FI MU

Faculty of Informatics
Masaryk University Brno

Qualitative Reachability in Stochastic BPA Games

by

Tomáš Brázdil
Václav Brožek
Antonín Kučera
Jan Obdržálek

FI MU Report Series **FIMU-RS-2009-01**

Publications in the FI MU Report Series are in general accessible via WWW:

`http://www.fi.muni.cz/reports/`

Further information can be obtained by contacting:

Faculty of Informatics
Masaryk University
Botanická 68a
602 00 Brno
Czech Republic

Qualitative Reachability in Stochastic BPA Games

Tomáš Brázdil Václav Brožek Antonín Kučera Jan Obdržálek
Faculty of Informatics, Masaryk University,
Botanická 68a, 60200 Brno,
Czech Republic.
{brazdil,brozek,kucera,obdrzalek}@fi.muni.cz

May 19, 2009

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** $\cap$ **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 to maximize the probability of all runs that stay in W so that it is $\rhd$-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 $(\rhd\varrho)$-won by player $\Box$ if its outcome is $\rhd$-related to ϱ; otherwise, the play is $(\rhd\varrho)$-won by player $\Diamond$. A strategy of player $\Box$ (or player $\Diamond$) is $(\rhd\varrho)$-*winning* if for every strategy of the other player, the corresponding play is $(\rhd\varrho)$-won by player $\Box$ (or by player $\Diamond$, respectively). A natural question is whether one of the two players always has a $(\rhd\varrho)$-winning strategy, i.e., whether the game is *determined*. The answer is somewhat subtle. A celebrated result of Martin [17] (see also [16]) implies that stochastic games with Borel winning conditions are *weakly determined*, i.e., each vertex v has a *value* given by

$$val(v) \quad = \quad \sup_{\sigma} \inf_{\pi} \mathcal{P}_v^{\sigma,\pi}(W) \quad = \quad \inf_{\pi} \sup_{\sigma} \mathcal{P}_v^{\sigma,\pi}(W) \tag{1}$$

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 $(\rhd\varrho)$-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 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 $\rho = \frac{1}{2}$. The problem whether player $\Box$ has a ($>\frac{1}{2}$)-winning strategy is known to be in $\mathbf{NP} \cap \mathbf{co\text{-}NP}$, but its membership to $\mathbf{P}$ is one of the long-standing open problems in algorithmic game theory [8, 19]. 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 analyzed. We refer to [9, 5, 7, 6, 20, 18] 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) [13, 14, 12, 11, 3]. 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., [15]) 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 [12], 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 $(\triangleright\rho)$-winning strategy for player $\Box$ and player $\Diamond$ can be decided in polynomial space. In the same paper [12], 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 [13]. 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 [4], 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 [13, 14, 12, 11, 3] 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 $(\triangleright\rho)$-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 [13] and use the corresponding algorithms as "black-box procedures" at appropriate places. We also rely on observations presented in [4] 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 $(\triangleright\rho)$-winning strategy (where $\rho \in \{0,1\}$) 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 $(\triangleright\rho)$-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.

To increase readability, most of the proofs of Section 4 have been postponed to appendix. 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 [17] (see also Theorem 3.2) 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* $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \geq \mathit{val}(v) - \varepsilon$ *for all* $\pi \in \Pi$*;*
- $\pi \in \Pi$ *is* ε-optimal *(or* ε-optimal minimizing*) if* $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \leq \mathit{val}(v) + \varepsilon$ *for all* $\sigma \in \Sigma$*.*

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

- $\sigma \in \Sigma$ *is* $(\rhd\rho)$-winning *if* $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \rhd \rho$ *for all* $\pi \in \Pi$*;*
- $\pi \in \Pi$ *is* $(\rhd\rho)$-winning *if* $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \not\rhd \rho$ *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 $(T, \rhd\rho)$ and every vertex v of a finitely branching stochastic game, one of the two players has a $(\rhd\rho)$-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 (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 $\mathcal{V}_i(v) \in \mathbb{N}_0$. For $v \in T$ we put $\mathcal{V}_i(v) = 1$. Otherwise ($v \in V \setminus T$) we define $\mathcal{V}_i(v)$ inductively as follows: $\mathcal{V}_0(v) = 0$; $\mathcal{V}_{i+1}(v)$ is equal either to $\max\{\mathcal{V}_i(u) \mid v \mapsto u\}$, $\min\{\mathcal{V}_i(u) \mid v \mapsto u\}$, or $\sum_{v \overset{x}{\mapsto} u} x \cdot \mathcal{V}_i(u)$, depending on whether $v \in V_\Box$, $v \in V_\Diamond$, or $v \in V_\bigcirc$, respectively.

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 $\mathcal{P}_v^{\sigma,\pi_i}(\mathit{Reach}_i(T)) \le \mathcal{V}_i(v) \le \mathcal{P}_v^{\sigma_i,\pi}(\mathit{Reach}_i(T))$.

Lemma 3.1. *There is a MD strategy $\bar{\pi} \in \Pi$ such that for every $v \in V$ and every $\sigma \in \Sigma$ we have that $\mathcal{P}_v^{\sigma,\bar{\pi}}(\mathit{Reach}(T)) \le \mathcal{V}(v)$.*

Proof. Consider an arbitrary $u \in V$. By finite branching there must be some $u \longrightarrow s$ such that for infinitely many $i \in \mathbb{N}$:

$$\mathcal{V}_i(s) = \min_{u\longrightarrow t} \mathcal{V}_i(t) \tag{2}$$

We define $\bar{\pi}$ by setting $\bar{\pi}(u) = s$.

We first prove that for every $u \in V$, $\sigma \in \Sigma$ and every $k \in \mathbb{N}_0$ there is some $j \in \mathbb{N}$ such that

$$\mathcal{P}_u^{\sigma,\bar{\pi}}(\mathit{Reach}_k(R)) \le \mathcal{V}_{j+k}(u) \tag{3}$$

The proof is by induction on k. For $k = 0$ the statement is clear. Assume that $k = h + 1$ for some $h \in \mathbb{N}_0$ and fix an arbitrary $u \in V$ and $\sigma \in \Sigma$. Assume first that $u \in V_\Diamond$. Denote $s = \bar{\pi}(u)$. From the inductive hypothesis for $k = h$ there is some j' such that $\mathcal{P}_s^{\sigma,\bar{\pi}}(\mathit{Reach}_h(R)) \le \mathcal{V}_{j'+h}(s)$. From (2) there must be some $j \ge j'$ such that $\mathcal{V}_{j+h}(s) = \min_{u\longrightarrow t} \mathcal{V}_{j+h}(t)$. Then

$$\begin{aligned}
\mathcal{P}_u^{\sigma,\bar{\pi}}(\mathit{Reach}_{h+1}(R)) &= \mathcal{P}_s^{\sigma,\bar{\pi}}(\mathit{Reach}_h(R)) && (4)\\
&\le \mathcal{V}_{j'+h}(s) && (5)\\
&\le \mathcal{V}_{j+h}(s) && (6)\\
&= \mathcal{V}_{j+h+1}(u) && (7)
\end{aligned}$$

since (4) follows from the definition of $\bar{\pi}$, (5) is the inductive hypothesis, $j \ge j'$ implies (6) and (7) follows from the definition of the sequence $\mathcal{V}_i$.

Now we finish the proof of (3) for the cases where $u \in V_\Box \cup V_\bigcirc$. Denote $S = \{s \in V \mid u \longrightarrow s\}$. Let $u \in V_\bigcirc$ first. For every $s \in S$ let $p_s = \mathit{Prob}(u \longrightarrow s)$. The inductive hypothesis for $k = h$ delivers j_s for every $s \in S$ such that $\mathcal{P}^{\sigma,\bar{\pi}}_s(\mathit{Reach}_h(R)) \le \mathcal{V}_{j_s+h}(s)$. Due to the finite branching we can set $j = \max_{s\in S} j_s$ and get

$$\mathcal{P}^{\sigma,\bar{\pi}}_u(\mathit{Reach}_{h+1}(R)) = \sum_{s\in S} p_s \cdot \mathcal{P}^{\sigma,\bar{\pi}}_s(\mathit{Reach}_h(R)) \le \sum_{s\in S} p_s \cdot \mathcal{V}_{j+h}(s) = \mathcal{V}_{j+h+1}(u)$$

Similarly for $u \in V_\Box$

$$\mathcal{P}^{\sigma,\bar{\pi}}_u(\mathit{Reach}_{h+1}(R)) \le \max_{s\in S} \mathcal{P}^{\sigma,\bar{\pi}}_s(\mathit{Reach}_h(R)) \le \max_{s\in S} \mathcal{V}_{j+h}(s) = \mathcal{V}_{j+h+1}(u)$$

The lemma follows from (3), $\mathcal{P}^{\sigma,\bar{\pi}}_u(\mathit{Reach}(R)) = \sup_{k\in\mathbb{N}} \mathcal{P}^{\sigma,\bar{\pi}}_u(\mathit{Reach}_k(R))$, and $\mathcal{V}_{j+k}(u) \le \mathcal{V}(u)$. □

Theorem 3.2. *Every $v \in V$ has a value and $\mathit{val}(v) = \mathcal{V}(v)$.*

Proof. One can easily verify that

$$\mathcal{V}(v) \quad \le \quad \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_v(\mathit{Reach}(T)) \quad \le \quad \inf_{\pi\in\Pi} \sup_{\sigma\in\Sigma} \mathcal{P}^{\sigma,\pi}_v(\mathit{Reach}(T)) \tag{8}$$

Now take $\bar{\pi}$ from Lemma 3.1, which satisfies $\sup_{\sigma\in\Sigma} \mathcal{P}^{\sigma,\bar{\pi}}_v(\mathit{Reach}(T)) \le \mathcal{V}(v)$. Together with the previous inequality we get that

$$\mathcal{V}(v) = \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_v(\mathit{Reach}(T)) = \inf_{\pi\in\Pi} \sup_{\sigma\in\Sigma} \mathcal{P}^{\sigma,\pi}_v(\mathit{Reach}(T)) = \sup_{\sigma\in\Sigma} \mathcal{P}^{\sigma,\bar{\pi}}_v(\mathit{Reach}(T)) = \mathit{val}(v)$$

□

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

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

Lemma 3.4. *For every fixed vertex $v \in V$, we have that*

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

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

Now we can state and prove the promised determinacy theorem.

