Qualitative Determinacy and Decidability of Stochastic Games with Signals

NATHALIE BERTRAND,, INRIA, Centre Bretagne Atlantique, Rennes, France
BLAISE GENEST,, CNRS, IRISA, Rennes, France
and HUGO GIMBERT,, CNRS, LaBRI, Bordeaux, France

We consider the standard model of finite two-person zero-sum stochastic games with signals. We are interested in the existence of almost-surely winning or positively winning strategies, under reachability, safety, Büchi or co-Büchi winning objectives. We prove two *qualitative determinacy* results. First, in a reachability game either player 1 can achieve almost-surely the reachability objective, or player 2 can achieve surely the complementary safety objective, or both players have positively winning strategies. Second, in a Büchi game if player 1 cannot achieve almost-surely the Büchi objective, then player 2 can ensure positively the complementary co-Büchi objective. We prove that players only need strategies with *finite-memory*. The number of memory states ranges from nil (memoryless) to doubly-exponential, with matching upper and lower bounds. Together with the qualitative determinacy results, we also provide fix-point algorithms for deciding which player has an almost-surely winning or a positively winning strategy and for computing an associated finite memory strategy. Complexity ranges from EXPTIME to 2EXPTIME, with matching lower bounds. Our fix-point algorithms also enjoy a better complexity in the cases where one of the players is better informed than her opponent.

Categories and Subject Descriptors: B.6.3 [**Logic Design**]: Design Aids

General Terms: Automatic synthesis

Additional Key Words and Phrases: Stochastic Games, Controller Synthesis

1. INTRODUCTION

Numerous advances in algorithmics of stochastic games have recently been made [de Alfaro et al. 2007; de Alfaro and Henzinger 2000; Chatterjee et al. 2004; Chatterjee et al. 2005; Gimbert and Horn 2008; Horn 2008], motivated in part by application in controller synthesis and verification of open systems. Open systems can be viewed as two-player games between the system and its environment. At each round of the game, both players independently and simultaneously choose actions and the two choices together with the current state of the game determine transition probabilities to the next state of the game. Properties of open systems are modeled as objectives of the games [de Alfaro and Henzinger 2000; Grädel et al. 2002], and strategies in these games represent either controllers of the system or behaviors of the environment.

Most algorithms for stochastic games suffer from the same restriction: they are designed for games where players can fully observe the state of the system (e.g. concurrent games [de Alfaro et al. 2007; de Alfaro and Henzinger 2000] and stochastic games with perfect information [Condon 1992; Horn 2008]). The full observation hypothesis can hinder interesting applications in controller synthesis, where the synthesized con-

This work was supported by the ANR projet "DOTS" and the ESF Research Networking Programme project "GAMES2".

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

troller should be robust against all possible behaviours of its environnment. In most controllable open systems full monitoring for the controller is not implementable in practice, for example the controller of a driverless subway system cannot directly observe an hardware failure and is only informed about failures detected by the monitoring system, including false alarm due to sensors failures. Also giving full information to the environnment is not realistic since for example a cryptographic protocol is hardly robust against attacks of an adversary that has full information about the encryption keys.

In the present paper, we consider *stochastic games with signals*, that are a standard tool in game theory to model games under partial observation [Sorin 2002; Rosenberg et al. 2003; Renault 2007]. When playing a stochastic game with signals, players cannot observe the actual state of the game, nor the actions played by themselves or their opponent: the only source of information of a player are private signals he receives throughout the play. Stochastic games with signals subsume standard stochastic games [Shapley 1953], repeated games with incomplete information [Aumann 1995], games with imperfect monitoring [Rosenberg et al. 2003], concurrent games [de Alfaro and Henzinger 2000] and deterministic games with imperfect information on one side [Reif 1979; Chatterjee et al. 2007]. Players make their decisions based upon the sequence of signals they receive: a strategy is hence a mapping from finite sequences of private signals to probability distributions over actions.

From the algorithmic point of view, stochastic games with signals are considerably harder to deal with than stochastic games with full observation. While *values* of the latter games are computable [de Alfaro and Henzinger 2000; Chatterjee et al. 2005], simple questions like 'is there a strategy for player 1 which guarantees winning with probability more than $\frac{1}{2}$?' are *undecidable* even for the restricted class of stochastic games with a single signal and a single player and for a reachability objective [Paz 1971]. Also, for this restricted classed corresponding to Rabin's probabilistic automata [Rabin 1963], the value 1 problem is undecidable [Gimbert and Oualhadj 2010]. Because of these undecidability results, rather than *quantitative* properties (i.e. questions about values), we focus in the present paper on *qualitative* properties of stochastic games with signals.

We study the following qualitative questions about stochastic games with signals, equipped with reachability, safety, Büchi objectives:

(i) Does player 1 have an *almost-surely winning strategy*, i.e. a strategy which guarantees the objective to be achieved with probability 1, whatever the strategy of player 2?
(ii) Does player 2 have a *positively winning strategy*, i.e. a strategy which guarantees the opposite objective to be achieved with strictly positive probability, whatever the strategy of player 1?

Obviously, given an objective, properties (i) and (ii) cannot hold simultaneously. We obtain the following results:

(1) Either property (i) holds or property (ii) holds; in other words these games are *qualitatively determined*.
(2) Players only need strategies with *finite-memory*. The number of memory states needed ranges from nil (memoryless) to doubly-exponential.
(3) Questions (i) and (ii) are decidable. We provide fix-point algorithms for computing all initial states that satisfy (i) or (ii), together with the corresponding finite-memory strategies. The complexity of the algorithms ranges from EXPTIME to 2EXPTIME.

These three results are detailed in Theorems 4.1, 4.6, 6.3 and 6.4. We prove that these results are tight and robust in several aspects. Games with co-Büchi objectives are absent from this picture, since they are neither qualitatively determined (see Section 5.2) nor decidable (as shown in [Baier et al. 2008; Chatterjee et al. 2010]).

Another element of surprise, is that for winning positively a safety or co-Büchi objective, a player *needs* a memory with a doubly-exponential number of states, and the corresponding decision problem is 2EXPTIME-complete. This result departs from what was previously known [Reif 1979; Chatterjee et al. 2007], where both the number of memory states and the complexity are simply exponential. Our contributions also reveal a nice property of *reachability* games, that Büchi games do not enjoy: Every initial state is either *almost-surely winning* for player 1, *surely winning* for player 2 or *positively winning* for both players.

Our results strengthen and generalize in several ways results that were previously known for concurrent games [de Alfaro et al. 2007; de Alfaro and Henzinger 2000] and deterministic games with imperfect information on one side [Reif 1979; Chatterjee et al. 2007].

First, the framework of stochastic games with signals strictly encompasses all the settings of [Reif 1979; de Alfaro et al. 2007; de Alfaro and Henzinger 2000; Chatterjee et al. 2007]. In concurrent games there is no signaling structure at all, and in deterministic games with imperfect information on one side [Chatterjee et al. 2007] transitions are deterministic and player 2 observes everything that happens in the game, including results of random choices of his (or her) opponent.

We believe that the extension of results of [Chatterjee et al. 2007] to games with imperfect information on both sides is mandatory to perform controller synthesis on real-life systems. For example, if the adversarial environment of a cryptographic protocol has full information then in particular it is informed about the cyphering keys used by the system, and the protocol can hardly be robust against attacks of such an omniscient adversary.

Second, we prove that Bchi games are qualitatively determined: when player 1 cannot win almost-surely a Bchi game then her opponent can win positively. This was not known previously, even for games with imperfect information on one side: in [Reif 1979; Chatterjee et al. 2007] algorithms are given for deciding whether the imperfectly informed player has an almost-surely winning strategy for a Büchi (or reachability) objective, however, no results (e.g.: strategy for the opponent) are given in case this player has no such strategy. Our qualitative determinacy result (1) is a radical generalization of the same result for concurrent games [de Alfaro and Henzinger 2000, Th.2], using different techniques. Interestingly, for concurrent games, qualitative determinacy holds for every omega-regular objectives [de Alfaro and Henzinger 2000], while for games with signals we show that it fails already for co-Büchi objectives. Interestingly also, stochastic games with signals and a reachability objective have a value [Renault and Sorin 2008] but this value is not computable [Paz 1971], whereas it is computable for concurrent games with omega-regular objectives [de Alfaro and Majumdar 2001]. The use of randomized strategies is mandatory for achieving determinacy results, this also holds for stochastic games without signals [Shapley 1953; de Alfaro et al. 2007] and even matrix games [von Neumann and Morgenstern 1944], which contrasts with [Berwanger et al. 2008; Reif 1979] where only deterministic strategies are considered.

We believe that qualitative determinacy is a crucial property of stochastic games when used for controller synthesis, because it allows for incremental design and refinement of systems models and controllers. In case a model of a system (say the door control system of the driverless subway in Paris) does not have a correct controller, qualitative determinacy can be used to compute a winning strategy of the environnement. Such an environnement strategy can be used to perform simulation and get error traces, and we believe this can be of great help for the system designers. In case where the environment strategy is not implementable on the actual system (for example if it requires a passenger to walk through a closed door) then the corresponding restrictions on the environment behaviour should be added to the model of the system. Otherwise (for example if backward moves of the subway are possible while doors are opened), the system itself should be patched to defeat this particular environment strategy (for example with extra conditions for the backward move to be allowed). Without qualitative determinacy, the designer is left with no feedback when the algorithm answers that there is no winning strategy for the system, and this is a serious limitation to the industrial use of automatic controller synthesis.

Also qualitative determinacy has a strong theoretical interest. The study of zero-sum stochastic games is usually focused on the existence of the *value* of games: the value is the threshold payoff which is a minimal income for player 1 and a maximal loss for player 2, when playing with optimal strategies. The existence of a value is a clue that the strategy sets of the players are rich enough to let them play efficiently, for example deterministic strategies which do not use random coin tosses are too restrictive to play a repeated rock-paper-scissors game. The synthesis of almost-surely winning strategies is not related to the notion of value since there are games with value 1 but no almost-surely winning strategies. To our opinion, qualitative determinacy is the key notion of determinacy for almost-surely winning strategies and the key criterion to check that the players are equipped with the suitable set of strategies.

Our results about randomized finite-memory strategies (2), stated in Theorem 4.6, are either brand new or generalize previous work. It was shown in [Chatterjee et al. 2007] that for deterministic games where player 2 is perfectly informed, strategies with a finite memory of exponential size are sufficient for player 1 to achieve a Büchi objective almost-surely. We prove the same result holds for the whole class of stochastic games with signals. Moreover we prove that for player 2 a doubly-exponential number of memory states is necessary and sufficient to achieve positively the complementary co-Büchi objective.

Concerning algorithmic results (3) (see details in Theorem 6.3 and 6.4) we give a fix-point based algorithm for deciding whether a player has an almost-surely winning strategy for a Büchi objective. If it is the case, a strategy for achieving almost-surely the Büchi objective (with an exponential number of memory states) can be derived easily. If it is not the case, a strategy (with a doubly exponential number of memory states) for player 2 to prevent the Büchi objective with positive probability can be derived easily. Our algorithm with 2EXPTIME complexity is optimal since the problem is indeed 2EXPTIME-hard (see Theorem 8.1). The same algorithm is also optimal, and with an EXPTIME complexity, under the hypothesis that player 2 has *more information* than player 1. This generalizes the EXPTIME-completeness result of [Chatterjee et al. 2007], in the case where player 2 has perfect information. Last our algorithm also runs in EXPTIME when player 1 has full information. In both subcases, player 2 needs only exponential memory (see Proposition 8.2).

Recently, a refined version of Büchi objectives was introduced: Instead of requiring infinitely many visits to accepting, it asks that the limit average of visits to accepting states is positive [Tracol 2011]. Considering this winning condition for the restricted class of probabilistic automata (which correspond to single player games where the player is blind) makes the positively-winning ser of states computable, contrary to probabilistic automata equipped with a standard Büchi condition. However, whether such a condition can be ensured almost surely is still undecidable. Therefore, games with this objective are not qualitatively determined, similarly to co-Büchi games.

Some of our results have been concurrently obtained in [Gripon and Serre 2009] whose contribution is weaker than ours in several aspects: no determinacy result is provided, nothing is said about strategies used by player 2 nor the memory he needs, and the algorithm provided is enumerative rather than fix-point based. Moreover it was shown in [Chatterjee and Doyen 2011] that the main result of [Gripon and Serre 2009] is incorrect. The complexity results of [Gripon and Serre 2009] hold in a different model, with the additional assumption that actions are visible to the players, which is a corollary of earliest results of the authors of the present paper [Bertrand et al. 2009]. A corrected version of [Gripon and Serre 2009] can be found in [Gripon and Serre 2011]. In the present paper we do not assume visibility of the actions.

Some results have recently appeared in [Chatterjee and Doyen 2011], and they are in sharp contrast with our results, in particular, it is shown in [Chatterjee and Doyen 2011] that a memory of non-elementary size is needed to win stochastic games wih invisible actions, while in the present paper we claim that doubly-exponential memory is sufficient. The point is that the class of finite-memory strategy used in [Chatterjee and Doyen 2011] is more restricted than ours because they require the choices of actions and the memory updates to be deterministic, while in the present paper we assume that a player can roll dices to update the memory of its strategy and can play using lotteries rather than fixed actions. In a nutshell, mimicking randomness with deterministic counters can be very costly.

The paper is organized as follows. In Section 2 we introduce partial observation games, in Section 3.1 we define the notion of qualitative determinacy. The main results are stated in Section 4: qualitative determinacy, memory complexity and algorithmic complexity. We then prove the determinacy results in Section 5, and give the fix-point algorithm to compute the winning sets of states for each player in Section 6. In Section 7, we discuss the memory needed by strategies. Section 8 establishes lower bounds in general as well as for special cases.

This paper is an extended version of [Bertrand et al. 2009]. In particular, there are two novelties: we provide a direct proof of qualitative determinacy and our results hold in the general case where the players canot observe their actions. Moreover complete proofs are provided.

2. STOCHASTIC GAMES WITH SIGNALS.

We consider the standard model of finite two-person zero-sum stochastic games with signals [Sorin 2002; Rosenberg et al. 2003; Renault 2007]. These are stochastic games where players cannot observe the actual state of the game, nor the actions played by themselves and their opponent; their only source of information are private signals they receive throughout the play. However, since the players know which game they are playing, their private signals give them some clues about the information hidden from them. Stochastic games with signals subsume standard stochastic games [Shapley 1953], repeated games with incomplete information [Aumann 1995], games with imperfect monitoring [Rosenberg et al. 2003], games with imperfect information [Chatterjee et al. 2007; Gripon and Serre 2009] and partial-observation stochastic games [Chatterjee et al. 2013].

Notations. Given a finite set K, we denote by $\Delta(K) = \{\delta : K \to [0,1] \mid \sum_k \delta(k) = 1\}$ the set of probability distributions on K. For every distribution $\delta \in \Delta(K)$, we denote $\mathrm{supp}(\delta) = \{k \in K \mid \delta(k) > 0\}$ its support. For every state $k \in K$, we denote $\mathbf{1}_k$ the unique distribution whose support is the singleton $\{k\}$.

States, actions, signals and arenas. Two players called 1 and 2 have opposite goals and play for an infinite sequence of steps, choosing actions and receiving signals. Players observe the signals they receive but they cannot observe the actual state of the game, nor the actions that are played nor the signals received by their opponent. We assume player 1 to be of female type and player 2 to be of male type. We borrow notations from [Renault 2007]. Initially, the game is in a state $k_0 \in K$ chosen according to an initial distribution $\delta \in \Delta(K)$ known by both players; the initial state is k_0 with probability $\delta(k_0)$. At each step $n \in \mathbb{N}$, players 1 and 2 choose some actions $i_n \in I$ and $j_n \in J$. They respectively receive signals $c_n \in C$ and $d_n \in D$, and the game moves to a new state k_{n+1}. This happens with probability $p(k_{n+1}, c_n, d_n \mid k_n, i_n, j_n)$ given by fixed transition probabilities $p : K \times I \times J \to \Delta(K \times C \times D)$, known by both players. The tuple (K, I, J, C, D, p) is called an arena.

Plays and strategies. A finite play is a sequence $(k_0, i_0, j_0, c_1, d_1, k_1, \ldots, c_n, d_n, k_n) \in (KIJCD)^*K$ such that for every $0 \leq m < n$, $p(k_{m+1}, c_{m+1}, d_{m+1} \mid k_m, i_m, j_m)) > 0$. An infinite play is a sequence in $(KIJCD)^\omega$ such that each prefix in $(KIJCD)^*K$ is a finite play.

Players observe and remember the sequence of signals they receive. At each step of the game, the action chosen by each player depends on the sequence of signals he has received so far. In words, the strategy of a player associates with an initial distribution and each finite sequence of signals, a probability distribution over actions called a *lottery* and the action actually played is randomly chosen according to this lottery. Formally, a (behavioral) strategy of player 1 is a mapping $\sigma : \Delta(K) \times C^* \to \Delta(I)$. If the initial distribution is δ and player 1 has seen signals $c_1, \ldots, c_n$ then he plays action i with probability $\sigma(\delta, c_1, \ldots, c_n)(i)$. Strategies for player 2 are defined symmetrically, and denoted τ.

