Second-order cone programming approach for elliptically distributed joint probabilistic constraints with dependent rows

Jianqiang Cheng∗ Michal Houda† Abdel Lisser‡
May 21, 2014

Abstract
In this paper, we investigate the problem of linear joint probabilistic constraints. We assume that the rows of the constraint matrix are dependent and the dependence is driven by a convenient Archimedean copula. Further we assume the distribution of the constraint rows to be elliptically distributed, covering normal, t, or Laplace distributions. Under these and some additional conditions, we prove the convexity of the investigated set of feasible solutions. We also develop two approximation schemes for this class of stochastic programming problems based on second-order cone programming, which provides lower and upper bounds. Finally, a computational study on randomly generated data is given to illustrate the tightness of these bounds.

Keywords: chance constrained programming; Archimedean copulas; elliptical distributions; convexity; second-order cone programming.

1 Introduction

Consider a linear optimization problem with uncertainty

\[
\begin{align*}
\min \ c^T x & \quad \text{subject to} \quad Tx \leq h, \ x \in X,
\end{align*}
\]

where \(X \subseteq \mathbb{R}^n\) is a deterministic closed convex set, \(c \in \mathbb{R}^n\), \(h = (h_1, \ldots, h_K)^T \in \mathbb{R}^K\) deterministic vectors, and \(T \in \mathbb{R}^{K \times n}\) is an uncertain (unknown) matrix with rows \(t_1, \ldots, t_K\); \(n, K\) are structural elements of the optimization problem (1). If \(X\) is polyhedral, and a realization of the data matrix \(T\) is known and fixed in advance, (1) is simply a linear programming problem. To deal with uncertainty of the data matrix, various methods were developed in optimization theory: classical sensitivity analysis, parametric programming, or robust optimization methods. We concentrate on stochastic chance-constrained programming approach. We assume that \(T\) is a random matrix with a known distribution. Furthermore, the uncertain constraints of the problem (1) are required to be satisfied with a known fixed probability level \(p \in [0; 1]\). The linear chance-constrained problem with random matrix reads as

\[
\begin{align*}
\min \ c^T x & \quad \text{subject to} \quad \mathbb{P}\{Tx \leq h\} \geq p, \ x \in X.
\end{align*}
\]

Note that the restriction on the random matrix only is not crucial. Indeed, if \(h\) is also random, one can introduce a new decision vector \(x' := (x^T, 1)^T\) and a random matrix \(T' = (T, -h)\); the new uncertain constraint takes the form of \(T'x \leq 0\) which falls within the introduced frame.

*LRI, Université Paris Sud, 91405, Orsay, France (cheng@lri.fr)
†LRI, Université Paris Sud, 91405, Orsay, France (houda@efjcu.cz).
‡LRI, Université Paris Sud, 91405, Orsay, France (lisser@lri.fr)
Denote $X(p)$ the feasible set of (2), i.e.,

$$X(p) := \{ x \in X \mid \mathbb{P}(Tx \leq h) \geq p \}.$$  

(3)

One of the main points of our investigation is to find equivalent or approximative formulation of $X(p)$ suitable for theoretical and numerical purposes – in particular, formulations carrying suitable convexity properties. Our approach is based on exploring convexity properties of another type of chance-constrained problems, namely nonlinear chance-constrained problems with random right-hand side

$$\min c^T x \quad \text{subject to} \quad \mathbb{P}\{ x \in X \mid g_k(x) \geq \xi, \ k = 1, \ldots, K \} \geq p,$$

(4)

where $\xi := (\xi_1, \ldots, \xi_K)$ is a random vector with known distribution, and $g_k(\cdot)$ are deterministic continuous functions. For notational purposes we define the feasible set of problem (4) by

$$M(p) := \{ x \in X \mid \mathbb{P}\{ g_k(x) \geq \xi, \ k = 1, \ldots, K \} \geq p \};$$

(5)

later we also provide additional conditions on $\xi$ and $g_k$ in order to obtain desired results.

1.1 Survey of literature

A problem involving probability constraints was first formulated by Charnes et al. [1958] and developed by author’s subsequent papers. Since these earliest results, it was recognized that these problems of probabilistically (or chance) constrained programming are hard to treat, both from theoretical and computational point of view. Van de Panne and Popp [1963] proposed a solution method for a problem of type (2) with one-row normally distributed constraint, transformed to a nonlinear constraint similar to (16). At the same time, Kataoka [1963] investigated a problem with normally distributed individual probabilistic constraints with random right-hand side; in the discussion he noticed that constraints with random matrices (but still individual) are also covered by his approach.

Convexity is widely considered as a considerable difficulty when investigating chance-constrained problems. Apart from simple problems presented above, chance constrained problems often lead to a feasible solution set which is not convex. Various techniques and conditions were developed to encompass this issue. As an introducing citation: a model with joint probabilistic constraints with independent random right-hand side was treated by Miller and Wagner [1965]. The convexity of their problem is ensured if the probability distribution possesses a property of decreasing reversed hazard function (increasing hazard function for their maximization form of the problem). Jagannathan [1974] extended the result to the dependent case, and considered also the case of random constraint matrix with normally distributed independent rows.

The essential step was made by Prékopa [1971] introducing the notion of logarithmically concave probability measure. He proved a general theorem which allowed him to introduce many convenient probability distributions (multivariate normal, Wishart, beta, Dirichlet), and to derive convexity for feasible sets of problems with dependent random right-hand side following these probability laws. The concept was further generalized by Borell [1975] and Brascamp and Lieb [1976] to $r$-concave ($\alpha$-concave) measures and functions (namely densities and distribution functions). A generalized definition of $r$-concave function on a set, suitable also for discrete distributions, was proposed by Dentcheva et al. [2000]. We refer to Prékopa [1995], Prékopa [2003] (Chapter 5 in Ruszczyński and Shapiro [2003]), and Chapter 4 of Shapiro et al. [2009] for an exhaustive study and bibliographical references concerning convexity theory in probabilistic programming.

Despite this considerable progress, the problem of convexity remains to present a big challenge of stochastic programming, especially for the problem of type (2) with random matrix. The are
more or less successful extensions to the grounds; for example, Prékopa et al. [2011] have recently extended the classical result of Prékopa [1974], asserting that the problem is convex if the rows are independently normally distributed and the covariances matrices of the rows are constant multiples of each other. A more promising direction was started by Henrion [2007] giving the complete structural description (not only the convexity) of a one-row linear chance constraint. Henrion and Strugarek [2008] introduced a notion of $r$-decreasing density, succeeded to relate this new notion with $r$-concavity of constraint function $g_k$ of (4), and proved the convexity of the set $M(p)$ for the case of independent random variables. The result was applied also for the problem of convexity of $X(p)$ with normally distributed independent rows, advancing so significantly the classical results. The result for right-hand side has then been extended towards dependency by Houda [2008] (see also Houda [2009]) using a variation to the strong mixing coefficient, and by Henrion and Strugarek [2011] and van Ackooij [2013] using the theory of copulas. In our paper, we pursue this last direction: we prove the convexity of the set $X(p)$ for (even non-normal) distributions falling within the class of elliptical distributions, and use the broad class of Archimedean copulas to represent the dependency of the rows in the problem (2).

Using elliptical distributions (as underlying class of probability measures), and using copulas to represent the structural dependency, is very rare in chance-constrained programming. For the ellipticity, Henrion [2007] restricts his consideration to a very special case of one-row constraint only; Calafiore and El Ghaoui [2006] used a similar notion of $Q$-radial distribution to develop a second-order cone constraint, but again for a one-row chance constraint only. Concerning the copulas, up to our knowledge and beyond the reference mentioned above, copula theory is used only in the context of generating scenarios for multistage stochastic optimization programs. In our paper, we exploit together both notions to reformulate the problem (2) as the problem of convex optimization, and to propose an approximation scheme for this problem using the second-order cone programming method.

Second-order cone programming (SOCP) is a subclass of convex optimization in which the problem constraint set is the intersection of an affine linear manifold and the Cartesian product of second-order (Lorenz) cone. For the details on we refer the reader to Alizadeh and Goldfarb [2003] and the monograph Boyd and Vandenberghe [2004]. The method was proposed by Cheng and Lisser [2012] and Cheng et al. [2012] for the problem of type (2) where the rows of the constraint matrix are assumed to be independent and normally distributed.

This paper is organized as follows: we start with some insights into the theory of copulas and elliptical distributions. A convexity result for right-hand sided problem (4) is given in Section 3, whereas the main convexity result and two approximation schemes for (2) are formulated in Section 4. The numerical study is conducted in Section 5 and conclusions are given in the last section.

2 Preliminaries

2.1 Dependence

The notion of copula is known for years in probability theory and mathematical statistics to describe the dependency structure of a random vector. We will use this notion to characterize dependence between constraint rows of the problems (2) and (4). In this section we mention only some basic facts about copulas needed for our following investigation; we refer to the book Nelsen [2006] for a complete introduction to the theory. Before the definition of the copula, we also recall the (usual) definition of the one-dimensional quantile function.

**Definition 2.1.** Given a distribution function $F : \mathbb{R} \to [0; 1]$, its **quantile function** is defined as

$$F^{(-1)}(\pi) := \inf \{ t \in \mathbb{R} \mid F(t) \geq \pi \}.$$
Definition 2.2. A copula is the distribution function \( C : [0; 1]^K \rightarrow [0; 1] \) of some \( K \)-dimensional random vector whose marginals are uniformly distributed on \([0; 1]\).

Proposition 2.1 (Sklar’s theorem). For any \( K \)-dimensional distribution function \( F : \mathbb{R}^K \rightarrow [0; 1] \) with marginals \( F_1, \ldots, F_K \), there exists a copula \( C \) such that

\[
\forall z \in \mathbb{R}^K \quad F(z) = C(F_1(z_1), \ldots, F_K(z_K)).
\]  
(6)

If, moreover, \( F_k \) are continuous, then \( C \) is uniquely given by

\[
C(u) = F(F_1^{(-1)}(u_1), \ldots, F_K^{(-1)}(u_K)).
\]  
(7)

Otherwise, \( C \) is uniquely determined on range \( F_1 \times \cdots \times \text{range } F_K \).

Sklar’s theorem provides a direct link between a copula and the joint distribution function of some random vector. By Sklar’s theorem, the copula represents a convenient tool to describe the dependence structure of the random vector considered; in addition, this description separates from description of the marginal distribution functions (marginals in short) of the vector. A special instance of copulas is the so-called independent (product) copula representing simply the independence of marginals:

\[
C_\Pi(u) := \prod_{k=1}^K u_k.
\]

Every copula (as a function) is bounded. Particularly important for our future investigation is the following upper bound.

Proposition 2.2 (Fréchet–Hoeffding upper bound). For any copula \( C \) and any \( u = (u_1, \ldots, u_K) \in [0; 1]^K \) it holds that

\[
C(u) \leq C_M(u) := \min_{k=1,\ldots,K} u_k.
\]

\( C_M \) is also a copula and it is known under the name of comonotone (maximum) copula. It represent the completely positive dependence of the marginals. In many cases it is found to be a limiting case of some sequence of copulas.

