Markov Decision Processes with Sure Parity and Multiple Reachability Objectives

Raphaël Berthon, Joost-Pieter Katoen and Tobias Winkler

September 27, 2024

Markov Decision Processes with Sure Parity and Multiple Reachability Objectives

Raphaël Berthon✉, Joost-Pieter Katoen, and Tobias Winkler *

RWTH Aachen University, 52062 Aachen, Germany
berthon@cs.rwth-aachen.de
katoen@cs.rwth-aachen.de
tobias.winkler@cs.rwth-aachen.de

Abstract. This paper considers the problem of finding strategies that satisfy a mixture of sure and threshold objectives in Markov decision processes. We focus on a single ω-regular objective expressed as parity that must be surely met while satisfying n reachability objectives towards sink states with some probability thresholds too. We consider three variants of the problem: (a) strict and (b) non-strict thresholds on all reachability objectives, and (c) maximizing the thresholds with respect to a lexicographic order. We show that (a) and (c) can be reduced to solving parity games, and (b) can be solved in EXPTIME. Strategy complexities as well as algorithms are provided for all cases.

Keywords: MDPs · Parity · Reachability · Multi-objective

1 Introduction

Markov decision processes (MDPs) [6,35] are prominent models for strategic planning and decision making in face of stochastic uncertainty. An important, yet intricate, problem is to determine if and how a combination of *multiple* properties, or *objectives*, is realizable in a given MDP. As objectives may be *conflicting*, it does not suffice to analyze each of them independently [20,37,4]. Instead, *trade-offs* between the objectives have to be taken into account. In this paper, we combine objectives of different nature: *Sure objectives* must be fulfilled on *all* possible executions of the MDP, even on those with probability 0. Thus, sure objectives do not depend on the exact transition probabilities; in fact, they can be analyzed by replacing the probabilities with an adversary. *Threshold objectives*, on the other hand, have to be satisfied with some probability of at least (or greater than) a given constant. Various combinations of sure and threshold objectives have been investigated in prior work [23,1,13,8,22,7]. Here, we focus on MDPs with a single sure ω-regular objective expressed as a *parity* condition together with n *reachability* threshold objectives towards sink states.

* Berthon and Katoen are supported by the DFG Grant Nr. 520530521 "POMPOM", Katoen and Winkler are supported by the EU's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement Nr. 101008233 "MISSION", Winkler is additionally supported by the DFG RTG 2236 "UnRAVeL".

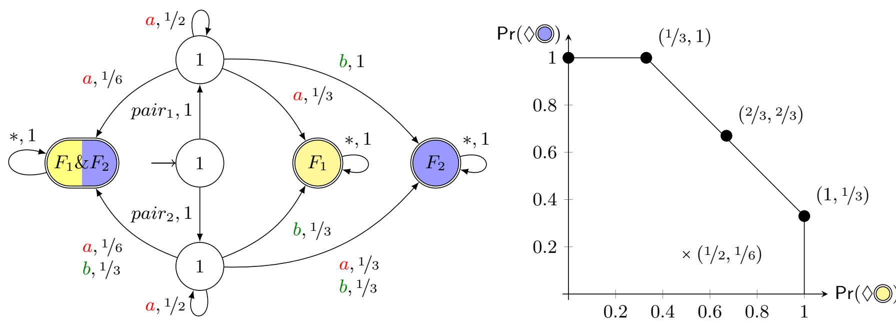

Fig. 1. An MDP with $Act = \{a, b, pair_1, pair_2, *\}$, target sets F_1, F_2, and parity condition ρ assigning 0 to the sinks and 1 to the non-sinks. *Right:* The Pareto frontier.

Running example In a game show a contestant plays a gamble to win either a bike, a surfboard, or both prizes. The gamble is as follows: The contestant must choose one out of two pairs of 6-sided dice. Each pair consists of a green and a red die. The four dice are all different; each of their faces shows either the bike, the board, both bike and board, or the symbol $\circlearrowright$ ("repeat"):

$pair_1$	*red*: $3 \times \circlearrowright, 1 \times$ both, $2 \times$ bike	*green*: $6 \times$ board
$pair_2$	*red*: $3 \times \circlearrowright, 1 \times$ both, $2 \times$ board	*green*: $2 \times$ both, $2 \times$ bike, $2 \times$ board

After committing to a pair of dice, the contestant rolls one die from their pair. The green die immediately ends the gamble with the resulting prize(s). The red die either ends the gamble or, in case of $\circlearrowright$, allows the contestant to roll again (the same die or the other one). However, since the show is broadcast on live TV, there is an additional rule: The gamble may not be prolonged indefinitely, i.e., the contestant may try the red die at most an arbitrary, but *a priori* fixed number of times. Clearly, optimal strategies depend on how much the contestant prefers one prize over the other. The MDP in Figure 1 models this gamble. Prizes are encoded as reachability of sinks ($F_1 \mathrel{\widehat{=}}$ bike, $F_2 \mathrel{\widehat{=}}$ board); the additional constraint is a sure parity condition.

This paper We study the following three problems: Given a finite MDP, decide if it is possible to satisfy a sure parity objective and, at the same time, n sink reachability objectives with (a) all *strict*, or (b) all *non-strict* probability thresholds. In addition, we consider the problem (c) of checking existence of a *lexicographically* maximal achievable threshold vector w.r.t. a given linear order on the reachability objectives. In all cases, we are also interested in computing witnessing strategies if they exist. These problems are challenging, both computationally and conceptually. Already for two reachability objectives (without any sure objective) the set of achievable thresholds —the *Pareto frontier*— is a convex polytope with superpolynomially many vertices in general [25]. The problems we study are more general than this and add further subtleties: While

Table 1. Some existing and our new (in bold) results on multi-objective MDPs.

Sure objective	*Probabilistic objective(s)*	*Complexity*	*Memory*	*Reference*
1 sure parity	-	$\mathsf{NP} \cap \mathsf{coNP}$	finite	[24]
1 sure parity	1 threshold parity	$\mathsf{NP} \cap \mathsf{coNP}$	infinite	[8]
1 sure parity	n almost-sure parity	$\mathsf{NP} \cap \mathsf{coNP}$	infinite	[7]
n sure parity	n almost-sure parity	$\mathsf{P}^{\mathsf{NP}} (= \Delta_2^{\mathsf{P}})$	infinite	[7]
1 sure parity	1 threshold reach	$\mathsf{NP} \cap \mathsf{coNP}$	finite	[8]
-	n threshold sink reach	PTIME	finite	[25]
1 sure parity	n **strict threshold sink reach** (a)	$\mathsf{NP} \cap \mathsf{coNP}$	**finite**	**[Thm. 1]**
1 sure parity	n **threshold sink reach** (c)	EXPTIME	**finite**	**[Thm. 3]**
-	lexicographic Streett	PTIME	finite	[17]
1 sure parity	**lexicographic sink reach** (b)	$\mathsf{NP} \cap \mathsf{coNP}$	**finite**	**[Thm. 2]**

it is easy to show that the thresholds in the *interior* of the Pareto frontier are satisfiable with *any* sure parity objective (Section 3), identifying exactly which points on the *boundary* of the frontier are achievable is quite involved (Section 5).

Contributions Our three main results are summarized in bold in Table 1. (a) Checking if a sure parity objective and n *strict* sink reachability thresholds are achievable simultaneously is in $\mathsf{NP} \cap \mathsf{coNP}$ (Section 3). This is done via a reduction to parity games, admitting a quasi-polynomial algorithm [14].

(c) We propose an algorithm that finds a strategy ensuring a sure parity objective while also maximizing the probability of reaching n sinks w.r.t. a *lexicographic* order (Section 4). It relies on a concept we call *projection*, a notion also used in prior work [25,4,17]. Our algorithm solves polynomially many parity games in sequence, hence the problem is (again) in $\mathsf{NP} \cap \mathsf{coNP}$.

(b) We present an algorithm for finding a strategy satisfying a sure parity objective and n sink reachability objectives with *non-strict* thresholds (Section 5). Our algorithm alternates between computing Pareto frontiers, making projections, and pruning states not satisfying the sure parity objective; to our knowledge, this idea is new. Its time complexity is exponential in the size of the MDP, as it relies on computing exact Pareto frontiers.

