FI MU

Faculty of Informatics
Masaryk University Brno

The Satisfiability Problem for Probabilistic CTL

by

Tomáš Brázdil
Vojtěch Forejt
Jan Křetínský
Antonín Kučera

FI MU Report Series **FIMU-RS-2008-18**

June 2008

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

The Satisfiability Problem for Probabilistic CTL

Tomáš Brázdil   Vojtěch Forejt   Jan Křetínský   Antonín Kučera

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

Abstract

We study the satisfiability problem for qualitative PCTL (Probabilistic Computation Tree Logic), which is obtained from "ordinary" CTL by replacing the EX, AX, EU, and AU operators with their qualitative counterparts $X^{>0}$, $X^{=1}$, $U^{>0}$, and $U^{=1}$, respectively. As opposed to CTL, qualitative PCTL does not have a small model property, and there are even qualitative PCTL formulae which have only infinite-state models. Nevertheless, we show that the satisfiability problem for qualitative PCTL is **EXPTIME**-complete and we give an exponential-time algorithm which for a given formula φ computes a finite description of a model (if it exists), or answers "not satisfiable" (otherwise). We also consider the finite satisfiability problem and provide analogous results. That is, we show that the finite satisfiability problem for qualitative PCTL is **EXPTIME**-complete, and every finite satisfiable formula has a model of an exponential size which can effectively be constructed in exponential time. Finally, we give some results about the quantitative PCTL, where the numerical bounds in probability constraints can be arbitrary rationals between 0 and 1. We prove that the problem whether a given quantitative PCTL formula φ has a model of the branching degree at most k, where $k \geq 2$ is an arbitrary but fixed constant, is highly undecidable. We also show that every satisfiable formula φ has a model with branching degree at most $|\varphi| + 2$. However, this does not yet imply the undecidability of the satisfiability problem for quantitative PCTL, and we in fact conjecture the opposite.

The presented work has been supported by the research center Institute for Theoretical Computer Science (ITI), project No. 1M0545.

1 Introduction

Probabilistic CTL (PCTL) [13] is a probabilistic extension of the well-known branching-time logic CTL [7] obtained by replacing the existential and universal path quantifiers with the probabilistic operator, which allows to quantify the probability of all runs that satisfy a given path formula. More precisely, the syntax of PCTL is built upon atomic propositions, using Boolean connectives and operators "next" and "until" of the form $X^{\bowtie\rho}\varphi$ and $\varphi_1 U^{\bowtie\rho}\varphi_2$, respectively, where $\bowtie$ is a numerical comparison such as $\leq$ or $>$, and $\rho \in [0,1]$ is a rational constant. We also use the standard abbreviations $F\varphi$ and $G\varphi$ to denote the path formulae $\mathtt{tt}U\varphi$ and $\neg F\neg\varphi$. A simple example of a PCTL formula is $G^{=1}(a \Rightarrow F^{\geq 0.2}b)$ which says "in each reachable state that satisfies a, the probability of visiting a state satisfying b is at least 0.2". Formally, PCTL formulae are interpreted over Markov chains where each state is assigned a subset of atomic propositions that are valid in a given state.

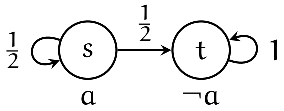
In this paper, we study the satisfiability problem for the *qualitative fragment* of PCTL, which is obtained by restricting the probabilistic operator to its qualitative forms (i.e., the constant ρ in $X^{\bowtie\rho}\varphi$ and $\varphi_1 U^{\bowtie\rho}\varphi_2$ can be just 0 or 1). Since the syntax of PCTL includes negation, we need to consider only the probability constraints >0 and $=1$ (for example, the formula $X^{<1}\varphi$ is equivalent to $\neg X^{=1}\varphi$). Hence, there are only four modal operators $X^{>0}$, $X^{=1}$, $U^{>0}$, and $U^{=1}$. At first glance, they seem to be closely related to standard CTL operators EX, AX, EU, and AU, respectively. To a large extent, this is true for $X^{>0}$, $X^{=1}$, $U^{>0}$, but the properties of $U^{=1}$ and AU are very different, which leads to the phenomena described in the next paragraphs.

First, let us recall known results about the satisfiability problem for CTL and related logics. For CTL, the problem is known to be **EXPTIME**-complete [8]. In the same paper [8], it is also shown that CTL has a *small model property*, i.e., every satisfiable CTL formula φ has a finite-state model whose size is exponential in φ. For the logic CTL*, the satisfiability problem is **2-EXPTIME**-complete [9, 20]. The **2-EXPTIME** lower bound holds even for the weaker logic CTL$^+$ [16]. The complexity of the satisfiability and validity problems for other fragments of CTL and CTL* (such as the existential and universal fragments) has also been studied (see, e.g., [19]). The satisfiability for the modal μ-calculus is **EXPTIME**-complete [2, 12], and even this powerful logic has the small model property [17]. The satisfiability and validity for some fragments of the modal μ-calculus have been studied in greater depth in [14]. To the best of our knowledge, the satisfiability problem for probabilistic CTL has not yet been examined. Nonetheless, there are some related results about PCTL model-checking (both for infinite- and finite-state systems, see e.g. [6, 15, 11, 10, 5]) and strategy synthesis for Markov decision processes with branching-time objectives [1, 18, 3, 4].

As we already noted, the qualitative PCTL formulae seem to be rather similar to "ordinary" CTL formulae. One may even be tempted to think that a qualitative PCTL formula φ is satisfiable iff the corresponding CTL formula φ' is satisfiable, where φ' is obtained from φ by replacing each occurrence of $X^{>0}$, $X^{=1}$, $U^{>0}$, and $U^{=1}$ with EX, AX, EU, and AU, respectively. This is not true; for example, the qualitative PCTL formula

$$\varphi \;\equiv\; a \wedge \big(G^{=1}(a \Rightarrow X^{>0}a)\big) \wedge \big(F^{=1}\neg a\big)$$

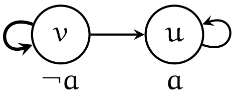
has the following model:



Note that $s \models \varphi$ because the probability of all runs initiated in s which eventually visit t is equal to 1. However, the corresponding CTL formula

$$\varphi' \;\equiv\; a \wedge \big(AG(a \Rightarrow EXa)\big) \wedge \big(AF\neg a\big)$$

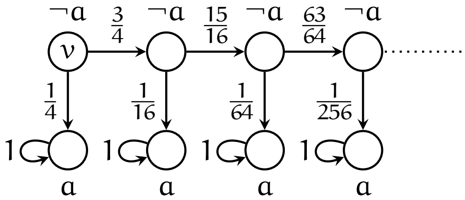
is not satisfiable. Further, qualitative PCTL does not have the small model property, and there are even satisfiable qualitative PCTL formulae that only have *infinite-state* models. A simple example of such a formula is $G^{>0}(\neg a \wedge F^{>0}a)$. Intuitively, this formula does not have a finite-state model, because for every finite-state Markov chain M there is a fixed constant $\varepsilon > 0$ such that every state of M which satisfies $F^{>0}a$ also satisfies $F^{\geq\varepsilon}a$. This means that the probability of all runs satisfying the path formula $G(\neg a \wedge F^{>0}a)$ is zero, hence $G^{>0}(\neg a \wedge F^{>0}a)$ does not hold. On the other hand, $G^{>0}(\neg a \wedge F^{>0}a)$ has an infinite-state model, which admits a simple symbolic description in a form of the following *marked graph*:



In general, a marked graph is a finite binary graph where each node has at least one out-going transition, and some transitions are "marked" (in the above figure, the only marked transition is the loop on v, which is indicated by a thick arrow). Each node v in a given marked graph $\mathcal{G}$ determines a unique infinite-state Markov chain $M_{\mathcal{G}}$ obtained by unfolding the structure of $\mathcal{G}$ into an infinite tree (with the root v), where the probabilities of outgoing transitions at each state s of $M_{\mathcal{G}}$ are determined as follows:

- if all outgoing transitions of s are either marked or non-marked, then all of them have the same probability p;
- otherwise, the probability of all marked transitions is p_1, the probability of all non-marked transitions is p_2, and the total probability of all marked transitions is $1 - 1/4^{d+1}$, where d is the distance of s from the root of $M_{\mathcal{G}}$ (note that p_1 and p_2 are uniquely determined by these conditions).

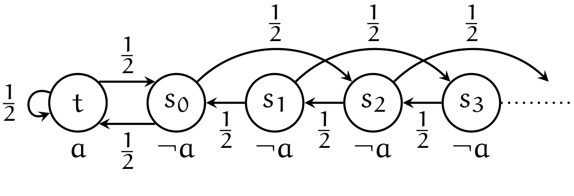
For example, the initial part of the chain $M_{\mathcal{G}}$, where $\mathcal{G}$ is the marked graph above, looks as follows (for simplicity, the loop on u is not unfolded):



Observe that $v \models G^{>0}(\neg a \wedge F^{>0} a)$, because the only run which satisfies the formula $G(\neg a \wedge F^{>0} a)$ has a positive probability.

We prove that *every* satisfiable qualitative PCTL formula φ has a model which can be represented by a marked graph whose size is exponential in $|\varphi|$, and we design an exponential-time algorithm which for a given φ computes a suitable marked graph if it exists, and outputs "unsatisfiable" otherwise. Hence, the satisfiability problem for qualitative PCTL is in **EXPTIME** and we also give the matching lower bound (the lower bound is proved by standard techniques). Since the logic PCTL contains negation and **EXPTIME** is closed under complement, the validity problem for qualitative PCTL is also **EXPTIME**-complete.

One may also ask whether the use of exponentially small probabilities in the way indicated above is indeed necessary. For example, the mentioned formula $G^{>0}(\neg a \wedge F^{>0} a)$ has another infinite-state model, where the probability of every transition is exactly $\frac{1}{2}$. The model looks as follows:



We have that $s_0 \models G^{>0}(\neg a \wedge F^{>0}a)$. To see this, realize that the probability of all runs initiated in s_0 which do not visit t is positive (this is a standard result of Markov chain theory; the exact value of this probability is $(3-\sqrt{5})/2$). However, the use of "exponentially small probabilities" is unavoidable in some cases. For example, one can easily show that the formula $G^{=1}(X^{>0}a) \wedge G^{>0}\neg a$ does not have a model where the probabilities of all transitions are uniformly bounded from below. However, the formula $G^{=1}(X^{>0}a) \wedge G^{>0}\neg a$ is satisfiable, which is witnessed by the marked graph for the formula $G^{>0}(\neg a \wedge F^{>0}a)$ constructed earlier.

Since some qualitative PCTL formulae have only infinite-state models, we also consider the *finite satisfiability* problem, where we ask whether a given qualitative PCTL formula has a *finite-state* model. We obtain similar results as for general satisfiability. We show that the existence of a finite-state model implies the existence of a model whose size is exponential in the size of a given formula, and we give an exponential-time algorithm which computes such a model if it exists, and outputs "not finite satisfiable" otherwise. Hence, the finite satisfiability/validity problems for qualitative PCTL are also in **EXPTIME**, and in fact **EXPTIME**-complete.

Finally, we give some results concerning the satisfiability problem for general PCTL. We show that the problem whether a given PCTL formula has a model where the branching degree is bounded by a fixed k is *highly undecidable* for every $k \geq 2$ (note that the k is *not* a part of the problem instance, but a fixed parameter—for a different choice of k we have a different problem, and each of these infinitely many problems is highly undecidable). Then, we show that every satisfiable PCTL formula φ has a model with branching degree at most $|\varphi| + 2$. At first glance, one may be tempted to think that these two results imply the undecidability of the satisfiability problem for the general PCTL, but in fact it is *not* the case. Despite a reasonable amount of effort, we did not manage to extend the undecidability proof to the satisfiability problem for PCTL, and the difficulties seem to be fundamental (at least, the undecidability proof techniques developed in [3, 5] specifically for probabilistic systems seem insufficient). On the other hand, there are some structural regularities in PCTL models which suggest that the problem might in fact be decidable. We present the decidability hypothesis as an open conjecture which surely deserves further attention.

2 Definitions

In this paper, we use $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Q}$, and $\mathbb{R}$ to denote the sets of positive integers, non-negative integers, rational numbers, and real numbers, respectively. We also use the standard notation for intervals of real numbers, writing, e.g., $(0, 1]$ to denote the set $\{x \in \mathbb{R} \mid 0 < x \leq 1\}$.

The set of all finite words over a given alphabet Σ is denoted Σ^*, and the set of all infinite words over Σ is denoted Σ^ω. We also use Σ^+ to denote the set $\Sigma^* \setminus \{\varepsilon\}$ where ε is the empty word. The length of a given $w \in \Sigma^* \cup \Sigma^\omega$ is denoted $len(w)$, where the length of an infinite word is ω. Given a word (finite or infinite) over Σ, the individual letters of w are denoted $w(0), w(1), \ldots$.

Let $V \neq \emptyset$, and let $\to \ \subseteq V \times V$ be a *total* relation (i.e., for every $v \in V$ there is some $u \in V$ such that $v \to u$). A *path* in V is a finite or infinite word $w \in V^+ \cup V^\omega$ such that $w(i{-}1) \to w(i)$ for every $1 \leq i < len(w)$. Sometimes we also write $s_0 \to \cdots \to s_n$ to denote the finite path $s_0, \cdots, s_n$, particularly in situations when the underlying relation $\to$ is not completely obvious from the context. We also use $\to^+$ to denote the transitive closure of $\to$, and $\to^*$ to denote the reflexive and transitive closure of $\to$. A *run* in V is an infinite path in V. The set of all runs that start with a given finite path w is denoted $Run(w)$. Let $U \subseteq V$. We say that U is *strongly connected* if $v \to^+ u$ for all $v, u \in U$ (from a graph-theoretic point of view, this definition is somewhat non-standard, because a singleton $\{s\}$ is strongly connected iff $s \to s$). Further, we say that U is a *strongly connected component (SCC)* if $U \neq \emptyset$ is a maximal strongly connected subset of V, and U is a *bottom SCC (BSCC)* if U is a SCC and for every $u \in U$ and every $u \to v$ we have that $v \in U$.

