Recursive Markov Decision Processes and Recursive Stochastic Games

Kousha Etessami[1] and Mihalis Yannakakis[2]

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

February 13, 2005

Abstract. We introduce Recursive Markov Decision Processes (RMDPs) and Recursive Simple Stochastic Games (RSSGs), and study the decidability and complexity of algorithms for their analysis and verification. These models extend Recursive Markov Chains (RMCs), introduced in [EY05a,EY05b] as a natural model for verification of probabilistic procedural programs and related systems involving both recursion and probabilistic behavior. RMCs define a class of denumerable Markov chains with a rich theory generalizing that of stochastic context-free grammars and multi-type branching processes, and they are also intimately related to probabilistic pushdown systems. RMDPs & RSSGs extend RMCs with one controller or two adversarial players, respectively. Such extensions are useful for modeling nondeterministic and concurrent behavior, as well as modeling a system's interactions with an environment.
We provide a number of upper and lower bounds for deciding, given an RMDP (or RSSG) A and probability p, whether player 1 has a strategy to force termination at a desired exit with probability at least p. We also address "qualitative" termination questions, where $p = 1$, and model checking questions.

1 Introduction

Markov Decision Processes (MDPs) are a fundamental formalism for modeling control optimization problems in sequential stochastic environments. They have found widespread applications in many fields (see, e.g., [Put94,FS02]). They have also been studied extensively in recent years for verification of probabilistic systems. Stochastic games generalize MDPs with multiple players, and in their 2-player zero-sum version are also known as Competitive MDPs (see [FV97]). Simple Stochastic Games (SSGs) [Con92] are a special class of 2-player zero-sum stochastic games, where the goal of one player is to reach a given terminal state, while the other aims to avoid it. SSGs generalize parity games and other important games for model checking, and the termination problem for finite SSGs already presents a well-known algorithmic challenge: it is in NP $\cap$ coNP, but no P-time algorithm is known (see, e.g., [Con92,EJ91,ZP96]).

Recursive Markov Chains (RMCs) were introduced and studied in our earlier work ([EY05a,EY05b]) as a natural model of probabilistic procedural programs and systems exhibiting both recursion and probabilistic behavior. There we provided strong upper and lower bounds for both reachability and ω-regular model checking questions for RMCs. Informally, a RMC consists of a (finite) collection of finite state Markov chains that can call each other in a potentially recursive manner. RMCs define a class of denumerable Markov chains with a rich theory generalizing that of Stochastic Context-Free Grammars (SCFGs) (see, e.g., [MS99]) and Multi-Type Branching Processes ([Har63]), both of which correspond to 1-exit RMCs: RMCs in which each component Markov chain has 1 terminating exit state

where it can return control back to a component that called it. RMCs are also intimately related to probabilistic Pushdown Systems (pPDSs), which have also been studied recently in connection to verification of probabilistic programs ([EKM04,BKS05]).

In the context of verification, it is quite natural and useful to extend RMCs with nondeterministic choice, where some states are controlled by the system while others exhibit probabilistic behavior. Indeed, finite MDPs have been studied extensively for verification of probabilistic systems, and optimized verification tools already exist for them (see, e.g.,[CY98,CY95,Var85,HSP83,BK98,dAKN$^+$00], and see [Kwi03] for a recent survey). SSGs extend MDPs further with a second (adversarial) player. Like non-probabilistic game graphs, they can also be used to model and analyze the interactions between a controlled (but probabilistic) system and an (adversarial) environment.

In this paper we focus on precisely such extensions of RMCs: we introduce *Recursive Markov Decision Processes* (RMDPs) and *Recursive Simple Stochastic Games* (RSSGs), which define natural classes of countable MDPs and SSGs, respectively. In the stochastic dynamic programming literature, MDPs are studied under many different reward criteria, such as average reward, discounted reward, etc. Our focus here is on verification of probabilistic systems, and for this purpose we study RMDPs and RSSGs under reachability criteria which are central to any analysis like model checking. In particular, we ask the *quantitative termination* question: given an RMDP (or RSSG) A and a probability p, is there a strategy for the controller where (regardless of the strategy used by the adversary, in the case of RSSGs) the process terminates at a desired exit with probability at least p (or at most p)? We also ask the *qualitative* question of whether the controller has a strategy to force termination with probability 1. Lastly, we address model checking questions.

Our positive results apply primarily to 1-exit RMDPs and 1-exit RSSGs, which correspond to controlled and game extensions, respectively, of both SCFGs and Multi-Type Branching Processes (MT-BPs). Branching processes are an important class of stochastic processes, dating back to the early work of Galton and Watson in the 19th century (they studied the single-type case, a subcase of 1-exit 1-entry 1-component RMCs), and continuing in the 20th century in the work of Kolmogorov, Sevastianov, Harris and others for MT-BPs and beyond (see, e.g., [Har63]). These have been used to model a wide variety of applications, including in population genetics ([Jag75]), nuclear chain reactions ([EU48]), and RNA modeling in computational biology (based on SCFGs) ([SBH$^+$94]). SCFGs are also fundamental models in statistical natural language processing (see, e.g., [MS99]). 1-exit RMDPs correspond to a controlled version of MT-BPs (and SCFGs): the reproduction of some types can be controlled, while the dynamics of other types is probabilistic as in ordinary MT-BPs. This model would also be suitable for analysis of evolutionary dynamics under worst-case (or best-case) assumptions for some types and probabilistic assumptions for others. Such controlled MT-BPs can be readily translated to 1-entry, 1-exit RMDPs, where the number of components is bounded by the number of types (a reverse translation is possible, but will not in general preserve the number of components, i.e., 1-entry, 1-exit RMDPS with a bounded number of components are more general than MT-BPs with a bounded number of types). Thus, our results on 1-exit RMDPs apply, among other things, to such controlled MT-BPs; these do not appear to have been studied in the rich Branching Process literature. Indeed, even some basic algorithmic problems about SCFGs and MT-BTs had received limited attention prior to our work in [EY05a,EY05b].
We now outline our main results in this paper:

- We first show that the Least Fixed Point solution of certain systems of nonlinear min/max equations capture optimal termination probabilities for 1-exit RMDPs & 1-exit

RSSGs. These equations generalize linear Bellman's equations for finite MDPs (see, e.g., [Put94,FV97]) and they also generalize the monotone systems of nonlinear polynomial equations for RMCs that we studied in ([EY05a]).

- We show a quite nontrivial *Stackless & Memoryless (S&M) Determinacy* result for 1-exit RSSG termination, whereas we observe this fails badly even for 2-exit RMDPs (namely, we observe that optimal strategies, of any kind, do not always exist for 2-exit RMDP termination, and one must make do with ϵ-optimal strategies).
- Using the equations, we show that quantitative termination for 1-exit RMDPs and 1-exit RSSGs is decidable in PSPACE. This matches our PSPACE upper bound for the special case of 1-exit RMCs in [EY05a] and, as shown there, it can not be improved without resolving a long standing open problem in the complexity of numerical computation, namely the square-root sum problem.
- Using S&M-determinacy, we show the qualitative termination problem (where $p = 1$) for 1-exit RMDPs can be decided in NP, and for 1-exit RSSGs can be decided in $\Sigma_2^P \cap \Pi_2^P$.
- Lastly, and unfortunately, we show that for multi-exit RMDPs and RSSGs the situation is far worse. Even qualitative termination for general RMDPs is undecidable, even when the number of exits in bounded by a fixed constant. And it is even undecidable, for any fixed $\epsilon > 0$, to distinguish whether the optimal value is 1 or is less than ϵ. So the optimal probabilities can not even be approximated in a strong sense, with any amount of resources. Furthermore, we show, also unfortunately, that the undecidability results apply already to qualitative model checking of 1-exit RMDPs, against regular, ω-regular, or LTL properties. Our undecidability results are derived from classic and recent undecidability results for Probabilistic Finite Automata (PFA) [Paz71,CL89,BC03]. We show that PFAs can be viewed as essentially a special case of multi-exit RMDPs.

Related work. Both MDPs and Stochastic Games have a vast literature, dating back to Bellman and Shapley (see, e.g., [Put94,FS02,FV97]). MDPs are studied in both finite state and infinite state variants. Verification of finite state MDPs, also called concurrent Markov chains, has been studied for a long time (see, e.g., [CY98,CY95,Var85,HSP83]). [CY98] provides efficient algorithms for ω-regular model checking of finite MDPs. Model checking tools like PRISM contain optimized implementations of branching-time model checkers for finite MDPs (see, e.g., [BK98,dAKN$^+$00,Kwi03]).

Our earlier work [EY05a,EY05b] developed the basic theory of RMCs and studied efficient algorithms for both their reachability analysis and model checking. We showed, among many results, that qualitative model checking of ω-regular properties for 1-exit RMCs can be decided in polynomial time in the size of the RMC, and that quantitative model checking of RMCs can be done in PSPACE in the size of the RMC. As mentioned, 1-exit RMCs correspond to both MT-BPs and SCFGs (see, e.g., [Har63] and [MS99]), while general RMCs are intimately related to probabilistic Pushdown Systems (pPDSs). Model checking questions for pPDSs, for both linear and branching time properties, have also been recently studied in [EKM04,BKS05]. [EE04] is an early survey paper which summarizes the results in only the papers [EKM04,EY05a,BKS05]. RMDPs and RSSGs are natural extensions of RMCs, introducing nondeterministic and game behavior. Countable state MDPs are studied extensively in the MDP literature (see, e.g., [Put94,FS02]), but the concise representations afforded by RMDPs and its algorithmic properties, appear not to have been studied.

