



HAL
open science

Joint Chance-Constrained Markov Decision Processes

V Varagapriya, Vikas Vikram Singh, Abdel Lisser

► **To cite this version:**

V Varagapriya, Vikas Vikram Singh, Abdel Lisser. Joint Chance-Constrained Markov Decision Processes. *Annals of Operations Research*, 2023, 322, pp.1013-1035. 10.1007/s10479-022-05025-3 . hal-04375382

HAL Id: hal-04375382

<https://hal.science/hal-04375382>

Submitted on 5 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Joint Chance-Constrained Markov Decision Processes

V Varagapriya¹, Vikas Vikram Singh¹ and Abdel Lisser²

¹Department of Mathematics, Indian Institute of Technology
Delhi, Hauz Khas, 110016, New Delhi, India.

²Universite Paris Saclay, CNRS, CentraleSupélec, 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. We assume that the random cost vectors follow multivariate elliptically symmetric distributions and dependence among the random constraints is driven by a Gumbel-Hougaard copula. We propose two second order cone programming problems whose optimal values give lower and upper 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. The proposed approximation methods are illustrated on randomly generated instances of queueing control problem as well as on well known class of Markov decision problems known as Garnets.

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, 3, 19]. MDPs can also be solved via policy iteration [25]. 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. For instance, in [1], the author studies the example of service and admission control in a queueing system in which the controller incurs three different costs at every stage. In [22], the authors study a variant of a machine replacement problem by considering the case when the controller incurs opportunity cost together with maintenance cost at every stage. In such cases, one type of running cost defines a long-run expected cost function which needs to be minimized and other types of running costs define the long-run expected cost functions on which constraints are imposed. This class of MDP is called a constrained Markov decision process (CMDP). For a finite state-action CMDP problem, when running costs and transition probabilities are stationary and exactly known, there exists a stationary randomized optimal policy which 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 [15]. The early literature on MDPs with uncertain parameters dates back to the 1970s, where MDPs under uncertain transition probabilities were studied [20, 23]. Previous studies have accounted for the uncertainties in MDPs under a robust optimization framework by considering the worst case realization of the uncertain parameters [12, 17, 24]. 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. [22]. Delage and Mannor [7] 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 [7]. 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 [7].

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 finite state-action 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 [4, 18] 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 are 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 [21]. 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 choice of elliptical distributions is due to the fact that it is a generalization of normal distributions and contains a list of important symmetric distributions [9] and they are closed under linear transformations. The Gumbel-Hougaard copula is considered because it is separable in nature. Using the fact that the elliptical distributions are closed under a linear transformation and Gumbel-Hougaard copula is separable, we propose a deterministic non-convex optimization problem which is equivalent to the JCCMDP problem. We approximate the equivalent non-convex problem by two SOCP problems whose optimal values give lower and upper 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 and on random CMDPs. 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) $p : \mathcal{K} \rightarrow \wp(S)$ 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) $c = (c(s, a))_{(s, a) \in \mathcal{K}}$ 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) $d^k = (d^k(s, a))_{(s, a) \in \mathcal{K}}$, $k = 1, 2, \dots, K$, 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 vector stochastic process $(\mathbb{X}_t, \mathbb{A}_t)_{t=0}^\infty$, 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 $c(s_0, a_0)$, $d^k(s_0, a_0)$, $k = 1, 2, \dots, K$, 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 $h_t = (s_0, a_0, s_1, a_1, \dots, s_{t-1}, a_{t-1}, s_t)$ 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 $f_t : H_t \rightarrow \wp(A)$, 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, \dots) 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 (for details see Section 2.1.6 of [19]), 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 $k = 1, 2, \dots, K$, where $(1 - \alpha)$ is a normalization constant. The set $\{g_\alpha(\gamma, f^h; s, a)\}_{(s,a)}$ is the occupation measure defined by

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

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

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

for all $k = 1, 2, \dots, K$. 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

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

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

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

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

We make the following assumption on the average cost criterion throughout the paper.

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, 19].*

Assumption 1 ensures that the accumulation points of $\{\{g_{ea}^n(\gamma, f^h; s, a)\}_{(s,a)}\}_n$ is unique and independent of the initial distribution γ . Therefore, the expected average costs are well defined. We do not need Assumption 1 in the case of discounted cost criterion. It is well known that the CMDP problem (1) for $\mathbf{v} \in \{\alpha, ea\}$ 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 Theorem 3.2 (resp. Theorem 4.2) of [1] that the set of occupation measures defined with respect to stationary policies for the discounted (resp. average) cost criterion and the set $\mathcal{Q}_\alpha(\gamma)$ (resp. $\mathcal{Q}_{ea}(\gamma)$) are equal. For $\mathbf{v} \in \{\alpha, ea\}$, the set $\mathcal{Q}_{\mathbf{v}}(\gamma)$ is defined as

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

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

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

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 (if it is zero, we choose $f^*(s)$ arbitrarily from $\wp(A(s))$).

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 are random vectors defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote the running cost vectors corresponding to objective function and constraints by \hat{c} and \hat{d}^k , $k = 1, 2, \dots, K$, respectively. Therefore, for each realization $\omega \in \Omega$, $\hat{c}(s, a, \omega)$, $\hat{d}^k(s, a, \omega)$ represents the real valued costs which are incurred at state s when an action a is chosen. We assume that the uncertainty in running costs is stationary in nature, i.e., the random cost vectors \hat{c} and \hat{d}^k , $k = 1, 2, \dots, K$, do not vary with time. For a given $f^h \in F_{HD}$, the costs corresponding to the objective function and the constraints of (1), denoted by, $\hat{C}_v(\gamma, f^h)$, $\hat{D}_v^k(\gamma, f^h)$, $k = 1, 2, \dots, K$, are random variables and it makes (1) a stochastic optimization problem. 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

$$\begin{aligned} & \min_{t, f^h \in F_{HD}} t \\ \text{s.t. } & \mathbb{P}\left(\hat{C}_v(\gamma, f^h) \leq t\right) \geq p_0, \\ & \mathbb{P}\left(\hat{D}_v^k(\gamma, f^h) \leq \xi_k, k = 1, 2, \dots, K\right) \geq p_1. \end{aligned} \quad (3)$$

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

Proposition 1. *Let Assumption 1 holds for average cost criterion. Then, 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*

$$\begin{aligned} & \min_{t, \rho} t \\ \text{s.t. } & (i) \mathbb{P}\left(\sum_{(s,a) \in \mathcal{K}} \rho(s, a) \hat{c}(s, a) \leq t\right) \geq p_0, \\ & (ii) \mathbb{P}\left(\sum_{(s,a) \in \mathcal{K}} \rho(s, a) \hat{d}^k(s, a) \leq \xi_k, k = 1, 2, \dots, K\right) \geq p_1, \\ & (iii) \rho \in \mathcal{Q}_v(\gamma). \end{aligned} \quad (4)$$

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 (if it is zero, we choose $f^*(s)$ arbitrarily from $\wp(A(s))$).

