

# Algorithmic Verification of Recursive Probabilistic State Machines



Etessami, K., & Yannakakis, M. (2005). Algorithmic Verification of Recursive Probabilistic State Machines. In TACAS. (pp. 253-270). Springer Berlin / Heidelberg.

# Algorithmic verification of recursive probabilistic state machines

Kousha Etessami[1] and Mihalis Yannakakis[2]

[1] School of Informatics, University of Edinburgh
[2] Department of Computer Science, Columbia University

**Abstract.** Recursive Markov Chains (RMCs) ([EY04]) are a natural abstract model of procedural probabilistic programs and related systems involving recursion and probability. They succinctly define a class of denumerable Markov chains that generalize multi-type branching (stochastic) processes. In this paper, we study the problem of model checking an RMC against a given $\omega$-regular specification. Namely, given an RMC $A$ and a Büchi automaton $B$, we wish to know the probability that an execution of $A$ is accepted by $B$. We establish a number of strong upper bounds, as well as lower bounds, both for *qualitative* problems (is the probability $= 1$, or $= 0$?), and for *quantitative* problems (is the probability $\geq p$?, or, approximate the probability to within a desired precision). Among these, we show that qualitative model checking for general RMCs can be decided in PSPACE in $|A|$ and EXPTIME in $|B|$, and when $A$ is either a single-exit RMC or when the total number of entries and exits in $A$ is bounded, it can be decided in polynomial time in $|A|$. We then show that quantitative model checking can also be done in PSPACE in $|A|$, and in EXPSPACE in $|B|$. When $B$ is deterministic, all our complexities in $|B|$ come down by one exponential.
For lower bounds, we show that the qualitative model checking problem, even for a fixed RMC, is already EXPTIME-complete. On the other hand, even for simple reachability analysis, we showed in [EY04] that our PSPACE upper bounds in $A$ can not be improved upon without a breakthrough on a well-known open problem in the complexity of numerical computation.

## 1 Introduction

Recursive Markov Chains (RMCs) are a natural abstract model of procedural probabilistic programs. They succinctly define a natural class of denumerable markov chains that generalize multi-type branching (stochastic) processes. Informally, an RMC consists of a collection of finite state component Markov chains (MC) that can call each other in a potentially recursive manner. Each component MC has a set of *nodes* (ordinary states), a set of *boxes* (each mapped to a component MC), a well-defined interface consisting of a set of *entry* and *exit* nodes (nodes where it may start and terminate), and a set of probabilistic transitions connecting the nodes and boxes. A transition to a box specifies the entry node and models the invocation of the component MC associated with the box; when (and if) the component MC terminates at an exit, execution of the calling MC resumes from the corresponding exit of the box.

RMCs are a probabilistic version of Recursive State Machines (RSMs) ([AEY01,BGR01]). RSMs and closely related models like Pushdown Systems (PDSs) have been studied extensively in recent research on model checking and program analysis, because of their applications to verification of sequential programs with procedures (see, e.g., [EHRS00,BR00]). Recursive Markov Chains generalize other well-studied models involving probability and recursion: *Stochastic Context-Free Grammars* (SCFGs) have been extensively studied, mainly

in natural language processing (NLP) (see [MS99]). *Multi-Type Branching Processes* (MT-BPs), are an important family of stochastic processes with many applications in a variety of areas (see, e.g., [Har63]). Both SCFG's and MT-BP's are essentially equivalent to *single-exit* RMC's: the special case of RMC's in which all components have one exit. Probabilistic models of programs and systems are of interest for several reasons. First, a program may use randomization, in which case the transition probabilities reflect the random choices of the algorithm. Second, we may want to model and analyse a program or system under statistical conditions on its behaviour (e.g., based on profiling statistics or on statistical assumptions), and to determine the induced probability of properties of interest.

We introduced RMCs in ([EY04]), where we developed some of their basic theory and focused on algorithmic reachability analysis: what is the probability of reaching a given state starting from another? In this paper, we study the more general problem of model checking an RMC against an $\omega$-regular specification: given an RMC $A$ and a Büchi automaton $B$, we wish to know the probability that an execution of $A$ is accepted by $B$. The techniques we develop in this paper for model checking go far beyond what was developed in [EY04] for reachability analysis.

General RMCs are intimately related to probabilistic Pushdown Systems (pPDSs), and there are efficient translations between RMCs and pPDSs. There has been some recent work on model checking of pPDSs ([EKM04,BKS04]). As we shall describe shortly, our results yield substantial improvements, when translated to the setting of pPDSs, on the best algorithmic upper and lower bounds known for $\omega$-regular model checking of pPDSs.

We now outline the main results in this paper. We are given an RMC $A$ and a property in the form of a (non-deterministic) Büchi automaton (BA) $B$, whose alphabet corresponds to (labels on) the vertices of $A$. Let $P_A(L(B))$ denote the probability that an execution of $A$ is accepted by $B$ (i.e., satisfies the property). The *qualitative* model checking problems are: (1) determine whether almost all executions of $A$ satisfy the property $B$ (i.e. is $P_A(L(B)) = 1$?, this corresponds to $B$ being a desirable correctness property), and (2) whether almost no executions of $A$ satisfy $B$ (i.e. is $P_A(L(B)) = 0$?, corresponding to $B$ being an undesirable error property). In the *quantitative* model checking problems we wish to compare $P_A(L(B))$ to a given rational threshold $p$, i.e., is $P_A(L(B)) \geq p$?, or alternatively, we may wish to approximate $P_A(L(B))$ to within a given number of bits of precision. Note that in general $P_A(L(B))$ may be irrational or may not even be expressible by radicals [EY04]. Hence it cannot be computed exactly.

We show that the qualitative model checking problems can be solved in PSPACE in $|A|$ and EXPTIME in $|B|$. More specifically, in a first phase the algorithm analyzes the RMC $A$ by itself (using PSPACE). In a second phase it analyses further $A$ in conjunction with $B$, using polynomial time in $A$ and exponential time in $B$. If the automaton $B$ is deterministic then the time is polynomial in $B$. Furthermore, if $A$ is a single-exit RMC (which corresponds to SCFG's and MT-BP's), then the first phase, and hence the whole algorithm, can be done in polynomial time in $A$. Another such case that we can model-check qualitatively in polynomial time in $A$ is when the total number of entries and exits in $A$ is bounded (we call them Bd-RMCs). In terms of probabilistic program abstractions, this class of RMC's corresponds to programs with a bounded number of different procedures, each of which has a bounded number of input/output parameter values. The internals of the components of the RMCs (i.e., the procedures) can be arbitrarily large and complex.

For quantitative model checking, we show that deciding whether $P_A(L(B)) \geq p$ for a given rational $p \in [0, 1]$ can be decided in PSPACE in $|A|$, and in EXPSPACE in $|B|$. When

Qualitative:

| | reachability | det. Büchi | nondet. Büchi |
|---|---|---|---|
| 1-exit | P | P | P in RMC, EXPTIME in Büchi |
| Bd | P | P | P in RMC, EXPTIME in Büchi |
| general | PSPACE | PSPACE | PSPACE in RMC, EXPTIME in Büchi |

Quantitative:

| | reachability | det. Büchi | nondet. Büchi |
|---|---|---|---|
| 1-exit | PSPACE | PSPACE | PSPACE in RMC, EXPSPACE in Büchi |
| Bd | P | P in RMC for fixed Büchi | P in RMC, for fixed Büchi |
| general | PSPACE | PSPACE | PSPACE in RMC, EXPSPACE in Büchi |

**Fig. 1.** Complexity of Qualitative and Quantitative problems

$B$ is deterministic, the space is polynomial in both $A$ and $B$. Moreover, for $A$ a Bd-RMC, and when $B$ is fixed, there is an algorithm that runs in P-time in $|A|$; however, in this case (unlike the others) the exponent of the polynomial depends on $B$. Table 1 summarizes our complexity upper bounds.

For lower bounds, we prove that the qualitative model checking problem, even for a fixed, single entry/exit RMC, is already EXPTIME-complete. On the other hand, even for reachability analysis, we showed in [EY04] that our PSPACE upper bounds in $A$, even for the quantitative 1-exit problem, and the general qualitative problem, can not be improved without a breakthrough on the complexity of the *square root sum* problem, a well-known open problem in the complexity of numerical computation (see Section 2.2).
Due to space limitations, we have removed almost all proofs from this paper.

**Related Work** Model checking of flat Markov chains has received extensive attention both in theory and practice (eg. [CY95,Kwi03,PZ93,Var85]). It is known that model checking of a Markov chain $A$ with respect to a Büchi automaton $B$ is PSPACE-complete, and furthermore the probability $P_A(L(B))$ can be computed exactly in time polynomial in $A$ and exponential in $B$. Recursive Markov chains were introduced recently in [EY04], where we developed some of their basic theory and investigated the termination and reachability problems; we summarize the main results in Section 2.2. Recursion introduces a number of new difficulties that are not present in the flat case. For example, in the flat case, the qualitative problems depend only on the structure of the Markov chain (which transitions are present) and not on the precise values of the transition probabilities; this is not any more the case for RMC's and numerical issues have to be dealt with even in the qualitative problem. Furthermore, unlike the flat case, the desired probabilities cannot be computed exactly.

The closely related model of probabilistic Pushdown Systems (pPDS) was introduced and studied recently in [EKM04,BKS04]. They largely focus on model checking against branching-time properties, but they also study deterministic ([EKM04]) and non-deterministic ([BKS04]) Büchi automaton specifications. There are efficient (linear time) translations between RMCs and pPDSs, similar to translations between RSMs and PDSs (see [AEY01,BGR01]). Our upper bounds, translated to pPDSs, improve those obtained in [EKM04,BKS04] by an exponential factor in the general setting, and by more for specific classes like single-exit and Bd-RMCs. Specifically, [BKS04], by extending results in [EKM04], show that qualitative model checking for pPDSs can be done in PSPACE in the size of the pPDS and 2-EXPSPACE in the size of the Büchi automaton, while quantitative model checking can be decided in EXPTIME in the size of the pPDS and in 3-EXPTIME in the size of the Büchi automaton. They do not

obtain stronger complexity results for the class of pBPAs (equivalent to single-exit RMCs). Also, the class of Bd-RMCs has no direct analog in pPDSs, as the total number of entries and exits of an RMC gets lost in translation to pPDSs.

The rest of this paper is organized as follows. In Section 2 we give the necessary definitions and background on RMC's from [EY04]. In Section 3 we show how to construct from an RMC $A$ a flat Markov chain $M'_A$ which in some sense summarizes the recursion in the trajectories of $A$; this chain plays a central role analogous to that of the "summary graph" for Recursive State machines [AEY01,BGR01]. In Section 4 we address the qualitative model checking problems, presenting both upper and lower bounds. Section 5 addresses the quantitative model checking problem; a fundamental "unique fixed point theorem" is proved for RMC's, and is used to develop our quantitative algorithms.

## 2 Definitions and Background

A *Recursive Markov Chain (RMC)*, $A$, is a tuple $A = (A_1, \ldots, A_k)$, where each *component chain* $A_i = (N_i, B_i, Y_i, En_i, Ex_i, \delta_i)$ consists of:

- A set $N_i$ of *nodes*
- A subset of *entry* nodes $En_i \subseteq N_i$, and a subset of *exit* nodes $Ex_i \subseteq N_i$.
- A set $B_i$ of *boxes*.
  Let $B = \cup_{i=1}^{k} B_i$ be the (disjoint) union of all boxes of $A$.
- A mapping $Y_i : B_i \mapsto \{1, \ldots, k\}$ assigns a component to every box.
  Let $Y = \cup_{i=1}^{k} Y_i$ be $Y : B \mapsto \{1, \ldots, k\}$ where $Y|_{B_i} = Y_i$, for $1 \leq i \leq k$.
- To each box $b \in B_i$, we associate a set of *call ports*, $Call_b = \{(b, en) \mid en \in En_{Y(b)}\}$, and a set of *return ports*, $Return_b = \{(b, ex) \mid ex \in Ex_{Y(b)}\}$.
- A transition relation $\delta_i$, where transitions are of the form $(u, p_{u,v}, v)$ where:
  1. the source $u$ is either a non-exit node $u \in N_i \setminus Ex_i$, or a return port $u = (b, ex) \in Return_b$, where $b \in B_i$.
  2. The destination $v$ is either a non-entry node $v \in N_i \setminus En_i$, or a call port $v = (b, en) \in Call_b$, where $b \in B_i$.
  3. $p_{u,v} \in \mathbb{R}_{>0}$ is the probability of transition from $u$ to $v$. (We assume $p_{u,v}$ is rational.)
  4. *Consistency of probabilities*: for each $u$, $\sum_{\{v' \mid (u, p_{u,v'}, v') \in \delta_i\}} p_{u,v'} = 1$, unless $u$ is a call port or exit node; neither have outgoing transitions, in which case $\sum_{v'} p_{u,v'} = 0$.

We will use the term *vertex* of $A_i$ to refer collectively to its set of nodes, call ports, and return ports, and we denote this set by $Q_i$, and we let $Q = \bigcup_{i=1}^{k} Q_i$ be the set of all vertices of the RMC $A$. That is, the transition relation $\delta_i$ is a set of probability-weighted directed edges on the set $Q_i$ of vertices of $A_i$. Let $\delta = \cup_i \delta_i$ be the set of all transitions of $A$.

An RMC $A$ defines a global denumerable markov chain $M_A = (V, \Delta)$ as follows. The global *states* $V \subseteq B^* \times Q$ are pairs of the form $\langle \beta, u \rangle$, where $\beta \in B^*$ is a (possibly empty) sequence of boxes and $u \in Q$ is a *vertex* of $A$. More precisely, the states $V \subseteq B^* \times Q$ and transitions $\Delta$ are defined inductively as follows:

1. $\langle \epsilon, u \rangle \in V$, for $u \in Q$. ($\epsilon$ denotes the empty string.)
2. if $\langle \beta, u \rangle \in V$ and $(u, p_{u,v}, v) \in \delta$, then $\langle \beta, v \rangle \in V$ and $(\langle \beta, u \rangle, p_{u,v}, \langle \beta, v \rangle) \in \Delta$.
3. if $\langle \beta, (b, en) \rangle \in V$ and $(b, en) \in Call_b$, then $\langle \beta b, en \rangle \in V$, & $(\langle \beta, (b, en) \rangle, 1, \langle \beta b, en \rangle) \in \Delta$.
4. if $\langle \beta b, ex \rangle \in V$ and $(b, ex) \in Return_b$, then $\langle \beta, (b, ex) \rangle \in V$ & $(\langle \beta b, ex \rangle, 1, \langle \beta, (b, ex) \rangle) \in \Delta$.
5. nothing else is in $V$ and $\Delta$.

Item 1 corresponds to the possible initial states, item 2 corresponds to a transition within a component, item 3 is when a new component is entered via a box, item 4 is when the process exits a component and control returns to the calling component.

