Percentile Queries in Multi-dimensional Markov Decision Processes

Mickael Randour[1(✉)], Jean-François Raskin[2], and Ocan Sankur[2]

[1] LSV, CNRS and ENS Cachan, Cachan, France
mickael.randour@gmail.com
[2] Département d'Informatique, Université Libre de Bruxelles (U.L.B.), Brussels, Belgium

Abstract. Markov decision processes (MDPs) with multi-dimensional weights are useful to analyze systems with multiple objectives that may be conflicting and require the analysis of trade-offs. In this paper, we study the complexity of percentile queries in such MDPs and give algorithms to synthesize strategies that enforce such constraints. Given a multi-dimensional weighted MDP and a quantitative payoff function f, thresholds v_i (one per dimension), and probability thresholds α_i, we show how to compute a single strategy to enforce that for all dimensions i, the probability of outcomes ρ satisfying $f_i(\rho) \geq v_i$ is at least α_i. We consider classical quantitative payoffs from the literature (sup, inf, lim sup, lim inf, mean-payoff, truncated sum, discounted sum). Our work extends to the quantitative case the multi-objective model checking problem studied by Etessami et al. [16] in unweighted MDPs.

1 Introduction

Markov decision processes (MDPs) are central mathematical models for reasoning about (optimal) strategies in *uncertain environments*. For example, if rewards (given as numerical values) are assigned to actions in an MDP, we can search for a strategy (policy) that resolves the nondeterminism in a way that the *expected mean reward* of the actions taken by the strategy along time is maximized. See for example [23] for a solution to this problem. If we are risk-averse, we may want to search instead for strategies that ensure that the mean reward along time is larger than a given value with a high probability, i.e., a probability that exceeds a given threshold. See for example [17] for a solution.

Recent works are exploring several natural extensions of those problems. First, there is a series of works that investigate MDPs with multi-dimensional weights [6,12] rather than single-dimensional as it is traditionally the case. Multi-dimensional MDPs are useful to analyze systems with *multiple objectives* that are potentially conflicting and make necessary the analysis of trade-offs. For instance, we may want to build a control strategy that both ensures some good

Work partially supported by ERC starting grant inVEST (FP7-279499) and European project CASSTING (FP7-ICT-601148).

D. Kroening and C.S. Păsăreanu (Eds.): CAV 2015, Part I, LNCS 9206, pp. 123–139, 2015.
DOI: 10.1007/978-3-319-21690-4_8

quality of service and minimizes the energy consumption. Second, there are works that aim at synthesizing strategies enforcing *richer properties*. For example, we may want to construct a strategy that both ensures some minimal threshold with certainty (or probability one) and a good expectation [7]. An illustrative survey of such extensions can be found in [25].

Our paper participates in this general effort by providing algorithms and complexity results on the synthesis of strategies that enforce *multiple percentile constraints*. A *multi-percentile query* and the associated synthesis problem is as follows: given a multi-dimensionally weighted MDP M and an initial state s_{init}, synthesize a unique strategy σ such that it satisfies the conjunction of q constraints $\mathcal{Q} \coloneqq \bigwedge_{i=1}^{q} \mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}\left[f_{l_i} \geq v_i\right] \geq \alpha_i$, where each l_i refers to a dimension of the weight vectors, v_i is a value threshold, and α_i a probability threshold, and f a payoff function. Each constraint i expresses that the strategy ensures probability at least α_i to obtain payoff at least v_i in dimension l_i.

In this paper, we consider seven payoff functions: sup, inf, limsup, liminf, mean-payoff, truncated sum and discounted sum. This wide range covers most classical functions: our exhaustive study provides a *complete picture* for the new multi-percentile framework and we focus on establishing meta-theorems and connections whenever possible. Some of our results are obtained by reduction to the previous work of [16], but for mean-payoff, truncated sum and discounted sum, that are *non-regular payoffs*, we need to develop original techniques.

Consider some examples. In a stochastic shortest path problem, we may want a strategy ensuring that the probability to reach the target within d time units exceeds 50 percent: this is a single-constraint percentile query. With a *multi-constraint percentile query*, we can impose richer properties, for instance, enforcing that the duration is less than d_1 in 50 percent of the cases, and less than d_2 in 95 percent of the cases, with $d_1 < d_2$. We may also consider multi-dimensional systems. If in the model, we add information about fuel consumption, we may also enforce that we arrive within d time units in 95 percent of the cases, and that in half of the cases the fuel consumption is below some threshold c.

Contributions. We study percentile problems for a range of payoffs: we establish algorithms and prove complexity and memory bounds. Our algorithms solve multi-constraint multi-dimensional queries, but we also study interesting subclasses such as the single-dimensional case. We present an overview of our results in Table 1. For all payoff functions but the discounted sum, they only require *polynomial time in the size of the model* when the query size is fixed. In most applications, the query size is typically small while the model can be very large. So our algorithms have clear potential to be useful in practice.

(A) We show the PSPACE-hardness of the multiple reachability problem with exponential dependency on the query size (Theorem 2), and the PSPACE-completeness of the almost-sure case, refining the results of [16]. We also give a polynomial-time algorithm for *nested* target sets (Theorem 3).

(B) For inf, sup, lim inf and lim sup, we establish a polynomial-time algorithm for single-dimension (Theorem 5), and an algorithm that is only exponential in

Table 1. Some results for percentile queries. Here $\mathcal{F} = \{\inf, \sup, \liminf, \limsup\}$, $\overline{\mathsf{MP}}$ (resp. $\underline{\mathsf{MP}}$) stands for sup. (resp. inf.) mean-payoff, SP for shortest path, and DS for discounted sum. Parameters M and $\mathcal{Q}$ resp. represent model size and query size; $\mathrm{P}(x)$, $\mathrm{E}(x)$ and $\mathrm{P}_{ps}(x)$ resp. denote polynomial, exponential and pseudo-polynomial time in parameter x. All results without reference are new.

	Single-constraint	Single-dim. Multi-constraint	Multi-dim. Multi-constraint
Reachability	P [23]	$\mathrm{P}(M)\cdot\mathrm{E}(\mathcal{Q})$ [16], PSPACE-h.	—
$f \in \mathcal{F}$	P [10]	P	$\mathrm{P}(M)\cdot\ \mathrm{E}(\mathcal{Q})$
			PSPACE-h.
$\overline{\mathsf{MP}}$	P [23]	P	P
$\underline{\mathsf{MP}}$	P [23]	$\mathrm{P}(M)\cdot\mathrm{E}(\mathcal{Q})$	$\mathrm{P}(M)\cdot\mathrm{E}(\mathcal{Q})$
SP	$\mathrm{P}(M)\cdot\mathrm{P}_{ps}(\mathcal{Q})$ [21]	$\mathrm{P}(M)\cdot\mathrm{P}_{ps}(\mathcal{Q})$ (one target)	$\mathrm{P}(M)\cdot\mathrm{E}(\mathcal{Q})$
	PSPACE-h. [21]	PSPACE-h. [21]	PSPACE-h. [21]
ε-gap DS	$\mathrm{P}_{ps}(M, \mathcal{Q}, \varepsilon)$	$\mathrm{P}_{ps}(M, \varepsilon)\cdot\mathrm{E}(\mathcal{Q})$	$\mathrm{P}_{ps}(M, \varepsilon)\cdot\mathrm{E}(\mathcal{Q})$
	NP-h.	NP-h.	PSPACE-h.

the query size for the general case (Theorem 6). We prove PSPACE-hardness for sup (Theorem 7), and give a polynomial time algorithm for lim sup (Theorem 8).

