Distributionally Robust Chance Constrained Games under Wasserstein Ball

Tian Xia[a], Jia Liu[a], Abdel Lisser[b]

[a] *School of Mathematics and Statistics, Xi'an Jiaotong University, 710049, Xi'an, P. R. China,*
[b] *CentraleSupelec, Laboratoire des Signaux et des Systemes, 91190 Gif-sur-Yvette, France,*

Abstract

This paper considers distributionally robust chance constrained games with a Wasserstein distance based uncertainty set. We assume that the center of the uncertainty set is an elliptical distribution. We derive a tractable reformulation and an efficient solution approach to the Nash equilibrium of distributionally robust chance constrained games. Numerical results show the price and benefit of the robust model compared with the non-robust model.

Keywords: Chance constrained games, Wasserstein ball, Elliptical distribution, Nash equilibrium

---

1. Introduction

The literature on the existence of an equilibrium in game theory started since the paper by John von Neumann [1], who showed the existence of a mixed strategy saddle point equilibrium for a two-player zero-sum matrix game. Then in 1950, John Nash [2] showed the existence of a mixed strategy Nash equilibrium for a finite strategic game, which brought the research on the existence of equilibria in game theory to a new stage. Since John von Neumann and John Nash, the traditional games with deterministic payoffs of the players have been widely studied. However, real world problems are significantly subject to uncertainties. Therefore, games with random payoffs are of increasing concern in game theory. Recently, Singh et al. [3, 4, 5, 6] initiated the studies on chance constrained games. By using chance constrained optimization models, we actually consider the random games where the players' payoffs are obtained with a certain confidence. The chance constrained payoffs represent the maximal threshold such that random payoffs are not less than the threshold with a large probability, e.g., 95% or 90%. In a traditional chance constrained game, each player should know the exact distribution of the random return/payoff, or specify an a-priori distribution before making the decision. However, due to the imperfectness of the historical data, and incompleteness of the information collection, the estimated a-priori distribution may be biased from the true distribution. If the player does not consider the ambiguity/impreciseness of the distribution and just use the traditional chance constraints, he might over-estimate the payoff at the equilibrium and makes an inefficient decision in the game[7]. Thus, to reduce the potential loss in extreme cases, it is natural to consider the ambiguity set of the distribution in the decision-making model and use the distributionally robust optimization (DRO) approach to make a decision against the worst-case distribution. The literature on distributionally robust chance constrained games (DRCCG) mainly focuses on the existence of Nash equilibrium when the information to each player is characterized by different kinds of uncertainty sets. For example, Singh et al. [4] considered DRCCG with a moment-type uncertainty set. Xu and Zhang [8] considered the convergence of the sample average approximation in DRCCG. Peng et al. [6] studied DRCCG under the divergence distance based uncertainty set and showed the existence of Nash equilibrium.

Wasserstein ball is also an important kind of uncertainty set widely used in distributionally robust optimization. Compared with phi-divergence distance based models, the Wasserstein distance based model allows that the support of the worst-case distribution is not necessarily contained in the support of the dataset/reference distribution. Thus, the utilization of Wasserstein distance ensures players make use of any a-priori information of support, which other models may ignore. Most of the studies on Wasserstein ball focus on the data-driven case [9, 10, 11, 12, 13, 14], where the reference distribution is a discrete distribution. Xie [12] gave an exact reformulation of the distributionally robust chance constraint with a data-driven reference distribution. Liu et al. [15] studied distributionally robust chance constrained geometric optimization. The discrete distributions based data-driven reformulations were generalized by [16, 17] to Polish spaces and continuous distributions. Shen and Jiang [18] considered the distributionally robust chance constraint where the reference distribution in the Wasserstein ball is a Gaussian distribution. Peng et al. [14] studied distributionally robust games with expected utility functions and data-driven Wasserstein ball. To the best of our knowledge, DRCCG under Wasserstein ball has not been studied in the literature.

In this paper, we study DRCCG under the Wasserstein ball. We consider the reference distribution as an elliptical distribution, i.e., a distribution from a large family of continuous distributions. A data-driven Wasserstein ball and a continuous reference distribution play different rules in distributionally robust optimization. The former allows us to calibrate and evaluate the size of the ambiguity set. The latter can be viewed as an adjustment of the over-optimism of the decision maker's a-priori distribution information. The radius of the ball reflects the strength of confidence in his a-priori information. In many applications, the decision makers ignore the fact that the a-priori distributions are not Gaussian.For instance, wind power and electric load forecasting errors are generally not Gaussian distributed in power system scheduling problems. The stock return rates are often regarded as high kurtosis and fat-tailed[19]. The elliptical distributions are a broad family of probability distributions that generalize the multivariate normal distribution, which thus play an important role in stochastic games.

As far as we know, this paper provides the first contribution which considers an elliptical reference distribution in a Wasserstein ball-based distributionally robust game. We propose a new approach which leads to the condition of the convexity of the chance constrained payoff with the Wasserstein ball and derive an efficient solution method to the equilibrium problem of this kind of games.

The paper is organized as follows. We introduce the DRCCG model under Wasserstein ball in Section 2. We derive the reformulation of DRCCG in Section 3.1 and Section 3.2, and show the existence of a Nash equilibrium in Section 3.3. We propose an optimization approach to find Nash equilibrium of DRCCG in Section 3.4. We carry out numerical tests under some popular distributions from the elliptical distribution family in Section 4.