A *probability distribution* over a finite or countably infinite set X is a function $f : X \to [0, 1]$ such that $\sum_{x \in X} f(x) = 1$. A probability distribution f over X is *positive* if $f(x) > 0$ for every $x \in X$, and *uniform* if $f(x) = f(y)$ for all $x, y \in X$. A σ-*algebra* over a set Ω is a set $\mathcal{F} \subseteq 2^\Omega$ that includes Ω and is closed under complement and countable union. A *probability space* is a triple $(\Omega, \mathcal{F}, \mathcal{P})$ where Ω is a set called *sample space*, $\mathcal{F}$ is a σ-algebra over Ω whose elements are called *events* (or *measurable* sets), and $\mathcal{P} : \mathcal{F} \to [0, 1]$ is a *probability measure* 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 $\mathcal{P}(\Omega)$=1.

Definition 2.1 (Markov Chain). *A* Markov chain *is a triple* $M = (St, \to, Prob)$ *where* St *is a finite or countably infinite set of* states, $\to \ \subseteq St \times St$ *is a total* transition relation*, and*

Prob is a function which to each state $s \in St$ *assigns a positive probability distribution over the outgoing transitions of* s. *As usual, we write* $s \xrightarrow{x} t$ *when* $s \rightarrow t$ *and* x *is the probability of* $s \rightarrow t$.

When defining the semantics of PCTL (see below), we need to measure the probability of certain sets of runs. Formally, to every $s \in St$ we associate the probability space $(Run(s), \mathcal{F}, \mathcal{P})$ where $\mathcal{F}$ is the σ-algebra generated by all *basic cylinders* $Run(w)$ where w is a finite path starting with s, and $\mathcal{P} : \mathcal{F} \rightarrow [0,1]$ is the unique probability measure such that $\mathcal{P}(Run(w)) = \Pi_{i=1}^{len(w)-1} x_i$ where $w(i-1) \xrightarrow{x_i} w(i)$ for every $1 \leq i < len(w)$. If $len(w) = 1$, we put $\mathcal{P}(Run(w)) = 1$. Hence, only certain subsets of $Run(s)$ are measurable, but in this paper we only deal with "safe" subsets that are guaranteed to be in $\mathcal{F}$.

Definition 2.2. *Let* $Ap = \{a, b, c, \ldots\}$ *be a countably infinite set of* atomic propositions. *The syntax of PCTL state and path formulae is defined by the following abstract syntax equations:*

$$\begin{aligned} \varphi &::= a \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \mathcal{P}^{\bowtie\varrho}\psi \\ \psi &::= X\varphi \mid \varphi_1 U \varphi_2 \end{aligned}$$

Here a *ranges over Ap,* $\bowtie$ *is a comparison (i.e.,* $\bowtie \in \{<, >, \leq, \geq, =, \neq\}$*), and* $\varrho \in [0,1]$ *is a rational constant. The* qualitative fragment *of PCTL is obtained by restricting* ϱ *to* 0 *and* 1 *(to prevent a confusion between PCTL and qualitative PCTL, we sometimes refer to "quantitative PCTL" instead of PCTL).*

In the rest of this paper, "PCTL formula" means "PCTL state formula". Since the probabilistic operator $\mathcal{P}^{\bowtie\varrho}$ is always bound to exactly one modal connective, we simplify the syntax by writing $X^{\bowtie\varrho}\varphi$ instead of $\mathcal{P}^{\bowtie\varrho}(X\varphi)$, and $\varphi_1 U^{\bowtie\varrho} \varphi_2$ instead of $\mathcal{P}^{\bowtie\varrho}(\varphi_1 U \varphi_2)$. For every PCTL formula φ, the symbol $\hat{\varphi}$ denotes either the formula ξ, or $\neg\varphi$, depending on whether φ is of the form $\neg\xi$, or not, respectively.

Let $M = (St, \rightarrow, Prob)$ be a Markov chain, and let $\nu : St \rightarrow 2^{Ap}$ be a *valuation*. The validity of PCTL formulae in the states of M is defined inductively as follows:

$$\begin{array}{lll} M, s \models^{\nu} a & \text{iff} & a \in \nu(s) \\ M, s \models^{\nu} \neg\varphi & \text{iff} & M, s \not\models^{\nu} \varphi \\ M, s \models^{\nu} \varphi_1 \wedge \varphi_2 & \text{iff} & M, s \models^{\nu} \varphi_1 \text{ and } M, s \models^{\nu} \varphi_2 \\ M, s \models^{\nu} X^{\bowtie\varrho}\varphi & \text{iff} & \mathcal{P}(\{w \in Run(s) \mid M, w(1) \models^{\nu} \varphi\}) \bowtie \varrho \\ M, s \models^{\nu} \varphi_1 U^{\bowtie\varrho} \varphi_2 & \text{iff} & \mathcal{P}(\{w \in Run(s) \mid \exists j \geq 0 : M, w(j) \models^{\nu} \varphi_2 \\ & & \quad \text{and } \forall 0 \leq i < j : M, w(i) \models^{\nu} \varphi_1\}) \bowtie \varrho \end{array}$$

A PCTL formula φ is *satisfiable* if $M, s \models^{\nu} \varphi$ for some M, s, and ν. The formula φ is *finite satisfiable* if $M, s \models^{\nu} \varphi$ for some finite-state M. The formula φ is *valid* if $M, s \models^{\nu} \varphi$ for all M, s, and ν.

3 A Solution for Qualitative PCTL

In this section, we solve the satisfiability and the finite satisfiability problems for qualitative PCTL. To emphasize (and identify) the main difference between qualitative PCTL and CTL, we follow the traditional approach based on filtration through Fischer-Ladner closure [12]. In the case of CTL, the main problem with this technique is the introduction of new cycles which can "spoil" universally quantified formulae such as $AF\varphi$ [8]. In the case of qualitative PCTL, these new cycles are not a problem because, roughly speaking, they are either harmless or they are eventually left with probability 1. On the other hand, the *invalidity* of a formula $\varphi_1 U^{=1} \varphi_2$ cannot be witnessed just by a single run which violates the path formula $\varphi_1 U \varphi_2$, because this single run can have zero probability. This means that the heart of our construction for qualitative PCTL is actually rather different from the one for CTL.

We start by recalling a folklore observation which is used quite frequently in the proofs of our main results.

Lemma 3.1. *Let $M = (St, \rightarrow, Prob)$ be a Markov chain, $\nu : St \rightarrow 2^{Ap}$ a valuation, and $\varphi_1 U^{=1} \varphi_2$ a qualitative PCTL formula. If $M, s \not\models^\nu \varphi_1 U^{=1} \varphi_2$ for a given $s \in St$, then there are two possibilities:*

(a) There is a finite path $s{=}s_0 \rightarrow \cdots \rightarrow s_n$ such that $M, s_i \models^\nu \hat{\varphi}_2$ for all $0 \leq i \leq n$, and $M, s_n \models^\nu \hat{\varphi}_1$.

(b) $\mathcal{P}(\mathcal{R}) > 0$, where $\mathcal{R}$ is the set of all $w \in Run(s)$ such that $w(i) \models^\nu \varphi_1 \wedge \hat{\varphi}_2$ for all $i \in \mathbb{N}_0$. If M is finite-state, this is equivalent to the existence of a finite path $s{=}s_0 \rightarrow \cdots \rightarrow s_n$ and a BSCC α of M such that $s_n \in \alpha$ and for every state t that appears in the path or in α we have that $M, t \models^\nu \varphi_1 \wedge \hat{\varphi}_2$.

Proof. The first part follows from the fact that every run $s_0 s_1 \cdots$ is either of one of the forms specified in (a) and (b), or satisfies $\varphi_1 U \varphi_2$.

Let M be finite. If there is a path and a BSCC α as described above, then $\mathrm{Run}(s_0 \cdots s_n) \subseteq \mathcal{R}$, hence $0 < \prod_{i=0}^{n-1} \mathrm{Prob}(s_i \rightarrow s_{i+1}) \leq \mathcal{P}(\mathcal{R})$. On the other hand, let all BSCCs reachable from s contain a state in which $\varphi_1 \wedge \hat{\varphi}_2$ does not hold or are reachable only via paths containing such a state. Since a run reaches some BSCC with the probability 1 and then visits all states in this BSCC again with probability 1, we conclude that $\mathcal{P}(\mathcal{R}) = 0$. □

A proof of Lemma 3.1 is simple and standard. Now we recall the Fischer-Ladner closure [12].

Definition 3.2. *Let ψ be a qualitative PCTL formula. The* closure *of ψ, denoted $Cl(\psi)$, is the least set C of PCTL formulae such that $\psi \in C$ and the following conditions are satisfied:*

- *if $\varphi \in C$, then $\hat{\varphi} \in C$*
- *if $\varphi_1 \wedge \varphi_2 \in C$, then $\varphi_1, \varphi_2 \in C$*
- *if $X^{>0}\varphi \in C$, then $\varphi \in C$*
- *if $X^{=1}\varphi \in C$, then $\varphi \in C$*
- *if $\varphi_1 U^{>0} \varphi_2 \in C$, then $\varphi_1, \varphi_2, X^{>0}(\varphi_1 U^{>0} \varphi_2) \in C$*
- *if $\varphi_1 U^{=1} \varphi_2 \in C$, then $\varphi_1, \varphi_2, X^{=1}(\varphi_1 U^{=1} \varphi_2) \in C$*
- *if $\varphi_1 U^{=1} \varphi_2 \in C$, then $\varphi_1 U^{>0} \varphi_2 \in C$*

Definition 3.2 mimics the variant of Fischer-Ladner closure used in [8] for the logic CTL. The only notable difference is the last item, where we require that if $\varphi_1 U^{=1} \varphi_2$ is in the closure, then $\varphi_1 U^{>0} \varphi_2$ is also there. The exact purpose of this rule is clarified in the proofs of our main results.

In our next definition we identify certain formulae that should be satisfied "together" in a given state.

Definition 3.3. *A set $S \subseteq Cl(\psi)$ is* eligible *if for every $\varphi \in Cl(\psi)$ we have that φ or $\hat{\varphi}$ belongs to S, and the following conditions hold:*

- *if $\varphi \in S$, then $\hat{\varphi} \notin S$*
- *if $\varphi_1 \wedge \varphi_2 \in S$, then $\varphi_1, \varphi_2 \in S$*
- *if $\neg(\varphi_1 \wedge \varphi_2) \in S$, then $\hat{\varphi}_1 \in S$ or $\hat{\varphi}_2 \in S$*
- *if $\varphi_1 U^{>0} \varphi_2 \in S$, then $\varphi_2 \in S$ or $\varphi_1, X^{>0}(\varphi_1 U^{>0} \varphi_2) \in S$*
- *if $\neg(\varphi_1 U^{>0} \varphi_2) \in S$, then $\hat{\varphi}_1, \hat{\varphi}_2 \in S$ or $\hat{\varphi}_2, \neg X^{>0}(\varphi_1 U^{>0} \varphi_2) \in S$*
- *if $\varphi_1 U^{=1} \varphi_2 \in S$, then $\varphi_2 \in S$ or $\varphi_1, X^{=1}(\varphi_1 U^{=1} \varphi_2) \in S$*
- *if $\neg(\varphi_1 U^{=1} \varphi_2) \in S$, then $\hat{\varphi}_1, \hat{\varphi}_2 \in S$ or $\hat{\varphi}_2, \neg X^{=1}(\varphi_1 U^{=1} \varphi_2) \in S$*

Definition 3.4. *A* pseudo-structure *for a qualitative PCTL formula ψ is a pair $\mathcal{A} = (A, \rightarrow)$, where A is a set of eligible subsets of $Cl(\psi)$ and $\rightarrow \subseteq A \times A$ is a total relation.*

Every pseudo-structure $\mathcal{A} = (A, \rightarrow)$ for a formula ψ determines a unique Markov chain $M_{\mathcal{A}} = (A, \rightarrow, Prob)$ where *Prob* assigns a uniform probability distribution to every state. Further, to each $\varphi \in Cl(\psi)$ we associate a fresh atomic proposition $[\varphi]$ and define a valuation ν over A such that $[\varphi] \in \nu(S)$ iff $\varphi \in S$. In the following, we write

- $\mathcal{A}, S \models X^{\bowtie\rho}\varphi$ instead of $M_{\mathcal{A}}, S \models^{\nu} X^{\bowtie\rho}[\varphi]$
- $\mathcal{A}, S \models \neg X^{\bowtie\rho}\varphi$ instead of $M_{\mathcal{A}}, S \models^{\nu} \neg X^{\bowtie\rho}[\varphi]$
- $\mathcal{A}, S \models \varphi_1 U^{\bowtie\rho}\varphi_2$ instead of $M_{\mathcal{A}}, S \models^{\nu} [\varphi_1]U^{\bowtie\rho}[\varphi_2]$
- $\mathcal{A}, S \models \neg(\varphi_1 U^{\bowtie\rho}\varphi_2)$ instead of $M_{\mathcal{A}}, S \models^{\nu} \neg([\varphi_1]U^{\bowtie\rho}[\varphi_2])$

where "$\bowtie\rho$" is of the form ">0" or "$=1$".

As we already mentioned, the invalidity of a formula $\varphi_1 U^{=1}\varphi_2$ cannot be witnessed just by a single run that violates the path formula $\varphi_1 U \varphi_2$. A suitable witness for $\neg(\varphi_1 U^{=1}\varphi_2)$ is identified in our next definition.