Note that the choice of a strategy determines which lotteries over actions are played, and players do know the lotteries they choose, but *we do not assume that players can observe the actions they have actually played*. This contrasts with the model of strategies used in other computer science papers about stochastic games with partial information [Chatterjee et al. 2007; Bertrand et al. 2009; Gripon and Serre 2009; Chatterjee and Doyen 2011; Chatterjee et al. 2013].

We use K_n, I_n, J_n, C_{n+1} and D_{n+1} to denote the random variables corresponding respectively to n-th state, action of player 1, action of player 2, signal of player 1 and signal of player 2.

In the usual way, an initial distribution δ and two strategies σ and τ define a probability measure $\mathbb{P}_\delta^{\sigma,\tau}$ on the set of infinite plays, equipped with the σ-algebra generated by cylinders, that is, sets of infinite plays that extend a common prefix finite play. The probability measure $\mathbb{P}_\delta^{\sigma,\tau}$ is the only probability measure over $(KIJCD)^\omega$ such that for every $k \in K$, $\mathbb{P}_\delta^{\sigma,\tau}(K_0 = k) = \delta(k)$ and for every $n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}_\delta^{\sigma,\tau}&(K_{n+1}, C_{n+1}, D_{n+1} \mid K_0, I_0, J_0, C_1, D_1, K_1, \ldots, C_n, D_n, K_n) \\ &= p(K_{n+1}, C_{n+1}, D_{n+1} \mid K_n, I_n, J_n) \ , \end{aligned} \tag{1}$$

where we use standard notations for conditional probability measures.

Finite memory strategies. Behavioral strategies are infinite objects, and to design algorithms that compute and manipulate strategies it is convenient to use finite memory strategies which can be defined finitely. A finite memory strategy for player 1 is described by a finite set M called the memory, a strategic function $\sigma_M : M \to \Delta(I)$, an update function $\mathrm{upd}_M : M \times C \to \Delta(M)$ and initialization function $\mathrm{init}_M : \mathcal{P}(K) \to \Delta(M)$. After a sequence of signals $c_1 \cdots c_n \in C^*$ the probability for the memory state to be m

is defined recursively by

$$p(m \mid \delta, c_1 \cdots c_n) = \begin{cases} \mathrm{init}_M(\mathrm{supp}(\delta))(m) & \text{if } n = 0, \\ \sum_{m' \in M} p(m' \mid \delta, c_1 \cdots c_{n-1}) \cdot \mathrm{upd}_M(m', c_n)(m) & \text{otherwise.} \end{cases}$$

In order to play with a finite-memory strategy, a player proceeds as follows. He initializes the memory of σ to $\mathrm{init}_M(L)$, where $L = \mathrm{supp}(\delta)$ is the support of the initial distribution δ. When the memory is in state $m \in M$, he plays action i with probability $\sigma_M(m)(i)$ and after receiving signal c, the new memory state is m' with probability $\mathrm{upd}_M(m, c)(m')$.

Note that random events such as transitions between memory states or results of lotteries over actions are unknown to the opponent, whose only source of information are the signals he receives.

With every finite-memory strategy is associated a behavioral strategy which induces the same probability distributions against every possible behavioral strategy of the opponent. Formally, the behavioral strategy associated with the finite memory strategy given by M, σ_M, upd_M and init_M is defined by:

$$\sigma(\delta)(c_1 \cdots c_n)(i) = \sum_{m \in M} p(m \mid \delta, c_1 \cdots c_n) \cdot \sigma_M(m)(i) \ .$$

We say a finite-memory strategy is *deterministic* if the distributions given by the three mappings $\mathrm{init}_M, \mathrm{upd}_M$ and σ_M are indicator functions: the initial memory state is uniquely determined by the support of the initial distribution, and the memory states and actions are uniquely determined by the sequence of signals. The power of deterministic finite memory strategies has been studied in [Chatterjee and Doyen 2011] and it appears that such strategies are sufficient to win almost-surely a partial observation reachability game when possible, but they may require much more memory than randomized finite-memory strategies, see Remark 6.2.

Winning conditions and almost-surely or positively winning strategies, distributions and supports. The goal of player 1 is described by a measurable set of infinite plays Win called the *winning condition*. When player 1 and 2 use strategies σ and τ and the initial distribution is δ, then player 1 wins the game with probability:

$$\mathbb{P}_{\delta}^{\sigma,\tau}(\mathrm{Win}) \ .$$

Player 1 wants to maximize this probability, while player 2 wants to minimize it. The best situation for player 1 is when she has an almost-surely winning strategy.

Definition 2.1 (*Almost-surely winning strategy*). A strategy σ for player 1 is *almost-surely winning* from an initial distribution δ if

$$\forall \tau, \mathbb{P}_{\delta}^{\sigma,\tau}(\mathrm{Win}) = 1 \ . \tag{2}$$

When such an almost-surely strategy σ exists, the initial distribution δ is said to be almost-surely winning (for player 1).

A less enjoyable situation for player 1 is when she only has a positively winning strategy.

Definition 2.2 (*Positively winning strategy*). A strategy σ for player 1 is *positively winning* from an initial distribution δ if

$$\forall \tau, \mathbb{P}_{\delta}^{\sigma,\tau}(\mathrm{Win}) > 0 \ . \tag{3}$$

When such a strategy σ exists, the initial distribution δ is said to be positively winning (for player 1).

Symmetrically, a strategy τ for player 2 is positively winning if it guarantees $\forall\sigma, \mathbb{P}_{\delta}^{\sigma,\tau}(\mathrm{Win}) < 1$.

The worst situation for player 1 is when his (or her) opponent has an almost-surely winning strategy τ, which thus ensures $\mathbb{P}_{\delta}^{\sigma,\tau}(\mathrm{Win}) = 0$ for all strategies σ chosen by player 1.

Note that whether a distribution δ is almost-surely or positively winning depends only on its support, because $\mathbb{P}_{\delta}^{\sigma,\tau}(\mathrm{Win}) = \sum_{k\in K} \delta(k) \cdot \mathbb{P}_{\delta}^{\sigma,\tau}(\mathrm{Win} \mid K_0 = k)$. As a consequence, we will say that a support $L \subseteq K$ is almost-surely or positively winning for a player if there exists a distribution with support L which has the same property.

Reachability, safety, Büchi and co-Büchi games. Motivated by applications in logic and controller synthesis [Grädel et al. 2002], we are especially interested in *reachability, safety, Büchi and co-Büchi conditions*. These four winning conditions use a subset $T \subseteq K$ of *target states* in their definition.

The reachability condition stipulates that T should be visited at least once,

$$\mathrm{Reach} = \{\exists n \in \mathbb{N}, K_n \in T\} \ .$$

The safety condition is complementary:

$$\mathrm{Safe} = \{\forall n \in \mathbb{N}, K_n \notin T\} \ .$$

For the Büchi condition the set of target states has to be visited infinitely often,

$$\mathrm{Büchi} = \{\forall m \in \mathbb{N}, \exists n \geq m, K_n \in T\} \ .$$

And the co-Büchi condition is complementary:

$$\mathrm{CoBüchi} = \{\exists m \in \mathbb{N}, \forall n \geq m, K_n \notin T\} \ .$$

An example Consider the game depicted on Fig. 1. The objective of player 1 is to

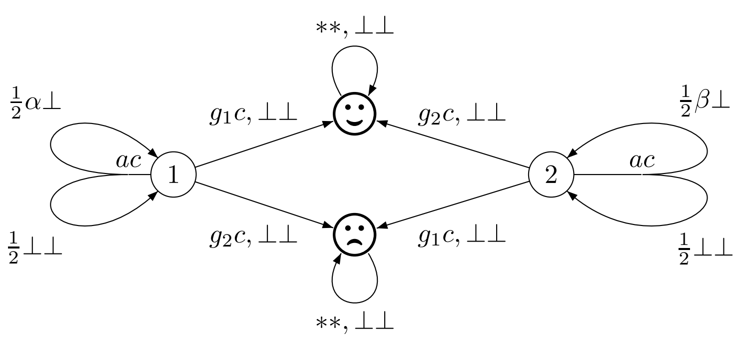

Fig. 1. When the initial state is chosen uniformly at random between states 1 and 2, Player 1 has a strategy to reach ☺ almost surely.

reach the ☺-state. The initial distribution is $\delta(1) = \delta(2) = \frac{1}{2}$ and $\delta(☺) = \delta(☹) = 0$. Player 1 plays with actions $I = \{a, g_1, g_2\}$, where g_1 and g_2 mean respectively 'guess 1' and 'guess 2', while player 2 plays with actions $J = \{c\}$ (that is, player 2 has no choice but playing always c). Player 1 receives signals $C = \{\alpha, \beta, \perp\}$ and player 2 is 'blind', she always receives the same signal $D = \{\perp\}$. This game can thus be seen as a one-player stochastic game with signals.

Transitions probabilities represented on Fig. 1 are interpreted in the following way. When the game is in state 1, player 1 plays a and player 2 plays c, then player 1 receives signal α or $\perp$, each with probability $\frac{1}{2}$, player 2 receives signal $\perp$ and the game stays in state 1. In state 2 when action of player 1 is a and action of player 2 is c, player 1 cannot receive signal α but instead she may receive signal β. When 'guessing the state' i.e. playing action g_i in state $j \in \{1, 2\}$, player 1 wins the game if $i = j$ (she guesses the correct state) whereas if $i \neq j$ player 1 loses the game which is stuck in the absorbing state ☹. The star symbol $*$ stands for any action.

In this game, player 1 has a strategy to reach ☺ almost surely. Her strategy is to keep playing action a as long as she keeps receiving signal $\perp$. The day player 1 receives signal α or β, she plays respectively action g_1 or g_2. This strategy is almost-surely winning because the probability for player 1 to receive signal $\perp$ forever is 0. This almost-surely winning strategy can be represented by a deterministic finite-memory strategy with three memory states $M = \{m_a, m_1, m_2\}$, whose initial mapping is constant equal to the Dirac distribution on m_a and whose transitions are depicted on Figure 2.

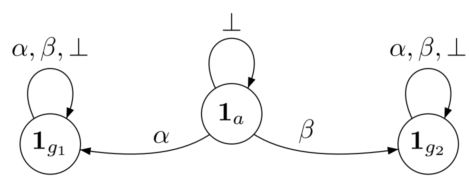

Fig. 2. A three states almost-surely winning deterministic finite memory strategy for the game of Figure 1. The initial distribution is the Dirac distribution on the middle state. States are labelled by the distribution to be played, all of them are Dirac distributions on this example.

Another example Interestingly, whether player 1 can win positively or not depends not only of his own signalling structure but also of the signaling structure of its opponent. This is illustrated by the game depicted on Figure 3.

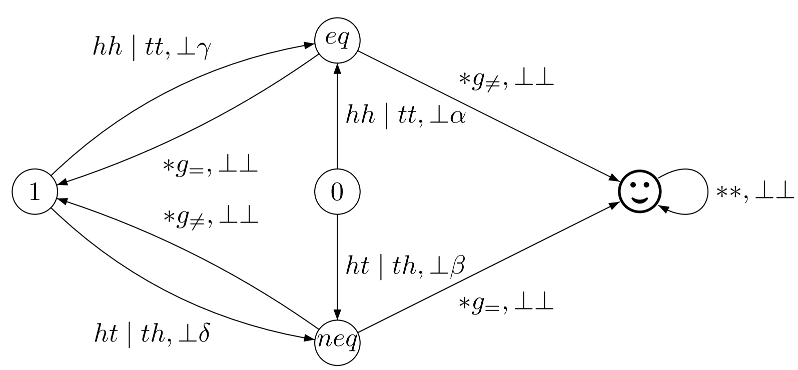

Fig. 3. Player 1 wins almost-surely, positively or not, depending on the signals for player 2.

The game starts in state 0, and both players choose heads (h) or tails (t). If they agree the game moves to state eq, otherwise to state neq. The behaviour is similar from state 1, but the signals received by player 2 might be different. Player 1 is blind and can only count the number of steps so far. The objective for player 1 is to reach the ☺-state, and she succeeds if player 2 makes a wrong guess: either he plays $g_{\neq}$ from state eq or he plays $g_{=}$ from states neq. Depending of the signals α, β, γ and δ, received by player 2, the game will be almost-surely winning, positively winning, or winning with probability zero for player 1.

Assume first that all signals α, β, γ and δ are distinct. Then, player 2 always knows when the play enters states eq and neq and can play accordingly, in order to avoid the ☺-state. Therefore player 2 has a surely winning strategy for her safety objective, and player 1 wins with probability 0.

Assume now that $\alpha = \beta$, but γ and δ are distinct. Informally, after the first move, player 2 cannot distinguish if the play is in state eq or neq. His best choice is then to play uniformly at random $g_=$ and $g_{\neq}$. Later, if the game reaches state 1, since $\gamma \neq \delta$, player 2 will be able to avoid the ☺-state, whatever player 1 does. For both players, in the first move, the best choice is to play uniformly at random heads or tails, so that in this case, player 1 wins with probability $1/2$.

Last, assume that $\alpha = \beta$ and $\gamma = \delta$, so that player 2 can never distinguish between states eq or neq. The best strategy for player 1 is to always choose uniformly at random heads or tails. Against this strategy, and whatever player 2 does, every other move, the probability is half to move to the ☺-state, so that player 1 wins almost-surely.

3. DETERMINACY AND BELIEFS

In this section, we introduce the important tools that will be relevant in the sequel of the paper. We start with the notion of qualitative determinacy, and compare it to the more common value determinacy. Then, we introduce beliefs and beliefs of beliefs that formalize the knowledge players have while playing the game.

3.1. Qualitative determinacy vs value determinacy.

If an initial distribution is positively winning for player 1 then by definition it is *not* almost-surely winning for his opponent player 2. A natural question is whether the converse implication holds.

Definition 3.1 (*Qualitative determinacy*). A winning condition Win is *qualitatively determined* if for every stochastic game with signals equipped with Win, every initial distribution is either almost-surely winning for player 1 or positively winning for player 2.

Qualitative determinacy is similar to but different from the usual notion of *(value) determinacy* which refers to the existence of a *value*. Actually both qualitative determinacy and value determinacy are formally expressed by a quantifier inversion. On one hand, qualitative determinacy rewrites as:

$$(\forall\sigma\ \exists\tau\ \mathbb{P}_\delta^{\sigma,\tau}(\mathrm{Win}) < 1) \implies (\exists\tau\ \forall\sigma\ \mathbb{P}_\delta^{\sigma,\tau}(\mathrm{Win}) < 1) \quad .$$

On the other hand, the game has a value if:

$$\sup_\sigma \inf_\tau \mathbb{P}_\delta^{\sigma,\tau}(\mathrm{Win}) \geq \inf_\tau \sup_\sigma \mathbb{P}_\delta^{\sigma,\tau}(\mathrm{Win}) \quad .$$

Both the converse implication of the first equation and the converse inequality of the second equation are obvious.

While *value determinacy* is a classical notion in game theory [Shapley 1953; Mertens and Neyman 1982], to our knowledge the notion of *qualitative determinacy* appeared only recently in the context of omega-regular concurrent games [de Alfaro et al. 2007; de Alfaro and Henzinger 2000] and stochastic games with perfect information [Horn 2008].

The existence of an almost-surely winning strategy ensures that the value of the game is 1, but the converse is not true. Actually it can even hold that player 2 has a positively winning strategy while at the same time the value of the game is 1. For example, consider the game depicted on Fig. 4, which is a slight modification of the one from Fig. 1 (only signals of player 1 and transitions probabilities differ). Player 1 has

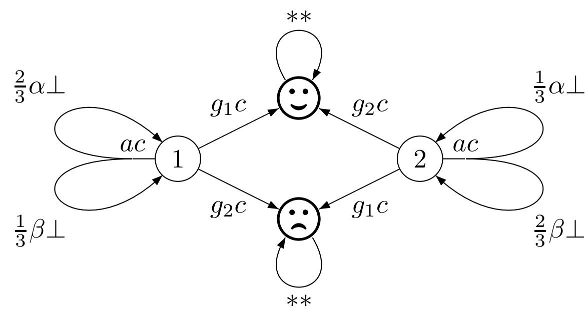

Fig. 4. A reachability game with value 1 where player 2 has a positively winning strategy.

signals $\{\alpha, \beta\}$ and similarly to the game on Fig. 1, her goal is to reach the target state ☺ by guessing correctly whether the initial state is 1 or 2. On one hand, player 1 can guarantee a winning probability as close to 1 as she wants: she plays a for a long time and compares how often she received signals α and β. If signal α was more frequent, then she plays action g_1, otherwise she plays action g_2. Of course, the longer player 1 plays a's the more accurate the prediction will be. On the other hand, the only strategy available to player 2 (always playing c) is positively winning, because any sequence of signals in $\{\alpha, \beta\}^*$ can be generated with positive probability from both states 1 and 2.

