



HAL
open science

Chance constrained problem and its applications

Shen Peng

► **To cite this version:**

Shen Peng. Chance constrained problem and its applications. Optimization and Control [math.OC]. Université Paris Saclay (COMUE); Xi'an Jiaotong University, 2019. English. NNT : 2019SACLS153 . tel-02303045

HAL Id: tel-02303045

<https://theses.hal.science/tel-02303045v1>

Submitted on 2 Oct 2019

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

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

Optimisation stochastique avec contraintes en probabilités et applications

Thèse de doctorat de Xi'an Jiaotong University et de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

Ecole doctorale n°580 Sciences et Technologies de l'Information et de la
Communication (STIC)
Spécialité de doctorat : Informatique

Thèse présentée et soutenue à Orsay, le 17 juin 2019, par

M. Shen PENG

Composition du Jury :

M. Alexandre CAMINADA Professeur, Université de Nice	Président, Rapporteur
M. Mounir HADDOU Professeur, INSA Rennes	Rapporteur
M. Yacine CHITOUR Professeur, CentraleSupélec	Examineur
Mme Janny LEUNG Professeure, Chinese University of Hong Kong (Shenzhen)	Examinatrice
Mme Francesca MAGGIONI Professeure, University of Bergamo	Examinatrice
M. Vikas Vikram SINGH Professeur assistant, Indian Institute of Technology Delhi	Examineur
M. Abdel LISSER Professeur, Université Paris-Sud	Directeur de thèse
M. Zhiping CHEN Professeur, Xi'an Jiaotong University	Co-directeur de thèse

Chance Constrained Problem and Its Applications

A dissertation submitted to
Université Paris Sud
in fulfillment of the requirements
for the Ph.D. degree in
Computer Science

By

Shen PENG

Supervisor: Prof. Abdel LISSER

Co-supervisor: Prof. Zhiping CHEN

June 2019

Acknowledgements

First of all, I would like to thank my supervisors, Prof. Abdel Lisser and Prof. Zhiping Chen, who illuminate my way to the science.

Prof. Zhiping Chen is my supervisor in Xi'an Jiaotong University. Thanks to his consistent support and encouragement, I can navigate many obstacles in pursuing the Ph.D degree. He has spent a lot of time and efforts on my study and research. He not only taught me scientific knowledge but also sharing a lot of experience on life and research. I have earned a lot from him both on how to be a researcher and how to research.

Prof. Abdel Lisser is my supervisor in Université Paris Sud. Thanks to his open-door policy, I can always discuss any problem with him and obtain the nutritious suggestions. He is not just an supervisor to my research, a mentor from whom I learnt, but also a friend. I feel extremely lucky and want to sincerely say thank you from the bottom of my heart.

My special thanks then go to Dr. Jia Liu of Xi'an Jiaotong University and Dr. Vikas Vikram Singh of Indian Institute of Technology Delhi for their positive discussion and helpful advice on my research work.

I also would like to give my gratitude to all of my friends and colleagues both in China and France from whom I have benefited enormously. Particularly, I thank Hehuan Shi who helped me solving a lot of registration problems in France. I thank Dr. Jihong Yu, Dr. Rongrong Zhang, Dr. Ruqi Huang, Dr. Chuan Xu, Zhoudan Lyu and Xingchao Cui for sharing a lot of great moments in Paris. I also thank Dr. Zongxin Li, Liyuan Wang, Xinkai Zhuang, Zhe Yan, Jie Jiang, Fei Yu, He Hu, Yu Mei, Wei Wang, Xin Wei, Rui Xie, Zhujia Xu, Ming Jin and Tianyao Chen for sharing many wonderful time with me and making my life rich and colorful.

Last but not least, I would like to devote the most deep gratitude to each member of my family: my parents, my parents-in-law and my wife for their unconditional love, support and encouragement. Especially, I want to express the most special gratitude to my wife Yingping Li. Thanks for being at my side for all these years.

Résumé

L'incertitude est une propriété naturelle des systèmes complexes. Les paramètres de certains modèles peuvent être imprécis ; la présence de perturbations aléatoires est une source majeure d'incertitude pouvant avoir un impact important sur les performances du système. L'optimisation sous contraintes en probabilités est une approche naturelle et largement utilisée pour fournir des décisions robustes dans des conditions d'incertitude. Dans cette thèse, nous étudierons systématiquement les problèmes d'optimisation avec contraintes en probabilités dans les cas suivants :

En tant que base des problèmes stochastiques, nous passons d'abord en revue les principaux résultats de recherche relatifs aux contraintes en probabilités selon trois perspectives: les problèmes liés à la convexité en présence de contraintes probabilistes, les reformulations et les approximations de ces contraintes, et les contraintes en probabilités dans le cadre de l'optimisation distributionnellement robuste.

Pour les problèmes d'optimisation géométriques stochastiques, nous étudions les programmes avec contraintes en probabilités géométriques rectangulaires jointes. A l'aide d'hypothèses d'indépendance des variables aléatoires elliptiquement distribuées, nous déduisons une reformulation des programmes à contraintes géométriques rectangulaires jointes. Comme la reformulation n'est pas convexe, nous proposons de nouvelles approximations convexes basées sur la transformation des variables ainsi que des méthodes d'approximation linéaire par morceaux. Nos résultats numériques montrent que nos approximations sont asymptotiquement serrées.

