On the Complexity of Reachability in Parametric Markov Decision Processes

Tobias Winkler
RWTH Aachen University, Germany
tobias.winkler1@rwth-aachen.de

Sebastian Junges 📍
RWTH Aachen University, Germany
sebastian.junges@cs.rwth-aachen.de

Guillermo A. Pérez 📍
University of Antwerp, Belgium
guillermoalberto.perez@uantwerpen.be

Joost-Pieter Katoen 📍
RWTH Aachen University, Germany
katoen@cs.rwth-aachen.de

──── **Abstract** ────

This paper studies parametric Markov decision processes (pMDPs), an extension to Markov decision processes (MDPs) where transitions probabilities are described by polynomials over a finite set of parameters. Fixing values for all parameters yields MDPs. In particular, this paper studies the complexity of finding values for these parameters such that the induced MDP satisfies some reachability constraints. We discuss different variants depending on the comparison operator in the constraints and the domain of the parameter values. We improve all known lower bounds for this problem, and notably provide ETR-completeness results for distinct variants of this problem. Furthermore, we provide insights in the functions describing the induced reachability probabilities, and how pMDPs generalise concurrent stochastic reachability games.

1 Introduction

Markov decision processes (MDPs) are *the* model to reason about sequential processes under (stochastic) uncertainty and non-determinism. Markov chains (MCs) are MDPs without non-determinism. Often, probability distributions in these models are difficult to assess precisely during design time of a system. This shortcoming has led to interval MCs [15, 35, 50, 54] and interval MDPs (also known as Bounded-parameter MDPs) [27, 42, 58], which allow for interval-labelled transitions. Analysis under interval Markov models is often too pessimistic: The actual probabilities on the transitions are considered to be non-deterministically and *locally* chosen. Intuitively, consider the probability of a coin-flip yielding heads in some stochastic environment. In interval models, the probability may vary with the local memory state of an agent acting in this environment. Such behaviour is unrealistic. *Parametric*

MCs/MDPs [19, 23, 28, 39] (pMCs, pMDPs) overcome this limitation by adding dependencies (or couplings) between various transitions – they add global restrictions to the selection of the probability distributions. Intuitively, the probability of flipping heads can be arbitrary, but should be independent of an agent's local memory. Such couplings are similar to restrictions on schedulers in decentralised/partially observable MDPs, considered in e.g., [5, 26, 51].

Technically, pMDPs label their transitions with polynomials over a finite set of parameters. Fixing all parameter values yields MDPs. The synthesis problem considered in this paper asks to find parameter values such that the induced MDPs satisfy reachability constraints. Such reachability constraints state that the probability – under some/all possible ways to resolve non-determinism in the MDP – to reach a target state is (strictly) above or below a threshold. A sample synthesis problem is thus: "Are there parameter values such that for all possible ways to resolve the non-determinism, the probability to reach a target state exceeds $\frac{1}{2}$?" Variants of the synthesis problem are obtained by varying the reachability constraints, and the domain of the parameter values. Parameter synthesis is supported by the model checkers PRISM [38] and Storm [22], and dedicated tools PARAM [29] and PROPhESY [21]. The complexity of the decision problems corresponding to parameter synthesis is mostly open.

This paper significantly extends complexity results for parameter synthesis in pMCs and pMDPs. Table 1 on page 5 gives an overview of new results: Most prominently, it establishes ETR-completeness of reachability problems for pMCs with non-strict comparison operators, and establishes NP-hardness for pMCs with strict comparison operators. For pMDPs with demonic non-determinism, it establishes ETR-completeness for any comparison operator. For angelic non-determinism, mostly the synthesis problems are equivalent to their pMC counterparts. When considering pMDPs with a fixed number of variables, we establish uniform NP upper bounds for parameter synthesis under angelic or demonic non-determinism. These results are partially based on properties of pMDPs scattered in earlier work, and use a strong connection between polynomial inequalities and parameter synthesis.

Finally, pMDPs are interesting generalisations of other models: [37] shows that parameter synthesis in pMCs is equivalent to the synthesis of finite-state controllers (with a-priori fixed bounds) of partially observable MDPs (POMDPs) [46] under reachability constraints. Thus, as a side product we improve complexity bounds [10, 56] for (a-priori fixed) memory bounded strategies in POMDPs. In this paper, we show how pMDPs generalise concurrent stochastic reachability games [12, 20, 52]. We finish the paper by drawing some connections with robust schedulers, i.e. the question of how to optimally resolve non-determinism taking into account the uncertainty in the stochastic dynamics. Proofs are given in the related technical report.

Related work.   Various results in this paper extend work by Chonev [16], who studied a model of augmented interval Markov chains. These coincide with parametric Markov chains. The work also builds upon results by Hutschenreiter *et al.* [33], in particular upon the result that pMCs with an a-priori fixed number of parameters can be checked in P. Furthermore, they study the complexity of PCTL model checking of pMCs. The complexity of finite-state controller synthesis in POMDPs has been studied in [10, 56]. Some of the proofs for ETR-completeness presented here reuse ideas from [48].

Methods (and implementations) to analyse pMCs by computing their characteristic *solution function* are considered in [19, 21, 23–25, 29, 33, 34]. Sampling-based approaches to find feasible instantiations in pMDPs are considered by [14, 28], while [3, 18] utilise optimisation methods. Finally, [44] presents a method to prove the absence of solutions in pMDPs by iteratively considering simple stochastic games [17]. Some other works on Markov models with structurally equivalent yet parameterised dynamics include [8, 9, 13, 53]. Parameter synthesis with statistical guarantees has been explored in, e.g., [6]. Further work on parameter synthesis in Markov models has been surveyed in [36].

2   Preliminaries

Let X be a finite set of variables. Let $\mathbb{Q}[X]$ and $\mathbb{Q}(X)$ denote the set of all rational-coefficient polynomial and rational functions on X, respectively. A rational function f/g can be represented as a pair (f, g) of polynomials. In turn, a polynomial can be represented as a sum of terms, where each term is given by a *coefficient* and a *monomial*. The (total) degree of a polynomial is the maximum over the sum of the exponents in the monomials. A polynomial is quadratic (respectively, quadric), if its total degree is two (four) or less. For a rational function $f(x_1, \ldots, x_k) \in \mathbb{Q}(X)$ and an instantiation $val \colon X \to \mathbb{R}$ we write $f[val]$ for the value $f(val(x_1), \ldots, val(x_k))$. We use $\bowtie$ to denote either of $\{\leq, <, \geq, >\}$ and $\rhd$ for either $\{\geq, >\}$ (and $\lhd$ analogously). With $\overline{\bowtie}$, we denote the complement, e.g. $\overline{\leq} \, = >$.

Consider a finite set S. Let $\mathrm{Distr}(S)$ denote the set of all distributions over S, and $\mathrm{supp}(\delta) \subseteq S$ the support $\{s \in S \mid \delta(s) > 0\}$ of distribution $\delta \in \mathrm{Distr}(S)$.

2.1   Parametric Markov models

▶ **Definition 1** (pMDP). *A parametric Markov Decision Process $\mathcal{M}$ is a tuple (S, X, Act, s_ι, P) with S a (finite) set of* states*, X a finite set of* parameters*, Act a finite set of* actions*, $s_\iota \in S$ the initial state, and $P \colon S \times Act \times S \to \mathbb{Q}[X] \cup \mathbb{R}$ the probabilistic transition function.*

Parameter-free pMDPs coincide with standard MDPs, as in [43]. We define $Act(s) = \{a \in Act \mid \exists s' \in S.\ P(s, a, s') \neq 0\}$. If $|Act(s)| = 1$ for all $s \in S$, then $\mathcal{M}$ is a *parametric Markov chain* (pMC). We denote its transitions with $P(s, s')$ and omit the actions.

A pMDP is *simple* if and only if non-constant probabilities labelling transitions (s, a, s') are of the form x or $1 - x$, and the sum of outgoing transitions from a state-action pair always is (equivalent to) 1. Formally, simple pMDPs satisfy the following two properties:
-   $P(s, a, s') \in \{x, 1 - x \mid x \in X\} \cup \mathbb{R}$ for all $s, s' \in S$ and $a \in Act$; and
-   $\sum_{s' \in S} P(s, a, s') = 1$ for all $s \in S$ and $a \in Act(s)$.