(C) In the mean-payoff case, we distinguish $\overline{\mathsf{MP}}$ defined by the limsup of the average weights, and $\underline{\mathsf{MP}}$ by their liminf. For the former, we give a polynomial-time algorithm for the general case (Theorem 10). For the latter, our algorithm is polynomial in the model size and exponential in the query size (Theorem 11).

(D) The truncated sum function computes the *sum* of weights until a target is reached. It models *shortest path* problems. We prove the multi-dimensional percentile problem to be undecidable when both negative and positive weights are allowed (Theorem 12). Therefore, we concentrate on the case of non-negative weights, and establish an algorithm that is polynomial in the model size and exponential in the query size (Theorem 13). We derive from recent results that even the single-constraint percentile problem is PSPACE-hard [21].

(E) Discounted sum turns out to be linked to a long-standing open problem, not known to be decidable (Lemma 8). Nevertheless, we give an algorithm for an approximation called ε-gap percentile problem. It guarantees correct answers up to an arbitrarily small zone of uncertainty (Theorem 14). We prove this problem is PSPACE-hard in general, and NP-hard for single-constraint queries. According to a very recent preprint by Haase and Kiefer [20], our reduction even proves PP-hardness of single-contraint queries, which suggests that the problem does not belong to NP at all otherwise the polynomial hierarchy would collapse.

We systematically study the memory requirement of strategies. We build our algorithms using different techniques. Here are a few of them. For inf and sup payoff functions, we reduce percentile queries to multiple reachability queries, and rely on the algorithm of [16]: those are the easiest cases. For lim inf, lim sup and $\overline{\mathsf{MP}}$, we additionally need to resort to maximal end-component decomposition of MDPs. For the following cases, there is no simple reduction to existing problems and we need non-trivial techniques to establish algorithms. For $\underline{\mathsf{MP}}$, we use linear programming techniques to characterize winning strategies, borrowing ideas from [6,16]. For shortest path and discounted sum, we consider unfoldings of the MDP, with particular care to bound their sizes, and for the latter, to analyze the cumulative error due to necessary roundings.

Related Work. There are works that study multi-dimensional MDPs: for discounted sum, see [12], and for mean-payoff, see [6,17]. In the latter papers, the following threshold problem is studied: given a threshold vector $\boldsymbol{v}$ and a probability threshold ν, does there exist a strategy σ such that $\mathbb{P}^{\sigma}_{s}[\boldsymbol{r} \geq \boldsymbol{v}] \geq \nu$, where $\boldsymbol{r}$ denotes the mean-payoff vector. The work [17] solves it for the single dimensional case, and the multi-dimensional for the *non-degenerate* case (w.r.t. solutions of a linear program). A general algorithm was given in [6]. This problem asks for a bound on the *joint probability* of the thresholds, i.e., the probability of satisfying *all* constraints simultaneously. In contrast, we bound the *marginal probabilities* separately, which may allow for more modeling flexibility. Maximizing the *expectation vector* was considered in [6]. An approach unifying the probability and expectation views for mean-payoff was recently presented in [11].

Multiple reachability objectives in MDPs were considered in [16]: given an MDP and multiple targets T_i, thresholds α_i, decide if there exists a strategy that forces each T_i with a probability larger than α_i. This work is the closest to our work and we show here that their problem is inter-reducible with our problem for the sup measure. In [16] the complexity results are given only for model size and not for query size: *we refine those results and answer questions left open.*

Several works consider percentile queries but only for *one* dimension and *one* constraint (while we consider multiple constraints and dimensions) and particular payoff functions. Single-constraint queries for lim sup and lim inf were studied in [10]. The threshold probability problem for truncated sum was studied for either all non-negative or all non-positive weights in [22,26]. *Quantile queries* in the single-constraint case were studied for the shortest path with non-negative weights in [29], and for energy-utility objectives in [1]. It has been recently extended to *cost problems* [21], in a direction orthogonal to ours. For fixed horizon, [32] studies maximization of the expected discounted sum subject to a single percentile constraint. Still for the discounted case, there are works studying *threshold problems* [30,31] and *value-at-risk problems* [5]. All can be related to single-constraint percentiles queries.

A long version of this paper with full details is available online [24].

2 Preliminaries

A finite *Markov decision process* (MDP) is a tuple $M = (S, A, \delta)$ where S is the finite set of *states*, A is the finite set of *actions* and $\delta\colon S \times A \to \mathcal{D}(S)$ is a partial function called the *probabilistic transition function*, where $\mathcal{D}(S)$ denotes the set of rational probability distributions over S. The set of actions that are available in a state $, \in S$ is denoted by $A(s)$. We use $\delta(s, a, s')$ as a shorthand for $\delta(s, a)(s')$. An *absorbing state* s is such that for all $a \in A(s)$, $\delta(s, a, s) = 1$. We assume w.l.o.g. that MDPs are *deadlock-free*: for all $s \in S$, $A(s) \neq \emptyset$ (if not the case, we simply replace the deadlock by an absorbing state with a unique action). An MDP where for all $s \in S$, $|A(s)| = 1$ is a fully-stochastic process called a *Markov chain*.

A *weighted* MDP is a tuple $M = (S, A, \delta, w)$, where w is a *d-dimension weight function* $w\colon A \to \mathbb{Z}^d$. For any $l \in \{1, \ldots, d\}$, we denote $w_l\colon A \to \mathbb{Z}$ the projection of w to the l-th dimension, i.e., the function mapping each action a to the l-th element of vector $w(a)$. A *run* of M is an infinite sequence $s_1 a_1 \ldots a_{n-1} s_n \ldots$ of states and actions such that $\delta(s_i, a_i, s_{i+1}) > 0$ for all $i \geq 1$. Finite prefixes of runs are called *histories*.

Fix an MDP $M = (S, A, \delta)$. An *end-component* (EC) of M is an MDP $C = (S', A', \delta')$ with $S' \subseteq S$, $\emptyset \neq A'(s) \subseteq A(s)$ for all $s \in S'$, and $\mathsf{Supp}(\delta(s, a)) \subseteq S'$ for all $s \in S', a \in A'(s)$ (here $\mathsf{Supp}(\cdot)$ denotes the support), $\delta' = \delta|_{S' \times A'}$ and such that C is *strongly connected*, i.e., there is a run between any pair of states in S'. The union of two ECs with non-empty intersection is an EC; one can thus define *maximal* ECs. We let $\mathsf{MEC}(M)$ denote the set of maximal ECs of M, computable in polynomial time [14].

