Joint Chance-Constrained Markov Decision Processes

Abdel Lisser, Vikas Vikram Singh, V Varagapriya

Joint Chance-Constrained Markov Decision Processes

V Varagapriya[1], Vikas Vikram Singh[1] and Abdel Lisser[2]

[1]Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, 110016, New Delhi, India.
[2]Universite Paris Saclay, CNRS, CentraleSupelec, Laboratoire des Signaux et Systemes, Bat Breguet, 3 Rue Joliot Curie, Gif-sur-Yvette, 91190, France.

Contributing authors: varagapriyav@gmail.com; vikassingh@iitd.ac.in; abdel.lisser@l2s.centralesupelec.fr;

Abstract

We consider a finite state-action uncertain constrained Markov decision process under discounted and average cost criteria. The running costs are defined by random variables and the transition probabilities are known. The uncertainties present in the objective function and the constraints are modelled using chance constraints. The dependence among the random constraint vectors is driven by a Gumbel-Hougaard copula. We propose two second order cone programming problems whose optimal values give upper and lower bounds of the optimal value of the uncertain constrained Markov decision process. As an application, we study a stochastic version of a service and admission control problem in a queueing system and illustrate the proposed approximation methods on randomly generated instances of different sizes.

Keywords: Constrained Markov decision process, Joint chance constraint, Second order cone programming problem, Copula, Elliptical distribution, Queueing problem.

1 Introduction

Markov decision process (MDP) is a decision making framework for a stochastic system which evolves over finite/infinite horizon. The system can have a finite or infinite number of states, and at each state, a controller chooses an action from a set of available actions at that state. The system moves from one state to another state according to a controlled Markov chain. At every stage, the system incurs cost whenever a state is visited and an action is taken. The objective is to find a policy which minimizes the expected cost incurred over a period of time. The discounted and average cost criteria are most commonly used in the literature. For a finite state-action MDP with known transition probabilities and running costs, there exists a stationary deterministic optimal policy and it can be obtained by using dynamic programming and Linear programming (LP) based methods [1, 4, 18]. In many real-life examples, there are more than one running costs incurred at every stage whenever the Markov chain visits a state and an action is taken [1–3]. One type of running cost defines an expected cost function which needs to be minimized and other types of running costs define the expected cost functions on which constraints are imposed. This class of MDP is called a constrained Markov decision process (CMDP). When running costs and transition probabilities are stationary and exactly known, there exists a stationary randomized optimal policy for a finite state-action CMDP problem and it can be obtained by solving an equivalent LP problem [1].

In practice, the MDP model parameters such as running costs and transition probabilities are not exactly known. They are either learnt with experience or from historical data, thereby leading to errors in the optimal policies [13]. The early literature on MDPs with uncertain parameters dates back to the 1970s, where MDPs under uncertain transition probabilities were studied [19, 22]. Previous studies have accounted for the uncertainties in MDPs under a robust optimization framework by considering the worst case realization of the uncertain parameters [11, 15, 23]. Recently, the equivalent reformulations of robust CMDP problem, with uncertain costs and known transition probabilities, for various uncertainty sets are proposed by Varagapriya et al. [21]. Delage and Mannor [8] observed that robust optimization problems can give highly conservative optimal solutions. As an alternative approach, they considered an uncertain MDP problem under discounted cost criterion where either running costs or transition probabilities are defined as random variables and proposed a chance constraint programming based formulation. The chance constraint based MDP is equivalent to a second order cone programming (SOCP) problem when the running cost vector follows a normal distribution and transition probabilities are known [8]. When the prior transition probabilities follow Dirichlet distribution and running costs are known, the uncertain MDP problem becomes intractable and optimal policies can be computed using approximation methods [8].

To the best of our knowledge, CMDPs where either running costs or transition probabilities are defined using random variables are not considered in the

literature. In this paper, we consider a CMDP problem under both discounted and average cost criteria where running cost vectors are random vectors and transition probabilities are known. We use chance constraint programming [5, 17] to model the uncertainties present in the objective function and the constraints of the CMDP problem. For a given policy, the chance constraint based objective function of the CMDP problem gives a minimum long run expected cost which is incurred with at least a given probability level. The random constraints of the CMDP problem is defined as a joint chance constraint which guarantees that the random constraints are jointly satisfied with at least a given probability level. We call this uncertain CMDP as a joint chance-constrained Markov decision process (JCCMDP). In general, the random constraint vectors present in the joint chance constraint can be dependent random vectors. In order to study the dependence structure of random variables, the concept of copula was introduced by Abe Sklar in 1959 [20]. We consider the case when running cost vectors follow multivariate elliptically symmetric distribution and dependence among random constraint vectors is driven by a Gumbel-Hougaard copula. The deterministic equivalent of the JCCMDP problem is a non-convex optimization problem. We approximate the latter by proposing two SOCP problems whose optimal values give upper and lower bounds of the optimal value of the JCCMDP problem. As an application, we study the problem of admission and service control in a queueing system and perform numerical experiments by considering various instances of different sizes.

The rest of the paper is organized as follows. Section 2 introduces the definition of a CMDP. Section 3 presents a JCCMDP problem and its SOCP based approximations. Section 4 gives the results of numerical experiments carried out on the queueing control problem. We conclude the paper in Section 5.

2 Constrained Markov decision processes

We define an infinite horizon CMDP by the following objects [1]:

(i) S is a finite set of states whose generic element is denoted by s.
(ii) A is a finite set of actions and $A(s)$ denotes the set of actions available at state s.
(iii) $\mathcal{K} = \{(s,a) \mid s \in S, a \in A(s)\}$ is the set of all state-action pairs; $|\mathcal{K}|$ denotes the cardinality of $\mathcal{K}$.
(iv) $\gamma = (\gamma(s))_{s\in S}$ is a probability distribution according to which an initial state is chosen.
(v) $p : \mathcal{K} \to \wp(S)$ is a transition probability function where $\wp(S)$ is the set of probability distributions on S; $p(s'|s,a)$ is the probability of moving from state s to s' when the controller chooses an action $a \in A(s)$.
(vi) $c = \big(c(s,a)\big)_{(s,a)\in\mathcal{K}}$ denotes the running cost vector associated with the objective function, where $c(s,a)$ is the cost incurred at state s when an action $a \in A(s)$ is chosen.

(vii) $d^k = \left(d^k(s,a)\right)_{(s,a)\in\mathcal{K}}$, $k = 1,2,\ldots,K$, denote the running cost vectors associated with the constraints, where $d^k(s,a)$ is the cost incurred at state s when an action $a \in A(s)$ is chosen.

(viii) $(\xi_k)_{k=1}^K$ are the bounds for the constraints.

