Optimal Strategies in Infinite-state Stochastic Reachability Games

Václav Brožek *
LFCS, School of Informatics
University of Edinburgh
Edinburgh, Scotland, UK
Vaclav.Brozek@ed.ac.uk

We consider perfect-information reachability stochastic games for 2 players on infinite graphs. We identify a subclass of such games, and prove two interesting properties of it: first, Player Max always has optimal strategies in games from this subclass, and second, these games are strongly determined. The subclass is defined by the property that the set of all values can only have one accumulation point – 0. Our results nicely mirror recent results for finitely-branching games, where, on the contrary, Player Min always has optimal strategies. However, our proof methods are substantially different, because the roles of the players are not symmetric. We also do not restrict the branching of the games. Finally, we apply our results in the context of recently studied One-Counter stochastic games.

1 Introduction

Two-player turn-based zero-sum stochastic games, simply called "games" in this text, evolve randomly in discrete *transitions* from one of countably many *states* to another. The winning condition is some property of such infinite evolutions. Each state is either owned by Player Max, Player Min, or it is stochastic, and has a fixed set, possibly infinite, of available outgoing transitions. The states and transitions define a *game graph*, an infinite path in this graph is called a *run*. The set of runs comes with a product topology over the discrete state space, i.e., open sets are generated by sets of runs sharing a common finite prefix. In stochastic states, the successor is sampled according to a fixed distribution, whereas players choose successors in states they own, based on the history of the play so far. This induces a probabilistic measure for Borel-measurable sets of runs in a natural way.

A winning condition is a set W of runs. A run from W is won by Player Max, the other runs are won by Player Min (the games are zero-sum). For Borel measurable sets W, a fixed pair (σ,π) of strategies for Player Max and Min, respectively, and an initial state, s, the probability that Max wins is denoted by $\mathbb{P}_s^{\sigma,\pi}[W]$. The *value* of the game in s, denoted by $Val(s)$, is defined as

$$Val(s) := \sup_{\sigma}\inf_{\pi}\mathbb{P}_s^{\sigma,\pi}[W] = \inf_{\pi}\sup_{\sigma}\mathbb{P}_s^{\sigma,\pi}[W]\,. \tag{1}$$

The above equality, a consequence of a more general, Blackwell-determinacy result of Martin [12], implies that for every $\varepsilon > 0$ both of the players have so called ε-*optimal* strategies, σ_ε and π_ε, such that $\inf_\pi \mathbb{P}_s^{\sigma_\varepsilon,\pi}[W] \geq Val(s) - \varepsilon$, and $\sup_\sigma \mathbb{P}_s^{\sigma,\pi_\varepsilon}[W] \leq Val(s) + \varepsilon$. This may not be true for the case when $\varepsilon = 0$, where the optimal (i.e., 0-optimal) strategies may not exist for neither of the players.

We consider a stronger notion of determinacy than (1), and call a game *strongly determined* if for every state s, every ν, $0 \leq \nu \leq 1$, and $\rhd \in \{>,\geq\}$ either Player Max has a strategy $\bar{\sigma}$ such that $\forall\pi$:

*Author supported by the Newton International Fellowship of Royal Society.

Giovanna D'Agostino, Salvatore La Torre (Eds.):
Proceedings of the Second International Symposium on
"Games, Automata, Logics and Formal Verification" (GandALF 2011)
EPTCS 54, 2011, pp. 60–73, doi:10.4204/EPTCS.54.5

© V. Brožek
This work is licensed under the Creative Commons
Attribution-Noncommercial-Share Alike License.

$\mathbb{P}_s^{\bar{\sigma},\pi}[W] \rhd v$, or Player Min has a strategy $\bar{\pi}$ such that $\forall \sigma : \mathbb{P}_s^{\sigma,\bar{\pi}}[W] \not\rhd v$. Denote $L := \sup_\sigma \inf_\pi \mathbb{P}_s^{\sigma,\pi}[W]$ and $R := \inf_\pi \sup_\sigma \mathbb{P}_s^{\sigma,\pi}[W]$, then if Max has a strategy $\bar{\sigma}$ such that $\forall \pi : \mathbb{P}_s^{\bar{\sigma},\pi}[W] \geq v$ then $v \leq L$. Similarly, if Player Min has a strategy $\bar{\pi}$ such that $\forall \sigma : \mathbb{P}_s^{\sigma,\bar{\pi}}[W] \leq v$ then $v \geq R$. By strong determinacy, $\forall v : \neg(R > v > L)$, thus $R \leq L$. $L \leq R$ follows from definitions, thus strong determinacy implies determinacy. On the other hand, it is easy to see that the existence of ε-optimal strategies for both players implies strong determinacy for cases where $|v - Val(s)| \geq 2\varepsilon$, the players simply use their ε-optimal strategies to win. This works even for $\varepsilon = 0$, thus whenever both players have optimal strategies, the game is strongly determined (for all v). To sum up the relation between the key three notions: Every game with a Borel winning condition is determined in the sense of (1), some of these games are strongly determined, and some of the strongly determined games are those admitting optimal strategies for both players. Example 1 and [7, Fig. 1] show that both the inclusions are proper. More precisely, in the game from Example 1, which we show later, Player Min has only one (trivial) strategy, thus the game is strongly determined. However, there is a state r_0, such that for every fixed strategy of Max the probability of winning is strictly below $Val(r_0)$. The game from [7, Fig. 1], is composed of two halves, one of which is essentially equivalent to the game in Example 1, and the other is a similar game adopted for Min (infinite branching needed). As a consequence, neither Player Max in the first half, nor Min in the second half have optimal strategies. Thus, fixing a strategy of one player first, which is ε-optimal, the other player may choose an $\varepsilon/2$-optimal strategy to beat the first player. As a consequence, no player has a winning strategy.

We are especially interested in the situation when W is an open set, and call such games *open* as well. This includes all *reachability* conditions, where W is the set of all runs visiting a state from a distinguished set of target states, T. For reachability, results of [7, 6] imply (see Corollary 1) that Player Min has always optimal strategies if every state, s, owned by Min has at least one successor, t, such that $Val(s) = Val(t)$. This is always the case in finitely-branching games, where all states have only finite number of successors. On the other hand, even in very simple reachability games where every state has at most 2 successors, Player Max may not have an optimal strategy (cf. Example 1). Our main result gives a condition sufficient for the existence of optimal strategies for Player Max.

Theorem 1. *Let $\mathscr{G}$ be an open stochastic game. Player Max has an optimal strategy in all states, if*

$$\text{the set } V_\varepsilon := \{Val(s) \mid s \text{ is a state of } \mathscr{G} \wedge Val(s) \geq \varepsilon\} \text{ is finite for every } \varepsilon > 0. \qquad (*)$$

In particular, $\mathscr{G}$ is not assumed to be finitely-branching. Condition $(*)$ is just saying that the set $V := \{Val(s) \mid s \text{ is a state}\}$ has no accumulation points, or the only such point is 0. It is a trivial task to construct a game where none of the players owns a single state, i.e., a Markov chain, and where the set V contains other accumulation points than 0. In Markov chains, however, each player has only one, trivial, strategy, which must thus be the optimal one. This shows that $(*)$ is not necessary. However, there are at least two reasons for which $(*)$ is interesting: First, we identify a class of recently studied infinite-state stochastic games which satisfy the assumption of Theorem 1, and for which the existence of optimal strategies for Max was not known before. This class, properly described later, consists of games generated by One-Counter automata [3, 2, 4], which satisfy a certain additional property, which can be tested algorithmically. As a special case, this class involves a maximizing variant of Solvency Games [1].

Second, in Examples 1 and 2, we show games where Player Max lacks optimal strategies. These games are rather simple, and violate $(*)$ only "very slightly", in particular, they (1) are finitely-branching, and in fact have both the out-degree and in-degree of the game graph bounded by 2, (2) do not contain states of Player Min at all, (3) all transition probabilities in stochastic states are uniformly distributed, and (4) V has only one accumulation point. This point is 1 in Example 1, and $1/2$ in Example 2. In the

latter case, the accumulation point is approached only from above, and $V \cap [0,1/2) = \{0\}$. Thus it is not possible to weaken the assumption $(*)$ in Theorem 1 by allowing other accumulation points than 0.

