On Strong Determinacy of Countable Stochastic Games

Citation for published version:
Mayr, R, Kiefer, S, Shirmohammadi, M & Wojtczak, D 2017, On Strong Determinacy of Countable Stochastic Games. in *2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS).* Institute of Electrical and Electronics Engineers, 2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science, Reykjavik, Iceland, 20/06/17. https://doi.org/10.1109/LICS.2017.8005134

Digital Object Identifier (DOI):
10.1109/LICS.2017.8005134

Published In:
2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)

On Strong Determinacy of Countable Stochastic Games

Stefan Kiefer*, Richard Mayr†, Mahsa Shirmohammadi*, Dominik Wojtczak‡
*University of Oxford, UK
†University of Edinburgh, UK
‡University of Liverpool, UK

***Abstract*—We study 2-player turn-based perfect-information stochastic games with countably infinite state space. The players aim at maximizing/minimizing the probability of a given event (i.e., measurable set of infinite plays), such as reachability, Büchi, ω-regular or more general objectives.**

These games are known to be weakly determined, i.e., they have value. However, strong determinacy of threshold objectives (given by an event $\mathcal{E}$ and a threshold $c \in [0, 1]$) was open in many cases: is it always the case that the maximizer or the minimizer has a winning strategy, i.e., one that enforces, against all strategies of the other player, that $\mathcal{E}$ is satisfied with probability $\geq c$ (resp. $< c$)?

We show that almost-sure objectives (where $c = 1$) are strongly determined. This vastly generalizes a previous result on finite games with almost-sure tail objectives. On the other hand we show that $\geq 1/2$ (co-)Büchi objectives are not strongly determined, not even if the game is finitely branching.

Moreover, for almost-sure reachability and almost-sure Büchi objectives in finitely branching games, we strengthen strong determinacy by showing that one of the players must have a memoryless deterministic (MD) winning strategy.

***Index Terms*—stochastic games, strong determinacy, infinite state space**

I. Introduction

Stochastic games. Two-player stochastic games [16] are adversarial games between two players (the maximizer $\Box$ and the minimizer $\Diamond$) where some decisions are determined randomly according to a pre-defined distribution. Stochastic games are also called $2\frac{1}{2}$-player games in the terminology of [8], [7]. Player $\Box$ tries to maximize the expected value of some payoff function defined on the set of plays, while player $\Diamond$ tries to minimize it. In concurrent stochastic games, in every round both players each choose an action (out of given action sets) and for each combination of actions the result is given by a pre-defined distribution. In the subclass of turn-based stochastic games (also called simple stochastic games) only one player gets to choose an action in every round, depending on which player owns the current state.

We study 2-player turn-based perfect-information stochastic games with *countably infinite* state spaces. We consider objectives defined via predicates on plays, not general payoff functions. Thus the expected payoff value corresponds to the probability that a play satisfies the predicate.

Standard questions are whether a game is determined, and whether the strategies of the players can without restriction be chosen to be of a particular type, e.g., MD (memoryless deterministic) or FR (finite-memory randomized).

Finite-state games vs. Infinite-state games. Stochastic games with finite state spaces have been extensively studied [23], [9], [11], [17], [8], both w.r.t. their determinacy and the strategy complexity (memory requirements and randomization). E.g., strategies in *finite* stochastic parity games can be chosen memoryless deterministic (MD) [10], [7], [6]. These results have a strong influence on algorithms for deciding the winner of stochastic games, because such algorithms often use a structural property that the strategies can be chosen of a particular type (e.g., MD or finite-memory).

More recently, several classes of finitely presented infinite-state games have been considered as well. These are often induced by various types of automata that use infinite memory (e.g., unbounded pushdown stacks, unbounded counters, or unbounded fifo-queues). Most of these classes are still finitely branching. Stochastic games on infinite-state probabilistic recursive systems (i.e., probabilistic pushdown automata with unbounded stacks) were studied in [13], [14], [12], and stochastic games on systems with unbounded fifo-queues were studied in [1]. However, most these works used techniques that are specially adapted to the underlying automata model, not a general analysis of infinite-state games. Some results on general stochastic games with countably infinite state spaces were presented in [19], [4], [18], [5] though many questions remained open (see our contributions further below).

It should be noted that many standard results and proof techniques from finite games do *not* carry over to countably infinite games. E.g.,

- Even if a state has value, an optimal strategy need not exist, not even for reachability objectives [19].
- Some strong determinacy properties (see below) do not hold, not even for reachability objectives [4], [18] (while in finite games they hold even for parity objectives [8]).
- The memory requirements of optimal strategies are different. In finite games, optimal strategies for parity objectives can be chosen memoryless deterministic [8]. In contrast, in countably infinite games (even if finitely branching) optimal strategies for reachability objectives, where they exist, require infinite memory [19].

One of the reasons underlying this difference is the following. Consider the values of the states in a game w.r.t. a certain objective. If the game is finite then there are only finitely

Extended version of material presented at LICS 2017. arXiv.org - CC BY 4.0.

Objective	> 0	$> c$	$\geq c$	$= 1$
Reachability	✓(MD)	✓(MD)	✓(¬FR)	✓(**MD**)
Büchi	✔(¬FR)	✖	✖	✔(**MD**)
Borel	✔(¬FR)	✖	✖	✔(¬FR)

(a) Finitely branching games

Objective	> 0	$> c$	$\geq c$	$= 1$
Reachability	✓(MD)	×	×	✔(¬FR)
Büchi	✔(¬FR)	×	×	✔(¬FR)
Borel	✔(¬FR)	×	×	✔(¬FR)

(b) Infinitely branching games

TABLE I: Summary of determinacy and memory requirement properties for reachability, Büchi and Borel objectives and various probability thresholds. The results for safety and co-Büchi are implicit, e.g., > 0 Büchi is dual to to $= 1$ co-Büchi. Similarly, (Objective, $> c$) is dual to (¬Objective, $\geq c$). The results hold for every constant $c \in (0, 1)$. Tables Ia and Ib show the results for finitely branching and infinitely branching countable games, respectively. "✓(MD)" stands for "strongly MD-determined", "✓(¬FR)" stands for "strongly determined but not strongly FR-determined" and $\times$ stands for "not strongly determined". New results are in boldface. (All these objectives are weakly determined by [20].)

many such values, and in particular there exists some minimal nonzero value (unless all states have value zero). This property does *not carry over* to infinite games. Here the set of states is infinite and the infimum over the nonzero values can be zero. As a consequence, even for a reachability objective, it is possible that all states have value > 0, but still the value of some states is < 1. Such phenomena appear already in infinite-state Markov chains like the classic Gambler's ruin problem with unfair coin tosses in the player's favor (e.g., 0.6 win and 0.4 lose). The value, i.e., the probability of ruin, is always > 0, but still < 1 in every state except the ruin state itself; cf. [15] (Chapt. 14).

Weak determinacy. Using Martin's result [21], Maitra & Sudderth [20] showed that stochastic games with Borel payoffs are *weakly determined*, i.e., all states have value. This very general result holds even for concurrent games and general (not necessarily countable) state spaces. They work in the framework of finitely additive probability theory (under weak assumptions on measures) and only assume a finitely additive law of motion. Also their payoff functions are general bounded Borel measurable functions, not necessarily predicates on plays.

Strong determinacy. Given a predicate $\mathcal{E}$ on plays and a constant $c \in [0, 1]$, strong determinacy of a threshold objective $(\mathcal{E}, \rhd c)$ (where $\rhd \in \{>, \geq\}$) holds iff either the maximizer or the minimizer has a winning strategy, i.e., a strategy that enforces (against any strategy of the other player) that the predicate $\mathcal{E}$ holds with probability $\rhd c$ (resp. $\not\rhd c$). In the case of $(\mathcal{E}, = 1)$, one speaks of an almost-sure $\mathcal{E}$ objective. If the winning strategy of the winning player can be chosen MD (memoryless deterministic) then one says that the threshold objective is strongly MD determined. Similarly for other types of strategies, e.g., FR (finite-memory randomized).

Strong determinacy in finite games. Strong determinacy for almost-sure objectives $(\mathcal{E}, = 1)$ (and for the dual positive probability objectives $(\mathcal{E}, > 0)$) is sometimes called *qualitative determinacy* [17]. In [17, Theorem 3.3] it is shown that *finite* stochastic games with Borel *tail* (i.e., prefix-independent) objectives are qualitatively determined. (We'll show a more general result for countably infinite games and general objectives; see below.) In the special case of parity objectives, even strong MD determinacy holds for any threshold $\rhd c$ [8].

Strong determinacy in infinite games. It was shown in [4], [18], [5] that in finitely branching games with countable state spaces reachability objectives with any threshold $\rhd c$ with $c \in [0, 1]$, are strongly determined. However, the player $\Box$ strategy may need infinite memory [19], and thus reachability objectives are not strongly MD determined. Strong determinacy does *not* hold for infinitely branching reachability games with thresholds $\rhd c$ with $c \in (0, 1)$; cf. Figure 1 in [4].

Our contribution to determinacy. We show that almost-sure Borel objectives are strongly determined for games with *countably infinite* state spaces. (In particular this even holds for infinitely branching games; cf. Table I.) This removes both the restriction to finite games and the restriction to tail objectives of [17, Theorem 3.3], and solves an open problem stated there. (To the best of our knowledge, strong determinacy was open even for almost-sure reachability objectives in infinitely branching countable games.)

On the other hand, we show that, for countable games, $\rhd c$ (co-)Büchi objectives are not strongly determined for any $c \in (0, 1)$, not even if the game graph is finitely branching.

Our contribution to strategy complexity. While $\rhd c$ reachability objectives in finitely branching countable games are not strongly MD determined in general [19], we show that strong MD determinacy holds for many interesting subclasses. In finitely branching games, it holds for strict inequality $> c$ reachability, almost-sure reachability, and in all games where either player $\Box$ does not have any value-decreasing transitions or player $\Diamond$ does not have any value-increasing transitions.

Moreover, we show that almost-sure Büchi objectives (but not almost-sure co-Büchi objectives) are strongly MD determined, provided that the game is finitely branching.

Table I summarizes all properties of strong determinacy and memory requirements for Borel objectives and subclasses on countably infinite games.

II. Preliminaries

A *probability distribution* over a countable (not necessarily finite) set S is a function $f : S \to [0, 1]$ s.t. $\sum_{s \in S} f(s) = 1$. We use $\mathsf{supp}(f) = \{s \in S \mid f(s) > 0\}$ to denote the *support* of f. Let $\mathcal{D}(S)$ be the set of all probability distributions over S.

We consider $2\frac{1}{2}$-player games where players have perfect information and play in turn for infinitely many rounds.

Games $\mathcal{G} = (S, (S_\Box, S_\Diamond, S_\bigcirc), \longrightarrow, P)$ are defined such that the countable set of *states* is partitioned into the set $S_\Box$ of states of player $\Box$, the set $S_\Diamond$ of states of player $\Diamond$ and *random states* $S_\bigcirc$. The relation $\longrightarrow \subseteq S \times S$ is the transition relation. We write $s \longrightarrow s'$ if $(s, s') \in \longrightarrow$, and we assume that each state s has a *successor* state s' with $s \longrightarrow s'$. The probability function $P : S_\bigcirc \to \mathcal{D}(S)$ assigns to each random state $s \in S_\bigcirc$ a probability distribution over its successor states. The game $\mathcal{G}$ is called *finitely branching* if each state has only finitely many successors; otherwise, it is *infinitely branching*. Let $\odot \in \{\Box, \Diamond\}$. If $S_\odot = \emptyset$, we say that player $\odot$ is *passive*, and the game is a *Markov decision process (MDP)*. A Markov chain is an MDP where both players are passive.

The stochastic game is played by two players $\Box$ (maximizer) and $\Diamond$ (minimizer). The game starts in a given initial state s_0 and evolves for infinitely many rounds. In each round, if the game is in state $s \in S_\odot$ then player $\odot$ chooses a successor state s' with $s \longrightarrow s'$; otherwise the game is in a random state $s \in S_\bigcirc$ and proceeds randomly to s' with probability $P(s)(s')$.

Strategies. A *play* w is an infinite sequence $s_0 s_1 \cdots \in S^\omega$ of states such that $s_i \longrightarrow s_{i+1}$ for all $i \geq 0$; let $w(i) = s_i$ denote the i-th state along w. A *partial play* is a finite prefix of a play. We say that (partial) play w *visits* s if $s = w(i)$ for some i, and that w starts in s if $s = w(0)$. A *strategy* of the player $\Box$ is a function $\sigma : S^* S_\Box \to \mathcal{D}(S)$ that assigns to partial plays $ws \in S^* S_\Box$ a distribution over the successors $\{s' \in S \mid s \longrightarrow s'\}$. Strategies $\pi : S^* S_\Diamond \to \mathcal{D}(S)$ for the player $\Diamond$ are defined analogously. The set of all strategies of player $\Box$ and player $\Diamond$ in $\mathcal{G}$ is denoted by $\Sigma_\mathcal{G}$ and $\Pi_\mathcal{G}$, respectively (we omit the subscript and write Σ and Π if $\mathcal{G}$ is clear). A (partial) play $s_0 s_1 \cdots$ is induced by strategies (σ, π) if $s_{i+1} \in \mathsf{supp}(\sigma(s_0 s_1 \cdots s_i))$ for all $s_i \in S_\Box$, and if $s_{i+1} \in \mathsf{supp}(\pi(s_0 s_1 \cdots s_i))$ for all $s_i \in S_\Diamond$.