A *strategy* σ is a function $(SA)^*S \to \mathcal{D}(A)$ such that for all $h \in (SA)^*S$ ending in s, we have $\mathsf{Supp}(\sigma(h)) \subseteq A(s)$. The set of all strategies is Σ. We consider *finite-* and *infinite-memory* strategies as strategies that can be encoded by Moore machines with finite or infinite states respectively. An MDP M, initial state s, and a strategy σ determines a Markov chain M_s^σ on which a unique probability measure is defined. Here, M_s^σ is defined on the state space that is product of M and that of the Moore machine encoding σ. Given an event $E \subseteq (SA)^\omega$, we denote by $\mathbb{P}_{M,s}^\sigma[E]$ the probability of runs of M_s^σ whose projection to M is in E. That is the probability of achieving event E when the MDP M is executed with initial state s and strategy σ.

Let $\mathsf{Inf}(\rho)$ denote the random variable representing the disjoint union of states and actions that occur infinitely often in the run ρ. By an abuse of notation, we see $\mathsf{Inf}(\rho)$ as a sub-MDP M' if it contains exactly the states and actions of M'. It was shown that for any MDP M, state s, strategy σ, $\mathbb{P}_{M,s}^\sigma[\mathsf{Inf} \text{ is an EC}] = 1$ [14].

Multiple Reachability. Given a subset T of states, let $\Diamond T$ be the *reachability objective w.r.t. T*, defined as the set of runs visiting a state of T at least once.

The *multiple reachability* problem consists, given MDP M, state s_{init}, target sets $T_1, \ldots, T_q$, and probabilities $\alpha_1, \ldots, \alpha_q \in [0, 1] \cap \mathbb{Q}$, in deciding whether there exists a strategy $\sigma \in \Sigma$ such that $\bigwedge_{i=1}^q \mathbb{P}_{M,s_{\mathsf{init}}}^\sigma[\Diamond T_i] \geq \alpha_i$. The *almost-sure multiple reachability* problem restricts to $\alpha_1 = \ldots = \alpha_q = 1$.

Percentile Problems. We consider *payoff functions* among inf, sup, lim inf, lim sup, mean-payoff, truncated sum (shortest path) and discounted sum. For any run $\rho = s_1 a_1 s_2 a_2 \ldots$, dimension $l \in \{1, \ldots, d\}$, and weight function w,

– $\inf_l(\rho) = \inf_{j \geq 1} w_l(a_j)$, $\sup_l(\rho) = \sup_{j \geq 1} w_l(a_j)$,

- $\liminf_l(\rho) = \liminf_{j\to\infty} w_l(a_j)$, $\limsup_l(\rho) = \limsup_{j\to\infty} w_l(a_j)$,
- $\underline{\mathsf{MP}}_l(\rho) = \liminf_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} w_l(a_j)$, $\overline{\mathsf{MP}}_l(\rho) = \limsup_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} w_l(a_j)$,
- $\mathsf{DS}_l^{\lambda_l}(\rho) = \sum_{j=1}^{\infty} \lambda_l^j \cdot w_l(a_j)$, with $\lambda_l \in]0,1[\cap \mathbb{Q}$ a rational discount factor,
- $\mathsf{TS}_l^T(\rho) = \sum_{j=1}^{n-1} w_l(a_j)$ with s_n the first visit of a state in $T \subseteq S$. If T is never reached, then we assign $\mathsf{TS}_l^T(\rho) = \infty$.

For any payoff function f, $f_l \geq v$ defines the runs ρ that satisfy $f_l(\rho) \geq v$. A *percentile constraint* is of the form $\mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}[f_l \geq v] \geq \alpha$, where σ is to be synthesized given threshold value v and probability α. We study *multi-constraint percentile queries* requiring to simultaneously satisfy q constraints each referring to a possibly different dimension. Formally, given a d-dimensional weighted MDP M, initial state $s_{\mathsf{init}} \in S$, payoff function f, dimensions $l_1, \ldots, l_q \in \{1, \ldots, d\}$, value thresholds $v_1, \ldots, v_q \in \mathbb{Q}$ and probability thresholds $\alpha_1, \ldots, \alpha_q \in [0,1] \cap \mathbb{Q}$ the *multi-constraint percentile problem* asks if there exists a strategy $\sigma \in \Sigma$ such that query $\mathcal{Q} := \bigwedge_{i=1}^{q} \mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}\big[f_{l_i} \geq v_i\big] \geq \alpha_i$ holds. We can actually solve queries $\exists?\,\sigma,\ \bigvee_{i=1}^{m} \bigwedge_{j=1}^{n_i} \mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}\big[f_{l_{i,j}} \geq v_{i,j}\big] \geq \alpha_{i,j}$. We present our results for conjunctions of constraints only since the latter is equivalent to verifying the disjuncts independently: in other terms, to $\bigvee_{i=1}^{m} \exists\sigma \bigwedge_{j=1}^{n_i} \mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}\big[f_{l_{i,j}} \geq v_{i,j}\big] \geq \alpha_{i,j}$.

We distinguish *single-dimensional percentile problems* ($d = 1$) from *multi-dimensional* ones ($d > 1$). We assume w.l.o.g. that $q \geq d$ otherwise one can simply neglect unused dimensions. sFor some cases, we will consider the ε*-relaxation* of the problem, which consists in ensuring each value $v_i - \varepsilon$ with probability α_i.

We assume binary encoding of constants, and define the *model size* $|M|$ as the size of the representation of M, and the *query size* $|\mathcal{Q}|$ that of the query. *Problem size* refers to the sum of the two. We study memory needs for strategies w.r.t. different classes of queries; but randomization is always necessary as shown in the next lemma.

Lemma 1. *Randomized strategies are necessary for multi-dimensional percentile queries for any payoff function.*

3 Multiple Reachability and Contraction of MECs

Multiple reachability. An algorithm to solve this problem was given in [16] based on a linear program (LP) of size polynomial in the model and exponential in the query; whereas restricting the target sets to absorbing states yields a polynomial-size LP. We will use this LP later in Fig. 1 in Sect. 5.

Theorem 1 [16]**.** *Memoryless strategies suffice for multiple reachability with absorbing target states, and can be computed in polynomial time. With arbitrary targets, exponential-memory strategies (in query size) can be computed in time polynomial in the model and exponential in the query.*

In this section, we improve over this result by showing that the case of almost-sure multiple reachability is PSPACE-complete, with a recursive algorithm and a reduction from QBF satisfiability. This also shows the PSPACE-hardness of the general problem. Moreover, we show that exponential memory is required for strategies, following a construction of [13].

Theorem 2. *The almost-sure multiple reachability problem is* PSPACE*-complete, and strategies need exponential memory in the query size.*

Despite the above lower bounds, it turns out that the polynomial time algorithm for the case of absorbing targets can be extended; we identify a subclass of the multiple reachability problem that admits a polynomial-time solution. In the *nested multiple reachability* problem, the target sets are nested, *i.e.* $T_1 \subseteq T_2 \subseteq \ldots \subseteq T_q$. The memory requirement for strategies is reduced as well to linear memory. Intuitively, we use $q+1$ copies of the original MDP, one for each target set, plus one last copy. The idea is then to travel between those copies in a way that reflects the nesting of target sets whenever a target state is visited. The crux to obtain a polynomial-time algorithm is then to reduce the problem to a multiple reachability problem *with absorbing states* over the MDP composed of the $q+1$ copies, and to benefit from the reduced complexity of this case.