Theorem 3.5 (Determinacy)**.** *Let $v \in V$ and let $(T, \rhd\rho)$ be a (quantitative) reachability objective. Then one of the two players has a $(\rhd\rho)$-winning strategy in v.*

Proof. It suffices to prove the following:

$$\forall \pi \in \Pi \ \exists \sigma \in \Sigma : \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \rhd \rho \quad \Rightarrow \quad \exists \sigma \in \Sigma \ \forall \pi \in \Pi : \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \rhd \rho \quad (9)$$

If $\rhd$ is $>$ or $\mathit{val}(v) \neq \rho$ then this easily follows from Corollary 3.3. Indeed:

- For $\mathit{val}(v) > \rho$, choose $\varepsilon > 0$ such that $\mathit{val}(v) - \varepsilon > \rho$ and any ε-optimal maximizing strategy $\bar{\sigma} \in \Sigma$. Observe that $\bar{\sigma}$ satisfies both sides of (9) in the place of σ.
- For $\mathit{val}(v) < \rho$, observe that none of the two sides of (9) is satisfied for π being the optimal minimizing strategy.
- If $\rhd$ is $>$, observe that the left-hand side of (9) implies that $\mathit{val}(v) > \rho$ simply by choosing π to be the optimal minimizing strategy.

For the constraint ≥ 0 the statement is trivial. Now suppose that $\rhd$ is $\geq$ and $\rho = \mathit{val}(v) > 0$, and assume that the left-hand side of (9) holds. In the following we restrict the set of edges of G so that whenever $u \mapsto u'$, then $\mathit{val}(u') = \mathit{val}(u)$. Note that $\mathit{val}(u) = \max_{u \mapsto u'} \mathit{val}(u')$ for $u \in V_\Box$, and hence the restriction leaves the edge relation total.

Now we prove that the left-hand side in (9) still holds even if we restrict ourselves to those strategies $\sigma \in \Sigma$ which select the edges from the restricted edge relation. Indeed, assume the contrary, i.e. that there is some $\pi \in \Pi$ such that for every finite path $w \in \mathit{FPath}(G, v)$ visiting some $u \in V_\Box$ as the last vertex, whenever $\mathcal{P}_v^{\sigma,\pi}(\mathit{Run}(G(\sigma,\pi)), w) > 0$ for some $\sigma \in \Sigma$ then either $\sigma(w,t) > 0$ for some $u \mapsto t$, $\mathit{val}(t) < \mathit{val}(u)$ or $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) < \rho$. Modifying π to start behaving like the optimal minimizing strategy in all such t with the history w, we see that there cannot be any strategy σ such that $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \geq \rho$. This is a contradiction with our assumption that the left-hand side in (9) holds in the original (unrestricted) game.

As a consequence, $\mathit{val}(v)$ is not changed by restricting the edges, since the left-hand side in (9) implies $\mathit{val}(v) \geq \rho$ and it could not get increased by restricting the strategies of player $\Box$.

Due to Lemma 3.4, for every $u \in V$ we can fix a strategy $\sigma_u \in \Sigma$ and $n_u \in \mathbb{N}$ such that $\forall \pi \in \Pi : \mathcal{P}_u^{\sigma_s,\pi}(\mathit{Reach}_{n_u}(T)) > \mathit{val}(u)/2$. For every $k \in \mathbb{N}_0$, let $B(k)$ be the set of all $t \in V$ such that t is reachable from v in G via a path of length exactly k which does not visit T. Observe that $B(k)$ is finite because G is finitely branching. Further, for every $i \in \mathbb{N}_0$ we define a bound $m_i \in \mathbb{N}$ inductively as follows: $m_0 = 1$, and

$m_{i+1} = m_i + \max\{n_t \mid t \in B(m_i)\}$. Now we define a strategy $\sigma \in \Sigma$ which turns out to be $(\rhd\rho)$-winning. For every $w \in V^*V_\square$ such that $m_i \leq |w| < m_{i+1}$ we put $\sigma(w) = \sigma_t(tw_2)$, where $w = w_1tw_2$, $|w_1| = m_i - 1$ and $t \in V$. Now it is easy to check that for every $i \in \mathbb{N}$ and every strategy $\pi \in \Pi$ we have that $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_{m_i}(T)) > (1 - \frac{1}{2^i})\rho$. This means that the strategy σ is $(\rhd\rho)$-winning.

$\square$

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* $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\square, \Gamma_\Diamond, \Gamma_\bigcirc), \mathit{Prob})$ *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* $X \hookrightarrow \alpha$*,* $(\Gamma_\square, \Gamma_\Diamond, \Gamma_\bigcirc)$ *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 $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\square, \Gamma_\Diamond, \Gamma_\bigcirc), \mathit{Prob})$ determines a unique stochastic game $G_\Delta = (\Gamma^*, \mapsto, (\Gamma_\square\Gamma^*, \Gamma_\Diamond\Gamma^*, \Gamma_\bigcirc\Gamma^* \cup \{\varepsilon\}), \mathit{Prob}_\Delta)$ 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_Δ 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 reachability objectives $(T, \rhd\rho)$ 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., [10]). 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* $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), \mathit{Prob})$ *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 $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), \mathit{Prob})$ 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 $\mathit{Reach}(\varepsilon)$).

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

In the forthcoming subsections we examine the sets $[T]_\Box^{\triangleright\rho}$ 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 $(\triangleright\rho)$-winning strategies for both players are effectively computable. The associated upper complexity bounds are essentially the same as above.

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

We start by observing that the sets $[T]_{\Box}^{\geq 0}$ 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* $\mathscr{A} = [T]_{\Box}^{\geq 0} \cap \Gamma$ *and* $\mathscr{B} = [T_\varepsilon]_{\Box}^{\geq 0} \cap \Gamma$. *Then* $[T]_{\Box}^{\geq 0} = \mathscr{B}^*\mathscr{A}\Gamma^*$ *and* $[T_\varepsilon]_{\Box}^{\geq 0} = \mathscr{B}^*\mathscr{A}\Gamma^* \cup \mathscr{B}^*$. *Consequently,* $[T]_{\Diamond}^{>0} = \Gamma^* \setminus [T]_{\Box}^{\geq 0} = (\mathscr{B} \setminus \mathscr{A})^* \cup (\mathscr{B} \setminus \mathscr{A})^*(\Gamma \setminus \mathscr{B})\Gamma^*$.

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:*

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

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 $[T]_{\Box}^{\geq 0}$ and $[T]_{\Diamond}^{>0}$ are computable in polynomial time. Thus, we obtain the following theorem:

Theorem 4.6. *The membership to* $[T]_{\Box}^{\geq 0}$ *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* $[T]_{\Box}^{\geq 0}$ *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 $[T]_{\Box}^{\geq 0}$, 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 $\hat{X} \xhookrightarrow{1} XZ$, $Z \xhookrightarrow{1} Z$, and setting $\hat{T} = T$). Due to Theorem 3.5, we have that $\mathscr{C} = \Gamma \smallsetminus ([T]_\Diamond^{=1} \cap \Gamma)$, and thus we can compute also the set $\mathscr{C}$. Since $\mathscr{B} = \Gamma \smallsetminus (\mathscr{A} \cup \mathscr{C})$ (again by Theorem 3.5), 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) $Y \in [T_\varepsilon]_\Diamond^{>0}$;

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

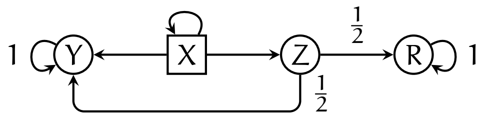
The set of all witnesses is denoted W.

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 $[\varepsilon]_\Diamond^{=1}$ is in **NP** $\cap$ **co-NP** due to

[13] (this is the only place where we use the decision algorithm for qualitative termination designed in [13]).

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 $\mathsf{V} \in \Gamma$ from which player $\Diamond$ can enforce visiting a witness with a positive probability (i.e., $\mathsf{V} \in Att(W)$ iff $\exists \pi \in \Pi \;\; \forall \sigma \in \Sigma \;:\; \mathcal{P}_{\mathsf{V}}^{\sigma,\pi}(Reach(W\Gamma^*)) > 0$). However, this is not true, as it is demonstrated in the following example:

Example 4.10. *Consider a stochastic BPA game* $\hat{\Delta} = (\{\mathsf{X},\mathsf{Y},\mathsf{Z},\mathsf{R}\}, \hookrightarrow, (\{\mathsf{X}\}, \emptyset, \{\mathsf{Y},\mathsf{Z},\mathsf{R}\}), Prob)$, *where* $\mathsf{X} \hookrightarrow \mathsf{X}$, $\mathsf{X} \hookrightarrow \mathsf{Y}$, $\mathsf{X} \hookrightarrow \mathsf{Z}$, $\mathsf{Y} \overset{1}{\hookrightarrow} \mathsf{Y}$, $\mathsf{Z} \overset{1/2}{\hookrightarrow} \mathsf{Y}$, $\mathsf{Z} \overset{1/2}{\hookrightarrow} \mathsf{R}$, $\mathsf{R} \overset{1}{\hookrightarrow} \mathsf{R}$, *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 $\mathscr{A} = \{\mathsf{X},\mathsf{Y},\mathsf{Z}\}$, $C = W = \{\mathsf{Y}\}$, *but* $Att(\{\mathsf{Y}\}) = \{\mathsf{Z},\mathsf{Y}\}$.

In Example 4.10, the problem is that player $\Box$ can use a strategy which always selects the rule $\mathsf{X} \hookrightarrow \mathsf{X}$ 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 $Att(\{\mathsf{Y}\})$, or he performs the loop $\mathsf{X} \hookrightarrow \mathsf{X}$ forever. The second possibility can be analyzed by "cutting off" the set $Att(\{\mathsf{Y}\})$ and recomputing the set of all witnesses together with its attractor in the resulting stochastic BPA game, which contains only X and the rule $\mathsf{X} \hookrightarrow \mathsf{X}$. 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 $\mathsf{X} \in \Gamma$ and every $\pi \in \Pi$ we have that $\mathcal{P}_{\mathsf{X}}^{\sigma,\pi}(Reach(\mathsf{T}_\varepsilon)) = 1$. 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 3.5 (here, the sets $[\varepsilon]^{=1}_{\Diamond}$ and $[\varepsilon]^{=1}_{\Box}$ are considered in the game Δ_C). Due to [12], 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 $\mathcal{P}^{\sigma_T,\pi}(\mathit{Reach}(\varepsilon)) = 1$. 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 $\mathcal{P}^{\sigma,\pi}_X(\mathit{Reach}(T_\varepsilon)) = 1$. By analyzing the play $G_\Delta(\sigma,\pi)$, one can show that there is a set of runs $V \subseteq \mathit{Run}(G_\Delta(\sigma,\pi), X)$ 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 $\hookrightarrow \setminus \hookrightarrow_C$ 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 $v \in \mathit{FPath}(X)$ such that

(B) $\mathcal{P}(U) > 0$, and each $w \in U$ satisfies the following: w starts with v, the rules of $\hookrightarrow \setminus \hookrightarrow_C$ 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. $\Box$

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" $\mathit{Att}(W)$.

