a strategy for positive reachability along with almost-sure safety is an almost-sure winning strategy for the reachability objective (details in [13, Theorem 4]). By Remark 3, the results of Theorem 2 and

Corollary 1 for almost-sure winning also hold for Büchi objectives.

Theorem 2. *In one-sided partial-observation stochastic games with player 1 perfect and player 2 partial, the following assertions hold:*

1) *Both pure almost-sure and pure positive winning strategies for reachability objectives for player 1 require memory of non-elementary size in general.*
2) *Non-elementary size memory is sufficient for pure strategies to ensure positive probability reachability along with almost-sure safety for player 1; and hence for pure positive winning strategies for reachability objectives for player 1 non-elementary memory bound is optimal.*
3) *Non-elementary size memory is sufficient for pure strategies to ensure almost-sure reachability for player 1; and hence for pure almost-sure winning strategies for reachability objectives for player 1 non-elementary memory bound is optimal.*

Corollary 1. *In one-sided partial-observation stochastic games with player 1 perfect and player 2 partial, the problem of deciding the existence of pure almost-sure and positive winning strategies for reachability objectives for player 1 can be solved in non-elementary time complexity.*

Discussion about the surprising non-elementary memory bound. We now discuss the surprising non-elementary memory bound for positive winning with reachability objectives for pure strategies in player-1 perfect player-2 partial stochastic games, comparing it with other related questions. We consider four related questions: two are related to stochasticity in transitions and strategies, and the other two are related to the information of the players (see also Fig. 4).

1) *Question 1.* If we consider player-1 perfect player-2 partial deterministic games with reachability objective, then for positive winning pure memoryless strategies are sufficient. This follows from the results of [36] because in deterministic games positive winning coincides with sure winning, and the results of [36] shows (see [16] for an explicit proof) that for sure winning the observation of player 2 is irrelevant. Hence the problem is same as sure winning in perfect-information deterministic games with reachability objective for which pure memoryless strategies exist.
2) *Question 2.* If we consider player-1 perfect player-2 partial stochastic games with reachability objective, but instead of pure strategies consider randomized strategies, then memoryless strategies are sufficient. It follows from [6] that if there is a randomized strategy to ensure reachability with positive probability, then the randomized memoryless strategy that plays all actions uniformly at random is also a positive winning strategy.
3) *Question 3.* If we consider perfect-information stochastic games (both players have perfect information) with reachability objective, then for positive winning pure memoryless strategies are sufficient. This follows from a more general result of [17] that in perfect-information stochastic games with reachability objective, pure memoryless optimal strategies exist.
4) *Question 4.* If we consider player-1 partial player-2 perfect stochastic games with reachability objective, then for positive winning exponential memory pure strategies are sufficient (by Theorem 1).

Observe that the question we study is a natural extension of the above questions: (1) adding stochasticity to the transition as compared to question 1; (2) restricting strategies to pure strategies as compared to randomized strategies of question 2; (3) player 2 is less informed as compared to question 3; and (4) player 1 is more informed and player 2 is less informed as compared to question 4. Our results show the natural variant of question 1 and question 2 obtained by adding stochasticity to transitions or removing stochasticity from strategies, and the variant of question 3 and question 4 by making player 1 most well informed lead to a sunrising memory bound for strategies (non-elementary complete memory bound, whereas for all the related questions memoryless or exponential-size memory strategies are sufficient). See also Fig. 4.

V. Two-sided Games

We show the existence of finite-memory pure strategies for positive and almost-sure winning in two-sided games.

Positive reachability with almost-sure safety. We show that to ensure positive reachability along with almost-sure safety, finite-memory strategies suffice. The proof is in two parts: (1) we show that if there is an infinite-memory strategy σ, then the strategy ensures positive reachability within a finite number N of steps and almost-sure safety (the result is shown by a simple extension of König's Lemma [28]), and (2) then a finite-memory strategy plays like σ for N steps and then switches to a strategy for almost-sure safety (and for almost-sure safety finite-memory strategies suffice [12]). See [13, Theorem 5] for details.

Almost-sure reachability. The proof to show that finite-memory strategies suffice for almost-sure winning is analogous to the proof of the previous section for player 1 perfect and player 2 partial, where an almost-sure winning strategy is constructed by repeatedly playing finite-memory strategies (of [13, Theorem 5]) for positive reachability along with almost-sure safety in a recharging scheme. See [13, Theorem 6] for details.