To emphasize the amount of memory required to implement a strategy, we present an equivalent formulation of strategies. A strategy of player $\odot$ can be implemented by a probabilistic transducer $\mathsf{T} = (\mathsf{M}, \mathsf{m}_0, \pi_u, \pi_s)$ where M is a countable set (the memory of the strategy), $\mathsf{m}_0 \in \mathsf{M}$ is the initial memory mode and S is the input and output alphabet. The probabilistic transition function $\pi_u : \mathsf{M} \times S \to \mathcal{D}(\mathsf{M})$ updates the memory mode of the transducer. The probabilistic successor function $\pi_s : \mathsf{M} \times S_\odot \to \mathcal{D}(S)$ outputs the next successor, where $s' \in \mathsf{supp}(\pi_s(\mathsf{m}, s))$ implies $s \longrightarrow s'$. We extend π_u to $\mathcal{D}(\mathsf{M}) \times S \to \mathcal{D}(\mathsf{M})$ and π_s to $\mathcal{D}(\mathsf{M}) \times S_\odot \to \mathcal{D}(S)$, in the natural way. Moreover, we extend π_u to paths by $\pi_u(\mathsf{m}, \varepsilon) = \mathsf{m}$ and $\pi_u(\mathsf{m}, s_0 \cdots s_n) = \pi_u(\pi_u(s_0 \cdots s_{n-1}, \mathsf{m}), s_n)$. The strategy $\tau_\mathsf{T} : S^* S_\odot \to \mathcal{D}(S)$ induced by the transducer T is given by $\tau_\mathsf{T}(s_0 \cdots s_n) := \pi_s(s_n, \pi_u(s_0 \cdots s_{n-1}, \mathsf{m}_0))$.

Strategies are in general *history dependent* (H) and *randomized* (R). An H-strategy $\tau \in \{\sigma, \pi\}$ is *finite memory* (F) if there exists some transducer T with memory M such that $\tau_\mathsf{T} = \tau$ and $|\mathsf{M}| < \infty$; otherwise τ *requires infinite memory*. An F-strategy is *memoryless* (M) (also called *positional*) if $|\mathsf{M}| = 1$. For convenience, we may view M-strategies as functions $\tau : S_\odot \to \mathcal{D}(S)$. An R-strategy τ is *deterministic* (D) if π_u and π_s map to Dirac distributions; it implies that $\tau(w)$ is a Dirac distribution for all partial plays w. All combinations of the properties in $\{\mathrm{M}, \mathrm{F}, \mathrm{H}\} \times \{\mathrm{D}, \mathrm{R}\}$ are possible, e.g., MD stands for memoryless deterministic. HR strategies are the most general type.

Probability Measure and Events. To a game $\mathcal{G}$, an initial state s_0 and strategies (σ, π) we associate the standard probability space $(s_0 S^\omega, \mathcal{F}, \mathcal{P}_{\mathcal{G}, s_0, \sigma, \pi})$ w.r.t. the induced Markov chain. First one defines a topological space on the set of infinite plays $s_0 S^\omega$. The *cylinder sets* are the sets $s_0 s_1 \ldots s_n S^\omega$, where $s_1, \ldots, s_n \in S$ and the open sets are arbitrary unions of cylinder sets, i.e., the sets $Y S^\omega$ with $Y \subseteq s_0 S^*$. The Borel σ-algebra $\mathcal{F} \subseteq 2^{s_0 S^\omega}$ is the smallest σ-algebra that contains all the open sets.

The probability measure $\mathcal{P}_{\mathcal{G}, s_0, \sigma, \pi}$ is obtained by first defining it on the cylinder sets and then extending it to all sets in the Borel σ-algebra. If $s_0 s_1 \ldots s_n$ is not a partial play induced by (σ, π) then let $\mathcal{P}_{\mathcal{G}, s_0, \sigma, \pi}(s_0 s_1 \ldots s_n S^\omega) = 0$; otherwise let $\mathcal{P}_{\mathcal{G}, s_0, \sigma, \pi}(s_0 s_1 \ldots s_n S^\omega) = \prod_{i=0}^{n-1} \tau(s_0 s_1 \ldots s_i)(s_{i+1})$, where τ is such that $\tau(ws) = \sigma(ws)$ for all $ws \in S^* S_\Box$, $\tau(ws) = \pi(ws)$ for all $ws \in S^* S_\Diamond$, and $\tau(ws) = P(s)$ for all $ws \in S^* S_\bigcirc$. By Carathéodory's extension theorem [2], this defines a unique probability measure $\mathcal{P}_{\mathcal{G}, s_0, \sigma, \pi}$ on the Borel σ-algebra $\mathcal{F}$.

We will call any set $\mathcal{E} \in \mathcal{F}$ an *event*, i.e., an event is a measurable (in the probability space above) set of infinite plays. Equivalently, one may view an event $\mathcal{E}$ as a Borel measurable payoff function of the form $\mathcal{E} : s_0 S^\omega \to \{0, 1\}$. Given $\mathcal{E}' \subseteq S^\omega$ (where potentially $\mathcal{E}' \not\subseteq s_0 S^\omega$) we often write $\mathcal{P}_{\mathcal{G}, s_0, \sigma, \pi}(\mathcal{E}')$ for $\mathcal{P}_{\mathcal{G}, s_0, \sigma, \pi}(\mathcal{E}' \cap s_0 S^\omega)$ to avoid clutter.

Objectives. Let $\mathcal{G} = (S, (S_\Box, S_\Diamond, S_\bigcirc), \longrightarrow, P)$ be a game. The objectives of the players are determined by events $\mathcal{E}$. We write $\neg\mathcal{E}$ for the dual objective defined as $\neg\mathcal{E} = S^\omega \setminus \mathcal{E}$.

Given a target set $\mathcal{T} \subseteq S$, the *reachability objective* is defined by the event

$$\mathtt{Reach}(\mathcal{T}) = \{s_0 s_1 \cdots \in S^\omega \mid \exists i.\, s_i \in \mathcal{T}\}.$$

Moreover, $\mathtt{Reach}_n(\mathcal{T})$ denotes the set of all plays visiting $\mathcal{T}$ in the first n steps, i.e., $\mathtt{Reach}_n(\mathcal{T}) = \{s_0 s_1 \cdots \mid \exists i \leq n.\, s_i \in \mathcal{T}\}$. The *safety objective* is defined as the dual of reachability: $\mathtt{Safety}(\mathcal{T}) = \neg\mathtt{Reach}(\mathcal{T})$.

For a set $\mathcal{T} \subseteq S$ of states called *Büchi states*, the *Büchi objective* is the event

$$\mathtt{Büchi}(\mathcal{T}) = \{s_0 s_1 \cdots \in S^\omega \mid \forall i\, \exists j \geq i.\, s_j \in \mathcal{T}\}.$$

The *co-Büchi objective* is defined as the dual of Büchi.

Note that the objectives of player $\Box$ (maximizer) and player $\Diamond$ (minimizer) are dual to each other. Where player $\Box$ tries to maximize the probability of some objective $\mathcal{E}$, player $\Diamond$ tries to maximize the probability of $\neg\mathcal{E}$.

III. Determinacy

A. Optimal and ϵ-Optimal Strategies; Weak and Strong Determinacy

Given an objective $\mathcal{E}$ for player $\Box$ in a game $\mathcal{G}$, state s has *value* if

$$\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}_{\mathcal{G},s,\sigma,\pi}(\mathcal{E}) = \inf_{\pi\in\Pi}\sup_{\sigma\in\Sigma}\mathcal{P}_{\mathcal{G},s,\sigma,\pi}(\mathcal{E}).$$

If s has value then $\mathtt{val}_{\mathcal{G}}(s)$ denotes the value of s defined by the above equality. A game with a fixed objective is called *weakly determined* iff every state has value.

Theorem 1 (follows immediately from [20])**.** *Countable stochastic games (as defined in Section II) are weakly determined.*

Theorem 1 is an immediate consequence of a far more general result by Maitra & Sudderth [20] on weak determinacy of (finitely additive) games with general Borel payoff objectives.

For $\epsilon \geq 0$ and $s \in S$, we say that

- $\sigma \in \Sigma$ is *ϵ-optimal (maximizing)* iff $\mathcal{P}_{\mathcal{G},s,\sigma,\pi}(\mathcal{E}) \geq \mathtt{val}_{\mathcal{G}}(s) - \epsilon$ for all $\pi \in \Pi$.
- $\pi \in \Pi$ is *ϵ-optimal (minimizing)* iff $\mathcal{P}_{\mathcal{G},s,\sigma,\pi}(\mathcal{E}) \leq \mathtt{val}_{\mathcal{G}}(s) + \epsilon$ for all $\sigma \in \Sigma$.

A 0-optimal strategy is called *optimal*. An optimal strategy for the player $\Box$ is *almost-surely* winning if $\mathtt{val}_{\mathcal{G}}(s) = 1$. Unlike in finite-state games, optimal strategies need not exist in countable games, not even for reachability objectives in finitely branching MDPs [3], [4].

However, since our games are weakly determined by Theorem 1, for all $\epsilon > 0$ there exist ϵ-optimal strategies for both players.

For an objective $\mathcal{E}$ and $\rhd \in \{\geq, >\}$ and threshold $c \in [0,1]$, we define *threshold objectives* $(\mathcal{E}, \rhd c)$ as follows.

- $[\mathcal{E}]^{\rhd c}_{\Box\,\mathcal{G}}$ is the set of states s for which there exists a strategy σ such that, for all $\pi \in \Pi$, we have $\mathcal{P}_{\mathcal{G},s,\sigma,\pi}(\mathcal{E}) \rhd c$.
- $[\mathcal{E}]^{\not\rhd c}_{\Diamond\,\mathcal{G}}$ is the set of states s for which there exists a strategy π such that, for all $\sigma \in \Sigma$, we have $\mathcal{P}_{\mathcal{G},s,\sigma,\pi}(\mathcal{E}) \not\rhd c$.

We omit the subscript $\mathcal{G}$ where it is clear from the context. We call a state s *almost-surely winning* for the player $\Box$ iff $s \in [\mathcal{E}]^{\geq 1}_{\Box}$.

By the duality of the players, a $(\mathcal{E}, \geq c)$ objective for player $\Box$ corresponds to a $(\neg\mathcal{E}, > 1-c)$ objective from player $\Diamond$'s point of view. E.g., an almost-sure Büchi objective for player $\Box$ corresponds to a positive-probability co-Büchi objective for player $\Diamond$. Thus we can restrict our attention to reachability, Büchi and general (Borel set) objectives, since safety is dual to reachability, and co-Büchi is dual to Büchi, and Borel is self-dual.

A game $\mathcal{G}$ with threshold objective $(\mathcal{E}, \rhd c)$ is called *strongly determined* iff in every state s either player $\Box$ or player $\Diamond$ has a winning strategy, i.e., iff $S = [\mathcal{E}]^{\rhd c}_{\Box} \uplus [\mathcal{E}]^{\not\rhd c}_{\Diamond}$.

Strong determinacy depends on the specified threshold $\rhd c$. Strong determinacy for almost-sure objectives $(\mathcal{E}, = 1)$ (and for the dual positive probability objectives $(\mathcal{E}, > 0)$) is sometimes called *qualitative determinacy* [17]. In [17, Theorem 3.3] it is shown that *finite* stochastic games with *tail* objectives are qualitatively determined. An objective $\mathcal{E}$ is called *tail* if for all $w_0 \in S^*$ and all $w \in S^\omega$ we have $w_0 w \in \mathcal{E} \Leftrightarrow w \in \mathcal{E}$, i.e., a tail objective is independent of finite prefixes. The authors of [17] express "hope that [their qualitative determinacy theorem] may be extended beyond the class of finite simple stochastic tail games". We fulfill this hope by generalizing their theorem from finite to countable games and from tail objectives to arbitrary objectives:

Theorem 2. *Stochastic games, even infinitely branching ones, with almost-sure objectives are strongly determined.*

Theorem 2 does not carry over to thresholds other than 0 or 1; cf. Theorem 3.