Lorsque les distributions de probabilité ne sont pas connues à l'avance ou que la reformulation des contraintes probabilistes est difficile à obtenir, des bornes obtenues à partir des contraintes en probabilités peuvent être très utiles. Par conséquent, nous développons quatre bornes supérieures pour les contraintes probabilistes individuelles, et jointes dont les vecteur-lignes de la matrice des contraintes sont indépendantes. Sur la base de l'inégalité unilatérale de Chebyshev, de l'inégalité de Chernoff, de l'inégalité de Bernstein et de l'inégalité de Hoeffding, nous proposons des approximations déterministes des contraintes probabilistes. En outre, quelques conditions suffisantes dans lesquelles les approximations susmentionnées sont convexes et solvables de manière efficace sont déduites. Pour réduire davantage la complexité des calculs, nous reformulons les approximations sous forme de problèmes d'optimisation convexes solvables basés sur des approximations linéaires et tangentielles par morceaux. Enfin, des expériences numériques sont menées afin de montrer la qualité des approximations déterminis-

tes étudiées sur des données générées aléatoirement.

Dans certains systèmes complexes, la distribution des paramètres aléatoires n'est que partiellement connue. Pour traiter les incertitudes complexes en termes de distribution et de données d'échantillonnage, nous proposons un ensemble d'incertitude basé sur des données obtenues à partir de distributions mixtes. L'ensemble d'incertitude basé sur les distributions mixtes est construit dans la perspective d'estimer simultanément des moments d'ordre supérieur. Ensuite, à partir de cet ensemble d'incertitude, nous proposons une reformulation du problème robuste avec contraintes en probabilités en utilisant des données issues d'échantillonnage. Comme la reformulation n'est pas un programme convexe, nous proposons des approximations nouvelles et convexes serrées basées sur la méthode d'approximation linéaire par morceaux sous certaines conditions. Pour le cas général, nous proposons une approximation DC pour dériver une borne supérieure et une approximation convexe relaxée pour dériver une borne inférieure pour la valeur de la solution optimale du problème initial. Nous établissons également le fondement théorique de ces approximations. Enfin, des expériences numériques sont effectuées pour montrer que les approximations proposées sont pratiques et efficaces.

Nous considérons enfin un jeu stochastique à n joueurs non-coopératif. Lorsque l'ensemble de stratégies de chaque joueur contient un ensemble de contraintes linéaires stochastiques, nous modélisons les contraintes linéaires stochastiques de chaque joueur sous la forme de contraintes en probabilité jointes. Pour chaque joueur, nous supposons que les vecteurs lignes de la matrice définissant les contraintes stochastiques sont indépendants les uns des autres. Ensuite, nous formulons les contraintes en probabilité dont les variables aléatoires sont soit normalement distribuées, soit elliptiquement distribuées, soit encore définies dans le cadre de l'optimisation distributionnellement robuste. Sous certaines conditions, nous montrons l'existence d'un équilibre de Nash pour ces jeux stochastiques.

Abstract

Uncertainty is a natural property of complex systems. Imprecise model parameters and random disturbances are major sources of uncertainties which may have a severe impact on the performance of the system. The target of optimization under uncertainty is to provide profitable and reliable decisions for systems with such uncertainties. Ensuring reliability means satisfying specific constraints of such systems. An appropriate treatment of inequality constraints of a system influenced by uncertain variables is required for the formulation of optimization problems under uncertainty. Therefore, chance constrained optimization is a natural and widely used approaches for this purpose. Moreover, the topics around the theory and applications of chance constrained problems are interesting and attractive.

Chance constrained problems have been developed for more than four decades. However, there are still some important issues requiring non-trivial efforts to solve. In view of this, we will systematically investigate chance constrained problems from the following perspectives.

(1) As the basis for chance constrained problems, we first review some main research results about chance constraints in three perspectives: convexity of chance constraints, reformulations and approximations for chance constraints and distributionally robust chance constraints. Since convexity is fundamental for chance constrained problems, we introduce some basic mathematical definitions and theories about the convexity of large classes of chance constrained problems. Then, we state some tractable convex reformulations and approximations for chance constrained problems. For distributionally robust optimization, we illustrate moments based uncertainty set, and distance based uncertainty set and show their applications in distributionally robust chance constrained problems, respectively.

(2) For stochastic geometric programs, we first review a research work about joint chance constrained geometric programs with normal distribution and independent assumptions. As an extension, when the stochastic geometric program has rectangular constraints, we formulate it as a joint rectangular geometric chance constrained program. With elliptically distributed and pairwise independent assumptions for stochastic parameters, we derive a reformulation of the joint rectangular geometric chance constrained programs. As the reformulation is not convex, we propose new convex approximations based on the variable transformation together with piecewise linear approximation methods. Our numerical results show that our approximations are asymptotically tight.

(3) When the probability distributions are not known in advance or the reformulation for chance constraints is hard to obtain, bounds on chance constraints

can be very useful. Therefore, we develop four upper bounds for individual and joint chance constraints with independent matrix vector rows. Based on the one-side Chebyshev inequality, Chernoff inequality, Bernstein inequality and Hoeffding inequality, we propose deterministic approximations for chance constraints. In addition, various sufficient conditions under which the aforementioned approximations are convex and tractable are derived. Therefore, to reduce further computational complexity, we reformulate the approximations as tractable convex optimization problems based on piecewise linear and tangent approximations. Finally, based on randomly generated data, numerical experiments are discussed in order to identify the tight deterministic approximations.

