Qualitative Reachability in Stochastic BPA Games

Tomas Brazdil, Vaclav Brozek, Antonin Kucera, Jan Obdrzalek

Symposium on Theoretical Aspects of Computer Science 2009 (Freiburg), pp. 207–218
www.stacs-conf.org

QUALITATIVE REACHABILITY IN STOCHASTIC BPA GAMES

TOMÁŠ BRÁZDIL [1] AND VÁCLAV BROŽEK [2] AND ANTONÍN KUČERA [2] AND JAN OBDRŽÁLEK [2]

[1] Institut für Informatik, TU München, Boltzmannstr. 3, 85748 Munich, Germany.
E-mail address: brazdil@fi.muni.cz
URL: http://www7.in.tum.de/

[2] Faculty of Informatics, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic.
E-mail address: brozek@fi.muni.cz
E-mail address: kucera@fi.muni.cz
E-mail address: obdrzalek@fi.muni.cz
URL: http://www.fi.muni.cz/

Abstract. We consider a class of infinite-state stochastic games generated by stateless pushdown automata (or, equivalently, 1-exit recursive state machines), where the winning objective is specified by a regular set of target configurations and a qualitative probability constraint '>0' or '=1'. The goal of one player is to maximize the probability of reaching the target set so that the constraint is satisfied, while the other player aims at the opposite. We show that the winner in such games can be determined in **NP** ∩ **co-NP**. Further, we prove that the winning regions for both players are regular, and we design algorithms which compute the associated finite-state automata. Finally, we show that winning strategies can be synthesized effectively.

1. Introduction

Stochastic games are a formal model for discrete systems where the behavior in each state is either controllable, adversarial, or stochastic. Formally, a stochastic game is a directed graph G with a denumerable set of vertices V which are split into three disjoint subsets $V_\Box$, $V_\Diamond$, and $V_\bigcirc$. For every $v \in V_\bigcirc$, there is a fixed probability distribution over the outgoing edges of v. We also require that the set of outgoing edges of every vertex is nonempty. The game is initiated by putting a token on some vertex. The token is then moved from vertex to vertex by two players, $\Box$ and $\Diamond$, who choose the next move in the vertices of $V_\Box$ and $V_\Diamond$, respectively. In the vertices of $V_\bigcirc$, the outgoing edges are chosen according to the associated fixed probability distribution. A *quantitative winning objective* is specified by some Borel set W of infinite paths in G and a probability constraint $\rhd\varrho$, where $\rhd \in \{>, \geq\}$ is a comparison and $\varrho \in [0,1]$. An important subclass of quantitative winning objectives are *qualitative winning objectives* where the constant ϱ must be either 0 or 1. The goal of player $\Box$ is

1998 ACM Subject Classification: G.3 PROBABILITY AND STATISTICS, F.1 COMPUTATION BY ABSTRACT DEVICES.
Key words and phrases: Stochastic games, reachability, pushdown automata.
All authors are supported by the research center Institute for Theoretical Computer Science (ITI), project No. 1M0545. T. Brázdil is also supported by the Alexander von Humboldt Foundation, and J. Obdržálek by grant GAČR 201/08/0308.

to maximize the probability of all runs that stay in W so that it is $\triangleright$-related to ϱ, while player $\Diamond$ aims at the opposite. A *strategy* specifies how a player should play. In general, a strategy may or may not depend on the history of a play (we say that a strategy is *history-dependent (H)* or *memoryless (M)*), and the edges may be chosen deterministically or randomly (*deterministic (D)* and *randomized (R)* strategies). In the case of randomized strategies, a player chooses a probability distribution on the set of outgoing edges. Note that deterministic strategies can be seen as restricted randomized strategies, where one of the outgoing edges has probability 1. Each pair of strategies (σ, π) for players $\Box$ and $\Diamond$ determines a *play*, i.e., a unique Markov chain obtained from G by applying the strategies σ and π in the natural way. The *outcome* of a play initiated in v is the probability of all runs initiated in v that are in the set W, denoted $\mathcal{P}_v^{\sigma,\pi}(W)$. We say that a play is $(\triangleright\varrho)$-won by player $\Box$ if its outcome is $\triangleright$-related to ϱ; otherwise, the play is $(\triangleright\varrho)$-won by player $\Diamond$. A strategy of player $\Box$ (or player $\Diamond$) is $(\triangleright\varrho)$-*winning* if for every strategy of the other player, the corresponding play is $(\triangleright\varrho)$-won by player $\Box$ (or by player $\Diamond$, respectively). A natural question is whether one of the two players always has a $(\triangleright\varrho)$-winning strategy, i.e., whether the game is *determined*. The answer is somewhat subtle. A celebrated result of Martin [18] (see also [17]) implies that stochastic games with Borel winning conditions are *weakly determined*, i.e., each vertex v has a *value* given by

$$val(v) \quad = \quad \sup_{\sigma} \inf_{\pi} \mathcal{P}_v^{\sigma,\pi}(W) \quad = \quad \inf_{\pi} \sup_{\sigma} \mathcal{P}_v^{\sigma,\pi}(W) \tag{1.1}$$

Here σ and π ranges over the set of all strategies for player $\Box$ and player $\Diamond$, respectively. However, the players do not necessarily have *optimal* strategies that would guarantee the outcome $val(v)$ or better against every strategy of the opponent. On the other hand, it follows directly from the above equation that each player has an ε-optimal strategy (see Definition 2.3) for every $\varepsilon > 0$. This means that if $\varrho \neq val(v)$, then one of the two players has a $(\triangleright\varrho)$-winning strategy for the game initiated in v. The situation when $\varrho = val(v)$ is more problematic, and to the best of our knowledge, the literature does not yet offer a general answer. Let us also note that for *finite-state* stochastic games and the "usual" classes of quantitative/qualitative Borel objectives (such as Büchi, Rabin, Street, etc.), the determinacy follows from the existence of optimal strategies (hence, the sup and inf in Equation 1.1 can be safely replaced with max and min, respectively). For classes of infinite-state stochastic games (such as stochastic BPA games considered in this paper), optimal strategies do not necessarily exist and the associated determinacy results must be proven by other methods.

Algorithmic issues for stochastic games with quantitative/qualitative winning objectives have been studied mainly for finite-state stochastic games. A lot of attention has been devoted to quantitative *reachability objectives*, even in the special case when $\varrho = \frac{1}{2}$. The problem whether player $\Box$ has a $(>\frac{1}{2})$-winning strategy is known to be in **NP** $\cap$ **co-NP**, but its membership to **P** is one of the long-standing open problems in algorithmic game theory [9, 20]. Later, more complicated qualitative/quantitative ω-regular winning objectives (such as Büchi, co-Büchi, Rabin, Street, Muller, etc.) were considered, and the complexity of the corresponding decision problems was analysed. We refer to [10, 6, 8, 7, 21, 19] for more details. As for infinite-state stochastic games, the attention has so far been focused on stochastic games induced by lossy channel systems [1, 2] and by pushdown automata (or, equivalently, recursive state machines) [14, 15, 13, 12, 4]. In the next paragraphs, we discuss the latter model in greater detail because these results are closely related to the results presented in this paper.