We consider a discrete-time infinite horizon controlled Markov chain $(\mathbb{X}_t, \mathbb{A}_t)_{t=0}^{\infty}$, where $(\mathbb{X}_t, \mathbb{A}_t)$ denotes the state-action pair at time t. At time $t = 0$, the state of the Markov chain is $s_0 \in S$ according to an initial distribution γ and the controller chooses an action $a_0 \in A(s_0)$. As a consequence, running costs $c(s_0,a_0)$, $d^k(s_0,a_0)$, $k = 1,2,\ldots,K$, are incurred. The Markov chain moves to state s_1 at time $t = 1$ with probability $p(s_1|s_0,a_0)$ and the same thing repeats at s_1 and it continues for the infinite horizon. In general, the decision at time t may depend on the history $h_t = (s_0,a_0,s_1,a_1,\ldots,s_{t-1},a_{t-1},s_t)$ at time t. Let H_t be the set of all possible histories at time t. A history dependent decision rule f_t at time t is a function $f_t : H_t \to \wp(A)$, where for every $h_t \in H_t$ with final state s_t, $f_t(h_t) \in \wp(A(s_t))$. A sequence of history dependent decision rules $f^h = (f_t)_{t=0}^{\infty}$ is called a history dependent policy. When each f_t in the sequence $(f_t)_{t=0}^{\infty}$ depends only on the state at time t, it is called a Markovian policy. A Markovian policy $(f_t)_{t=0}^{\infty}$ is called a stationary policy if there exists a decision rule f such that $f_t = f$ for all t. We write a stationary policy $(f,f,\ldots)$ by f with abuse of notation, and define $f = (f(s))_{s\in S}$ such that $f(s) \in \wp(A(s))$ for every $s \in S$. According to a stationary policy f, whenever the Markov chain visits state s, the controller chooses an action from $A(s)$ according to the decision rule $f(s)$. We denote the set of all history dependent and stationary policies by F_{HD} and F_S, respectively.

The policy f^h and initial distribution γ define a probability measure $P_\gamma^{f^h}$ over the state and action trajectories, and $E_\gamma^{f^h}$ denotes the expectation operator corresponding to probability measure $P_\gamma^{f^h}$. For a given history dependent policy $f^h \in F_{HD}$ and an initial distribution γ, the expected discounted costs at a discount factor $\alpha \in (0,1)$ are defined as [1]

$$\begin{aligned} C_\alpha(\gamma, f^h) &= (1-\alpha)\sum_{t=0}^{\infty}\alpha^t E_\gamma^{f^h} c(\mathbb{X}_t,\mathbb{A}_t) \\ &= \sum_{s\in S}\sum_{a\in A(s)} g_\alpha(\gamma, f^h;\, s,a)c(s,a), \\ D_\alpha^k(\gamma, f^h) &= (1-\alpha)\sum_{t=0}^{\infty}\alpha^t E_\gamma^{f^h} d^k(\mathbb{X}_t,\mathbb{A}_t) \\ &= \sum_{s\in S}\sum_{a\in A(s)} g_\alpha(\gamma, f^h;\, s,a)d^k(s,a), \end{aligned}$$

for all $k = 1, 2, \ldots, K$, where $(1 - \alpha)$ is a normalization constant. The set $\{g_\alpha(\gamma, f^h;\ s, a)\}_{(s,a)}$ is the occupation measure defined by

$$g_\alpha(\gamma, f^h;\ s, a) = (1-\alpha)\sum_{t=0}^{\infty} \alpha^t P_\gamma^{f^h}(\mathbb{X}_t = s, \mathbb{A}_t = a),\ \forall\ (s,a) \in \mathcal{K}.$$

For a given history dependent policy f^h and an initial distribution γ, the expected average costs are defined as [1]

$$\begin{aligned} C_{ea}(\gamma, f^h) &= \limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} E_\gamma^{f^h} c(\mathbb{X}_t, \mathbb{A}_t) \\ &= \sum_{s\in S} \sum_{a\in A(s)} g_{ea}(\gamma, f^h;\ s, a) c(s, a), \\ D_{ea}^k(\gamma, f^h) &= \limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} E_\gamma^{f^h} d^k(\mathbb{X}_t, \mathbb{A}_t) \\ &= \sum_{s\in S} \sum_{a\in A(s)} g_{ea}(\gamma, f^h;\ s, a) d^k(s, a), \end{aligned}$$

for all $k = 1, 2, \ldots, K$. The set $\{g_{ea}(\gamma, f^h;\ s, a)\}_{(s,a)}$ is an occupation measure obtained from the accumulation points of $\{\{g_{ea}^n(\gamma, f^h;\ s, a)\}_{(s,a)}\}_n$, where

$$g_{ea}^n(\gamma, f^h;\ s, a) = \frac{1}{n} \sum_{t=0}^{n-1} P_\gamma^{f^h}(\mathbb{X}_t = s, \mathbb{A}_t = a),\ \forall\ (s,a) \in \mathcal{K}.$$

The optimal policy of a CMDP problem can be obtained by solving the following optimization problem

$$\begin{aligned} \min_{f^h \in F_{HD}}\ & C_{\mathfrak{v}}(\gamma, f^h) \\ \text{s.t.}\ & D_{\mathfrak{v}}^k(\gamma, f^h) \le \xi_k,\ \forall\ k = 1, 2, \ldots, K, \end{aligned} \tag{1}$$

where $\mathfrak{v} = \alpha$ and $\mathfrak{v} = ea$ represent a CMDP problem with discounted cost criterion and average cost criterion, respectively.

Assumption 1. *For the CMDP problem with average cost criterion, we assume that the CMDP is unichain, i.e., under any stationary deterministic policy, the Markov chain is aperiodic and irreducible (with possibly some transient states) [1, 18].*

It is well known that the CMDP problem (1) for $\mathfrak{v} \in \{\alpha, ea\}$ can be restricted to the class of stationary policies without loss of optimality when the running costs and transition probabilities are stationary, i.e., they depend only on states and actions but time. This follows from the fact that the set of occupation measures defined with respect to history dependent policies and the set of

occupation measures defined with respect to stationary policies are equal (see Theorems 3.1 and 4.1 of [1]). Moreover, it follows from Theorems 3.2 and 4.2 of [1] that the set of occupation measures defined with respect to stationary policies for the discounted cost criterion is equal to the set $\mathcal{Q}_\alpha(\gamma)$, and for the average cost criterion it is equal to the set $\mathcal{Q}_{ea}(\gamma)$. For $\mathfrak{v} \in \{\alpha, ea\}$, the set $\mathcal{Q}_\mathfrak{v}(\gamma)$ is defined as