2. DRCCG under Wasserstein ball

2.1. Introduction to chance constrained games

We consider a n-player strategic game. Let $I = \{1, 2, ...n\}$ denotes the set of players. For each $i \in I$, A_i represents a finite action set of player i and its generic element is denoted by a_i. The vector $a = (a_1, a_2, ...a_n)$ denotes the action profile of the game. The set of all action profiles is denoted by the product set $A = \times_{i=1}^{n} A_i$. We denote $A_{-i} = \times_{j=1;j\neq i}^{n} A_j$, and $a_{-i} = (a_1, a_2, .., a_{i-1}, a_{i+1}, .., a_n) \in A_{-i}$. Let X_i be the set of mixed strategies of player i which is a subset of all probability distributions over the action set A_i. A mixed strategy $\tau_i \in X_i$ is represented by $\tau_i = (\tau_i(a_i))_{a_i \in A_i}$, where $\tau_i(a_i) \geqslant 0$ is the probability with which player i chooses an action a_i and $\sum_{a_i \in A_i} \tau_i(a_i) = 1$. The set of all mixed strategy profiles is denoted by $X = \times_{i=1}^{n} X_i$ and a mixed strategy profile $\tau = (\tau_i)_{i\in I}$ is a generic element of X. Denote $X_{-i} = \times_{j=1;j\neq i}^{n} X_j$ and $\tau_{-i} \in X_{-i}$ is a vector of mixed strategies $\tau_j, j \neq i$. We define $(\hat{\tau}_i, \tau_{-i})$ to be a strategy profile where player i uses the strategy $\hat{\tau}_i$ and each other player $j, j \neq i$ uses strategy τ_j. Denote $l_i(a) \in \mathbb{R}$ the payoff of player i for an action profile a, and $r_i \in \mathbb{R}^{|A|}$ the payoff vector of the game where $|A|$ is the cardinality of set A. For such mixed strategies games, Nash [2] showed that there always exists a Nash equilibrium.

In some practical applications, the players' payoffs are random due to several uncertainty sources. Thus, it is natural to study stochastic games by using stochastic optimization approach. Let $(\Omega, \mathcal{F}, P)$ be a probability space where $l_i : \Omega \to \mathbb{R}$ for each $i \in I$. We can view the random payoff vector r_i as a measurable function $r_i(\omega) = (l_i(a, \omega))_{a\in A} : \Omega \to \mathbb{R}^{|A|}$, whose distribution is F. Then for a given strategy $\tau \in X$ and a scenario $\omega \in \Omega$, the random payoff $\mathcal{R}_i$ of player i is

$$\mathcal{R}_i(\tau, \omega) = \sum_{a\in A} \left(l_i(a, \omega) \prod_{j\in I} \tau_j(a_j) \right). \tag{1}$$

We denote vector $\eta^\tau = (\eta^\tau(a))_{a\in A}$ for short, where $\eta^\tau(a) = \prod_{i\in I} \tau_i(a_i)$. Then $\mathcal{R}_i = r_i^\top \eta^\tau$.

Considering the randomness of the payoff $\mathcal{R}_i$, Singh et al. [3] study the following chance constrained payoff:

$$u_i^{\alpha_i}(\tau) = \sup \left\{ v_i | \mathbb{P}_F((r_i)^\top \eta^\tau \geqslant v_i) \geqslant \alpha_i \right\}, \tag{2}$$

which is the highest level of the player's payoff that he can attain with at least a specified level of confidence $\alpha_i \in (0,1)$. The confidence level is given in advance and known to all players. That is, the game we study is non-cooperative with complete information.

When the distribution of r_i is known and follows a multivariate normal distribution with a mean vector μ_i and a positive definite covariance matrix Σ_i, the chance constrained payoff equals

$$u_i^{\alpha_i}(\tau) = \sum_{a \in A} \left(\mu_i(a) \prod_{j \in I} \tau_j(a_j) \right) + \|\Sigma_i^{\frac{1}{2}} \eta^\tau\|_2 \phi^{-1}(1-\alpha_i),$$

where $\phi^{-1}(\cdot)$ is the quantile function of a standard normal distribution $N(0,1)$ [3].

2.2. Distributionally robust chance constrained games

In many practical situations, the probability distribution of r_i is not completely known to the players. Instead, the players know uncertainty sets $\mathcal{D}_i, i \in I$, where $\mathcal{D}_i$ is the set of all possible distributions of r_i. We assume that the uncertainty set of each player is known to all players in the game, and the players consider the worst case of their payoffs. Thus, player i holds distributionally robust chance constrained payoff function:

$$u_i^{\alpha_i}(\tau) = \sup \left\{ v_i | \inf_{F \in \mathcal{D}_i} \mathbb{P}_F((r_i)^\top \eta^\tau \geqslant v_i) \geqslant \alpha_i \right\}, \tag{3}$$

The game is called a distributionally robust chance constrained game (DRCCG) and was studied in [4, 6]. The set of best response strategies of player $i, i \in I$ against a given strategy profile τ_{-i} is

$$\mathrm{BR}_i^{\alpha_i}(\tau_{-i}) = \left\{ \tau_i^* \in X_i \mid u_i^{\alpha_i}(\tau_i^*, \tau_{-i}) \geqslant u_i^{\alpha_i}(\tau_i, \tau_{-i}), \forall \tau_i \in X_i \right\}.$$

Definition 1. *A strategy profile $\tau^* \in X$ is said to be a Nash equilibrium of a DRCCG for a given α, if for all $i \in I$, the following inequality holds,*

$$u_i^{\alpha_i}(\tau_i^*, \tau_{-i}^*) \geqslant u_i^{\alpha_i}(\tau_i, \tau_{-i}^*), \forall \tau_i \in X_i. \tag{4}$$

2.3. Wasserstein distance based uncertainty set

Since the true probability distribution of $r_i, i \in I$ is unknown, we replace the true distribution with a reference one which we might derive from some empirical data.

Thus, we consider the uncertainty set $\mathcal{D}_i, i \in I$ as the neighbourhood of the reference distribution $\hat{F}_i$:

$$\mathcal{D}_i = \{F : d_w(F, \hat{F}_i) \leqslant \delta_i\},$$

where $\delta_i > 0$ is a pre-specified radius of the uncertainty set $\mathcal{D}_i$. $d_w(\cdot, \cdot)$ is the Wasserstein distance and the set $\mathcal{D}_i$ is known as a Wasserstein ball.

Definition 2. *The Wasserstein distance d_w is defined by*

$$d_w(F, \hat{F}_i) = \inf_{\pi \in \Pi} \left\{ \int_{\Omega \times \Omega} \|\xi_1 - \xi_2\| \pi(d\xi_1, d\xi_2) \right\},$$

where Π is the space of all joint distributions of ξ_1 and ξ_2 with marginals F and $\hat{F}_i$ respectively, $\|\cdot\|$ is the norm defined as $\|\cdot\| := \|\Sigma_i^{-\frac{1}{2}}(\cdot)\|_2$ where Σ_i is a positive definite covariance matrix of the reference distribution $\hat{F}_i$. We denote $\|\cdot\|_*$ the dual norm of $\|\cdot\|$, i.e., $\|\cdot\|_* = \|\Sigma_i^{\frac{1}{2}}(\cdot)\|_2$ for each i.