A *pushdown automaton (PDA)* (see, e.g., [16]) is equipped with a finite control unit and an unbounded stack. The dynamics is specified by a finite set of rules of the form $pX \hookrightarrow q\alpha$, where p, q are control states, X is a stack symbol, and α is a (possibly empty) sequence of stack symbols. A rule of the form $pX \hookrightarrow q\alpha$ is applicable to every configuration of the form $pX\beta$ and produces the configuration $q\alpha\beta$. If there are several rules with the same left-hand side, one of them must be chosen, and the choice is appointed to player $\Box$, player $\Diamond$, or it is randomized. Technically, the set of all left-hand sides (i.e., pairs of the form pX) is split into three disjoint subsets $H_\Box$, $H_\Diamond$, and $H_\bigcirc$, and for all $pX \in H_\bigcirc$ there is a fixed probability distribution over the set of all rules of the form $pX \hookrightarrow q\alpha$. Thus, each PDA induces the associated infinite-state stochastic game where the vertices are PDA configurations and the edges are determined in the natural way. An important subclass of PDA is obtained by restricting the number of control states to 1. Such PDA are also known as *stateless* PDA or (mainly in concurrency theory) as BPA. PDA and BPA correspond to *recursive state machines (RSM)* and *1-exit RSM* respectively, in the sense that their descriptive powers are equivalent, and there are effective linear-time translations between the corresponding models.

In [13], the quantitative and qualitative *termination objective* for PDA and BPA stochastic games is examined (a terminating run is a run which hits a configuration with the empty stack; hence, termination is a special form of reachability). For BPA, it is shown that the vector of optimal values ($val(X), X \in \Gamma$), where Γ is the stack alphabet, forms the least solution of an effectively constructible system of min-max equations. Moreover, both players have optimal MD strategies which depend only on the top-of-the-stack symbol of a given configuration (such strategies are called SMD, meaning Stackless MD). Hence, stochastic BPA games with quantitative/qualitative termination objectives are determined. Since the least solution of the constructed equational system can be encoded in first order theory of the reals, the existence of a ($\triangleright\varrho$)-winning strategy for player $\Box$ and player $\Diamond$ can be decided in polynomial space. In the same paper [13], the $\Sigma_2^P \cap \Pi_2^P$ upper complexity bound for the subclass of qualitative termination objectives is established. As for PDA games, it is shown that for every fixed $\varepsilon > 0$, the problem to distinguish whether the optimal value $val(pX)$ is equal to 1 or less than ε, is undecidable. The $\Sigma_2^P \cap \Pi_2^P$ upper bound for stochastic BPA games with qualitative termination objectives was improved to **NP** $\cap$ **co-NP** in [15]. In the same paper, it is also shown that the quantitative reachability problem for finite-state stochastic games (see above) is efficiently reducible to the qualitative termination problem for stochastic BPA games. Hence, the **NP** $\cap$ **co-NP** upper bound cannot be improved without a major breakthrough in algorithmic game theory. In the special case of stochastic BPA games where $H_\Diamond = \emptyset$ or $H_\Box = \emptyset$, the qualitative termination problem is shown to be in **P** (observe that if $H_\Diamond = \emptyset$ or $H_\Box = \emptyset$, then a given BPA induces an infinite-state Markov decision process and the goal of the only player is to maximize or minimize the termination probability, respectively). The results for Markov decision processes induced by BPA are generalized to (arbitrary) qualitative *reachability objectives* in [5], retaining the **P** upper complexity bound. In the same paper, it is also noted that the properties of reachability objectives are quite different from the ones of termination (in particular, there is no apparent way how to express the vector of optimal values as a solution of some recursive equational system, and the SMD determinacy result (see above) does not hold).

Our contribution: In this paper, we continue the study initiated in [14, 15, 13, 12, 4] and solve the qualitative reachability problem for unrestricted stochastic BPA games. Thus, we obtain a substantial generalization of the previous results.

We start by resolving the determinacy issue in Section 3, and this part of our work actually applies to arbitrary *finitely branching* stochastic games, where each vertex has only finitely many successors (BPA stochastic games are finitely branching). We show that finitely branching stochastic games with quantitative/qualitative reachability objectives are determined, i.e., in every vertex, one of the two players has a ($\triangleright\varrho$)-wining strategy. This is a consequence of several observations that are specific for reachability objectives and perhaps interesting on their own.

The main results of our paper, presented in Section 4, concern stochastic BPA games with qualitative reachability objectives. In the context of BPA, a reachability objective is specified by a *regular* set T of target configurations. We show that the problem of determining the winner in stochastic BPA games with qualitative reachability objectives is in **NP** $\cap$ **co-NP**. Here we rely on the previously discussed results about qualitative termination [15] and use the corresponding algorithms as "black-box procedures" at appropriate places. We also rely on observations presented in [5] which were used to solve the simpler case with only one player. However, the full (two-player) case brings completely new complications that need to be tackled by new methods and ideas. Many "natural" hypotheses turned out to be incorrect (some of the interesting cases are documented by examples in Section 4). We also show that the sets of all configurations where player $\Box$ and player $\Diamond$ have a $(\rhd\varrho)$-winning strategy (where $\varrho \in \{0,1\}$) is effectively regular and the corresponding finite-state automata are effectively constructible by a deterministic polynomial-time algorithm with **NP** $\cap$ **co-NP** oracle. Finally, we also give an algorithm which *computes* a $(\rhd\varrho)$-winning strategy if it exists. These strategies are randomized and memoryless, and they are also *effectively regular* in the sense that their functionality can effectively be encoded by finite-state automata (see Definition 4.3). Hence, winning strategies in stochastic BPA games with qualitative reachability objectives can be effectively implemented.

Due to space constraints, most of the proofs had to be omitted and can be found in the full version of this paper [3]. In the main body of the paper, we try to sketch the key ideas and provide some intuition behind the presented technical constructions.

2. Basic Definitions

In this paper, the set of all positive integers, non-negative integers, rational numbers, real numbers, and non-negative real numbers are denoted $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{R}^{\geq 0}$, respectively. For every finite or countably infinite set S, the symbol S^* denotes the set of all finite words over S. The length of a given word u is denoted $|u|$, and the individual letters in u are denoted $u(0), \cdots, u(|u|-1)$. The empty word is denoted ε, where $|\varepsilon| = 0$. We also use S^+ to denote the set $S^* \setminus \{\varepsilon\}$. For every finite or countably infinite set M, a binary relation $\rightarrow \subseteq M \times M$ is *total* if for every $m \in M$ there is some $n \in M$ such that $m \rightarrow n$. A *path* in $\mathcal{M} = (M, \rightarrow)$ is a finite or infinite sequence $w = m_0, m_1, \ldots$ such that $m_i \rightarrow m_{i+1}$ for every i. The *length* of a finite path $w = m_0, \ldots, m_i$, denoted *length*(w), is $i+1$. We also use $w(i)$ to denote the element m_i of w, and w_i to denote the path $m_i, m_{i+1}, \ldots$ (by writing $w(i) = m$ or w_i we implicitly impose the condition that *length*$(w) \geq i{+}1$). A given $n \in M$ is *reachable* from a given $m \in M$, written $m \rightarrow^* n$, if there is a finite path from m to n. A *run* is an infinite path. The sets of all finite paths and all runs in $\mathcal{M}$ are denoted *FPath*$(\mathcal{M})$ and *Run*$(\mathcal{M})$, respectively. Similarly, the sets of all finite paths and runs that start in a given $m \in M$ are denoted *FPath*$(\mathcal{M}, m)$ and *Run*$(\mathcal{M}, m)$, respectively.