Some states of $M_A$ are *terminating*, having no outgoing transitions. These are precisely the states $\langle \epsilon, ex \rangle$, where $ex$ is an exit. We want to view $M_A$ as a proper markov chain, so we consider terminating states to be *absorbing* states, with a self-loop of probability 1.

A *trace* (or *trajectory*) $t \in V^\omega$ of $M_A$ is an infinite sequence of states $t = s_0 s_1 s_2 \ldots$. such that for all $i \geq 0$, there is a transition $(s_i, p_{s_i,s_{i+1}}, s_{i+1}) \in \Delta$, with $p_{s_i,s_{i+1}} > 0$. Let $\Omega \subseteq V^\omega$ denote the set of traces of $M_A$. For a state $s = \langle \beta, v \rangle \in V$, let $Q(s) = v$ denote the vertex at state $s$. Generalizing this to traces, for a trace $t \in \Omega$, let $Q(t) = Q(s_0)Q(s_1)Q(s_2)\ldots \in Q^\omega$. We will consider $M_A$ with *initial states* from $Init = \{\langle \epsilon, v \rangle \mid v \in Q\}$. More generally we may have a probability distribution $p_{\text{init}} : V \mapsto [0,1]$ on initial states (we usually assume $p_{\text{init}}$ has support only in $Init$, and we always assume it has finite support). This induces a probability distribution on traces generated by random walks on $M_A$. Formally, we have a probability space $(\Omega, \mathcal{F}, \mathbf{Pr}_\Omega)$, parametrized by $p_{\text{init}}$, where $\mathcal{F} = \sigma(\mathcal{C}) \subseteq 2^\Omega$ is the $\sigma$-field generated by the set of *basic cylinder sets*, $\mathcal{C} = \{C(x) \subseteq \Omega \mid x \in V^*\}$, where for $x \in V^*$ the cylinder at $x$ is $C(x) = \{t \in \Omega \mid t = xw,\ w \in V^\omega\}$. The probability distribution $\mathbf{Pr}_\Omega : \mathcal{F} \mapsto [0,1]$ is determined uniquely by the probabilities of cylinder sets, which are given as follows:

$$\mathbf{Pr}_\Omega(C(s_0 s_1 \ldots s_n)) = p_{\text{init}}(s_0) p_{s_0,s_1} p_{s_1,s_2} \ldots p_{s_{n-1},s_n}$$

See, e.g., [Bil95]. RMCs where every component has at most one exit are called *1-exit* RMCs. RMCs where the total number of entries and exits is bounded by a constant $c$, (i.e., $\sum_{i=1}^{k} |En_i| + |Ex_i| \leq c$) are called *bounded total entry-exit* RMCs (Bd-RMCs, for short).

### 2.1 The central questions for model checking of RMCs.

We first define reachability probabilities that play an important role in our analysis. Given a vertex $u \in Q_i$ and an exit $ex \in Ex_i$, both in the same component $A_i$, let $q^*_{(u,ex)}$ denote the probability of eventually reaching the state $\langle \epsilon, ex \rangle$, starting at the state $\langle \epsilon, u \rangle$. Formally, we have $p_{\text{init}}(\langle \epsilon, u \rangle) = 1$, and $q^*_{(u,ex)} \doteq \mathbf{Pr}_\Omega(\{t = s_0 s_1 \ldots \in \Omega \mid \exists\, i\,, s_i = \langle \epsilon, ex \rangle\})$. As we shall see, the probabilities $q^*_{(u,ex)}$ will play an important role in obtaining other probabilities.

Recall that a Büchi automaton $B = (\Sigma, S, q_0, R, F)$, has an alphabet $\Sigma$, a set of states $S$, an initial state $q_0 \in S$, a transition relation $R \subseteq S \times \Sigma \times S$, and a set of accepting states $F \subseteq S$. A *run* of $B$ is a sequence $\pi = q_0 v_0 q_1 v_1 q_2 \ldots$ of alternating states and letters such that for all $i \geq 0$ $(q_i, v_i, q_{i+1}) \in R$. The $\omega$-word associated with run $\pi$ is $w_\pi = v_0 v_1 v_2 \ldots \in \Sigma^\omega$. The run $\pi$ is *accepting* if for infinitely many $i$, $q_i \in F$. Define the $\omega$-language $L(B) = \{w_\pi \mid \pi \text{ is an accepting run of } B\}$. Note that $L(B) \subseteq \Sigma^\omega$. Let $\mathcal{L} : Q \mapsto \Sigma$, be a given $\Sigma$-labelling of the vertices $v$ of RMC $A$. $\mathcal{L}$ naturally generalizes to $\mathcal{L} : Q^\omega \mapsto \Sigma^\omega$: for $w = v_0 v_1 v_2 \ldots \in Q^\omega$, $\mathcal{L}(w) = \mathcal{L}(v_0)\mathcal{L}(v_1)\mathcal{L}(v_2)\ldots$. Given RMC $A$, with initial state $s_0 = \langle \epsilon, u \rangle$, and given a BA $B$ over the alphabet $\Sigma$, let $P_A(L(B))$ denote the probability that a trace of $M_A$ is in $L(B)$. More precisely: $P_A(L(B)) \doteq \mathbf{Pr}_\Omega(\{t \in \Omega \mid \mathcal{L}(Q(t)) \in L(B)\})$. One needs to show that the sets $\{t \in \Omega \mid \mathcal{L}(Q(t)) \in L(B)\}$ are measurable (in $\mathcal{F}$). This is not difficult (see similar proofs in [CY95,Var85]). The *model checking* problems for $\omega$-regular properties of RMCs are:

(1) The *qualitative* model checking problems: Is $P_A(L(B)) = 1$? Is $P_A(L(B)) = 0$?
(2) The *quantitative* model checking problems: given $p \in [0,1]$, is $P_A(L(B)) \geq p$? Also, we may wish to approximate $P_A(L(B))$ to within a given number of bits of precision.

Note that if we have a routine for the problem $P_A(L(B)) \geq p?$, then we can approximate $P_A(L(B))$ to within $i$ bits of precision using binary search with $i$ calls to the routine. Thus, for quantitative model checking it suffices to address the first problem.

Note that probabilistic reachability is a special case of model checking: Given a vertex $u$ of the RMC $A$ and a subset of vertices $F$, the probability that the RMC starting at $u$ visits some vertex in $F$ (in some stack context) is equal to $P_A(L(B))$, where we let the labelling $\mathcal{L}$ map vertices in $F$ to 1 and the other vertices to 0, and $B$ is the 2-state automaton that accepts strings that contain a 1. Similarly, for the *repeated reachability* problem, where we are interested whether a trajectory from $u$ visits infinitely often a vertex of $F$, we can let $B$ be the (2-state deterministic) automaton that accepts strings with an infinite number of 1's.

To simplify the descriptions of our results, we assume henceforth that $\Sigma = Q$, the vertices of $A$. This is w.l.o.g. since the problem can be reduced to this case by relabelling the RMC $A$ and modifying the automaton $B$ (see, e.g., [CY95]), however care must be taken when measuring complexity separately in the RMC, $A$, and in the BA, $B$, since typically $B$ and $\Sigma$ are small in relation to $A$. Our complexity results hold with respect to the given inputs $A$, $B$.

### 2.2 Basic RMC theory & reachability analysis (from [EY04])

We recall some of the basic theory of RMCs developed in [EY04], where we studied reachability analysis. Considering the probabilities $q^*_{(u,ex)}$ as unknowns, we can set up a system of (non-linear) polynomial equations, such that the probabilities $q^*_{(u,ex)}$ are the *Least Fixed Point* (LFP) solution of this system. Use a variable $x_{(u,ex)}$ for each unknown probability $q^*_{(u,ex)}$. We will often find it convenient to index the variables $x_{(u,ex)}$ according to a fixed order, so we can refer to them also as $x_1, \ldots, x_n$, with each $x_{(u,ex)}$ identified with $x_j$ for some $j$. We thus have a vector of variables: $\mathbf{x} = (x_1\ x_2 \ldots x_n)^T$.

**Definition 1.** *Given RMC $A = (A_1, \ldots, A_k)$, define the system of polynomial equations, $S_A$, over the variables $x_{(u,ex)}$, where $u \in Q_i$ and $ex \in Ex_i$, for $1 \leq i \leq k$. The system contains one equation $x_{(u,ex)} = P_{(u,ex)}(\mathbf{x})$, for each variable $x_{(u,ex)}$. $P_{(u,ex)}(\mathbf{x})$ denotes a multivariate polynomial with positive rational coefficients. There are 3 cases, based on the "type" of vertex $u$:*

1. *Type I: $u = ex$. In this case: $x_{(ex,ex)} = 1$.*
2. *Type II: either $u \in N_i \setminus \{ex\}$ or $u = (b, ex')$ is a return port. In these cases:*
$$x_{(u,ex)} = \sum_{\{v \mid (u,p_{u,v},v) \in \delta\}} p_{u,v} \cdot x_{(v,ex)}.$$
3. *Type III: $u = (b, en)$ is a call port. In this case:*
$$x_{((b,en),ex)} = \sum_{ex' \in Ex_{Y(b)}} x_{(en,ex')} \cdot x_{((b,ex'),ex)}$$

*In vector notation, we denote $S_A = (x_j = P_j(\mathbf{x}) \mid j = 1, \ldots, n)$ by: $\mathbf{x} = P(\mathbf{x})$.*

Given $A$, we can construct $\mathbf{x} = P(\mathbf{x})$ in P-time: $P(\mathbf{x})$ has size $O(|A|\theta^2)$, where $\theta$ denotes the maximum number of exits of any component. For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, define $\mathbf{x} \preceq \mathbf{y}$ to mean that $x_j \leq y_j$ for every coordinate $j$. For $D \subseteq \mathbb{R}^n$, call a mapping $H : \mathbb{R}^n \mapsto \mathbb{R}^n$ *monotone* on $D$, if: for all $\mathbf{x}, \mathbf{y} \in D$, if $\mathbf{x} \preceq \mathbf{y}$ then $H(\mathbf{x}) \preceq H(\mathbf{y})$. Define $P^1(\mathbf{x}) = P(\mathbf{x})$, and $P^k(\mathbf{x}) = P(P^{k-1}(\mathbf{x}))$, for $k > 1$. Let $\mathbf{q}^* \in \mathbb{R}^n$ denote the $n$-vector of probabilities $q^*_{(u,ex)}$, using the same indexing as used for $\mathbf{x}$. Let $\mathbf{0}$ denote the all 0 $n$-vector. Define $\mathbf{x}^0 = \mathbf{0}$, and $\mathbf{x}^k = P(\mathbf{x}^{k-1}) = P^k(\mathbf{0})$, for $k \geq 1$. The map $P : \mathbb{R}^n \mapsto \mathbb{R}^n$ is monotone on $\mathbb{R}^n_{\geq 0}$.

**Theorem 1.** *([EY04], see also [EKM04])* $\mathbf{q}^* \in [0,1]^n$ *is the* Least Fixed Point *solution,* $\mathrm{LFP}(P)$*, of* $\mathbf{x} = P(\mathbf{x})$*. Thus,* $\mathbf{q}^* = P(\mathbf{q}^*)$ *and* $\mathbf{q}^* = \lim_{k\to\infty} \mathbf{x}^k$*, and for all* $k \geq 0$*,* $\mathbf{x}^k \preceq \mathbf{x}^{k+1} \preceq \mathbf{q}^*$*, and for all* $\mathbf{q}' \in \mathbb{R}^n_{\geq 0}$*, if* $\mathbf{q}' = P(\mathbf{q}')$*, then* $\mathbf{q}^* \preceq \mathbf{q}'$*.*

There are already 1-exit RMCs for which the probability $q^*_{(en,ex)}$ is irrational and not "solvable by radicals" ([EY04]). Thus, we can't compute probabilities exactly.

Given a system $x = P(x)$, and a vector $q \in [0,1]^n$, consider the following sentence in the *Existential Theory of Reals* (which we denote by $\boldsymbol{ExTh}(\mathbb{R})$):

$$\varphi \equiv \exists x_1, \ldots, x_m \bigwedge_{i=1}^{m} P_i(x_1, \ldots, x_m) = x_i \wedge \bigwedge_{i=1}^{m} 0 \leq x_i \wedge \bigwedge_{i=1}^{m} x_i \leq q_i$$

$\varphi$ is true precisely when there is some $z \in \mathbb{R}^m$, $0 \preceq z \preceq q$, and $z = P(z)$. Thus, if we can decide the truth of this sentence, we could tell whether $q^*_{(u,ex)} \leq p$, for some rational $p$, by using the vector $q = (1, \ldots, p, 1, \ldots)$. We will rely on decision procedures for $\boldsymbol{ExTh}(\mathbb{R})$. It is known that $\boldsymbol{ExTh}(\mathbb{R})$ can be decided in PSPACE and in exponential time, where the time exponent depends (linearly) only on the number of variables; thus for a fixed number of variables the algorithm runs in polynomial time [Can88,Ren92,BPR96]. As a consequence:

**Theorem 2.** *([EY04]) Given RMC A and rational value* $\rho$*, there is a PSPACE algorithm to decide whether* $q^*_{(u,ex)} \leq \rho$*, with running time* $O(|A|^{O(1)} \cdot 2^{O(m)})$ *where* $m$ *is the number of variables in the system* $x = P(x)$ *for A. Moreover* $q^*_{(u,ex)}$ *can be approximated to within* $j$ *bits of precision within PSPACE and with running time at most* $j$ *times the above.*

For Bd-RMCs, as shown in [EY04] it is possible to construct efficiently a system of equations in a *bounded* number of variables, whose LFP yields the entry-exit probabilities $q^*_{(en,ex)}$. Since $\boldsymbol{ExTh}(\mathbb{R})$ is decidable in P-time when the number of variables is bounded, this yields:

**Theorem 3.** *([EY04]) Given a Bd-RMC A and a rational value* $p \in [0,1]$*, there is P-time algorithm that decides for a vertex* $u$ *and exit* $ex$*, whether* $q^*_{(u,ex)} \geq p$ *(or* $< p$*).*

For single-exit RMCs (SCFGs) the qualitative termination (exit) problem can be solved efficiently, using graph theoretic techniques and an eigenvalue characterization.

**Theorem 4.** *([EY04]) There is P-time algorithm that for a 1-exit RMC, vertex* $u$ *and exit* $ex$*, decides which of the following holds:(1)* $q^*_{(u,ex)} = 0$*,(2)* $q^*_{(u,ex)} = 1$*,or (3)*$0 < q^*_{(u,ex)} < 1$*.*

Hardness, such as NP-hardness, is not known for RMC reachability. However, in [EY04] we gave strong evidence of "difficulty": the square-root sum problem is P-time reducible to deciding whether $q^*_{(u,ex)} \geq p$, in a single-exit RMC, and to deciding whether $q^*_{(u,ex)} = 1$ for a 2-exit RMC (see also [BKS04]. *Square-root sum* is the following decision problem: given $(d_1, \ldots, d_n) \in \mathbb{N}^n$ and $k \in \mathbb{N}$, decide whether $\sum_{i=1}^n \sqrt{d_i} \leq k$. It is solvable in PSPACE, but it has been a major open problem since the 1970's (see, e.g., [GGJ76,Tiw92]) whether it is solvable even in NP.