▶ **Definition 2** (Instantiation). *Let $\mathcal{M} = (S, X, Act, s_\iota, P)$ be a pMDP. An instantiation $val \colon X \to \mathbb{R}$ is* well-defined *if the induced functions $P(s, a, \cdot)$ are distributions over S, i.e.*

$$\forall s, s' \in S, \forall a \in Act.\ 0 \leq P(s, a, s')[val] \in \mathbb{R} \wedge \sum_{\hat{s} \in S} P(s, a, \hat{s})[val] = 1.$$

Let $\mathcal{M}[val]$ denote the parameter-free MDP in which $P(s, a, s')$ has been replaced by $P(s, a, s')[val]$. We denote with $P_{=0}^{val} \coloneqq \{(s, a, s') \in S \times Act \times S \mid P(s, a, s') \neq 0 \wedge P(s, a, s')[val] = 0\}$ the transitions of $\mathcal{M}$ that become 0 in $\mathcal{M}[val]$. A well-defined instantiation val is *graph-preserving* if the topology of the pMDP is preserved, i.e. if $P_{=0}^{val} = \emptyset$.

The (well-defined) parameter space $\mathcal{P}_{\mathcal{M}}^{\mathrm{wd}}$ for $\mathcal{M}$ is $\{val \colon X \to \mathbb{R} \mid val \text{ is well defined}\}$ and the graph-preserving parameter space $\mathcal{P}_{\mathcal{M}}^{\mathrm{gp}} \coloneqq \{val \colon X \to \mathbb{R} \mid val \text{ is graph-preserving}\}$. In simple pMDPs, the well-defined (respectively, graph-preserving) parameter space is the set of instantiations $val \colon X \to [0, 1]$ (respectively, $val \colon X \to (0, 1)$). We omit the subscript from $\mathcal{P}_{\mathcal{M}}^{\mathrm{gp}}$ and $\mathcal{P}_{\mathcal{M}}^{\mathrm{wd}}$ when the pMDP $\mathcal{M}$ is understood from the context.

A *graph-consistent region* R for $\mathcal{M}$ is a subset of $\mathcal{P}^{\mathrm{wd}}$ such that all instantiations in R induce the same graph, i.e., $P_{=0}^{val} = P_{=0}^{val'}$ for all $val, val' \in R$.

▶ **Remark 3.** For any simple pMDP, $\mathcal{P}^{\mathrm{wd}}$ can be partitioned into $3^{|X|}$ many graph-consistent regions R. For any graph-consistent region, we can (in linear time) construct a simple pMDP $\mathcal{M}'$ such that the graph-consistent region R corresponds to $\mathcal{P}_{\mathcal{M}'}^{\mathrm{gp}}$. Essentially, the construction merely removes the transitions $P_{=0}^{val}$ for $val \in R$, and adjusts the probabilities of some other transitions to 1 to ensure simplicity. A complete construction is given in [36].

Reachability, schedulers and induced Markov chains.   Consider a parameter-free MC $\mathcal{M}$ and a state s_0. A *run* of $\mathcal{M}$ from s_0 is an infinite sequence of states $s_0 s_1 \ldots$ such that $P(s_i, s_{i+1}) > 0$ for all $i \geq 0$. We denote by Runs^{s_0} the set of all runs of $\mathcal{M}$ that start with the state s_0. The *probability* of a measurable *event* $E \subseteq \mathrm{Runs}^{s_0}$ is defined using a standard cylinder construction [2, 43]. Let $\mathrm{Pr}_{\mathcal{M}}(\Diamond T)$ denote the probability to *eventually reach T* from the initial state of $\mathcal{M}$; and $\mathrm{Pr}_{\mathcal{M}}(s \to \Diamond T)$ denote the probability to eventually reach T starting from state s. We omit the subscript $\mathcal{M}$ if it is clear from the context.

To define reachability in pMDPs, we need to eliminate the non-determinism. We do so by means of a *scheduler* (a.k.a. a *policy* or *strategy*).

▶ **Definition 4** (Scheduler). *A randomised (memoryless) scheduler is a function $\sigma \colon S \to$ $\mathrm{Distr}(Act)$ s.t. $\mathrm{supp}(\sigma(s)) \subseteq Act(s)$. A scheduler is deterministic if $|\mathrm{supp}(\sigma(s))| = 1$ (i.e. $\sigma(s)$ is Dirac) for every $s \in S$. We refer to deterministic schedulers as* schedulers.

We denote the set of randomised schedulers with $R\Sigma$, and (deterministic) schedulers with Σ.

For pMDP $\mathcal{M} = (S, X, Act, s_\iota, P)$ and $\sigma \in R\Sigma_{\mathcal{M}}$, the *induced pMC* $\mathcal{M}^\sigma$ is defined as (S, X, s_ι, P') with $P'(s, s') = \sum_{a \in Act} \sigma(s)(a) \cdot P(s, a, s')$. For simple pMDPs, the induced pMC of a deterministic scheduler is simple. Under randomised schedulers, the induced pMC can be transformed into a simple pMC (e.g. [37]). We abbreviate $\mathrm{Pr}_{\mathcal{M}^\sigma}$ by $\mathrm{Pr}_{\mathcal{M}}^\sigma$.

▶ **Remark 5.** Deterministic schedulers dominate randomised schedulers for reachability properties [43], i.e. for each MDP there exists a deterministic scheduler σ s.t. $\mathrm{Pr}_{\mathcal{M}}^\sigma(\Diamond T) = \sup_{\sigma' \in R\Sigma} \mathrm{Pr}_{\mathcal{M}}^{\sigma'}(\Diamond T)$. Therefore, in the remainder, we focus on deterministic schedulers.

▶ **Definition 6** (Solution function). *For a pMC $\mathcal{M}$ and a state s, let the* solution function $sol_s^{\mathcal{M},T} \colon \mathcal{P}_{\mathcal{M}}^{\mathrm{wd}} \to [0,1]$ *be defined as* $sol_s^{\mathcal{M},T}[val] := \mathrm{Pr}_{\mathcal{M}[val]}(s \to \Diamond T)$. *For a pMDP $\mathcal{M}$, let* $minsol_s^{\mathcal{M},T}[val] := \min_{\sigma \in \Sigma} sol_s^{\mathcal{M}^\sigma,T}[val] = \min_{\sigma \in \Sigma} \mathrm{Pr}_{\mathcal{M}[val]}^\sigma(s \to \Diamond T)$. *We define* $maxsol_s^{\mathcal{M},T}$ *analogously as the maximum.*

Let $sol^{\mathcal{M},T}$ denote $sol_{s_\iota}^{\mathcal{M},T}$ with the convention that T is omitted whenever it is clear from the context. On $\mathcal{P}^{\mathrm{gp}}$, $sol^{\mathcal{M}}$ is described by a rational function over the parameters [19, 39], and is computable in $\mathcal{O}\left(poly(|S| \cdot d)^{|X|}\right)$, where d is the maximal degree of polynomials in $\mathcal{M}$'s transitions [33]. The number of resulting monomials is polynomial in $|S|$ and d but exponential in $|X|$. Furthermore, the degree of f and g in the resulting function f/g is upper-bounded by $\ell(d)$ – where ℓ is a linear function.[1] For acyclic pMCs, $sol^{\mathcal{M}}$ is described by a polynomial.

2.2 Existential theory of the reals

Many results in this paper are based on results from the existential theory of the reals [4]. We give a brief recap. We consider the first-order theory of the reals: the set of all valid sentences in the first-order language $(\mathbb{R}, +, \cdot, 0, 1, <)$. The existential theory of the reals restricts the language to (purely) existentially quantified sentences. The complexity of deciding membership, i.e. whether a sentence is (true) in the theory of the reals, is in PSPACE [7] and NP-hard. A careful analysis of its complexity is given in [45]. In particular, deciding membership for sentences with an a-priori fixed upper bound on the number of variables is in polynomial time. ETR denotes the complexity class [48] of problems with a polynomial-time many-one reduction to deciding membership in the existential theory of the reals.

---

[1] Importantly, this means that if the coefficients and exponents were written in binary for the given pMC then linearly more bits suffice to do the same for the computed rational function.

Table 1 The complexity landscape for reachability in simple pMDPs. All problems are in ETR.

		Fixed # parameters	Arbitrary # parameters	
			well-defined	graph-preserving
pMC	$\exists\text{Reach}^{\geq/\leq}$	in P [33]	— ETR-complete [Thm. 27] —	
	$\exists\text{Reach}^{>}$	"	NP-hard [Thm. 33]	$\exists\text{Reach}_{\text{wd}}^{>}$-complete [Lem. 32]
	$\exists\text{Reach}^{<}$	"	NP-hard [Thm. 33]	$\exists\text{Reach}_{\text{wd}}^{>}$-complete [Lem. 9]
pMDP	$\exists\exists\text{Reach}^{\geq/\leq}$	in NP [Thm. 13]	— ETR-complete — (trivial)	
	$\exists\exists\text{Reach}^{>}$	"	— $\exists\text{Reach}_{\text{wd}}^{>}$-complete [Lem. 34, Cor. 36] —	
	$\exists\exists\text{Reach}^{<}$	"	$\exists\text{Reach}_{\text{wd}}^{<}$-complete [Lem. 34]	$\exists\text{Reach}_{\text{wd}}^{>}$-hard (trivial)
	$\exists\forall\text{Reach}^{\bowtie}$	in NP [Thm. 14]	— ETR-complete [Thm. 37] —	

3 Problem landscape

In this section, we introduce the family of decision problems of our main interest. Let a *simple* pMDP $\mathcal{M}$ with all constants rational, and a set T of target states be the given input. We analyse the decision problems according to whether the set X of parameters from $\mathcal{M}$ has bounded size – with a-priori fixed bound – or arbitrary size.