3. Reformulation of DRCCG under elliptical reference distributions

3.1. Reformulation of DRCCG

In this section, we concentrate on the reformulation of the payoff function $u_i^{\alpha_i}(\tau), i \in I$ for given τ and α, when the uncertainty set $\mathcal{D}_i$ is a Wasserstein ball centered at an elliptical distribution.

We introduce the definitions of Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) of a random variable,

$$\mathrm{VaR}_{\alpha,F}(X) := \inf\left\{x : \mathbb{P}_F[X \leqslant x] \geqslant \alpha\right\},$$

$$\mathrm{CVaR}_{\alpha,F}(X) := \min_{\gamma}\left\{\gamma + \frac{1}{1-\alpha}\mathbb{E}_F[(X-\gamma)^+]\right\}.$$

It is well known that both VaR and CVaR are translation invariant, monotone and positive homogeneous. Additionally, CVaR is subadditive, which is a coherent risk measure.

Next, we focus on the reformulation of $u_i^{\alpha_i}(\tau)$ in (3).

Proposition 1. *For each player $i \in I$, $\delta_i > 0$ and $\alpha_i \in (0,1)$, suppose that the reference distribution $\hat{F}_i$ is continuous, then the distributional robust optimal (DRO) payoff (3) is equal to the optimal value of the following optimization problem:*

$$\max_{\hat{\alpha}_i, v_i \in \mathbb{R}} \quad v_i \tag{5a}$$

$$\text{s.t.} \quad \delta_i + (1-\alpha_i)\mathrm{CVaR}_{\alpha_i,\hat{F}_i}(-f_i(\tau, r_i)) \leqslant \sup_{\hat{\alpha}_i \in [0,1]} \left\{ (1-\hat{\alpha}_i)\mathrm{CVaR}_{\hat{\alpha}_i,\hat{F}_i}(-f_i(\tau, r_i)) \right\}, \tag{5b}$$

$$0 \geqslant \mathrm{VaR}_{\alpha_i,\hat{F}_i}(-f_i(\tau, r_i)) \tag{5c}$$

where $f_i(\tau, r_i) = \frac{-v_i + (r_i)^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}}$, and Σ_i is the covariance matrix of the reference distribution $\hat{F}_i$.

Proof. By Corollary 1 of [18], the distributionally robust chance constraint $\inf_{F \in \mathcal{D}_i} \mathbb{P}_F((r_i)^\top \eta^\tau \geqslant v_i) \geqslant \alpha_i$ is equivalent to the following group of inequalities,

$$\frac{\delta_i}{1-\alpha_i} + \mathrm{CVaR}_{\alpha_i,\hat{F}_i}(-f_i(\tau, r_i)) \leqslant \frac{1}{1-\alpha_i}\mathbb{E}_{\hat{F}_i}\left[(-f_i(\tau, r_i))^+\right], \tag{6a}$$

$$0 \geqslant \mathrm{VaR}_{\alpha_i,\hat{F}_i}(-f_i(\tau, r_i)). \tag{6b}$$

where $f_i(\tau, r_i) = \frac{-v_i + (r_i)^\top \eta^\tau}{\|\eta^\tau\|_*} = \frac{-v_i + (r_i)^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}}$, Σ_i is the covariance matrix of the reference distribution $\hat{F}_i$. Therefore the payoff (3) is equal to the maximal v_i subjecting to (6a) and (6b). Meanwhile,

$$\begin{aligned}
(6a) \;\Leftrightarrow\; & \frac{\delta_i}{1-\alpha_i} + \mathrm{CVaR}_{\alpha_i,\hat{F}_i}(-f_i(\tau, r_i)) \leqslant \frac{1}{1-\alpha_i}\mathbb{E}_{\hat{F}_i}\left[-f_i(\tau, r_i) \cdot \mathbf{1}\{-f_i(\tau, r_i) \geqslant 0\}\right], \\
\Leftrightarrow\; & \frac{\delta_i}{1-\alpha_i} + \mathrm{CVaR}_{\alpha_i,\hat{F}_i}(-f_i(\tau, r_i)) \leqslant \frac{1}{1-\alpha_i}\sup_{t_i \in \mathbb{R}}\mathbb{E}_{\hat{F}_i}\left[-f_i(\tau, r_i) \cdot \mathbf{1}\{-f_i(\tau, r_i) \geqslant t_i\}\right], \\
\Leftrightarrow\; & \delta_i + (1-\alpha_i)\mathrm{CVaR}_{\alpha_i,\hat{F}_i}(-f_i(\tau, r_i)) \leqslant \sup_{t_i \in \mathbb{R}}\mathbb{E}_{\hat{F}_i}\left[-f_i(\tau, r_i) \cdot \mathbf{1}\{-f_i(\tau, r_i) \geqslant t_i\}\right], \\
\Leftrightarrow\; & \delta_i + (1-\alpha_i)\mathrm{CVaR}_{\alpha_i,\hat{F}_i}(-f_i(\tau, r_i)) \leqslant \sup_{t \in \mathbb{R}}\left\{(1-g_i(t))\mathrm{CVaR}_{g_i(t),\hat{F}_i}(-f_i(\tau, r_i))\right\}, \\
, \\
\Leftrightarrow\; & \delta_i + (1-\alpha_i)\mathrm{CVaR}_{\alpha_i,\hat{F}_i}(-f_i(\tau, r_i)) \leqslant \sup_{\hat{\alpha}_i \in [0,1]}\left\{(1-\hat{\alpha}_i)\mathrm{CVaR}_{\hat{\alpha}_i,\hat{F}_i}(-f_i(\tau, r_i))\right\},
\end{aligned}$$

where $g_i(t) = 1 - \mathbb{P}_{\hat{F}_i}[-f_i(\tau, r_i) \geqslant t]$.