Now we recall basic notions of probability theory. Let A be a finite or countably infinite set. A *probability distribution* on A is a function $f : A \rightarrow \mathbb{R}^{\geq 0}$ such that $\sum_{a \in A} f(a) = 1$. A distribution f is *rational* if $f(a) \in \mathbb{Q}$ for every $a \in A$, *positive* if $f(a) > 0$ for every $a \in A$, and *Dirac* if $f(a) = 1$ for some $a \in A$. The set of all distributions on A is denoted $\mathcal{D}(A)$.

A *σ-field* over a set X is a set $\mathcal{F} \subseteq 2^X$ that includes X and is closed under complement and countable union. A *measurable space* is a pair $(X, \mathcal{F})$ where X is a set called *sample space* and $\mathcal{F}$ is a σ-field over X. A *probability measure* over a measurable space $(X, \mathcal{F})$ is a function $\mathcal{P} : \mathcal{F} \rightarrow \mathbb{R}^{\geq 0}$ such that, for each countable collection $\{X_i\}_{i \in I}$ of pairwise disjoint elements of $\mathcal{F}$, $\mathcal{P}(\bigcup_{i \in I} X_i) = \sum_{i \in I} \mathcal{P}(X_i)$, and moreover $\mathcal{P}(X) = 1$. A *probability space* is a triple $(X, \mathcal{F}, \mathcal{P})$ where $(X, \mathcal{F})$ is a measurable space and $\mathcal{P}$ is a probability measure over $(X, \mathcal{F})$.

Definition 2.1. A *Markov chain* is a triple $\mathcal{M} = (M, \longrightarrow, Prob)$ where M is a finite or countably infinite set of *states*, $\longrightarrow \subseteq M \times M$ is a total *transition relation*, and *Prob* is a function which to each $s \in M$ assigns a positive probability distribution over the set of its outgoing transitions.

In the rest of this paper, we write $s \xrightarrow{x} t$ whenever $s \longrightarrow t$ and $Prob((s,t)) = x$. Each $w \in FPath(\mathcal{M})$ determines a *basic cylinder* $Run(\mathcal{M}, w)$ which consists of all runs that start with w. To every $s \in M$ we associate the probability space $(Run(\mathcal{M}, s), \mathcal{F}, \mathcal{P})$ where $\mathcal{F}$ is the σ-field generated by all basic cylinders $Run(\mathcal{M}, w)$ where w starts with s, and $\mathcal{P} : \mathcal{F} \to \mathbb{R}^{\geq 0}$ is the unique probability measure such that $\mathcal{P}(Run(\mathcal{M}, w)) = \Pi_{i=0}^{m-1} x_i$ where $w = s_0, \cdots, s_m$ and $s_i \xrightarrow{x_i} s_{i+1}$ for every $0 \leq i < m$ (if $m = 0$, we put $\mathcal{P}(Run(\mathcal{M}, w)) = 1$).

Definition 2.2. A *stochastic game* is a tuple $G = (V, \mapsto, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ where V is a finite or countably infinite set of *vertices*, $\mapsto \subseteq V \times V$ is a total *edge relation*, $(V_\Box, V_\Diamond, V_\bigcirc)$ is a partition of V, and $Prob$ is a *probability assignment* which to each $v \in V_\bigcirc$ assigns a positive probability distribution on the set of its outgoing transitions. We say that G is *finitely branching* if for each $v \in V$ there are only finitely many $u \in V$ such that $v \mapsto u$.

A stochastic game is played by two players, $\Box$ and $\Diamond$, who select the moves in the vertices of $V_\Box$ and $V_\Diamond$, respectively. Let $\odot \in \{\Box, \Diamond\}$. A *strategy* for player $\odot$ is a function which to each $wv \in V^* V_\odot$ assigns a probability distribution on the set of outgoing edges of v. The set of all strategies for player $\Box$ and player $\Diamond$ is denoted Σ and Π, respectively. We say that a strategy τ is *memoryless (M)* if $\tau(wv)$ depends just on the last vertex v, and *deterministic (D)* if $\tau(wv)$ is a Dirac distribution for all wv. Strategies that are not necessarily memoryless are called *history-dependent (H)*, and strategies that are not necessarily deterministic are called *randomized (R)*. Hence, we can define the following four classes of strategies: MD, MR, HD, and HR, where MD $\subseteq$ HD $\subseteq$ HR and MD $\subseteq$ MR $\subseteq$ HR, but MR and HD are incomparable.

Each pair of strategies $(\sigma, \pi) \in \Sigma \times \Pi$ determines a unique *play* of the game G, which is a Markov chain $G(\sigma, \pi)$ where V^+ is the set of states, and $wu \xrightarrow{x} wuu'$ iff $u \mapsto u'$ and one of the following conditions holds:

- $u \in V_\Box$ and $\sigma(wu)$ assigns x to $u \mapsto u'$, where $x > 0$;
- $u \in V_\Diamond$ and $\pi(wu)$ assigns x to $u \mapsto u'$, where $x > 0$;
- $u \in V_\bigcirc$ and $u \overset{x}{\mapsto} u'$.

Let $T \subseteq V$ be a set of *target* vertices. For each pair of strategies $(\sigma, \pi) \in \Sigma \times \Pi$ and every $v \in V$, let $\mathcal{P}_v^{\sigma,\pi}(Reach(T))$ be the probability of all $w \in Run(G(\sigma, \pi), v)$ such that w visits some $u \in T$ (technically, this means that $w(i) \in V^* T$ for some $i \in \mathbb{N}_0$). We say that a given $v \in V$ *has a value* if $\sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}_v^{\sigma,\pi}(Reach(T)) = \inf_{\pi\in\Pi} \sup_{\sigma\in\Sigma} \mathcal{P}_v^{\sigma,\pi}(Reach(T))$. If v has a value, then $val(v)$ denotes the *value of* v defined by this equality. Since the set of all runs that visit a vertex of T is obviously Borel, we can apply the powerful result of Martin [18] (see also Theorem 3.3) and conclude that *every* $v \in V$ has a value.

Definition 2.3. Let $\varepsilon \geq 0$. We say that

- $\sigma \in \Sigma$ is *ε-optimal* (or *ε-optimal maximizing*) if $\mathcal{P}_v^{\sigma,\pi}(Reach(T)) \geq val(v) - \varepsilon$ for all $\pi \in \Pi$;
- $\pi \in \Pi$ is *ε-optimal* (or *ε-optimal minimizing*) if $\mathcal{P}_v^{\sigma,\pi}(Reach(T)) \leq val(v) + \varepsilon$ for all $\sigma \in \Sigma$.

A 0-optimal strategy is called *optimal*. A *(quantitative) reachability objective* is a pair $(T, \rhd\varrho)$ where $T \subseteq V$ and $\rhd\varrho$ is a probability constraint, i.e., $\rhd \in \{>, \geq\}$ and $\varrho \in [0, 1]$. If $\varrho \in \{0, 1\}$, then the objective is *qualitative*. We say that

- $\sigma \in \Sigma$ is *$(\rhd\varrho)$-winning* if $\mathcal{P}_v^{\sigma,\pi}(Reach(T)) \rhd \varrho$ for all $\pi \in \Pi$;
- $\pi \in \Pi$ is *$(\rhd\varrho)$-winning* if $\mathcal{P}_v^{\sigma,\pi}(Reach(T)) \not\rhd \varrho$ for all $\sigma \in \Sigma$.

