IST AUSTRIA

Institute of Science and Technology

The Complexity of Partial-observation Stochastic Parity Games With Finite-memory Strategies

Krishnendu Chatterjee and Laurent Doyen and Sumit Nain and Moshe Y. Vardi

Technical Report No. IST-2013-141-v1+1
Deposited at 12 Sep 2013 08:12
http://repository.ist.ac.at/141/1/main-tech-rpt.pdf

IST Austria (Institute of Science and Technology Austria)
Am Campus 1
A-3400 Klosterneuburg, Austria

The Complexity of Partial-observation Stochastic Parity Games With Finite-memory Strategies

Krishnendu Chatterjee[†] Laurent Doyen[§] Sumit Nain[‡] Moshe Y. Vardi[‡]
[†] IST Austria
[§] CNRS, LSV, ENS Cachan
[‡] Rice University, USA

Abstract

We consider two-player partial-observation stochastic games where player 1 has partial observation and player 2 has perfect observation. The winning condition we study are ω-regular conditions specified as parity objectives. The qualitative analysis problem given a partial-observation stochastic game and a parity objective asks whether there is a strategy to ensure that the objective is satisfied with probability 1 (resp. positive probability). While the qualitative analysis problems are known to be undecidable even for very special cases of parity objectives, they were shown to be decidable in 2EXPTIME under finite-memory strategies. We improve the complexity and show that the qualitative analysis problems for partial-observation stochastic parity games under finite-memory strategies are EXPTIME-complete; and also establish optimal (exponential) memory bounds for finite-memory strategies required for qualitative analysis.

1 Introduction

Partial-observation stochastic games. Partial-observation stochastic games are played between two players (player 1 and player 2) on a graph with finite state space. The game is played for infinitely many rounds where in each round either player 1 chooses a move or player 2 chooses a move, and the successor state is determined by a probabilistic transition function. Player 1 has partial observation where the state space is partitioned according to observations that she can observe i.e., given the current state, the player can only view the observation of the state (the partition the state belongs to), but not the precise state. Player 2, the adversary to player 1, has perfect observation and can observe the precise state.

The class of ω-regular objectives. An objective specifies the desired set of behaviors (or paths) for the controller. In verification and control of stochastic systems an objective is typically an ω-regular set of paths. The class of ω-regular languages extends classical regular languages to infinite strings, and provides a robust specification language to express all commonly used specifications [18]. In a parity objective, every state of the game is mapped to a non-negative integer priority and the goal is to ensure that the minimum priority visited infinitely often is even. Parity objectives are a canonical way to define such ω-regular specifications.

Qualitative analysis. Given a partial-observation stochastic game with a parity objective and a start state, the *qualitative analysis* asks whether the objective can be ensured with probability 1 (*almost-sure winning*) or positive probability (*positive winning*).

Known results and our contribution. The qualitative analysis problems for stochastic games with parity objectives are undecidable [1]. However, in many practical applications the more relevant question is the existence of finite-memory strategies. The qualitative analysis problems for partial-observation stochastic parity games were shown to be decidable with 2EXPTIME complexity for finite-memory strategies [16]; and the exact complexity of the problems were open which we settle in this work. Our contributions are as follows: for the qualitative analysis problems for partial-observation stochastic parity games under finite-memory strategies we show that (i) the problems are EXPTIME-complete; and (ii) if there is a finite-memory almost-sure (resp. positive) winning strategy, then there is a strategy that uses at most exponential memory (matching the exponential lower bound known for the simpler case of reachability and safety objectives). Thus we establish both optimal computational and strategy complexity results.

Related works. The undecidability of the qualitative analysis problem for partial-observation stochastic parity games with infinite-memory strategies follows from [1]. For partially observable Markov decision processes (POMDPs), which is a special case of partial-observation stochastic games where player 2 does not have any choices, the qualitative analysis problem for parity objectives with finite-memory strategies was shown to be EXPTIME-complete [3]. For partial-observation stochastic games the almost-sure winning problem was shown to be EXPTIME-complete for Büchi objectives (both for finite-memory and infinite-memory strategies) [7, 4]. Finally, for partial-observation stochastic parity games the almost-sure winning problem under finite-memory strategies was shown to be decidable in 2EXPTIME in [16].

2 Partial-observation Stochastic Parity Games

We consider partial-observation stochastic parity games where player 1 has partial observation and player 2 has perfect observation. We will consider parity objectives, and for almost-sure winning under finite-memory strategies for player 1 present a polynomial reduction to sure winning in three-player parity games where player 1 has partial observation, player 3 has perfect observation and is helpful towards player 1, and player 2 has perfect observation and is adversarial to player 1. A similar reduction also works for positive winning. We will then show how to solve the sure-winning problem for three-player games using alternating parity tree automata. Thus the steps are as follows:

1. Reduction of partial-observation stochastic parity games for almost-sure winning with finite-memory strategies to three-player parity games sure-winning problem (with player 1 partial, other two perfect, player 1 and player 3 existential, and player 2 adversarial).
2. Solving the sure winning problem for three-player parity games using alternating parity tree automata.

In this section we present the details of the first step. The second step is given in the following section.

2.1 Basic definitions

We start with basic definitions related to partial-observation stochastic parity games.

Partial-observation stochastic games. We will consider slightly different notation (though equivalent) to the classical definitions, but the slightly different notation helps for more elegant and explicit reduction. We consider partial-observation stochastic games as a tuple $G = (S_1, S_2, S_P, A_1, \delta, E, \mathcal{O}, \mathsf{obs})$ as follows: $S = S_1 \cup S_2 \cup S_P$ is the state space partitioned into player-1 states (S_1), player-2 states (S_2), and probabilistic states (S_P); and A_1 is a finite set of actions for player 1. Since player 2 has perfect observation, she will choose edges instead of actions. The transition function is as follows: $\delta : S_1 \times A_1 \to S_2$ that given a player-1 state in S_1 and an action in A_1 gives the next state in S_2 (which belongs to player 2); and $\delta : S_P \to \mathcal{D}(S_1)$ given a probabilistic state gives the probability distribution over the set of player-1 states. The set of edges is as follows: $E = \{(s, t) \mid s \in S_P, t \in S_1, \delta(s)(t) > 0\} \cup E'$, where $E' \subseteq S_2 \times S_P$. The observation set $\mathcal{O}$ and observation mapping obs are standard, i.e., $\mathsf{obs} : S \to \mathcal{O}$. Note that player 1 plays after every three steps (every move of player 1 is followed by a move of player 2, then a probabilistic choice). In other words, first player 1 chooses an action, then player 2 chooses an edge, and then there is a probability distribution over states where player 1 again chooses and so on.

Three player non-stochastic turn-based games. We consider three-player partial-observation (non-stochastic turn-based) games as a tuple $G = (S_1, S_2, S_3, A_1, \delta, E, \mathcal{O}, \mathsf{obs})$ as follows: S is the state space partitioned into player-1 states (S_1), player-2 states (S_2), and player-3 states (S_3); and A_1 is a finite set of actions for player 1. The transition function is as follows: $\delta : S_1 \times A_1 \to S_2$ that given a player-1 state in S_1 and an action in A_1 gives the next state (which belongs to player 2). The set of edges is as follows: $E \subseteq (S_2 \cup S_3) \times S$. Hence in these games player 1 chooses an action, and the other players have perfect observation and choose edges. We will only consider the sub-class where player 1 plays in every k-steps, for a fixed k. The observation set $\mathcal{O}$ and observation mapping obs are again standard.

Plays and strategies. A *play* in a partial-observation stochastic game is an infinite sequence of states $s_0 s_1 s_2 \ldots$ such that the following conditions hold for all $i \geq 0$: (i) if $s_i \in S_1$, then there exists $a_i \in A_1$ such that $s_{i+1} = \delta(s_i, a_i)$; and (ii) if $s_i \in (S_2 \cup S_P)$, then $(s_i, s_{i+1}) \in E$. The function obs is extended to sequences $\rho = s_0 \ldots s_n$ of states in the natural way, namely $\mathsf{obs}(\rho) = \mathsf{obs}(s_0) \ldots \mathsf{obs}(s_n)$. A strategy for a player is a recipe to extend the prefix of a play. Formally, player-1 strategies are functions $\sigma : S^* \cdot S_1 \to A_1$; and player-2 (and analogously player-3 strategies) are functions: $\pi : S^* \cdot S_2 \to S$ such that for all $w \in S^*$ and $s \in S_2$ we have $(s, \pi(w \cdot s)) \in E$. We will consider only observation-based strategies for player 1, i.e., for two play prefixes ρ and ρ' if the corresponding observation sequences match ($\mathsf{obs}(\rho) = \mathsf{obs}(\rho')$), then the strategy must choose the same action ($\sigma(\rho) = \sigma(\rho')$); and the other players have all strategies. The notations for three-player games are similar.