One of the most prominent copula is the one defined by Sklar’s theorem as

\[
C_\Sigma(u) := \Phi^\Sigma\left( \Phi^{(-1)}(u_1), \ldots, \Phi^{(-1)}(u_K) \right)
\]

where \( \Phi^\Sigma \) is the distribution function of \( K \)-variate normal distribution with zero mean, unit one-dimensional variances, and covariance matrix \( \Sigma \), and \( \Phi^{(-1)}(u_k) \) are standard one-dimensional normal quantiles. This copula is called Gaussian copula and it is the only copula that can represent the joint normal distribution. Unfortunately, it is not Archimedean (see below; it even cannot be formulated by an analytical expression). Hence, the results provided by our paper do not apply to such matrices \( T \) which are driven by Gaussian copulas, except the independent case.

Definition 2.3. A copula \( C \) is called Archimedean if there exists a continuous strictly decreasing function \( \psi : [0; 1] \rightarrow [0; +\infty] \), called generator of \( C \), such that \( \psi(1) = 0 \) and

\[
C(u) = \psi^{(-1)}\left( \sum_{i=1}^n \psi(u_i) \right).
\]  
(8)

If \( \lim_{u \rightarrow 0} \psi(u) = +\infty \) then \( C \) is called a strict Archimedean copula and \( \psi \) is called a strict generator.

The inverse \( \psi^{(-1)} \) of the generator function is continuous and strictly decreasing on \([0; \psi(0)]\) (the value of \( \psi(0) \) considered as \( +\infty \) if the copula is strict). Sometimes, \( \psi^{(-1)} \) is defined as the generalized inverse on the whole positive half-line \([0; +\infty] \) by setting \( \psi^{(-1)}(s) = 0 \) for \( s \geq \psi(0) \), losing that the strictness of the decrease property, but such a definition is not needed through the context of our paper. To determine if a continuous strictly decreasing function \( \psi \) is a copula generator we introduce the following notion of \( K \)-monotonic function.
Table 1: Table of selected Archimedean copulas with completely monotonic inverse generators.

<table>
<thead>
<tr>
<th>Copula family</th>
<th>Parameter $\theta$</th>
<th>Generator $\psi_\theta(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent (product)</td>
<td>$-$</td>
<td>$-\ln t$</td>
</tr>
<tr>
<td>Gumbel–Hougaard</td>
<td>$\theta \geq 1$</td>
<td>$(-\ln t)^\theta$</td>
</tr>
<tr>
<td>Frank</td>
<td>$\theta &gt; 0$</td>
<td>$-\ln\left(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1}\right)$</td>
</tr>
<tr>
<td>Clayton</td>
<td>$\theta &gt; 0$</td>
<td>$\frac{1}{\theta}(t^{-\theta} - 1)$</td>
</tr>
<tr>
<td>Joe</td>
<td>$\theta \geq 1$</td>
<td>$-\ln[1 - (1 - t)^\theta]$</td>
</tr>
</tbody>
</table>

Definition 2.4. A real function $f : \mathbb{R} \to \mathbb{R}$ is called $K$-monotonic on an open interval $I \subseteq \mathbb{R}$ for some $K \geq 2$ if it is differentiable up to the order $K - 2$, the derivatives satisfy

$$(-1)^k \frac{d^k}{dt^k} f(t) \geq 0 \quad \forall k = 0, 1, \ldots, K - 2 \quad \text{and} \quad \forall t \in I,$$

and the function $(-1)^{K-2} \frac{d^{K-2}}{dt^{K-2}} f(t)$ is nonincreasing and convex in $I$. If the condition (9) is valid for all $K \geq 2$ then $f$ is called completely monotonic on $I$.

Proposition 2.3 (McNeil and Nešlehová [2009]). Let $\psi : [0; 1] \to [0; +\infty]$ be a strictly decreasing function with $\psi(1) = 0$. Then it is the generator of a $K$-dimensional Archimedean copula if and only if $\psi^{-1}$ is $K$-monotonic on $(0; \psi(0))$.

Corollary 2.4. Any copula generator is convex.

Archimedean copulas are exceptionally suitable for applications due to their separable and relatively simple formulation through a one-dimensional function, usually analytic with a small number of parameters (one or two). Many such copulas were already provided by the literature. For example, the book Nelsen [2006] provides a table of 22 one-parameter Archimedean copulas. The most cited Archimedean copula families are given in Table 2.1, together with their parameter ranges and generator formulas.

Focus our attention to the set $M(p)$ defined by (5), i.e., the set of feasible solutions of the chance-constrained problem with random right-hand side. Assume (for each $k = 1, \ldots, K$) that the elements $\xi_k$ of $\xi$ have continuous distribution functions $F_k$, and the whole vector $\xi$ has joint distribution induced by a copula $C$. With these assumptions, we can rewrite the set $M(p)$ as

$$M(p) = \{x \in X \mid C\left(F_1(g_1(x)), \ldots, F_K(g_K(x))\right) \geq p\}.$$  

If $C$ is Archimedean we can obtain the following equivalent description of $M(p)$.

Let

$$M_I(p) = \left\{x \in X \mid \exists y_k \geq 0 : \psi[F_k(g_k(x))] \leq \psi(p) y_k \text{ for } k = 1, \ldots, K, \sum_{k=1}^{K} y_k = 1\right\}.$$  

Lemma 2.5. If the copula $C$ is Archimedean with the generator $\psi$ then $M(p) = M_I(p)$.

Proof. It is easily seen that

$$M(p) = \left\{x \in X \mid \psi^{-1}\left(\sum_{k=1}^{K} \psi[F_k(g_k(x))]\right) \geq p\right\} = \left\{x \in X \mid \sum_{k=1}^{K} \psi[F_k(g_k(x))] \leq \psi(p)\right\}$$  

5
as the generator of an Archimedean copula is strictly decreasing function, and noting that $\psi(p) \leq \psi(0)$ if the generator is not strict (the inverse $\psi^{-1}$ is strictly decreasing on $[0; \psi(0)]$). Introducing auxiliary nonnegative variables $y = (y_1, \ldots, y_K)$ with $\sum_k y_k = 1$, the inequality in (12) is equivalent to
\[
\sum_{k=1}^K \psi[F_k(g_k(x))] \leq \psi(p) \sum_{k=1}^K y_k \text{ for some } y_k \geq 0 \text{ with } \sum_{k=1}^K y_k = 1. \tag{13}
\]
Denote
\[
M_I(p) = \left\{ x \in X \mid \exists y_k \geq 0 : \psi[F_k(g_k(x))] \leq \psi(p)y_k \text{ for } k = 1, \ldots, K, \sum_{k=1}^K y_k = 1 \right\};
\]
we will show that $M(p) = M_I(p)$. Without lost of generality we assume $p < 1$ (the case $p = 1$ is obvious).

The inclusion $M_I(p) \subseteq M(p)$ is seen immediately: it suffices to sum up the individual inequalities in $M_I(p)$. For the opposite direction, consider $x \in M(p)$. The Fréchet–Hoeffding upper bound (Proposition 2.2) implies, for the Archimedean $C$, that
\[
F_k(g_k(x)) \geq \min_{j=1,\ldots,K} F_j(g_j(x)) \geq \psi^{-1} \left( \sum_{j=1}^K \psi[F_j(g_j(x))] \right) \geq p \forall k = 1, \ldots, K
\]
thus,
\[
\psi[F_k(g_k(x))] \leq \psi(p) \forall k = 1, \ldots, K.
\]
Now, define
\[
y_k := \frac{\psi[F_k(g_k(x))]}{\psi(p)} \text{ for } k = 1, \ldots, K - 1, \quad y_K := 1 - \sum_{k=1}^{K-1} y_k.
\]
The definition of $y_k$ implies directly that $\psi[F_k(g_k(x))] \geq \psi(p)y_k$ for $k = 1, \ldots, K - 1$. Furthermore,
\[
\psi[F_K(g_K(x))] \geq \psi(p) - \sum_{k=1}^{K-1} \psi[F_k(g_k(x))] = \psi(p)\left( 1 - \sum_{k=1}^{K-1} y_k \right) = \psi(p)y_K.
\]
Hence, $x \in M_I(p)$. \hfill \Box

**Remark 2.1.** It is worth to note that the auxiliary variables $y_k$ should have their values greater than zero for random vector with unbounded support (e.g. normally distributed). Indeed, $y_k = 0$ signifies $F_k(g_k(x)) = 1$, i.e., the constraint $g_k(x) \geq \xi_k$ should be satisfied almost surely. In fact, the feasible set of the original problem (4) is empty in such case. We will avoid these pathological case from our further consideration by a general assumption that the feasible sets $M(p)$ or $X(p)$ are nonempty.

### 2.2 Elliptically symmetric random vectors

The concept of elliptically (or radially) symmetric random vectors was introduced in the theory of probability in the early seventies of the 20th century to extend the class of multivariate normal distributions. A thorough survey of basic results and properties of elliptical distributions can be found in the book Fang et al. [1990].

**Definition 2.5.** An $s$-dimensional random vector $\xi$ is said to have an elliptically symmetric distribution if its characteristic function is given by
\[
\phi(z) := \mathbb{E}e^{iz^T \xi} = e^{i z^T \mu} \varphi(z^T \Sigma z)
\]
where $\varphi$ is some scalar function (called characteristic generator), $\mu$ some vector (location parameter), and $\Sigma$ a matrix with rank $r$ (scale matrix).
Then rows of the problem or matrix) regardless the actual value of
about the “row dependence” of the problem or matrix and scale parameters to values different than
butions, together with their characteristic generators and radial densities. Setting the location
Remark 2.3.
definition and the language of Paindaveine [2012].

\[ \phi(t) = \frac{c_s}{\sqrt{\det \Sigma}} g_s \left( \sqrt{(z - \mu)^T \Sigma^{-1} (z - \mu)} \right) \]  

(14)

where \( g_s : \mathbb{R}^+ \to \mathbb{R}^+ \) is a so-called radial density, \( c_s > 0 \) is a normalization factor ensuring that \( f_\xi \) integrates to one, and the matrix \( \Sigma \) is required to have a full rank, i.e., to be positive definite (we denote this by \( \Sigma \succ 0 \) in the sequel). Among many properties of elliptical distributions we note that the class of elliptical distributions is closed under affine transformations: if \( \xi \sim \text{Ellip}_s(\mu, \Sigma, \varphi) \) then for any \( (r \times s) \)-matrix \( L \) and any \( r \)-vector \( b \), the distribution of \( L\xi + b \) is \( \text{Ellip}_p(L\mu + b, L\Sigma L^T, \varphi) \).

Remark 2.2. The definition of \( g_s \) is unique only up to a multiplicative constant. In this view, different equivalent formulations for elliptical density appear in the literature, mostly using the notion of density generators \( t \mapsto c_s g_s(\sqrt{t}) \) instead of radial densities. Here, we have adopted the definition and the language of Paindaveine [2012].