3. The Determinacy of Stochastic Games with Reachability Objectives

In this section we show that finitely-branching stochastic games with quantitative/qualitative reachability objectives are *determined* in the sense that for every quantitative reachability objective $(T, \rhd\varrho)$ and every vertex v of a finitely branching stochastic game, one of the two players has a $(\rhd\varrho)$-winning strategy.

For the rest of this section, let us fix a finitely branching game $G = (V, \mapsto, (V_\Box, V_\Diamond, V_\bigcirc), \mathit{Prob})$ and a set of target vertices T. Also, for every $n \in \mathbb{N}_0$ and a pair of strategies $(\sigma, \pi) \in \Sigma \times \Pi$, let $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_n(T))$ be the probability of all runs $w \in \mathit{Run}(G(\sigma,\pi), v)$ such that w visits some $u \in T$ in at most n transitions (clearly, $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) = \lim_{n\to\infty} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_n(T))$).

To keep this paper self-contained, we start by giving a simple proof of Martin's weak determinacy result (Equation 1.1) for the special case of finitely-branching games with reachability objectives. For every $v \in V$ and $i \in \mathbb{N}_0$, we define $\mathcal{V}_i(v) \in \mathbb{N}_0$ inductively as follows: $\mathcal{V}_0(v)$ is equal either to 1 or 0, depending on whether $v \in T$ or not, respectively. $\mathcal{V}_{i+1}(v)$ (for $v \notin T$) is equal either to $\max\{\mathcal{V}_i(u) \mid v \mapsto u\}$, $\min\{\mathcal{V}_i(u) \mid v \mapsto u\}$, or $\sum_{v \xrightarrow{x} u} x \cdot \mathcal{V}_i(u)$, depending on whether $v \in V_\Box$, $v \in V_\Diamond$, or $v \in V_\bigcirc$, respectively. (For $v \in T$ we put $\mathcal{V}_{i+1}(v) = 1$.) Further, put $\mathcal{V}(v) = \lim_{i\to\infty} \mathcal{V}_i(v)$ (note that the limit exists because the sequence $\mathcal{V}_0(v), \mathcal{V}_1(v), \ldots$ is non-decreasing and bounded). A straightforward induction on i reveals that

$$\mathcal{V}_i(v) \quad = \quad \max_{\sigma\in\Sigma} \min_{\pi\in\Pi} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_i(T)) \quad = \quad \min_{\pi\in\Pi} \max_{\sigma\in\Sigma} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_i(T))$$

Also observe that, for every $i \in \mathbb{N}_0$, there are fixed HD strategies $\sigma_i \in \Sigma$ and $\pi_i \in \Pi$ such that for every $\pi \in \Pi$ and $\sigma \in \Sigma$ we have that $\mathcal{P}_v^{\sigma,\pi_i}(\mathit{Reach}_i(T)) \leq \mathcal{V}_i(v) \leq \mathcal{P}_v^{\sigma_i,\pi}(\mathit{Reach}_i(T))$.

Theorem 3.1. *Every $v \in V$ has a value and $\mathit{val}(v) = \mathcal{V}(v)$.*

Proof. One can easily verify that

$$\mathcal{V}(v) \quad \leq \quad \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \quad \leq \quad \inf_{\pi\in\Pi} \sup_{\sigma\in\Sigma} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \tag{3.1}$$

Hence, it suffices to show that, for every $v \in V$, player $\Diamond$ has a $(>\mathcal{V}(v))$-winning HD strategy $\bar{\pi}$ in v.

For every $i \in \mathbb{N}$, let W_i be the set of all $w \in V^* V_\Diamond$ such that $w(0) = v$, $\mathit{length}(w) = i$, and $w(i) \mapsto w(i{+}1)$ for every $0 \leq i < \mathit{length}(w)$. The strategy $\bar{\pi}$ is defined inductively, together with an auxiliary set $\Pi_i \subseteq \Pi$. We start by putting $\Pi_1 = \{\pi_i \mid i \in \mathbb{N}_0\}$. Now assume that Π_i has already been defined. For every $wu \in W_i$, let us fix an edge $u \mapsto u'$ such that $\pi(wu)(u \mapsto u') = 1$ for infinitely many $\pi \in \Pi_i$ (observe that there must be such an edge because G is finitely branching). We put $\bar{\pi}(wu)(u \mapsto u') = 1$ and $\Pi_{i+1} = \{\pi \in \Pi_i \mid \pi(wu)(u \mapsto u') = 1\}$.

We claim that for every $\sigma \in \Sigma$ we have that $\mathcal{P}_v^{\sigma,\bar{\pi}}(\mathit{Reach}(T)) \leq \mathcal{V}(v)$. Assume the opposite. Then there is $\bar{\sigma} \in \Sigma$ such that $\mathcal{P}_v^{\bar{\sigma},\bar{\pi}}(\mathit{Reach}(T)) = \varrho > \mathcal{V}(v)$. Further, there is some $k \in \mathbb{N}$ such that $\mathcal{P}_v^{\bar{\sigma},\bar{\pi}}(\mathit{Reach}_k(T)) > \mathcal{V}(v) + (\varrho - \mathcal{V}(v))/2$. It follows directly from the definition of $\bar{\pi}$ that there is some $m \in \mathbb{N}, m > k$ such that $\pi_m \in \Pi_m$ and $\bar{\pi}(w) = \pi_m(w)$ for every $w \in W_m$. Hence, $\mathcal{P}_v^{\bar{\sigma},\pi_m}(\mathit{Reach}_m(T)) > \mathcal{V}(v) + (\varrho - \mathcal{V}(v))/2 > \mathcal{V}(v)$, which contradicts the definition of $\mathcal{V}$. ∎

The characterization of $\mathit{val}(v)$ as a limit of $\mathcal{V}_i(v)$ has the following important consequence:

Lemma 3.2. *For every fixed vertex $v \in V$, we have that*

$$\forall \varepsilon{>}0 \;\; \exists \sigma \in \Sigma \;\; \exists n \in \mathbb{N} \;\; \forall \pi \in \Pi \;:\; \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_n(T)) > \mathit{val}(v) - \varepsilon$$

Proof. It suffices to choose a sufficiently large $n \in \mathbb{N}$ and put $\sigma = \sigma_n$. ∎

Note that from the proof of Theorem 3.1 we obtain a HD strategy $\overline{\pi} \in \Pi$ such that $\forall v \in V$ and $\forall \sigma \in \Sigma$ we have that $\mathcal{P}_v^{\sigma,\overline{\pi}}(Reach(T)) \leq val(v)$. This result can be strengthened to MD strategies.