The main ingredients of the proof of Theorem 2 are transfinite induction, weak determinacy of stochastic games (Theorem 1), the concept of a "reset" strategy from [17], and Lévy's zero-one law. The principal idea of the proof is to construct a transfinite sequence of subgames, by removing parts of the game that player $\Box$ cannot risk entering. This approach is used later in this paper as well, for Theorems 5 and 11.

Example 1. *We explain this approach using the reachability game in Figure 1 as an example. Each state has value* 1 *in this game, except those labeled with* 0*. However, only the states labeled with $\bot$ are almost-surely winning for player $\Box$. To see this, consider a player $\Box$ state labeled with* 1*. In order to reach $\mathcal{T}$, player $\Box$ eventually needs to take a transition to a* 0*-labeled state, which is not almost-surely winning. This means that the* 1*-labeled states are not almost-surely winning either. Hence, player $\Box$ cannot risk entering them if the player wants to win almost surely. Continuing this style of reasoning, we infer that the* 2*-labeled states are not almost-surely winning, and so on. This implies that the ω-labeled states are not almost-surely winning, and so on. The only almost-surely winning player $\Box$ state is the $\bot$-labeled state at the bottom of the figure, and the only winning strategy is to take the direct transition to the target in the bottom-left corner.*

Proof of Theorem 2. The first step of the proof is to transform the game and the objective so that the objective can in some respects be treated like a tail objective. Let $\hat{\mathcal{G}}$ be a stochastic game with countable state space $\hat{S}$ and objective $\hat{\mathcal{E}}$. We convert the game graph to a forest by encoding the history in the states. Formally we proceed as follows. The state space, S, of the new game, $\mathcal{G}$, consists of the partial plays in $\hat{\mathcal{G}}$, i.e., $S \subseteq \hat{S}^*\hat{S}$. Observe that S is countable. For any $\odot \in \{\Box, \Diamond, \bigcirc\}$ we define $S_\odot := \{w\hat{s} \in S \mid \hat{s} \in \hat{S}_\odot\}$. A transition is a transition of $\mathcal{G}$ iff it is of the form $w\hat{s} {\longrightarrow} w\hat{s}\hat{s}'$ where $w\hat{s} \in S$ and $\hat{s} {\longrightarrow} \hat{s}'$ is a transition in $\hat{\mathcal{G}}$. The probabilities in $\mathcal{G}$ are defined in the obvious way. For $\hat{s} \in \hat{S}$ we define an objective $\mathcal{E}_{\hat{s}}$ so that a play in $\mathcal{G}$ starting from the singleton $\hat{s} \in S$ satisfies $\mathcal{E}_{\hat{s}}$ iff the corresponding play from $\hat{s} \in \hat{S}$ in $\hat{\mathcal{G}}$ satisfies $\hat{\mathcal{E}}$. Since strategies in $\mathcal{G}$ (for singleton initial states in $\hat{S}$) carry over to strategies in $\hat{\mathcal{G}}$, it suffices to prove our determinacy result for $\mathcal{G}$.

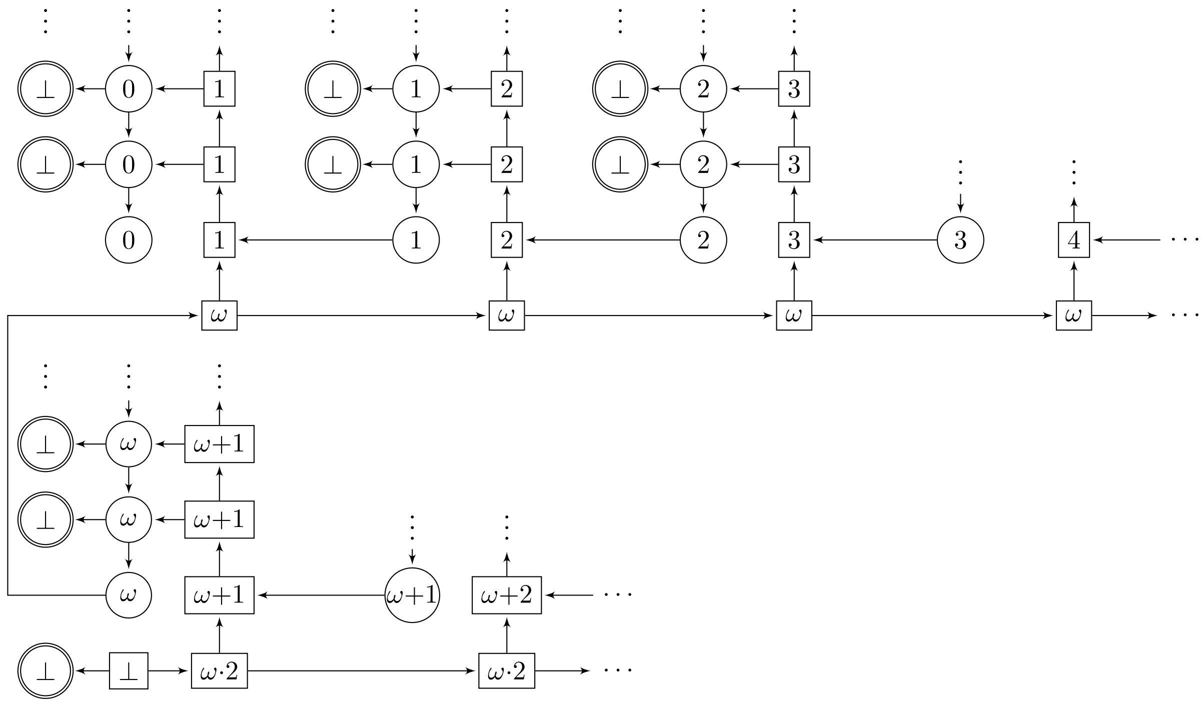

Fig. 1: A finitely branching reachability game where the states of player □ are drawn as squares and the random states as circles. Player ◇ is passive in this game. The states with double borders form the target set $\mathcal{T}$; those states have self-loops which are not drawn in the figure. For each random state, the distribution over the successors is uniform. Each state is labeled with an ordinal, which indicates the *index* of the state. In particular, the example shows that transfinite indices are needed.

Let us inductively extend the definition of $\mathcal{E}_s$ from $s = \hat{s} \in \hat{S}$ to arbitrary $s \in S$. For any transition $s{\longrightarrow}s'$ in $\mathcal{G}$, define $\mathcal{E}_{s'} := \{x \in s'S^\omega \mid sx \in \mathcal{E}_s\}$. This is well-defined as the transition graph of $\mathcal{G}$ is a forest. For any $s \in S$, the event $\mathcal{E}_s$ is also measurable. By this construction we obtain the following property: If a play y in $\mathcal{G}$ visits states $s, s' \in S$ then the suffix of y starting from s satisfies $\mathcal{E}_s$ iff the suffix of y starting from s' satisfies $\mathcal{E}_{s'}$. This property is weaker than the tail property (which would stipulate that all $\mathcal{E}_s$ are equivalent), but it suffices for our purposes.

In the remainder of the proof, when $\mathcal{G}'$ is (a subgame of) $\mathcal{G}$, we write $\mathcal{P}_{\mathcal{G}',s,\sigma,\pi}(\mathcal{E})$ for $\mathcal{P}_{\mathcal{G}',s,\sigma,\pi}(\mathcal{E}_s)$ to avoid clutter. Similarly, when we write $\mathtt{val}_{\mathcal{G}'}(s)$ we mean the value with respect to $\mathcal{E}_s$.

In order to characterize the winning sets of the players, we construct a transfinite sequence of subgames $\mathcal{G}_\alpha$ of $\mathcal{G}$, where $\alpha \in \mathbb{O}$ is an ordinal number, by stepwise removing certain states that are losing for player □, along with their incoming transitions. Thus some subgames $\mathcal{G}_\alpha$ may contain states without any outgoing transitions (i.e., dead ends). Such dead ends are always considered as losing for player □. (Formally, one might add a self-loop to such states and remove from the objective all plays that reach these states.)

Let S_α denote the state space of the subgame $\mathcal{G}_\alpha$. We start with $\mathcal{G}_0 := \mathcal{G}$. Given $\mathcal{G}_\alpha$, denote by D_α the set of states $s \in S_\alpha$ with $\mathtt{val}_{\mathcal{G}_\alpha}(s) < 1$. For any $\alpha \in \mathbb{O} \setminus \{0\}$ we define $S_\alpha := S \setminus \bigcup_{\gamma<\alpha} D_\gamma$.

Since the sequence of sets S_α is non-increasing and $S_0 = S$ is countable, it follows that this sequence of games $\mathcal{G}_\alpha$ converges (i.e., is ultimately constant) at some ordinal β where $\beta \leq \omega_1$ (the first uncountable ordinal). That is, we have $\mathcal{G}_\beta = \mathcal{G}_{\beta+1}$. Note in particular that $\mathcal{G}_\beta$ does not contain any dead ends. (However, its state space S_β might be empty. In this case it is considered to be losing for player □.)

We define the *index*, $I(s)$, of a state s as the smallest ordinal α with $s \in D_\alpha$, and as $\perp$ if such an ordinal does not exist. For all states $s \in S$ we have:

$$I(s) = \perp \;\Leftrightarrow\; s \in S_\beta \;\Leftrightarrow\; \mathtt{val}_{\mathcal{G}_\beta}(s) = 1$$

We show that states s with $I(s) \in \mathbb{O}$ are in $[\mathcal{E}]^{<1}_{\Diamond\,\mathcal{G}}$, and states s with $I(s) = \perp$ are in $[\mathcal{E}]^{=1}_{\Box\,\mathcal{G}}$.

Strategy $\hat{\pi}_s$: For each $s \in S$ with $I(s) \in \mathbb{O}$ we construct a player ◇ strategy $\hat{\pi}_s$ such that $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}_s}(\mathcal{E}) < 1$ holds for all player □ strategies σ. The strategy $\hat{\pi}_s$ is defined inductively over the index $I(s)$.

Let $s \in S$ with $I(s) = \alpha \in \mathbb{O}$. In game $\mathcal{G}_\alpha$ we have $\mathtt{val}_{\mathcal{G}_\alpha}(s) < 1$. So by weak determinacy (Theorem 1) there is a strategy $\hat{\pi}_s$ with $\mathcal{P}_{\mathcal{G}_\alpha,s,\sigma,\hat{\pi}_s}(\mathcal{E}) < 1$ for all σ. (For example, one may take a $(1-\mathtt{val}_{\mathcal{G}_\alpha}(s))/2$-optimal player ◇ strategy). We extend $\hat{\pi}_s$ to a strategy in $\mathcal{G}$ as follows. Whenever the play enters a state $s' \notin S_\alpha$ (hence $I(s') < \alpha$) then $\hat{\pi}_s$ switches to the previously defined strategy $\hat{\pi}_{s'}$. (One could show that only player $\Box$ can take a transition leaving S_α, although this is not needed at the moment.)

We show by transfinite induction on the index that $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}_s}(\mathcal{E}) < 1$ holds for all player $\Box$ strategies σ and for all states $s \in S$ with $I(s) \in \mathbb{O}$.

For the induction hypothesis, let α be an ordinal for which this holds for all states s with $I(s) < \alpha$. For the inductive step, let $s \in S$ be a state with $I(s) = \alpha$, and let σ be an arbitrary player $\Box$ strategy in $\mathcal{G}$.

Suppose that the play from s under the strategies $\sigma, \hat{\pi}_s$ always remains in S_α, i.e., the probability of ever leaving S_α under $\sigma, \hat{\pi}_s$ is zero. Then any play in $\mathcal{G}$ under these strategies coincides with a play in $\mathcal{G}_\alpha$, so we have $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}_s}(\mathcal{E}) = \mathcal{P}_{\mathcal{G}_\alpha,s,\sigma,\hat{\pi}_s}(\mathcal{E}) < 1$, as desired. Now suppose otherwise, i.e., the play from s under $\sigma, \hat{\pi}_s$, with positive probability, enters a state $s' \notin S_\alpha$, hence $I(s') < \alpha$. By the induction hypothesis we have $\mathcal{P}_{\mathcal{G},s',\sigma',\hat{\pi}_{s'}}(\mathcal{E}) < 1$ for any σ'. Since the probability of entering s' is positive, we conclude $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}_s}(\mathcal{E}) < 1$, as desired.