We also treat strategy complexity for each case. Our results are a further step towards a solution for general combinations of sure and probabilistic objectives. An extended version of the results with detailed proofs can be found in [9].

Related work Previous research [7] on mixtures of sure and probabilistic objectives focused on *qualitative* thresholds, i.e., >0 and $=1$. Here, we also allow *quantitative* thresholds strictly between 0 and 1. We rely on some results of [8] that studied combining one sure parity and *one threshold parity* objective. This problem was shown to be in $\mathsf{NP} \cap \mathsf{coNP}$, via a reduction to parity games with weights that can be solved in quasi-polynomial time [38]. The main difference to [8] is that we consider *multiple reachability* threshold objectives. The setting of one sure parity and one almost-sure parity has been further studied in [22] where it was shown that the restriction to finite memory strategies is still in $\mathsf{NP} \cap \mathsf{coNP}$ for MDPs, and coNP-complete for stochastic games.

The seminal paper [25] shows that computing the Pareto frontier for a mixture of either reachability or ω-regular objectives can be reduced to solving linear programs. An efficient technique, value iteration, is exploited by tools such as PRISM [33], Storm [30], and MultiGain [12]. The work [19] considers MDPs with two different kinds of stochastic mean-payoff objectives, and supports computing the Pareto frontier. *Percentile queries*, multiple threshold constraints that must each be satisfied with some probability, were studied in [36]. In [10,29], multiple reachability conditions associated to the expected or accumulated cost to reach a target are considered.

Lexicographic optimization is a widely employed principle in multi-objective decision making [34,40]. The idea is that a strategy should prioritize a primary objective while still doing best possible for a secondary objective, etc. The work of [17] imposes a lexicographic order on multiple, possibly conflicting, reachability, safety and ω-regular objectives. Reinforcement learning with lexicographic ω-regular conditions is studied in [28]. To the best of our knowledge, lexicographic optimization in MDPs together with a sure condition has not been studied yet.

Other approaches have been considered. Combinations of parity and mean-payoff [1], and parity and weighted games [38] have been studied in prior work. An alternative way to combining objectives is *strategy logic* (SL) [16], an extension of CTL that can express formulas involving the change of strategies. A probabilistic SL has been defined in [2].

2 Preliminaries

We write $\mathbb{N} = \{1, 2, 3, \ldots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For $n \in \mathbb{N}$, we let $[n] = \{1, \ldots, n\}$, and $[n]_0 = [n] \cup \{0\}$. Vectors $\mathbf{v} \in \mathbb{R}^n$ are written in bold. For $\mathbf{v} \in [0, 1]^n$ and $i \in [n]$, we denote by v_i the i-th component of $\mathbf{v}$. Given $\mathbf{v}, \mathbf{u} \in \mathbb{R}^n$, their *dot product* is defined as $\mathbf{v} \cdot \mathbf{u} = \sum_{i \in [n]} v_i \cdot u_i$. The symbol $\mathbf{e}_i \in \{0, 1\}^n$ is the unit vector where the i-th component is 1 and all others are 0. The componentwise order on $\mathbb{R}^n$ is denoted with $\leq$. Given a finite set A, a *(probability) distribution* on A is a function $f : A \to [0, 1]$ such that $\sum_{a \in A} f(a) = 1$. $\mathcal{D}(A)$ denotes the set of all distributions on A. We define the *support* $\mathsf{supp}(f) = \{a \in A \mid f(a) > 0\}$.

2.1 MDPs, Strategies, and Objectives

A *Markov decision process* (MDP) is a tuple $\mathit{\Gamma} = (S, \mathit{Act}, \mathbf{P})$ where $S \neq \emptyset$ is a countable set of *states*, $\mathit{Act} \neq \emptyset$ is a finite set of *actions*, and $\mathbf{P} : S \times \mathit{Act} \times S \to [0, 1]$ is a *transition probability function* satisfying $\sum_{s' \in S} \mathbf{P}(s, a, s') \in \{0, 1\}$ for all $s \in S, a \in \mathit{Act}$. If the sum is 1 for a state-action pair s, a, then a is *enabled* at s. We write $\mathit{Act}(s)$ for the set of all actions enabled at s, and require that $\mathit{Act}(s) \neq \emptyset$. An MDP $\mathit{\Gamma}$ is called *finite* if S is finite. A state $s \in S$ is called a *sink* if for all $a \in \mathit{Act}(s)$ we have $\mathbf{P}(s, a, s) = 1$. For technical convenience we assume

$|Act(s)| = 1$ for all sinks $s \in S$. Note that we may consider the same MDP with different initial states since the latter is not fixed in our definition. See Figure 1 for a finite example MDP.

A (discrete-time) *Markov chain* (MC) is an MDP with $|Act| = 1$. We omit Act from the definition of MC and just write $\mathcal{M} = (S, \mathbf{P})$. We also identify $\mathbf{P}$ with a function of type $S \times S \to [0, 1]$.

A *strategy* for an MDP Γ is a state machine $\sigma = (Q, q_\iota, \delta, o)$ where Q is a countable set of memory modes, $q_\iota \in Q$ is the initial mode, $\delta\colon Q \times S \to Q$ is a transition function, and $o\colon Q \times S \to \mathcal{D}(Act)$ is an output function with $\mathsf{supp}(o(q, s)) \subseteq Act(s)$ for all $q \in Q, s \in S$. σ is called *finite-memory* if $|Q| < \infty$, *memoryless* if $|Q| = 1$, and *deterministic* if $|\mathsf{supp}(o(q, s))| = 1$ for all $q \in Q, s \in S$.

Given an MDP $\Gamma = (S, Act, \mathbf{P})$ with strategy $\sigma = (Q, q_\iota, \delta, o)$, we define the *induced MC* $\Gamma[\sigma] = (S \times Q, \mathbf{P}^\sigma)$ where $\mathbf{P}^\sigma((s, q), (s', q')) = \mathbf{P}(s, o(q, s), s')$ if $\delta(q, s) = q'$, and otherwise $\mathbf{P}^\sigma((s, q), (s', q')) = 0$. In the following, we only consider finite MDPs, but when considering an infinite-memory strategy, the resulting MC is countably infinite. In the context of algorithms, we always assume that the probabilities in the given MDPs, strategies, and probability thresholds are *rational* numbers encoded as numerator-denominator pairs in binary.

Given an MC $\mathcal{M} = (S, \mathbf{P})$ with a distinguished initial state $s \in S$, we consider the σ-algebra $\mathcal{F}$ generated by the *cylinder sets* $\{\pi S^\omega \mid \pi \in S^*\}$ and the associated probability measure $\mathsf{Pr}_s^{\mathcal{M}}\colon \mathcal{F} \to [0, 1]$ which is uniquely defined by requiring that for all $\pi = s_0 \dots s_k \in S^+$, $k \geq 0$, we have $\mathsf{Pr}_s^{\mathcal{M}}(\pi S^\omega) = \prod_{i=0}^{k-1} \mathbf{P}(s_i, s_{i+1})$ if $s_0 = s$, and $\mathsf{Pr}_s^{\mathcal{M}}(\pi S^\omega) = 0$ if $s_0 \neq s$. See, e.g. [5, Chapter 10] for more details. The sets in $\mathcal{F}$ are called *measurable*. Further, we define $\mathsf{Paths}_s^{\mathcal{M}} = \{s_0 s_1 \ldots \in S^\omega \mid s_0 = s \land \forall i \geq 0\colon \mathbf{P}(s_i, s_{i+1}) > 0\}$.