3.2. Beliefs and their properties

The proofs of our qualitative results are based on the notion of *belief* of a player, which is the set of possible states of the game, according to the signals received by the player. Since there are finitely many beliefs, using strategies based on beliefs and beliefs of beliefs is the key for the decidability of most of our decision problems.

Definition 3.2 (*Belief*). From an initial set of states $L \subseteq K$, the belief of player 1 after having played the lottery $\delta \in \Delta(I)$ and received signal c is:

$$\mathcal{B}_1(L, \delta, c) =$$
$$\{k \in K \mid \exists l \in L, i \in \mathrm{supp}(\delta), j \in J, d \in D \text{ such that } p(k, c, d \mid l, i, j) > 0\} \ .$$

Since $\mathcal{B}_1(L, \delta, c)$ only depends on the support of δ, for a non-empty set $I_0 \subseteq I$ we define $\mathcal{B}_1(L, I_0, c) = \mathcal{B}_1(L, \delta, c)$ where δ is any distribution with support I_0.

The belief of player 1 after having played lotteries with support $I_0, I_1, \ldots, I_n$ and received a sequence of signals $c_1, \ldots, c_n$ is defined inductively by:

$$\mathcal{B}_1(L, I_0, c_1, I_1, c_2, \ldots, I_n, c_n) = \mathcal{B}_1(\mathcal{B}_1(L, I_0, c_1, \ldots, I_{n-1}, c_{n-1}), I_n, c_n).$$

Beliefs of player 2 are defined similarly.

Given σ and τ and an initial distribution δ we denote $\mathcal{B}_1^n$ the random variable defined by $\mathcal{B}_1^n = \mathcal{B}_1(L, I_0, C_1, I_1, C_2, \ldots, I_n, C_n)$ where L is the support of δ and $I_k = \mathrm{supp}(\sigma(\delta_L, C_0, C_1, \ldots, C_n))$ is the support of the k-th lottery played by player 1.

To give a complete picture of stochastic games with signals, and to compare with existing work on games with imperfect information, we will at some places consider restricted classes of games, based on their signalling structures.

Definition 3.3 (*Particular signalling structures*). Player 2 is perfectly informed if $|\mathcal{B}_2(L, J_0, d_1, J_1, d_2, \ldots, J_n, d_n)| = 1$, for every $L, J_0, d_1, J_1, d_2, \ldots, J_n, d_n$.
He is better informed than player 1 if for every $L, J_0, d_1, J_1, d_2, \ldots, J_n, d_n$, $L', I_0, c_1, I_1, c_2, \ldots, I_n, c_n$, and every play $l_0, i_0, j_0, c_1, d_1, \cdots c_n, d_n$ with $l_0 \in L \subseteq L'$, and $i_k \in I_k$,$j_k \in J_k$ for all k, we have $\mathcal{B}_2(L, J_0, d_1, J_1, d_2, \ldots, J_n, d_n) \subseteq \mathcal{B}_1(L', I_0, c_1, I_1, c_2, \ldots, I_n, c_n)$.

Symmetric definitions hold for the other player.

The following lemma lists some useful properties of beliefs.

LEMMA 3.4. *For every strategies σ and τ and every initial distribution $\delta \in \Delta(K)$ and for every $n \in \mathbb{N}$,*

$$\mathbb{P}_\delta^{\sigma,\tau}\left(K_n \in \mathcal{B}_1^n\right) = 1 \ . \tag{4}$$

Assume that τ is the strategy τ_{rand} which always plays the uniform distribution over J. Then for every $n \in \mathbb{N}$ and $c_1 \cdots c_n \in C^*$, for every state $k \in K$,

$$\begin{aligned} \mathbb{P}_\delta^{\sigma,\tau_{\mathsf{rand}}}&\left(C_1 = c_1, \ldots, C_n = c_n\right) > 0 \\ &\wedge k \in \mathcal{B}_1(I_0, c_1, I_1, c_2, I_2, \ldots, c_n) \\ &\iff \mathbb{P}_\delta^{\sigma,\tau_{\mathsf{rand}}}\left(K_{n+1} = k, C_1 = c_1, \ldots, C_n = c_n\right) > 0 \ , \end{aligned} \tag{5}$$

where $I_0 = \mathrm{supp}(\delta)$, $I_1 = \mathrm{supp}(\sigma(\delta, c_1))$, $I_2 = \mathrm{supp}(\sigma(\delta, c_1, c_2)), \ldots$ are the supports of the lotteries played by player 1 when she sees signals $c_1, \ldots, c_n$.

Assume that σ is almost-surely winning for player 1 from δ for a Büchi or a co-Büchi game. Then for every $n \in \mathbb{N}$ and strategy τ,

$$\mathbb{P}_\delta^{\sigma,\tau}\left(\mathcal{B}_1^n \textit{ is almost surely winning for player } 1\right) = 1 \ . \tag{6}$$

PROOF. To prove (4) for $n = 1$ we use the fact that for every state $k \in K$ and signal $c \in C$,

$$\left(\mathbb{P}_\delta^{\sigma,\tau}\left(K_1 = k \wedge C_1 = c\right) > 0\right) \implies \left(k \in \mathcal{B}_1(\mathrm{supp}(\delta), \sigma(\delta), c)\right) \ , \tag{7}$$

which is straightforward using (1) and the definition of beliefs. The proof of (4) for any $n \in \mathbb{N}$ goes along the same lines.

Property (5) is straightforward to prove in one-player games i.e. in games where player 2 has only one action available. Proving (5) in games where player 2 plays τ_{rand} is similar.

We prove (6) by contradiction. Otherwise there exists a support L which is not almost surely winning and $n \in \mathbb{N}$ such that $\mathbb{P}_\delta^{\sigma,\tau}\left(\mathcal{B}_1^n = L\right) > 0$. But then let τ' be the strategy for player 2 that plays like τ the first n steps and then forget all signals of player 2 and switches to a strategy positively winning for player 2 from the initial uniform distribution δ_L over L. Then τ' wins positively against strategy σ, a contradiction. □

We will also rely on the notion of *belief of belief*, called here *2-belief*, which, roughly speaking, represents for one player the set of possible beliefs for his (or her) adversary.

Definition 3.5 (*2-Belief*). From an initial set $\mathcal{L} \subseteq K \times \mathcal{P}(K)$ of pairs composed of a state and a belief for player 2, the 2-belief of player 1 after having played a lottery with support I_0, and received signal c is the subset of $K \times \mathcal{P}(K)$ defined by:

$$\begin{aligned} \mathcal{B}_1^{(2)}(\mathcal{L}, I_0, c) = \{(k, \mathcal{B}_2(L, J_0, d)) \mid\ &\exists (l, L) \in \mathcal{L}, \exists J_0 \subseteq J \\ &\exists i \in I_0, j \in J_0 \text{ such that } p(k, c, d \mid l, i, j) > 0\} \ . \end{aligned}$$

From an initial set $\mathcal{L} \subseteq K \times \mathcal{P}(K)$ of pairs composed of a state and a belief for player 2, the 2-belief of player 1 after having played lotteries with support $I_0, I_1, \ldots, I_n$ and received a sequence of signals $c_1, \ldots, c_n$ is defined inductively by:

$$\mathcal{B}_1^{(2)}(\mathcal{L}, I_0, c_1, I_1, c_2, \ldots, I_n, c_n) = \mathcal{B}_1^{(2)}\left(\mathcal{B}_1^{(2)}\left(\mathcal{L}, I_0, c_1, \ldots, I_{n-1}, c_{n-1}\right), I_n, c_n\right) .$$

In the sequel we use 2-beliefs to build positively winning strategies for player 2 in Büchi games.

Finally let us present a property that describes the effect of shifting time on the probability measure induced by two strategies; we use this *shifting lemma* several times in our proofs.

LEMMA 3.6 (SHIFTING LEMMA). *Let $f : K^\omega \to \{0, 1\}$ be the indicator function of a measurable event, δ be an initial distribution and σ and τ two strategies. Then:*

$$\mathbb{P}_\delta^{\sigma,\tau}\left(f(K_1, K_2, \ldots) = 1 \mid C_1 = c, D_1 = d\right) = \mathbb{P}_{\delta_{cd}}^{\sigma_c,\tau_d}\left(f(K_0, K_1, \ldots) = 1\right),$$

where $\forall k \in K, \delta_{cd}(k) = \mathbb{P}_\delta^{\sigma,\tau}\left(K_1 = k \mid C_1 = c, D_1 = d\right)$ and $\sigma_c(\delta_{cd}, c_2c_3\cdots c_n) = \sigma(\delta, cc_2c_3\cdots c_n)$ and $\tau_d(\delta_{cd}, d_2d_3\cdots d_n) = \sigma(\delta, dd_2d_3\cdots d_n)$.

PROOF. Using the definition of the probability measure $\mathbb{P}_\delta^{\sigma,\tau}()$, this holds when f is the indicator function of a finite union of cylinders, and the class of events that satisfy this property is a monotone class. □

Based on the notions of belief and 2-beliefs, we introduce the following families of strategies with finite memory, that will be sufficient to win stochastic games with signals either positively or almost-surely.

Definition 3.7 (*Belief strategies and 2-belief strategies*). A *belief strategy* of player 1 is a strategy where the memory states are $\mathcal{P}(K) \times \mathcal{P}(I)$ and the action of the update function on the component $\mathcal{P}(K)$ is the belief $\mathcal{B}_1$ of 1.

A 2-*belief strategy* of player 1 is a strategy where the memory states are $\mathcal{P}(K \times \mathcal{P}(K))$ and the update function is the 2-belief $2\mathcal{B}_1$ of 1.

4. MAIN RESULTS.

In this section we state our main contributions, and the proofs can be found in the next Sections.

4.1. Qualitative Determinacy.

We start with qualitative determinacy, which holds, except for co-Büchi objectives.

THEOREM 4.1. *Stochastic games with signals with reachability, safety and Büchi winning conditions are qualitatively determined.*

The proof can be found in Section 5.

Since reachability and safety games are dual, a consequence of Theorem 4.1, is that in a reachability game, every initial distribution is either almost-surely winning for player 1, *almost-surely* winning for player 2, or positively winning for both players. If player 2 wins almost-surely a reachability game, it trivially implies that the safety condition is satisfied by all consistent plays, in other words player 2 wins surely.

Büchi games do not share this nice feature because co-Büchi games are not qualitatively determined. An example of a co-Büchi game which is not determined will be discussed in Section 5.2.

4.2. Algorithmic complexity of deciding the winner.

We now turn to the result concerning the (time) complexity to decide stochastic games with signals, starting with the easy case of safety games.

PROPOSITION 4.2. *In a safety game, deciding whether player 1 has an almost-surely winning strategy from a given initial distribution is EXPTIME-complete. If player 1 is perfectly informed, the above decision problem is in PTIME.*

Proposition 4.2 is easy to establish after observing that from an initial distribution where player 1 wins almost surely, player 1 can also win *surely*. We then obtain the result by applying [Reif 1979; Chatterjee et al. 2007] which tackle surely winning in partially observable games.

Beside the determinacy result stated in Theorem 4.1, the main contribution of this article concerns the complexity of deciding reachability and Büchi games, for which we will establish the following theorem:

THEOREM 4.3. *In reachability and Büchi games deciding whether player 1 has an almost-surely winning strategy from a given initial distribution is 2EXPTIME-complete.*

Concerning winning positively a *safety or co-Büchi game*, one can use Theorem 4.1 and the determinacy property: player 2 has a positively winning strategy in the above game if and only if player 1 has no almost-surely winning strategy. Therefore, deciding when player 2 has a positively winning strategy can also be done, with the same complexity. The proof of the upper bound of Theorem 4.3 can be found in Section 6. The lower bound can be found in Theorem 8.1.

For particular signalling structures, the complexity is better than 2EXPTIME, as established in [Chatterjee et al. 2007] when player 2 is perfectly informed. This reduced complexity holds for even less restricted cases, namely when player 1 is perfectly informed, or when player 2 is better informed than player 1 (see Definition 3.3).

THEOREM 4.4. *For reachability and Büchi games where player 1 is perfectly informed or player 2 is better informed than player 1, deciding whether player 1 has an almost-surely winning strategy from a given initial distribution is EXPTIME-complete.*

The upper bound in Theorem 4.4 is shown in Proposition 8.2. The winning states can be computed by the same fix-point algorithm used for Theorem 4.3 without any change. The lower bound derives from [Chatterjee et al. 2007].

4.3. Complexity of strategies

Then, we describe the number of memory states needed by strategies to win.

The doubly exponential complexity of Theorem 4.3 is surprising. The main explanation to this complexity is that a player may need doubly exponential memory to win positively. As a matter of fact, algorithmic complexity of these games is highly related to the memory needed by winning strategies: finite memory is sufficient to win every decidable game we consider in this paper. We give the precise tight memory requirements in Table 5.

First, as already mentioned for Proposition 4.2, almost surely winning safety games is equivalent with surely winning safety games. Hence, results of [Reif 1979; Chatterjee et al. 2007] can be applied, giving the exponential upper-bound for the memory size needed for (almost-)surely winning safety games. More precisely, belief-based strategies are sufficient to win (almost-)surely safety games.

The upper-bound on memory for almost-surely winning reachability and Büchi games can be derived from the proof of the determinacy of reachability and Büchi

	Almost-surely	**Positively**
Reachability	exponential	memoryless
Safety	exponential	doubly-exponential
Büchi	exponential	*not computable*
Co-Büchi	*not computable*	doubly-exponential

Fig. 5. Tight memory requirement for strategies.

games (see Corollary 6.1). Here again, belief-based strategies are sufficient to win almost-surely reachability and Büchi games. This is not very surprising since similar strategies were used in [Chatterjee et al. 2007] where this result was used for games where player 2 has perfect information.

PROPOSITION 4.5 (BELIEF STRATEGIES ARE SUFFICIENT TO WIN ALMOST-SURELY). *In safety, reachability and Büchi games, a player has an almost-surely winning belief strategy from states where this player has an almost-surely winning strategy. There are games with signals for which strategies with an exponential number of memory states are necessary for a player to win almost-surely.*

We now turn to the memory needed to win positively. First, memoryless strategies playing uniformly at random are sufficient to win positively reachability games.

The element of surprise is the amount of memory needed for winning positively co-Büchi and safety games. In these situations, it is still enough for a player to use a strategy with finite-memory but, surprisingly perhaps, an exponential memory size is not enough to win positively. Actually, one can derive from the proof of Theorem 6.4 in section 6.3 that doubly-exponential size memory is sufficient to win positively a safety or co-Büchi objectives (for player 2). More precisely, 2-belief strategies are sufficient for positively winning safety and co-Büchi games. This result cannot be derived from the memory requirements for player 1 to win almost-surely, nor from the work in [Gripon and Serre 2009]. The lower bound on memory for winning positively a safety or co-Büchi game is established in Proposition 7.1.

THEOREM 4.6 (2-BELIEF STRATEGIES ARE SUFFICIENT TO WIN POSITIVELY). *In safety and co-Büchi games, a player has a positively winning 2-belief strategy from states where this player has a positively winning strategy. There are games with signals for which strategies with a doubly exponential number of memory states are necessary for a player to win positively.*

Last, the undecidability results for positively winning Büchi games or almost-surely winning co-Büchi games is a consequence of [Baier et al. 2008] and [Chatterjee et al. 2010]; in both cases, memory considerations are less relevant.

5. DETERMINACY OF STOCHASTIC GAMES WITH SIGNALS.

5.1. Determinacy of reachability, Büchi and safety games.

The goal of this subsection is to prove Theorem 4.1, that states the qualitative determinacy of reachability, Büchi and safety games. Notice that the qualitative determinacy of Büchi games implies the qualitative determinacy of reachability games, since any reachability game can be turned into an equivalent Büchi one by making all target states absorbing. Qualitative determinacy of safety games is rather easy to establish, so we omit the proof here. Proving qualitative determinacy of Büchi games is harder and we provide full details.

5.1.1. The canonical strategy. We define a canonical strategy σ_{can} for player 1 and prove that this strategy is almost-surely winning from any initial distribution which is not positively winning for player 2. This strategy is quite simple to define, as follows. Let $\mathcal{L} \subseteq \mathcal{P}(K)\backslash\{\emptyset\}$ be the set of supports that are positively winning for player 2, and let $\overline{\mathcal{L}} = \mathcal{P}(K) \setminus (\mathcal{L} \cup \{\emptyset\})$, then for every $L \in \overline{\mathcal{L}}$ we define the set

$$SS(L) = \left\{ I' \subseteq I \mid I' \neq \emptyset \wedge \forall c \in C, \mathcal{B}_1(L, I', c) \in \{\emptyset\} \cup \overline{\mathcal{L}} \right\} \ ,$$

and σ_{can} is the strategy that picks actions in $SS(B)$ at random, when B is the current belief of player 1.

Definition 5.1. The canonical strategy σ_{can} is a finite-memory strategy whose memory states are:

$$M = \{(B, I') \mid B \in \overline{\mathcal{L}} \wedge I' \in SS(B)\} \ .$$

The strategy σ_{can} uses the memory states to keep track of the current belief of player 1 and of the next set of actions she should play. Initially from a distribution δ with support $L \in \overline{\mathcal{L}}$, the memory state is (L, I') with I' chosen uniformly at random in $SS(L)$. When the memory state is (B, I') then player 1 plays the uniform distribution $\sigma_{\mathsf{can}}(B, I') = \delta_{I'}$ over I'. When player 1 receives some signal $c \in C$ then the new memory state is chosen uniformly at random in $\{B'\} \times SS(B')$ where $B' = \mathcal{B}_1(B, I', c)$.