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* $X \hookrightarrow X$*,* $X \hookrightarrow Y$*,* $X \hookrightarrow ZY$*,* $Y \stackrel{1}{\hookrightarrow} Y$*,* $Z \stackrel{1/2}{\hookrightarrow} X$*,* $Z \stackrel{1/2}{\hookrightarrow} R$*,* $R \stackrel{1}{\hookrightarrow} R$*, and* $\hat{\Gamma}_T = \{R\}$*. 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* $X \hookrightarrow X$*,* $Z \stackrel{1/2}{\hookrightarrow} X$*,* $Z \stackrel{1/2}{\hookrightarrow} R$*,* $R \stackrel{1}{\hookrightarrow} R$*. 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 [13]. 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 *Lecture Notes in Computer Science*, 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 Transactions on Computational Logic*, 9(1), 2007.

[3] T. Brázdil, V. Brožek, V. Forejt, and A. Kučera. Reachability in recursive Markov decision processes. In *Proceedings of CONCUR 2006*, volume 4137 of *Lecture Notes in Computer Science*, pages 358–374. Springer, 2006.

[4] T. Brázdil, V. Brožek, V. Forejt, and A. Kučera. Reachability in recursive Markov decision processes. *Information and Computation*, 206(5):520–537, 2008.

[5] K. Chatterjee, L. de Alfaro, and T. Henzinger. The complexity of stochastic Rabin and Streett games. In *Proceedings of ICALP 2005*, volume 3580 of *Lecture Notes in Computer Science*, pages 878–890. Springer, 2005.

[6] K. Chatterjee, M. Jurdziński, and T. Henzinger. Simple stochastic parity games. In *Proceedings of CSL'93*, volume 832 of *Lecture Notes in Computer Science*, pages 100–113. Springer, 1994.

[7] K. Chatterjee, M. Jurdziński, and T. Henzinger. Quantitative stochastic parity games. In *Proceedings of SODA 2004*, pages 121–130. SIAM, 2004.

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

[9] L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. *Journal of Computer and System Sciences*, 68:374–397, 2004.

[10] J. Esparza, A. Kučera, and S. Schwoon. Model-checking LTL with regular valuations for pushdown systems. *Information and Computation*, 186(2):355–376, 2003.

[11] K. Etessami, D. Wojtczak, and M. Yannakakis. Recursive stochastic games with positive rewards. In *Proceedings of ICALP 2008, Part I*, volume 5125 of *Lecture Notes in Computer Science*, pages 711–723. Springer, 2008.

[12] K. Etessami and M. Yannakakis. Recursive Markov decision processes and recursive stochastic games. In *Proceedings of ICALP 2005*, volume 3580 of *Lecture Notes in Computer Science*, pages 891–903. Springer, 2005.

[13] 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 *Lecture Notes in Computer Science*, pages 634–645. Springer, 2006.

[14] K. Etessami and M. Yannakakis. Recursive concurrent stochastic games. In *Proceedings of ICALP 2006*, volume 4052 of *Lecture Notes in Computer Science*, pages 324–335. Springer, 2006.

[15] J.E. Hopcroft and J.D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, 1979.

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

[17] D.A. Martin. The determinacy of blackwell games. *Journal of Symbolic Logic*, 63(4):1565–1581, 1998.

[18] W. Thomas. Infinite games and verification. In *Proceedings of CAV 2003*, volume 2725 of *Lecture Notes in Computer Science*, pages 58–64. Springer, 2003.

[19] N. Vieille. Stochastic games: Recent results. *Handbook of Game Theory*, pages 1833–1850, 2002.

[20] I. Walukiewicz. A landscape with games in the background. In *Proceedings of LICS 2004*, pages 356–366. IEEE Computer Society Press, 2004.

A Proofs of Section 4

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. We also assume that for each $R \in \Gamma_T$, the only rule where R appears on the left-hand side is $R \hookrightarrow R$.

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, $[T]_\Diamond^{=0} = \Gamma^* \setminus [T]_\Box^{>0} = (\mathscr{B} \setminus \mathscr{A})^* \cup (\mathscr{B} \setminus \mathscr{A})^*(\Gamma \setminus \mathscr{B})\Gamma^*$.*

Proof. We start by introducing some notation. For every strategy $\sigma \in \Sigma$ and every $\alpha \in \Gamma^*$, let

- $\sigma[-\alpha]$ be a strategy such that for every finite sequence of configurations $\gamma_1, \ldots, \gamma_n, \gamma$, where $n \geq 0$ and $\gamma \in \Gamma_\Box\Gamma^*$, and every edge $\gamma \mapsto \delta$ we have that $\sigma[-\alpha](\gamma_1, \ldots, \gamma_n, \gamma)(\gamma \mapsto \delta) = \sigma(\gamma_1\alpha, \ldots, \gamma_n\alpha, \gamma\alpha)(\gamma\alpha \mapsto \delta\alpha)$
- $\sigma[+\alpha]$ be a strategy such that for every finite sequence of configurations $\gamma_1\alpha, \ldots, \gamma_n\alpha, \gamma\alpha$, where $n \geq 0$ and $\gamma\alpha \in \Gamma_\Box\Gamma^*$, and every edge $\gamma\alpha \mapsto \delta\alpha$ we have that $\sigma[+\alpha](\gamma_1\alpha, \ldots, \gamma_n\alpha, \gamma\alpha)(\gamma\alpha \mapsto \delta\alpha) = \sigma(\gamma_1, \ldots, \gamma_n, \gamma)(\gamma \mapsto \delta)$

By induction on the length of $\alpha \in \Gamma^*$, we prove that $\alpha \in [T]_\Box^{>0}$ iff $\alpha \in \mathscr{B}^*\mathscr{A}\Gamma^*$. For $\alpha = \varepsilon$, both sides of the equivalence are false. Now assume that the equivalence holds for all configurations of length k and consider an arbitrary $X\alpha \in \Gamma^+$ where $|\alpha| = k$. If $X\alpha \in [T]_\Box^{>0}$ then there are two possibilities:

- There is a strategy $\sigma \in \Sigma$ such that for all $\pi \in \Pi$, the probability of reaching T without prior reaching α is positive in the play $G_\Delta(\sigma, \pi)$ initiated in $X\alpha$. Then $\sigma[-\alpha]$ is $(T, >0)$-winning in X, which means $X \in [T]_\Box^{>0}$, i.e., $X \in \mathscr{A}$.
- There is a strategy $\sigma \in \Sigma$ such that for all $\pi \in \Pi$, the probability of reaching T is positive in the play $G_\Delta(\sigma, \pi)$ initiated in $X\alpha$, but for some $\hat{\pi} \in \Pi$, the configuration α is always reached before reaching T. In this case, consider again the strategy $\sigma[-\alpha]$. Then σ' is $(T_\varepsilon, >0)$-winning in X, which means $X \in [T_\varepsilon]_\Box^{>0}$, i.e., $X \in \mathscr{B}$. Moreover, observe that the strategy σ is $(T, >0)$-winning in α. Thus, $\alpha \in [T]_\Box^{>0}$ and by induction hypothesis we obtain $\alpha \in \mathscr{B}^*\mathscr{A}\Gamma^*$.

In both cases, we obtained $X\alpha \in \mathscr{B}^*\mathscr{A}\Gamma^*$. If $X\alpha \in \mathscr{B}^*\mathscr{A}\Gamma^*$, we can again distinguish two possibilities:

- $X \in \mathscr{A}$ and there is a $(T, >0)$-winning strategy $\sigma \in \Sigma$ for the initial configuration X. Then the strategy $\sigma[+\alpha]$ is $(T, >0)$-winning for $X\alpha$. Thus, $X\alpha \in [T]_{\square}^{>0}$.

- $X \in \mathscr{B}$ and $\alpha \in \mathscr{B}^* \mathscr{A} \Gamma^*$. Then we have a $(T_\varepsilon, >0)$-winning strategy $\sigma_1 \in \Sigma$ for X. By induction hypothesis, there is a $(T, >0)$-winning strategy $\sigma_2 \in \Sigma$ for α. We construct a strategy σ' which behaves like $\sigma_1[+\alpha]$ until α is reached, and from that point on it behaves like σ_2. Obviously, σ' is $(T, >0)$-winning, which means that $X\alpha \in [T]_{\square}^{>0}$.

The proof of $[T_\varepsilon]_{\square}^{>0} = \mathscr{B}^* \mathscr{A} \Gamma^* \cup \mathscr{B}^*$ is similar. $\square$

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:*

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

Proof. For every $i \in \mathbb{N}_0$, let $(A_i, B_i) = F^i(\emptyset, \emptyset)$, and let $(\mathscr{A}_F, \mathscr{B}_F) = (\bigcup_{i \in \mathbb{N}} A_i, \bigcup_{i \in \mathbb{N}} B_i)$ be the least fixed-point of F. We show that $(\mathscr{A}_F, \mathscr{B}_F) = (\mathscr{A}, \mathscr{B})$.

We start with the "$\subseteq$" direction. We use the following notation:

- for every $X \in \mathscr{A}_F$, let $I_A(X)$ be the least $i \in \mathbb{N}$ such that $X \in A_i$;

- for every $X \in \mathscr{B}_F$, let $I_B(X)$ be the least $i \in \mathbb{N}$ such that $X \in B_i$;

- for every $\alpha Y \in \mathscr{B}_F^* \mathscr{A}_F$, let $I(\alpha Y) = \max(\{I_A(Y)\} \cup \{I_B(Z) \mid Z \text{ appears in } \alpha\})$;

- for every $\beta \in \Gamma^*$, let $\textit{price}(\beta) = \min\{I(\gamma) \mid \gamma \text{ is a prefix of } \beta, \gamma \in \mathscr{B}_F^* \mathscr{A}_F\}$, where $\min(\emptyset) = \infty$.

For every $X \in \mathscr{A}_F \cap \Gamma_\square$, we fix some $X \hookrightarrow \alpha$ (the "A-rule") such that $\textit{price}(\alpha) < I_A(X)$. It follows directly from the definition of F that there must by such a rule. Similarly, for every $X \in \mathscr{B}_F \cap \Gamma_\square$, we fix some $X \hookrightarrow \alpha$ (the "B-rule") such that either $\textit{price}(\alpha) < I_B(X)$, or $\alpha \in \mathscr{B}_F^*$ and $I_B(Y) < I_B(X)$ for every Y of α.

Now consider a MD strategy $\sigma \in \Sigma$ which for a given $X\alpha \in \mathscr{B}_F^* \mathscr{A}_F \Gamma^* \cap \Gamma_\square \Gamma^*$ selects

- the A-rule of X if $X \in \mathscr{A}_F$ and $I_A(X) = \mathit{price}(X\alpha)$;
- the B-rule of X otherwise.

We claim that σ is (>0)-winning with respect to T in every configuration of $\mathscr{B}_F^* \mathscr{A}_F \Gamma^*$. In particular, this means that $\mathscr{A}_F \subseteq \mathscr{A}$. To see this, realize that for every $\pi \in \Pi$, the play $G_\Delta(\sigma, \pi)$ contains a path along which every transition either decreases the price, or maintains the price but decreases either the length or replaces the first symbol with a sequence of symbols whose I_B value is strictly smaller. Hence, this path must inevitably visit T after performing a finite number of transitions.