An *objective* for an MDP Γ is a measurable[1] set of paths $\Omega \subseteq S^\omega$. A *reachability* objective for Γ is of the form $\Diamond F = \{\pi \in S^\omega \mid \exists k \geq 0\colon \pi(k) \in F\}$ where $F \subseteq S$. *Bounded reachability* objectives have the form $\Diamond^{\leq B} F = \{\pi \in S^\omega \mid \exists k \in [B]_0\colon \pi(k) \in F\}$, for some $B \in \mathbb{N}_0$. A *parity* objective for Γ is defined via a *priority function* $\rho\colon S \to [k]_0$, where $k \in \mathbb{N}_0$. For $\pi \in S^\omega$, let $\inf(\pi) = \{s \in S \mid \forall i \geq 0\colon \exists j \geq i\colon \pi(j) = s\}$ be the set of states visited infinitely often on π. Then the parity objective defined by ρ is $\{\pi \in S^\omega \mid \max \rho(\inf(\pi)) \text{ is even}\}$. In the following, we identify the function ρ with the objective it defines.

2.2 Multi-Objectives and Pareto Frontiers

A *multi-objective (MO) formula* for MDP Γ is a syntactic object $\varphi = \bigwedge_{i=1}^n atom_i$ with $atom_i \in \{\mathtt{S}(\Omega), \mathtt{Pr}_{\sim p}(\Omega) \mid \Omega \text{ an objective for } \Gamma, \sim \in \{>, \geq\}, p \in [0, 1]\}$. An MC $\mathcal{M}$ with a distinguished initial state s satisfies a *threshold* constraint $\mathtt{Pr}_{\sim p}(\Omega)$ if $\mathsf{Pr}_s^{\mathcal{M}}(\Omega) \sim p$, a *sure* constraint $\mathtt{S}(\Omega)$ if $\mathsf{Paths}_s^{\mathcal{M}} \subseteq \Omega$, and it satisfies the formula φ (in symbols: $s \models_{\mathcal{M}} \varphi$) if it satisfies $atom_i$ for all $i \in [n]$. For an MDP Γ with strategy σ we write $s, \sigma \models_\Gamma \varphi$ if $s \models_{\Gamma[\sigma]} \varphi$, and $s \models_\Gamma \varphi$ if $s, \sigma \models_\Gamma \varphi$ for *some* strategy σ for Γ. In this paper, we only consider formulas of the form

[1] Measurability is actually only important for probabilistic objectives, not for sure objectives. However, all concrete objectives considered in this paper are measurable.

$\mathtt{S}(\rho) \wedge \mathtt{Pr}_{\sim p_1}(\Diamond F_1) \wedge \ldots \wedge \mathtt{Pr}_{\sim p_n}(\Diamond F_n)$, i.e., conjunctions of one sure parity and $n \geq 1$ reachability objectives, either all with strict or non-strict thresholds.

We now define Pareto frontiers. Given an MO formula φ for MDP Γ containing n threshold constraints $\mathtt{Pr}_{\sim p_1}(\Omega_1), \ldots, \mathtt{Pr}_{\sim p_n}(\Omega_n)$ (in this order), we write $\varphi(p_1, \ldots, p_n) = \varphi(\mathbf{p})$ to emphasize the dependency of φ on the threshold vector $\mathbf{p} \in [0,1]^n$. We also write $\varphi(\mathbf{x})$ without further qualifying $\mathbf{x}$ to indicate that the thresholds are *variables* $\mathbf{x} = (x_1, \ldots, x_n)$. We define the set of *achievable threshold vectors* as $\mathit{Ach}(\Gamma, s, \varphi(\mathbf{x})) = \{\mathbf{p} \in [0,1]^n \mid s \models_\Gamma \varphi(\mathbf{p})\} \subseteq [0,1]^n$. Note that $\mathit{Ach}(\Gamma, s, \varphi(\mathbf{x}))$ is *downward-closed* as $\sim \in \{\geq, >\}$, and *convex* since a convex combination $c \cdot \mathbf{p} + (1-c) \cdot \mathbf{p}'$, $c \in (0,1)$, of achievable threshold vectors $\mathbf{p}, \mathbf{p}'$ is achieved by a strategy that plays the strategy for $\mathbf{p}$ with probability c and the one for $\mathbf{p}'$ with probability $1-c$, see, e.g. [25]. Given a set of vectors X, we define $\mathrm{gen}(X)$, the subspace generated by X, as the intersection of all subspaces of $\mathbb{R}^n$ containing X. It is the smallest subspace containing X. For the next definition recall that the *boundary* ∂X of a set $X \subseteq [0,1]^n$ is defined as $\overline{X} \setminus \operatorname{int} X$, where $\overline{X}$ is the closure of X and $\operatorname{int} X$ is the interior, i.e., the largest open subset of X.

Definition 1 (Pareto frontier). *Let $\Gamma = (S, \mathit{Act}, \mathbf{P})$ be an MDP, $\varphi(\mathbf{x})$ an MO formula for Γ, and $s \in S$. We define $\mathit{Pareto}(\Gamma, s, \varphi(\mathbf{x})) = \partial \mathit{Ach}(\Gamma, s, \varphi(\mathbf{x}))$.*

The above definition is similar to the one from [3]; other authors define the Pareto frontier in a slightly different way, e.g., as the $\leq$-maxima of $\mathit{Ach}(\Gamma, s, \varphi(\mathbf{x}))$ [26].

The Pareto frontier is the boundary of a convex polytope of dimension at most n [25]. Such a polytope P has faces of lower dimension, from 0 (a vertex) to $n-1$. These faces are defined as follows: given a hyperplane H intersecting P, the polytope $H \cap P$ is a face of P iff P lies fully on one of the two closed half-spaces defined by H. When considering the polytope associated to a Pareto frontier (and by generalization the Pareto frontier itself), we can freely separate points between those strictly in the interior, those on the border, and those in the exterior. In what follows, we only consider faces defined by intersection with hyperplanes whose normal vectors only have non-negative components. More on convex polytopes can be found in [27].

Example 1. Consider the MDP Γ in Figure 1 on page 2 and the formula $\varphi(\mathbf{x}) = \mathtt{Pr}_{\geq x_1}(\Diamond F_1) \wedge \mathtt{Pr}_{\geq x_2}(\Diamond F_2)$ (we ignore parity). Let s be the marked initial state.

- $\mathbf{p}_1 = (1, 1/3)$ is achievable from s by choosing pair_1, and then playing a repeatedly to reach F_1 with $2/3$, and both F_1 and F_2 with probability $1/3$.
- $\mathbf{p}_2 = (1/3, 1)$ is achievable from s by choosing pair_2, and then a repeatedly to reach F_2 with $2/3$ and both F_1 and F_2 with probability $1/3$.
- As mentioned earlier, a convex combination of two achievable points is achievable by following one of the two strategies with suitable probabilities. However, in this specific example, the vector $\mathbf{p}_3 = (2/3, 2/3) = 0.5 \cdot \mathbf{p}_1 + 0.5 \cdot \mathbf{p}_2$ is achievable with a *deterministic* strategy as well: First choose pair_2 in s and then b to reach state F_1 with $1/3$, state F_2 with $1/3$, and both F_1 and F_2 with probability $1/3$. These points will be relevant in Sections 4 and 5.

– The above strategies are all *Pareto-optimal*, but there are also sub-optimal strategies, e.g., choosing $pair_1$ in s and then b leads to reaching F_1 and F_2 with probability 0 and 1, respectively. This is sub-optimal as $(0,1) \leq \mathbf{p}_2$.

We consider *clean* MDPs throughout the rest of the paper:

Definition 2 (Clean MDP). *Let $\Gamma = (S, Act, \mathbf{P})$ be an MDP. Γ is* clean *...*

– *... w.r.t. a parity objective $\rho\colon S \to [k]_0$ if for all $s \in S$, we have $s \models_\Gamma \mathtt{S}(\rho)$, i.e., ρ is* surely *satisfiable from every state s.*
– *... w.r.t. target sets $F_1, \dots, F_n \subseteq S$ if for all $s \in S$, we have $s \models_\Gamma \mathtt{Pr}_{\geq 1}(\Diamond F)$, where $F = \bigcup_{i=1}^n F_i$, and every state in F is a sink.*

Example 2. The MDP from Figure 1 is clean w.r.t. ρ because from every state, there is a strategy that surely reaches a sink with priority 0. For instance, from the topmost state, the rightmost sink is reachable by playing action b. The MDP is also clean w.r.t. F_1, F_2 because from every state there exists a strategy reaching $F = F_1 \cup F_2$ with probability one, and because F contains sink states only.