Proof Fix a realization $\omega \in \Omega$. Let (t, f^h) be a feasible solution of (3). It follows from Theorems 3.1 and 4.1 of [1] that the set of occupation measures corresponding to history dependent strategies and stationary strategies are the same, i.e., there exists $f \in F_S$ such that $g_v(\gamma, f^h; s, a) = g_v(\gamma, f; s, a)$ for all $(s, a) \in \mathcal{K}$. Hence, $\hat{C}_v(\gamma, f^h, \omega) = \hat{C}_v(\gamma, f, \omega)$ and $\hat{D}_v^k(\gamma, f^h, \omega) = \hat{D}_v^k(\gamma, f, \omega)$, $k = 1, 2, \dots, K$. Thus, we can restrict the JCCMDP problem (3) to the class of stationary policies and obtain the following equivalent problem

$$\begin{aligned} & \min_{t, f \in F_S} t \\ \text{s.t. } & \mathbb{P}\left(\hat{C}_v(\gamma, f) \geq t\right) \geq p_0, \\ & \mathbb{P}\left(\hat{D}_v^k(\gamma, f) \leq \xi_k, k = 1, 2, \dots, K\right) \geq p_1. \end{aligned} \quad (5)$$

Now, we show that (4) and (5) are equivalent. Let (t, f) be a feasible point of (5). Define a vector ρ such that $\rho(s, a) = g_v(\gamma, f; s, a)$ for all $(s, a) \in \mathcal{K}$. It follows from Theorems 3.2 and 4.2 of [1] that $\rho \in \mathcal{Q}_v(\gamma)$ and $\hat{C}_v(\gamma, f, \omega) = \sum_{(s,a) \in \mathcal{K}} \rho(s, a) \hat{c}(s, a, \omega)$

and $\hat{D}_v^k(\gamma, f, \omega) = \sum_{(s,a) \in \mathcal{K}} \rho(s, a) \hat{d}^k(s, a, \omega)$, $k = 1, 2, \dots, K$. Hence, (t, ρ) is a feasible point of (4). Conversely, if (t, ρ) is a feasible solution of (4), then there exists $f \in F_S$ defined as $f(s, a) = \frac{\rho(s, a)}{\sum_{a \in A(s)} \rho(s, a)}$, whenever the denominator is nonzero (if it is zero, we choose $f(s)$ arbitrarily from $\wp(A(s))$).

For such an f , we have $\hat{C}_v(\gamma, f, \omega) = \sum_{(s,a) \in \mathcal{K}} \rho(s, a) \hat{c}(s, a, \omega)$ and $\hat{D}_v^k(\gamma, f, \omega) = \sum_{(s,a) \in \mathcal{K}} \rho(s, a) \hat{d}^k(s, a, \omega)$, $k = 1, 2, \dots, K$.

This implies that (t, f) is a feasible solution of (5). Thus, (5) is equivalent to (4). \square

The problem (4) is an LP problem with a joint chance constraint [14]. In general, such problems are hard to solve. This is because, for any given ρ , finding the feasibility of the joint chance constraint requires multi-dimensional integration, which is NP-hard [10]. Moreover, the feasible set of this constraint is not convex. In order to obtain an equivalent deterministic optimization problem, we need the information of the distributions of random variables $\rho^T \hat{c}$ and $\rho^T \hat{d}^k$. These random variables are linear combinations of random running cost vectors \hat{c} and \hat{d}^k . Therefore, we assume that the running cost vectors \hat{c} , \hat{d}^k follow multivariate elliptically symmetric distributions. The class of elliptically symmetric distributions is a generalization of the normal distribution [9] and it is closed under linear transformation (Theorem 2.16 of [9]). Thus, the distribution of running cost vectors extends to the associated long-run expected costs. Further, we assume the dependence among the random constraints present under the joint chance constraint of (4) is driven by Gumbel-Hougaard copula. The separable nature of this copula enables us to express the joint chance

constraint as a set of individual chance constraints. 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 [6, 13, 16]) A K -dimensional copula, where $K \geq 2$, 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 $X = (X_1, X_2, \dots, X_K)$ and a copula is given by the Sklar's theorem [16].

Theorem 1 (Sklar's theorem [16]). *For a given K -dimensional distribution function $\hat{\Phi}$ and all its one dimensional marginals $\hat{\Phi}_1, \hat{\Phi}_2, \dots, \hat{\Phi}_K$, there exists a copula \mathcal{C} that satisfies*

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

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

where $\text{Range } \hat{\Phi}_i$ denotes the set of all possible values $\hat{\Phi}_i$ can take for $i = 1, 2, \dots, K$. In this paper, we consider a particular class of copulas, namely, the Gumbel-Hougaard family of copulas, denoted by \mathcal{C}_θ , defined as in [6]

$$\mathcal{C}_\theta(u) = \exp \left\{ - \left[\sum_{k=1}^K (-\ln(u_k))^\theta \right]^{\frac{1}{\theta}} \right\}, \quad (6)$$

where $u = (u_1, u_2, \dots, u_K)^T \in [0, 1]^K$, $\theta \geq 1$ and \exp is the exponential function. 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, \dots, 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

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

where the scale matrix Σ is positive definite, $\det(\Sigma)$ is the determinant of Σ , 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

By using Sklar's theorem given in Section 3.1, the joint chance constraint (ii) of (4) can be equivalently written as

$$\begin{aligned} & \mathbb{P}\left(\sum_{(s,a) \in \mathcal{K}} \rho(s,a) \hat{d}^k(s,a) \leq \xi_k, k = 1, 2, \dots, K\right) \\ &= \mathcal{C}_{\theta, \rho}\left(\mathbb{P}(\rho^T \hat{d}^1 \leq \xi_1), \mathbb{P}(\rho^T \hat{d}^2 \leq \xi_2), \dots, \mathbb{P}(\rho^T \hat{d}^K \leq \xi_K)\right) \geq p_1, \end{aligned}$$

where $\mathcal{C}_{\theta, \rho}$ denotes the copula that defines the joint distribution of the random vector $(\rho^T \hat{d}^1, \rho^T \hat{d}^2, \dots, \rho^T \hat{d}^K)^T$. We make the following assumption on this copula.

Assumption 2. *There exists a K -dimensional Gumbel-Hougaard copula, \mathcal{C}_θ given by (6), such that $\mathcal{C}_{\theta, \rho} = \mathcal{C}_\theta$ for all $\rho \in \mathcal{Q}_v(\gamma)$.*

Proposition 2. *Let Assumption 2 holds, and $\hat{c} \sim \mathcal{E}_{|\mathcal{K}|}(\mu_{\hat{c}}, \Sigma_{\hat{c}}, \varphi_{\hat{c}})$, $\hat{d}^k \sim \mathcal{E}_{|\mathcal{K}|}(\mu_{\hat{d}^k}, \Sigma_{\hat{d}^k}, \varphi_{\hat{d}^k})$, $k = 1, 2, \dots, K$, such that $\Sigma_{\hat{c}}, \Sigma_{\hat{d}^k}$, $k = 1, 2, \dots, K$, are positive definite matrices. Then, the JCCMDP problem (4) is equivalent to the following non-linear programming problem,*

$$\begin{aligned} & \min_{t, \rho, (y_k)_{k=1}^K} t \\ \text{s.t. } & (i) \quad \rho^T \mu_{\hat{c}} + \Phi_{\hat{c}}^{-1}(p_0) \|\Sigma_{\hat{c}}^{\frac{1}{2}} \rho\|_2 \leq t, \\ & (ii) \quad \rho^T \mu_{\hat{d}^k} + \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{\frac{1}{\theta}}}) \|\Sigma_{\hat{d}^k}^{\frac{1}{2}} \rho\|_2 \leq \xi_k, \quad \forall k = 1, 2, \dots, K, \quad (7) \\ & (iii) \quad \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad \forall k = 1, 2, \dots, K, \\ & (iv) \quad \rho \in \mathcal{Q}_v(\gamma), \end{aligned}$$

where $p_0, p_1 \in (0.5, 1)$ and $\Phi_{\hat{c}}^{-1}(\cdot)$, $\Phi_{\hat{d}^k}^{-1}(\cdot)$ are the quantile functions of spherical distributions corresponding to \hat{c} , \hat{d}^k , respectively.