Theorem 3.3. *There is a MD strategy $\pi \in \Pi$ such that for every $v \in V$ and every $\sigma \in \Sigma$ we have that $\mathcal{P}_v^{\sigma,\pi}(Reach(T)) \leq val(v)$. That is, π is an optimal minimizing strategy in every vertex.*

Theorem 3.4 (Determinacy)**.** *Let $v \in V$ and let $(T, \rhd\varrho)$ be a (quantitative) reachability objective. Then one of the two players has a $(\rhd\varrho)$-winning strategy in v.*

Proof outline. We prove that if player $\Diamond$ does not have a $\rhd\varrho$-winning strategy, then player $\Box$ has a $\rhd\varrho$-winning strategy. That is, we prove the implication

$$\forall \pi \in \Pi \;\; \exists \sigma \in \Sigma \; : \; \mathcal{P}_v^{\sigma,\pi}(Reach(T)) \rhd \varrho \quad \Rightarrow \quad \exists \sigma \in \Sigma \;\; \forall \pi \in \Pi \; : \; \mathcal{P}_v^{\sigma,\pi}(Reach(T)) \rhd \varrho \tag{3.2}$$

If $\rhd$ is $>$ or $val(v) \neq \varrho$, then this follows easily by Theorem 3.3. For the constraint ≥ 0 the statement is trivial. Now suppose that $\rhd$ is $\geq$ and $\varrho = val(v) > 0$, and assume that the left-hand side in (3.2) holds. Observe that we can safely restrict the set of edges available to player $\Box$ to those $u \mapsto u'$ where $val(u') = val(u)$. Using the left-hand side of (3.2), one can show that for every $s \in V$, the value $val(s)$ stays unchanged in the new game obtained by applying this restriction. Due to Lemma 3.2, to every $s \in V$ in the new game we can associate a strategy $\sigma_s \in \Sigma$ and $n_s \in \mathbb{N}$ such that for every $\pi \in \Pi$ we have that $\mathcal{P}_s^{\sigma_s,\pi}(Reach_{n_s}(T)) > val(s)/2$. The $\geq\varrho$-winning strategy σ for player $\Box$ is obtained by "iterating" the strategies σ_s in the following sense: we start with σ_v, and after performing a path w of length n_v, we change the strategy to σ_s where s is the last vertex visited by w. The strategy σ_s is used for the next n_s transition, and then we perform another "iteration". Observe that each round of this "iteration" decreases the probability that T is *not* reached by a factor of $1/2$, independently of the strategy of player $\Diamond$. ∎

4. Qualitative Reachability in Stochastic BPA Games

Stochastic BPA games correspond to stochastic games induced by stateless pushdown automata or 1-exit recursive state machines (see Section 1). A formal definition follows.

Definition 4.1. A *stochastic BPA* game is a tuple $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), Prob)$ where Γ is a finite *stack alphabet*, $\hookrightarrow \subseteq \Gamma \times \Gamma^{\leq 2}$ is a finite set of *rules* (where $\Gamma^{\leq 2} = \{w \in \Gamma^* : |w| \leq 2\}$) such that for each $X \in \Gamma$ there is some rule $X \hookrightarrow \alpha$, $(\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc)$ is a partition of Γ, and $Prob$ is a *probability assignment* which to each $X \in \Gamma_\bigcirc$ assigns a rational positive probability distribution on the set of all rules of the form $X \hookrightarrow \alpha$.

A *configuration* of Δ is a word $\alpha \in \Gamma^*$, which can intuitively be interpreted as the current stack content where the leftmost symbol of α is on top of the stack. Each stochastic BPA game $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), Prob)$ determines a unique stochastic game $G_\Delta = (\Gamma^*, \mapsto, (\Gamma_\Box\Gamma^*, \Gamma_\Diamond\Gamma^*, \Gamma_\bigcirc\Gamma^* \cup \{\varepsilon\}), Prob_\Delta)$ where the transitions of $\mapsto$ are determined as follows: $\varepsilon \mapsto \varepsilon$, and $X\beta \mapsto \alpha\beta$ iff $X \hookrightarrow \alpha$. The probability assignment $Prob_\Delta$ is the natural extension of $Prob$, i.e., $\varepsilon \overset{1}{\mapsto} \varepsilon$ and for all $X \in \Gamma_\bigcirc$ we have that $X\beta \overset{x}{\mapsto} \alpha\beta$ iff $X \overset{x}{\hookrightarrow} \alpha$.

In this section we consider stochastic BPA games with qualitative termination objectives $(T, \rhd\varrho)$ where $T \subseteq \Gamma^*$ is a *regular* set of configurations. For technical convenience, we define the size of T as the size of the minimal deterministic finite-state automaton $\mathscr{A}_T = (Q, q_0, \delta, F)$ which recognizes the *reverse* of T (if we view configurations as stacks, this corresponds to bottom-up direction). Note that the automaton $\mathscr{A}_T$ can be simulated on-the-fly in Δ by employing standard techniques (see, e.g., [11]). That is, the stack alphabet is extended to $\Gamma \times Q$ and the rules are adjusted accordingly (for example, if $X \hookrightarrow YZ$, then for every $q \in Q$ the extended BPA game has a rule $(X,q) \hookrightarrow (Y,r)(Z,q)$ where $\delta(q,Z) = r$). Note that the on-the-fly simulation of $\mathscr{A}_T$ in Δ does not affect the way how the game is played, and the size of the extended game in polynomial in $|\Delta|$ and $|\mathscr{A}_T|$. The main advantage of this simulation is that the information whether a current configuration belongs to T or not can now be deduced just by looking at the symbol on top of the stack. This leads to an important technical simplification in the definition of T:

Definition 4.2. We say that $T \subseteq \Gamma^*$ is *simple* if $\varepsilon \notin T$ and there is $\Gamma_T \subseteq \Gamma$ such that for every $X\alpha \in \Gamma^+$ we have that $X\alpha \in T$ iff $X \in \Gamma_T$.

Note that the requirement $\varepsilon \notin T$ in the previous definition is not truly restrictive, because each BPA can be equipped with a fresh bottom-of-the-stack symbol which cannot be removed. Hence, we can safely restrict ourselves just to simple sets of target configurations. All of the obtained results (including the complexity bounds) are valid also for regular sets of target configurations.

Since stochastic BPA games have infinitely many vertices, even memoryless strategies are not necessarily finitely representable. It turns out that the winning strategies for both players in stochastic BPA games with qualitative reachability objectives are (effectively) *regular* in the following sense:

Definition 4.3. Let $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), Prob)$ be a stochastic BPA game, and let $\odot \in \{\Box, \Diamond\}$. We say that a strategy τ for player $\odot$ is *regular* if there is a deterministic finite-state automaton $\mathscr{A}$ over the alphabet Γ such that, for every $X\alpha \in \Gamma_\odot\Gamma^*$, the value of $\tau(X\alpha)$ depends just on the control state entered by $\mathscr{A}$ after reading the reverse of $X\alpha$ (i.e., the automaton $\mathscr{A}$ reads the stack bottom-up).

For the rest of this section, we fix a stochastic BPA game $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), Prob)$ and a simple set T of target configurations. Since we are interested just in reachability objectives, we can safely assume that for every $R \in \Gamma_T$, the only rule where R appears on the left-hand side is $R \hookrightarrow R$ (this assumption simplifies the formulation of some claims). We use T_ε to denote the set $T \cup \{\varepsilon\}$, and we also slightly abuse the notation by writing ε instead of $\{\varepsilon\}$ at some places (particularly in expressions such as $Reach(\varepsilon)$).

For a given set $C \subseteq \Gamma^*$ and a given qualitative probability constraint $\rhd\varrho$, we use $[C]_\Box^{\rhd\varrho}$ and $[C]_\Diamond^{\rhd\varrho}$ to denote the set of all $\alpha \in \Gamma^*$ from which player $\Box$ and player $\Diamond$ has a ($\rhd\varrho$)-winning strategy in the game Δ with the reachability objective $(C, \rhd\varrho)$, respectively. Observe that $[C]_\Box^{\rhd\varrho} = \Gamma^* \smallsetminus [C]_\Diamond^{\rhd\varrho}$ due to the determinacy results presented in Section 3.