Similar arguments show that σ is (>0)-winning with respect to T_ε in every configuration of $\mathscr{B}_F^* \mathscr{A}_F \Gamma^* \cup \mathscr{B}_F^*$. In particular, this means that $\mathscr{B}_F \subseteq \mathscr{B}$.

Now we prove the "$\supseteq$" direction, i.e. $\mathscr{A}_F \supseteq \mathscr{A}$ and $\mathscr{B}_F \supseteq \mathscr{B}$. Let us define the $\mathscr{A}$-*norm* of a given $X \in \Gamma$, $N_A(X)$, to be the least n such that for some $\sigma \in \Sigma$ and for all $\pi \in \Pi$ there is a path in $G_\Delta(\sigma, \pi)$ of length at most n from X to T. Similarly, define the $\mathscr{B}$-*norm* of a given $X \in \Gamma$, $N_B(X)$, to be the least n such that for some $\sigma \in \Sigma$ and for all $\pi \in \Pi$ there is a path in $G_\Delta(\sigma, \pi)$ of length at most n from X to T_ε. (If there are no such paths, then we put $N_A(X) = \infty$ (resp. $N_B(X) = \infty$)).

It follows from König's lemma and the fact that the game is finitely branching that $N_A(X)$ is finite for every $X \in \mathscr{A}$, and $N_B(X)$ is finite for every $X \in \mathscr{B}$. Also note that for all $X \in \Gamma$ we have $N_A(X) \geq N_B(X)$.

We show by induction on n that every $X \in \mathscr{A}$ s.t. $N_A(X) = n$ belongs to A_n and that every $X \in \mathscr{B}$ s.t. $N_B(X) = n$ belongs to B_n.

- $X \in \mathscr{A}$: For $X \in \Gamma_\Box$ (for $X \in \Gamma_\Diamond$) some transition (every transition) has the form $X \hookrightarrow \beta Y \gamma$ where $\beta \in \mathscr{B}^*$, $Y \in \mathscr{A}$, $N_A(Y) < n$ and $N_B(Z) < n$ for all Z which appear in β. By induction, $\beta \in B_{n-1}^*$ and $Y \in A_{n-1}$. Hence, $X \in A_n$.
- $X \in \mathscr{B}$: For $X \in \Gamma_\Box$ (for $X \in \Gamma_\Diamond$) some transition (every transition) of the form $X \hookrightarrow \bar\beta$ satisfies one of the following conditions:
  - $\bar\beta = \beta Y \gamma$ where $\beta \in \mathscr{B}^*$, $Y \in \mathscr{A}$, $N_A(Y) < n$ and $N_B(Z) < n$ for all Z which appear in β. By induction, $\beta \in B_{n-1}^*$ and $Y \in A_{n-1}$. Hence, $X \in A_n \subseteq B_n$.
  - $\bar\beta \in \mathscr{B}^*$ where $N_B(Z) < n$ for all Z which appear in $\bar\beta$. By induction, $\bar\beta \in B_{n-1}^*$, and hence $X \in B_n$.

□

Theorem 4.6 *The membership to* $[T]^{\geq 0}_{\Box}$ *and* $[T]^{>0}_{\Diamond}$ *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* $[T]^{\geq 0}_{\Box}$ *and* $[T]^{>0}_{\Diamond}$*, respectively.*

Proof. Due to Proposition 4.5, it only remains to show that the (>0)-winning strategies for both players are regular and effectively constructible in polynomial time. Observe that the strategy σ defined in the proof of Proposition 4.5 is regular and (>0)-winning for player $\Box$. Moreover, σ is regular, which can be seen from the following automaton construction: We start with the observation that $(\mathscr{A}, \mathscr{B}) = (\bigcup_{i\geq 0}^{2|\Gamma|} A_i, \bigcup_{i\geq 0}^{2|\Gamma|} B_i)$. This is since F is monotone and the longest chain in $2^\Gamma \times 2^\Gamma$ has length $2|\Gamma| + 1$. We create a finite-state automaton which has one state q_i for each $i, 0 \leq i \leq 2|\Gamma|$, and a starting state q_∞. It reads the configuration stack from the bottom. When it reads $X \in \mathscr{A}$ in state q_i, it compares $j = I_A(X)$ with i. If $j < i$ it changes state to q_j. Otherwise, and also upon reading a symbol $X \in \mathscr{B} \smallsetminus \mathscr{A}$, the next state is q_k, where $k = \max(I_B(X), i)$. Finally upon reading symbol from $\Gamma \smallsetminus \mathscr{B}$ the next state is q_∞. It can easily be seen that for a configuration $X\alpha = \mathscr{B}^* \mathscr{A} \Gamma^* \cap \Gamma_\Box \Gamma^*$ the automaton ends up in a state q_i, where $i = \mathit{price}(X\alpha)$. Selecting the correct rule is then obvious.

It remains to prove that there is a regular MD strategy π for player $\Diamond$, which is (>0)-winning in every configuration of $[T]^{>0}_{\Diamond} = (\mathscr{B} \smallsetminus \mathscr{A})^* \cup (\mathscr{B} \smallsetminus \mathscr{A})^*(\Gamma \smallsetminus \mathscr{B})\Gamma^*$. However in this case such a strategy for $X \in (\mathscr{B} \smallsetminus \mathscr{A})$ just selects some (fixed) rule $X \hookrightarrow \beta$, where $\beta \in (\mathscr{B} \smallsetminus \mathscr{A})^* \cup (\mathscr{B} \smallsetminus \mathscr{A})^*(\Gamma \smallsetminus \mathscr{B})\Gamma^*$, and similarly for $X \in (\Gamma \smallsetminus \mathscr{B})$ some rule $X \hookrightarrow \beta$, where $\beta \in (\mathscr{B} \smallsetminus \mathscr{A})^*(\Gamma \smallsetminus \mathscr{B})\Gamma^*$. It is clear that such transitions must always exist. Then such strategy π is obviously (>0)-winning for player $\Diamond$ and moreover it depends only on the top of the stack, and therefore is effectively regular. $\Box$

In the proof of Proposition 4.11, we use the following corollary:

Corollary A.1. *Let* $\sigma \in \Sigma$ *be a strategy of player* $\Box$ *such that for every configuration* $\alpha \in \Gamma^*$ *we have that* $\sigma(\alpha)$ *is the uniform probability distribution over the outgoing transitions of* α*. Then for each* $X \in [T]^{\geq 0}_{\Box} \cap \Gamma$ *(or* $X \in [T_\varepsilon]^{\geq 0}_{\Box} \cap \Gamma$*) and each* $\pi \in \Pi$ *there is a path* w *from* X *to* T *(to* T_ε*, resp.) in* $G_\Delta(\sigma, \pi)$ *such that*

1. *the length of* w *is at most* $2^{2|\Gamma|}$*;*

2. *the length of all configurations entered along* w *is at most* $2|\Gamma|$*.*

Proof. Let us consider the sets A_i and B_i from the above proof of Proposition 4.5. Note that $[T]_{\Box}^{\geq 0} \cap \Gamma = \bigcup_{i=0}^{2|\Gamma|} A_i$ and $[T_\varepsilon]_{\Box}^{\geq 0} \cap \Gamma = \bigcup_{i=0}^{2|\Gamma|} B_i$. We prove by induction on i that for each $X \in A_i$ (or $X \in B_i$) and each $\pi \in \Pi$ there is a path w from X to T (to T_ε, resp.) in $G_\Delta(\sigma, \pi)$ such that

1. the length of w is at most 2^i;
2. the length of all configurations entered along w is at most i.

The case $i = 0$ is trivial. Assume that $i \geq 1$. If $X \in A_i \cap (\Gamma_\Box \cup \Gamma_\bigcirc)$, then by the definition of A_i, there is a transition $X \hookrightarrow \gamma$ such that $\gamma \in \Gamma_T \cup A_{i-1}\Gamma \cup B_{i-1}A_{i-1} \cup A_{i-1}$. By induction hypothesis, there is a path w' from γ to T in $G_\Delta(\sigma, \pi)$ of length at most $2^i + 2^i = 2^{i+1}$ such that the length of all configurations entered by w' is at most $\max\{i+1, i\} = i+1$. The rest follows from the fact that σ plays uniformly in X whenever $X \in \Gamma_\Box$. Similarly, if $X \in A_i \cap \Gamma_\Diamond$, then all transitions have the form $X \hookrightarrow \gamma$ where $\gamma \in \Gamma_T \cup A_{i-1}\Gamma \cup B_{i-1}A_{i-1} \cup A_{i-1}$, and we obtain the desired result by induction. The case $X \in B_i$ is similar. $\square$

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 $[T]_\Diamond^{=1} = \mathscr{B}^*\mathscr{A}\Gamma^* \cup \mathscr{B}^*$.*

Proof. By applying Theorem 3.5, we see that $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{C}$ are pairwise disjoint and their union is Γ. Using Theorem 3.5 again, it suffices to prove that $[T]_\Box^{=1} = \mathscr{B}^*\mathscr{C}\Gamma^*$.

We proceed by induction on the length of $\alpha \in \Gamma^*$ and prove that $\alpha \in [T]_\Box^{=1}$ iff $\alpha \in \mathscr{B}^*\mathscr{C}\Gamma^*$ (we use the notation $\sigma[-\alpha]$ and $\sigma[+\alpha]$ that was introduced in the proof of Proposition 4.4). For $\alpha = \varepsilon$, both sides of the equivalence are false. Now assume that the equivalence holds for all configurations of length k, and consider an arbitrary $X\alpha \in \Gamma^+$ where $|\alpha| = k$. If $X\alpha \in [T]_\Box^{=1}$, we distinguish two possibilities:

- There is a strategy $\sigma \in \Sigma$ such that for all $\pi \in \Pi$, the probability of reaching T without prior reaching α is 1 from $X\alpha$ in $G_\Delta(\sigma, \pi)$. Then $\sigma[-\alpha]$ is $(T, {=}1)$-winning in X, which means that $X \in [T]_\Box^{=1}$, i.e., $X \in \mathscr{C}$.

- There is a strategy $\sigma \in \Sigma$ such that for all $\pi \in \Pi$, the probability of reaching T in the play $G_\Delta(\sigma, \pi)$ initiated in $X\alpha$ is 1, but for some $\hat{\pi} \in \Pi$, the configuration α is reached with a positive probability before reaching T. In this case, consider again the strategy $\sigma[-\alpha]$, which is $(T_\varepsilon, {=}1)$-winning in X, which means $X \in [T_\varepsilon]_\Box^{=1}$, i.e., $X \in \mathscr{B}$. Moreover, observe that the strategy σ is $(T, {=}1)$-winning in α. Hence, $\alpha \in [T]_\Box^{=1}$ and by applying induction hypothesis we obtain $\alpha \in \mathscr{B}^*\mathscr{C}\Gamma^*$.