Some details are needed to prove that the canonical strategy is well-defined, because we have to establish that the set of memory states M is stable by the function updating memory states, i.e. the first component B stays in $\overline{\mathcal{L}}$ and that the update of the memory state is always possible i.e. $\{B'\} \times SS(B')$ is never empty.

First, let $B' = \mathcal{B}_1(B, I', c)$ in Definition 5.1. Then $B' \in \overline{\mathcal{L}} \cup \{\emptyset\}$ because of the very definition of $SS(L)$ and $B' \neq \emptyset$ because B' is equal to the belief of player 1 hence according to (4) in Lemma 3.4 it is never empty.

As a consequence, the beliefs of player 1 when playing σ_{can} are always in $\overline{\mathcal{L}}$, and for every distribution δ whose support is in $\overline{\mathcal{L}}$,

$$\forall \tau, \mathbb{P}_{\delta}^{\sigma_{\mathsf{can}}, \tau} \left(\forall n \in \mathbb{N}, \mathcal{B}_1^n \in \overline{\mathcal{L}} \right) = 1 \ . \tag{8}$$

Second, $SS(B') \neq \emptyset$ in Definition 5.1 because of the following lemma.

LEMMA 5.2.

$$\forall L \in \overline{\mathcal{L}}, SS(L) \neq \emptyset \ . \tag{9}$$

PROOF. The proof is by contradiction, assume that $SS(L) = \emptyset$ for some $L \in \overline{\mathcal{L}}$. Then in particular the singleton sets are not in $SS(L)$ thus for every action $i \in I, \{i\} \notin SS(L)$. That is, for every action $i \in I$ there exists a signal $c \in C$ such that $\mathcal{B}_1(L, i, c) \neq \emptyset$ and $\mathcal{B}_1(L, i, c) \in \mathcal{L}$. Then by definition of beliefs:

$$\exists l \in L, \exists j \in J, \text{ such that } \sum_{k \in K, d \in D} p(k, c, d \mid l, i, j) > 0 \ .$$

Let τ_{rand} be the strategy of player 2 which plays always the uniform distribution over all actions of player 2. We prove:

$$\forall \sigma, \mathbb{P}_{\delta_L}^{\sigma, \tau_{\mathsf{rand}}} \left(\mathcal{B}_1^1 \in \mathcal{L} \right) > 0 \ . \tag{10}$$

Let σ a strategy for player 1 and $I' = \mathrm{supp}(\sigma(\delta_L))$ and $i \in I'$. Since $\tau_{\mathsf{rand}}(j) > 0$ for every j there is non-zero probability that player 1 receives a signal c such that $\mathcal{B}_1(L, i, c) \in \mathcal{L}$. Since $i' \in I$ and by definition of beliefs, $\mathcal{B}_1(L, i, c) \subseteq \mathcal{B}_1(L, I', c)$ and since $\mathcal{L}$ is upward-closed we get $\mathcal{B}_1(L, I', c) \in \mathcal{L}$. This proves (10).

To get the contradiction, we define a strategy τ' for player 2 which is positively winning from δ_L. Let τ' be the strategy for player 2 which plays τ_{rand} for the first round then selects at random some non-empty support $B \in \mathcal{L}$ and plays from the second round till the end a positively winning strategy τ_B from the initial distribution δ_B, formally $\tau'(\delta_L) = \delta_J$ and $\tau'(\delta_L, d_1 d_2 \cdots d_n) = \tau_B(\delta_B, d_2 \cdots d_n)$. According to (10), there exists $B \in \mathcal{L}$ and $c \in C$ such that :

$$\mathbb{P}^{\sigma,\tau'}_{\delta_L}\left(\mathcal{B}^1_1 = B, C_1 = c\right) > 0 \tag{11}$$

and we fix such a B and c. Since τ_B is positively winning from δ_B, there exists $l_c \in B$ such that:

$$\mathbb{P}^{\sigma_c,\tau_B}_{\delta_B}(\text{Win} \mid K_0 = l_c) > 0\ . \tag{12}$$

By definition of τ', $\mathbb{P}^{\sigma,\tau_{\mathsf{rand}}}_{\delta_L}\left(\mathcal{B}^1_1 = B\right) = \mathbb{P}^{\sigma,\tau'}_{\delta_L}\left(\mathcal{B}^1_1 = B\right) > 0$ by (11) hence according to (5) in Lemma 3.4 we get

$$\mathbb{P}^{\sigma,\tau'}_{\delta_L}\left(K_1 = l_c \mid \mathcal{B}^1_1 = B\right) > 0\ . \tag{13}$$

Since Büchi and coBüchi winning conditions are invariant by shifting, the shifting lemma and the definition of τ' implies:

$$\mathbb{P}^{\sigma,\tau'}_{\delta_L}\left(\text{Win} \mid \mathcal{B}^1_1 = B\right) \geq \frac{1}{|\mathcal{L}|}\mathbb{P}^{\sigma_c,\tau_B}_{\delta_c}\left(\text{Win} \mid \mathcal{B}^1_1 = B\right) \cdot \mathbb{P}^{\sigma,\tau_{\mathsf{rand}}}_{\delta_L}\left(C_1 = c \mid \mathcal{B}^1_1 = B\right), \tag{14}$$

using the same notations than in the shifting lemma and $\delta_c(k) = \mathbb{P}^{\sigma,\tau_{\mathsf{rand}}}_{\delta_L}(K_1 = k \mid C_1 = c)$. We prove that the right handside of the above inequality is strictly positive. According to (11) and (13), $\delta_c(l_c) > 0$, hence according to (12), $\mathbb{P}^{\sigma_c,\tau_B}_{\delta_c}\left(\text{Win} \mid \mathcal{B}^1_1 = B\right) > 0$. Moreover according to (11), $\mathbb{P}^{\sigma,\tau_{\mathsf{rand}}}_{\delta_L}\left(C_1 = c \mid \mathcal{B}^1_1 = B\right) > 0$ hence the right handside of (14) is strictly positive and $\mathbb{P}^{\sigma,\tau'}_{\delta_L}\left(\text{Win} \mid \mathcal{B}^1_1 = B\right) > 0$ and with (11) we deduce $\mathbb{P}^{\sigma,\tau'}_{\delta_L}(\text{Win}) > 0$. Since this holds for every strategy σ, the strategy τ' is positively winning from support δ_L thus $L \in \mathcal{L}$, a contradiction with $L \in \overline{\mathcal{L}}$. This completes the proof of (9). $\square$

Now that σ_{can} is well-defined, we can complete the proof of Theorem 4.1.

5.1.2. Proof of Theorem 4.1. Recall that we consider a Büchi winning condition. We have to show that the canonical strategy σ_{can} is almost-surely winning from every support in $\overline{\mathcal{L}}$.

The first step is to prove by contradiction that for every $L \in \overline{\mathcal{L}}$,

$$\forall k \in L, \forall \tau, \mathbb{P}^{\sigma_{\mathsf{can}},\tau}_{\delta_L}(\text{Safe} \mid K_0 = k) < 1\ . \tag{15}$$

Otherwise there would be some $L \in \overline{\mathcal{L}}$ and $k \in L$ and strategy τ such that

$$\mathbb{P}^{\sigma_{\mathsf{can}},\tau}_{\delta_L}(\text{Safe} \mid K_0 = k) = 1\ . \tag{16}$$

Now we show how to use τ to build a strategy positively winning from L, which will contradict the definition of $\overline{\mathcal{L}}$. Of course τ itself is not necessarily a positively winning strategy from L because it is positively winning against σ_{can} but there is no reason for τ to be positively winning against all strategies σ. Instead we define a strategy τ' as follows. Let τ' be the finite-memory strategy that tosses a fair coin at the beginning of the play and:

- plays like τ forever if the result of this initial coin toss is head,
- otherwise at each step τ' tosses a fair coin and as long as the result is tail τ' plays the uniform distribution over J. If the result of one of the coin tosses is head, τ' switches to a new behaviour: τ' selects randomly a support $B \in \overline{\mathcal{L}}$ and a strategy τ_B positively winning from B, forgets past signals and switches definitively to τ_B.

Now that τ' is defined, we prove it is positively winning from L. Let E be the event $E = \{\forall n \in \mathbb{N}, I_n \in SS(\mathcal{B}_1^n)\}$, where $I_n = \mathrm{supp}(\sigma(\delta_L, C_1 \cdots C_n))$ is the set of actions possibly played by σ at step n. Then for every strategy σ:

- Either $\mathbb{P}_{\delta_L}^{\sigma,\tau'}(E \mid K_0 = k) = 1$. In this case we prove that $\mathbb{P}_{\delta_L}^{\sigma,\tau'}(\mathrm{Safe} \mid K_0 = k) \geq \frac{1}{2}$, because for every finite play p,
$$\left(\mathbb{P}_{\delta_L}^{\sigma,\tau}(p \mid K_0 = k) > 0\right) \implies \left(\mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau}(p \mid K_0 = k) > 0\right) \implies (p \in \mathrm{Safe}) \ ,$$
where the first implication holds because by definition of σ_{can} and E, all actions played by σ with positive probability can be played by σ as well, while the second implication is from (16). This implies $\mathbb{P}_{\delta_L}^{\sigma,\tau}(\mathrm{Safe} \mid K_0 = k) = 1$ thus we get $\mathbb{P}_{\delta_L}^{\sigma,\tau'}(\mathrm{Safe} \mid K_0 = k) \geq \frac{1}{2}$ by definition of τ' (which coincides with τ when the initial coin toss is head).
- Or $\mathbb{P}_{\delta_L}^{\sigma,\tau'}(E \mid K_0 = k) < 1$. Then by definition of E there exists $n \in \mathbb{N}$ such that $\mathbb{P}_{\delta_L}^{\sigma,\tau'}(I_n \notin SS(\mathcal{B}_1^n) \mid K_0 = k) > 0$ then by definition of τ_{rand} and $SS()$ there exists $B \in \mathcal{L}$ such that $\mathbb{P}_{\delta_L}^{\sigma,\tau_{\mathsf{rand}}}\left(\mathcal{B}_1^{n+1} = B \mid K_0 = k\right) > 0$. By definition of τ', there is a possibility that the initial coin toss is tail and moreover the random moment and random support chosen by player 2 are precisely $n+1$ and B. As a consequence we get that $\mathbb{P}_{\delta_L}^{\sigma,\tau'}(\neg\,\mathrm{Win} \mid K_0 = k) > 0$.

In both cases, for every σ, $\mathbb{P}_{\delta_L}^{\sigma,\tau'}(\neg\,\mathrm{Win} \mid K_0 = k) > 0$ and since $k \in L$, $\mathbb{P}_{\delta_L}^{\sigma,\tau'}(\neg\,\mathrm{Win}) > 0$, thus τ' is positively winning from L and since $L \in \overline{\mathcal{L}}$, this contradicts the definition of $\overline{\mathcal{L}}$. As a consequence we get (15) by contradiction.

The second step to establish that σ_{can} is almost-surely winning is to prove that there exists $N \in \mathbb{N}$ such that:
$$\forall L \in \overline{\mathcal{L}}, \forall k \in L, \forall \tau, \mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau}(\exists n \leq N, K_n \in T \mid K_0 = k) > \frac{1}{N} \ . \tag{17}$$
We proceed by contradiction. Otherwise for every $N \in \mathbb{N}$ there exists some $L_N \in \overline{\mathcal{L}}, k_N \in L_N$ and a strategy τ_N such that
$$\mathbb{P}_{\delta_{L_N}}^{\sigma_{\mathsf{can}},\tau_N}(\exists n \leq N, K_n \in T \mid K_0 = k_N) \leq \frac{1}{N} \ . \tag{18}$$
Since $\overline{\mathcal{L}}$ and L are finite we can assume w.l.o.g. that the sequences $(L_N)_N$ and $(k_N)_N$ are constant and equal to some L and k. By compacity of the set of strategies of player 2 (for the Borel topology) we can assume that τ_n converges to some strategy τ. According to (18), for every $N_0 \in \mathbb{N}$,
$$\forall N \geq N_0, \mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau_N}(\exists n \leq N_0, K_n \in T \mid K_0 = k) \leq \frac{1}{N} \ ,$$
since the event $\{\exists n \leq N_0, K_n \in T\}$ only depends on a cylinder of length N_0, and by definition of the convergence of τ_N to τ we get:
$$\forall N \geq N_0, \mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau}(\exists n \leq N_0, K_n \in T \mid K_0 = k) \leq \frac{1}{N} \ ,$$
hence $\mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau}(\exists n \leq N_0, K_n \in T \mid K_0 = k) = 0$, and this holds for every $N_0 \in \mathbb{N}$ thus $\mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau}(\mathrm{Safe} \mid K_0 = k) = 1$ which contradicts (15). By contradiction we get (17).

According to (8) the belief of player 1 when using σ_{can} is always in $\overline{\mathcal{L}}$, thus according to (4)of Lemma 3.4 and the fact that $\overline{\mathcal{L}}$ is downward-closed we get that $\forall N_0 \in \mathbb{N}$ and

$c_1 \ldots c_{N_0} \in C^*$ and $d_1 \ldots d_{N_0} \in D^*$ and $k \in K$:

$\left\{k \in K \mid \mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau}\left(K_{N_0} = k \mid C_1 \cdots C_{N_0} = c_1 \ldots c_{N_0} \wedge D_1 \cdots D_{N_0} = d_1 \ldots d_{N_0}\right) > 0\right\} \in \overline{\mathcal{L}} \cup \{\emptyset\}$.

Together with the shifting Lemma and (17), for every $L \in \overline{\mathcal{L}}$ and $k \in L$ and strategy τ, $\forall N_0 \in \mathbb{N}$,

$$\mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau}\left(\exists n \in \mathbb{N} \text{ such that } N_0 \leq n \leq N_0 + N \text{ and } K_n \in T \mid K_0, K_1, \ldots, K_{N_0}\right) > \frac{1}{N} \ ,$$

which implies

$$\forall \tau, \mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau}\left(\forall n \leq k \cdot (N+1), K_n \notin T\right) \leq \left(1 - \frac{1}{N}\right)^k \ ,$$

and according to Borel-Cantelli Lemma

$$\forall L \in \overline{\mathcal{L}}, \forall \tau, \mathbb{P}_{\delta_L}^{\sigma_{\mathsf{can}},\tau}\left(\text{Büchi}\right) = 1 \ . \tag{19}$$

This completes the proof that supports not positively winning for player 2 are almost-surely winning for player 1, since σ_{can} is almost-surely winning from every support in $\bar{\mathcal{L}}$. □

5.2. Nondeterminacy of co-Büchi games.

In contrast with Büchi games, not all co-Büchi games are qualitatively determined: a counter-example is represented on Fig. 6. Similar examples can be used to prove that stochastic Büchi games with signals do not have a value [Renault and Sorin 2008]. In this game, player 1 observes everything, player 2 is blind (he only observes his own actions), and player 1's objective is to avoid the ☹-state from some moment on. The initial state is ☹.

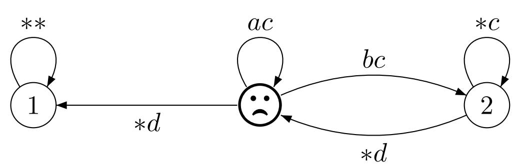

Fig. 6. Co-Büchi games are not qualitatively determined.

On one hand, player 1 does not have an almost-surely winning strategy for the co-Büchi objective. By contradiction, assume that there exists a strategy σ for player 1 that is almost-surely winning. To win against the strategy where player 2 plays c forever, with probability 1 the strategy σ should eventually play action b in order the play does not stay stuck in state ☹. Since σ is fixed, there exists a number n of rounds after which player 1 has played at least one b with probability arbitrarily close to 1. Consider the strategy of player 2 which plays c for the first $n-1$ actions, and then plays one d. Although player 2 is blind, obviously he can play such a strategy which requires only counting the number of actions he played since the beginning of the play. After $n-1$ actions played, with probability arbitrarily close to 1, the game is in state 2 and playing a d puts the game back in state ☹. Now, consider the following strategy for player 2: it plays c till the probability following σ to be in state 2 is $1-\varepsilon/2$, then it plays a d, then plays c till the probability following σ (assuming that state 1 was not reached) to be in state 2 is $1-\varepsilon/4$, etc, playing c to reach state 2 with probability

$1-\varepsilon/2^i$ after the $i-1$-th d player 2 played. The probability under these two strategies to eventually reach state 1 is $\sum_{i=1}^{\infty}\varepsilon/2^i = \varepsilon$. That is, there is at most probability ε to achieve the co-Büchi objective to avoid ☹ eventually. A contradiction with σ is almost-surely winning.