(4) In some complex systems, the distribution of the random parameters is only known partially. To deal with the complex uncertainties in terms of the distribution and sample data, we propose a data-driven mixture distribution based uncertainty set. The data-driven mixture distribution based uncertainty set is constructed from the perspective of simultaneously estimating higher order moments. Then, with the mixture distribution based uncertainty set, we derive a reformulation of the data-driven robust chance constrained problem. As the reformulation is not a convex program, we propose new and tight convex approximations based on the piecewise linear approximation method under certain conditions. For the general case, we propose a DC approximation to derive an upper bound and a relaxed convex approximation to derive a lower bound for the optimal value of the original problem, respectively. We also establish the theoretical foundation for these approximations. Finally, simulation experiments are carried out to show that the proposed approximations are practical and efficient.

(5) We consider a stochastic n -player non-cooperative game. When the strategy set of each player contains a set of stochastic linear constraints, we model the stochastic linear constraints of each player as a joint chance constraint. For each player, we assume that the row vectors of the matrix defining the stochastic constraints are pairwise independent. Then, we formulate the chance constraints with the viewpoints of normal distribution, elliptical distribution and distributionally robustness, respectively. Under certain conditions, we show the existence of a Nash equilibrium for the stochastic game.

Contents

1	Introduction	1
1.1	Convexity	2
1.2	Reformulations and approximations for Chance Constraints	3
1.3	Distributionally Robust Chance Constrained Problems	4
1.4	Contribution and Outline of the Dissertation	7
2	Chance Constraints	9
2.1	α -concave measures and convexity of chance constraints	9
2.2	Reformulations and approximations for chance constraints	14
2.2.1	Individual chance constraints	14
2.2.2	Normally distributed joint linear chance constraints	16
2.2.3	Integer programming approaches for joint linear chance constraints	17
2.2.4	Convex approximations for chance constraints	18
2.2.5	Sample average approximations for chance constraints . . .	19
2.3	Distributionally robust chance constraints	20
2.3.1	Distributionally robust individual chance constraints with known mean and covariance	20
2.3.2	Distributionally robust joint chance constraints with known mean and covariance	21
2.3.3	Distributionally robust joint chance constraints with uncertain mean and covariance	22
2.3.4	Distributionally robust chance constrained problem based on ϕ -divergence	25
2.3.5	Distributionally robust chance constrained problem based on Wasserstein distance	27
2.4	Conclusion	29
3	Geometric Chance Constrained Programs	31
3.1	Stochastic geometric optimization with joint chance constraints .	31
3.1.1	Introduction	31
3.1.2	Stochastic geometric optimization under Gaussian distribution	32
3.1.3	Numerical experiments	37
3.2	Joint rectangular geometric chance constrained programs	38
3.2.1	Introduction	39
3.2.2	Elliptically distributed stochastic geometric problems . . .	40

3.2.3	Convex approximations of constraints (3.35) and (3.36) . . .	43
3.2.4	Numerical experiments	50
3.3	Conclusion	55
4	Bounds for Chance Constrained Problems	57
4.1	Introduction	57
4.2	Chebyshev and Chernoff Bounds	58
4.2.1	Chebychev bounds	58
4.2.2	Chernoff bounds	61
4.3	Bernstein and Hoeffding Bounds	65
4.3.1	Bernstein bounds	65
4.3.2	Hoeffding bounds	66
4.4	Computational Results	68
4.5	Conclusion	72
5	Data-Driven Robust Chance Constrained Problems: A Mixture Model Approach	75
5.1	Mixture Distribution Based Uncertainty Set	75
5.1.1	Data-Driven Confidence Region for Moments	77
5.1.2	Data-Driven Confidence Region for Mixture Weights . . .	81
5.2	Data-Driven Robust Chance Constrained Problems	82
5.2.1	Reformulation of Problem (\mathbf{P})	82
5.2.2	Convex Approximation	84
5.2.3	DC reformulation	89
5.3	Numerical Results	95
5.3.1	Convex approximation	95
5.3.2	DC approximation	97
5.4	Conclusion	98
6	Chance Constrained Stochastic Game Theory	99
6.1	The model	99
6.2	Existence of Nash equilibrium with normal distribution	100
6.3	Existence of Nash equilibrium with elliptical distribution	107
6.4	Existence of Nash equilibrium for distributionally robust model .	111
6.4.1	ϕ -divergence uncertainty set	112
6.4.2	Moments based uncertainty set I	114
6.4.3	Moments based uncertainty set II	116
6.4.4	Moments based uncertainty set III	117
6.5	Conclusion	119
7	Conclusions and Prospects	121
7.1	Conclusions	121
7.2	Prospects	122
	Bibliography	125

Chapter 1

Introduction

Consider the following constrained optimization problem:

$$\begin{aligned} \min_{\mathbf{x}} \quad & g(\mathbf{x}) \\ \text{s.t.} \quad & c_1(\mathbf{x}, \boldsymbol{\xi}) \leq 0, \dots, c_d(\mathbf{x}, \boldsymbol{\xi}) \leq 0, \\ & \mathbf{x} \in \mathbf{X}, \end{aligned} \tag{1.1}$$

where $\mathbf{X} \in \mathbb{R}^m$ is a deterministic set, $\mathbf{x} \in \mathbf{X}$ is the decision vector, $\boldsymbol{\xi}$ is a k -dimensional parameter vector, $g(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$ and $c_i(\mathbf{x}, \boldsymbol{\xi}) : \mathbb{R}^{n+k} \rightarrow \mathbb{R}, i = 1, \dots, d$, are real value functions. Furthermore, we assume that $g(\mathbf{x})$ and $c_i(\mathbf{x}, \boldsymbol{\xi}), i = 1, \dots, d$ are convex in \mathbf{x} and \mathbf{X} is a compact and convex set. Then, problem (1.1) is a constrained convex optimization problem. This kind of problem has broad applications in communications and networks, product design, system control, statistics, and finance, and it can be solved efficiently.