2 Basics

A *Recursive Simple Stochastic Game (RSSG)*, A, is a tuple $A = (A_1, \ldots, A_k)$, where each *component graph* $A_i = (N_i, B_i, Y_i, En_i, Ex_i, \mathtt{pl}_i, \delta_i)$ consists of:

- A set N_i of *nodes*. Let $N = \cup_{i=1}^k N_i$ be the (disjoint) union of all nodes of A.
- A distinguished subset of *entry* nodes $En_i \subseteq N_i$, and a disjoint subset of *exit* nodes $Ex_i \subseteq N_i$. Let $En = \cup_{i=1}^k En_i$ and $Ex = \cup_{i=1}^k Ex_i$.
- 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 $Y_i : B_i \mapsto \{1, \ldots, k\}$ that assigns to every box (the index of) of a component. Let $Y = \cup_{i=1}^k Y_i$ be the map $Y : B \mapsto \{1, \ldots, k\}$ 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*, $Return_b = \{(b, ex) \mid ex \in Ex_{Y(b)}\}$. Let $Call^i = \cup_{b \in B_i} Call_b$ and let $Call = \cup_{i=1}^k Call^i$ denote all calls in A. Similarly, define $Return^i$ and $Return$.
- We let $Q_i = N_i \cup Call^i \cup Return^i$, denote collectively the nodes, call ports, and return ports, We will use the term *vertex* of A_i to refer to elements of Q_i. We let $Q = \bigcup_{i=1}^k Q_i$ be the set of all vertices of the RSSG A.
- A mapping $\mathtt{pl}_i : Q_i \mapsto \{0, 1, 2\}$ that assigns to every vertex a <u>player</u> (Player 0 represents "chance" or "nature"). We assume $\mathtt{pl}_i(ex) = 0$ for all $ex \in Ex_i$. Let $\mathtt{pl} = \cup_{i=1}^k \mathtt{pl}_i$ denote the map $\mathtt{pl} : Q \mapsto \{0, 1, 2\}$ where $\mathtt{pl}|_{Q_i} = \mathtt{pl}_i$, for $1 \leq i \leq k$.
- A transition relation $\delta_i \subseteq (Q_i \times (\mathbb{R} \cup \{\bot\}) \times Q_i)$, where for each tuple $(u, x, v) \in \delta_i$, the source $u \in (N_i \setminus Ex_i) \cup Return^i$, the destination $v \in (N_i \setminus En_i) \cup Call^i$, and x is either (i) a real number $p_{u,v} \in [0, 1]$ (the transition probability) if $\mathtt{pl}(u) = 0$, or (ii) $x = \bot$ if $\mathtt{pl}(u) = 1$ or 2. For computational purposes we assume that the given probabilities $p_{u,v}$ are rational. Furthermore they must satisfy the consistency property: for every $u \in \mathtt{pl}^{-1}(0)$, $\sum_{\{v' \mid (u, p_{u,v'}, v') \in \delta_i\}} p_{u,v'} = 1$, unless u is a call port or exit node, neither of which have outgoing transitions, in which case by default $\sum_{v'} p_{u,v'} = 0$.
  Let $\delta = \cup_i \delta_i$ be the set of all transitions of A.

An RSSG A defines a global denumerable Simple Stochastic Game (SSG) $M_A = (V = V_0 \cup V_1 \cup V_2, \Delta, \mathtt{pl})$ as follows. The global *states* $V \subseteq B^* \times Q$ of M_A 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 Δ are defined inductively as follows:

1. $\langle \epsilon, u \rangle \in V$, for $u \in Q$. (ϵ denotes the empty string.)
2. if $\langle \beta, u \rangle \in V$ & $(u, x, v) \in \delta$, then $\langle \beta, v \rangle \in V$ and $(\langle \beta, u \rangle, x, \langle \beta, v \rangle) \in \Delta$.
3. if $\langle \beta, (b, en) \rangle \in V$, with $(b, en) \in Call_b$, then $\langle \beta b, en \rangle \in V$ & $(\langle \beta, (b, en) \rangle, 1, \langle \beta b, en \rangle) \in \Delta$.
4. if $\langle \beta b, ex \rangle \in V$, & $(b, ex) \in Return_b$, then $\langle \beta, (b, ex) \rangle \in V$ & $(\langle \beta b, ex \rangle, 1, \langle \beta, (b, ex) \rangle) \in \Delta$.

Item 1 corresponds to the possible initial states, item 2 corresponds to control staying within a component, item 3 is when a new component is entered via a box, item 4 is when control exits a box and returns to the calling component. The mapping $\mathtt{pl} : V \mapsto \{0, 1, 2\}$ is given as follows: $\mathtt{pl}(\langle \beta, u \rangle) = \mathtt{pl}(u)$ if u is in $Q \setminus (Call \cup Ex)$, and $\mathtt{pl}(\langle \beta, u \rangle) = 0$ if $u \in Call \cup Ex$. The set of vertices V is partitioned into V_0, V_1, and V_2, where $V_i = \mathtt{pl}^{-1}(i)$.

We consider M_A with various *initial states* of the form $\langle \epsilon, u \rangle$, denoting this by M_A^u. Some states of M_A are *terminating states* and have no outgoing transitions. These are states $\langle \epsilon, ex \rangle$, where ex is an exit node. If we wish to view M_A as a non-terminating SSG, we can consider the terminating states as absorbing states of M_A, with a self-loop of probability 1.

An RSSG where $V_2 = \emptyset$ is called a *Recursive Markov Decision Process* (RMDP); an RSSG where $V_1 \cup V_2 = \emptyset$ is called a *Recursive Markov Chain* ([EY05a,EY05b]); an RSSG

where $V_0 \cup V_2 = \emptyset$ is called a *Recursive Graph* ([AEY01]); an RSSG where $V_0 = \emptyset$ is called a *Recursive Game Graph* (see [ATM03,Ete04]).

We use *1-exit RSSG* to refer to RSSGs where every component has 1 exit. Without loss of generality, we can assume that every component has 1 entry, because multi-entry RSSGs can be transformed to equivalent 1-entry RSSGs with at most polynomial blowup (similar to RSM transformations, see [AEY01]). However, this is decidedly not the case for exits: indeed 1-exit RSSGs form an important sub-class of RSSGs and are the main focus of our upper bounds in this paper.

Our most basic goal is to answer termination questions for RSSGs of the form: *"Does player 1 have a strategy to force the game to terminate (i.e., reach node $\langle \epsilon, ex \rangle$), starting at $\langle \epsilon, u \rangle$, with probability $\geq p$, regardless of how player 2 plays?"*.

First, some definitions: a *strategy* σ for player i, $i \in \{1, 2\}$, is a function $\sigma : V^* V_i \mapsto V$, where, given the history $ws \in V^* V_i$ of play so far, with $s \in V_i$ (i.e., it is player i's turn to play a move), $\sigma(w) = s'$ determines the next move of player i, where $(s, \bot, s') \in \Delta$.

Let Ψ_i denote the set of all strategies for player i. Given a start node u, a strategy $\sigma \in \Psi_1$ for player 1, and a strategy $\tau \in \Psi_2$ for player 2, we define a new Markov chain (with initial state u) $M_A^{u,\sigma,\tau} = (\mathcal{S}, \Delta')$. The states $\mathcal{S} \subseteq \langle \epsilon, u \rangle V^*$ of $M_A^{u,\sigma,\tau}$ are non-empty sequences of states of M_A, which must begin with $\langle \epsilon, u \rangle$. Inductively, if $ws \in \mathcal{S}$, then: (0) if $s \in V_0$ and $(s, p_{s,s'}, s') \in \Delta$ then $wss' \in \mathcal{S}$ and $(ws, p_{s,s'}, wss') \in \Delta'$; (1) if $s \in V_1$ and $\sigma(ws) = s'$ (where $(s, \bot, s') \in \Delta$) then $wss' \in \mathcal{S}$ and $(ws, 1, wss') \in \Delta'$; (2) if $s \in V_2$ and $\tau(ws) = s'$ (where $(s, \bot, s') \in \Delta$) then $wss' \in \mathcal{S}$ and $(ws, 1, wss') \in \Delta'$.

Given initial vertex u, and final exit ex in the same component, and given strategies $\sigma \in \Psi_1$ and $\tau \in \Psi_2$, for $k \geq 0$, let $q^{k,\sigma,\tau}_{(u,ex)}$ be the probability that, in $M_A^{u,\sigma,\tau}$, starting at initial state $\langle \epsilon, u \rangle$, we will reach a state $w\langle \epsilon, ex \rangle$ in at most k "steps" (i.e., where $|w| \leq k$). Let $q^{*,\sigma,\tau}_{(u,ex)} = \lim_{k \to \infty} q^{k,\sigma,\tau}_{(u,ex)}$ be the probability of ever terminating at ex, i.e., reaching $\langle \epsilon, ex \rangle$. (Note, the limit exists: it is a monotonically non-decreasing sequence bounded by 1). Let $\mathbf{q}^k_{(u,ex)} = \max_{\sigma \in \Psi_1} \min_{\tau \in \Psi_2} q^{k,\sigma,\tau}_{(u,ex)}$ and let $\mathbf{q}^*_{(u,ex)} = \sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} q^{*,\sigma,\tau}_{(u,ex)}$. Next, for a strategy $\sigma \in \Psi_1$, let $q^{k,\sigma}_{(u,ex)} = \min_{\tau \in \Psi_2} q^{k,\sigma,\tau}_{(u,ex)}$, and let $q^{*,\sigma}_{(u,ex)} = \inf_{\tau \in \Psi_2} q^{*,\sigma,\tau}_{(u,ex)}$. Lastly, given instead a strategy $\tau \in \Psi_2$, let $q^{k,\cdot,\tau}_{(u,ex)} = \max_{\sigma \in \Psi_1} q^{k,\sigma,\tau}_{(u,ex)}$, and let $q^{*,\cdot,\tau}_{(u,ex)} = \sup_{\sigma \in \Psi_1} q^{*,\sigma,\tau}_{(u,ex)}$.

From very general determinacy results (Martin's "Blackwell determinacy" [Mar98] is one such result which applies to all two-player zero-sum stochastic games with countable state spaces) it follows that the games M_A are *determined*, meaning that $\sup_{\sigma \in \Psi_1} \inf_{\tau \in \Psi_2} q^{*,\sigma,\tau}_{(u,ex)} = \inf_{\tau \in \Psi_2} \sup_{\sigma \in \Psi_1} q^{*,\sigma,\tau}_{(u,ex)}$. Of course, finite SSGs are even *memorylessly determined* ([Con92]), meaning that the strategies of either player can be restricted to *memoryless strategies* which ignore the history prior to the current position, without harming the optimal outcome for that player. As we shall see, 1-exit RSSGs exhibit memoryless determinacy in an even stronger sense, namely, the strategy is also independent of the call stack. This fails badly for multi-exit RMDPs, as we will see. We are interested in the following questions:

(1) The *qualitative* termination problem: Is $\mathbf{q}^*_{(u,ex)} = 1$?
(2) The *quantitative* termination problems: Given $r \in [0, 1]$, is $q^*_{(u,ex)} \geq r$? Is $q^*_{(u,ex)} = r$? We may also wish to compute or approximate the exact probabilities $q^*_{(u,ex)}$.