In the forthcoming subsections we examine the sets $[T]_\Box^{\rhd\varrho}$ for the two meaningful qualitative probability constraints >0 and $=1$ (observe that $[T]_\Box^{\geq 0} = \Gamma^*$ and $[T]_\Box^{>1} = \emptyset$). We show that the membership to $[T]_\Box^{>0}$ and $[T]_\Box^{=1}$ is in **P** and **NP** $\cap$ **co-NP**, respectively. The same holds for the sets $[T]_\Diamond^{>0}$ and $[T]_\Diamond^{=1}$, respectively. Further, we show that all of these sets are effectively regular, and that ($\rhd\varrho$)-winning strategies for both players are effectively computable. The associated upper complexity bounds are essentially the same as above.

4.1. The Set $[T]_\Box^{>0}$

We start by observing that the sets $[T]_\Box^{>0}$ and $[T]_\Diamond^{>0}$ are regular, and the associated finite-state automata have a fixed number of control states. A proof of this observation is actually straightforward.

Proposition 4.4. *Let* $\mathscr{A} = [T]_\Box^{>0} \cap \Gamma$ *and* $\mathscr{B} = [T_\varepsilon]_\Box^{>0} \cap \Gamma$. *Then* $[T]_\Box^{>0} = \mathscr{B}^*\mathscr{A}\Gamma^*$ *and* $[T_\varepsilon]_\Box^{>0} = \mathscr{B}^*\mathscr{A}\Gamma^* \cup \mathscr{B}^*$. *Consequently,* $[T]_\Diamond^{>0} = \Gamma^* \smallsetminus [T]_\Box^{>0} = (\mathscr{B} \smallsetminus \mathscr{A})^* \cup (\mathscr{B} \smallsetminus \mathscr{A})^*(\Gamma \smallsetminus \mathscr{B})\Gamma^*$.

Our next proposition says how to compute the sets $\mathscr{A}$ and $\mathscr{B}$.

Proposition 4.5. *The pair $(\mathscr{A}, \mathscr{B})$ is the least fixed-point of the function $F : (2^\Gamma \times 2^\Gamma) \to (2^\Gamma \times 2^\Gamma)$, where $F(A, B) = (\hat{A}, \hat{B})$ is defined as follows:*

$$\begin{array}{rcl} \hat{A} & = & \Gamma_T \cup A \cup \{X \in \Gamma_\Box \cup \Gamma_\bigcirc \mid \textit{there is } X \hookrightarrow \beta \textit{ such that } \beta \in B^* A \Gamma^*\} \\ & \cup & \{X \in \Gamma_\Diamond \mid \textit{for all } X \hookrightarrow \beta \textit{ we have that } \beta \in B^* A \Gamma^*\} \\ \hat{B} & = & \Gamma_T \cup B \cup \{X \in \Gamma_\Box \cup \Gamma_\bigcirc \mid \textit{there is } X \hookrightarrow \beta \textit{ such that } \beta \in B^* A \Gamma^* \cup B^*\} \\ & \cup & \{X \in \Gamma_\Diamond \mid \textit{for all } X \hookrightarrow \beta \textit{ we have that } \beta \in B^* A \Gamma^* \cup B^*\} \end{array}$$

Since the least fixed-point of the function F defined in Proposition 4.5 is computable in polynomial time, the finite-state automata recognizing the sets $[T]_\Box^{>0}$ and $[T]_\Diamond^{>0}$ are computable in polynomial time. Thus, we obtain the following theorem:

Theorem 4.6. *The membership to $[T]_\Box^{>0}$ and $[T]_\Diamond^{>0}$ is decidable in polynomial time. Both sets are effectively regular, and the associated finite-state automata are constructible in polynomial time. Further, there are regular strategies $\sigma \in \Sigma$ and $\pi \in \Pi$ constructible in polynomial time that are* (>0)*-winning in every configuration of $[T]_\Box^{>0}$ and $[T]_\Diamond^{>0}$, respectively.*

4.2. The Set $[T]_\Box^{=1}$

The results presented in this subsection constitute the very core of this paper. The problems are more complicated than in the case of $[T]_\Box^{>0}$, and several deep observations are needed to tackle them. We start by showing that the sets $[T]_\Box^{=1}$ and $[T]_\Diamond^{=1}$ are regular.

Proposition 4.7. *Let $\mathscr{A} = [T_\varepsilon]_\Diamond^{=1} \cap \Gamma$, $\mathscr{B} = [T_\varepsilon]_\Box^{=1} \cap [T]_\Diamond^{=1} \cap \Gamma$, $\mathscr{C} = [T]_\Box^{=1} \cap \Gamma$. Then $[T]_\Box^{=1} = \mathscr{B}^* \mathscr{C} \Gamma^*$ and $[T]_\Diamond^{=1} = \mathscr{B}^* \mathscr{A} \Gamma^* \cup \mathscr{B}^*$.*

Proposition 4.7 can be proven by a straightforward induction on the length of configurations. Observe that if there is an algorithm which computes the set $\mathscr{A} = [T_\varepsilon]_\Diamond^{=1} \cap \Gamma$ for an arbitrary stochastic BPA game, then this algorithm can also be used to compute the set $[T]_\Diamond^{=1} \cap \Gamma$ (this is because $X \in [T]_\Diamond^{=1}$ iff $\hat{X} \in [\hat{T}_\varepsilon]_\Diamond^{=1}$, where $[\hat{T}_\varepsilon]_\Diamond^{=1}$ is considered in a stochastic BPA game $\hat{\Delta}$ obtained from Δ by adding two fresh stochastic symbols $\hat{X}, Z$ together with the rules $\hat{X} \overset{1}{\hookrightarrow} XZ$, $Z \overset{1}{\hookrightarrow} Z$, and setting $\hat{T} = T$). Due to Theorem 3.4, we have that $\mathscr{C} = \Gamma \setminus ([T]_\Diamond^{=1} \cap \Gamma)$, and thus we can compute also the set $\mathscr{C}$. Since $\mathscr{B} = \Gamma \setminus (\mathscr{A} \cup \mathscr{C})$ (again by Theorem 3.4), we can also compute the set $\mathscr{B}$. Hence, the core of the problem is to design an algorithm which computes the set $\mathscr{A}$.

In the next definition we introduce the crucial notion of a *terminal* set of stack symbols, which plays a key role in our considerations.

Definition 4.8. A set $M \subseteq \Gamma$ is *terminal* if the following conditions are satisfied:

- $\Gamma_T \cap M = \emptyset$;
- for every $Z \in M \cap (\Gamma_\Box \cup \Gamma_\bigcirc)$ and every rule of the form $Z \hookrightarrow \alpha$ we have that $\alpha \in M^*$;
- for every $Z \in M \cap \Gamma_\Diamond$ there is a rule $Z \hookrightarrow \alpha$ such that $\alpha \in M^*$.

Since $\emptyset$ is terminal and the union of two terminal sets is terminal, there is the greatest terminal set that will be denoted C in the rest of this section. Also note that C determines a unique stochastic BPA game Δ_C obtained from Δ by restricting the set of stack symbols to C and including all rules $X \hookrightarrow \alpha$ where $X, \alpha \in C^*$. The set of rules of Δ_C is denoted $\hookrightarrow_C$. The probability of stochastic rules in Δ_C is the same as in Δ.

Definition 4.9. A stack symbol $Y \in \Gamma$ is a *witness* if one of the following conditions is satisfied:

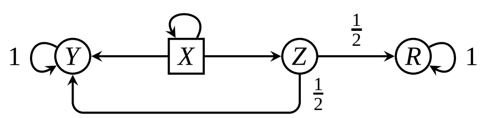
(1) $Y \in [T_\varepsilon]^{>0}_{\Diamond}$;
(2) $Y \in C$ and $Y \in [\varepsilon]^{=1}_{\Diamond}$, where the set $[\varepsilon]^{=1}_{\Diamond}$ is computed in Δ_C.