It remains for us to fix an encoding for rational functions. Henceforth, we assume the coefficients, exponents and constants are all given as binary-encoded integer pairs.

Decision problems. The first problem is the existence of so-called *robust parameter values* or lack thereof. More precisely, the question is whether some instantiation of $\mathcal{M}$ is such that its maximal or minimal probability of eventually reaching T compares with $\frac{1}{2}$ in some desired way. In symbols, for $\mathcal{Q}_1, \mathcal{Q}_2 \in \{\exists, \forall\}$ and $\bowtie \in \{\leq, <, >, \geq\}$, let

$$\mathcal{Q}_1 \mathcal{Q}_2 \text{Reach}_{\text{wd}}^{\bowtie} \overset{\text{def}}{\iff} \mathcal{Q}_1 \; val \in \mathcal{P}^{\text{wd}}, \mathcal{Q}_2 \; \sigma \in \Sigma. \; \Pr_{\mathcal{M}[val]}^{\sigma}(\lozenge T) \bowtie \frac{1}{2}$$

be the problem of interest. We write $\mathcal{Q}_1 \mathcal{Q}_2 \text{Reach}_{\text{gp}}^{\bowtie}$ whenever $\mathcal{Q}_1$ quantifies over graph-preserving instantiations. We write $\mathcal{Q}_1 \mathcal{Q}_2 \text{Reach}_*^{\bowtie}$ to denote both the wd and gp variants. Furthermore, if $\mathcal{M}$ is a pMC we omit the second quantifier, e.g. $\exists\text{Reach}_*^{<}$. Table 1 surveys the results.

▶ **Proposition 7.** *For every* $\mathcal{Q}_1, \mathcal{Q}_2 \in \{\exists, \forall\}$ *and* $\bowtie \in \{\leq, <, >, \geq\}$, $\mathcal{Q}_1 \mathcal{Q}_2 \text{Reach}_*^{\bowtie}$ *are decidable in ETR.*

Proof. Both $\exists\forall\text{Reach}_*^{\bowtie}$ and $\exists\exists\text{Reach}_*^{\bowtie}$ are in ETR (for an encoding, see Appendix B in the technical report). It follows that $\exists\text{Reach}_*^{\bowtie}$ are also in ETR. ◀

Problems with fixed threshold. In the above-defined problems, we have fixed a threshold of $\frac{1}{2}$. This is no loss of generality as any *given* rational threshold can be reduced to $\frac{1}{2}$:

▶ Remark 8. An arbitrary threshold $0 < \lambda < 1$, $\lambda \in \mathbb{Q}$, is reducible to $\frac{1}{2}$ by the constructions depicted in Fig. 1: If $\lambda \leq \frac{1}{2}$ then we prepend a transition with probability $p = 2\lambda$ to the initial state and with probability $1 - p$ to a sink state. Otherwise, if $\lambda > \frac{1}{2}$, we prepend a transition with probability $q = 2(1 - \lambda)$ to the initial state and $1 - q$ to the target state. Conversely, the $\frac{1}{2}$ threshold may analogously be reduced to an arbitrary threshold $0 < \lambda < 1$.

(a) $\lambda \leq \frac{1}{2}$ **(b)** $\lambda > \frac{1}{2}$

Figure 1 Reductions to reachability threshold $\lambda = \frac{1}{2}$, cf. Remark 8.

Considerations for the comparison relations.

▶ **Lemma 9.** *For every* $\mathcal{Q}_1, \mathcal{Q}_2 \in \{\exists, \forall\}$*, there are polynomial-time Karp reductions*

- *among the problems* $\mathcal{Q}_1 \mathcal{Q}_2 \mathrm{Reach}_{\mathrm{gp}}^{>}$ *and* $\mathcal{Q}_1 \mathcal{Q}_2 \mathrm{Reach}_{\mathrm{gp}}^{<}$ *and*
- *among the problems* $\mathcal{Q}_1 \mathcal{Q}_2 \mathrm{Reach}_{\mathrm{gp}}^{\geq}$ *and* $\mathcal{Q}_1 \mathcal{Q}_2 \mathrm{Reach}_{\mathrm{gp}}^{\leq}$.

The above claim only holds when restricted to graph-preserving parameter spaces.

Semi-continuity.   The following theorem formalises an observation in [37, Thm. 5].

▶ **Theorem 10.** *For each simple pMC* $\mathcal{M}$*, the function* $\mathrm{sol}^{\mathcal{M}}$ *is lower semi-continuous, and continuous on* $\mathcal{P}_{\mathcal{M}}^{\mathrm{gp}}$*. For acyclic simple pMCs* $\mathcal{M}$*,* $\mathrm{sol}^{\mathcal{M}}$ *is continuous on* $\mathcal{P}_{\mathcal{M}}^{\mathrm{wd}}$*.*

Continuity on $\mathcal{P}_{\mathcal{M}}^{\mathrm{gp}}$ follows as $\mathrm{sol}^{\mathcal{M}}$ is a rational function bounded by $[0, 1]$ on all well-defined points [44]. Graph non-preserving instantiations might yield additional sink states in the induced MC, therefore, the probability may drop when changing a parameter instantiation, e.g. $p = 0$ with an single outgoing transition with probability p and a self-loop with probability $1 - p$. The semi-continuity is the main reason that we do not have symmetric entries for upper and lower bounds in Table 1.

The following result follows immediately from properties of (semi-)continuous functions.

▶ **Corollary 11.** *For all pMDPs, the functions* $\mathrm{minsol}^{\mathcal{M}}$ *and* $\mathrm{maxsol}^{\mathcal{M}}$ *are lower semi-continuous and have a minimum. For acyclic pMDPs, these functions are continuous.*

▶ **Corollary 12.** *For acyclic pMCs,* $\mathrm{sol}^{\mathcal{M}}$ *is described by a polynomial even on* $\mathcal{P}_{\mathcal{M}}^{\mathrm{wd}}$*.*

4   Fixing the number of parameters

In this section, we assume that the number of parameters is fixed. We focus ourselves on graph-preserving instantiations, as the analysis of pMDP $\mathcal{M}$ and $\mathcal{P}_{\mathcal{M}}^{\mathrm{wd}}$ corresponds to analysing constantly many pMDPs $\mathcal{M}'$ on $\mathcal{P}_{\mathcal{M}'}^{\mathrm{gp}}$, cf. Rem. 3.

Upper bounds.   Below, we establish NP membership for all variants.

▶ **Lemma 13.** *In the fixed parameter case,* $\exists\exists\mathrm{Reach}_{*}^{\bowtie}$ *is in NP.*

Proof. Guess a memoryless scheduler. Construct the induced pMC, and verify it in P.   ◀

▶ **Theorem 14.** *In the fixed parameter case,* $\exists\forall\mathrm{Reach}_{*}^{\bowtie}$ *is in NP.*

In the non-parametric case, a scheduler σ of an MDP is called *minimal* if it minimises $\mathrm{Pr}^{\sigma}(\lozenge T)$, i.e. if $\sigma \in \mathrm{argmin}_{\sigma' \in \Sigma} \mathrm{Pr}^{\sigma'}(\lozenge T)$. Consider the probabilities $x_s = \mathrm{Pr}^{\sigma}(s \to \lozenge T)$ for $s \in S$. It is well-known (see, e.g., [43]) that σ is minimal if and only if $x_s \leq \sum_{s' \in S} P(s, a, s') \cdot x_{s'}$ holds for all $s \in S$ and $a \in Act(s)$. (There is a similar condition for *maximal* schedulers.)

The minimality criterion can be lifted to the parametric case: Suppose $R \subseteq \mathcal{P}^{\mathrm{wd}}$ is a graph-consistent region and let $f_s = sol_s^{\mathcal{M}^{\sigma}}$. Then σ is *somewhere* minimal on R if and only if there exists some $val \in R$ such that

$$f_s[val] \leq \sum_{s' \in S} P(s, a, s') \cdot f_{s'}[val] \tag{1}$$

for all $s \in S$ and $a \in Act$. (For *everywhere* minimal strategies, a universal quantification over val yields the correct criterion).