$$\mathcal{Q}_\mathfrak{v}(\gamma) = \Big\{\rho \in \mathbb{R}^{|\mathcal{K}|} \mid \sum_{(s,a)\in\mathcal{K}} \rho(s,a)\big(\delta(s,s') - \eta p(s'|s,a)\big) = (1-\eta)\gamma(s'),\ \forall\, s' \in S,$$
$$\sum_{(s,a)\in\mathcal{K}} \rho(s,a) = 1,\ \rho(s,a) \geq 0,\ \forall\, (s,a) \in \mathcal{K}\Big\},$$

where $\delta(s,s')$ is the Kronecker delta and $\eta \in (0,1]$. If $\eta < 1$, we take it as the discount factor α and $\mathcal{Q}_\mathfrak{v}(\gamma) = \mathcal{Q}_\alpha(\gamma)$, while if $\eta = 1$, $\mathcal{Q}_\mathfrak{v}(\gamma) = \mathcal{Q}_{ea}(\gamma)$. Therefore, the CMDP problem (1) is equivalent to the LP problem

$$\begin{aligned} \min_{\rho}\ & \sum_{s\in S}\sum_{a\in A(s)} \rho(s,a)c(s,a) \\ \text{s.t.}\ & \sum_{s\in S}\sum_{a\in A(s)} \rho(s,a)d^k(s,a) \leq \xi_k,\ \forall\ k = 1,2,\ldots,K, \\ & \rho \in \mathcal{Q}_\mathfrak{v}(\gamma). \end{aligned} \tag{2}$$

If ρ^* is an optimal solution of (2), the optimal stationary policy f^* of the CMDP problem is defined as $f^*(s,a) = \frac{\rho^*(s,a)}{\sum_{a\in A(s)}\rho^*(s,a)}$ for all $s \in S, a \in A(s)$, provided the denominator is non-zero $\big($if it is zero, we choose $f^*(s)$ arbitrarily from $\wp(A(s))\big)$.

3 CMDP under chance constraints

In real-life scenario, the model parameters such as running costs and transition probabilities are subject to errors and are not exactly known. In this paper, we assume that the transition probabilities are stationary and exactly known while the running cost vectors c, d^k, $k = 1,2,\ldots,K$, are random vectors defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The uncertainty in running costs is stationary in nature, i.e., the random cost vectors c, d^k, $k = 1,2,\ldots,K$, do not vary with time. For a given $f^h \in F_{HD}$, the costs $C_\mathfrak{v}(\gamma, f^h), D^k_\mathfrak{v}(\gamma, f^h)$, $k = 1,2,\ldots,K$, corresponding to the objective function and the constraints are random variables. We consider the case where the controller is interested in the minimum long run expected cost which is incurred with at least p_0 probability, and the random constraints are jointly satisfied with at least p_1 probability. This leads to the following joint chance constraint programming based formulation of the

CMDP problem

$$\begin{aligned} &\min_{t, f^h \in F_{HD}} \quad t \\ \text{s.t.} \quad &\mathbb{P}\Big(C_{\mathfrak{v}}(\gamma, f^h) \le t\Big) \ge p_0, \\ &\mathbb{P}\Big(D^k_{\mathfrak{v}}(\gamma, f^h) \le \xi_k,\ k = 1, 2, \ldots, K\Big) \ge p_1. \end{aligned} \tag{3}$$

We call (3) a joint chance-constrained Markov decision process (JCCMDP) problem.

Proposition 1. *The JCCMDP problem* (3) *can be restricted to the class of stationary policies without loss of optimality. Moreover, it can be equivalently written using the decision vector* (t, ρ) *as*

$$\begin{aligned} &\min_{t, \rho} \ t \\ \text{s.t.} \quad &\mathbb{P}\Bigg(\sum_{(s,a)\in\mathcal{K}} \rho(s,a)c(s,a) \le t\Bigg) \ge p_0, \\ &\mathbb{P}\Bigg(\sum_{(s,a)\in\mathcal{K}} \rho(s,a)d^k(s,a) \le \xi_k,\ k = 1, 2, \ldots, K\Bigg) \ge p_1, \\ &\rho \in \mathcal{Q}_{\mathfrak{v}}(\gamma). \end{aligned} \tag{4}$$

If ρ^* *is a part of an optimal solution vector of* (4)*, the optimal stationary policy* f^* *of the JCCMDP problem is given by* $f^*(s,a) = \frac{\rho^*(s,a)}{\sum_{a\in A(s)} \rho^*(s,a)}$ *for all* $s \in S, a \in A(s)$*, provided the denominator is non-zero* $\big($*if it is zero, we choose* $f^*(s)$ *arbitrarily from* $\wp(A(s))\big)$*.*

Proof The proof follows from the fact that the set of occupation measures for history dependent strategies is the same as $\mathcal{Q}_{\mathfrak{v}}(\gamma)$ as discussed in Section 2. □

We consider the case where the running cost vectors c, d^k, $k = 1, 2 \ldots, K$, follow multivariate elliptically symmetric distributions and the dependence among the random constraints present under the joint chance constraint of (4) is driven by Gumbel-Hougaard copula. We first review a few definitions and results on copulas and elliptical distributions, which will be used in subsequent analysis.

3.1 Preliminaries

Definition 1 (Copula [7, 12, 14]) A K-dimensional copula, where $K \ge 2$, is a distribution function on $[0,1]^K$ such that all its one dimensional marginals follow a uniform distribution on $[0,1]$.

The relation between the joint distribution of a random vector $X = (X_1, X_2, \ldots, X_K)$ and a copula is given by the Sklar's theorem [14].

Theorem 1 (Sklar's theorem [14])**.** *For a given K-dimensional distribution function $\hat{\Phi}$ and all its one dimensional marginals $\hat{\Phi}_1, \hat{\Phi}_2, \ldots, \hat{\Phi}_K$, there exists a copula $\mathcal{C}$ that satisfies*

$$\hat{\Phi}(v_1, v_2, \ldots, v_K) = \mathcal{C}\big(\hat{\Phi}_1(v_1), \hat{\Phi}_2(v_2), \ldots, \hat{\Phi}_K(v_K)\big),\ \forall\, v_1, v_2, \ldots, v_K \in \mathbb{R}.$$

When all the marginals are continuous, $\mathcal{C}$ is unique throughout its domain; else it is unique on $\times_{i=1}^{K} Range\, \hat{\Phi}_i$.

In this paper, we consider a particular class of copulas, namely, the Gumbel-Hougaard family of copulas, denoted by $\mathcal{C}_\theta$, defined as [7]

$$\mathcal{C}_\theta(u) = \exp\left\{ -\left[\sum_{k=1}^{K} \big(-\ln(u_k) \big)^\theta \right]^{\frac{1}{\theta}} \right\}, \tag{5}$$

where $u = (u_1, u_2, \ldots, u_K)^T \in [0,1]^K$ and $\theta \geq 1$. When $\theta = 1$, $\mathcal{C}_1(u) = \prod_{k=1}^{K} u_k$ is the product copula. The joint distribution function of a random vector X defined by a product copula implies that the random variables $X_1, X_2, \ldots, X_K$ are independent.