Finite-memory strategies. A player-1 strategy uses *finite-memory* if it can be encoded by a deterministic transducer $\langle \mathsf{M}, m_0, \sigma_u, \sigma_n \rangle$ where M is a finite set (the memory of the strategy), $m_0 \in \mathsf{M}$ is the initial memory value, $\sigma_u : \mathsf{M} \times \mathcal{O} \to \mathsf{M}$ is the memory-update function, and $\sigma_n : \mathsf{M} \times \mathcal{O} \to A_1$ is the next-move function. The *size* of the strategy is the number $|\mathsf{M}|$ of memory values. If the current observation is o, and the current memory value is m, then the strategy chooses the next action $\sigma_n(m, o)$, and the memory is updated to $\sigma_u(m, o)$. Formally, $\langle \mathsf{M}, m_0, \sigma_u, \sigma_n \rangle$ defines the strategy σ such that $\sigma(\rho \cdot q) = \sigma_n(\widehat{\sigma}_u(m_0, \mathsf{obs}(\rho)), \mathsf{obs}(s))$ for all $\rho \in S^*$ and $s \in S_1$, where $\hat{\sigma}_u$ extends σ_u to sequences of observations as expected. This definition extends to infinite-memory strategies by dropping the assumption that the set M is finite.

Parity objectives. An *objective* for Player 1 in G is a set $\varphi \subseteq S^\omega$ of infinite sequences of states. A play ρ *satisfies* the objective φ if $\rho \in \varphi$. For a play $\rho = s_0 s_1 \ldots$ we denote by $\mathsf{Inf}(\rho)$ the set of states that occur infinitely often in ρ, that is, $\mathsf{Inf}(\rho) = \{s \mid s_j = s \text{ for infinitely many } j\text{'s}\}$. For $d \in \mathbb{N}$, let $p : S \to \{0, 1, \ldots, d\}$ be a *priority function*, which maps each state to a nonnegative integer priority. The *parity* objective $\mathsf{Parity}(p)$ requires that the minimum priority that occurs infinitely often be even. Formally, $\mathsf{Parity}(p) = \{\rho \mid \min\{p(s) \mid s \in \mathsf{Inf}(\rho)\} \text{ is even}\}$. Parity objectives are a canonical way to express ω-regular objectives [18].

Almost-sure winning and positive winning. An *event* is a measurable set of plays. For a partial-observation stochastic game, given strategies σ and π for the two players, the probabilities of events are uniquely defined [19]. For a parity objective $\mathsf{Parity}(p)$, we denote by $\mathbb{P}_s^{\sigma,\pi}(\mathsf{Parity}(p))$ the probability that $\mathsf{Parity}(p)$ is satisfied by the play obtained from the starting state s when the strategies σ and π are used. The *almost-sure* (resp. *positive*) winning problem under finite-memory strategies asks, given a partial-observation stochastic game, a parity objective $\mathsf{Parity}(p)$, and a starting state s, whether there exists a finite-memory observation-based strategy σ for player 1 such that against all strategies π for player 2 we have $\mathbb{P}_s^{\sigma,\pi}(\mathsf{Parity}(p)) = 1$ (resp. $\mathbb{P}_s^{\sigma,\pi}(\mathsf{Parity}(p)) > 0$). The almost-sure and positive winning problems are also referred to as the qualitative analysis problems for stochastic games.

Sure winning in three player games. In three player games once the starting state s and strategies σ, π, and τ of the three players are fixed we obtain a unique play, which we denote as $\rho_s^{\sigma,\pi,\tau}$. In three player games we consider the following *sure* winning problem: given a parity objective $\mathsf{Parity}(p)$, sure winning is ensured if there exists a finite-memory observation-based strategy σ for player 1, such that in the two-player perfect-observation game obtained after fixing σ, player 3 can esnure the parity objective against all strategies of player 2. Formally, the sure winning problem asks whether there exist a finite-memory observation-based strategy σ for player 1 and a strategy τ for player 3, such that for all strategies π for player 2 we have $\rho_s^{\sigma,\pi,\tau} \in \mathsf{Parity}(p)$.

REMARK 1. *We remark that for the model of partial-observation stochastic games studied in literature the* *two players simultaneously choose actions, and a probabilistic transition function determine the probability distribution of the next state. In our model, the game is turn-based and the probability distribution is chosen only in probabilistic states. However, it follows from the results of [5] that the models are equivalent: by the results of [5, Section 3.1] the interaction of the players and probability can be separated without loss of generality; and [5, Theorem 4] shows that in presence of partial observation, concurrent games can be reduced to turn-based games in polynomial time.*

REMARK 2. *In this work we only consider pure strategies. In partial-observation games, randomized strategies are also relevant as they are more powerful than pure strategies. However, for finite-memory strategies the almost-sure and positive winning problem for randomized strategies can be reduced in polynomial time to the problem for finite-memory pure strategies [4, 16]. Hence without loss of generality we only consider pure strategies.*

2.2 Reduction of partial-observation stochastic games to three player games In this section we present a polynomial-time reduction for the almost-sure winning problem in partial-observation stochastic parity games to the sure winning problem in three player parity games.

Reduction. Let us denote by $[d]$ the set $\{0, 1, \ldots, d\}$. Given a partial-observation stochastic parity game graph $G = (S_1, S_2, S_P, A_1, \delta, E, \mathcal{O}, \mathsf{obs})$ with a parity objective defined by priority function $p : S \to [d]$ we construct a 3-player game graph $\overline{G} = (\overline{S}_1, \overline{S}_2, \overline{S}_3, A_1, \overline{\delta}, \overline{E}, \mathcal{O}, \overline{\mathsf{obs}})$ together with priority function $\overline{p}$. The construction is specified as follows.

1. For every nonprobabilistic state $s \in S_1 \cup S_2$, there is a corresponding state $\overline{s} \in \overline{S}$ such that
   - $\overline{s} \in \overline{S}_1$ if $s \in S_1$, else $\overline{s} \in \overline{S}_2$;
   - $\overline{p}(\overline{s}) = p(s)$ and $\overline{\mathsf{obs}}(\overline{s}) = \mathsf{obs}(s)$;
   - $\overline{\delta}(\overline{s}, a) = \overline{t}$ where $t = \delta(s, a)$, for $s \in S_1$ and $a \in A_1$; and
   - $(\overline{s}, \overline{t}) \in \overline{E}$ iff $(s, t) \in E$, for $s \in S_2$.
2. Every probabilistic state $s \in S_P$ is replaced by the gadget shown in Figure 1 and Figure 2. In the figure, square-shaped states are player-2 states (in $\overline{S}_2$), and circle-shaped (or ellipsoid-shaped) states are player-3 states (in $\overline{S}_3$). Formally, from the state $\overline{s}$ with priority $p(s)$ and observation $\mathsf{obs}(s)$ (i.e., $\overline{p}(\overline{s}) = p(s)$ and $\overline{\mathsf{obs}}(\overline{s}) = \mathsf{obs}(s)$) the players play the following three-step game in $\overline{G}$.
   - First, in state $\overline{s}$ player 2 chooses a successor $(\widetilde{s}, 2k)$, for $2k \in \{0, 1, \ldots, p(s) + 1\}$.
   - For every state $(\widetilde{s}, 2k)$, we have $\overline{p}((\widetilde{s}, 2k)) = p(s)$ and $\overline{\mathsf{obs}}((\widetilde{s}, 2k)) = \mathsf{obs}(s)$. For $k \geq 1$, in state $(\widetilde{s}, 2k)$ player 3 chooses between two successors: state $(\widehat{s}, 2k-1)$ with priority $2k-1$ and same observation as s, or state $(\widehat{s}, 2k)$ with priority $2k$ and same observation as s, (i.e., $\overline{p}((\widehat{s}, 2k-1)) = 2k-1$, $\overline{p}((\widehat{s}, 2k)) = 2k$, and $\overline{\mathsf{obs}}((\widehat{s}, 2k-1)) = \overline{\mathsf{obs}}((\widehat{s}, 2k)) = \mathsf{obs}(s)$). The state $(\widetilde{s}, 0)$ has only one successor $(\widehat{s}, 0)$, with $\overline{p}((\widehat{s}, 0)) = 0$ and $\overline{\mathsf{obs}}((\widehat{s}, 0)) = \mathsf{obs}(s)$.
   - Finally, in each state $(\widehat{s}, k)$ the choice is between all states $\overline{t}$ such that $(s, t) \in E$, and it belongs to player 3 (i.e., in $\overline{S}_3$) if k is odd, and to player 2 (i.e., in $\overline{S}_2$) if k is even. Note that every state in the gadget has the same observation as the original state.

We denote by $\overline{G} = \mathrm{Tr}_{\mathrm{as}}(G)$ the 3-player game, where player 1 has partial-observation, and both player 2 and player 3 have perfect-observation, obtained from a partial-observation stochastic game. Also observe that in $\overline{G}$ there are exactly four steps between two player 1 moves.