The set of all witnesses is denoted W.

Observe that the problem whether $Y \in W$ for a given $Y \in \Gamma$ is decidable in **NP**$\cap$**co-NP**, because Condition (1) is decidable in **P** due to Theorem 4.6, the set C is computable in polynomial time, and the membership to $[\varepsilon]^{=1}_{\Diamond}$ is in **NP** $\cap$ **co-NP** due to [15] (this is the only place where we use the decision algorithm for qualitative termination designed in [15]).

Obviously, $W \subseteq \mathscr{A}$. One may be tempted to think that the set $\mathscr{A}$ is just the *attractor* of W, denoted $Att(W)$, which consists of all $V \in \Gamma$ from which player $\Diamond$ can enforce visiting a witness with a positive probability (i.e., $V \in Att(W)$ iff $\exists \pi \in \Pi \;\; \forall \sigma \in \Sigma \; : \; \mathcal{P}^{\sigma,\pi}_V(Reach(W\Gamma^*)) > 0$). However, this is not true, as it is demonstrated in the following example:

Example 4.10. Consider a stochastic BPA game $\hat{\Delta} = (\{X, Y, Z, R\}, \hookrightarrow, (\{X\}, \emptyset, \{Y, Z, R\}), Prob)$, where $X \hookrightarrow X$, $X \hookrightarrow Y$, $X \hookrightarrow Z$, $Y \overset{1}{\hookrightarrow} Y$, $Z \overset{1/2}{\hookrightarrow} Y$, $Z \overset{1/2}{\hookrightarrow} R$, $R \overset{1}{\hookrightarrow} R$, and the set T_Γ contains just R. The game is initiated in X, and the relevant part of $G_{\hat{\Delta}}$ (reachable from X) is shown in the following figure:



Observe that $\mathscr{A} = \{X, Y, Z\}$, $C = W = \{Y\}$, but $Att(\{Y\}) = \{Z, Y\}$.

In Example 4.10, the problem is that player $\Box$ can use a strategy which always selects the rule $X \hookrightarrow X$ with probability one, and player $\Diamond$ has no way to influence this. Nevertheless, observe that player $\Box$ has essentially two options: he either enters a symbol of $Att(\{Y\})$, or he performs the loop $X \hookrightarrow X$ forever. The second possibility can be analyzed by "cutting off" the set $Att(\{Y\})$ and recomputing the set of all witnesses together with its attractor in the resulting stochastic BPA game, which contains only X and the rule $X \hookrightarrow X$. Observe that X is a witness in this game, and hence it can be safely added to the set $\mathscr{A}$. Thus, the computation of the set $\mathscr{A}$ for the stochastic BPA game $\hat{\Delta}$ is completed.

For general stochastic BPA games, the algorithm for computing the set $\mathscr{A}$ proceeds by initiating $\mathscr{A}$ to $\emptyset$ and then repeatedly computing the set $Att(W)$, setting $\mathscr{A} := \mathscr{A} \cup Att(W)$, and "cutting off" the set $Att(W)$ from the game. This goes on until the game or the set $Att(W)$ becomes empty. The way how $Att(W)$ is "cut off" from the current game is described below. First, let us present an important (and highly non-trivial) result which states the following:

Proposition 4.11. *If $\mathscr{A} \neq \emptyset$, then $W \neq \emptyset$.*

Proof outline. We show that if $W = \emptyset$, then there is a MR strategy $\sigma \in \Sigma$ such that for every $X \in \Gamma$ and every $\pi \in \Pi$ we have that $\mathcal{P}^{\sigma,\pi}_X(Reach(T_\varepsilon)) = 1$. In particular, this means that $\mathscr{A} = \emptyset$.

Since $W = \emptyset$, the condition of Definition 4.9 does not hold for any $Y \in \Gamma$, which in particular means that for all $Y \in C$ we have that $Y \notin [\varepsilon]^{=1}_{\Diamond}$, i.e., $Y \in [\varepsilon]^{=1}_{\Box}$ by Theorem 3.4 (here, the sets $[\varepsilon]^{=1}_{\Diamond}$ and $[\varepsilon]^{=1}_{\Box}$ are considered in the game Δ_C). Due to [13], there exists a SMD strategy σ_T for player $\Box$ in Δ_C such that for every $Y \in C$ and every strategy π of player $\Diamond$ in Δ_C we have that $\mathcal{P}^{\sigma_T,\pi}(Reach(\varepsilon)) = 1$. Now we define the promised MR strategy $\sigma \in \Sigma$ as follows: for a given $X\alpha \in \Gamma_\Box\Gamma^*$, we put $\sigma(X\alpha) = \sigma_T(X\alpha)$ if $X\alpha$ starts with some $\beta \in C^*$ where $|\beta| > |\Delta|$. Otherwise, $\sigma(X\alpha)$ returns the uniform probability distribution over the outgoing transitions of $X\alpha$.

Now, let us fix some strategy $\pi \in \Pi$. Our goal is to show that $\mathcal{P}^{\sigma,\pi}_X(Reach(T_\varepsilon)) = 1$. By analyzing the play $G_\Delta(\sigma,\pi)$, one can show that there is a set of runs $V \subseteq Run(G_\Delta(\sigma,\pi),X)$ and a set of rules $\hookrightarrow_V \subseteq \hookrightarrow$ such that

(A) $\mathcal{P}(V) > 0$, $\hookrightarrow_V \subseteq \hookrightarrow_C$, and for every $w \in V$ we have that w does not visit T_ε and the set of rules that are used infinitely often in w is exactly $\hookrightarrow_V$.

Observe that each $w \in V$ has a finite prefix v_w such that the rules of $\hookrightarrow \smallsetminus \hookrightarrow_C$ are used only in v_w. Further, we can partition the runs of V into countably many sets according to this prefix. One of these sets must have a positive probability, and hence we can conclude that there is $U \subseteq V$ and a finite path $v \in FPath(X)$ such that

(B) $\mathcal{P}(U) > 0$, and each $w \in U$ satisfies the following: w starts with v, the rules of $\hookrightarrow \smallsetminus \hookrightarrow_C$ are used only in the prefix v of w, and the length of every configuration of w visited after the prefix v is at least as large as the length of the last configuration in the prefix v (the last condition still requires a justification which is omitted in here).

We show that $\mathcal{P}(U) = 0$, which is a contradiction. Roughly speaking, this is achieved by observing that, after performing the prefix v, the strategies σ and π can be "simulated" by strategies σ' and π' in the game G_{Δ_C} so that the set of runs U is "projected" onto the set of runs U' in the play $G_{\Delta_C}(\sigma',\pi')$ where $\mathcal{P}(U) = \mathcal{P}(U')$. Then, it is shown that $\mathcal{P}(U') = 0$. This is because the strategy σ' is "sufficiently similar" to the strategy σ_T (see above), and hence the probability of visiting ε in $G_{\Delta_C}(\sigma',\pi')$ is 1. From this we get $\mathcal{P}(U') = 0$, because U' consists only of infinite runs, which cannot visit ε. The arguments are subtle and rely on several auxiliary technical observations. ■

In other words, the non-emptiness of $\mathscr{A}$ is always certified by at least one witness of W, and hence each stochastic BPA game with a non-empty $\mathscr{A}$ can be made smaller by "cutting off" $Att(W)$.