Some remarks about clean MDPs are in order: (i) One can *clean* an MDP w.r.t. parity by identifying and removing states that violate $\mathtt{S}(\rho)$. The latter can be done by solving the 2-player deterministic parity game obtained by replacing the randomness in the MDP by an antagonistic player. Note that deciding the winner in a parity game (and hence checking if $s \models_\Gamma \mathtt{S}(\rho)$ holds for state s) is in $\mathsf{NP} \cap \mathsf{coNP}$ [14], and even in $\mathsf{UP} \cap \mathsf{coUP}$ [31], but is not known to be in PTIME. (ii) Reachability towards *sinks* only is a more severe restriction. We make it because simultaneous almost-sure reachability of n general target sets in an MDP is already PSPACE-complete [36]. Intuitively, this is because strategies have to remember which targets were already seen. Contrarily, sink reachability often admits more practical complexities as shown in Sections 3 and 4 (also see, e.g., [17] and [39]) and is still of practical interest. We leave a solution for general reachability for future work, since it would likely further improvements on the techniques we introduce.

3 Sure Parity and n Strict Reachability Thresholds

We study MO formulas of the form $\mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{>p_i}(\Diamond F_i)$ in this section, i.e., with *strict* thresholds only. Non-strict thresholds are more involved, see Section 5. We start by stating the main result of this section. Note that it is formulated for MDPs that are clean w.r.t. the parity objective ρ and the target sets F_i. The assumption of being clean w.r.t. parity can be dropped, but this incurs the additional complexity of solving a parity game (see Definition 2 and subsequent remarks), and hence leads to an $\mathsf{NP} \cap \mathsf{coNP}$ complexity bound on the associated decision problem.

Theorem 1. *Let Γ be a clean MDP w.r.t. parity objective ρ and target sets $F_1, \ldots, F_n \subseteq S$. Further, let $\mathbf{p} \in [0,1]^n$, and let $s \in S$ be a state. Then:*

1. *The decision problem $\exists \sigma\colon s, \sigma \models_\Gamma \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{>p_i}(\Diamond F_i)$ is in* PTIME.
2. *A witness strategy σ using at most $2^{\mathsf{poly}(|\Gamma|+nD)}$ memory, where D is the bit-complexity of the rational numbers in $\mathbf{p}$, can be effectively constructed for the* Yes*-instances.*

Proof (sketch). Using Corollary 3.5 of [25], we can test if $\varphi(\mathbf{p}) = \bigwedge_{i=1}^n \mathtt{Pr}_{>p_i}(\Diamond F_i)$ is achievable (note that we have dropped the parity objective). If it is not, then the answer is clearly No. Otherwise, there exists a memoryless but possibly randomized strategy achieving $\varphi(\mathbf{p})$. As the inequalities in $\varphi(\mathbf{p})$ are strict, it can be shown that the reachability thresholds can be guaranteed after playing the strategy for some *finite* but exponential number of steps. As parity objectives are prefix-independent, we can then simply switch to a memoryless deterministic winning strategy for parity. For the latter argument to work, it is crucial that the MDP is clean w.r.t. the parity objective ρ, i.e., that it is possible to satisfy $\mathtt{S}(\rho)$ from every state of the MDP.

Example 3. Reconsider the MDP Γ from Figure 1 with initial state s, yellow target F_1 and blue target F_2. To surely satisfy ρ, we must visit non-sink states only finitely often (this is a co-Büchi condition). We show that $s \models_\Gamma \mathtt{S}(\rho) \wedge \mathtt{Pr}_{>1/2}(\Diamond F_1) \wedge \mathtt{Pr}_{>1/6}(\Diamond F_2)$, achieving a value strictly greater than $\mathbf{p}_4 = (1/2, 1/6)$ which is strictly inside the Pareto frontier. To achieve this objective, we take the following strategy: we first play action $pair_1$, then a twice. By doing so, we have probability $1/4$ of reaching the leftmost state (contained in both F_1 and F_2), and probability $1/2$ of reaching the state only fulfilling F_1. If we now play action b, we satisfy condition ρ surely. We end up reaching F_1 with probability $3/4 > 1/2$ and reaching F_2 with probability $1/2 > 1/6$. Note that in general, thresholds that are achievable with strict inequalities are located strictly inside the Pareto frontier.

4 Sure Parity and Lexicographic Reachability

We are now interested in surely satisfying a parity objective while maximizing the probability of reaching n target sets in lexicographic order. Towards this goal we define the notion of *projection* in Definition 3, a concept also used extensively in Section 5. We then propose an algorithm using projection and prove it correct.

Recall that the *lexicographic order* on $[0,1]^n$ is the total order defined as $\mathbf{x} <_{lex} \mathbf{y}$ iff there is $k \in [n]$ such that (i) $x_k < y_k$ and (ii) $x_i = y_i$ for all $i \in [k-1]$. In the following, the order of our target sets $F_1, \ldots, F_n$ is relevant: For all $i, j \in [n]$, F_i appears before F_j iff F_i is more important than F_j.

One of the difficulties is that when considering the set of achievable points, the lexicographic supremum may not be achievable, i.e., the lexicographic maximum may not exist. We now formally give the main result of this section.

Theorem 2. *Let MDP Γ be clean w.r.t. parity objective ρ and target sets $F_1, \ldots, F_n \subseteq S$, and let $s \in S$ be a state. Then:*

- *It is decidable if $\mathbf{p}^* = \max_{lex}\{\mathbf{p} \in [0,1]^n \mid s \models_\Gamma \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq p_i}(\Diamond F_i)\}$ exists by solving $\mathcal{O}(\mathsf{poly}(n))$ many parity games (hence the problem is in $\mathsf{NP} \cap \mathsf{coNP}$ [11]).*
- *A witnessing strategy using at most $2|\Gamma||\rho|$ memory can be effectively constructed for the* Yes*-instances.*

Our approach considers every target set F_i one by one, following the lexicographic order. The general idea is to successively remove all transitions that do not achieve the maximal probability to reach F_i. Thus, after having pruned transitions w.r.t. the first i target sets, any strategy that maximizes the probability to reach the set $F = F_1 \cup \ldots \cup F_n$ also maximizes the probabilities of reaching $F_1, \ldots, F_i$ lexicographically [18]. In order to find the maximal probability to achieve a single objective, we adapt the notion of *projection* from [25,26]. The main difference is that we keep reachability objectives, instead of converting them into reward objectives, enabling us to use existing results [8] on combinations of sure and almost-sure objectives.

We define the MDP $\Gamma^{\pi\mathbf{v}}$, the projection of MDP Γ on a non-zero vector $\mathbf{v} \in [0,1]^n$ where we can freely assume $\|\mathbf{v}\|_1 = 1$. Intuitively, to obtain $\Gamma^{\pi\mathbf{v}}$, we consider a k-dimensional face of the Pareto frontier of $\bigwedge_{i=1}^n \Pr_{\geq p_i}(\Diamond F_i)$, maximal in the direction $\mathbf{v}$. This is thus an intersection with a hyperplane, and defines a face of dimension k. We remove all available actions that are used in none of the strategies achieving this face of dimension k, i.e. we remove all non-optimal actions when trying to maximize in the direction $\mathbf{v}$. Our purpose is to obtain a new MDP, in which every strategy that almost-surely reaches a final state in F also maximizes the probability to reach these states weighted with the direction $\mathbf{v}$. We remark that in this new MDP, the parity condition ρ may not be surely satisfied from every state; we will thus need to address this condition later.