▶ **Lemma 15.** *In the fixed parameter case, checking whether a given strategy is somewhere (resp. everywhere) minimal (resp. maximal) on $\mathcal{P}^{\mathrm{wd}}$ is in P.*

Proof sketch. Condition (1) can be reformulated as the ETR formula with $|X|$ many variables

$$\Psi = \exists val \colon \Phi_R(val) \longrightarrow \Phi_{\sigma}(val) \tag{2}$$

where $\Phi_R(val)$ is a formula which is true if and only if $val \in R$ and

$$\Phi_{\sigma}(val) = \bigwedge_{s \in S} \bigwedge_{a \in Act} \left(g_s[val] \cdot \prod_{s' \neq s} h_{s'}[val] \ \leq \ \sum_{s' \in S} P(s, a, s') \cdot g_{s'}[val] \cdot \prod_{s'' \neq s'} h_{s''}[val] \right) \tag{3}$$

where $g_s/h_s = f_s$ for $g_s, h_s \in \mathbb{Q}[val]$. (W.l.o.g. it holds that $h_s[val] > 0$ for all $val \in R$.)   ◀

Proof sketch of Thm. 14. Consider $\bowtie \,=\, \geq$: Guess a somewhere minimal scheduler. Check its minimality similar to Lem. 15, but extended to simultaneously ensure that the induced pMC satisfies the threshold. The other relations in $\bowtie$ are analogous.   ◀

Sets of optimal schedulers. For the problems $\forall\forall\mathrm{Reach}_*^{\bowtie}$ and $\forall\exists\mathrm{Reach}_*^{\bowtie}$ (with fixed parameters) we already have coNP-membership (as we considered their complements before). It is tempting to assume that their NP-membership can be established analogous to above, relying on *everywhere* optimal schedulers which, according to Lem. 15, can also be verified in polynomial time. However, such schedulers do not necessarily exist. What we need instead is a *set* of somewhere optimal schedulers covering the entire parameter space – a so called *optimal-scheduler set (OSS)*.

▶ **Definition 16** (Optimal scheduler set). *A set $\Omega \subseteq \Sigma$ is called an* optimal scheduler set *(OSS) on $R \subseteq \mathcal{P}^{\mathrm{wd}}$ if*

$$\forall val \in R, \exists \sigma \in \Omega. \ \mathrm{Pr}^{\sigma}_{\mathcal{M}[val]}(\lozenge T) = \max_{\sigma' \in \Sigma} \mathrm{Pr}^{\sigma'}_{\mathcal{M}[val]}(\lozenge T),$$

i.e. Ω contains a maximal scheduler for every point in the region R. The notion can be analogously defined for minimal schedulers.

An OSS of minimal cardinality is called a *minimal optimal scheduler set* (MOSS). For many applications it is appropriate to describe a region R via a quantifier-free ETR-formula Φ_R with $|X|$ free variables such that $Sat(\Phi_R) = R$. In that case, we have the following:

▶ **Theorem 17.** *In the fixed parameter case, checking whether a given $\Omega \subseteq \Sigma$ constitutes an OSS on $R = Sat(\Phi_R)$ can be done in time polynomial in the size of $\mathcal{M}$, Ω and Φ_R.*

Proof. For every $\sigma \in \Omega$, we construct the formulas Φ_{σ} as in (3) in polynomial time. Then, we check whether the fixed-parameter ETR-formula

$$\exists val \colon \Phi_R(val) \longrightarrow \bigwedge_{\sigma \in \Omega} \neg\Phi_{\sigma}(val)$$

is unsatisfiable (also in polynomial time). If yes, return true and otherwise false.   ◀

Checking whether a set is a MOSS then additionally requires consideration of all the subsets with one element removed (again in polynomial time).

▶ **Lemma 18.** *If the size of a MOSS on $\mathcal{P}^{\mathrm{wd}}$ is polynomially bounded for fixed-parameter pMDPs, then $\exists\exists\mathrm{Reach}_*^{\bowtie}$ and $\exists\forall\mathrm{Reach}_*^{\bowtie}$ are in coNP.*

The proof considers the complement of $\exists\forall\mathrm{Reach}_*^{\bowtie}$, that is $\forall\exists\mathrm{Reach}_*^{\bowtie}$. Under the assumption in the lemma, it now suffices to guess a MOSS and verify $\exists\mathrm{Reach}_*^{\bowtie}$ on the induced pMCs in polynomial time, showing that the complement is in NP.

In the arbitrary parameter case, we obtain an exponential lower bound on the MOSS size:

▶ **Lemma 19.** *There exists a family $(\mathcal{M}_n)_{n\in\mathbb{N}}$ of simple pMDPs with $n+2$ states s.t. $|\Omega| \geq 2^n$ for any OSS Ω on $\mathcal{P}^{\mathrm{wd}}$, i.e., the size of a MOSS can grow exponentially in the pMDP's size.*

This lemma, and what follows below, consider the unbounded parameter case, i.e., from now on, parameters are part of the input.

5   The expressiveness of simple pMCs

We investigate the relation between polynomial inequalities and the $\exists\mathrm{Reach}$ problems. The first lemma in this section is a key ingredient for our complexity analysis later on.

▶ **Lemma 20** (Chonev's trick [16, Remark 7])**.** *Let $f \in \mathbb{Q}[X]$ be a polynomial, $\mu \in \mathbb{Q}$ and $0 < \lambda < 1$. There exists a simple acyclic pMC $\mathcal{M}$ with a target state T such that for all $val \colon X \to [0,1]$ and all comparison relations $\bowtie \in \{<, \leq, \geq, >, =\}$ it holds that*

$$f[val] \bowtie \mu \iff \mathrm{Pr}_{\mathcal{M}[val]}(\Diamond T) \bowtie \lambda.$$

Moreover, if d is the total degree of f, t the number of terms in f and κ a bound on the (bit-)size of the coefficients and the thresholds μ, λ, then $\mathcal{M}$ can be constructed in time $\mathcal{O}(poly(d,t,\kappa))$.

▶ **Example 21.** Consider the inequality $-2x^2y + y \geqslant 5$. We reformulate this to: $2 \cdot ((1-x)xy + (1-x)y + (1-y) - 1) + y \geqslant 5$ and then to $2 \cdot (1-x)xy + 2 \cdot (1-x)y + 2 \cdot (1-y) + y \geqslant 7$. Observe that both sides now only contain positive coefficients. Furthermore, observe that we wrote the left-hand side as sum of products over $\{x, 1-x, y, 1-y\}$. After rescaling (with $\frac{1}{8}$), we can construct the pMC $\mathcal{M}$ depicted in Fig. 2a and set $\lambda = \frac{7}{8}$.

Checking a bound on a given polynomial over X thus is equivalent to checking a bound on a reachability probability in a simple acyclic pMC over X. For the fixed parameter case, this gives rise to the following equivalence relating arbitrary pMCs to simple acyclic pMCs.

▶ **Theorem 22.** *For any non-simple pMC $\mathcal{M}$ with $\mathcal{P}_{\mathcal{M}}^{\mathrm{gp}} = (0,1)^X$ there exists a simple acyclic pMC $\mathcal{M}'$ such that*

$$\{val \in \mathcal{P}_{\mathcal{M}}^{\mathrm{gp}} \mid \mathrm{Pr}_{\mathcal{M}[val]}(\Diamond T) \bowtie \lambda\} = \{val \in \mathcal{P}_{\mathcal{M}'}^{\mathrm{gp}} \mid \mathrm{Pr}_{\mathcal{M}'[val]}(\Diamond T) \bowtie \lambda\}.$$

In the fixed parameter case, $\mathcal{M}'$ can be computed in polynomial time.

The proof is constructive: one first computes the (rational function) $sol_{\mathcal{M}}$, reformulates that as a polynomial constraint, and casts that into a simple acyclic pMC using Lemma 20.