Observation sequence mapping. Note that since in our partial-observation games first player 1 plays, then player 2, followed by probabilistic states, repeated ad infinitum, wlog, we can assume that for every observation

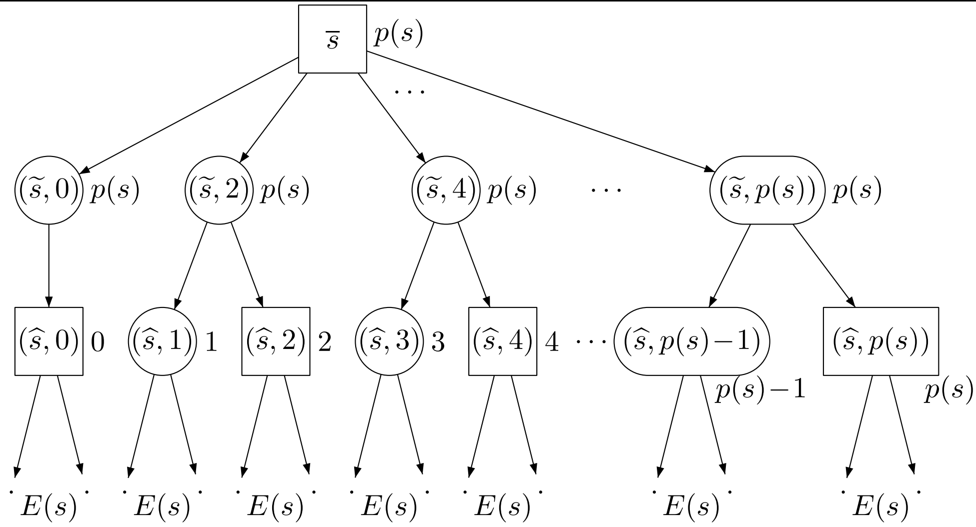

Figure 1: Reduction gadget when $p(s)$ is even.

$o \in \mathcal{O}$ we have either (i) $\mathsf{obs}^{-1}(o) \subseteq S_1$; or (ii) $\mathsf{obs}^{-1}(o) \subseteq S_2$; or (i) $\mathsf{obs}^{-1}(o) \subseteq S_P$. Thus we partition the observations as $\mathcal{O}_1$, $\mathcal{O}_2$, and $\mathcal{O}_P$. Given an observation sequence $\kappa = o_0 o_1 o_2 \ldots o_n$ in G corresponding to a finite prefix of a play, we inductively define the sequence $\overline{\kappa} = \overline{h}(\kappa)$ in $\overline{G}$ as follows: (i) $\overline{h}(o_0) = o_0$ if $o_0 \in \mathcal{O}_1 \cup \mathcal{O}_2$, else $o_0 o_0 o_0$; (ii) $\overline{h}(o_0 o_1 \ldots o_n) = \overline{h}(o_0 o_1 \ldots o_{n-1}) o_n$ if $o_n \in \mathcal{O}_1 \cup \mathcal{O}_2$, else $\overline{h}(o_0 o_1 \ldots o_{n-1}) o_n o_n o_n$. Intuitively the mapping takes care of the two extra step of the gadgets introduced for probabilistic states. The mapping is a bijection, and hence given an observation sequence $\overline{\kappa}$ of a play prefix in $\overline{G}$ we consider the inverse play prefix $\kappa = \overline{h}^{-1}(\overline{\kappa})$ such that $\overline{h}(\kappa) = \overline{\kappa}$.

Strategy mapping. Given an observation-based strategy $\overline{\sigma}$ in $\overline{G}$ we consider a strategy $\sigma = \mathrm{Tr}_{\mathrm{as}}(\overline{\sigma})$ as follows: for an observation sequence κ corresponding to a play prefix in G we have $\sigma(\kappa) = \overline{\sigma}(\overline{h}(\kappa))$. The strategy σ is observation-based (since $\overline{\sigma}$ is observation-based). The inverse mapping ${\mathrm{Tr}_{\mathrm{as}}}^{-1}$ of strategies from G to $\overline{G}$ is analogous. Note that for σ in G we have $\mathrm{Tr}_{\mathrm{as}}({\mathrm{Tr}_{\mathrm{as}}}^{-1}(\sigma)) = \sigma$. Let $\overline{\sigma}$ be a finite-memory strategy with memory M for player 1 in the game $\overline{G}$. The strategy $\overline{\sigma}$ can be considered as a memoryless strategy, denoted as $\overline{\sigma}^* = \mathsf{MemLess}(\overline{\sigma})$, in $\overline{G} \times \mathsf{M}$ (the synchronous product of $\overline{G}$ with M). Given a strategy (pure memoryless) $\overline{\pi}$ for player 2 in the 2-player game $\overline{G} \times \mathsf{M}$, a strategy $\pi = \mathrm{Tr}_{\mathrm{as}}(\overline{\pi})$ in the partial-observation stochastic game $G \times \mathsf{M}$ is defined as follows:

$$\pi((s,m)) = (t,m'), \text{ if and only if } \overline{\pi}((\overline{s},m)) = (\overline{t},m'); \text{ for all } s \in S_2.$$

End component and the key property. Given an MDP, a set U is an end component in the MDP if the sub-graph induced by U is strongly connected, and for all probabilistic states in U all out-going edges end up in U (i.e., U is closed for probabilistic states). The key property about MDPs that will be used in our proofs is a result established by [8, 9] that given an MDP, for all strategies, with probability 1 the set of states visited inifinitely often is an end component. The key property will allow us to analyze end components of MDPs and from properties of the end component conclude properties about all strategies.

The key lemma. We are now ready to present our main lemma that establishes the correctness of the reduction. Since the proof of the lemma is long we will split the proof into two parts.

LEMMA 2.1. *Given a partial-observation stochastic parity game G with parity objective* $\mathsf{Parity}(p)$*, let $\overline{G} = \mathrm{Tr}_{\mathrm{as}}(G)$ be the 3-player game with the modified parity objective* $\mathsf{Parity}(\overline{p})$ *obtained by our reduction. Consider*

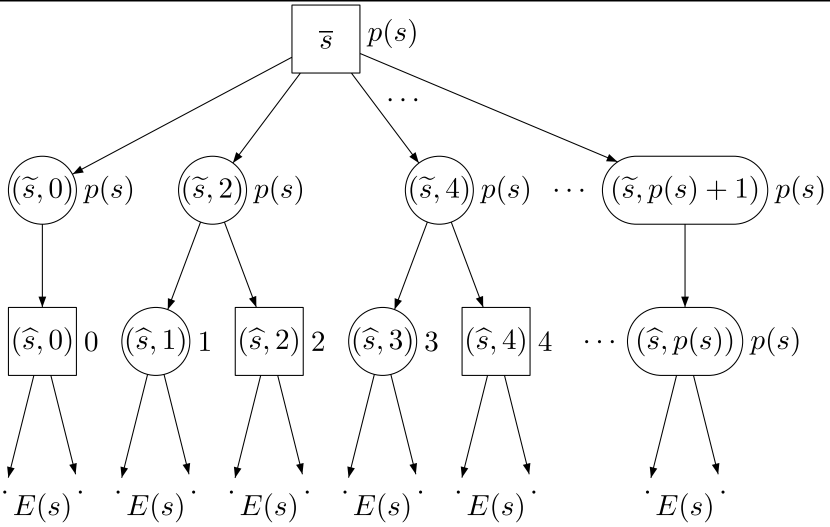

Figure 2: Reduction gadget when $p(s)$ is odd.

a finite-memory strategy $\overline{\sigma}$ with memory M *for player 1 in $\overline{G}$. Let us denote by $\overline{G}_{\overline{\sigma}}$ the perfect-observation two-player game played over $\overline{G} \times$ M by player 2 and player 3 after fixing the strategy $\overline{\sigma}$ for player 1. Let*

$$\overline{U}_1^{\overline{\sigma}} = \{(\overline{s}, m) \in \overline{S} \times \mathsf{M} \mid \textit{player 3 has a sure winning strategy for the objective } \mathsf{Parity}(\overline{p}) \textit{ from } (\overline{s}, m) \textit{ in } \overline{G}_{\overline{\sigma}}\};$$

and let $\overline{U}_2^{\overline{\sigma}} = (\overline{S} \times \mathsf{M}) \setminus \overline{U}_1^{\overline{\sigma}}$ be the set of sure winning states for player 2 in $\overline{G}_{\overline{\sigma}}$. Consider the strategy $\sigma = \mathrm{Tr}_{\mathrm{as}}(\overline{\sigma})$, and the sets $U_1^{\sigma} = \{(s, m) \in S \times \mathsf{M} \mid (\overline{s}, m) \in \overline{U}_1^{\overline{\sigma}}\}$; and $U_2^{\sigma} = (S \times \mathsf{M}) \setminus U_1^{\sigma}$. The following assertions hold.

1. *For all $(s, m) \in U_1^{\sigma}$, for all strategies π of player 2 we have $\mathbb{P}_{(s,m)}^{\sigma,\pi}(\mathsf{Parity}(p)) = 1$.*
2. *For all $(s, m) \in U_2^{\sigma}$, there exists a strategy π of player 2 such that $\mathbb{P}_{(s,m)}^{\sigma,\pi}(\mathsf{Parity}(p)) < 1$.*

We first present the proof for part 1 and then for part 2.