However, in many practical problems, the parameter $\boldsymbol{\xi}$ of problem (1.1) might be uncertain. If we ignore the uncertainty, such as using the expectation of $\boldsymbol{\xi}$, the obtained optimal solution might be infeasible with high probability.

To take the uncertainty into consideration, we can formulate the problem as a chance constrained optimization problem:

$$\begin{aligned} \min_{\mathbf{x}} \quad & g(\mathbf{x}) \\ \text{s.t.} \quad & \mathbb{P}_F \{c_1(\mathbf{x}, \boldsymbol{\xi}) \leq 0, \dots, c_d(\mathbf{x}, \boldsymbol{\xi}) \leq 0\} \geq \epsilon, \\ & \mathbf{x} \in \mathbf{X}, \end{aligned} \tag{1.2}$$

where $\boldsymbol{\xi}$ is a random vector defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$, F is the joint probability distribution function of the random vector $\boldsymbol{\xi}$ and $\epsilon \in (0, 1)$ is the tolerance probability, say, 0.95 or 0.99. Therefore, the solution to problem (1.2) is guaranteed to be a feasible solution to the original problem (1.1) with a probability at least ϵ . Problem (1.2) is called a joint chance constrained problem, and the chance constraint is called a joint chance constraint. When $d = 1$, problem (1.2) is called an individual chance constrained problem because it requires only a single constraint to be satisfied with probability ϵ .

Chance (probabilistic) constrained optimization problems were firstly proposed by Charnes et al.[17] and Miller and Wagner [75]. In the paper of Charnes et al. [17], chance constraints are imposed individually on each constraint involv-

ing random variables. The paper by Miller and Wagner [75] takes the chance constraint jointly on the stochastic constraints but handles only independent random variables appearing on the right hand sides of the stochastic constraints. Prékopa [87] initiated the research on this topic, where the chance constraint is taken jointly for the stochastic constraints and the random variables involved are stochastically dependent, in general. Applications of probabilistic constraints are substantial in engineering and finance. For an overview on the theory, solution and applications of chance constrained optimization problems, one can refer to the monographs [13], [89] and [96].

1.1 Convexity

A fundamental issue for chance constrained problems is the convexity of chance constraints, which is widely considered as a considerable difficulty. Both from theoretical and computational perspectives, it was recognized that chance constrained problems are hard to treat. Van de Panne and Popp [79] proposed a solution method for a linear chance constrained problem with one-row normally distributed constraint, transformed to a nonlinear convex constraint. Kataoka [59] studied an individual chance constrained problem with random right-hand side following a normal distribution.

Apart from the above simple problems, chance constrained problems often lead to a non-convex feasible solution set. Various conditions and techniques were developed to handle this issue. Miller and Wagner [75] considered a joint chance constrained model with independent random right-hand side. The convexity of their problem is ensured when the probability distribution possesses a property of decreasing reversed hazard function. Jagannathan [56] extended this result to the dependent case, and also considered a linear joint chance constrained problem with normally distributed independent rows.

Prékopa [88] made an essential step by proving the convexity of the feasible solution set of linear chance constrained problems with dependent random right-hand sides in the chance constraint when the probability distribution is logarithmically concave. This kind of distributions include normal, Wishart and Dirichlet distributions. Borell [9] and Brascamp and Lieb [11] generalized logarithmic concavity to a concept called r -concave measures. A generalized definition of r -concave function on set, which is also suitable for discrete distributions, was proposed by Dentcheva et al. [28]. Concretely, Calafiore and El Ghaoui [15] reformulated an individual chance constraint as a second order cone constraint by using a similar notion of Q -radial distribution.

Despite this progress, the problem of convexity remains to be a big challenge for chance constrained problems, especially for joint chance constrained problems. However, there are still some extensions about the convexity. Prékopa et al. [90] asserted that a joint linear chance constrained problem is convex if the rows are independently normally distributed and the covariances matrices of the rows are constant multiples of each other. Henrion [49] gave a completely structural description of the feasible solution set defined by individual linear chance constraints, which can be seen as a more promising direction. Following this

direction, by using the r -concavity and introducing a notion of r -decreasing function, Henrion and Strugarek [50] further proved the convexity of joint chance constraints with independent random variables separated from decision vectors. For the dependent case with random right-hand side, Henrion and Strugarek [51] and Ackooij [1] extended the result by using the theory of copulas, respectively, while Houda [54] considered the dependent case by using a variation to the mixing coefficient. In addition, handling the dependence of the random vectors by copulas, Cheng et al. [23] proposed a mixed integer linear reformulation and provide an efficient semidefinite relaxation of 0–1 quadratic programs with joint probabilistic constraints. Lagoa et al. [62] proved the convexity when the random vector have a log-concave symmetric distribution. Recently, Lubin et al. [66] showed the convexity of two-sided chance constrained program when ξ follows normal (or log-concave) distribution.

1.2 Reformulations and approximations for Chance Constraints

Generally, the probability associated with chance constraints is difficult to compute due to the multiple integrals. For this reason, many equivalent reformulations for chance constraints or their approximations have been proposed. When the random vector ξ in individual linear chance constraints follows normal distribution, elliptical distribution or radial distribution, the chance constraint in the individual linear chance constrained problem can be reformulated as a second order conic programming (SOCP) constraint ([15, 49, 56, 88]). For the joint linear chance constrained problem with normally distributed coefficients and independent matrix rows, Cheng and Lissner [24] proposed SOCP approximations by using piecewise linear and piecewise tangent approximations. And for the joint chance constrained geometric programming problem with independent normally distributed parameters, Liu et al [65] derived asymptotically tight geometric programming approximations based on variable transformation and piecewise linear approximations. Using Archimedean copula, Cheng et al. [22] considered elliptically distributed joint linear chance constraints with dependent rows and derived SOCP approximation schemes. Luedtke and Ahmed [68] constructed a mixed integer linear programming reformulation for joint linear chance constrained problems when the random vector has finite support.