Definition 3 (Projection). *Let $\Gamma = (S, Act, \mathbf{P})$ be clean w.r.t. $F_1, \ldots, F_n \subseteq S$. The* projection *$\Gamma^{\pi\mathbf{v}}$ of Γ in direction $\mathbf{v} \geq \mathbf{0}$ with $\|\mathbf{v}\|_1 = 1$, is defined in two steps: (1) Let $\Gamma' = (S', Act, \mathbf{P}')$ be an MDP where*

- *$S' = S \cup \{\bot\}$ where $\bot$ is a fresh sink state, and*
- *$\mathbf{P}'$ is defined similar to $\mathbf{P}$ with the following modifications (let $F = \bigcup_{i=1}^n F_i$):*
  - *For all $s \in S \setminus F$, $a \in Act(s)$, and $s' \in F$, we set $\mathbf{P}'(s,a,s') = \mathbf{P}(s,a,s') \cdot \sum_{i:s' \in F_i} v_i$ and $\mathbf{P}'(s,a,\bot) = 1 - \sum_{s'' \in S} \mathbf{P}'(s,a,s'')$.*
  - *$\mathbf{P}'(\bot, a, \bot) = 1$, where $a \in Act$ is arbitrary.*

(2) For each state $s \in S'$, let $y_s = \max_\sigma \mathsf{Pr}_s^{\Gamma'[\sigma]}(\Diamond F)$ be the maximum probability to reach F from s in Γ'. The MDP $\Gamma^{\pi\mathbf{v}}$ is then obtained from Γ' by removing all actions $a \in Act(s)$ that do not satisfy $y_s = \sum_{s' \in S'} \mathbf{P}'(s,a,s') \cdot y_{s'}$.

Example 4. The MDP in Figure 2 results from projecting the MDP from Figure 1 on $\mathbf{v} = (0,1)$ ($\bot$ is not reachable). Only the actions reaching the blue target F_2 with maximal probability remain.

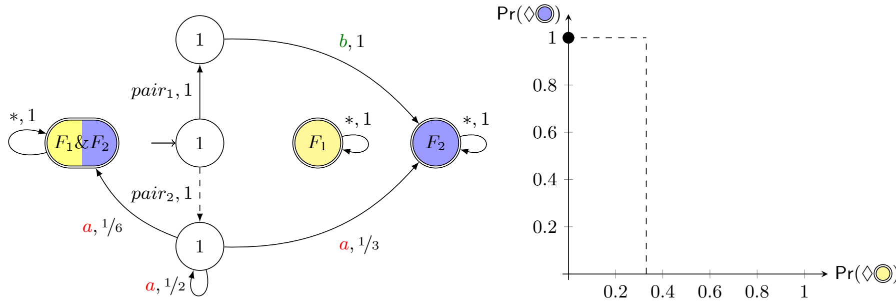

Fig. 2. The MDP from Figure 1 projected on $\mathbf{v} = (0, 1)$. *Right:* The set of achievable points w.r.t. $\mathtt{S}(\rho) \wedge \mathtt{Pr}_{\geq x_1}(\Diamond F_1) \wedge \mathtt{Pr}_{\geq x_2}(\Diamond F_2)$ is $[0, 1/3) \times [0, 1) \cup \{(0, 1)\}$; the lexicographic maximum for order F_2, F_1 is thus $\mathbf{p}^* = (1, 0)$.

Algorithm 1 Sure parity and lexicographic reachability

Input: MDP Γ – clean w.r.t. parity objective ρ and $F_1, \ldots, F_n \subseteq S$, a state $s \in S$
Output: If $\mathbf{p}^* = \max_{lex}\{\mathbf{p} \in [0,1]^n \mid s \models_\Gamma \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq p_i}(\Diamond F_i)\}$ exists, then the output is a witness strategy σ, otherwise the output is `false`.

$\Gamma_0 \leftarrow \Gamma$
$F \leftarrow \bigcup_{i=1}^n F_i$
for i from 1 to n **do**
  Compute $\Gamma_{i-1}^{\pi \mathbf{e}_i}$ ▷ See Definition 3.
  $\Gamma_i \leftarrow$ result of pruning all states not satisfying $\mathtt{S}(\rho) \wedge \mathtt{Pr}_{=1}(\Diamond F)$ in $\Gamma_{i-1}^{\pi \mathbf{e}_i}$.
end for
if s is not a state of Γ_n **then return** `false`
else return σ such that $s, \sigma \models_{\Gamma_n} \mathtt{S}(\rho) \wedge \mathrm{Pr}_{=1}(\Diamond F)$ ▷ By Theorem 2.
end if

Given Γ, $F_1, \ldots, F_n$ and $\mathbf{v}$, it is clear from Definition 3 that we can construct the projection $\Gamma^{\pi \mathbf{v}}$ in polynomial time. Note that strategies in $\Gamma^{\pi \mathbf{v}}$ are still valid in Γ, but the converse is not necessarily the case as projection removes actions.

Lemma 1. *(Key property of projection) Let $\Gamma = (S, Act, \mathbf{P})$ be clean w.r.t. $F_1, \ldots, F_n, \perp \subseteq S$, and let $\mathbf{v} \geq \mathbf{0}$, $\|\mathbf{v}\|_1 = 1$. Then for all strategies σ of $\Gamma^{\pi \mathbf{v}}$, the following holds: $s, \sigma \models_{\Gamma^{\pi \mathbf{v}}} \mathtt{Pr}_{=1}(\Diamond F)$ iff there exists $\mathbf{x} \in [0,1]^n$ such that (i) $\mathbf{x} \cdot \mathbf{v}$ is maximal among the achievable $\mathbf{x}$, and (ii) $s, \sigma \models_\Gamma \bigwedge_{i=1}^n \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$.*

5 Sure Parity and n non-Strict Reachability Thresholds

Finally, we consider the case of one sure parity condition and multiple *non-strict* threshold reachability objectives, i.e., formulas like $\mathtt{S}(\rho) \bigwedge_{i=1}^n \mathtt{Pr}_{\geq p_i}(\Diamond F_i)$. We do not impose a lexicographic ordering on the target sets. Our main result is:

Theorem 3. *Let MDP Γ be clean w.r.t. parity objective ρ and target sets $F_1, \ldots, F_n \subseteq S$. Further, let $\mathbf{p} \in [0,1]^n$, and let $s \in S$ be a state. Then:*

1. *The decision problem $\exists \sigma \colon s, \sigma \models_\Gamma \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq p_i}(\Diamond F_i)$ is in* $\mathsf{EXPTIME}$.
2. *A witness strategy σ using at most $2^{\mathsf{poly}(|\Gamma|+nD)}$ memory, where D is the bit-complexity of the rational numbers in $\mathbf{p}$, can be effectively constructed for the* Yes*-instances.*

We solved the case where $\mathbf{p}$ is strictly inside the Pareto frontier in Section 3. It remains to show how to achieve $\mathtt{S}(\rho) \bigwedge_{i=1}^n \mathtt{Pr}_{\geq p_i}(\Diamond F_i)$ when $\mathbf{p}$ is exactly on the frontier. We first consider the case where $\mathbf{p}$ is a vertex of the Pareto frontier, that we will then use as a base case for an arbitrary point $\mathbf{p}$ of the frontier. We sketch the proof in the remainder of the section.

Since the Pareto frontier does not depend on the sure objective, to determine whether $\mathbf{p}$ is exactly on the Pareto frontier, it suffices to check if $s \models_\Gamma \bigwedge_{i=1}^n \mathtt{Pr}_{\geq p_i}(\Diamond F_i)$ and $s \not\models_\Gamma \bigwedge_{i=1}^n \mathtt{Pr}_{> p_i}(\Diamond F_i)$. The first formula checks if $\mathbf{p}$ is achievable, the second checks whether it is on the boundary. Hence the main difficulty is to decide whether adding a sure parity condition keeps the achievability of a point. We illustrate this in the following example.

Example 5. In the MDP of Figure 1, playing $pair_1$ then a forever gives a total probability of 1 of reaching F_1 and probability $1/3$ of reaching F_2. This strategy does not surely satisfy the parity condition though, since there exists a path that visits the uppermost state, labelled 1, forever. Playing $pair_1$ and then b once surely reaches F_2. It is thus possible to satisfy $\mathtt{S}(\rho) \wedge \mathtt{Pr}_{=1}(\Diamond F_2)$, but not $\mathtt{S}(\rho) \wedge \mathtt{Pr}_{=1}(\Diamond F_1) \wedge \mathtt{Pr}_{\geq 1/3}(\Diamond F_2)$. Still, for every $\varepsilon > 0$, we can achieve $\mathtt{S}(\rho) \wedge \mathtt{Pr}_{\geq 1-\varepsilon}(\Diamond F_1) \wedge \mathtt{Pr}_{\geq 1/3-\varepsilon}(\Diamond F_2)$ by Theorem 1, using a finite-memory strategy.