Theorem 3. *The nested multiple reachability problem can be solved in polynomial time. Strategies have memory linear in the query size, which is optimal.*

Contraction of MECs. In order to solve percentile queries, we sometimes reduce our problems to multiple reachability by first contracting MECs of given MDPs, which is a known technique [14]. We define a transformation of MDP M to represent the events $\mathsf{Inf}(\rho) \subseteq C$ for $C \in \mathsf{MEC}(M)$ as fresh states. Intuitively, all states of a MEC will now lead to an absorbing state that will abstract the behavior of the MEC.

Consider M with $\mathsf{MEC}(M) = \{C_1, \ldots, C_m\}$. We define MDP M' from M as follows. For each C_i, we add state s_{C_i} and action a^* from each state $s \in C_i$ to s_{C_i}. All states s_{C_i} are absorbing, and $A(s_{C_i}) = \{a^*\}$. The probabilities of events $\mathsf{Inf}(\rho) \subseteq C_i$ in M are captured by the reachability of states s_{C_i} in M', as follows. We use the classical temporal logic symbols $\Diamond$ and $\Box$ to represent the *eventually* and *always* operators respectively.

Lemma 2. *Let M be an MDP and $\mathsf{MEC}(M) = \{C_1, \ldots, C_m\}$. For any strategy σ for M, there exists a strategy τ for M' such that for all $i \in \{1, \ldots, m\}$, $\mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}[\Diamond\Box C_i] = \mathbb{P}^{\tau}_{M',s_{\mathsf{init}}}[\Diamond s_{C_i}]$. Conversely, for any strategy τ for M' such that $\sum_{i=1}^{m} \mathbb{P}^{\tau}_{M',s_{\mathsf{init}}}[\Diamond s_{C_i}] = 1$, there exists σ such that for all i, $\mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}[\Diamond\Box C_i] = \mathbb{P}^{\tau}_{M',s}[\Diamond s_{C_i}]$.*

Under some hypotheses, solving multi-constraint percentile problems on ECs yield the result for all MDPs, by the transformation of Lemma 2. We prove a general theorem and then derive particular results as corollaries. Informally, for prefix-independent payoff functions, if for any EC, there is a strategy that is optimal in each dimension, and if optimal values are computable in polynomial time, then the percentile problem can be solved in polynomial time.

Theorem 4. *Consider all prefix-independent payoff functions f such that for all strongly connected MDPs M, and all $(l_i, v_i)_{1 \le i \le q} \in \{1, \ldots, d\} \times \mathbb{Q}$, there exists a strategy σ such that $\forall i \in \{1, \ldots, d\}, \mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}[f_{l_i} \ge v_i] \ge \sup_{\tau} \mathbb{P}^{\tau}_{M,s_{\mathsf{init}}}[f_{l_i} \ge v_i]$. If the value $\sup_{\tau}$ is computable in polynomial time for strongly connected MDPs, then the multi-constraint percentile problem for f is decidable in polynomial time. Moreover, if strategies achieving $\sup_{\tau}$ for strongly connected MDPs use $\mathcal{O}(g(M,q))$ memory, then the overall strategy use $\mathcal{O}(g(M,q))$ memory.*

The hypotheses are crucial. Essentially, we require payoff functions that are prefix-independent and for which strategies can be combined easily inside MECs (in the sense that if two constraints can be satisfied independently, they can be satisfied simultaneously). Prefix-independence also implies that we can forget about what happens before a MEC is reached. Hence, by using the MEC contraction, we can reduce the percentile problem to multiple reachability for absorbing target states.

4 Inf, Sup, LimInf, LimSup Payoff Functions

We give polynomial-time algorithms for the *single*-dimensional multi-constraint percentile problems. For inf, sup we reduce the problem to nested multiple reachability, while lim inf and lim sup are solved by applying Theorem 4.

Theorem 5. *The single-dimensional multi-constraint percentile problems can be solved in polynomial time in the problem size for* inf, sup, lim inf, *and* lim sup *functions. Computed strategies use memory linear in the query size for* inf *and* sup, *and constant memory for* lim inf *and* lim sup.

We are now interested in the multi-dimensional case. We show that all multi-dimensional cases can be solved in time polynomial in the model size and exponential in the query size by a reduction to multiple LTL objectives studied in [16]. Our algorithm actually solves a more general class of queries, where the payoff function can be different for each query.

Given an MDP M, for all $i \in \{1 \ldots q\}$ and value v_i, we denote $A^{\ge v_i}_{l_i}$ the set of actions of M whose rewards are at least v_i. We fix an MDP M. For any constraint $\phi_i \equiv f(w_{l_i}) \ge v_i$, we define an LTL formula denoted Φ_i as follows. For $f_{l_i} = \inf$, $\Phi_i = \Box A^{\ge v_i}_{l_i}$, for $f_{l_i} = \sup$, $\Phi_i = \Diamond A^{\ge v_i}_{l_i}$, for $f_{l_i} = \liminf$, $\Phi_i = \Diamond\Box A^{\ge v_i}_{l_i}$, and for $f_{l_i} = \limsup$, $\Phi_i = \Box\Diamond A^{\ge v_i}_{l_i}$. The percentile problem is then reduced to queries of the form $\wedge_{i=1}^{q} \mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}[\Phi_i] \ge \alpha_i$, for which an algorithm was given in [16] that takes time polynomial in $|M|$ and doubly exponential in q. We improve this complexity since our formulae have bounded sizes.

Theorem 6. *The multi-dimensional percentile problems for* sup, inf, lim sup *and* lim inf *can be solved in time polynomial in the model size and exponential in the query size, yielding strategies with memory exponential in the query.*

The problem is PSPACE-hard for sup as shown in the following theorem.

Theorem 7. *The multi-dimensional percentile problem is* PSPACE*-hard for* sup.

Nevertheless, the complexity can be improved for lim sup functions, for which we give a polynomial-time algorithm by an application of Theorem 4.

Theorem 8. *The multi-dimensional percentile problem for* lim sup *is solvable in polynomial time. Computed strategies use constant-memory.*

The exact query complexity of the lim inf and inf cases are left open.

5 Mean-Payoff

We consider the multi-constraint percentile problem both for $\underline{\mathsf{MP}}$ and $\overline{\mathsf{MP}}$. We will see that strategies require infinite memory in both cases, in which case it is known that the two payoff functions differ. The *single-constraint* percentile problem was first solved in [17]. The case of multiple dimensions was mentioned as a challenging problem but left open. We solve this problem thus generalizing the previous work.

The Single-Dimensional Case. We start with a polynomial-time algorithm for the single-dimensional case obtained by an application of Theorem 4.