The procedure which "cuts off" the symbols $Att(W)$ is not completely trivial. A naive idea of removing the symbols of $Att(W)$ together with the rules where they appear (this was used for the stochastic BPA game of Example 4.10) does not always work. This is illustrated in the following example:

Example 4.12. Consider a stochastic BPA game $\hat{\Delta} = (\{X,Y,Z,R\}, \hookrightarrow, (\{X\},\emptyset,\{Y,Z,R\}), Prob)$, where $X \hookrightarrow X$, $X \hookrightarrow Y$, $X \hookrightarrow ZY$, $Y \overset{1}{\hookrightarrow} Y$, $Z \overset{1/2}{\hookrightarrow} X$, $Z \overset{1/2}{\hookrightarrow} R$, $R \overset{1}{\hookrightarrow} R$, and $T_{\hat{\Gamma}} = \{R\}$. The game is initiated in X. We have that $\mathscr{A} = \{Y\}$ (observe that $X,Z,R \in [T_\varepsilon]^{=1}_{\Box}$, because the strategy σ of player $\Box$ which always selects the rule $X \hookrightarrow ZY$ is (=1)-winning). We have that $C = W = Att(W) = \{Y\}$. If we remove Y together with all rules where Y appears, we obtain the game $\Delta' = (\{X,Z,R\}, \hookrightarrow, (\{X\},\emptyset,\{Z,R\}), Prob)$, where $X \hookrightarrow X$, $Z \overset{1/2}{\hookrightarrow} X$, $Z \overset{1/2}{\hookrightarrow} R$, $R \overset{1}{\hookrightarrow} R$. In the game Δ', X becomes a witness and hence the algorithm would incorrectly put X into $\mathscr{A}$.

Hence, the "cutting" procedure must be designed more carefully. Intuitively, we do not remove rules of the form $X \hookrightarrow ZY$ where $Y \in Att(W)$, but change them into $X \hookrightarrow Z'Y$, where the symbol Z' behaves like Z but it cannot reach ε. Thus, we obtain the following theorem:

Theorem 4.13. *The membership to $[T]^{=1}_{\Box}$ and $[T]^{=1}_{\Diamond}$ is decidable in **NP** $\cap$ **co-NP**. Both sets are effectively regular, and the associated finite-state automata are constructible by a deterministic polynomial-time algorithm with **NP** $\cap$ **co-NP** oracle. Further, there is a regular strategy $\sigma \in \Sigma$ that is (=1)-winning in every configuration of $[T]^{=1}_{\Box}$. Moreover, the strategy σ is constructible by a deterministic polynomial-time algorithm with **NP** $\cap$ **co-NP** oracle.*

Note that in Theorem 4.13, we do not claim the existence (and constructability) of a regular (=1)-winning strategy π for player $\Diamond$. Actually, such a strategy *does* effectively exist, but we only managed to find a relatively complicated and technical proof which, in our opinion, is of little practical interest (we do not see any natural reason for implementing a strategy which guarantees that the probability of visiting T is strictly less than 1). Hence, this proof is not included in the paper.

5. Conclusions

We have solved the qualitative reachability problem for stochastic BPA games, retaining the same upper complexity bounds that have previously been established for termination [15]. One interesting question which remains unsolved is the decidability of the problem whether $val(\alpha) = 1$ for a given BPA configuration α (we can only decide whether player $\Box$ has a (=1)-winning strategy, which is sufficient but not necessary for $val(\alpha) = 1$). Another open problem is quantitative reachability for stochastic BPA games, where the methods presented in this paper seem insufficient.

References

[1] P. Abdulla, N.B. Henda, L. de Alfaro, R. Mayr, and S. Sandberg. Stochastic games with lossy channels. In *Proceedings of FoSSaCS 2008*, volume 4962 of *LNCS*, pages 35–49. Springer, 2008.
[2] C. Baier, N. Bertrand, and Ph. Schnoebelen. Verifying nondeterministic probabilistic channel systems against ω-regular linear-time properties. *ACM Trans. on Comp. Logic*, 9(1), 2007.
[3] T. Brázdil, V. Brožek, A. Kučera, and J. Obdržálek. Qualitative reachability in stochastic BPA games. Technical report FIMU-RS-2009-01, Faculty of Informatics, Masaryk University, 2009.
[4] T. Brázdil, V. Brožek, V. Forejt, and A. Kučera. Reachability in recursive Markov decision processes. In *Proceedings of CONCUR 2006*, volume 4137 of *LNCS*, pages 358–374. Springer, 2006.
[5] T. Brázdil, V. Brožek, V. Forejt, and A. Kučera. Reachability in recursive Markov decision processes. *I&C*, 206(5):520–537, 2008.
[6] K. Chatterjee, L. de Alfaro, and T. Henzinger. The complexity of stochastic Rabin and Streett games. In *Proceedings of ICALP 2005*, volume 3580 of *LNCS*, pages 878–890. Springer, 2005.
[7] K. Chatterjee, M. Jurdziński, and T. Henzinger. Simple stochastic parity games. In *Proceedings of CSL'93*, volume 832 of *LNCS*, pages 100–113. Springer, 1994.
[8] K. Chatterjee, M. Jurdziński, and T. Henzinger. Quantitative stochastic parity games. In *Proceedings of SODA 2004*, pages 121–130. SIAM, 2004.
[9] A. Condon. The complexity of stochastic games. *I&C*, 96(2):203–224, 1992.
[10] L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. *JCSS*, 68:374–397, 2004.
[11] J. Esparza, A. Kučera, and S. Schwoon. Model-checking LTL with regular valuations for pushdown systems. *I&C*, 186(2):355–376, 2003.
[12] K. Etessami, D. Wojtczak, and M. Yannakakis. Recursive stochastic games with positive rewards. In *Proceedings of ICALP 2008, Part I*, volume 5125 of *LNCS*, pages 711–723. Springer, 2008.
[13] K. Etessami and M. Yannakakis. Recursive Markov decision processes and recursive stochastic games. In *Proceedings of ICALP 2005*, volume 3580 of *LNCS*, pages 891–903. Springer, 2005.
[14] K. Etessami and M. Yannakakis. Efficient qualitative analysis of classes of recursive Markov decision processes and simple stochastic games. In *Proceedings of STACS 2006*, volume 3884 of *LNCS*, pages 634–645. Springer, 2006.
[15] K. Etessami and M. Yannakakis. Recursive concurrent stochastic games. In *Proceedings of ICALP 2006*, volume 4052 of *LNCS*, pages 324–335. Springer, 2006.
[16] J.E. Hopcroft and J.D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, 1979.
[17] A. Maitra and W. Sudderth. Finitely additive stochastic games with Borel measurable payoffs. *Int. Jour. of Game Theory*, 27:257–267, 1998.
[18] D.A. Martin. The determinacy of blackwell games. *Journal of Symbolic Logic*, 63(4):1565–1581, 1998.
[19] W. Thomas. Infinite games and verification. In *Proceedings of CAV 2003*, volume 2725 of *LNCS*, pages 58–64. Springer, 2003.
[20] N. Vieille. Stochastic games: Recent results. *Handbook of Game Theory*, pages 1833–1850, 2002.
[21] I. Walukiewicz. A landscape with games in the background. In *Proceedings of LICS 2004*, pages 356–366. IEEE, 2004.