As a practical algorithm for numerically computing the probabilities $q^*_{(u,ex)}$, it was proved in [EY04] that multi-dimensional Newton's method converges monotonically to the LFP of $\mathbf{x} = P(\mathbf{x})$, and constitutes a rapid acceleration of iterating $P^k(\mathbf{0})$, $k \to \infty$ (see [EY04]).

## 3 The Conditioned Summary Chain $M'_A$

For an RMC $A$, suppose we somehow have the probabilities $q^*_{(u,ex)}$ "in hand". Based on these, we construct a *conditioned summary chain*, $M'_A$, a finite markov chain that will allow us to answer repeated reachability questions. Extensions of $M'_A$ will later be a key to model checking RMCs. Since probabilities $q^*_{(u,ex)}$ are potentially irrational, we can not compute $M'_A$ exactly. However, $M'_A$ will be important in our correctness arguments, and we will in fact be able to compute the "structure" of $M'_A$, i.e., what transitions have non-zero probability. The structure of $M'_A$ will be sufficient for answering various "qualitative" questions.

We will assume, w.l.o.g., that each RMC has one initial state $s_0 = \langle \epsilon, en_{\text{init}} \rangle$, with $en_{\text{init}}$ the only entry of some component that does not contain any exits. Any RMC can readily be converted to an "equivalent" one in this form, while preserving relevant probabilities.

Before describing $M'_A$, let us recall from [AEY01], the construction of a "summary graph", $H_A = (Q, E_{H_A})$, which ignores probabilities and is based only on information about reachability in the underlying RSM of $A$. Let $R$ be the binary relation between entries and exits of components such that $(en, ex) \in R$ precisely when there exists a path from $\langle \epsilon, en \rangle$ to $\langle \epsilon, ex \rangle$, in the underlying graph of $M_A$. The edge set $E_{H_A}$ is defined as follows. For $u, v \in Q$, $(u, v) \in E_{H_A}$ iff one of the following holds:

1. $u$ is not a call port, and $(u, p_{u,v}, v) \in \delta$, for $p_{u,v} > 0$.
2. $u = (b, en)$ is a call port, and $(en, ex) \in R$, and $v = (b, ex)$ is a return port.
3. $u = (b, en)$ is a call port, and $v = en$ is the corresponding entry.

For each vertex $v \in Q_i$, let us define the probability of *never exiting*: $\text{ne}(v) = 1 - \sum_{ex \in Ex_i} q^*_{(v,ex)}$. Call a vertex $v$ *deficient* if $\text{ne}(v) > 0$, i.e. there is a nonzero probability that if the RMC starts at $v$ it will never terminate (reach an exit of the component).

We define $M'_A = (Q_{M'_A}, \delta_{M'_A})$ as follows. The set of states $Q_{M'_A}$ of $M'_A$ is the set of deficient vertices: $Q_{M'_A} = \{v \in Q \mid \text{ne}(v) > 0\}$. For $u, v \in Q_{M'_A}$, there is a transition $(u, p'_{u,v}, v)$ in $\delta_{M'_A}$ if and only if one of the following conditions holds:

1. $u, v \in Q_i$ and $(u, p_{u,v}, v) \in \delta_i$, and $p'_{u,v} = \frac{p_{u,v} \cdot \text{ne}(v)}{\text{ne}(u)}$.
2. $u = (b, en) \in \mathit{Call}_b$ and $v = (b, ex) \in \mathit{Return}_b$ and $q^*_{(en,ex)} > 0$, and $p'_{u,v} = \frac{q^*_{(en,ex)} \text{ne}(v)}{\text{ne}(u)}$.
3. $u = (b, en) \in \mathit{Call}_b$ and $v = en$, and $p'_{u,v} = \frac{\text{ne}(v)}{\text{ne}(u)}$. We call these transitions, from a call port to corresponding entry, special <u>*red*</u> transitions.

Note that in all three cases, $p'_{u,v}$ is well-defined (the denominator is nonzero) and it is positive. Recall that we assumed that the initial vertex $en_{\text{init}}$ is the entry of a component $A_0$, and $A_0$ has no exits. Thus for all $v \in Q_0$, $\text{ne}(u) = 1$, and thus $Q_0 \subseteq Q_{M'_A}$, and if $(u, p_{u,v}, v) \in \delta_0$, then $(u, p_{u,v}, v) \in \delta_{M'_A}$.

**Proposition 1.** *Probabilities on transitions out of each state in $Q_{M'_A}$ sum to 1.*

$M'_A$ is an ordinary (flat) Markov chain. Let $(\Omega', \mathcal{F}', \mathbf{Pr}_{\Omega'})$ denote the probability space on traces of $M'_A$. We now define a mapping $\rho : \Omega \mapsto \Omega' \cup \{\star\}$, that maps every trace $t$ of the original (infinite) Markov chain $M_A$, either to a unique trajectory $\rho(t) \in \Omega'$ of the MC $M'_A$, or to the special symbol $\star$. Trajectories mapped to $\star$ will be precisely those that go through missing vertices $u \in Q$ that are not in $Q_{M'_A}$, i.e., with $\text{ne}(u) = 0$. We will show that the total probability of all these trajectories is 0, i.e., that $\mathbf{Pr}_\Omega(\rho^{-1}(\star)) = 0$, and

moreover, that $M'_A$ preserves the probability measure of $M_A$: for all $D \in \mathcal{F}'$, $\rho^{-1}(D) \in \mathcal{F}$, and $\mathbf{Pr}_{\Omega'}(D) = \mathbf{Pr}_{\Omega}(\rho^{-1}(D))$.

We define $\rho$ in two phases. We first define, as a precursor to $\rho(t)$, a map $\rho^H : \Omega \mapsto Q^\omega$, where every trajectory $t \in \Omega$ is mapped to an infinite path $\rho^H(t)$ in the summary graph $H_A$. Thereafter, we let $\rho(t) = \rho^H(t)$ if all vertices of $\rho^H(t)$ are in $M'_A$, and let $\rho(t) = \star$ otherwise. We define $\rho^H$ for a trace $t = s_0 s_1 \dots s_i \dots$, sequentially based on prefixes of $t$, as follows. By assumption, $s_0 = \langle \epsilon, en_{\text{init}} \rangle$. $\rho^H$ maps $s_0$ to $en_{\text{init}}$. Suppose $s_i = \langle \beta, u \rangle$, and, inductively, suppose that $\rho^H$ maps $s_0 \dots s_i$ to $e_{\text{init}} \dots u$. First, suppose $u$ is not a call port, and that $s_{i+1} = \langle \beta, v \rangle$, then $s_0 \dots s_i s_{i+1}$ maps to $e_{\text{init}} \dots uv$. Next, suppose $u = (b, en)$ is a call port and $s_{i+1} = \langle \beta b, en \rangle$. If the trace eventually returns from this call, i.e. there exists $j > i+1$, such that $s_j = \langle \beta b, ex \rangle$ and $s_{j+1} = \langle \beta, (b, ex) \rangle$, and such that each of of the states $s_{i+1} \dots s_j$, have $\beta b$ as a prefix of the call stack, then $s_0 \dots s_j$ is mapped by $\rho^H$ to $en_{\text{init}} \dots u(b, ex)$. If the trace never returns from this call, then $s_0 \dots s_i s_{i+1}$ maps to $en_{\text{init}} \dots u\ en$. This concludes the definition of $\rho^H$. We show that the mapping $\rho$ is measure preserving.

**Lemma 1.** $\mathbf{Pr}_{\Omega}(\rho^{-1}(\star)) = 0$.

**Lemma 2.** *For all $D \in \mathcal{F}'$, $\rho^{-1}(D) \in \mathcal{F}$ and $\mathbf{Pr}_{\Omega}(\rho^{-1}(D)) = \mathbf{Pr}_{\Omega'}(D)$.*

Let $H'_A = (Q_{H'_A}, E_{H'_A})$ be the underlying directed graph of $M'_A$. In other words, the states $Q_{H'_A} = Q_{M'_A}$, and $(u, v) \in E_{H'_A}$ iff $(u, p'_{u,v}, u) \in \delta_{M'_A}$. We will show that we can compute $H'_A$ in P-time for single-exit RMCs and Bd-RMCs, and in PSPACE for arbitrary RMCs. The basic observation is that the structure of $M'_A$ depends only on qualitative facts about the probabilities $q^*_{(en,ex)}$ and $\text{ne}(u)$, for $u \in Q$.

**Proposition 2.** *For a RMC $A$ (respectively, single-exit or Bd-RMC), and $u \in Q$, we can decide whether* $\text{ne}(u) > 0$ *in PSPACE (respectively, P-time).*

*Proof.* Suppose $u$ is in a component $A_i$ where $Ex_i = \{ex_1, \dots, ex_k\}$. Clearly, $\text{ne}(u) > 0$ iff $\sum_{j=1}^{k} q^*_{(u,ex_j)} < 1$. Consider the following sentence, $\varphi$, in $\boldsymbol{ExTh}(\mathbb{R})$.

$$\varphi \equiv \exists x_1, \dots, x_n \bigwedge_{i=1}^{n} P_i(x_1, \dots, x_n) = x_i \wedge \bigwedge_{i=1}^{n} 0 \leq x_i \wedge \sum_{j=1}^{k} x_{(u,ex_j)} < 1$$

Since $\mathbf{q}^*$ is the LFP solution of $\mathbf{x} = P(\mathbf{x})$, $\varphi$ is true in the reals if and only if $\sum_{j=1}^{k} q^*_{(u,ex_j)} < 1$. This query can be answered in PSPACE. In the special case of a single-exit RMC, we have $Ex_i = \{ex_1\}$, and $\text{ne}(u) > 0$ iff $q^*_{(u,ex_1)} < 1$. As mentioned in section 2.2, this can be answered in P-time for single-exit RMCs ([EY04]). Similarly, for Bd-RMCs the question can be answered in P-time by the techniques developed in [EY04]. □

Once we determine the deficient vertices of $A$, the structure of $M'_A$ can be determined in polynomial time.

**Corollary 1.** *For a RMC $A$ (respectively, single-exit or Bd-RMC), we can compute $H'_A$ in PSPACE (respectively, in polynomial time).*

*Proof.* Recall that $u \in Q_{H'_A}$ precisely when $u \in Q$ and $\text{ne}(u) > 0$. Thus we can determine the set of nodes with the said complexities, respectively. The transitions of type 1 and 3 in the definition of $M'_A$ are immediately determined. For the type 2 transitions, where $u = (b, en)$ and $v = (b, ex)$, in order to determine whether to include the corresponding summary edge $(u, v)$ we need to decide whether $q^*_{(en,ex)} > 0$. This can be done in polynomial time by invoking the reachability algorithm for RSM's [AEY01,BGR01]. □

## 4 Qualitative Model checking

**Upper bounds.** Given an RMC $A = (A_1, \ldots, A_k)$ and a (nondeterministic) Büchi automaton $B = (\Sigma, S, q_0, R, F)$ whose alphabet $\Sigma$ is the vertex set of $A$, we wish to determine whether $P_A(L(B)) = 1$, $= 0$, or is in-between. We will construct a finite Markov chain $M'_{A,B}$ such that $P_A(L(B))$ is equal to the probability that a trajectory of $M'_{A,B}$ starting from a given initial state reaches one of a designated set of "accepting" bottom SCCs.

First, let $B' = (\Sigma, 2^S, \{q_0\}, R', F')$ be the deterministic automaton obtained by the usual subset construction on $B$. In other words, the states of $B'$ are subsets $T \subseteq S$, and the transition function $R' : (2^S \times \Sigma) \mapsto 2^S$ is given by: $R'(T_1, v) = \{q' \in S \mid \exists q \in T_1 \; s.t. \; (q, v, q') \in R\}$. (We are making no claim that $L(B) = L(B')$.)

Next we define the standard *product* RMC, $A \otimes B'$, of the RMC $A$, and the deterministic Büchi automaton $B'$. $A \otimes B'$ has the same number of components as $A$. Call these $A'_1, \ldots, A'_k$. The vertices in component $A'_i$ are pairs $(u, T)$, where $u \in Q_i$ and $T \in 2^S$, and $(u, T)$ is an entry (exit) iff $u$ is an entry (exit). The transitions of $A'_i$ are as follows: there is a transition $((u, T), p_{u,v}, (v, R'(T, v)))$ in $A'_i$ iff there is a transition $(u, p_{u,v}, v)$ in $A_i$.

Define $M'_{A,B}$ as $M'_{A,B} = M'_{A \otimes B'}$. Thus $M'_{A,B}$ is the conditioned summary chain of RMC $A \otimes B'$. For qualitative analysis on $M'_{A,B}$, we need the underlying graph $H'_{A,B}$. Importantly for the complexity of our algorithms, we do not have to explicitly construct $A \otimes B'$ to obtain $H'_{A,B}$. Observe that states of $M'_{A,B} = (Q \times 2^S, \delta_{M'_{A,B}})$ are pairs $(v, T)$ where $v$ is a state of $M'_A$, and $T$ a state of $B'$. The initial state of $M'_{A,B}$ is $(v_0, \{q_0\})$, where $v_0$ is the initial state of $M'_A$ and $q_0$ of $B$. The transitions of $M'_{A,B}$ from a state $(v, T)$ are as follows:

- Case 1: $v$ is not a call port. Then for every transition $(v, p'_{v,v'}, v') \in \delta_{M'_A}$, we have a corresponding transition $((v, T), p'_{v,v'}, (v', R'(T, v'))) \in \delta_{M'_{A,B}}$.
- Case 2: $v$ is a call port, $v = (b, en)$ where $v$ is vertex in component $A_i$ and box $b$ is mapped to component $A_j$. If there is a *red* transition $(v, p_{v,en}, en) \in \delta_{M'_A}$ then there is a *red* transition $((v, T), p_{v,en}, (en, R'(T, en)) \in \delta_{M'_{A,B}}$ with the same probability.
- Case 3: If $v$ has a summary transition $(v, p_{v,v'}, v')$ in $M'_A$, where $v' = (b, ex)$, then we have summary transitions of the form $((v, T), p'', (v', T'))$ in $M'_{A,B}$ to states of the form $(v', T')$ iff there exists a path in $M_A$ from $\langle \epsilon, en \rangle$ to $\langle \epsilon, ex \rangle$ which, viewed as a string, drives $B'$ from $T$ to $T'$; the probability $p''$ of the transition is $p'' = p' \cdot ne(v')/ne(v)$ where $p'$ is the probability of all such $v$-$v'$ paths that drive $B'$ from $T$ to $T'$.