Theorem 9. *The single dimensional multi-constraint percentile problems for payoffs* $\underline{\mathsf{MP}}$ *and* $\overline{\mathsf{MP}}$ *are equivalent and solvable in polynomial time. Computed strategies use constant memory.*

Percentiles on Multi-dimensional $\overline{\mathsf{MP}}$. Let $\mathbb{E}^{\sigma}_{M,s_{\mathsf{init}}}[\overline{\mathsf{MP}}_i]$ be the *expectation* of $\overline{\mathsf{MP}}_i$ under strategy σ, and $\mathsf{Val}^*_{M,s_{\mathsf{init}}}(\overline{\mathsf{MP}}_i) = \sup_\sigma \mathbb{E}^{\sigma}_{M,s_{\mathsf{init}}}[\overline{\mathsf{MP}}_i]$, computable in polynomial time [23]. We solve the problem inside ECs, then apply Theorem 4. It is known that for strongly connected MDPs, for each i, some strategy σ satisfies $\mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}[\overline{\mathsf{MP}}_i = \mathsf{Val}^*_{M,s_{\mathsf{init}}}(\overline{\mathsf{MP}}_i)] = 1$, and that for all strategies τ, $\mathbb{P}^{\tau}_{M,s_{\mathsf{init}}}[\overline{\mathsf{MP}}_i > v] = 0$ for all $v > \mathsf{Val}^*_{M,s_{\mathsf{init}}}(\overline{\mathsf{MP}}_i)$. By switching between these optimal strategies for each dimension, with growing intervals, we prove that for strongly connected MDPs, a single strategy can simultaneously optimize $\overline{\mathsf{MP}}_i$ on *all* dimensions.

Lemma 3. *For any strongly connected MDP M, there is an infinite-memory strategy σ such that $\forall i \in \{1, \ldots, d\}$, $\mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}[\overline{\mathsf{MP}}_i \geq \mathsf{Val}^*_{M,s_{\mathsf{init}}}(\overline{\mathsf{MP}}_i)] = 1$.*

Thanks to the above lemma, we fulfill the hypotheses of Theorem 4, and we obtain the following theorem.

Theorem 10. *The multi-dimensional percentile problem for* $\overline{\mathsf{MP}}$ *is solvable in polynomial time. Strategies use infinite-memory, which is necessary.*

Percentiles on Multi-dimensional $\underline{\mathsf{MP}}$. In contrast with the $\overline{\mathsf{MP}}$ case, our algorithm for $\underline{\mathsf{MP}}$ is more involved, and requires new techniques. In fact, the case of end-components is already non-trivial for $\underline{\mathsf{MP}}$, since there is no single strategy that satisfies all percentile constraints in general, and one cannot hope to apply Theorem 4 as we did in previous sections. We rather need to consider the set of strategies σ_I satisfying *maximal* subsets of percentile constraints; these are called *maximal strategies*. We then prove that any strategy satisfying all percentile queries can be written as a *linear combination* of maximal strategies, that is, there exists a strategy which chooses and executes each σ_I following a probability distribution.

For general MDPs, we first consider each MEC separately and write down the linear combination with unknown coefficients. We know that any strategy in a MDP eventually stays forever in a MEC. Thus, we adapt the linear program of [16] that encodes the reachability probabilities with multiple targets, which are the MECs here. We combine these reachability probabilities with the unknown linear combination coefficients, and obtain a linear program (Fig. 1), which we prove to be equivalent to our problem.

Single EC. Fix a strongly connected d-dimensional MDP M and pairs of thresholds $(v_i, \alpha_i)_{1 \leq i \leq q}$. We denote each event by $A_i \equiv \underline{\mathsf{MP}}_i \geq v_i$. In [6], the problem of maximizing the *joint* probability of the events A_i was solved in polynomial time. In particular, we have the following for strongly connected MDPs.

Lemma 4 [6]**.** *If M is strongly connected, then there exists σ such that $\mathbb{P}^{\sigma}_{M,s}[\wedge_{1 \leq i \leq q} A_i] > 0$ if, and only if there exists σ' such that $\mathbb{P}^{\sigma'}_{M,s}[\wedge_{1 \leq i \leq q} A_i] = 1$. Moreover, this can be decided in polynomial time, and for positive instances, for any $\varepsilon > 0$, a memoryless strategy τ can be computed in polynonomial time in M, $\log(v_i)$ and $\log(\frac{1}{\varepsilon})$, such that $\mathbb{P}^{\tau}_{M,s}[\wedge_{1 \leq i \leq q} \underline{\mathsf{MP}}_i \geq v_i - \varepsilon] = 1$.*

We give an overview of our algorithm. Using Lemma 4, we define strategy σ_I achieving $\mathbb{P}^{\sigma_I}_{M,s}[\wedge_{i \in I} A_i] = 1$ for any maximal subset $I \subseteq \{1, \ldots, q\}$ for which such a strategy exists. Then, to build a strategy for the multi-constraint problem, we look for a linear combination of these σ_I: given $\sigma_{I_1}, \ldots, \sigma_{I_m}$, we choose each $i_0 \in \{1, \ldots, m\}$ following a probability distribution to be computed, and we run $\sigma_{I_{i_0}}$.

We now formalize this idea. Let $\mathcal{I}$ be the set of maximal I (for set inclusion) such that some σ_I satisfies $\mathbb{P}^{\sigma_I}_{M,s}[\wedge_{i \in I} A_i] = 1$. Note that for all $I \in \mathcal{I}$, and $j \notin I$, $\mathbb{P}^{\sigma_I}_{M,s}[\wedge_{i \in I} A_i \wedge A_j] = 0$. Assuming otherwise would contradict the maximality of I, by Lemma 4. We consider the events $\mathcal{A}_I = \wedge_{i \in I} A_i \wedge_{i \notin I} \neg A_i$ for maximal I.

We are looking for a non-negative family $(\lambda_I)_{I \in \mathcal{I}}$ whose sum equals 1 with $\forall i \in \{1, \ldots, q\}, \sum_{I \in \mathcal{I} \text{ s.t. } i \in I} \lambda_I \geq \alpha_i$. This will ensure that if each σ_I is chosen with probability λ_I (among the set $\{\sigma_I\}_{I \in \mathcal{I}}$); with probability at least α_i, some strategy satisfying A_i with probability 1 is chosen. So each A_i is satisfied with probability at least α_i. This can be written in the matrix notation as

$$\mathcal{M}\boldsymbol{\lambda} \geq \boldsymbol{\alpha}, 0 \leq \boldsymbol{\lambda}, \mathbf{1} \cdot \boldsymbol{\lambda} = 1, \quad (1)$$

where $\mathcal{M}$ is a $q \times |\mathcal{I}|$ matrix with $\mathcal{M}_{i,I} = 1$ if $i \in I$, and 0 otherwise.

Lemma 5. *For any strongly connected MDP M, and an instance $(v_i, \alpha_i)_{1 \leq i \leq q}$ of the multi-constraint percentile problem for $\underline{\mathsf{MP}}$, (1) has a solution if, and only if there exists a strategy σ satisfying the multi-constraint percentile problem.*

$$\mathbf{1}_{s_{\mathsf{init}}}(s) + \sum_{s' \in S, a \in A(s')} y_{s',a}\delta(s', a, s) = \sum_{a \in A'(s)} y_{s,a}, \qquad \forall s \in S, \tag{2}$$

$$\sum_{s \in S_{\mathrm{MEC}}} y_{s,a^*} = 1, \tag{3}$$

$$\sum_{s \in C} y_{s,a^*} = \sum_{I \in \mathcal{I}^C} \lambda_I^C, \quad \forall C \in \mathsf{MEC}(M), \tag{4}$$

$$\lambda_I^C \geq 0, \quad \forall C \in \mathsf{MEC}(M), \forall I \in \mathcal{I}^C, \tag{5}$$

$$\sum_{C \in \mathsf{MEC}(M)} \sum_{I \in \mathcal{I}^C : i \in I} \lambda_I^C \geq \alpha_i, \qquad \forall i = 1 \ldots d. \tag{6}$$

Fig. 1. Linear program (L) for the multi-constraint percentiles for $\underline{\mathsf{MP}}$.