The goal of the rest of this section is to prove a result which is, in a sense, a stronger version of Lemma 20. In particular, we want to describe polynomials by $sol^{\mathcal{M}}$ for an acyclic pMC. We call a polynomial $f \in \mathbb{Q}[X]$ *adequate* if $0 < f[val] < 1$ for all $val \colon X \to (0,1)$ and $0 \leq f[val] \leq 1$ for all $val \colon X \to [0,1]$. Note that $sol^{\mathcal{M}}$ *is an adequate polynomial if $\mathcal{M}$ is both simple and acyclic*, and there is no acyclic pMC $\mathcal{M}$ (with a single parameter) such that $sol_{\mathcal{M}}$ is *not* adequate – except where $sol_{\mathcal{M}} = 0$ or $sol_{\mathcal{M}} = 1$.
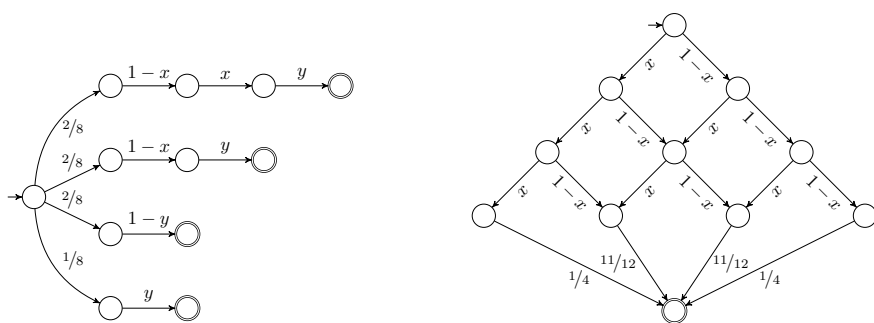
(a) The reachability probability in $\mathcal{M}[val]$ is at least $\frac{7}{8}$ iff $(-2x^2y + y)[val] \geq 5$.

(b) The reachability probability from source to target equals $f = 2x \cdot (1 - x) + \frac{1}{4}$.

■ **Figure 2** Examples for the strong connection between polynomial (inequalities) and pMCs. Transitions to the sink are not depicted for conciseness.

▶ **Theorem 23.** *Let $f \in \mathbb{Q}[x]$ be a (univariate) adequate polynomial. There exists a simple acyclic pMC $\mathcal{M}$ with a target state T such that $f = sol^{\mathcal{M}}$.*

Our construction of a pMC for some adequate polynomial is based on the following result:

▶ **Lemma 24** (Handelman's theorem [30]). *Let $\beta_1 \geq 0, \ldots, \beta_\ell \geq 0$ be linear constraints that define a compact convex polyhedron $P \subseteq \mathbb{R}^n$ with interior. If a polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ is strictly positive on P, then f may be written as*

$$f = \sum_{i=1}^{k} \lambda_i h_i \tag{4}$$

where $h_i = \beta_1^{e_{i,1}} \cdot \ldots \cdot \beta_\ell^{e_{i,\ell}}$ for some natural exponents $e_{i,j}$ and real coefficients $\lambda_i > 0$.

Form (4) is called a *Handelman representation* of f w.r.t. $\beta_1, \ldots, \beta_\ell$. The next lemma states the existence of a specific Handelman representation which we can map to a pMC.

▶ **Lemma 25.** *Let $f \in \mathbb{Q}[x]$ be an adequate polynomial. There exists an $n \geq 0$ such that*

$$f = \sum_{k=0}^{n} p_k \cdot \binom{n}{k} \cdot x^{n-k} \cdot (1 - x)^k \qquad \text{with } p_k \in [0, 1] \text{ for all } 0 \leq k \leq n \tag{5}$$

▶ **Example 26.** Consider $f = 2x \cdot (1 - x) + \frac{1}{4}$ which is strictly positive on $[0, 1]$ and already in a Handelman representation. Following the proof of Lem. 25, we find that

$$f = \frac{1}{4}\binom{0}{3}x^3 + \frac{11}{12}\binom{1}{3}x^2(1-x) + \frac{11}{12}\binom{2}{3}x(1-x)^2 + \frac{1}{4}\binom{3}{3}(1-x)^3$$

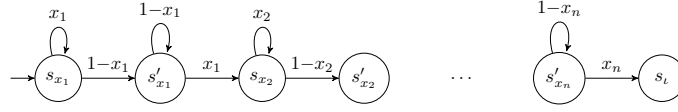The construction (described in the proof of Thm. 23) yields the pMC depicted in Fig. 2b.

6 The complexity of reachability in pMCs

We improve lower bounds for ∃Reach problems. The results depend on the comparison type:

Nonstrict inequalities. This paragraph is devoted to proving the following theorem:

▶ **Theorem 27.** *∃Reach$_*^{\leq}$, ∃Reach$_*^{\geq}$ are all ETR-complete (even for acyclic pMCs).*

Figure 3 Gadget for the proof of Lemma 32.

▶ **Definition 28.** *The decision problem* modified-closed-bounded-4-feasibility *(mb4FEAS-c) asks: Given a (non-negative) quadric polynomial f, $\exists val\colon X \to [0,1]$ s.t. $f[val] \leq 0$? The* modified-open-bounded-4-feasibility *(mb4FEAS-o) is analogously defined with val ranging over $(0,1)$.*

This problem easily reduces to its $\geq$-variant by multiplying f with -1.

▶ **Lemma 29.** *The problems mb4FEAS-c and mb4FEAS-o are ETR-hard.*

Essentially, one reduces from the existence of common roots of quadratic polynomials lying in a unit ball, which is known to be ETR-complete [47, Lemma 3.9]. The reduction to mb4FEAS follows the reduction[2] between unconstrained variants (i.e., variants in which the position of the root is not constrained) of the same decision problems [48, Lemma 3.2].

▶ Remark 30. Observe that there may be exactly one satisfying assignment to mb4FEAS-o/c, which may be irrational. In contrast, if there exists a satisfying assignment for $f > 0$, then there exist infinitely many satisfying (rational) assignments. To the best of our knowledge, the complexity of a variant of mb4FEAS-o/c with strict bounds is open. Therefore, we have no ETR-hardness for $\exists$Reach with strict bounds. In general, *conjunctions* of strict inequalities are also ETR-complete [48]. We exploit this in the proof of Thm. 37 on page 12.

Proof of Thm. 27. The reduction from mb4FEAS-c to $\exists$Reach$_{\mathrm{wd}}^{\leq}$ is a straightforward application of Lemma 20 with $\mu = 0$ and $\lambda = \frac{1}{2}$. For $\exists$Reach$_{\mathrm{gp}}^{\leq}$, we reduce from the open variant and notice that as the construction in Lemma 20 preserves all satisfying instantiations $val\colon X \to [0,1]$ it, in particular, also preserves them on the graph-preserving parameter space. For $\geq$, we apply Lemma 20 on $-f$. ◀

The tight complexity class shows that the assumption of simplicity is not a real restriction. Furthermore, a similar construction can be used for (sufficiently large[3], linear) subsets of the parameter space. In particular, methods [18, 44] targeted at a variant of $\exists$Reach considering a so-called ϵ-preserving parameter space ($[\epsilon, 1-\epsilon]^k$) target an ETR-complete problem.

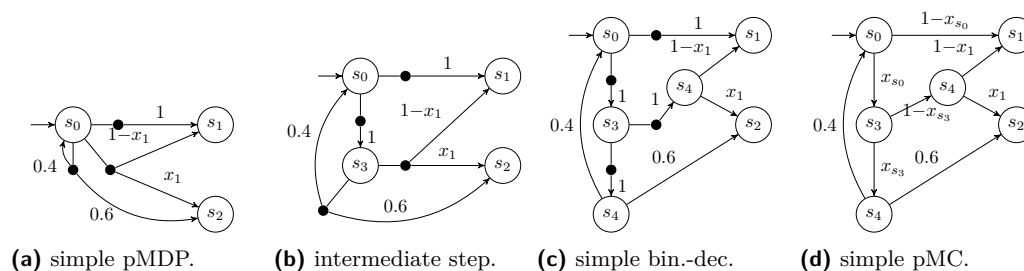**Strict inequalities.** In this paragraph, the main result is:

▶ **Theorem 31.** $\exists$Reach$_*^{\gtrless}$ *and* $\exists$Reach$_*^{\lessgtr}$ *are NP-hard.*

The gadget in Fig. 3 ensures that for any graph non-preserving instantiation, the probability to reach the target is 0, while it does not affect reachability probabilities for graph-preserving instantiations. Together with semi-continuity of the solution function, we deduce that assuming graph-preservation is equivalent to not making this assumption:

▶ **Lemma 32.** *There are polynomial-time Karp reductions among* $\exists$Reach$_{gp}^{>}$ *and* $\exists$Reach$_{\mathrm{wd}}^{>}$.

---

[2] Essentially the polynomial f in mb4FEAS is constructed by taking the sum-of-squares of the quadratic polynomials, and further operations are adequately shifting the polynomial.

[3] The bounds should be at least δ apart, where δ requires at most single-exponentially many bits.

(a) simple pMDP. **(b)** intermediate step. **(c)** simple bin.-dec. **(d)** simple pMC.

Figure 4 From simple pMDP to simple pMC.