More generally, we can ask model checking questions, where, given a Σ-labeling of vertices, and e.g., an ω-automaton $\mathcal{B}$ over Σ, we ask what is the supremum probability with which player 1 can force the satisfaction of property $\mathcal{B}$? We refrain from formal definitions due to space (see,e.g., [CY98,EY05b]). Our results for model checking will be negative: undecidability, stemming from the undecidability of termination problems for general RMDPs.

3 The System of Nonlinear Min-Max Equations for 1-exit RSSGs

We generalize the monotone nonlinear systems of equations for RMCs ([EY05a]) to monotone nonlinear min-max systems for 1-exit RSSGs. If we consider the values $q^*_{(u,ex)}$ as unknowns, we can set up a system of *nonlinear min-max* equations, such that the termination probabilities must be a solution of the system, and in fact the *Least Fixed Point* solution. Let us use a variable $x_{(u,ex)}$ for each unknown $q^*_{(u,ex)}$. We will often find it convenient to index the variables $x_{(u,ex)}$ according to a fixed order (say lexicographical), 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. In this way we obtain a vector of variables: $\mathbf{x} = (x_1\ x_2 \ldots x_n)^T$.

Definition 1. *Given 1-exit RSSG $A = (A_1, \ldots, A_k)$, we define a system of polynomial/min-max 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 of the form $x_{(u,ex)} = P_{(u,ex)}(\mathbf{x})$, for each variable $x_{(u,ex)}$. Here $P_{(u,ex)}(\mathbf{x})$ denotes a multivariate min-max/polynomial with positive coefficients over the variables in $\mathbf{x}$. There are 5 cases to distinguish, based on what "type" of vertex u is:*

1. *Type I: $u = ex$. In this case: $x_{(ex,ex)} = 1$.*
2. *Type II: $\mathtt{pl}(u) = 0$ & $u \in (N_i \backslash \{ex\}) \cup \mathrm{Return}^i$. Then $x_{(u,ex)} = \sum_{\{v|(u,p_{u,v},v)\in\delta\}} p_{u,v} x_{(v,ex)}$. (If u has no outgoing transitions, this equation is by definition $x_{(u,ex)} = 0$.)*
3. *Type III: $u = (b, en)$ is a call port. In this case $x_{((b,en),ex)} = x_{(en,ex')} \cdot x_{((b,ex'),ex)}$, where $ex' \in Ex_{Y(b)}$ is the unique exit of $A_{Y(b)}$.*
4. *Type IV: $\mathtt{pl}(u) = 1$ and $u \in (N_i \backslash \{ex\}) \cup \mathrm{Return}^i$. Then $x_{(u,ex)} = \max_{\{v|(u,\perp,v)\in\delta\}} x_{(v,ex)}$. (If u has no outgoing transitions, we define the maximum of an empty set to be 0.)*
5. *Type V: $\mathtt{pl}(u) = 2$ and $u \in (N_i \backslash \{ex\}) \cup \mathrm{Return}^i$. Then $x_{(u,ex)} = \min_{\{v|(u,\perp,v)\in\delta\}} x_{(v,ex)}$. (If u has no outgoing transitions, we define the minimum of an empty set to be 0.)*

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

Given 1-exit RSSG A, we can easily construct $\mathbf{x} = P(\mathbf{x})$ in linear time. We now identify a particular solution to $\mathbf{x} = P(\mathbf{x})$, called the *Least Fixed Point* (LFP) solution, which gives precisely the termination game values. For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, define the partial-order $\mathbf{x} \preceq \mathbf{y}$ to mean $x_j \leq y_j$ for every coordinate j. For $D \subseteq \mathbb{R}^n$, we 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 define $P^k(\mathbf{x}) = P(P^{k-1}(\mathbf{x}))$, for $k > 1$. Let $\mathbf{q}^* \in \mathbb{R}^n$ denote the n-vector $q^*_{(u,ex)}$ (using the same indexing as used for $\mathbf{x}$). For $k \geq 0$, let $\mathbf{q}^k$ denote, similarly, the n-vector $q^k_{(u,ex)}$. Let $\mathbf{0}$ ($\mathbf{1}$) denote the n-vector consisting of 0 (respectively, 1) in every coordinate. Define $\mathbf{x}^0 = \mathbf{0}$, and for $k \geq 1$, define $\mathbf{x}^k = P(\mathbf{x}^{k-1}) = P^k(\mathbf{0})$.

Theorem 1. *Let $\mathbf{x} = P(\mathbf{x})$ be the system S_A associated with 1-exit RSSG A.*

1. *The map $P : \mathbb{R}^n \mapsto \mathbb{R}^n$ is monotone on $\mathbb{R}^n_{\geq 0}$. Hence, for all $k \geq 0$, $\mathbf{0} \preceq \mathbf{x}^k \preceq \mathbf{x}^{k+1}$.*
2. *For all $k \geq 0$, $\mathbf{q}^k \preceq \mathbf{x}^{k+1} \preceq \mathbf{q}^{2^k}$.*
3. *$\mathbf{q}^* = P(\mathbf{q}^*)$. In other words, $\mathbf{q}^*$ is a* fixed point *of the map P.*
4. *For all $k \geq 0$, $\mathbf{x}^k \preceq \mathbf{q}^*$.*
5. *For all $\mathbf{q}' \in \mathbb{R}^n_{\geq 0}$, if $\mathbf{q}' = P(\mathbf{q}')$, then $\mathbf{q}^* \preceq \mathbf{q}'$.*
   In other words, $\mathbf{q}^*$ is the Least Fixed Point, LFP(P), *of $P : \mathbb{R}^n_{\geq 0} \mapsto \mathbb{R}^n_{\geq 0}$.*
6. *$\mathbf{q}^* = \lim_{k\to\infty} \mathbf{x}^k = \lim_{k\to\infty} \mathbf{q}^k$.*

All assertions are proved in the appendix. Their proofs are similar to those of an analogous theorem in [EY05a] for nonlinear systems associated with RMCs, but some parts are more tricky because of the players. We sketch here the idea for part (5). Consider any fixpoint $\mathbf{q}'$ of the equations, i.e., $\mathbf{q}' = P(\mathbf{q}')$. Let τ' be the (S&M) strategy for player 2 that always picks, at any state $\langle \beta, u \rangle$, for vertex $u \in \mathtt{pl}^{-1}(2)$, the particular successor v of u such that $v = \arg\min_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}'_{(v,ex)}$ (breaking ties, say, lexicographically). Then we prove a lemma stating that, for all strategies $\sigma \in \Psi_1$ of player 1, and for all $k \geq 0$, $\mathbf{q}^{k,\sigma,\tau'} \preceq \mathbf{q}'$. The lemma is proved by induction on k, for each entry (u, ex), going through the 5 type cases of the vertex u. The lemma implies that $\mathbf{q}^{*,\sigma,\tau'} = \lim_{k\to\infty} \mathbf{q}^{k,\sigma,\tau'} \preceq \mathbf{q}'$. This holds for any strategy $\sigma \in \Psi_1$. Therefore, $\sup_{\sigma\in\Psi_1} q^{*,\sigma,\tau'}_{(u,ex)} \leq \mathbf{q}'_{(u,ex)}$, for every vertex u. Thus, by the determinacy of RSSG games, it follows that $\mathbf{q}^*_{(u,ex)} = \inf_{\tau\in\Psi_2} \sup_{\sigma\in\Psi_1} q^{*,\sigma,\tau}_{(u,ex)} \leq \sup_{\sigma\in\Psi_1} q^{*,\sigma,\tau'}_{(u,ex)} \leq \mathbf{q}'_{(u,ex)}$, for all vertices u. In other words, $\mathbf{q}^* \preceq \mathbf{q}'$. □

4 S&M Determinacy

We now identify a very restricted kind of strategy that suffices as an optimal strategy in 1-exit RSSGs. Call a strategy *Stackless & Memoryless (S&M)* if it is not only independent of the history of the game, but also independent of the current call stack, i.e., only depends on the current vertex. (See also [ATM03], where such strategies are called *modular* strategies.)

Corollary 1. *In every 1-exit RSSG termination game, player 2 (the minimizer) has an optimal S&M strategy.*

Proof. Consider the strategy τ' in the proof of part (5) of Theorem 1, chosen not for just any fixed point $\mathbf{q}'$, but for $\mathbf{q}^*$ itself. □

Far less trivially, we establish next that player 1 (the maximizer) also has an optimal S&M strategy and thus the game is *S&M-determined*, meaning both players have optimal S&M strategies. (Note that the game is not symmetric with respect to the two players.)

Theorem 2. *Every 1-exit RSSG termination game is S&M-determined.*

Although the statement is intuitive, the proof is quite nontrivial and delicate; the full proof is given in the appendix. We sketch the approach here. By Corollary 1, we only need to show that player 1 has an optimal S&M strategy. Let σ be any S&M strategy for player 1, and let $\mathbf{q}^{*,\sigma} = \inf_{\tau\in\Psi_2} \mathbf{q}^{*,\sigma,\tau}$. If $\mathbf{q}^{*,\sigma}$ is a fixpoint of the equations then it follows that it is the least fixpoint and σ is optimal. On the other hand, it can be shown that $\mathbf{q}^{*,\sigma}$ satisfies all the equations except possibly for some type IV equations. We argue that if u is such a vertex (belonging to player 1) whose equation is violated, then switching to another strategy σ' where u picks another successor leads to a strictly better strategy than σ (for any strategy of player 2). This is the heart of the proof. We parameterize the game with respect to the value t at vertex u, and we express the optimal values of the other vertices z (for all strategies τ of player 2) as functions $f_z(t)$. We then analyze carefully the properties of these functions and their fixpoints, and conclude that switching the choice at vertex u leads to a strategy σ' that has at least as great value as σ at every vertex, and strictly better at u. We repeat the process until we arrive at a S&M strategy σ^* whose probabilities satisfy all the equations, and hence it is optimal. We refer to the appendix for the details. □

We remark that already for the 2-exit RMDP termination problem, not only are there no optimal S&M strategies for player 1, there are in general no optimal strategies at all!

