

# Reachability in Recursive Markov Decision Processes<sup>\*</sup>

Tomáš Brázdil, Václav Brožek, Vojtěch Forejt, and Antonín Kučera

Faculty of Informatics, Masaryk University,  
Botanická 68a, 60200 Brno,  
Czech Republic.  
{brazdil,xbrozek,xforejt,kucera}@fi.muni.cz

**Abstract.** We consider a class of infinite-state Markov decision processes generated by stateless pushdown automata. This class corresponds to  $1\frac{1}{2}$ -player games over graphs generated by BPA systems or (equivalently) 1-exit recursive state machines. An *extended reachability objective* is specified by two sets  $S$  and  $T$  of *safe* and *terminal* stack configurations, where the membership to  $S$  and  $T$  depends just on the top-of-the-stack symbol. The question is whether there is a suitable strategy such that the probability of hitting a terminal configuration by a path leading only through safe configurations is equal to (or different from) a given  $x \in \{0, 1\}$ . We show that the qualitative extended reachability problem is decidable in polynomial time, and that the set of all configurations for which there is a winning strategy is effectively regular. More precisely, this set can be represented by a deterministic finite-state automaton with a fixed number of control states. This result is a generalization of a recent theorem by Etessami & Yannakakis which says that the qualitative termination for 1-exit RMDPs (which exactly correspond to our  $1\frac{1}{2}$ -player BPA games) is decidable in polynomial time. Interestingly, the properties of winning strategies for the extended reachability objectives are quite different from the ones for termination, and new observations are needed to obtain the result. As an application, we derive the **EXPTIME**-completeness of the model-checking problem for  $1\frac{1}{2}$ -player BPA games and qualitative PCTL formulae.

## 1 Introduction

$1\frac{1}{2}$ -player games (or Markov decision processes) are a fundamental model in the area of system design and control optimization [11, 8]. Formally, a  $1\frac{1}{2}$ -player game  $G$  is a directed graph where the vertices are split into two disjoint subsets  $V_{\square}$  and  $V_{\circ}$ . For every  $v \in V_{\circ}$ , there is a fixed probability distribution over the set of its outgoing transitions. A *play* is initiated by putting a token on some vertex. The token is then moved from vertex to vertex by one “real” player  $\square$  (controller) and one “virtual” player  $\circ$  (stochastic environment), who are responsible for selecting outgoing transitions in the vertices of  $V_{\square}$  and  $V_{\circ}$ , respectively. Player  $\circ$  does not make a real choice, but selects his next move randomly according to the fixed probability distribution over the outgoing transitions. A *strategy* specifies how player  $\square$

---

<sup>\*</sup> Supported by the research center Institute for Theoretical Computer Science (ITI), project No. 1M0545.

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 transitions may be chosen deterministically or randomly (*deterministic* ( $D$ ) and *randomized* ( $R$ ) strategies). In the case of randomized strategies, player  $\square$  chooses a probability distribution on the set of outgoing transitions. Note that deterministic strategies can be seen as restricted randomized strategies, where one of the outgoing transitions has probability 1. Each strategy  $\sigma$  determines a unique Markov chain  $G(\sigma)$  where the states are finite paths in  $G$ , and  $wu \rightarrow wuu'$  with probability  $x$  iff  $(u, u')$  is a transition in the game and  $x$  is the probability chosen by player  $\square$ , or the fixed probability of the transition  $(u, u')$  when  $u \in V_{\square}$ . A *winning objective* for player  $\square$  is some property of Markov chains that is to be achieved. A *winning strategy* is a strategy that achieves the objective. In the context of “classical” MDP theory, winning objectives are typically related to long-time characteristics such as the expected total reward, the expected reward per transition, etc. [11, 8]. In the context of formal verification, winning objectives are often specified as formulae of suitable temporal logics and their probabilistic variants such as PCTL or PCTL\* [9]. For games with finitely many vertices, the corresponding decision algorithms have been designed [9, 2, 1] and also implemented in verifications tools such as PRISM (see, e.g., [10]). Recently, the scope of this study has been extended to a class of *infinite-state* games generated by *recursive state machines* ( $RSM$ ) [6, 7]. Intuitively, a RSM is a finite collection of finite-state automata which can call each other in a recursive fashion, maintaining the (unbounded) stack of activation records. RSM are semantically equivalent to *pushdown automata* ( $PDA$ ), and there are effective linear-time translations between the two models. A given RSM can be encoded in PDA syntax by storing the collection of finite-state automata in the control unit, and the recursive calls/returns are modeled by pushing/popping symbols onto/from the stack. An important subclass of RSM are *1-exit RSM*, where each finite-state automaton in the collection terminates in exactly one state. This means that no information can be returned back to the caller. In PDA terms, this means that whenever a given stack symbol  $X$  is popped from the stack, the same control state  $p_X$  is entered. Hence, the finite-state control unit can be encoded directly into the stack alphabet and simulated in top-of-the-stack symbol. Thus, 1-exit RSM can effectively be represented as *stateless* PDA, which are also denoted BPA in the context of concurrency theory.

Now we briefly summarize some of the results presented in [6, 7]. To be able to give a clear comparison with our work, we reformulate these results in PDA/BPA terminology. A *termination objective* is specified by two control states  $p, q$  and one stack symbol  $X$  of a given PDA. The task of player  $\square$  is to maximize/minimize the probability of hitting  $q\varepsilon$  from  $pX$  (each “head”  $rY$  in a given PDA is either probabilistic or non-deterministic; transitions from probabilistic heads are chosen randomly according to a fixed distribution, while the transitions from non-deterministic heads can be chosen by player  $\square$ ). In the case of BPA, there are no control states and the termination objective is specified simply by the stack symbol which is to be removed.

In [6, 7], it has been shown that optimal minimizing/maximizing strategies in general  $1\frac{1}{2}$ -player PDA games with termination objectives do not always exist, and that the problem whether there is a strategy such that termination is achieved

with probability 1 is undecidable. The situation is different for  $1\frac{1}{2}$ -player BPA games, where the optimal minimizing/maximizing strategies do exist, and can be constructed so that they depend only on top-of-the-stack symbol of a given configuration. Hence, the optimal strategies are stackless, memoryless, and deterministic (SMD). Furthermore, the corresponding minimal/maximal termination probabilities are expressible as the least solution of an effectively constructible system of non-linear recursive equations. Since the least solution of this system can effectively be expressed in first-order theory of the reals, this entails the decidability of the *quantitative* termination problem, i.e., the question whether the minimal/maximal achievable termination probability is bounded by a given constant. For the *qualitative* subcase (i.e., the problem whether the minimal/maximal achievable termination probability is equal to one), polynomial-time algorithms have been designed. On the other hand, in [6] it was noted that model-checking  $1\frac{1}{2}$ -player BPA games against *qualitative LTL objectives* is already undecidable.

**Our contribution:** In this paper we consider  $1\frac{1}{2}$ -player BPA games with *extended reachability objectives (EROs)*. An ERO is specified by two sets of *safe* and *terminal* stack symbols. A configuration is safe/terminal iff its top-of-the-stack symbol is safe/terminal. A run  $w$  satisfies a given ERO iff  $w$  visits a terminal configuration and all configurations preceding this visit are safe. The goal of player  $\square$  is to minimize/maximize the probability of all runs satisfying a given ERO. Note that termination objectives can easily be encoded as EROs (this may require a new bottom-of-the-stack symbol). However, the properties of EROs are surprisingly different from the ones of termination objectives (in contrast, methods for termination can easily be extended to EROs for fully probabilistic PDA [4]). We show that optimal maximizing strategies may not exist at all, and even if they do exist, they are not necessarily SMD. The optimal minimizing strategies are guaranteed to exist, but are not necessarily SMD. The method of expressing the minimal/maximal termination probability by a system of non-linear equations used in [6] cannot be easily extended to EROs, and the reasons seem to be fundamental.

At the core of our paper are results about *qualitative* EROs. We show that the sets of all configurations for which there exists a strategy such that the probability of all runs satisfying a given ERO is equal to zero (equal to one, larger than zero, less than one, resp.) are regular and the corresponding finite-state automata can be constructed in polynomial time. In our algorithms, we use the results about qualitative termination as “black boxes” and concentrate on problems that are specific to EROs. We note that the subcase “equal to one”, and particularly the subcase “less than one”, require non-trivial methods and observations.

As an application, we design an exponential-time model-checking algorithm for  $1\frac{1}{2}$ -player BPA games and the qualitative fragment of the logic PCTL. More precisely, our algorithm is polynomial in the size of a given BPA and exponential in the size of a given formula (hence, the algorithm becomes polynomial for each fixed formula). Since there is a matching **EXPTIME** lower bound [3], we yield the **EXPTIME**-completeness of the problem. As a consequence we also obtain the **EXPTIME**-completeness of the model-checking problem for fully probabilistic BPA and qualitative PCTL (fully probabilistic BPA correspond to a subclass of  $1\frac{1}{2}$ -player BPA games where all heads are probabilistic). This problem has been studied in [4, 3], but the best known upper complexity bound was **EXSPACE**.