Remark 2.3. In Table 2 we provide a short selection of prominent multivariate elliptical distributions, together with their characteristic generators and radial densities. Setting the location and scale parameters to values different than \( \mu = 0 \) and \( \Sigma = I_s \), we can easily obtain the non-standardized versions of these well-known distributions. Note that the Cauchy distribution is a special case of the \( t \) distribution with \( \nu = 1 \). The star * denotes an expression which is too involved to be mentioned in the table. Concerning the multivariate \( t \) distribution we refer to the book Kotz and Nadarajah [2004]; for the logistic distribution, see Volodin [1999].

<table>
<thead>
<tr>
<th>Law</th>
<th>Characteristic generator ( \varphi(t) )</th>
<th>Radial density ( g_s(t) )</th>
<th>Normalizing constant ( c_s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>normal</td>
<td>( \exp{-\frac{1}{2} t} )</td>
<td>( \exp{-\frac{1}{2} t^2} )</td>
<td>( (2\pi)^{-s/2} )</td>
</tr>
<tr>
<td>( t )</td>
<td>*</td>
<td>( (1 + \frac{1}{2} t^2)^{-\left(s+\nu\right)/2} )</td>
<td>( (\nu \pi)^{-s/2} \frac{\Gamma\left((s+\nu)/2\right)}{\Gamma\left(\nu/2\right)} )</td>
</tr>
<tr>
<td>Cauchy</td>
<td>( \exp{-\sqrt{t}} )</td>
<td>( (1 + t^2)^{-\left(s+1\right)/2} )</td>
<td>( \pi^{-\left(s+1\right)} \frac{\Gamma\left(\frac{1}{2}(s + 1)\right)}{\Gamma(\frac{1}{2})} )</td>
</tr>
<tr>
<td>Laplace</td>
<td>( (1 + \frac{1}{2} t) ) (^{-1} )</td>
<td>( \exp{-\sqrt{2}t} )</td>
<td>( \pi^{-\frac{s}{2}} \frac{\Gamma\left(\frac{s}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} )</td>
</tr>
<tr>
<td>logistic</td>
<td>( \frac{2\pi \sqrt{\pi}}{e^{\pi \sqrt{t} - e^{-\pi \sqrt{t}}} \sqrt{\pi}} )</td>
<td>( e^{-t^2} )</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 2: Table of selected multivariate elliptically symmetric distributions.

We’ll write \( \xi \sim \text{Ellip}_s(\mu, \Sigma; \varphi) \), and eventually drop the dimensionality index \( s \) where no confusion can happen. Not all elliptically symmetric distributions have density, but if they have some, it must be of the form

(15)

The following result is a special case of Lemma 2.2 in Henrion [2007].

**Proposition 2.6.** Assume \( p \in (0; 1) \), \( \xi \sim \text{Ellip}(\mu, \Sigma; \varphi) \) where \( \Sigma \succ 0 \), and denote

\[ Y(p) := \left\{ x \in X \mid \mathbb{P}\{\xi^Tx \leq h\} \geq p \right\}. \]

Then

(16)

where \( \Psi \) is one-dimensional distribution function induced by the characteristic function \( \varphi(t) = \varphi(t^2) \). In particular, \( \Psi \) does not depend on \( x \).

We extend this idea to the multi-row case introducing a copula to describe the dependence structure of the optimization problem. It can be taught independently of the variable \( x \) if the rows are elliptically distributed, as stated by the next lemma. This lemma allows us to speak about the “row dependence” of the problem or matrix \( T \) (i.e., as the dependence between the rows of the problem or matrix) regardless the actual value of \( x \).
Lemma 2.7. If \( t_k^T \sim \text{Ellip}(\mu_k, \Sigma_k; \varphi_k) \) with \( \Sigma_k \succ 0 \) then the copula joining random variables \( t_k^T x \) is independent of \( x \).

Proof. If \( x = 0 \) then \( t_k^T x = 0 \) are scalar for every \( k \). Hence, suppose \( x \neq 0 \) and define

\[
\xi_k(x) := \frac{t_k^T x - \mu_k^T x}{\sqrt{x^T \Sigma_k x}}.
\]

The random variable \( \xi_k(x) \) has the distribution function \( \Psi_k \), defined as in Proposition 2.6, independent of \( x \). In other words, there exists a random variable \( \xi_k \) (independent of \( x \)) with the distribution function \( \Psi_k \) such that

\[
t_k^T x = t(x) := \mu_k^T x + \xi_k \sqrt{x^T \Sigma_k x}.
\]

(17)

The joint distribution function of the vector \( \xi := (\xi_1, \ldots, \xi_K)^T \) is given, by Sklar’s Theorem, by

\[
\mathbb{P}\{\xi \leq y\} = C(\Psi_1(y_1), \ldots, \Psi_K(y_K))
\]

for \( y \in \mathbb{R}^K \) where \( C \) is uniquely given on range \( \Psi_1 \times \ldots \times \text{range} \Psi_K \). For fixed \( x \), the mapping \( t_x(\cdot) \) defined by (17) is strictly increasing (affine) transformation. By Theorem 2.4.3 of Nelsen [2006], the joint distribution function of \( (t_x(\xi_1), \ldots, t_x(\xi_K))^T \) is driven by the copula \( C \) (of \( \xi \)). But \( C \) was defined independently of \( x \), hence the assertion of the lemma is fulfilled. \( \square \)

As the final result of this section, we obtain an equivalent representation of the feasible set \( X(p) \) of the problem (2) if the distribution of the rows is normal and the copula joining these rows is Archimedean. In particular, the problem with independent normally distributed rows now comes as a special case of this theorem.

Theorem 2.8. Suppose, in (2), that \( t_k^T \sim \text{Ellip}(\mu_k, \Sigma_k; \varphi_k) \) where \( \Sigma_k \succ 0 \). If \( p \in (0; 1) \) then the feasible set of the problem (2) can be equivalently written as

\[
X(p) = \left\{ x \in X \mid \exists y_k \geq 0 : \mu_k^T x + \Psi_k^{-1}(\psi^{-1}(y_k \psi(p))) \sqrt{x^T \Sigma_k x} \leq h_k, \ k = 1, \ldots, K, \sum_k y_k = 1 \right\}
\]

(18)

where \( \Psi_k \) are one-dimensional distribution functions induced by characteristic functions of the form \( \phi_k(t) = \varphi_k(t^2) \), and \( \psi \) is the generator of an Archimedean copula describing the dependence properties of the rows of the matrix \( T \).

Proof. If \( x = 0 \), the equivalence is obvious. Suppose so that \( x \neq 0 \) (say \( x \not\in X \)) and denote

\[
\xi_k(x) := \frac{t_k^T x - \mu_k^T x}{\sqrt{x^T \Sigma_k x}}, \quad g_k(x) := \frac{h_k - \mu_k^T x}{\sqrt{x^T \Sigma_k x}},
\]

(19)

then

\[
X(p) = \left\{ x \in X \mid \mathbb{P}[T x \leq h] \geq p \right\}
\]

\[
= \left\{ x \in X \mid \mathbb{P}[t_k^T x \leq h_k, \ k = 1, \ldots, K] \geq p \right\}
\]

\[
= \left\{ x \in X \mid \mathbb{P}[\xi_k(x) \leq g_k(x), \ k = 1, \ldots, K] \geq p \right\}.
\]

(20)

According to the calculus rule for elliptical distributions, namely \( \phi_x(t) = e^{itd} \cdot \phi_{-t}(ct) \), the characteristic function of \( \xi_k(x) \) is

\[
\phi_{\xi_k(x)}(t) = \exp \left\{ -it \frac{\mu_k^T x}{\sqrt{x^T \Sigma x}} \cdot \phi_{t_k^T} \left( \frac{x}{\sqrt{x^T \Sigma x}} \right) \right\}.
\]
The characteristic function of $t_k^T$ is $\phi_{t_k^T}(z) = e^{iz^T\mu_k(z^T\Sigma_k)}$, so $\phi_{\xi_k(x)}(t) = \varphi_k(t^2)$. It follows that the distribution function of $\xi_k(x)$ is $\Psi_k$, independent of $x$. Returning to (20), and applying Lemma 2.5, we have

$$X(p) = \left\{ x \in X \mid \exists y_k \geq 0 : \psi[\Psi_k(g_k(x))] \leq y_k\psi(p), \ k = 1, \ldots, K, \sum_{k=1}^{K} y_k = 1 \right\}$$

$$= \left\{ x \in X \mid \exists y_k \geq 0 : g_k(x) \leq \Psi_k(\psi^{-1}(\psi(p)y_k)), \ k = 1, \ldots, K, \sum_{k=1}^{K} y_k = 1 \right\}$$

$$= \left\{ x \in X \mid \exists y_k \geq 0 : \mu_k^T x + \Psi_k^{-1}\left(\psi^{-1}(y_k\psi(p))\right) \sqrt{x^T\Sigma_k x} \leq h_k, \ k = 1, \ldots, K, \sum_{k} y_k = 1 \right\}.$$

$\square$

## 3 Convexity

### 3.1 Generalized convexity notions

To deal with convexity properties of the sets $M(p)$ and $X(p)$, we introduce two particular convexity notions: $r$-concave function and $r$-decreasing density.

**Definition 3.1** (Prékopa [1995], Chapter 4 of Shapiro et al. [2009]). A function $f: \mathbb{R}^s \to (0; +\infty)$ is called $r$-concave for some $r \in [-\infty; +\infty]$ if $\text{dom} f$ is convex and

$$f(\lambda x + (1 - \lambda) y) \geq [\lambda f^r(x) + (1 - \lambda) f^r(y)]^{1/r} \quad (21)$$

is fulfilled for all $x, y \in \text{dom} f$ and all $\lambda \in [0; 1]$. The cases $r = -\infty, 0, +\infty$ are to be interpreted by continuity.

The case $r = 1$ is the concavity in its usual sense. The case $r = 0$ corresponds to the so-called log-concavity, i.e., to the case in which the function $\ln f$ is concave. The case $r = -\infty$ is known as the quasi-concavity and the corresponding right-hand side of (21) takes the form of $\min\{f(x), f(y)\}$. If $f$ is $r$-concave for some $r$, then it is also $r'$-concave for all $r' \leq r$; in particular, all $r$-concave functions are quasi-concave.

**Definition 3.2** (Henrion and Strugarek [2008]). A function $f: \mathbb{R} \to \mathbb{R}$ is called $r$-decreasing for some $r \in \mathbb{R}$ with the threshold $t^*(r) > 0$ if it is continuous on $(0; +\infty)$ and the function $t \mapsto t^r f(t)$ is strictly decreasing for all $t > t^*(r)$.

The threshold $t^*(r)$ depends on the value of $r$; we conserve this explicit dependence in the notation as it becomes important in our later statements. If a function $f$ is nonnegative and $r$-decreasing for some $r$, then it is $r'$-decreasing for all $r' \leq r$. In particular, if $r > 0$ then $f(t)$ is $0$-decreasing, hence strictly decreasing for $t > t^*(0)$. The table of prominent one-dimensional $r$-decreasing densities together with their thresholds was given in Henrion and Strugarek [2008]. By the following proposition, we add some elliptical densities to this list.