The first equivalence is by the definition of $[\cdot]^+$. The second equivalence comes from the fact that t_i reaches its optimal value at 0. The third equivalence is by multiplying

$(1-\alpha_i)$ to both sides of the inequality. The forth equivalence is a reformulation by the definition of CVaR. The last equivalence is due to the replacement of variable t by its quantile $\hat{\alpha}_i = g_i(t)$, which works for the assumption that $\hat{F}_i$ is continuous such that $g_i(t)$ is a continuous function.

$\square$

3.2. Reformulation of DRCCG under elliptical reference distributions

In this section, each player i uses an elliptical reference distribution and considers a worst-case payoff over a Wasserstein ball centered at an elliptical distribution.

Definition 3 ([20])**.** *A d-dimensional vector $X \in \mathbb{R}^d$ follows an elliptical distribution $\mathbb{E}_d(\mu, \Sigma, \psi)$ if the probability density function (PDF) is $f(x) = |\Sigma|^{-\frac{1}{2}} g((x-\mu)^\top \Sigma^{-1}(x-\mu))$, where $\mu \in \mathbb{R}^d$ is the location parameter, $\Sigma \in \mathbb{R}^{d\times d}$ is the dispersion matrix, ψ is the characteristic generator and $g : \mathbb{R}_+ \to 0$ is the density generator such that the Fourier transform of $g(|x|^2)$, as a generalized function, is equal to $\psi(|\xi|^2)$.*

By [20], for any matrix $A \in \mathbb{R}^{N\times d}$ and any vector $b \in \mathbb{R}^N$, we have $AX + b \sim \mathbb{E}_N(A\mu + b, A^\top \Sigma A, \psi)$.

Lemma 1 ([21, Theorem 1])**.** *For an elliptical distributed vector $\zeta \sim E_d(\mu, \Sigma, \psi)$, a real vector $w \in \mathbb{R}^d$, and $\alpha \in (0,1)$,*

$$\mathrm{CVaR}_\alpha(w^\top \zeta) = w^\top \mu + \sqrt{w^\top \Sigma w}\, T_\alpha, \tag{7}$$

where

$$T_\alpha = \frac{\pi^{\frac{d-1}{2}}}{2\alpha\Gamma(\frac{d+1}{2})} \int_{q_\alpha^2}^{\infty} (u - q_\alpha^2)^{\frac{d-1}{2}} g(u) du, \;\; q_\alpha = \frac{-w^\top \mu + \mathrm{VaR}_\alpha(w^\top \zeta)}{\sqrt{w^\top \Sigma w}},$$

and T_α denotes the value of $\mathrm{CVaR}_\alpha(w^\top \tilde{\zeta})$ when $\tilde{\zeta} \sim E_d(0, \Sigma, \psi)$ and $w^\top \Sigma w = 1$.

Remark 1. *There exists a slight difference between* (7) *and the definition in [21] because here* CVaR *is defined with respect to the loss function rather than the reward function in [21].*

Assumption 1. *For each player $i \in I$, the confidence level $\alpha_i \in (0,1)$ and the uncertainty set $\mathcal{D}_i$ is a Wasserstein ball centered at an elliptical distribution $E_{|A|}(\mu_i, \Sigma_i, \psi_i)$ with a radius $\delta_i > 0$.*

Theorem 1. *Given Assumption 1, the DRO payoff function* (3) *is equal to the following reformulation:*

$$u_i^{\alpha_i}(\tau) = (\eta^\tau)^\top \mu_i + \beta_i \sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}, \tag{8}$$

where

$$\beta_i = \sup_{\hat{\alpha}_i \in (\alpha_i, 1)} \frac{1}{(\hat{\alpha}_i - \alpha_i)} \left[(1-\hat{\alpha}_i) T_{\hat{\alpha}_i} - (1-\alpha_i) T_{\alpha_i} - \delta_i\right]. \tag{9}$$

Proof. By Proposition 1, we know that the DRO payoff function (3) is equal to the optimal value of the optimization problem (5a)-(5c). We consider the reformulation of (5b).

In (5b), the CVaR value is evaluated under the reference distribution $\hat{F}_i$ of r_i. By the translation invariance property of CVaR and Lemma 1, we have

$$\begin{aligned} &\mathrm{CVaR}_{\alpha_i, \hat{F}_i}(-f_i(\tau, r_i)) = \mathrm{CVaR}_{\alpha_i, \hat{F}_i}\left(-\frac{-v_i + (r_i)^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}}\right) \\ &= \mathrm{CVaR}_{\alpha_i, \hat{F}_i}\left(\frac{-(r_i - \mu_i)^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}}\right) + \frac{v_i - \mu_i^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}} = T_{\alpha_i} + \frac{v_i - \mu_i^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}} \end{aligned} \tag{10}$$

where $\frac{-(r_i-\mu_i)^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}}$ follows a uni-variate elliptical distribution $E_1(0, 1, \psi_i)$.

Taking (10) into (5b), we have

$$\begin{aligned} &(5\mathrm{b}) \Leftrightarrow \\ &\sup_{\hat{\alpha}_i} \left\{ (\hat{\alpha}_i - \alpha_i)\mu_i^\top \eta^\tau + \sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau} \left[(1-\hat{\alpha}_i) T_{\hat{\alpha}_i} - (1-\alpha_i) T_{\alpha_i} - \delta_i\right] - (\hat{\alpha}_i - \alpha_i) v_i \right\} \geqslant 0. \end{aligned} \tag{11}$$

Also we have $\hat{\alpha}_i \geqslant \alpha_i$ by (5c), because otherwise

$$(1-\alpha_i)\mathrm{CVaR}_{\alpha_i, \hat{F}_i}(-f_i(\tau, r_i)) \geqslant (1-\hat{\alpha}_i)\mathrm{CVaR}_{\hat{\alpha}_i, \hat{F}_i}(-f_i(\tau, r_i)).$$

Thus combined with (5c), we have

$$(5\mathrm{b}) - (5\mathrm{c}) \Rightarrow v_i \leqslant \mu_i^\top \eta^\tau + \sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau} \sup_{\hat{\alpha}_i \geqslant \alpha_i} \left\{ \frac{\left[(1-\hat{\alpha}_i) T_{\hat{\alpha}_i} - (1-\alpha_i) T_{\alpha_i} - \delta_i\right]}{\hat{\alpha}_i - \alpha_i} \right\}, \tag{12}$$

that is, $(5\mathrm{b}) - (5\mathrm{c}) \Rightarrow (8)$.