Definition 2 (Elliptical distribution [9]) A K-dimensional random vector X is said to follow a multivariate elliptically symmetric distribution if its characteristic function has the form $E(e^{ib^T X}) = e^{ib^T \mu} \varphi(b^T \Sigma b)$, where $\mu \in \mathbb{R}^K$ is the location parameter, the matrix $\Sigma \in \mathbb{R}^{K \times K}$ is the scale parameter and φ is the characteristic generator. We denote it as $X \sim \mathcal{E}_K(\mu, \Sigma, \varphi)$.

If $X \sim \mathcal{E}_K(\mu, \Sigma, \varphi)$ and the density function of X, denoted by ϕ exists, then it has the form

$$\phi(x) = \frac{1}{\sqrt{\det(\Sigma)}} g\big((x-\mu)^T \Sigma^{-1} (x-\mu)\big),$$

where the scale matrix Σ is positive definite and the function g is the density generator. In such a case, we also write $X \sim \mathcal{E}_K(\mu, \Sigma, g)$. From Theorem 2.16 of [9], a linear combination of the components of an elliptically distributed random vector follows a univariate elliptical distribution. Thus, if $X \sim \mathcal{E}_K(\mu, \Sigma, \varphi)$, for a $b \in \mathbb{R}^K$, $b^T X \sim \mathcal{E}_1(b^T \mu, b^T \Sigma b, \varphi)$. Moreover, if Σ is a positive definite matrix, then $\frac{b^T X - b^T \mu}{\|\Sigma^{\frac{1}{2}} b\|_2}$ is a spherically distributed random variable with characteristic generator φ.

3.2 Second order cone approximations

The deterministic optimization problem equivalent to the JCCMDP problem (4), under certain conditions on random cost vectors c and d^k, is given by the following Proposition 2.

Proposition 2. *Under the assumptions that*
(i) the joint distribution of the random vector $(\rho^T d^1, \rho^T d^2, \ldots, \rho^T d^K)^T$ *is defined using the Gumbel-Hougaard family of copulas given by (5), and*
(ii) $c \sim \mathcal{E}_{|\mathcal{K}|}(\mu_c, \Sigma_c, \varphi)$, $d^k \sim \mathcal{E}_{|\mathcal{K}|}(\mu_{d^k}, \Sigma_{d^k}, \varphi)$, $k = 1, 2, \ldots, K$, *such that* Σ_c, Σ_{d^k}, $k = 1, 2, \ldots, K$, *are positive definite matrices,*
the JCCMDP problem (4) is equivalent to the following non-linear programming problem

$$
\begin{aligned}
&\min_{t,\rho,(y_k)_{k=1}^K} \quad t \\
\text{s.t.} \quad &\rho^T \mu_c + \Phi^{-1}(p_0) \|\Sigma_c^{\frac{1}{2}} \rho\|_2 \le t, \\
&\rho^T \mu_{d^k} + \Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}}) \|\Sigma_{d^k}^{\frac{1}{2}} \rho\|_2 \le \xi_k, \ \forall \ k = 1, 2, \ldots, K, \\
&\sum_{k=1}^{K} y_k = 1, \\
&y_k \ge 0, \ \forall \ k = 1, 2, \ldots, K, \\
&\rho \in \mathcal{Q}_{\mathfrak{v}}(\gamma),
\end{aligned}
\tag{6}
$$

where $p_0, p_1 \in (0.5, 1)$ *and* $\Phi^{-1}(\cdot)$ *is the quantile function of a spherical distribution.*

Proof The proof follows from Lemma 2 of [16] and Lemma 2.2 of [10]. □

Remark 1. *If* $K = 1$, *the optimization problem (6) is an SOCP problem which implies that the JCCMDP problem (4) is equivalent to an SOCP problem.*

For $K \ge 2$, the optimization problem (6) is a non-convex optimization problem. We approximate (6) by two SOCP problems whose optimal values give lower and upper bounds on the optimal value of the JCCMDP problem. Our approximations hold under the following assumption.

Assumption 2. *For each* $k = 1, 2, \ldots, K$, *the quantile function* $\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})$ *is a convex function of* y_k.

For the spherical distributions listed in Table 1, Assumption 2 holds true. This follows from the fact that $\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})$ is a composition of functions $p_1^{y_k^{\frac{1}{\theta}}}$

and $\Phi^{-1}(y_k)$, where former is a convex function of $y_k \in [0,1]$ and latter is non-decreasing and convex function of $y_k \in [p_1, 1]$. The convexity of $\Phi^{-1}(y_k)$ for the probability distributions listed in Table 1 can be verified by checking the following condition

$$(\Phi^{-1})''(y) = \frac{-\phi'(\Phi^{-1}(y))}{(\phi(\Phi^{-1}(y)))^3} \geq 0, \ \forall \ y \in [p_1, 1),$$

where $\phi(\cdot)$ is the corresponding probability density function. The convexity of $\Phi^{-1}(y_k)$ on $[p_1, 1]$ holds because it is a continuous function at $y_k = 1$. We

Table 1 List of some spherical distributions.

Probability Distribution	**Density Function** ($\phi(u)$)
Normal	$ce^{\frac{-u^2}{2}}$
Pearson Type VII	$c\left(1+\frac{u^2}{s}\right)^{-N}$; $N > \frac{1}{2}$, $s > 0$
t	$c\left(1+\frac{u^2}{m}\right)^{\frac{-(1+m)}{2}}$; $m > 0$ an integer
Cauchy	$c\left(1+\frac{u^2}{s}\right)^{-1}$; $s > 0$
Logistic	$\frac{ce^{-u^2}}{\left(1+e^{-u^2}\right)^2}$
Laplace	$ce^{-\lvert u\rvert}$

The constant $c > 0$ is the normalization factor that ensures the density function gives 1 on integration.

approximate the functions $\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})$, $k = 1, 2, \ldots, K$, using piecewise tangent and linear approximations to obtain two SOCP problems whose optimal values give the lower and upper bounds for the optimal value of (6).