**Proposition 3.1.** The following one-dimensional elliptical distributions have $r$-decreasing densities for some $r$:

1. normal distribution, for $r > 0$ with the threshold $t^*(r) = \frac{1}{2} \left( \mu + \sqrt{\mu^2 + 4r\sigma^2} \right)$;
2. Student’s $t$ distribution with $\nu$ degrees of freedom, for $0 < r < \nu + 1$ with the threshold $t^*(r) = \sqrt{\frac{\nu r}{\nu + 1 - r}}$;
3. Laplace (double exponential) distribution, for all $r > 0$ with the threshold $t^*(r) = \frac{\sigma^2}{\sqrt{2}}$. 


Proof. To test \( r \)-decreasing property for a differentiable elliptical density we have only to check if the derivative is strictly negative for \( t > t^*(r) \). Due to a special form (14) of the elliptical density, this is equivalent to check

\[
\left( \frac{\mu}{\sigma} + \dot{t} \right) g_s'(\dot{t}) + r g_s(\dot{t}) < 0
\]  

(22)

for all \( \dot{t} > \dot{t}^*(r) \geq \mu \), using substitution \( \dot{t} := \frac{t - \mu}{\sigma} \) for \( t \geq \mu \). By the backward substitution, the resulting threshold will have the form

\[
t^*(r) := \mu + \sigma \ddot{t}^*.
\]  

(23)

1. A proof for the normal distribution was given in Henrion and Strugarek [2008] as Proposition 4.1.

2. The derivative of the radial density for a \( t \) distribution (even non-standardized) reads as

\[
g'_s(t) = -\frac{s + \nu}{\nu + t^2} g_s(t).
\]

Condition (22) translates to

\[
-\dot{t}^2(s + \nu - r) - \frac{\mu}{\sigma}(s + \nu) + r \nu < 0.
\]

The optimal threshold is calculated by (23) as

\[
t^*(r) = \mu \left( 1 + \frac{s + \nu}{2\sigma(s + \nu - r)} \right) + \sqrt{\frac{1}{4} \left( \frac{s + \nu - \mu}{s + \nu - r} \right)^2 + \frac{r \nu}{s + \nu - r} - \sigma^2}
\]

for \( r < s + \nu \). For the standardized univariate \( t \) distribution use \( \mu = 0, \sigma = 1, s = 1 \) to get the result.

3. The condition (22) for the Laplace distribution reduces to

\[
-\left( \frac{\mu}{\sigma} + \ddot{t}^* \right) \sqrt{2} + r < 0
\]

which translates to the optimal \( \ddot{t}^* = \frac{r}{\sqrt{2}} - \frac{\mu}{\sigma} \); the formula for \( t^*(r) \) is then easily computed by (23).

\[\square\]

### 3.2 Convexity of the problem with random right-hand side

We start our investigation by recalling the following technical proposition which joins \((r + 1)\)-decreasing density function and the convexity of the associated distribution function.

**Proposition 3.2** (Henrion and Strugarek [2008], Lemma 3.1). Let \( F : \mathbb{R} \rightarrow [0; 1] \) be a distribution function with \((r + 1)\)-decreasing density for some \( r > 0 \) and threshold \( t^*(r+1) > 0 \). Then the function \( z \mapsto F(z^{-1/r}) \) is concave on \((0; (t^*(r+1))^{-r})\). Moreover, \( F(t) < 1 \) for all \( t \in \mathbb{R} \).

The next theorem provides the result for convexity of the problem (4) with random right-hand side where the dependence of the rows is driven by an Archimedean copula.

**Theorem 3.3.** Consider the set \( M(p) \) and the following assumptions for \( k = 1, \ldots, K \):

1. there exist \( r_k > 0 \) such that \( g_k \) are \((-r_k)\)-concave on their domains,
2. the marginal distribution functions $F_k$ have $(r_k + 1)$-decreasing densities with the thresholds $t_k^*(r_k + 1)$, and

3. the copula $C$ joining the components of the random vector $\xi$ is Archimedean.

Then $M(p)$ is convex for all $p > p^* := \max_k F_k[t_k^*(r_k + 1)]$.

Proof. Let $p > p^*$, $\lambda \in [0; 1]$, and $x, y \in M(p)$. We have to show that $\lambda x + (1 - \lambda) y \in M(p)$, that is

$$C\left(F_1[g_1(\lambda x + (1 - \lambda) y)], \ldots, F_K[g_K(\lambda x + (1 - \lambda) y)]\right) = \psi^{-1}\left(\sum_{k=1}^{K} \psi\left(F_k[g_k(\lambda x + (1 - \lambda) y)]\right)\right) \geq p$$

or, equivalently,

$$\sum_{k=1}^{K} \psi\left(F_k[g_k(\lambda x + (1 - \lambda) y)]\right) \leq \psi(p)$$

(c.f. the proof of Lemma 2.5). As $x \in M(p)$, together with Proposition 2.2 we have

$$p \leq C\left(F_1[g_1(x)], \ldots, F_K[g_K(x)]\right) \leq \min_k F_k[g_k(x)] < 1.$$

Adding the definition of $p^*$ implies

$$0 \leq F_k[t_k^*(r_k + 1)] < p \leq F_k[g_k(x)] < 1. \tag{24}$$

for each $k = 1, \ldots, K$.

Note that, for $\pi \in (0; 1)$, $F_k^{(-1)}(\pi)$ is a real number. Having a density, $F_k$ are continuous by assumption 2 and the following implication holds for all $t \in \mathbb{R}$ and $\pi \in (0; 1)$

$$F_k(t) < \pi \Rightarrow t < F_k^{(-1)}(\pi).$$

Knowing that $F_k^{(-1)}$ is increasing, this implication and (24) provide

$$0 < t_k^*(r_k + 1) < F_k^{(-1)}(p) \leq F_k^{(-1)}(F_k[g_k(x)]) \leq g_k(x).$$

In particular, $g_k(x) > 0$ and $0 < g_k(x)^{-r_k} < (t_k^*(r_k + 1))^{-r_k}$ as negative power is a decreasing function on $(0; +\infty)$. Of course, the same inequalities apply also to $y \in M(p)$.

By assumption 1, $g_k$ is $(-r_k)$-concave, that is

$$g_k(\lambda x + (1 - \lambda) y) \geq \left[\lambda g_k^{-r_k}(x) + (1 - \lambda) g_k^{-r_k}(y)\right]^{-1/r_k}$$

for $\lambda \in [0; 1]$. $F_k$ is nondecreasing, hence

$$F_k[g_k(\lambda x + (1 - \lambda) y)] \geq F_k\left(\left[\lambda g_k^{-r_k}(x) + (1 - \lambda) g_k^{-r_k}(y)\right]^{-1/r_k}\right).$$

Proposition 3.2 implies that

$$F_k[g_k(\lambda x + (1 - \lambda) y)] \geq \lambda F_k[g_k(x)] + (1 - \lambda) F_k[g_k(y)],$$

i.e., the composition $F_k \circ g_k$ is a concave function on $M(p)$.

The Archimedean generator $\psi$ is a strictly decreasing convex function on $[0; \psi(0)]$ (which is enough as $\psi(F_k[\cdot]) \leq \psi(p) < \psi(0)$). Hence

$$\psi\left(F_k[g_k(\lambda x + (1 - \lambda) y)]\right) \leq \lambda \psi\left(F_k[g_k(x)]\right) + (1 - \lambda) \psi\left(F_k[g_k(y)]\right).$$
According to Lemma 2.5, for \(x, y \in M(p)\) there exists \(z_{k}^{(x)}, z_{k}^{(y)} \geq 0\) with \(\sum z_{k}^{(x)} = \sum z_{k}^{(y)} = 1\) such that

\[
\psi\left(F_{k}[g_{k}(x)]\right) \leq \psi(p)z_{k}^{(x)} \quad \text{and} \quad \psi\left(F_{k}[g_{k}(y)]\right) \leq \psi(p)z_{k}^{(y)}.
\]

Hence, we conclude on

\[
\sum_{k=1}^{K} \psi\left(F_{k}[g_{k}(\lambda x + (1-\lambda)y)]\right) \leq \sum_{k=1}^{K} \lambda \psi\left(F_{k}[g_{k}(x)]\right) + (1-\lambda) \sum_{k=1}^{K} \psi\left(F_{k}[g_{k}(y)]\right)
\]

\[
\leq \sum_{k=1}^{K} \left(\lambda \psi(p)z_{k}^{(x)} + (1-\lambda)\psi(p)z_{k}^{(y)}\right) = \psi(p).
\]

\[\square\]

### 3.3 Concavity of the standardization functions

Our aim is to prove the convexity of the set \(X(p)\). Partial result was given for independent normally distributed rows by Henrion and Strugarek [2008]. This result can be reconsidered using logexp-concave copulas introduced by Henrion and Strugarek [2011] (for example, this applies for Gumbel–Hougaard copula). However, we provide a generalized result which encompasses the whole class of elliptical distributions and all Archimedean copulas. In some steps, we also provide some improvements of the basic results from Henrion and Strugarek [2008]. It is still worth to note that Gaussian copulas do not belong to the class of investigated copulas as they are neither Archimedean nor generally logexp-concave (except the independent case); the convexity of linear chance-constrained problem with normally distributed technology matrix with dependent rows remain still an open question.

The following convexity lemma is an improved version of Lemma 5.1 of Henrion and Strugarek [2008].

**Lemma 3.4.** Given \(\mu \in \mathbb{R}^{n}\) and \(\Sigma > 0\), then for any \(r > 1\) the function

\[
f(x) := \left(\frac{\sqrt{x^{T}\Sigma x}}{h - \mu^{T}x}\right)^{r},
\]

defined on the domain \(\text{dom } f := \{x \mid h - \mu^{T}x > 0\}\), is convex on the set

\[
\Omega := \left\{x \mid h - \mu^{T}x > \frac{r + 1}{r - 1} \lambda_{\min}^{-1/2}\|\mu\|\sqrt{x^{T}\Sigma x}\right\}
\]

where \(\lambda_{\min}\) is the smallest eigenvalue of \(\Sigma\).