Strategy $\hat{\sigma}$: For each $s \in S$ with $I(s) = \bot$ (and thus $s \in S_\beta$) we construct a player $\Box$ strategy $\hat{\sigma}$ such that $\mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathcal{E}) = 1$ holds for all player $\Diamond$ strategies π. We first observe that if $s_1 \longrightarrow s_2$ is a transition in $\mathcal{G}$ with $s_1 \in S_\Diamond \cup S_\bigcirc$ and $I(s_2) \neq \bot$ then $I(s_1) \neq \bot$. Indeed, let $I(s_2) = \alpha \in \mathbb{O}$, thus $\mathtt{val}_{\mathcal{G}_\alpha}(s_2) < 1$; if $s_1 \in S_\alpha$ then $\mathtt{val}_{\mathcal{G}_\alpha}(s_1) < 1$ and thus $I(s_1) = \alpha$; if $s_1 \notin S_\alpha$ then $I(s_1) < \alpha$. It follows that only player $\Box$ could ever leave the state space S_β, but our player $\Box$ strategy $\hat{\sigma}$ will ensure that the play remains in S_β forever. Recall that $\mathcal{G}_\beta$ does not contain any dead ends and that $\mathtt{val}_{\mathcal{G}_\beta}(s) = 1$ for all $s \in S_\beta$. For all $s \in S_\beta$, by weak determinacy (Theorem 1) we fix a strategy σ_s with $\mathcal{P}_{\mathcal{G}_\beta,s,\sigma_s,\pi}(\mathcal{E}) \geq 2/3$ for all π.

Fix an arbitrary state $s_0 \in S_\beta$ as the initial state. For a player $\Box$ strategy σ, define mappings $X_1^\sigma, X_2^\sigma, \ldots : s_0 S^\omega \to [0,1]$ using conditional probabilities:

$$X_i^\sigma(w) := \inf_{\pi \in \Pi_{\mathcal{G}_\beta}} \mathcal{P}_{\mathcal{G}_\beta,s_0,\sigma,\pi}(\mathcal{E} \mid E_i(w)),$$

where $E_i(w)$ denotes the event containing the plays that start with the length-i prefix of $w \in s_0 S^\omega$. Thanks to our "forest" construction at the beginning of the proof, $X_i^\sigma(w)$ depends, in fact, only on the i-th state visited by w.

For some illustration, a small value of $X_i^\sigma(w)$ means that considering the length-i prefix of w, player $\Diamond$ has a strategy that makes $\mathcal{E}$ unlikely at time i. Similarly, a large value of $X_i^\sigma(w)$ means that at time i (when the length-i prefix has been "uncovered") the probability of $\mathcal{E}$ using σ is large, regardless of the player $\Diamond$ strategy.

In the following we view X_i^σ as a random variable (taking on a random value depending on a random play).

We define our almost-surely winning player $\Box$ strategy $\hat{\sigma}$ as the limit of inductively defined strategies $\hat{\sigma}_0, \hat{\sigma}_1, \ldots$. Let $\hat{\sigma}_0 := \sigma_{s_0}$. Using the definition of σ_{s_0} we get $X_1^{\hat{\sigma}_0} \geq 2/3$. For any $k \in \mathbb{N}$, define $\hat{\sigma}_{k+1}$ as follows. Strategy $\hat{\sigma}_{k+1}$ plays $\hat{\sigma}_k$ as long as $X_i^{\hat{\sigma}_k} \geq 1/3$. This could be forever. Otherwise, let i denote the smallest i with $X_i^{\hat{\sigma}_k} < 1/3$, and let s be the i-th state of the play. At that time, $\hat{\sigma}_{k+1}$ switches to strategy σ_s, implying $X_i^{\hat{\sigma}_{k+1}} \geq 2/3$. This switch of strategy is referred to as a "reset" in [17], where the concept is used similarly. For any k, strategy $\hat{\sigma}_k$ performs at most k such resets. Define $\hat{\sigma}$ as the limit of the $\hat{\sigma}_k$, i.e., the number of resets performed by $\hat{\sigma}$ is unbounded.

In order to show that $\hat{\sigma}$ is almost surely winning, we first argue that $\hat{\sigma}$ almost surely performs only a finite number of resets. Suppose $w \in S^\omega$ and k, i are such that a k-th reset happens after visiting the i-th state in w. As argued above, we have $X_i^{\hat{\sigma}_k}(w) \geq 2/3$. Towards a contradiction assume that player $\Diamond$ has a strategy π_1 to cause yet another reset with probability $p_1 > 1/2$, i.e.,

$$p_1 := \mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma}_k,\pi_1}(R \mid E_i(w)) > 1/2\,,$$

where R denotes the event of another reset after time i. If another reset occurs, say at time j, then $X_j^{\hat{\sigma}_k}(w) < 1/3$, and then player $\Diamond$ can switch to a strategy π_2 to force $\mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma}_k,\pi_2}(\mathcal{E} \mid E_j(w)) \leq 1/3$. Hence:

$$p_2 := \mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma}_k,\pi_2}(\mathcal{E} \mid R \wedge E_i(w)) \leq 1/3$$

Let $\pi_{1,2}$ denote the player $\Diamond$ strategy combining π_1 and π_2. Then it follows:

$$\begin{aligned}\mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma}_k,\pi_{1,2}}(\mathcal{E} \wedge R \mid E_i(w)) &= p_1 \cdot p_2 \qquad \text{and}\\ \mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma}_k,\pi_{1,2}}(\mathcal{E} \wedge \neg R \mid E_i(w)) &\leq \mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma}_k,\pi_{1,2}}(\neg R \mid E_i(w))\\ &= 1 - p_1\end{aligned}$$

Hence we have:

$$\begin{aligned}\mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma}_k,\pi_{1,2}}(\mathcal{E} \mid E_i(w)) &\leq (p_1 \cdot p_2) + (1 - p_1) \leq 1 - \frac{2}{3}p_1\\ &< 1 - \frac{2}{3} \cdot \frac{1}{2} = \frac{2}{3}\,,\end{aligned}$$

contradicting $X_i^{\hat{\sigma}_k}(w) \geq 2/3$. So at time i, the probability of another reset is bounded by $1/2$. Since this holds for every reset time i, we conclude that almost surely there will be only finitely many resets under $\hat{\sigma}$, regardless of π.

Now we can show that $\mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma},\pi}(\mathcal{E}) = 1$ holds for all π. Fix π arbitrarily. For $k \in \mathbb{N}$ define Q_k as the event that exactly k resets occur. Let us write $\mathcal{P}_k = \mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma}_k,\pi}$ to avoid clutter. By Lévy's zero-one law (see, e.g., [25, Theorem 14.2]), for any k, we have $\mathcal{P}_k$-almost surely that either

$$(\mathcal{E} \vee \neg Q_k) \wedge \lim_{i\to\infty} \mathcal{P}_k(\mathcal{E} \vee \neg Q_k \mid E_i(w)) = 1$$

or

$$(\neg\mathcal{E} \wedge Q_k) \wedge \lim_{i\to\infty} \mathcal{P}_k(\mathcal{E} \vee \neg Q_k \mid E_i(w)) = 0$$

holds. Let w be a play that satisfies the second option. In particular, $w \in Q_k$, so there exists $i_0 \in \mathbb{N}$ with $X_i^{\hat{\sigma}_k}(w) \geq 1/3$ for all $i \geq i_0$. It follows that $\mathcal{P}_k(\mathcal{E} \mid E_i(w)) \geq 1/3$ holds for all $i \geq i_0$. But that contradicts the fact that $\lim_{i\to\infty} \mathcal{P}_k(\mathcal{E} \vee \neg Q_k \mid E_i(w)) = 0$. So plays satisfying the second option do not actually exist.

Hence we conclude $\mathcal{P}_k(\mathcal{E} \vee \neg Q_k) = 1$, thus $\mathcal{P}_k(\neg\mathcal{E} \wedge Q_k) = 0$. Since the strategies $\hat{\sigma}$ and $\hat{\sigma}_k$ agree on all finite prefixes of all plays in Q_k, the probability measures $\mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma},\pi}$ and $\mathcal{P}_k$ agree on all subevents of Q_k. It follows $\mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma},\pi}(\neg\mathcal{E} \wedge Q_k) =$

0. We have shown previously that the number of resets is almost surely finite, i.e., $\mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma},\pi}(\bigvee_{k\in\mathbb{N}} Q_k) = 1$. Hence we have:

$$\begin{aligned}\mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma},\pi}(\neg\mathcal{E}) &= \mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma},\pi}\Big(\neg\mathcal{E} \wedge \bigvee_{k\in\mathbb{N}} Q_k\Big) \\ &\leq \sum_{k\in\mathbb{N}} \mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma},\pi}(\neg\mathcal{E} \wedge Q_k) \\ &= 0\end{aligned}$$

Thus, $\mathcal{P}_{\mathcal{G}_\beta,s_0,\hat{\sigma},\pi}(\mathcal{E}) = 1$. Since $\hat{\sigma}$ is defined on $\mathcal{G}_\beta$, this strategy never leaves S_β. Since only player $\Box$ might have transitions that leave S_β, we conclude $\mathcal{P}_{\mathcal{G},s_0,\hat{\sigma},\pi}(\mathcal{E}) = 1$. $\Box$

B. Reachability and Safety

It was shown in [4] and [18] (and also follows as a corollary from [5]) that finitely branching games with reachability objectives with any threshold $\rhd c$ with $c \in [0,1]$ are strongly determined. In contrast, strong determinacy does not hold for infinitely branching reachability games with thresholds $\rhd c$ with $c \in (0,1)$; cf. Figure 1 in [4]. However, by Theorem 2, strong determinacy does hold for almost-sure reachability and safety objectives in infinitely branching games. By duality, this also holds for reachability and safety objectives with threshold >0. (For almost-sure safety (resp. >0 reachability), this could also be shown by a reduction to non-stochastic 2-player reachability games [26].)

C. Büchi and co-Büchi

Let $\mathcal{E}$ be the Büchi objective (the co-Büchi objective is dual). Again, Theorem 2 applies to almost-sure and positive-probability Büchi and co-Büchi objectives, so those games are strongly determined, even infinitely branching ones.

However, this does not hold for thresholds $c \in (0,1)$, not even for finitely branching games:

Theorem 3. *Threshold (co-)Büchi objectives $(\mathcal{E}, \rhd c)$ with thresholds $c \in (0,1)$ are not strongly determined, even for finitely branching games.*

A fortiori, threshold parity objectives are not strongly determined, not even for finitely branching games. We prove Theorem 3 using the finitely branching game in Figure 2. It is inspired by an infinitely branching example in [4], where it was shown that threshold reachability objectives in infinitely branching games are not strongly determined.

Proof sketch of Theorem 3. The game in Figure 2 is finitely branching, and we consider the Büchi objective. The infinite choice for player $\Diamond$ in the example of [4] is simulated with an infinite chain $s'_0 s'_1 s'_2 \cdots$ of Büchi states in our example. All states $s'_0 s'_1 s'_2 \cdots$ are finitely branching and belong to player $\Diamond$. The crucial property is that player $\Diamond$ can stay in the states s'_i for arbitrarily long (thus making the probability of reaching the state t arbitrarily small) but not forever. Since the states s'_i are Büchi states, plays that stay in them forever satisfy the Büchi objective surely, something that player $\Diamond$ needs to avoid. So a player $\Diamond$ strategy must choose a transition $s'_i \longrightarrow r'_i$ for some $i \in \mathbb{N}$, resulting in a faithful simulation of infinite branching from s'_0 to some state r'_i, just like in the reachability game in [4].

From the fact that $\mathtt{val}_{\mathcal{G}}(r_i) = 1 - 2^{-i}$ and $\mathtt{val}_{\mathcal{G}}(r'_i) = 2^{-i}$, we deduce the following properties of this game:

- $\mathtt{val}_{\mathcal{G}}(s_0) = 1$, but there exists no optimal strategy starting in s_0. The value is witnessed by a family of ϵ-optimal strategies σ_i: traversing the ladder $s_0 s_1 \cdots s_i$ and choosing $s_i \longrightarrow r_i$.
- $\mathtt{val}_{\mathcal{G}}(s'_0) = 0$, but there exists no optimal minimizing strategy starting in s'_0; however, in analogy with s_i, there are ϵ-optimal strategies.
- $\mathtt{val}_{\mathcal{G}}(i) = \frac{1}{2}$. We argue below that neither player has an optimal strategy starting in i. It follows that $i \notin [\mathcal{E}]_{\Box}^{\geq\frac{1}{2}} \uplus [\mathcal{E}]_{\Diamond}^{\not\geq\frac{1}{2}}$ for the Büchi condition φ. So neither player has a winning strategy, neither for $(\mathcal{E}, \geq 1/2)$ nor for $(\mathcal{E}, >1/2)$. Indeed, consider any player $\Box$ strategy σ. Following σ, once the game is in s_0, Büchi states cannot be visited with probability more than $\frac{1}{2}\cdot(1-\epsilon)$ for some fixed $\epsilon > 0$ and all strategies π. Player $\Diamond$ has an $\frac{\epsilon}{2}$-optimal strategy π starting in s'_0. Then we have:

$$\mathcal{P}_{\mathcal{G},i,\sigma,\pi}(\mathcal{E}) \leq \frac{1}{2}\cdot(1-\epsilon) + \frac{1}{2}\cdot\frac{\epsilon}{2} < \frac{1}{2},$$

  so σ is not optimal. One can argue symmetrically that player $\Diamond$ does not have an optimal strategy either.