In order to solve chance constrained problems efficiently, we need both the convexity of the corresponding feasible set and efficient computability of the considered probability [76]. This combination is rare, and very few are the cases in which a chance constraint can be processed efficiently (see [28, 62, 91]). Whenever this is the case, tractable approximations of chance constraints can be very useful in practice.

A computationally tractable approximation of chance constrained problems is given by the scenario approach, based on Monte Carlo sampling techniques. With this kind of techniques, one can recur to approximate solutions based on constraint sampling. And the constraint consists in taking into account only

a finite set of constraints, chosen at random among the possible continuum of constraint instances of the uncertainty. The attractive feature of this method is to provide explicit bounds on the measure of the original constraints that are possibly violated by the randomized solution. The properties of the solutions provided by this approach, called scenario approach, have been studied in [12, 16, 40], where it has been shown that most of the constraints of the original problem are satisfied provided that the number of samples is sufficiently large. The constraint sampling method has been extensively studied within the chance constraint approach through different directions by [36] and [78]. More concretely, Nemirovski and Shapiro [77] solved joint linear chance constrained problems through scenario approximation and studied the conditions under which the solution of the approximation problem is feasible for the original problem with high probability. By using the sample approximation method, Luedtke and Ahmed [67] derived the minimal sample size and the probability requirement such that a solution of the approximate problem is feasible for the original joint chance constrained problem. Based on Monte Carlo sampling techniques, Hong, Yang and Zhang [53] proposed a difference of convex (DC) functions approximation for the joint nonlinear chance constrained problem, and solved it by a sequence of convex approximations.

Besides the scenario approximation approach, Ben-Tal and Nemirovski [6] proposed a conservative convex approximation, which includes the quadratic approximation for individual linear chance constraints. Nemirovski and Shapiro [76] provided the CVaR approximation and Bernstein approximation for individual chance constraints.

An alternative to the above approximated approaches consists in providing bounds based on using deterministic analytical approximations of chance constraints. For the case of individual chance constraint, the bounds are mainly based on extensions of Chebyshev inequality together with the first two moments [8, 52, 85]. For joint chance constraints, deterministic equivalent approximations have been widely studied in [23, 24, 25, 65, 99].

In the literature, several bounding techniques have been proposed for two-stage and multistage stochastic programs with expectation (see for instance [8, 84]). This class of problems brings computational complexity which increases exponentially with the size of the scenario tree, representing a discretization of the underlying random process. Even if a large discrete tree model is constructed, the problem might be untractable due to the curse of dimensionality. In this situation, easy-to-compute bounds have been proposed in the literature (see for instance [2, 37, 70, 71, 72]) by solving small size problems.

1.3 Distributionally Robust Chance Constrained Problems

In many practical situations, one can only obtain the partial information about the probability distribution of ξ . If we replace the real distribution by an estimated one, the obtained optimal solution may be infeasible in practice with high

probability [108]. For this reason, the distributionally robust chance constrained problems are proposed in [15, 36, 108].

$$\begin{aligned} \min_{\mathbf{x}} \quad & g(\mathbf{x}) \\ \text{s.t.} \quad & \inf_{F \in D} \mathbb{P}_F \{c_1(\mathbf{x}, \boldsymbol{\xi}) \leq 0, \dots, c_d(x, \boldsymbol{\xi}) \leq 0\} \geq \epsilon, \\ & \mathbf{x} \in \mathbf{X}, \end{aligned} \tag{1.3}$$

where D is an uncertainty set of the joint probability distribution F .

Early uncertainty sets of distribution are based on exact moments information of the random parameter. With this kind of uncertainty sets, El Ghaoui et al. [35], Popescu [86] and Chen et al. [18] studied the distributionally robust counterparts of many risk measures and expected value functions. Calafiore and El Ghaoui [15] reformulated a distributionally robust individual linear chance constrained problem as a SOCP problem with known mean and covariance. Zymler et al. [108] approximated the distributionally robust joint chance constrained problem as a tractable semidefinite programming (SDP) problem, and the authors prove that the proposed SDP is a reformulation when the chance constraint is individual. Li et al. [63] studied a single chance constraint with known first and second moments and unimodality of the distributions.

Considering the uncertainties in terms of the distribution and of the first two order moments, Delage and Ye [27] introduced the uncertainty set characterized by an elliptical constraint and a linear matrix inequality. Based on that, they transformed expected utility problems into optimization problems with linear matrix inequality constraints. Cheng et al. [21] established a distributionally robust chance constrained knapsack problem where the first order moment is fixed and the second order moment in the uncertainty set is contained in an elliptical set. Yang and Xu [103] showed that the distributionally robust chance constrained optimization is tractable if the uncertainty set is characterized by its mean and variance in a given set, and the constraint function is concave with respect to the decision variables and quasi-convex with respect to the uncertain parameters. In addition, Zhang et al. [104] developed a SOCP reformulation for a distributional family subject to an elliptical constraint and a linear matrix inequality. Moreover, Xie and Ahmed [102] showed that a distributionally robust joint chance constrained optimization problem is convex when the uncertainty set is specified by convex moment constraints.

Except for the above two kinds of uncertainty sets, Hanasusanto et al. [48] studied distributionally robust chance constraints where the distribution of uncertain parameters belongs to an uncertainty set characterized by the mean, the support and the dispersion information instead of the second order moment.