We may thus turn our attention to well-defined parameter spaces: The decision problem

$$\exists \mathrm{Reach}_{\mathrm{wd}}^{\geq 1} \stackrel{\mathrm{def}}{\iff} \exists val \in \mathcal{P}_{\mathcal{M}}^{\mathrm{wd}}.\ \mathrm{Pr}_{\mathcal{M}[val]}(\lozenge T) \geq 1$$

is NP-complete [16, Thm. 3]. A more refined analysis of the 3SAT-reduction yields:

▶ **Theorem 33.** $\exists \mathrm{Reach}_{\mathrm{wd}}^{>}$ *and* $\exists \mathrm{Reach}_{\mathrm{wd}}^{<}$ *are NP-hard.*

Proof of Thm. 31. Lem. 32, Thm. 33, and Lem. 9 together imply Thm. 31. ◀

This concludes our complexity analysis for pMCs.

7 The complexity of reachability in pMDPs

7.1 Exists-exists reachability

By definition, every pMC is a pMDP. Conversely, from any pMDP we can construct a pMC such that their $\exists\exists\mathrm{Reach}_{\mathrm{wd}}^{\bowtie}$ problems coincide. A similar construction relates pMCs to the existence of optimal randomised memoryless strategies in partially observable MDPs [37].

▶ **Lemma 34.** *There are polynomial-time Karp reductions among* $\exists \mathrm{Reach}_{\mathrm{wd}}^{\bowtie}$ *and* $\exists\exists\mathrm{Reach}_{\mathrm{wd}}^{\bowtie}$.

We outline the steps in Fig. 4 and in the description below.

Binary-decision pMDPs. The first step of the translation consists in restricting the non-determinism resolved by a scheduler to (at most) two options from every state. A binary-decision pMDP is a pMDP such that $|Act(s)| \leq 2$ for all states $s \in S$ and if $|Act(s)| = 2$ then $\forall a \in Act(s), \forall s' \in S,\ P(s, a, s') \in \{0, 1\}$. Any pMDP can be transformed (in polynomial time) into a binary-decision pMDP by introducing auxiliary states and simulating k-ary non-deterministic choice using a binary-tree-like scheme in which all non-Dirac transitions are pushed to the leaves (see, e.g., [37, 44, 49]). Such a construction preserves simplicity.

From non-determinism to parameters. For a given binary-decision pMDP $\mathcal{M}$, we may replace all non-determinism by parameters, inspired by [28, 37]. We introduce fresh variables $X_S = \{x_s \mid s \in S\}$. In $\mathcal{M}$, for any state s with $Act(s) = \{a, a'\}$ we replace
- the unique transition $P(s, a, s') = 1$ by $P(s, a, s') = x_s$
- the unique transition $P(s, a', s') = 1$ by $P(s, a', s') = 1 - x_s$.

The outcome is a simple pMC $\mathcal{M}'$. To translate instantiations into schedulers, and vice versa, it is helpful to consider randomised schedulers. Observe that, by Rem. 5, instantiations which translate into such schedulers are always dominated by deterministic ones.

Using the previously described construction, we obtain the following.
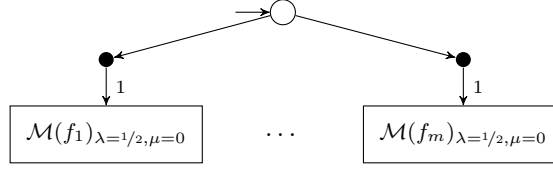
■ **Figure 5** Construction for the proof of Thm. 37.

▶ **Lemma 35.** *For all simple pMDPs $\mathcal{M}$ one can construct in polynomial time a (linearly larger) simple pMC $\mathcal{M}'$ s.t.*

$$\left(\exists val \in \mathcal{P}_{\mathcal{M}}^{\mathrm{wd}}, \exists \sigma \in \Sigma. \ \mathrm{Pr}_{\mathcal{M}[val]}^{\sigma}(\Diamond T) \bowtie \frac{1}{2} \right) \iff \left(\exists val \in \mathcal{P}_{\mathcal{M}'}^{\mathrm{wd}}. \ \mathrm{Pr}_{\mathcal{M}'[val]}(\Diamond T) \bowtie \frac{1}{2} \right).$$

▶ **Corollary 36.** *There are polynomial-time Karp reductions among $\exists\exists\mathrm{Reach}_{\mathrm{gp}}^{>}$ and $\exists\mathrm{Reach}_{\mathrm{wd}}^{>}$.*

Proof. Minor adaptions in the proofs of Lemma 34 and Lemma 32. ◀

7.2 Exists-forall reachability

Contrary to pMCs, we obtain ETR-completeness in pMDPs for any comparison relation:

▶ **Theorem 37.** *$\exists\forall\mathrm{Reach}_{*}^{\bowtie}$ are all ETR-complete (even for acyclic pMDPs with a single non-deterministic state).*

For the strict relations, we use a different problem to reduce from.

▶ **Definition 38.** *The decision problem* bounded-conjunction-of-inequalities *(bcon4INEQ-c) asks: Given a family of quadric polynomials $f_1, \ldots, f_m$, $\exists val : X \to [0,1]$ s.t. $\bigwedge_{i=1}^{m} f_i[val] < 0$? The open variant (bcon4INEQ-o) can be defined analogously.*

By a reduction from mb4FEAS (adapted from [48, Thm 4.1]):

▶ **Lemma 39.** *The bcon4INEQ-o/c problems are ETR-hard.*

Proof sketch of Thm. 37. ETR-hardness for non-strict inequalities follows from Thm. 27. For strict inequalities, we reduce from bcon4INEQ-o/c: Generalise the construction from Thm. 27: Build a pMDP as in Fig. 5 with pMCs $\mathcal{M}(f_i)_{\lambda=1/2, \mu=0}$ created by Lemma 20. ◀

Relation to stochastic games

We argue that pMDPs are – in a sense – a generalisation of *Concurrent Stochastic Reachability Games* (CSRG), a model which has been extensively studied [11, 12, 20, 31, 52].

Playing a stochastic game. A CSRG is a two-player game $\mathcal{G}$ played on a finite set S of states. The objective of player I is to reach a target states $T \subseteq S$ while player II has to avoid ever reaching a state in T. A play of $\mathcal{G}$ begins in an initial state s_{ι} and proceeds as follows: In state s, both players I and II *concurrently* select an action $a \in A_s$ (resp. $b \in B_s$), the finite set of actions available to player I (resp. II) in state s. The game then picks a successor state s' according to a fixed probability distribution $P(\cdot|s, a, b)$ over S, and the play continues in s'. The transition from s to s' is called a *round* of $\mathcal{G}$. Player I wins $\mathcal{G}$ once a state in T is reached. Otherwise, if a target is never reached, then II wins.

A *strategy* σ of a player is, in essence, a scheduler. However, strategies in a CSRG map state-action sequences $s_0(a_0, b_0) \ldots s_{k-1}(a_{k-1}, b_{k-1}) s_k$ to a probability distribution over the actions A_{s_k} (resp. B_{s_k}) available in the current state s_k. We call σ a *stationary* strategy if it does not depend on the history but only on the current state, i.e. it is a randomised memoryless scheduler. Let Σ^i denote the set of stationary strategies for player $i \in \{\mathrm{I}, \mathrm{II}\}$.

Instantiations and MDPs. The *instantiation* $\mathcal{G}^\sigma$ of $\mathcal{G}$ with a stationary strategy for player I is the structure obtained by forcing player I to follow σ. Notice that $\mathcal{G}^\sigma$ is a finite MDP $\mathcal{M}$. Its transition probability function $P_\mathcal{M}$ is obtained by letting

$$P_\mathcal{M}(s, b, s') = \sum_{a \in A_s} \sigma(a|s) P(s'|s, a, b) \tag{6}$$

for all $s, s' \in S$ and actions $b \in B_s$ of player II. (Instantiations are defined completely symmetrically for strategies of player II.) Conversely, every MDP may be viewed as a CSRG where $|A_s| = 1$ (or $|B_s| = 1$) for all $s \in S$, i.e. one of the players does never have any choice.

Value of a CSRG. Let $\mathrm{Pr}_\mathcal{G}^{\sigma, \tau}(\Diamond T)$ be the probability that T is reached if player I plays according to σ and player II according to τ. The *value* of $\mathcal{G}$ is defined as $V(\mathcal{G}) := \sup_\sigma \inf_\tau \mathrm{Pr}_\mathcal{G}^{\sigma, \tau}(\Diamond T)$ where the sup and inf range over all strategies of both players respectively. Intuitively, it is the maximal winning probability of player I that can be guaranteed against all strategies of player II. The existence of stationary optimal strategies for player II [32, 41] allows us to encode a CSRG in a pMDP by replacing the universal player II with parameters:

▶ **Theorem 40.** *For any given CSRG $\mathcal{G}$, there exists a simple pMDP $\mathcal{M}$ such that*

$$V(\mathcal{G}) = \min_{\tau \in \Sigma^{\mathrm{II}}} \max_{\sigma \in \Sigma^{\mathrm{I}}} \mathrm{Pr}_\mathcal{G}^{\sigma, \tau}(\Diamond T) = \min_{val \in \mathcal{P}^{\mathrm{wd}}} \max_{\sigma \in \Sigma} \mathrm{Pr}_{\mathcal{M}[val]}^\sigma(\Diamond T)$$

and $\mathcal{M}$ can be computed in polynomial time (in the size of $\mathcal{G}$).