Proof. [(of Lemma 2.1: part 1).] Consider a finite-memory strategy $\overline{\sigma}$ for player 1 with memory M in the game $\overline{G}$. Once the strategy $\overline{\sigma}$ is fixed we obtain the two-player finite-state perfect-observation game $\overline{G}_{\overline{\sigma}}$ (between player 3 and the adversary player 2). Recall the sure winning sets

$$\overline{U}_1^{\overline{\sigma}} = \{(\overline{s}, m) \in \overline{S} \times \mathsf{M} \mid \text{player 3 has a sure winning strategy for the objective } \mathsf{Parity}(\overline{p}) \text{ from } (\overline{s}, m) \text{ in } \overline{G}_{\overline{\sigma}}\}$$

for player 3, and $\overline{U}_2^{\overline{\sigma}} = (\overline{S} \times \mathsf{M}) \setminus \overline{U}_1^{\overline{\sigma}}$ for player 2, respectively, in $\overline{G}_{\overline{\sigma}}$. Let $\sigma = \mathrm{Tr}_{\mathrm{as}}(\overline{\sigma})$ be the corresponding strategy in G. We denote by $\overline{\sigma}^* = \mathsf{MemLess}(\overline{\sigma})$ and σ^* the corresponding memoryless strategies of $\overline{\sigma}$ in $\overline{G} \times$ M and σ in $G \times$ M, respectively. We will show that all states in U_1^{σ} are almost-sure winning, i.e., given σ, for all $(s, m) \in U_1^{\sigma}$, for all strategies π for player 2 in G we have $\mathbb{P}_{(s,m)}^{\sigma,\pi}(\mathsf{Parity}(p)) = 1$ (recall $U_1^{\sigma} = \{(s, m) \in S \times \mathsf{M} \mid (\overline{s}, m) \in \overline{U}_1^{\overline{\sigma}}\}$). We will also consider explicitly the MDP $(G \times \mathsf{M} \upharpoonright U_1^{\sigma})_{\sigma^*}$ to analyze strategies of player 2 on the synchronous product, i.e., we consider the player-2 MDP obtained after fixing the memoryless strategy σ^* in $G \times$ M, and then restrict the MDP to the set U_1^{σ}.

Two key components. The proof will have two key components. First, we argue that all end components in the MDP restricted to U_1^{σ} are winning for player 1 (have min priority even). Second we argue that given the starting state (s,m) is in U_1^σ, almost-surely the set of states visited inifnitely often is an end component in U_1^σ against all strategies of player 2. This two key components establish the desired result.

Winning end components. Our first goal is to show that every end component C in the player-2 MDP $(G \times \mathsf{M} \upharpoonright U_1^\sigma)_{\sigma^*}$ is winning for player 1 for the parity objective, i.e., the minimum priority of C is even. We argue that if there is an end component C in $(G \times \mathsf{M} \upharpoonright U_1^\sigma)_{\sigma^*}$ that is winning for player 2 for the parity objective (i.e., minimum priority of C is odd), then against any memoryless player-3 strategy $\overline{\tau}$ in $\overline{G}_{\overline{\sigma}}$, player 2 can construct a cycle in the game $(\overline{G} \times \mathsf{M} \upharpoonright \overline{U_1^{\sigma}})_{\overline{\sigma}^*}$ that is winning for player 2 (i.e., minimum priority of the cycle is odd) (note that given the strategy $\overline{\sigma}$ is fixed, we have finite-state perfect-observation parity games, and hence in the enlarged game we can restrict ourselves to memoryless strategies for player 3). This will give a contradiction because player 3 has a sure winning strategy from the set $\overline{U_1^{\sigma}}$ in the 2-player parity game $\overline{G}_{\overline{\sigma}}$. Towards contradiction, let C be an end component in $(G \times \mathsf{M} \upharpoonright U_1^\sigma)_{\sigma^*}$ that is winning for player 2, and let its minimum odd priority be $2r-1$, for some $r \in \mathbb{N}$. Then there is a memoryless strategy π' for player 2 in the MDP $(G \times \mathsf{M} \upharpoonright U_1^\sigma)_{\sigma^*}$ such that C is a bottom scc (or a terminal scc) in the Markov chain graph of $(G \times \mathsf{M} \upharpoonright U_1^\sigma)_{\sigma^*,\pi'}$. Let $\overline{\tau}$ be a memoryless for player 3 in $(\overline{G} \times \mathsf{M} \upharpoonright \overline{U_1^{\sigma}})_{\overline{\sigma}^*}$. Given $\overline{\tau}$ for player 3 and strategy π' for player 2 in $G \times \mathsf{M}$, we construct a strategy $\overline{\pi}$ for player 2 in the game $(\overline{G} \times \mathsf{M} \upharpoonright \overline{U_1^{\sigma}})_{\overline{\sigma}^*}$ as follows. For a player-2 state in C, the strategy $\overline{\pi}$ follows the strategy π', i.e., for a state $(s,m) \in C$ with $s \in S_2$ we have $\overline{\pi}((\overline{s},m)) = (\overline{t},m')$ where $(t,m') = \pi'((s,m))$. For a probabilistic state in C we define the strategy as follows (i.e., we now consider a state $(s,m) \in C$ with $s \in S_P$):

- if for some successor state $((\widetilde{s},2\ell),m')$ of $(\overline{s},m)$, the player-3 strategy $\overline{\tau}$ chooses a successor $((\widehat{s},2\ell-1),m'') \in C$ at the state $((\widetilde{s},2\ell),m')$, for $\ell < r$, then the strategy $\overline{\pi}$ chooses at state $(\overline{s},m)$ the successor $((\widetilde{s},2\ell),m')$; and

- otherwise the strategy $\overline{\pi}$ chooses at state $(\overline{s},m)$ the successor $((\widetilde{s},2r),m')$, and at $((\widehat{s},2r),m'')$ it chooses a successor shortening the distance (i.e., chooses a successor with smaller breadth-first-search distance) to a fixed state $(\overline{s}^*,m)$ of priority $2r-1$ of C (such a state (s^*,m) exists in C since C is strongly connected and has minimum priority $2r-1$); and for the fixed state of priority $2r-1$ the strategy chooses a successor $(\overline{s},m)$ such that $(s,m) \in C$.

Consider an arbitrary cycle in the subgraph $(\overline{G} \times \mathsf{M} \upharpoonright \overline{C})_{\overline{\sigma},\overline{\pi},\overline{\tau}}$ where $\overline{C}$ is the set of states in the gadgets of states in C. There are two cases.

- If there is at least one state $((\widehat{s},2\ell-1),m)$, with $\ell \leq r$ on the cycle, then the minimum priority on the cycle is odd, as even priorities smaller than $2r$ are not visited by the construction as C does not contain states of even priorities smaller than $2r$.

- Otherwise, in all states choices shortening the distance to the state with priority $2r-1$ are taken and hence the cycle must contatin a priority $2r-1$ state and all other priorities on the cycle are $\geq 2r-1$, so $2r-1$ is the minimum priority on the cycle.

Hence a winning end component for player 2 in the MDP contradicts that player 3 has a sure winning strategy in $\overline{G}_{\overline{\sigma}}$ from $\overline{U_1^{\sigma}}$. Thus it follows that all end components are winning for player 1 in $(G \times \mathsf{M} \upharpoonright U_1^\sigma)_{\sigma^*}$.

Almost-sure reachability to winning end-components. Finally, we consider the probability of staying in U_1^σ. For every probabilistic state $(s,m) \in (S_P \times \mathsf{M}) \cap U_1^\sigma$, all of its successors must be in U_1^σ. Otherwise, player 2 in the state $(\overline{s},m)$ of the game $\overline{G}_{\overline{\sigma}}$ can choose the successor $(\widetilde{s},0)$ and then a successor to its winning set $\overline{U_2^{\sigma}}$. This will again contradict the assumption that $(\overline{s},m)$ belong to the sure winning states $\overline{U_1^{\sigma}}$ for player 3 in $\overline{G}_{\overline{\sigma}}$. Similarly, for every state $(s,m) \in (S_2 \times \mathsf{M}) \cap U_1^\sigma$ we must have all its successors are in U_1^σ. For all states $(s,m) \in (S_1 \times \mathsf{M}) \cap U_1^\sigma$, the strategy σ chooses a successor in U_1^σ. Hence for all strategies π, for all states

$(s,m) \in U_1^\sigma$, the objective $\mathsf{Safe}(U_1^\sigma)$ is ensured almost-surely (in fact surely), and hence with probability 1 the set of states visited infinitely often is an end component in U_1^σ (by key property of MDPs). Since every end component in $(G \times \mathsf{M} \upharpoonright U_1^\sigma)_{\sigma^*}$ has even minimum priority, it follows that the strategy σ is an almost-sure winning strategy for the parity objective $\mathsf{Parity}(p)$ for player 1 from all states $(s,m) \in U_1^\sigma$. This concludes the proof for first part of the lemma. ∎

We now present the proof for the second part.

Proof. [(of Lemma 2.1:part 2).] Consider a memoryless sure winning strategy $\overline{\pi}$ for player 2 in $\overline{G}_{\overline{\sigma}}$ from the set $\overline{U}_2^{\overline{\sigma}}$. Let us consider the strategies $\sigma = \mathrm{Tr}_{\mathrm{as}}(\overline{\sigma})$ and $\pi = \mathrm{Tr}_{\mathrm{as}}(\overline{\pi})$, and consider the Markov chain $G_{\sigma,\pi}$. Our proof will show the following two properties to establish the claim: (1) in the Markov $G_{\sigma,\pi}$ all bottom sccs (the recurrent classes) in U_2^σ have odd minimum priority; and (2) from all states in U_2^σ some recurrent class in U_2^σ is reached with positive probability. This will establish the desired result of the lemma.