Finally, let us note that since model-checking  $1\frac{1}{2}$ -player BPA games against qualitative LTL properties is already undecidable [6], our result cannot be extended to the qualitative fragment of PCTL\*.

## 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.

We start by recalling basic notions of probability theory. Let  $A$  be a finite or countably infinite set. A *probability distribution* on  $A$  is a function  $f : A \rightarrow [0, 1]$  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  $\sigma$ -*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  $\sigma$ -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 *probabilistic 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  $(X, \mathcal{F})$ .

**Markov chains.** A *Markov chain* is a triple  $\mathcal{M} = (M, \rightarrow, Prob)$  where  $M$  is a finite or countably infinite set of *states*,  $\rightarrow \subseteq M \times M$  is a *transition relation* such that for every  $s \in M$  there is some transition  $s \rightarrow t$ , 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 also write  $s \xrightarrow{x} t$  instead of  $Prob(s \rightarrow t) = x$ . A *path* in  $\mathcal{M}$  is a finite or infinite sequence  $w = s_0, s_1, \dots$  of states such that  $s_i \rightarrow s_{i+1}$  for every  $i$ . We also use  $w(i)$  to denote the state  $s_i$  of  $w$ , and  $w_i$  to denote the path  $s_i, s_{i+1}, \dots$  (by writing  $w(i) = s$  or  $w_i$  we implicitly impose the condition that the length of  $w$  is at least  $i + 1$ ). A *run* is an infinite path. The sets of all finite paths and all runs of  $\mathcal{M}$  are denoted  $FPath(\mathcal{M})$  and  $Run(\mathcal{M})$ , respectively. Similarly, the sets of all finite paths and runs that start in a given  $s \in M$  are denoted  $FPath(\mathcal{M}, s)$  and  $Run(\mathcal{M}, s)$ , respectively.

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 probabilistic space  $(Run(\mathcal{M}, s), \mathcal{F}, \mathcal{P})$  where  $\mathcal{F}$  is the  $\sigma$ -field generated by all basic cylinders  $Run(\mathcal{M}, w)$  where  $w$  starts with  $s$ , and  $\mathcal{P} : \mathcal{F} \rightarrow [0, 1]$  is the unique probability function such that  $\mathcal{P}(Run(\mathcal{M}, w)) = \prod_{i=0}^{m-1} x_i$  where  $w = s_0, \dots, s_m$  and  $s_i \xrightarrow{x_i} s_{i+1}$  for every  $0 \leq i < m$  (if  $m = 0$ , we put  $\mathcal{P}(Run(\mathcal{M}, w)) = 1$ ).

For all  $S, T \subseteq M$  and  $s \in M$ , we define the sets

- $Run(\mathcal{M}, s, SUT) = \{w \in Run(\mathcal{M}, s) \mid \exists j \geq 0 : w(j) \in T \wedge \forall i < j : w(i) \in S\}$
- $Run(\mathcal{M}, s, FT) = \{w \in Run(\mathcal{M}, s) \mid \exists j \geq 0 : w(j) \in T\}$
- $Run(\mathcal{M}, s, \neg FT) = \{w \in Run(\mathcal{M}, s) \mid \forall j \geq 0 : w(j) \notin T\}$
- $Run(\mathcal{M}, s, \mathcal{X}S) = \{w \in Run(\mathcal{M}, s) \mid w(1) \in S\}$

**Games, strategies, and objectives.** A  $1\frac{1}{2}$ -player game (or *Markov decision process*) is a tuple  $G = (V, E, (V_\square, V_\circ), Prob)$  where  $V$  is a finite or countably infinite set of *vertices*,  $E \subseteq V \times V$  is a set of *transitions*,  $(V_\square, V_\circ)$  is a partition of  $V$ , and  $Prob$  is a *probability assignment* which to each  $v \in V_\circ$  assigns a positive probability distribution on the set of its outgoing transitions. For technical convenience, we assume that each vertex has at least one outgoing transition. We say that  $G$  is *finitely-branching* if for each  $v \in V$  there are only finitely many  $u \in V$  such that  $(v, u) \in E$ .

The game is played by a player  $\square$  who selects the moves in the  $V_\square$  vertices, and a “virtual” player  $\circ$  who selects the moves in the  $V_\circ$  vertices according to the corresponding probability distribution.

A *strategy* for player  $\square$  is a function  $\sigma$  which to each  $wv \in V^*V_\square$  assigns a probability distribution on the set of outgoing transitions of  $v$ . We say that a strategy  $\sigma$  is *memoryless* ( $M$ ) if  $\sigma(wv)$  depends just on the last vertex  $v$ , and *deterministic* ( $D$ ) if  $\sigma(wv)$  is a Dirac distribution for each  $wv \in V^*V_\square$ . 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.

*Remark 1.* Each MD strategy  $\sigma$  determines a unique function  $f_\sigma : V_\square \rightarrow V$ . Conversely, each function  $f : V_\square \rightarrow V$  such that  $(v, f(v)) \in E$  for every  $v \in V_\square$  determines a unique MD strategy  $\sigma_f$ .

Each strategy  $\sigma$  for player  $\square$  determines a unique *play* of the game  $G$ , which is a Markov chain  $G(\sigma)$  where  $V^+$  is the set of states, and  $wu \xrightarrow{x} wuu'$  iff  $(u, u') \in E$  and one of the following conditions holds:

- $u \in V_\circ$  and  $Prob(u, u') = x$ ;
- $u \in V_\square$  and  $\sigma(wu)$  assigns  $x$  to  $(u, u')$ .

For every  $w \in Run(G(\sigma))$  and every  $i \in \mathbb{N}_0$ , we define  $w[i]$  to be the last vertex of  $w(i)$  (realize that  $w(i)$  is a finite sequence of vertices of the game  $G$ ).

**The logic PCTL.** The logic PCTL, the probabilistic extension of CTL, was introduced by Hansson & Jonsson in [9]. Let  $Ap = \{p, q, \dots\}$  be a countably infinite set of *atomic propositions*. The syntax of PCTL formulae is given by the following abstract syntax equation:

$$\Phi ::= p \mid \Phi_1 \wedge \Phi_2 \mid \neg \Phi \mid \mathcal{X}^{\bowtie \varrho} \Phi \mid \Phi_1 \mathcal{U}^{\bowtie \varrho} \Phi_2$$

Here  $p \in Ap$ ,  $\varrho \in [0, 1]$ , and  $\bowtie \in \{\leq, <, \geq, >, =, \neq\}$ .

Let  $G = (V, E, (V_\square, V_\circ), Prob)$  be a  $1\frac{1}{2}$ -player game, and let  $\nu : Ap \rightarrow 2^V$  be a *valuation*. The semantics of PCTL is defined below.

$$\begin{aligned} \llbracket p \rrbracket^\nu &= \nu(p) \\ \llbracket \Phi_1 \wedge \Phi_2 \rrbracket^\nu &= \llbracket \Phi_1 \rrbracket^\nu \cap \llbracket \Phi_2 \rrbracket^\nu \\ \llbracket \neg \Phi \rrbracket^\nu &= V \setminus \llbracket \Phi \rrbracket^\nu \\ \llbracket \mathcal{X}^{\bowtie \varrho} \Phi \rrbracket^\nu &= \{u \in V \mid \forall \sigma \in HR : \mathcal{P}(Run(G(\sigma), u, \mathcal{X} \llbracket \Phi \rrbracket^\nu)) \bowtie \varrho\} \\ \llbracket \Phi_1 \mathcal{U}^{\bowtie \varrho} \Phi_2 \rrbracket^\nu &= \{u \in V \mid \forall \sigma \in HR : \mathcal{P}(Run(G(\sigma), u, \llbracket \Phi_1 \rrbracket^\nu \mathcal{U} \llbracket \Phi_2 \rrbracket^\nu)) \bowtie \varrho\} \end{aligned}$$

The  $\mathcal{F}^{\boxtimes e}$  and  $G^{\boxtimes e}$  operators are defined in the standard way:  $\mathcal{F}^{\boxtimes e}\Phi$  stands for  $\mathbf{tt} \mathcal{U}^{\boxtimes e}\Phi$ , and  $G^{\boxtimes e}\Phi$  stands for  $\mathbf{tt} \mathcal{U}^{\boxtimes 1-e} \neg\Phi$ , where  $\boxtimes$  is  $<$ ,  $>$ ,  $\leq$ ,  $\geq$ ,  $=$ , or  $\neq$ , depending on whether  $\boxtimes$  is  $>$ ,  $<$ ,  $\geq$ ,  $\leq$ ,  $=$ , or  $\neq$ , respectively.