3.2.1 Lower bound approximation

For every k, we approximate the term $\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})$ in (6) by choosing N tangent points $y_k^1 < y_k^2 < \ldots < y_k^N$ from the interval $(0, 1]$. We denote the Taylor series expansion around the point y_k^i by $\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})_i$ and define it as

$$\begin{aligned}
\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})_i &= \Phi^{-1}(p_1^{y_k^{i\,\frac{1}{\theta}}}) + (y_k - y_k^i)\frac{d}{dy_k}\left(\Phi^{-1}(p_1^{y_k^{i\,\frac{1}{\theta}}})\right) \\
&= \Phi^{-1}(p_1^{y_k^{i\,\frac{1}{\theta}}}) - \frac{(y_k^i)^{\frac{1}{\theta}} p_1^{y_k^{i\,\frac{1}{\theta}}} \ln(p_1)}{\theta\phi(\Phi^{-1}(p_1^{y_k^{i\,\frac{1}{\theta}}}))} + \frac{p_1^{y_k^{i\,\frac{1}{\theta}}} \ln(p_1)(y_k^i)^{\frac{1}{\theta}-1} y_k}{\theta\phi(\Phi^{-1}(p_1^{y_k^{i\,\frac{1}{\theta}}}))} \\
&= \bar{a}_k^i + \bar{b}_k^i y_k, \ \forall\, i = 1, 2, \ldots, N,
\end{aligned}$$

$$\text{where } \bar{a}_k^i = \Phi^{-1}(p_1^{y_k^i \frac{1}{\theta}}) - \frac{(y_k^i)^{\frac{1}{\theta}} p_1^{y_k^i \frac{1}{\theta}} \ln(p_1)}{\theta\phi(\Phi^{-1}(p_1^{y_k^i \frac{1}{\theta}}))}, \; \bar{b}_k^i = \frac{p_1^{y_k^i \frac{1}{\theta}} \ln(p_1)(y_k^i)^{\frac{1}{\theta}-1}}{\theta\phi(\Phi^{-1}(p_1^{y_k^i \frac{1}{\theta}}))},$$

and $\phi(\cdot)$ is the probability density function of the spherical distribution whose cumulative distribution function is $\Phi(\cdot)$.

Theorem 2. *Let Assumption 2 holds. The lower bound for the optimal value of the equivalent JCCMDP problem (6) is given by the optimal value of the following SOCP problem*

$$\begin{aligned}
&\min_{t,\rho,(x_{kj})_{k=1,\,j=1}^{K\;\;|\mathcal{K}|},(z_{kj})_{k=1,\,j=1}^{K\;\;|\mathcal{K}|}} t \\
\text{s.t.}\;\; &\rho^T\mu_c + \Phi^{-1}(p_0)\|\Sigma_c^{\frac{1}{2}}\rho\|_2 \le t, \\
&\rho^T\mu_{d^k} + \|\Sigma_{d^k}^{\frac{1}{2}} z_k\|_2 \le \xi_k, \;\; \forall\; k = 1,2,\ldots,K, \\
&z_{kj} \ge (\bar{a}_k^i\rho_j + \bar{b}_k^i x_{kj}), \;\; \forall\; k = 1,2,\ldots,K,\; j = 1,2,\ldots,|\mathcal{K}|,\; i = 1,2,\ldots,N, \\
&\sum_{k=1}^{K} x_{kj} = \rho_j, \; \forall\, j = 1,2,\ldots,|\mathcal{K}|, \\
&x_{kj} \ge 0, \; \forall\; k = 1,2,\ldots,K,\; j = 1,2,\ldots,|\mathcal{K}|, \\
&\rho \in \mathcal{Q}_{\mathfrak{v}}(\gamma),
\end{aligned} \tag{7}$$

where ρ_j is the j-th component of $\rho \in \mathbb{R}^{|\mathcal{K}|}$ and $z_k = (z_{kj})_{j=1}^{|\mathcal{K}|}$ for every $k = 1,2,\ldots,K$.

Proof It is sufficient to show that for every feasible solution of (6) there exists a feasible solution of (7). The proof is similar to Theorem 3.1.1 of [6]. Suppose the vector $\left(t,\rho,(y_k)_{k=1}^K\right)$ is a feasible solution of (6). Under Assumption 2, we obtain $\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}}) \ge \max\limits_{1\le i\le N}(\bar{a}_k^i + \bar{b}_k^i y_k)$ for all k. Therefore,

$$\rho^T\mu_{d^k} + \max_{1\le i\le N}(\bar{a}_k^i + \bar{b}_k^i y_k)\|\Sigma_{d^k}^{\frac{1}{2}}\rho\|_2 \le \xi_k, \;\; \forall\; k = 1,2,\ldots,K.$$

The vectors z_k, $k = 1,2,\ldots K$, where $z_{kj} = \max\limits_{1\le i\le N}(\bar{a}_k^i\rho_j + \bar{b}_k^i y_k\rho_j)$ for all $k = 1,2,\ldots,K$, $j = 1,2,\ldots,|\mathcal{K}|$, satisfies

$$\begin{aligned}
&\rho^T\mu_{d^k} + \|\Sigma_{d^k}^{\frac{1}{2}} z_k\|_2 \le \xi_k, \; \forall\; k = 1,2,\ldots,K, \\
&z_{kj} \ge \bar{a}_k^i\rho_j + \bar{b}_k^i y_k\rho_j, \; \forall\; k = 1,2,\ldots,K,\; j = 1,2,\ldots,|\mathcal{K}|.
\end{aligned}$$

Then, the vector $\left(t,\rho,(x_{kj})_{k=1,\,j=1}^{K\;\;|\mathcal{K}|},(z_{kj})_{k=1,\,j=1}^{K\;\;|\mathcal{K}|}\right)$ where

$$x_{kj} = y_k\rho_j, \; \forall\, k = 1,2,\ldots,K,\; j = 1,2,\ldots,|\mathcal{K}|,$$

is a feasible solution for the problem (7). □

3.2.2 Upper bound approximation

For every k, we approximate the term $\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})$ in (6) by choosing N interpolation points $y_k^1 < y_k^2 < \ldots < y_k^N$ from the interval $(0,1]$. We denote the linear interpolating polynomial passing through the points $(y_k^i, \Phi^{-1}(p_1^{y_k^{i\frac{1}{\theta}}}))$ and $(y_k^{i+1}, \Phi^{-1}(p_1^{y_k^{i+1\frac{1}{\theta}}}))$ by $\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})_i$ and define it as

$$\begin{aligned}
\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})_i &= \Phi^{-1}(p_1^{y_k^{i\frac{1}{\theta}}}) + \frac{y_k - y_k^i}{y_k^{i+1} - y_k^i}\Big(\Phi^{-1}(p_1^{y_k^{i+1\frac{1}{\theta}}}) - \Phi^{-1}(p_1^{y_k^{i\frac{1}{\theta}}})\Big) \\
&= \frac{y_k^{i+1}\Phi^{-1}(p_1^{y_k^{i\frac{1}{\theta}}}) - y_k^i\Phi^{-1}(p_1^{y_k^{i+1\frac{1}{\theta}}})}{y_k^{i+1} - y_k^i} + \frac{\big(\Phi^{-1}(p_1^{y_k^{i+1\frac{1}{\theta}}}) - \Phi^{-1}(p_1^{y_k^{i\frac{1}{\theta}}})\big)y_k}{y_k^{i+1} - y_k^i} \\
&= \hat{a}_k^i + \hat{b}_k^i y_k, \ \forall\, i = 1, 2, \ldots, N-1,
\end{aligned}$$