$M'_{A,B}$ is a well-defined Markov chain, which is a refinement of $M'_A$. That is, every trajectory of $M'_{A,B}$ projected on the first component is a trajectory of $M'_A$ and the projection preserves probabilities. We can define a mapping $\sigma$ from the trajectories $t$ of the original (infinite) Markov chain $M_A$ to the trajectories of $M'_{A,B}$, or the special symbol $\star$, in a similar manner as we defined the mapping $\rho$ from trajectories of $M$ to $M'_A$. For a trajectory $t$ of $M_A$, it is easy to see that if $\rho(t) \neq \star$ then also $\sigma(t) \neq \star$. Thus, with probability 1 a trajectory of $M_A$ is mapped to one of $M'_{A,B}$. Furthermore, we can show along similar lines the analogue of Lemma 2, i.e. the mapping $\sigma$ preserves probabilities.

Consider a product graph (without probabilities) $M'_A \otimes B$ between the Markov chain $M'_A$ and the given nondeterministic BA $B$ (not $B'$) as follows: The product has nodes $(v, q)$, for all vertices $v$ of $M'_A$ and states $q$ of $B$, and an edge $(v, q) \to (v', q')$ if either (i) $v \to v'$ is an ordinary edge or a red edge of $M'_A$ and $q$ has a transition to $q'$ on input $v'$, or (ii) $v \to v'$ is a summary edge and the RMC has a path from $v$ to $v'$ that corresponds to a run of $B$ from

$q$ to $q'$; if the run goes through an accepting state then we mark the edge $(v,q) \to (v',q')$ as an *accepting* edge. Also, call a node $(v,q)$ *accepting* if $q \in F$ is an accepting state of $B$.

With every transition (edge) of $M'_{A,B}$ and every edge of $M'_A \otimes B$ we associate a string $\gamma$ over $\Sigma$ (the vertex set of $A$) that caused the edge to be included; i.e., if edge $(v,T) \to (v',T')$ of $M'_{A,B}$ (respectively, edge $(v,q) \to (v',q')$ of $M'_A \otimes B$) corresponds to an ordinary or red edge of $M'_A$ then $\gamma = v'$. If it corresponds to a summary edge then we let $\gamma$ be any string that corresponds to a $v - v'$ path that drives $B'$ from $T$ to $T'$ (resp., for which $B$ has a path from $q$ to $q'$; if the edge $(v,q) \to (v',q')$ is marked as accepting then we pick a path that goes through an accepting state of $B$). In the case of a summary edge, there may be many strings $\gamma$ as above; we just pick anyone of them.

Let $t$ be any trajectory of $M_A$ starting from $\langle \epsilon, v \rangle$, for some vertex $v$ of $M'_A$ and let $r$ be a corresponding run of $B$ starting from a state $q$. With probability 1, $t$ maps to a trajectory $t' = \rho(t)$ of $M'_A$. The mapping $\rho$ can be extended to pairs $(t,r)$, where $r$ is a run of $B$ on $t$, i.e., the pair $(t,r)$ is mapped to a run (path) $r' = \rho(t,r)$ of $M'_A \otimes B$. If $r$ is an accepting run of $B$ then $r'$ goes infinitely often through an accepting node or an accepting edge. The converse does not hold necessarily: a non-accepting run $r$ of $B$ corresponding to a trajectory $t$ may be mapped to a run $r'$ of $M'_A \otimes B$ that traverses infinitely often an accepting edge.

If $B$ is a deterministic BA, then $M'_{A,B}$ and $M'_A \otimes B$ are clearly the same (except that in $M'_A \otimes B$ we did not include the probabilities of the edges). In this case, the analysis is simpler. Let us say that a bottom strongly connected component (SCC) of $M'_{A,B}$ (and $M'_A \otimes B$) is *accepting* iff it contains an accepting node or an accepting edge.

**Theorem 5.** *For a RMC $A$ and a deterministic BA $B$, the probability $P_A(L(B))$ that a trajectory of $A$ is accepted by $B$ is equal to the probability that a trajectory of $M'_{A,B}$ starting from the initial node $(v_0, q_0)$ reaches an accepting bottom SCC.*

Suppose now that $B$ is nondeterministic. We will follow the approach of [CY95] for flat Markov chains, except that here we have to deal with recursive calls and with the summary edges of the constructed Markov chain $M'_{A,B}$ which correspond to sets of paths in the original chain $M_A$ rather than single steps. This complicates things considerably.

Let $v$ be a vertex of $M'_A$ and $q \in F$ an accepting state of $B$. Let $D(v,q)$ be the subgraph of $M'_{A,B}$ induced by the node $(v,\{q\})$ and all nodes reachable from it . We say that the pair $(v,q)$ is *special of type 1* if some bottom SCC $C$ of $D(v,q)$ contains a state $(v,T)$ with $q \in T$. We associate with such a pair $(v,q)$ a string $\gamma(v,q) \in \Sigma^*$ that is the concatenation of the strings associated with the edges of $D(v,q)$ on a path from $(v,\{q\})$ to a node of $C$. (There may be many such paths; just pick anyone.)

Let $v = (b,en)$ be a vertex of $M'_A$ that is a call port of a box $b$ of $A$ and let $q \notin F$ be a non-accepting state of $B$. Define a graph $D(v,q)$ as follows. The graph contains a root node $vq$ and a subgraph of $M'_{A,B}$ consisting of the nodes reachable from $vq$ after we add the following edges. We add an edge from $vq$ to a node $(v',\{q'\})$ of $M'_{A,B}$, where $v' = (b,ex)$ is a return port of the same box $b$ as $v$, iff there is a path $\gamma$ from $\langle \epsilon, en \rangle$ to $\langle \epsilon, ex \rangle$ such that $B$ has a run from $q$ to $q'$ on $\gamma$ that goes through an accepting state; we label the edge $vq \to (v',\{q'\})$ with such a string $\gamma$. The graph $D(v,q)$ consists of the root $vq$ and the subgraph of $M'_{A,B}$ induced by all the nodes that are reachable from $vq$ after adding the above edges. We call the pair $(v,q)$ *special of type 2* if some bottom SCC $C$ of $D(v,q)$ contains a state $(v,T)$ with $q \in T$. As in the previous case, we associate with the pair $(v,q)$ a string $\gamma(v,q) \in \Sigma^*$ that is the concatenation of the strings associated with the edges of $D(v,q)$ on a path from $vq$ to a node of $C$. Special pairs have the following important properties.

**Lemma 3.** *Suppose $(v, q)$ is special and that RMC $A$ starts at $\langle \epsilon, v \rangle$ and first performs the transitions in $\gamma(v, q)$. Then with probability 1 such a trajectory $t$ of the RMC is accepted by $B$ with initial state $q$. Specifically, there is a corresponding accepting run $r$ of $B$ such that $\rho(t, r)$ is a run of $M'_A \otimes B$ starting from $(v, q)$ that infinitely repeats node $(v, q)$ if $(v, q)$ is special of type 1, or repeats an accepting edge out of $(v, q)$ if $(v, q)$ is special of type 2.*

**Lemma 4.** *Suppose there is non-zero probability that a trajectory of the RMC $A$ starting at any vertex $u \in M'_A$ has a corresponding run in $M'_A \otimes B$ starting from any node $(u, p)$ which repeats an accepting state $(v, q)$ infinitely often or repeats an accepting edge $(v, q) \to (v', q')$ infinitely often. Then $(v, q)$ is special.*

**Proposition 3.** *$P_A(L(B)) > 0$ iff from $(v_0, q_0)$ in $M'_A \otimes B$ we can reach a special $(v, q)$.*

Call a bottom SCC of the flat Markov chain $M'_{A,B}$ *accepting* if it contains a state $(v, T)$ and $T$ contains some $q$ such that $(v, q)$ is special; otherwise call the bottom SCC *rejecting*.

**Theorem 6.** *$P_A(L(B))$ is equal to the probability that a trajectory of $M'_{A,B}$ starting from the initial state $(v_0, \{q_0\})$ reaches an accepting bottom SCC.*

It follows that $P_A(L(B)) = 1$ iff all the bottom SCCs of $M'_{A,B}$ reachable from $(v_0, \{q_0\})$ are accepting, and $P_A(L(B)) = 0$ iff no reachable bottom SCC is accepting (or equivalently by Proposition 3, there is no path in $M'_A \otimes B$ from $(v_0, \{q_0\})$ to any special node $(v, q)$).

As with $M'_A$ and $H'_A$, let $H'_{A,B}$ denote the underlying directed graph of $M'_{A,B}$. For the qualitative problem, we only need (1) to construct $H'_{A,B}$ and thus only need to know which nodes and edges are present, and (2) to determine which pairs $(v, q)$ are special, and hence which bottom SCCs are accepting. Thus we first have to identify the vertices $u$ of the RMC $A$ for which $\mathrm{ne}(u) > 0$, which can be done in PSPACE for general RMCs and P-time for single-exit RMCs and for Bd-RMCs. Then, the edges of $H'_{A,B}$ can be determined by the standard reachability algorithm for RSMs ([AEY01]). This works by first constructing the genuine product of the underlying RSM of $A$ (ignoring probabilities on transitions) together with the Büchi automaton $B'$. This defines a new RSM $A \otimes B'$ (no probabilities), whose size is polynomial in $A$ and $B'$, and thus is exponential in the original non-deterministic Büchi automaton $B$. The time required for reachability analysis for RSMs is polynomial ([AEY01]). Thus, once we have identified the deficient vertices of the RMC, the rest of the construction of $H'_{A,B}$ takes time polynomial in $A$ and $B'$.

To determine which pairs $(v, q)$ are special, we construct for each candidate pair $(v, q)$ the graph $D(v, q)$. For a pair $(v, q)$ with $q \in F$, this is immediate from $H'_{A,B}$. For a pair $(v, q)$ with $q \notin F$ and $v = (b, en)$ a call port of a box $b$, we test for each return port $v' = (b, ex)$ of the box and each state $q'$ of $B$ whether there should be an edge $vq \to (v', \{q'\})$; this involves a call to the RSM algorithm of [AEY01] to determine whether there is a path in the RSM $A \otimes B$ from $(en, q)$ to $(ex, q')$ (with empty stack) that goes through a vertex whose second component is an accepting state of $B$. Once we determine these edges, we can construct $D(v, q)$. This takes time polynomial in $A$ and $B'$. Then compute the SCCs of $D(v, q)$, examine the bottom SCCs and check if one of them contains $(v, T)$ with $q \in T$.

Finally, once we have identified the special pairs, we examine the reachable bottom SCCs of $H'_{A,B}$ and determine which ones are accepting and which are rejecting. The dependence of the time complexity on the size of the given RMC $A$ is polynomial except for the identification of the vertices $u$ for which $\mathrm{ne}(u) > 0$. The dependence on $|B|$ is exponential because of the subset construction. If $B$ is deterministic to begin with, we avoid the exponential blow-up and thus have polynomial complexity in $B$. Thus we have:

**Theorem 7.** *Given a RMC $A$ and a Büchi automaton $B$, we can decide whether $P_A(L(B)) = 0$, $P_A(L(B)) = 1$, or $0 < P_A(L(B)) < 1$ in PSPACE in $A$, and EXPTIME in $B$. For a 1-exit RMC or Bd-RMC, the time complexity is polynomial in $A$. Furthermore, if $B$ is deterministic, the dependence of the time complexity on $|B|$ is also polynomial.*

**Lower Bounds.** We show conversely that the exponential time complexity of qualitative model checking for a nondeterministic BA is in general unavoidable.

**Theorem 8.** *The qualitative problem of determining whether a given RMC $A$ satisfies a property specified by a Büchi automaton $B$ with probability $= 1$, (i.e., whether $P_A(L(B)) = 1$)) is EXPTIME-complete. Furthermore, this holds even if the RMC is fixed and each component has one entry and one exit. Moreover, the qualitative "emptiness" problem, namely determining whether $P_A(L(B)) = 0$, is also EXPTIME-complete, again even when the RMC is fixed and each component has one entry and one exit.*

## 5 Quantitative model checking

As we have mentioned, the transition probabilities of the chain $M'_{A,B}$ cannot be computed exactly, but instead have to be determined implicitly. To do quantitative model checking in PSPACE in $|A|$, it will be crucial to use $\boldsymbol{ExTh}(\mathbb{R})$ to uniquely identify $\mathrm{LFP}(P)$ for the systems $x = P(x)$. The folowing key theorem enables this.

**Theorem 9.** *(The Unique Fixed Point Theorem) The set of equations $x = P(x)$ has a unique fixed point that satisfies $\sum_{ex} x_{(u,ex)} < 1$ for every deficient vertex $u$, and $\sum_{ex} x_{(u,ex)} \leq 1$ for every other vertex $u$. (This fixed point, of course, is $q^* = \mathrm{LFP}(P)$.)*

**Theorem 10.** *Given RMC, $A$, and BA, $B$, and a rational value $p \in [0,1]$, we can decide whether $P_A(L(B)) \geq p$ in PSPACE in $|A|$ and in EXPSPACE in $B$, specifically in space $O(|A|^{c_1} 2^{c_2|B|})$ for some constants $c_1, c_2$. Furthermore, if $B$ is deterministic we can decide this in PSPACE in both $A$ and $B$.*

*Proof.* We make crucial use of Theorem 9, and we combine this with use of the summary chain $M'_{A,B}$, and queries to $\boldsymbol{ExTh}(\mathbb{R})$. Observe that by Theorem 6, all we need to do is "compute" the probability that a trajectory of $M'_{A,B}$, starting from the initial state $(v_0, \{q_0\})$ reaches an accepting bottom SCC. We can not compute $M'_{A,B}$ exactly, however, we will be able to identify the transition probabilities uniquely inside a $\boldsymbol{ExTh}(\mathbb{R})$ query, and will, inside the same query identify the probability of reaching an accepting bottom SCC.

Let $\mathbf{q}^* = \mathrm{LFP}(P)$ be the solution vector of probabilities for the system $\mathbf{x} = P(\mathbf{x})$ associated with RMC $A$. Recall that by Proposition 2, we can compute in PSPACE in $|A|$ the set $Q' = \{u \in Q \mid \mathrm{ne}(u) > 0\}$ of deficient vertices. We do this as a first step. Consider next the following quantifier-free formula, where $c(u)$ is the index of the component of a vertex $u$:

$$\varphi_1(\mathbf{x}) \equiv \mathbf{x} = P(\mathbf{x}) \ \wedge \ 0 \preceq \mathbf{x} \ \wedge \bigwedge_{u \in Q'} \sum_{ex \in Ex_{c(u)}} x_{(u,ex)} < 1 \ \wedge \bigwedge_{u \in Q \setminus Q'} \sum_{ex \in Ex_{c(u)}} x_{(u,ex)} = 1$$

By Theorem 9, the only solution vector $\mathbf{x}$ in $\mathbb{R}^n$ for which $\varphi_1(\mathbf{x})$ holds true is $\mathbf{q}^*$. In other words, $\varphi_1$ uniquely identifies $\mathrm{LFP}(P)$.