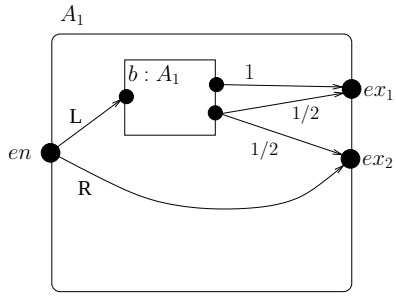

Fig. 1. *Maximizing termination probability at ex_1 in a 2-exit RMDP:* no optimal strategy exists.

This is illustrated by Figure 1. In this 2-exit RMDP with one component containing a box mapped to the same component, it can easily be verified that the *supremum* probability of terminating at exit ex_1 starting from en is 1. However, there is no optimal strategy for player 1 that actually achieves this. The deeper the call stack is made by player 1, the higher the probability of termination at ex_1. However, at some depth, player 1 finally has to decide to come out, otherwise it will never terminate. Specifically, for $n \geq 0$, the strategy $L^n R$ terminates at ex_1 with probability $(1 - \frac{1}{2^n})$. Note also that any S&M strategy for player 1 would yield probability 0 of terminating at ex_1, so such strategies are all the worst possible.

5 Qualitative & Quantitative termination: 1-exit RMDPs & RSSGs

Using Corollary 1 and Theorem 2, and results from [EY05a], we can show the following results for qualitative termination of 1-exit RMDPs and 1-exit RSSGs:

Theorem 3.

1. *We can decide in polynomial time if the value of a 1-exit RSGG termination game (or the optimal termination probability of a maximizing or minimizing RMDP) is 0.*
2. *We can decide in NP whether the maximum probability of termination in a maximizing 1-exit RMDP is exactly* 1*, and in coNP whether the minimum probability of termination in a minimizing 1-exit RMDP is exactly* 1*.*
3. *Deciding whether a 1-exit RSSG termination game has value exactly 1 is in $\Sigma_2^P \cap \Pi_2^P$.*

Part (1) is done via a fixpoint algorithm. Parts (2) and (3) involve guessing the optimal S&M strategy (or strategies) and verifying the optimality with the appropriate complexity; see the appendix. □

We next show that the central quantitative termination questions for 1-exit RMDPs and 1-exit RSSGs can be answered in PSPACE by appealing to algorithms for deciding the *Existential Theory of the Reals*, $\boldsymbol{ExTh}(\mathbb{R})$. A first-order sentence in the theory of reals is formed from quantifiers and boolean connectives over a vocabulary with "atomic predicates" of the form: $f_i(\mathbf{x})\Delta 0$, where f_i, $i = 1, \ldots, m$, are m distinct multi-variate polynomials with rational coefficients over the variables $\mathbf{x} = x_1, \ldots, x_n$, and where Δ is any comparison operator among $=, \neq, \geq, \leq, <, >$. The fragment that we will be concerned with is the *existential theory of reals*, $\boldsymbol{ExTh}(\mathbb{R})$. These consist of sentences in prenex form of the form: $\exists x_1, \ldots, x_n R(x_1, \ldots, x_n)$, where R is a boolean combination of "atomic predicates". Beginning with Tarski, algorithms for deciding the first-order theory of reals, $\boldsymbol{Th}(\mathbb{R})$, and its

fragments such as $\boldsymbol{ExTh}(\mathbb{R})$, have been deeply investigated. In the current state of the art, it is known that $\boldsymbol{ExTh}(\mathbb{R})$ can be decided in PSPACE [Can88,Ren92,BPR96]. Furthermore, it can be decided in exponential time where the exponent depends (linearly) only on the number of variables; hence for a fixed number of variables the time is polynomial.

Suppose we want to decide whether a vector $\mathbf{c} = [c_1, \ldots, c_n]^T$ of rational numbers is $LFP(P)$, where P is the system of fixpoint equations for a given 1-exit RSSG. Consider the sentence:

$$\varphi \equiv \exists x_1, \ldots, x_n \bigwedge_{i=1}^{n} (P_i(x_1, \ldots, x_n) = x_i) \wedge \bigwedge_{i=1}^{n} (x_i = c_i)$$

φ is true iff $\mathbf{c} = P(\mathbf{c})$. For type I, II, and III nodes, P_i is a polynomial. It remains to show how to encode, in arithmetic, the predicate "$P_i(x_1, \ldots, x_n) = x_i$" in the case (IV) where $P_i(x_1, \ldots, x_n) = \max_{j \in J} x_j$, and in the case (V) where $P_i(x_1, \ldots, x_n) = \min_{j \in J} x_j$, for some subset $J \subseteq \{1, \ldots, n\}$. For type IV nodes, note that $x_i = \max_{j \in J} x_j$ iff $\bigwedge_{j \in J} x_i \geq x_j \wedge (\bigvee_{j \in J} x_i \leq x_j)$. Likewise, for type V nodes, $x_i = \min_{j \in J} x_j$ iff $\bigwedge_{j \in J} x_i \leq x_j \wedge (\bigvee_{j \in J} x_i \geq x_j)$. Thus, we can encode the predicates $x_i = P_i(x_1, \ldots, x_n)$ as a boolean combination of quantifier-free predicates, and we can encode the sentence φ in $\boldsymbol{ExTh}(\mathbb{R})$. To guarantee that $\mathbf{c} = \mathrm{LFP}(P)$, we need in addition to check the following sentence:

$$\psi \equiv \exists x_1, \ldots, x_n \bigwedge_{i=1}^{n} (P_i(x_1, \ldots, x_n) = x_i) \wedge \bigwedge_{i=1}^{n} (0 \leq x_i) \wedge \bigvee_{i=1}^{n} (x_i < c_i)$$

ψ is false iff there is no solution $\mathbf{z} \in \mathbb{R}^n_{\geq 0}$ to $\mathbf{x} = P(\mathbf{x})$ such that $\mathbf{c} \not\preceq \mathbf{z}$. Hence, to decide whether $\mathbf{c} = \mathrm{LFP}(P)$, we only need two queries to a decision procedure for $\boldsymbol{ExTh}(\mathbb{R})$. Namely, we check that φ is true, and hence $\mathbf{c} = P(\mathbf{c})$, and that ψ is false, and hence $\mathbf{c} = \mathrm{LFP}(P)$.

If we only want to check an inequality $\mathbf{q}^*_j \leq c_j$, then let φ' be φ with the last conjunction of equations replaced by $\bigwedge_{i=1}^{n} (0 \leq x_i) \wedge (x_j \leq c_j)$. Note that all multi-variate polynomials in our systems $\mathbf{x} = P(\mathbf{x})$ have (multivariate) degree $d \leq 2$. Applying the result on $\boldsymbol{ExTh}(\mathbb{R})$, we obtain the following:

Theorem 4. *Given a 1-exit RSSG A and a vector of rational probabilities $\mathbf{c}$, there is a PSPACE algorithm to decide whether $\mathbf{q}^* = \mathbf{c}$, as well as to decide whether $\mathbf{q}^*_j \, \Delta c_j$, for any comparison operator Δ. Moreover, the running time is $O(|A|^{O(1)} \cdot 2^{O(n)})$ where n is the number of variables in $\mathbf{x} = P(\mathbf{x})$. Hence the running time is polynomial if n is bounded.*

As we have seen, $\boldsymbol{ExTh}(\mathbb{R})$ gives us a way to ask questions like: "Is there a solution to $\mathbf{x} = P(\mathbf{x})$ where $a \leq x_i \leq b$?", for any rational numbers a and b, and if we wish, with either inequality replaced by strict inequality. Since $\mathbf{0} \leq \mathrm{LFP}(P) \leq \mathbf{1}$, we can use such queries in a "binary search" to "narrow in" on the value of each coordinate of $\mathrm{LFP}(P)$. Via obvious modifications of sentences like ψ, we can gain one extra bit of precision on the exact value of c_i with each extra query to $\boldsymbol{ExTh}(\mathbb{R})$. So, if we want j bits of precision for each c_i, $i = 1, ...n$, we need to make $j \cdot n$ queries. The sizes of the queries do not vary by much: only with an additive factor of at most j-bits, to account for the constants a and b. This discussion yields:

Theorem 5. *Given 1-exit RSSG A and a number j in unary, there is an algorithm that approximates every coordinate of $\mathbf{q}^*$ to within j bits of precision in PSPACE. Moreover, the running time is $O(j \cdot |A|^{O(1)} \cdot 2^{O(n)})$, where n is the number of variables in $\mathbf{x} = P(\mathbf{x})$.*

6 Multi-exit RMDP termination and 1-exit RMDP model checking: undecidability

In this section we show that the qualitative termination problem for multi-exit RMDPs is already undecidable, and that this is so even when the number of exits is bounded by a constant. We also show that for 1-exit RMDPs the qualitative model checking problem, against ω-regular or LTL properties (specified, say, by a Büchi automaton) is also undecidable. It obviously follows that the same questions for the more general RSSGs are also undecidable.

Theorem 6. *Given a multi-exit RMDP, A, entry en and exit ex, it is undecidable whether $q^*_{(en,ex)} = 1$. This is so even when the number of exits in each component of A is bounded by a fixed constant. Furthermore, there is no algorithm that approximates the probability $q^*_{(en,ex)}$ within any constant (multiplicative) factor. In particular:*
1. For any fixed rational ϵ with $0 < \epsilon < 1$, given RMDP A with only one component such that either $q^*_{(en,ex)} > 1 - \epsilon$ or $q^*_{(en,ex)} < \epsilon$, it is undecidable to distinguish which is the case.
2. For any fixed rational ϵ with $0 < \epsilon < 1$, given a RMDP A with only two components such that either $q^*_{(en,ex)} = 1$ or $q^*_{(en,ex)} < \epsilon$, it is undecidable to distinguish which is the case.