On the other hand, player 2 does not have a positively winning strategy either. By contradiction, assume that such a positively winning strategy τ of player 2 exists. Recall that player 2 is blind hence the probability that strategy τ plays an action is independent of the strategy chosen by player 1. There are two cases. Either the probability that under τ, action d is eventually played at least once is 1. Then in this first case player 1 can defeat τ by playing always a and ensure state 1 to be reached with probability 1, a contradiction. Or in the second case, the probability that strategy τ plays d is $p < 1$. Then in this second case, consider the strategy σ_b of player 1 playing always b. As τ is winning with some probability $p' > 0$ against that strategy σ_b, the probability thay strategy τ plays infinitely often d is at least p'. Let $q = 1-p$ be the probability to never see d. We show that $q = 0$. As there is probability $p' > 0$ to see d infinitely often, there is a number n_1 such that there is probability at least $p'/2$ to see d among the first n_1 actions. For sure, $q \leq 1-p'/2$. Now d is also seen infinitely often with probability at least $p'/2$ after the n_1 first actions. There exists n_2 such that there is probability at least $p'/2$ to see at least one d among the actions between the n_1-th and the n_2-th. It implies that $q \leq (1-p'/2)\cdot(1-p'/2)$. We show by induction that $q \leq (1-p'/2)^i$ for all i, that is $q = 0$, a contradiction with $p = 1-q < 1$.

6. FIX-POINT ALGORITHMS.

6.1. Overview of the results and algorithms

As a corollary of the proof of qualitative determinacy (Theorem 4.1), we get a canonical strategy σ_{can} for player 1 (see Definition 5.1) to win almost-surely Büchi games.

COROLLARY 6.1. *If player 1 has an almost-surely winning strategy in a Büchi game then the canonical strategy σ_{can} is almost-surely winning and has memory states $\mathcal{P}(K)\times\mathcal{P}(I)$.*

The canonical strategy σ_{can} is defined as the finite memory strategy playing uniformly actions that surely lead to beliefs that are almost surely winning, and updating the memory uniformly at random to state (B', I') where, from belief B', all actions in I' also lead to almost surely winning beliefs. One can consider a family of canonical strategies where the set of winning beliefs is replaced with any subset X of $\mathcal{P}(K)$. A simple algorithm to decide for which player a game is winning can then be derived from Corollary 6.1: enumerate all such canonical strategies for $X \subseteq \mathcal{P}(K)$, and test each one to see if it is almost surely winning. As there is a doubly exponential number of sets $X \subseteq \mathcal{P}(K)$, this can be done in time doubly exponential. This is similar to the algorithm given in [Gripon and Serre 2009], although we use a different proof and we handle a larger class of games where players may not be able to observe the actions they play. This settles the upper bound for Theorem 4.3. The lower bounds are established in Theorem 8.1, proving that this enumeration algorithm is optimal for worst case complexity. While optimal in the worst case, this algorithm is however not efficient. For instance, if player 1 has no almost surely winning strategy, then this algorithm will enumerate every single of the doubly explonential many canonical strategies. Instead, we provide fix-point algorithms which do not enumerate every possible strategy in Theorem 6.3 for reachability games and Theorem 6.4 for Büchi games. Although they should perform better on games with particular structures, these fix-point algorithms still have a worst-case 2-EXPTIME complexity.

Remark 6.2. Notice that Corollary 6.1 only holds when memory can remember a stochastic choice, but it does not hold anymore if one restricts to finite-memory strategies with deterministic choice of actions and deterministic memory updates. Chatterjee and Doyen proved recently that with this restricted classes of finite-memory strategies, the memory needed may be a tower of exponential [Chatterjee and Doyen 2011]. In a nutshell, mimicking randomness with deterministic counters can be very costly.

We turn now to the (fix-points) algorithms which compute the set of supports that are almost-surely or positively winning for various objectives.

THEOREM 6.3 (DECIDING POSITIVE WINNING IN REACHABILITY GAMES). *In a reachability game each initial distribution δ is either positively winning for player 1 or surely winning for player 2, and this depends only on $\mathrm{supp}(\delta) \subseteq K$. The corresponding partition of $\mathcal{P}(K)$ is computable in time $\mathcal{O}\left(G \cdot 2^K\right)$, where G denotes the size of the description of the game.*

As detailed in subsection 6.2, the set of supports $\mathcal{L} \subseteq \mathcal{P}(K)$ surely-winning for player 2 are characterized as the largest fix-point of some monotonic operator $\Phi : \mathcal{P}(\mathcal{P}(K)) \to \mathcal{P}(\mathcal{P}(K))$. The operator Φ associates with $\mathcal{L} \subseteq \mathcal{P}(K)$ the set of supports $L \in \mathcal{L}$ that do not intersect target states and such that player 2 has an action which ensures that his next belief is in $\mathcal{L}$ as well, whatever action is chosen by player 1 and whatever signal player 2 receives:

$$\Phi(\mathcal{L}) = \{L \in \mathcal{L} \mid L \cap T = \emptyset \wedge \exists j \in J, \forall d \in D, \mathcal{B}_2(L, \{j\}, d) \in \emptyset \cup \mathcal{L}\} \ .$$

For $\mathcal{L} \subseteq \mathcal{P}(K)$, the value of $\Phi(\mathcal{L})$ is computable in time linear in $\mathcal{L}$ and in the description of the game, yielding the exponential complexity bound.

To decide whether player 1 wins almost-surely a Büchi game, we provide an algorithm which runs in doubly-exponential time. It uses the algorithm of Theorem 6.3 as a sub-procedure.

THEOREM 6.4 (DECIDING ALMOST-SURE WINNING IN BÜCHI GAMES). *In a Büchi game each initial distribution δ is either almost-surely winning for player 1 or positively winning for player 2, and this depends only on $\mathrm{supp}(\delta) \subseteq K$. The corresponding partition of $\mathcal{P}(K)$ is computable in time $\mathcal{O}\left(2^{2^G}\right)$, where G denotes the size of the description of the game. The algorithm computes at the same time the finite-memory strategies described in Corollary 6.1.*

The proof of Theorem 6.4 is detailed in subsection 6.3. We sketch here the main ideas.

First, suppose that from *every* support, player 1 can win the *reachability objective* with positive probability. Since this positive probability can be bounded from below (there are finitely many supports), repeating the same strategy can ensure that player 1 wins the Büchi condition with probability 1.

In general though player 1 is not in such an easy situation and there exists some support L which is *not* positively winning for him. Then by qualitative determinacy, the support L is winning surely for player 2 for his safety objective. In particular, it also means that player 2 has a surely winning strategy from L for her co-Büchi objective.

We prove that in case player 2 can *force with positive probability the belief of player* 1 to be L someday from another support L', then L' is positively winning as well for player 2. This is not completely obvious because in general player 2 cannot know exactly *when* the belief of player 1 is L (he can only compute the 2-Belief, letting him know all the possible beliefs player 1 can have). For winning positively from L', player 2 plays totally randomly until he guesses randomly that the belief of player 1 is L, at that moment he switches to a strategy surely winning from L. Such a strategy is far from being optimal, because player 2 plays randomly and in most cases he makes a wrong guess about the belief of player 1. However player 2 wins positively because there is a non zero probability that he guesses correctly at the right moment the belief of player 1.

Hence, player 1 should surely avoid her belief to be L or L' if she wants to win almost-surely. However, doing so player 1 may prevent the play from reaching target states, which may create another positively winning support for player 2, and so on. This is the basis of our fix-point algorithm.

Using these ideas, we prove that the set $\mathcal{L}_\infty \subseteq \mathcal{P}(K)$ of supports almost-surely winning for player 1 for the Büchi objective is the largest set of initial supports from which:

$$\text{player 1 has a strategy to win positively the reachability game while ensuring at the same time her belief to stay in } \mathcal{L}_\infty. \quad (\dagger)$$

Property (†) can be reformulated as a reachability condition in a new game whose states are states of the original game augmented with beliefs of player 1, kept hidden to player 2.

The fix-point characterization suggests the following algorithm for computing the set of supports positively winning for player 2: $\mathcal{P}(K)\setminus\mathcal{L}_\infty$ is the limit of the sequence $\emptyset = \mathcal{L}'_0 \subsetneq \mathcal{L}'_0 \cup \mathcal{L}''_1 \subsetneq \mathcal{L}'_0 \cup \mathcal{L}'_1 \subsetneq \mathcal{L}'_0 \cup \mathcal{L}'_1 \cup \mathcal{L}''_2 \subsetneq \ldots \subsetneq \mathcal{L}'_0 \cup \cdots \cup \mathcal{L}'_m = \mathcal{P}(K)\setminus\mathcal{L}_\infty$, where

(a) from supports in $\mathcal{L}''_{i+1}$ player 2 can surely guarantee the safety objective, under the hypothesis that player 1 beliefs stay outside $\mathcal{L}'_i$,
(b) from supports in $\mathcal{L}'_{i+1}$ player 2 can ensure with positive probability the belief of player 1 to be in $\mathcal{L}''_{i+1}$ someday, under the same hypothesis.

The overall strategy of player 2 positively winning for the co-Büchi objective consists in playing randomly for some time until he decides to pick up randomly a belief L of player 1 in some $\mathcal{L}''_i$. He forgets the signals he has received up to that moment and switches definitively to a strategy which guarantees (a). With positive probability, player 2 guesses correctly the belief of player 1 at the right moment, and future beliefs of player 1 will stay in $\mathcal{L}'_i$, in which case the co-Büchi condition holds and player 2 wins.

Property (†) can be formulated by mean of a fix-point according to Theorem 6.3, hence the set of supports positively winning for player 2 can be expressed using two nested fix-points. This should be useful for actually implementing the algorithm and for computing symbolic representations of winning sets.

6.2. Proof of Theorem 6.3

PROOF.
Let $\mathcal{L}_\infty \subseteq \mathcal{P}(K\setminus T)$ be the greatest fix-point of the monotonic operator $\Phi : \mathcal{P}(\mathcal{P}(K\setminus T)) \to \mathcal{P}(\mathcal{P}(K\setminus T))$ defined by:

$$\Phi(\mathcal{L}) = \{L \in \mathcal{L} \mid \exists J_L \subseteq J, \forall d \in D, \mathcal{B}_2(L, J_L, d) \in \mathcal{L} \cup \{\emptyset\}\} \ ,$$

in other words $\Phi(\mathcal{L})$ is the set of supports such that player 2 can play a lottery such that whatever signal d he might receive her new belief will still be in $\mathcal{L}$. Let σ_{rand} be the strategy for player 1 that plays randomly any action.

We are going to prove that:

(A) every support in $\mathcal{L}_\infty$ is surely winning for player 2,
(B) and σ_{rand} is positively winning from any support $L \subseteq K$ which is not in $\mathcal{L}_\infty$.

We start with proving (A). To win surely from any support $L \in \mathcal{L}_\infty$, player 2 uses the following finite-memory strategy τ: if the current belief of player 2 is $L \in \mathcal{L}_\infty$ then player 2 chooses a lottery with support J_L such that his next belief $\mathcal{B}_2(L, d)$ will be in $\mathcal{L}_\infty$ as well, whatever signal he receives. By definition of Φ and since $\mathcal{L}_\infty$ is a fix-point of Φ, there always exists such a support J_L, and this defines a finite memory strategy with memory $\mathcal{L}_\infty$ and update operator $\mathcal{B}_2$.

When playing with strategy τ, starting from a support in $\mathcal{L}_\infty$, the beliefs of player 2 stay in $\mathcal{L}_\infty$ and never intersect T because $\mathcal{L}_\infty \subseteq \mathcal{P}(K\backslash T)$. According to (4) in Lemma 3.4, this guarantees the play never visits T, whatever strategy is used by player 1.

We now prove (B). Remember that σ_{rand} is the memoryless strategy for player 1 that plays randomly any action. Once σ_{rand} is fixed, the game is a one-player game where only player 2 has choices to make: it is enough to prove (B) in the special case where the set of actions of player 1 is a singleton $I = \{i\}$. Let $\mathcal{L}_0 = \mathcal{P}(K\backslash T) \supseteq \mathcal{L}_1 = \Phi(\mathcal{L}_0) \supseteq \mathcal{L}_2 = \Phi(\mathcal{L}_1) \ldots$ and $\mathcal{L}_\infty$ be the limit of this sequence, the greatest fix-point of Φ. We prove that for any support $L \in \mathcal{P}(K)$, if $L \notin \mathcal{L}_\infty$ then:

$$L \text{ is positively winning for player } 1 \ . \tag{20}$$

If $L \cap T \neq \emptyset$, (20) is obvious. To deal with the case where $L \in \mathcal{P}(K\backslash T)$, we define for every $n \in \mathbb{N}$, $\mathcal{K}_n = \mathcal{P}(K\backslash T)\backslash\mathcal{L}_n$, and we prove by induction on $n \in \mathbb{N}$ that for every $L \in \mathcal{K}_n$, for every initial distribution δ_L with support L, for every strategy τ,

$$\mathbb{P}^{\tau}_{\delta_L}\left(\exists m, 2 \leq m \leq n+1, K_m \in T\right) > 0 \ . \tag{21}$$

For $n = 0$, (21) is obvious because $\mathcal{K}_0 = \emptyset$. Suppose that for some $n \in \mathbb{N}$, (21) holds for every $L \in \mathcal{K}_n$, and let $L \in \mathcal{K}_{n+1}$. If $L \in \mathcal{K}_n$ then by inductive hypothesis, (21) holds. Otherwise $L \in \mathcal{K}_{n+1}\backslash\mathcal{K}_n$ and by definition of $\mathcal{K}_{n+1}$,

$$L \in \mathcal{L}_n\backslash\Phi(\mathcal{L}_n) \ . \tag{22}$$

Let δ_L be an initial distribution with support L and τ a strategy for player 2. Let $J_0 \subseteq J$ be the suport of $\tau(\delta_L)$. According to (22), by definition of Φ, there exists a signal $d \in D$ such that $\mathcal{B}_2(L, J_0, d) \notin \mathcal{L}_n$. If $\mathcal{B}_2(L, J_0, d) \cap T \neq \emptyset$ then according to (5) in Lemma 3.4, $\mathbb{P}^{\tau}_{\delta_L}\left(K_2 \in T\right) > 0$. Otherwise $\mathcal{B}_2(L, J_0, d) \in \mathcal{P}(K\backslash T)\backslash\mathcal{L}_n = \mathcal{K}_n$ hence distribution $\delta_d(k) = \mathbb{P}^{\tau}_{\delta_L}\left(K_2 = k \mid D_1 = d\right)$ has its support in $\mathcal{K}_n$. By inductive hypothesis, for every strategy τ',

$$\mathbb{P}^{\tau'}_{\delta_d}\left(\exists m \in \mathbb{N}, 2 \leq m \leq n+1, K_m \in T\right) > 0$$

hence using the shifting lemma and the definition of δ_d,

$$\mathbb{P}^{\tau}_{\delta}\left(\exists m \in \mathbb{N}, 3 \leq m \leq n+2, K_m \in T\right) > 0 \ ,$$

which completes the proof of the inductive step.

To compute the partition of supports between those positively winning for player 1 and those surely winning for player 2, it is enough to compute the largest fix-point of Φ. Since Φ is monotonic, and each application of the operator can be computed in time linear in the size of the game (G) and the number of supports (2^K) the overall computation can be achieved in time $G \cdot 2^K$. To compute the strategy τ, it is enough to compute for each $L \in \mathcal{L}_\infty$ the support J_L which ensures $\mathcal{B}_2(L, J_L, d) \in \mathcal{L}_\infty$. $\square$

As a byproduct of the proof one obtains the following bounds on time and probabilities before reaching a target state, when player 1 uses the uniform memoryless strategy σ_{rand}. From an initial distribution positively winning for the reachability objective, for every strategy τ,

$$\mathbb{P}_{\delta}^{\sigma_{\mathsf{rand}},\tau}\left(\exists n \leq 2^{|K|}, K_n \in T\right) \geq \left(\frac{1}{p_{\min} \mid I \mid}\right)^{2^{|K|}} , \tag{23}$$

where $p_{\min}$ is the smallest non-zero transition probability.

6.3. Proof of Theorem 6.4

To establish Theorem 6.4, we start with formalizing what it means for player 1 to enforce her beliefs to stay outside a certain set.

Definition 6.5. Let $\mathcal{L} \subseteq \mathcal{P}(K)$ be a set of supports. We say that player 1 *can enforce her beliefs to stay outside* $\mathcal{L}$ if player 1 has a strategy σ such that for every strategy τ of player 2 and every initial distribution δ whose support is not in $\mathcal{L}$,

$$\mathbb{P}_{\delta}^{\sigma,\tau}\left(\forall n \in \mathbb{N}, \mathcal{B}_1^n \notin \mathcal{L}\right) = 1 \ . \tag{24}$$

Equivalently, for every $L \notin \mathcal{L}$, the set:

$$I(L) = \{I' \subseteq I \mid \forall c \in C, \mathcal{B}_1(L, I', c) = \emptyset \vee \mathcal{B}_1(L, I', c) \notin \mathcal{L}\} \ ,$$

is not empty.