No winning bottom scc for player 1 in U_2^σ. Assume towards contradiction that there is a bottom scc C contained in U_2^σ in the Markov chain $G_{\sigma,\pi}$ such that the minimum priority in C is even. From C we will construct a winning cycle (minimum priority is even) in $\overline{U}_2^{\overline{\sigma}}$ for player 3 in the game $\overline{G}_{\overline{\sigma}}$ given the strategy $\overline{\pi}$. This will contradict that $\overline{\pi}$ is a sure winning strategy for player 2 from $\overline{U}_2^{\overline{\sigma}}$ in $\overline{G}_{\overline{\sigma}}$. Let the minimum priority of C be $2r$ for some $r \in \mathbb{N}$. The idea is similar to the construction of part 1. Given C, and the strategies $\overline{\sigma}$ and $\overline{\pi}$, we construct a strategy $\overline{\tau}$ for player 3 in $\overline{G}$ as follows: For a probabilistic state (s,m) in C:

- if $\overline{\pi}$ chooses a state $((\widetilde{s}, 2\ell-2), m')$, with $\ell \leq r$, then $\overline{\tau}$ chooses the successor $((\widehat{s}, 2\ell-2), m')$;
- otherwise $\ell > r$ (i.e., $\overline{\pi}$ chooses a state $((\widetilde{s}, 2\ell-2), m')$ for $\ell > r$), then $\overline{\tau}$ chooses the state $((\widehat{s}, 2\ell-1), m')$, and then a successor to shorten the distance to a fixed state with priority $2r$ (such a state exists in C); and for the fixed state of priority $2r$, the strategy $\overline{\tau}$ chooses a successor in C.

Similar to the proof of part 1, we argue that we obtain a cycle with minimum even priority in the graph $(\overline{G} \times \mathsf{M} \upharpoonright \overline{U}_2^{\overline{\sigma}})_{\overline{\sigma},\overline{\pi},\overline{\tau}}$. Consider an arbitrary cycle in the subgraph $(\overline{G} \times \mathsf{M} \upharpoonright \overline{C})_{\overline{\sigma},\overline{\pi},\overline{\tau}}$ where $\overline{C}$ is the set of states in the gadgets of states in C. There are two cases.

- If there is at least one state $((\widehat{s}, 2\ell-2), m)$, with $\ell \leq r$ on the cycle, then the minimum priority on the cycle is even, as odd priorities strictly smaller than $2r+1$ are not visited by the construction as C does not contain states of odd priorities strictly smaller than $2r+1$.
- Otherwise, in all states choices shortening the distance to the state with priority $2r$ are taken and hence the cycle must contatin a priority $2r$ state and all other priorities on the cycle are $\geq 2r$, so $2r$ is the minimum priority on the cycle.

Thus we obtain cycles winning for player 3, and this contradicts that $\overline{\pi}$ is a sure winning strategy for player 2 from $\overline{U}_2^{\overline{\sigma}}$. Thus it follows that all recurrent classes in U_2^σ in the Markov chain $G_{\sigma,\pi}$ are winning for player 2.

Not almost-sure reachability to U_1^σ. We now argue that given σ and π there exists no state in U_2^σ such that U_1^σ is reached almost-surely. This would ensure that from all states in U_2^σ some recurrent class in U_2^σ is reached with positive probability and establish the desired claim since we have already shown that all recurrent classes in U_2^σ are winning for player 2. Given σ and π, let $X \subseteq U_2^\sigma$ be the set of states such the set U_1^σ is reached almost-surely from X, and assume towards contradiction that X is non-empty. This implies that from every state in X, in the Markov chain $G_{\sigma,\pi}$, there is a path to the set U_1^σ, and from all states in X the successors are in X. We construct a strategy $\overline{\tau}$ in the 3-player game $\overline{G}_{\overline{\sigma}}$ against strategy $\overline{\pi}$ exactly as the strategy constructed for winning bottom scc, with the following difference: instead of shortening distance the a fixed state of priority $2r$ (as for winning bottom scc's), in this case the strategy $\overline{\tau}$ shortens distance to $\overline{U}_1^{\overline{\sigma}}$. Formally, given X, the strategies $\overline{\sigma}$ and $\overline{\pi}$, we construct a strategy $\overline{\tau}$ for player 3 in $\overline{G}$ as follows: For a probabilistic state (s,m) in X:

- if $\overline{\pi}$ chooses a state $((\widetilde{s}, 2\ell), m')$, with $\ell \geq 1$, then $\overline{\tau}$ chooses the state $((\widehat{s}, 2\ell - 1), m')$, and then a successor to shorten the distance to the set $\overline{U}_1^{\overline{\sigma}}$ (such a successor exists since from all states in X the set $\overline{U}_1^{\overline{\sigma}}$ is reachable).

Against the strategy of player 3 in $\overline{G}_{\overline{\sigma}}$ either (i) $\overline{U}_1^{\overline{\sigma}}$ is reached in finitely many steps, or (ii) else player 2 infinitely often chooses successor states of the form $(\widetilde{s}, 0)$ with priority 0 (the minimum even priority), i.e., there is a cycle with a state $(\widetilde{s}, 0)$ which has priority 0. If priority 0 is visited infinitely often, then the parity objective is satisfied. This ensures that in $\overline{G}_{\overline{\sigma}}$ player 3 can ensure either to reach $\overline{U}_1^{\overline{\sigma}}$ in finitely many steps from some state in $\overline{U}_2^{\overline{\sigma}}$ against $\overline{\pi}$, or the parity objective is satisfied without reaching $\overline{U}_1^{\overline{\sigma}}$. In either case this implies that against $\overline{\pi}$ player 3 can ensure to satisfy the parity objective (by reaching $\overline{U}_1^{\overline{\sigma}}$ in finitely many steps and then playing a sure winning strategy from $\overline{U}_1^{\overline{\sigma}}$, or satisfying the parity objective without reaching $\overline{U}_1^{\overline{\sigma}}$ by visiting priority 0 infinitely often) from some state in $\overline{U}_2^{\overline{\sigma}}$, contradicting that $\overline{\pi}$ is a sure winning strategy for player 2 from $\overline{U}_2^{\overline{\sigma}}$. Thus we have a contradiction, and obtain the desired result. ∎

Lemma 2.1 establishes the desired correctness result as follows: (1) If $\overline{\sigma}$ is a finite-memory strategy such that in $\overline{G}_{\overline{\sigma}}$ player 3 has a sure winning strategy, then by part 1 of Lemma 2.1 we obtain that $\sigma = \mathrm{Tr}_{\mathrm{as}}(\overline{\sigma})$ is almost-sure winning. (2) Conversely, if σ is a finite-memory almost-sure winning strategy, then consider a strategy $\overline{\sigma}$ such that $\sigma = \mathrm{Tr}_{\mathrm{as}}(\overline{\sigma})$ (i.e., $\overline{\sigma} = \mathrm{Tr}_{\mathrm{as}}{}^{-1}(\sigma)$). By part 2 of Lemma 2.1, given the finite-memory strategy $\overline{\sigma}$, player 3 must have a sure winning strategy in $\overline{G}_{\overline{\sigma}}$, otherwise we will have a contradiction that σ is almost-sure winning. Thus we have the following theorem.

THEOREM 2.1. (POLYNOMIAL REDUCTION) *Given a partial-observation stochastic game graph G with a parity objective* $\mathsf{Parity}(p)$ *for player 1, we construct a three-player game $\overline{G} = \mathrm{Tr}_{\mathrm{as}}(G)$ with a parity objective* $\mathsf{Parity}(\overline{p})$, *where player 1 has partial-observation and the other two players have perfect-observation, in time $O((n+m)\cdot d)$, where n is the number of states of the game, m is the number of transitions, and d the number of priorities of the priority function p, such that the following assertion holds: there is a finite-memory almost-sure winning strategy σ for player 1 in G iff there exists a finite-memory strategy $\overline{\sigma}$ for player 1 in $\overline{G}$ such that in the game $\overline{G}_{\overline{\sigma}}$ obtained given $\overline{\sigma}$, player 3 has a sure winning strategy for* $\mathsf{Parity}(\overline{p})$. *The game graph $\mathrm{Tr}_{\mathrm{as}}(G)$ has $O(n \cdot d)$ states, $O(m \cdot d)$ transitions, and $\overline{p}$ has at most $d+1$ priorities.*

REMARK 3. *We have presented the details of the polynomial reduction for almost-sure winning, and now we discuss how a very similar reduction works for positive winning. We explain the key steps, and omit the proof as it is very similar to our proof for almost-sure winning. For clarity in presentation we use a priority -1 in the reduction, which is the least odd priority, and visiting the priority -1 infinitely often ensures loosing for player 1. Note that all priorities can be increased by 2 to ensure that priorities are nonnegative, but we use the priority -1 as it keeps the changes in the reduction for positive winning minimal as compared to almost-sure winning.*