We have two proofs: one is a direct reduction from the halting problem for 2-counter machines (see, e.g., [HU79]), and the second is a reduction from the emptiness problem for Probabilistic Finite Automata (PFAs) [Paz71]. The latter reduction is simpler and has the advantage that it connects the well-studied area of PFAs to RMDPs (we show that PFAs can be seen, in effect, as a special case of RMDPs) and also allows us to leverage the extensive research in that area. We sketch this reduction here. A PFA, $M = (Q, \Sigma, T, q_1, q_n)$, has a (finite) set Q of n states, a (finite) input alphabet Σ, a function T which maps every input letter $a \in \Sigma$ to a stochastic $n \times n$ matrix T_a, an initial state q_1, and an accepting state q_n. The automaton starts at the initial state q_1 and operates as follows for an input string $w \in \Sigma^*$: at each step, if the automaton is in state q_i and the next input letter is a, the automaton transitions to state q_j with probability equal to $T_a[i, j]$. The probability $P_M(w)$ that M accepts the string w is defined as the probability that the automaton is at the accepting state q_n after reading w. The language of a PFA is defined with respect to a given threshold λ: The language is $L(M, \lambda) = \{w \in \Sigma^* | P_M(w) > \lambda\}$. The *PFA emptiness problem* is to decide, given a PFA M and a threshold λ, whether $L(M, \lambda) = \emptyset$. This problem was shown to be undecidable originally in [Paz71]; subsequent proofs have established stronger undecidability properties [CL89,MHC03,BC03]. (The undecidability holds both whether we use strict or weak inequality in the definition of $L(M, \lambda)$.)

Let $M = (Q, \Sigma, T, q_1, q_n)$ be a PFA with n states, and let $p^*_M = \sup\{P_M(w) | w \in \Sigma^*\}$. Define a RMDP A that has one component (call it also A) with a single entry en, and n exits $ex_1, \ldots, ex_n$, one for each state of M. The entry en is a player 1 (max) node and has edges to the call ports of a set of $|\Sigma|$ boxes $b_a, a \in \Sigma$; all the boxes are of course mapped to the single component A. In addition en has an edge to the exit ex_1. The return ports of the boxes b_a are probabilistic vertices. Each return port (b_a, ex_i) has an edge to each exit ex_j with probability $T_a[i, j]$. This concludes the definition of the RMDP A. Starting from the entry en of A, the max player wants to maximize the probability $q^*_{(en,ex_n)}$ of terminating at exit ex_n. We can show that $q^*_{(en,ex_n)}$ is precisely $p^*_M = \sup\{P_M(w) | w \in \Sigma^*\}$ (see the appendix). Thus, for a threshold λ, the language $L(M, \lambda) = \emptyset$ iff $q^*_{(en,ex)} \leq \lambda$; this establishes the undecidability of the quantitative problem for RMDPs.

In [BC03] it is shown that the PFA emptiness problem is undecidable even for PFAs with only 2 letters and 46 states. It follows that the quantitative problem is undecidable

for RMDPs with one component and 46 exits. A smaller number of exits, namely 10, can be obtained using our alternative 2-counter proof (although the number of components in that reduction is not fixed). Note incidentally that it is open whether the PFA emptiness problem for two-state PFAs is decidable or not [BC03,MHC03]. If it is undecidable, it would follow that the termination problem for RMDPs with two exits is also undecidable.

The inapproximability result (1.) follows from our reduction combined with a result of [CL89,MHC03] for PFAs. The number of exits in this reduction (which is the number of states of the PFA of [CL89,MHC03]) is unbounded. With a more involved construction (omitted from this extended abstract) we can make the number of exits in (1.) bounded (the bound depends on ϵ), while letting the number of components be unbounded.

For the qualitative problem (is $q^*_{(en,ex)} = 1$?) we embed the RMDP A that is constructed above from a PFA M into another RMDP A'. (Note that the qualitative problem for PFAs, i.e. determining whether $p^*_M = 1$, is decidable; this follows from a result of [ACY95].) The RMDP A' has one more component, and the embedding has the property that it turns the termination probability from $1 - \epsilon$ into 1; see the appendix for details. □

We next show undecidability of the model checking problem for 1-exit RMDPs with respect to regular (ω-regular or finite) properties specified by a property automaton B.

Theorem 7. *The qualitative and quantitative model checking problems for 1-exit RMDPs are undecidable, and in fact there is a fixed property for which this holds. Furthermore, the maximum probability cannot be approximated within any constant factor.*

Sketch: The reductions are similar to the previous theorem. One difference is that instead of using the exits of the components to keep track of the states of the PFA, we have the states of the property automaton B simulate the PFA. The only catch is that B does not have probabilities and thus the transition probabilities of the PFA have to be reflected in the RMDP. The reduction incorporates these probabilities inside the RMDP in a manner that does not use the state of the PFA (since the RMDP does not track the PFA state). □

References

[ACY95] R. Alur, C. Courcoubetis, and M. Yannakakis. Distinguishing tests for nondeterministic and probabilistic machines. In *Proc. 20th STOC*, pages 363–372, 1995.

[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.

[ATM03] R. Alur, S. La Torre, and P. Madhusudan. Modular strategies for recursive game graphs. In *TACAS*, volume 2619 of *LNCS*, pages 363–378, 2003.

[BC03] V. D. Blondel and V. Canterini. Undecidable problems for probabilistic automata of fixed dimension. *Theory of Computing Systems*, 36:231–245, 2003.

[BK98] C. Baier and M. Kwiatkowska. Model checking for a probabilistic branching time logic with fairness. *Distributed Computing*, 11(3):125–155, 1998.

[BKS05] T. Brázdil, A. Kučera, and O. Stražovský. Decidability of temporal properties of probabilistic pushdown automata. In *Proc. of 22nd STACS'05*. Springer, 2005.

[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.

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

[CL89] A. Condon and R. Lipton. The complexity of space bounded interactive proofs. In *Proc. of 30th IEEE FOCS*, 1989.

[Con92] A. Condon. The complexity of stochastic games. *Inf. & Comp.*, 96(2):203–224, 1992.

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

[CY98] C. Courcoubetis and M. Yannakakis. Markov decision processes and regular events. *IEEE Trans. on Automatic Control*, 43(10):1399–1418, 1998.

[dAKN+00] L. de Alfaro, M. Kwiatkowska, G. Norman, D. Parker, and R. Segala. Symbolic model checking of probabilistic processes using MTBDDs and the kronecker representation. In *Proc. of 6th TACAS*, pages 395–410, 2000.

[EE04] J. Esparza and K. Etessami. Verifying probabilistic procedural programs. In *Proc. FSTTCS'04*, 2004. (Invited survey paper).

[EJ91] E. A. Emerson and C. S. Jutla. Tree automata, mu-calculus, and determinacy. In *Proc. of 32nd IEEE FOCS*, pages 368–377, 1991.

[EKM04] J. Esparza, A. Kučera, and R. Mayr. Model checking probabilistic pushdown automata. In *Proc. of 19th IEEE LICS'04*, 2004.

[Ete04] K. Etessami. Analysis of recursive game graphs using data flow equations. In *Proc. 5th VMCAI*, volume 2937 of *LNCS*, pages 282–296. Springer, 2004.

[EY05a] K. Etessami and M. Yannakakis. Recursive markov chains, stochastic grammars, and monotone systems of nonlinear equations. In *Proc. of 22nd STACS'05*. Springer, 2005. (Tech. Report, U. Edinburgh, June 2004).

[EY05b] K. Etessami and M. Yannakakis. Algorithmic verification of recursive probabilistic state machines. In *Proc. 11th Int. Conf. on Tools and Algorithms for the Construction and Analysis of Systems*, 2005.

[EU48] C. J. Everett and S. Ulam. Multiplicative systems, part i., ii, and iii. Technical Report 683,690,707, Los Alamos Scientific Laboratory, 1948.

[FS02] E. Feinberg & A. Shwartz, eds. *Handbook of Markov Decision Processes*. Kluwer, 2002.

[FV97] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, 1997.

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

[HSP83] S. Hart, M. Sharir, and A. Pnueli. Termination of probabilistic concurrent programs. *ACM Trans. on Programming Languages and Systems*, 5(3):356–380, 1983.

[HU79] J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Formal Languages and Computation*. Addison-Wesley, 1979.

[Jag75] P. Jagers. *Branching Processes with Biological Applications*. Wiley, 1975.

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

[Mar98] D. A. Martin. Determinacy of Blackwell games. *J. Symb. Logic*, 63(4):1565–1581, 1998.

[MHC03] O. Madani, S. Hanks, and A. Condon. On the undecidability of probabilistic planning and related stochastic optimization problems. *Artificial Intelligence*, 147(1):5–34, 2003.

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

[Paz71] A. Paz. *Introduction to Probabilistic Automata*. Academic Press, 1971.

[Put94] M. L. Puterman. *Markov Decision Processes*. Wiley, 1994.

[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.

[SBH+94] Y. Sakakibara, M. Brown, R Hughey, I.S. Mian, K. Sjolander, R. Underwood, and D. Haussler. Stochastic context-free grammars for tRNA modeling. *Nucleic Acids Research*, 22(23):5112–5120, 1994.

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

[ZP96] U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158(1-2):343–359, 1996.

A Proofs

A.1 Proof of Theorem 1

Proof. We prove each assertion in turn.