Definition 3.5. *Let ψ be a qualitative PCTL formula and $\mathcal{A} = (A, \rightarrow)$ a pseudo-structure for ψ. A* witness *for a formula $\neg(\varphi_1 U^{=1}\varphi_2) \in Cl(\psi)$ in $\mathcal{A}$ is a pseudo-structure $\mathcal{B} = (B, \hookrightarrow)$ where $\emptyset \neq B \subseteq A$ and $\hookrightarrow \subseteq \rightarrow$ such that*

- *$\mathcal{B}$ is strongly connected.*
- *For every $S \in B$ we have that $\hat{\varphi}_2 \in S$.*
- *For every $S \in B$ and every $\xi_1 U^{=1}\xi_2 \in S$ we have that $\mathcal{B}, S \models \xi_1 U^{=1}\xi_2$.*

Definition 3.6. *Let ψ be a qualitative PCTL formula. A* pseudo-model *for ψ is a pseudo-structure $\mathcal{A} = (A, \rightarrow)$ for ψ such that $\psi \in T$ for some $T \in A$, and every $S \in A$ satisfies the following conditions:*

(1) If $\xi \in S$, where ξ is of the form $X^{=1}\varphi$, $\neg X^{=1}\varphi$, $X^{>0}\varphi$, $\neg X^{>0}\varphi$, $\varphi_1 U^{>0}\varphi_2$, $\neg(\varphi_1 U^{>0}\varphi_2)$, or $\varphi_1 U^{=1}\varphi_2$, then $\mathcal{A}, S \models \xi$.

(2) If $\neg(\varphi_1 U^{=1}\varphi_2) \in S$, then one of the following conditions is satisfied:

(a) $\mathcal{A}, S \not\models \varphi_1 U^{=1}\varphi_2$.

(b) There is a witness $\mathcal{B} = (B, \hookrightarrow)$ for $\neg(\varphi_1 U^{=1} \varphi_2)$ and a finite path $S_0 \rightarrow \cdots \rightarrow S_n$ such that $S_0{=}S$, $S_n \in B$, and $\hat{\varphi}_2 \in S_i$ for every $0 \le i \le n$.

A pseudo-model is simple *if the condition (2) is always satisfied by item (a), i.e., no witness is employed.*

The next theorem says that the satisfiability of a given qualitative PCTL formula is always certified by a pseudo-model.

Theorem 3.7. *Let ψ be a qualitative PCTL formula. If ψ is satisfiable, then there is a pseudo-model $\mathcal{A} = (A, \mapsto)$ for ψ. Moreover, if ψ is finite-satisfiable, then $\mathcal{A}$ is simple.*

Proof. Let us fix a satisfiable qualitative PCTL formula ψ. Then there is a Markov chain $M = (St, \rightarrow, Prob)$, a valuation ν, and a state $s_\psi \in St$ such that $M, s_\psi \models^\nu \psi$. For every $s \in St$, let $[s] = \{\varphi \in Cl(\psi) \mid M, s \models^\nu \varphi\}$. We define $A = \{[s] \mid s \in St\}$, and $[s] \mapsto [t]$ iff there are some $s', t' \in St$ such that $[s] = [s']$, $[t] = [t']$, and $s' \rightarrow t'$. We show that $\mathcal{A} = (A, \mapsto)$ is a pseudo-model for ψ. Clearly, all elements of A are eligible, and $\psi \in [s_\psi]$. Now let $[s] \in A$ and $\xi \in [s]$. We verify the two conditions of Definition 3.6.

Condition (1). Let ξ be of the form $X^{=1}\varphi$. If $[s] \mapsto [t]$, then $s' \rightarrow t'$ for some some $[s'] = [s], [t'] = [t]$. Since $M, s' \models^\nu X^{=1}\varphi$, we have $M, t' \models^\nu \varphi$, hence $\varphi \in [t'] = [t]$. Similarly for $\neg X^{>0}\varphi$.

Let ξ be of the form $\varphi_1 U^{>0} \varphi_2$. Since $M, s \models^\nu \varphi_1 U^{>0} \varphi_2$, there is a path $s = s_0 \rightarrow \cdots \rightarrow s_n$, such that $M, s_i \models^\nu \varphi_1$ for $0 \le i < n$ and $M, s_n \models^\nu \varphi_2$. Therefore, there is a path $[s] = [s_0] \mapsto \cdots \mapsto [s_n]$, where $\varphi_1 \in [s_i]$ for $0 \le i < n$ and $\varphi_2 \in [s_n]$. Hence $\mathcal{A}, [s] \models \varphi_1 U^{>0} \varphi_2$. Similarly for $X^{>0}\varphi$ and $\neg X^{=1}\varphi$.

Let $\xi = \neg(\varphi_1 U^{>0} \varphi_2)$, and let us assume (for the sake of deriving a contradiction) that $\mathcal{A}, [s] \not\models \neg(\varphi_1 U^{>0} \varphi_2)$, i.e., $\mathcal{A}, [s] \models \varphi_1 U^{>0} \varphi_2$. Then there is is finite path $[s_0] \mapsto \cdots \mapsto [s_n]$, where $n \in \mathbb{N}_0$, $[s] = [s_0]$, $\varphi_2 \in [s_n]$, and $\varphi_1 \in [s_i]$ for all $0 \le i < n$. By induction on n we show that some $[s_i]$ in this finite path is not eligible, which is a contradiction.

- $n = 0$. Then we have that $\varphi_2 \in [s_0]$ and $\neg(\varphi_1 U^{>0} \varphi_2) \in [s_0]$, which means that $\varphi_2, \hat{\varphi}_2 \in [s_0]$, hence $[s_0]$ is not eligible.
- **Induction step.** Since $\neg(\varphi_1 U^{>0} \varphi_2) \in [s_0]$, there are two possibilities (see Definition 3.3): Either $\hat{\varphi}_1, \hat{\varphi}_2 \in [s_0]$, which means that $\hat{\varphi}_1, \varphi_1 \in [s_0]$ and hence $[s_0]$ is not eligible, or $\neg X^{>0}(\varphi_1 U^{>0} \varphi_2) \in [s_0]$, which means that $\neg(\varphi_1 U^{>0} \varphi_2) \in [s_1]$ and we can apply induction hypothesis.

Let $\xi = \varphi_1 U^{=1} \varphi_2$, and let us assume that $\mathcal{A}, [s] \not\models \varphi_1 U^{=1} \varphi_2$. According to Lemma 3.1, there are two possibilities:

(a) There is a finite path $[s_0] \mapsto \cdots \mapsto [s_n]$, where $n \in \mathbb{N}_0$, $[s] = [s_0]$, $\hat{\varphi}_2 \in [s_i]$ for all $0 \leq i \leq n$, and $\hat{\varphi}_1 \in [s_n]$. Since $\hat{\varphi}_1, \hat{\varphi}_2 \in [s_n]$, we have that $M, s_n \models^{\nu} \hat{\varphi}_1 \wedge \hat{\varphi}_2$, hence $M, s_n \models^{\nu} \varphi_1 U^{=0} \varphi_2$. By a straightforward induction on j we can show that $M, s_{n-j} \models^{\nu} \varphi_1 U^{<1} \varphi_2$ for all $0 \leq j \leq n$, which means that $\varphi_1 U^{=1} \varphi_2 \notin [s_0] = [s]$, and we have a contradiction.

(b) There is a BSCC α of $\mathcal{A}$ such that $\hat{\varphi}_2 \in [t]$ for every $[t] \in \alpha$, and a finite path $[s_0] \mapsto \cdots \mapsto [s_n]$ such that $n \in \mathbb{N}_0$, $[s] = [s_0]$, $[s_n] \in \alpha$, and $\hat{\varphi}_2 \in [s_i]$ for all $0 \leq i \leq n$. First, let us realize that if $[t] \in \alpha$, then for every $t' \in \mathit{St}$ such that $t \rightarrow^* t'$ we have that $M, t' \not\models^{\nu} \varphi_2$. Hence, $M, t \models^{\nu} \varphi_1 U^{=0} \varphi_2$. Since $[s_n] \in \alpha$, we have that $M, s_n \models^{\nu} \varphi_1 U^{=0} \varphi_2$. From this we obtain (similarly as in (a)) that $\varphi_1 U^{=1} \varphi_2 \notin [s_0] = [s]$, which is a contradiction.

Condition (2). Let $\neg(\varphi_1 U^{=1} \varphi_2) \in [s]$. First we show that if M is a finite-state Markov chain (i.e., the formula ψ is finite satisfiable), then $\mathcal{A}, [s] \not\models \varphi_1 U^{=1} \varphi_2$. Since $M, s \not\models \varphi_1 U^{=1} \varphi_2$ and M is a finite-state Markov chain, there are two possibilities (see Lemma 3.1):

(a) There is a finite path $s = s_0 \rightarrow \cdots \rightarrow s_n$, $n \in \mathbb{N}_0$, such that $M, s_i \models^{\nu} \hat{\varphi}_2$ for all $0 \leq i \leq n$, and $M, s_n \models^{\nu} \hat{\varphi}_1$. But then $[s] = [s_0] \mapsto \cdots \mapsto [s_n]$, $\hat{\varphi}_2 \in [s_i]$ for all $0 \leq i \leq n$, and $\hat{\varphi}_1 \in [s_n]$, which means that $\mathcal{A}, [s] \not\models \varphi_1 U^{=1} \varphi_2$ as needed.

(b) Since A is finite, there is a BSCC α of M such that $M, t \models^{\nu} \hat{\varphi}_2$ for every $t \in \alpha$, and a finite path $s = s_0 \rightarrow \cdots \rightarrow s_n$, $n \in \mathbb{N}_0$, such that $s_n \in \alpha$ and $M, s_i \models^{\nu} \hat{\varphi}_2$ for all $0 \leq i \leq n$. Consider the finite path $[s]=[s_0] \mapsto \cdots \mapsto [s_n]$. Clearly $\hat{\varphi}_2 \in [s_i]$ for all $0 \leq i \leq n$. We show that $\mathcal{A}, [s_n] \models \neg(\varphi_1 U^{>0} \varphi_2)$, which implies that $\mathcal{A}, [s] \not\models \varphi_1 U^{=1} \varphi_2$ as needed.

Since $M, t \models^{\nu} \neg(\varphi_1 U^{>0} \varphi_2)$ for every $t \in \alpha$, we have that $\neg(\varphi_1 U^{>0} \varphi_2) \in [t]$ for every $t \in \alpha$ (here we rely on the last rule of Definition 3.2 which guarantees that $\neg(\varphi_1 U^{>0} \varphi_2) \in \mathit{Cl}(\psi)$). In particular, $\neg(\varphi_1 U^{>0} \varphi_2) \in [s_n]$. By applying the analysis of Condition (1), we can conclude that $\mathcal{A}, [s_n] \models \neg(\varphi_1 U^{>0} \varphi_2)$.

Now consider the general case when the Markov chain M is not necessarily finite-state. Since $M, s \not\models \varphi_1 U^{=1} \varphi_2$, there are two possibilities (see Lemma 3.1):

(a) There is a finite path $s = s_0 \rightarrow \cdots \rightarrow s_n$, $n \in \mathbb{N}_0$, such that $M, s_i \models^{\nu} \hat{\varphi}_2$ for all $0 \leq i \leq n$, and $M, s_n \models^{\nu} \hat{\varphi}_1$. Then we obtain $\mathcal{A}, [s] \not\models \varphi_1 U^{=1} \varphi_2$ as in (a) above.

(b) $\mathcal{P}(\mathcal{R}) > 0$, where $\mathcal{R} = \{w \in \mathit{Run}(s) \mid M, w(i) \models \varphi_1 \wedge \hat{\varphi}_2\}$. For every $w \in \mathcal{R}$, the *type* of w is the pseudo-structure $\mathcal{B} = (B, \hookrightarrow)$ where

- $[t] \in B$ iff there are infinitely many $i \in \mathbb{N}_0$ such that $[w(i)] = [t]$;
- $[t] \hookrightarrow [u]$ iff there are infinitely many $i \in \mathbb{N}_0$ such that $[w(i)] = [t]$, $[w(i{+}1)] = [u]$, and $w(i) \rightarrow w(i{+}1)$.

For every type $\mathcal{B}$, let $\mathcal{R}(\mathcal{B}) = \{w \in \mathcal{R} \mid \text{the type of } w \text{ is } \mathcal{B}\}$. Since there are only finitely many types, $\mathcal{R} = \bigcup_{\mathcal{B}} \mathcal{R}(\mathcal{B})$, and $\mathcal{P}(\mathcal{R}) > 0$, there must be a type $\mathcal{B}$ such that $\mathcal{P}(\mathcal{R}(\mathcal{B})) > 0$. For the rest of this proof, let us fix such a type $\mathcal{B} = (B, \hookrightarrow)$. We claim that $\mathcal{B}$ is a witness for $\neg(\varphi_1 U^{=1} \varphi_2)$. Clearly $\mathcal{B}$ is strongly connected and $\hat{\varphi}_2 \in [t]$ for every $[t] \in B$. It remains to show that for every $[t] \in B$ and every $\xi_1 U^{=1} \xi_2 \in [t]$ we have that $\mathcal{B}, [t] \models \xi_1 U^{=1} \xi_2$. Suppose that $\mathcal{B}, [t] \not\models \xi_1 U^{=1} \xi_2$. Since $\mathcal{B}$ has finitely many states and is strongly connected, there are two possibilities (see Lemma 3.1):

(a) There is a finite path $[t_0] \hookrightarrow \cdots \hookrightarrow [t_n]$, where $n \in \mathbb{N}_0$, $[t] = [t_0]$, $\hat{\xi}_2 \in [t_i]$ for all $0 \leq i \leq n$, and $\hat{\varphi}_1 \in [t_n]$. Then we obtain $\xi_1 U^{=1} \xi_2 \notin [t_0] = [t]$ in the same way as in the paragraph Condition (1) (a).

(b) $\xi_1, \hat{\xi}_2 \in [u]$ for every $[u] \in B$ (note that $\mathcal{B}$ is strongly connected). We show that $\xi_1 U^{=1} \xi_2 \notin [t]$. For this we need the observation formulated in the next paragraph.