Now (1) has size $O(q \cdot 2^q)$, and each subset I can be checked in time polynomial in the model size. The computation of $\mathcal{I}$, the set of maximal subsets, can be carried out in a top-down fashion; one might thus avoid enumerating all subsets in practice. We get the following result.

Lemma 6. *For strongly connected MDPs, the multi-dimensional percentile problem for* $\underline{\mathsf{MP}}$ *can be solved in time polynomial in* M *and exponential in* q*. Strategies require infinite-memory in general. On positive instances,* 2^q*-memory randomized strategies can be computed for the* ε*-relaxation of the problem in time polynomial in* $|M|, 2^q, \max_i \big(\log(v_i), \log(\alpha_i)\big), \log(\frac{1}{\varepsilon})$.

General MDPs. Given MDP M, let us consider M' given by Lemma 2. We start by analyzing each maximal EC C of M as above, and compute the sets $\mathcal{I}^C$ of maximal subsets. We define a variable λ_I^C for each $I \in \mathcal{I}^C$, and also $y_{s,a}$ for each state s and action $a \in A'(s)$. Recall that $A'(s) = A(s) \cup \{a^*\}$ for states s that are inside a MEC, and $A'(s) = A(s)$ otherwise. Let S_{MEC} be the set of states of M that belong to a MEC. We consider the linear program (L) of Fig. 1.

The linear program follows the ideas of [6,16]. Note that the first two lines of (L) corresponds to the multiple reachability LP of [16] for absorbing target states. The equations encode strategies that work in two phases. Variables $y_{s,a}$ correspond to the expected number of visits of state-action s, a in the first phase. Variable y_{s,a^*} describes the probability of switching to the second phase at state s. The second phase consists in surely staying in the current MEC, so we require $\sum_{s \in S_{\mathrm{MEC}}} y_{s,a^*} = 1$ (and we will have $y_{s,a^*} = 0$ if s does not belong to a MEC). In the second phase, we immediately switch to some strategy σ_I^C where C denotes the current MEC. Thus, variable λ_I^C corresponds to the probability with which we enter the second phase in C and switch to strategy σ_I^C (see (4)). Intuitively, given a solution $(\lambda_I)_I$ computed for one EC by (1), we have the correspondence $\lambda_I^C = \sum_{s \in C} y_{s,a^*} \cdot \lambda_I$. The interpretation of (6) is that each event A_i is satisfied with probability at least α_i.

Lemma 7. *The LP (L) has a solution if, and only if the multi-constraint percentiles problem for* $\underline{\mathsf{MP}}$ *has a solution. Moreover, the equation has size polynomial in* M *and exponential in* q*. From any solution of (L) randomized finite memory strategies can be computed for the* ε*-relaxation problem.*

Theorem 11. *The multi-dimensional percentile problem for* $\underline{\mathsf{MP}}$ *can be solved in time polynomial in the model, and exponential in the query. Infinite-memory strategies are necessary, but exponential-memory (in the query) suffices for the* ε*-relaxation and can be computed with the same complexity.*

6 Shortest Path

We study shortest path problems in MDPs, which generalize the classical graph problem. In MDPs, the problem consists in finding a strategy ensuring that a target set is reached with bounded truncated sum with high probability. This problem has been studied in the context of games and MDPs (e.g., [2,7,15]). We consider percentile queries of the form $\mathcal{Q} := \bigwedge_{i=1}^{q} \mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}\big[\mathsf{TS}^{T_i}_{l_i} \leq v_i\big] \geq \alpha_i$ (inner inequality $\leq$ is more natural but $\geq$ could be used by negating all weights). Each constraint i may relate to a different target set $T_i \subseteq S$.

Arbitrary Weights. We prove that without further restriction, the multi-dimensional percentile problem is undecidable, even for a fixed number of dimensions. Our proof is inspired by the approach of Chatterjee et al. for the undecidability of two-player multi-dimensional total-payoff games [8] but requires additional techniques to adapt to the stochastic case.

Theorem 12. *The multi-dimensional percentile problem is undecidable for the truncated sum payoff function, for MDPs with both negative and positive weights and four dimensions, even with a unique target set.*

Non-negative Weights. In the light of this result, we will restrict our setting to non-negative weights (we could equivalently consider non-positive weights with inequality $\geq$ inside percentile constraints). We first discuss recent related work.

Quantiles and Cost Problems. In [29], Ummels and Baier study *quantile queries* over non-negatively weighted MDPs. They are equivalent to minimizing $v \in \mathbb{N}$ in a single-constraint percentile query $\mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}\big[\mathsf{TS}^{T} \leq v\big] \geq \alpha$ such that there still exists a satisfying strategy, for some fixed α. Very recently, Haase and Kiefer extended quantile queries by introducing *cost problems* [21]. They can be seen as single-constraint percentile queries where inequality $\mathsf{TS}^{T} \leq v$ is replaced by an arbitrary Boolean combination of inequalities φ. Hence, it can be written as $\mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}\big[\mathsf{TS}^{T} \models \varphi\big] \geq \alpha$. Cost problems are studied on single-dimensional MDPs and all the inequalities relate to the same target T, in contrast to our setting which allows both for multiple dimensions and multiple target sets. The single probability threshold bounds the probability of the whole event φ.

Both settings are incomparable. Still, our queries share common subclasses with cost problems: atomic formulae φ exactly correspond to our single-constraint queries. Moreover, cost problems for such formulae are inter-reducible with quantile queries [21, Proposition 2]. Cost problems with atomic formulae are PSPACE-hard, so this also holds for *single-constraint* percentile queries. The best known algorithm in this case is in EXPTIME. In the following, we establish an algorithm that still only requires exponential time while allowing for *multi-constraint multi-dimensional multi-target* percentile queries.

Main Results. Our main contributions for the shortest path are as follows.

Theorem 13. *The percentile problem for the shortest path with non-negative weights can be solved in time polynomial in the model size and exponential in the query size (exponential in the number of constraints and pseudo-polynomial in the largest threshold). The problem is* PSPACE*-hard even for single-constraint queries. Exponential-memory strategies are sufficient and in general necessary.*

Sketch of Algorithm. Consider a d-dimensional MDP M and a q-query percentile problem, with potentially different targets for each query. Let v_{max} be the maximum of the thresholds v_i. Because weights are non-negative, extending a finite history never decreases the sum of its weights. Thus, any history ending with a sum exceeding v_{max} in all dimensions is surely losing under any strategy.

Based on this, we build an MDP M' by unfolding M and integrating the sum for each dimension in states of M'. We ensure its finiteness thanks to the above observation and we reduce its overall size to a *single*-exponential by defining a suitable equivalence relation between states of M': we only care about the current sum in each dimension, and we can forget about the actual path that led to it. Precisely, the states of M' are in $S \times \{0, \dots, v_{\mathsf{max}} + 1\}^d$. Now, for each constraint, we compute a set of target states in M' that exactly captures all runs satisfying the inequality of the constraint. Thus, we are left with a multiple reachability problem on M': we look for a strategy σ' that ensures that each of these sets R_i is reached with probability α_i. This query can be answered in time polynomial in $|M'|$ but exponential in the number of sets R_i, i.e., in q (Theorem 1).