As noted before, *both* players having optimal strategies implies strong determinacy. But even for finitely-branching reachability games strong determinacy still holds, although Player Max may not have optimal strategies, and only Player Min always does [7]. Interestingly, we show here that under $(*)$, where Max has optimal strategies, and Min may not have such, strong determinacy survives.

Theorem 2. *Let $\mathscr{G}$ be an open stochastic game satisfying $(*)$. Then $\mathscr{G}$ is strongly determined.*

Related work and open questions. Blackwell games are more general than our stochastic games, players there choose their moves simultaneously, not knowing the concurrent choice of the opponent. A famous determinacy result in the sense of (1) for Blackwell games is given in [12]. Finitely-branching reachability games have been studied as a theoretical background for some algorithmic results concerning BPA games (i.e., games with graphs generated by stateless pushdown automata) in [7, 6]. Finite-state reachability stochastic games were studied in [8]. In view of existence of optimal strategies and strong determinacy, finite-state games are not interesting: optimal strategies always exist there. However, the precise complexity of associated computational problems for these games is a long-standing and interesting open problem.

Theorem 2 and the results from [7, 6] give us two classes of strongly determined games: games satisfying $(*)$, and finitely-branching games, respectively. Neither of these two classes is contained in the other. The most interesting question in our opinion is whether the following conjecture is true; and if it is not, for which, as weak as possible, restrictions on W and/or $\mathscr{G}$ it becomes true.

Conjecture 1. *Let $\mathscr{G}$ be a stochastic game, and W a winning condition, such that Player Max (or Player Min) has an optimal strategy in every state of $\mathscr{G}$. Then $\mathscr{G}$ is strongly determined.*

We do not even know whether the conjecture is true for all games where W is a reachability condition. Other open questions include finding new interesting classes of games where one of the players is guaranteed to have optimal strategies.

Outline of the paper. We briefly formalise the necessary notions, and recall some important known facts in Section 2. In Section 3 we prove Theorem 1 in the special case of games without Player Min. Both theorems are then proved in full generality in Section 4. Finally, in Section 5 we briefly explain what are One Counter games, and apply our results to them.

2 Preliminaries

As noted in the Introduction, we use the simple term "games" for our special kind of games (Definition 2). Because we do not speak about other games here, we hope the reader will excuse us for this inaccuracy.

Definition 1. A *game graph*, $G = (S, \to, \delta)$, has a countable set S of states, partitioned into sets S_0, S_1, S_2 of stochastic states, states of Player Max, and Player Min, respectively; a countable *transition relation* $\to \subseteq S \times S$ such that $\forall r \in S : \exists s \in S : r \to s$; and a probability weight function $\delta : S_0 \times S \to [0,1]$ such that for all $r \in S_0$ we have $\sum_{r \to s} \delta(r,s) = 1$.

A *run* is an infinite path in a game graph. For a finite path w, we denote the states it visits by $w(0), w(1), \ldots, w(k)$, and call $k = len(w)$ the *length* of w. $Run(w)$ is the set of all runs extending w. Unions of sets of the form $Run(w)$ are called open sets, they are open in the product topology over the discrete spaces S. Closing the set of open sets under complements and countable union defines the set of (Borel-)measurable sets.

Definition 2. A *game*, $\mathscr{G}$, is given by a game graph, G, and a Borel-measurable set of runs, W, called the *winning condition*. If there is some $T \subseteq S$ so that $W = \bigcup\{Run(w) \mid w \text{ ends in } T\}$ then W is a *reachability* condition, and $\mathscr{G}$ is called a *reachability game*.

A strategy for Player Max is a function assigning to every finite path (called a *history*) ending in a state $s \in S_1$ a distribution over the successors of s. Similarly, a strategy for Min is defined for histories ending in S_2. A strategy is memoryless, if it only depends on the last state of the history.

Fixing a pair of strategies, (σ, π), for Max and Min, respectively, we assign to every finite path, w, the product, $\rho^{\sigma,\pi}(w)$, of weights on the edges along w given by δ, σ, and π. Fixing also an initial state, s, we define a probability measure $\mathbb{P}_s^{\sigma,\pi}[\cdot]$ by $\mathbb{P}_s^{\sigma,\pi}[Run(w)] := 0$ for w not starting in s, $\mathbb{P}_s^{\sigma,\pi}[Run(w)] := \rho^{\sigma,\pi}(w)$ for w starting in s, and extending this to complement and union to satisfy the axioms of a probability measure. The uniqueness of this construction is a standard fact, see, e.g., [13, p. 30].

The definition of the value, $Val(\cdot)$, given in (1), has thus been formalised. For $\varepsilon \geq 0$, a strategy, σ, for Max is ε-optimal in a state s if $\mathbb{P}_s^{\sigma,\pi}[W] \geq Val(s) - \varepsilon$ for all strategies, π, for Min. The ε-optimal strategies for Min are defined analogously. We call 0-optimal strategies just optimal.

2.1 Technical Assumptions

Although a game graph, in general, may have an arbitrary structure, we can always transform it to be a forest, without changing the properties of the game, by keeping track of the history inside the states. More precisely, given a game $\mathscr{G} = (G, W)$, $G = (S, \to, \delta)$, consider a game $\mathscr{G}' = (G', W')$, $G' = (S', \hookrightarrow, \delta')$, where the states in S' are just finite sequences of states from S. In particular, $S \subseteq S'$, and whenever $r \to s$ in $\mathscr{G}$ then $wr \hookrightarrow wrs$ in $\mathscr{G}'$. Projecting the states of S' to their last component induces a map, ϕ, from paths in G' to paths in G. We set $W' := \phi^{-1}(W)$. The map ϕ also induces a map, Φ, from strategies in $\mathscr{G}$ to strategies in $\mathscr{G}'$, by sending histories through ϕ. Naturally, the partition of S', and the weight function δ' are both derived from S and δ by projecting states from S' to the last component.

It is easy to verify that for every $s \in S$, if we restrict the game graphs of $\mathscr{G}$ and $\mathscr{G}'$ to states reachable from s, then ϕ is clearly bijective and preserves measurability in both directions. Also Φ is bijective, and for all measurable $A \subseteq Run(s)$, and all pairs (σ, π) of strategies: $\mathbb{P}_s^{\sigma,\pi}[A] = \mathbb{P}_s^{\Phi(\sigma),\Phi(\pi)}[\phi(A)]$. As a consequence, $Val(s)$ is the same in $\mathscr{G}$ and $\mathscr{G}'$ for all $s \in S$, and the sets of all values in $\mathscr{G}$ and in $\mathscr{G}'$ are equal. Also, W is open iff W' is a reachability condition. Every strategy in $\mathscr{G}'$ is memoryless, because G' is a forest. Finally, once we have a reachability objective, with the target set T, we may clearly assume without loss of generality, that all states in T are absorbing. This shows that to prove Theorems 1 and 2 we may safely assume the following:

Assumption 1. *The game graph is always a forest, all strategies are memoryless, and the winning condition is a reachability condition specified by some target set $T \subseteq S$, such that for all $t \in T$ the only edge leaving t is $t \to t$.*

2.2 Known Results for Reachability Games

We state here some known results to be used later. The following gives a characterisation of values, and allows us to characterise the existence of optimal strategies for Min.