In the example in Figure 2, the game branches from state i to s_0 and s'_0 with probability $1/2$ respectively. However, the above argument can be adapted to work for probabilities c and $1-c$ for every constant $c \in (0,1)$. $\Box$

IV. Memory Requirements

In this section we study how much memory is needed to win objectives $(\mathcal{E}, \rhd c)$, depending on $\mathcal{E}$ and on the constraint $\rhd c$.

We say that an objective $(\mathcal{E}, \rhd c)$ is *strongly MD-determined* iff for every state s either

- there exists an MD-strategy σ such that, for all $\pi \in \Pi$, we have $\mathcal{P}_{\mathcal{G},s,\sigma,\pi}(\mathcal{E}) \rhd c$, or
- there exists an MD-strategy π such that, for all $\sigma \in \Sigma$, we have $\mathcal{P}_{\mathcal{G},s,\sigma,\pi}(\mathcal{E}) \not\rhd c$.

If a game is strongly MD-determined then it is also strongly determined, but not vice-versa. Strong FR-determinacy is defined analogously.

A. Reachability and Safety Objectives

Let $\mathcal{T} \subseteq S$ and $(\mathtt{Reach}(\mathcal{T}), \rhd c)$ be a threshold reachability objective. (Safety objectives are dual to reachability.)

Let us briefly discuss infinitely branching reachability games. If $c \in (0,1)$ then strong determinacy does not hold; cf. Figure 1 in [4]. Objectives $(\mathtt{Reach}(\mathcal{T}), \geq 1)$ are strongly determined (Theorem 2), but not strongly FR-determined, because player $\Diamond$ needs infinite memory (even if player $\Box$ is passive) [19]. Objectives $(\mathtt{Reach}(\mathcal{T}), > 0)$ correspond to non-stochastic 2-player reachability games, which are strongly MD-determined [26].

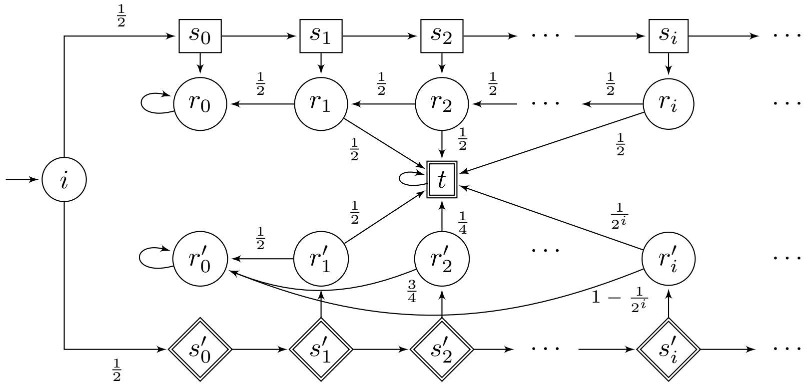

Fig. 2: A finitely branching game where the states of players $\Box$ and $\Diamond$ are drawn as squares and diamonds, respectively; random states $s \in S_{\bigcirc}$ are drawn as circles. The states s_i' and state t (double borders) are Büchi states, all other states are not. The value of the initial state i is $\frac{1}{2}$, for the Büchi objective $\mathcal{E}$. However, $i \notin [\mathcal{E}]_{\Box}^{\geq \frac{1}{2}} \uplus [\mathcal{E}]_{\Diamond}^{\not> \frac{1}{2}}$, meaning that neither player has a winning strategy, neither for the objective $(\mathcal{E}, \geq 1/2)$ nor for $(\mathcal{E}, > 1/2)$.

In the rest of this subsection we consider finitely branching reachability games. It is shown in [4], [18] that finitely branching reachability games are strongly determined, but the winning $\Box$ strategy constructed therein uses infinite memory. Indeed, Kučera [19] showed that infinite memory is necessary in general:

Theorem 4 (follows from Proposition 5.7.b in [19])**.** *Finitely branching reachability games with* $(\mathtt{Reach}(\mathcal{T}), \geq c)$ *objectives are not strongly FR-determined for* $c \in (0, 1)$.

The example from [19] that proves Theorem 4 has the following properties:

(1) player $\Box$ has *value-decreasing* (see below) transitions;
(2) player $\Diamond$ has *value-increasing* (see below) transitions;
(3) threshold $c \neq 0$ and $c \neq 1$;
(4) nonstrict inequality: $\geq c$.

Given a game $\mathcal{G}$, we call a transition $s \longrightarrow s'$ *value-decreasing (resp., value-increasing)* if $\mathtt{val}_{\mathcal{G}}(s) > \mathtt{val}_{\mathcal{G}}(s')$ (resp., $\mathtt{val}_{\mathcal{G}}(s) < \mathtt{val}_{\mathcal{G}}(s')$). If player $\Box$ (resp., player $\Diamond$) controls a transition $s \longrightarrow s'$, i.e., $s \in S_{\Box}$ (resp., $s \in S_{\Diamond}$), then the transition cannot be value-increasing (resp., value-decreasing). We write $RVI(\mathcal{G})$ for the game obtained from $\mathcal{G}$ by removing the value-increasing transitions controlled by player $\Diamond$. Note that this operation does not create any dead ends in finitely branching games, because at least one transition to a successor state with the same value will always remain for such games.

We show that a reachability game is strongly MD-determined if any of the properties listed above is not satisfied:

Theorem 5. *Finitely branching games* $\mathcal{G}$ *with reachability objectives* $(\mathtt{Reach}(\mathcal{T}), \rhd c)$ *are strongly MD-determined, provided that at least one of the following conditions holds.*

(1) player $\Box$ does not have value-decreasing transitions, or
(2) player $\Diamond$ does not have value-increasing transitions, or
(3) almost-sure objective: $\rhd = \geq$ and $c = 1$, or
(4) strict inequality: $\rhd = >$.

Remark 1. Condition (1) or (2) of Theorem 5 is trivially satisfied if the corresponding player is passive, i.e., in MDPs. It was already known that MD strategies are sufficient for safety and reachability objectives in countable finitely branching MDPs ([22], Section 7.2.7). Theorem 5 generalizes this result.

Remark 2. Theorem 5 does not carry over to stochastic reachability games with an arbitrary number of players, not even if the game graph is finite. Instead multiplayer games can require infinite memory to win. Proposition 4.13 in [24] constructs an 11-player finite-state stochastic reachability game with a pure subgame-perfect Nash equilibrium where the first player wins almost surely by using infinite memory. However, there is no finite-state Nash equilibrium (i.e., an equilibrium where all players are limited to finite memory) where the first player wins with positive probability. That is, the first player cannot win with only finite memory, not even if the other players are restricted to finite memory.

The rest of the subsection focuses on the proof of Theorem 5. We will need the following result from [4]:

Lemma 6. (Theorem 3.1 in [4]) *If* $\mathcal{G}$ *is a finitely branching reachability game then there is an MD strategy* $\pi \in \Pi$ *that is optimal minimizing in every* $\Diamond$ *state (i.e.,* $\mathtt{val}_{\mathcal{G}}(\pi(s)) = \mathtt{val}_{\mathcal{G}}(s)$*).*

One challenge in proving Theorem 5 is that an optimal minimizing player $\Diamond$ MD strategy according to Lemma 6 is not necessarily winning for player $\Diamond$, even for almost-sure reachability and even if player $\Diamond$ has a winning strategy. Indeed, consider the game in Figure 2, and add a new player $\Diamond$ state u and transitions $u \longrightarrow s_0$ and $u \longrightarrow t$. For the reachability objective $\mathtt{Reach}(\{t\})$, we then have $\mathtt{val}_{\mathcal{G}}(u) = \mathtt{val}_{\mathcal{G}}(s_0) = \mathtt{val}_{\mathcal{G}}(t) = 1$, and the player $\Diamond$ MD strategy π with $\pi(u) = t$ is optimal minimizing. However, $\Diamond$ is not winning from u w.r.t. the almost-sure objective $(\mathtt{Reach}(\{t\}), \geq 1)$. Instead the winning strategy is π' with $\pi'(u) = s_0$.

By the following lemma (from [4]), player $\Box$ has for every state an ϵ-optimal strategy that needs to be defined only on a finite horizon:

Lemma 7. (Lemma 3.2 in [4]) *If $\mathcal{G}$ is a finitely branching game with reachability objective* $\mathtt{Reach}(\mathcal{T})$ *then:*

$$\forall s \in S\ \forall \epsilon > 0\ \exists \sigma \in \Sigma\ \exists n \in \mathbb{N}\ \forall \pi \in \Pi\,. \\ \mathcal{P}_{\mathcal{G},s,\sigma,\pi}(\mathtt{Reach}_n(\mathcal{T})) > \mathtt{val}_{\mathcal{G}}(s) - \epsilon\,,$$

where $\mathtt{Reach}_n(\mathcal{T})$ *denotes the event of reaching $\mathcal{T}$ within at most n steps.*

Towards a proof of item (1) of Theorem 5, we prove the following lemma:

Lemma 8. *Let $\mathcal{G}$ be a finitely branching game with reachability objective* $\mathtt{Reach}(\mathcal{T})$. *Suppose that player $\Box$ does not have any value-decreasing transitions. Then there exists a player $\Box$ MD strategy $\hat{\sigma}$ that is optimal in all states. That is, for all states s and for all player $\Diamond$ strategies π we have* $\mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathtt{Reach}(\mathcal{T})) \geq \mathtt{val}_{\mathcal{G}}(s)$.

Proof. In order to construct the claimed MD strategy $\hat{\sigma}$, we define a sequence of modified games $\mathcal{G}_i$ in which the strategy of player $\Box$ is already fixed on a finite subset of the state space. We will show that the value of any state remains the same in all the $\mathcal{G}_i$, i.e., $\mathtt{val}_{\mathcal{G}_i}(s) = \mathtt{val}_{\mathcal{G}}(s)$ for all s. Fix an enumeration $s_1, s_2, \ldots$ that includes every state in S infinitely often. Let $\mathcal{G}_0 := \mathcal{G}$.

Given $\mathcal{G}_i$ we construct $\mathcal{G}_{i+1}$ as follows. We use Lemma 7 to get a strategy σ_i and $n_i \in \mathbb{N}$ s.t. $\mathcal{P}_{\mathcal{G}_i,s_i,\sigma_i,\pi}(\mathtt{Reach}_{n_i}(\mathcal{T})) > \mathtt{val}_{\mathcal{G}_i}(s_i) - 2^{-i}$. From the finiteness of n_i and the assumption that $\mathcal{G}$ is finitely branching, we obtain that $Env_i := \{s \mid s_i {\longrightarrow}^{\leq n_i} s\}$ is finite. Consider the subgame $\mathcal{G}'_i$ with finite state space Env_i. In this subgame there exists an optimal MD strategy σ'_i that maximizes the reachability probability for every state in Env_i. In particular, σ'_i achieves the same approximation in $\mathcal{G}'_i$ as σ_i in $\mathcal{G}_i$, i.e., $\mathcal{P}_{\mathcal{G}'_i,s_i,\sigma'_i,\pi}(\mathtt{Reach}(\mathcal{T})) > \mathtt{val}_{\mathcal{G}_i}(s_i) - 2^{-i}$. Let Env'_i be the subset of states s in Env_i with $\mathtt{val}_{\mathcal{G}'_i}(s) > 0$. Since Env'_i is finite, there exist $n'_i \in \mathbb{N}$ and $\lambda > 0$ with $\mathcal{P}_{\mathcal{G}'_i,s,\sigma'_i,\pi}(\mathtt{Reach}_{n'_i}(\mathcal{T})) \geq \lambda$ for all $s \in Env'_i$ and all $\pi \in \Pi_{\mathcal{G}'_i}$.

We now construct $\mathcal{G}_{i+1}$ by modifying $\mathcal{G}_i$ as follows. For every player $\Box$ state $s \in Env'_i$ we fix the transition according to σ'_i, i.e., only transition $s{\longrightarrow}\sigma'_i(s)$ remains and all other transitions from s are deleted. Since all moves from $\Box$ states in Env'_i have been fixed according to σ'_i, the bounds above for $\mathcal{G}'_i$ and σ'_i now hold for $\mathcal{G}_{i+1}$ and any $\sigma \in \Sigma_{\mathcal{G}_{i+1}}$. That is, we have $\mathcal{P}_{\mathcal{G}_{i+1},s_i,\sigma,\pi}(\mathtt{Reach}(\mathcal{T})) > \mathtt{val}_{\mathcal{G}_i}(s_i) - 2^{-i}$ and $\mathcal{P}_{\mathcal{G}_{i+1},s,\sigma,\pi}(\mathtt{Reach}_{n'_i}(\mathcal{T})) \geq \lambda$ for all $s \in Env'_i$ and all $\sigma \in \Sigma_{\mathcal{G}_{i+1}}$ and all $\pi \in \Pi_{\mathcal{G}_{i+1}}$.