If $X\alpha \in \mathscr{B}^*\mathscr{C}\Gamma^*$, there are two possibilities:

- $X \in \mathscr{C}$ and there is a $(T, =1)$-winning strategy $\sigma \in \Sigma$ for X. Then $\sigma[+\alpha]$ is $(T, =1)$-winning for $X\alpha$. Thus, $X\alpha \in [T]_{\Box}^{=1}$.

- $X \in \mathscr{B}$ and $\alpha \in \mathscr{B}^*\mathscr{C}\Gamma^*$. Then there is a $(T_\varepsilon, =1)$-winning strategy $\sigma_1 \in \Sigma$ for X. By applying induction hypothesis, there is a $(T, =1)$-winning strategy $\sigma_2 \in \Sigma$ for α. Now we can set up a $(T, =1)$-winning strategy for $X\alpha$, which behaves like $\sigma[+\alpha]$ until α is reached, and from that point on it behaves like σ_2. Hence, $X\alpha \in [T]_{\Box}^{=1}$.

$\square$

A.1 A Proof of Proposition 4.11

Our proof of Proposition 4.11 is obtained as a consequence of the following (more general) result:

Lemma A.2. *If there are no witnesses in Δ, then there is a MR strategy $\sigma \in \Sigma$ such that for every strategy $\pi \in \Pi$ and every $X \in \Gamma$ we have that $\mathcal{P}_X^{\sigma,\pi}(Reach(T_\varepsilon)) = 1$. Moreover, the strategy σ can be effectively represented by a finite automaton of size polynomial in $|\Delta|$.*

Note that Lemma A.2 indeed implies Proposition 4.11, because the existence of the strategy σ implies that $\mathscr{A} = \emptyset$.

The proof of Lemma A.2 proceeds as follows: First, we define the strategy $\sigma \in \Sigma$ using two auxiliary strategies σ_U and σ_T (see below). Then, we show that assuming that there are no witnesses in Γ, no strategy $\pi \in \Pi$ can win against σ in an arbitrary $X \in \Gamma$, i.e., $\mathcal{P}_X^{\sigma,\pi}(Reach(T_\varepsilon)) = 1$.

The auxiliary strategies σ_U and σ_T are defined as follows. The former, σ_U, is the unique MR strategy assigning to every configuration the uniform distribution over its outgoing transitions. The strategy σ_U has the following important property:

Lemma A.3. *There is $\xi > 0$ such that for every $X \in \Gamma$ and every $\pi \in \Pi$ there is a path w from X to a configuration of T_ε in $G_\Delta(\sigma_U, \pi)$ satisfying the following: The stack height along w is bounded by $2|\Gamma|$ and the probability of following w in $G_\Delta(\sigma_U, \pi)$ is at least ξ.*

Proof. It follows from the fact that $W = \emptyset$ and from Theorem 3.5 that for every strategy $\bar{\pi} \in \Pi$ and every $X \in \Gamma$ there is a path $v_{\bar{\pi}}$ from X to T_ε in $G_\Delta(\sigma_U, \bar{\pi})$. Moreover, by Corollary A.1, the length of $v_{\bar{\pi}}$ is bounded by $2^{\mathcal{O}(|\Delta|)}$ and the stack height along $v_{\bar{\pi}}$ is bounded by $2|\Gamma|$. Hence, for every $X \in \Gamma$ and every *deterministic* $\bar{\pi} \in \Pi$, the path $v_{\bar{\pi}}$

satisfies the following: The stack height along $\nu_{\bar{\pi}}$ is bounded by $2|\Gamma|$ and the probability of following $\nu_{\bar{\pi}}$ in $G_\Delta(\sigma_U, \bar{\pi})$ is at least $(\frac{\mu}{|\hookrightarrow|})^{2^{\mathcal{O}(|\Delta|)}}$ where μ is the least probability weight assigned by *Prob*.

Let $\pi \in \Pi$ be an arbitrary strategy. Let $\bar{\pi}$ be the deterministic strategy that always chooses the transition chosen by π with maximal probability. Note that the maximal probability is always at least $\frac{1}{|\hookrightarrow|}$. It follows that the probability of following $\nu_{\bar{\pi}}$ in $G_\Delta(\sigma_U, \pi)$ is at least $\xi := (\frac{\mu}{|\hookrightarrow|})^{2^{\mathcal{O}(|\Delta|)}}$. $\square$

Now we define σ_T. According to Theorem 3.5 applied to G_{Δ_C}, $W = \emptyset$ implies

$$\forall Y \in C \;\; \exists \sigma \in \Sigma_C \;\; \forall \pi \in \Pi_C \; : \; \mathcal{P}_Y^{\sigma,\pi}(\mathit{Reach}(\varepsilon)) = 1$$

By the results of [12], there is a SMD strategy $\sigma_T \in \Sigma_C$ satisfying

$$\forall Y \in C \;\; \forall \pi \in \Pi_C \; : \; \mathcal{P}_Y^{\sigma_T,\pi}(\mathit{Reach}(\varepsilon)) = 1$$

Now we are ready to define the MR strategy $\sigma \in \Sigma$. Note that it suffices to specify the values of σ for the individual configurations (i.e. histories of length 1) because σ is MR. For every configuration $\alpha \in \Gamma^*$, we define $\sigma(\alpha)$ as follows: If $\alpha = \beta\gamma$ such that $\beta \in C^*$ and $|\beta| > 2|\Gamma|$, then we put $\sigma(\alpha) = \sigma_T(\beta)\gamma$. Otherwise, we put $\sigma(\alpha) = \sigma_U(\alpha)$.

Now let us fix some strategy $\pi \in \Pi$. Our aim is to show that $\mathcal{P}_X^{\sigma,\pi}(\mathit{Reach}(T_\varepsilon)) = 1$ for all $X \in \Gamma$. Suppose the converse, i.e., there is some $X \in \Gamma$ such that $\mathcal{P}_X^{\sigma,\pi}(\mathit{Reach}(T_\varepsilon)) < 1$.

Lemma A.4. *There is a set of runs $V \subseteq \mathit{Run}(G_\Delta(\sigma, \pi), X)$ and a set of rules $\hookrightarrow_V \subseteq \hookrightarrow$ such that:*

1. $\mathcal{P}_X^{\sigma,\pi}(V) > 0$,
2. *no run in V visits T_ε,*
3. *for every run w in V the set of rules used infinitely often in w is $\hookrightarrow_V$,*
4. $\hookrightarrow_V \subseteq \hookrightarrow_C$

Proof. It follows from our choice of X that there exists a set of runs V' satisfying (1) and (2). We can classify every run w of V' according to the set of rules that are used in w infinitely many times. Since there are only finitely many rules, there is a subset V of V' and a set of rules $\hookrightarrow_V$ that satisfy (1) and (3). Because $V \subseteq V'$, we obtain that V and $\hookrightarrow_V$ satisfy (1) and (2) and (3).

It remains to prove that V and $\hookrightarrow_V$ satisfy (4). Let $L \subseteq \Gamma$ be the set of all left-hand sides of all rules from $\hookrightarrow_V$. To show that $\hookrightarrow_V \subseteq \hookrightarrow_C$, it suffices to prove that

(a) for all $X \in (L \setminus C) \cap (\Gamma_{\bigcirc} \cup \Gamma_{\Box})$, all rules for X are in $\hookrightarrow_V$;

(b) the right-hand sides of rules in $\hookrightarrow_V$ are in $(L \cup C)^*$.

Indeed, (a) and (b) imply that $L \cup C$ is a terminal set. Hence, $L \cup C = C$ by the maximality of C and $\hookrightarrow_V \subseteq \hookrightarrow_C$.

Note that (a) follows from the fact that the strategy σ plays uniformly in all configurations with the head from $(L \setminus C) \cap \Gamma_{\Box}$. Then every transition rule for $X \in (L \setminus C) \cap (\Gamma_{\bigcirc} \cup \Gamma_{\Box})$ has probability of being chosen greater than some fixed non-zero bound, which means that such a rule is in $\hookrightarrow_V$.

Now we prove (b). Let $(X, \gamma) \in \hookrightarrow_V$. If $\gamma = \varepsilon$ then $\gamma \in (L \cup C)^*$. If either $\gamma = P$, or $\gamma = PQ$, then surely $P \in L$ since configurations with P on the top of the stack occur infinitely many times in all runs of V. If $Q \in C$, then we are done. Assume that $Q \notin C$. Note that the strategy σ plays uniformly in all configurations of the form $\beta Q \alpha$ where $|\beta| \leq 2|\Gamma|$. By Lemma A.3, there is $0 < \xi < 1$ such that for every configuration of the form $PQ\alpha$ there is a path w from $PQ\alpha$ to $T \cup \{Q\alpha\}$ in $G_{\Delta}(\sigma, \pi)$ satisfying the following:

- all configurations in w are of the form $\beta Q \alpha$ where $|\beta| \leq 2|\Gamma|$;
- the probability of following w in $G_{\Delta}(\sigma, \pi)$ is at least ξ.

It follows that $Q \in L$ because every run from V contains infinitely many occurrences of configurations of the form $PQ\alpha$ and no run of V reaches T. □

Lemma A.5. *There is a finite path v and a set of runs $U \subseteq V$ such that*

1. $\mathcal{P}_X^{\sigma,\pi}(U) > 0$,
2. *every run of U has a prefix v;*
3. *for every run w of U, rules from $\hookrightarrow \setminus \hookrightarrow_C$ can occur only in the prefix v of w;*
4. *every $w \in U$ satisfies the following: every configuration reached by w after the initial path v has the stack height greater than or equal to the stack height of the last configuration reached by v.*

Proof. For every finite path v initiated in X there is a set of runs $V_v \subseteq V$ satisfying (2), (3) and (4) (where we use V_v instead of U). It follows from Lemma A.4 that $V = \bigcup_v V_v$. Since there are countably many finite paths, there is a v such that V_v has a non-zero probability. It suffices to put $U = V_v$. □

We prove that $\mathcal{P}_X^{\sigma,\pi}(U) = 0$ and obtain a contradiction. Let $X\alpha$ be the last configuration reached by v (formally, the last state of v is a string of configurations and $X\alpha$ is the last configuration of this string). Note that by Lemma A.5, part (3), $X \in C$. Let $\sigma' \in \Sigma$ be a MR strategy in G_{Δ_C} such that for every $\beta \in C^+$ we have $\sigma'(\beta) = \sigma(\beta\alpha)$. We define a strategy π' in G_{Δ_C} as follows: Let us execute, in parallel, a play in G_{Δ_C} initiated in X and a play in G_Δ initiated in $X\alpha$. Player $\Box$ plays according to σ' in G_{Δ_C}, and according to σ in G_Δ; player $\Diamond$ plays according to π in G_Δ, and in every step chooses simultaneously the same distribution on transition rules[1] in Δ_C as in Δ until the following happens: a transition of $\hookrightarrow \setminus \hookrightarrow_C$ is assigned a non-zero probability by π. From this moment on, both players play according to arbitrary (but fixed) strategies in both games.