Fact 1 (cf. [7, Theorem 3.1])**.** *Let $\mathscr{G} = (G, W)$, $G = (S, \to, \delta)$ be a game, with $W = \bigcup\{Run(w) \mid w \text{ ends in } T\}$. The least fixed point of the following (Bellman) functional $\mathscr{V} : (S \to [0,1]) \to (S \to [0,1])$*

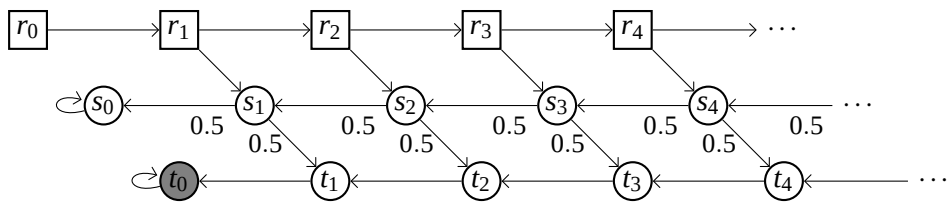

Figure 1: A reachability game where Player Max (□ states) has no optimal strategy.

exists and is equal to $Val(\cdot)$.

$$\mathscr{V}(f)(s) = \begin{cases} 1 & \text{if } s \in T \\ \sup\{f(r) \mid s \to r\} & \text{if } s \in S_1 \setminus T \\ \inf\{f(r) \mid s \to r\} & \text{if } s \in S_2 \setminus T \\ \sum_{s \to r} \delta(s,r) \cdot f(r) & \text{if } s \in S_0 \setminus T \end{cases}$$

Corollary 1 (cf. [7, Theorem 3.1])**.** *Let $\mathscr{G}$ be a game as in Fact 1. Let $G' = (S, \hookrightarrow, \delta)$ be a subgraph of G where $\hookrightarrow$ is a subset of $\to$, and if there is a pair $r, s \in S$ such that $r \to s$ and $r \not\to s$ then $r \in S_2$ and there is some $s' \in S$ such that $r \hookrightarrow s'$ and $Val(s') \leq Val(s)$ in $\mathscr{G}$. Let $\mathscr{G}' = (G', W)$. Then the values are the same in $\mathscr{G}$ and $\mathscr{G}'$.*

As a consequence, a strategy, π, for Min is optimal iff for all $r \in S_2$ it chooses with positive probability only successors $s \in S$ satisfying $Val(r) = Val(s)$.

Proof. Let $\mathscr{V}'$ be the Bellman functional associated with $\mathscr{G}'$. Observe that the values in $\mathscr{G}$ form a fixed point of $\mathscr{V}'$, thus for all $s \in S$, $Val(s)$ in $\mathscr{G}'$ is equal to or less than $Val(s)$ in $\mathscr{G}'$. Moreover, it cannot be less, because Player Max has the same set of strategies in $\mathscr{G}'$ as in $\mathscr{G}$, whereas Player Min does not get more strategies in $\mathscr{G}'$. To derive the consequence, remove all edges not used by π. □

Note that the situation is not symmetric for Player Max. Consider games without Player Min, and with out-degree and in-degree bounded by 2. In particular, this implies that every state, r, of Player Max has at least one successor, s, with $Val(r) = Val(s)$. Even in these games, Player Max may lack optimal strategies, as illustrated in the following classical (see, e.g., [4, p. 871],[5, Example 6]) example.

Example 1. Consider the reachability game from Figure 1. Its game graph, G, has the set $S := \{r_i, s_i, t_i \mid i \geq 0\}$ of states, partitioned by $S_0 = \{s_i, t_i \mid i \geq 0\}$, $S_1 = \{r_i \mid i \geq 0\}$, and $S_2 = \emptyset$. Transitions are $s_0 \to s_0$, $t_0 \to t_0$, and $r_{i-1} \to r_i$, $r_i \to s_i$, $s_i \to s_{i-1}$, $s_i \to t_i$, and $t_i \to t_{i-1}$ for $i > 0$. Probabilities are always uniform. The target set is $T = \{t_0\}$. Clearly, $Val(s_i) = 1 - 2^{-i}$ for all $i \geq 0$. Thus $Val(r_i) = 1$ for all $i \geq 0$: for every $N > 0$, choosing the transition $r_i \to r_{i+1}$ for $i < N$, and the transition $r_i \to s_i$ for $i \geq N$, is a 2^{-N}-optimal strategy for Max. Yet Max has no optimal strategy in any r_i, $i \geq 0$: no strategy reaching some s_j is optimal, and, on the other hand, never reaching s_j means never reaching t.

3 Games without Player Min

Proposition 1. *Let $\mathscr{G} = (G, W)$ be a stochastic game, where $G = (S, \to, \delta)$ and $S_2 = \emptyset$,* [1] *and W is open. If $(*)$ from Theorem 1 is satisfied then Player Max has an optimal strategy in all states.*

[1] These are also sometimes called (minimizing) Markov Decision Processes (MDPs), see, e.g., [3, 2, 13].

We fix the game $\mathscr{G}$ from Proposition 1 in the rest of this section, devoted to proving the proposition. By Assumption 1, G is a forest, and there is $T \subseteq S$ such that $W = \bigcup\{Run(w) \mid w \text{ ends in } T\}$ and for all $t \in T$ there is only one transition: $t \to t$. The proof is by contradiction, in three steps. First, we prove that if there is a state with no optimal strategy, then there must be a state from which winning with probability sufficiently close to the optimum implies the need to use some value decreasing transition. A transition $r \to s$ is value decreasing if $Val(r) > Val(s)$. Second, we will argue that the potential "damage" caused by this transition is positive and bounded away from 0, independently of the actual strategy. Third, we show that $(*)$ implies that the potential "damage" factor is indeed bounding the probability of reaching T away from the value, which is a contradiction with the definition of the value.

We introduce a random variable, L (for "loss"). For a run, ω, a *losing* index is every i, such that $\omega(i) \in S_1$ and $Val(\omega(i)) > Val(\omega(i+1))$. If there is no losing index for ω, we set $L(\omega) := 0$. Otherwise, there is the least losing index, i, and we set $L(\omega) := Val(\omega(i)) > 0$. Finally, we say that a state $s \in S$ is *losing* if there is some $\delta_s > 0$ such that for every δ_s-optimal strategy, σ, in s, we have $\mathbb{P}_s^\sigma[L > 0] > 0$.

Lemma 1. *Assume $(*)$. If $\exists s \in S$ such that $\forall \sigma : \mathbb{P}_s^\sigma[W] < Val(s)$ then there is also some losing state.*

Proof. By contradiction. Assume there is no losing state, we construct an optimal strategy in every state. Define a subset $\hookrightarrow$ of the transition relation $\to$ of $\mathscr{G}$, by setting for every pair $r, s \in S$: $r \hookrightarrow s$ iff $r \to s$ and either $r \in S_0$, or $Val(r) = Val(s)$. Observe that $(*)$ implies that for all $r \in S_1$ there is at least one s such that $r \hookrightarrow s$ and $Val(r) = Val(s)$. Thus $\hookrightarrow$ is total and $G' = (S, \hookrightarrow, \delta)$ is a game graph. Without losing states, for every $r \in S$ and every $\varepsilon > 0$ there is some ε-optimal strategy, σ, such that $\mathbb{P}_s^\sigma[L > 0] = 0$, i.e., σ does not use value-decreasing transitions. This strategy works in $\mathscr{G}' = (G', W)$ as well, winning with the same probability, as in $\mathscr{G}$. The values in $\mathscr{G}$ and $\mathscr{G}'$ are thus the same.

Consider now $\mathscr{G}'$. Denote by $FP_k(s)$ the set of all finite paths of length k starting in s. Due to the last sentence in Assumption 1, and because $\hookrightarrow$ preserves value, the following is true in $\mathscr{G}'$:

$$\forall k \geq 0 : \forall \sigma : \forall s \in S : Val(s) = \sum_{w \in FP_k(s)} \mathbb{P}_s^\sigma[Run(w)] \cdot Val(w(k)). \tag{2}$$

For all $s \in S$ fix a $1/4 \cdot Val(s)$-optimal strategy σ_s. After some $n_s \geq 0$ of steps, T must be reached from s under σ_s with probability at least $Val(s)/2$, as $\mathbb{P}_s^{\sigma_s}[W] = \lim_{k\to\infty} \mathbb{P}_s^\sigma[\{Run(w) \mid len(w) \leq k \wedge w(k) \in T\}]$.