PROOF. The equivalence is straightforward. In one direction, let σ be a strategy with the property (24), $L \notin \mathcal{L}$, δ_L a distribution with support L, and i an action such that $\sigma(\delta_L)(i) > 0$. Then according to (24), $\operatorname{supp}(\sigma(\delta_L)) \in I(L)$ hence $I(L)$ is not empty. In the other direction, if $I(L)$ is not empty for every $L \notin \mathcal{L}$ then consider the finite-memory strategy σ for player 1 which plays a lottery whose support is in $I(L)$ when the belief of player 1 is L. Then by definition of $I(L)$, property (24) holds. □

We need also the notion of $\mathcal{L}$-games.

Definition 6.6 (*$\mathcal{L}$-games*). Let $\mathcal{L}$ be an upward-closed set of supports such that player 1 can enforce her beliefs to stay outside $\mathcal{L}$. For every support $L \notin \mathcal{L}$, let $I(L)$ be the set of supports given by Definition 6.5. The $\mathcal{L}$-game has same actions, transitions and signals than the original partial observation game, only the winning condition changes: player 1 wins if the play reaches a target state and moreover player 1 is restrcticted to use lotteries whose support is in $I(L)$ whenever her belief is L. Formally given an initial distribution δ with support L and two strategies σ and τ the winning probability of player 1 is:

$$\mathbb{P}_{\delta}^{\sigma,\tau}\left(\exists n, K_n \in T \text{ and } \forall n, \operatorname{supp}(\sigma(\delta, C_1, \ldots, C_n)) \in I(\mathcal{B}_1^n)\right). \tag{25}$$

Note that strictly speaking, $\mathcal{L}$-games are not stochastic games because the winning probability of player 1 is not defined as $\mathbb{P}_{\delta}^{\sigma,\tau}(W)$ for some set W independent of σ and τ. However the definition of positively winning or almost-surely winning strategies is easy to extend to $\mathcal{L}$-games, using the very same conditions that for stochastic games.

Actually, winning positively an $\mathcal{L}$-game amounts to winning positively a reachability game with state space $\mathcal{P}(K) \times K$, as shown by the following lemma and its proof.

PROPOSITION 6.7 ($\mathcal{L}$-GAMES). *Let $\mathcal{L} \subseteq \mathcal{P}(K)$ be a set of supports such that $\mathcal{L}$ is upward-closed and player 1 can enforce her beliefs to stay outside $\mathcal{L}$.*

(i) In the $\mathcal{L}$-game, every support is either positively winning for player 1 or surely winning for player 2. We denote $\mathcal{L}''$ the set of supports that are not in $\mathcal{L}$ and are surely winning for player 2 in the $\mathcal{L}$-game.

(ii) Assume $\mathcal{L}''$ is empty. Then every support not in $\mathcal{L}$ is almost-surely winning for player 1, both in the $\mathcal{L}$-game and also for the Büchi objective in game G. Moreover, the

strategy $\sigma_{\mathcal{L}}$ for player 1 which consists in chosing randomly any support $I_0 \in I(L)$ and play the uniform lottery on I_0 when her belief is L is almost-surely winning in the $\mathcal{L}$-game.

(iii) Assume $\mathcal{L}''$ is not *empty. Then player 2 has a strategy τ to win surely the $\mathcal{L}$-game from any support in $\mathcal{L}''$, and τ has finite memory $\mathcal{P}((\mathcal{P}(K)\backslash\mathcal{L}) \times K)$.*

(iv) There is an algorithm running in time doubly-exponential time in the size of G to compute $\mathcal{L}''$ and, in case (iii) holds, strategy τ.

The proof is based on Theorem 6.3.

PROOF. We define a reachability game which is a synchronized product of the original game G with beliefs of player 1, with a few modifications.

This new reachability game is denoted $G_{\mathcal{L}}$. Intuitively in $G_{\mathcal{L}}$ player 1 is restricted to play uniform distributions over his set of actions in G and moreover when player 1 has a belief L, he is forced to choose among the set of actions in $I(L)$ because otherwise the play is sent to a losing sink state $\perp$. The state space is $K \times (\mathcal{P}(K)\backslash\mathcal{L}) \cup \{\perp\}$, where $\{\perp\}$ is a sink state. The second component is used to store the current beief of player 1. Target states of $G_{\mathcal{L}}$ are those whose first component is a target state of the initial game G. Actions and signals of player 2 in $G_{\mathcal{L}}$ are the same as in G plus one new signal $\perp$. The actions of player 1 in $G_{\mathcal{L}}$ are all the non-empty set of his actions in G i.e. $\mathcal{P}(I)\backslash\emptyset$ and her set of signals in $G_{\mathcal{L}}$ is $C \times \mathcal{P}(I) \cup \{\perp\}$. Thus player 1 can choose a lottery over various supports of uniform distributions over I, then one of the uniform distribution is randomly selected according to the lottery and player 1 receives a signal with two components, the first behind inherited from the signalling structure of G and the second is the support of the uniform distribution that was selected. Formally the non-zero values of the transition function $p_{\mathcal{L}}$ of $G_{\mathcal{L}}$ are defined for every $I_0 \subseteq I$ and $k, k' \in K$ and $B \subseteq K$ by:

$$
\begin{aligned}
&p_{\mathcal{L}}(\perp, \perp, \perp \mid (k, B), I_0, j) = 1 \text{ if } I_0 \notin I(B),\\
&p_{\mathcal{L}}(\perp, \perp, \perp \mid \perp, I_0, j) = 1,\\
&p_{\mathcal{L}}((k', \mathcal{B}_1(B, I_0, c)), (c, I_0), d \mid (k, B), I_0, j) = \frac{1}{|I_0|} \sum_{i_0 \in I_0} p(k', c, d \mid k, i_0),
\end{aligned}
$$

and all transition probabilities in $G_{\mathcal{L}}$ not specified by these three equations are equal to 0.

To get (i) and (iii) we apply Theorem 6.3 to the reachability game $G_{\mathcal{L}}$: every support is either positively winning for 1 or surely winning for 2 in $G_{\mathcal{L}}$.

In the first case, the support is also almost-surely winning in the $\mathcal{L}$-game because player 1 can transform his canonical winning strategy $\sigma_{\mathcal{L}}$ in $G_{\mathcal{L}}$ with finite memory $\mathcal{P}(K)\times\mathcal{P}(\mathcal{P}(I))$ to an almost surely winning strategy σ in G with finite memory $\mathcal{P}(K)\times\mathcal{P}(\mathcal{P}(I))\times\mathcal{P}(I)$ When the memory state is $(B, \mathcal{I}_\prime, I_0)$ strategy σ plays the uniform lottery over I_0, receives signal $c \in C$ and the new memory state is $(\mathcal{B}_1(B, I_0, c), \mathcal{I}_0', I_0')$ where the two first components are updated according to $\sigma_{\mathcal{L}}$ with signal (c, I_0) and I_0' is choosen uniformly at random in $\mathcal{I}_0'$.

In the second case, we notice that the very same strategy τ which is surely winning in $G_{\mathcal{L}}$ is surely winning in the $\mathcal{L}$-game as well, because in particular it is surely winning against the totally randomized strategy σ_{rand} in $G_{\mathcal{L}}$ thus by construction of $G_{\mathcal{L}}$ the same holds in the $\mathcal{L}$-game as well. According to Theorem 6.3, for τ we can choose for τ a finite memory strategy whose memory are the beliefs of player 2 in $G_{\mathcal{L}}$. Since the state space of $G_{\mathcal{L}}$ is $K \times (\mathcal{P}(K)\backslash\mathcal{L}) \cup \{\perp\}$ and player 2 can forget about state $\perp$ because it is a sink state, we get (iii).

The computability of $\mathcal{L}''$ and σ and τ stated in (iv) is a consequence of Theorem 6.3 applied to $G_{\mathcal{L}}$.

Now we suppose $\mathcal{L}''$ is empty and prove (ii). According to Theorem 6.3, any support not in $\mathcal{L}$ is positively winning for player 1 in $G_{\mathcal{L}}$ and moreover the strategy σ_{rand} which consists in playing randomly any action is positively winning for player 1. In the game $G_{\mathcal{L}}$, when the belief of player 1 is L, playing a lottery whose support is not in $I(L)$ leads immediatly to a non-accepting sink state, hence strategy $\sigma_{\mathcal{L}}$ which consists in playing randomly any action in $I(L)$ is positively winning as well, from any initial distribution whose support is not in $\mathcal{L}$.

To prove (ii) it is enough to show that for every initial distribution δ whose support is not in $\mathcal{L}$,

$$\sigma_{\mathcal{L}} \text{ is almost-surely winning for player } 1 \text{ from } \delta \ . \tag{26}$$

According to (23) applied to the game $G_{\mathcal{L}}$, there exists $N \in \mathbb{N}$ such that for every strategy τ and every distribution δ whose support is not in $\mathcal{L}$,

$$\mathbb{P}_{\delta}^{\sigma_{\mathcal{L}},\tau}\left(\exists n \leq N, K_n \in T\right) \geq \frac{1}{N} \ . \tag{27}$$

Moreover, since $\sigma_{\mathcal{L}}$ has finite memory and since $\sigma_{\mathcal{L}}$ guarantees the belief of player 1 we can use the shifting Lemma to obtain:

$$\forall m \in \mathbb{N}, \mathbb{P}_{\delta}^{\sigma_{\mathcal{L}},\tau}\left(\exists m \leq n \leq m+N, K_n \in T\right) \geq \frac{1}{N} \ , \tag{28}$$

hence using the Borel-Cantelli lemma:

$$\mathbb{P}_{\delta}^{\sigma_{\mathcal{L}},\tau}\left(\exists^{\infty} n, K_n \in T\right) = 1 \ .$$

This completes the proof that $\sigma_{\mathcal{L}}$ is almost-surely winning from any support $L \notin \mathcal{L}$ for the Büchi condition. It proves (26), and hence (ii). □

The following proposition provides a fix-point characterization of almost-surely winning supports for player 1.

PROPOSITION 6.8 (FIX-POINT CHARACTERIZATION OF ALMOST-SURELY WINNING SUPPORTS). *Let $\mathcal{L} \subseteq \mathcal{P}(K)$ be a set of supports. Suppose player 1 can enforce her beliefs to stay outside $\mathcal{L}$. Then,*

(i) either every support $L \notin \mathcal{L}$ is almost-surely winning for player 1 and her Büchi objective,

(ii) or there exists a set of supports $\mathcal{L}' \subseteq \mathcal{P}(K)$ and a strategy τ^* for player 2 such that:
  - *(a) $\mathcal{L}'$ is not empty and does not intersect $\mathcal{L}$,*
  - *(b) player 1 can enforce her beliefs to stay outside $\mathcal{L} \cup \mathcal{L}'$,*
  - *(c) for every strategy σ and initial distribution δ with support in $\mathcal{L}'$,*

$$\mathbb{P}_{\delta}^{\sigma,\tau^*}\left(\forall n \geq 2^K, K_n \notin T \mid \forall n, \mathcal{B}_1^n \notin \mathcal{L}\right) > 0 \ . \tag{29}$$

There exists an algorithm running in time doubly-exponential in the size of G for deciding which of cases (i) or (ii) holds. In case (i) holds, the strategy $\sigma_{\mathcal{L}}$ for player 1 which consists in playing randomly any action in $I(L)$ when her belief is L is almost-surely winning for the Büchi objective. In case (ii) holds, the algorithm computes at the same time $\mathcal{L}'$ and a finite memory strategy τ^* with memory $\mathcal{P}(\mathcal{L}' \times K)\backslash\{\emptyset\}$ such that (29) *holds.*

PROOF. Let $\mathcal{L}''$ be the set of supports surely winning for player 2 in the $\mathcal{L}$-game. Let τ_{rand} be the memoryless strategy for player 2 playing randomly any action. Let $\mathcal{L}'$ be the set of supports L such that $L \notin \mathcal{L}$ and,

$$\forall \sigma, \mathbb{P}_{\delta_L}^{\sigma,\tau_{\text{rand}}}\left(\exists n \leq 2^K, \mathcal{B}_1^n \in \mathcal{L}'' \cup \mathcal{L}\right) > 0 \ , \tag{30}$$

where δ_L is the uniform distribution on L.

We start with proving that if $\mathcal{L}''$ is empty then case (i) of Proposition 6.8 holds. Since player 1 can enforce her beliefs to stay outside $\mathcal{L}$, then $\mathcal{L}'$ is empty as well. Moreover, according to (ii) of Proposition 6.7, every support not in $\mathcal{L}$ is almost-surely winning for player 1 for the Büchi condition, hence we are in case (i) of Proposition 6.8.

Suppose now that $\mathcal{L}''$ is *not* empty, Then we prove (ii)(a), (ii)(b) and (ii)(c) of Proposition 6.8.

First (ii)(a) is obvious because since $\mathcal{L}'' \subseteq \mathcal{L}'$, then $\mathcal{L}'$ is not empty either

Now we prove property (ii)(b) holds: player 1 can enforce his beliefs to stay outside $\mathcal{L} \cup \mathcal{L}'$. There exists σ such that (30) does not hold, i.e. σ enforce the 2^K first belief of player 1 to stay outside $\mathcal{L}'' \cup \mathcal{L}$. We can modify σ such that this holds for all steps of the game. For that, player 1 can use strategy σ' which plays like σ, and as soon as player 1 has twice the same belief L, she forgets every signal she received between the two occurences of L and keeps playing with σ. Since therer are less than 2^K different possible beliefs for player 1, the strategy σ' guarantees the belief of player 1 to stay outside $\mathcal{L}'' \cup \mathcal{L}$ forever.

Description of the positively winning strategy τ^* for player 2. It remains to prove (ii)(c). According to (iii) of Proposition 6.7, there exists a strategy τ' for player 2 which is surely winning in the $\mathcal{L}$-game from any support in $\mathcal{L}''$.

We define a strategy τ^* for player 2 such that (29) holds. At each step, player 2 throws a coin. As long as the result is "tail", then player 2 plays randomly any action: he keeps playing with τ_{rand}. If the result is "head" then player 2 picks randomly a support $L \in \mathcal{L}''$ (actually he guesses the belief of player 1), forgets all his signals up to now and switches definitively to strategy τ' with initial support L.

Intuitively, what matters with strategy τ^* is that the opponent player 1 does not know whether he faces strategy τ' or strategy τ_{rand}, because everything is possible with strategy τ_{rand}.

Let us prove that τ^* guarantees property (29) to hold.

We start with proving for every strategy σ of player 1 and δ an initial distribution whose support is in $L \in \mathcal{L}'$, there exists a support $L'' \in \mathcal{L}''$, $N \leq 2^K$ and signals $c_1^0, \ldots, c_N^0$ and actions sets $I_1^0, \ldots, I_N^0$ such that:

$$\forall l \in L'', \delta''(l) = \mathbb{P}_\delta^{\sigma,\tau^*}(K_n = L \wedge A) > 0 \ , \tag{31}$$

where A denotes the event $\{I_1 = I_1^0, C_1 = c_1^0, \ldots, C_n = c_N^0\}$.

By definition of $\mathcal{L}'$ and τ_{rand}, there exists signals $c_1^0, \ldots, c_N^0$, actions sets $I_1^0, \ldots, I_N^0$ and a support $L'' \in \mathcal{L}''$ such that $L'' = \mathcal{B}_1(L, I_1^0, c_1^0, \ldots, I_N^0, c_N^0)$, $N \leq 2^K$ and $\mathbb{P}_\delta^{\sigma,\tau_{\mathsf{rand}}}\left(\mathcal{B}_1^N = L'' \wedge A\right) > 0$. Then, according to (5) of Lemma 3.4 $\forall l \in L'', \mathbb{P}_\delta^{\sigma,\tau_{\mathsf{rand}}}(K_n = l \wedge A) > 0$. Since by definition of τ^*, there is positive probability that τ plays like τ_{rand} up to stage N, then we get (31).

Now we can complete the proof of (29). Since τ' is surely winning in the $\mathcal{L}$-game from $L'' \in \mathcal{L}''$, it guarantees that:

$$\forall \sigma, \mathbb{P}_{\delta''}^{\sigma,\tau'}(\forall n \in \mathbb{N}, K_n \not\subseteq T \mid \forall n \in \mathbb{N}, I_n \in I(\mathcal{B}_1^n)) = 1 \ .$$

There is positive probability that at stage n, τ^* switches to strategy τ' in initial state L''. By definition of beliefs,

$$\mathcal{B}_1(L'', I_1, C_1, \ldots, I_n, C_n) = \mathcal{B}_1(L, I_1^0, c_1^0, \ldots, c_N^0, I_1, C_1, \ldots, I_n, C_n)$$

hence according to (31) and the shifting lemma,

$$\forall \sigma, \mathbb{P}_\delta^{\sigma,\tau^*}(\forall n \geq N, K_n \not\subseteq T \wedge A \mid \forall n \geq N, I_n \in I(\mathcal{B}_1^n)) > 0 \ . \tag{32}$$

According to the definition of $I(L)$, for every σ and $n \in \mathbb{N}$,

$$\mathbb{P}_{\delta}^{\sigma,\tau_{\mathsf{rand}}}\left(\mathcal{B}_1^n \in \mathcal{L} \mid I_n \notin I(\mathcal{B}_1^n)\right) > 0, \quad .$$

and since there is positive probability that τ plays like τ_{rand} up to stage n, the same holds for τ, hence:

$$\mathbb{P}_{\delta}^{\sigma,\tau^*}\left(\forall n \in \mathbb{N}, I_n \in I(\mathcal{B}_1^n) \mid \forall n \in \mathbb{N}, \mathcal{B}_1^n \notin \mathcal{L}\right) > 0 \ .$$

This last equation together with (32) proves (29), which completes the proof of (ii)(c) of Proposition 6.8.