Now we show that the values of all states s in $\mathcal{G}_{i+1}$ are still the same as in $\mathcal{G}_i$. Since our games are weakly determined, it suffices to show that player $\Box$ has an ϵ-optimal strategy from s in $\mathcal{G}_{i+1}$ for every $\epsilon > 0$. Let π be an arbitrary $\Diamond$ strategy from s in $\mathcal{G}_{i+1}$. Let s be a state and σ be an $\epsilon/2$-optimal $\Box$ strategy from s in $\mathcal{G}_i$. We now define a $\Box$ strategy σ' from s in $\mathcal{G}_{i+1}$. If the game does not enter Env'_i then σ' plays exactly as σ (which is possible since outside Env'_i no transitions have been removed). If the game enters Env'_i then it will reach the target from within Env'_i with probability $\geq \lambda$. Moreover, if the game stays inside Env'_i forever then it will almost surely reach the target, since $(1-\lambda)^\infty = 0$. Otherwise, it exits Env'_i at some state $s' \notin Env'_i$ (strictly speaking, at a distribution of such states). If this was the k-th visit to Env'_i then, from s', σ' plays an $\epsilon/2^{k+1}$-optimal strategy w.r.t. $\mathcal{G}_i$ (with the same modification as above if it visits Env'_i again). We can now bound the error of σ' from s as follows. The set of plays which visit Env'_i infinitely often contribute no error, since they almost surely reach the target by $(1-\lambda)^\infty = 0$. Since all transitions are at least value-preserving in $\mathcal{G}$ and hence in $\mathcal{G}_i$, the error of the plays which visit Env'_i at most j times is bounded by $\sum_{k=1}^{j} \epsilon/2^k$. Therefore, the error of σ' from s in $\mathcal{G}_{i+1}$ is bounded by ϵ and thus $\mathtt{val}_{\mathcal{G}_{i+1}}(s) = \mathtt{val}_{\mathcal{G}_i}(s)$.

Finally, we can construct the player $\Box$ MD winning strategy $\hat{\sigma}$ as the limit of the MD strategies σ'_i, which are all compatible with each other by the construction of the games $\mathcal{G}_i$. We obtain $\mathcal{P}_{\mathcal{G},s_i,\hat{\sigma},\pi}(\mathtt{Reach}(\mathcal{T})) > \mathtt{val}_{\mathcal{G}}(s_i) - 2^{-i}$ for all $i \in \mathbb{N}$. Let $s \in S$. Since $s = s_i$ holds for infinitely many i, we conclude Thus $\mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathtt{Reach}(\mathcal{T})) \geq \mathtt{val}_{\mathcal{G}}(s)$ as required. ☐

Towards a proof of items (2) and (3) of Theorem 5, we consider the operation $RVI(\mathcal{G})$, defined before the statement of Theorem 5. The following lemma shows that in reachability games all value-increasing transitions of player $\Diamond$ can be removed without changing the value of any state (although the outcome of the threshold reachability game may change in general).

Lemma 9. *Let $\mathcal{G}$ be a finitely branching reachability game and $\mathcal{G}' := RVI(\mathcal{G})$. Then for all $s \in S$ we have $\mathtt{val}_{\mathcal{G}'}(s) = \mathtt{val}_{\mathcal{G}}(s)$. Thus $RVI(\mathcal{G}') = \mathcal{G}'$.*

Proof. Since only $\Diamond$ transitions are removed, we trivially have $\mathtt{val}_{\mathcal{G}'}(s) \geq \mathtt{val}_{\mathcal{G}}(s)$. For the other inequality observe that the optimal minimizing strategy of Lemma 6 never takes any value-increasing transition and thus also guarantees the value in $\mathcal{G}'$. Thus also $\mathtt{val}_{\mathcal{G}'}(s) \leq \mathtt{val}_{\mathcal{G}}(s)$. ☐

Lemma 9 is in sharp contrast to Example 1 on page 4, which showed that the removal of value-*decreasing* transitions can change the value of states and can cause further transitions to become value-decreasing.

Similar to the proof of Theorem 2, the proof of the following lemma considers a transfinite sequence of subgames, where each subgame is obtained by removing the value-decreasing transitions from the previous subgames.

Lemma 10. *Let $\mathcal{G}$ be a finitely branching game with reachability objective* $\mathtt{Reach}(\mathcal{T})$. *Then there exist a player $\Box$ MD strategy $\hat{\sigma}$ and a player $\Diamond$ MD strategy $\hat{\pi}$ such that for all states $s \in S$, if $\mathcal{G} = RVI(\mathcal{G})$ or $\mathtt{val}_{\mathcal{G}}(s) = 1$, then the following is true:*

$$\forall \pi \in \Pi_{\mathcal{G}} : \mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathtt{Reach}(\mathcal{T})) \geq \mathtt{val}_{\mathcal{G}}(s) \quad \textit{or} \\ \forall \sigma \in \Sigma_{\mathcal{G}} : \mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Reach}(\mathcal{T})) < \mathtt{val}_{\mathcal{G}}(s).$$

Proof. We construct a transfinite sequence of subgames $\mathcal{G}_\alpha$, where $\alpha \in \mathbb{O}$ is an ordinal number, by stepwise removing certain transitions. Let $\longrightarrow_\alpha$ denote the set of transitions of the subgame $\mathcal{G}_\alpha$.

First, let $\mathcal{G}_0 := RVI(\mathcal{G})$. Since $\mathcal{G}$ is assumed to have no dead ends, it follows from the definition of RVI that $\mathcal{G}_0$ does not contain any dead ends either. In the following, we only remove transitions of player $\Box$. The resulting games $\mathcal{G}_\alpha$ with $\alpha > 0$ may contain dead ends, but these are always considered to be losing for player $\Box$. (Formally, one might add a dummy loop at these states.) For each $\alpha \in \mathbb{O}$ we define a set D_α as the set of transitions that are controlled by player $\Box$ and that are value-decreasing in $\mathcal{G}_\alpha$. For any $\alpha \in \mathbb{O} \setminus \{0\}$ we define $\longrightarrow_\alpha := \longrightarrow \setminus \bigcup_{\gamma<\alpha} D_\gamma$.

Since the sequence of sets $\longrightarrow_\alpha$ is non-increasing and we assumed that our game $\mathcal{G}$ has only countably many states and transitions, it follows that this sequence of games $\mathcal{G}_\alpha$ converges at some ordinal β where $\beta \le \omega_1$ (the first uncountable ordinal). I.e., we have $\mathcal{G}_\beta = \mathcal{G}_{\beta+1}$. In particular there are no value-decreasing player $\Box$ transitions in $\mathcal{G}_\beta$, i.e., $D_\beta = \emptyset$.

The removal of transitions of player $\Box$ can only decrease the value of states, and the operation RVI is value preserving by Lemma 9. Thus $\mathtt{val}_{\mathcal{G}_\beta}(s) \le \mathtt{val}_{\mathcal{G}_\alpha}(s) \le \mathtt{val}_{\mathcal{G}}(s)$ for all $\alpha \in \mathbb{O}$. We define the *index* of a state s by $I(s) := \min\{\alpha \in \mathbb{O} \mid \mathtt{val}_{\mathcal{G}_\alpha}(s) < \mathtt{val}_{\mathcal{G}}(s)\}$, and as $\bot$ if the set is empty.

Strategy $\hat{\sigma}$: Since $\mathcal{G}_\beta$ does not have value-decreasing transitions, we can invoke Lemma 8 to obtain a player $\Box$ MD strategy $\hat{\sigma}$ with $\mathcal{P}_{\mathcal{G}_\beta,s,\hat{\sigma},\pi}(\mathtt{Reach}(\mathcal{T})) \ge \mathtt{val}_{\mathcal{G}_\beta}(s) = \mathtt{val}_{\mathcal{G}}(s)$ for all π and for all s with $I(s) = \bot$. We show that, if $I(s) = \bot$ and either $\mathtt{val}_{\mathcal{G}}(s) = 1$ or $\mathcal{G} = RVI(\mathcal{G})$, then also in $\mathcal{G}$ we have $\mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathtt{Reach}(\mathcal{T})) \ge \mathtt{val}_{\mathcal{G}}(s)$. The only potential difference in the game on $\mathcal{G}$ is that π could take a $\Diamond$ transition, say $s' \longrightarrow s''$, that is present in $\mathcal{G}$ but not in $\mathcal{G}_\beta$. Since all $\Diamond$ transitions of $\mathcal{G}_0$ are kept in $\mathcal{G}_\beta$, such a transition would have been removed in the step $\mathcal{G}_0 := RVI(\mathcal{G})$. We show that this is impossible.

For the first case suppose that s satisfies $I(s) = \bot$ and $\mathtt{val}_{\mathcal{G}}(s) = 1$. It follows $\mathtt{val}_{\mathcal{G}_\beta}(s) = 1$. Since $\mathcal{G}_\beta$ does not have value-decreasing transitions, we have $\mathtt{val}_{\mathcal{G}_\beta}(s') = \mathtt{val}_{\mathcal{G}_\beta}(s'') = 1$, hence $\mathtt{val}_{\mathcal{G}}(s') = \mathtt{val}_{\mathcal{G}}(s'') = 1$, so the transition $s' \longrightarrow s''$ is not value-increasing in $\mathcal{G}$. Hence the transition is present in $\mathcal{G}_0$, hence also in $\mathcal{G}_\beta$.

For the second case suppose $\mathcal{G} = RVI(\mathcal{G})$. Since $\mathcal{G}$ does not contain any value-increasing transitions, the transition $s' \longrightarrow s''$ is not value-increasing in $\mathcal{G}$. So it is present in $\mathcal{G}_0$, and thus also in $\mathcal{G}_\beta$.

It follows that under $\hat{\sigma}$ the play remains in the states of $\mathcal{G}_\beta$ and only uses transitions that are present in $\mathcal{G}_\beta$, regardless of the strategy π. In this sense, all plays under $\hat{\sigma}$ on $\mathcal{G}$ coincide with plays on $\mathcal{G}_\beta$. Hence $\mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathtt{Reach}(\mathcal{T})) = \mathcal{P}_{\mathcal{G}_\beta,s,\hat{\sigma},\pi}(\mathtt{Reach}(\mathcal{T})) \ge \mathtt{val}_{\mathcal{G}}(s)$.

Strategy $\hat{\pi}$: It now suffices to define a player $\Diamond$ MD strategy $\hat{\pi}$ so that we have $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Reach}(\mathcal{T})) < \mathtt{val}_{\mathcal{G}}(s)$ for all σ and for all s with $I(s) \in \mathbb{O}$. This strategy $\hat{\pi}$ is defined as follows.

- If $I(s) = \alpha$ then $\hat{\pi}(s) = s'$ where s' is an arbitrary but fixed successor of s where transition $s \longrightarrow s'$ is present in $\mathcal{G}_\alpha$ and $\mathtt{val}_{\mathcal{G}_\alpha}(s) = \mathtt{val}_{\mathcal{G}_\alpha}(s')$ and $I(s') = I(s) = \alpha$. This exists by the assumption that $\mathcal{G}$ is finitely branching and the definition of $\mathcal{G}_\alpha$. In particular, since the transition $s \longrightarrow s'$ is present in $\mathcal{G}_\alpha$, it is not value-increasing in the game $\mathcal{G}$; otherwise it would have been removed in the step from $\mathcal{G}$ to $\mathcal{G}_0$.
- If $I(s) = \bot$, $\hat{\pi}$ plays the optimal minimizing MD strategy on $\mathcal{G}$ from Lemma 6, i.e., we have $\hat{\pi}(s) = s'$ where s' is an arbitrary but fixed successor of s in $\mathcal{G}$ with $\mathtt{val}_{\mathcal{G}}(s) = \mathtt{val}_{\mathcal{G}}(s')$.

Considering both cases, it follows that strategy $\hat{\pi}$ is optimal minimizing in $\mathcal{G}$.

Let s_0 be an arbitrary state with $I(s_0) \in \mathbb{O}$. To show that $\mathcal{P}_{\mathcal{G},s_0,\sigma,\hat{\pi}}(\mathtt{Reach}(\mathcal{T})) < \mathtt{val}_{\mathcal{G}}(s_0)$ holds for all σ, let σ be any strategy of player $\Box$. Let $\alpha \neq \bot$ be the smallest index among the states that can be reached with positive probability from s_0 under the strategies $\sigma, \hat{\pi}$. Let s_1 be such a state with index α. In the following we write σ also for the strategy σ after a partial play leading from s_0 to s_1 has been played.