Note that every run $w \in U$ induces a sequence of transition rules $t_1, t_2, \ldots$ where $w(i) \longrightarrow w(i+1)$ is induced by t_i for every $i \geq 1$. Now the probability that a sequence of rules induced by a run of U is executed in $G_{\Delta_C}(\sigma', \pi')$ is equal to $\mathcal{P}_X^{\sigma,\pi}(U)$. Observe that if a sequence of transitions induced by a run of U is executed in G_{Δ_C}, then ε is never reached. However, we show that the probability of reaching ε in $G_{\Delta_C}(\sigma', \pi')$ is 1, and thus that $\mathcal{P}_X^{\sigma,\pi}(U) = 0$.

Observe that the strategy σ' works as follows: In $\beta \in C^+$ such that $|\beta| \leq 2|\Gamma|$, the strategy σ' plays uniformly. Otherwise, it chooses the same transitions as σ_T. We show that there is $0 < \xi < 1$ such that for every β satisfying $|\beta| \leq 2|\Gamma|$ the probability of reaching ε from β in $G_{\Delta_C}(\sigma', \pi')$ is at least ξ. Indeed, by Lemma A.3, there would be such a ξ if player $\Box$ were playing uniformly in all configurations. However, playing according to σ_T in configurations of the form γ where $|\gamma| > 2|\Gamma|$ could only increase the probability of reaching ε. Now note that almost all runs of $Run(G_{\Delta_C}(\sigma', \pi'), X)$ contain configurations of the form $\beta \in C^*$ where $|\beta| \leq 2|\Gamma|$ infinitely many times. It follows that almost all runs of $Run(G_{\Delta_C}(\sigma', \pi'), X)$ reach ε.

A.2 A Proof of Theorem 4.13

Without the loss of generality, we adopt some additional assumptions about the considered stochastic BPA game Δ. First, we assume that $\Gamma_T = \{R\}$ (this is clearly no restriction). Further, we assume that for every $X, Y \in \Gamma$, there is at most one rule of the form $X \hookrightarrow Y\alpha$, and for every $X \in \Gamma_\Box \cup \Gamma_\Diamond$ there are at least two rules of the form $X \hookrightarrow \beta$ (every BPA can

[1]Strictly formally, the strategy σ assigns a distribution on transitions from a configuration $Y\gamma$. However, every transition from $Y\gamma$ is induced by a transition rule and we can consider such a distribution to be a distribution on transition rules.

be put into this form in polynomial time by renaming symbols in conflicting rules). We put $\Gamma' = \{X' \mid X \in \Gamma\}$, and for a given $\odot \in \{\bigcirc, \Diamond, \Box\}$, we put $\Gamma'_\odot = \{X' \mid X \in \Gamma_\odot\}$.

We construct a sequence of stochastic BPA games where the stack alphabet is always a subset of $\Gamma \cup \Gamma'$ (the elements of $\Gamma \cup \Gamma'$ are denoted by $\bar{X}, \bar{Y}, \ldots$). We use the convention, that the set of stack symbols of such BPA is completely determined by its rules, i.e., a symbol $\bar{X}$ is a symbol of a given BPA iff $\bar{X}$ occurs in some transition rule of the BPA (if $\bar{X}$ does not occur on the left-hand side of any rule, we treat $\bar{X}$ as if there is a rule $\bar{X} \hookrightarrow \bar{X}$). Also, every symbol $\bar{X} \in \Gamma \cup \Gamma'$ belongs to player $\odot$, where $\odot \in \{\bigcirc, \Diamond, \Box\}$, iff $\bar{X} \in \bar{\Gamma}_\odot = \Gamma_\odot \cup \Gamma'_\odot$.

We start with Δ and apply the following two procedures **Init** and **Main**.

Init: For every $X \in \Gamma$ do the following:

- if $X \in \Gamma_\Diamond \cup \Gamma_\bigcirc$ and $X \hookrightarrow \varepsilon$, then add a rule $X' \hookrightarrow X'$;
- else
  - add $X' \hookrightarrow YZ'$ for every rule of the form $X \hookrightarrow YZ$;
  - add $X' \hookrightarrow Z'$ for every rule of the form $X \hookrightarrow Z$.
- $\Gamma_T := \{R, R'\}$

Main:

1. Set $\mathcal{W} := \emptyset$.
2. Compute the greatest set W of witnesses;
   - if $W = \emptyset$, then stop and return $\mathcal{W}$.
3. For every rule $\bar{X} \hookrightarrow \bar{Z}\bar{Y}$ such that $\bar{Y} \in W$, erase $\bar{X} \hookrightarrow \bar{Z}\bar{Y}$ and add $\bar{X} \hookrightarrow \bar{Z}'\bar{Y}$.
4. For every $\bar{X}$ such that $\bar{X} \hookrightarrow \bar{Y}\alpha$ where $\bar{Y} \in W$, do the following:
   - if $\bar{X} \in \bar{\Gamma}_\Diamond \cup \bar{\Gamma}_\bigcirc$, erase all rules with the left-hand side $\bar{X}$;
   - if $\bar{X} \in \bar{\Gamma}_\Box$, erase $\bar{X} \hookrightarrow \bar{Y}\alpha$.
5. Erase all remaining occurrences of symbols of W, i.e., for every $\bar{Y} \in W$
   - erase all rules of the form $\bar{Y} \hookrightarrow \alpha$;
   - erase every rule of the form $\bar{X} \hookrightarrow \bar{Z}\bar{Y}$ and add $\bar{X} \hookrightarrow \bar{Z}$.

6. Set $\mathcal{W} := \mathcal{W} \cup W$ and go to (2).

Let W_i be the set W computed in the i-th iteration of the loop at lines (2)–(6) in **Main**. Let $\bar{\Delta}$ be the BPA resulting from the application of the procedure **Init** to Δ. In what follows, symbols of Γ are denoted $X, Y, Z, K, \ldots$, symbols of Γ' are denoted $X', Y', Z', K', \ldots$, and symbols of $\Gamma \cup \Gamma'$ are denoted $\bar{X}, \bar{Y}, \bar{Z}, \bar{K}, \ldots$.

We show that the procedure **Main** computes the set $\mathscr{A}$ of Proposition 4.7 for the stochastic BPA game $\bar{\Delta}$. This surely suffices because $\bar{\Delta}$ subsumes the original Δ. Hence, we can also compute the sets $\mathscr{B}$ and $\mathscr{C}$ of Proposition 4.7, and thus the sets $[T]_{\Box}^{=1}$ and $[T]_{\Diamond}^{=1}$.

Proposition A.6. *In the stochastic BPA game $\bar{\Delta}$ we have that $\mathcal{W} = [T_\varepsilon]_{\Diamond}^{=1} \cap (\Gamma \cup \Gamma')$.*

Proof. First, we prove that $\mathcal{W} \subseteq [T_\varepsilon]_{\Diamond}^{=1}$ in $\bar{\Delta}$. Let us denote $U = [T_\varepsilon]_{\Diamond}^{=1}$ and $V = [T]_{\Diamond}^{=1}$ in $\bar{\Delta}$. By induction on n, we prove that $\bigcup_{i=1}^{n} W_i \subseteq U$ in $\bar{\Delta}$. By the definition of $\bar{\Delta}$, we obtain that if $X' \in \mathcal{W}$, then $X \in V$.

For $n = 1$, the inclusion follows directly from the definition of the set of witnesses and the definition of $\bar{\Delta}$.

Now let $n \geq 2$. Let $\bar{\Delta}_n$ by the BPA game obtained in $n-1$ iterations of the above algorithm (note that W_n is computed in $\bar{\Delta}_n$). For fixed $X, Y \in \Gamma$, we say that all rules of the form $\bar{X} \hookrightarrow \bar{Y}\alpha$, where $\bar{X} \in \{X, X'\}$ and $\bar{Y} \in \{Y, Y'\}$, *correspond* each other. Due to our assumptions about Δ and the definition of $\bar{\Delta}_n$, each rule $\bar{X} \hookrightarrow \alpha$ of $\bar{\Delta}_n$ corresponds to a *unique* rule of the form $X \hookrightarrow \beta$ (here $X \in \Gamma$) of $\bar{\Delta}$. Also, every rule $X \hookrightarrow \alpha$ of $\bar{\Delta}$ corresponds to at most one rule of the form $X \hookrightarrow \beta$ (here $X \in \Gamma$) of $\bar{\Delta}_n$ and to at most one rule of the form $X' \hookrightarrow \gamma$ (here $X' \in \Gamma'$) of $\bar{\Delta}_n$. In the case of rules that belong to player $\bigcirc$, the corresponding rules have always the same probability.

Claim A.7. *The following is true:*

1. *If $X \in \Gamma$ is a symbol of $\bar{\Delta}_n$, then*

   (a) *for every rule $X \hookrightarrow \bar{Y}\bar{Z}$ in $\bar{\Delta}_n$ we have $\bar{Y}, \bar{Z} \in \Gamma$ and $X \hookrightarrow \bar{Y}\bar{Z}$ in $\bar{\Delta}$;*

   (b) *for every rule $X \hookrightarrow \bar{Y}$ in $\bar{\Delta}_n$ we have either $\bar{Y} \in \Gamma$ and $X \hookrightarrow \bar{Y}$ in $\bar{\Delta}$, or $\bar{Y} = K' \in \Gamma'$ in which case $X \hookrightarrow KZ$ in $\bar{\Delta}$ for some $Z \in U$;*

   (c) *for every rule $X \hookrightarrow \varepsilon$ in $\bar{\Delta}_n$ there is a rule $X \hookrightarrow \varepsilon$ in $\bar{\Delta}$;*

2. *If $X' \in \Gamma'$ is a symbol of $\bar{\Delta}_n$, then*

(a) *for every rule* $X' \hookrightarrow \bar{Y}\bar{Z}$ *in* $\bar{\Delta}_n$ *we have that* $\bar{Y} \in \Gamma$ *and* $\bar{Z} = K' \in \Gamma'$ *and* $X \hookrightarrow \bar{Y}K$ *in* $\bar{\Delta}$*;*

(b) *for every rule* $X' \hookrightarrow \bar{Y}$ *in* $\bar{\Delta}_n$ *we have that* $\bar{Y} = K' \in \Gamma'$*, and either* $X \hookrightarrow K$ *in* $\bar{\Delta}$*, or* $X \hookrightarrow KZ$ *in* $\bar{\Delta}$ *where* $Z \in V$*;*

Proof. The cases (1) (a), (1) (c), and (2) (a) follow directly from the construction of the algorithm (the rules considered in these cases were not modified by the procedure **Main**).

ad (1) (b) The case $\bar{Y} \in \Gamma$ is clear. Assume that $\bar{Y} = K' \in \Gamma'$. There is $k < n$ such that $X \hookrightarrow KZ$ is a rule of $\bar{\Delta}_k$ (let k be the greatest number with this property). By the construction of the algorithm, $Z \in W_k$, and hence, by induction, $Z \in U$.

ad (2) (b) It is easy to show that $\bar{Y} = K' \in \Gamma'$ (in the the algorithm, the rules of this form can be produced only in the procedure **Init**, or in the step (5) of **Main** where $\bar{Z}$ is already in Γ'). If there is a rule $X \hookrightarrow K$ in $\bar{\Delta}$, then we are done. Otherwise, there is the greatest $k < n$ such that $X' \hookrightarrow KZ'$ is in $\bar{\Delta}_k$ and $Z' \in W_k$. It follows, by induction, that $Z' \in U$, and hence $Z \in V$.