Proof Since, $\hat{c} \sim \mathcal{E}_{|\mathcal{K}|}(\mu_{\hat{c}}, \Sigma_{\hat{c}}, \varphi_{\hat{c}})$, the constraint (i) of (4) is equivalent to constraint (i) of (7) (see Lemma 2.2 of [11]). It follows from Lemma 1 of [6] that the joint chance constraint (ii) of (4) can be equivalently written as

$$\begin{aligned} \mathbb{P}(\rho^T \hat{d}^k \leq \xi_k) &\geq p_1^{y_k^{\frac{1}{\theta}}}, \quad \forall k = 1, 2, \dots, K, \\ \sum_{k=1}^K y_k &= 1, \\ y_k &\geq 0, \quad \forall k = 1, 2, \dots, K. \end{aligned} \quad (8)$$

We provide the proof to make our paper self-contained. Let ρ be a feasible point of the joint chance constraint (ii) of (4). Define $\bar{y}_k, y_k, k = 1, 2, \dots, K$, such that $\mathbb{P}(\rho^T \hat{d}^k \leq \xi_k) = p_1^{\bar{y}_k^{\frac{1}{\theta}}}$, i.e., $\bar{y}_k = \left(\frac{\ln \mathbb{P}(\rho^T \hat{d}^k \leq \xi_k)}{\ln(p_1)} \right)^\theta$ for all k , and $y_k = \frac{\bar{y}_k}{\sum_{k=1}^K \bar{y}_k}$ for all k . From the joint chance constraint (ii) of (4), we have

$$\begin{aligned} p_1 &\leq \mathbb{P}\left(\sum_{(s,a) \in \mathcal{K}} \rho(s,a) \hat{d}^k(s,a) \leq \xi_k, \quad k = 1, 2, \dots, K \right) \\ &= \mathcal{C}_\theta \left(\mathbb{P}(\rho^T \hat{d}^1 \leq \xi_1), \mathbb{P}(\rho^T \hat{d}^2 \leq \xi_2), \dots, \mathbb{P}(\rho^T \hat{d}^K \leq \xi_K) \right) \\ &= \mathcal{C}_\theta \left(p_1^{\bar{y}_1^{\frac{1}{\theta}}}, p_1^{\bar{y}_2^{\frac{1}{\theta}}}, \dots, p_1^{\bar{y}_K^{\frac{1}{\theta}}} \right) \\ &= \exp \left\{ - \left[\sum_{k=1}^K \left(-\ln(p_1^{\bar{y}_k^{\frac{1}{\theta}}}) \right)^\theta \right]^{\frac{1}{\theta}} \right\} \\ &= \exp \left\{ - \left[\sum_{k=1}^K \bar{y}_k \left(\ln \frac{1}{p_1} \right)^\theta \right]^{\frac{1}{\theta}} \right\} \\ &= \exp \left\{ \left(\sum_{k=1}^K \bar{y}_k \right)^{\frac{1}{\theta}} \ln(p_1) \right\} \\ &= p_1^{\left(\sum_{k=1}^K \bar{y}_k \right)^{\frac{1}{\theta}}}, \end{aligned}$$

where the first equality follows by Sklar's theorem. The condition $p_1 \leq p_1^{\left(\sum_{k=1}^K \bar{y}_k \right)^{\frac{1}{\theta}}}$ implies that $\sum_{k=1}^K \bar{y}_k \leq 1$. Then, from the definition of y_k , we have $\bar{y}_k \leq y_k$ for all k , which in turn implies that

$$\mathbb{P}(\rho^T \hat{d}^k \leq \xi_k) \geq p_1^{y_k^{\frac{1}{\theta}}}, \quad \forall k = 1, 2, \dots, K.$$

The other two constraints of (8) follow from the definition of y_k .

Conversely, suppose $(\rho, (y_k)_{k=1}^K)$ is feasible for (8). By Sklar's theorem and the first constraint of (8),

$$\begin{aligned} &\mathbb{P}\left(\sum_{(s,a) \in \mathcal{K}} \rho(s,a) \hat{d}^k(s,a) \leq \xi_k, \quad k = 1, 2, \dots, K \right) \\ &= \mathcal{C}_\theta \left(\mathbb{P}(\rho^T \hat{d}^1 \leq \xi_1), \mathbb{P}(\rho^T \hat{d}^2 \leq \xi_2), \dots, \mathbb{P}(\rho^T \hat{d}^K \leq \xi_K) \right) \\ &\geq \mathcal{C}_\theta \left(p_1^{y_1^{\frac{1}{\theta}}}, p_1^{y_2^{\frac{1}{\theta}}}, \dots, p_1^{y_K^{\frac{1}{\theta}}} \right) \end{aligned}$$

$$= \exp \left\{ - \left[\sum_{k=1}^K (-\ln(p_1^{y_k^{\frac{1}{\theta}}}))^\theta \right]^{\frac{1}{\theta}} \right\} = p_1.$$

Hence, ρ is a feasible solution of (ii) of (4). Since, $\hat{d}^k \sim \mathcal{E}_{|\mathcal{K}|}(\mu_{\hat{d}^k}, \Sigma_{\hat{d}^k}, \varphi_{\hat{d}^k})$, $k = 1, 2, \dots, K$, it follows from Lemma 2.2 of [11] that the set of constraints in (8) can be equivalently written as

$$\begin{aligned} \rho^T \mu_{\hat{d}^k} + \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{\frac{1}{\theta}}}) \|\Sigma_{\hat{d}^k}^{\frac{1}{2}} \rho\|_2 &\leq \xi_k, \quad \forall k = 1, 2, \dots, K, \\ \sum_{k=1}^K y_k &= 1, \\ y_k &\geq 0, \quad \forall k = 1, 2, \dots, K. \end{aligned}$$

Thus, (4) is equivalent to (7). \square

Remark 1. When $K = 1$, the constraint (iii) of (7) gives $y_1 = 1$ and (7) reduces to the following SOCP problem

$$\begin{aligned} \min_{t, \rho} \quad & t \\ \text{s.t.} \quad & \rho^T \mu_{\hat{c}} + \Phi_{\hat{c}}^{-1}(p_0) \|\Sigma_{\hat{c}}^{\frac{1}{2}} \rho\|_2 \leq t, \\ & \rho^T \mu_{\hat{d}^1} + \Phi_{\hat{d}^1}^{-1}(p_1) \|\Sigma_{\hat{d}^1}^{\frac{1}{2}} \rho\|_2 \leq \xi_1, \\ & \rho \in \mathcal{Q}_v(\gamma). \end{aligned}$$

For $K \geq 2$, the optimization problem (7) is a non-convex optimization problem due to the product terms $\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{\frac{1}{\theta}}}) \|\Sigma_{\hat{d}^k}^{\frac{1}{2}} \rho\|_2$ present in its constraints. We approximate (7) 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 3. For each $k = 1, 2, \dots, K$, the quantile function $\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{\frac{1}{\theta}}})$ is a convex function of y_k .