5.1 Vertex of the Pareto frontier

We first consider the easier case where we want to achieve a point which is a vertex of the Pareto frontier. We assume $\mathbf{p}$ to be a vertex of the Pareto frontier. Our proof relies on projection (Definition 3). Indeed, since $\mathbf{p}$ is a vertex, there exists some vector $\mathbf{v}$ such that $\mathbf{p}$ is the *unique* point of the Pareto frontier maximizing $\mathbf{p} \cdot \mathbf{v}$. We obtain the following lemma.

Lemma 2. *Suppose that the MDP Γ is clean w.r.t. parity objective ρ and target sets $F_1, \ldots, F_n \subseteq S$. Further, let $\mathbf{p} \in [0,1]^n$, and let $s \in S$ be a state. If $\mathbf{p}$ is a vertex of the Pareto frontier of $\bigwedge_{i=1}^n \mathtt{Pr}_{\geq p_i}(\Diamond F_i)$ from s, then we can decide if $s \models_\Gamma \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq p_i}(\Diamond F_i)$ and if so give a finite-memory strategy.*

5.2 Arbitrary Point of the Pareto Frontier

We now consider any arbitrary point $\mathbf{p}$ of the Pareto frontier. Since $\mathbf{p}$ may be contained in a k-dimensional face of the frontier (with $k > 0$; $k = 0$ means that $\mathbf{p}$ is a vertex, see Section 5.1), projecting on this face will not be sufficient

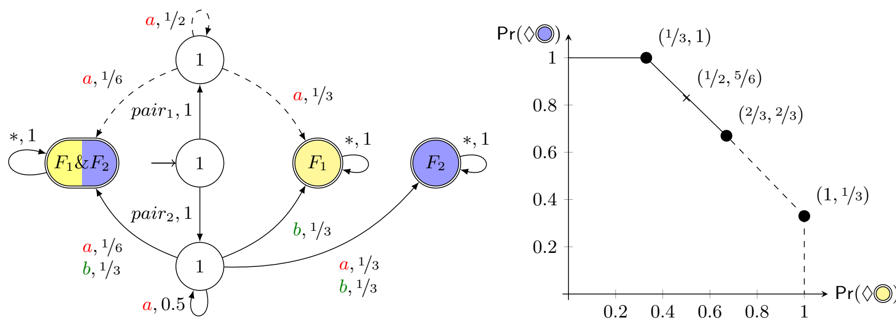

Fig. 3. The MDP of Figure 1 after projection on $\mathbf{v} = (1, 1)$.

to obtain $\mathbf{p}$. Similar to Algorithm 1, we iterate projections and state removal, thereby reducing the dimension of the Pareto frontier until either reducing $\mathbf{p}$ to a vertex, or entering a situation where we cannot project anymore. We remark that the latter only happens in specific cases. To properly define these cases, given a Pareto frontier P, we consider the smallest vector space containing P. We show that we cannot project any more when $\mathbf{p}$ is an interior point of P for this subspace, denoted $\mathbf{p} \in \mathrm{int}_{\mathrm{gen}(P)}(P)$.

Example 6. After projecting the MDP of Figure 1 on $\mathbf{v} = (1, 1)$, we obtain the MDP in Figure 3 (including both the dashed and the solid transitions). Since from the uppermost state, no strategy surely satisfies the parity condition, $\mathtt{S}(\rho) \wedge \mathtt{Pr}_{=1}(\Diamond F)$ does not hold. We thus prune the dashed transitions, and the Pareto frontier is now restricted between $(1/3, 1)$ and $(2/3, 2/3)$, see Figure 3 (right). To achieve e.g. $\mathbf{p} = (1/2, 5/6)$, which is strictly inside this line segment, it suffices to play $pair_2$, then a once, and then finally b once.

To obtain our result, we get the following lemma. After projecting on a given vector, and removing any state refuting $\mathtt{S}(\rho) \wedge \mathtt{Pr}_{=1}(\Diamond F)$ we obtain some polytope P; any point of P that is a topologically interior point in the smallest vector space containing P is achievable. Formally:

Lemma 3. *Let the MDP Γ be clean w.r.t. parity objective ρ and target sets $F_1, \ldots, F_n \subseteq S$. Further, let $\mathbf{v} \in [0,1]^n$, and $s \in S$. Let Γ_ρ be obtained by taking the MDP $\Gamma^{\pi \mathbf{v}}$ and pruning all states that refute $\mathtt{S}(\rho) \wedge \mathtt{Pr}_{=1}(\Diamond F)$. Let $B \subseteq [0,1]^n$ be the set of $\leq$-maximal points of the Pareto frontier of Γ_ρ from s. Then: For every $\mathbf{x} \in \mathrm{int}_{\mathrm{gen}(B)}(B)$, we have $s \models_{\Gamma_\rho} \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^{n} \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$, and we can compute a strategy that achieves this.*

The proof of this lemma is quite involved. Figure 4 provides some intuition on the proof. If $\mathbf{x}$ is inside an m-dimensional surface B, we can find $m+1$ elements of B such that $\mathbf{x}$ is within their convex hull. These are y^1, y^2, y^3 in Figure 4. For every y^j, we can find a strategy satisfying $\bigwedge_{i \in [n]} \mathrm{Pr}_{\geq y^j_i}(\Diamond F_i)$. By

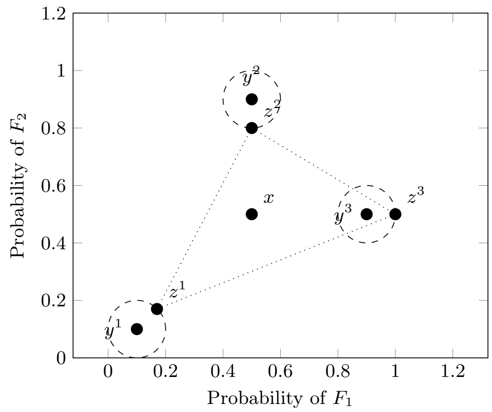

Fig. 4. To obtain x on the two-dimension plane, we first take three points y^1, y^2, y^3, then find ε small enough for x to be within the convex hull of any ε-approximation z^1, z^2, z^3 of y^1, y^2, y^3.

playing such a strategy for sufficiently many steps, then switching to a strategy satisfying $\mathtt{S}(\rho) \wedge \Pr_{=1}(\Diamond F)$, we can ε-approximate the y^j while staying inside B. Points achieved by such approximation are denoted z^1, z^2, z^3 in Figure 4. It then remains to show that if ε is small enough, $\mathbf{x}$ is within the convex hull of z^1, z^2, z^3 and thus can be achieved.

We can now give our result stating that we can verify the achievability of an arbitrary point of the Pareto frontier.

Lemma 4. *Suppose that the MDP Γ is clean w.r.t. parity objective ρ and target sets $F_1, \ldots, F_n \subseteq S$. Further, let $s \in S$ be a state, and*
$\mathbf{x} = (x_i)_{i\in[n]} \in [0,1]^n$ *on the border of the Pareto frontier of* $\bigwedge_{i\in[n]} \Pr_{\geq x_i}(\Diamond F_i)$

- *Checking whether $s \models_\Gamma \mathtt{S}(\rho) \wedge \bigwedge_{i\in[n]} \Pr_{\geq x_i}(\Diamond F_i)$ is decidable;*
- *if* Yes *a witnessing strategy with at most $2|\Gamma||\,\rho\,|$ memory can be effectively computed.*

Proof. We show that Algorithm 2 answers Lemma 4. The main differences with Alg. 1 is that we must find the vector on which we project, and at the end of the loop of Alg. 2 we have to split between the cases where $\mathbf{x}$ is a vertex or not.