**Proof.** If \(r > 1\) and \(f(x) > 0\), the Hessian of \(f\) calculates as

\[
\nabla^{2}f(x) = r f(x)(h - \mu^{T}x)^{-2}(x^{T}\Sigma x)^{-2}A
\]

where

\[
A := (h - \mu^{T}x)^{2}(x^{T}\Sigma x)\Sigma + 2r(h - \mu^{T}x)(x^{T}\Sigma x)\Sigma x\mu^{T}
\]

\[
- (2 - r)(h - \mu^{T}x)^{2}(\Sigma x)(\Sigma x)^{T} + (1 + r)(x^{T}\Sigma x)^{2}\mu^{T}.
\]

To check the positive definiteness of \(\nabla^{2}f(x)\), it is enough to check the positive definiteness of the matrix \(A\). To do this, consider \(z \in \mathbb{R}^{n}, z \neq 0\) and check the positivity of \(z^{T}Az\). A technical calculation provides

\[
z^{T}Az = (h - \mu^{T}x)^{2}(x^{T}\Sigma x)(z^{T}\Sigma z) + 2(h - \mu^{T}x)(x^{T}\Sigma x)(z^{T}\Sigma x)(\mu^{T}z)
\]

\[
- \frac{3 - r}{r + 1}(h - \mu^{T}x)^{2}(z^{T}\Sigma x)^{2} + \frac{1}{r + 1}\left[(r - 1)(h - \mu^{T}x)(z^{T}\Sigma x) + (r + 1)(x^{T}\Sigma x)(\mu^{T}z)\right]^{2}
\]

\[
12
\]
The last term of \( z^T Az \) is nonnegative; also note that \( \frac{3-r}{r+1} < 1 \) if \( r > 1 \). By the Cauchy–Schwarz inequality and known facts of linear algebra,

\[
(z^T \Sigma x)^2 \leq (z^T \Sigma z)(x^T \Sigma x), \quad \mu^T z \geq -\|\mu\|\|z\|, \quad \lambda_{\text{min}}\|z\|^2 \leq z^T \Sigma z.
\]

Hence, we are allowed to continue by

\[
z^T Az \geq \left( 1 - \frac{3-r}{r+1} \right) (h - \mu^T x)^2 (x^T \Sigma x)(z^T \Sigma z) - 2(h - \mu^T x)(x^T \Sigma x)(z^T \Sigma x)\|\mu\|\|z\|
\]

\[
\geq \left( 1 - \frac{3-r}{r+1} \right) (h - \mu^T x)(x^T \Sigma x) \left[(h - \mu^T x)(z^T \Sigma z) - \frac{r+1}{r-1}\sqrt{z^T \Sigma z}\sqrt{x^T \Sigma x}\|\mu\|\|z\|\right]
\]

\[
\geq \left( 1 - \frac{3-r}{r+1} \right) (h - \mu^T x)(x^T \Sigma x)\sqrt{z^T \Sigma z}\|z\| \left[(h - \mu^T x)\sqrt{\lambda_{\text{min}}} - \frac{r+1}{r-1}\sqrt{x^T \Sigma x}\|\mu\|\right] > 0
\]

by the definition of \( \Omega \).

\[\square\]

**Corollary 3.5.** For any \( r_k > 1 \), the functions \( g_k(x) \) defined by (19) are \((-r_k)\)-concave on every convex subset of the set

\[
\Omega^{(k)} = \left\{ x \neq 0 \mid h_k - \mu_k^T x > \frac{r_k+1}{r_k-1} \lambda_{\text{min}}^{-1/2} \sqrt{x^T \Sigma_k x} \right\}.
\]

**Proof.** The function \( g_k(x) \) is positive on \( \Omega^{(k)} \) but not well defined for \( x = 0 \) and excluding zero from \( \Omega^{(k)} \) may hurt its convexity. Restricting the domain to any convex subset of \( \Omega^{(k)} \) we are ready to apply Lemma 3.4.

\[\square\]

Compared to Lemma 5.1 of Henrion and Strugarek [2008], formulated only for the case \( r_k = 2 \), Corollary 3.5 provides a better bound \( (\frac{\lambda_{\text{min}}}{\lambda_{\text{min}}})^{1/2} \) times smaller right-hand side in the definition of \( \Omega^{(k)} \).

If \( \mu_k = 0 \), even better result can be found by the following lemma.

**Lemma 3.6.** If \( \mu_k = 0 \) and \( h_k > 0 \) then the function \( g_k(x) \) defined by (19) is \((-1)\)-concave on any convex subset of \( \mathbb{R}^n \setminus \{0\} \).

**Proof.** If \( h_k > 0 \), the function \( f(x) = \frac{\sqrt{x^T \Sigma x}}{h_k} \) is positive and convex on any convex subset \( S \) of \( \mathbb{R}^n \setminus \{0\} \). Its Hessian calculates as

\[
\nabla^2 f(x) = \frac{(x^T \Sigma x) \Sigma - \Sigma xx^T \Sigma}{h_k (x^T \Sigma x)^{3/2}}
\]

and

\[
z^T \nabla^2 f(x) = \frac{z^T \Sigma xx^T \Sigma z - (z^T \Sigma x)^2}{h_k (x^T \Sigma x)^{3/2}} \geq 0
\]

for \( x \neq 0 \) and any \( z \in \mathbb{R}^n \) by the Cauchy–Schwarz inequality. Therefore \( g_k(x) = \frac{h_k}{\sqrt{x^T \Sigma_k x}} \) is \((-1)\)-concave on \( S \).

\[\square\]

## 4 Problems with random matrices — main result

### 4.1 Convex reformulation

We are now ready to focus on the reformulation of the feasible set \( X(p) \) of the problem (2). In the main theorem of this paper we introduce sufficient conditions under which the set \( X(p) \) is convex.

**Theorem 4.1.** Consider problem (2) where
1. rows $T_k$ of the matrix $T$ have elliptically symmetric distributions with parameters $(\mu_k, \Sigma_k, \varphi_k)$ where $\Sigma_k$ are positive definite matrices; denote by $\Psi_k$ the (standardized row) distribution functions generated by characteristic functions of the form $\varphi_k(t)$;

2. the joint distribution function of $\Psi_k$ is driven by an Archimedean copula with the generator $\psi$.

Then the problem (2) can be equivalently written as

$$\min c^T x \quad \text{subject to}$$

$$\mu_k^T x + \Psi_k^{-1}(\psi^{-1}(y_k\psi(p))) \sqrt{x^T \Sigma_k x} \leq h_k$$

$$\sum_k y_k = 1$$

$$x \in X, \ y_k \geq 0 \ \text{with} \ k = 1, \ldots, K.$$

Moreover, if

3. the densities associated with $\Psi_k$ are (at least) $(r_k + 1)$-decreasing with thresholds $t^*_k(r_k + 1) > 0$ for some $r_k > 1$, or at least 2-decreasing if $\mu_k = 0$;

4. the probability level $p$ satisfies

$$p > p^* := \max_k \left\{ \Psi_k \left( \max \left\{ t^*_k(r_k + 1), \frac{r_k + 1}{r_k - 1}(\lambda_{\min}^{(k)})^{-1/2}\|\mu_k\| \right\} \right) \right\}$$

where $\lambda_{\min}^{(k)}$ are lowest eigenvalues of the matrices $\Sigma_k$.

then the problem is convex.

Proof. The first part of the theorem has already been proven as Theorem 2.8. For the convexity result, we partially reproduce the proof of Theorem 5.1 of Henrion and Strugarek [2008] but with modifications concerning our use of copulas, elliptical distributions, and improved thresholds.

We consider separately the cases $0 \in X(p)$ and $0 \notin X(p)$. Note that $p > 0$ by assumption 4 (as a particular case), hence the property $0 \in X(p)$ is equivalent to $h_k \geq 0$ for all $k = 1, \ldots, K$.

Consider first the case $0 \notin X(p)$. For $x \neq 0$ denote again

$$\xi_k(x) := \frac{t^*_k x - \mu_k^T x}{\sqrt{x^T \Sigma_k x}}, \quad g_k(x) := \frac{h_k - \mu_k^T x}{\sqrt{x^T \Sigma_k x}}.$$

Referring to the proof of Theorem 2.8, recall that the (one-dimensional) random variables $\xi_k(x)$ have elliptical distributions with the distribution functions $\Psi_k$ not depending on $x$, and the feasible set can be rewritten as

$$X(p) = \{ x \in \mathbb{R}^n \mid P\{\xi_k(x) \leq g_k(x), \ k = 1, \ldots, K \} \geq p \}.$$

Denote

$$u^*_k := \frac{r_k + 1}{r_k - 1}(\lambda_{\min}^{(k)})^{-1/2}\|\mu_k\|,$$

$$\Omega_k^{(0)} := \left\{ x \in \mathbb{R}^n \mid h_k - \mu_k^T x > u^*_k(r_k)\sqrt{x^T \Sigma_k x} \right\},$$

$$\Omega_k^{(k)} := \Omega_k^{(0)} \setminus \{0\}.$$

Together with assumption 4, it can be shown that

$$X(p) \subseteq \Omega_k^{(k)}.$$
Indeed, let \( x \in X(p) \) be arbitrary. Then
\[
\Psi_k(g_k(x)) \geq \min_{k=1,\ldots,K} \Psi_k(g_k(x)) \geq C(\Psi_1(g_1(x)), \ldots, \Psi_K(g_K(x))) \geq p > \Psi_k(u_k^*).
\]
Due to assumption 3, \( \Psi_k \) is strictly increasing at least if \( p > \Psi_k(t_k^*(r_k + 1)) \) which is ensured by assumption 4. Hence, \( g_k(x) > u_k^* \) and thus \( x \in \Omega_0^{(k)} \).

If \( \mu_k \neq 0 \), the distribution function \( \Psi_k \) has an \((r_k + 1)\)-decreasing density with \( r_k > 1 \); therefore assumption 1 of Theorem 3.3 is satisfied with this \( r_k \). In the case \( \mu_k = 0 \) it is true even for \( r_k = 1 \).

A more attention should be deserved to \((-r_k)\)-concavity of the functions \( g_k \) as it is ensured only on a convex subset of \( \Omega^{(k)} \) (not on the whole domain of \( g_k \) which may be nonconvex). However, we will first show that if \( x, y \in X(p) \) then \( \lambda x + (1 - \lambda)y \neq 0 \). Indeed, if \( \lambda x + (1 - \lambda)y = 0 \) for some \( \lambda \in (0; 1) \) then
\[
x = -\frac{1 - \lambda}{\lambda} y.
\]
Since \( h_k - \mu_k^T x > 0 \) for any \( x \in X(p) \) (see Lemma 3.4 and (27)) it follows that
\[
\mu_k^T x = -\frac{1 - \lambda}{\lambda} \mu^T y < h_k, \quad \text{hence} \quad \mu^T y > -\frac{\lambda}{1 - \lambda} h_k;
\]

simultaneously \( \mu^T y < h_k \) as also \( y \in X(p) \). It follows that
\[
|\mu_k^T y| < \min \left\{ h_k; \frac{\lambda}{1 - \lambda} h_k \right\}. \tag{28}
\]
But, as \( 0 \not\in X(p) \) by assumption, at least one \( h_k \) is negative which contradicts (28).

Taking \( \text{conv} \ X(p) \) (the convex hull of \( X(p) \)) we have shown that \( \text{conv} \ X(p) \subset \Omega^{(k)} \), hence \( g_k \) is \((-r_k)\)-concave on \( \text{conv} \ X(p) \). It is clear from the proof of Theorem 3.3 that it is enough to conclude on the convexity of \( X(p) \), provided \( p > \Psi_k[t_k^*(r_k + 1)] \).