□

Claim A.8. *The following is true:*

1. *If* $X \in \Gamma$ *is a symbol of* $\bar{\Delta}_n$ *and*

   (a) *if* X *belongs either to* $\Diamond$*, or to* $\bigcirc$*, and if there are no rules leaving* X *in* $\bar{\Delta}_n$*, then there is a rule* $X \hookrightarrow \alpha$ *in* $\bar{\Delta}$ *such that* $\alpha \in U$*.*

   (b) *if* X *belongs to* $\Box$ *and if there is a rule* $X \hookrightarrow \alpha$ *in* $\bar{\Delta}$ *for which there is no corresponding rule of the form* $X \hookrightarrow \beta$ *in* $\bar{\Delta}_n$*, then* $\alpha \in U$*.*

2. *If* $X' \in \Gamma'$ *is a symbol of* $\bar{\Delta}_n$ *and*

   (a) *if* X' *belongs either to* $\Diamond$*, or to* $\bigcirc$*, and if there are no rules leaving* X' *in* $\bar{\Delta}_n$*, then there is a rule* $X \hookrightarrow \alpha$ *in* $\bar{\Delta}$ *such that* $\alpha \in V$*.*

   (b) *if* X' *belongs to* $\Box$ *and if there is a rule* $X \hookrightarrow \alpha$ *in* $\bar{\Delta}$ *for which there is no corresponding rule of the form* $X' \hookrightarrow \beta$ *in* $\bar{\Delta}_n$*, then* $\alpha \in V$*.*

Proof. For the cases (1) (a) and (b): Let $k < n$ be the greatest number such that $\bar{\Delta}_k$ contains a rule of the form $X \hookrightarrow \beta$ (case (a)), or a corresponding rule $X \hookrightarrow \beta$ to $X \hookrightarrow \alpha$ (case (b)). Then there are the following cases:

- $X \in W_k$, in which case $X \in U$ by induction. If X belongs either to $\Diamond$, or to $\bigcirc$, then there is a rule $X \hookrightarrow \gamma$ for some $\gamma \in U$ by the definition of U. If X belongs to $\Box$, then for all rules of the form $X \hookrightarrow \gamma$ we have $\gamma \in U$.

- $\beta = Y\gamma$ where $Y \in W_k$. Then $X \hookrightarrow Y\gamma$, and by induction $Y\gamma \in U$;

- $\beta = Y'$ where $Y' \in W_k$. Then, by induction, $Y' \in U$ and hence $Y \in V$. Also there is $\ell < k$ such that $X \hookrightarrow YZ$ for some $Z \in \Gamma$ is a rule of Δ_ℓ (assume that ℓ is the greatest number with this property). Then $Z \in W_\ell$, which implies that $Z \in U$. Hence, $X \hookrightarrow YZ$ where $YZ \in U$.

For the cases (2) (a) and (b): If $X \hookrightarrow \varepsilon$ (case (a)), or $\alpha = \varepsilon$ (case (b)), then we are done. Otherwise, let $k < n$ be the greatest number such that $\bar{\Delta}_k$ contains a rule of the form $X' \hookrightarrow \beta$ (case (a)), or a corresponding rule $X' \hookrightarrow \beta$ to $X \hookrightarrow \alpha$ (case (b)). Then there are the following cases:

- $X' \in W_k$, in which case $X \in V$ by induction. If X belongs either to $\Diamond$, or to $\bigcirc$, then there is a rule $X \hookrightarrow \gamma$ for some $\gamma \in V$ by the definition of V. If X belongs to $\Box$, then for all rules of the form $X \hookrightarrow \gamma$ we have $\gamma \in V$.

- $\beta = YZ'$ where $Y \in W_k$. Then $X \hookrightarrow YZ$, and by induction $YZ \in U$;

- $\beta = Y'$ where $Y' \in W_k$. Then, by induction, $Y' \in U$ and hence $Y \in V$. If $X \hookrightarrow Y$ in $\bar{\Delta}$, then we are done. Otherwise, there is $\ell < k$ such that $X' \hookrightarrow YZ'$ for some $Z' \in \Gamma'$ is a rule of Δ_ℓ (assume that ℓ is the greatest number with this property). Then $Z' \in W_\ell$, which implies that $Z \in V$. Hence, $X \hookrightarrow YZ$ where $YZ \in V$.

$\Box$

Let $\bar{X} \in W_n$ and let $\pi \in \Pi$ be a $(T_\varepsilon, =1)$-winning strategy for $\bar{X}$ in $\bar{\Delta}_n$. Let us define a new strategy $\bar{\pi}$ which is $(T_\varepsilon, =1)$-winning for $\bar{X}$ in $\bar{\Delta}$.

Let us execute, in parallel, plays in both $\bar{\Delta}$ and $\bar{\Delta}_n$ initiated in $\bar{X}$ as follows: At each stage of the execution, π chooses a transition of $\bar{\Delta}_n$ and $\bar{\pi}$ chooses the transition of $\bar{\Delta}$ corresponding to the transition chosen by π, and players $\Box$ and $\bigcirc$ choose a transition in $\bar{\Delta}$ and then the corresponding transition in $\bar{\Delta}_n$, until one of the following events occurs:

i. the current configuration of $\bar{\Delta}$ belongs either to $\Diamond$, or to $\bigcirc$, and there are no transitions in $\bar{\Delta}_n$.

ii. the current configuration of $\bar{\Delta}$ belongs to $\Box$, and player $\Box$ chooses a transition for which there is no corresponding transition in $\bar{\Delta}_n$.

If i. occurs, then the current configuration of $\bar{\Delta}$ is $[T_\varepsilon]^{=1}_{\Diamond}$-winning in $\bar{\Delta}$ (as we show below), and if ii. occurs, then the next configuration is $[T_\varepsilon]^{=1}_{\Diamond}$-winning in $\bar{\Delta}$. In both cases i. and ii. $\bar{\pi}$ continues as one of these $(T_\varepsilon, =1)$-winning strategies.

Assume first that one of the events i. and ii. occurs at some point. Then we have the following cases:

- If the current configuration of $\bar{\Delta}_n$ has the form $Y\beta$ where $Y \in \Gamma$, then the current configuration of $\bar{\Delta}$ has the form $Y\bar{\beta}$ for some $\bar{\beta}$. Now if i. occurs, then by Claim A.8 (1) (a) there is a rule $Y \hookrightarrow \alpha$ in $\bar{\Delta}$ such that $\alpha \in U$, and hence $Y\bar{\beta}$ is $[T_\varepsilon]^{=1}_{\Diamond}$-winning. On the other hand, if ii. occurs, then by Claim A.8 (1) (b) the rule chosen by $\Box$ has the form $Y \hookrightarrow \alpha$ where $\alpha \in U$, and hence the next configuration $\alpha\bar{\beta}$ of $Y\bar{\beta}$ is $[T_\varepsilon]^{=1}_{\Diamond}$-winning.

- If the current configuration of $\bar{\Delta}_n$ has the form $Y'\beta$ where $Y' \in \Gamma'$, then the current configuration of $\bar{\Delta}$ has the form $\bar{Y}\bar{\beta}$ for some $\bar{\beta}$ and $\bar{Y} \in \{Y, Y'\}$. By Claim A.7, Y' on the top of the stack of $\bar{\Delta}_n$ means that there are the following cases:

  - The current configuration has been reached along the following sequence of configurations: $Z\beta, K'\beta, \ldots, Y'\beta$. By Claim A.7 (1) (b), $\bar{\beta} \in U$.

    Now if i. occurs, then by Claim A.8 (2) (a) there is a rule $Y \hookrightarrow \alpha$ in $\bar{\Delta}$ such that $\alpha \in V$. Hence, $\bar{Y}\bar{\beta}$ is $[T_\varepsilon]^{=1}_{\Diamond}$-winning.

    On the other hand, if ii. occurs, then by Claim A.8 (2) (b), the rule chosen by $\Box$ is of the form $X \hookrightarrow \alpha$ where $\alpha \in V$. Thus $\alpha\bar{\beta} \in U$, and the next configuration $\alpha\bar{\beta}$ of $\bar{Y}\bar{\beta}$ is $[T_\varepsilon]^{=1}_{\Diamond}$-winning.

  - The current configuration has been reached along the following sequence of configurations: $\bar{X}, K', \ldots, Y'$ (here $\bar{X} \in \Gamma'$ is the initial symbol). By Claim A.7 (2) (b), $\bar{\beta} \in V$.

    Now if i. occurs, then by Claim A.8 (2) (a) there is a rule $Y \hookrightarrow \alpha$ in $\bar{\Delta}$ such that $\alpha \in V$. Hence, $\bar{Y}\bar{\beta}$ is $[T]^{=1}_{\Diamond}$-winning. However, because $\bar{X} \in \Gamma'$, we have that $\bar{\beta} \in \Gamma^*\Gamma'$. It follows that $\bar{Y}\bar{\beta}$ is $[T_\varepsilon]^{=1}_{\Diamond}$-winning.

    On the other hand, if ii. occurs, then by Claim A.8 (2) (b), the rule chosen by $\Box$ is of the form $Y \hookrightarrow \alpha$ where $\alpha \in V$. Thus $\alpha\bar{\beta} \in V$. Similarly as above, $\alpha\bar{\beta} \in \Gamma^*\Gamma'$, which means that the next configuration $\alpha\bar{\beta}$ of $\bar{Y}\bar{\beta}$ is $[T_\varepsilon]^{=1}_{\Diamond}$-winning.

Assume that neither i., nor ii. occurs anywhere. Note that corresponding transitions belonging to player $\bigcirc$ have the same probability. It follows, that if neither i., nor ii. occurs, then the probability of avoiding T_ε is the same in both plays, and hence non-zero in the play played according to $\bar{\pi}$ in $\bar{\Delta}$. On the other hand, $\bar{\pi}$ wins every play where either i., or ii. occurs, and hence $\bar{\pi}$ is $(T_\varepsilon, =1)$-winning for $\bar{X}$ in $\bar{\Delta}$. This proves that $\mathcal{W} \subseteq [T_\varepsilon]_\Diamond^{=1}$ in $\bar{\Delta}$.