During iteration i, we first find a vector orthogonal to the face of P_i that $\mathbf{x}$ belongs to. To do so, we may need to fully compute the Pareto frontier of Γ_{i-1}. In line 4 we project the current MDP Γ_{i-1} on vector $\mathbf{v}_i$, obtaining $\Gamma_{i-1}^{\pi\mathbf{v}_i}$. By Lemma 1, a strategy satisfies $\Pr_{=1}(\Diamond F)$ in $\Gamma_{i-1}^{\pi\mathbf{e}_i}$ iff it maximizes p such that $\Pr_{\geq p}(\Diamond F_i)$ in Γ_{i-1}. We then prune all states that do not satisfy the conjunction with the parity condition $\mathtt{S}(\rho) \wedge \Pr_{=1}(\Diamond F)$.

If $\mathbf{x}$ is an interior point, it follows by Lemma 3 that we can find a strategy σ such that $s, \sigma \models_{\Gamma_i} \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$. Otherwise, we start the loop again.

Algorithm 2 Sure parity and n non-strict reachability threshold objectives

Input: MDP Γ – clean w.r.t. ρ and $F_1, \ldots, F_n \subseteq S$, a state $s \in S$, a vector $\mathbf{x} \in [0,1]^n$
Output: A strategy σ such that $s, \sigma \models_\Gamma \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$ if it exists, else `false`.

1: Set $\Gamma_0 = \Gamma$ and $i = 1$.
2: **while** $\mathbf{x}$ is not a vertex of P_i, the Pareto frontier of Γ_i **do**
3: Get $\mathbf{v}_i$ a vector orthogonal to the smallest face of P_i that $\mathbf{x}$ belongs to.
4: Compute $\Gamma_i^{\pi \mathbf{v}_i}$ from s. ▷ By Lemma 1.
5: Set Γ_i by taking $\Gamma_{i-1}^{\pi \mathbf{v}_i}$ and pruning all states that do not satisfy $\mathtt{S}(\rho) \wedge \mathtt{Pr}_{\geq 1}(\Diamond F)$.
6: If $\mathbf{x}$ is an interior point of P_i return σ s.t. $s, \sigma \models_{\Gamma_i} \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$ ▷ By Lemma 3.
7: $i := i + 1$
8: **end while**
9: Check if there exists σ such that $s \models_{\Gamma_i} \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$. ▷ By Lemma 2.
10: If such a σ does not exist then return `false`, else return σ.

Since every time we project, it is onto a face of the Pareto frontier of dimension smaller than the current Pareto frontier, we can only take the loop at most n times. After this we go to line 9, and since $\mathbf{x}$ is now a vertex of the Pareto frontier, we can use Lemma 2 to find whether there exists a strategy such that $s \models_{\Gamma_i} \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$.

If the output is `false`, we remark that since projection in line 4 keeps all states belonging to strategies such that $\Pr_{\geq p}(\Diamond F_i)$ in Γ_{i-1} (by Lemma 1), it keeps all states belonging to strategies such that $\bigwedge_{i=1}^n \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$. Step 5 may prune states used in strategies such that $\bigwedge_{i=1}^n \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$, but then by definition these strategies did not satisfy $\mathtt{S}(\rho)$. Hence, since none of the pruned states belonged to a strategies such that $s \models_{\Gamma_i} \mathtt{S}(\rho) \wedge \bigwedge_{i=1}^n \mathtt{Pr}_{\geq x_i}(\Diamond F_i)$, the algorithm is correct.

We show that Algorithm 2 solves at most a polynomial number of parity game of size polynomial in $|\Gamma|$.

Indeed, to obtain $|\Gamma_{i-1}^{\pi \mathbf{v}_i}|$, we have to compute the Pareto frontier of $|\Gamma_{i-1}|$. This new MDP only has one more state than Γ and at most $|\Gamma|$ additional transitions (that may lead from states originally in Γ to the new state $\perp$). We then remove once all states that do not satisfy $\mathtt{S}(\rho) \wedge \Pr_{=1}(\Diamond F)$, which can be done by solving a polynomial number of parity games. Every time we do this step, we project onto a face of the Pareto frontier of dimension smaller than the current Pareto frontier and this can only happen at most n times. Thus we end up solving a number polynomial in n of parity game of size polynomial in Γ, and compute Pareto frontiers for MDPs with n objectives at most n times.

We output `true` iff we find a strategy σ that is a solution of $s, \sigma \models_{\Gamma_i} \mathtt{S}(\rho) \wedge \Pr_{=1}(\Diamond F)$ iff it is a solution of $s \models_\Gamma \mathtt{S}(\rho) \wedge \bigwedge_{i \in [n]} \Pr_{\geq x_i}(\Diamond F_i)$, and we can find strategies satisfying the left hand formula that use $2|\Gamma||\rho|$ memory, as proved in Theorem 2.

Example 7. For the MDP from Figure 1, with initial state s, and where F_1 is the yellow target and F_2 is the blue target, we check if $s \models_\Gamma \mathtt{S}(\rho) \wedge \Pr_{\geq 2/3}(\Diamond F_1) \wedge \Pr_{\geq 2/3}(\Diamond F_2)$. Point $(2/3, 2/3)$ is on the Pareto frontier but not a vertex of it, and following Algorithm 2, a vector orthogonal to it is $(1,1)$. After projection on $(1,1)$, we obtain the MDP in Figure 3 (left), with both full and dashed transitions. As in Example 6, no strategy satisfies surely the parity condition in the uppermost state, we prune the dashed transitions, restricting the Pareto frontier to between $(1/3, 1)$ and $(2/3, 2/3)$. Now $(2/3, 2/3)$ is a vertex of the new Pareto frontier, Lemma 2 tells us to project on vector $(2,1)$, and so we prune transition a from the lowermost state, only keeping transition b. Since we can satisfy the parity condition from the lowermost state, we obtain that the strategy playing b twice satisfies $\mathtt{S}(\rho) \wedge \Pr_{\geq 2/3}(\Diamond F_1) \wedge \Pr_{\geq 2/3}(\Diamond F_2)$.

6 Conclusion and Future Work

Combining sure parity and n reachability threshold objectives can be done via a reduction to parity games in the case of strict thresholds and when maximizing the threshold lexicographically, and in exponential time with non-strict thresholds. Finite-memory strategies suffice in all cases. One direction for future work is to implement our algorithms in the probabilistic model checker Storm [30]. Further open problems include the case where targets are not sinks, and the study of one sure parity and n parity threshold objectives. However, the exact memory required for one sure and *one* almost-sure parity is already unknown. It seems worthwhile to investigate if 1-bit Markov strategies suffice, as they do in countable MDPs with parity objectives [32]. In [1,8], the solution of sure parity and almost-sure reachability in MDPs relies on a reduction to a game with a *fair* opponent. Results in [15] concern *stochastic games* with a fair opponent, and may thus help extending the results from [1,8] to stochastic games. Another possible extension is to consider combinations of multiple objectives in *partially observable* MDPs (POMDPs), as in [21].