As a direct consequence, we obtain CSRG-hardness.

▶ **Corollary 41.** *Determining whether $V(\mathcal{G}) \trianglelefteq \lambda$, for $\lambda \in \mathbb{Q}$, reduces to $\exists \forall \mathrm{Reach}_{\mathrm{wd}}^{\trianglelefteq}$.*

It follows from [11, Thms. 6 and 12] that optimal rational instantiations may be complex.

▶ **Theorem 42.** *There are pMDPs for which rational optimal and ε-optimal parameter values minimising the value $\max_{\sigma \in \Sigma} \mathrm{Pr}_{\mathcal{M}[val]}^\sigma(\Diamond T)$ require exponentially-many bits to be written as a binary-encoded integer-pair.*

8 Robust reachability

In this section, we briefly consider pMDPs in which we focus on obtaining (robust) schedulers rather than (robust) parameter values: We swap the quantification order from the $\mathcal{Q}_1 \mathcal{Q}_2 \mathrm{Reach}$ problem. Intuitively, we ask whether some scheduler gives guarantees on the maximal or minimal probability of all instantiations of the pMDP eventually reaching T. Formally, for each $\mathcal{Q}_1, \mathcal{Q}_2 \in \{\exists, \forall\}$ and $\bowtie \in \{\leq, <, >, \geq\}$, let

$$\mathcal{Q}_1 \mathcal{Q}_2 \mathrm{RobReach}_{\mathrm{wd}}^\bowtie \stackrel{\mathrm{def}}{\Longleftrightarrow} \mathcal{Q}_1 \sigma \in \Sigma, \ \mathcal{Q}_2 val \in \mathcal{P}^{\mathrm{wd}}. \ \mathrm{Pr}_{\mathcal{M}[val]}^\sigma(\Diamond T) \bowtie \frac{1}{2}.$$

We adopt the same conventions as for the $\mathcal{Q}_1 \mathcal{Q}_2 \mathrm{Reach}$ problem when considering graph-preserving instantiations. Variants which use the same quantifier twice, or consider pMCs yield the same results as for $\mathrm{Reach}^\bowtie$, and are therefore omitted.

Robust strategies have been widely studied in the field of operations research (see, e.g., [40, 57]) and are the main focus of *reinforcement learning* [55]. It is known that the robust-reachability problem as defined above is not the most general question one can ask. Indeed, we restrict our attention to *memoryless schedulers* while, in general, optimal robust schedulers require memory and randomisation [1].

Our interest in the robust-reachability problem is twofold. First, it naturally corresponds to the quantifier-swapped version of the reachability problem. Second, memoryless schedulers are desirable in practice for their comprehensibility and ease of implementation.

▶ **Theorem 43.** *In the fixed parameter case,* $\exists\forall \text{RobReach}_*^{</>}$ *are NP-complete. NP-hardness holds even for acyclic pMDPs with a single parameter.*

Proof sketch. Membership in NP is analogous to Lem. 13. NP-hardness is based on a reduction from 3-SAT, with a construction similar to Fig. 5. ◀

▶ **Proposition 44.** *The decision problems* $\exists\forall \text{RobReach}_*^{\bowtie}$ *are NP-hard and coNP-hard, and in PSPACE. For non-strict inequalities, the problems are coETR-hard.*

Proof. NP-hardness follows from Thm. 43, coNP/coETR-hardness follows from Thm. 31, Thm. 33, and Thm. 27, respectively. Iterating over all (finitely many) schedulers, check each scheduler in ETR or in coETR (and thus in PSPACE). ◀

Consequently, it is unlikely that either of the problems are in ETR or coETR, as then ETR and coETR would coincide (which is not impossible, but unlikely [48]).

9 Conclusions

We have studied the complexity of various reachability problems for *simple* pMCs and pMDPs. All the problems we have considered are easily seen to be solvable in PSPACE via reductions to the existential theory of the reals. We have complemented this observation with lower bounds, i.e. ETR hardness for several versions of the problem both for (tree-like) pMCs and pMDPs. These lower bounds naturally extend to general pMCs and pMDPs, and to expected reward measures.

We have given an NP decision procedure for pMDPs with a fixed number of parameters. The exact complexity of pMDP reachability problems with this restriction remains open, and our upper bounds do not straightforwardly generalise beyond simple pMDPs (see Rem. 3).

Finally, we have established a tight connection between polynomials and pMCs (even beyond [16]). However, our results do not allow us to conclude whether there always are "small" pMCs for every polynomial. Such a result would provide more evidence of ETR being the right framework to solve problems for our parametric models.

─── **References** ───

1 Sebastian Arming, Ezio Bartocci, and Ana Sokolova. SEA-PARAM: exploring schedulers in parametric MDPs. In *QAPL@ETAPS*, volume 250 of *EPTCS*, pages 25–38, 2017.

2 Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT Press, 2008.

3 Ezio Bartocci, Radu Grosu, Panagiotis Katsaros, C. R. Ramakrishnan, and Scott A. Smolka. Model Repair for Probabilistic Systems. In *TACAS*, volume 6605 of *LNCS*, pages 326–340. Springer, 2011.

4 Saugata Basu, Richard Pollack, and Marie-Françoise Roy. Existential theory of the reals. *Algorithms in Real Algebraic Geometry*, pages 505–532, 2006.

5   Daniel S. Bernstein, Robert Givan, Neil Immerman, and Shlomo Zilberstein. The Complexity of Decentralized Control of Markov Decision Processes. *Math. Oper. Res.*, 27(4):819–840, 2002.

6   Luca Bortolussi and Simone Silvetti. Bayesian Statistical Parameter Synthesis for Linear Temporal Properties of Stochastic Models. In *TACAS (2)*, volume 10806 of *LNCS*, pages 396–413. Springer, 2018.

7   John F. Canny. Some Algebraic and Geometric Computations in PSPACE. In *STOC*, pages 460–467. ACM, 1988.

8   Milan Ceska, Frits Dannenberg, Nicola Paoletti, Marta Kwiatkowska, and Lubos Brim. Precise parameter synthesis for stochastic biochemical systems. *Acta Inf.*, 54(6):589–623, 2017.

9   Krishnendu Chatterjee. Robustness of Structurally Equivalent Concurrent Parity Games. In *FOSSACS*, volume 7213 of *LNCS*, pages 270–285. Springer, 2012.

10   Krishnendu Chatterjee, Martin Chmelik, and Jessica Davies. A Symbolic SAT-Based Algorithm for Almost-Sure Reachability with Small Strategies in POMDPs. In *AAAI*, pages 3225–3232. AAAI Press, 2016.

11   Krishnendu Chatterjee, Kristoffer Arnsfelt Hansen, and Rasmus Ibsen-Jensen. Strategy Complexity of Concurrent Safety Games. In *MFCS*, volume 83 of *LIPIcs*, pages 55:1–55:13. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.

12   Krishnendu Chatterjee and Thomas A. Henzinger. A survey of stochastic omega-regular games. *J. Comput. Syst. Sci.*, 78(2):394–413, 2012.

13   Taolue Chen, Yuan Feng, David S. Rosenblum, and Guoxin Su. Perturbation Analysis in Verification of Discrete-Time Markov Chains. In *CONCUR*, volume 8704 of *LNCS*, pages 218–233. Springer, 2014.

14   Taolue Chen, Ernst Moritz Hahn, Tingting Han, Marta Z. Kwiatkowska, Hongyang Qu, and Lijun Zhang. Model Repair for Markov Decision Processes. In *TASE*, pages 85–92. IEEE Computer Society, 2013.

15   Taolue Chen, Tingting Han, and Marta Z. Kwiatkowska. On the complexity of model checking interval-valued discrete time Markov chains. *Inf. Process. Lett.*, 113(7):210–216, 2013.

16   Ventsislav Chonev. Reachability in Augmented Interval Markov Chains. *CoRR*, abs/1701.02996, 2017. `arXiv:1701.02996`.

17   Anne Condon. *Computational models of games.* ACM distinguished dissertations. MIT Press, 1989.

18   Murat Cubuktepe, Nils Jansen, Sebastian Junges, Joost-Pieter Katoen, and Ufuk Topcu. Synthesis in pMDPs: A Tale of 1001 Parameters. In *ATVA*, volume 11138 of *LNCS*, pages 160–176. Springer, 2018.