Recall that $\text{ne}(u) = 1 - \sum_{ex \in Ex_{c(u)}} q^*_{(u,ex)}$. Now, let $\mathbf{y}$ be a vector of variables indexed by vertices of $A$, and let $\varphi_2(\mathbf{x}, \mathbf{y}) \equiv \bigwedge_{u \in Q} y_u = 1 - \sum_{ex \in Ex_{c(u)}} x_{(u,ex)}$. The only vector of reals $(\mathbf{x}, \mathbf{y})$ that satisfies $\varphi_1 \wedge \varphi_2$ is the one where $x_{(u,ex)} = q^*_{(u,ex)}$ and $y_u = \text{ne}(u)$.

Recall the construction of $M'_{A,B}$. The states of $M'_{A,B}$ are pairs $(v, T)$, where $v \in Q'$, and $T \subseteq S$ is a set of states of $B$. The transitions of $M'_{A,B}$ come in three varieties.

Case 1: $v$ is not a call port, and $(v, p'_{v,v'}, v') \in \delta_{M'_A}$. Then we have a corresponding transition $((v, T), p'_{v,v'}, (v', R'(T, v'))) \in \delta_{M'_{A,B}}$, where $p'_{v,v'} = p_{v,v'} \text{ne}(v')/\text{ne}(v)$, and thus $p'_{v,v'} \text{ne}(v) = p_{v,v'} \text{ne}(v')$. Associate a variable $z_{v,v'}$ with each such probability $p'_{v,v'}$, and define the formula: $\varphi_3(\mathbf{y}, \mathbf{z}) \equiv \bigwedge_{(v,v') \in \text{Case1}} z_{v,v'} y_v = p_{v,v'} y_{v'}$.

Case 2: $v$ is a call port, $v = (b, en)$ where $v$ is vertex in component $A_i$ and box $b$ is mapped to component $A_j$, and $v' = en$, and there is a *red* transition $(v, p'_{v,v'}, v') \in \delta_{M'_A}$. Then there is a *red* transition $((v, T), p'_{v,v'}, (v', R'(T, v')) \in \delta_{M'_{A,B}}$ with the same probability. Here $p'_{v,v'} = \text{ne}(v')/\text{ne}(v)$, and thus $p'_{v,v'} \text{ne}(v) = \text{ne}(v')$. Associate a variable $z_{v,v'}$ with each such probability $p'_{v,v'}$, and define: $\varphi_4(\mathbf{y}, \mathbf{z}) \equiv \bigwedge_{(v,v') \in \text{Case2}} z_{v,v'} y_v = y_{v'}$.

Case 3: $v$ is a call port that has a summary transition $(v, p'_{v,v'}, v')$ in $M'_A$ to a vertex $v' = (b, ex)$, then we have summary transitions of the form $((v, T), p'', (v', T'))$ in $M'_{A,B}$ to the following set of states of the form $(v', T')$: If there exists a path of $M_A$ that starts at the entry $en$ of $A_j$ and ends at the exit $ex$ (with empty call stack) which, viewed as a string drives $B'$ from $T$ to $T'$, then we include the edge $((v, T), p'_{(v,T),(v',T')}, (v', T'))$ in $\delta_{M'_{A,B}}$, where $p'_{(v,T),(v',T')} = q^*_{((en,T),(ex,T'))} \cdot ne(v')/ne(v)$, and where $q^*_{((en,T),(ex,T'))}$ is the probability of reaching $\langle \epsilon, (ex, T') \rangle$ from $\langle \epsilon, (en, T) \rangle$ in the product RMC $A \otimes B'$. First, compute $A \otimes B'$ and its associated equations $\mathbf{w} = P^{\otimes}(\mathbf{w})$ explicitly. Note that $|A \otimes B'| = O(|A||B'|)$. Let $Q^{\otimes}$ be the set of vertices of $A \otimes B'$. We can compute the set $Q'^{\otimes}$ of vertices $v$ of $A \otimes B'$, for which $\text{ne}(v) > 0$ in PSPACE in $|A \otimes B'|$. Consider now the quantifier-free formula:

$$\varphi_5(\mathbf{w}) \equiv \mathbf{w} = P^{\otimes}(\mathbf{w}) \wedge 0 \preceq \mathbf{w} \wedge \bigwedge_{u \in Q'^{\otimes}} \sum_{ex \in Ex_{c(u)}} w_{(u,ex)} < 1 \wedge \bigwedge_{u \in Q^{\otimes} \setminus Q'^{\otimes}} \sum_{ex \in Ex_{c(u)}} w_{(u,ex)} = 1$$

By Theorem 9, $\text{LFP}(P^{\otimes})$, is the only vector in $\mathbb{R}^n$ for which $\varphi_5(\mathbf{w})$ holds true. In other words, $\varphi_5$ uniquely identifies $\text{LFP}(P^{\otimes})$. Now, associate a variable $z_{(v,T),(v',T')}$ with each probability $p'_{(v,T),(v',T')}$, where $v = (b, en)$ and $v' = (b, ex)$, and define: $\varphi_6(\mathbf{y}, \mathbf{w}, \mathbf{z}) \equiv \bigwedge_{((v,T),(v',T')) \in \text{Case3}} z_{(v,T),(v',T')} y_v = w_{((en,T),(ex,T'))} y_{v'}$.

Observe that $\bigwedge_{j=1}^{6} \varphi_j$ has a unique solution, and the values of variables $\mathbf{z}$ in this solution identify the probabilities $p'$ on transitions of $M'_{A,B}$. By the qualitative methods of section 4, we compute the underlying graph $H'_{A,B}$ of $M'_{A,B}$, and we compute the SCCs of $H'_{A,B}$ that contain either an accepting node or an accepting edge.

Let us define a revised finite markov chain, $M''_{A,B}$, in which we remove all SCCs in $M'_{A,B}$ that contain an accepting node or edge, and replace them by a new absorbing node $v^*$, with a probability 1 transition to itself. Furthermore, in $M''_{A,B}$ we also remove all nodes that can not reach $v^*$, and all transitions into those nodes. (Technically, some nodes of $M''_{A,B}$ may no longer have full probability on the transitions leaving them, but that is ok for our purposes.)

Now, recall from standard markov chain theory (see, e.g., [Bil95]) that for such a finite (sub)markov chain $M''_{A,B}$, there is a *linear* system of equations $\mathbf{t} = F(\mathbf{t})$, over variables $t_{u,v^*}$, where $u$ is any node of $M''_{A,B}$, and where the coefficients in the linear system $F(\mathbf{t})$ are the probabilties $p'$ on transitions of $M''_{A,B}$. such that the least fixed point solution, $\text{LFP}(F)$, of $\mathbf{t} = F(\mathbf{t})$ assigns to variable $t_{u,v^*}$ the probability that $v^*$ is reachable from $u$. (In particular, one of the linear equations is $t_{v^*,v^*} = 1$.) Moreover, because we have eliminated from $M''_{A,B}$

all nodes that can not reach $v^*$, LFP($F$) is the *unique* solution to this linear system. Thus consider the formula: $\varphi_7(\mathbf{w}, \mathbf{t}) \equiv \mathbf{t} = F(\mathbf{t})$. Thus the quantifier-free formula $\bigwedge_{j=1}^{7} \varphi_j$ has a unique solution in the reals, and the values assigned to variables $t_{(u,v^*)}$ in this solution identify the probability of reaching an accepting SCC from node $u$ in $M'_{A,B}$.

For initial node $u^* = (v_0, \{q_0\})$ of $M'_{A,B}$, and rational $p \in [0,1]$, the following $\boldsymbol{ExTh}(\mathbb{R})$ sentence, $\psi$, is true in $\mathbb{R}$ iff $P_A(L(B)) \geq p$: $\psi \equiv \exists \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}, \mathbf{t} \ \bigwedge_{j=1}^{7} \varphi_j \ \wedge \ t_{u^*,v^*} \geq p$. □

**Theorem 11.** *For a fixed BA, $B$, given a Bd-RMC, $A$, and a rational value $p \in [0,1]$, we can decide whether $P_A(L(B)) \geq p$ in time polynomial in $|A|$.*

*Proof.* (idea) The proof is a modification of Theorem 10. We extend a technique developed in [EY04] to this setting. We use variables only for the entry-exit pairs of $A$ and $A \otimes B'$, express all the other variables as rational functions of those, and then transform the system to a system of constraints of polynomials in a bounded number of variables. □

## References


[AEY01] R. Alur, K. Etessami, and M. Yannakakis. Analysis of recursive state machines. In *Proc. of 13th Int. Conf. on Computer-Aided Verification*, pages 304–313, 2001.

[BGR01] M. Benedikt, P. Godefroid, and T. Reps. Model checking of unrestricted hierarchical state machines. In *Proc. of ICALP'01*, volume 2076 of *LNCS*, pages 652–666, 2001.

[Bil95] P. Billingsley. *Probability and Measure*. J. Wiley and Sons, 3rd edition, 1995.

[BKS04] T. Brázdil, A. Kučera, and O. Stražovský. Decidability of temporal properties of probabilistic pushdown automata. Technical report, 2004. In preparation.

[BPR96] S. Basu, R. Pollack, and M. F. Roy. On the combinatorial and algebraic complexity of quantifier elimination. *J. of the ACM*, 43(6):1002–1045, 1996.

[BR00] T. Ball and S. Rajamani. Bebop: A symbolic model checker for boolean programs. In *SPIN'2000*, volume 1885 of *LNCS*, pages 113–130, 2000.

[Can88] J. Canny. Some algebraic and geometric computations in PSPACE. In *Prof. of 20th ACM STOC*, pages 460–467, 1988.

[CY95] C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *Journal of the ACM*, 42(4):857–907, 1995.

[EHRS00] J. Esparza, D. Hansel, P. Rossmanith, and S. Schwoon. Efficient algorithms for model checking pushdown systems. In *12th CAV*, volume 1855, pages 232–247. Springer, 2000.

[EKM04] Javier Esparza, Antonín Kučera, and Richard Mayr. Model checking probabilistic pushdown automata. In *LICS 2004*, 2004.

[EY04] K. Etessami and M. Yannakakis. Recursive markov chains, stochastic grammars, and monotone systems of non-linear equations. Technical report, School of Informatics, University of Edinburgh, 2004. *Submitted for publication.*

[GGJ76] M. R. Garey, R. L. Graham, and D. S. Johnson. Some NP-complete geometric problems. In *8th ACM Symp. on Theory of Computing*, pages 10–22, 1976.

[Har63] T. E. Harris. *The Theory of Branching Processes*. Springer-Verlag, 1963.

[Kwi03] M. Kwiatkowska. Model checking for probability and time: from theory to practice. In *Proc. 18th IEEE LICS*, pages 351–360, 2003.

[MS99] C. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, 1999.

[PZ93] A. Pnueli and L. D. Zuck. Probabilistic verification. *Inf. and Comp.*, 103(1):1–29, 1993.

[Ren92] J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. parts i,ii, iii. *J. of Symbolic Computation*, pages 255–352, 1992.

[Tiw92] P. Tiwari. A problem that is easier to solve on the unit-cost algebraic ram. *Journal of Complexity*, pages 393–397, 1992.

[Var85] M. Vardi. Automatic verification of probabilistic concurrent finite-state programs. In *Proc. of 26th IEEE Symp. on Foundations of Comp. Sci.*, pages 327–338, 1985.

# A Proofs

### Proof of propostion 1

*Proof.* We split into cases. Case 1: $u$ is any vertex in $Q_{M'_A}$ other than a call port. In this case, $\sum_v p'_{u,v} = \sum_v \frac{p_{u,v}\,\mathrm{ne}(v)}{\mathrm{ne}(u)}$. Note that $\mathrm{ne}(u) = \sum_v p_{u,v}\,\mathrm{ne}(v)$. Hence $\sum p'(u,v) = 1$. Case 2: Suppose $u$ is a call port $u = (b, en)$ in $A_i$, and box $b$ is mapped to component $A_j$. Starting at $u$, the trace will never exit $A_i$ iff either it never exits the box $b$ (which happens with probability $\mathrm{ne}(en)$) or it exits $b$ through some return vertex $v = (b, ex)$ and from there it does not manage to exit $A_i$ (which has probability $q^*_{(en,ex)}\,\mathrm{ne}((b, ex))$). That is, $\mathrm{ne}((b, en)) = \mathrm{ne}(en) + \sum_{ex \in Ex_j} q^*_{(en,ex)}\,\mathrm{ne}((b, ex))$. Dividing both sides by $\mathrm{ne}((b, en))$, we have $1 = \mathrm{ne}(en)/\,\mathrm{ne}(u) + \sum_{ex} q^*_{(en,ex)}\,\mathrm{ne}((b, ex))/\,\mathrm{ne}((b, en))$, which is the sum of the probabilities of the edges out of $(b, u)$. □

### Proof of Lemma 1

*Proof.* Let $D = \rho^{-1}(\star)$. We can partition $D$ according to the first failure. For $t \in D$, let $\rho^H(t) = w_0 w_1 \ldots \in Q^\omega$. Let $i \geq 0$ be the least index such that $w_i \in Q_{H_A}$ but $w_{i+1} \notin Q_{H_A}$ (such an index must exist). We call $w' = w_0 \ldots w_{i+1}$ a *failure prefix*. Let $C(w') = \{w \in \Omega' \mid w = w'w'' \text{ where } w'' \in Q^\omega\}$ be the cylinder at $w'$, inside $\mathcal{F}'$. Let $D[w'] = \rho^H(C(w'))$.

We claim $\mathbf{Pr}_\Omega(D[w']) = 0$ for all such "failure" prefixes, $w'$. (To be completely formal, we have to first argue that $D[w'] \in \mathcal{F}$, but this is not difficult to establish: $D[w']$ can be shown to be a countable union of cylinders in $\mathcal{F}$.)

By definition, $\mathrm{ne}(w_i) > 0$, but $\mathrm{ne}(w_{i+1}) = 0$. We distinguish cases, based on what type of vertex $w_i$ and $w_{i+1}$ are.

Case 1: Suppose $w_i \in Q$ is not a call port. In this case, $(w_i, w_{i+1}) \in E_{H_A}$ reflect an ordinary non-zero edge in the RMC A. A trajectory $t \in D[w']$, is one that reaches $\langle \beta, w_i \rangle$ then moves to $\langle \beta, w_{i+1} \rangle$ and then never exits the component of $w_i$ and $w_{i+1}$, i.e., retains $\beta$ as a prefix of the call stack. (This follows by the definition of $\rho^H$, and the fact that in $H_A$ there are no edges out of exit vertices). Since $\mathrm{ne}(w_{i+1}) = 0$ the probability of such a trajectories $t$ is 0, i.e., $\mathbf{Pr}_\Omega(D[w']) = 0$.