1. That P is monotone on $\mathbb{R}^n_{\geq 0}$ follows immediately from the fact that all coefficients in the polynomials P_j defining P are non-negative, and the fact that, if $x \preceq y$, then clearly $\min_{i \in I} x_i \leq \min_{i \in I}(y_i)$, and $\max_{i \in I} x_i \leq \max_{i \in I} y_i$, for any subset $I \subseteq \{1, \ldots, n\}$. Thus, if $\mathbf{0} \preceq \mathbf{x} \preceq \mathbf{y}$ then $\mathbf{0} \preceq P(\mathbf{x}) \preceq P(\mathbf{y})$. By induction on $k \geq 0$, $\mathbf{0} \preceq \mathbf{x}^k \preceq \mathbf{x}^{k+1}$.
2. By induction on $k \geq 0$. For $k = 0$: $\mathbf{x}^1 = P(\mathbf{0})$ is an n-vector where $P_{(u,ex)}(\mathbf{0}) = 1$ if $u = ex$, and $P_{(u,ex)}(\mathbf{0}) = 0$ otherwise. Hence, for each (u, ex), $\mathbf{x}^1_{(u,ex)} = \mathbf{q}^0{}_{(u,ex)}$, the probability of reaching $\langle \epsilon, ex \rangle$ from $\langle \epsilon, u \rangle$ in at most 0 steps. Hence, also clearly, $\mathbf{x}^1_{(u,ex)} \preceq \mathbf{q}^{2^0}_{(u,ex)}$.
   Inductively, suppose $\mathbf{q}^k \preceq \mathbf{x}^{k+1} \preceq \mathbf{q}^{2^k}$. Consider $\mathbf{x}^{k+2}_{(u,ex)}$. There are 5 cases, based on what type of vertex u is:
   (a) Type I. If $u = ex$, then clearly $\mathbf{q}^k_{(u,e)} = \mathbf{q}^{k+1}_{(u,ex)} = 1$. Note that since $P_{(ex,ex)}(\mathbf{x}) = 1$, $\mathbf{x}^k_{(ex,ex)} = P^k_{(ex,ex)}(\mathbf{0}) = 1$, for all $k \geq 1$. Thus $\mathbf{q}^{k+1}_{(u,ex)} = \mathbf{x}^{k+2}$.
   (b) Type II. In this case, $\mathbf{q}^{k+1}_{(u,ex)} = \sum_v p_{u,v}\, \mathbf{q}^k_{(v,ex)}$. Thus, by inductive hypothesis

   $$\mathbf{x}^{k+2}_{(u,ex)} = P_{(u,ex)}(\mathbf{x}^{k+1}) = \sum_v p_{u,v}\, \mathbf{x}^{k+1}_{(v,ex)} \geq \sum_v p_{u,v}\, \mathbf{q}^k{}_{(v,ex)} = (\mathbf{q}^{k+1})_{(u,ex)}$$

   Likewise, by inductive hypothesis

   $$\mathbf{x}^{k+2}_{(u,ex)} = \sum_v p_{u,v}\, \mathbf{x}^{k+1}_{(v,ex)} \leq \sum_v p_{u,v}\, \mathbf{q}^{2^k}{}_{(v,ex)} \leq (\mathbf{q}^{2^{k+1}})_{(u,ex)}$$

   (c) Type III. Here, $u = (b, en) \in Call_b$, and $\mathbf{q}^{k+1}_{(u,ex)} \leq \mathbf{q}^k{}_{(en,ex')} \cdot \mathbf{q}^k_{((b,ex'),ex)} \leq \mathbf{q}^{2k}_{(u,ex)}$, where $ex' \in Ex_{Y(b)}$ is the unique exit node of $A_{Y(b)}$. To see the two inequalities hold, first note that in order to get to $\langle ex \rangle$ from $\langle u \rangle$ in at most $k+1$ steps, we must first get from the entry en of the component labeling box b to exit ex' in at most some number $m \leq k$ steps, and then get from that box-exit $\langle (b, ex') \rangle$ to $\langle ex \rangle$ in at most $m' \leq k$ steps, such that, $m + m' \leq k + 1$. In the formula for the upper bound, we have relaxed the requirements and only require that each of m and m' is $\leq k$. Hence the first inequality. For the second inequality, $\mathbf{q}^k{}_{(en,ex')} \cdot \mathbf{q}^k_{((b,ex'),ex)} \leq \mathbf{q}^{2k}_{(u,ex)}$, observe that one way to get from $\langle u \rangle$ to $\langle ex \rangle$ in at most $2k$ steps is to get from $\langle en \rangle$ to $\langle ex' \rangle$ in at most k steps, and then to get from $\langle (b, ex') \rangle$ to $\langle ex \rangle$ in at most k steps. Thus $\mathbf{q}^k{}_{(en,ex')} \cdot \mathbf{q}^k_{((b,ex'),ex)} \leq \mathbf{q}^{2k}_{(u,ex)}$. Now, by the inductive assumption, $\mathbf{q}^{2^k} \succeq \mathbf{x}^{k+1} \succeq \mathbf{q}^k$. Hence, using the inequality, and substituting, we get

   $$\mathbf{q}^{k+1}{}_{(u,ex)} \leq \mathbf{x}^{k+1}_{(en,ex')}\, \mathbf{x}^{k+1}_{((b,ex'),ex)} = P(\mathbf{x}^{k+1})_{(u,ex)} = \mathbf{x}^{k+2}{}_{(u,ex)}.$$

   We also get

   $$\mathbf{x}^{k+2}{}_{(u,ex)} = \mathbf{x}^{k+1}_{(en,ex')}\, \mathbf{x}^{k+1}_{((b,ex'),ex)} \leq \mathbf{q}^{2^k}_{(ex,ex')} \mathbf{q}^{2^k}_{((b,ex'),ex)} \leq \mathbf{q}^{2^{k+1}}_{((b,ex'),ex)}.$$

   (d) Type IV: In this case, it is easy to see that $\mathbf{q}^{k+1}_{(u,ex)} = \max_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}^k_{(v,ex)}$. Thus, by inductive hypothesis, $\mathbf{q}^{k+1}_{(u,ex)} = \max_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}^k_{(v,ex)} \leq \max_{\{v|(u,\perp,v)\in\delta\}} \mathbf{x}^{k+1}_{(v,ex)} = \mathbf{x}^{k+2}_{u,ex}$. Likewise, $\mathbf{x}^{k+2}_{u,ex} = \max_{\{v|(u,\perp,v)\in\delta\}} \mathbf{x}^{k+1}_{(v,ex)} \leq \max_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}^{2^k}_{(v,ex)} \leq \mathbf{q}^{2^{k+1}}_{(u,ex)}$.

(e) Type V: Same as IV: $\mathbf{q}^{k+1}_{(u,ex)} = \min_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}^{k}_{(v,ex)} \leq \min_{\{v|(u,\perp,v)\in\delta\}} \mathbf{x}^{k+1}_{(v,ex)} = \mathbf{x}^{k+2}_{u,ex}$. Again, like case IV, $\mathbf{x}^{k+2}_{u,ex} \leq \mathbf{q}^{2^{k+1}}_{(u,ex)}$.

We have established assertion (2).

3. Assertion (3) follows from the definition of $\mathbf{q}^*$. Suppose $\mathbf{q}^* \neq P(\mathbf{q}^*)$. The equations for vertices u of type I, II, and III, can be used to define precisely the probabilities $\mathbf{q}^*{}_{(u,ex)}$ in terms of other probabilities $\mathbf{q}^*{}_{(v,ex)}$. Thus, the only possibility is that $\mathbf{q}^*_{(u,ex)} \neq P_{(u,ex)}(\mathbf{q}^*)$ for some vertex u of Type IV or V.
Suppose u is of Type IV, i.e. a "max" vertex. Then, clearly, $\mathbf{q}^*_{(u,ex)} \geq \mathbf{q}^*_{(v,ex)}$ for any neighbor of u, with $(u,\perp,v) \in \delta$, because if $\mathbf{q}^*_{(u,ex)} < \mathbf{q}^*_{(v,ex)}$, then player 1 could play the transition $(u,\perp,v)$ at the beginning of the game M_A starting at u and improve its payoff. Likewise, $\mathbf{q}^*_{(u,ex)} \leq \mathbf{q}^*_{(v,ex)}$, for some neighbor v, because otherwise, no matter what initial move player 1 makes from u, its payoff would be less than the purported $\mathbf{q}^*_{(u,ex)}$. Similarly, suppose u is of Type V, i.e., a "min" vertex. Then, again, $\mathbf{q}^*_{(u,ex)} \leq \mathbf{q}^*_{(v,ex)}$ for any neighbor of u, with $(u,\perp,v) \in \delta$, because if $\mathbf{q}^*_{(u,ex)} > \mathbf{q}^*_{(v,ex)}$, then player 2 can switch to a strategy which, starting at u, moves initially to v, and then regardless of how player 1 plays, player 2 would have a strategy to limit the payoff to $\mathbf{q}^*_{(v,ex)} < \mathbf{q}^*_{(u,ex)}$, a contradiction. Likewise, $\mathbf{q}^*_{(u,ex)} \geq \mathbf{q}^*_{(v,ex)}$, for some neighbor v, because otherwise, no matter what initial move player 2 makes from u, player 1 can play in such a way that, no matter what player 2 does 1's ultimate payoff would be strictly greater than the purported $\mathbf{q}^*_{(u,ex)}$. Hence $\mathbf{q}^*$ is a fixed-point of P.
4. Note that P is monotonic, and that $\mathbf{q}^*$ is a fixed-point of P. Since $\mathbf{x}^0 = 0 \preceq \mathbf{q}^*$, it follows, by induction on $k \geq 0$, that $\mathbf{x}^k \preceq \mathbf{q}^*$, for all $k \geq 0$.
5. Consider any fixpoint $\mathbf{q}'$ of the equations, i.e., where $\mathbf{q}' = P(\mathbf{q}')$. We shall argue that $\mathbf{q}^* \preceq \mathbf{q}'$. Let τ' be the (stationary) strategy for player 2 that always picks, at any state $\langle\beta, u\rangle$, for vertex $u \in \mathtt{pl}^{-1}(2)$, the particular successor v of u such that $v = \arg\min_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}'_{(v,ex)}$ (breaking ties, say, lexicographically).

Lemma 1. *For all strategies $\sigma \in \Psi_1$ of player 1, and for all $k \geq 0$, $\mathbf{q}^{k,\sigma,\tau'} \preceq \mathbf{q}'$.*

Proof. By induction, similar to the proof of assertion (2). The base case $\mathbf{q}^{0,\sigma,\tau'} \preceq \mathbf{q}'$ is trivial.

(a) Type I. If $u = ex$, then for all $k \geq 0$, clearly $\mathbf{q}^{k,\sigma,\tau'}_{(ex,ex)} = \mathbf{q}'_{(ex,ex)} = 1$.