Various natural fragments of PCTL can be obtained by restricting the PCTL syntax to certain modal connectives and/or certain operator/number combinations. The *qualitative* fragment of PCTL is obtained by restricting the allowed operator/number combinations to ' $\boxtimes 0$ ' and ' $\boxtimes 1$ '. Hence,  $a\mathcal{U}^{<1}b \vee \mathcal{F}^{>0}c$  is a qualitative PCTL formula.

**BPA games.** A  $1\frac{1}{2}$ -player BPA game is a tuple  $\Delta = (\Gamma, \hookrightarrow, (\Gamma_{\square}, \Gamma_{\circ}), Prob)$  where  $\Gamma$  is a finite *stack alphabet*,  $\hookrightarrow \subseteq \Gamma \times \Gamma^{\leq 2}$  is a set of *rules* (where  $\Gamma^{\leq 2} = \{w \in \Gamma^* : |w| \leq 2\}$ ) such that for each  $X \in \Gamma$  there is some  $X \hookrightarrow \alpha$ ,  $(\Gamma_{\square}, \Gamma_{\circ})$  is a partition of  $\Gamma$ , and  $Prob$  is a *probability assignment* which to each  $X \in \Gamma_{\circ}$  assigns a rational positive probability distribution on the set of all rules of the form  $X \hookrightarrow \alpha$ .

Each  $1\frac{1}{2}$ -player BPA game  $\Delta = (\Gamma, \hookrightarrow, (\Gamma_{\square}, \Gamma_{\circ}), Prob)$  determines a unique  $1\frac{1}{2}$ -player game  $G_{\Delta} = (\Gamma^*, E_{\Delta}, (\Gamma_{\square}\Gamma^*, \Gamma_{\circ}\Gamma^* \cup \{\varepsilon\}), Prob_{\Delta})$  where the transitions of  $E_{\Delta}$  are determined as follows:  $\varepsilon \rightarrow \varepsilon$ , and  $X\beta \rightarrow \alpha\beta$  iff  $X \hookrightarrow \alpha$ . The probability assignment  $Prob_{\Delta}$  is the natural extension of  $Prob$ , i.e.,  $Prob_{\Delta}(X\beta \rightarrow \alpha\beta) = Prob(X \hookrightarrow \alpha)$ , and  $Prob_{\Delta}(\varepsilon \rightarrow \varepsilon) = 1$ . Note that  $G_{\Delta}$  is finitely branching.

Given a configuration  $X\alpha \in \Gamma^*$ , we put  $head(X\alpha) = X$ .

### 3 $1\frac{1}{2}$ -Player BPA Games with Extended Reachability Objectives

In this section we present several results about  $1\frac{1}{2}$ -player BPA games with extended reachability objectives.

**Definition 2.** Let  $G = (V, E, (V_{\square}, V_{\circ}), Prob)$  be an (arbitrary)  $1\frac{1}{2}$ -player game. An extended reachability objective (ERO) is a pair  $(S, T)$ , where  $S, T \subseteq V$  are the subsets of safe and terminal vertices.

Let  $(S, T)$  be an ERO. For every HR strategy  $\sigma$  and every  $u \in V$  we define the  $\sigma$ -value of  $u$ , denoted  $Val^{\sigma}(u)$ , as follows:

$$Val^{\sigma}(u) = \mathcal{P}(Run(G(\sigma), u, SUT))$$

Furthermore, we define the upper and lower value of  $u$ , denoted  $Val^{+}(u)$  and  $Val^{-}(u)$ , as the sup and inf of the set  $\{Val^{\sigma}(u) \mid \sigma \in HR\}$ , respectively.

If the player  $\square$  wants to maximize (or minimize) the value of a certain vertex  $u$ , he uses a maximizing (or minimizing) strategy. An *optimal maximizing* (or *optimal minimizing*) strategy for a vertex  $u$  is a strategy  $\sigma$  such that  $Val^{\sigma}(u)$  is equal to  $Val^{+}(u)$  (or to  $Val^{-}(u)$ , resp.).

It has been shown in [2] that optimal maximizing/minimizing strategies always exist in  $1\frac{1}{2}$ -player games with *finitely* many vertices; moreover, there are efficiently constructible optimal maximizing/minimizing MD strategies. This does *not* hold for games with infinitely many vertices—one can easily give an example of a game

with countably many vertices where an optimal minimizing strategy does not exist for a certain vertex  $u$ , and

$$1 = \inf_{\sigma \in \text{MD}} \{\mathcal{P}(\text{Run}(G(\sigma), u, \text{SUT}))\} > \inf_{\sigma \in \text{HR}} \{\mathcal{P}(\text{Run}(G(\sigma), u, \text{SUT}))\} = 0$$

We are primarily interested in finitely-branching games, where the situation is somewhat specific. The following proposition gives a full characterization:

**Proposition 3.** *Let  $G = (V, E, (V_{\square}, V_{\circ}), \text{Prob})$  be a finitely-branching  $1\frac{1}{2}$ -player game,  $u \in V$ , and  $(S, T)$  an ERO. Then*

- (a) *there is an optimal minimizing MD strategy for  $u$ ;*
- (b)  $\sup_{\sigma \in \text{MD}} \{\mathcal{P}(\text{Run}(G(\sigma), u, \text{SUT}))\} = \sup_{\sigma \in \text{HR}} \{\mathcal{P}(\text{Run}(G(\sigma), u, \text{SUT}))\}$
- (c) *if there is an optimal maximizing HR strategy for  $u$ , then there is also an optimal maximizing MD strategy for  $u$ .*

Proposition 3 admits the non-existence of an optimal maximizing strategy for  $u$ . This can indeed happen, even for  $1\frac{1}{2}$ -player BPA games (see also [6]):

*Example 4.* Let  $\Delta = (\{X, A, D\}, \hookrightarrow, (\{X\}, \{A, D\}), \text{Prob})$  be a  $1\frac{1}{2}$ -player BPA game, where

$$X \hookrightarrow XA, X \hookrightarrow \varepsilon, A \xrightarrow{1/2} D, A \xrightarrow{1/2} \varepsilon, D \xrightarrow{1} D.$$

Let  $S = \{X, A, D\}^*$  and  $T = \{D\}\{A\}^*$ . One can easily verify that  $\text{Val}^+(X) = 1$ . However, for every HR strategy  $\sigma$  we have that  $\text{Val}^{\sigma}(X) < 1$ .

In the rest of this section we restrict our attention to  $1\frac{1}{2}$ -player BPA games. Due to Proposition 3, from now on we can safely consider just MD strategies because they are equivalently powerful as HR strategies in the context of extended reachability objectives.

To simplify our notation, for the rest of this section we fix a  $1\frac{1}{2}$ -player BPA game  $\Delta = (\Gamma, \hookrightarrow, (\Gamma_{\square}, \Gamma_{\circ}), \text{Prob})$ . Realize that all of the previously introduced game-theoretic notions (strategy, upper/lower value, etc.) apply to  $G(\Delta)$ , not directly to  $\Delta$ . In particular, the vertices of  $G(\Delta)$  are stack configurations of  $\Gamma^*$ , which means that MD strategies generally depend on the whole sequence of symbols which form a given vertex. An MD strategy  $\sigma$  is *stackless (SMD)* if it depends just on the top-of-the-stack symbol of a given vertex.

A *termination objective* is an ERO where  $S = \Gamma^*$  and  $T = \{\varepsilon\}$ . In [6, 7], it has been shown that  $1\frac{1}{2}$ -player BPA games with termination objectives have the following properties:

- (a) There are optimal SMD minimizing and maximizing strategies.
- (b) For each  $X \in \Gamma$ , the values  $\text{Val}^+(X)$  and  $\text{Val}^-(X)$  are expressible as the least solution of an effectively constructible system of non-linear equations. This allows to express the values  $\text{Val}^+(X)$  and  $\text{Val}^-(X)$  in  $(\mathbb{R}, +, *, \leq)$ , i.e., first-order arithmetic of the reals.
- (c) The problems whether  $\text{Val}^+(\alpha) = x$ , where  $x \in \{0, 1\}$ , and whether  $\text{Val}^-(\alpha) = x$ , where  $x \in \{0, 1\}$ , are solvable in polynomial time.

In this paper we consider  $1\frac{1}{2}$ -player BPA games with more general EROs, where the sets  $S$  and  $T$  are *simple*:

**Definition 5.** A set  $M \subseteq \Gamma^*$  is simple iff there is a characteristic set  $C(M) \subseteq \Gamma$  such that  $M = \bigcup_{Y \in C(M)} \{Y\}\Gamma^*$ . An ERO  $(S, T)$  is simple if  $S$  and  $T$  are simple.