For the spherical distributions listed in Table 1, Assumption 3 holds true. This follows from the fact that $\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{\frac{1}{\theta}}})$ is a composition of functions $p_1^{y_k^{\frac{1}{\theta}}}$ and $\Phi_{\hat{d}^k}^{-1}(y_k)$, 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 $\Phi_{\hat{d}^k}^{-1}(y_k)$ for the probability distributions listed in Table 1 can be verified by checking the following condition for all k ,

$$(\Phi_{\hat{d}^k}^{-1})''(y) = \frac{-\phi'_{\hat{d}^k}(\Phi_{\hat{d}^k}^{-1}(y))}{(\phi_{\hat{d}^k}(\Phi_{\hat{d}^k}^{-1}(y)))^3} \geq 0, \quad \forall y \in [p_1, 1),$$

where $\phi_{\hat{d}^k}(\cdot)$ is the probability density function of the spherical distribution whose cumulative distribution function is $\Phi_{\hat{d}^k}(\cdot)$. The convexity of $\Phi_{\hat{d}^k}^{-1}(y_k)$ for all 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 ($\phi_{\hat{d}^k}(u)$)
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	$\frac{ce^{-u^2}}{(1+e^{-u^2})^2}$
Laplace	$ce^{- u }$

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

approximate the functions $\Phi_{\hat{d}^k}^{-1}(p_1^{\frac{1}{\theta}})$, $k = 1, 2, \dots, K$, 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 (7).

3.2.1 Lower bound approximation

For every k , we approximate the term $\Phi_{\hat{d}^k}^{-1}(p_1^{\frac{1}{\theta}})$ in (7) by choosing N tangent points $y_k^1 < y_k^2 < \dots < y_k^N$ from the interval $(0, 1]$. We denote the Taylor series expansion around the point y_k^i by $\Phi_{\hat{d}^k}^{-1}(p_1^{\frac{1}{\theta}})_i$ and define it as

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

$$\text{where } \bar{a}_k^i = \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}}) - \frac{(y_k^i)^{\frac{1}{\theta}} p_1^{y_k^i \frac{1}{\theta}} \ln(p_1)}{\theta \phi_{\hat{d}^k}(\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}}))}, \bar{b}_k^i = \frac{p_1^{y_k^i \frac{1}{\theta}} \ln(p_1) (y_k^i)^{\frac{1}{\theta}-1}}{\theta \phi_{\hat{d}^k}(\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}}))}.$$

Theorem 2. *Let Assumption 3 holds. A lower bound for the optimal value of the equivalent JCCMDP problem (7) is given by the optimal value of the following SOCP problem*

$$\begin{aligned}
& \min_{t, \rho, (x_{kj})_{k=1, j=1}^{K, |\mathcal{K}|}, (z_{kj})_{k=1, j=1}^{K, |\mathcal{K}|}} t \\
\text{s.t. } & \rho^T \mu_{\hat{c}} + \Phi_{\hat{c}}^{-1}(p_0) \|\Sigma_{\hat{c}}^{\frac{1}{2}} \rho\|_2 \leq t, \\
& \rho^T \mu_{\hat{d}^k} + \|\Sigma_{\hat{d}^k}^{\frac{1}{2}} z_k\|_2 \leq \xi_k, \quad \forall k = 1, 2, \dots, K, \\
& z_{kj} \geq (\bar{a}_k^i \rho_j + \bar{b}_k^i x_{kj}), \quad \forall k = 1, 2, \dots, K, j = 1, 2, \dots, |\mathcal{K}|, i = 1, 2, \dots, N, \\
& \sum_{k=1}^K x_{kj} = \rho_j, \quad \forall j = 1, 2, \dots, |\mathcal{K}|, \\
& x_{kj} \geq 0, \quad \forall k = 1, 2, \dots, K, j = 1, 2, \dots, |\mathcal{K}|, \\
& \rho \in \mathcal{Q}_v(\gamma),
\end{aligned} \tag{9}$$

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 $k = 1, 2, \dots, K$.

Proof It is sufficient to show that for every feasible solution of (7) there exists a feasible solution of (9). The proof is similar to Theorem 3.1.1 of [5]. Suppose the vector $(t, \rho, (y_k)_{k=1}^K)$ is a feasible solution of (7). Under Assumption 3, we obtain $\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{\frac{1}{\theta}}}) \geq \max_{1 \leq i \leq N} (\bar{a}_k^i + \bar{b}_k^i y_k)$ for all k . Therefore,

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

The vectors $z_k, k = 1, 2, \dots, K$, where $z_{kj} = \max_{1 \leq i \leq N} (\bar{a}_k^i \rho_j + \bar{b}_k^i y_k \rho_j)$ for all $k = 1, 2, \dots, K, j = 1, 2, \dots, |\mathcal{K}|$, satisfy

$$\begin{aligned}
& \rho^T \mu_{\hat{d}^k} + \|\Sigma_{\hat{d}^k}^{\frac{1}{2}} z_k\|_2 \leq \xi_k, \quad \forall k = 1, 2, \dots, K, \\
& z_{kj} \geq \bar{a}_k^i \rho_j + \bar{b}_k^i y_k \rho_j, \quad \forall k = 1, 2, \dots, K, j = 1, 2, \dots, |\mathcal{K}|, i = 1, 2, \dots, N.
\end{aligned}$$

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

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

is a feasible solution for the problem (9). \square

3.2.2 Upper bound approximation

For every k , we approximate the term $\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{\frac{1}{\theta}}})$ in (7) by choosing N interpolation points $y_k^1 < y_k^2 < \dots < y_k^N$ from the interval $(0, 1]$, such that the point y_k^1

is sufficiently close to 0. We denote the linear interpolating polynomial passing through the points $(y_k^i, \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}}))$ and $(y_k^{i+1}, \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{i+1} \frac{1}{\theta}}))$ by $\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}})_i$ and define it as

$$\begin{aligned} \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}})_i &= \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}}) + \frac{y_k - y_k^i}{y_k^{i+1} - y_k^i} \left(\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{i+1} \frac{1}{\theta}}) - \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}}) \right) \\ &= \frac{y_k^{i+1} \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}}) - y_k^i \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{i+1} \frac{1}{\theta}})}{y_k^{i+1} - y_k^i} + \frac{(\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{i+1} \frac{1}{\theta}}) - \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}})) y_k}{y_k^{i+1} - y_k^i} \\ &= \hat{a}_k^i + \hat{b}_k^i y_k, \quad \forall i = 1, 2, \dots, N-1, \end{aligned}$$

$$\text{where } \hat{a}_k^i = \frac{y_k^{i+1} \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}}) - y_k^i \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{i+1} \frac{1}{\theta}})}{y_k^{i+1} - y_k^i}, \quad \hat{b}_k^i = \frac{\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^{i+1} \frac{1}{\theta}}) - \Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}})}{y_k^{i+1} - y_k^i}.$$

Theorem 3. *Let Assumption 3 holds. An upper bound for the optimal value of the equivalent JCCMDP problem (7) is given by the optimal value of the following SOCP problem*

$$\begin{aligned} & \min_{t, \rho, (y_k)_{k=1}^K, (z_k)_{k=1}^K} t \\ \text{s.t. } & \rho^T \mu_{\hat{c}} + \Phi_{\hat{c}}^{-1}(p_0) \|\Sigma_{\hat{c}}^{\frac{1}{2}} \rho\|_2 \leq t, \\ & \rho^T \mu_{\hat{d}^k} + z_k V_k^* \leq \xi_k, \quad \forall k = 1, 2, \dots, K, \\ & z_k \geq \hat{a}_k^i + \hat{b}_k^i y_k, \quad \forall k = 1, 2, \dots, K, i = 1, 2, \dots, N-1, \\ & \sum_{k=1}^K y_k = 1, \\ & y_k \geq 0, \quad \forall k = 1, 2, \dots, K, \\ & \rho \in \mathcal{Q}_v(\gamma), \end{aligned} \tag{10}$$

where ρ_j is the j -th component of $\rho \in \mathbb{R}^{|\mathcal{K}|}$, $(\Sigma_{\hat{d}^k}^{\frac{1}{2}})_j$ is the j -th column of the matrix $\Sigma_{\hat{d}^k}^{\frac{1}{2}}$ and, $V_k^* = \max_{\bar{\rho} \in \mathcal{Q}_v(\gamma)} \sum_{j=1}^{|\mathcal{K}|} \bar{\rho}_j \|(\Sigma_{\hat{d}^k}^{\frac{1}{2}})_j\|_2$ for every $k = 1, 2, \dots, K$.