Theorem 3. *In two-sided partial-observation stochastic games finite memory is sufficient (and non-elementary memory is required in general) for pure strategies both for positive and almost-sure winning for reachability objectives for player 1.*

VI. Equivalence of Randomized Action-invisible and Pure Strategies

In this section, we show that for two-sided partial-observation games, the problem of almost-sure winning with randomized action-invisible strategies is inter-reducible with the problem of almost-sure winning with pure strategies. The reductions are polynomial in the number of states in the

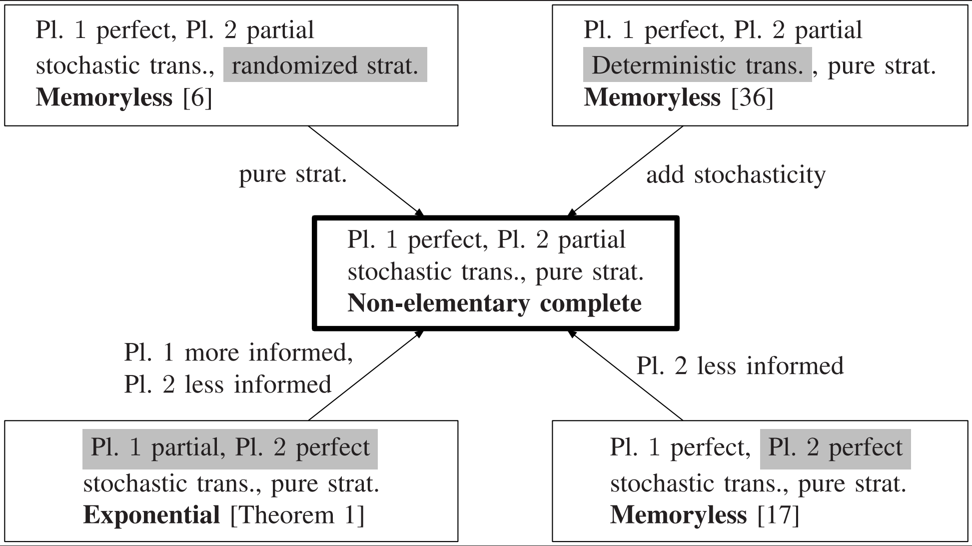

Fig. 4. The surprising non-elementary bound for memory of pure strategies in one-sided partial-observation stochastic games for player 1 perfect and player 2 partial for positive winning with reachability objectives (Theorem 2).

game (the reduction from randomized to pure strategies is exponential in the number of actions).

Reduction of randomized action-invisible strategies to pure strategies. We give a reduction for almost-sure winning for randomized action-invisible strategies to pure strategies. Given a stochastic game G we will construct another stochastic game H such that there is a randomized action-invisible almost-sure winning strategy in G iff there is a pure almost-sure winning strategy in H. The idea of the reduction is as follows: in the game H, player 1 can choose a non-empty subset $A \subseteq A_1$ of actions, and the probabilistic transition function in H is as follows: for $q \in Q$, $A \subseteq A_1$ and $b \in A_2$, we have $\delta_H(q, A, b)(q') = \frac{1}{|A|} \cdot \sum_{a \in A} \delta(q, a, b)(q')$. The observation mapping is same as in G. Further details in [13, Section 6.1] establish the following theorem and corollary.

Theorem 4. *Given a two-sided (resp. one-sided) partial-observation stochastic game G with a reachability objective we can construct in time polynomial in the size of the game and exponential in the size of the action sets a two-sided (resp. one-sided) partial-observation stochastic game H such that there exists a randomized action-invisible almost-sure winning strategy in G iff there exists a pure almost-sure winning strategy in H.*

For positive winning, randomized memoryless strategies are sufficient (both for action-visible and action-invisible) and the problem is PTIME-complete for one-sided and EXPTIME-complete for two-sided [6]. The above theorem along with Theorem 1 gives us the following corollary.

Corollary 2. *Given one-sided partial-observation stochastic games with player 1 partial and player 2 perfect, the following assertions hold for reachability objectives for player 1. (1) Exponential memory is sufficient for randomized action-invisible strategies for almost-sure winning. (2) The existence of a randomized action-invisible almost-sure winning strategy can be decided in time exponential in the state space of the game and exponential in the size of the action sets. (3) The problem of deciding the existence of a randomized action-invisible almost-sure winning strategy is EXPTIME-complete.*

Reduction of pure strategies to randomized action-invisible strategies. We present a reduction for almost-sure winning with pure strategies to randomized action-invisible strategies. Given a stochastic game G we construct another stochastic game H such that there exists a pure almost-sure winning strategy in G iff there exists a randomized almost-sure winning strategy in H. The idea of the reduction is to force player 1 to play a pure strategy in H. The game H simulates G and requires player 1 to repeat each action played (i.e., to play each action two times). Then, if player 1 uses randomization, he has to repeat the actions chosen randomly in the previous step. Since the actions are invisible, this can be achieved only if the support of the randomized actions is a singleton, i.e., the strategy is pure. Note that the reduction works for randomized strategies with actions invisible, and not when the actions are visible (details in [13, Section 6.2]).