Case 2: $w_i = (b, en)$ is a call port, and $w_{i+1} = (b, ex)$. Thus $(w_i, w_{i+1}) \in E_{H_A}$ is a "summary edge", within some component $A_k$. Again, $\mathrm{ne}(w_i) > 0$, but $\mathrm{ne}(w_{i+1}) = 0$. Any trajectory $t \in D[w']$, reaches $\langle \beta, w_i \rangle$, then sometime later reaches $\langle \beta, w_{i+1} \rangle$, having always retained $\beta$ as a prefix of the call stack in between, and thereafter it never exits the component of $w_i$ and $w_{i+1}$. (Again, similar to case 1, this follows by definition of $\rho^H$, and $H_A$.) But since $\mathrm{ne}(w_{i+1}) = 0$, this $\mathbf{Pr}_\Omega(D[w']) = 0$.

Case 3: $w_i = (b, en)$ and $w_{i+1} = en$. In other words, $(w_i, w_{i+1})$ is an edge of $E_{H_A}$ were we move from a call port to the corresponding entry $en$ of the component $A_j$, where $Y(b) = j$. Thus a trajectory $t \in D[w']$ enters component $A_j$ at entry $en$, on step i+1, and never exits this component thereafter. Note again, however, that $\mathrm{ne}(w_{i+1}) = 0$. Thus, $\mathbf{Pr}_\Omega(D[w']) = 0$.

Now note that $D = \bigcup_{w'} D[w']$, where the union is over all failure prefixes, $w' \in Q^*$. Note that this is a countable union of sets, each having probability 0, thus $\mathbf{Pr}_\Omega(D) = 0$. □

### Proof of Lemma 2

*Proof.* It suffices, by standard facts about probability measure, to prove the claim for cylinders $C(w') \in \Omega'$, where $w' = w_0, \ldots w_k$. We use induction on $k$. The base case ($k = 0$) follows from Lemma 1. Namely, $C(\epsilon) = \Omega'$, and $\rho^{-1}(\Omega') = \Omega \setminus \rho^{-1}(\star)$. Thus $\mathbf{Pr}_\Omega(\rho^{-1}(\Omega')) = 1 - \mathbf{Pr}_\Omega(\rho^{-1}(\star)) = 1$.

For the induction step, suppose that the claim hold for the prefix $w' = w_0 w_1 \ldots w_k$. Let $D[w'] = \rho^{-1}(C(w'))$. Define the event $J_{(i,y)} \in \mathcal{F}$ to be $J_{(i,y)} = \{t \in \Omega \mid \rho(t) = w_0 \ldots w_i \ldots \text{, and } w_i = y\}$.

Note that, by the definition of conditional probability, $\mathbf{Pr}_\Omega(D[w'w_{k+1}]) = \mathbf{Pr}_\Omega(D[w'])\, \mathbf{Pr}_\Omega(J_{k+1,w_{k+1}} \mid D[w'])$.

We want to show that $\mathbf{Pr}_\Omega(D[w'w_{k+1}]) = \mathbf{Pr}_{\Omega'}(C(w'w_{k+1}))$. We distinguish three cases, based on what type of edge $(w_k, w_{k+1})$ is in $H_A$, as in the proof of Lemma 1.

Case 1: $w_k$ is not a call port. Thus $(w_k, w_{k+1}) \in E_{H_A}$ is an ordinary edge, inside some component $A_i$ of $A$. Consider the trajectories $t \in D[w'w_{k+1}]$. After some prefix, the trajectory arrives at a vertex $\langle \beta, w_k \rangle$, and subsequently never reaches an exit, i.e., retains $\beta$ as a prefix of the call stack. The conditional probability $\mathbf{Pr}_\Omega(J_{k+1,w_{k+1}} \mid D[w'])$, is the probability that the $(k+1)$-st step of $\rho(t)$ is $w_{k+1}$, given that the prefix of $\rho(t)$ is $w_0 w_1, \ldots w_k$. Note that this conditional probability is independent of the call stack $\beta$, and that this process has the markov property, so that it is also indepdenent of how we arrive at $w_k$. Let $\mathrm{NE}(u) \in \mathcal{F}$ be the event that, starting at a node $\langle \beta, u \rangle$, we will never reach an exit. i.e., $\beta \in B^+$ will forever remain on the call stack.

Since $w_k$ is not a call port, and using the markovian property, we seen that:

$$\begin{aligned}
\mathbf{Pr}_\Omega(J_{k+1,w_{k+1}} \mid D[w']) &= \mathbf{Pr}_\Omega(J_{k+1,w_{k+1}} \mid J_{k,w_k}) \\
&= \mathbf{Pr}_\Omega(J_{1,w_{k+1}} \mid J_{0,w_k}), \text{ (now assuming } p_{\mathrm{init}}(\langle \epsilon, w_k \rangle) = 1) \\
&= \mathbf{Pr}_\Omega(J_{1,w_{k+1}} \cap \mathrm{NE}(w_k)) / \mathbf{Pr}_\Omega(\mathrm{NE}(w_k)) \\
&= \mathbf{Pr}_\Omega(J_{1,w_{k+1}} \cap \mathrm{NE}(w_{k+1})) / \operatorname{ne}(w_k) \\
&= \mathbf{Pr}_\Omega(J_{1,w_{k+1}})\, \mathbf{Pr}_\Omega(\mathrm{NE}(w_{k+1})) / \operatorname{ne}(w_k) \\
&= p_{w_k,w_{k+1}} \operatorname{ne}(w_{k+1}) / \operatorname{ne}(w_k)
\end{aligned}$$

Therefore, $\mathbf{Pr}_\Omega(D[w'w_{k+1}]) = \mathbf{Pr}_\Omega(D[w'])p_{w_k,w_{k+1}} \operatorname{ne}(w_{k+1}) / \operatorname{ne} w_k$. By the induction hypothesis, and the construction of $M'_A$, $\mathbf{Pr}_{\Omega'}(C(w'w_{k+1})) = \mathbf{Pr}_{\Omega'}(C(w'))p'_{w_k,w_{k+1}}$
$= \mathbf{Pr}_\Omega(D[w'])p_{w_k,w_{k+1}} \operatorname{ne}(w_{k+1}) / \operatorname{ne} w_k = \mathbf{Pr}_\Omega(D[w'w_{k+1}])$.

Cases 2: $w_k = (b, en)$ is a call port, and $w_{k+1} = (b, ex)$ is a return port. In this case, similar to case 1, we have:

$$\begin{aligned}
\mathbf{Pr}_\Omega(J_{k+1,w_{k+1}} \mid D[w']) &= \mathbf{Pr}_\Omega(J_{1,w_{k+1}} \cap \mathrm{NE}(w_{k+1})) / \operatorname{ne}(w_k), \quad (\text{assuming } p_{\mathrm{init}}(\langle \epsilon, w_k \rangle) = 1) \\
&= \mathbf{Pr}_\Omega(J_{1,w_{k+1}}) \operatorname{ne}(w_{k+1}) / \operatorname{ne}(w_k) \\
&= q^*_{(en,ex)} \operatorname{ne}(w_{k+1}) / \operatorname{ne}(w_k)
\end{aligned}$$

Again, $\mathbf{Pr}_\Omega(D[w'w_{k+1}]) = \mathbf{Pr}_\Omega(D[w'])q^*_{(w_k,w_{k+1})} \operatorname{ne}(w_{k+1}) / \operatorname{ne}(w_k)$, and by induction, $\mathbf{Pr}_{\Omega'}(C(w'w_{k+1})) = \mathbf{Pr}_{\Omega'}(C(w'))p'_{w_k,w_{k+1}} = \mathbf{Pr}_\Omega(D[w'])q^*_{(w_k,w_{k+1})} \operatorname{ne}(w_{k+1}) / \operatorname{ne}(w_k) = \mathbf{Pr}_\Omega(D[w'w_{k+1}])$.

Cases 3: $w_k = (b, en)$ is a call port, and $w_{k+1} = en$ is the corresponding entry. In this case,

$$\begin{aligned}\mathbf{Pr}_\Omega(J_{k+1,w_{k+1}} \mid D[w']) &= \mathbf{Pr}_\Omega(J_{1,w_{k+1}} \mid J_{0,w_k}) \\ &= \mathbf{Pr}_\Omega(J_{1,w_{k+1}} \cap \mathrm{NE}(w_k))/\,\mathbf{Pr}_\Omega(\mathrm{NE}(w_k)),\ (\text{assuming } p_{\mathrm{init}}(\langle \epsilon, w_k\rangle) = 1) \\ &= \mathbf{Pr}_\Omega(J_{1,w_{k+1}})/\,\mathrm{ne}(w_k),\ (\text{because } \mathrm{NE}(w_k) \subseteq J_{1,w_{k+1}}) \\ &= \mathbf{Pr}_\Omega(\mathrm{NE}(w_{k+1}))/\,\mathrm{ne}(w_k) = \mathrm{ne}(w_{k+1})/\,\mathrm{ne}(w_k)\end{aligned}$$

Again, $\mathbf{Pr}_\Omega(D[w'w_{k+1}]) = \mathbf{Pr}_\Omega(D[w'])\,\mathrm{ne}(w_{k+1})/\,\mathrm{ne}(w_k)$, and $\mathbf{Pr}_{\Omega'}(C(w'w_{k+1})) = \mathbf{Pr}_{\Omega'}(C(w'))p'_{w_k,w_{k+1}} = \mathbf{Pr}_\Omega(D[w'])\,\mathrm{ne}(w_{k+1})/\,\mathrm{ne}(w_k) = \mathbf{Pr}_\Omega(D[w'w_{k+1}])$. □

**Proof of Lemma 3**

*Proof.* We construct the accepting run $r$ of $B$ and run $r'$ of $M'_A \otimes B$ one segment at a time. Suppose that $(v, q)$ is special of type 1. Then $\gamma(v, q)$ corresponds to a path in $D(v, q)$ (and $M'_{A,B}$) from $(v, \{q\})$ to a node of a bottom SCC $C$ that contains a state $(v, T)$ with $q \in T$. Consider a trajectory $t$ of the RMC that starts with $\gamma$ and the corresponding a trajectory $t'$ of $M'_{A,B}$ starting from $(v, \{q\})$. With probability 1, $t'$ exists (i.e. $t$ maps to a trajectory of of $M'_{A,B}$ starting from $(v, \{q\})$), and $t'$ goes to the bottom SCC $C$ and visits infinitely often all the states of $C$. For every visit to the state $(v, T)$ there is a nonzero probability that in the following steps the trajectory $t'$ will perform the transitions of $\gamma(v, q)$. Hence, with probability 1, at some finite step $i$, $t'$ visits $(v, T)$ and in the following steps the trajectory $t$ performs $\gamma(v, q)$. Let $i$ be the first time this happens. Since $q \in T$, the prefix of $t$ up to step $i$ has a corresponding run in $B$ from $q$ to $q$ and in $M'_A \otimes B$ from $(v, q)$ to $(v, q)$. This constitutes the first segment of the constructed run $r$.

At step $i$, the trajectory $t$ is at vertex $v$ and the suffix from this point on starts again with the sequence $\gamma(v, q)$ of transitions. Since we have a Markov process we can repeat the argument for the remainder of $T$ and construct the second and subsequent segments of $r$. In general, if $E_k$ denotes the event that the procedure succeeds in constructing $k$ segments, then the probability of $E_{k+1}$ conditioned on $E_k$ is 1. Therefore, the probability of $\cap_k E_k$ is also 1, and thus the required accepting run $r$ will be constructed with probability 1.

Suppose that $(v, q)$ is special of type 2 and let $vq \to (v', \{q'\})$ be the first edge (an accepting edge) in $D(v, q)$ of the path corresponding to $\gamma(v, q)$ that leads from the root $vq$ to the bottom SCC $C$ that contains $(v, T)$ with $q \in T$. Let $\alpha$ be the label of this edge; then $\gamma(v, q) = \alpha\beta$ for some $\beta$. The argument is similar to the case of type 1. Consider a trajectory $t$ of the RMC starting from $v$ with the transitions of $\gamma(v, q)$, and let $t = \alpha\tau$. After the prefix $\alpha$, the trajectory $t$ is at vertex $v'$ (with empty stack, i.e the chain $M_A$ is at vertex $\langle \epsilon, v'\rangle$). The remaining trajectory $\tau$ starts with $\beta$. With probability 1, $\tau$ maps to a trajectory $\tau'$ of $M'_{A,B}$ starting from state $(v', \{q'\})$, and since $\tau$ starts with $\beta$, $\tau'$ goes to the bottom SCC $C$. As in case 1, the trajectory hits with probability 1 infinitely often all the states of $C$, and furthermore there is a finite time $i$ at which it reaches $(v, T)$ and the following suffix of $t$ starts again with $\gamma(v, q)$. We can map now the prefix of $t$ up to step $i$ to a run of $B$ from $q$ that goes first to $q'$ passing on the way through an accepting state of $B$ (this path corresponds to the prefix $\alpha$) and then continues and reaches state $q$ again at time $i$; the corresponding path of $M'_A \otimes B$ follows first the edge to $(v', q')$ and then goes on to reach $(v, q)$. This constitutes the first segment of the constructed run $r$. As in case 2, we can then repeat the process to construct the subsequent segments, and the process will succeed with probability 1. □

**Proof of Lemma 4**

*Proof.* Suppose that an accepting state $(v, q)$ is not special. With probability 1, a trajectory $t$ of the RMC that starts at $v$ corresponds to a trajectory $t'$ of $M'_{A,B}$ that starts at $(v, \{q\})$ and reaches a bottom SCC $C$ of $M'_{A,B}$ (and of $D(v, q)$). Since $(v, q)$ is not special, there is no state $(v, T)$ of $C$ with $q \in T$. Therefore, every run of $M'_A \otimes B$ starting at $(v, q)$ that corresponds to $t$ does not visit $(v, q)$ after $t'$ reaches $C$, hence, repeats $(v, q)$ only finitely often.

Suppose that $t$ starts at a vertex $u \in M'_A$ and corresponds to a run of $M'_A \otimes B$ starting at a node $(u, p)$ that visits $(v, q)$ infinitely often. Let $i$ be the first step at which the run visits $(v, q)$. The suffix of $t$ from this point on corresponds to a run of $M'_A \otimes B$ starting from $(v, q)$ that visits $(v, q)$ infinitely often. By our above argument, the probability that a trajectory of the RMC has this property is equal to 0, and by the Markov property it follows that the probability that $t$ has such a suffix is also 0.

Consider an accepting edge $(v, q) \to (v', q')$ and suppose that $(v, q)$ is not special. The graph $D(v, q)$ contains an edge $vq \to (v', \{q'\})$. Since $(v, q)$ is not special, no bottom SCC contains any state $(v, T)$ with $q \in T$. Suppose that a trajectory $t$ of the RMC starting at $v'$ corresponds to a run of $M'_A \otimes B$ starting at $(v', q')$ that traverses the edge $(v, q) \to (v', q')$ infinitely often. With probability 1, $t$ corresponds to a trajectory of $M'_{A,B}$ starting from $(v', \{q'\})$ that reaches a bottom SCC $C$ of $D(q, v)$. Since no such bottom SCC contains a state $(v, T)$ with $q \in T$ it follows that every run of $M'_A \otimes B$ from $(v', q')$ that corresponds to $t$ does not visit $(v, q)$ after some point, and hence does not traverse the edge.