It remains to prove that $([T_\varepsilon]_\Diamond^{=1} \cap (\Gamma \cup \Gamma')) \subseteq \mathcal{W}$ in $\bar{\Delta}$. Let $\bar{X} \in [T_\varepsilon]_\Diamond^{=1}$ in $\bar{\Delta}$. Let $\bar{\pi} \in \Pi$ be a $(T_\varepsilon, =1)$-winning strategy for $\bar{X}$ in $\bar{\Delta}$. Let us define a new strategy π that is $(T_\varepsilon, =1)$-winning for $\bar{X}$ in $\bar{\Delta}_n$.

Let us execute, in parallel, plays in both $\bar{\Delta}$ and $\bar{\Delta}_n$ initiated in $\bar{X}$ as follows: At each stage of the execution, $\bar{\pi}$ chooses a transition of $\bar{\Delta}$ and π chooses the transition of $\bar{\Delta}_n$ corresponding to the transition chosen by $\bar{\pi}$, and players $\Box$ and $\bigcirc$ choose a transition in $\bar{\Delta}_n$ and the corresponding transition in $\bar{\Delta}$, until the following event occurs: The current configuration of $\bar{\Delta}_n$ does not have outgoing transitions, and the head of this configuration is not in Γ_T. In this case there is only one transition from the current configuration to itself, and such a play is won by player $\Diamond$. Assume that this event does not occur. It follows from Claim A.7 (using simple induction on the number of stages of the play) that if X is the head of the current configuration of $\bar{\Delta}$, then either X, or X' is the head of the current configuration of $\bar{\Delta}_n$. Observe that by Claim A.7, there is always a corresponding transition in $\bar{\Delta}$ to a transition in $\bar{\Delta}_n$. The probability of avoiding T_ε is the same in both plays, and hence non-zero in the play played according to π in $\bar{\Delta}_n$. Hence, $([T_\varepsilon]_\Diamond^{=1} \cap (\Gamma \cup \Gamma')) \subseteq \mathcal{W}$ in $\bar{\Delta}$. $\Box$

The only claim of Theorem 4.13 which has not yet been proven is the effective constructability of a regular $(=1)$-winning strategy for player $\Box$. This proof is given in the following subsection.

A.3 $(T_\varepsilon, =1)$-winning strategy for player $\Box$

Let us first assume that for every $X, Y \in \Gamma$ there is at most one rule of the form $X \hookrightarrow Y\alpha$. Let us consider $n \geq 1$ such that the above algorithm stops in $n+1$-th iteration (which means that $W_{n+1} = \emptyset$). Let us consider $\bar{\Delta}_n$. By Lemma A.2, there is a strategy $\bar{\sigma}$ which is $(T_\varepsilon, =1)$-winning for every symbol $\bar{X}$ of $\bar{\Delta}_n$, because there are no witnesses in $\bar{\Delta}_n$.

Let us define a strategy σ in Δ as follows: Let us execute, in parallel, plays in $\bar{\Delta}_n$ and Δ initiated in X' and in X, respectively, as follows: At each stage, player Box plays in $\bar{\Delta}_n$

according to $\bar{\sigma}$ and in Δ chooses the corresponding transition (thus defining the σ). Each of the players $\Diamond$ and $\bigcirc$ chooses a transition in Δ and then the corresponding transition in $\bar{\Delta}_n$.

Claim A.9. *Assume that the current configuration of Δ is α and the current configuration of $\bar{\Delta}_n$ is β. Then*

$$\beta = \beta_1 X_1' \beta_2 X_2' \cdots \beta_k X_k' \beta_{k+1}$$

for some $X_1, \ldots, X_k \in [T]_{\Box}^{=1} \cap \Gamma$ in Δ and $\beta_1, \ldots, \beta_{k+1} \in [T_\varepsilon]_{\Box}^{=1}$ in Δ, and

$$\alpha = \beta_1 X_1 \alpha_1 Y_1 \beta_2 X_2 \alpha_2 Y_2 \cdots \beta_k X_k \alpha_k Y_k \beta_{k+1}$$

for some $\alpha_1, \ldots, \alpha_k \in [T]_{\Diamond}^{=1}$ in Δ and $Y_1, \ldots, Y_k \in [T_\varepsilon]_{\Diamond}^{=1} \cap \Gamma$ in Δ.

Proof. The fact that β has the desired form follows immediately from the definition of the algorithm and Claim A.7 (by induction, X has the desired form, and the form is preserved by transitions of $\bar{\Delta}_n$). Moreover, all symbols of $\bar{\Delta}_n$ are in $\bar{\Gamma} \setminus \mathcal{W}$, and hence by Proposition A.6, $X_1, \ldots, X_k \in [T]_{\Box}^{=1} \cap \Gamma$ and $\beta_1, \ldots, \beta_{k+1} \in [T_\varepsilon]_{\Box}^{=1}$. Similarly, by Claim A.7, the configuration α has the above form. $\Box$

Thus, by Claim A.9, the strategy σ is correctly defined (note that if Y is a head of the current configuration of Δ, then either X, or X' is the head of the current configuration of $\bar{\Delta}_n$). It is easy to see that the probability of reaching T_ε is the same in both plays, and hence the probability of reaching T_ε is 1 in the play played by σ in Δ.

In the rest we show that σ is effectively regular. Let $\mathscr{A}$ be the finite state automaton implementing the strategy $\bar{\sigma}$. The alphabet of $\mathscr{A}$ is a subset $\Sigma \subseteq \Gamma \cup \Gamma'$. Let us denote $Q_{\mathscr{A}}$ the set of states of $\mathscr{A}$. Given two states $p, q \in Q_{\mathscr{A}}$, we write $p \xrightarrow{\bar{X}}_{\mathscr{A}} q$ whenever the automaton $\mathscr{A}$ moves from p to q after reading $\bar{X}$.

We define a new automaton $\mathscr{B}$ over the alphabet Γ such that the set of states of $\mathscr{B}$ contains $Q_{\mathscr{A}}$. For every pair of configurations β and α of the form described in Claim A.9 we have the following: There is a state $q \in Q_{\mathscr{A}}$ such that $\mathscr{B}$ enters q after reading α, and $\mathscr{A}$ enters q after reading β. Basically, the automaton reads the configuration α bottom up. While reading β_{k+1}, the automaton $\mathscr{B}$ behaves similarly as $\mathscr{A}$. Once the symbol Y_k occurs, the automaton $\mathscr{B}$ (being in a state p) changes its state to a newly added state a_p. In a_p, the automaton $\mathscr{B}$ waits till X_k occurs. Then $\mathscr{B}$ proceeds to a state q such that $p \xrightarrow{X_k'} q$. Consequently, this process is repeated with β_k, Y_{k-1} and X_{k-1}, etc. By Claim A.9, only the configurations of the form described in Claim A.9 are relevant with respect to a winning strategy. The strategy based on $\mathscr{B}$ is described below.

The automaton $\mathscr{B}$ is formally defined as follows:

- the alphabet of $\mathscr{B}$ is Γ
- the set of states of $\mathscr{B}$ consists of all states of $Q_{\mathscr{A}}$, and moreover, for every state $q \in Q_{\mathscr{A}}$, we add an auxiliary state a_q
- the transition function of $\mathscr{B}$ is defined as follows:
  - for every $p, q \in Q_{\mathscr{A}}$ and every $X \in \Sigma \cap \Gamma$, we put $p \xrightarrow{X}_{\mathscr{B}} q$ iff $p \xrightarrow{X}_{\mathscr{A}} q$;
  - for every $p \in Q_{\mathscr{A}}$ and every $Y \in [T_\varepsilon]^{=1}_{\Diamond} \cap \Gamma$ we put $p \xrightarrow{Y}_{\mathscr{B}} a_p$
  - for every $p \in Q_{\mathscr{A}}$ and every $Z \in [T]^{=1}_{\Diamond} \cap \Gamma$, we put $a_p \xrightarrow{Z}_{\mathscr{B}} a_p$
  - for every $q \in Q_{\mathscr{A}}$ and every $X \in [T]^{=1}_{\Box} \cap \Gamma$, we put $a_p \xrightarrow{X}_{\mathscr{B}} q$ iff $p \xrightarrow{X'}_{\mathscr{A}} q$

Observe that for every configuration of $\bar{\Delta}_n$ of the form $\bar{Y}\beta$ where $\bar{Y} \in \bar{\Gamma}_\Box$ the state q entered by $\mathscr{A}$ after reading $\bar{Y}\beta$ determines uniquely the head $\bar{Y}$ (we denote $\bar{h}_q = \bar{Y}$) and a distribution $\bar{d}_q$ on the outgoing transitions of $\bar{Y}$. Let $X\alpha \in \Gamma_\Box\Gamma^*$ and let $q \in Q_{\mathscr{A}}$ be the state reached by $\mathscr{B}$ after reading $X\alpha$. Note that either $\bar{h}_q = X$, or $\bar{h}_q = X'$. The distribution $\bar{d}_q$ induces a distribution d_q on outgoing transitions of X as follows: for every rule $X \hookrightarrow \gamma$, we define $d_q(X \hookrightarrow \gamma)$ to be the probability assigned by $\bar{d}_q$ to the unique rule of $\bar{\Delta}_n$ corresponding to $X \hookrightarrow \gamma$ (and 0 if there is no such a corresponding rule). Consequently, we have that $\sigma(X\alpha) = d_q$.

Finally, let us drop the above assumption about uniqueness of transitions. First, we transform the BPA game as follows: For every $Y, Z \in \Gamma$, add a fresh new symbol of the form Y_Z to $\Gamma_\bigcirc$. Moreover, remove every rule of the form $X \hookrightarrow YZ$ and add rules $X \hookrightarrow Y_Z$ and $Y_Z \xhookrightarrow{1} YZ$. Let Δ' be the resulting BPA game. It is easy to see that Δ' satisfies the uniqueness assumption. Hence, there is a regular strategy σ' that is $(T_\varepsilon, =1)$-winning for every $X \in [T_\varepsilon]^{=1}_{\Box}$ in Δ'. Let σ be a strategy in Δ defined as follows: Let $X\alpha \in \Gamma_\Box\Gamma^*$. For every rule of the form $X \hookrightarrow \gamma$, where $|\gamma| \leq 1$, we define $\sigma(X\alpha)(X \hookrightarrow \gamma) = \sigma'(X\alpha)(X \hookrightarrow \gamma)$. For every rule of the form $X \hookrightarrow YZ$ we define $\sigma(X\alpha)(X \hookrightarrow YZ) = \sigma'(X\alpha)(X \hookrightarrow Y_Z)$. It is easy to see that σ is $(T_\varepsilon, =1)$-winning for every $X \in [T_\varepsilon]^{=1}_{\Box}$ in Δ. Clearly, σ is regular with the same finite state automaton as σ'.