Proof We show that for every feasible solution of (10) there exists a feasible solution of (7). Suppose the vector $(t, \rho, (y_k)_{k=1}^K, (z_k)_{k=1}^K)$ is a feasible solution of (10). It is enough to show that ρ is a feasible solution of constraint (ii) of (7). Under Assumption 3, we obtain $\Phi_{\hat{d}^k}^{-1}(p_1^{y_k^i \frac{1}{\theta}}) \leq \max_{1 \leq i \leq N-1} (\hat{a}_k^i + \hat{b}_k^i y_k)$ for all k . For an arbitrarily fixed

$k \in \{1, 2, \dots, K\}$, consider

$$\begin{aligned} \|\Sigma_{\hat{d}^k}^{\frac{1}{2}} \rho\|_2 &= \|\rho_1(\Sigma_{\hat{d}^k}^{\frac{1}{2}})_1 + \rho_2(\Sigma_{\hat{d}^k}^{\frac{1}{2}})_2 + \dots + \rho_{|\mathcal{K}|}(\Sigma_{\hat{d}^k}^{\frac{1}{2}})_{|\mathcal{K}|}\|_2 \\ &\leq \rho_1\|(\Sigma_{\hat{d}^k}^{\frac{1}{2}})_1\|_2 + \rho_2\|(\Sigma_{\hat{d}^k}^{\frac{1}{2}})_2\|_2 + \dots + \rho_{|\mathcal{K}|}\|(\Sigma_{\hat{d}^k}^{\frac{1}{2}})_{|\mathcal{K}|}\|_2 \\ &\leq \max_{\bar{\rho} \in \mathcal{Q}_v(\gamma)} \sum_{j=1}^{|\mathcal{K}|} \bar{\rho}_j \|(\Sigma_{\hat{d}^k}^{\frac{1}{2}})_j\|_2 = V_k^*. \end{aligned}$$

The first inequality holds by triangular inequality of $\|\cdot\|_2$, while the second inequality holds because $\rho \in \mathcal{Q}_v(\gamma)$. Thus, V_k^* is an upper bound of $\|\Sigma_{\hat{d}^k}^{\frac{1}{2}} \rho\|_2$. Consequently, the following inequality holds true

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

Therefore,

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

Hence, the vector $(t, \rho, (y_k)_{k=1}^K)$ is a feasible solution for the problem (7). \square

4 Numerical experiments

We perform numerical experiments using MOSEK solver of 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. We study our approximations on a queueing control problem [1] in Section 4.1, and on randomly generated CMDPs in Section 4.2. In all our numerical experiments, we fix $p_0, p_1 = 0.95$, $\alpha = 0.99$, and assume γ to be a uniformly distributed vector. For our approximations, we vary N , pick uniformly spaced points between 10^{-5} and 1 and keep them as both tangent and interpolation points for all the constraints. We note that the interpolation point y_k^1 is sufficiently close to 0.

4.1 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, $S = \{0, 1, \dots, L\}$, 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 $A^1 \subseteq [a_{min}^1, a_{max}^1]$ and $A^2 \subseteq [a_{min}^2, a_{max}^2]$, where $0 < a_{min}^1 \leq a_{max}^1 < 1$ and $0 \leq a_{min}^2 \leq a_{max}^2 < 1$. 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 $\hat{c} \sim \mathcal{N}(\mu_{\hat{c}}, \Sigma_{\hat{c}})$, $\hat{d}^k \sim \mathcal{N}(\mu_{\hat{d}^k}, \Sigma_{\hat{d}^k})$, $k = 1, 2$, where $\hat{c} = (\hat{c}(s))_{s \in S}$, $\hat{d}^1 = (\hat{d}^1(a^1))_{a^1 \in A^1}$ and $\hat{d}^2 = (\hat{d}^2(a^2))_{a^2 \in A^2}$. We consider the case where the number of states is 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 \hat{c}, \hat{d}^2 to be normally distributed random vectors and \hat{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 $\xi_2 = 9$, $\mu_{\hat{c}} = (0, 1, 2, \dots, 9)^T$, the matrix $\Sigma_{\hat{c}} \in \mathbb{R}^{10 \times 10}$ with all diagonal values 0.9 and off-diagonal values 0.35, and

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

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 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

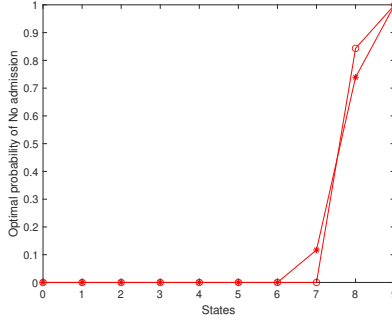


Fig. 1 The solid lines marked with '*' and 'o' denote the optimal probability of not admitting a customer into the queue for discounted and average cost problems, respectively.

lower bound and upper bound for the optimal value of the JCCMDP problem (7) by solving the SOCP problems (9) and (10), respectively. For $L = n - 1$ ($|S| = n$), we take $\mu_{\hat{c}} = (0, 1, 2, \dots, n - 1)^T$, the matrix $\Sigma_{\hat{c}} \in \mathbb{R}^{n \times n}$ with all diagonal values 0.9 and off-diagonal values 0.35, $\mu_{\hat{d}^1} = (4.32, 9.1875, 10.83)^T$, $\mu_{\hat{d}^2} = (10.00, 8.50, 7.60)^T$, and

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

The mean vector $\mu_{\hat{c}}$, where $\mu_{\hat{c}}(s) = s$, captures the fact that the holding cost increases with s . We take the mean vectors $\mu_{\hat{d}^1}$ and $\mu_{\hat{d}^2}$ such that $\mu_{\hat{d}^1}(a^1) = 3(1 + a^1)^2$ and $\mu_{\hat{d}^2}(a^2) = 10 - 3a^2$. It is motivated from the fact that the service cost increases with a^1 and no admission cost decreases with a^2 . We take $\xi_1 = 11.30$, $\xi_2 = 11.35$. We solve the SOCP problems (9) and (10) for different values of θ and $|S|$. The higher values of θ indicate stronger dependence between the random constraints 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. In the average cost criterion, for all states except $s = 0, 1$, the corresponding ρ becomes significantly low, such that it has a negligible effect on the optimal value. Thus, the optimal value remains almost the same with the number of states. We define the gap between the lower and upper bounds for the optimal value of the JCCMDP problem (4) by

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

From Tables 2 and 3, we observe that the gap ranges from 0 to 1.3565. The CPU time analysis shows that the SOCP problems (9) and (10) can be solved efficiently for our example.

Table 2 Queuing control problem: Discounted cost criterion.