Consider now the case \( 0 \in X(p) \), i.e., all \( h_k \geq 0 \). Suppose \( x, y \in X(p) \) arbitrary, we have to check \( x_\lambda := \lambda x + (1 - \lambda)y \in X(p) \) for all \( \lambda \in [0; 1] \). Four cases may happen:

1. Case \( x = y = 0 \) leads to \( x_\lambda = 0 \in X(p) \) by assumption.
2. Case \( x \neq 0, y = 0 \): by Proposition 5.1 of Henrion and Strugarek [2008], the set \( X(p) \) is star-shaped with respect to the origin, hence, \( x_\lambda = \lambda x \in X(p) \). Note that the proposition remains valid for an arbitrary distribution.
3. Case \( x = 0, y \neq 0 \): \( x_\lambda \in X(p) \) by the same argument.
4. Case \( x \neq 0, y \neq 0 \): either \( x_\lambda = 0 \in X(p) \) by assumption, or \( x_\lambda \neq 0 \) and we can proceed as in the first part of the proof to state the \((-2)\)-concavity of the function \( g_k \) on a convex subset of \( \Omega^{(k)} \) leading to \( x_\lambda \in X(p) \) and the desired convexity result.

For specific distributions, we can provide explicit probability levels replacing the implicit inequality (26) by an explicit one using the known formulas for the thresholds \( t^*(r) \). The following corollary provide such level for normally distributed rows.

**Corollary 4.2.** Suppose that the rows \( t_k^* \) of the matrix \( T \) have multivariate normal distributions with parameters \((\mu_k, \Sigma_k > 0)\), and the joint distribution of \( t_k^* x \) is driven by an Archimedean copula. Then the set \( X(p) \) is convex if
\[
p > p^* = \Phi \left( \frac{1}{2} \max_k \left\{ \frac{\|\mu_k\|}{\sqrt{\lambda_{\min}^{(k)}}} + \sqrt{\frac{8 + \|\mu_k\|^2}{\lambda_{\min}^{(k)}}} \right\} \right),
\]
where \( \Phi \) is the one-dimensional standard normal distribution function.
Proof. According to Proposition 3.1, the standard normal distribution function is \((r_k + 1)\)-concave for any \(r_k > 0\) with the threshold \(t_k(r_k + 1) = \sqrt{r_k + 1}\) which is an increasing function of \(r_k\). If \(\mu_k = 0\) the result is immediate. With fixed \(\mu_k \neq 0\) and \(\lambda_{\min}^{(k)} \neq 0\) the function \(\frac{r_k+1}{r_k} \lambda_{\min}^{(k)} \|\mu_k\|^2\) is decreasing for \(r_k > 1\). To find the optimal value of \(p^*\), it is sufficient to solve the equation

\[
\sqrt{r_k + 1} = \frac{r_k + 1}{r_k - 1} \cdot \frac{\|\mu_k\|^2}{\lambda_{\min}^{(k)}}
\]  

for some \(r_k > 1\). Squaring (29) we obtain a quadratic equation whose rightmost solution is

\[
r_k^* = 1 + \frac{\|\mu_k\|^2}{2 \lambda_{\min}^{(k)}} + \frac{\|\mu_k\|^2}{2 \lambda_{\min}^{(k)}} \sqrt{8 + \frac{\|\mu_k\|^2}{\lambda_{\min}^{(k)}}}.
\]

Note that (29) is equivalent to

\[
\sqrt{r_k + 1} = (r_k - 1) \frac{\lambda_{\min}^{(k)}}{\|\mu_k\|}
\]

if \(r_k > 1\). Plugging the optimal \(r_k^*\) to the right-hand side of this last equation, we obtain the desired value of \(p^*\). The set \(X(p)\) is convex by Theorem 4.1.

To illustrate the tightness of this last bound on \(p\), we note that for the case \(\mu_k = 0\) the bound provided by Corollary 4.2 is \(p^* = \Phi(\sqrt{2}) \approx 0.921\); for \(\mu_k \sqrt{\lambda_{\min}^{(k)}} = 1\) it is still \(p^* = \Phi(2) \approx 0.977\). Both these result are less conservative than the result provided by Henrion and Strugarek [2008].

Example 4.1. Let \(t_k^T \sim N(\mu_k, \Sigma_k > 0)\) be the (mutually) independent random rows of a matrix \(T\). Then the problem (2) can be equivalently formulated as

\[
\min c^T x \quad \text{subject to } \mu_k^T x + \Phi^{-1}(p^{\lambda_k\theta}) \sqrt{x^T \Sigma_k x} \leq h_k,
\]

\[
\sum_k y_k = 1
\]

\[
x \in X, \quad y_k \geq 0 \quad \text{with } k = 1, \ldots, K,
\]

where \(\Phi\) is one-dimensional standard normal distribution. Indeed, recall that \(\psi(t) = -\ln t\) for the independent copula, hence \(\psi^{-1}(s) = e^{-s}\) and \(\psi^{-1}(y_k \psi(p)) = p^{\lambda_k\theta}\). This reformulation was given for example in Cheng and Lisser [2012].

Example 4.2. Let \(t_k^T \sim N(\mu_k, \Sigma_k > 0)\) be the random rows of a matrix \(T\) whose dependence is driven by the Gumbel–Hougaard copula with some given parameter \(\theta \geq 1\). Then the problem (2) can be equivalently formulated as

\[
\min c^T x \quad \text{subject to } \mu_k^T x + \Phi^{-1}(p^{\lambda_k^{1/\theta}}) \sqrt{x^T \Sigma_k x} \leq h_k,
\]

\[
\sum_k y_k = 1
\]

\[
x \in X, \quad y_k \geq 0 \quad \text{with } k = 1, \ldots, K,
\]

where \(\Phi\) is one-dimensional standard normal distribution. Indeed, recall that \(\psi(t) = (-\ln t)^\theta\) for the Gumbel–Hougaard copula, hence \(\psi^{-1}(s) = e^{-s^{1/\theta}}\) and \(\psi^{-1}(y_k \psi(p)) = p^{\lambda_k^{1/\theta}}\). Note the extreme values of the parameter \(\theta\): the case \(\theta = 1\) is that of Example 4.1 with independent rows, whereas if \(\theta \to +\infty\) then \(\Phi^{-1}(p^{\lambda_k^{1/\theta}}) \to \Phi^{-1}(p)\) — this is the problem with individual probabilistic constraints.
4.2 Special case: normal distribution and Gumbel–Hougaard copula

If the underlying distribution is normal and the dependence of the rows is driven by the Gumbel–Hougaard copula we can get even better (although implicit) threshold on the probability \( p \) for the convexity of the feasible set \( X(p) \).

Lemma 4.3. Let \( \theta \geq 1, r > 1, p \geq \frac{1}{2} \). Denote \( \Phi \) the standard normal distribution function and \( f_\Phi \) its associated density. If

\[
p \geq \Phi \left( \frac{r + 1}{\sqrt{r + 1 + \left( \frac{1-\theta}{2p \ln p} f_\Phi(\sqrt{r + 1}) \right)^2 + \frac{1-\theta}{2p \ln p} f_\Phi(\sqrt{r + 1})}} \right)
\]

then the function

\[
G(y) := \left[ \Phi^{-1} \left( p^{\theta/\theta} \right) \right]^{-r}
\]

is concave on \((0; 1)\).

Proof. Let \( h(y) := \left[ \Phi^{-1} \left( p^{\theta/\theta} \right) \right]^{-1} \). As \( p \geq \frac{1}{2} \), \( h(y) \) is positive on \((0; 1)\). The second derivative of the function \( G(y) = h(y)^r \) is

\[
G''(y) = r(r - 1)h(y)^{-2}h'(y)^2 + r h(y)^{r-1}h''(y).
\]

To check \( G''(y) < 0 \) it is enough to check

\[
(r - 1)h'(y)^2 + h(y)h''(y) < 0. \tag{30}
\]

Using the relations

\[
[\Phi^{-1}(t)]' = \frac{1}{f_\Phi(F^{-1}(t))} \quad \text{and} \quad \frac{f_\Phi(t)}{f_\Phi(t)} = -t
\]

valid for standard normal distribution we calculate

\[
h'(y) = -\frac{1}{\theta} \ln p \, p^{\theta/\theta} \, y^{\theta-1} h(y)^2 \, f_\Phi(h(y)^{-1})^{-1}
\]

\[
h''(y) = h'(y) \left\{ \frac{1}{\theta} \ln p \, y^{\theta-1} + \left( \frac{1}{\theta} - 1 \right) \frac{1}{y} + h'(y) \left( 2h(y)^{-1} - h(y)^{-3} \right) \right\}.
\]

Plugging these into (30), simplifying it and noting that \( h'(y) > 0 \) we conclude that \( G(y) \) is concave if

\[
W(y) := h(y) \left[ \frac{1}{\theta} \ln p \, y^{\theta-1} + \left( \frac{1}{\theta} - 1 \right) \frac{1}{y} + h'(y) \left[ r + 1 - h(y)^{-2} \right] \right] < 0.
\]

The first term of \( W(y) \) is negative; hence a sufficient condition for \( W(y) \) being negative is

\[
h(y)^{-1} \geq \sqrt{r + 1} \tag{31}
\]

for all \( y \in (0; 1) \). Note that this conditions is exactly the condition \( p > \Phi \left( t^*(r + 1) \right) = \Phi \left( \sqrt{r + 1} \right) \) provided by (26) in Theorem 4.1, as \( \inf_{y \in (0; 1)} \, p^{\theta/\theta} = p \).

If (31) is not satisfied, we can compensate the positivity of the second term of \( W(y) \) by the first (negative) term or some part of it. We will proceed by verifying

\[
h(y) \left( \frac{1}{\theta} - 1 \right) \frac{1}{y} + h'(y) \left[ r + 1 - h(y)^{-2} \right] < 0
\]
that is
\[
\frac{(1 - \theta) f_\Phi(h(y)^{-1})}{\ln p \ p^{\theta/y} y^{1/y} h(y)} - (r + 1) + h(y)^{-2} > 0
\]

The function \( y \mapsto p^{\theta/y} y^{1/y} \) is non-decreasing on the interval \((0; 1]\) and the density \( f_\Phi(t) \) is decreasing if \( t \geq 0 \). Together with assumption that \( h(y)^{-1} < \sqrt{r + 1} \) we obtain
\[
\frac{(1 - \theta) f_\Phi(h(y)^{-1})}{\ln p \ p^{\theta/y} y^{1/y} h(y)} - (r + 1) + h(y)^{-2} \geq \frac{(1 - \theta) f_\Phi(\sqrt{r + 1})}{p \ln p} h(y)^{-1} - (r + 1) + h(y)^{-2}
\]

To check the positivity of these expressions it is enough to verify
\[
\left( h(y)^{-1} + \frac{(1 - \theta) f_\Phi(\sqrt{r + 1})}{2p \ln p} \right)^2 - \left( \frac{(1 - \theta) f_\Phi(\sqrt{r + 1})}{2p \ln p} \right)^2 - (r + 1) > 0
\]
that is
\[
h(y)^{-1} > \sqrt{r + 1 + \left( \frac{(1 - \theta) f_\Phi(\sqrt{r + 1})}{2p \ln p} \right)^2 - \frac{(1 - \theta) f_\Phi(\sqrt{r + 1})}{2p \ln p}} - \frac{r + 1}{\sqrt{r + 1 + \left( \frac{(1 - \theta) f_\Phi(\sqrt{r + 1})}{2p \ln p} \right)^2 + \frac{(1 - \theta) f_\Phi(\sqrt{r + 1})}{2p \ln p}}}.
\]

Finally, \( h(y)^{-1} := \Phi^{-1}(p^{\theta/y}) \) together with \( p^{\theta/y} \geq p \) for all \( y \in (0; 1] \) we conclude on the condition
\[
p \geq \Phi\left( \frac{r + 1}{\sqrt{(r + 1) + \left( \frac{1 - \theta}{2p \ln p} f_\Phi(\sqrt{r + 1}) \right)^2 + \frac{1 - \theta}{2p \ln p} f_\Phi(\sqrt{r + 1})}} \right)
\]
ensuring that \( G''(y) < 0 \).