Suppose that a trajectory $t$ starts at a vertex $u \in M'_A$ and corresponds to a run of $M'_A \otimes B$ starting at a node $(u, p)$ that visits the edge $(v, q) \to (v', q')$ infinitely often. The argument is similar to the type 1 case. Consider the first time that the edge is traversed and write $t$ as $t = \alpha\tau$, where the prefix $\alpha$ corresponds to the run from $(u, p)$ to $(v', q')$ ending with the traversal of the edge. The suffix $\tau$ corresponds to a run starting from $(v', q')$ that repeats the edge infinitely often. ¿From the above argument, the probability that a trajectory $\tau$ of the RMC starting at $v'$ has this property is 0, hence the probability that $t$ has such a suffix is also 0. □

**Proof of Theorem 6**

*Proof.* With probability 1 a trajectory $t$ of the RMC maps to a trajectory $t' = \sigma(t)$ of $M'_{A,B}$ which reaches a bottom SCC $C$.

If $C$ is not accepting then there is no special pair $(v, q)$ such that $C$ contains a state $(v, T)$ with $q \in T$. Then every run of $M'_A \otimes B$ starting from $(v_0, q_0)$ that corresponds to $t$ visits special nodes only finitely many times. It follows that with probability 1 $t$ is not accepted by $B$.

If $C$ is an accepting bottom SCC, then there is a special pair $(v, q)$ such that $C$ contains a state $(v, T)$ with $q \in T$. The trajectory will visit $(v, T)$ infinitely often, and at every visit there is nonzero probability that the RMC will execute next the sequence $\gamma(v, q)$. Hence, with probability 1 this will occur at some finite point. Then the trajectory $t$ will be accepted by $B$ with probability 1. □

**Proof of Proposition 3**

*Proof.* Suppose that a trajectory $t$ of the RMC starting at $v_0$ is accepted by $B$ (starting at $q_0$), With probability 1, $t$ has a corresponding run in $M'_A \otimes B$ starting at $(v_0, q_0)$ that repeats

infinitely often some accepting state $(v, q)$ or some accepting edge $(v, q) \to (v', q')$. It follows from the preceding lemma that $(v, q)$ must be special, and obviously $(v_0, q_0)$ can reach $(v, q)$.

Conversely, suppose that $(v_0, q_0)$ can reach the special pair $(v, q)$ in the graph $M'_A \otimes B$ and let $\alpha$ be the label of such a path from $(v_0, q_0)$ to $(v, q)$. With nonzero probability, the RMC will execute first the sequence of transitions $\alpha\gamma(v, q)$. If this occurs, then from that point on with probability 1 the trajectory will correspond to an accepting run of $B$. □

**Proof of Theorem 8**

*Proof.* We begin by proving hardness for deciding whether $P_A(L(B)) = 1$, where both $A$ and $B$ are part of the input. The case where $A$ is fixed, and the case for qualitative emptiness, $P_A(L(B)) \stackrel{?}{=} 0$, are variations on the same proof, and we sketch them at the end.

The reduction is from the acceptance problem for alternating linear space bounded Turing machines. As is well known, $\text{ASPACE}(S(n)) = \cup_{c>0}\text{DTIME}(c^{S(n)})$. There is a fixed linear space bounded alternating Turing machine, $T$, such that the problem of deciding whether $T$ acccepts a given input of length $n$ is EXPTIME-complete. We can assume wlog that $T$ has one tape, and uses space $n$. The tape contains initially the given input $x$. Recall that an alternating TM has four types of states: existential, universal, accepting and rejecting. We assume wlog that the TM has two possible moves from each existential and universal state, and it halts when it is in an accepting or rejecting state. Let $\Gamma$ be the tape alphabet, $Q$ the set of states and $\Delta = \Gamma \cup (Q \times \Gamma)$ the extended tape alphabet. A configuration of the TM is expressed as usual as a string of length $n$ where the $i$th symbol is $(q, X) \in (Q \times \Gamma)$ (we will usually write $qX$ instead of $(q, X)$) if the head is on the tape cell $i$, the state is $q$ and the tape symbol is $X$, and otherwise the $i$th symbol is the tape symbol $X$ in cell $i$. The type of a configuration (existential, universal etc) is determined by the type of the state. A *computation* is a sequence of configurations starting from the initial one, according to the transition rules of the TM. We assume wlog that all computations of the TM halt.

There is a natural game associated with an alternating TM between two players, an existential player E and a universal player U. The positions of the game correspond to the configurations. Player E moves at the existential configurations and player U at the universal ones. Accepting configurations are winning positions for player E, and rejecting configurations for player U. An input $x$ is accepted by the TM iff the existential player E has a winning strategy from the initial configuration corresponding to $x$.

We will construct a RMC, $A$, and a BA, $B$, so that $A$ satisfies $B$ with probability 1 iff $x$ is not accepted by $T$, i.e. E does not have a winning strategy.

Let us first mention that the only thing that will matter about $A$, is its "structure", i.e., which edges have non-zero probability. We thus describe these edges without defining the probabilities explicitly: any probabilites that sum to 1 will do.

The RMC $A$ has an initial component $C_0$ and a component $C(q, X)$ for each state $q \in Q$ and tape symbol $x \in \Gamma$. The automaton $B$ has an initial state $s_0$, a final state $f$ which is the only accepting state, and a state $(\delta, i)$ for each $\delta \in \Delta$, and $i = 1, \ldots, n$. The alphabet of $B$ is the vertex set of $A$.

Let $q_0$ be the initial state of the TM $T$, and let $x = x_1 \cdots x_n$ be the input. Component $C_0$ of $A$ has an edge from its entry to a node $u_0$, an edge from $u_0$ to a box that is mapped to $C(q_0, x_1)$ and an edge from the exit of the box to an absorbing node $v_0$ that has a self-loop.

Component $C(q, X)$, where $q$ is an existential state and $X \in \Gamma$, is structured as follows. Suppose that the two moves of the TM when it is in state $q$ and reads $X$ are $(p_k, Y_k, D_k), k =$

$1, 2$, where $p_k \in Q$ is the next state, $Y_k$ is the symbol written over $X$, and $D_k = L/R$ (left/right) is the direction of the head movement. For each $i = 1, .., n$, $k = 1, 2$, and $Z \in \Gamma$, the component has a set of nodes $u[q, X, i, k, Z]$, $v[q, X, i, k, Z]$, and a set of boxes $b[q, X, i, k, Z]$, each mapped to component $C(p_k, Z)$. The entry of the component $C(q, X)$ has edges to each of the nodes $u[q, X, i, k, Z]$, every node $u[q, X, i, k, Z]$ has an edge to the call port of the corresponding box $b[q, X, i, k, Z]$, the return port of each such box has an edge to the corresponding node $v[q, X, i, k, Z]$, and each of these nodes has an edge to the exit of the component.

Component $C(q, X)$, where $q$ is a universal state and $X \in \Gamma$, is structured as follows. Let again the two moves of the TM for $q$ and $X$ be $(p_k, Y_k, D_k), k = 1, 2$. For each $i = 1, .., n$, $k = 1, 2$, and $Z \in \Gamma$, the component has again a set of nodes $u[q, X, i, k, Z]$, $v[q, X, i, k, Z]$, and a set of boxes $b[q, X, i, k, Z]$ mapped to $C(p_k, Z)$, and has in addition one more node $w[q, X]$. The entry of the component $C(q, X)$ has edges to each of the nodes $u[q, X, i, 1, Z]$, every node $u[q, X, i, 1, Z]$ has an edge to the call port of the corresponding box $b[q, X, i, 1, Z]$, the return port of each such box has an edge to the corresponding node $v[q, X, i, 1, Z]$, and each of these has an edge to node $w[q, X]$. Node $w[q, X]$ has edges to all the nodes $u[q, X, i, 2, Z]$, every node $u[q, X, i, 2, Z]$ has an edge to the call port of the corresponding box $b[q, X, i, 2, Z]$, the return port of each such box has an edge to the corresponding node $v[q, X, i, 2, Z]$, and each of these has an edge to the exit of the component.

Component $C(q, X)$, where $q$ is a halting (accepting or rejecting) state and $X \in \Gamma$ has an edge from its entry to a node $u[q, X]$ and from $u[q, X]$ to the exit of the component.

The transitions of the automaton $B$ are as follows. The initial state $s_0$ of $B$ transitions on input $u_0$ to the set of states $\{(q_0 x_1, 1), (x_2, 2), \ldots, (x_n, n)\}$. There are no other transitions out of $s_0$. The final state $f$ transitions to itself on every input.

Let $q$ be an existential or universal state and suppose that the two moves of the TM when it is in state $q$ and reads $X$ are $(p_k, Y_k, D_k), k = 1, 2$. On input $u[q, X, i, k, Z]$, a state $(\delta, j)$ of $B$ has exactly one transition, as follows: If $j = i$ and $\delta \neq qX$ then it transitions to $f$; else, if $j = i$ and $\delta = qX$ then it transitions to state $(Y_k, i)$; else, if $((j = i + 1$ and $D_k = R)$ or $(j = i - 1$ and $D_k = L))$ and $\delta = Z$ then it transitions to $(p_k Z, j)$; else, if $((j = i + 1$ and $D_k = R)$ or $(j = i - 1$ and $D_k = L))$ and $\delta \neq Z$ then it transitions to $f$; else, it transitions to itself, $(\delta, j)$. On input $v[q, X, i, k, Z]$, a state $(\delta, j)$ of $B$ has the following transition: If $j = i$ then it transitions to $(qX, i)$; else, if $((j = i + 1$ and $D_k = R)$ or $(j = i - 1$ and $D_k = L))$ then it transitions to $(Z, j)$; else, it transitions to itself, $(\delta, j)$. All states have a self-loop on input $w[q, X]$, $v_0$, as well as for all the vertices that are entries and exits of boxes.

Let $q$ be a halting state of the TM. On input $u[q, X]$, a state $(\delta, j)$ of $B$ transitions to itself if $\delta \in \Gamma$ or ($\delta = qX$ and $q$ is accepting), and it transitions to $f$ otherwise.

This concludes the definition of the RMC $A$ and the BA automaton $B$. Note that $A$ has a bounded number of components (independent of the length of the input $x$), and every component has one entry and one exit. Note also that all the transitions of $B$ are deterministic except for the transition of the initial state $s_0$ on input $u_0$.

Consider a path of the RMC, and look at the corresponding set $P$ of states of $B$ at each step. At $u_0$, the set $P$ contains one state $(\delta, i)$ for each $i = 1, \ldots, n$ corresponding to the initial configuration of the TM. ¿From then on, it is easy to check that $P$ always contains *at most* one state $(\delta, i)$ for each $i$, and either these states form a configuration of the TM or $P$ contains $f$. Once $f$ is included in $P$, then it will stay there forever and any continuation of the path will be accepted by $B$.

Call a path of the RMC *valid* if the set $P$ at the end (and during the path) does not contain $f$. Consider the game tree $G$ of the game corresponding to the TM on the given input $x$: The nodes of the tree are the configurations reached by the TM in its computation, the root is the initial configuration, the children of each node are the two successor configurations, and the leaves correspond to halting configurations. An existential strategy corresponds to a subtree $G_E$ of $G$ that contains one child of each (reachable) existential configuration (nodes that are not reachable any more from the root are not included in $G_E$). We consider the two children of each node as being ordered according to the indexing ($k = 1, 2$) of the two moves of the configuration. We claim that every valid path of the RMC corresponds to a prefix of the depth-first-search traversal of an existential game tree $G_E$, where all the leaves in the prefix are accepting; and conversely every such prefix of a DFS traversal corresponds to a valid path. Note that when a valid path is at the entry of an existential component $C(q, X)$, in order for it to continue to be valid it must move to a node $u[q, X, i, k, Z]$ such that $i$ is the current position of the head, $q$ and $X$ must be the current state and symbol at cell $i$, and $Z$ must be the symbol in the tape cell where the head moves next according to move $k = 1$ or 2 of the TM. That is, there are precisely two valid choices corresponding to the two possible moves of the existential player. The transitions of $B$ are defined so that the states of the new current set $P$ form the next configuration as the path of the RMC moves to the box corresponding to the move of the TM. When the path exits the box, if it is still valid, then the set $P$ is the same as when the path entered the box. After the node $v[q, X, i, k, Z]$, the set $P$ is updated to restore the configuration as it was when the component $C(q, x)$ was called. For a universal component $C(q, X)$ there is only one correct choice if the path is to remain valid. If the path exits the component remaining valid, it means that it never went through a rejecting component, i.e., the corresponding subtree of $G_E$ that was traversed has only accepting leaves. If $x$ is accepted by the TM, then the existential player has a winning strategy, hence there is a valid path of the RMC that reaches node $v_0$ of $C_0$ and stays there forever. Thus, with positive probability the RMC follows this path which is not accepted by $B$. On the other hand, if $x$ is not accepted by the TM, then every path becomes eventually invalid (either because it reaches a rejecting component or because one of its transitions does not correspond to a TM move), hence is accepted by $B$; thus the acceptance probability is 1.

We are done with the proof that checking $P_A(L(B)) = 1$ is EXPTIME-hard. By Theorem 7, the problem is also EXPTIME-complete.

We now sketch how a variation of the same proof shows that probabilistic emptiness ($P_A(L(B)) > 0$?) is also EXPTIME-complete.

For each component except $C_0$, add a direct path from entry to exit $en \to r \to ex$ through a new node $r$ where the first edge has probability $> 1/2$. Every state of the BA, $B$, goes to $f$ on these intermediate nodes. (The purpose of these paths is to make sure that every component exits with probability 1 - but these are not valid paths). Remove the self loop of $v_0$, add new nodes $y_0$,$z_0$ to $C_0$, and edges $v_0 \to y_0 \to z_0 \to u_0$ with probability 1. Also add a new state $g$ to $B$ which is the only accepting state ($f$ is not accepting anymore). On input $y_0$, all states of $B$ die except for $f$ that goes to $g$. On $z_0$, $g$ goes to the initial state $s_0$.

By the previous proof , (1) if input $x$ is accepted by the ATM, the old RMC had a path $p$ from the initial vertex to $v_0$ such that the corresponding set of states of the BA at the end (for all possible runs) did not include $f$. (2) If $x$ is not accepted by the ATM, then for every trajectory of the old RMC, the BA has a run that gets to $f$.