Theorem 5. *Given a two-sided partial-observation stochastic game G with a reachability objective we can construct in time polynomial in the size of the game and size of the action sets a two-sided partial-observation stochastic game H such that there exists a pure almost-sure winning strategy in G iff there exists a randomized action-invisible almost-sure winning strategy in H.*

Belief-based strategies are not sufficient. We illustrate our reduction with the following example that shows belief-based (belief-only) randomized action-invisible strategies are not sufficient for almost-sure reachability in one-sided partial-observation games (player 1 partial and player 2 perfect), showing that a remark (without proof) of [16, p.4] and the result and construction of [25, Theorem 1] are wrong.

Example 2. *We illustrate the reduction on the example of Fig. 1. The result of the reduction is given in Fig. 2. Remember that Example 1 showed that belief-based pure strategies are not sufficient for almost-sure winning. We show that belief-based randomized strategies are not sufficient for almost-sure winning in the game of Fig. 2. First, in $\{q_1, q_2\}$ player 1 has to play pure since he has to be able to repeat the same action to avoid reaching a sink state ☹ with positive probability. Now, the argument is the same as in Example 1: playing always the same action (either a or b) in $\{q_1, q_2\}$ is not even positive winning as player 2 can choose either q_2 or q_1.* ■

Note that our reduction preserves the structure and memory of almost-sure winning strategies, hence the non-elementary lower bound given in Theorem 3 for pure strategies also holds for randomized action-invisible strategies.

Corollary 3. *For one-sided partial-observation stochastic games, with player 1 partial and player 2 perfect, belief-based randomized action-invisible strategies are not sufficient for almost-sure winning for reachability objectives. For two-sided partial-observation stochastic games, memory of non-elementary size is necessary in general for almost-sure winning for randomized action-invisible strategies for reachability objectives.*

Acknowledgement. This work was partially supported by FWF Grant No P 23499-N23, FWF NFN Grant No S11407-N23 (RiSE), ERC Start grant (279307: Graph Games), and Microsoft faculty fellows award.

References

[1] M. Abadi, L. Lamport, and P. Wolper. Realizable and unrealizable specifications of reactive systems. In *Proc. of ICALP*, LNCS 372, pages 1–17. Springer, 1989.

[2] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49:672–713, 2002.

[3] C. Baier, N. Bertrand, and M. Größer. On decision problems for probabilistic Büchi automata. In *Proc. of FoSSaCS*, LNCS 4962, pages 287–301. Springer, 2008.

[4] C. Baier, N. Bertrand, and M. Größer. The effect of tossing coins in omega-automata. In *Proc. of CONCUR*, LNCS 5710, pages 15–29. Springer, 2009.

[5] C. Baier and M. Größer. Recognizing omega-regular languages with probabilistic automata. In *Proc. of LICS*, pages 137–146, 2005.

[6] N. Bertrand, B. Genest, and H. Gimbert. Qualitative determinacy and decidability of stochastic games with signals. In *Proc. of LICS*, pages 319–328, 2009.

[7] D. Berwanger and L. Doyen. On the power of imperfect information. In *Proc. of FSTTCS*, Dagstuhl Seminar Proceedings 08004. IBFI, 2008.

[8] P. Cerný, K. Chatterjee, T. A. Henzinger, A. Radhakrishna, and R. Singh. Quantitative synthesis for concurrent programs. In *Proc. of CAV*, LNCS 6806, pages 243–259. Springer, 2011.

[9] R. Chadha, A. P. Sistla, and M. Viswanathan. On the expressiveness and complexity of randomization in finite state monitors. *Journal of the ACM*, 56:1–44, 2009.

[10] R. Chadha, A. P. Sistla, and M. Viswanathan. Power of randomization in automata on infinite strings. In *Proc. of CONCUR*, LNCS 5710, pages 229–243. Springer, 2009.

[11] R. Chadha, A. P. Sistla, and M. Viswanathan. Model checking concurrent programs with nondeterminism and randomization. In *Proc. of FSTTCS*, volume 8 of *LIPIcs*, pages 364–375, 2010.

[12] K. Chatterjee and L. Doyen. The complexity of partial-observation parity games. In *Proc. of LPAR*, LNCS 6397, pages 1–14. Springer, 2010.

[13] K. Chatterjee and L. Doyen. Partial-observation stochastic games: How to win when belief fails. *CoRR*, abs/1107.2141, July 2011.