(b) Type II. Let σ' be the strategy defined by $\sigma'(\beta) = \sigma(\langle\epsilon, u\rangle\beta)$ for all $\beta \in V^*$. Then,

$$\mathbf{q}^{k+1,\sigma,\tau'}_{(u,ex)} = \sum_v p_{u,v}\, \mathbf{q}^{k,\sigma',\tau'}_{(v,ex)} \leq \sum_v p_{u,v}\, \mathbf{q}^{'}{}_{(v,ex)} = (\mathbf{q}^{'})_{(u,ex)}.$$

(c) Type III. In this case, $u = (b, en) \in Call_b$, and $\mathbf{q}^{k+1,\sigma,\tau'}_{(u,ex)} \leq max_\rho \mathbf{q}^{k,\rho,\tau'}_{(en,ex')} \cdot max_\rho \mathbf{q}^{k,\rho,\tau'}_{((b,ex'),ex)}$, where $ex' \in Ex_{Y(b)}$ is the unique exit node of $A_{Y(b)}$. Now, by the inductive assumption, $\mathbf{q}^{k,\rho,\tau'} \preceq \mathbf{q}'$ for all ρ. Moreover, since $\mathbf{q}' = P(\mathbf{q}')$, $\mathbf{q}^{'}_{(u,ex)} = \mathbf{q}^{'}{}_{(en,ex')} \cdot \mathbf{q}^{'}_{((b,ex'),ex)}$
Hence, using these inequalities and substituting, we get

$$\mathbf{q}^{k+1,\sigma,\tau'}{}_{(u,ex)} \leq \mathbf{q}^{'}_{(en,ex')}\, \mathbf{q}^{'}_{((b,ex'),ex)} = \mathbf{q}^{'}{}_{(u,ex)}.$$

(d) Type IV: In this case, starting at $\langle\epsilon, u\rangle$, whatever player 1's strategy σ is, initially it has to move to some neighbor $\langle\epsilon, v\rangle$ from which the probability of termination in at most k steps is precisely $\mathbf{q}^{k,\sigma',\tau'}_{(v,ex)}$. Thus $\mathbf{q}^{k+1,\sigma,\tau'}_{(u,ex)} \leq \max_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}^{k,\sigma',\tau'}_{(v,ex)}$. By the inductive hypothesis $\mathbf{q}^{k,\sigma',\tau'}_{(v,ex)} \leq \mathbf{q}^{'}_{(v,ex)}$, so we are done by induction.

(e) Type V: Since $\mathbf{q}' = P(\mathbf{q}')$, we know that $\mathbf{q}'_{(u,ex)} = \min_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}'_{(v,ex)}$. We also know that $\tau'(u) = v$, where $v = \arg\min_{\{v|(u,\perp,v)\in\delta\}} q'_{(v,ex)}$. But then, by the inductive hypothesis, $\mathbf{q}^{k+1,\sigma,\tau'}_{(u,ex)} = \mathbf{q}^{k,\sigma',\tau'}_{(v,ex)} \leq \mathbf{q}'_{(v,ex)} = \min_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}'_{(v,ex)} = \mathbf{q}'_{(u,ex)}$.

□

Now, by the lemma, $\mathbf{q}^{*,\sigma,\tau'} = \lim_{k\to\infty} \mathbf{q}^{k,\sigma,\tau'} \preceq \mathbf{q}'$. This holds for any strategy $\sigma \in \Psi_1$. Therefore, $\sup_{\sigma\in\Psi_1} q^{*,\sigma,\tau'}_{(u,ex)} \leq \mathbf{q}'_{(u,ex)}$, for every vertex u. Thus, by the determinacy of RSSG games, we have established that $\mathbf{q}^*_{(u,ex)} = \inf_{\tau\in\Psi_2} \sup_{\sigma\in\Psi_1} q^{*,\sigma,\tau}_{(u,ex)} \leq \mathbf{q}'_{(u,ex)}$, for all vertices u. In other words, $\mathbf{q}^* \preceq \mathbf{q}'$.

6. Finally, observe that $\lim_{k\to\infty} \mathbf{x}^k$ exists and is bounded within $[0,1]^n$. The sequence x^k, $k \to \infty$ is monotonically non-decreasing, and by definition $\lim_{k\to\infty} \mathbf{x}^k$ is a fixed point of $\mathbf{x} = P(\mathbf{x})$. By part (4), $\lim_{k\to\infty} \mathbf{x}^k \preceq \mathbf{q}^*$. Thus, by part (5), $\lim_{k\to\infty} \mathbf{x}^k = \mathbf{q}^*$. On the other hand, since $\mathbf{q}^k \preceq \mathbf{x}^{k+1} \preceq \mathbf{q}^{2^k}$, we have $\lim_{k\to\infty} \mathbf{x}^k = \lim_{k\to\infty} \mathbf{q}^k$.

□

A.2 Proof of Theorem 2

Proof. By Corollary 1, we only need to show that player 1 has an optimal S&M strategy. Let σ be any S&M strategy for player 1. Consider $\mathbf{q}^{*,\sigma} = \inf_{\tau\in\Psi_2} \mathbf{q}^{*,\sigma,\tau}$. First, let us note that if $\mathbf{q}^{*,\sigma} = P(\mathbf{q}^{*,\sigma})$ then $\mathbf{q}^{*,\sigma} = \mathbf{q}^*$. This is so because, by Theorem 1, $\mathbf{q}^* \preceq \mathbf{q}^{*,\sigma}$, and on the other hand, σ is just one strategy for player 1, and for every vertex u, $\mathbf{q}^*_{(u,ex)} = \sup_{\sigma'\in\Psi_1} \mathbf{q}^{*,\sigma'}_{(u,ex)} \geq \mathbf{q}^{*,\sigma}_{(u,ex)}$. Now, we claim that, for all vertices u that do not belong to player 1 (i.e., such that u is not of type IV) $\mathbf{q}^{*,\sigma}_{(u,ex)}$ satisfies its equation in $\mathbf{x} = P(\mathbf{x})$. In other words, $\mathbf{q}^{*,\sigma}_{(u,ex)} = P_{(u,ex)}(\mathbf{q}^{*,\sigma})$. To see this, note that for vertices u of types I,II, and III, no choice of either player is involved, thus the equation holds by definition of $\mathbf{q}^{*,\sigma}$. For nodes of Type V, which belong to player 2 (the minimizer), we have the equation $\mathbf{x}_{(u,ex)} = \min_{\{v|(u,\perp,v)\in\delta\}} \mathbf{x}_{(v,ex)}$. But note that the best player 2 can do against strategy σ, starting at $\langle\epsilon, u\rangle$, is to move to a neighboring vertex v such that $v = \arg\min_{\{v|(u,\perp,v)\in\delta\}} \mathbf{q}^{*,\sigma}_{(v,ex)}$.

Thus, the only equations that may fail are those of Type IV, of the form $\mathbf{x}_{(u,ex)} = \max_{\{v|(u,\perp,v)\in\delta\}} \mathbf{x}_{(v,ex)}$. Suppose $\sigma(u) = v$, for some neighbor v. Clearly then, $\mathbf{q}^{*,\sigma}_{(u,ex)} = \mathbf{q}^{*,\sigma}_{(v,ex)}$. Thus, $\mathbf{q}^{*,\sigma}_{(u,ex)} \leq \max_{\{v'|(u,\perp,v')\in\delta\}} \mathbf{q}^{*,\sigma}_{(v',ex)}$. Thus equality fails iff there is another vertex $w \neq v$, with $(u,\perp,w) \in \delta$, such that $\mathbf{q}^{*,\sigma}_{(v,ex)} < \mathbf{q}^{*,\sigma}_{(w,ex)}$. Consider such a vertex w, and consider now a revised S&M strategy σ', which is identical to σ, except that $\sigma'(u) = w$.

Next, consider a parameterized 1-exit RSSG, $A(t)$, which is identical to A, except that all edges out of vertex u are removed, and replaced by a single edge labeled by probability variable t to the exit of the same component. Fixing the value t determines an RSSG, $A(t)$.

Note that if we restrict the S&M strategies σ or σ' to all vertices other than u, then they both define the same S&M strategy for the RSSG $A(t)$. Define $q^{*,\sigma,\tau,t}_{(z,ex)}$ to be the probability of eventually reaching $\langle\epsilon, ex\rangle$ starting from $\langle\epsilon, z\rangle$ in the Markov chain $M^{z,\sigma,\tau}_{A(t)}$. Now, for each vertex z, define the function $f_z(t) = \inf_{\tau\in\Psi_2} q^{*,\sigma,\tau,t}_{(z,ex)}$. In other words, $f_z(t)$ is the infimum of probabilities, over all strategies of player 2, that from $\langle\epsilon, z\rangle$ we terminate in z's corresponding exit $\langle\epsilon, ex\rangle$ in $A(t)$. This probability is parameterized by t.

Now, letting $t_1 = \mathbf{q}^{*,\sigma}_{(u,ex)}$, observe that $f_z(t_1) = \mathbf{q}^{*,\sigma}_{(z,ex)}$ for every z. This is so because, any strategy for minimizing the probability of exiting from z would, upon hitting a state $\langle\beta, u\rangle$, be best off minimizing the probability of exiting from $\langle\beta, u\rangle$, i.e., reaching $\langle\beta, ex'\rangle$,

regardless the context (for, without exiting from this context, it could not hope to terminate in the empty context).

Note that, by Corollary 1, in the RSSG termination game on $A(t)$, for any value of t, and any start vertex z, player 2 has an optimal S&M strategy $\tau_{z,t}$, such that $\tau_{z,t} = \arg\min_{\tau \in \Psi_2} q^{*,\sigma,\tau}_{(z,ex)}$. Let $g_{(z,\tau)}(t) = q^{*,\sigma,\tau}_{(z,ex)}$. Note that $f_z(t) = \min_\tau g_{z,\tau}(t)$, where the minimum is over S&M strategies. Now, note that the function $g_{z,\tau}(t)$ is the probability of reaching an exit in an RMC starting from a particular vertex. Thus, by [EY05a], $g_{z,\tau}(t) = (\lim_{k\to\infty} R^k(\mathbf{0}))_{(z,ex)}$ for a polynomial system $\mathbf{x} = R(\mathbf{x})$ with non-negative coefficients, but with the additional feature that the variable t appears as one of the coefficients. Since this limit can be described by a power series in the variable t with non-negative coefficients, $g_{z,\tau}(t)$ has the following properties: it is a continuous, differentiable, and nondecreasing function of $t \in [0,1]$, with continuous and nondecreasing derivative, $g'_{z,\tau}(t)$ and since the limit defines probabilities we also know that for $t \in [0,1]$, $g_{z,\tau}(t) \in [0,1]$. Thus $g_{z,\tau}(0) \geq 0$ and $g_{z,\tau}(1) \leq 1$.