It is well known that the distributions of financial security returns are often skewed with high leptokurtic ([29]). This suggests to consider higher moments (especially skewness and kurtosis) in some realistic models. A tractable way to take higher order moments into consideration is to adopt the mixture distribution framework. Simply speaking, a mixture distribution is a convex combination of some component distributions. In this aspect, Hanasusanto et al. [46] considered a risk-averse newsvendor model where the demand distribution is supposed to be a

mixture distribution with known weights and unknown component distributions. Zhu and Fukushima [107] considered a robust portfolio selection problem with the finite normal mixture distribution, where the unknown mixture weights are restricted in linear or elliptical uncertainty sets.

However, for quite a few complex decision making problems under uncertainty, the precise knowledge about the necessary partial information is rarely available in reality. Hence, a strategy based on erroneous inputs might be infeasible or exhibit poor performance when implemented. To deal with such complex decision problems, a new framework called 'data-driven' robust optimization is proposed. This new scheme directly uses the observations of random variables as the inputs to the mathematical programming model and is now receiving particular attention in operations research community.

Basically, there are two kinds of data-driven approaches for distributionally robust optimizations. The first one estimates the parameters in an uncertainty set through statistical methods. For example, from the parametric estimation perspective, Delage and Ye [27] showed that the uncertainty set characterizing the uncertainties of first and second order moments can be estimated through statistical confidence regions, which can be calculated directly from the data. Bertsimas et al. [7] proposed a systematic scheme for estimating uncertainty sets from data by using statistical hypothesis tests. And they proved that robust optimization problems over each of their uncertainty sets are generally tractable. However, as far as we know, the tractable data-driven uncertainty sets can only characterize the first and the second moments. From the non-parametric perspective, Zhu et al.[106] showed that the data-driven uncertainty sets of mixture distribution weights can be estimated through Bayesian learning. Recently, Gupta [43] proposed a near-optimal Bayesian uncertainty set, which is smaller than uncertainty sets based on statistical confidence regions.

The second kind of data-driven approaches directly characterizes the uncertainty of a distribution through a probabilistic distance, and thus forms a new type of data-driven uncertainty sets. One of the most frequently used data-driven uncertainty sets is characterized by the distance function based on probability density, such as ϕ -divergence and Wasserstein distance. By utilizing sophisticated probabilistic measures, Ben-Tal et al.[4] proposed a class of data-driven uncertainty sets based on ϕ -divergences. Hu and Hong [55] studied distributionally robust individual chance constrained optimization problems where the uncertainty set of the probability distribution is defined by the Kullback-Leibler divergence, which is a special case of ϕ -divergences, the authors showed that the distributionally robust chance-constrained problem can be reformulated as the chance constrained problem with an adjusted confidence level. Jiang and Guan [57] considered a family of density-based uncertainty sets based on ϕ -divergence and proved that a distributionally robust joint linear chance constraint is equivalent to a chance constraint with a perturbed risk level. Recently, Esfahani and Kuhn [38] and Zhao and Guan [105] showed that under certain conditions, the distributionally robust expected utility optimization problem with Wasserstein distance is tractable. On the other hand, Hanasusanto et al. [47] showed that the distributionally robust joint linear chance constrained program with the uncertainty set characterized by Wasserstein distance is strongly NP-hard.

1.4 Contribution and Outline of the Dissertation

Motivated by the literature review, we conduct a systematic research in this dissertation on chance constrained problem and its applications. The main work of this dissertation is stated as following:

(1) First of all, as the basic knowledge of chance constrained problems, we review some standard work about chance constraints in Chapter 2. We firstly state some mathematical theories about convexity of large classes of chance constrained problems. Then, we review some standard research results about convex reformulations and solvable approximations for chance constraints. These results are fundamental for solving chance constrained problems. At last, we introduce moments based uncertainty set and distance based uncertainty set, which are important and popular uncertainty sets in research about distributionally robust chance constrained problems.

(2) Chance constrained geometric programs play an important role in many practical problems. Therefore, as a first step, we first review the recent results about geometric programs with joint chance constraints, where the stochastic parameters are normally distributed and independent of each other. To extend this work, we discuss joint rectangular geometric chance constrained programs under elliptical distribution with independent components. With standard variable transformation, a convex reformulation of rectangular geometric chance constrained programs can be derived. As for the quantile function of elliptical distribution in the reformulation, we propose convex approximations with piecewise linear approximation method, which are asymptotically tight approximations. This result is discussed in Chapter 3.

(3) As equivalent reformulations of chance constrained problems are hard to obtained in most cases, finding bounds of chance constraints is a tractable and feasible approach for solving chance constrained problems. From this viewpoint, we develop bounds for individual and joint chance constrained problems with independent matrix vector rows. The deterministic bounds of chance constraints are based on the one-side Chebyshev inequality, Chernoff inequality, Bernstein and Hoeffding inequalities, respectively. Under various sufficient conditions related to the confidence parameter value, we show that the aforementioned approximations are convex and tractable. To obtain the approximations by applying probability inequalities, Chebyshev inequality requires the knowledge of the first and second moments of the random variables while Bernstein and Hoeffding ones require their mean and their support. On the contrary, Chernoff inequality requires only the moment generating function of the random variables. In order to reduce further the computational complexity of the problem, we propose approximations based on piecewise linear and tangent. This work is illustrated precisely in Chapter 4.