where $\hat{a}_k^i = \dfrac{y_k^{i+1}\Phi^{-1}(p_1^{y_k^{i\frac{1}{\theta}}}) - y_k^i\Phi^{-1}(p_1^{y_k^{i+1\frac{1}{\theta}}})}{y_k^{i+1} - y_k^i}$, $\hat{b}_k^i = \dfrac{\Phi^{-1}(p_1^{y_k^{i+1\frac{1}{\theta}}}) - \Phi^{-1}(p_1^{y_k^{i\frac{1}{\theta}}})}{y_k^{i+1} - y_k^i}$.

Theorem 3. *Let Assumption 2 holds. The upper bound for the optimal value of the equivalent JCCMDP problem (6) is given by the optimal value of the following SOCP problem*

$$\begin{aligned}
&\min_{t,\rho,(y_k)_{k=1}^K,(z_k)_{k=1}^K} \quad t \\
\text{s.t.} \quad &\rho^T\mu_c + \Phi^{-1}(p_0)\|\Sigma_c^{\frac{1}{2}}\rho\|_2 \le t, \\
&\rho^T\mu_{d^k} + z_k\Big(\max_{1\le j\le|\mathcal{K}|}\|(\Sigma_{d^k}^{\frac{1}{2}})_j\|_2\Big) \le \xi_k, \ \forall\ k = 1, 2, \ldots, K, \\
&z_k \ge \hat{a}_k^i + \hat{b}_k^i y_k, \ \forall\ k = 1, 2, \ldots, K,\ i = 1, 2, \ldots, N-1, \\
&\sum_{k=1}^K y_k = 1, \\
&y_k \ge 0, \ \forall\ k = 1, 2, \ldots, K, \\
&\rho \in \mathcal{Q}_{\mathfrak{v}}(\gamma),
\end{aligned} \tag{8}$$

where ρ_j is the j-th component of $\rho \in \mathbb{R}^{|\mathcal{K}|}$ and $(\Sigma_{d^k}^{\frac{1}{2}})_j$ is the j-th column of the matrix $\Sigma_{d^k}^{\frac{1}{2}}$.

Proof It is sufficient to show that for every feasible solution of (8) there exists a feasible solution of (6). Suppose the vector $\big(t, \rho, (y_k)_{k=1}^K, (z_k)_{k=1}^K\big)$ is a feasible solution of (8). Under Assumption 2, we obtain $\Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}}) \le \max_{1\le i\le N-1}(\hat{a}_k^i + \hat{b}_k^i y_k)$ for all k. For

an arbitrarily fixed $k \in \{1, 2, \ldots, K\}$, consider

$$\begin{aligned}\|\Sigma_{d^k}^{\frac{1}{2}}\rho\|_2 &= \|\rho_1(\Sigma_{d^k}^{\frac{1}{2}})_1 + \rho_2(\Sigma_{d^k}^{\frac{1}{2}})_2 + \ldots + \rho_{|\mathcal{K}|}(\Sigma_{d^k}^{\frac{1}{2}})_{|\mathcal{K}|}\|_2 \\ &\leq \rho_1\|(\Sigma_{d^k}^{\frac{1}{2}})_1\|_2 + \rho_2\|(\Sigma_{d^k}^{\frac{1}{2}})_2\|_2 + \ldots + \rho_{|\mathcal{K}|}\|(\Sigma_{d^k}^{\frac{1}{2}})_{|\mathcal{K}|}\|_2 \\ &\leq \max_{1\leq j\leq|\mathcal{K}|}\|(\Sigma_{d^k}^{\frac{1}{2}})_j\|_2.\end{aligned}$$

The first inequality holds by triangular inequality of $\|\cdot\|_2$, while the second inequality holds by the last two conditions of the set $\mathcal{Q}_{\mathfrak{v}}(\gamma)$. Consequently, the following inequality holds true

$$\begin{aligned}\rho^T\mu_{d^k} + \Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})\|\Sigma_{d^k}^{\frac{1}{2}}\rho\|_2 &\leq \rho^T\mu_{d^k} + \max_{1\leq i\leq N-1}(\hat{a}_k^i + \hat{b}_k^i y_k)\Big(\max_{1\leq j\leq|\mathcal{K}|}\|(\Sigma_{d^k}^{\frac{1}{2}})_j\|_2\Big) \\ &\leq \rho^T\mu_{d^k} + z_k\Big(\max_{1\leq j\leq|\mathcal{K}|}\|(\Sigma_{d^k}^{\frac{1}{2}})_j\|_2\Big).\end{aligned}$$

Therefore,

$$\rho^T\mu_{d^k} + \Phi^{-1}(p_1^{y_k^{\frac{1}{\theta}}})\|\Sigma_{d^k}^{\frac{1}{2}}\rho\|_2 \leq \xi_k, \ \forall\ k = 1, 2, \ldots, K.$$

Hence, the vector $\big(t, \rho, (y_k)_{k=1}^K\big)$ is a feasible solution for the problem (6). □

4 Queueing control problem

We consider a stochastic version of a discrete-time service and admission control problem in a single queue introduced by Altman (see Chapter 5 of [1]). The system has two controllers, namely, service and admission controllers. We consider them as a single controller because their joint objective is to minimize the expected cost incurred in the long run. The states represent the queue length with buffer size L. Thus, $S = \{0, 1, \ldots, L\}$, where $s = 0$ indicates that the queue is empty, while $s = L$ indicates that the queue is full. The action set of the controller is $A^1 \times A^2$, where A^1 and A^2 are finite sets of probability values. If $(a^1, a^2) \in A^1 \times A^2$, then a^1 and a^2 denote the probabilities with which the controller finishes the service and admits a customer into the system, respectively. We assume that $A^1 \subseteq [a^1_{min}, a^1_{max}]$ and $A^2 \subseteq [a^2_{min}, a^2_{max}]$, where $0 < a^1_{min} \leq a^1_{max} < 1$ and $0 \leq a^2_{min} \leq a^2_{max} < 1$. Moreover, when $s = L$, we assume no admission takes place. The transition probabilities from [1] are given by

$$p(s'|s, (a^1, a^2)) = \begin{cases} a^1(1-a^2) & 1 \leq s \leq L-1, s' = s-1, \\ a^1a^2 + (1-a^1)(1-a^2) & 1 \leq s \leq L-1, s' = s, \\ (1-a^1)a^2 & 0 \leq s \leq L-1, s' = s+1, \\ 1-(1-a^1)a^2 & s' = s = 0, \\ 1-a^1 & s = L, s' = L, \\ a^1 & s = L, s' = L-1. \end{cases}$$