By Lemma 1 and the translation invariance of VaR, we have

$$\mathrm{VaR}_{\alpha_i, \hat{F}_i}(-f_i(\tau, r_i)) = \mathrm{VaR}_{\alpha_i, \hat{F}_i}\left(\frac{-(r_i - \mu_i)^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}}\right) + \frac{v_i - \mu_i^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}} = q_{\alpha_i} + \frac{v_i - \mu_i^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}},$$

Table 1: The values of T_{α_i} and q_{α_i} under different elliptical distributions

Distribution	T_{α_i}	$-q_{\alpha_i}$
Gaussian	$\frac{e^{-\frac{q_{1-\alpha_i}^2}{2}}}{(1-\alpha_i)\sqrt{2\pi}}$	$-\Phi^{-1}(\alpha_i)$
Laplace	$\frac{1}{\sqrt{2}}(1-\ln{(2-2\alpha_i)})$	$\frac{1}{\sqrt{2}}\ln{(2-2\alpha_i)}$
Logistic	$\frac{\sqrt{3}}{\pi}\ln\frac{\alpha_i^{-\frac{\alpha_i}{1-\alpha_i}}}{1-\alpha_i}$	$-\frac{\sqrt{3}}{\pi}\ln{(\frac{\alpha_i}{1-\alpha_i})}$

and

$$(5c) \Leftrightarrow v_i \leqslant \mu_i^\top \eta^\tau - \sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau} q_{\alpha_i}, \tag{13}$$

where by Lemma 1, q_{α_i} here is just the quantile at α_i of the elliptical distribution.

Suppose that (8) holds. Let $Y = \frac{-(r_i-\mu_i)^\top \eta^\tau}{\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}}$. Then, by the definition of CVaR, we have

$$(1-\alpha_i)\mathrm{CVaR}_{\alpha_i,\hat{F}_i}(Y) - (1-\hat{\alpha}_i)\mathrm{CVaR}_{\hat{\alpha}_i,\hat{F}_i}(Y) = \mathbb{E}_{\hat{F}_i}\left[Y \cdot \mathbf{1}\{Y \in [\mathrm{VaR}_{\hat{\alpha}_i,\hat{F}_i}(Y), \mathrm{VaR}_{\alpha_i,\hat{F}_i}(Y)]\}\right]$$

$$= \int_{\mathrm{VaR}_{\alpha_i,\hat{F}_i}(Y)}^{\mathrm{VaR}_{\hat{\alpha}_i,\hat{F}_i}(Y)} y \mathrm{d}\hat{F}_i(y) = \int_{\alpha_i}^{\hat{\alpha}_i} \mathrm{VaR}_{s,\hat{F}_i}(Y)\mathrm{d}s \geqslant (\alpha_i - \hat{\alpha}_i)\mathrm{VaR}_{\alpha_i,\hat{F}_i}(Y).$$

The last equality is obtained by the change of variable $s = \hat{F}_i(y)$, which induces $\beta_i \leqslant -q_{\alpha_i}$. Hence (8) $\Rightarrow$ (13), and (8) $\Rightarrow$ (5c). With the combination of (11), we have (8) $\Rightarrow$ (5b). Thus we get (8) $\Rightarrow$ (5b), (5c) which completes the proof.

□

From Theorem 1 we know that given a confidence level α_i for each player $i \in I$, and any random variable r_i elliptically distributed, we can solve the DRO payoff function $u_i^{\alpha_i}$ analytically according to (8).

By [22], we can compute for each player i, the value T_{α_i} and $-q_{\alpha_i}$ under three widely used standard elliptical distributions (the values are shown in Table 1), where $\Phi(\cdot)$ is the cdf of the standard Gaussian distribution. With the values of T_{α_i} and $-q_{\alpha_i}$ under specified elliptical distributions, we can derive the values of β_i in Theorem 1.

3.3. Existence of the Nash equilibrium

Assumption 2. *For each player $i \in I$, we assume that the radius of the uncertainty set $\mathcal{D}_i$ satisfies $\delta_i \geqslant \max_{\hat{\alpha}_i \in (\alpha_i, 1)} [(1-\hat{\alpha}_i)T_{\hat{\alpha}_i} - (1-\alpha_i)T_{\alpha_i}]$ such that the parameter β_i in Theorem 1 meets the condition $\beta_i \leqslant 0$.*

Remark 2. *For each $i \in I$, as $\hat{\alpha}_i \in (\alpha_i, 1)$ in β_i, the condition $\delta_i \geqslant \max_{\hat{\alpha}_i \in (\alpha_i, 1)} [(1-\hat{\alpha}_i)T_{\hat{\alpha}_i} - (1-\alpha_i)T_{\alpha_i}]$ in Assumption 2 is necessary and sufficient for $\beta_i \leqslant 0$.*

Theorem 2. *Given Assumptions 1,2, the DRO payoff function $u_i^{\alpha_i}(\cdot, \tau_{-i})$ is a concave function of τ_i for every $\tau_{-i} \in X_{-i}$.*

Proof. From the definition of η^τ, we have

$$\eta^\tau = (\eta^\tau(a))_{a\in A} = \left(\prod_{j\in I} \tau_j(a_j)\right)_{a\in A} = \left(\tau_i(a_i)\prod_{j\neq i} \tau_j(a_j)\right)_{a\in A}.$$

For a given $\tau_{-i} \in X_{-i}$ and for each $a \in A$, each element of η^τ is the form expressed as $\tau_i(a_i)K_a$, where $K_a = \prod_{j\neq i} \tau_j(a_j) \in \mathbb{R}$ is a known coefficient. Therefore, every element of η^τ is both convex and concave of τ_i for every $\tau_{-i} \in X_{-i}$.