(4) For data-driven robust optimization problems, to catch the fat-tailness, skewness and high leptokurticness properties of the random variables, we construct a data-driven mixture distribution based uncertainty set, which can characterize higher order moments information from data. Such a data-driven uncertainty set can more efficiently match non-Gaussian characters of real random

variables. Then, with the data-driven mixture distribution based uncertainty set, we study a distributionally robust individual linear chance constrained problem. Under certain conditions of parameters, the distributionally robust chance constrained problem can be reformulated as a convex programming problem. As the convex equivalent reformulation contains a quantile function, we further propose two approximations leading to tight upper and lower bounds. Moreover, under much weaker conditions, the distributionally robust chance constrained problem can be reformulated as a DC programming problem. In this case, we propose a sequence convex approximation method to find a tight upper bound, and use relaxed convex approximation method to find a lower bound. This work is shown in Chapter 5.

(5) As an application of chance constrained problems, we consider an n -player non-cooperative game with stochastic strategy sets, which is constructed by a set of stochastic linear constraints. For each player, the stochastic linear constraint is formulated as a joint chance constraint. In addition, we further assume that each random vector of the matrix defining stochastic linear constraints is pairwise independent. Under certain conditions, we propose new convex reformulations for the joint chance constraints, following normally, elliptically distributed and distributionally robust framework, respectively. We show that there exists a Nash equilibrium of such a chance constrained game if the payoff function of each player satisfies certain assumptions. This part of work is stated in Chapter 6.

In Chapter 7, we conclude the main work of this dissertation and develop a discussion on open issues and questions and future work.

Chapter 2

Chance Constraints

In this chapter, we summarize some standard work about chance constraints. We first introduce some mathematical theories which can be used to prove the convexity of large classes of chance constrained problems. Then, we state some convex reformulations and approximations for chance constraints which lead to chance constrained problems tractable. Finally, we discuss distributionally robust chance constrained problems with moments based uncertainty set and distance based uncertainty set.

2.1 α -concave measures and convexity of chance constraints

Logconcave measure was introduced by Prékopa [88] in the stochastic programming framework but they became widely used also in statistics, convex geometry, mathematical analysis, economics, etc.

Definition 2.1. A function $f(z) \geq 0, z \in \mathbb{R}^n$, is said to be logarithmically concave (logconcave), if for any z_1, z_2 and $0 < \lambda < 1$ we have the inequality

$$f(\lambda z_1 + (1 - \lambda)z_2) \geq [f(z_1)]^\lambda [f(z_2)]^{(1-\lambda)}. \quad (2.1)$$

If $f(z) > 0$ for $z \in \mathbb{R}^n$, then this means that $\log f(z)$ is a convex function in \mathbb{R}^n .

Definition 2.2. A probability measure defined on the Borel sets of \mathbb{R}^n is said to be logarithmically concave (logconcave) if for any convex subsets of \mathbb{R}^n : A, B and $0 < \lambda < 1$ we have the inequality

$$\mathbb{P}(\lambda A + (1 - \lambda)B) \geq [\mathbb{P}(A)]^\lambda [\mathbb{P}(B)]^{(1-\lambda)}, \quad (2.2)$$

where $\lambda A + (1 - \lambda)B = \{z = \lambda x + (1 - \lambda)y | x \in A, y \in B\}$.

Theorem 2.1. If the probability measure \mathbb{P} is absolutely continuous with respect to Lebesgue measure and is generated by a logconcave probability density function then the measure \mathbb{P} is logconcave.

In [88], Prékopa also proved the following simple consequences of Theorem 2.1.

Theorem 2.2. *If \mathbb{P} is a logconcave probability distribution and $A \in \mathbb{R}^n$ is a convex set, then $\mathbb{P}(A + x)$, $x \in \mathbb{R}^n$, is a logconcave function.*

Theorem 2.3. *If $\xi \in \mathbb{R}^n$ is a random variable, whose probability distribution is logconcave, then the probability distribution function $F(x) = \mathbb{P}(\xi \leq x)$ is a logconcave function in \mathbb{R}^n .*

Theorem 2.4. *If $n = 1$ in Theorem 2.3 then also $1 - F(x) = \mathbb{P}(\xi > x)$ is a logconcave function in \mathbb{R}^1 .*

With these fundamental theorems, some convexity results of chance constraints can be derived. Consider the following minimization problems:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & \mathbb{P}\{g_1(x) \geq \beta_1, \dots, g_m(x) \geq \beta_m\} \geq \epsilon, \end{aligned} \quad (2.3)$$

where β_1, \dots, β_m are random variables, $\epsilon \in (0, 1)$ is a prescribed probability and $g_1(x), \dots, g_m(x)$ are convex functions in the whole space \mathbb{R}^n .

With Theorem 2.1-Theorem 2.4, Prékopa proved that the function $h(x) = \mathbb{P}\{g_1(x) \geq \beta_1, \dots, g_m(x) \geq \beta_m\}$ is logarithmic concave in the whole space \mathbb{R}^n . By taking the logarithm of both sides of the constraint (2.3), we can obtain a convex problem.

As generalizations of logconcave measures, Borell [9] and Brascamp and Lieb [11] introduced the α -concave measure.