Hence, since $g'_{z,\tau}(t)$ is non-decreasing, if for some $t \in [0,1]$, $g_{z,\tau}(t) > t$, then for all $t' < t$, $g_{z,\tau}(t') > t'$. (To see this, note that if $g_{z,\tau}(t) > t$ and $g'_{z,\tau}(t) \geq 1$, then for all $t'' > t$, $g_{z,\tau}(t'') > t''$, which contradicts the fact that $g_{z,\tau}(1) = 1$. Thus $g'_{z,\tau}(t') < 1$ for all $t' \leq t$, and since $g_{z,\tau}(t) > t$, we also have $g_{z,\tau}(t') > t'$ for all $t' < t$. Similarly, if $g_{z,\tau}(t) < t$ for some t, then $g_{z,\tau}(t'') < t''$ for all $t'' \in [t,1)$. To see this, note that if for some $t'' > t$, $t'' < 1$, $g_{z,\tau}(t'') = t''$, then since $g'_{z,\tau}$ is non-decreasing and $g_{z,\tau}(t) < t$, it must be the case that $g'_{z,\tau}(t'') > 1$. But then $g_{z,\tau}(1) > 1$, which is a contradiction.

It follows that $f_z(t)$ also has the same properties, namely: if $f_z(t) > t$ at some point $t \in [0,1]$ then $g_{z,\tau}(t) > t$ for all τ, and hence for all $t' < t$ and for all τ, $g_{z,\tau}(t') > t'$, and thus also $f_z(t') > t'$ for all $t' \in [0,t]$. On the other hand, if $f_z(t) < t$ at $t \in [0,1]$, then there must be some τ' such that $g_{z,\tau'}(t) < t$. Hence $g_{z,\tau'}(t'') < t''$, for all $t'' \in [t,1)$, and hence also $f_z(t'') < t''$ for all $t'' \in [t,1)$.

Let $t_2 = q^{*,\sigma}_{(w,ex)}$. We know $t_2 = q^{*,\sigma}_{(w,ex)} = f_w(t_1) > t_1 = q^{*,\sigma}_{(v,ex)}$. Therefore $f_w(t) > t$ for all $t < t_1$ by the above property. And in fact, $f_w(t) > t$ for all $t < t_2$, because f_w is nondecreasing. Therefore, the least fixed point (i.e., least solution) of $f_w(t) = t$ is $\geq t_2$. Now if we switch strategy σ to σ', where $\sigma'(u) = w$, then $q^{*,\sigma'}_{(w,ex)}$ is a fixed point, t_3, of $f_w(t) = t$, so $t_3 \geq t_2 > t_1$ and $q^{*,\sigma'}_{(z,ex)} = f_z(t_3) \geq q^{*,\sigma}_{(z,ex)}$ for all z, with strict inequality for u: $q^{*,\sigma'}_{(u,ex)} = t_3 > q^{*,\sigma}_{(u,ex)} = t_1$. Thus, switching to the new S&M strategy σ', we get $\mathbf{q}^{*,\sigma'}$ which dominates $\mathbf{q}^{*,\sigma}$, and is strictly greater in some coordinate. But there are only a finite number of S&M strategy, thus repeating this process we must eventually get to S&M strategy σ^* that can't be improved in this way. Thus $\mathbf{q}^{*,\sigma^*} = P(\mathbf{q}^{*,\sigma^*})$, and hence by our earlier claim $\mathbf{q}^{*,\sigma^*} = \mathbf{q}^*$. Thus, player 1 has an optimal S&M strategy. □

A.3 Proof of Theorem 3

Proof. For part 1, we compute the set T of vertices u for which $q^*_{(u,ex)} = 0$ as follows. Initialize set $S := Ex$. Repeat the following until there is no change in the set S:

- If a probabilistic vertex or a player 1 vertex u has a successor in S then add u to S.
- If a player 2 vertex u has all successors in S then add u to S.
- If $u = (b, en)$ is a call port of box b and both en (the entry of the corresponding component) and the return port (b, ex') of the box are in S then add u to S.

When the process finishes, we let $T = Q - S$. This is the set of vertices that have value 0. Player 2's strategy is to pick for each $u \in T$ a successor that is also in T. Then the vertices that are in T have no way of reaching their exits, and these are all such vertices.

The proof of parts 2 and 3 follows from the S&M determinacy of 1-exit RSSG termination games. For 1-exit RMDPs, for maximization, we guess the S&M strategy which gives probability 1. This leaves us to check termination with probability 1 in a 1-exit RMC. In [EY05a] it was shown that this can be achieved in polynomial time. For minimization, we can guess the S&M strategy that gives probability < 1, and verify that this is the case using the algorithm of [EY05a]. Note that we cannot guess the probabilities themselves (the least fixpoint of the equations) because they may not even be rational.

For 1-exit RSSGs, since the game is S&M determined, we can check whether the value of the game is 1 both in Σ_2^P and in Π_2^P. To check that the value is 1 in Σ_2^P, guess a S&M strategy σ for the max player, and then verify that for every S&M strategy τ of the min player the termination probability in the RMC $G[\sigma, \tau]$ obtained from the given game G by fixing the strategies to σ and τ is 1. Note that this guarantees that the min player has no strategy whatsoever (even no strategy with memory) that achieves value less than 1 against σ. To check that the value is 1 in Π_2^P, or equivalently to check that the value is < 1 in Σ_2^P, guess a S&M strategy τ for the min player and verify that for every strategy σ of the max player, the probability of termination in the RMC $G[\sigma, \tau]$ is less than 1. □

A.4 Proof-sketch of Theorem 6

We will reduce from the emptiness problem for PFA. Given a PFA M, let $p_M^* = \sup\{P_M(w) | w \in \Sigma^*\}$. We will construct a RMDP A with entry en and exit ex such that $p_M^* = q_{(en,ex)}^*$. Thus, for a threshold λ, the language $L(M, \lambda) = \emptyset$ iff $q_{(en,ex)}^* \leq \lambda$; this establishes the undecidability of the quantitative problem for RMDPs. For the qualitative problem ($q_{(en,ex)}^* = 1$?) we will embed A into another RMDP A'. (Note that the qualitative problem for PFAs, i.e. determining whether $p_M^* = 1$, is decidable; this follows from a result of [ACY95].)

Let $M = (Q, \Sigma, T, q_1, q_n)$ be a PFA with n states. Define a RMDP A that has one component (call it also A) with a single entry en, and n exits $ex_1, \ldots, ex_n$, one for each state of M. The entry en is a max node and has edges to the call ports of a set of $|\Sigma|$ boxes $b_a, a \in \Sigma$; all the boxes are of course mapped to the single component A. In addition en has an edge to the exit ex_1. The return ports of the boxes b_a are probabilistic vertices. Each return port (b_a, ex_i) has an edge to each exit ex_j with probability $T_a[i, j]$. This concludes the definition of the RMDP A. Starting from the entry en of A, the max player wants to maximize the probability $q_{(en,ex_n)}^*$ of terminating at exit ex_n. We claim that $q_{(en,ex_n)}^*$ is precisely $p_M^* = \sup\{P_M(w) | w \in \Sigma^*\}$.

For any word $w \in \Sigma^*$, consider the strategy σ_w of the max player which at the i-th step chooses the edge to the box corresponding to the ith letter of the reverse word w^R for $i = 1, \ldots, |w|$; at step $|w| + 1$, the max player chooses the edge to the exit ex_1. From that point on, all the actions are probabilistic. It is easy to see then that by the construction, the sequence of probabilistic actions corresponds to a run of the PFA M on the input string w, and the run ends at the accepting state q_n iff the RMDP terminates at the exit ex_n of the top component. Thus, the probability $q_{(en,ex_n)}^{*,\sigma_w}$ of termination at ex_n under strategy σ is $P_M(w)$. It follows that $p_M^* \leq q_{(en,ex_n)}^*$.

Conversely, consider a strategy σ of the max player. At each step, when the process is at an entry of a box, the max player has to either choose a letter of Σ and transition to the call port of the corresponding box or decide to exit the current box. In the latter case,

the control passes to the probabilistic player and stays with him for the remainder of the game until the process reaches an exit of the top component. If $q^{*,\sigma}_{(en,ex_n)} > 0$, then the max player has to choose the edge to the exit at some point (otherwise the process will never terminate). Let w be the reverse of the sequence of actions of the max player. Then $\sigma = \sigma_w$ and $q^{*,\sigma}_{(en,ex_n)} = P_M(w)$. Therefore, $q^*_{(en,ex_n)} \leq p^*_M$ and the two quantities are equal.

In [BC03] it is shown that the PFA emptiness problem is undecidable even for PFAs with only 2 letters and 46 states. It follows that the quantitative problem is undecidable for RMDPs with one component and 46 exits. A smaller number of exits, namely 10, can be obtained using our alternative 2-counter proof, although the number of components in that reduction is not fixed.

In [CL89,MHC03] it is shown that for every $\epsilon > 0$ it is undecidable for a given PFA M to distinguish between the case that M accepts some word with probability $> 1 - \epsilon$ and the case that M accepts no word with probability $< \epsilon$. This fact together with our reduction implies statement (1.) of the theorem. The number of exits in this reduction is unbounded. With a more involved construction (omitted from this extended abstract) we can make the number of exits bounded (the bound depends on ϵ), while letting the number of components be unbounded.

For the qualitative RMDP problem, given a PFA M and threshold λ define a RMDP A' as follows. Assume that $\lambda \leq 1/2$; a similar construction applies if $\lambda > 1/2$. The RMDP A' has two components A_1, A_2. Component A_2 is the same as A of the above construction. A_1 has a single entry en and exit ex. The entry en is probabilistic and has an edge with probability $p = 1/(2 - 2\lambda)$ to a box b_2 mapped to A_2 and an edge with probability $1 - p$ to the exit ex. The return port (b_2, ex_n) of b_2 has a probability 1 edge to the exit ex, while all the other return ports (b_2, ex_i) have probability 1 edges to the entry of a box b_1 mapped to A_1. The return port (b_1, ex) has a probability 1 edge to another box b'_1 mapped also to A_1 and the return port (b'_1, ex) of b'_1 has a probability 1 edge to the exit ex of A_1. This concludes the definition of A'. It can be shown that the supremum termination probability $q^*_{(en,ex)}$ in A' (over all strategies of the max player) is 1 if and only if $p^*_M \geq \lambda$. For the inapproximability result, set $\lambda = 1/2$ (i.e. $p = 1$) in the above construction. If $p^*_M \geq 1/2$ then $q^*_{(en,ex)} = 1$, while if $p^*_M < \delta = \epsilon/(1 + \epsilon)$ then $q^*_{(en,ex)} < \delta/(1 - \delta) = \epsilon$. $\square$