The properties (a)–(c) stated above do not hold for BPA games with simple EROs. In particular, note the following:

- (A) An optimal minimizing SMD strategy may not exist (though there must be an optimal minimizing MD strategy by Proposition 3). An optimal maximizing strategy may not exist at all (see Example 4), and the existence of an optimal maximizing strategy does not imply the existence of an optimal maximizing SMD strategy.
- (B) The system of non-linear recursive equations which was used in [6] for termination objectives cannot be immediately generalized to simple EROs. Intuitively, the reason is that the optimal minimizing/maximizing strategy in a configuration  $X\alpha$  does not depend just on  $X$  but also on  $\alpha$ , and a small modification of  $\alpha$  may lead to a completely different optimal strategy. This is because one has to “balance” between the probability of termination and the probability of hitting a terminal configuration for each stack symbol, depending on what is achievable for the symbols stored below in the stack.
- (C) For a given  $\alpha \in \Gamma^*$ , the problems whether  $Val^-(\alpha) = 0$ , whether  $Val^+(\alpha) = 0$ , and whether  $Val^-(\alpha) = 1$  are solvable in polynomial time. The decidability of the problem whether  $Val^+(\alpha) = 1$  is left open. Nevertheless, we show that the problem whether there is an optimal maximizing strategy  $\sigma$  such that  $Val^\sigma(\alpha) = 1$  is decidable in polynomial time (remember that  $Val^+(\alpha)$  can be 1 even if no optimal maximizing strategy exists).

The property (A) is demonstrated in the following example:

*Example 6.* Let  $\Delta = (\{X, R\}, \hookrightarrow, (\{X, R\}, \emptyset), Prob)$  be a  $1\frac{1}{2}$ -player BPA game, where  $X \hookrightarrow XR$ ,  $X \hookrightarrow \varepsilon$ ,  $R \hookrightarrow R$ . Let us consider an ERO  $(S, T)$  where  $C(S) = \{X\}$  and  $C(T) = \{R\}$ . Then  $Val^+(X) = 1$  and there an optimal maximizing MD strategy, but there is no optimal maximizing SMD strategy.

Let  $\Delta' = (\{X, Y, Z, H, R\}, \hookrightarrow, (\{X, Y, Z, R\}, \{H\}), Prob)$  be a  $1\frac{1}{2}$ -player BPA game, where

$$X \hookrightarrow YR, Y \hookrightarrow H, Y \hookrightarrow \varepsilon, R \hookrightarrow R, Z \hookrightarrow Z, H \xleftrightarrow{1/2} YZ, H \xleftrightarrow{1/2} R.$$

Let us consider an ERO  $(S, T)$  where  $C(S) = \{X, Y, Z, H\}$  and  $C(T) = \{R\}$ . Then  $Val^-(X) = 1/2$  and there an optimal minimizing MD strategy, but there is no optimal minimizing SMD strategy.

Now we present a sequence of results from which (C) follows as a simple consequence, and which allow to design the model-checking algorithm for  $1\frac{1}{2}$ -player games and qualitative PCTL formulae presented in Section 4.

For the rest of this section, let us fix a simple ERO  $(S, T)$ . Let  $\Delta' = (\Gamma, \rightsquigarrow, (\Gamma_\square, \Gamma_\circ), Prob')$  be a modification of the game  $\Delta$  obtained by replacing all rules of the form  $P \hookrightarrow \alpha$ , where  $P \in C(T) \cup (\Gamma \setminus C(S))$ , with a single rule  $P \rightsquigarrow P$  (the other rules are preserved). One can easily verify that for every strategy  $\sigma$  and every  $\alpha \in \Gamma^*$  we have that  $\mathcal{P}(G_\Delta(\sigma), \alpha, SUT) = \mathcal{P}(G_{\Delta'}(\sigma), \alpha, FT)$ , and this fact is heavily used in the proofs of subsequent theorems where we freely “shift” between  $\Delta$  and  $\Delta'$ .



**Theorem 7.** Let  $[SU^{>0}T]$  be the set of all  $\alpha \in \Gamma^*$  for which there is a strategy  $\sigma$  such that  $\mathcal{P}(G_{\Delta}(\sigma), \alpha, SUT) > 0$ . Then there are  $\mathcal{A}, \mathcal{B} \subseteq \Gamma$  computable in polynomial time such that  $[SU^{>0}T] = \mathcal{A}^* \mathcal{B} \Gamma^*$ .

*Proof.* Let  $\mathcal{A} = \{X \in \Gamma \mid X \rightsquigarrow^* \varepsilon\}$  and  $\mathcal{B} = \{X \in \Gamma \mid X \rightsquigarrow^* R\beta, \text{ where } R \in C(T) \text{ and } \beta \in \Gamma^*\}$ . Now observe that  $\alpha \in [SU^{>0}T]$  iff  $\alpha \rightsquigarrow^* R\beta$  for some  $R \in C(T)$  and  $\beta \in \Gamma^*$  iff  $\alpha \in \mathcal{A}^* \mathcal{B} \Gamma^*$ . The sets  $\mathcal{A}, \mathcal{B}$  can be computed using standard algorithms for PDA reachability.  $\square$

**Theorem 8.** Let  $[SU^{=0}T]$  be the set of all  $\alpha \in \Gamma^*$  for which there is a strategy  $\sigma$  such that  $\mathcal{P}(G_{\Delta}(\sigma), \alpha, SUT) = 0$ . Then there are  $\mathcal{A}, \mathcal{B} \subseteq \Gamma$  computable in polynomial time such that  $[SU^{=0}T] = \mathcal{B}^* \cup (\mathcal{B}^* \mathcal{A} \Gamma^*)$ .

*Proof.* We define the sets  $\mathcal{A}$  and  $\mathcal{B}$  as follows:

- $X \in \mathcal{A}$  iff there is a strategy  $\sigma$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma), X, \mathcal{F}(T \cup \{\varepsilon\}))) = 0$
- $X \in \mathcal{B}$  iff there is a strategy  $\sigma$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma), X, \mathcal{F}T)) = 0$

It is easy to verify that  $\mathcal{A}$  and  $\mathcal{B}$  satisfy the property that  $[SU^{=0}T] = \mathcal{B}^* \cup \mathcal{B}^* \mathcal{A} \Gamma^*$ .

We show that the sets  $\mathcal{A}$  and  $\mathcal{B}$  can be computed as the greatest fixpoint of a monotonic function  $\Theta : 2^{\Gamma} \times 2^{\Gamma} \rightarrow 2^{\Gamma} \times 2^{\Gamma}$ , where  $\Theta(M, N) = (M', N')$  is defined as follows:

- $X \in M'$  iff  $X \in M \setminus C(T)$  and the following conditions are satisfied:
  - If  $X \in \Gamma_{\square}$ , then there is a rule of one of the following forms:  $X \rightsquigarrow Y$  where  $Y \in M$ , or  $X \rightsquigarrow YZ$  where either  $Y \in M$ , or  $Y \in N$  and  $Z \in M$ .
  - If  $X \in \Gamma_{\circ}$ , then all rules of the form  $X \rightsquigarrow \alpha$  satisfy either  $\alpha = Y$  where  $Y \in M$ , or  $\alpha = YZ$  where either  $Y \in M$ , or  $Y \in N$  and  $Z \in M$ .
- $X \in N'$  iff  $X \in N \setminus C(T)$  and the following conditions are satisfied:
  - If  $X \in \Gamma_{\square}$ , then there is a rule of one of the following forms:  $X \rightsquigarrow \varepsilon$ , or  $X \rightsquigarrow Y$  where  $Y \in N \cup M$ , or  $X \rightsquigarrow YZ$  where either  $Y \in M$ , or  $Y, Z \in N \cup M$ .
  - If  $X \in \Gamma_{\circ}$ , then all rules of the form  $X \rightsquigarrow \alpha$  satisfy either  $\alpha = \varepsilon$ , or  $\alpha = Y$  where  $Y \in N \cup M$ , or  $\alpha = YZ$  where either  $Y \in M$ , or  $Y, Z \in N \cup M$ .

It is easy to show that:

- (1)  $(\mathcal{A}, \mathcal{B})$  is a fixpoint of  $\Theta$ .
- (2) If  $(C, D)$  is a fixpoint of  $\Theta$ , then  $C \subseteq \mathcal{A}$  and  $D \subseteq \mathcal{B}$ .

Hence, the sets  $\mathcal{A}$  and  $\mathcal{B}$  can be computed in polynomial time by a simple iterative algorithm.  $\square$

**Theorem 9.** Let  $[SU^{=1}T]$  be the set of all  $\alpha \in \Gamma^*$  for which there is a strategy  $\sigma$  such that  $\mathcal{P}(G_{\Delta}(\sigma), \alpha, SUT) = 1$ . Then there are  $\mathcal{A}, \mathcal{B}, \mathcal{C} \subseteq \Gamma$  computable in polynomial time such that  $[SU^{=1}T] = (\mathcal{B} \cup \mathcal{C})^* \mathcal{A} \Gamma^*$ .