For all $s \in S$ we finally construct a strategy σ for $\mathscr{G}'$, optimal in s. Because the values are the same in $\mathscr{G}$ and $\mathscr{G}'$, and every strategy for $\mathscr{G}'$ is also a strategy for $\mathscr{G}$, this will finish the proof of the lemma. The strategy σ starts in s according to σ_s, and follows it for n_s steps. After that, having arrived to some state r, it switches to σ_r and follows it for other n_r steps. This is repeated ad infinitum. The invariant (2), and the choice of n_r and σ_r for $r \in S$, guarantee that after the m-th stage of the above repetitive process, T has actually been reached with probability $(1 - 2^{-m}) \cdot Val(s)$, proving that σ is optimal. □

For every losing state, $s \in S$, and every constant $\varepsilon > 0$ we define $\ell_s^\varepsilon := \inf\{\mathbb{E}_s^\sigma[L] \mid \sigma \text{ is } \varepsilon\text{-optimal in } s\}$. Since $\ell_s^\varepsilon \leq \ell_s^\zeta \leq 1$ for $\varepsilon \geq \zeta$, the limit $\ell_s := \lim_{\varepsilon\to 0} \ell_s^\varepsilon$ exists.

Lemma 2. *Assume $(*)$. For every losing state, s, in $\mathscr{G}$ we have $\ell_s > 0$.*

Proof. By contradiction. Assume that s is losing and $\ell_s = 0$. To every strategy σ which may possibly use value-decreasing transitions $r \to r'$ where $Val(r) > Val(r')$ we consider a strategy $\bar{\sigma}$, which copies the moves of σ until a value-decreasing transition is chosen. From that point on, just before the value-decreasing transition, the strategy $\bar{\sigma}$ keeps choosing arbitrary successors with the only requirement that they preserve the value, i.e., whenever $\bar{\sigma}$ chooses a transition $s \to s'$ with a positive probability, $Val(s) = Val(s')$. Such a choice always exists, because $\sup_{s\to s'} Val(s') = Val(s)$, and either $Val(s) = 0$, in which

case $Val(s') = 0$ for all s', $s \to s'$, or $Val(s) > 0$, and by $(*)$ $Val(s) > 0$ cannot be an accumulation point, so there is some s', $s \to s'$ with $Val(s) = Val(s')$. Observe that for every σ, $\mathbb{P}_s^{\sigma}[W] - \mathbb{P}_s^{\bar{\sigma}}[W] \leq \mathbb{E}_{\sigma}^{s}[L]$. As a consequence, due to $\ell_s = 0$, $Val(s) = \sup\{\mathbb{P}_s^{\bar{\sigma}}[W] \mid \sigma \text{ is some strategy}\}$. This contradicts s being losing, since $\mathbb{P}_s^{\bar{\sigma}}[L > 0] = 0$ for every σ. □

Proof of Proposition 1. By contradiction. Assume $(*)$, and that there is some $r \in S$ with no strategy optimal in r. By Lemma 1, there is a losing state, $s \in S$. By Lemma 2, $\ell_s > 0$. Choose some $\varepsilon > 0$ such that $\ell_s^{\varepsilon} \geq \ell_s/2 > 0$. Thus under every ε-optimal strategy, σ, with some positive probability, $p > 0$, a state $r \in S_1$ with $Val(r) \geq \ell_s^{\varepsilon}$ is visited, and some transition $r \hookrightarrow r'$ with $Val(r') < Val(r)$ is taken. Observe that $(*)$ gives us the following "value-gap":

$$\delta := \inf\{|Val(r) - Val(r')| \mid r, r' \in S, Val(r) \neq Val(r'), Val(r) \geq \ell_s^{\varepsilon}\} > 0.$$

This allows us to bound p independently of σ, since $\ell_s^{\varepsilon} \leq \mathbb{E}_{\sigma}^{s}[L] \leq p \cdot 1 + (1-p)(\ell_s^{\varepsilon} - \delta)$ and hence

$$p \geq \frac{\delta}{1 + \delta - \ell_s^{\varepsilon}} > 0.$$

Thus for every strategy, σ, we have that $Val(s) - \mathbb{P}_s^{\sigma}[W] \geq \min\{\varepsilon, \delta \cdot p\} > 0$. This clearly contradicts the definition of $Val(s)$. The proof is finished. □

4 Reachability Games

In this section we prove Theorems 1 and 2. Let us fix a game $\mathscr{G} = (G, W)$, where $G = (S, \to, \delta)$, satisfying Assumption 1. Also assume that W is open, and thus there is $T \subseteq S$ such that $W = \bigcup\{Run(w) \mid w \text{ ends in } T\}$. We call a state s *safe* if $\forall\sigma$ for Max : $\exists\pi_{\sigma}$ for Min : $\mathbb{P}_s^{\sigma,\pi_{\sigma}}[W] = 0$. The following lemma states the strong determinacy restricted to states with value 0, and will be useful in proving each of both theorems.

Lemma 3. *If $\mathscr{G}$ satisfies $(*)$ then for every safe $s \in S$: $\exists\pi$ for Min : $\forall\sigma$ for Max : $\mathbb{P}_s^{\sigma,\pi}[W] = 0$.*

Proof. We cut off some choices for Min in the game graph G of $\mathscr{G}$, and obtain its sub-graph G', so that all states reachable in G' from s have value 0 in $\mathscr{G}' = (G', W)$. In particular, no run can satisfy W. Because the choices of Max remain unrestricted in G', this ensures that the probability of W is 0 in $\mathscr{G}$ as well. Let us proceed in more detail.

Observe that every safe state has value 0, so no safe state is in T. Also, observe that for every safe $r \in S_0 \cup S_1$ and $s \in S$, if $r \to s$ then s is safe. Likewise, if $r \in S_2$ is safe, then there must be a safe s such that $r \to s$. Fix a safe s, and define G' as the smallest sub-graph of G containing s and satisfying that if r is in G', then so is every safe successor r' of r in G. As shown above, G' is a game graph, the probability assignment δ from G is valid in G' as well, and all states in G' are safe. Hence, no paths in G' visit T, and the value of every state in $\mathscr{G}'$ is 0. Fix an arbitrary strategy π for Min in $\mathscr{G}' = (G', W)$, then $\mathbb{P}_s^{\sigma,\pi}[W] = 0$ for all σ of Max in $\mathscr{G}'$. All transitions out of safe states of Max were preserved in G', and π is also a strategy in $\mathscr{G}$, so we have $\mathbb{P}_s^{\sigma,\pi}[W] = 0$ also for every σ of Max in $\mathscr{G}$. □

4.1 Proof of Theorem 1

Lemma 4. *If $\mathscr{G}$ satisfies $(*)$, then for all $s \in S$ we have: $\forall\pi$ for Min : $\exists\sigma$ for Max : $\mathbb{P}_s^{\sigma,\pi}[W] \geq Val(s)$.*

Proof. For every (memoryless, due to Assumption 1) strategy π of Player Min, we denote by $\mathscr{G}_\pi$ the game where the choices of Player Min are resolved using π. Formally, $\mathscr{G}_\pi = (G', W)$, where $G' = (S', \hookrightarrow, \delta')$, and (1) $S' = S$ but comes with a different partition: $S'_0 = S_0 \cup S_2$, $S'_1 = S_1$, $S'_2 = \emptyset$, (2) the relation $\hookrightarrow \subseteq \rightarrow$ is given by $r \hookrightarrow s$ iff $r \rightarrow s$ and either $r \in S_0 \cup S_1$, or $r \in S_2$ and $\pi(r)(s) > 0$, and (3) $\delta' = \delta \cup \pi$. For every strategy σ for Player Max, and every $s \in S$ the measure $\mathbb{P}_s^{\sigma,\pi}[\cdot]$ in $\mathscr{G}$ obviously coincides with $\mathbb{P}_s^{\sigma}[\cdot]$ in $\mathscr{G}_\pi$. Thus we may apply Proposition 1 to all $\mathscr{G}_\pi$ to derive the lemma. □

Consider now the following game $\mathscr{H} = (H, W)$, which is a slight modification of $\mathscr{G}$. The set of states of $H = (S, \hookrightarrow, \delta_H)$ is S, the same as in G, and with the same partition. There is a transition $r \hookrightarrow s$ iff exactly one of these three situations occurs: $Val(r) = 0$ in $\mathscr{G}$, and $s = r$; or $Val(r) > 0$, $r \in S_0$ and $r \rightarrow s$; or $Val(r) > 0$, $r \notin S_0$, $r \rightarrow s$, and $Val(r) = Val(s)$ in $\mathscr{G}$. In other words, in H we made all states with value 0 absorbing, and only left value preserving transitions for players. Finally, δ_H is the only probability weight function which coincides with δ on stochastic states with positive value.