Suppose that the play from s_1 under the strategies $\sigma, \hat{\pi}$ always remains in $\mathcal{G}_\alpha$. Strategy $\hat{\pi}$ might not be optimal minimizing in $\mathcal{G}_\alpha$ in general. However, we show that it is optimal minimizing in $\mathcal{G}_\alpha$ from all states with index $\ge \alpha$. Let s be a $\Diamond$ state with index $I(s) = \alpha' \ge \alpha$. By definition of $\hat{\pi}$ we have $\hat{\pi}(s) = s'$ where the transition $s \longrightarrow s'$ is present in $\mathcal{G}_{\alpha'}$ with $\mathtt{val}_{\mathcal{G}_{\alpha'}}(s) = \mathtt{val}_{\mathcal{G}_{\alpha'}}(s')$ and $I(s') = I(s) = \alpha'$. In the case where $\alpha' = \alpha$ this directly implies that the step $s \longrightarrow s'$ is optimal minimizing in $\mathcal{G}_\alpha$. The remaining case is that $\alpha' > \alpha$. Here, by definition of the index, $\mathtt{val}_{\mathcal{G}}(s) = \mathtt{val}_{\mathcal{G}_\alpha}(s)$ and $\mathtt{val}_{\mathcal{G}}(s') = \mathtt{val}_{\mathcal{G}_\alpha}(s')$. Since the transition $s \longrightarrow s'$ is present in $\mathcal{G}_{\alpha'}$, it is also present in $\mathcal{G}_0$ and $\mathcal{G}_\alpha$. Since $\mathcal{G}_0 = RVI(\mathcal{G})$, this transition is not value-increasing in $\mathcal{G}$. Also, it is not value-decreasing in $\mathcal{G}$, because it is a $\Diamond$ transition. Therefore $\mathtt{val}_{\mathcal{G}}(s) = \mathtt{val}_{\mathcal{G}}(s')$, and thus $\mathtt{val}_{\mathcal{G}_\alpha}(s) = \mathtt{val}_{\mathcal{G}_\alpha}(s')$. Also in this case the step $s \longrightarrow s'$ is optimal minimizing in $\mathcal{G}_\alpha$.

So the only possible exceptions where strategy $\hat{\pi}$ might not be optimal minimizing in $\mathcal{G}_\alpha$ are states with index $< \alpha$. Since we have assumed above that such states cannot be reached under $\sigma, \hat{\pi}$, it follows that $\mathcal{P}_{\mathcal{G},s_1,\sigma,\hat{\pi}}(\mathtt{Reach}(\mathcal{T})) \le \mathtt{val}_{\mathcal{G}_\alpha}(s_1) < \mathtt{val}_{\mathcal{G}}(s_1)$.

Now suppose that the play from s_1 under $\sigma, \hat{\pi}$, with positive probability, takes a transition, say $s_2 \longrightarrow s_3$, that is not present in $\mathcal{G}_\alpha$. Then this transition was value-decreasing for some game $\mathcal{G}_{\alpha'}$ with $\alpha' < \alpha$: that is, $\mathtt{val}_{\mathcal{G}_{\alpha'}}(s_2) > \mathtt{val}_{\mathcal{G}_{\alpha'}}(s_3)$. Since the indices of both s_2 and s_3 are $\ge \alpha > \alpha'$, we have $\mathtt{val}_{\mathcal{G}}(s_2) = \mathtt{val}_{\mathcal{G}_{\alpha'}}(s_2) > \mathtt{val}_{\mathcal{G}_{\alpha'}}(s_3) = \mathtt{val}_{\mathcal{G}}(s_3)$. Hence the transition $s_2 \longrightarrow s_3$ is value-decreasing in $\mathcal{G}$. Since $\hat{\pi}$ is optimal minimizing in $\mathcal{G}$, we also have $\mathcal{P}_{\mathcal{G},s_1,\sigma,\hat{\pi}}(\mathtt{Reach}(\mathcal{T})) < \mathtt{val}_{\mathcal{G}}(s_1)$.

Since $\hat{\pi}$ is optimal minimizing in $\mathcal{G}$, we conclude that we have $\mathcal{P}_{\mathcal{G},s_0,\sigma,\hat{\pi}}(\mathtt{Reach}(\mathcal{T})) < \mathtt{val}_{\mathcal{G}}(s_0)$. $\Box$

We are now ready to prove Theorem 5.

Proof of Theorem 5. Let $\mathcal{G}$ be a finitely branching game with reachability objective $(\mathtt{Reach}(\mathcal{T}), \rhd c)$. Let $s_0 \in S$ be an arbitrary initial state.

Suppose $\mathtt{val}_{\mathcal{G}}(s_0) < c$. Then player $\Diamond$ wins with the MD strategy from Lemma 6.

Suppose $\mathtt{val}_{\mathcal{G}}(s_0) > c$. Let $\delta := \mathtt{val}_{\mathcal{G}}(s_0) - c > 0$. By Lemma 7 there are a strategy $\sigma \in \Sigma$ and $n \in \mathbb{N}$ such that $\mathcal{P}_{\mathcal{G},s_0,\sigma,\pi}(\mathtt{Reach}_n(\mathcal{T})) > \mathtt{val}_{\mathcal{G}}(s_0) - \frac{\delta}{2} > c$ holds for all $\pi \in \Pi$. The strategy σ plays on the subgame $\mathcal{G}'$ with state space $S' = \{s' \in S \mid s \longrightarrow^{\leq n} s'\}$, which is finite since $\mathcal{G}$ is finitely branching. Therefore, there exists an MD strategy σ' with $\mathcal{P}_{\mathcal{G}',s_0,\sigma',\pi}(\mathtt{Reach}(\mathcal{T})) \geq \mathcal{P}_{\mathcal{G},s_0,\sigma,\pi}(\mathtt{Reach}_n(\mathcal{T}))$. Since $S' \subseteq S$, the strategy σ' also applies in $\mathcal{G}$, hence $\mathcal{P}_{\mathcal{G},s_0,\sigma',\pi}(\mathtt{Reach}(\mathcal{T})) \geq \mathcal{P}_{\mathcal{G}',s_0,\sigma',\pi}(\mathtt{Reach}(\mathcal{T}))$. By combining the mentioned inequalities we obtain that $\mathcal{P}_{\mathcal{G},s_0,\sigma',\pi}(\mathtt{Reach}(\mathcal{T})) > c$ holds for all $\pi \in \Pi$. So the MD strategy σ' is winning for player $\Box$.

It remains to consider the case $\mathtt{val}_{\mathcal{G}}(s_0) = c$. Let us discuss the four cases from the statement of Theorem 5 individually.

(4) If $\rhd = >$ then player $\Diamond$ wins with the MD strategy from Lemma 6.

So for the remaining cases it suffices to consider the threshold objective $(\mathtt{Reach}(\mathcal{T}), \geq \mathtt{val}_{\mathcal{G}}(s_0))$.

(1) If player $\Box$ does not have value-decreasing transitions then player $\Box$ wins with the MD strategy from Lemma 8.
(2) If player $\Diamond$ does not have value-increasing transitions then Lemma 10 supplies either player $\Box$ or player $\Diamond$ with an MD winning strategy.
(3) If $c = \mathtt{val}_{\mathcal{G}}(s_0) = 1$ then, again, Lemma 10 supplies either player $\Box$ or player $\Diamond$ with an MD winning strategy.

This completes the proof of Theorem 5. $\Box$

B. Büchi and co-Büchi Objectives

Let $\mathcal{E}$ be the Büchi objective. (The co-Büchi objective is dual.) Quantitative Büchi objectives $(\mathcal{E}, \rhd c)$ with $c \in (0,1)$ are not strongly determined, not even for finitely branching games (Theorem 3), but positive probability $(\mathcal{E}, >0)$ and almost-sure $(\mathcal{E}, \geq 1)$ Büchi objectives are strongly determined (Theorem 2).

However, $(\mathcal{E}, >0)$ objectives are not strongly FR-determined, even in finitely branching systems. Even in the special case of finitely branching MDPs (where player $\Diamond$ is passive and the game is trivially strongly determined), player $\Box$ may require infinite memory to win [18].

In infinitely branching games, the almost-sure Büchi objective $(\mathcal{E}, \geq 1)$ is not strongly FR-determined, because it subsumes the almost-sure reachability objective; cf. Subsection IV-A.

In contrast, in finitely branching games, the almost-sure Büchi objective $(\mathcal{E}, \geq 1)$ is strongly MD-determined, as the following theorem shows:

Theorem 11. *Let $\mathcal{G}$ be a finitely branching game with objective* $\mathtt{Büchi}(\mathcal{T})$. *Then there exist a player $\Box$ MD strategy $\hat{\sigma}$ and a player $\Diamond$ MD strategy $\hat{\pi}$ such that for all states $s \in S$:*

$$\forall \pi \in \Pi_{\mathcal{G}} : \mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathtt{Büchi}(\mathcal{T})) = 1 \quad \textit{or}$$
$$\forall \sigma \in \Sigma_{\mathcal{G}} : \mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Büchi}(\mathcal{T})) < 1.$$

Hence finitely branching almost-sure Büchi games are strongly MD-determined.

For the proof we need the following lemmas, which are variants of Lemmas 6 and 8 for the objective $\mathtt{Reach}^+(\mathcal{T})$, which is defined as:

$$\mathtt{Reach}^+(\mathcal{T}) := \{s_0 s_1 \cdots \in S^\omega \mid \exists\, i \geq 1.\, s_i \in \mathcal{T}\}$$

The difference to $\mathtt{Reach}(\mathcal{T})$ is that $\mathtt{Reach}^+(\mathcal{T})$ requires a path to $\mathcal{T}$ that involves at least one transition.

Lemma 12. *Let $\mathcal{G}$ be a finitely branching game with objective* $\mathtt{Reach}^+(\mathcal{T})$. *Then there is an MD strategy $\pi \in \Pi$ that is optimal minimizing in every state.*

Proof. Outside $\mathcal{T}$, the objectives $\mathtt{Reach}(\mathcal{T})$ and $\mathtt{Reach}^+(\mathcal{T})$ coincide, so outside $\mathcal{T}$, the MD strategy π from Lemma 6 is optimal minimizing for $\mathtt{Reach}^+(\mathcal{T})$. Any $s \in \mathcal{T} \cap S_\Diamond$ with $\mathtt{val}_{\mathcal{G}}(s) < 1$ must have a transition $s \longrightarrow s'$ with $s' \notin \mathcal{T}$ and $\mathtt{val}_{\mathcal{G}}(s) = \mathtt{val}_{\mathcal{G}}(s')$, where the value is always meant with respect to $\mathtt{Reach}^+(\mathcal{T})$. Set $\pi(s) := s'$. Then π is optimal minimizing in every state, as desired. $\Box$

Lemma 13. *Let $\mathcal{G}$ be a finitely branching game with objective* $\mathtt{Reach}^+(\mathcal{T})$. *Suppose player $\Box$ does not have value-decreasing transitions. Then there is an MD strategy $\sigma \in \Sigma$ that is optimal maximizing in every state.*

Proof. Outside $\mathcal{T}$, the objectives $\mathtt{Reach}(\mathcal{T})$ and $\mathtt{Reach}^+(\mathcal{T})$ coincide, so outside $\mathcal{T}$, the MD strategy σ from Lemma 8 is optimal maximizing for $\mathtt{Reach}^+(\mathcal{T})$. Any $s \in \mathcal{T} \cap S_\Box$ must have a transition $s \longrightarrow s'$ with $s' \in \mathcal{T}$ or $\mathtt{val}_{\mathcal{G}}(s) = \mathtt{val}_{\mathcal{G}}(s')$, where the value is always meant with respect to $\mathtt{Reach}^+(\mathcal{T})$. Set $\sigma(s) := s'$. Then σ is optimal maximizing in every state, as desired. $\Box$

With this at hand, we prove Theorem 11.

Proof of Theorem 11. We proceed similarly to the proof of Theorem 2. In the present proof, whenever we write $\mathtt{val}_{\mathcal{G}'}(s)$ for a subgame $\mathcal{G}'$ of $\mathcal{G}$, we mean the value of state s with respect to $\mathtt{Reach}^+(\mathcal{T} \cap S')$, where $S' \subseteq S$ is the state space of $\mathcal{G}'$.