θ States	N = 5			N = 10			N = 20									
	Lower Bound (LB)		Upper Bound (UB)	Lower Bound (LB)		Upper Bound (UB)	Lower Bound (LB)		Upper Bound (UB)	Gap(%)						
	Opt.	CPU(s)	Opt.	CPU(s)	Opt.	CPU(s)	Opt.	CPU(s)	Opt.	CPU(s)						
1	500	177.4043	68.25	178.639	11.8281	0.696	177.4071	75.7813	178.5013	12.9375	0.6168	177.408	97.7344	178.3905	11.9219	0.5538
	1000	422.0911	342.1875	427.8169	31.6875	1.3565	422.2185	353.9531	427.266	27.5938	1.1955	422.226	412.1875	427.2463	39.5156	1.189
	1500	671.2592	943.6094	677.8156	56.7656	0.9767	671.4757	877.5938	677.2248	62.8438	0.8562	671.5145	980.75	677.2431	65.8281	0.8531
	2000	920.8955	1908.6	927.8149	85.8281	0.7514	921.2717	2004.6	927.2242	87.7813	0.6461	921.3652	2318.6	927.2425	111.9219	0.6379
3	500	177.3957	63.2188	177.3996	9.9063	0.0022	177.3955	69.7656	177.3988	11.125	0.0019	177.3959	83.4531	177.3984	11.4688	0.0014
	1000	420.2786	299.2813	423.1495	33.875	0.6831	420.2828	310.9688	423.1099	28.7031	0.6727	420.2915	345.9844	423.0939	32.1719	0.6668
	1500	669.3847	829.7344	673.0587	58.5	0.5489	669.4349	892.6406	672.9414	62.3594	0.5238	669.4366	951.6875	672.9065	72.1094	0.5183
	2000	919.0816	1809.2	923.0578	96.6875	0.4326	919.1713	1930.1	922.9407	110.0625	0.4101	919.1847	2105.8	922.9059	104.6875	0.4048

Table 3 Queuing control problem: Average cost criterion.

θ States	N = 5			N = 10			N = 20									
	Lower Bound (LB)		Upper Bound (UB)	Lower Bound (LB)		Upper Bound (UB)	Lower Bound (LB)		Upper Bound (UB)	Gap(%)						
	Opt.	CPU(s)	Opt.	CPU(s)	Opt.	CPU(s)	Opt.	CPU(s)	Opt.	CPU(s)						
1	500	1.5606	73.75	1.5607	16.6094	0.0064	1.5606	77.2188	1.5607	14.7656	0.0064	1.5606	92.7031	1.5607	14.2188	0.0064
	1000	1.5606	331.9063	1.5607	33.8594	0.0064	1.5606	353.1406	1.5607	32.7656	0.0064	1.5606	412.4063	1.5607	31.875	0.0064
	1500	1.5606	980.5313	1.5607	71.875	0.0064	1.5606	927.5313	1.5606	103.5938	0	1.5606	975.5781	1.5607	72.125	0.0064
	2000	1.5606	1776.8	1.5606	272.3438	0	1.5605	2195.8	1.5607	119.5313	0.0128	1.5606	2231.2	1.5606	136	0
3	500	1.5606	65.2188	1.5607	12.75	0.0064	1.5606	68.7969	1.5607	12.6563	0.0064	1.5606	84.3906	1.5607	11.3125	0.0064
	1000	1.5606	294	1.5607	30.5938	0.0064	1.5606	320.2656	1.5607	33.0625	0.0064	1.5606	346.2656	1.5607	31.6406	0.0064
	1500	1.5606	882.6875	1.5607	104.1563	0.0064	1.5606	933.2031	1.5607	88.8125	0.0064	1.5606	941	1.5607	68.625	0.0064
	2000	1.5606	1973.1	1.5606	176.5625	0	1.5605	1948.7	1.5606	124.0781	0.0064	1.5605	2099	1.5606	136.4219	0.0064

4.2 Random CMDPs

We generate random finite CMDPs with normally distributed cost vectors to study the efficiency of our proposed methods. Previous literature has conducted experiments on randomly generated finite MDPs known as Garnets [2, 8]. We construct our system similar to the description in [8], using the parameter tuple $(|S|, |A|, |B_f|)$, where $|S|$ and $|A|$ are the number of states and actions, respectively, while $|B_f|$ is the branching factor that indicates the number of states reachable from a given state-action pair. For each state-action pair (s, a) , we draw $|B_f|$ states using the function $\mathbf{sort}(\mathbf{randsample}(|S|, |B_f|))$ and denote the corresponding states as $(s^i)_{i=1}^{|B_f|}$. To fix the non-zero probability values for these states, we randomly generate $|B_f| - 1$ values in the interval $(0, 1)$ using the function $\mathbf{sort}(\mathbf{rand}(|B_f| - 1, 1))$ and denote them as $(q^i)_{i=1}^{|B_f|-1}$. We take $q^0 = 0$ and $q^{|B_f|} = 1$. Following the rule given in Section 4.1 of [8], the transition probabilities are given by

$$p(s'|s, a) = \begin{cases} q^i - q^{i-1} & \text{if } s' = s^i, i = 1, 2, \dots, |B_f|, \\ 0 & \text{otherwise.} \end{cases}$$

For simplicity, we generate only irreducible transition probability matrices. This ensures that Assumption 1 is trivially satisfied. We fix $\theta = 6$, $|S| = 500$,

and $|A|, K = 10$. Thus, $\hat{c} \sim \mathcal{N}(\mu_{\hat{c}}, \Sigma_{\hat{c}})$, $\hat{d}^k \sim \mathcal{N}(\mu_{\hat{d}^k}, \Sigma_{\hat{d}^k})$, $k = 1, 2, \dots, 10$ are random vectors of dimension 5000×1 . We randomly generate $\mu_{\hat{c}}$ in $(30, 200)$, $\mu_{\hat{d}^k}$ in $(30, 150)$, and ξ_k in $(75, 120)$ for $k = 1, 2, \dots, 10$ using the function **rand**. We generate the square roots of $\Sigma_{\hat{c}}, \Sigma_{\hat{d}^k}, k = 1, 2, \dots, 10$, each of dimension 5000×5000 using the function **sprandsym**(5000, 0.0005, **3rand**(5000, 1)). The generated matrices are symmetric with density 0.0005 and randomly generated eigenvalues in the interval $(0, 3)$. We consider two cases: (i) $|B_f| = 100$ (20% of the states are reachable from each state-action pair), (ii) $|B_f| = 400$ (80% of the states are reachable from each state-action pair). We solve the SOCP problems (9) and (10), and summarize the results in Tables 4 and 5 for the case of discounted and average cost criteria, respectively.

In both the tables, the average gap between the lower and upper bounds reduces as N increases, irrespective of $|B_f|$. Furthermore, the average gaps for $|B_f| = 100$ are larger than the corresponding gaps for $|B_f| = 400$, while the gaps obtained for the discounted cost criterion are smaller than the corresponding gaps for the average cost criterion. For the data we generate, the gap ranges from 1.5155 to 12.9035. This indicates that gaps obtained are data-dependent. The average CPU time to solve the problem (10) is less than (9), possibly since the latter problem has K additional second order constraints. Furthermore, on average it takes more time to solve the problems with $|B_f| = 400$ than with $|B_f| = 100$ and more time to solve the discounted cost criterion than its average cost criterion counterpart. However, we conclude that both our approximations can be solved efficiently for both the cost criteria.