For every $u \in \mathit{St}$ such that $[u] \in B$ we define $\mathit{Run}(u, \mathcal{B})$ as the set of all $w \in \mathit{Run}(u)$ such that $[w(i)] \hookrightarrow [w(i{+}1)]$ for all $i \in \mathbb{N}_0$. We claim that for every $u \in \mathit{St}$ such that $[u] \in B$ there is some $u' \in \mathit{St}$ such that $[u] = [u']$ and $\mathcal{P}(\mathit{Run}(u', \mathcal{B})) > 0$. Suppose that this condition is violated by some $u \in \mathit{St}$. Let $\mathcal{H}$ be the set of all finite paths in M initiated in s and terminated in some $u' \in \mathit{St}$ where $[u'] = [u]$. Since every run of $\mathcal{R}(\mathcal{B})$ eventually visits some u' such that $[u'] = [u]$, we have that $\mathcal{R}(\mathcal{B}) \subseteq \bigcup_{v \in \mathcal{H}} \big(v \cdot \mathit{Run}(u_v, \mathcal{B})\big)$, where u_v is the last state of v and $v \cdot \mathit{Run}(u_v, \mathcal{B})$ the set of all runs of the form $v\bar{w}$ where $u_v\bar{w} \in \mathit{Run}(u_v, \mathcal{B})$. Hence, $\mathcal{P}(\mathcal{R}(\mathcal{B})) \leq \sum_{v \in \mathcal{H}} \mathcal{P}(\mathit{Run}(v)) \cdot \mathcal{P}(\mathit{Run}(u_v, \mathcal{B}))$. Since $\mathcal{P}(\mathit{Run}(u_v, \mathcal{B}) = 0$ for every $v \in \mathcal{H}$, we obtain $\mathcal{P}(\mathcal{R}(\mathcal{B})) = 0$, which is a contradiction.

Due to the observation formulated in the previous paragraph, there is some $t' \in \mathit{St}$ such that $[t'] = [t]$ and $\mathcal{P}(\mathit{Run}(t', \mathcal{B})) > 0$. Since $M, w(i) \not\models^{\nu} \xi_2$

for every $w \in Run(t', \mathcal{B})$ and $i \in \mathbb{N}_0$, we obtain that $M, t' \not\models^{\nu} \xi_1 U^{=1} \xi_2$, hence $\xi_1 U^{=1} \xi_2 \notin [t'] = [t]$. □

According to Theorem 3.7, every satisfiable qualitative PCTL formula has a finite pseudo-model. Now we show that this pseudo-model can be turned into a "real" model, which can be infinite but always admits a finite description in the form of a *marked graph*.

Definition 3.8. *A* marked graph *is a triple $\mathcal{G} = (G, \hookrightarrow, L)$ where G is a finite set of* nodes, *$\hookrightarrow \subseteq G \times G$ is a total relation, and $L \subseteq \hookrightarrow$ a subset of* marked *transitions.*

Each marked graph $\mathcal{G} = (G, \hookrightarrow, L)$ determines a unique Markov chain $M_{\mathcal{G}} = (G^+, \to, Prob)$ where

- for all $w \in G^*$ and $v \in G$, $wv \to \bar{w}$ iff $\bar{w} = wvv'$ for some $v' \in G$ such that $v \hookrightarrow v'$. We say that $wv \to wvv'$ is *marked* iff $v \hookrightarrow v'$ is marked;
- for every $w \in G^+$, $Prob(w)$ is a uniform distribution if none or all outgoing transitions of w are marked. Otherwise, $Prob$ assigns the same probability p to all non-marked transitions and the same probability p' to all marked transitions so that $\sum_{w \to w' \in L} p$ is equal to $1 - 1/4^{len(w)}$.

In other words, each marked graph $\mathcal{G}$ is a finite representation of an infinite-state Markov chain $M_{\mathcal{G}}$ obtained by "unfolding" the structure of $\mathcal{G}$ and assigning larger and larger probabilities to marked transitions.

Remark 3.9. *Let $\mathcal{G} = (G, \hookrightarrow, L)$ be a marked graph. For every run $w \in Run(u)$ of $\mathcal{G}$ there is a corresponding run $\bar{w} \in Run(u)$ of $M_{\mathcal{G}}$, where $\bar{w}(k) = w(0) \cdots w(k)$ for every $k \in \mathbb{N}_0$. Further, each $\eta : G \to 2^{Ap}$ can be naturally extended to a valuation $\eta' : G^+ \to 2^{Ap}$ where $\eta'(wv) = \eta(v)$ for every $w \in G^*$ and $v \in G$. For the sake of simplicity, we introduce the following notation:*

- *for every PCTL formula φ and every $\eta : G \to 2^{Ap}$, we write $\mathcal{G}, v \models^{\eta} \varphi$ iff $M_{\mathcal{G}}, v \models^{\eta'} \varphi$*
- *for every set of runs $\mathcal{L} \subseteq Run(u)$ in $\mathcal{G}$ we put $\bar{\mathcal{L}} = \{\bar{w} \mid w \in \mathcal{L}\}$ and define $\mathcal{P}(\mathcal{L}) = \mathcal{P}(\bar{\mathcal{L}})$ whenever $\bar{\mathcal{L}}$ is measurable.*

Let v be a finite path in $\mathcal{G}$ initiated in some $S \in G$. One can easily prove that for every $\mathcal{L} \subseteq Run(v)$ such that $\mathcal{P}(\mathcal{L})$ is defined we have that $\mathcal{P}(\mathcal{L}) = \mathcal{P}(Run(v)) \cdot \mathcal{P}(\{w \mid vw \in \mathcal{L}\})$, and this observation is frequently used in the proof of Theorem 3.10.

Theorem 3.10. *Let ψ be a qualitative PCTL formula. If there is a pseudo-model $\mathcal{A}$ for ψ, then there is a marked graph $\mathcal{G} = (G, \hookrightarrow, L)$ whose size is exponential in $|\varphi|$, a valuation $\eta : G \to 2^{Ap}$, and $v \in G$ such that $\mathcal{G}, v \models^{\eta} \psi$. Moreover, if $\mathcal{A}$ is simple, then $L = \emptyset$ and ψ has a finite-state model whose size is exponential in $|\psi|$.*

Proof. Let $\mathcal{A} = (A, \mapsto)$ be a pseudo-model for ψ. If $\mathcal{A}$ is simple, we can put $\mathcal{G} = (A, \mapsto, \emptyset)$ and $\eta(S) = \{p \in Ap \mid p \in S\}$ for every $S \in A$. It is easy to verify that for every $S \in A$ and $\varphi \in S$ we have that $\mathcal{G}, S \models^{\eta} \varphi$. (A proof is a straightforward induction on the structure of φ, where all subcases follow immediately from Definition 3.6).

If $\mathcal{A}$ is not simple, we proceed as follows. Let $\mathcal{B}_i = (B_i, \rightsquigarrow_i)$, $1 \le i \le m$, be a family of pseudo-structures such that

- for every $S \in A$ and every $\neg(\varphi_1 U^{=1} \varphi_2) \in S$ there is a suitable $\mathcal{B}_i$ and a finite path $S = S_0 \mapsto \cdots \mapsto S_n$ such that $\hat{\varphi}_2 \in S_j$ for every $0 \le j \le n$, $S_n \in B_i$, and $\mathcal{B}_i$ is a witness for $\neg(\varphi_1 U^{=1} \varphi_2)$ in $\mathcal{A}$;
- each $\mathcal{B}_i$ is a witness for some $\neg(\varphi_1 U^{=1} \varphi_2) \in Cl(\psi)$ in $\mathcal{A}$.

Since we need at most one witness for every $S \in A$ and every $\neg(\varphi_1 U^{=1} \varphi_2) \in S$, we can safely assume that $m \le |A| \cdot |\psi|$.

The nodes of $\mathcal{G}$ are obtained by taking the disjoint union of A and all B_i, $1 \le i \le m$. Formally, we put $G = \bigcup_{i=0}^{m} B_i \times \{i\}$, where $B_0 = A$. The $\hookrightarrow$ and L are defined as follows: if $S \mapsto T$, then

- $(S, 0) \hookrightarrow (T, i)$ for every $0 \le i \le m$ such that $T \in B_i$;
- for every $1 \le i \le m$ such that $S, T \in B_i$ and $S \rightsquigarrow_i T$ we have that $(S, i) \hookrightarrow (T, i)$ and this transition is marked;
- for every $1 \le i \le m$ such that $S \in B_i$ and $T \notin B_i$ we have that $(S, i) \hookrightarrow (T, 0)$.

Both $\hookrightarrow$ and L contain only those transitions which can be derived by the above rule. Note that the size of $\mathcal{G}$ is exponential in $|\psi|$.

Before continuing with the main proof, we need to formulate auxiliary observations about $\mathcal{G}$. We say that a run w of $\mathcal{G}$ *stays at* i, where $0 \le i \le m$, if $w(k) \in B_i \times \{i\}$ for every $k \in \mathbb{N}_0$. We say that a run w *enters* i, where $0 \le i \le m$, if w is of the form $u\bar{w}$, where $\bar{w}$ stays at i. Now observe that

(i) for every $(S, i) \in G$ where $i \geq 1$, the probability of all $w \in \mathit{Run}((S, i))$ that stay at i is at least 2/3. To see this, realize that the probability of all $w \in \mathit{Run}((S, i))$ such that a non-marked transition is performed in w is bounded by $\sum_{k=1}^{\infty} 1/4^k = 1/3$. This follows directly from the definition of $M_{\mathcal{G}}$.

(ii) for every $(S, i) \in G$, the probability of all $w \in \mathit{Run}((S, i))$ such that w does *not* enter any j, where $0 \leq j \leq m$, is zero. This is an immediate consequence of the previous observation.

(iii) for every $(S, i) \in G$ and every $\xi_1 U^{=1} \xi_2 \in S$, the conditional probability of all $w \in \mathit{Run}((S, i))$ such that $\mathcal{G}, w \models \xi_1 U \xi_2$, under the condition that w stays at i, is equal to 1. This follows from Definition 3.5.

Now we continue with the main proof. Let η be a valuation given by $\eta((S, i)) = \{p \in \mathit{Ap} \mid p \in S\}$ for every $0 \leq i \leq m$. We show that for every $\varphi \in \mathit{Cl}(\psi)$ and every $(S, i) \in G$ we have that $\varphi \in S$ iff $\mathcal{G}, (S, i) \models^{\eta} \varphi$ (from now on, the η in $\models^{\eta}$ is omitted). We proceed by induction on the structure of φ.

- The cases when φ is of the form p, $\neg\varphi_1$, $\varphi_1 \wedge \varphi_2$, $X^{>0}\varphi_1$, or $X^{=1}\varphi_1$ follow immediately.

- Let φ be of the form $\varphi_1 U^{>0} \varphi_2$. If $\varphi_1 U^{>0} \varphi_2 \in S$, then $\mathcal{A}, S \models \varphi_1 U^{>0} \varphi_2$, and there is a finite path $S{=}S_0 \mapsto \cdots \mapsto S_n$, where $\varphi_1 \in S_j$ for every $0 \leq j < n$ and $\varphi_2 \in S_n$. Hence, $(S, i){=}(S_0, i_0) \hookrightarrow \cdots \hookrightarrow (S_n, i_n)$, where $0 \leq i_j \leq m$ for every $0 \leq j \leq n$ (this follows from the definition of $\mathcal{G}$). Further, $\mathcal{G}, (S_j, i_j) \models \varphi_1$ for every $0 \leq j < n$ and $\mathcal{G}, (S_n, i_n) \models \varphi_2$ (by induction hypothesis), hence $\mathcal{G}, (S, i) \models \varphi_1 U^{>0} \varphi_2$ as required. Similarly, we show that if $\mathcal{G}, (S, i) \models \varphi_1 U^{>0} \varphi_2$, then $\mathcal{A}, S \models \varphi_1 U^{>0} \varphi_2$, hence $\varphi_1 U^{>0} \varphi_2 \in S$ as needed.

- Let φ be of the form $\varphi_1 U^{=1} \varphi_2$. We start with the "$\Rightarrow$" direction, i.e., we prove that if $\mathcal{G}, (S, i) \not\models \varphi_1 U^{=1} \varphi_2$ for some $(S, i) \in G$, then $\varphi_1 U^{=1} \varphi_2 \notin S$.

  Since $\mathcal{G}, (S, i) \not\models \varphi_1 U^{=1} \varphi_2$, there are two possibilities (see Lemma 3.1):

  (a) There is a finite path $(S, i){=}(S_0, i_0) \hookrightarrow \cdots \hookrightarrow (S_n, i_n)$ such that $\mathcal{G}, (S_j, i_j) \models \hat{\varphi}_2$ for every $0 \leq j \leq n$ and $\mathcal{G}, (S_n, i_n) \models \hat{\varphi}_1$. But then $S{=}S_0 \mapsto \cdots \mapsto S_n$ where $\mathcal{A}, S_j \models \hat{\varphi}_2$ for every $0 \leq j \leq n$ and $\mathcal{A}, S_n, \models \hat{\varphi}_1$ (by induction hypothesis), which means $\mathcal{A}, S \not\models \varphi_1 U^{=1} \varphi_2$, and hence $\varphi_1 U^{=1} \varphi_2 \notin S$ by definition.

(b) We have $\mathcal{P}(\mathcal{R}) > 0$, where $\mathcal{R} = \{w \in \mathit{Run}((S,i)) \mid \mathcal{G}, w(k) \models \varphi_1 \wedge \hat{\varphi}_2\}$. Let us assume that $\varphi_1 U^{=1} \varphi_2 \in S$, for a contradiction. Observe that for every run (S,i)=$(S_0,i_0),(S_1,i_1),\cdots$ of $\mathcal{R}$ and every $k \in \mathbb{N}_0$ we have that $\varphi_1 U^{=1}\varphi_2 \in S_k$. This can be shown by induction k:

- **$\mathbf{k = 0}$**. This is immediate because $\varphi_1 U^{=1}\varphi_2 \in S$.
- **Induction step:** Let us assume $\varphi_1 U^{=1}\varphi_2 \in S_{k-1}$ and $(S_{k-1}, i_{k-1}) \hookrightarrow (S_k, i_k)$. Since we have $\mathcal{G},(S_{k-1},i_{k-1}) \not\models \varphi_2$, we obtain $\varphi_2 \notin S_{k-1}$ by induction hypothesis (here we consider the "outer" structural induction). Since S_{k-1} is eligible, we have that $X^{=1}(\varphi_1 U^{=1}\varphi_2) \in S_{k-1}$ (see Definition 3.3). Since $(\mathcal{A}, \mapsto)$ is a pseudo-model and $S_{k-1} \mapsto S_k$, we obtain $\varphi_1 U^{=1}\varphi_2 \in S_k$ as needed.