Key steps. *First we observe that in the reduction gadgets for almost-sure winning, player 2 would never choose the leftmost edge to state $(\widetilde{s}, 0)$ from $\overline{s}$ in the cycles formed, but only use them for reachability to cycles. Intuitively, the leftmost edge corresponds to edges which must be choosen only finitely often and ensures positive reachability to the desired end components in the stochastic game. For positive winning these edges need to be in control of player 3, but must be allowed to be taken only finitely often. Thus for positive winning, the gadget is modified as follows: (i) we omit the leftmost edge from the state $\overline{s}$; (ii) we add an additional player-3 state $\widehat{s}$ in the beginning, which has an edge to $\overline{s}$ and an edge to $(\widehat{s}, 0)$; and (iii) the state $(\widehat{s}, 0)$ is assigned priority -1. Figure 3 presents a pictorial illustration of the gadget of the reduction for positive winning. Note that in the reduction for positive winning the finite reachability through the leftmost edge is in control of player-3, but it has the worst odd priority and must be used only finitely often. This essentially corresponds to reaching winning end components in finitely*

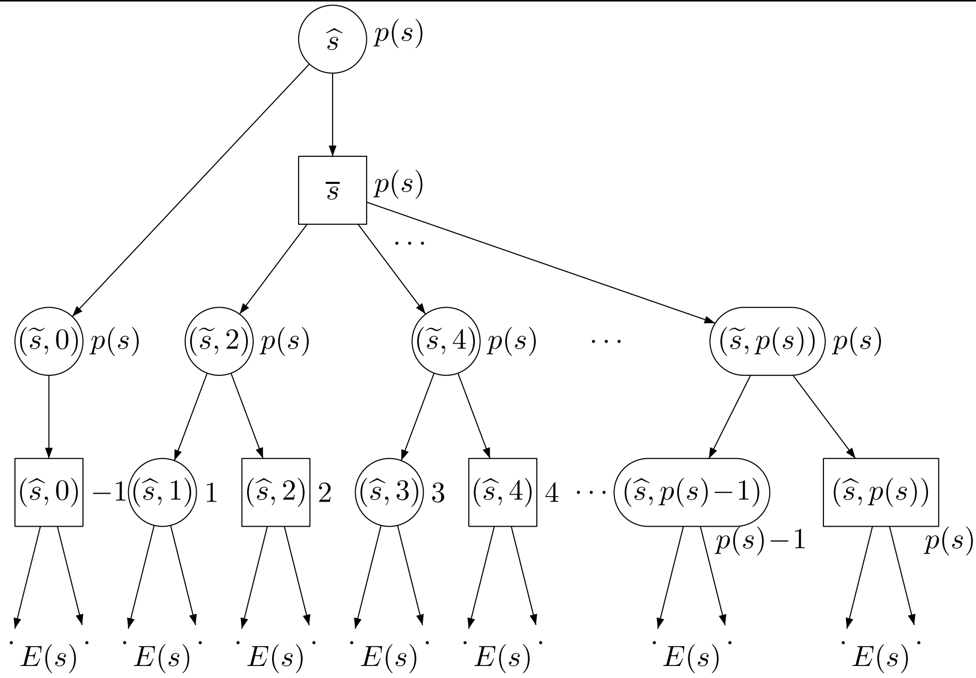

Figure 3: Reduction gadget for positive winning when $p(s)$ is even.

many steps in the stochastic game. In the game obtained after the reduction, the three-player game is surely winning iff player 1 has a finite-memory positive winning strategy in the partial-observation stochastic game.

In this section we established polynomial reductions of the qualitative analysis problems for partial-observation stochastic parity games under finite-memory strategies to the sure winning problem in three-player games (player 1 partial, both the other players perfect, and player 1 and 3 existential, player 2 adversarial). The following section shows that the sure winning problem for three-player games is EXPTIME-complete by reduction to alternating parity tree automata.

3 Solving Sure Winning for Three-player Parity Games

In this section we will present the solution for sure winning in three-player non-stochastic parity games. We start with the basic definitions.

3.1 Basic definitions We first present a model of partial-observation concurrent 3-player games, where player 1 has partial observation, and player 2 and player 3 have perfect observation. Player 1 and Player 3 have the same objective and they play against player 2. We will also show that three-player turn-based games model (of Section 2) can be treated as a special case of this model.

Partial-observation three-player concurrent games. Given alphabets A_i of actions for player i ($i = 1, 2, 3$), a partial-observation three-player concurrent game (for brevity, *3-player game* in sequel) is a tuple $G = \langle S, s_0, \delta, \mathcal{O}, \mathsf{obs} \rangle$ where:

- S is a finite set of states;

- $s_0 \in S$ is the initial state;

- $\delta : S \times A_1 \times A_2 \times A_3 \to S$ is a deterministic transition function that, given a current state s, and actions $a_1 \in A_1$, $a_2 \in A_2$, $a_3 \in A_3$ of the players, gives the successor state $s' = \delta(s, a_1, a_2, a_3)$ of s; and

- $\mathcal{O}$ is a finite set of observations and obs is the observation mapping (as in Section 2).

Modeling turn-based games. A three-player turn-based game will be a special case of the model three-player concurrent games. Formally, we consider a three-player turn-based game as a tuple $\langle S_1, S_2, S_3, A_1, \delta, E\rangle$ where $\delta : S_1 \times A_1 \to S_2$ is the transition function for player 1, and $E \subseteq (S_2 \cup S_3) \times S$ is a set of edges. Since player 2 and player 3 have perfect observation, we consider that $A_2 = S$ and $A_3 = S$, that is player 2 and player 3 choose directly a successor in the game. The transition function $\overline{\delta}$ for an equivalent concurrent version is as follows (i) for $s \in S_1$, for all $a_2 \in A_2$ and $a_3 \in A_3$, we have $\overline{\delta}(s, a_1, a_2, a_3) = \delta(s, a_1)$; (ii) for $s \in S_2$, for all $a_1 \in A_1$ and $a_3 \in A_3$, for $a_2 = s'$ we have $\overline{\delta}(s, a_1, a_2, a_3) = s'$ if $(s, s') \in E$, else $\overline{\delta}(s, a_1, a_2, a_3) = s_{\mathsf{good}}$, where s_{good} is a special state in which player 2 loses (the objective of player 1 and 3 is satisfied if player 2 chooses an edge that is not in E); and (iii) for $s \in S_3$, for all $a_1 \in A_1$ and $a_2 \in A_2$, for $a_3 = s'$ we have $\overline{\delta}(s, a_1, a_2, a_3) = s'$ if $(s, s') \in E$, else $\overline{\delta}(s, a_1, a_2, a_3) = s_{\mathsf{bad}}$, where s_{bad} is a special state in which player 2 wins (the objective of player 1 and 3 is violated if player 3 chooses an edge that is not in E). The set $\mathcal{O}$ and the mapping obs are obvious.

Strategies. Define the set Σ of *strategies* $\sigma : \mathcal{O}^+ \to A_1$ of player 1 that, given a sequence of past observations, return an action for player 1. Equivalently, we sometimes view a strategy of player 1 as a function $\sigma : S^+ \to A_1$ satisfying $\sigma(\rho) = \sigma(\rho')$ for all $\rho, \rho' \in S^+$ such that $\mathsf{obs}(\rho) = \mathsf{obs}(\rho')$, and say that σ is *observation-based*. A strategy of player 2 (resp, player 3) is a function $\pi : S^+ \to A_2$ (resp., $\tau : S^+ \to A_3$) without any restriction. We denote by Π, Γ the set of strategies of player $2, 3$ respectively.

Sure winning. Given strategies σ, π, τ of the three players in G, the *outcome play* from s_0 is the infinite sequence $\rho_{s_0}^{\sigma,\pi,\tau} = s_0 s_1 \dots$ such that for all $j \geq 0$, we have $s_{j+1} = \delta(s_j, a_j, b_j, c_j)$ where $a_j = \sigma(s_0 \dots s_j)$, $b_j = \pi(s_0 \dots s_j)$, and $c_j = \tau(s_0 \dots s_j)$. Given a game $G = \langle S, s_0, \delta, \mathcal{O}, \mathsf{obs}\rangle$ and a parity objective $\varphi \subseteq S^\omega$, the sure winning problem asks to decide if $\exists \sigma \in \Sigma \cdot \forall \pi \in \Pi \cdot \exists \tau \in \Gamma : \rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$. It will follow from our result that if the answer to the sure winning problem is yes, then there exists a witness finite-memory strategy σ for player 1.

3.2 Alternating Tree Automata In this section we recall the definitions of alternating tree automata, and present the solution of the sure winning problem for three-player games with parity objectives by a reduction to the emptiness problem of alternating tree automata with parity acceptance conditions.

Trees. We follow some definitions and notation of [11]. Given a finite sequence $w = s_0 \dots s_n \in \Omega^+$ over a finite set Ω, let $\mathsf{last}(w) = s_n$ be the last element of w.

A Ω-labeled tree (T, V) consists of a prefix-closed set $T \subseteq \mathbb{N}^*$ (i.e., if $x \cdot d \in T$ with $x \in \mathbb{N}^*$ and $d \in \mathbb{N}$, then $x \in T$), and a mapping $V : T \to \Omega$ that assigns to each node of T a letter in Ω. Given $x \in \mathbb{N}^*$ and $d \in \mathbb{N}$ such that $x \cdot d \in T$, we call $x \cdot d$ the *successor* in direction d of x. The degree $\deg(x)$ of a node $x \in T$ is the number of successors of x in T. The node ε is the *root* of the tree. An *infinite path* in T is an infinite sequence $\pi = d_1 d_2 \dots$ of directions $d_i \in \mathbb{N}$ such that every finite prefix of π is a node in T.