5 Conclusion

We consider a JCCMDP problem under discounted and average cost criteria. 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 lower and upper bounds on the optimal value of the JCCMDP problem. We illustrate our results on queueing control problem and on a class of randomly generated CMDP problems by considering various small and large instances. The gap between the lower and upper bounds shows the effectiveness of the SOCP approximations. The CPU 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.

Table 4 Random CMDPs: Discounted cost criterion.

Instance	$ B_j = 100$									$ B_j = 400$								
	$N = 5$			$N = 10$			$N = 20$			$N = 5$			$N = 10$			$N = 20$		
	LB	UB	Gap(%)	LB	UB	Gap(%)	LB	UB	Gap(%)	LB	UB	Gap(%)	LB	UB	Gap(%)	LB	UB	Gap(%)
1	45.6958	46.5055	1.7719	45.6958	46.4991	1.7579	45.6958	46.498	1.7555	45.6793	46.5135	1.8262	45.6793	46.5073	1.8126	45.6793	46.5063	1.8104
2	46.9667	48.8036	3.9111	46.9667	48.8026	3.9089	46.9668	48.7993	3.9017	47.0762	48.9294	3.9366	47.0762	48.928	3.9336	47.0763	48.9278	3.9343
3	47.5001	50.7706	6.8852	47.5002	50.6258	6.5802	47.5002	50.5977	6.521	47.7092	50.8135	6.5067	47.7093	50.6863	6.2309	47.7094	50.6608	6.1862
4	46.9037	50.0895	6.7922	46.9039	50.0866	6.7856	46.904	50.086	6.7841	46.9617	50.1586	6.8075	46.962	50.1543	6.7976	46.9621	50.1526	6.7938
5	48.4094	51.6753	6.7464	48.4095	51.67	6.7352	48.4095	51.6686	6.7324	48.4266	51.6696	6.6967	48.4267	51.6685	6.6942	48.4268	51.6661	6.6891
6	50.2195	55.6935	10.9001	50.2199	55.6925	10.8973	50.2201	55.6895	10.8909	50.4922	56.0198	10.9474	50.4926	56.0186	10.9442	50.4927	56.0169	10.9406
7	46.8407	48.0525	2.5871	46.8407	48.041	2.5625	46.8407	48.0407	2.5619	46.8432	48.0463	2.5684	46.8432	48.0352	2.5447	46.8432	48.0351	2.5444
8	47.8961	51.8746	8.3065	47.8964	51.8643	8.2843	47.8965	51.8631	8.2816	47.8878	51.8113	8.1931	47.8881	51.8008	8.1705	47.8882	51.7989	8.1663
9	45.9614	47.0868	2.4486	45.9614	47.0773	2.4279	45.9615	47.0736	2.4196	46.1181	47.2549	2.465	46.1181	47.245	2.4435	46.1181	47.2416	2.4361
10	47.9666	52.4435	9.3334	47.9668	52.4081	9.2591	47.9668	52.3971	9.2362	48.0967	52.5019	9.159	48.0969	52.4673	9.0867	48.097	52.4564	9.0638
11	47.3897	50.392	6.3353	47.39	50.3697	6.2876	47.39	50.3565	6.2598	47.5184	50.5069	6.2891	47.5186	50.484	6.2405	47.5187	50.4695	6.2098
12	47.6423	51.1958	7.4587	47.6425	50.9342	6.9092	47.6425	50.9223	6.8842	47.6322	51.2244	7.5415	47.6324	50.9769	7.0215	47.6324	50.9651	6.9967
13	44.6431	45.4249	1.7512	44.6431	45.4087	1.7149	44.6431	45.4058	1.7084	44.6522	45.3395	1.5392	44.6522	45.3318	1.522	44.6522	45.3289	1.5155
14	47.0049	49.3541	4.9978	47.005	49.3476	4.9837	47.0051	49.3482	4.9848	47.1141	49.5231	5.1131	47.1141	49.5162	5.0985	47.1141	49.516	5.098
15	48.7016	52.5524	7.9069	48.7018	52.5262	7.8527	48.7019	52.525	7.85	48.9692	52.8337	7.8917	48.9694	52.8105	7.8439	48.9694	52.8071	7.8369
16	49.9025	54.1964	8.6046	49.9028	54.1793	8.5697	49.9029	54.1771	8.565	49.9596	54.3096	8.707	49.9599	54.293	8.6732	49.9601	54.2906	8.6679
17	46.5361	48.3116	3.8153	46.5362	48.2901	3.7689	46.5361	48.2849	3.7579	46.6312	48.5094	4.0278	46.6312	48.4822	3.9604	46.6312	48.4742	3.9523
18	49.9315	56.3042	12.7629	49.9319	56.2062	12.5657	49.932	56.1629	12.4788	49.9556	56.4016	12.9035	49.9561	56.2919	12.6827	49.9563	56.251	12.6004
19	47.641	49.8784	4.6964	47.6411	49.856	4.6491	47.6411	49.8535	4.6439	47.6382	49.8758	4.6971	47.6383	49.8497	4.6421	47.6383	49.8488	4.6402
20	47.2805	51.1291	8.1399	47.2808	50.7363	7.3085	47.2808	50.7212	7.2765	47.1849	50.9132	7.9015	47.1851	50.5678	7.169	47.1852	50.5531	7.1376
Av. CPU time	924.9644	65.6149	-	945.365	66.5758	-	1101.2228	76.8516	-	959.7674	82.7281	-	992.1094	80.3164	-	1141.3742	92.493	-
Av. Gap	-	-	6.3076	-	-	6.1904	-	-	6.1747	-	-	6.2859	-	-	6.1765	-	-	6.161

The last two rows of the table give the average CPU time taken to solve the SOCP problems (against the columns with optimal values) and the average gap between the lower and upper bounds (against the columns with Gap(%)).

Conflict of Interest

The authors declare that there is no conflict of interest.

References

- [1] Altman, E.: *Constrained Markov Decision Processes*. Chapman and Hall/CRC, London (1999)
- [2] Archibald, T.W., McKinnon, K.I.M., Thomas, L.C.: On the generation of Markov decision processes. *Journal of the Operational Research Society* **46**(3), 354–361 (1995)
- [3] Bertsekas, D.P.: *Dynamic Programming and Optimal Control*, 2nd edn. Athena Scientific, Belmont, MA, USA (2000)
- [4] Charnes, A., Cooper, W.W.: Chance-constrained programming. *Management Science* **6**, 73–79 (1959)
- [5] Cheng, J., Lissner, A.: A second-order cone programming approach for linear programs with joint probabilistic constraints. *Operations Research*

Table 5 Random CMDPs: Average cost criterion.