[14] K. Chatterjee, L. Doyen, H. Gimbert, and T. A. Henzinger. Randomness for free. In *Proc. of MFCS*, LNCS 6281, pages 246–257. Springer, 2010.

[15] K. Chatterjee, L. Doyen, and T. A. Henzinger. Qualitative analysis of partially-observable Markov decision processes. In *Proc. of MFCS*, LNCS 6281, pages 258–269. Springer, 2010.

[16] 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.

[17] A. Condon. The complexity of stochastic games. *Information and Computation*, 96(2):203–224, 1992.

[18] L. de Alfaro and T. A. Henzinger. Interface automata. In *Proc. of FSE*, pages 109–120. ACM Press, 2001.

[19] L. de Alfaro, T. A. Henzinger, and O. Kupferman. Concurrent reachability games. *Theor. Comput. Sci.*, 386(3):188–217, 2007.

[20] M. De Wulf, L. Doyen, and J.-F. Raskin. A lattice theory for solving games of imperfect information. In *Proc. of HSCC*, LNCS 3927, pages 153–168. Springer, 2006.

[21] D. L. Dill. *Trace Theory for Automatic Hierarchical Verification of Speed-independent Circuits*. The MIT Press, 1989.

[22] L. Doyen and J.-F. Raskin. Antichains algorithms for finite automata. In *Proc. of TACAS*, LNCS 6015, pages 2–22. Springer, 2010.

[23] E. A. Emerson and C. Jutla. Tree automata, mu-calculus and determinacy. In *Proc. of FOCS*, pages 368–377, 1991.

[24] H. Gimbert and Y. Oualhadj. Probabilistic automata on finite words: Decidable and undecidable problems. In *Proc. of ICALP (2)*, LNCS 6199, pages 527–538. Springer, 2010.

[25] V. Gripon and O. Serre. Qualitative concurrent stochastic games with imperfect information. In *Proc. of ICALP (2)*, LNCS 5556, pages 200–211. Springer, 2009.

[26] T. A. Henzinger and P. W. Kopke. Discrete-time control for rectangular hybrid automata. *Theor. Comp. Science*, 221:369–392, 1999.

[27] A. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.

[28] D. König. *Theorie der endlichen und unendlichen Graphen*. Akademische Verlagsgesellschaft, Leipzig, 1936.

[29] H. Kress-Gazit, G. E. Fainekos, and G. J. Pappas. Temporal-logic-based reactive mission and motion planning. *IEEE Transactions on Robotics*, 25(6):1370–1381, 2009.

[30] O. Kupferman and M. Y. Vardi. Synthesis with incomplete informatio. In *Advances in Temporal Logic*, pages 109–127. Kluwer Academic Publishers, 2000.

[31] C. H. Papadimitriou and J. N. Tsitsiklis. The complexity of Markov decision processes. *Math. of Op. Research*, 12:441–450, 1987.

[32] A. Paz. *Introduction to probabilistic automata*. Academic Press, 1971.

[33] A. Pnueli and R. Rosner. On the synthesis of a reactive module. In *Proc. of POPL*, pages 179–190. ACM Press, 1989.

[34] M. O. Rabin. Probabilistic automata. *Inf. & Cont.*, 6:230–245, 1963.

[35] 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.

[36] J. H. Reif. Universal games of incomplete information. In *Proc. of STOC*, pages 288–308. ACM, 1979.

[37] J. H. Reif. The complexity of two-player games of incomplete information. *JCSS*, 29:274–301, 1984.

[38] J. H. Reif and G. L. Peterson. A dynamic logic of multiprocessing with incomplete information. In *Proc. of POPL*, pages 193–202. ACM, 1980.

[39] D. Rosenberg, E. Solan, and N. Vieille. Stochastic games with a single controller and incomplete information. *SIAM J. Control and Optimization*, 43(1):86–110, 2004.

[40] L. S. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. USA*, 39:1095–1100, 1953.

[41] S. Sorin. *A first course in zero-sum repeated games*. Springer, 2002.

[42] W. Thomas. Languages, automata, and logic. In *Handbook of Formal Languages*, volume 3, Beyond Words, chapter 7, pages 389–455. Springer, 1997.

[43] M. Tracol, C. Baier, and M. Größer. Recurrence and transience for probabilistic automata. In *Proc. of FSTTCS*, volume 4 of *LIPIcs*, pages 395–406, 2009.

[44] M. Y. Vardi. Automatic verification of probabilistic concurrent finite-state systems. In *Proc. of FOCS*, pages 327–338, 1985.

[45] M. T. Vechev, E. Yahav, and G. Yorsh. Inferring synchronization under limited observability. In *Proc. of TACAS*, LNCS 5505, pages 139–154. Springer, 2009.