For every $0 \leq j \leq m$, let $\mathcal{R}_j = \{w \in \mathcal{R} \mid w \text{ enters } j\}$. Due to observation (ii) above, $\mathcal{P}(\mathcal{R}) = \sum_{j=0}^{m} \mathcal{P}(\mathcal{R}_j)$, and hence there is some j such that $\mathcal{P}(\mathcal{R}_j) > 0$. For the rest of this paragraph, let us fix such a j. Let $\mathcal{H}$ be the set of all finite paths initiated in (S,i). For every $u \in \mathcal{H}$, let $\mathcal{R}_j(u) = \{w \in \mathcal{R}_j \mid w = u\bar{w} \text{ where } \bar{w} \text{ stays at } j\}$. Since $\mathcal{P}(\mathcal{R}_j) > 0$ and $\mathcal{R}_j = \bigcup_{u \in \mathcal{H}} \mathcal{R}_j(u)$, there is some $u \in \mathcal{H}$ such that $\mathcal{P}(\mathcal{R}_j(u)) > 0$. For the rest of this paragraph, we fix such a u. Let $u = u'(S_k, i_k)$, and let $\mathcal{L} = \{\bar{w} \mid u'\bar{w} \in \mathcal{R}_j(u)\}$. Since $\mathcal{P}(\mathcal{R}_j(u)) > 0$, we have that $\mathcal{P}(\mathcal{L}) > 0$. Since $\varphi_1 U^{=1}\varphi_2 \in S_k$ and $\mathcal{G}, \bar{w} \not\models \varphi_1 U \varphi_2$ for every $\bar{w} \in \mathcal{L}$, we obtain a contradiction with observation (iii).

It remains to prove the "$\Leftarrow$" direction. We show that if $\varphi_1 U^{=1}\varphi_2 \notin S$, then $\mathcal{G},(S,i) \not\models \varphi_1 U^{=1}\varphi_2$ for every $0 \leq i \leq m$ such that $(S,i) \in G$. According to Definition 3.6, two cases arise:

(a) $\mathcal{A}, S \not\models \varphi_1 U^{=1}\varphi_2$. Since A is finite-state, there are two possibilities (see Lemma 3.1):

i. There is a finite path S=$S_0 \mapsto \cdots \mapsto S_n$ where $\mathcal{A}, S_j \models \hat{\varphi}_2$ for every $0 \leq j \leq n$ and $\mathcal{A}, S_n, \models \hat{\varphi}_1$. From this we easily obtain that $\mathcal{G},(S,i) \not\models \varphi_1 U^{=1}\varphi_2$.

ii. There is a BSCC α of $\mathcal{A}$ such that $\mathcal{A}, T \models \hat{\varphi}_2$ for every $T \in \alpha$, and a finite path $S = S_0 \mapsto \cdots \mapsto S_n$, $n \in \mathbb{N}_0$, such that $S_n \in \alpha$ and $\mathcal{A}, S_i \models \hat{\varphi}_2$ for all $0 \leq i \leq n$. Then (S,i)=$(S_0,i_0) \hookrightarrow \cdots \hookrightarrow (S_n,i_n)$ where $\mathcal{G},(S_n,i_n) \models \varphi_1 U^{=0}\varphi_2$. From this we get $\mathcal{G},(S,i) \not\models \varphi_1 U^{=1}\varphi_2$.

(b) There is a witness $\mathcal{B}_k = (B_k, \rightsquigarrow_k)$ and a finite path $S = S_0 \mapsto \cdots \mapsto S_n \in B_k$, where $\hat{\varphi}_2 \in S_j$ for every $0 \leq j \leq n$, and $\hat{\varphi}_2 \in T$ for every $T \in B_k$. Hence also

$(S, i) = (S_0, j_0) \hookrightarrow \cdots \hookrightarrow (S_n, j_n)$ where $j_n = k$. Let $\mathcal{R} = \{w \in \mathrm{Run}((S, i)) \mid w = (S_0, j_0) \cdots (S_n, j_n)u$, where u stays at $k\}$. Since $\mathcal{G}, w(\ell) \models \hat{\varphi}_2$ for every $w \in \mathcal{R}$ and $\ell \in \mathbb{N}_0$, and $\mathcal{P}(\mathcal{R}) > 0$ using observation (i), we receive $\mathcal{G}, (S, i) \not\models \varphi_1 U^{=1} \varphi_2$. □

Note that the construction of the marked graph $\mathcal{G}$ in Theorem 3.10 is effective (provided that the family $\mathcal{B}_i$ of witness has already been computed). Hence, it remains to show how to compute a (simple) pseudo-model for a given qualitative PCTL formula (if it exists) together with the associated family of witnesses.

Theorem 3.11. *Let ψ be a qualitative PCTL formula. The existence of a (simple) pseudo-model for ψ is decidable in time exponential in $|\psi|$. Moreover, if a (simple) pseudo-model for ψ exists, it can be effectively constructed in time exponential in $|\psi|$.*

Proof. Let ψ be a qualitative PCTL formula. An algorithm for constructing a (finite) pseudo-model for ψ is given in Figure 1. The algorithm executes either the line `10a` or `10b` (not both), depending on whether the constructed pseudo-model is to be simple or not, respectively. We show that the algorithm has the required properties. This is done in three steps. We show that

(a) if the algorithm returns some $\mathcal{A} = (A, \mapsto)$, then $\mathcal{A}$ is a (simple) pseudo-model for ψ;

(b) if ψ is (finite) satisfiable, then the algorithm returns a (simple) pseudo-model for ψ;

(c) the algorithm terminates in time which is exponential in $|\psi|$.

Step (a). It suffices to verify the conditions stated in Definition 3.6. Let $S \in A$ and $\xi \in S$. If ξ is of the form $X^{=1}\varphi$ or $\neg X^{>0}\varphi$, then $\mathcal{A}, S \models \xi$ because only the "safe" outgoing edges of S satisfy the conditions given at line 3 and line 4, respectively. Similarly, if ξ is of the form $X^{>0}\varphi$, $\neg X^{=1}\varphi$, or $\varphi_1 U^{>0} \varphi_2$, then $\mathcal{A}, S \models \xi$ because otherwise S would have to be deleted from A at line 7.

Now let $\xi \equiv \neg(\varphi_1 U^{>0} \varphi_2)$. We need to show that $\mathcal{A}, S \models \xi$. Suppose the converse, i.e., $\mathcal{A}, S \models \varphi_1 U^{>0} \varphi_2$. Then there is a finite path $S = S_0 \mapsto \cdots \mapsto S_n$, where $\varphi_2 \in S_n$ and $\varphi_1 \in S_i$ for all $0 \leq i < n$. By induction on n we show that some S_i in this finite path is not eligible, which is a contradiction.

- $n = 0$. Then $\varphi_2 \in S_0$ and $\neg(\varphi_1 U^{>0} \varphi_2) \in S_0$, which means that $\varphi_2, \hat{\varphi}_2 \in S_0$, hence S_0 is not eligible.

Input: A qualitative PCTL formula ψ.

Output: A (simple) pseudo-model $\mathcal{A} = (A, \mapsto)$ if ψ is (finite) satisfiable, *unsatisfiable* otherwise.

```
1:    $A$ := the set of all eligible subsets of $Cl(\psi)$
2:    $\mapsto := A \times A$
3:    for all $S \in A, X^{=1} \in S$ do
          delete all edges $S \mapsto T$ where $\varphi \notin T$
4:    for all $S \in A, \neg X^{>0} \in S$ do
          delete all edges $S \mapsto T$ where $\varphi \in T$
5:    repeat
6:        for all $S \in A, \xi \in S$ (in any order) do
7:            if $\xi \equiv X^{>0}\varphi$ or $\xi \equiv \neg X^{=1}\varphi$ or $\xi \equiv \varphi_1 U^{>0} \varphi_2$ then
                  if $\mathcal{A}, S \not\models \xi$ then $A := A \smallsetminus \{S\}$
8:            if $\xi \equiv \varphi_1 U^{=1} \varphi_2$ then
                  for every BSCC $B$ of $(A, \mapsto)$ (in any order) do
                      if $\varphi_1, \hat{\varphi}_2, \varphi_1 U^{=1} \varphi_2 \in T$ for every $T \in B$ then
                          $A := A \smallsetminus B$
                  done
9:            if $\xi \equiv \neg(\varphi_1 U^{=1} \varphi_2)$ then
                  if there is no finite path $S = S_0 \mapsto \cdots \mapsto S_n$ where $\neg(\varphi_1 U^{>0} \varphi_2) \in S_n$ and
                  $\varphi_1, \hat{\varphi}_2 \in S_i$ for all $0 \leq i < n$ then
10a:                  $A := A \smallsetminus \{S\}$
10b:              if there is no witness $(B, \hookrightarrow)$ for $\neg(\varphi_1 U^{=1} \varphi_2)$ in $(A, \mapsto)$ such that there
                  is a finite path $S = S_0 \mapsto \cdots \mapsto S_n$ where $S_n \in B$ and $\hat{\varphi}_2 \in S_i$ for all
                  $0 \leq i \leq n$ then
                      $A := A \smallsetminus \{S\}$
11:           repeat
                  $\mapsto := \mapsto \cap (A \times A)$
                  $A := A \smallsetminus \{S \in A \mid S$ has no outgoing edges$\}$
              until $A$ does not change
          done
12:   until $(A, \mapsto)$ does not change
13:   if $\psi \in S$ for some $S \in A$ then
          return $\mathcal{A} = (A, \mapsto)$
      else return unsatisfiable
```

Figure 1: An algorithm for constructing a (simple) pseudo-model.

- **Induction step.** Since $\neg(\varphi_1 U^{>0} \varphi_2) \in S_0$, there are two possibilities (see Definition 3.3): Either $\hat{\varphi}_1, \hat{\varphi}_2 \in S_0$, which means that $\hat{\varphi}_1, \varphi_1 \in S_0$ and hence S_0 is not eligible, or $\neg X^{>0}(\varphi_1 U^{>0} \varphi_2) \in S_0$, which means that $\neg(\varphi_1 U^{>0} \varphi_2) \in S_1$ and we can apply induction hypothesis.

The next case is $\xi \equiv \varphi_1 U^{=1} \varphi_2$. Again, we need to show that $\mathcal{A}, S \models \xi$. Suppose the converse, i.e., $\mathcal{A}, S \models \neg(\varphi_1 U^{=1} \varphi_2)$. According to Lemma 3.1, there are two possibilities:

- There is a finite path $S = S_0 \mapsto \cdots \mapsto S_n$ such that $\hat{\varphi}_1 \in S_n$ and $\hat{\varphi}_2 \in S_i$ for all $0 \le i \le n$. Similarly as above, we can show (by a straightforward induction on n) that some S_i is not eligible.

- There is a BSCC B of $\mathcal{A} = (A, \mapsto)$ and a finite path $S = S_0 \mapsto \cdots \mapsto S_n$ such that $S_n \in B$ and $\varphi_1, \hat{\varphi}_2 \in T$ for every $T \in A$ which appears in the path or in B. A simple induction reveals that $\varphi_1 U^{=1} \varphi_2 \in S_n$, and in fact $\varphi_1 U^{=1} \varphi_2 \in T$ for every $T \in B$. Hence, the BSCC B was deleted from A at line 8, which is a contradiction.

Finally, let us consider the case when $\xi \equiv \neg(\varphi_1 U^{=1} \varphi_2)$. According to Definition 3.6, we need to verify that $\mathcal{A}, S \models \neg(\varphi_1 U^{=1} \varphi_2)$ or there is a suitable witness for $\neg(\varphi_1 U^{=1} \varphi_2)$. The latter possibility is considered only if the algorithm is supposed to construct a pseudo-model that is not necessarily simple. Let us assume that $\mathcal{A}, S \models \varphi_1 U^{=1} \varphi_2$. But then there cannot be any finite path $S = S_0 \mapsto \cdots \mapsto S_n$ such that $\neg(\varphi_1 U^{>0} \varphi_2) \in S_n$ and $\varphi_1, \hat{\varphi}_2 \in S_i$ for all $0 \le i \le n$, which means that the condition of line 9 is satisfied. If the algorithm is supposed to construct a simple pseudo-model, we obtain a contradiction because S is deleted from A at line 10a. Otherwise, the algorithm proceeds with line 10b, which verifies the existence of a suitable witness for $\neg(\varphi_1 U^{=1} \varphi_2)$. If there was no witness, S would have been deleted from A at line 10b.

Step (b). Let us assume that ψ is (finite) satisfiable. By Theorem 3.7, there is a (simple) pseudo-model $\mathcal{A}' = (A', \rightsquigarrow)$ for ψ. We show that $\mathcal{A}' \subseteq \mathcal{A}$ (taken componentwise) is an invariant of the main **repeat-until** loop at lines 5-12. Here we consider each of the **if** statements individually and show that no element of $\mathcal{A}'$ can be deleted from the current $\mathcal{A}$, assuming that $\mathcal{A}' \subseteq \mathcal{A}$.

If $\xi = \varphi_1 U^{>0} \varphi_2$ and $\mathcal{A}', S \models \xi$, then there is a path $S = S_0 \mapsto \cdots \mapsto S_n$ in $\mathcal{A}'$ with $\varphi_1 \in S_i$ for $0 \le i < n$ and $\varphi_2 \in S_n$. Since $\mathcal{A}' \subseteq \mathcal{A}$ by the induction hypothesis, this path is also a path in $\mathcal{A}$, hence $\mathcal{A}, S \models \xi$. Similarly for ξ of the forms $X^{>0}$ and $\neg X^{=1}$.