Description of the algorithm. To terminate the proof of Proposition 6.8, we have to describe the doubly-exponential algorithm. This algorithm is a fix-point algorithm, actually there are two embedded fix-points, since this algorithm uses twice as sub-procedures the algorithm provided by Theorem 6.3 on game $G_{\mathcal{L}}$ defined in the proof of Proposition 6.7.

The algorithm of Proposition 6.7, property (iv) is used for computing $\mathcal{L}''$, and σ or τ'.

In case $\mathcal{L}''$ is empty, the algorithm simply outputs strategy $\sigma_{\mathcal{L}}$ described in (ii) of Proposition 6.8. In case $\mathcal{L}''$ is not empty, the algorithm computes the set of supports $\mathcal{L}'$ defined by (30), from which player 2 can force the belief of player 1 to be in $\mathcal{L}'' \cup \mathcal{L}$ someday with positive probability. For computing $\mathcal{L}'$, we have to fix a strategy τ_{rand} in the game $G_{\mathcal{L}}$ and check whether player 1 has a strategy for avoiding surely his beliefs to be in $\mathcal{L}' \cup \mathcal{L}$, which can be done running the algorithm of Proposition 6.8 to the game $G_{\mathcal{L}}$.

Once $\mathcal{L}'$ has been computed, the algorithm outputs strategy τ^* described above. □

The proof of Theorem 6.4 describes the composition of the various finite memory strategies of Proposition 6.8 in order to obtain a strategy for player 2 which is positively winning and has finite memory $\mathcal{P}(\mathcal{P}(K) \times K)$.

PROOF OF THEOREM 6.4. According to Proposition 6.8, starting with $\mathcal{L}_0 = \emptyset$, there exists a sequence $\mathcal{L}'_0, \mathcal{L}'_1, \ldots, \mathcal{L}'_n$ of disjoint non-empty sets of supports such that for every $m \leq n$,

— if $0 \leq m < M$ then $\mathcal{L}_m = \mathcal{L}'_0 \cup \cdots \cup \mathcal{L}'_{m-1}$, matches case (ii) of Proposition 6.8. We denote τ_m the corresponding finite memory strategy.
— $\mathcal{L}_M$ matches case (i) of Proposition 6.8.

Then according to Proposition 6.8, the set of supports positively winning for player 2 is exactly $\mathcal{L}_M$, and supports that are not in $\mathcal{L}_M$ are almost-surely winning for player 1. This proves qualitative determinacy.

The sequence $\mathcal{L}'_0, \mathcal{L}'_1, \ldots, \mathcal{L}'_n$ is computable in doubly-exponential time, because each application of Proposition 6.8 involves running the doubly exponential-time algorithm, and the length of the sequence is at most doubly-exponential in the size of the game.

The only thing that remains to prove is the existence and computability of a positively winning strategy τ^+ for player 2, with finite memory $\mathcal{P}(\mathcal{P}(K) \times K)$. Strategy τ consists in playing randomly any action as long as a coin gives result "head". When the coin gives result "tail", then strategy τ^+ chooses randomly an integer $0 \leq m < M$ and a support $L \in \mathcal{L}'_m$ and switches to strategy τ_m. Since each strategy τ_m has memory $\mathcal{P}(\mathcal{L}'_m \times K) \backslash \{\emptyset\}$ and the $\mathcal{L}'_m$ are distincts, strategy τ^+ has memory $\mathcal{P}(\mathcal{P}(K) \times K)$ with $\emptyset$ used as the initial memory state.

We prove that τ^+ is positively winning for player 2 from $\mathcal{L}_M$. Let σ be a strategy for player 1, $L \in \mathcal{L}_M$ and δ an initial distribution with support L. Let m_0 be the smallest index m such that

$$\mathbb{P}_{\delta}^{\sigma,\tau^+}\left(\exists n \in \mathbb{N}, \mathcal{B}_1^n \in \mathcal{L}'_m\right) > 0 \ .$$

Since $L \in \mathcal{L}_M$ and $\mathcal{L}_M = \bigcup_{m<M} \mathcal{L}'_m$, the set in the definition of m_0 is non-empty and m_0 is well defined. Let $n_0 \in \mathbb{N}$ and $c_1, c_2, \ldots, c_{n_0} \in C^{n_0}$ and action sets $I_1^0, I_2^0, \ldots, I_{n_0}^0$ such that $\mathcal{B}_1(L, I_1^0, c_1, \ldots, I_{n_0}^0, c_{n_0}) \in \mathcal{L}'_{m_0}$ and

$$\mathbb{P}_\delta^{\sigma,\tau^+}(E) > 0 \ ,$$

where $E = \{C_1 = c_1, \ldots, C_{n_0} = c_{n_0} \wedge I_1 = I_1^0, \ldots, I_n = I_n^0\}$. According to the definition of τ^+, there is positive probability that τ^+ plays randomly until step n_0 hence according to (5), for every state $l \in \mathcal{B}_1(L, c_1, \ldots, c_{n_0})$,

$$\mathbb{P}_\delta^{\sigma,\tau^+}(E \text{ and } K_n = l) > 0 \ . \tag{33}$$

According to the definition of τ^+ again, there is positive probability that τ^+ switches to strategy τ_{m_0} at instant n_0. Since $\mathcal{B}_1(L, I_1^0, c_1, \ldots, I_{n_0}^0, c_{n_0}) \in \mathcal{L}'_{n_0}$ then according to (33) and to (29) of Proposition 6.8,

$$\mathbb{P}_\delta^{\sigma,\tau^+}\left(\forall n \geq 2^K, K_n \notin T \mid \forall n \geq n_0, \mathcal{B}_1^n \notin \mathcal{L}_{m_0}\right) > 0 \ . \tag{34}$$

By definition of m_0 and since $\mathcal{L}_{m_0} = \mathcal{L}'_0 \cup \cdots \cup \mathcal{L}'_{m_0-1}$,

$$\mathbb{P}_\delta^{\sigma,\tau^+}(\forall n \in \mathbb{N}, \mathcal{B}_1^n \notin \mathcal{L}_{m_0}) = 1 \ ,$$

then together with (34),

$$\mathbb{P}_\delta^{\sigma,\tau^+}\left(\forall n \geq 2^K, K_n \notin T\right) > 0 \ ,$$

which proves that τ^+ is positively winning for the co-Büchi condition. □

7. LOWER BOUND ON MEMORY NEEDED BY STRATEGIES.

7.1. Overview of the proof

We show in this section that a doubly-exponential memory is necessary to win positively safety (and hence co-Büchi) games. We construct, for each integer n, a reachability game, whose number of state is polynomial in n and such that player 2 has a positively winning strategy for his safety objective. This game, called $\mathsf{guess\_my\_set}_n$, is described on Fig. 7. The objective of player 2 is to stay away from ☺, while player 1 tries to reach ☺.

We prove that every positively winning strategy of player 2 (for the safety objective) uses a memory at least doubly exponential in n. This property is stated precisely in Proposition 7.1. Prior to that, we describe the high level structure of the game $\mathsf{guess\_my\_set}_n$ for a fixed $n \in \mathbb{N}$.

Idea of the game. The game $\mathsf{guess\_my\_set}_n$ is divided into three rounds. In the first round, player 1 generates a set $X \subsetneq \{1, \ldots, n\}$ of size $|X| = n/2$. There are $\binom{n}{n/2}$ possibilities of such sets X. Player 2 is blind in this round and has no action to play.

In the second round, player 1 announces by her actions $\frac{1}{2}\binom{n}{n/2}$ (pairwise different) sets of size $n/2$ which are different from X. Player 2 has no action to play in that round, but he observes the actions of player 1 (and hence the sets announced by player 1).

In the third round, player 2 can announce by his action up to $\frac{1}{2}\binom{n}{n/2}$ sets of size $n/2$. Player 1 observes actions of player 2. If player 2 succeeds in finding the set X, the game restarts from scratch. Otherwise, the game goes to state ☺ and player 1 wins.

It is worth noticing that in order to implement the game $\mathsf{guess\_my\_set}_n$ in a compact way, we allow player 1 to cheat, and rely on probabilities to always have a chance to catch player 1 cheating, in which case the game is sent to the sink state s, and player 1 loses. That is, player 1 has to play following the rules without cheating else she cannot

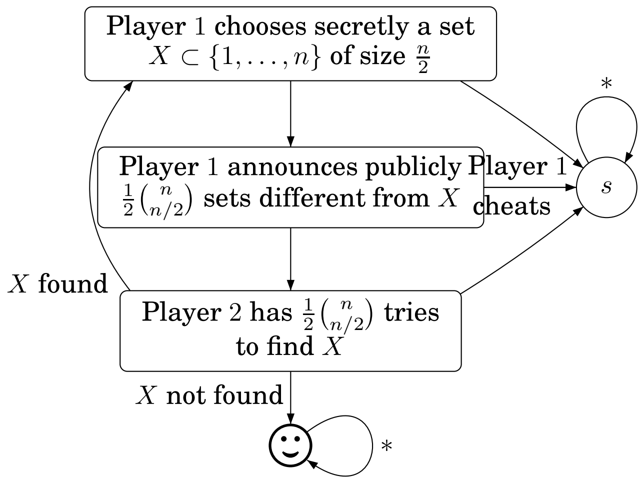

Fig. 7. A game where player 2 needs doubly-exponential memory to avoid the ☺-state.

win almost-surely her reachability objective. However we do not need to allow player 2 to cheat. Notice that player 1 is better informed than player 2 in this game.

Concise encoding. We now explain the important features of our encoding of the game $\mathsf{guess\_my\_set}_n$, to prove that it can be encoded with a number of states polynomial in n. The formal construction can be found in the next subsection. There are three problems to be solved, that we sketch here. First, remembering set X in the state of the game would ask for an exponential number of states. Instead, we use a fairly standard technique: recall at random a single element $x \in X$. In order to check that a set Y of size $n/2$ is different from the set X of size $n/2$, we challenge player 1 to point out some element $y \in Y \setminus X$. We ensure by construction that $y \in Y$, for instance by asking it when Y is given. This way, if player 1 cheats, then she will give $y \in X$, leaving a non zero probability that $y = x$. If $y = x$, player 1 is definitely cheating: the game goes to some sink state s where player 1 looses.

The second problem is to make sure that player 1 generates an exponential number of pairwise different sets $X_1, X_2, \ldots, X_{\frac{1}{2}\binom{n}{n/2}}$. Notice that the game cannot recall even one of these sets. Instead, player 1 generates the sets in some total order, denoted $<$, and thus it suffices to check only one inequality each time a set X_{i+1} is given, namely $X_i < X_{i+1}$. It is done in a similar but more involved way as before, by remembering randomly two elements of X_i instead of one.

The last problem is to count up to $\frac{1}{2} \cdot \binom{n}{n/2}$ with a logarithmic number of bits. Again, we ask player 1 to increment a counter, while remembering only one of the bits and punishing her if she increments the counter wrongly.

PROPOSITION 7.1. *Player 2 has a finite-memory strategy with* $3 \times 2^{\frac{1}{2} \cdot \binom{n}{n/2}}$ *different memory states to win positively* $\mathsf{guess\_my\_set}_n$.

No finite-memory strategy of player 2 with less than $2^{\frac{1}{2} \cdot \binom{n}{n/2}}$ *memory states wins positively* $\mathsf{guess\_my\_set}_n$.

PROOF. The first claim is quite straightforward. Player 2 remembers in which round he is (3 different possibilities). In round 2, player 2 remembers all the sets proposed by player 1 ($2^{\frac{1}{2} \cdot \binom{n}{n/2}}$ possibilities). Between round 2 and round 3, player 2 inverses his memory to remember the sets player 1 did not propose (still $2^{\frac{1}{2}\cdot\binom{n}{n/2}}$ possibilities). Then he proposes each of these sets, one by one, in round 3, deleting the set from his memory after he proposed it. Let us assume first that player 1 does not cheat and plays fair. Then all the sets of size $n/2$ are proposed (since there are $2 \cdot \frac{1}{2} \cdot \binom{n}{n/2}$ such sets), that is X has been found and the game starts another round without entering state ☺. Else, if player 1 cheats at some point, then the probability to reach the sink state s is non zero, and player 2 also wins *positively* his safety objective.

The second claim is not hard to show either. Assume by contradiction that there exists a finite memory strategy τ winning positively for player 2, and with stricly less than $2^{\frac{1}{2}\cdot\binom{n}{n/2}}$ states of memory. We construct a counter strategy for player 1. First, this strategy never cheats, else the game would enter the sink state with positive probability. It actually takes all its decision at random, following the rules: it chooses the secret set X at random in round 1, then in round 2, it proposes pairwise different sets $Y \neq X$ in a lexicographical way and uniformly at random. Now, consider for each round of the game the end of round 2, when player 1 has proposed a set A of $\frac{1}{2} \cdot \binom{n}{n/2}$ sets. The distribution over memory states of player 2 at that point only depends upon A and upon the distribution of his memory state at the begining of the round. Let us fix a round (thus the initial distribution over memory states is fixed) and denote by m_A the distribution of memory states of player 2 after A has been disclosed. As there are less than $2^{\frac{1}{2}\cdot\binom{n}{n/2}}$ memory states, there exists at least one memory state x and two sets $B \neq C$ such that $m_B(x) \neq 0$ and $m_C(x) \neq 0$. Let $\overline{D}$ be the complement of $D \in \{B, C\}$ among the sets of $n/2$ elements. Now, $\overline{B} \cup \overline{C}$ has strictly more than $\frac{1}{2} \cdot \binom{n}{n/2}$ sets of $n/2$ elements, as $B \neq C$. Hence, there exists at least one set $Y \in \overline{B} \cup \overline{C}$ which is proposed by player 2 with probability stricly less than 1 after memory state x. Without loss of generality, we can assume that $Y \notin B$ (the other case $Y \notin C$ is symmetrical). Now, there is a non zero probability that player 1 chooses secretly set Y in that round, and discloses B. Hence there is a non zero probability that player 2 does not propose set Y at step 3, in which case ☺ is reached in that round. It is easy to see that there is a uniform lower bound $p > 0$ on the probability to reach ☺ at round i. As it is true for every round, there is probability 1 to eventually reach ☺, and hence there is probability 0 to stay away from ☺, a contradiction with the fact that τ is positively winning. This achieves the proof that no finite-memory strategy of player 2 with less than $2^{\frac{1}{2}\cdot\binom{n}{n/2}}$ states of memory is positively winning. □

Before giving the formal construction of $\mathsf{guess\_my\_set}_n$, we explain how to concisely encode counting up to an exponential number with a polynomial number of states.

7.2. Exponential number of steps

Let $y_1 \cdots y_n$ be the binary encoding of a number y exponential in n (y_n being the parity of x). Here is a reachability game that the player needs to play for ny steps to surely win. Intuitively, the player needs to enumerate one by one the successors of 0 until reaching $y_1 \cdots y_n$ in order to win. Let say $x'_1 \cdots x'_n$ is the binary encoding of x', the successor of counter x. In order to check that the player does not cheat, the bit x'_i for a random i is secretly remembered. The value of x'_i can be easily computed on the fly while reading $x_i \ldots x_n$. Indeed, $x'_i = x_i$ iff there exists some $k > i$ with $x_k = 0$.

The set of signals is the same as the set of actions. This set is $\{0, 1, 2\}$. An action $a \in \{0, 1\}$ stands for the value of bits, while $a = 2$ stands for the fact that the player is claiming to have reached x.

The state space is basically the following: $(i, b, j, b', j', c)_{i,j,j' \leq n, x, x' \in \{0,1\}}$. The signification of such a state is that the player will give bit x_i, b, j are the check to make to the current number (checking that $x_j = b$), b', j' are the check to make to the successor of x ($x'_{j'} = b'$) , and c indicates whether there is a carry (correcting b' in case $c = 1$ at the end of the current number ($i = n$)). The initial distribution is the uniform distribution on $(0, 0, k, 0, 1)$ (checking that the initial number generated is indeed 0). If the player plays 2, then if $y_j = b$ the game goes to the goal state, else it goes to a sink state s.

We have the stochastic transition $p((i, b, j, b', j', c), a, \mathrm{s}) = 1$ if $i = j$ and $a \neq b$. Else, if $i \neq n$, the stochastic transition is $p((i, b, j, b', j', c), a, (i + 1, b, j, b', j', c \wedge a)) = \frac{1}{2}$ (the current bit will not be checked, and the carry is 1 if both c and a are 1), and the stochastic transition is $p((i, b, j, b', j', c), a, (i + 1, b, j, a, i, 1)) = 1/2$. At last, for $i = n$, we have the stochastic transition $p((i, b, j, b', j', c), a, (1, b' \wedge c, j', a, 1, 1)) = 1$ (the bit of the next number becomes the bit for the current configuration, taking care of the carry c). Clearly, if the player does not play y_n steps of the game, then it means she did not compute accurately the successor at one step, hence she has a chance to get caught and lose. That is, the probability to reach the goal state is not 1.