*Proof.* We define the sets  $\mathcal{A}, \mathcal{B}$  and  $\mathcal{C}$  as follows:

- $X \in \mathcal{A}$  iff there is a strategy  $\sigma$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma)), X, \mathcal{F}T) = 1$
- $X \in \mathcal{B}$  iff there is a strategy  $\sigma$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma)), X, \mathcal{F}(T \cup \{\varepsilon\})) = 1$  and  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma)), X, \mathcal{F}T) > 0$

–  $X \in \mathcal{C}$  iff there is a strategy  $\sigma$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma)), X, \mathcal{F}\{\varepsilon\}) = 1$

Observe that  $\mathcal{C}$  can be computed in polynomial time using the algorithm of [7]. Moreover, it is easy to show that  $[SU^{=1}T] = (\mathcal{B} \cup \mathcal{C})^* \mathcal{A} \Gamma^*$  and  $[SU^{=1}(T \cup \{\varepsilon\})] = (\mathcal{B} \cup \mathcal{C})^* \cup (\mathcal{B} \cup \mathcal{C})^* \mathcal{A} \Gamma^*$ .

We prove that the sets  $\mathcal{A}$  and  $\mathcal{B}$  can be computed in polynomial time. The proof proceeds as follows: first we define a monotone function  $\Theta : 2^\Gamma \times 2^\Gamma \rightarrow 2^\Gamma \times 2^\Gamma$  and show that  $(\mathcal{A}, \mathcal{B})$  is the greatest fixpoint of  $\Theta$ . Second, we show how to compute  $\Theta$  (and hence also its greatest fixpoint) in polynomial time.

In order to define the function  $\Theta$  we need to introduce some notation. Let  $R, H \subseteq \Gamma$ . For every MD strategy  $\pi$ , we define two predicates  $Q_{R,H}$  and  $Q'_{R,H}$  over  $FPath(G_{\Delta'}(\pi), X)$  as follows: Given a path  $u \in FPath(G_{\Delta'}(\pi), X)$  of length  $n \geq 0$ , the predicate  $Q_{R,H}(u)$  (or  $Q'_{R,H}(u)$ ) holds iff  $u[n] \in T$  and for all  $0 \leq i < n$  such that  $u[i] \in \Gamma \circ \Gamma^*$  we have that *all* successors of  $u(i)$  are of the form  $(H \cup \mathcal{C})^* R \Gamma^*$  (or  $(H \cup \mathcal{C})^* R \Gamma^* \cup (H \cup \mathcal{C})^*$ , resp.). Now, we put  $\Theta(R, H) = (R', H')$  where

$$\begin{aligned} R' &= \{X \in R \cup C(T) \mid \exists \pi \exists u \in FPath(G_{\Delta'}(\pi), X), Q_{R \cup C(T), H \cup C(T)}(u) = \text{true}\} \\ H' &= \{X \in H \cup C(T) \mid \exists \pi \exists u \in FPath(G_{\Delta'}(\pi), X), Q'_{R \cup C(T), H \cup C(T)}(u) = \text{true}\} \end{aligned}$$

It follows directly from the definition that  $\Theta$  is monotone. It remains to show that  $(\mathcal{A}, \mathcal{B})$  is the greatest fixpoint of  $\Theta$ . First, we prove that  $(\mathcal{A}, \mathcal{B})$  is a fixpoint.

Let  $\Theta(\mathcal{A}, \mathcal{B}) = (\mathcal{A}', \mathcal{B}')$ . Since  $\mathcal{A}' \subseteq \mathcal{A}$  and  $\mathcal{B}' \subseteq \mathcal{B}$  by definition of  $\Theta$ , it suffices to show the opposite inclusions. Let  $X \in \mathcal{A}$  and let  $\sigma$  be a strategy which witnesses that  $X \in \mathcal{A}$ . Let us consider a path of minimal length in  $G_{\Delta'}(\sigma)$  from  $X$  to a configuration of  $T$ . Since every configuration reachable from  $X$  along a path which does not visit  $T$  belongs to  $[SU^{=1}T] = (\mathcal{B} \cup \mathcal{C})^* \mathcal{A} \Gamma^*$ , we can conclude  $X \in \mathcal{A}'$ . Similarly, we can show that  $\mathcal{B} \subseteq \mathcal{B}'$  which implies that  $\Theta(\mathcal{A}, \mathcal{B}) = (\mathcal{A}, \mathcal{B})$ .

Now, suppose that  $(R, H)$  is a fixpoint of  $\Theta$ . We prove that  $R \subseteq \mathcal{A}$  and  $H \subseteq \mathcal{B}$ . For every  $Y \in R$  (or  $Y \in H$ ), let us fix a path  $u_Y$  which witnesses that  $Y \in R$  (or  $Y \in H$ , resp.). It follows from the definition of  $\Theta$  that if  $Y \in R$  (or  $Y \in H$ ) then all successors of all stochastic configurations that appear on  $u_Y$  are of the form  $(H \cup \mathcal{C})^* R \Gamma^*$  (or  $(H \cup \mathcal{C})^* R \Gamma^* \cup (H \cup \mathcal{C})^*$ , resp.). Note that for every configuration of the form  $(H \cup \mathcal{C})^* R \Gamma^*$  there is a strategy which forces almost all runs to reach a configuration of the form  $R \Gamma^* \cup (H(H \cup \mathcal{C})^* R \Gamma^*)$ . Let us consider a strategy  $\pi$  which from configurations of the form  $\{X\} \Gamma^*$  where  $X \in R$  (or configurations of the form  $\{X\} (H \cup \mathcal{C})^* R \Gamma^*$  where  $X \in H$ ) follows the path  $u_X$  and for all successors of stochastic configurations on  $u_X$  strives to reach configurations of the form  $R \Gamma^* \cup (H(H \cup \mathcal{C})^* R \Gamma^*)$  with probability 1. Now, observe that almost every run of  $\text{Run}(G_{\Delta'}(\pi), X)$ , where  $X \in R$ , enters configurations with a head  $Y \in R \cup H$  infinitely often, which implies that almost every run takes a path  $u_Y$  for some  $Y \in R \cup H$ . It follows that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\pi), X, \mathcal{F}T)) = 1$ . Similarly, we prove that for  $X \in H$  there is a strategy  $\pi$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\pi), X, \mathcal{F}(T \cup \{\varepsilon\}))) = 1$  and  $\mathcal{P}(\text{Run}(G_{\Delta'}(\pi), X, \mathcal{F}T)) > 0$ .

So, we proved that  $(\mathcal{A}, \mathcal{B})$  is the greatest fixpoint of  $\Theta$ . Now we indicate how to compute  $\Theta(R, H)$  in polynomial time.

We can consider  $\Delta'$  as a non-probabilistic BPA (just ignoring the probabilities on transitions from the stochastic configurations). Observe that there is a path  $u$  with the properties stated in the definition of  $R'$  (or  $H'$ ) iff in the non-probabilistic

BPA  $\Delta'$  the set  $T$  is reachable via configurations with all successors in  $(H \cup C)^* R \Gamma^*$  (or  $(H \cup C)^* R \Gamma^* \cup (H \cup C)^*$ , resp.). This variant of reachability problem is decidable in polynomial time for BPA processes using standard techniques.  $\square$

**Theorem 10.** *Let  $[SU^{<1}T]$  be the set of all  $\alpha \in \Gamma^*$  for which there is a strategy  $\sigma$  such that  $\mathcal{P}(G_{\Delta'}(\sigma), \alpha, SU T) < 1$ . Then there are  $\mathcal{A}, \mathcal{B} \subseteq \Gamma$  computable in polynomial time such that  $[SU^{<1}T] = \mathcal{A}^* \cup (\mathcal{A}^* \mathcal{B} \Gamma^*)$ .*

*Proof.* Let us define the sets  $\mathcal{A}$  and  $\mathcal{B}$  as follows:

- $X \in \mathcal{A}$  iff there is a strategy  $\sigma$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma), X, \mathcal{F}\{\varepsilon\})) > 0$
- $X \in \mathcal{B}$  iff there is a strategy  $\sigma$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma), X, \neg \mathcal{F}(T \cup \{\varepsilon\}))) > 0$

It is easy to prove that  $\mathcal{A}$  and  $\mathcal{B}$  satisfy the desired property  $[SU^{<1}T] = \mathcal{A}^* \cup (\mathcal{A}^* \mathcal{B} \Gamma^*)$ . It follows from Theorem 7 that the membership to  $\mathcal{A}$  is decidable in polynomial time. For the rest of this proof we fix some  $X \in \Gamma$  and examine the conditions under which  $X \in \mathcal{B}$ .