Moreover, we know that $\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}$ is convex of each element of η^τ, and each element of η^τ is convex of τ_i for every $\tau_{-i} \in X_{-i}$. Then, $\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}$ is convex of each element of η^τ. By Assumption 2, $\beta_i \leqslant 0$, and thus $\beta_i\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}$ is concave of each element of η^τ.

Based on the above results, $(\eta^\tau)^\top \mu + \beta_i\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau}$ is concave of τ_i for every $\tau_{-i} \in X_{-i}$. Therefore, the payoff function $u_i^{\alpha_i}(\cdot, \tau_{-i})$ is a concave function of τ_i for every $\tau_{-i} \in X_{-i}$. $\square$

Lemma 2 ([4, Theorem 1])**.** *For a given confidence level vector $\alpha \in [0,1]^n$, assume that for each $i \in I$,*

1. *the payoff function of player i, $u_i^{\alpha_i} : X_i \times X_{-i} \to \mathbb{R}$ defined by (3) is a continuous function.*
2. *the payoff function $u_i^{\alpha_i}(\cdot, \tau_{-i})$ is a concave function of τ_i for every $\tau_{-i} \in X_{-i}$.*

Then, there always exists a mixed strategy Nash equilibrium of a DRCCG at a confidence level α.

Theorem 3. *Consider an n-player finite strategic game where the payoff vector $r_i = (r_i(a))_{a \in A}$ of each player $i, i \in I$ is a random vector, given Assumptions 1,2, then there always exists a mixed strategy Nash equilibrium.*

Proof. Given Assumptions 1,2, by Theorem 2, $u_i^{\alpha_i}(\cdot, \tau_{-i}), i \in I$ is a concave function of τ_i for every $\tau_{-i} \in X_{-i}$. From (8), $u_i^{\alpha_i}(\cdot, \tau_{-i}), i \in I$ is a continuous function of τ. That is, both conditions of Lemma 2 are satisfied. Thus, there exists a mixed strategy Nash equilibrium. □

3.4. Mathematical programming formulation to compute the Nash equilibrium

To compute the Nash equilibrium of a DRCCG, we solve the following convex program, whose global maximizer is the Nash equilibrium of the DRCCG.

Remark 3. *Consider the DRCCG whose reference distribution is defined in Theorem 1, for each player $i \in I$, and any action profile $a_i \in A_i$,. The expected payoff $u_i^{\alpha_i}$ for a fixed strategy τ_{-i} can be reformulated by the following optimization program*

$$\min_{\tau_i} \quad -\beta_i \sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau} - \mu_i^\top \eta^\tau \tag{14a}$$

$$\text{s.t.} \quad \sum_{a_i \in A_i} \tau_i(a_i) = 1, \tag{14b}$$

$$\tau_i(a_i) \geqslant 0. \tag{14c}$$

The dual problem of (14) *is*

$$\max_{\lambda_i, v_i} \quad \lambda_i \tag{15a}$$

$$\text{s.t.} \quad \lambda_i \leqslant \sum_{a_{-i} \in A_{-i}} \prod_{j \in I; j \neq i} \tau_j(a_j) \left[-\beta_i \sqrt{(v_i)^\top \Sigma_i v_i} - \mu_i(a_i, a_{-i}) \right], \tag{15b}$$

$$\|v_i\| \leqslant 1. \tag{15c}$$

As (14) is an SOCP problem, strong duality holds if (14) has a finite optimal value, i.e., the optimal value of (14) is equal to the optimal value of (15). Thus, the optimal policy τ_i^* of the i-th player can be found by solving the following set of equations

$$\begin{cases} -\beta_i \sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau} - \mu_i^\top \eta^\tau = \lambda_i, & \text{(16a)} \\ (14b) - (14c), (15b) - (15c). & \text{(16b)} \end{cases} \tag{16}$$

Alternatively, we can solve an optimization problem instead of solving the equilibrium equations by penalizing the violation in the objective function ,i.e.,

$$\max_{\lambda_i, v_i, \tau_i,\ i=1,\ldots,n} \quad \sum_{i\in I}\left[\lambda_i + \beta_i\sqrt{(\eta^\tau)^\top \Sigma_i \eta^\tau} + \mu_i^\top \eta^\tau\right] \tag{17a}$$

$$\text{s.t.} \qquad (14\text{b}) - (14\text{c}), (15\text{b}) - (15\text{c}),\ i = 1,\ldots,n. \tag{17b}$$

4. Numerical experiments

In this section, we carry out a series of numerical tests under three different kinds of reference distributions. We compute the Nash equilibrium and the corresponding payoff of DRCCG by solving the mathematical program (17). Through 100 randomly generated groups of Gaussian distributions, we compare the performances of the Nash equilibrium of different radius δ_i and observe the robustness of our model.

We consider a two-players DRCCG example introduced in [4], where $I = \{1,2\}, A_1 = \{1,2,3\}, A_2 = \{1,2,3\}$. The mean vectors for both players are $\mu_1 = (10,9,11,8,12,10,7,8,13)^\top$, $\mu_2 = (9,7,8,9,10,10,10,9,8)^\top$ and the covariance matrices for both players are

$$\Sigma_1 = \begin{pmatrix} 6 & 4 & 3 & 3 & 2 & 3 & 4 & 2 & 4 \\ 4 & 6 & 3 & 4 & 3 & 3 & 3 & 2 & 3 \\ 3 & 3 & 8 & 4 & 2 & 3 & 3 & 2 & 4 \\ 3 & 4 & 4 & 6 & 2 & 3 & 3 & 3 & 2 \\ 2 & 3 & 2 & 2 & 6 & 2 & 4 & 3 & 3 \\ 3 & 3 & 3 & 3 & 2 & 6 & 3 & 3 & 4 \\ 4 & 3 & 3 & 3 & 4 & 3 & 8 & 4 & 3 \\ 2 & 2 & 2 & 3 & 3 & 3 & 4 & 6 & 4 \\ 4 & 3 & 4 & 2 & 3 & 4 & 3 & 4 & 8 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 6 & 3 & 3 & 3 & 3 & 2 & 4 & 3 & 2 \\ 3 & 6 & 3 & 3 & 2 & 2 & 3 & 3 & 4 \\ 3 & 3 & 6 & 3 & 3 & 3 & 4 & 3 & 4 \\ 3 & 3 & 3 & 6 & 3 & 2 & 2 & 3 & 3 \\ 3 & 2 & 3 & 3 & 6 & 4 & 2 & 2 & 3 \\ 2 & 2 & 3 & 2 & 4 & 6 & 3 & 3 & 4 \\ 4 & 3 & 4 & 2 & 2 & 3 & 6 & 3 & 2 \\ 3 & 3 & 3 & 3 & 2 & 3 & 3 & 6 & 3 \\ 2 & 4 & 4 & 3 & 3 & 4 & 2 & 3 & 6 \end{pmatrix},$$