If at state $s \in S$, the controller chooses an action (a^1, a^2), it incurs running costs $c(s)$, $d^1(a^1)$ and $d^2(a^2)$. The running cost $c(s)$ is the holding cost incurred due

to customers waiting in the queue. Therefore, it only depends on the size of the queue. The holding cost is zero for an empty queue, i.e., $c(0) = 0$. The running cost $d^1(a^1)$ is a service cost incurred when a customer is served with probability a^1 and $d^2(a^2)$ denotes the cost when a customer is not admitted into the system with $(1-a^2)$ probability. The service cost $d^1(a^1)$ increases with a^1 and no admission cost $d^2(a^2)$ decreases with a^2. These costs do not depend on the size of the queue. The controller aims to minimize the expected holding cost in the long run by keeping the long run expected service and admission costs below certain threshold values. Usually the costs in the queueing system are not realized before the decisions are taken. Therefore, the running cost vectors c , d^1 and d^2 are better modelled using random vectors. We assume that they follow normal distribution and denote them as $c \sim \mathcal{N}(\mu_c, \Sigma_c)$, $d^k \sim \mathcal{N}(\mu_{d^k}, \Sigma_{d^k})$, $k = 1, 2$, where $c = (c(s))_{s \in S}$, $d^1 = (d^1(a^1))_{a^1 \in A^1}$ and $d^2 = (d^2(a^2))_{a^2 \in A^2}$.

4.1 Numerical experiments

We perform numerical experiments, using CVX package in MATLAB optimization toolbox, on an Intel(R) 64-bit Core(TM) i5-8250U CPU @ 1.60GHz with 8.0 GB RAM machine. In all the numerical experiments, we fix $p_0, p_1 = 0.95$, $\alpha = 0.99$ and assume γ to be a uniformly distributed vector. We consider the case where the number of states are 10, i.e., $L = 9$, and the controller finishes the service with probability $a^1 = 0.75$. The admission controller has two possible actions: 'No Admission ($a^2 = 0$)' or 'Admission with probability 0.8 ($a^2 = 0.8$)'. We take c, d^2 to be normally distributed random vectors and d^1 to be fixed, thus $K = 1$. Hence, from Remark 1, it follows that the JCCMDP problem (4) reduces to an SOCP problem. We take $\xi_2 = 9$, $\mu_c = (0, 1, 2, \ldots, 9)^T$, the matrix $\Sigma_c \in \mathbb{R}^{10 \times 10}$ with all diagonal values 0.9 and off-diagonal values 0.35, and

$$\mu_{d^2} = \begin{pmatrix} 10.0 \\ 7.60 \end{pmatrix}, \quad \Sigma_{d^2} = \begin{pmatrix} 0.80 & 0.24 \\ 0.24 & 0.61 \end{pmatrix}.$$

The optimal values of the JCCMDP problem with the above data for discounted and average cost cases are 5.7963 and 6.2296, respectively. Figure 1 shows the optimal probability of not admitting a customer into the queue at various states. For both the cases, we conclude that the optimal probability of 'No Admission' is 1 at the last state.

We consider a relatively large system where the action set of the controller is given by $A^1 \times A^2 = \{(0.2, 0.75, 0.9), (0, 0.5, 0.8)\}$. Thus, $K = 2$, and we compute the lower bound and upper bound for the optimal value of the JCCMDP problem (4) by solving the SOCP problems (7) and (8), respectively. For $L = n - 1$ ($|S| = n$), we take $\mu_c = (0, 1, 2, \ldots, n-1)^T$, the matrix $\Sigma_c \in \mathbb{R}^{n \times n}$ with all diagonal values 0.9 and off-diagonal values 0.35, $\mu_{d^1} = (4.32, 9.1875, 10.83)^T$,

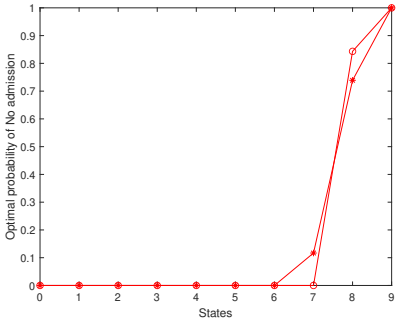

Fig. 1 The solid lines marked with $'*'$ and $'o'$ denote the optimal probability of not admitting a customer into the queue for discounted and average cost problems, respectively.

$\mu_{d^2} = (10.00, 8.50, 7.60)^T$, and

$$\Sigma_{d^1} = \begin{pmatrix} 0.15 & 0.05 & 0.10 \\ 0.05 & 0.10 & 0.15 \\ 0.10 & 0.15 & 0.40 \end{pmatrix}, \quad \Sigma_{d^2} = \begin{pmatrix} 0.80 & 0.35 & 0.24 \\ 0.35 & 0.70 & 0.20 \\ 0.24 & 0.20 & 0.61 \end{pmatrix}.$$

The mean vector μ_c, where $\mu_c(s) = s$, captures the fact that the holding cost increases with s. We take the mean vectors μ_{d^1} and μ_{d^2} such that $\mu_{d^1}(a^1) = 3(1 + a^1)^2$ and $\mu_{d^2}(a^2) = 10 - 3a^2$. It is motivated from the fact that the service cost increases with a^1 and no admission cost decreases with a^2. We take $\xi_1 = 11.30$, $\xi_2 = 11.35$, the sets of tangent and linear points are $\{0.01, 0.15, 0.45\}$ and $\{10^{-5}, 10^{-3}, 0.15, 1\}$, respectively, such that the interpolation point y_k^1 is sufficiently close to 0. We solve the SOCP problems (7) and (8) for various values of θ and $|S|$. The higher values of θ indicates stronger dependence between the random vectors $d^1(a^1)$ and $d^2(a^2)$ driven by Gumbel-Hougaard copula. The numerical results are summarized in Tables 2 and 3 for the case of discounted and average cost criteria, respectively. The lower bound and upper bound for the optimal value of JCCMDP problem is given in column 3 and column 5 of each table, respectively. The computational time to solve (7) and (8) is given in columns 4 and 6, respectively. The last column of both the Tables represent the gap between the upper and lower bound for the optimal value of the JCCMDP problem (4), where the gap is defined by

$$\text{Gap}(\%) = \frac{\text{Opt.(UB)-Opt.(LB)}}{\text{Opt.(LB)}} \times 100.$$

It is clear from Tables 2 and 3 that in most of the cases the gap is very small. Only in a few instances it goes up to 2.15%. This shows that the SOCP problems (7) and (8) are good approximations of the JCCMDP problem. The time analysis shows that the SOCP problems (7) and (8) can be solved efficiently.