**Theorem 4.4.** Suppose the rows of the matrix \( T \) are normally distributed with means \( \mu_k \) and covariance matrices \( \Sigma_k > 0 \), and the joint distribution of these rows is driven by the Gumbel–Hougaard copula with parameter \( \theta \geq 1 \). If \( p > \Phi(\max\{u^*, v^*\}) \) where
\[
u^* := \max_{k=1,\ldots,K} \frac{r_k + 1}{r_k} \left( \frac{\lambda^{(k)}}{\lambda_{\min}^{(k)}} \right)^{-1/2} \|\mu_k\|
\]
\[
v^* := \max_{k=1,\ldots,K} \frac{r_k + 1}{\sqrt{(r_k + 1) + \left( \frac{1 - \theta}{2p \ln p} f_\Phi(\sqrt{r + 1}) \right)^2 + \frac{1 - \theta}{2p \ln p} f_\Phi(\sqrt{r + 1})}}
\]
for some \( r_k > 1 \) then the set \( X(p) \) is convex.

**Proof.** Now easy as \( G(y) - g_k(x)^{-r_k} \) is concave.

### 4.3 Approximation Method

Formulation (25) of problem (2) is still not a second-order cone program due to decision variables appearing as arguments to the (nonlinear) quantile functions \( \Psi_k^{(-)} \). To resolve the issue, we formulate lower approximation to the problem (25) by using favorable properties of Archimedean generators. We also provide an upper approximation based on sequential updating of the variables \( y_k \).
4.3.1 Lower bound: piecewise tangent approximation

We first formulate an auxiliary convexity lemma which gives us the possibility to find this approximation.

**Lemma 4.5.** If

1. \( \Psi \) is a distribution function induced by the characteristic function \( \phi(t) = \varphi(t^2) \) where \( \varphi \) is the characteristic generator of an elliptical distribution,

2. the associated density is 0-decreasing with some threshold \( t^*(0) > 0 \),

3. \( p > p^* = \Psi(t^*(0)) \), and

4. \( \psi \) is the generator of an Archimedean copula,

then the function

\[
y \mapsto \Psi^{-1} \left( \psi^{-1}(y \psi(p)) \right)
\]

is convex on \([0; 1]\).

**Proof.** Corollary 2.4 claims that \( \psi \) is strictly decreasing convex function on \([0; 1]\), hence \( \psi^{-1} \) is strictly decreasing convex function on \([0; \psi(0)]\) with values in \([p; 1]\). The second assumption implies the concavity of \( \Psi(\cdot) \) on \((t^*(0), +\infty)\), hence the convexity of \( \Psi^{-1}(\cdot) \) on \((p^*; 1)\). Together with the third assumption, and the fact that \( \Psi^{-1} \) is a quantile function, hence non-decreasing, the assertion of lemma is proved.

The proposed approximation technique follows the outline appearing in Cheng and Lisser [2012] and Cheng et al. [2012]. For each variable \( y_k \), consider a partition of the interval \((0; 1]\) in the form \( 0 < y_{k1} < \ldots < y_{kJ} \leq 1 \).

**Theorem 4.6.** Consider the problem

\[
\begin{align*}
\min & \ c^T x \quad \text{subject to} \\
& \mu_k^T x + \sqrt{z^T \Sigma_k z} \leq h_k, \\
& \Sigma_k w^k = x, \\
& x \in X, \ w^k \geq 0, \ z^k \geq 0 \ 	ext{with } k = 1, \ldots, K, \ j = 1, \ldots, J,
\end{align*}
\]

(33)

where

\[
\begin{align*}
& a_{kj} := H_k(y_{kj}) - b_{kj}y_{kj}, \\
& b_{kj} := \frac{\psi(p)}{f_k(H(y_{kj})) \psi' \left( \psi^{-1}(y_{kj} \psi(p)) \right)}, \\
& H_k(y) := \Psi_k^{-1} \left( \psi^{-1}(y \psi(p)) \right),
\end{align*}
\]

and \( f_k \) is the density function associated with the distribution function \( \Psi_k \). If \( X \subset \mathbb{R}^n_+ \) then the optimal value of (33) is a lower bound for the optimal value of the problem (2).

\[ ^{1} \text{The number } J \text{ of partition points can differ for each row } k \text{ but, to simplify the notation and without loss of generality, we consider this number to be the same for each } k. \]
Proof. Fix a row $k$. The first order Taylor approximations of $H_k(y)$ at each point $y_{kj}$ of the partition are given by

$$T_{H_k(y_{kj})}(y) = H(y_{kj}) + H'(y_{kj}) (y - y_{kj}).$$

Using the simple fact that

$$[\psi^{(-1)}(y_{kj}\psi(p))]' = \Psi_k \left( \Psi_k^{(-1)} \left[ \psi^{(-1)}(y_{kj}\psi(p)) \right] \right)'$$

we obtain explicitly the derivative $H'_k(y_{kj})$ in terms of $\Psi_k = f_k$ and we can continue by

$$T_{H_k(y_{kj})}(y) = H_k(y_{kj}) + \frac{\psi(p)}{f_k(H_k(y_{kj}))} \cdot (\psi^{(-1)})'(y_{kj}\psi(p)) \cdot (y - y_{kj}).$$

The derivative $(\psi^{(-1)})'$ is obtained similar way; finally we have

$$T_{H_k(y_{kj})}(y) = H_k(y_{kj}) + \frac{\psi(p)}{f_k(H_k(y_{kj}))} \cdot (\psi^{(-1)})'(y_{kj}\psi(p)) \cdot (y - y_{kj}) =: a_{kj} + b_{kj}y,$$

where $a_{kj}$ and $b_{kj}$ are given by Theorem 4.6. According to Lemma 4.5, the function $H_k(y)$ is convex on $(0; 1]$ hence the piecewise-linear function $\{a_{kj} + b_{kj}y\}$ is a lower bound for $H_k(y)$.

Introducing auxiliary decision vectors $z^k$ fulfilling (33), and $w^k := y_kx$, the final problem formulation is an SOCP problem and the proof is then completed. $\Box$

Remark 4.1. The linear functions $a_{kj} + b_{kj}y$ are tangent to the (quantile) function $H_k$ at the partition points; hence the origin of the name tangent approximation. This approximation leads to an outer bound for feasible solution set $X(p)$.

4.3.2 Upper bound: sequential approximation

The idea of the sequential method was taken from ? when the variables $y_k$, $k = 1, \ldots, K$ are fixed, the problem (25) becomes an SOCP problem. Therefore, we propose a sequential approximation method that iteratively adjusts the parameters $y_k$ and solves the updating SOCP problem until no further improvement is achieved.

Algorithm 1 (Sequential Approximation Procedure).

Step 1 (Initialization). Let $y^1 = (y^1_1, \ldots, y^1_K) \in \mathbb{R}_+^K$ be initial scaling parameters, i.e.,

$$\sum_{k=1}^K y^1_k = 1, \quad y^1_k \geq 0.$$

Set the iteration counter to $t := 1$.

Step 2 (Update). Solve problem (25) with $y_k$ fixed to $y^t_k$ and let $x^t$ and $f^t$ denote an optimal solution and the optimal value, respectively. Let

$$y_k(x) = \frac{b_k - \mu_k x}{\sqrt{x^t \Sigma_k x}} \quad \text{and} \quad \tilde{y}_k^t = \frac{\psi(\Psi_k(y^t_k))}{\psi(p)}.$$

Set

$$y^{t+1} \leftarrow U(y^t; \tilde{y}^t)$$

for some update policy $U : \Delta^K \times [0, 1]^K \to \Delta^K$ where

$$\Delta^K := \left\{ y \in \mathbb{R}^K \mid y \geq 0, \sum_{k=1}^K y_k = 1 \right\}$$

is the probability simplex.
Step 3 (Stopping criterion). If
\[ \left| \frac{f^t - f^{t-1}}{f^{t-1}} \right| \leq \varepsilon \]

where \( \varepsilon \) is a tolerance parameter, stop and return \( x^t, f^t, \) and \( y^t \). Otherwise set \( t := t + 1 \) and go to Step 2 (Update).

**Theorem 4.7.** If the problem (25) is bounded, has a feasible solution with the initial values \( y^1 \), and \( U(y; \bar{y}) \geq \bar{y} \) for all \( y \geq \bar{y} \), then Algorithm 1 terminates in a finite number of steps, and the returned value \( f^T \) is an upper bound of the problem (25).

**Proof.** First of all, we prove that the sequence of values \( f^t \) produced by the algorithm is non-increasing and it is sufficient to show that the solution \( x_t \) is still feasible in the problem (25) when \( y \) changes from \( y^t \) to \( y^{t+1} \). Indeed, as \( \bar{y}_k \) captures the minimum amount that \( y_k \) should be to allow \( x^t \) to be feasible with respect to the \( k \)-th constraint, i.e., \( \mu_k x + \Psi_k^{-1} \left( \psi_k^{-1}(y_k \psi(p)) \right) \sqrt{x^T \sum_k x} \leq h_k \), and \( \Psi_k^{-1} \left( \psi_k^{-1}(y_k \psi(p)) \right) \) is non-increasing on the interval \((0; 1]\), then \( x^t \) remains feasible with \( y^{t+1} = U(y^t; \bar{y}^t) \geq \bar{y}^t \).