Because of the new paths to the exits that we have added, every component exits with probability 1 (this requires a proof, which we omit, but follows from basic facts about RMCs,

see [EY04]). hence, infinitely often (i.o.), the trajectory will go to $u_0$, traverse a path, come out at $v_0$, go to $y_0$,$z_0$, back to $u_0$, and again all over. If the state set of the BA includes $f$ when the path arrives at $v_0$, then it will go next to $g$, then reset to the initial state and start again. Therefore, if $x$ is not accepted by the ATM, this will happen every time, hence $g$ will appear i.o. and the probability of acceptance $P_A(L(B) = 1$.

If $x$ is accepted by the ATM, and in some iteration the RMC follows the path $p$ as above then the BA will die when the path reaches $y_0$. Every time the process returns to $u_0$ and tries again, there is positive probability that it will follow the path $p$, so eventually this will happen at some point with probability 1. When it happens, the BA will die and hence will not accept the trajectory. Thus, in this case $P_A(L(B)) = 0$.

Next, we briefly sketch how we actually only need a fixed RMC, whose size does not depend on the size of the input tape of the ATM. Here is the modification Drop the index $i$ from the $u$ and $v$ nodes of $A$, and add a self loop to these nodes. Basically, the RMC is going to guess what is the correct $i$ which will be the number of times it loops at the node $u$ (and $v$). The BA states keep track of how many times the RMC goes around the loop at $u[...]$. In other words, the BA states have now, besides extended tape symbol and cell number another counter ($\leq n$) - if the counter exceeds n then transition to $f$. In fact if the BA is at state $(qX, i, j)$ and the counter $j$ passes $i$ (without leaving $u[]$) or it leaves before $i$, then the state goes to $f$ and the game is in effect over. If it leaves exactly at the correct time, then $(qX, i, i)$ makes the right transition to $(Y, i, 0)$. For the other states $(\delta, i, j)$, first if $\delta$ has a state and is not $qX$ then go to $f$ right away. If state $= (\delta, k, i)$ when the RMC moves out of $u[]$ and $k \neq i$ the state assumes that the RMC moved at the right time (i.e. tape head is at cell $i$) and acts accordingly: for example if the head is supposed to move left and new state $= p$, new symbol (in new position)$= Z$, then $(\delta, k, i)$ transitions to $(\delta, k, 0)$ if $k \neq i - 1$, to $f$ if $k = i - 1$ but $\delta \neq Z$, and to $(pZ, k, 0)$ otherwise. The moves at $v[]$ that restore the state are similar.

□

**Proof of the Unique Fixed Point Theorem**

**Lemma 5.** *Let $G$ be a finite Markov chain on state set $V$, and let $D$ be a subset of states such that each state $u \in D$ has a transition with probability at least $p > 0$ to a dead (absorbing) state $d$. Then for every positive integer $N$, the probability that, a trajectory of $M$ starting at any state visits at least $N$ times a state of $D$ and is not absorbed in the dead state $d$, is at most $(1 - p)^N$.*

*Proof.* Every time the chain visits a state in $D$, with probability at least $p$ it transitions to $d$, and survives with probability at most $1 - p$ (continues without being absorbed in $d$). Hence if it visits $D$ $N$ times, the probability that it survives is $\leq (1 - p)^N$. More formally, we prove by induction on $N$. The basis, $N = 0$, is trivial. Suppose the claim holds for $N - 1$. Let $E_i(s)$ be the event that $G$ starting from state $s$ survives $i$ visits to $D$ Then $P(E_N(s)) = \sum_{u \in D} P(u \text{ is the first visited state of } D)P(E_N(u))$. Now, $P(E_N(u)) = \sum_{v \neq d} p_{u,v} P(E_{N-1}(v))$. By induction $P(E_{N-1}(v)) \leq (1 - p)^{N-1}$ for all $v$, and $\sum_{v \neq d} p_{u,v} \leq 1 - p$ since $u \in D$. Therefore, $P(E_N(u)) \leq (1 - p)^N$, and hence $P(E_N(s)) \leq (1 - p)^N$. □

Consider now a RMC $A$. First, we can determine in (polynomial time) the vertex-exit pairs $(u, ex)$ for each component such that the probability $q^*_{(u,ex)} = 0$. Introduce variables $x_{u,ex}$ only for the remaining pairs. Note that if a vertex $u$ cannot exit its component, i.e. $q^*_{(u,ex)} = 0$ for all $ex$ then there is no variable involving $u$. Consider the set of fixpoint equations $x = P(x)$

(we drop the terms that involved "missing" variables). The least fixpoint $q^*$ is the true vector of probabilitities of each vertex $u$ reaching exit $ex$ (with empty stack). Let us say that a vertex $u$ is *deficient* if $\sum_{ex} q^*_{(u,ex)} < 1$, i.e. $ne(u) > 0$; otherwise $u$ is *full*. Note that by the qualitative analysis, we can determine which vertices are deficient and which are full in PSPACE.

**Theorem 12.** *(Unique Fixed Point Theorem) The set of equations $x = P(x)$ has a unique fixed point that satisfies $\sum_{ex} x_{(u,ex)} < 1$ for every deficient vertex $u$, and $\sum_{ex} x_{(u,ex)} \leq 1$ for every other vertex $u$. (This, of course, is $q^* = \mathrm{LFP}(P)$.)*

*Proof.* Suppose that there is another fixpoint $y$, besides $q^*$, that satisfies the constraints on $\sum_{ex} x_{(u,ex)}$. Since $q^*$ is the least fixpoint we have $q^* \leq y$. If $u$ is a full vertex then $\sum_{ex} y_{(u,ex)} \leq 1 = \sum_{ex} q^*_{(u,ex)}$ and $q^* \leq y$ imply that $y_{(u,ex)} = q^*_{(u,ex)}$ for every $ex$.

We will show below that $y$ agrees with $q^*$ also on the deficient vertices. Let $(u, ex)$ be a pair such that $y_{(u,ex)} > q^*_{(u,ex)}$. We will derive a contradiction.

Let $x_{(u,ex)} = f_1(x)$ be the equation for variable $x_{(u,ex)}$ in the system $x = P(x)$. The right hand side $f_1(x)$ is a sum of monomials and possibly a constant term. If $u$ is not a call port then each monomial is of the form $p_{u,v}x_{(v,ex)}$, where $v$ is a successor of $u$, and if $u = (b, en)$ is a call port of a box $b$ then each monomial is of the form $x_{en,ex'}x_{(b,ex'),ex}$ where $ex'$ is an exit of the component corresponding to box $b$; in the latter case we consider the variables of the monomial as ordered. We will rewrite iteratively the right hand side $f_1(x)$ as follows. In the $i$th iteration we have an expression $f_i(x)$ which is the sum of a constant term (possibly 0) and of a set of *ordered* monomials; i.e. each monomial has a constant coefficient and the product of a sequence of variables (with possible repetitions allowed) in a specific order. We take every nonconstant monomial and replace the leftmost variable of the monomial by the right hand side of its equation in the system $x = P(x)$. We combine like terms (treated again as ordered monomials) and let $f_{i+1}(x)$ be the resulting expression.

Observe first that both fixpoints, $q^*$ and $y$ satisfy the equation $x_{(u,ex)} = f_n(x)$ for all $n$. Second, we claim that $f_n(x)$ is related to the (infinite) Markov chain $M_A$ corresponding to the RMC $A$ in the following way. Let $Z_n$ be the state at time $n$ of the chain $M_A$ with initial state $\langle \epsilon, u \rangle$. Note that if the chain hits $\langle \epsilon, ex \rangle$ at some time $t$ then it stays there forerever, i.e. $Z_n = \langle \epsilon, ex \rangle$ for all $n \geq t$.

**Lemma 6.** *The constant term of $f_n(x)$ is equal to $Prob(Z_n = \langle \epsilon, ex \rangle)$. Furthermore, for each state $\langle \beta, v \rangle$ where $\beta = b_1 \dots b_j$ is a sequence of boxes and $v$ is a vertex such that $Prob(Z_n = \langle \beta, v \rangle) > 0$, and for every sequence $\gamma = w_1, \dots, w_j$ of exits of the components corresponding to the boxes such that the variables with indices $(v, w_j), ((b_j, w_j), w_{j-1}), \dots ((b_2, w_2), w_1), ((b_1, w_1), ex)$ exist, the expression $f_n(x)$ has an ordered monomial $Prob(Z_n = \langle \beta, v \rangle)x_{(v,w_j)}x_{((b_j,w_j),w_{j-1})} \cdots x_{((b_2,w_2),w_1)}x_{((b_1,w_1),ex)}$. If $\beta$ is the empty string $\epsilon$ then the monomial is simply $Prob(Z_n = \langle \epsilon, v \rangle)x_{(v,ex)}$. These are all the monomials of $f_n(x)$*

*Proof.* By induction, starting with $f_0(x) = x_{(u,ex)}$. The basis is trivial: $Prob(Z_0 = \langle \epsilon, u \rangle) = 1$. For the induction step, consider a monomial of $f_n(x)$ corresponding to the state $\langle \beta, v \rangle$ and a sequence $\gamma$ of exits to the boxes (if $\beta$ is nonempty). If $v$ is an exit and $\beta = \epsilon$, then $v$ must be $ex$ (because for other exits the variable does not exist since it is 0), and $x_{v,ex}$ will be replaced by 1, increasing the constant term. If $v$ is an exit and $\beta \neq \epsilon$, then $v$ must be $w_j$ (again because otherwise the variable does not exist). In this case we will replace also $x_{(v,w_j)}$ by 1, which corresponding to the chain $M_A$ moving from state $\langle b_1 \dots b_j, v \rangle$ to state $\langle b_1 \dots b_{j-1}, (b_j, w_j) \rangle$, i.e. returning from the call of box $b_j$ to the return port $(b_j, w_j)$.

If $v$ is not a call port (or an exit) then the equation for the leftmost variable $x_{(v,w_j)}$ is $\sum_{v'} p_{v,v'} x_{(v',w_j)}$ where the sum ranges over all successors $v'$ of $v$ for which the variable $x_{(v',w_j)}$ exists. In particular, if $\beta = \epsilon$, then $x_{(v,ex)} = \sum_{v'} p_{v,v'} x_{(v,ex)}$. Note also that $Prob(Z_{n+1} = \langle \beta, v' \rangle | Z_n = \langle \beta, v \rangle) = p_{v,v'}$.

Finally, if $v = (b', v')$ is a call port of a box $b'$ corresponding to some component $A_k$ with an entry $v'$, then we will replace the leftmost variable $x_{(v,w_j)}$ with $\sum_{w'} x_{(v',w')} x_{((b',w'),w_j)}$ where the sum ranges over all exits $w'$ of $A_k$ for which both variables $x_{(v',w')}, x_{((b',w'),w_j)}$ exist. This corresponds to the chain moving with probability 1 from state $\langle \beta, v \rangle$ to state $\langle \beta b', v' \rangle$, and including all feasible extensions $w'\gamma$ of $\gamma$. □

Let $N$ be any fixed positive integer and consider $n$ going to infinity. We can write $f_n(x)$ as the sum of three terms $c_n, g_n(x), h_n(x)$, where $c_n = Prob(Z_n = \langle \epsilon, ex \rangle)$ is the constant term. A monomial
$Prob(Z_n = \langle \beta, v \rangle) x_{(v,w_j)} x_{((b_j,w_j),w_{j-1})} \dots x_{((b_2,w_2),w_1)} x_{((b_1,w_1),ex)}$. corresponding to a state $\langle \beta, v \rangle$, and a sequence $\gamma = w_1, \dots, w_j$ of exits is included in the second term $g_n(x)$ iff at most $N$ of the vertices $v, (b_j, w_j) \dots (b_2, w_2)(b_1, w_1)$ are deficient; otherwise it is included in $h_n(x)$. Clearly, as $n \to \infty$, the first term $c_n \to q^*_{(u,ex)}$. For $q^*$, the second and third term $g_n(q^*), h_n(q^*)$ tend to 0. Consider these two terms for $y$.

Let $r$ be the minimum component in $q^*$. Then clearly $y \leq 1 \leq q^*/r$. Since in every monomial of the second term at most $N$ of the vertices are defficient, and since $q^*$ and $y$ have the same value for each pair whose first component is a full vertex, it follows that the value of each monomial of $g_n(x)$ evaluated at $y$ is bounded from above by the value of the monomial evaluated at $q^*$ divided by $r^N$. Hence $g_n(y) \leq g_n(q^*)/r^N$. Since $N$ is fixed and $g_n(q^*) \to 0$ as $n \to \infty$, it follows that also $g_n(y) \to 0$ as $n \to \infty$.

Consider all the monomials in the third term $h_n(x)$ corresponding to a state $\langle \beta, v \rangle$ of $M_A$. Let $G$ be the layered Markov chain that has a source node $v$, then it has $j$ layers (numbered from $j$ down to 1) and finally it has a sink node $ex$. Each layer $i$ contains a node labelled $w_i$ for each exit $w_i$ of the component corresponding to the box $b_i$. In addition there is a dead state $d$. Nodes $ex$ and $d$ have self-loops with probability 1. There is a transition from $v$ to a node $w_j$ in layer $j$ with probability $y_{(v,w_j)}$ iff the corresponding variable $x_{(v,w_j)}$ exists. For each pair of nodes $w_i, w_{i-1}$ in successive layers, $i, i-1$ there is a transition from node $w_i$ of layer $i$ to node $w_{i-1}$ of layer $i-1$ with probability $y_{((b_i,w_i),w_{i-1})}$ if the corresponding variable exists. Finally there is a transition from each node $w_1$ of layer 1 to the sink $ex$ with probability $y_{((b_1,w_1),ex)}$ (if the variable exists). Note that the probabilities of the above transitions out of a node of $G$ sum to less than 1 iff the corresponding vertex $v$ or $(b_i, w_i)$ of the RMC is deficient. Let $D$ be the set of these 'deficient' nodes of $G$. For every deficient node add a transition to $d$ with the missing probability. Let $U$ be the set of deficient vertices of the RMC, and let $p = \min\{1 - \sum_{ex'} y_{(u',ex')} | u' \in U\}$. Note that $p > 0$. Each deficient node of $G$ has a transition to $d$ with probability at least $p$.

By our construction of $G$, every monomial of $h_n(y)$ involving the state $\langle \beta, v \rangle$ corresponds to a path in $G$ from $v$ to $ex$ that goes through at least $N$ deficient nodes; the value of the monomial is equal to $Prob(Z_n = \langle \beta, v \rangle)$ times the probability of the path in $G$. The lemma implies then that the contribution to $h_n(y)$ of the set of monomials for state $\langle \beta, v \rangle$ is at most $Prob(Z_n = \langle \beta, v \rangle)(1-p)^N$. Therefore, $h_n(y) \leq (1-p)^N$. Since $(1-p) < 1$ and $N$ is an arbitrary integer, the right hand side can be made arbitrarily small, in particular strictly smaller that $y_{(u,ex)} - q^*_{(u,ex)}$. This contradicts the fact that $y_{(u,ex)} = f_n(y)$ for all $n$, and hence $q^*_{(u,ex)} = \lim_{n \to \infty} f_n(y)$. □