Remark 1. Percentile problems with unique target are solvable in time polynomial in the number of constraints but still exponential in the number of dimensions.

For single-dimensional queries with a unique target set (but still potentially multi-constraint), our algorithm remains pseudo-polynomial as it requires polynomial time in the thresholds values (i.e., exponential in their encoding).

Corollary 1. *The single-dimensional percentile problem with a unique target set can be solved in pseudo-polynomial time.*

Lower Bound. By equivalence with cost problems for atomic cost formulae, it follows from [21, Theorem 7] that no truly-polynomial-time algorithm exists for the single-constraint percentile problem unless P = PSPACE.

Memory. The upper bound is by reduction to multiple reachability over an exponential unfolding. The lower bound is via reduction from multiple reachability.

7 Discounted Sum

The *discounted sum* models that short-term rewards or costs are more important than long-term ones. It is well-studied in automata [3] and MDPs [9,12,23]. We consider queries of the form $\mathcal{Q} \coloneqq \bigwedge_{i=1}^{q} \mathbb{P}^{\sigma}_{M,s_{\mathsf{init}}}\big[\mathsf{DS}^{\lambda_i}_{l_i} \geq v_i\big] \geq \alpha_i$, for discount factors $\lambda_i \in \,]0,1[\, \cap \mathbb{Q}$ and the usual thresholds. That is, we study multi-dimensional MDPs and possibly distinct discount factors for each constraint.

Our setting encompasses a simpler question which is still not known to be decidable. Consider the *precise discounted sum problem*: given a rational t, and a rational discount factor $\lambda \in \,]0,1[$, does there exist an infinite binary sequence $\tau = \tau_1\tau_2\tau_3\ldots \in \{0,1\}^{\omega}$ such that $\sum_{j=1}^{\infty} \lambda^j \cdot \tau_j = t$? In [4], this problem is related to several long-standing open questions, such as decidability of the *universality problem for discounted-sum automata* [3]. A slight generalization to paths in graphs is also mentioned by Chatterjee et al. as a key open problem in [9].

Lemma 8. *The precise discounted sum problem can be reduced to an almost-sure percentile problem over a two-dimensional MDP with only one state.*

This suggests that answering percentile problems would require an important breakthrough. In the following, we establish a conservative algorithm that, in some sense, can approximate the answer.

The ε-gap Problem. Our algorithm takes as input a percentile query and an arbitrarily small *precision factor* $\varepsilon > 0$ and has three possible outputs: Yes, No and Unknown. If it answers Yes, then a satisfying strategy exists and can be synthesized. If it answers No, then no such strategy exists. Finally, the algorithm may output Unknown for a specified "zone" close to the threshold values involved in the problem and of width which depends on ε. It is possible to incrementally reduce the uncertainty zone, but it cannot be eliminated as the case $\varepsilon = 0$ would answer the precise discounted sum problem, which is not known to be decidable.

We actually solve an *ε-gap problem*, a particular case of *promise problems* [19], where the set of inputs is partitioned in three subsets: yes-inputs, no-inputs and the rest of them. The promise problem then asks to answer Yes for all yes-inputs and No for all no-inputs, while the answer may be arbitrary for the remaining inputs. In our setting, the set of inputs for which no guarantee is given can be taken arbitrarily small, parametrized by value $\varepsilon > 0$: this is an ε-gap problem. This notion is formalized in Theorem 15.

Related Work: Single-Constraint Case. There are papers considering models related to *single-constraint* percentile queries. Consider a single-dimensional MDP and a single-constraint query, with thresholds v and α. The *threshold problem* fixes v and maximizes α [30,31]. The *value-at-risk problem* fixes α and maximizes v [5]. This is similar to *quantiles* in the shortest path setting [29]. Paper [5] is the first to provide an exponential-time algorithm to approximate the optimal value v^* under a fixed α in the general setting. The authors also rely on approximation. While we do not consider optimization, we do extend the setting to *multi-constraint*, *multi-dimensional*, *multi-discount* problems, and we are able to remain in the same complexity class, namely EXPTIME.

Main Results. Our main contributions for the discounted sum are as follows.

Theorem 14. *The ε-gap percentile problem for the discounted sum can be solved in time pseudo-polynomial in the model size and the precision factor, and exponential in the query size: polynomial in the number of states, the weights, the discount factors and the precision factor, and exponential in the number of constraints. It is* PSPACE*-hard for two-dimensional MDPs and already* NP*-hard for single-constraint queries. Exponential-memory strategies are both sufficient and in general necessary to satisfy ε-gap percentile queries.*

Cornerstones of the Algorithm. Our approach is similar to the shortest path: we want to build an unfolding capturing the needed information w.r.t. the discounted sums, and then reduce the percentile problem to a multiple reachability problem over this unfolding. However, several challenges have to be overcome.

First, we need a *finite* unfolding. This was easy in the shortest path due to non-decreasing sums and corresponding upper bounds. Here, it is not the case as we put no restriction on weights. Nonetheless, thanks to the discount factor, weights contribute less and less to the sum along a run. In particular, cutting all runs after a pseudo-polynomial length changes the overall sum by at most $\varepsilon/2$.

Second, we reduce the overall size of the unfolding. For the shortest path we took advantage of integer labels to define equivalence. Here, the space of values taken by the discounted sums is too large for a straightforward equivalence. To reduce it, we introduce a *rounding* scheme of the numbers involved. This idea is inspired by [5]. We bound the error due to cumulated roundings by $\varepsilon/2$.

So, we control the amount of information lost to guarantee exact answers except inside an arbitrarily small ε-zone. Given a q-constraint query $\mathcal{Q}$ for thresholds v_i, α_i, dimensions l_i and discounts λ_i, we define the *x-shifted query $\mathcal{Q}_x$*, for $x \in \mathbb{Q}$, as the exact same problem for thresholds $v_i + x$, α_i, dimensions l_i and discounts λ_i. Our algorithm satisfies the following theorem, which formalizes the ε-gap percentile problem mentioned in Theorem 14.

Theorem 15. *There is an algorithm that, given an MDP, a percentile query $\mathcal{Q}$ for the discounted sum and a precision factor $\varepsilon > 0$, solves the following ε-gap problem in exponential time. It answers*

- *Yes if there is a strategy satisfying the $(2 \cdot \varepsilon)$-shifted percentile query $\mathcal{Q}_{2\cdot\varepsilon}$;*
- *No if there is no strategy satisfying the $(-2 \cdot \varepsilon)$-shifted percentile query $\mathcal{Q}_{-2\cdot\varepsilon}$;*
- *and arbitrarily otherwise.*

Lower Bounds. The ε-gap percentile problem is PSPACE-hard by reduction from subset-sum games [28]. Two tricks are important. First, counterbalancing the discount effect via adequate weights. Second, simulating an equality constraint. This cannot be achieved directly because it requires to handle $\varepsilon = 0$. Still, by choosing weights carefully we restrict possible discounted sums to integer values only. Then we choose the thresholds and $\varepsilon > 0$ such that no run can take a value within the uncertainty zone. This circumvents the limitation due to uncertainty. For single-constraint ε-gap problems, we prove NP-hardness, even for Markov chains. Our proof is by reduction from the K-th largest subset problem [18], inspired by [7, Theorem 11]. A recent, not yet published, paper by Haase and Kiefer [20] claims that this K-th largest subset problem is actually PP-complete. If this claim holds, then it suggests that the single-constraint problem does not belong to NP at all, otherwise the polynomial hierarchy would collapse to P^{NP} by Toda's theorem [27].