References

1. Shaull Almagor, Orna Kupferman, and Yaron Velner. Minimizing Expected Cost Under Hard Boolean Constraints, with Applications to Quantitative Synthesis. In *CONCUR*, volume 59 of *LIPIcs*, pages 9:1–9:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.
2. Benjamin Aminof, Marta Kwiatkowska, Bastien Maubert, Aniello Murano, and Sasha Rubin. Probabilistic Strategy Logic. In *IJCAI*, pages 32–38. ijcai.org, 2019.
3. Pranav Ashok, Krishnendu Chatterjee, Jan Kretínský, Maximilian Weininger, and Tobias Winkler. Approximating Values of Generalized-Reachability Stochastic Games. In *LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8-11, 2020*, pages 102–115. ACM, 2020. `doi:10.1145/3373718.3394761`.
4. Christel Baier, Clemens Dubslaff, and Sascha Klüppelholz. Trade-off analysis meets probabilistic model checking. In *CSL-LICS*, pages 1:1–1:10. ACM, 2014.
5. Christel Baier and Joost-Pieter Katoen. *Principles of Model Checking*. MIT press, 2008.
6. Richard Bellman. A Markovian decision process. *Journal of mathematics and mechanics*, pages 679–684, 1957.
7. Raphaël Berthon, Shibashis Guha, and Jean-François Raskin. Mixing Probabilistic and non-Probabilistic Objectives in Markov Decision Processes. In *LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8-11, 2020*, pages 195–208. ACM, 2020. `doi:10.1145/3373718.3394805`.
8. Raphaël Berthon, Mickael Randour, and Jean-François Raskin. Threshold Constraints with Guarantees for Parity Objectives in Markov Decision Processes. In *ICALP*, volume 80 of *LIPIcs*, pages 121:1–121:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. `doi:10.4230/LIPIcs.ICALP.2017.121`.
9. Raphaël Berthon, Joost-Pieter Katoen, and Tobias Winkler. Markov Decision Processes with Sure Parity and Multiple Reachability Objectives, 2024. `arXiv:2408.01212`.
10. Patricia Bouyer, Mauricio González, Nicolas Markey, and Mickael Randour. Multi-weighted markov decision processes with reachability objectives. In *GandALF*, volume 277 of *EPTCS*, pages 250–264, 2018.
11. Gilles Brassard. A note on the complexity of cryptography (corresp.). *IEEE Transactions on Information Theory*, 25(2):232–233, 1979.
12. Tomás Brázdil, Krishnendu Chatterjee, Vojtech Forejt, and Antonín Kucera. MultiGain: A Controller Synthesis Tool for MDPs with Multiple Mean-Payoff Objectives. In *TACAS*, volume 9035 of *Lecture Notes in Computer Science*, pages 181–187. Springer, 2015.
13. Véronique Bruyère, Emmanuel Filiot, Mickael Randour, and Jean-François Raskin. Meet your expectations with guarantees: Beyond worst-case synthesis in quantitative games. *Information and Computation*, 254:259–295, 2017. `doi:10.1016/j.ic.2016.10.011`.
14. Cristian S. Calude, Sanjay Jain, Bakhadyr Khoussainov, Wei Li, and Frank Stephan. Deciding parity games in quasipolynomial time. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017*, pages 252–263. ACM, 2017. `doi:10.1145/3055399`.
15. Pablo F. Castro, Pedro R. D'Argenio, Ramiro Demasi, and Luciano Putruele. Playing Against Fair Adversaries in Stochastic Games with Total Rewards. In *CAV*

(2), volume 13372 of *Lecture Notes in Computer Science*, pages 48–69. Springer, 2022.

16. Krishnendu Chatterjee, Thomas A Henzinger, and Nir Piterman. Strategy logic. *Information and Computation*, 208(6):677–693, 2010.
17. Krishnendu Chatterjee, Joost-Pieter Katoen, Stefanie Mohr, Maximilian Weininger, and Tobias Winkler. Stochastic games with lexicographic objectives. *Formal Methods in System Design*, 2023. doi:10.1007/s10703-023-00411-4.
18. Krishnendu Chatterjee, Joost-Pieter Katoen, Maximilian Weininger, and Tobias Winkler. Stochastic Games with Lexicographic Reachability-Safety Objectives. In *CAV (2)*, volume 12225 of *Lecture Notes in Computer Science*, pages 398–420. Springer, 2020. doi:10.1007/978-3-030-53291-8_21.
19. Krishnendu Chatterjee, Zuzana Kretínská, and Jan Kretínský. Unifying Two Views on Multiple Mean-Payoff Objectives in Markov Decision Processes. *Log. Methods Comput. Sci.*, 13(2), 2017. doi:10.23638/LMCS-13(2:15)2017.
20. Krishnendu Chatterjee, Rupak Majumdar, and Thomas A. Henzinger. Markov Decision Processes with Multiple Objectives. In *STACS*, volume 3884 of *Lecture Notes in Computer Science*, pages 325–336. Springer, 2006.
21. Krishnendu Chatterjee, Petr Novotný, Guillermo A. Pérez, Jean-François Raskin, and Dorde Zikelic. Optimizing Expectation with Guarantees in POMDPs. In *AAAI*, pages 3725–3732. AAAI Press, 2017.
22. Krishnendu Chatterjee and Nir Piterman. Combinations of Qualitative Winning for Stochastic Parity Games. In *30th International Conference on Concurrency Theory, CONCUR 2019*, volume 140 of *LIPIcs*, pages 6:1–6:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. URL: http://www.dagstuhl.de/dagpub/978-3-95977-121-4.
23. Lorenzo Clemente and Jean-François Raskin. Multidimensional beyond worst-case and almost-sure problems for mean-payoff objectives. In *30th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2015, Kyoto, Japan, July 6-10, 2015*, pages 257–268. IEEE Computer Society, 2015. doi:10.1109/LICS.2015.33.
24. E. Allen Emerson, Charanjit S. Jutla, and A. Prasad Sistla. On model checking for the μ-calculus and its fragments. *Theor. Comput. Sci.*, 258(1-2):491–522, 2001. doi:10.1016/S0304-3975(00)00034-7.
25. Kousha Etessami, Marta Z. Kwiatkowska, Moshe Y. Vardi, and Mihalis Yannakakis. Multi-Objective Model Checking of Markov Decision Processes. *Logical Methods in Computer Science*, 4(4), 2008.
26. Vojtech Forejt, Marta Z. Kwiatkowska, and David Parker. Pareto Curves for Probabilistic Model Checking. In *ATVA*, volume 7561 of *Lecture Notes in Computer Science*, pages 317–332. Springer, 2012.
27. Branko Grünbaum, Victor Klee, Micha A Perles, and Geoffrey Colin Shephard. *Convex polytopes*, volume 16. Springer, 1967.
28. Ernst Moritz Hahn, Mateo Perez, Sven Schewe, Fabio Somenzi, Ashutosh Trivedi, and Dominik Wojtczak. Good-for-MDPs Automata for Probabilistic Analysis and Reinforcement Learning. In *TACAS (1)*, volume 12078 of *Lecture Notes in Computer Science*, pages 306–323. Springer, 2020.
29. Arnd Hartmanns, Sebastian Junges, Joost-Pieter Katoen, and Tim Quatmann. Multi-cost bounded tradeoff analysis in MDP. *J. Autom. Reason.*, 64(7):1483–1522, 2020.
30. Christian Hensel, Sebastian Junges, Joost-Pieter Katoen, Tim Quatmann, and Matthias Volk. The probabilistic model checker Storm. *Int. J. Softw. Tools Technol. Transf.*, 24(4):589–610, 2022.

31. Marcin Jurdzinski. Deciding the winner in parity games is in UP $\cap$ co-UP. *Inf. Process. Lett.*, 68(3):119–124, 1998. doi:10.1016/S0020-0190(98)00150-1.
32. Stefan Kiefer, Richard Mayr, Mahsa Shirmohammadi, and Patrick Totzke. Strategy Complexity of Parity Objectives in Countable MDPs. In *CONCUR*, volume 171 of *LIPIcs*, pages 39:1–39:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.
33. 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.
34. Shuwa Miura, Kyle Hollins Wray, and Shlomo Zilberstein. Heuristic Search for SSPs with Lexicographic Preferences over Multiple Costs. In *SOCS*, pages 127–135. AAAI Press, 2022.
35. Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* Wiley Series in Probability and Statistics. Wiley, 1994.
36. Mickael Randour, Jean-François Raskin, and Ocan Sankur. Percentile queries in multi-dimensional Markov decision processes. *Formal Methods in System Design*, 50(2):207–248, 2017. doi:10.1007/s10703-016-0262-7.
37. Diederik M. Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. A Survey of Multi-Objective Sequential Decision-Making. *J. Artif. Intell. Res.*, 48:67–113, 2013. doi:10.1613/jair.3987.
38. Sven Schewe, Alexander Weinert, and Martin Zimmermann. Parity Games with Weights. *Log. Methods Comput. Sci.*, 15(3), 2019. doi:10.23638/LMCS-15(3:20)2019.
39. Tobias Winkler and Maximilian Weininger. Stochastic Games with Disjunctions of Multiple Objectives. In *GandALF*, volume 346 of *EPTCS*, pages 83–100, 2021. doi:10.4204/EPTCS.346.6.
40. Kyle Hollins Wray, Shlomo Zilberstein, and Abdel-Illah Mouaddib. Multi-Objective MDPs with Conditional Lexicographic Reward Preferences. In *AAAI*, pages 3418–3424. AAAI Press, 2015.