Lemma 5. *If $\mathscr{G}$ satisfies $(*)$, then H is a game graph, and the values are the same in $\mathscr{G}$ and $\mathscr{H}$.*

Proof. We refine the modifications from above into three steps, obtaining game graphs $H_0 = G$, H_1, H_2, and $H_3 = H$. We will show for each $i \in \{1,2,3\}$ that H_i is a game graph, and that the values are the same in $\mathscr{H}_i = (H_i, W)$ as they are in $\mathscr{G}$. All the graphs constructed have the same set of states, S, and the same partition, as G, and the same weight function, δ_H, as H.

$H_1 = (S, \mapsto, \delta_H)$, and $r \mapsto s$ iff $Val(r) = 0$ in $\mathscr{G}$, and $s = r$, or $Val(r) > 0$ and $r \rightarrow s$. H_1 is clearly a game graph, because $\mapsto$ is total. The values did not change, because each absorbing loop outside of T has value 0. Moreover, every $r \in S_2$ has always a successor with the same value. Indeed, if $Val(r) = 0$ then r itself is its own successor in G_1; if $Val(r) > 0$ then $\inf_{r \mapsto s} Val(s) = Val(r)$, and by $(*)$, since $Val(r) > 0$ cannot be an accumulation point, there is some s, $r \mapsto s$ with $Val(r) = Val(s)$. By Corollary 1, Min has optimal strategies in $\mathscr{H}_1$.

$H_2 = (S, \rightsquigarrow, \delta_H)$, and $r \rightsquigarrow s$ iff $r \mapsto s$ and either $Val(r) = 0$ in $\mathscr{G}$, or $r \notin S_2$, or (if $Val(r) > 0$ and $r \in S_2$) $Val(r) = Val(s)$ in $\mathscr{G}$. Because Min has always value-preserving transitions in $\mathscr{H}_1$, H_2 is clearly a game graph, and by Corollary 1 all strategies of Min in $\mathscr{H}_2$ are optimal. Fix one such π for Min, and an arbitrary $s \in S$. By Lemma 4 there is a σ for Max in $\mathscr{G}$ (and thus also in $\mathscr{H}_2 = (H_2, W)$) such that $\mathbb{P}_s^{\sigma,\pi}[W] \geq Val(s)$. Because π is optimal, σ cannot choose value-decreasing transitions. Thus, even when only using edges in $\hookrightarrow$, i.e., from $H_3 = H$, we still obtain that $\inf_\pi \sup_\sigma \mathbb{P}_s^{\sigma,\pi}[W] = Val(s)$. Thus also the graph H is a game graph, and the values in $\mathscr{H}$ and $\mathscr{G}$ are the same. □

Lemma 6. *If $\mathscr{G}$ satisfies $(*)$, then Player Max has an optimal strategy, σ, in $\mathscr{H}$.*

Proof. We first describe σ, then we prove that it is optimal. In every state, s, there is some $1/2 \cdot Val(s)$-optimal strategy, τ_s, for Max. We call a history (i.e., a finite path), w, starting in some state s, and ending in r, *lazy*, if $Val(r) > 0$ and $\inf_\pi \mathbb{P}_s^{\tau_s,\pi}[W \mid Run(w)] = 0$. Observe that each history, w, can be uniquely split into a sequence of sub-paths, divided by single states, $w = s_0 w_0 s_1 w_1 s_2 \cdots s_k w_k$, $k \geq 1$, $s_i \in S$, $w_i \in S^*$, such that for all $i < k$, $s_i w_i s_{i+1}$ is lazy, and for all $i \leq k$, $s_i w_i$ is not lazy. We call k the *laziness index* of w, written $laz(w)$ and $s_k w_k$ the *non-lazy* suffix of w. We now define σ for a history w with a non-lazy suffix $s_k w_k$ by $\sigma(w) := \tau_{s_k}(s_k w_k)$.

Now we prove that σ is optimal. To do so, we need to extend the laziness index to runs. For a run, ω, we set $Laz(\omega) := \sup\{laz(w) \mid \omega \in Run(w)\} \in \mathbb{N} \cup \{\infty\}$. Thus we defined a random variable, Laz. We prove the following claim, which clearly implies the statement of the lemma:

$$\forall s \in S : \forall \pi \text{ for Min} : \forall k \geq 0 : \mathbb{P}_s^{\sigma,\pi}[W \wedge Laz \leq k] \geq Val(s) \cdot (1 - 2^{-k}). \tag{3}$$

By induction on k. Fix some $s \in S$, and a strategy, π, for Min. Clearly, (3) is true for $k = 0$. Also it is true when $Val(s) = 0$. Assume thus $Val(s) > 0$ and $k = \ell + 1$ for some $\ell \geq 0$. We set L to be the set of all finite paths, w, such that $laz(w) = k$ and the non-lazy suffix only consists of one state. Denote by $last(w)$ the last state of w. Observe that, by the definition of σ and τ_{s_k},

$$\forall w \in L : \forall \pi \text{ for Min} : \mathbb{P}_s^{\sigma,\pi}[W \mid Run(w)] \geq 1/2 \cdot Val(last(w)). \tag{4}$$

Let Λ be any prefix-free set of finite paths such that $\mathbb{P}_s^{\sigma,\pi}[\bigcup_{w\in\Lambda} Run(w)] = 1$. Because $\mathscr{H}$ only contains value-preserving edges for players, we have

$$Val(s) = \sum_{w\in\Lambda} \mathbb{P}_s^{\sigma,\pi}[Run(w)] \cdot Val(last(w)). \tag{5}$$

We have $p := \mathbb{P}_s^{\sigma,\pi}[W \wedge Laz \leq \ell] \geq Val(s) \cdot (1 - 2^{-\ell})$, by the inductive hypothesis. We also have $q := \sum_{w\in L} \mathbb{P}_s^{\sigma,\pi}[Run(w)] \cdot Val(last(w)) = Val(s) - p$, by (5). By (4), $\mathbb{P}_s^{\sigma,\pi}[W \wedge Laz = k] \geq q \cdot 1/2$. Finally,

$$\begin{aligned}
\mathbb{P}_s^{\sigma,\pi}[W \wedge Laz \leq k] &= \mathbb{P}_s^{\sigma,\pi}[W \wedge Laz \leq \ell] + \mathbb{P}_s^{\sigma,\pi}[W \wedge Laz = k] \\
&= p + q \cdot 1/2 = p + (Val(s) - p) \cdot 1/2 = p/2 + Val(s)/2 \\
&\geq (2^{-1} - 2^{-(\ell+1)}) \cdot Val(s) + Val(s) \cdot 2^{-1} = (1 - 2^{-(\ell+1)}) \cdot Val(s).
\end{aligned}$$

□

Proof of Theorem 1. Consider the strategy σ from Lemma 6. It partially defines a strategy in $\mathscr{G}$. To complete its definition, we now specify it for histories containing a transition of the form $r \to s$, where $r \in S_2$ and $Val(s) > Val(r)$, by requiring σ to behave as a $1/2 \cdot (Val(s) - Val(r))$-optimal strategy since that point. Fix an initial state, s, and consider an arbitrary strategy, π, of Min. If π is optimal, then it is also valid in $\mathscr{H}$, and $\mathbb{P}_s^{\sigma,\pi}[W] = Val(s)$ by Lemmata 5 and 6. For a non-optimal π it is easy to verify that $\mathbb{P}_s^{\sigma,\pi}[W] > Val(s)$ by both the definition of σ, and Lemmata 5 and 6. □

4.2 Proof of Theorem 2

If both players have optimal strategies, the game is strongly determined. However, even under Condition $(*)$, Player Min may not always have an optimal strategy, because of states with value 0, without value-preserving transition for Min available. See the game in [7, Fig. 1] restricted to states reachable from s, for an example. Theorem 2 is a direct consequence of Lemma 7 and Lemma 9, where the former lemma deals with all "easy cases", and the latter "patches" the above deficiency by using Lemma 3 to deal with states with value 0, and "restoring" the optimal strategies for both players in the rest.