Second, since by the definition of the update policy each term of the sequence of \( y^t \) is feasible according to this problem, then every value of \( f^t \) is an upper bound on the optimal value of problem (25). Moreover, given that the sequence is nonincreasing, and that it is bounded by the optimal value of problem (25) which is bounded, this implies that the sequence \( \{f^t\} \) will converge to a finite limit \( f^\infty \). We are therefore guaranteed that for any fixed tolerance level \( \varepsilon \), after a finite number of steps \( T \), the difference
\[ \left| \frac{f^T - f^T}{f^T} \right| \leq \left| \frac{f^\infty - f^T}{l^*} \right| \leq \varepsilon \]

where \( l^* \) is a lower bound of \( |f^T| \), hence the algorithm should have terminated. \( \square \)

For the numerical implementation, we apply
\[ U(y; \bar{y}) = \bar{y} + \alpha (1 - (y - \bar{y})) \]

where
\[ \alpha = \frac{\sum_k (y_k - \bar{y}_k)}{\sum_k (1 - (y_k - \bar{y}_k))} \]
as the update policy. First of all, we can verify easily that this adjustment policy satisfies the required properties mentioned above in Theorem 4.7. As it is evident that both \( \alpha \) and \( 1 - (y - \bar{y}) \) are nonnegative, we have \( U(y; \bar{y}) \geq \bar{y} \geq 0 \). In addition, the vector generated through \( U(y; \bar{y}) \) should also satisfy the constraint \( \sum_k y_k = 1 \).
\[ 1^T U(y; \bar{y}) = \sum_k \bar{y}_k + \alpha \sum_k (1 - (y_k - \bar{y}_k)) = \sum_k \bar{y}_k + \sum_k (y_k - \bar{y}_k) = 1 \]

Second, we can see from this adjustment policy that under the current solution \( x^t \), more tighter one chance constraint \( k \) is, more margin is given to the corresponding \( \bar{y}_k^{t+1} \). The tightness of one chance constraint can be measured by \( y_k^t - \bar{y}_k^t \).

## 5 Numerical Study

In this section, we evaluate numerically the performance of our proposed methods. All the models considered later were solved using SeDuMi 1.3 (Sturm [1999]) with their default parameters on an Intel Core 2 Duo @ 2.26 GHz with 4.0 GB RAM.
We performed computational tests on a stochastic version of the resource constrained shortest path problem (RCSP) where the resources consumed by traversing the arc are assumed to be random. The RCSP can be mathematically formulated as

\[
\min c^T x \quad \text{subject to} \quad Tx \leq d, \ Mx = b, \ x \in \{0, 1\}^n
\]

where \(c \in \mathbb{R}^n\), \(M \in \mathbb{R}^{m \times n}\), which is the node-arc incidence matrix (\(T\)), and \(b \in \mathbb{R}^m\), where all elements are 0 except the \(s\)-th and the \(t\)-th which are 1 and \(-1\), respectively. \(T\) is a \(K \times n\) matrix, \(d\) is a positive vector of \(K\) elements. The corresponding linear relaxation of the stochastic RCSP with joint probabilistic constraints is written as

\[
\min c^T x \quad \text{subject to} \quad \mathbb{P}\{Tx \leq d\} \geq p, \ Mx = b, \ x \geq 0
\]

When \(t_k\) is uniformly distributed over the ellipsoid support \(S = \{\xi : (\xi - \mu_k)^T \Sigma_k^{-1} (\xi - \mu_k) \leq n + 3\}\), which is elliptically distributed, then \(\mathbb{E}t_k = \mu_k\) and \(\mathbb{E}(t_k - \mu_k)(t_k - \mu_k)^T = \Sigma_k\). Further, we assume the dependence of random rows of the matrix \(T\) to be driven by the Gumbel–Hougaard copula with some given parameter \(\theta \geq 1\) whose equivalent formulation can be seen in Example 4.2 (except the quantile function). The stochastic program is equivalent to the problem

\[
\min c^T x \quad \text{subject to} \quad \mu_k^T x + \sqrt{(n + 3)\Psi_b^{(-1)}(2p^{1/\theta} - 1)} \sqrt{x^T \Sigma_k x} \leq d_k
\]

\[
\sum_k y_k = 1, \ y_k \geq 0 \ \forall k = 1, \ldots, K, \ Mx = b, \ x \geq 0
\]

where \(\Psi_b^{(-1)}(\cdot)\) is the quantile function of the beta distribution with shape parameters \((1/2; n/2 + 1)\).

We perform our tests on five randomly generated instances of RCSP problem with \((n, m, K) = (21, 39, 10)\), while the input data for the models is randomly generated as follows:

- the cost \(c\) is uniformly generated on the interval \([1, 10]\);
- the mean \(\mu_k\) is uniformly generated on the interval \([5, 10]\);
- the entries of the covariance matrix \(\Sigma_k\) are generated by the MATLAB function \texttt{gallery(’randcorr’,n)*1};
- resource threshold parameter \(d_k\) is uniformly generated on the interval \([30, 100]\).

Confidence parameter is set to \(p = 0.9\). Moreover, we choose eleven tangent points

\[
y_{k1} = 0.01, \quad y_{ki} = 0.1 \ast (i - 1), \ i = 2, \ldots, 11
\]

for the tangent approximation (same for each \(k\)). For the sequential approximations, we set the initial parameter \(y_k\) to \(\frac{1}{K}\) and the tolerance parameter \(\varepsilon\) is set to \(10^{-4}\). Concerning the Gumbel–Hougaard copula, its dependence parameter \(\theta\) is set to 1 (independence), 2 (moderate dependence), and 5 (high dependence), respectively.

The numerical results over five instances for each choice of \(\theta\) are given in Table 3: column one gives the name of the instance, columns two and three show the optimal objective values \(V^{TA}\) of the tangent approximation and the corresponding CPU time, whereas columns four and five give the optimal objective values \(V^{SA}\) and the CPU time of the sequential approximation. The number of iterations for sequential approximations is given in sixth column, while the gap

\[
Gap = \frac{V^{SA} - V^{TA}}{V^{SA}} \cdot 100\%
\]

between the two approximations is given in the last column of the table.
<table>
<thead>
<tr>
<th>( \theta = 1 )</th>
<th>( V^{TA} )</th>
<th>CPU (s)</th>
<th>( V^{SA} )</th>
<th>CPU (s)</th>
<th>Iter</th>
<th>Gap(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inst01</td>
<td>19.5</td>
<td>14.4</td>
<td>19.5</td>
<td>3.9</td>
<td>15</td>
<td>0.0</td>
</tr>
<tr>
<td>Inst02</td>
<td>24.5</td>
<td>14.0</td>
<td>24.9</td>
<td>1.8</td>
<td>6</td>
<td>1.6</td>
</tr>
<tr>
<td>Inst03</td>
<td>19.6</td>
<td>13.9</td>
<td>20.1</td>
<td>2.4</td>
<td>9</td>
<td>2.5</td>
</tr>
<tr>
<td>Inst04</td>
<td>25.8</td>
<td>14.1</td>
<td>26.0</td>
<td>2.9</td>
<td>9</td>
<td>0.8</td>
</tr>
<tr>
<td>Inst05</td>
<td>28.1</td>
<td>14.6</td>
<td>29.3</td>
<td>9.7</td>
<td>34</td>
<td>4.1</td>
</tr>
<tr>
<td>Average</td>
<td>–</td>
<td>14.2</td>
<td>–</td>
<td>4.1</td>
<td>15</td>
<td>1.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \theta = 2 )</th>
<th>( V^{TA} )</th>
<th>CPU (s)</th>
<th>( V^{SA} )</th>
<th>CPU (s)</th>
<th>Iter</th>
<th>Gap(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inst01</td>
<td>19.5</td>
<td>13.8</td>
<td>19.5</td>
<td>2.6</td>
<td>10</td>
<td>0.0</td>
</tr>
<tr>
<td>Inst02</td>
<td>24.5</td>
<td>14.6</td>
<td>24.8</td>
<td>1.1</td>
<td>4</td>
<td>1.2</td>
</tr>
<tr>
<td>Inst03</td>
<td>19.6</td>
<td>13.4</td>
<td>20.0</td>
<td>1.5</td>
<td>5</td>
<td>2.0</td>
</tr>
<tr>
<td>Inst04</td>
<td>25.6</td>
<td>13.5</td>
<td>25.8</td>
<td>2.7</td>
<td>8</td>
<td>0.8</td>
</tr>
<tr>
<td>Inst05</td>
<td>28.1</td>
<td>14.3</td>
<td>29.1</td>
<td>5.3</td>
<td>18</td>
<td>3.4</td>
</tr>
<tr>
<td>Average</td>
<td>–</td>
<td>13.9</td>
<td>–</td>
<td>2.6</td>
<td>9</td>
<td>1.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \theta = 5 )</th>
<th>( V^{TA} )</th>
<th>CPU (s)</th>
<th>( V^{SA} )</th>
<th>CPU (s)</th>
<th>Iter</th>
<th>Gap(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inst01</td>
<td>19.5</td>
<td>13.4</td>
<td>19.5</td>
<td>1.3</td>
<td>5</td>
<td>0.0</td>
</tr>
<tr>
<td>Inst02</td>
<td>24.5</td>
<td>14.3</td>
<td>24.7</td>
<td>0.7</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>Inst03</td>
<td>19.6</td>
<td>12.9</td>
<td>19.8</td>
<td>1.0</td>
<td>3</td>
<td>1.0</td>
</tr>
<tr>
<td>Inst04</td>
<td>25.5</td>
<td>15.1</td>
<td>25.6</td>
<td>2.2</td>
<td>7</td>
<td>0.4</td>
</tr>
<tr>
<td>Inst05</td>
<td>28.1</td>
<td>14.3</td>
<td>28.7</td>
<td>2.8</td>
<td>8</td>
<td>2.1</td>
</tr>
<tr>
<td>Average</td>
<td>–</td>
<td>14.0</td>
<td>–</td>
<td>1.6</td>
<td>5</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 3: Computational results
In Table 3, we can first observe that for all the instances, the optimal values of the two approximations, that is $V^{TA}$ and $V^{SA}$, are nonincreasing which is coherent with the fact that the joint chance constraint becomes less restrictive as $\theta$ increases. Second, the CPU time for the tangent approximation is within 16 seconds and there is no big difference for the different value of $\theta$. Of course, the CPU time can be cut by reducing the number of the tangent points but may deteriorate the quality of solutions (Cheng and Lisser [2012]). But for the sequential approximation, the CPU time is within 5 seconds and decreases as $\theta$ increases, which is due to the fact that the number of iterations decreases as $\theta$ increases. Last but most importantly, column seven shows that the sequential approximation and the tangent approximation provide the competitive upper and lower bounds of the problem (25), as shown by the obtained small gaps (less than 4.2%), which also means that the gap between the optimal value of the mother problem and the optimal value of the two approximations is less than 2.1%. In addition, the gap is decreasing as $\theta$ increases, i.e., higher dependence leads to a lower gap.

6 Conclusion

In this paper, we studied the problem of linear joint probabilistic constraints where the distribution of the constraint rows is elliptically distributed and the dependence of the rows is depicted by a convenient Archimedean copula. We also investigated the convexity of the set of feasible solutions and provided improved convexity results. Despite the fact that the problem studied is generally nonconvex, two approximations are provided to solve the problem, one of which is a relaxed approximation whereas the other is conservative approximation. Finally, numerical experiments were given to indicate that the two proposed approximations are competitive.

References