Instance	$ B_j = 100$									$ B_j = 400$								
	$N = 5$			$N = 10$			$N = 20$			$N = 5$			$N = 10$			$N = 20$		
	LB	UB	Gap(%)	LB	UB	Gap(%)	LB	UB	Gap(%)	LB	UB	Gap(%)	LB	UB	Gap(%)	LB	UB	Gap(%)
1	45.695	46.5047	1.772	45.695	46.4982	1.7577	45.695	46.4971	1.7553	45.6792	46.5134	1.8262	45.6792	46.5071	1.8124	45.6792	46.5061	1.8102
2	46.9646	48.8013	3.9108	46.9646	48.8002	3.9085	46.9646	48.8002	3.9085	47.0759	48.9296	3.9377	47.076	48.9282	3.9345	47.076	48.928	3.9341
3	47.4974	50.7698	6.8896	47.4975	50.6238	6.582	47.4976	50.5957	6.5226	47.7091	50.8134	6.5067	47.7092	50.6896	6.2407	47.7092	50.6622	6.1896
4	46.9017	50.0882	6.794	46.9019	50.0852	6.7871	46.902	50.0844	6.7852	46.9612	50.1588	6.809	46.9615	50.1545	6.7992	46.9615	50.1546	6.7994
5	48.4008	51.6737	6.7488	48.407	51.6704	6.7416	48.407	51.6679	6.7364	48.4222	51.6657	6.6984	48.4224	51.6646	6.6957	48.4224	51.6624	6.6911
6	50.2145	55.6889	10.902	50.2149	55.6871	10.8976	50.215	55.6847	10.8926	50.4912	56.019	10.948	50.4916	56.0174	10.944	50.4918	56.0161	10.941
7	46.84	48.0515	2.5865	46.84	48.0402	2.5623	46.84	48.0403	2.5626	46.8431	48.046	2.5679	46.8431	48.0348	2.544	46.8431	48.0349	2.5442
8	47.8958	51.8755	8.3091	47.8961	51.8654	8.2873	47.8962	51.864	8.2842	47.8883	51.8122	8.1939	47.8886	51.8019	8.1717	47.8887	51.8002	8.1679
9	45.9592	47.0843	2.448	45.9592	47.0745	2.4267	45.9592	47.0715	2.4202	46.1179	47.2548	2.4652	46.1179	47.2449	2.4437	46.1179	47.2418	2.437
10	47.9638	52.4409	9.3343	47.964	52.4053	9.2977	47.9641	52.3991	9.2361	48.0961	52.5021	9.1608	48.0963	52.4671	9.0876	48.0963	52.4559	9.0643
11	47.3877	50.3894	6.3343	47.3879	50.3672	6.287	47.388	50.3556	6.2623	47.518	50.5066	6.2894	47.5182	50.4837	6.2408	47.5183	50.4718	6.2155
12	47.6421	51.1962	7.46	47.6422	50.9347	6.9109	47.6423	50.9229	6.8859	47.6325	51.2258	7.5438	47.6327	50.9782	7.0235	47.6327	50.9677	7.0015
13	44.6419	45.4248	1.7537	44.6419	45.4084	1.717	44.6419	45.4053	1.7101	44.6519	45.3393	1.5395	44.652	45.3315	1.5218	44.652	45.3295	1.5173
14	47.0033	49.3519	4.9967	47.0034	49.3461	4.9841	47.0035	49.3457	4.9830	47.1142	49.5244	5.1157	47.1143	49.5164	5.0985	47.1143	49.5162	5.098
15	48.698	52.5481	7.9061	48.6982	52.5228	7.8537	48.6983	52.5203	7.8483	48.9692	52.8335	7.8913	48.9694	52.8103	7.8435	48.9694	52.8079	7.8386
16	49.8999	54.1923	8.602	49.9003	54.1771	8.5707	49.9004	54.1729	8.5621	49.9587	54.3092	8.7082	49.9591	54.2925	8.6739	49.9592	54.29	8.6687
17	46.5338	48.3091	3.8151	46.5339	48.2887	3.771	46.5339	48.2823	3.7573	46.631	48.5096	4.0287	46.631	48.4824	3.9703	46.631	48.4744	3.9532
18	49.9295	56.3007	12.7604	49.93	56.2018	12.5612	49.9301	56.1594	12.476	49.955	56.4006	12.9028	49.9555	56.2931	12.6865	49.9556	56.25	12.6
19	47.6405	49.8774	4.6954	47.6406	49.8552	4.6486	47.6406	49.8525	4.6429	47.6382	49.8756	4.6967	47.6383	49.8495	4.6416	47.6384	49.8485	4.6393
20	47.2795	51.129	8.142	47.2797	50.7361	7.3105	47.2798	50.7216	7.2796	47.1843	50.9124	7.9011	47.1845	50.5679	7.1706	47.1845	50.5533	7.1396
Av. CPU time	736.1953	45.7406	-	823.9044	46.9141	-	955.171	54.1117	-	840.7188	59.4695	-	899.7156	62.1219	-	1019.4791	68.4539	-
Av. Gap	-	-	6.308	-	-	6.1933	-	-	6.1756	-	-	6.2865	-	-	6.1772	-	-	6.1625

The last two rows of the table give the average CPU time taken to solve the SOCP problems (against the columns with optimal values) and the average gap between the lower and upper bounds (against the columns with Gap(%)).

Letters, 325–328 (2012)

- [6] Cheng, J., Houda, M., Lisser, A.: Chance constrained 0–1 quadratic programs using copulas. *Optimization Letters* **9**(7), 1283–1295 (2015)
- [7] Delage, E., Mannor, S.: Percentile optimization for Markov decision processes with parameter uncertainty. *Operations Research* **58**(1), 203–213 (2010)
- [8] El Asri, L., Piot, B., Geist, M., Laroche, R., Pietquin, O.: Score-based inverse reinforcement learning. In: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems. AAMAS '16*, pp. 457–465. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2016)
- [9] Fang, K.-T., Kotz, S., Ng, K.W.: *Symmetric Multivariate and Related Distributions*. Chapman and Hall/CRC, London, New York (1990)
- [10] Geng, X., Xie, L.: Data-driven decision making in power systems with probabilistic guarantees: Theory and applications of chance-constrained optimization. *Annual Reviews in Control* **47** (2019)

- [11] Henrion, R.: Structural properties of linear probabilistic constraints. *Optimization* **56**, 425–440 (2007)
- [12] Iyengar, G.N.: Robust dynamic programming. *Mathematics of Operations Research* **30**(2), 257–280 (2005)
- [13] 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)
- [14] Luedtke, J., Ahmed, S., Nemhauser, G.: An integer programming approach for linear programs with probabilistic constraints. *Math. Program.* **122**(2), 247–272 (2010)
- [15] Mannor, S., Simester, D., Sun, P., Tsitsiklis, J.N.: Bias and variance approximation in value function estimates. *Management Science* **53**(2), 308–322 (2007)
- [16] Nelsen, R.B.: *An Introduction to Copulas* (Springer Series in Statistics). Springer, Berlin, Heidelberg (2006)
- [17] Nilim, A., El Ghaoui, L.: Robust control of Markov decision processes with uncertain transition matrices. *Operations Research* **53**(5), 780–798 (2005)
- [18] Prékopa, A.: *Stochastic Programming*. Kluwer Academic Publishers, Berlin (1995)
- [19] Puterman, M.L.: *Markov Decision Process*, 1st edn. John Wiley & Sons, USA (1994)
- [20] Satia, J.K., Lave Jr, R.E.: Markovian decision processes with uncertain transition probabilities. *Operations Research* **21**(3), 728–740 (1973)
- [21] Sklar, M.: Fonctions de répartition à n dimensions et leurs marges. *Publ. inst. statist. univ. Paris* **8**, 229–231 (1959)
- [22] Varagapriya, V., Singh, V.V., Lisser, A.: Constrained Markov decision processes with uncertain costs. *Operations Research Letters* **50**(2), 218–223 (2022)
- [23] White, C.C., Eldeib, H.K.: Markov decision processes with imprecise transition probabilities. *Operations Research* **42**(4), 739–749 (1994)
- [24] Wiesemann, W., Kuhn, D., Rustem, B.: Robust Markov decision processes. *Mathematics of Operations Research* **38**(1), 153–183 (2013)
- [25] Ye, Y.: The simplex and policy-iteration methods are strongly polynomial

for the Markov decision problem with a fixed discount rate. *Mathematics of Operations Research* **36**(4), 593–603 (2011)