Memory. For the precise discounted sum and generalizations, infinite memory is needed [9]. For ε-gap problems, the exponential upper bound follows from the algorithm while the lower bound is shown via a family of problems that emulate the ones used for multiple reachability (Theorem 2).

References

1. Baier, C., Daum, M., Dubslaff, C., Klein, J., Klüppelholz, S.: Energy-utility quantiles. In: Badger, J.M., Rozier, K.Y. (eds.) NFM 2014. LNCS, vol. 8430, pp. 285–299. Springer, Heidelberg (2014)
2. Bertsekas, D.P., Tsitsiklis, J.N.: An analysis of stochastic shortest path problems. Math. Oper. Res. **16**, 580–595 (1991)
3. Boker, U., Henzinger, T.A.: Exact and approximate determinization of discounted-sum automata. LMCS **10**(1), 1–33 (2014)
4. Boker, U., Henzinger, T.A., Otop, J.: The target discounted-sum problem. In: Proceedings of LICS. IEEE Computer Society (2015)
5. Brázdil, T., Chen, T., Forejt, V., Novotný, P., Simaitis, A.: Solvency Markov decision processes with interest. In: Proceedings of FSTTCS, LIPIcs, vol. 24, pp. 487–499. Schloss Dagstuhl - LZI (2013)
6. Brázdil, T., Brozek, V., Chatterjee, K., Forejt, V., Kucera, A.: Markov decision processes with multiple long-run average objectives. LMCS **10**(13), 1–29 (2014)
7. Bruyère, V., Filiot, E., Randour, M., Raskin, J.-F.: Meet your expectations with guarantees: beyond worst-case synthesis in quantitative games. In: Proceedings of STACS, LIPIcs, vol. 25, pp. 199–213. Schloss Dagstuhl - LZI (2014)
8. Chatterjee, K., Doyen, L., Randour, M., Raskin, J.-F.: Looking at mean-payoff and total-payoff through windows. In: Van Hung, D., Ogawa, M. (eds.) ATVA 2013. LNCS, vol. 8172, pp. 118–132. Springer, Heidelberg (2013)
9. Chatterjee, K., Forejt, V., Wojtczak, D.: Multi-objective discounted reward verification in graphs and MDPs. In: McMillan, K., Middeldorp, A., Voronkov, A. (eds.) LPAR-19 2013. LNCS, vol. 8312, pp. 228–242. Springer, Heidelberg (2013)
10. Chatterjee, K., Henzinger, T.A.: Probabilistic systems with limsup and liminf objectives. In: Archibald, M., Brattka, V., Goranko, V., Löwe, B. (eds.) ILC 2007. LNCS, vol. 5489, pp. 32–45. Springer, Heidelberg (2009)
11. Chatterjee, K., Komárková, Z., Kretínský, J.: Unifying two views on multiple mean-payoff objectives in Markov decision processes. In: Proceedings of LICS. IEEE Computer Society (2015)
12. Chatterjee, K., Majumdar, R., Henzinger, T.A.: Markov decision processes with multiple objectives. In: Durand, B., Thomas, W. (eds.) STACS 2006. LNCS, vol. 3884, pp. 325–336. Springer, Heidelberg (2006)
13. Chatterjee, K., Randour, M., Raskin, J.-F.: Strategy synthesis for multi-dimensional quantitative objectives. Acta Inform. **51**(3–4), 129–163 (2014)
14. de Alfaro, L.: Formal verification of probabilistic systems. Ph.D. thesis, Stanford University (1997)

15. de Alfaro, L.: Computing minimum and maximum reachability times in probabilistic systems. In: Baeten, J.C.M., Mauw, S. (eds.) CONCUR 1999. LNCS, vol. 1664, pp. 66–81. Springer, Heidelberg (1999)
16. Etessami, K., Kwiatkowska, M.Z., Vardi, M.Y., Yannakakis, M.: Multi-objective model checking of Markov decision processes. LMCS **4**(4), 1–21 (2008)
17. Filar, J.A., Krass, D., Ross, K.W.: Percentile performance criteria for limiting average Markov decision processes. IEEE Trans. Aut. Control **40**(1), 2–10 (1995)
18. Garey, Michael R., Johnson, David S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, New York (1979)
19. Goldreich, O.: On promise problems: a survey. In: Goldreich, O., Rosenberg, A.L., Selman, A.L. (eds.) Theoretical Computer Science. LNCS, vol. 3895, pp. 254–290. Springer, Heidelberg (2006)
20. Haase, C., Kiefer, S.: The complexity of the Kth largest subset problem and related problems. CoRR, abs/1501.06729 (2015)
21. Haase, C., Kiefer, S.: The odds of staying on budget. In: Halldórsson, M.M., Iwama, K., Kobayashi, N., Speckmann, B. (eds.) ICALP 2015. LNCS, vol. 9135, pp. 234–246. Springer, Heidelberg (2015)
22. Ohtsubo, Y.: Optimal threshold probability in undiscounted Markov decision processes with a target set. Appl. Math. Comput. **149**(2), 519–532 (2004)
23. Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming, 1st edn. Wiley, New York (1994)
24. Randour, M., Raskin, J.-F., Sankur, O.: Percentile queries in multi-dimensional Markov decision processes. CoRR, abs/1410.4801 (2014)
25. Randour, M., Raskin, J.-F., Sankur, O.: Variations on the stochastic shortest path problem. In: D'Souza, D., Lal, A., Larsen, K.G. (eds.) VMCAI 2015. LNCS, vol. 8931, pp. 1–18. Springer, Heidelberg (2015)
26. Sakaguchi, M., Ohtsubo, Y.: Markov decision processes associated with two threshold probability criteria. J. Control Theor. Appl. **11**(4), 548–557 (2013)
27. Toda, S.: PP is as hard as the polynomial-time hierarchy. SIAM J. Comput. **20**(5), 865–877 (1991)
28. Travers, S.D.: The complexity of membership problems for circuits over sets of integers. Theor. Comput. Sci. **369**(1–3), 211–229 (2006)
29. Ummels, M., Baier, C.: Computing quantiles in Markov reward models. In: Pfenning, F. (ed.) FOSSACS 2013 (ETAPS 2013). LNCS, vol. 7794, pp. 353–368. Springer, Heidelberg (2013)
30. White, D.J.: Minimizing a threshold probability in discounted Markov decision processes. J. Math. Anal. Appl. **173**(2), 634–646 (1993)
31. Wu, C., Lin, Y.: Minimizing risk models in Markov decision processes with policies depending on target values. J. Math. Anal. Appl. **231**(1), 47–67 (1999)
32. Xu, H., Mannor, S.: Probabilistic goal Markov decision processes. In: IJCAI, pp. 2046–2052 (2011)