For $\xi = \varphi_1 U^{=1} \varphi_2$, let R be a state from a BSCC B, where $\varphi_1, \hat{\varphi}_2, \varphi_1 U^{=1} \varphi_2 \in T$ for every $T \in B$. Suppose for a contradiction that R is a state of $\mathcal{A}'$. Then there is a path

$R = S_0 \mapsto \cdots \mapsto S_n$ in $\mathcal{A}'$ with $\varphi_1 \in S_i$ for $0 \le i < n$ and $\varphi_2 \in S_n$. Since $\mathcal{A}' \subseteq \mathcal{A}$ by the induction hypothesis, this path is also a path in $\mathcal{A}$. However, all paths from R are in B, which is a contradiction.

For $\xi = \neg(\varphi_1 U^{=1} \varphi_2)$ and $\xi \in S$, let S be a vertex of $\mathcal{A}'$. Firstly, let $\mathcal{A}'$ be simple, then $\mathcal{A}', S \not\models \xi$. According to Lemma 3.1, there are two possibilities. Either there is a path $S = S_0 \mapsto \cdots \mapsto S_n$ such that $\varphi_1, \hat{\varphi}_2 \in S_i$ for $0 \le i < n$ and $\hat{\varphi}_1, \hat{\varphi}_2 \in S_n$ whence $\neg(\varphi_1 U^{>0} \varphi_2) \in S_n$. Or there is a BSCC α and a path $S = S_0 \mapsto \cdots \mapsto S_n \in \alpha$ where all states on this path or in α contain φ_1 and $\hat{\varphi}_2$. As no state containing φ_2 is reachable from S_n we have $\neg(\varphi_1 U^{>0} \varphi_2) \in S_n$. In both cases, there is a path specified in step 9 and thus S is not deleted. If $\mathcal{A}'$ is not simple, a third possibility occurs, that there is a witness $\mathcal{B}$ for $\neg(\varphi_1 U^{=1} \varphi_2)$ reachable via a path with its states containing $\hat{\varphi}_2$. Since $\mathcal{A}' \subseteq \mathcal{A}$ this is a witness in $\mathcal{A}'$ reachable along the same path.

Clearly, states with no outgoing edges in $\mathcal{A}$ have no outgoing edges in $\mathcal{A}'$ and hence are not states of $\mathcal{A}'$.

Since A is initialized to the set of all eligible states and $\mapsto$ is initialized to $A \times A$ and the edges deleted in steps 3 and 4 clearly cannot be in $\mathcal{A}'$, the invariant $\mathcal{A}' \subseteq \mathcal{A}$ surely holds before executing the main **repeat-until** loop. Hence, we also have that $\mathcal{A}' \subseteq \mathcal{A}$ after this loop terminates.

Step (c). Since the model-checking problem for qualitative PCTL and finite-state Markov chains is decidable in polynomial time, all steps can be implemented in time which is polynomial in $|\psi|$ and the size of A (i.e., exponential in $|\psi|$). The existence of a suitable witness at line 10b can be decided as follows: suppose that $\neg(\varphi_1 U^{=1} \varphi_2) \in S$. First, initialize B to $\{T \in A \mid \hat{\varphi}_2 \in T\}$, and then do the following:

(A) Compute the strongly connected components $B_1, \ldots, B_n$ of B using the current $\mapsto$, and put $\mathcal{B}_i = (B_i, \hookrightarrow_i)$, where $S \hookrightarrow_i T$ iff $S, T \in B_i$ and $S \mapsto T$.

(B) Compute the set C of all $S \in B_i$, where $1 \le i \le n$, such that for some $\xi_1 U^{=1} \xi_2 \in S$ we have that $\mathcal{B}_i, S \not\models \xi_1 U^{=1} \xi_2$.

(C) Put $B := B \smallsetminus C$. If $C = \emptyset$, terminate. Otherwise, goto (A).

Obviously, the above procedure can be implemented in time which is polynomial in $|\psi|$ and the size of A. We show that every witness for $\neg(\varphi_1 U^{=1} \varphi_2)$ in the current $\mathcal{A} = (A, \mapsto)$ is contained in some SCC of the resulting B, and each of these SCCs itself is a witness for $\neg(\varphi_1 U^{=1} \varphi_2)$.

Firstly, we show that no state S of any witness $\mathcal{B}'$ for $\neg(\varphi_1 U^{=1} \varphi_2)$ in $\mathcal{A}'$ cannot be deleted in step (B). Let S be deleted for a contradiction. Since $\mathcal{A}' \subseteq \mathcal{A}$ there is $1 \leq i \leq n$ such that $\mathcal{B}' \subseteq \mathcal{B}_i$ and there is $\xi_1 U^{=1} \xi_2 \in S$ such that $\mathcal{B}_i, S \not\models \xi_1 U^{=1} \xi_2$. The possibility (a) in Lemma 3.1 cannot occur. Indeed, since we have already performed step 3 and eligibility guarantees that if $\xi_1 U^{=1} \xi_2 \in T$, then $\xi_2 \in T$ or $\xi_1, X^{=1}(\xi_1 U^{=1} \xi_2) \in T$. Hence, a simple induction reveals that there is no path $S = S_0 \hookrightarrow \cdots \hookrightarrow S_n$ such that $\hat{\xi}_2 \in S_i$ for $1 \leq i \leq n$ and $\hat{\xi}_1 \in S_n$. Hence, the possibility (b) takes place, i.e., $\xi_1, \hat{\xi}_2 \in T$ for every state T in $\mathcal{B}_i$ since $\mathcal{B}_i$ is the only reachable BSCC. Therefore, $\xi_1, \hat{\xi}_2 \in T$ for every state T in $\mathcal{B}'$, which means that $\mathcal{B}'$ is not a witness, a contradiction.

Secondly, we show that after termination every SCC $\mathcal{B}_i$ in the resulting B is a witness for $\neg(\varphi_1 U^{=1} \varphi_2)$. Clearly, $\mathcal{B}_i$ is strongly connected and satisfies that for every state S of $\mathcal{B}_i$ and every $\xi_1 U^{=1} \xi_2 \in S$ we have that $\mathcal{B}_i, S \models \xi_1 U^{=1} \xi_2$. Due to the initialization $B \subseteq \{T \in A \mid \hat{\varphi}_2 \in T\}$, and thus all states of $\mathcal{B}_i$ contain $\hat{\varphi}_2$. □

For the sake of completeness, we explicitly verify the corresponding lower complexity bound, although the proof is not very different from the non-probabilistic case.

Theorem 3.12. *The satisfiability problem and the finite-satisfiability problem are* ***EXPTIME****-hard.*

A direct corollary to the previously presented results is the following:

Theorem 3.13. *Let ψ be a qualitative PCTL formula. The problem whether ψ is satisfiable (or finite-satisfiable) is* ***EXPTIME****-complete. Moreover, if ψ is satisfiable (or finite satisfiable), then there is a marked graph (or a finite Markov chain) of size exponential in $|\psi|$ constructible in time exponential in $|\psi|$ which defines a model for ψ.*

4 Some Notes on Quantitative PCTL

In this section, we present some results about the satisfiability problem for quantitative PCTL formulae, which seem confusing at first glance, but which in fact indicate that this problem is more "fragile" and subtle than it looks.

First, we show that the satisfiability problem for quantitative PCTL is highly undecidable for a restricted class of models where the branching degree is bounded by a fixed constant $k \geq 2$. Our proof uses a technique for encoding the computations of non-deterministic Minsky machines, which was developed and used in [3] to show the high undecidability of $1\frac{1}{2}$-player games with PCTL objectives.

Theorem 4.1. *Let ψ be a quantitative PCTL formula, and let $k \geq 2$. The existence of a model for ψ where each state has at most k outgoing transitions is highly undecidable. Moreover, the existence of a* finite *model for ψ where each state has at most k outgoing transitions is undecidable.*

Note that this theorem does not allow to conclude that the satisfiability problem as such is undecidable for quantitative PCTL, because the branching degree of the model cannot be bounded by any fixed constant.

Finally, we present a result which reveals some kind of regularity in PCTL models. First we formulate a general lemma which will be used in the proof of our main result. Intuitively, the lemma says that a countable convex combination of vectors is expressible as a convex combination of a finite number of these vectors. Moreover, if n is the dimension of the vector space, $n+1$ vectors are sufficient (the lemma is a basic result in geometry; a self-contained proof is included in the appendix).

Lemma 4.2. *Let I be a countable set, $v \in \mathbb{R}^n$, and $u_i \in \mathbb{R}^n$ for every $i \in I$. If $v = \sum_{i \in I} a_i u_i$ where $a_i \geq 0$ and $\sum_{i \in I} a_i = 1$, then there is $J \subseteq I$ such that $|J| \leq n+1$, and for each $j \in J$ there is $b_j \geq 0$ such that $\sum_{j \in J} b_j = 1$ and $v = \sum_{j \in J} b_j u_j$.*

Now we show that every satisfiable PCTL formula ψ has a model where each state has at most $|\psi| + 2$ outgoing transitions. This result is non-trivial and uses a combination of geometrical and probabilistic arguments.

Theorem 4.3. *Every satisfiable PCTL formula ψ has a model where each state has at most $|\psi|+2$ outgoing transitions.*

Proof Sketch. Let ψ be a satisfiable formula. This means that there are $M = (St, \to, Prob)$, $s_{in} \in St$ and ν satisfying $M, s_{in} \models^{\nu} \psi$. We may safely assume that M is a (possibly infinite branching) tree rooted in s_{in}, i.e., that every state of M distinct from s_{in} has precisely one incoming transition and s_{in} has no incoming transitions (note that every model of ψ can be "unfolded" into a tree). We construct a model with branching degree bounded by $|\psi| + 2$. The proof is based on choosing suitable successors and assigning them appropriate probabilities, and pruning the others while keeping a model.

Let us denote $\mathcal{S}(\psi)$ the set of all state subformulae of ψ. Further, let

$$I = \mathcal{S}(\psi) \cup \{X\varphi \mid X^{\bowtie\rho}\varphi \in \mathcal{S}(\psi)\} \cup \{\varphi_1 U \varphi_2 \mid \varphi_1 U^{\bowtie\rho} \varphi_2 \in \mathcal{S}(\psi)\}$$

Let us consider vectors of dimension $|I|$ over real numbers with components indexed by elements of I. For every state s of M we define a vector $\vec{s} \in \mathbb{R}^{|I|}$ as follows:

$$\begin{aligned} \vec{s}_{X\varphi} = r &\quad \text{iff } M, s \models X^{=r}\varphi \\ \vec{s}_{\varphi_1 U \varphi_2} = r &\quad \text{iff } M, s \models \varphi_1 U^{=r} \varphi_2 \end{aligned}$$

and for $\varphi \in \mathcal{S}(\psi)$ we put

$$\vec{s}_\varphi = \begin{cases} 1 & \text{if } M, s \models^\nu \varphi; \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to verify that for every state s the vector $\vec{s}$ satisfies the following "local consistency" equations:

$$\vec{s}_{X\varphi} = \sum_{s \xrightarrow{x} t} x \cdot \vec{t}_\varphi$$

and

$$\vec{s}_{\varphi_1 U \varphi_2} = \sum_{s \xrightarrow{x} t} x \cdot \vec{t}_{\varphi_1 U \varphi_2}$$

for $\varphi_1 U \varphi_2 \in I$ such that $M, s \models^\nu \varphi_1$ and $M, s \not\models^\nu \varphi_2$. Note also that components of $\vec{s}$ not occurring in the above equations are determined by other components of $\vec{s}$. Hence, if we are to prune the transition relation of M we should strive to satisfy the above equations.

Let us denote

$$N(s) = \{X\varphi \in I\} \cup \{\varphi_1 U \varphi_2 \in I \mid M, s \models^\nu \varphi_1; M, s \not\models^\nu \varphi_2\}$$

For every t such that $s \rightarrow t$ we define a vector $p(t) \in \mathbb{R}^{|N(s)|}$ (indexed by elements of $N(s)$) as follows:

$$p(t)_{X\varphi} = \vec{t}_\varphi \text{ and } p(t)_{\varphi_1 U \varphi_2} = \vec{t}_{\varphi_1 U \varphi_2}$$

Let s be a state of M. Observe that now we may apply Lemma 4.2 to prune outgoing transitions from s while preserving the model. Indeed, it suffices to apply Lemma 4.2 to $\sum_{s \xrightarrow{x} t} x \cdot p(t)$ and obtain numbers $b_1, \ldots, b_k$ and successors $t_1, \ldots, t_k$ of s such that $\sum_{s \xrightarrow{x} t} x \cdot p(t) = \sum_{i=1}^{k} b_i \cdot p(t_i)$. Consequently, it suffices to modify transitions of M in such a way that $s \xrightarrow{b_i} t_i$. It is easy to verify that the resulting Markov chain is still a model of ψ. However, this pruning cannot be treated out for all states of M. Intuitively, the problem is that fulfilling $\varphi_1 U \varphi_2$ can be deferred indefinitely while still preserving the local consistency (a path formula $\varphi_1 U \varphi_2$ is *fulfilled* on a run w in n steps if φ_2 is satisfied in $w(n)$ and φ_1 is satisfied in $w(i)$ for all $0 \leq i < n$; a formula $\varphi_1 U \varphi_2$ is fulfilled if it is fulfilled in n steps for some n). Therefore, the chain M has to be pruned carefully so that a progress in fulfilling "until" formulae is ensured.

Let $\xi_1, \ldots, \xi_\ell$ be all formulae of I of the form $\varphi_1 U \varphi_2$. Observe that there is $n_1 \geq 0$ such that with a probability $r_s \geq \frac{1}{2}\vec{s}_{\xi_1}$ the formula ξ_1 is fulfilled in at most n_1 steps. Let us assign to every state t reachable from s in $k \leq n_1$ steps the probability r_t of fulfilling ξ_1 in less than $n_1 - k$ steps (i.e., up to the n_1'th level of the subtree rooted in s). First, assume that $\xi_1 \in N(s)$. Note that $r_s = \sum_{s \xrightarrow{x} t} x \cdot r_t$. Let us apply Lemma 4.2 to $\sum_{s \xrightarrow{x} t} x \cdot (p(t), r_t)$ and obtain numbers $b_1, \ldots, b_m$ and successors $t_1, \ldots, t_m$ of s such that $\sum_{s \xrightarrow{x} t} x \cdot (p(t), r_t) = \sum_{i=1}^{m} b_i \cdot (p(t_i), r_{t_i})$. Now we may safely prune the outgoing transitions from s so that $s \xrightarrow{b_i} t_i$. On the other hand, if $\xi_1 \notin N(s)$, we may ignore r_t and apply Lemma 4.2 to $\sum_{s \xrightarrow{x} t} x \cdot p(t)$ in the same way as above. We inductively repeat this pruning for all states reachable from s in at most n_1 steps. Observe that after this pruning the resulting chain is still a model of ψ.