One sufficient condition for  $X \in \mathcal{B}$  is the existence of some  $Y \in \Gamma \setminus C(T)$  and two strategies  $\sigma, \pi$  where  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma), X, \mathcal{F}(\{Y\}\Gamma^*))) > 0$  and  $\mathcal{P}(\text{Run}(G_{\Delta'}(\pi), Y, \mathcal{F}(T \cup \{\varepsilon\}))) = 0$ . The strategies  $\sigma$  and  $\pi$  can be combined into a single strategy  $\sigma'$  which behaves like  $\sigma$  until a configuration with the head  $Y$  is reached, and then it behaves like  $\pi$ . Obviously,  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma'), X, \neg \mathcal{F}(T \cup \{\varepsilon\}))) > 0$ . The existence of such  $Y, \sigma$  and  $\pi$  can be decided in polynomial time using Theorem 7 and Theorem 8. However, this condition is not necessary as the following example illustrates. Let us consider a  $1\frac{1}{2}$ -player BPA game with rules

$$A \xrightarrow{1/2} C, A \xrightarrow{1/2} B, B \xrightarrow{3/4} BB, B \xrightarrow{1/4} \varepsilon, C \xrightarrow{1} C.$$

Note that this example closely resembles one-dimensional (asymmetric) random walk. It can be shown that using the only (empty) strategy the probability of reaching  $\varepsilon$  or (a configuration with a head)  $C$  from  $A$  is less than 1 but every  $Y$  reachable from  $A$  reaches  $\varepsilon$  or  $C$  with a non-zero probability.

Hence, let us assume that there are no suitable  $Y, \sigma$  and  $\pi$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma), X, \mathcal{F}(\{Y\}\Gamma^*))) > 0$  and  $\mathcal{P}(\text{Run}(G_{\Delta'}(\pi), Y, \mathcal{F}(T \cup \{\varepsilon\}))) = 0$ . For now, let us assume that  $X \in \mathcal{B}$  and let us fix a strategy  $\sigma$  which witnesses that  $X \in \mathcal{B}$ .

*Claim (1).* There are  $Y \in \Gamma \setminus C(T)$  where  $\mathcal{P}(G_{\Delta'}(\sigma), X, \mathcal{F}(\{Y\}\Gamma^*)) > 0$ , sets  $A \subseteq \Gamma \setminus C(T)$  and  $R \subseteq \rightsquigarrow$ , a strategy  $\sigma'$ , and a set of runs  $V \subseteq \text{Run}(G_{\Delta'}(\sigma'), Y)$  such that

1.  $\mathcal{P}(V) > 0$ ;
2. for all  $w \in V$  and  $i \geq 0$  we have that  $\text{head}(w[i]) \in A$  (in particular,  $Y \in A$ ) and  $w(i) \rightarrow w(i+1)$  is induced by a transition rule of  $R$ ;
3. for every  $Z \in A$  there are infinitely many  $i \geq 0$  such that  $\text{head}(w[i]) = Z$  and for every rule  $Z \rightsquigarrow \alpha \in R$  there are infinitely many  $i \geq 0$  such that  $w(i) \rightarrow w(i+1)$  is induced by  $Z \rightsquigarrow \alpha$ .

*Proof (of Claim (1)).* For all  $A \subseteq \Gamma$  and  $R \subseteq \rightsquigarrow$ , we denote  $L_{A,R}$  the set of all  $w \in \text{Run}(G_{\Delta'}(\sigma), X, \neg \mathcal{F}(T \cup \{\varepsilon\}))$  such that exactly the heads of  $A$  and exactly the transitions induced by the rules of  $R$  occur infinitely often along  $w$ . Since there

are only finitely many sets  $L_{A,R}$ ,  $Run(G_{\Delta'}(\sigma), X, \neg\mathcal{F}(T \cup \{\varepsilon\})) = \bigcup L_{A,R}$ , and  $\mathcal{P}(Run(G_{\Delta'}(\sigma), X, \neg\mathcal{F}(T \cup \{\varepsilon\}))) > 0$ , we have that  $\mathcal{P}(L_{A,R}) > 0$  for some  $A$  and  $R$ . Let us fix such sets  $A$  and  $R$ .

For every  $v \in FPath(G_{\Delta'}(\sigma), X)$ , we denote  $U_v$  the set of all runs  $w$  of  $L_{A,R}$  such that  $v$  is a prefix of  $w$ ,  $w[|v|-1]$  is a *minimum* of  $w$  (i.e. for all  $j \geq |v|$  we have that the length of  $w[j]$  is not strictly less than the length of  $w[|v|-1]$ ), and for all  $i \geq |v|-1$  we have that  $head(w[i]) \in A$  and  $w(i) \rightarrow w(i+1)$  is induced by a rule of  $R$ . Clearly,  $L_{A,R} = \bigcup_{v \in FPath(G_{\Delta'}(\sigma), X)} U_v$ . It follows that  $0 < \mathcal{P}(L_{A,R}) \leq \sum_{v \in FPath(G_{\Delta'}(\sigma), X)} \mathcal{P}(U_v)$  and thus there is  $v \in FPath(G_{\Delta'}(\sigma), X)$  such that  $\mathcal{P}(U_v) > 0$ .

Let us assume that  $v[|v|-1] = Y\alpha$  and let  $w \in U_v$ . Since  $w[|v|-1]$  is a minimum of  $w$ , for all  $i \geq 0$  we have that  $w(i+|v|-1) = \beta_i\alpha$  where  $\beta_i \neq \varepsilon$ . We define  $\Theta(w) = \beta_0, \beta_0\beta_1, \beta_0\beta_1\beta_2, \dots$ . Now, let us consider a strategy  $\sigma'$  which, along a run  $\Theta(w)$  of  $\Theta(U_v)$ , behaves similarly as  $\sigma$  over  $w$  after the prefix  $v$  (ignoring the context  $\alpha$ ). It can be proved using standard arguments for Markov chains that  $\mathcal{P}(U_v) = \mathcal{P}(Run(G_{\Delta'}(\sigma), v)) \cdot \mathcal{P}(\Theta(U_v))$ , which implies that

$$\mathcal{P}(\Theta(U_v)) = \frac{\mathcal{P}(U_v)}{\mathcal{P}(Run(G_{\Delta'}(\sigma), v))} > 0$$

Now it is easy to verify that we can safely put  $V = \Theta(U_v)$ .  $\diamond$

For all  $A' \subseteq \Gamma$  and  $R' \subseteq \rightsquigarrow$ , we put  $\Delta'_{A',R'} = (A', R', (A' \cap \Gamma_{\square}, A' \cap \Gamma_{\circ}), Prob)$ .

*Claim (2).*  $\Delta'_{A,R}$  is a  $1\frac{1}{2}$ -player BPA game, and there is a strategy  $\pi$  for  $\Delta'_{A,R}$  such that  $\mathcal{P}(Run(G_{\Delta'_{A,R}}(\pi), Y, \mathcal{F}\{\varepsilon\})) < 1$ .

*Proof (of Claim (2)).* First we take a closer look at the rules in  $R$ . Let  $Z \in \Gamma_{\circ} \cap A$  and let  $Z \rightsquigarrow D$ . We claim that  $(Z, D) \in R$  and  $D \in A$ . Indeed, all runs of  $V$  enter a configuration having  $Z$  as its head infinitely often which implies that almost all runs of  $V$  take the rule  $Z \rightsquigarrow D$  infinitely often. Thus, if  $(Z, D) \notin R$  or  $D \notin A$ , then  $\mathcal{P}(V) = 0$ . Similarly, we prove that if  $Z \rightsquigarrow \varepsilon$  then  $(Z, \varepsilon) \in R$ .

Let us consider a rule  $Z \rightsquigarrow DE$ . By the same argument as above we conclude that  $(Z, DE) \in R$  and  $D \in A$ . We show that  $E \in A$ . Suppose the converse, i.e.,  $E \notin A$ . We claim that then for every  $\xi > 0$  there is a strategy  $\zeta$  such that  $\mathcal{P}(Run(G_{\Delta'}(\zeta), D, \mathcal{F}(T \cup \{\varepsilon\}))) < \xi$ , because otherwise  $\mathcal{P}(V) = 0$ . Indeed, if for every  $\zeta$  we have  $\mathcal{P}(Run(G_{\Delta'}(\zeta), D, \mathcal{F}(T \cup \{\varepsilon\}))) \geq \xi$  for a fixed  $\xi > 0$ , then almost all runs of  $V$  enter  $E$  infinitely often, because all runs of  $V$  enter a configuration with the head  $Z$  infinitely often and no run of  $V$  enters  $T$ . Now, we employ Proposition 3 (a) and conclude that there is a strategy  $\zeta'$  such that  $\mathcal{P}(Run(G_{\Delta'}(\zeta'), D, \mathcal{F}(T \cup \{\varepsilon\}))) = 0$ . However,  $D$  is reachable from  $X$  (using a suitable strategy) which contradicts our assumption. Hence,  $E \in A$ .

Now let us assume that  $Z \in \Gamma_{\square} \cap A$ . Observe that  $(Z, \alpha) \in R$  for at least one rule  $Z \rightsquigarrow \alpha$ . Let us assume that  $(Z, D) \in R$ . We have that  $Z \rightsquigarrow D$  is used infinitely often along every run of  $V$  and thus  $D \in A$ . Similarly, if  $(Z, DE) \in R$  then  $D \in A$  and using similar arguments as above (for  $Z \in \Gamma_{\circ} \cap A$ ) we can show that  $E \in A$ .