In order to characterize the winning sets of the players with respect to the objective $\mathtt{Büchi}(\mathcal{T})$, we construct a transfinite sequence of subgames $\mathcal{G}_\alpha$ of $\mathcal{G}$, where $\alpha \in \mathbb{O}$ is an ordinal number, by stepwise removing certain states, along with their incoming transitions. Let S_α denote the state space of the subgame $\mathcal{G}_\alpha$. We start with $\mathcal{G}_0 := \mathcal{G}$. Given $\mathcal{G}_\alpha$, define D^0_α as the set of states $s \in S_\alpha$ with $\mathtt{val}_{\mathcal{G}_\alpha}(s) < 1$, and for any $i \geq 0$ define D^{i+1}_α as the set of states $s \in \left(S_\alpha \setminus \bigcup_{j=0}^{i} D^j_\alpha\right) \cap (S_\Diamond \cup S_\bigcirc)$ that have a transition $s \longrightarrow s'$ with $s' \in D^i_\alpha$. The set $\bigcup_{i \in \mathbb{N}} D^i_\alpha$ can be seen as the backward closure of D^0_α under random transitions and transitions controlled by player $\Diamond$. For any $\alpha \in \mathbb{O} \setminus \{0\}$ we define $S_\alpha := S \setminus \bigcup_{\gamma < \alpha} \bigcup_{i \in \mathbb{N}} D^i_\gamma$.

Since the number of states never increases and S is countable, it follows that this sequence of games $\mathcal{G}_\alpha$ converges at some ordinal β where $\beta \leq \omega_1$ (the first uncountable ordinal). That is, we have $\mathcal{G}_\beta = \mathcal{G}_{\beta+1}$.

As in the proof of Theorem 2, some games $\mathcal{G}_\alpha$ may contain dead ends, which are always considered to be losing for player $\Box$. However, $\mathcal{G}_\beta$ does not contain dead ends. (If S_β is empty then player $\Box$ loses.) We define the *index*, $I(s)$, of a state s as the ordinal α with $s \in \bigcup_{i\in\mathbb{N}} D_\alpha^i$, and as $\bot$ if such an ordinal does not exist. For all states $s \in S$ we have:

$$I(s) = \bot \;\Leftrightarrow\; s \in S_\beta \;\Leftrightarrow\; \mathtt{val}_{\mathcal{G}_\beta}(s) = 1$$

In particular, player $\Box$ does not have value-decreasing transitions in $\mathcal{G}_\beta$. We show that states s with $I(s) \in \mathbb{O}$ are in $\left[\mathtt{Büchi}(\mathcal{T})\right]^{<1}_{\Diamond\,\mathcal{G}}$, and states s with $I(s) = \bot$ are in $\left[\mathtt{Büchi}(\mathcal{T})\right]^{=1}_{\Box\,\mathcal{G}}$, and in each case we give the claimed witnessing MD strategy.

Strategy $\hat{\pi}$: We define the claimed MD strategy $\hat{\pi}$ for all $s \in S_\Diamond$ with $I(s) = \alpha \in \mathbb{O}$ as follows. For all $s \in D_\alpha^0$, define $\hat{\pi}(s)$ as in the MD strategy from Lemma 12 for $\mathcal{G}_\alpha$ and $\mathtt{Reach}^+(\mathcal{T} \cap S_\alpha)$. For all $s \in D_\alpha^{i+1} \cap S_\Diamond$ for some $i \in \mathbb{N}$, define $\hat{\pi}(s) := s'$ such that $s \longrightarrow s'$ and $s' \in D_\alpha^i$.

In each $\mathcal{G}_\alpha$, strategy $\hat{\pi}$ coincides with the strategy from Lemma 12, except possibly in states $s \in S_\alpha$ with $\mathtt{val}_{\mathcal{G}_\alpha}(s) = 1$. It follows that $\hat{\pi}$ is optimal minimizing for all $\mathcal{G}_\alpha$ with $\alpha \in \mathbb{O}$.

We show by transfinite induction on the index that $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Büchi}(\mathcal{T})) < 1$ holds for all states $s \in S$ with $I(s) \in \mathbb{O}$ and for all player $\Box$ strategies σ. For the induction hypothesis, let α be an ordinal for which this holds for all states s with $I(s) < \alpha$. For the inductive step, let $s \in S$ be a state with $I(s) = \alpha$, and let σ be an arbitrary player $\Box$ strategy in $\mathcal{G}$.

- Let $s \in D_\alpha^0$. Suppose that the play from s under the strategies $\sigma, \hat{\pi}$ always remains in S_α, i.e., the probability of ever leaving S_α under $\sigma, \hat{\pi}$ is zero. Then any play in $\mathcal{G}$ under these strategies coincides with a play in $\mathcal{G}_\alpha$, so we have $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Reach}^+(\mathcal{T})) = \mathcal{P}_{\mathcal{G}_\alpha,s,\sigma,\hat{\pi}}(\mathtt{Reach}^+(\mathcal{T} \cap S_\alpha))$. Since $\hat{\pi}$ is optimal minimizing in $\mathcal{G}_\alpha$, we have $\mathcal{P}_{\mathcal{G}_\alpha,s,\sigma,\hat{\pi}}(\mathtt{Reach}^+(\mathcal{T} \cap S_\alpha)) \le \mathtt{val}_{\mathcal{G}_\alpha}(s) < 1$. Since $\mathtt{Büchi}(\mathcal{T}) \subseteq \mathtt{Reach}^+(\mathcal{T})$, we have $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Büchi}(\mathcal{T})) \le \mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Reach}^+(\mathcal{T}))$. By combining the mentioned equalities and inequalities we get $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Büchi}(\mathcal{T})) < 1$, as desired.
  Now suppose otherwise, i.e., the play from s under $\sigma, \hat{\pi}$, with positive probability, enters a state $s' \notin S_\alpha$, hence $I(s') < \alpha$. By the induction hypothesis we have $\mathcal{P}_{\mathcal{G},s',\sigma',\hat{\pi}}(\mathtt{Büchi}(\mathcal{T})) < 1$ for any σ'. Since the probability of entering s' is positive, we conclude $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Büchi}(\mathcal{T})) < 1$, as desired.
- Let $s \in D_\alpha^i$ for some $i \ge 1$. It follows from the definitions of D_α^i and of $\hat{\pi}$ that $\hat{\pi}$ induces a partial play of length $i+1$ from s to a state $s' \in D_\alpha^0$ (player $\Box$ does not play on this partial play). We have shown above that $\mathcal{P}_{\mathcal{G},s',\sigma,\hat{\pi}}(\mathtt{Büchi}(\mathcal{T})) < 1$. It follows that $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Büchi}(\mathcal{T})) < 1$, as desired.

We conclude that we have $\mathcal{P}_{\mathcal{G},s,\sigma,\hat{\pi}}(\mathtt{Büchi}(\mathcal{T})) < 1$ for all σ and all $s \in S$ with $I(s) \in \mathbb{O}$.

Strategy $\hat{\sigma}$: We define the claimed MD strategy $\hat{\sigma}$ for all $s \in S_\Box$ with $I(s) = \bot$ to be the MD strategy from Lemma 13 for $\mathcal{G}_\beta$ and $\mathtt{Reach}^+(\mathcal{T} \cap S_\beta)$. This definition ensures that player $\Box$ never takes a transition in $\mathcal{G}$ that leaves S_β. Random transitions and player $\Diamond$ transitions in $\mathcal{G}$ never leave S_β either: indeed, if $s' \in S$ with $I(s') = \alpha \in \mathbb{O}$ then $s' \in D_\alpha^i$ for some i, hence if $s \in S_\Diamond \cup S_\bigcirc$ and $s \longrightarrow s'$ then $I(s) \le \alpha$. We conclude that starting from S_β all plays in $\mathcal{G}$ remain in S_β, under $\hat{\sigma}$ and all player $\Diamond$ strategies.

Let $s \in S_\beta$, hence $\mathtt{val}_{\mathcal{G}_\beta}(s) = 1$. Let π be any player $\Diamond$ strategy. Since $\hat{\sigma}$ is optimal maximizing in $\mathcal{G}_\beta$, we have $\mathcal{P}_{\mathcal{G}_\beta,s,\hat{\sigma},\pi}(\mathtt{Reach}^+(\mathcal{T} \cap S_\beta)) = 1$. As argued above, S_β is not left even in $\mathcal{G}$, hence $\mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathtt{Reach}^+(\mathcal{T} \cap S_\beta)) = 1$.

Therefore $\mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathtt{Reach}^+(\mathcal{T} \cap S_\beta)) = 1$ holds for all $s \in S_\beta$ and all π. Since Büchi is repeated reachability, we also have $\mathcal{P}_{\mathcal{G},s,\hat{\sigma},\pi}(\mathtt{Büchi}(\mathcal{T})) = 1$ for all π and all $s \in S$ with $I(s) = \bot$. ☐

V. Conclusions and Open Problems

With the results of this paper at hand, let us review the landscape of strong determinacy for stochastic games. We have shown that almost-sure objectives are strongly determined (Theorem 2), even in the infinitely branching case.

Let us review the finitely branching case. Quantitative reachability games are strongly determined [18], [4], [5]. They are generally not strongly FR-determined [19], but they are strongly MD-determined under any of the conditions provided by Theorem 5. Almost-sure reachability games and even almost-sure Büchi games are strongly MD-determined (Theorems 5 and 11). Almost-sure co-Büchi games are generally not strongly FR-determined [18], even if player $\Box$ is passive, because player $\Diamond$ may need infinite memory to win. However, the following question is open: if a state is almost-surely winning for player $\Box$ in a co-Büchi game, does player $\Box$ also have a winning MD strategy?

The same question is open for infinitely branching almost-sure reachability games (these games are generally not strongly FR-determined either [19]). In fact, one can show that a positive answer to the former question implies a positive answer to the latter question.

Acknowledgements. This work was partially supported by the EPSRC through grants EP/M027287/1, EP/M027651/1, EP/P020909/1 and EP/M003795/1 and by St. John's College, Oxford.

References

[1] P. A. Abdulla, L. Clemente, R. Mayr, and S. Sandberg. Stochastic parity games on lossy channel systems. *Logical Methods in Computer Science*, 10(4:21), 2014.

[2] P. Billingsley. *Probability and Measure*. Wiley, New York, NY, 1995. Third Edition.

[3] T. Brázdil, V. Brožek, K. Etessami, A. Kučera, and D. Wojtczak. One-counter Markov decision processes. In *SODA'10*, pages 863–874. SIAM, 2010.

[4] T. Brázdil, V. Brožek, A. Kučera, and J. Obdrzálek. Qualitative reachability in stochastic BPA games. *Information and Computation*, 209, 2011.

[5] V. Brožek. Determinacy and optimal strategies in infinite-state stochastic reachability games. *TCS*, 493, 2013.

[6] K. Chatterjee, L. de Alfaro, and T. Henzinger. Strategy improvement for concurrent reachability games. In *QEST*, pages 291–300. IEEE Computer Society Press, 2006.

[7] K. Chatterjee, M. Jurdziński, and T. Henzinger. Simple stochastic parity games. In *CSL'03*, volume 2803 of *LNCS*, pages 100–113. Springer, 2003.

[8] K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Quantitative stochastic parity games. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '04, pages 121–130, Philadelphia, PA, USA, 2004. Society for Industrial and Applied Mathematics.

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

[10] L. de Alfaro and T. Henzinger. Concurrent omega-regular games. In *Proc. of LICS*, pages 141–156. IEEE, June 2000.

[11] L. de Alfaro, T. Henzinger, and O. Kupferman. Concurrent reachability games. In *FOCS*, pages 564–575. IEEE Computer Society Press, 1998.

[12] K. Etessami, D. Wojtczak, and M. Yannakakis. Recursive stochastic games with positive rewards. In *ICALP*, volume 5125 of *LNCS*. Springer, 2008.

[13] K. Etessami and M. Yannakakis. Recursive Markov decision processes and recursive stochastic games. In *ICALP'05*, volume 3580 of *LNCS*, pages 891–903. Springer, 2005.

[14] K. Etessami and M. Yannakakis. Recursive concurrent stochastic games. *LMCS*, 4, 2008.

[15] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley & Sons, second edition, 1966.

[16] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, 1997.

[17] H. Gimbert and F. Horn. Solving simple stochastic tail games. In *Proceedings of SODA*, pages 847–862. SIAM, 2010.

[18] J. Krčál. Determinacy and Optimal Strategies in Stochastic Games. Master's thesis, Masaryk University, School of Informatics, Brno, Czech Republic, 2009.

[19] A. Kučera. Turn-based stochastic games. In K. R. Apt and E. Grädel, editors, *Lectures in Game Theory for Computer Scientists*. Cambridge University Press, 2011.

[20] A. Maitra and W. Sudderth. Finitely additive stochastic games with Borel-measurable payoffs. *International Journal of Game Theory*, 27:257–267, 1998.

[21] D. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63:1565–1581, 1998.

[22] M. L. Puterman. *Markov Decision Processes*. Wiley, 1994.

[23] L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953.

[24] M. Ummels and D. Wojtczak. The complexity of Nash equilibria in stochastic multiplayer games. *LMCS*, 7(3:20), 2011.

[25] D. Williams. *Probability with Martingales*. Cambridge University Press, 1991.

[26] W. Zielonka. Infinite games on finitely coloured graphs with applications to automata on infinite trees. *TCS*, 200(1-2):135–183, 1998.