Alternating tree automata. Given a parameter $k \in \mathbb{N} \setminus \{0\}$, we consider input trees of rank k, i.e. trees in which every node has at most k successors. We present a definition of alternating tree automata (see e.g. [14, 11]) with the syntactic restriction that the states are associated to a fixed direction in the input tree. The restriction is for the sake of simplifying the presentation, and does not reduce the expressiveness of the class of automata (i.e., they recognize the regular languages of infinite trees with fixed finite rank). An *alternating tree automaton* over alphabet Ω is a tuple $\mathcal{A} = \langle S, s_0, \delta, \mathsf{dir}\rangle$ where:

- S is a finite set of states;

- $s_0 \in S$ is the initial state;

- $\delta : S \times \Omega \to \mathcal{B}^+(S)$ is a transition function where $\mathcal{B}^+(S)$ is the set of positive Boolean formulas over S, that is formulas built from elements in $S \cup \{\mathsf{true}, \mathsf{false}\}$ using the Boolean connectives $\wedge$ and $\vee$;
- $\mathsf{dir} : S \to \{0, \dots, k-1\}$ associates a fixed direction to each state.

Intuitively, the automaton is executed from the initial state s_0 and reads the input tree in a top-down fashion starting from the root ε. In state s, if $a \in \Omega$ is the letter that labels the current node x of the input tree, the behavior of the automaton is given by the formulas $\varphi = \delta(s, a)$ and the function dir. Informally, the automaton chooses a satisfying assignment of φ, i.e. a set $Q \subseteq S$ such that the formula φ is satisfied when the elements of Q are replaced by true, and the elements of $S \setminus Q$ are replaced by false. Then, for each $s' \in Q$ a copy of the automaton is spawned in state s', and proceeds the node $x \cdot d$ of the input tree, where $d = \mathsf{dir}(s')$ is the direction associated to s'. In particular, it requires that $x \cdot d$ belongs to the input tree. For example, if $\delta(s, a) = (s_1 \wedge s_2) \vee (s_3 \wedge s_4 \wedge s_5)$, and $\mathsf{dir}(s_1) = \mathsf{dir}(s_2) = \mathsf{dir}(s_3) = 0$ and $\mathsf{dir}(s_4) = \mathsf{dir}(s_5) = 1$, then the automaton should either spawn two copies that process the successor of x in direction 0 (i.e., the node $x \cdot 0$) and that enter the respective states s_1 and s_2, or spawn three copies of which one processes $x \cdot 0$ and enters state s_3, and the other two process $x \cdot 1$ and enter the states s_4 and s_5 respectively.

In a standard definition of alternating tree automata [14, 11], there is no fixed direction associated to each state of the automaton. Rather the transition function can specify a direction to proceed along with each state to enter (the transition relation is then of the form $\delta : S \times \Omega \to \mathcal{B}^+(S \times \{0, \dots, k-1\})$. And it is possible to specify several directions along with the same state, for instance $(s_1, 0) \wedge (s_1, 1)$ requires that the automaton spawn two copies in state s_1, one that proceeds direction 0 in the input tree, and one that proceeds direction 1. Hence our definition can be viewed as a syntactic restriction of the standard definition. However, the two definitions are equally powerful as alternating tree automata of the standard definition can be encoded in our definition as follows. For each state s, construct k copies $(s, 0), (s, 1), \dots, (s, k-1)$ of s (i.e., the transition relation in each copy is the same as in s), and assign direction $\mathsf{dir}(s, d) = d$ for each $0 \le d < k$.

Runs. The usual definition of a run of $\mathcal{A}$ over a Ω-labeled input tree (T, V) is a tree (T_r, r) labeled by elements of $T \times S$, where a node of T_r labeled by (x, s) corresponds to a copy of the automaton proceeding the node x of the input tree in state s. The root of T_r is labeled by (ε, s_0). We use a slightly richer definition: a run tree (T_r, r) is labeled by elements of $S^* \times S$, where a label $(\rho, s) = (s_1 \dots s_n, s)$ corresponds to a copy of the automaton that has visited the sequence of states $s_0 \cdot \rho = s_0 \dots s_n$ and is now proceeding the node $x = \mathsf{dir}(\rho) = \mathsf{dir}(s_1) \dots \mathsf{dir}(s_n)$ in state $s_n = s$. For $n = 0$, we assume that $\rho = \varepsilon$, and thus the automaton is in state s_0 and proceeds the root $x = \varepsilon$ of the input tree. Note that in all nodes of T_r except the root, the label (ρ, s) of the node satisfies $\mathsf{last}(\rho) = s$.

Formally, a run of $\mathcal{A}$ over an input tree (T, V) is a $(S^* \times S)$-labeled tree (T_r, r) such that $r(\varepsilon) = (\varepsilon, s_0)$ and for all $y \in T_r$, if $r(y) = (\rho, s)$, then the set $\{s' \mid \exists d \in \mathbb{N} : r(y \cdot d) = (\rho \cdot s', s')\}$ is a satisfying assignment for $\delta(s, V(\mathsf{dir}(\rho)))$. Hence we require that, given a node y labeled by (ρ, s), there is a satisfying assignment $Q \subseteq S$ for the formula $\delta(s, a)$ where $a = V(\mathsf{dir}(\rho))$ is the letter labeling the current node of the input tree, and for all states $s' \in Q$ there is a (successor) node $y \cdot d$ in T_r labeled by $(\rho \cdot s', s')$.

Given an accepting condition $\varphi \subseteq S^\omega$, we say that a run (T_r, r) is *accepting* if for all infinite paths $d_1 d_2 \dots$ of T_r, the sequence $s_0 s_1 s_2 \dots$ such that $r(d_i) = (\cdot, s_i)$ for all $i \ge 0$ is in φ. The *language* of $\mathcal{A}$ is the set $L_k(\mathcal{A})$ of all input trees of rank k over which there exists an accepting run of $\mathcal{A}$. The emptiness problem for alternating tree automata is to decide, given $\mathcal{A}$ and parameter k, whether $L_k(\mathcal{A}) = \emptyset$.

3.3 Solution of the Sure Winning Problem for Three-player Games We now present the solution of the sure winning problem for three-player games.

THEOREM 3.1. *Given a 3-player game $G = \langle S, s_0, \delta, \mathcal{O}, \mathsf{obs} \rangle$ and a $\{$safety, reachability, parity$\}$ objective φ, the problem of deciding whether*

$$\exists \sigma \in \Sigma \cdot \forall \pi \in \Pi \cdot \exists \tau \in \Gamma : \rho_{s_0}^{\sigma, \pi, \tau} \in \varphi$$

is EXPTIME-complete.

Proof. The EXPTIME-hardness follows from EXPTIME-hardness of two-player partial-observation games with reachability objective [17, 7] and safety objective [2].

We prove membership in EXPTIME by a reduction to the emptiness problem for alternating tree automata, which is solvable in EXPTIME for parity objectives [13, 14, 15]. The reduction is as follows. Given a game $G = \langle S, s_0, \delta, \mathcal{O}, \mathsf{obs} \rangle$ over alphabet of actions A_i ($i = 1, 2, 3$), we construct the alternating tree automaton $\mathcal{A} = \langle S', s'_0, \delta', \mathsf{dir} \rangle$ over alphabet Ω and parameter $k = |\mathcal{O}|$ where:

- $S' = S$, and $s'_0 = s_0$;
- $\Omega = A_1$;
- δ' is defined by $\delta'(s, a_1) = \bigvee_{a_3 \in A_3} \bigwedge_{a_2 \in A_2} \delta(s, a_1, a_2, a_3)$ for all $s \in S$ and $a_1 \in \Omega$;
- $\mathsf{dir} = \mathsf{obs}$ (strictly speaking, assuming observations in $\mathcal{O}$ are numbered $0, \ldots, k-1$, for each $s \in S$ the direction $\mathsf{dir}(s)$ is the number of the observation $\mathsf{obs}(s)$).

The acceptance condition φ of the automaton is same as the objective of the game G. We prove that $\exists \sigma \in \Sigma \cdot \forall \pi \in \Pi \cdot \exists \tau \in \Gamma : \rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$ if and only if $L_k(\mathcal{A}) \neq \emptyset$.