which are both positive definite. We consider three kinds of reference distributions in elliptical distributions family as the center of the Wasserstein ball in DRCCG, namely Gaussian, Logistic and Laplace distributions respectively. We find the coefficient $\beta_i, i \in I$ by solving the optimization problem (9) with the function "globalsearch" in MATLAB when the radius $\delta_i = 0.5, 10^{-1}, 10^{-5}, 0$. When $\delta_i = 0$, the robust model is reduced to the non-robust model. The coefficient of the non-robust model is $\beta_i = -q_{\alpha_i}$ according to the results in Section 3.2 of [3]. We find the Nash equilibrium by solving the optimization problem (17) with BARON solver in YALMIP tool box of MATLAB.

In Table 2, We list our computational results of the Nash equilibrium τ_i and the corresponding payoffs $u_i^{\alpha_i}$ for three different elliptical distributions under different radius δ_i and confidence probability coefficients α_i respectively. From Table 2, we observe that the value of β_i converges to $-q_{\alpha_i}$ when the radius goes to zero,which means the distributionally robust model gradually recovers the non-robust one. Correspondingly, the payoff $u_1^{\alpha_1}, u_2^{\alpha_2}$ both increase as δ_i converges to 0. In fact, when the radius of Wasserstein ball increases, the model covers more possible situations and thereby presents more robust results. The reduction of the payoffs can be seen as the trade off for the robustness.

Table 2: Nash equilibrium solution and payoff under different reference distributions, radii and tolerance probabilities

Distribution	(α_1, α_2)	δ_i	β_i	τ_1	τ_2	$u_1^{\alpha_1}$	$u_2^{\alpha_2}$
Robust-Gaussian	(0.95,0.95)	0.5	-12.0627	$(\frac{2731}{10000}, \frac{6166}{10000}, \frac{1104}{10000})$	$(\frac{3943}{10000}, \frac{2949}{10000}, \frac{3108}{10000})$	-12.4130	-12.8836
		10^{-1}	-4.0627	$(\frac{3922}{10000}, \frac{5257}{10000}, \frac{821}{10000})$	$(\frac{4178}{10000}, \frac{2252}{10000}, \frac{3570}{10000})$	2.2079	1.5976
		10^{-5}	-1.6588	(1,0,0)	$(\frac{7199}{10000}, 0, \frac{2801}{10000})$	6.5906	5.0894
Non robust-Gaussian	(0.95,0.95)	0	-1.6449	(1,0,0)	$(\frac{7219}{10000}, 0, \frac{2781}{10000})$	6.6184	5.1198
Robust-Laplace	(0.85,0.85)	0.5	-4.8894	$(\frac{3428}{10000}, \frac{5759}{10000}, \frac{812}{10000})$	$(\frac{3987}{10000}, \frac{2522}{10000}, \frac{3491}{10000})$	0.7406	0.1460
		10^{-1}	-2.1050	(1,0,0)	$(\frac{6712}{10000}, 0, \frac{3288}{10000})$	5.6730	4.1194
		10^{-5}	-0.8611	(1,0,0)	$(\frac{9667}{10000}, 0, \frac{333}{10000})$	7.9579	6.8917
Non robust-Laplace	(0.85,0.85)	0	-0.8513	(1,0,0)	$(\frac{9733}{10000}, 0, \frac{267}{10000})$	7.9685	6.9153
Robust-Logistic	(0.7,0.7)	0.5	-2.7774	$(\frac{7456}{10000}, \frac{2058}{10000}, \frac{486}{10000})$	$(\frac{5776}{10000}, \frac{669}{10000}, \frac{3556}{10000})$	4.3319	3.3341
		10^{-1}	-1.2856	(1,0,0)	$(\frac{7900}{10000}, 0, \frac{2100}{10000})$	7.3088	5.9140
		10^{-5}	-0.4744	(0,1,0)	$(0, \frac{5010}{10000}, \frac{4990}{10000})$	10.0532	8.9392
Non robust-Logistic	(0.7,0.7)	0	-0.4671	(0,1,0)	$(0, \frac{5031}{10000}, \frac{4969}{10000})$	10.0720	8.9555

For the radius $\delta_i = 0.5, 0.1$ and 0, we choose the results of Nash equilibrium τ_1, τ_2 and the corresponding payoffs $u_1^{\alpha_1}, u_2^{\alpha_2}$ under the standard Gaussian distribution with $\alpha_i = 0.95$. We randomly generate 100 groups of Gaussian distributions and take $\tau_1, \tau_2, u_1^{\alpha_1}, u_2^{\alpha_2}$ for which we compute the satisfaction probability $\mathbb{P}_{\mathcal{K}_i}((r_i)^\top \eta^\tau \geqslant u_i^{\alpha_i})$, where $\mathcal{K}_i$ is the i-th randomly generated Gaussian distribution, $i = 1, \ldots, 100$. Figure 1 shows the values of the satisfaction probability $\mathbb{P}_{\mathcal{K}_1}$ under 100 randomly generated Gaussian distributions for player 1 when $\delta_1 = 0.5, 0.1,$ and 0. The results of the players 1 and 2 are similar, thus we present only the player 1 results. From Figure 1, when $\delta_1 = 0.5, 0.1$, we see that $\mathbb{P}_{\mathcal{K}_1}$ for all 100 distributions are beyond 0.95. When $\delta_1 = 0$, i.e., the distributionally robust model is reduced to the non robust model, there are only 3 random distributions' satisfaction probabilities beyond 0.95. Through comparing the satisfaction probabilities of these three radii, we see the performance of robustness for the solution of Nash equilibrium $u_1^{\alpha_1}$. Therefore, it is clear that the robustness of the non robust model is by far less than the distributionally robust one.