Table 2 Lower and Upper bounds for Discounted cost case.

θ	States	Lower Bound (LB)		Upper Bound (UB)		Gap(%)
		Opt.	CPU(s)	Opt.	CPU(s)	
1	200	47.3240	15.4219	47.3260	10.1875	0.0042
	400	131.2503	57.1094	131.3343	23.1094	0.0639
	600	224.7818	142.8438	229.3581	72.2656	2.0358
	800	322.1396	347.7031	329.0680	161.7188	2.1507
3	200	47.3247	14.8125	47.3260	5.2500	0.0027
	400	131.2513	53.8281	131.2536	23.2500	0.0017
	600	224.7558	142.5938	224.8627	52.2031	0.0475
	800	321.4530	312.4063	323.9796	176.6094	0.7859
6	200	47.3250	13.0938	47.3260	5.0313	0.0021
	400	131.2518	49.5313	131.2533	19.4688	0.0011
	600	224.7562	140.2188	224.7802	47.6250	0.0106
	800	321.4326	314.4531	322.6907	158.7344	0.3914

Table 3 Lower and Upper bounds for Average cost case.

θ	States	Lower Bound (LB)		Upper Bound (UB)		Gap(%)
		Opt.	CPU(s)	Opt.	CPU(s)	
1	200	1.5604	19.3438	1.5607	10.8906	0.0192
	400	1.5604	61.9531	1.5607	26.1875	0.0192
	600	1.5604	166.5781	1.5607	64.7344	0.0192
	800	1.5604	340.1719	1.5607	120.9219	0.0192
3	200	1.5605	15.5625	1.5607	5.9688	0.0128
	400	1.5605	53.3750	1.5607	29.7344	0.0128
	600	1.5605	156.7969	1.5607	71.3594	0.0128
	800	1.5605	287.0469	1.5607	131.2813	0.0128
6	200	1.5606	18.2188	1.5607	9.8750	0.0064
	400	1.5606	56.7344	1.5607	24.6563	0.0064
	600	1.5606	150.8281	1.5607	61.7656	0.0064
	800	1.5606	276.4063	1.5607	135.9844	0.0064

5 Conclusion

We consider a CMDP problem with random costs under discounted and average cost criteria. We formulate it as a joint chance-constrained optimization problem where the dependence among the random cost constraint vectors is driven by a Gumbel-Hougaard copula. When the random cost vectors follow multivariate elliptically symmetric distributions, we propose SOCP based approximations which give upper and lower bounds on the optimal value of the JCCMDP problem. We illustrate our results on a queueing control problem by considering various small and large instances. The gap between the upper and lower bounds is very small in most of our experiments which shows the

effectiveness of the SOCP approximations. The time analysis shows that both the approximations can be solved efficiently for large instances.

Acknowledgement

The research of first author was supported by CSIR, India. The research of second and third author was supported by DST/CEFIPRA Project No. IFC/4117/DST-CNRS-5th call/2017-18/2 and CNRS Project No. AR/SB:2018-07-440, respectively.

References

[1] Altman, E.: Constrained Markov Decision Processes. Chapman and Hall/CRC, London (1999)

[2] Altman, E.: Asymptotic properties of constrained Markov decision processes. Zeitschrift für Operations Research **37**(2), 151–170 (1993)

[3] Archak, N., Mirrokni, V., Muthukrishnan, S.: Budget optimization for online advertising campaigns with carryover effects. In: Sixth Ad Auctions Workshop (2010)

[4] Bertsekas, D.P.: Dynamic Programming and Optimal Control, 2nd edn. Athena Scientific, Belmont, MA, USA (2000)

[5] Charnes, A., Cooper, W.W.: Chance-constrained programming. Management Science **6**, 73–79 (1959)

[6] Cheng, J., Lisser, A.: A second-order cone programming approach for linear programs with joint probabilistic constraints. Operations Research, 325–328 (2012)

[7] Cheng, J., Houda, M., Lisser, A.: Chance constrained 0–1 quadratic programs using copulas. Optimization Letters **9**(7), 1283–1295 (2015)

[8] Delage, E., Mannor, S.: Percentile optimization for Markov decision processes with parameter uncertainty. Operations Research **58**(1), 203–213 (2010)

[9] Fang, K.-T., Kotz, S., Ng, K.W.: Symmetric Multivariate and Related Distributions. Chapman and Hall/CRC, London, New York (1990)

[10] Henrion, R.: Structural properties of linear probabilistic constraints. Optimization **56**, 425–440 (2007)

[11] Iyengar, G.N.: Robust dynamic programming. Mathematics of Operations Research **30**(2), 257–280 (2005)

[12] Jaworski, P., Durante, F., Härdle, W.K., Rychlik, T.: Copula Theory and Its Applications. Proceedings of the Workshop Held in Warsaw, Poland, 25–26 September 2009, (2010)

[13] Mannor, S., Simester, D., Sun, P., Tsitsiklis, J.N.: Bias and variance approximation in value function estimates. Management Science **53**(2), 308–322 (2007)

[14] Nelsen, R.B.: An Introduction to Copulas (Springer Series in Statistics). Springer, Berlin, Heidelberg (2006)

[15] Nilim, A., El Ghaoui, L.: Robust control of Markov decision processes with uncertain transition matrices. Operations Research **53**(5), 780–798 (2005)

[16] Peng, S., Maggioni, F., Lisser, A.: Bounds for probabilistic programming with application to a blend planning problem. European Journal of Operational Research **297** (2021)

[17] Prékopa, A.: Stochastic Programming. Kluwer Academic Publishers, Berlin (1995)

[18] Puterman, M.L.: Markov Decision Process, 1st edn. John Wiley & Sons, USA (1994)

[19] Satia, J.K., Lave Jr, R.E.: Markovian decision processes with uncertain transition probabilities. Operations Research **21**(3), 728–740 (1973)

[20] Sklar, M.: Fonctions de répartition à n dimensions et leurs marges. Publ. inst. statist. univ. Paris **8**, 229–231 (1959)

[21] Varagapriya, V., Singh, V.V., Lisser, A.: Constrained Markov decision processes with uncertain costs. Operations Research Letters **50**(2), 218–223 (2022)

[22] White III, C.C., Eldeib, H.K.: Markov decision processes with imprecise transition probabilities. Operations Research **42**(4), 739–749 (1994)

[23] Wiesemann, W., Kuhn, D., Rustem, B.: Robust Markov decision processes. Mathematics of Operations Research **38**(1), 153–183 (2013)