The above arguments imply that  $\Delta'_{A,R}$  is a  $1\frac{1}{2}$ -player BPA game. Now, let us consider an arbitrary strategy  $\pi$  for  $\Delta'_{A,R}$  which behaves similarly as  $\sigma'$  over runs of  $V$ . The existence of such a strategy is guaranteed by Claim (1). Using

standard arguments for Markov chains, it can be proven that the probability of  $V$  remains the same (i.e., non-zero) in  $G_{\Delta'_{A,R}}(\pi)$  as in  $G_{\Delta'}(\sigma')$ , which implies that  $\mathcal{P}(G_{\Delta'_{A,R}}, Y, \mathcal{F}\{\varepsilon\}) < 1$ .  $\diamond$

So far, we have proved that if  $X \in \mathcal{B}$ , then

- there are  $Y \in \Gamma \setminus C(T)$ ,  $A \subseteq \Gamma \setminus C(T)$  and  $R \subseteq \rightsquigarrow$  such that  $Y \in A$  and  $\Delta'_{A,R}$  is a  $1\frac{1}{2}$ -player BPA game;
- there is a strategy  $\sigma$  satisfying  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma), Y, \mathcal{F}(\{Y\}\Gamma^*))) > 0$ ;
- there is a strategy  $\pi$  satisfying  $\mathcal{P}(\text{Run}(G_{\Delta'_{A,R}}(\pi), Y, \mathcal{F}\{\varepsilon\})) < 1$ .

Using the strategies  $\sigma$  and  $\pi$ , we can easily define a strategy  $\zeta$  for  $\Delta'$  which witnesses that  $X \in \mathcal{B}$  ( $\zeta$  behaves like  $\sigma$  until  $Y$  is reached and then behaves like  $\pi$ ).

Now, it is easy to see that if  $\Delta'_{A',R'}$  is a  $1\frac{1}{2}$ -player BPA game for some  $A' \supseteq A$  and  $R' \supseteq R$ , then the strategy  $\pi$  can be extended to a strategy  $\pi'$  in  $\Delta'_{A',R'}$  which satisfies  $\mathcal{P}(\text{Run}(G_{\Delta'_{A',R'}}(\pi'), Y, \mathcal{F}\{\varepsilon\})) = \mathcal{P}(\text{Run}(G_{\Delta'_{A,R}}(\pi), Y, \mathcal{F}\{\varepsilon\})) < 1$ . Furthermore, if for  $A', R'$  and  $A'', R''$  we have that  $\Delta'_{A',R'}$  and  $\Delta'_{A'',R''}$  are  $1\frac{1}{2}$ -player BPA games, then  $\Delta'_{A' \cup A'', R' \cup R''}$  is a  $1\frac{1}{2}$ -player BPA game.

Hence, in order to decide whether  $X \in \mathcal{B}$ , it suffices to compute the largest sets  $A \subseteq \Gamma \setminus C(T)$  and  $R \subseteq \rightsquigarrow$  for which  $\Delta'_{A,R}$  is a  $1\frac{1}{2}$ -player BPA game, and to decide whether there are  $Y \in A$ ,  $\sigma$ , and  $\pi$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'}(\sigma), X, \mathcal{F}(\{Y\}\Gamma^*))) > 0$  and  $\mathcal{P}(\text{Run}(G_{\Delta'_{A,R}}(\pi), Y, \mathcal{F}\{\varepsilon\})) < 1$ . The problem whether there is  $\pi$  such that  $\mathcal{P}(\text{Run}(G_{\Delta'_{A,R}}(\pi), Y, \mathcal{F}\{\varepsilon\})) < 1$  can be decided in polynomial time using the algorithm of [7]. The maximal sets  $A$  and  $R$  can be computed using a simple fixpoint algorithm.  $\square$

## 4 Model-checking Qualitative PCTL for $1\frac{1}{2}$ -player BPA Games

In this section we show that the results about  $1\frac{1}{2}$ -player BPA games with extended reachability objectives (see Section 3) can be used to design an essentially optimal model-checking algorithm for the qualitative fragment of PCTL and  $1\frac{1}{2}$ -player BPA games. For technical convenience, we restrict ourselves to *simple* valuations, where  $\nu(p)$  is a simple set for each  $p \in Ap$  (see Definition 5).

Infinite sets of stack configurations will be represented by deterministic finite-state automata (DFA) which read the stack bottom-up. Formally, a DFA is a tuple  $\mathcal{F} = (Q, \Sigma, \delta, \hat{q}, F)$  where  $Q$  is a finite set of *control states*,  $\Sigma$  is a finite *input alphabet*,  $\delta : (Q \times \Sigma) \rightarrow Q$  is a total *transition function*,  $\hat{q} \in Q$  is the *initial state*, and  $F \subseteq Q$  is a subset of *final states*. The function  $\delta$  is extended to the elements of  $Q \times \Sigma^*$  in the natural way. A word  $w \in \Sigma^*$  is *accepted* by  $\mathcal{F}$  iff  $\delta(q_0, w) \in F$ .

Let  $\Delta$  be a  $1\frac{1}{2}$ -player BPA game with the stack alphabet  $\Gamma$ , and let  $\mathcal{F}$  be a DFA with the input alphabet  $\Gamma$ . We say that a stack configuration  $\alpha \in \Gamma^*$  is *recognized* by  $\mathcal{F}$  iff the *reverse* of  $\alpha$  is accepted by  $\mathcal{F}$ . Note that stack configurations are traditionally written as words starting with the top-of-the-stack symbol, but for technical reasons we prefer to read them in the bottom-up (i.e., right to left) direction.

In the proof of our next theorem we use the standard technique of simulating DFA in the stack alphabet (see, e.g., [5]).

**Theorem 11.** Let  $\Delta = (\Gamma, \hookrightarrow, (\Gamma_{\square}, \Gamma_{\circ}), \text{Prob})$  by a  $1\frac{1}{2}$ -player BPA game. Let  $\nu$  be a simple valuation and  $\Phi$  a qualitative PCTL formula. Then there is a DFA  $\mathcal{F}_{\Phi}$  of size  $|\Delta| \cdot 2^{\mathcal{O}(|\Phi|)}$  constructible in time which is polynomial in  $|\Delta|$  and exponential in  $|\Phi|$  such that for all  $\alpha \in \Gamma^*$  we have that  $\alpha \models^{\nu} \Phi$  iff  $\alpha$  is recognized by  $\mathcal{F}_{\Phi}$ .

*Proof.* We proceed by induction on the structure of  $\Phi$ . The cases when  $\Phi \equiv p$ ,  $\Phi \equiv \Phi_1 \wedge \Phi_2$ , and  $\Phi \equiv \neg\Phi_1$  follow immediately.

Let  $\Phi \equiv \mathcal{X}^=1\Phi_1$ , and let  $\mathcal{F}_1 = (Q_1, \Gamma, \delta_1, \hat{q}, F_1)$  be the DFA associated with  $\Phi_1$ . The automaton  $\mathcal{F}$  associated with  $\Phi$  should then recognize exactly all  $\alpha \in \Gamma^*$  such that for every transition  $\alpha \rightarrow \beta$  we have that  $\beta$  is recognized by  $\mathcal{F}_1$ . Hence, we put  $\mathcal{F} = (Q_1 \cup Q'_1, \Gamma, \delta, \hat{r}, Q'_1)$ , where  $Q'_1 = \{q' \mid q \in Q_1\}$  and the transition function  $\delta$  is constructed as follows: Let  $q \in Q_1$ ,  $A \in \Gamma$ , and let  $t = \delta_1(q, A)$ . If for all rules  $A \hookrightarrow \gamma$  we have that  $\delta_1(q, \gamma^r) \in F_1$  (where  $\gamma^r$  denotes the reverse of  $\gamma$ ), then  $\delta(q, A) = \delta(q', A) = t'$ . Otherwise,  $\delta(q, A) = \delta(q', A) = t$ . The initial state  $\hat{r}$  of  $\mathcal{F}$  is either  $\hat{q}'$  or  $\hat{q}$ , depending on whether  $\varepsilon$  is recognized by  $\mathcal{F}_1$  or not, respectively.

The cases when  $\Phi \equiv \mathcal{X}^{<1}\Phi_1$ ,  $\Phi \equiv \mathcal{X}^{=0}\Phi_1$ , and  $\Phi \equiv \mathcal{X}^{>0}\Phi_1$  are handled similarly.