Lemma 7. *Assume that $\mathscr{G}$ satisfies $(*)$. Let $s \in S$, $0 \leq \nu \leq 1$, and $\rhd \in \{>, \geq\}$. Assume that either $Val(s) = 0$, or $\nu \neq Val(s)$, or $\rhd = \geq$. Then either Player Max has a strategy $\bar{\sigma}$ such that $\forall \pi : \mathbb{P}_s^{\bar{\sigma},\pi}[W] \rhd \nu$, or Player Min has a strategy $\bar{\pi}$ such that $\forall \sigma : \mathbb{P}_s^{\sigma,\bar{\pi}}[W] \not\rhd \nu$.*

Proof. The case when $\nu = Val(s) = 0$ is solved by Lemma 3. If $\nu < Val(s)$, we can choose any $1/2 \cdot (Val(s) - \nu)$-optimal strategy for Max as $\bar{\sigma}$. Similarly, if $\nu > Val(s)$, we can choose any $1/2 \cdot (\nu - Val(s))$-optimal strategy for Min as $\bar{\pi}$. If $\nu = Val(s)$ and $\rhd = \geq$, we can choose any optimal strategy for Max as $\bar{\sigma}$. Such a strategy exists due to Theorem 1. □

It remains to solve $v = Val(s) > 0$ and $\rhd = >$. We do two preprocessing steps on $\mathscr{G}$ to first obtain $\mathscr{G}'$, and then $\mathscr{H}$. In $\mathscr{H}$ both players will have optimal strategies, and we will be able to lift such a strategy for Min back to $\mathscr{G}$ iff Max does not have a strategy ing $\mathscr{G}$ to always win with probability $> Val(s)$.

We fix $s \in S$ with $Val(s) > 0$ and set $R := \{r \in S \mid Val(r) = 0 \wedge \exists \bar{\sigma} : \forall \pi : \mathbb{P}_r^{\bar{\sigma},\pi}[W] > 0\}$. Intuitively, if Max does not have a strategy to always win with probability $> Val(s)$, then Min can always respond to a strategy σ of Max with a π_σ, so that R is not visited at all from s under these strategies, and yet Max wins with probability at most $Val(s)$. Thus, if we cut off all states from R, producing the game $\mathscr{H}$, we obtain a valid game graph, and the values of states will not change.

Before we describe this formally, we observe that neither of the players benefits from using transitions which do not preserve the value. Let $\mathscr{G}' = (G', W)$, $G' = (S, \mapsto, \delta)$ be a game given by restricting the edges of G to value-preserving where possible: for all $r, r' \in S$ we require that $r \mapsto r'$ iff $r \to r'$ and either $r \in S_0 \cup R$, or $Val(r) = Val(r')$.

Lemma 8. *Assume that $\mathscr{G}$ satisfies $(*)$. Then the values in $\mathscr{G}$ and in $\mathscr{G}'$ are the same, and for all $s \in S$, each of the following is true in $\mathscr{G}'$ if it is true in $\mathscr{G}$:*

$$\forall \sigma \text{ for Max} : \exists \pi_\sigma \text{ for Min} : \mathbb{P}_s^{\sigma,\pi_\sigma}[W] \leq Val(s), \tag{6}$$

$$\forall \pi \text{ for Min} : \exists \sigma_\pi \text{ for Max} : \mathbb{P}_s^{\sigma_\pi,\pi}[W] > Val(s). \tag{7}$$

Proof. By Theorem 1, there is an optimal strategy, σ, for Min. This is also a strategy for $\mathscr{G}'$, thus for all $s \in S$, $Val(s)$ in $\mathscr{G}'$ is at least $Val(s)$ in $\mathscr{G}$. On the other hand, by Corollary 1, cutting off non-optimal edges leaving states from $S_2 \setminus R$ does not alter the values. Further, cutting off non-optimal edges from S_1 could only decrease the values. Thus, for all $s \in S$, the values in $\mathscr{G}'$ and $\mathscr{G}$ are equal.

Now we fix some $s \in S$, and prove that if (6) is true in $\mathscr{G}$ then it is true in $\mathscr{G}'$. Let σ be a strategy for Max in $\mathscr{G}'$, i.e., it is a strategy for $\mathscr{G}$ which does not use value-decreasing edges. If σ is optimal, then the strategy π_σ from (6) in $\mathscr{G}$ necessarily has to use value-preserving edges everywhere, and thus it is valid in $\mathscr{G}'$ as well. If σ is not optimal, consider again the response π_σ of Min to satisfy (6) in $\mathscr{G}$. If π_σ cannot be used directly in $\mathscr{G}'$, then there must be some $r \in S_2 \setminus R$ where π_σ chooses a successor r' with $Val(r') > Val(r)$. But because $r \notin R$, there must also be a successor r'' such that $Val(r'') = Val(r)$. We modify π_σ to a π'_σ, which chooses for all such r the value-preserving successor instead of r', and continues as a $1/2 \cdot (Val(r') - Val(r))$-optimal strategy in $\mathscr{G}'$. Clearly, $\mathbb{P}_s^{\sigma,\pi'_\sigma}[W] \leq \mathbb{P}_s^{\sigma,\pi_\sigma}[W]$ in $\mathscr{G}$, and since π'_σ is also a strategy in $\mathscr{G}'$, (6) is true in $\mathscr{G}'$ as well.

Finally, we prove that if (7) is true in $\mathscr{G}$ then it is true in $\mathscr{G}'$. Let π be a strategy in $\mathscr{G}'$. Fix the choices of π in $\mathscr{G}'$ outside of R to define a game $\mathscr{G}_\pi$. By Corollary 1, $\mathscr{G}_\pi$ has the same values as $\mathscr{G}$. Thus, optimal strategies of Max in $\mathscr{G}_\pi$ exist, because $\mathscr{G}_\pi$ satisfies $(*)$, and only choose edges preserving the value in $\mathscr{G}$. Consider the strategy σ_π witnessing (6) in $\mathscr{G}$. We now define a strategy σ'_π in $\mathscr{G}'$: it copies moves of σ_π in $\mathscr{G}'$, unless σ_π chooses some value-decreasing edge. In that case, instead of following σ_π, σ'_π immediately switches to some optimal strategy for $\mathscr{G}_\pi$. Since the values in $\mathscr{G}_\pi$ are the same as in $\mathscr{G}$, this only increases the probability of winning, thus $\mathbb{P}_s^{\sigma'_\pi,\pi}[W] \geq \mathbb{P}_s^{\sigma_\pi,\pi}[W]$. □

Lemma 9. *Assume that $\mathscr{G}$ satisfies $(*)$. For all $s \in S$ such that $Val(s) > 0$, if*

$$\forall \sigma \text{ for Max} : \exists \pi_\sigma \text{ for Min} : \mathbb{P}_s^{\sigma,\pi_\sigma}[W] \leq Val(s), \tag{8}$$

then

$$\exists \pi \text{ for Min} : \forall \sigma \text{ for Max} : \mathbb{P}_s^{\sigma,\pi}[W] \leq Val(s). \tag{9}$$

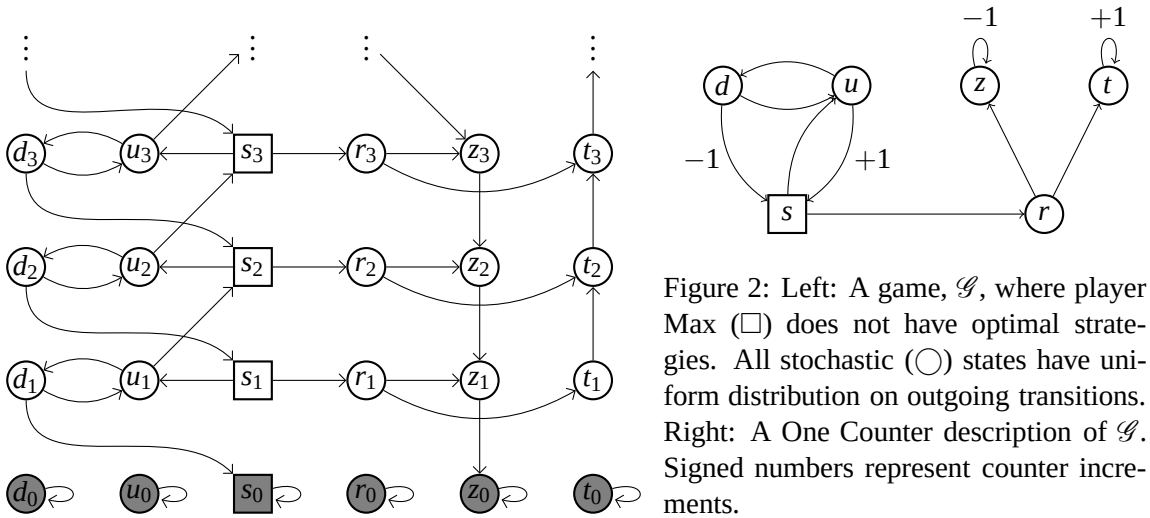

Figure 2: Left: A game, $\mathscr{G}$, where player Max (□) does not have optimal strategies. All stochastic (◯) states have uniform distribution on outgoing transitions. Right: A One Counter description of $\mathscr{G}$. Signed numbers represent counter increments.