Now let us repeat the above procedure for all states reachable from s in $n_1 + 1$ steps and for the formula ξ_2. We obtain $n_2 \geq n_1$ and a new Markov chain, a model of ψ, which has the property that ξ_1 and ξ_2 are fulfilled with probability at least $\frac{1}{2}\vec{s}_{\xi_1}$ and $\frac{1}{2}\vec{s}_{\xi_2}$, resp., in n_2 steps (starting in s). Note that while taking care of ξ_2 the part of the chain reachable from s in at most n_1 steps remains unaltered. Similarly, we carry out this process for the remaining formulae $\xi_3, \ldots, \xi_\ell$. We obtain a model M_1 of ψ such that for some $m_1 \geq 0$ all states reachable from s in at most m_1 steps have a branching degree bounded by $|\psi| + 2$. Moreover, every $\varphi_1 U \varphi_2 \in I$ is fulfilled in at most m_1 steps with probability at least $\frac{1}{2}\vec{s}_{\varphi_1 U \varphi_2}$.

Repeating the whole construction for states reachable from s in $m_1 + 1$ steps we obtain M_2 and $m_2 \geq m_1$ with similar properties as M_1 and m_1, resp., except that every $\varphi_1 U \varphi_2 \in I$ is fulfilled in at most m_2 steps with probability at least $\frac{3}{4}\vec{s}_{\varphi_1 U \varphi_2}$. Repeating this process ad infinitum we obtain a model M_∞ of ψ such that every state reachable from s has a branching degree bounded by $|\psi| + 2$.

Finally, to obtain a model for ψ with branching degree bounded by $|\psi| + 2$ it suffices to perform the construction of M_∞ for $s = s_{in}$. □

5 Conclusions, Future Work

We solved the satisfiability problem for qualitative PCTL. Although there are some similarities with the logic CTL, the actual properties of these two logics are rather different. Since qualitative PCTL formulae may have only infinite-state models, we also considered the finite satisfiability problem. Since some qualitative PCTL formulae may also require transition probabilities arbitrarily close to zero, another refinement of the satisfiability question might be to consider only models which are possibly infinite-state, but where all probability distributions are uniform. For example, the (satisfiable) formula $\mathrm{G}^{=1}(\mathrm{X}^{>0}\mathrm{p}) \wedge \mathrm{G}^{>0}\neg\mathrm{p}$ does not have this kind of model, while the formula $\mathrm{G}^{>0}(\neg\mathrm{p} \wedge \mathrm{F}^{>0}\mathrm{p})$ (which has only infinite-state models) has a model where the probabilities of all transitions are equal to $\frac{1}{2}$. Another direction for future work is to design a complete deductive system for qualitative PCTL. The decidability of the satisfiability problem for quantitative PCTL remains also open. It seems that proof techniques known to the authors are not sufficient to prove the undecidability. On the other hand, there is some indication that the problem might actually be decidable; of course, all the problems that have successfully been defeated in the qualitative case now rise with a new power.

References

[1] C. Baier, M. Größer, M. Leucker, B. Bollig, and F. Ciesinski. Controller synthesis for probabilistic systems. In *Proceedings of IFIP TCS'2004*, pages 493–506. Kluwer, 2004.

[2] B. Banieqbal and H. Barringer. Temporal logic with fixed points. In *Temporal Logic in Specification*, volume 398 of *Lecture Notes in Computer Science*, pages 62–74. Springer, 1987.

[3] T. Brázdil, V. Brožek, V. Forejt, and A. Kučera. Stochastic games with branching-time winning objectives. In *Proceedings of LICS 2006*, pages 349–358. IEEE Computer Society Press, 2006.

[4] T. Brázdil, V. Forejt, and A. Kučera. Controller synthesis and verification for Markov decision processes with qualitative branching time objectives. In *Proceedings of ICALP 2008*, Lecture Notes in Computer Science. Springer, 2008.

[5] T. Brázdil, A. Kučera, and O. Stražovský. On the decidability of temporal properties of probabilistic pushdown automata. In *Proceedings of STACS 2005*, volume 3404 of *Lecture Notes in Computer Science*, pages 145–157. Springer, 2005.

[6] C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *Journal of the Association for Computing Machinery*, 42(4):857–907, 1995.

[7] E.A. Emerson. Temporal and modal logic. *Handbook of Theoretical Computer Science*, B:995–1072, 1991.

[8] E.A. Emerson and J.Y. Halpern. Decision procedures and expressiveness in the temporal logic of branching time. In *Proceedings of STOC'82*, pages 169–180. ACM Press, 1982.

[9] E.A. Emerson and C.S. Jutla. Tree automata and the logics of programs. *SIAM-JC*, 29(1):132–158, 1986.

[10] J. Esparza, A. Kučera, and R. Mayr. Model-checking probabilistic pushdown automata. *Logical Methods in Computer Science*, 2(1:2):1–31, 2006.

[11] K. Etessami and M. Yannakakis. Recursive Markov chains, stochastic grammars, and monotone systems of non-linear equations. In *Proceedings of STACS 2005*, volume 3404 of *Lecture Notes in Computer Science*, pages 340–352. Springer, 2005.

[12] M. Fischer and R. Ladner. Propositional dynamic logic of regular programs. *Journal of Computer and System Sciences*, 18:194–211, 1979.

[13] H. Hansson and B. Jonsson. A logic for reasoning about time and reliability. *Formal Aspects of Computing*, 6:512–535, 1994.

[14] T. Henzinger, O. Kupferman, and R. Majumdar. On the universal and existential fragments of the μ-calculus. *Theoretical Computer Science*, 354(2):173–186, 2006.

[15] M. Huth and M.Z. Kwiatkowska. Quantitative analysis and model checking. In *Proceedings of LICS'97*, pages 111–122. IEEE Computer Society Press, 1997.

[16] J. Johannsen and M. Lange. CTL+ is complete for double exponential time. In *Proceedings of ICALP 2003*, volume 2719 of *Lecture Notes in Computer Science*, pages 767–775. Springer, 2003.

[17] D. Kozen. A finite-model theorem for the propositional μ-calculus. *Studia Logica*, 47(3):233–241, 1988.

[18] A. Kučera and O. Stražovský. On the controller synthesis for finite-state Markov decision processes. In *Proceedings of FST&TCS 2005*, volume 3821 of *Lecture Notes in Computer Science*, pages 541–552. Springer, 2005.

[19] O. Kupferman and M.Y. Vardi. An automata-theoretic approach to modular model checking. *ACM Transactiona on Programming Languages and Systems*, 22:87–128, 2000.

[20] M. Vardi and L. Stockmeyer. Improved upper and lower bounds for modal logics of programs. In *Proceedings of STOC'85*, pages 240–251. ACM Press, 1985.

A Appendix

Proof of Theorem 3.12

In what follows we show that the satisfiability problem for qualitative PCTL is **EXPTIME**-hard. We give a reduction from the acceptance problem for alternating LBA. Most of the definitions are taken from [3]. An alternating LBA is a tuple $\mathcal{T} = (Q, \mathcal{A}, \Gamma, q_0, \vdash, \dashv, \delta, P)$ where Q is a finite set of control states, $\mathcal{A}$ is a finite input alphabet, $\Gamma \supseteq \mathcal{A}$ is a finite tape alphabet, $q_0 \in Q$ is the initial control state, $\vdash, \dashv \in \Gamma$ are the left-end and the right-end markers, $\delta : Q \times \Gamma \to 2^{Q\times\Gamma\times\{L,R\}}$ is a transition function, and $P = (Q_\forall, Q_\exists, Q_{acc}, Q_{rej})$ is a partition of control states into universal, existential, accepting and rejecting states.

W.l.o.g. we assume that $Q \cap \Gamma = \emptyset$ and that $\delta(q, A)$ has exactly two elements (q_1, A_1, D_1), (q_2, A_2, D_2) where $q_1 \neq q_2$ for every $q \in Q$ and $A \in \Gamma$. A computational tree for $\mathcal{T}$ on a word $u \in \mathcal{A}^*$ is a tree T satisfying the following: the root of T is (labeled by) the initial configuration for u, and if N is a node of T labeled by a configuration with a control state q, then the following holds:

- if q is accepting or rejecting, then N is a leaf;
- if q is existential, then N has one successor labeled by a configuration reachable from the configuration of N in one step;
- if q is universal, then the successors of N are the two configurations reachable from the configuration of N in one step.

$\mathcal{T}$ accepts u iff there is a finite computational tree T for $\mathcal{T}$ on u such that all leafs of T are accepting. We can safely assume that all computational trees for $\mathcal{T}$ are finite. Configurations of $\mathcal{T}$ are written as words over the alphabet $\Xi = Q \cup \Gamma$ in the standard way; for example, the initial configuration for u is written as $q_0{\vdash}u{\dashv}$. Depending on the control state of a configuration, we call a configuration universal, existential, accepting or rejecting. Another standard result is that one can efficiently compute the set $\mathrm{Comp}(\mathcal{T}) \subseteq \Xi^6$ of all compatible 6-tuples such that for each configuration c (written as a word over Ξ)

we have that $c' \in \Xi^*$ is a one-step successor of c iff c' has the same length as c and for all $1 \leq i \leq |c| - 2$ we have that $(c(i), c(i+1), c(i+2), c'(i), c'(i+1), c'(i+2)) \in \mathrm{Comp}(\mathcal{T})$.

Let $\mathcal{T} = (Q, \mathcal{A}, \Gamma, q_0, \vdash, \dashv, \delta, P)$ be an alternating LBA and $u \in \mathcal{A}^*$ an input word of length z. We construct (in polynomial time) a qualitative PCTL formula φ such that φ is satisfiable iff $\mathcal{T}$ accepts u. The formula φ will be chosen so that each model must encode a computational tree in the way that will be explained later.

In the following, we denote r the length of a configuration of $\mathcal{T}$, i.e. $r = z + 3$. We also use $(X^{=1})^n$ to denote string of n iterations of $X^{=1}$. For example, $(X^{=1})^2 p$ stands for $X^{=1}X^{=1}p$. Very same notation is also used for other modal operators.

Let us fix fresh atomic propositions $\widehat{q}$ and $\overline{q}$ for each $q \in Q$, and an atomic proposition t for each $t \in \Xi$. Also, let us fix a distinguished proposition start. The formula φ is of the following form:

$$\varphi \equiv \mathit{Struct} \wedge \mathit{Init} \wedge \mathit{CState} \wedge \mathit{NState} \wedge \mathit{Compat} \wedge \mathit{UnivStates} \wedge \mathit{Acc}$$

The formula Struct encodes the syntactical structure of computational trees.

$$\mathit{Struct} \equiv \mathit{Props} \wedge \mathit{start} \wedge G^{=1}\big(\mathit{start} \Rightarrow (\mathit{Config} \wedge \mathit{States} \wedge \mathit{Delim})\big)$$

Here, the formula Props says that exactly one of atomic propositions occurring in φ holds in each state (it is straightforward to encode it in PCTL). The formula Delim says that the atomic proposition start occurs on every path after each $r + 3$ steps. Formally, $\mathit{Delim} \equiv \big((X^{=1})^{r+3}\mathit{start}\big)$. Formulae Config and States enforce a certain atomic propositions to hold between each two states in which start holds. Formally, Config and States are defined as follows:

$$\begin{array}{rcl} \mathit{Config} & \equiv & \big(\bigvee_{t\in\Xi}(X^{=1})^1 t\big) \wedge \ldots \wedge \big(\bigvee_{t\in\Xi}(X^{=1})^r t\big) \\ \mathit{States} & \equiv & \big(\bigvee_{q\in Q}(X^{=1})^{r+1}\widehat{q}\big) \wedge \big((X^{=1})^{r+2}\bigvee_{q\in Q}\overline{q}\big) \end{array}$$

To satisfy Struct, each run initiated in the initial state must be composed of sequences (we call them *chunks*) $s_0 s_1 \ldots s_{r+2}$ of states of the following form:

- $s_0 \models \mathit{start}$;
- $s_i \models t_i$ for all $1 \leq i \leq r$ where each t_i is from Ξ;
- $s_{r+1} \models \widehat{q}$ for some $q \in Q$ and
- $s_{r+2} \models \overline{q'}$ for some $q' \in Q$.

Intuitively, we intend each chunk to "encode" a configuration of $\mathcal{T}$, with redundant information about a current control state at $(r+2)$-th position, and information about a control state that will be used in the next chunk at $(r+3)$-th position.

The formula Init enforces the first chunk of every run to encode the initial configuration of $\mathcal{T}$. Formally, it is defined by

$$\mathit{Init} \equiv \big((X^{=1})^1 c(0)\big) \wedge \ldots \wedge \big((X^{=1})^r c(r-1)\big)$$

where $c = q_0 \vdash u \dashv$ is the initial configuration.

The formula CState ensures that a state that occurs at $(r+2)$-th position in a chunk is the same as the one that occurs in the configuration encoded by the chunk.

$$\mathit{CState} \equiv \bigwedge_{q \in Q} G^{=1}\big(q \Rightarrow (\neg \mathit{start} U^{=1} \widehat{q})\big)$$

The formula NState ensures that a state that occurs at $(r+3)$-th position in a chunk is equal to the one that occurs in the configuration encoded by the adjacent chunk.

$$\mathit{NState} \equiv \bigwedge_{q \in Q} G^{=1}\big(\overline{q} \Rightarrow (X^{=1})^2(\neg \mathit{start} U^{=1} q)\big)$$

The formula Compat ensures that if there are two adjacent chunks encoding configurations c and c', then c' is a one step successor of c.

$$\mathit{Compat} \equiv G^{=1} \bigwedge_{t_1 t_2 t_3 t_1' t_2' t_3' \in \Xi^6 \setminus \mathrm{Comp}(\mathcal{T})} \neg\big(\mathit{Current} \wedge (X^{>0})^{r+3} \mathit{Next}\big)$$

Here, the formula $\mathit{Current}$ is equal to $t_1 \wedge X^{=1} t_2 \wedge (X^{=1})^2 t_3$, the formula Next is equal to $t_1' \wedge X^{=1} t_2' \wedge (X^{=1})^2 t_3'$.