1. *Sure winning implies non-emptiness.* First, assume that for some $\sigma \in \Sigma$, we have $\forall \pi \in \Pi \cdot \exists \tau \in \Gamma : \rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$. Then, by fixing σ in the game G, we obtain a two-player perfect-information game with countably-infinite state space, which is determined [12]. Hence there also exists $\tau \in \Gamma$ such that $\forall \pi \in \Pi : \rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$. From σ, we define an input tree (T, V) where $T = \{0, \ldots, k-1\}^*$ and $V(\gamma) = \sigma(\mathsf{obs}(s_0) \cdot \gamma)$ for all $\gamma \in \mathcal{O}^*$ (we view σ as a function $\mathcal{O}^+ \to \Omega$, remember that $\Omega = A_1$). From τ, we define a $(S^* \times S)$-labeled tree (T_r, r) such that $r(\varepsilon) = (\varepsilon, s_0)$ and for all $y \in T_r$, if $r(y) = (\rho, s)$, then for $a_1 = \sigma(\mathsf{obs}(s_0 \cdot \rho)) = V(\mathsf{dir}(\rho))$, for $a_3 = \tau(s_0 \cdot \rho)$, for every s' in the set $Q = \{s' \mid \exists a_2 \in A_2 : s' = \delta(s, a_1, a_2, a_3)\}$, there is a successor $y \cdot d$ of y in T_r labeled by $r(y \cdot d) = (\rho \cdot s', s')$. Note that Q is a satisfying assignment for $\delta'(s, a_1)$ and $a_1 = V(\mathsf{dir}(\rho))$, hence (T_r, r) is a run of $\mathcal{A}$ over (T, V). For every infinite path ρ in (T_r, r), consider a strategy $\pi \in \Pi$ consistent with ρ. Then $\rho = \rho_{s_0}^{\sigma,\pi,\tau}$, hence $\rho \in \varphi$ and the run (T_r, r) is accepting, showing that $L_k(\mathcal{A}) \neq \emptyset$.

2. *Non-emptiness implies sure winning.* Second, assume that $L_k(\mathcal{A}) \neq \emptyset$. Let $(T, V) \in L_k(\mathcal{A})$ and (T_r, r) be an accepting run of $\mathcal{A}$ over (T, V). From (T, V), define a strategy σ of player 1 such that $\sigma(s_0 \cdot \rho) = V(\mathsf{dir}(\rho))$ for all $\rho \in S^*$. Note that σ is indeed a strategy of player 1 since $\sigma(\rho) = \sigma(\rho')$ for all $\rho, \rho' \in S^+$ such that $\mathsf{obs}(\rho) = \mathsf{obs}(\rho')$. From (T_r, r), we know that for all nodes $y \in T_r$ with $r(y) = (\rho, s)$, the set $Q = \{s' \mid \exists d \in \mathbb{N} : r(y \cdot d) = (\rho \cdot s', s')\}$ is a satisfying assignment of $\delta'(s, V(\mathsf{dir}(\rho)))$, hence there exists $a_3 \in A_3$ such that for all $a_2 \in A_2$, there is a successor of y labeled by $(\rho \cdot s', s')$ with $s' = \delta(s, a_1, a_2, a_3)$ and $a_1 = \sigma(s_0 \cdot \rho)$. Then define $\tau(s_0 \cdot \rho) = a_3$.

   Now, for all strategies $\pi \in \Pi$ the outcome $\rho_{s_0}^{\sigma,\pi,\tau}$ is a path in (T_r, r) hence $\rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$. Therefore $\exists \sigma \in \Sigma \cdot \exists \tau \in \Gamma \cdot \forall \pi \in \Pi : \rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$, and by determinacy [12] it follows that $\exists \sigma \in \Sigma \cdot \forall \pi \in \Pi \cdot \exists \tau \in \Gamma : \rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$.

The desired result follows. ∎

The emptiness problem for an alternating tree automaton $\mathcal{A}$ with parity condition can be solved by constructing an equivalent nondeterministic parity tree automaton $\mathcal{N}$ (such that $L_k(\mathcal{A}) = L_k(\mathcal{N})$), and then checking emptiness of $\mathcal{N}$. By the result of [15, Theorem 1.2] for binary trees, if $\mathcal{A}$ has n states and d priorities, then $\mathcal{N}$ has $n' = 2^{O(d \cdot n \cdot \log n)}$ states and $e = O(d \cdot n \cdot \log n)$ priorities. Note that the emptiness problem for input trees of rank k is equivalent to the emptiness problem for binary trees in an alternating automaton with $k \cdot n$ states. Finally, the emptiness of a nondeterministic parity tree automaton with m transitions and e priorities is equivalent to solving a two-player parity game [10], which can be done in time $m^{O(e)}$ [20]. Moreover, since memoryless strategies exist for parity games [10], if the nondeterministic parity tree automaton is nonempty, then it accepts a regular tree that can be encoded by a transducer of size m. Since m is at most quadratic in the size of the state space, the emptiness problem for alternating tree automaton with parity condition can be solved in time $2^{O((d \cdot k \cdot n \cdot \log k \cdot n)^2)}$, and it is sufficient to consider input trees encoded by transducers of size $2^{O(d \cdot k \cdot n \cdot \log k \cdot n)}$.

THEOREM 3.2. *Given a* 3*-player game* $G = \langle S, s_0, \delta, \mathcal{O}, \mathsf{obs} \rangle$ *with* n *states (and* $k \leq n$ *observations for player* 1*) and parity objective* φ *defined by* d *priorities, the problem of deciding whether*

$$\exists \sigma \in \Sigma \cdot \forall \pi \in \Pi \cdot \exists \tau \in \Gamma : \rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$$

can be solved in time $2^{O(d^2 \cdot n^4 \cdot \log^2 n)}$*. Moreover, memory of size* $2^{O(d \cdot n^2 \cdot \log n)}$ *is sufficient for player* 1*.*

REMARK 4. *Note that we considered the problem of deciding whether*

$$\exists \sigma \in \Sigma \cdot \forall \pi \in \Pi \cdot \exists \tau \in \Gamma : \rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$$

which is equivalent to

$$\exists \sigma \in \Sigma \cdot \exists \tau \in \Gamma \cdot \forall \pi \in \Pi : \rho_{s_0}^{\sigma,\pi,\tau} \in \varphi$$

because once the strategy for player 1 is given we have a perfect-observation game where we can switch the quantifiers of strategies due to determinacy. Second, by our reduction to alternating parity tree automata and the fact that if an alternating parity tree automaton is non-empty, there is a regular witness tree for non-emptiness it follows that strategies for player 1 can be restricted to finite-memory without loss of generality. This ensures that we can solve the problem of existence of finite-memory almost-sure strategies in parial-observation stochastic parity games (by Theorem 2.1 of Section 2 also in EXPTIME), and EXPTIME-completeness of the problem follows.

THEOREM 3.3. *Given a partial-observation stochastic game and a parity objective* φ *defined by* d *priorities, the problem of deciding whether there exists a finite-memory almost-sure (resp. positive) winning strategy for player 1 (i) is EXPTIME complete; (ii) can be decided time* $2^{O((d+1)^2 \cdot (n \cdot d)^4 \cdot \log^2 (n \cdot d))}$*. Moreover, if there is an almost-sure (resp. positive) winning strategy, there exists one that uses memory of size* $2^{O(d \cdot (n \cdot d)^2 \cdot \log (n \cdot d))}$*.*

REMARK 5. *As mentioned in Remark 2 the EXPTIME upper bound for qualtitative analysis of partial-observation stochastic parity games with finite-memory randomized strategies follows from Theorem 3.3. The EXPTIME lower bound and the exponential lower bound on memory requirement for finite-memory randomized strategies follows from the results of [7, 6] for reachability and safety objectives (even for POMDPs).*

References

[1] 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.

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

[3] K. Chatterjee, M. Chmelik, and M. Tracol. What is decidable about partially observable Markov decision processes with omega-regular objectives. In *Proceedings of CSL 2013: Computer Science Logic*, 2013.

[4] K. Chatterjee and L. Doyen. Partial-observation stochastic games: How to win when belief fails. In *Proceedings of LICS 2012: Logic in Computer Science*, pages 175–184. IEEE Computer Society Press, 2012.

[5] K. Chatterjee, L. Doyen, H. Gimbert, and T. A. Henzinger. Randomness for free. In *CoRR abs/1006.0673 (Full version)*, 2010. Conference version Proc. of MFCS, Springer, LNCS 6281, pages 246-257.

[6] 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.

[7] 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.

[8] C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *Journal of the ACM*, 42(4):857–907, 1995.

[9] L. de Alfaro. *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University, 1997. Technical Report STAN-CS-TR-98-1601.

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

[11] O. Kupferman, M. Y. Vardi, and P. Wolper. An automata-theoretic approach to branching-time model checking. *J. ACM*, 47(2):312–360, 2000.

[12] D.A. Martin. Borel determinacy. *Annals of Mathematics*, 102(2):363–371, 1975.

[13] D. E. Muller, A. Saoudi, and P. E. Schupp. Alternating automata. the weak monadic theory of the tree, and its complexity. In *Proc. of ICALP: Automata, Languages and Programming*, LNCS 226, pages 275–283. Springer, 1986.

[14] D. E. Muller and P. E. Schupp. Alternating automata on infinite trees. *Theor. Comput. Sci.*, 54:267–276, 1987.

[15] D. E. Muller and P. E. Schupp. Simulating alternating tree automata by nondeterministic automata: New results and new proofs of the theorems of Rabin, McNaughton and Safra. *Theor. Comput. Sci.*, 141(1&2):69–107, 1995.

[16] S. Nain and M. Y. Vardi. Solving partial-information stochastic parity games. In *LICS*, pages 341–348, 2013.

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

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

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

[20] W. Zielonka. Infinite games on finitely coloured graphs with applications to automata on infinite trees. *Theor. Comput. Sci.*, 200:135–183, 1998.