Proof. By Lemma 8, if (8) $\implies$ (9) in $\mathscr{G}'$ then the implication holds in $\mathscr{G}$ as well, and if $\mathscr{G}$ satisfies $(*)$ then so does $\mathscr{G}'$. Thus we focus on $\mathscr{G}'$ instead. We describe the modification of $\mathscr{G}'$, called $\mathscr{H}$, where we cut off R. By $Att(R)$ we denote the set of all states, r, such that in $\mathscr{G}'$ Max has a strategy, σ, such that for all π for Min, $\mathbb{P}_r^{\sigma,\pi}[\text{Reach } R] > 0$. Further, we consider the edge relation $\hookrightarrow$, which is simply the relation $\mapsto$ without edges leading to states from $Att(R)$.

We fix some s, $Val(s) > 0$, satisfying (8), and by S' we denote the subset of all $r \in S$ to which there is a path from s in the graph $(S, \hookrightarrow)$. Consider a game graph, $H = (S', \hookrightarrow, \delta)$, inheriting the partition of states from G. The edge relation is the $\hookrightarrow$ defined above, only restricted to $S' \times S'$. Observe that if $r \in S'_0 \cup S'_1$ and $r \to r'$ for some $r' \in S$, then $r' \in S'$. This is because $r \notin Att(R)$ implies $r' \notin Att(R)$ if r is not owned by Min. Similarly, for all $r \in S'_2$ there is a $r' \in S'$ such that $r \hookrightarrow r'$. Thus δ, restricted to S', is still a valid probability weight function, and H is a valid game graph. We abuse the letter W to denote a restriction of W to $\mathscr{H}$, and define a game $\mathscr{H} = (H, W)$.

Because all edges leaving states from $S_2 \setminus R$ were value-preserving in $\mathscr{G}'$, Corollary 1 yields that the values stay the same in $\mathscr{H}$ as they were in $\mathscr{G}'$, and there is an optimal strategy, $\bar{\pi}$ for Min in $\mathscr{H}$. This is also a strategy for $\mathscr{G}'$, and because the choices of Player Max were not affected when reducing $\mathscr{G}'$ to $\mathscr{H}$, we obtain, that for all σ for Max we have $\mathbb{P}_s^{\sigma,\bar{\pi}}[W] \leq Val(s)$ both in $\mathscr{H}$ and in $\mathscr{G}'$. This proves (9). □

5 One Counter Games

One Counter stochastic games (OC-SSGs), see, e.g., [3, 2, 4], are games played on transition graphs of one-counter automata. Such automata have a finite *control-state* unit, Q, and a set of rules, which are triples of the form (r, k, s) with $r, s \in Q$ and $k \in \{-1, 0, +1\}$. States of an OC-SSG are then of the form s_n where $s \in Q$ is a *control state*, and $n \geq 0$ is an integer, representing the *counter value*. Transitions are generated by setting $r_i \to s_j$ if $i > 0$ and there is a rule $(r, j-i, s)$. Moreover, states with counter 0 are made absorbing, $s_0 \to s_0$, to reflect that the system halts with the empty counter. The partition of states is induced by a partition of Q, and the probabilities of transitions out of stochastic states, are induced by probabilities on rules. OC-SSGs come with an implicit reachability objective, the set to be reached is the set $\{s_0 \mid s \in Q\}$ of states with counter 0. Because the system halts in 0 we also call this a *termination* winning condition.

Example 2. In the right-hand part of Figure 2 we give the one-counter automaton with the set $Q = \{s,u,d,r,z,t\}$ of control states. An unlabelled edge, like $s \to u$, represents a 0-rule, e.g., $(s,0,u)$. A label (± 1) represents the counter change, e.g., the loop $t \to t$ represents $(t,+1,t)$. The square-state s belongs to Max, other states are stochastic. The distributions on outgoing transitions are implicitly uniform in this example. In the left-hand part is the generated OC-SSG. Grey states are to be reached. Later in this section we will show that $Val(s_i) = \frac{2^i+1}{2^{i+1}}$, but no strategy of Player Max is optimal in s_i. Observe that $1/2 = \lim_{i\to\infty} \frac{2^i+1}{2^{i+1}}$ is an accumulation point in the set of all values.

Note that every OC-SSG has bounded out-degree and in-degree, in particular it is finitely branching. Thus Min has always optimal strategies in OC-SSGs. However, they may not always satisfy $(*)$, and Example 2 shows that in OC-SSGs, Max may have no optimal strategies. On the other hand, the structure of the accumulation points in the set of all values is well understood for OC-SSGs. To describe it, we need to introduce another winning objective.

In OC-SSGs there is an implicit boundary on the counter value – if it reaches zero, the system halts. However, we may also interpret the one-counter automaton as a directed graph on Q, with the rules as edges with *rewards*. This way we obtain a finite game graph. Accumulating those rewards along a run in such a game graph then corresponds to observing the counter in the OC-SSG, with the exception that the counter does not stop in 0 and may get negative. Adding the winning condition (for Max) that the $\liminf$ of the accumulated rewards be $-\infty$, we just defined $(LimInf = -\infty)$-games.

In [2, 4] it was shown that both players always have pure and memoryless optimal strategies in $(LimInf = -\infty)$-games, and the optimal value is always rational and computable. Observe that the termination values, $Val(s_n)$, for a fixed $s \in Q$, are non-increasing with increasing n. Thus their limit exists, and, in fact, it is an easy exercise to employ the results of [2, 4] to prove that the $(LimInf = -\infty)$-value of a control state, s, equals $\lim_{n\to\infty} Val(s_n)$. Intuitively this is because, with increasing the initial counter, n, the objective of reaching 0 becomes more and more similar to the $(LimInf = -\infty)$ objective. Thus the set of $(LimInf = -\infty)$-values of all states $s \in Q$ contains the set of all accumulation points of the termination values. It is also possible to decide in time polynomial in $|Q|$ whether a $(LimInf = -\infty)$-value, v, actually is an accumulation point, i.e., whether for all states, s, with $(LimInf = -\infty)$-value v the limit of termination values stabilises after finitely many steps.

Corollary 2. *Let $\mathscr{G}$ be an OC-SSG with the set Q of control states. If for every $s \in Q$ the $(LimInf = -\infty)$-value of s is 1 or 0, then Player Max has an optimal strategy for termination in $\mathscr{G}$.*

Proof. The limits of termination values are approached from above, because $Val(s_n) \geq Val(s_{n+1})$ for all $s \in Q$ and all $n \geq 0$. Thus, 1 is not an accumulation point, and we may apply Theorem 1. □

Note that the class of OC-SSGs satisfying the condition of Corollary 2 involves all OC-SSGs where the graph of rules is strongly connected, and one of the players is missing. This is because $(LimInf = -\infty)$ is a prefix independent objective, and the strong connectivity allows the only player to reach each control state almost surely, thus all control states have the same $(LimInf = -\infty)$-value. By results of [11, Theorem 3.2], such a common value can only be 0 or 1. In particular, Corollary 2 covers both the Solvency games, see [1], and their maximizing variant.