Now, let us consider the case when  $\Phi \equiv \Phi_1 \mathcal{U}^=1\Phi_2$ . Note that  $\alpha \models^{\nu} \Phi_1 \mathcal{U}^=1\Phi_2$  iff there is no MD strategy  $\sigma$  such that  $\mathcal{P}(\text{Run}(G_{\Delta}(\sigma), \alpha, \llbracket \Phi_1 \rrbracket \mathcal{U} \llbracket \Phi_2 \rrbracket)) < 1$ . Let  $\mathcal{F}_1 = (Q_1, \Gamma, \delta_1, \hat{q}, F_1)$  and  $\mathcal{F}_2 = (Q_2, \Gamma, \delta_2, \hat{r}, F_2)$  be the DFA associated with  $\Phi_1$  and  $\Phi_2$ . We construct another DFA  $\mathcal{F}$  which accepts exactly those  $\alpha \in \Gamma^*$  for which there *exists* an MD strategy  $\sigma$  such that  $\mathcal{P}(\text{Run}(G_{\Delta}(\sigma), \alpha, \llbracket \Phi_1 \rrbracket \mathcal{U} \llbracket \Phi_2 \rrbracket)) < 1$ . The desired DFA is then obtained simply by complementing the automaton  $\mathcal{F}$ . First we construct a  $1\frac{1}{2}$ -player BPA game  $\bar{\Delta}$  which is obtained from  $\Delta$  by encoding the automata  $\mathcal{F}_1, \mathcal{F}_2$  into the stack alphabet and simulating them “on-the-fly”. Formally,  $\bar{\Delta} = (\bar{\Gamma}, \bar{\hookrightarrow}, (\bar{\Gamma}_{\square}, \bar{\Gamma}_{\circ}), \text{Pr})$  where  $\bar{\Gamma} = (\Gamma \cup \{\varepsilon\}) \times Q_1 \times Q_2$ ,  $\bar{\Gamma}_{\square} = \Gamma_{\square} \times Q_1 \times Q_2$ ,  $\bar{\Gamma}_{\circ} = (\Gamma_{\circ} \cup \{\varepsilon\}) \times Q_1 \times Q_2$  and the transition relation  $\bar{\hookrightarrow}$  together with  $\text{Pr}$  are defined as follows ( $A, q$ , and  $r$  range over  $\Gamma, Q_1$ , and  $Q_2$ , respectively):

- $(A, q, r) \bar{\hookrightarrow} \varepsilon$  iff  $A \bar{\hookrightarrow} \varepsilon$ ;
- $(A, q, r) \bar{\hookrightarrow} (B, q, r)$  iff  $A \bar{\hookrightarrow} B$
- $(A, q, r) \bar{\hookrightarrow} (B, q', r')(C, q, r)$  iff  $A \bar{\hookrightarrow} BC$ ,  $\delta_1(q, C) = q'$  and  $\delta_2(r, C) = r'$ ;
- $(\varepsilon, q, r) \bar{\hookrightarrow} (\varepsilon, q, r)$ .

For every configuration  $\alpha \in \Gamma^*$  of the form  $A_n \cdots A_1$  there is a unique configuration  $[\alpha] \in \bar{\Gamma}^*$  of the form  $(A_n, q_n, r_n) \cdots (A_1, q_1, r_1)(\varepsilon, \hat{q}, \hat{r})$  where  $q_1 = \hat{q}$ ,  $r_1 = \hat{r}$ , and for all  $0 \leq i < n$  we have that  $\delta_1(q_i, A_i) = q_{i+1}$  and  $\delta_2(r_i, A_i) = r_{i+1}$ . Note that for every  $\alpha \in \Gamma^*$ , the subgraphs of  $G_{\Delta}$  and  $G_{\bar{\Delta}}$  which consist of all vertices reachable from  $\alpha$  and  $[\alpha]$  are isomorphic. Let  $S, T \subseteq \bar{\Gamma}^*$  be the simple sets where

- $C(S) = \{(x, q, r) \mid x \in \Gamma \cup \{\varepsilon\}, \delta_1(q, x) \in F_1, r \in Q_2\}$
- $C(T) = \{(x, q, r) \mid x \in \Gamma \cup \{\varepsilon\}, q \in Q_1, \delta_2(r, x) \in F_2\}$ .

Now it is easy to see that  $\{\alpha \in \Gamma^* \mid \exists \sigma : \mathcal{P}(\text{Run}(G_{\Delta}(\sigma), \alpha, \llbracket \Phi_1 \rrbracket \mathcal{U} \llbracket \Phi_2 \rrbracket)) < 1\}$  is equal to the set  $K = \{\alpha \in \Gamma^* \mid \exists \sigma : \mathcal{P}(\text{Run}(G_{\bar{\Delta}}(\sigma), [\alpha], \text{SUT})) < 1\}$ . By Theorem 10, there effectively exist the sets  $\mathcal{A}, \mathcal{B} \subseteq \bar{\Gamma}$  such that  $K = \{\alpha \in \Gamma^* \mid [\alpha] \in \mathcal{A}^* \cup (\mathcal{A}^* \mathcal{B} \bar{\Gamma}^*)\}$ . Hence, the automaton  $\mathcal{F}$  recognizing the set  $K$  can now be constructed as follows: we put  $\mathcal{F} = (Q, \Gamma, \delta, \hat{t}, F)$  where

- $Q = Q_1 \times Q_2 \times \{0, 1\}$ .
- For all  $A \in \Gamma$ ,  $q \in Q_1$ ,  $r \in Q_2$ , and  $i \in \{0, 1\}$  we put  $\delta((q, r, i), A) = (\delta_1(q, A), \delta_2(r, A), j)$ , where
  - $j = 0$  iff either  $i = 0$  and  $(q, r, A) \in \mathcal{A} \cup \mathcal{B}$ , or  $i = 1$  and  $(q, r, A) \in \mathcal{B}$ ;
  - $j = 1$  iff either  $i = 0$  and  $(q, r, A) \in \Gamma \setminus (\mathcal{A} \cup \mathcal{B})$ , or  $i = 1$  and  $(q, r, A) \in \Gamma \setminus \mathcal{B}$ .
- The initial state  $\hat{t}$  is either  $(\hat{q}, \hat{r}, 0)$  or  $(\hat{q}, \hat{r}, 1)$ , depending on whether  $(\varepsilon, \hat{q}, \hat{r}) \in \mathcal{A} \cup \mathcal{B}$  or not, respectively.
- $F = Q_1 \times Q_2 \times \{0\}$ .

The cases when  $\Phi \equiv \Phi_1 \mathcal{U}^{=0} \Phi_2$ ,  $\Phi \equiv \Phi_1 \mathcal{U}^{>0} \Phi_2$ , and  $\Phi \equiv \Phi_1 \mathcal{U}^{<1} \Phi_2$  are handled similarly, using Theorem 7, 8, and 9, respectively.

The complexity of the whole algorithm is easy to evaluate (it suffices to consider the worst subcase  $\Phi \equiv \Phi_1 \mathcal{U}^{\neq 0} \Phi_2$ ).  $\square$

Since the model-checking problem for qualitative PCTL and fully probabilistic BPA (i.e., the subclass of  $1\frac{1}{2}$ -player BPA games where  $\Gamma_{\square} = \emptyset$ ) is known to be **EXPTIME**-hard [3], we obtain the following:

**Corollary 12.** *The model-checking problem for qualitative PCTL and  $1\frac{1}{2}$ -player BPA games is **EXPTIME**-complete. For each fixed formula, the problem becomes solvable in polynomial time.*

*Acknowledgement.* We thank an anonymous reviewer for fixing a mistake in the proof of Theorem 9.

## References

1. C. Baier and M. Kwiatkowska. Model checking for a probabilistic branching time logic with fairness. *Distributed Computing*, 11(3):125–155, 1998.
2. A. Bianco and L. de Alfaro. Model checking of probabilistic and nondeterministic systems. In *Proceedings of FST&TCS'95*, vol. 1026 of *LNCS*, pp. 499–513. Springer, 1995.
3. T. Brázdil, A. Kučera, and O. Stražovský. On the decidability of temporal properties of probabilistic pushdown automata. In *Proceedings of STACS'2005*, vol. 3404 of *LNCS*, pp. 145–157. Springer, 2005.
4. J. Esparza, A. Kučera, and R. Mayr. Model-checking probabilistic pushdown automata. In *Proceedings of LICS 2004*, pp. 12–21. IEEE, 2004.
5. J. Esparza, A. Kučera, and S. Schwoon. Model-checking LTL with regular valuations for pushdown systems. *IEC*, 186(2):355–376, 2003.
6. K. Etessami and M. Yannakakis. Recursive Markov decision processes and recursive stochastic games. In *Proceedings of ICALP 2005*, vol. 3580 of *LNCS*. Springer, 2005.
7. K. Etessami and M. Yannakakis. Efficient qualitative analysis of classes of recursive Markov decision processes and simple stochastic games. In *Proceedings of STACS'2006*, LNCS. Springer, 2006.
8. E. Feinberg and A. Shwartz, editors. *Handbook of Markov Decision Processes*. Kluwer, 2002.
9. H. Hansson and B. Jonsson. A logic for reasoning about time and reliability. *Formal Aspects of Computing*, 6:512–535, 1994.
10. A. Hinton, M. Kwiatkowska, G. Norman, and D. Parker. PRISM: a tool for automatic verification of probabilistic systems. In *Proceedings of TACAS 2006*, LNCS. Springer, 2006. To appear.
11. M.L. Puterman. *Markov Decision Processes*. Wiley, 1994.