Formula $\mathit{UnivStates}$ says that whenever a universal control state occurs in the chunk, then a model branches so that both successive configurations are explored.

$$\mathit{UnivStates} \equiv \bigwedge_{q \in Q_\forall} \big(\widehat{q} \Rightarrow \bigvee_{p, p' \in Q, p \neq p'} (X^{>0} \overline{p} \wedge X^{>0} \overline{p'})\big)$$

Finally, the formula Acc ensures that an accepting state will be reached almost surely (note that because all computational trees are finite, this ensures that accepting state will be reached *surely*).

$$\mathit{Acc} \equiv F^{=1} \bigvee_{q \in Q_{acc}} \bar{q}$$

It remains to argue that the formula φ is satisfiable iff $\mathcal{T}$ accepts u. First, suppose that φ is satisfiable. Let M be a model of φ, let ν be a valuation and let s be a state of M such that $M, s \models^{\nu} \varphi$. Let w be a path of length $r+1$ initiated in s. We label root of T with a configuration c encoded by $w(1) \dots w(r)$ (i.e., the only configuration c satisfying $w(i+1) \models c(i)$ for all $1 \leq i \leq r$). Now let N be a node of T labeled with a configuration encoded by some path $s_1 \dots s_r$. We construct successors of N as follows.

- if N is labeled by an accepting configuration, then N is a leaf;
- if N is labeled by an existential configuration, then N has exactly one successor which is labeled by a configuration encoded by $v_1 \dots v_r$, where $s_r s_{r+1} s_{r+2} s_{r+3} v_1 \dots v_r$ is a path in M
- if N is labeled by a universal configuration, then N has exactly two successors which are labeled by a configurations encoded by $v_1 \dots v_r$ and $v'_1, \dots v'_r$ satisfying the following. There are states $q_1, q_2 \in Q$ (where $q_1 \neq q_2$) and paths $s_r s_{r+1} s_{r+2} s_{r+3} v_1 \dots v_r$ and $s_r s_{r+1} s'_{r+2} s'_{r+3} v'_1 \dots v'_r$ in M such that $s_{r+2} \models \overline{q_1}$ and $s'_{r+2} \models \overline{q_2}$. Observe that the existence of such sequences is ensured by the formula φ.

Note that there may be several possibilities how to choose sequences in the second and third step of the procedure. Depending on this choice, different computational trees may be produced. However, each of these trees is accepting.

The construction of a model for φ from an accepting computational tree is similar. The graph of the model is the computational tree, where each configuration c is replaced by the following graph:

- a path $sc_1 \cdots c_r tu$ where $s \models \mathit{start}, c_i \models c(i), t \models \hat{q}, u \models \overline{p}$ where q is an existential state in c and p is the state in the next configuration;
- a path $sc_1 \cdots c_r t$ with t having two successors u_1 and u_2 where $s \models \mathit{start}, c_i \models c(i), t \models \hat{q}, u_1 \models \overline{p_1}, u_2 \models \overline{p_2}$ where q is a universal state in c and p_i are the states in the next configurations and the transitions to the next configurations lead from their respective u_i's.
- a path $sc_1 \cdots c_r tu$ where $s \models \mathit{start}, c_i \models c(i), t \models \hat{q}, u \models \overline{q}$ where q is an accepting state in c and we add a transition $u \to s$.

All probability distributions are uniform. □

Proof of Theorem 4.1 (sketch)

We show how to modify the proof of high undecidability of the strategy synthesis problem for $1\frac{1}{2}$-games with PCTL objectives to prove high undecidability of the satisfiability of PCTL formulae.

A $1\frac{1}{2}$-player game is a tuple $G = (V, E, (V_\Box, V_\bigcirc), \mathit{Prob})$ where V is a finite set of *vertices*, $E \subseteq V \times V$ is a total *transition relation*, $V_\Box$ and $V_\bigcirc$ partition the set of vertices V into *non-deterministic* and *stochastic* vertices, respectively, and *Prob* assigns to each probabilistic vertex a positive probability distribution over the set of its outgoing transitions. A *strategy* is a function σ which to every $wv \in V^*V_\Box$ assigns a probability distribution over the set of outgoing transitions of v. A game G together with a strategy σ determine a Markov chain $G_\sigma = (V^+, \hookrightarrow, \mathit{Prob}')$, where $wv \hookrightarrow wvv'$ for all $w \in V^*$, $v, v' \in V$ such that $v \rightarrow v'$, and $\mathit{Prob}'(wv)$ is either $\mathit{Prob}(v)$ or $\sigma(wv)$, depending on whether v is stochastic or non-deterministic, respectively. Every valuation $\nu : V \rightarrow 2^{Ap}$ uniquely determines a valuation $\bar{\nu} : V^+ \rightarrow 2^{Ap}$ by $\bar{\nu}(wv) = \nu(v)$.

Let G be a game, v_{in} a vertex of G, and ν a valuation. It is shown in [3] that the problem whether there is a strategy σ such that $G_\sigma, v_{in} \models^{\bar{\nu}} \varphi$ is highly undecidable. The proof proceeds by reduction from the problem whether an initial instruction of a non-deterministic Minsky machine is executed infinitely many times. More concretely, given a Minsky machine, we construct a game G, a vertex v_{in}, a valuation ν, and a formula φ such that the Minsky machine executes the initial configuration infinitely many times iff there is a strategy σ satisfying $G_\sigma, v_{in} \models^{\bar{\nu}} \varphi$. The crucial observation is that the construction can be modified so that G, φ, and ν satisfy the following conditions:

1. every stochastic vertex has exactly k successors;

2. every non-deterministic vertex has $k+1$ successors and every strategy σ satisfying $G_\sigma, v_{in} \models^{\bar{\nu}} \varphi$ assigns non-zero probabilities precisely to k successors of every non-deterministic state.

Essentially, it suffices to add, for each stochastic (or non-deterministic) vertex v with n successors, a k-clique of stochastic vertices together with $k - n$ transitions (or $k - n + 1$ transitions, respectively) from v to different vertices of this clique. Furthermore, we modify the construction of the formula φ to enforce the condition 2.

Now we show how to encode G and φ to a single formula ψ such that ψ has a model iff there is σ satisfying $G_\sigma, v_{in} \models^{\bar{\nu}} \varphi$. For each vertex $v \in V$ we introduce a fresh atomic proposition p_v with the intended meaning of being in v. Given a vertex $v \in V_\Box$ we denote $Succ_k(v) = \{A \subseteq V \mid |A| = k, \forall u \in A : (v,u) \in E\}$ the set of all k-element sets of successors of v. In order to describe the game graph transitions and the valuation, we define a formula φ_v for each vertex $v \in V$ as follows (here At is the set of atomic propositions occurring in φ): For $v \in V_\Box$ we put

$$\varphi_v \equiv p_v \Rightarrow \Big(\bigvee_{A \in Succ_k(v)} \big(\bigwedge_{u \in A} X^{>0}(p_u) \wedge \bigwedge_{u \notin A} \neg X^{>0}(p_u)\big) \quad \wedge \bigwedge_{a \in \nu(v) \cap At} a \wedge \bigwedge_{a \in At \setminus \nu(v)} \neg a\Big)$$

For $v \in V_\bigcirc$ we put

$$\varphi_v \equiv p_v \Rightarrow \Big(\bigwedge_{(v,u) \in E} X^{=Prob(v \rightarrow u)}(p_u) \qquad \wedge \bigwedge_{a \in \nu(v) \cap At} a \wedge \bigwedge_{a \in At \setminus \nu(v)} \neg a\Big)$$

We set $\psi = p_{v_{in}} \wedge G^{=1}(\bigwedge_{v \in V} \varphi_v) \wedge G^{=1}(\bigvee_{v \in V}(p_v \wedge \bigwedge_{u \neq v} \neg p_u)) \wedge \varphi$ to state that we start in v, the transitions in a model have always to respect the game graph, the original valuation is respected, and φ is satisfied.

Assume that $G_\sigma, v_{in} \models^{\bar{\nu}} \varphi$ for a strategy σ. We may safely assume that for every $v \in V$ we have that $p_v \in \nu(v)$ and that $p_u \notin \nu(v)$ for $u \neq v$. It is easy to verify that G_σ is k-branching (due to 2.) and $G_\sigma, v_{in} \models^{\bar{\nu}} \psi$.

On the other hand, assume that $M, s_{in} \models^\mu \psi$ where $M = (S, \hookrightarrow, Prob')$ is k-branching. Note that every state of M satisfies precisely one of the propositions of the form p_v. Let us define a strategy σ as follows: Let $v_0 \cdots v_n \in V^* V_\Box$. If there is a (unique) path $s_0, \cdots, s_n$ in M such that $s_0 = s_{in}$ and $p_{v_i} \in \mu(s_i)$ for every $i \geq 0$, then for every u satisfying $p_u \in \mu(s)$ where $s_n \xrightarrow{x} s$ we define $\sigma(v_0 \cdots v_n)(v_n, u) = x$. Otherwise, we define $\sigma(v_0 \cdots v_n)$ arbitrarily. By the construction of ψ and the fact that M is k-branching, the strategy σ is well defined. In fact, G_σ corresponds to an "unfolding" of the model M. Hence $G_\sigma, v_{in} \models^{\bar{\nu}} \varphi$. □

Proof of Lemma 4.2

Let us at first assume that $I = \mathbb{N}$. A convex combination of vectors $v_1, \ldots, v_m \in \mathbb{R}^n$ is a sum $\sum_{i=1}^m a_i v_i$ where $a_i \geq 0$ and $\sum_{i=1}^m a_i = 1$. A countable convex combination of $v_1, v_2, \ldots \in \mathbb{R}^n$ is a sum $\sum_{i=1}^\infty a_i v_i$ where $a_i \geq 0$ and $\sum_{i=1}^\infty a_i = 1$. Given a set $A \subseteq \mathbb{R}^n$ of vectors we denote $\mathcal{CH}(A)$ and $\mathcal{CCH}(A)$ the sets of all convex combinations and countable convex combinations, respectively, of vectors of A. The core is to prove that

$$\mathcal{CCH}(\{u_i \mid i \in \mathbb{N}\}) \subseteq \bigcup_{j=1}^{\infty} \mathcal{CH}(\{u_i \mid i \leq j\})$$

Then v is an element of a polyhedron $\mathcal{CH}(\{u_i \mid i \leq j\})$ with vertices from the set $\{u_i \mid i \leq j\}$. The same holds when I is finite. The polyhedron can be triangulated, and

v lies in at least one of the n-simplices, hence it is a combination of $n+1$ points of $\{u_i \mid i \leq j\}$.

Firstly, let $v \in \mathcal{CCH}(\{u_i \mid i \in \mathbb{N}\})$ be an interior point. As v has a neighborhood in this interior, there are $n+1$ points $y_k \neq v$ from this neighborhood such that $\mathcal{CH}(\{y_k \mid k \leq n+1\})$ contains a neighborhood of v. We can take e.g. vertices of a regular n-simplex barycentered in v. Consider a function $f : (\mathbb{R}^n)^{n+1} \to \mathbb{R}$ defined as follows. $f(\delta_1, \dots, \delta_{n+1})$ is the radius of a maximal ball centered in v that is contained in $\mathcal{CH}(\{y_k + \delta_k \mid k \leq n+1\})$ or 0 if there is none. As $f(0, \dots, 0) > 0$ and f is continuous, there is $\varepsilon > 0$ such that narrowing of f to arguments lesser than ε is positive. In other words, there is $\varepsilon > 0$ such that if we take any point z_k from each ε-neighborhood of y_k, then $\mathcal{CH}(\{z_k \mid k \leq n+1\})$ always contains v.

Every y_k is an infinite sum of points. Consider the sequences of partial sums such that the weights are normalized so that the sum of weights is always 1. Since these sequences converge to y_k's, there is $m \in \mathbb{N}$ such that the m-th normalized partial sums lie in the respective ε-neighborhoods. Hence we receive the desired z_k's, each of which is a convex combination of m points from $\{u_i \mid i \in \mathbb{N}\}$. Altogether, v is a combination of at most $m \cdot (n+1)$ points.

Secondly, let $v \in \mathcal{CCH}(\{u_i \mid i \in \mathbb{N}\})$ be a boundary point. We consider a supporting hyperplane, i.e., a hyperplane containing v and determining an open half space not containing any point of the $\mathcal{CCH}(\{u_i \mid i \in \mathbb{N}\})$. It exists for $\mathcal{CCH}(\{u_i \mid i \in \mathbb{N}\})$ is clearly convex, and thus we can apply the supporting hyperplane theorem on its closure. The intersection of the hyperplane and $\mathcal{CCH}(\{u_i \mid i \in \mathbb{N}\})$ is $\mathcal{CCH}(\{u_i \mid i \in \mathbb{N},\ u_i \text{ lies in the hyperplane}\})$. Indeed, if the sum expressing a point p in the intersection included $a_k u_k$ where the distance between hyperplane and u_k is $\delta > 0$, then the distance between p and the hyperplane would be at least $a_k \delta$, since there are no points in the other half space, a contradiction.

By induction, either (i) v is an interior point of $\mathcal{CCH}(\{u_i \mid i \in K\})$ for some $K \subseteq \mathbb{N}$, a "face" of the original $\mathcal{CCH}(\{u_i \mid i \in \mathbb{N}\})$, of dimension $1 \leq \dim < n$, and we conclude by the above argument, or (ii) v is an element of a zero dimensional countable convex hull, ergo is one of the u_i's, whence the assertion. □