In Solvency games, a gambler has an initial positive amount of money, and in each step chooses one of finitely many actions. Each action is associated with a distribution on a finite set of integers. A number from this set is then sampled, and added to the sum of money owned by the gambler (it can be, however, negative), and the process ends only when the wealth becomes ≤ 0. This is easily modelled by one-player OC-SSGs (see [4]), and these have strongly connected graphs of rules, because the only state where the gambler chooses the action, is reachable from all other states. The natural scenario is,

obviously, with Player Min for these games, and there the existence of optimal strategies follows from the finite branching. However, the dual situation, with Player Max, is theoretically interesting as well, and we are not aware of any result prior to our Corollary 2, indicating the existence of optimal strategies for Player Max.

5.1 Analysis of Example 2

Consider an arbitrary $n \geq 1$. It is easy to see that $Val(r_n) = \frac{1}{2}$. Observe that starting in u_n, s_{n+1} is visited with probability $\sum_{i=0}^{\infty} 2^{-1-2i} = \frac{2}{3}$, and s_{n-1} with probability $\frac{1}{3}$.

Lemma 10. *For the unique strategy, σ, not using transitions $s_n \to r_n,\ n \geq 1$, we have $\mathbb{P}^{\sigma}_{s_i}[W] = 2^{-i}$.*

Proof. Clearly $\mathbb{P}^{\sigma}_{s_0}[W] = 1 = 2^{-0}$. Further, the assignment $x := \mathbb{P}^{\sigma}_{s_1}[W]$ is the least non-negative solution of the equation $x = \frac{2}{3} + \frac{x^2}{3}$, see, e.g., [9, Theorem 3.4] or [10, Theorem 1], which is $\frac{1}{2}$. Solving the recurrence $\mathbb{P}^{\sigma}_{s_i}[W] = \frac{2}{3} \cdot \mathbb{P}^{\sigma}_{s_{i-1}}[W] + \frac{1}{3} \cdot \mathbb{P}^{\sigma}_{s_{i+1}}[W]$, given the initial conditions for $i = 0, 1$, yields $\mathbb{P}^{\sigma}_{s_i}[W] = 2^{-i}$. □

Lemma 11. $Val(s_1) = \frac{3}{4}$.

Proof. First we prove $Val(s_1) \geq \frac{3}{4}$. For any n consider the memoryless strategy, σ_n, given by $\sigma_n(s_i)(u_i) = 1$ if $i < n$ and $\sigma_n(s_i)(r_i) = 1$ if $i \geq n$. Set $p_i := \mathbb{P}^{\sigma_i}_{s_1}[\text{Reach } s_i]$. Observe that p_i does not change if we define it using any σ_n with $n \geq i$, and that $1 - p_i = \mathbb{P}^{\sigma_n}_{s_1}[W \wedge \neg\text{Reach } s_i]$ for $n \geq i$. Moreover, $p_1 = 1$ and $p_{i+1} := \frac{2}{3} \cdot (p_i + (1 - p_i) \cdot p_{i+1})$. This uniquely determines that $p_i = \frac{2^{i-1}}{2^i - 1}$. Finally, observe that $\mathbb{P}^{\sigma_n}_{s_1}[W] = (1 - p_n) + p_n \cdot \frac{1}{2}$, thus $Val(s_1) \geq \lim_{n\to\infty}(1 - p_n) + p_n \cdot \frac{1}{2} = \frac{3}{4}$.

Now we prove that $Val(s_1) \leq \frac{3}{4}$ by proving $\mathbb{P}^{\sigma}_{s_1}[W] \leq \frac{3}{4}$ for all σ. Consider the following probabilities: $p_a := \mathbb{P}^{\sigma}_{s_1}[W \wedge \neg\text{Reach some } r_j]$, $p_b := \mathbb{P}^{\sigma}_{s_1}[W \wedge \text{Reach some } r_j]$, $p_c := \mathbb{P}^{\sigma}_{s_1}[\text{Reach some } r_j]$. Clearly $p_b = \frac{p_c}{2}$. Due to Lemma 10 applied to $i = 1$ we also have that $p_a \leq \frac{1}{2}$. Finally, $p_a + p_c \leq 1$ since the events are disjoint. We conclude that $\mathbb{P}^{\sigma}_{s_1}[W] = p_a + p_b \leq p_a + \frac{1}{2} \cdot (1 - p_a) = \frac{1}{2} \cdot p_a + \frac{1}{2} \leq \frac{3}{4}$. □

Lemma 12. $Val(s_i) = \frac{2^i+1}{2^{i+1}}$ *for all* $i \geq 0$.

Proof. The case $i = 0$ is trivial, and $i = 1$ is Lemma 11. Solving the recurrence $Val(s_i) = \frac{2}{3} \cdot Val(s_{i-1}) + \frac{1}{3} \cdot Val(s_{i+1})$, given the initial conditions for $i = 0, 1$, yields $Val(s_i) = \frac{2^i+1}{2^{i+1}}$. □

In particular, for all $i \geq 1$, $Val(s_i) > Val(r_i)$, thus no optimal strategy may use transitions $s_n \to r_n,\ n \geq 1$. By Lemma 10, there are no optimal strategies in s_i.

References

[1] N. Berger, N. Kapur, L. J. Schulman & V. Vazirani (2008): *Solvency Games*. In: *Proc. of FSTTCS'08*. doi:10.4230/LIPIcs.FSTTCS.2008.1741.

[2] T. Brázdil, V. Brožek & K. Etessami (2010): *One-Counter Simple Stochastic Games*. In: *Proc. of FSTTCS'10*. pp. 108–119, doi:10.4230/LIPIcs.FSTTCS.2010.108.

[3] T. Brázdil, V. Brožek, K. Etessami & A. Kučera (2011): *Approximating the Termination Value of One-Counter MDPs and Stochastic Games*. In: *Proc. ICALP 2011*. Accepted, to appear.

[4] T. Brázdil, V. Brožek, K. Etessami, A. Kučera & D. Wojtczak (2010): *One-Counter Markov Decision Processes*. In: *ACM-SIAM SODA*. pp. 863–874. Full tech report: CoRR, abs/0904.2511, 2009. http://arxiv.org/abs/0904.2511.

[5] T. Brázdil, V. Brožek, V. Forejt & A. Kučera (2008): *Reachability in Recursive Markov Decision Processes*. *Inform. and Comput.* 206(5), pp. 520–537, doi:10.1016/j.ic.2007.09.002.

[6] T. Brázdil, V. Brožek, A. Kučera & J. Obdržálek (2009): *Qualitative Reachability in Stochastic BPA Games*. In: *Proc. 26th STACS*. pp. 207–218, doi:10.4230/LIPIcs.STACS.2009.1837.

[7] T. Brázdil, V. Brožek, A. Kučera & J. Obdržálek (2011): *Qualitative Reachability in Stochastic BPA Games*. *Inform. and Comput.* doi:10.1016/j.ic.2011.02.002. Accepted, to appear soon. Preliminary version available at http://arxiv.org/abs/1003.0118.

[8] A. Condon (1992): *The Complexity of Stochastic Games*. *Inform. and Comput.* 96, pp. 203–224.

[9] J. Esparza, A. Kučera & R. Mayr (2004): *Model-Checking Probabilistic Pushdown Automata*. In: *Proc. of 19th LICS'04*. IEEE, pp. 12–21.

[10] K. Etessami & M. Yannakakis (2005): *Recursive Markov Chains, Stochastic Grammars, and Monotone Systems of Nonlinear Equations*. In: *Proc. of 22nd STACS'05*. Springer, pp. 340–352.

[11] H. Gimbert & F. Horn (2010): *Solving Simple Stochastic Tail Games*. In: *ACM-SIAM Symposium on Discrete Algorithms (SODA10)*. pp. 847–862.

[12] D. A. Martin (1998): *The Determinacy of Blackwell Games*. *The Journal of Symbolic Logic* 63(4), pp. 1565–1581.

[13] M. L. Puterman (1994): *Markov Decision Processes*. J. Wiley and Sons.