19   Conrado Daws. Symbolic and Parametric Model Checking of Discrete-Time Markov Chains. In *ICTAC*, volume 3407 of *LNCS*, pages 280–294. Springer, 2004.

20   Luca de Alfaro, Thomas A. Henzinger, and Orna Kupferman. Concurrent reachability games. *Theor. Comput. Sci.*, 386(3):188–217, 2007.

21   Christian Dehnert, Sebastian Junges, Nils Jansen, Florian Corzilius, Matthias Volk, Harold Bruintjes, Joost-Pieter Katoen, and Erika Ábrahám. PROPhESY: A PRObabilistic ParamEter SYnthesis Tool. In *CAV (1)*, volume 9206 of *LNCS*, pages 214–231. Springer, 2015.

22   Christian Dehnert, Sebastian Junges, Joost-Pieter Katoen, and Matthias Volk. A Storm is Coming: A Modern Probabilistic Model Checker. In *CAV (2)*, volume 10427 of *LNCS*, pages 592–600. Springer, 2017.

23   Karina Valdivia Delgado, Scott Sanner, and Leliane Nunes de Barros. Efficient solutions to factored MDPs with imprecise transition probabilities. *Artif. Intell.*, 175(9-10):1498–1527, 2011.

24   Antonio Filieri, Giordano Tamburrelli, and Carlo Ghezzi. Supporting Self-Adaptation via Quantitative Verification and Sensitivity Analysis at Run Time. *IEEE Trans. Software Eng.*, 42(1):75–99, 2016.

25   Paul Gainer, Ernst Moritz Hahn, and Sven Schewe. Accelerated Model Checking of Parametric Markov Chains. In *ATVA*, volume 11138 of *LNCS*, pages 300–316. Springer, 2018.

26   Sergio Giro, Pedro R. D'Argenio, and Luis María Ferrer Fioriti. Distributed probabilistic input/output automata: Expressiveness, (un)decidability and algorithms. *Theor. Comput. Sci.*, 538:84–102, 2014.

27   Robert Givan, Sonia Leach, and Thomas Dean. Bounded-parameter Markov decision processes. *Artif. Intell.*, 122(1-2):71–109, 2000.

28   Ernst Moritz Hahn, Tingting Han, and Lijun Zhang. Synthesis for PCTL in Parametric Markov Decision Processes. In *NASA Formal Methods*, volume 6617 of *LNCS*, pages 146–161. Springer, 2011.

29   Ernst Moritz Hahn, Holger Hermanns, and Lijun Zhang. Probabilistic reachability for parametric Markov models. *STTT*, 13(1):3–19, 2010.

30   David Handelman. Representing polynomials by positive linear functions on compact convex polyhedra. *Pacific Journal of Mathematics*, 132(1):35–62, 1988.

31   Kristoffer Arnsfelt Hansen, Michal Koucký, and Peter Bro Miltersen. Winning Concurrent Reachability Games Requires Doubly-Exponential Patience. In *LICS*, pages 332–341. IEEE Computer Society, 2009.

32   CJ Himmelberg, Thiruvenkatachari Parthasarathy, TES Raghavan, and FS Van Vleck. Existence of p-equilibrium and optimal stationary strategies in stochastic games. *Proceedings of the American Mathematical Society*, 60(1):245–251, 1976.

33   Lisa Hutschenreiter, Christel Baier, and Joachim Klein. Parametric Markov Chains: PCTL Complexity and Fraction-free Gaussian Elimination. In *GandALF*, volume 256 of *EPTCS*, pages 16–30, 2017.

34   Nils Jansen, Florian Corzilius, Matthias Volk, Ralf Wimmer, Erika Ábrahám, Joost-Pieter Katoen, and Bernd Becker. Accelerating Parametric Probabilistic Verification. In *QEST*, volume 8657 of *LNCS*, pages 404–420. Springer, 2014.

35   Bengt Jonsson and Kim Guldstrand Larsen. Specification and Refinement of Probabilistic Processes. In *LICS*, pages 266–277. IEEE Computer Society, 1991.

36   Sebastian Junges, Erika Abraham, Christian Hensel, Nils Jansen, Joost-Pieter Katoen, Tim Quatmann, and Matthias Volk. Parameter Synthesis for Markov Models. *CoRR*, abs/1903.07993, 2019. `arXiv:1903.07993`.

37   Sebastian Junges, Nils Jansen, Ralf Wimmer, Tim Quatmann, Leonore Winterer, Joost-Pieter Katoen, and Bernd Becker. Finite-State Controllers of POMDPs using Parameter Synthesis. In *UAI*, pages 519–529. AUAI Press, 2018.

38   Marta Z. Kwiatkowska, Gethin Norman, and David Parker. PRISM 4.0: Verification of probabilistic real-time systems. In *CAV*, volume 6806 of *LNCS*, pages 585–591. Springer, 2011.

39   Ruggero Lanotte, Andrea Maggiolo-Schettini, and Angelo Troina. Parametric probabilistic transition systems for system design and analysis. *Formal Asp. Comput.*, 19(1):93–109, 2007.

40   Arnab Nilim and Laurent El Ghaoui. Robust Control of Markov Decision Processes with Uncertain Transition Matrices. *Operations Research*, 53(5):780–798, 2005.

41   Thiruvenkatachari Parthasarathy. Discounted and positive stochastic games. *Bulletin of the American Mathematical Society*, 77(1):134–136, 1971.

42   Alberto Puggelli, Wenchao Li, Alberto L. Sangiovanni-Vincentelli, and Sanjit A. Seshia. Polynomial-Time Verification of PCTL Properties of MDPs with Convex Uncertainties. In *CAV*, volume 8044 of *LNCS*, pages 527–542. Springer, 2013.

43   Martin L. Puterman. *Markov Decision Processes*. Wiley, 1995.

44   Tim Quatmann, Christian Dehnert, Nils Jansen, Sebastian Junges, and Joost-Pieter Katoen. Parameter Synthesis for Markov Models: Faster Than Ever. In *ATVA*, volume 9938 of *LNCS*, pages 50–67, 2016.

45   James Renegar. On the Computational Complexity and Geometry of the First-Order Theory of the Reals, Part I: Introduction. Preliminaries. The Geometry of Semi-Algebraic Sets. The Decision Problem for the Existential Theory of the Reals. *J. Symb. Comput.*, 13(3):255–300, 1992.

46   Stuart J. Russell and Peter Norvig. *Artificial Intelligence - A Modern Approach.* Pearson Education, 2010.

47   Marcus Schaefer. *Realizability of Graphs and Linkages*, pages 461–482. Springer New York, 2013.

48   Marcus Schaefer and Daniel Stefankovic. Fixed Points, Nash Equilibria, and the Existential Theory of the Reals. *Theory Comput. Syst.*, 60(2):172–193, 2017.

49   Roberto Segala and Andrea Turrini. Comparative Analysis of Bisimulation Relations on Alternating and Non-Alternating Probabilistic Models. In *QEST*, pages 44–53. IEEE Computer Society, 2005.

50   Koushik Sen, Mahesh Viswanathan, and Gul Agha. Model-Checking Markov Chains in the Presence of Uncertainties. In *TACAS*, volume 3920 of *LNCS*, pages 394–410. Springer, 2006.

51   Sven Seuken and Shlomo Zilberstein. Formal models and algorithms for decentralized decision making under uncertainty. *Autonomous Agents and Multi-Agent Systems*, 17(2):190–250, 2008.

52   Lloyd S. Shapley. Stochastic Games. *PNAS*, 39(10):1095–1100, 1953.

53   Eilon Solan. Continuity of the value of competitive Markov decision processes. *Journal of Theoretical Probability*, 16(4):831–845, 2003.

54   Jeremy Sproston. Qualitative Reachability for Open Interval Markov Chains. In *RP*, volume 11123 of *LNCS*, pages 146–160. Springer, 2018.

55   Richard S. Sutton and Andrew G. Barto. *Reinforcement learning - an introduction.* Adaptive computation and machine learning. MIT Press, 1998.

56   Nikos Vlassis, Michael L. Littman, and David Barber. On the Computational Complexity of Stochastic Controller Optimization in POMDPs. *TOCT*, 4(4):12:1–12:8, 2012.

57   Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem. Robust Markov Decision Processes. *Math. Oper. Res.*, 38(1):153–183, 2013.

58   Di Wu and Xenofon D. Koutsoukos. Reachability analysis of uncertain systems using bounded-parameter Markov decision processes. *Artif. Intell.*, 172(8-9):945–954, 2008.