7.3. Implementing $\mathsf{guess\_my\_set}_n$ with a polynomial size game.

We now turn to the formal definition of $\mathsf{guess\_my\_set}_n$, with a number of states polynomial in n. From each state (but from the sink state s), player 1 can restart the game from the begining. In this case, we will say that it performs another round of the game.

The first round of the game is fairly standard, it consists in asking player 1 (who wants to reach the sink state ☺) for a set X of $n/2$ pairwise different numbers below n. The states of that round of the game is encoded in the form (x, i), where x is a number in set X that the system remembers (hidden for both players), and $i \leq n - 2$ is the size of X so far. Player 1 actions in that round are $\{0, \ldots, n\}$. Signals of player 1 are the same as her actions. There is no action nor signal for player 2.

If player 1 proposes again number x to be in her secret set, with x the number that the system remembered as in being in the secret set already, we have the stochastic transition $p((x, i), x, s) = 1$ (she is caught cheating and sent to her loosing state). Whenever $y \neq x$ is proposed by player 1, we have the stochastic transition $p((x, i), y, (x, i + 1)) = 1/2$ (the number y is accepted as the i-th number in the set, the number of numbers in the set is increased by 1, and the memory x is not updated), or the transition $p((x, i), y, (y, i + 1) = 1/2$ (the number y is accepted as the i-th number in the set, the number of numbers in the set is increased by 1, and the memory x is updated to y). If player 1 plays 0, it means that she has given $n/2$ number, the system checks that the current state is indeed of the form $(x, n/2)$, and the next round of the game begins. If the current state is not $(x, n/2)$, then the state changes to s (the stochastic transition is $p((x, i), 0, s) = 1$)and player 1 looses.

The number x in the memory of the system at the end of round 1 will be used and remembered all along this round of the game in the other rounds. We turn now to the second round, where player 1 gives $\frac{1}{2} \cdot \binom{n}{n/2}$ sets Y different to X. First, in order to be sure that every set Y she proposes is never X, player 1 is asked to give one number in $Y \setminus X$ (this number is not observed by player 2). If player 1 gives x, then the system knows that she is cheating, and the game goes to the sink state s from which player 1 loses. Since player 1 does not know what x is, playing any number in X is dangerous and ensures that the probability of the play reaching the sink state s is stricly positive. Hence if it cheats by proposing X again, it cannot reach its goal almost surely. The way the sets are announced by player 1 is the following. First, player 1 is asked whether number 1 belongs to the set it is annoucing (she plays a if yes, a' if not, and a'' if it is and furthermore it is the biggest number which will change compared to the following set). Player 2 has no choice of action to play. The observation of player 1 and 2 is the same as the action of player 1, that is player 2 is informed of the sets announced by player 1.

Second, the game needs to ensure that each set is different. For that, it asks player 1 to generate the sets in lexicographic order (if Y is given before Y', then there exists $(i, j) \in Y \times Y'$ such that $i < j$ and for all $k \in X$ with $k > i$, $k \in X'$ and $k \neq j$). Also, player 1 should announce with her action what is the biggest number i of current set Y which will not appear in the next set. The game remembers i, plus one number $j \in Y$ with $j > i$ (if any) (it can be done with polynomially many states). The game checks whether the next set Y' contains j, plus a number $i' \in Y'$ with $i < i'$ and $i' \neq j$. This number i' exists as the sets are given in lexicographic order and i is not in the next set. Again, since player 1 does not know the number j chosen, if she cheats and changes a number $k > i$ of Y, then there is always a chance that the game remembers that number and catches her cheating, in which case the game goes to the sink state s. To be sure that player 1 gives $\frac{1}{2} \cdot \binom{n}{n/2}$ sets, she plays the game of section 7.2 step by step, advancing to the successor of the current counter only when a set Y is proposed. Furthermore, when she has finished giving $\frac{1}{2} \cdot \binom{n}{n/2}$, she goes to the third round.

The third round ressembles the second round: player 2 proposes $\frac{1}{2} \cdot \binom{n}{n/2}$ sets instead of player 1, and player 1 observes these sets. For each set Y proposed by player 2, player 1 has to give an event in $X \setminus Y$ (this is not observed by player 2). This is ensured in the same way as in round 2. Recall that player 1 has always a reset action to restart the game from step 1, but in the sink state s. That is, if $Y = X$, player 1 can ends the round, and restart the game with a new set X in the following round.

After each set proposed by player 2, the game of section 7.2 advances to its next step. Once there has been $\frac{1}{2} \cdot \binom{n}{n/2}$ sets Y proposed with the proof by player 1 that $X \neq Y$, then player 1 goes to the goal state ☺ and wins.

8. COMPLEXITY LOWER-BOUND AND SPECIAL CASES.

In this section we show that our algorithms are optimal regarding complexity. Furthermore, we show that these algorithms enjoy better complexity in restricted cases, generalizing some known algorithms [Reif 1979; Chatterjee et al. 2007] to more general subcases, while keeping the same complexity as these restricted cases.

The special cases that we consider regard inclusion between knowledges of players. Recall that if at each moment of the game the belief of a player x is included in the one of his or her opponent $3 - x$, then player x is said to have more information (or to be better informed) than its opponent. It is in particular the case when for every transition, the signal of player 1 contains the signal of player 2.

8.1. Complexity lower bound for reachability and Büchi games.

We prove here that the problem of knowing whether the initial support of a reachability game or a Büchi game is almost-surely winning for player 1 is 2EXPTIME-complete. The lower bound even holds when player 1 is more informed than player 2.

THEOREM 8.1. *In a reachability or Büchi game, deciding whether player 1 has an almost-surely winning strategy is 2EXPTIME-hard, even if player 1 is more informed than player 2.*

We provide a proof for reachability games. The lower-bound of course extends to Büchi games since any reachability game can be turned into an equivalent Büchi one by making target states absorbing.

PROOF. We do a reduction from the membership problem for EXPSPACE alternating Turing machines. Let $\mathcal{M}$ be an EXPSPACE alternating Turing machine, and w be an input word of length n. From $\mathcal{M}$ and w we build a stochastic game with partial observation such that player 1 can achieve almost-surely a reachability objective if and only if w is accepted by $\mathcal{M}$. The idea of the game is that player 2 describes an execution of $\mathcal{M}$ on w, that is, he enumerates the tape contents of successive configurations. Moreover he chooses the rule to apply when the state of $\mathcal{M}$ is universal, whereas player 1 is responsible for choosing the rule in existential states. When the Turing machine reaches its final state, the play is won by player 1. In this game, if player 2 really implements some execution of $\mathcal{M}$ on w, player 1 has a surely winning strategy if and only if w is accepted by $\mathcal{M}$. Indeed, if all executions on w reach the final state of $\mathcal{M}$, then whatever the choices player 2 makes in universal states, player 1 can properly choose rules to apply in existential states in order to reach a final configuration of the Turing machine. On the other hand, if some execution on w does not lead to the final state of $\mathcal{M}$, player 1 is not sure to reach a final configuration and win the game.

This reasoning holds under the assumption that player 2 effectively describes the execution of $\mathcal{M}$ on w consistent with the rules chosen by both players. However, player 2 could cheat when enumerating successive configurations of the execution. He would for instance do so, if w is indeed accepted by $\mathcal{M}$, in order to have a chance not to lose the game. To prevent player 2 from cheating (or at least to prevent him from cheating too often), it would be convenient for the game to remember the tape contents, and check that in the next configuration, player 2 indeed applied the chosen rule. However, the game can remember only a logarithmic number of bits, while the configurations have a number of bits exponential in n. Instead, we ask player 1 to pick any position k of the tape, and to announce it to the game (player 2 does not know k), which is described by a linear number of bits. The game keeps the the letter at this position together with the previous and next letter on the tape. This allows the game to compute the letter a at position k of the *next* configuration. As player 2 describes the next configuration, player 1 will annouce to the game that position k has been reached again. The game will thus check that the letter player 2 gives is indeed a. This way, the game has a positive probability to detect that player 2 is cheating. If so, the game goes to a sink state which is winning for player 1. To increase the probability for player 1 of observing player 2 cheating, player 1 has the possibility to restart the whole execution from the beginning whenever she wants. In particular, she will do so when an execution lasts longer than 2^{2^n} steps. This way, if player 2 cheats infinitely often, player 1 will detect it with probability one, and will win the game almost-surely.

So far, we described a deterministic game satisfying that if w is accepted by $\mathcal{M}$, player 1 has a mixed strategy to reach her winning state almost surely, and without cheating (that is, denonciating player 2 only if he was cheating).

We now have to take into account that player 1 could cheat: she could point a certain position of the tape contents at a given step, and point somewhere else in the next step. To avoid this kind of behaviour, or at least refrain it, a piece of information about the position pointed by player 1 is kept secret (to both players) in the state of the game. More precisely, a bit of the binary encoding of the letter position on the tape, and the position of this bit itself is randomly chosen among the at most n possible positions. If player 1 is caught cheating (that is, if the bits at the position remembered differ between both step), the game goes to a sink state losing for player 1. This way, when player 1 decides to cheat, there is a positive probability that she loses the game. At this stage, the game is stochastic (a bit and a position are remembered randomly in states of the game), player 1 does not have full information (she does not know which bit is remembered in the state), but she has more information than player 2 (the latter does not know what letter player 1 decided to memorize). This game satisfies the following: w is accepted by $\mathcal{M}$ if and only if player 1 has mixed winning strategy which ensures reaching a goal state almost surely. Note that in the game described above player 1 does not have full information but has more information than player 2. □

8.2. Special cases.

A first straightforward result is that in a safety game where player 1 has full information, deciding whether she has an almost-surely winning strategy is in PTIME.

Now, consider a Büchi game. In general, as shown in the previous section, deciding whether the game is almost-surely winning for player 1 is 2EXPTIME-complete. However, it is already known that when player 2 has a full observation of the game the problem is EXPTIME-complete only [Chatterjee et al. 2007]. We show that our algorithm keeps the same EXPTIME upper-bound even in the more general case where player 2 is more informed than player 1, as well as in the case where player 1 fully observes the state of the game.

PROPOSITION 8.2. *In a Büchi game where either player 2 has more information than player 1 or player 1 has full observation, deciding whether player 1 has an almost-surely winning strategy or not (in which case player 2 has a positively winning strategy) can be done in exponential time.*

PROOF.
The reason for the single EXPTIME complexity in these special cases is that in both cases, there are at most an exponential number of 2-beliefs for player 2. If player 1 has complete information, then player 1 beliefs are singleton state. Thus there are at most an exponential number 2 belief of player 2 (the sets of beliefs of player 1).

If player 2 has more information than player 1, then player 2 can simulate player 1 and know exactly player 1 belief. That is, the 2-belief of player 2 is exactly the singleton made of the belief of player 1. As there are exponentially many beliefs for player 1, there are exponentially many 2 beliefs for player 2. □

Note that the latter proposition does not hold when player 1 has more information than player 2. Indeed in the game presented for the lower-bound, in the proof of Theorem 8.1, player 1 does have more information than player 2; yet she does not have full information.

9. CONCLUSION.

We considered stochastic games with signals and established two determinacy results. First, a reachability game is either almost-surely winning for player 1, surely winning for player 2 or positively winning for both players. Second, a Büchi game is either almost-surely winning for player 1 or positively winning for player 2. We gave algorithms for deciding in doubly-exponential time which case holds and for computing winning strategies with finite memory. Further, we showed that both the memory and the algorithmic complexities are tight.

Changing the notion of reaching a Büchi objective with positive probability for notion where the frequency at which a target state is visited does not converge towards 0 [Tracol 2011] leads to decidability in the subclass of Probabilistic Finite Automaton [Rabin 1963]. It would be interesting to extend this result to two players with signals.

REFERENCES

AUMANN, R. J. 1995. *Repeated Games with Incomplete Information*. MIT Press.

BAIER, C., BERTRAND, N., AND GRÖSSER, M. 2008. On decision problems for probabilistic Büchi automata. In *Proc. of FOSSACS'08*. LNCS, vol. 4972. Springer, 287–301.

BERTRAND, N., GENEST, B., AND GIMBERT, H. 2009. Qualitative determinacy and decidability of stochastic games with signals. In *LICS*. IEEE Computer Society, 319–328.

BERWANGER, D., CHATTERJEE, K., DOYEN, L., HENZINGER, T. A., AND RAJE, S. 2008. Strategy construction for parity games with imperfect information. In *Proc. of CONCUR'08*. LNCS, vol. 5201. Springer, 325–339.

CHATTERJEE, K., DE ALFARO, L., AND HENZINGER, T. A. 2005. The complexity of stochastic Rabin and Streett games. In *Proc. of ICALP'05*. LNCS, vol. 3580. Springer, 878–890.

CHATTERJEE, K. AND DOYEN, L. 2011. Partial-observation stochastic games: How to win when belief fails. *CoRR abs/1107.2141*.

CHATTERJEE, K., DOYEN, L., GIMBERT, H., AND HENZINGER, T. A. 2010. Randomness for free. In *Proc. of MFCS'10*. LNCS, vol. 6281. Springer, 246–257.

CHATTERJEE, K., DOYEN, L., AND HENZINGER, T. A. 2013. A survey of partial-observation stochastic parity games. *Formal Methods in System Design 43,* 2, 268–284.

CHATTERJEE, K., DOYEN, L., HENZINGER, T. A., AND RASKIN, J.-F. 2007. Algorithms for omega-regular games of incomplete information. *Logical Methods in Computer Science 3,* 3.

CHATTERJEE, K., JURDZINSKI, M., AND HENZINGER, T. A. 2004. Quantitative stochastic parity games. In *Proc. of SODA'04*. SIAM, 121–130.

CONDON, A. 1992. The complexity of stochastic games. *Information and Computation 96*, 203–224.

DE ALFARO, L. AND HENZINGER, T. A. 2000. Concurrent omega-regular games. In *Proc. of LICS'00*. IEEE, 141–154.

DE ALFARO, L., HENZINGER, T. A., AND KUPFERMAN, O. 2007. Concurrent reachability games. *Theoretical Computer Science 386,* 3, 188–217.

DE ALFARO, L. AND MAJUMDAR, R. 2001. Quantitative solution of omega-regular games. In *Proc. of STOC'01*. ACM, 675–683.

GIMBERT, H. AND HORN, F. 2008. Simple stochastic games with few random vertices are easy to solve. In *Proc. of FOSSACS'08*. LNCS, vol. 4972. Springer, 5–19.

GIMBERT, H. AND OUALHADJ, Y. 2010. Probabilistic automata on finite words: Decidable and undecidable problems. In *Proc. of the 37th International Colloquium on Automata, Languages and Programming (ICALP '10)*. Lecture Notes in Computer Science, vol. 6199. Springer, 527–538.

GRÄDEL, E., THOMAS, W., AND WILKE, T. 2002. *Automata, Logics and Infinite Games*. LNCS, vol. 2500. Springer.

GRIPON, V. AND SERRE, O. 2009. Qualitative concurrent stochastic games with imperfect information. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming (ICALP'09)*. Lecture Notes in Computer Science, vol. 5556. Springer, 200–211.

GRIPON, V. AND SERRE, O. 2011. Qualitative concurrent stochastic games with imperfect information. *CoRR abs/0902.2108*.

HORN, F. 2008. Random games. Ph.D. thesis, Université Denis-Diderot.

MERTENS, J.-F. AND NEYMAN, A. 1982. Stochastic games have a value. In *Proc. of the National Academy of Sciences USA*. Vol. 79. 2145–2146.

PAZ, A. 1971. *Introduction to probabilistic automata*. Academic Press.

RABIN, M. O. 1963. Probabilistic automata. *Information and Control 6,* 3, 230–245.

REIF, J. H. 1979. Universal games of incomplete information. In *Proc. of STOC'79*. ACM, 288–308.

RENAULT, J. 2007. The value of repeated games with an informed controller. Tech. rep., CEREMADE, Paris. Jan.

RENAULT, J. AND SORIN, S. 2008. Personal Communications.

ROSENBERG, D., SOLAN, E., AND VIEILLE, N. 2003. Stochastic games with imperfect monitoring. Tech. Rep. 1376, Northwestern University. July.

SHAPLEY, L. S. 1953. Stochastic games. In *Proc. of the National Academy of Sciences USA*. Vol. 39. 1095–1100.

SORIN, S. 2002. *A first course on zero-sum repeated games*. Springer.

TRACOL, M. 2011. Recurrence and transience for finite probabilistic tables. *Theor. Comput. Sci 412,* 12-14, 1154–1168.

VON NEUMANN, J. AND MORGENSTERN, O. 1944. *Theory of games and economic behavior*. Princeton University Press.

Received ; revised ; accepted