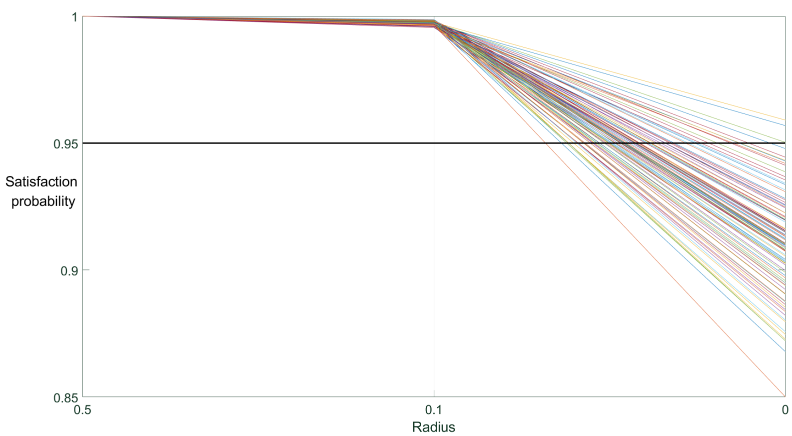

Figure 1: Values of satisfaction probability $\mathbb{P}_{\mathcal{K}_i}$, $i = 1, \dots, 100$, under 100 randomly generated Gaussian distributions

Conclusion

In this paper, we study DRCCG under Wasserstein ball, where the reference distribution is an elliptical distribution. We prove the existence of a Nash equilibrium of DRCCG and propose an optimization approach to compute the Nash equilibrium. By compuaring the out-of-sample performances under some randomly generated distributions, we examine the the robustness of the DRCCG compared with the non-robust model. Considering skewed and non-linearly dependent reference distribution is a promising topic for further research.

Acknowledgement

The authors are grateful to the editors and anonymous reviewers for their extremely insightful and detailed comments and suggestions, which have helped us to improve the paper significantly in both content and style. This research was supported by National Key R&D Program of China under No. 2022YFA1004000 and National Natural Science Foundation of China under No. 11991023 and 11901449.

References

[1] John Von Neumann. On the theory of games of strategy. *Contributions to the Theory of Games*, 4:13–42, 1959.

[2] John F Nash Jr. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.

[3] Vikas Vikram Singh, Oualid Jouini, and Abdel Lisser. Existence of nash equilibrium for chance-constrained games. *Operations Research Letters*, 44(5):640–644, 2016.

[4] Vikas Vikram Singh, Oualid Jouini, and Abdel Lisser. Distributionally robust chance-constrained games: existence and characterization of nash equilibrium. *Optimization Letters*, 11(7):1385–1405, 2017.

[5] Vikas Vikram Singh and Abdel Lisser. A characterization of nash equilibrium for the games with random payoffs. *Journal of Optimization Theory and Applications*, 178(3):998–1013, 2018.

[6] Shen Peng, Abdel Lisser, Vikas Vikram Singh, Nalin Gupta, and Eshan Balachandar. Games with distributionally robust joint chance constraints. *Optimization Letters*, 15(6):1931–1953, 2021.

[7] James E Smith and Robert L Winkler. The optimizer's curse: Skepticism and post-decision surprise in decision analysis. *Management Science*, 52(3):311–322, 2006.

[8] Huifu Xu and Dali Zhang. Stochastic nash equilibrium problems: sample average approximation and applications. *Computational Optimization and Applications*, 55(3):597–645, 2013.

[9] Esfahani Mohajerin Peyman and Daniel Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1-2):1–52, 2018.

[10] Chaoyue Zhao and Yongpei Guan. Data-driven risk-averse stochastic optimization with wasserstein metric. *Operations Research Letters*, 46(2):262–267, 2018.

[11] Zhi Chen, Daniel Kuhn, and Wolfram Wiesemann. Data-driven chance constrained programs over wasserstein balls. *arXiv preprint arXiv:1809.00210*, 2018.

[12] Weijun Xie. On distributionally robust chance constrained programs with wasserstein distance. *Mathematical Programming*, 186(1):115–155, 2021.

[13] Ran Ji and Miguel A Lejeune. Data-driven distributionally robust chance-constrained optimization with wasserstein metric. *Journal of Global Optimization*, 79(4):779–811, 2021.

[14] Guanze Peng, Tao Zhang, and Quanyan Zhu. A data-driven distributionally robust game using wasserstein distance. In *International Conference on Decision and Game Theory for Security*, pages 405–421. Springer, 2020.

[15] Jia Liu, Abdel Lisser, and Zhiping Chen. Distributionally robust chance constrained geometric optimization. *Mathematics of Operations Research*, 2022.

[16] Jose Blanchet, Kang Yang, and Karthyek Murthy. Robust wasserstein profile inference and applications to machine learning. *Journal of Applied Probability*, 56(3):830–857, 2019.

[17] Gao Rui and Anton J Kleywegt. Distributionally robust stochastic optimization with wasserstein distance. *arXiv preprint arXiv:1604.02199*, 2016.

[18] Haoming Shen and Ruiwei Jiang. Convex chance-constrained programs with wasserstein ambiguity. *arXiv preprint arXiv:2111.02486*, 2021.

[19] Frank J Fabozzi, Svetlozar T Rachev, and Christian Menn. *Fat-tailed and skewed asset return distributions: implications for risk management, portfolio selection, and option pricing*. John Wiley & Sons, 2005.

[20] Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press, 2015.

[21] Dobrislav Dobrev, Travis D Nesmith, and Dong Hwan Oh. Accurate evaluation of expected shortfall for linear portfolios with elliptically distributed risk factors. *Journal of Risk and Financial Management*, 10(1):5, 2017.

[22] Valentyn Khokhlov. Conditional value-at-risk for elliptical distributions. *Evropskỳ časopis ekonomiky a managementu*, 2(6):70–79, 2016.