Definition 2.3. *A nonnegative function $f(x)$ defined on a convex set $\Omega \subset \mathbb{R}^n$ is said to be α -concave, where $\alpha \in [-\infty, +\infty]$, if for all $x, y \in \Omega$ and all $\lambda \in [0, 1]$ the following inequality holds true:*

$$f(\lambda x + (1 - \lambda)y) \geq m_\alpha(f(x), f(y), \lambda),$$

where $m_\alpha : \mathbb{R}_+ \times \mathbb{R}_+ \times [0, 1] \rightarrow \mathbb{R}$ is defined as follows:

$$m_\alpha(a, b, \lambda) = 0 \text{ if } ab = 0,$$

and if $a > 0$, $b > 0$, $0 \leq \lambda \leq 1$, then

$$m_\alpha(a, b, \lambda) = \begin{cases} a^\lambda b^{1-\lambda} & \text{if } \alpha = 0, \\ \max\{a, b\} & \text{if } \alpha = \infty, \\ \min\{a, b\} & \text{if } \alpha = -\infty, \\ (\lambda a^\alpha + (1 - \lambda)b^\alpha)^{1/\alpha} & \text{otherwise.} \end{cases}$$

In the case of $\alpha = 0$, the function f is called logarithmically concave or log-concave because $\ln f(\cdot)$ is a concave function. In the case of $\alpha = 1$, the function f is simply concave. In the case of $\alpha = -\infty$, the function f is quasi-concave.

By the definition of the α -concave function, we have the following theorems.

Theorem 2.5. *Let f be a concave function defined on a convex set $C \subset \mathbb{R}^s$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ be a nonnegative nondecreasing α -concave function, $\alpha \in [-\infty, \infty]$. Then the function $g \circ f$ is α -concave.*

Theorem 2.6. *If the functions $f_i : \mathbb{R}^n \rightarrow \mathbb{R}_+, i = 1, \dots, m$, are α_i -concave and α_i are such that $\sum_{i=1}^m \alpha_i^{-1} > 0$, then the function $g : \mathbb{R}^{nm} \rightarrow \mathbb{R}_+$, defined as $g(x) = \prod_{i=1}^m f_i(x_i)$ is γ -concave with $\gamma = (\sum_{i=1}^m \alpha_i^{-1})^{-1}$.*

We point out that for two Borel measurable sets A, B in \mathbb{R}^s , the Minkowski sum $A + B = \{x + y : x \in A, y \in B\}$ is Lebesgue measurable in \mathbb{R}^s .

Definition 2.4. *A probability measure \mathbb{P} defined on the Lebesgue measurable subsets of a convex set $\Omega \subset \mathbb{R}^s$ is said to be α -concave if for any Borel measurable sets $A, B \subset \Omega$ and for all $\lambda \in [0, 1]$, we have the inequality*

$$\mathbb{P}(\lambda A + (1 - \lambda)B) \geq m_\alpha(\mathbb{P}(A), \mathbb{P}(B), \lambda),$$

where $\lambda A + (1 - \lambda)B = \{\lambda x + (1 - \lambda)y : x \in A, y \in B\}$.

Then, as shown [96], we have the following theorems about convexity of chance constraints.

Theorem 2.7. *Let the functions $g_i : \mathbb{R}^n \rightarrow \mathbb{R}^s, i = 1, \dots, m$, be quasi-concave. If $Z \in \mathbb{R}^s$ is a random vector that has an α -concave probability distribution, then the function*

$$G(x) = \mathbb{P}\{g_i(x, Z) \geq 0, i = 1, \dots, m\}$$

is α -concave on the set

$$D = \{x \in \mathbb{R}^n : \exists z \in \mathbb{R}^s \text{ such that } g_i(x, z) \geq 0, i = 1, \dots, m\}.$$

As a consequence, we obtain convexity statements for sets described by probabilistic constraints.

Theorem 2.8. *Assume that the functions $g_i(\cdot, \cdot), i = 1, \dots, m$, are quasi-concave jointly in both arguments and that $Z \in \mathbb{R}^s$ is a random vector that has an α -concave probability distribution. Then the following set is convex and closed:*

$$X_0 = \{x \in \mathbb{R}^n : \mathbb{P}\{g_i(x, Z) \geq 0, i = 1, \dots, m\} \geq \epsilon\}.$$

We consider the case of a separable mapping g when the random quantities appear only on the right-hand side of the inequalities.

Theorem 2.9. *Let the mapping $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be such that each component g_i is a concave function. Furthermore, assume that the random vector Z has independent components and the one-dimensional marginal distribution functions $F_{Z_i}, i = 1, \dots, m$, are α_i -concave. Furthermore, let $\sum_{i=1}^k \alpha_i^{-1} > 0$. Then the set*

$$X_0 = \{x \in \mathbb{R}^n : \mathbb{P}\{g(x) \geq Z\} \geq \epsilon\}$$

is convex.

For elliptical distributions, Henrion [49] gave a convexity result of the feasible solution set defined by individual linear chance constraints.

A multivariate distribution is said to be elliptical if its characteristic function $\phi(t)$ is of the form $\phi(t) = e^{it^T \mu} \psi(t^T \Sigma t)$ for a specified μ , positive-definite matrix

Σ , and the characteristic generator function ψ . For an s -dimensional random vector ξ , we denote $\xi \sim \text{Ellip}(\mu, \Sigma; \phi)$ if ξ has an elliptical distribution.

When the density functions exist, they have the following structure:

$$f(x) = \frac{c}{\sqrt{\det \Sigma}} g\left(\sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)}\right), \quad (2.4)$$

where function $g : \mathbb{R}_+ \rightarrow \mathbb{R}_{++}$ is the so-called radial density, x is an n -dimensional random vector with median vector μ (which is also the mean vector if exists), Σ is a positive definite matrix which is proportional to the covariance matrix if exists, and $c > 0$ is a normalization factor ensuring that f integrates to one.

Law	Characteristic generator	Radial density
normal	$\exp\{-\frac{1}{2}t\}$	$\exp\{-\frac{1}{2}t^2\}$
t	*	$(1 + \frac{1}{\nu}t^2)^{-(n+\nu)/2}$
Cauchy	$\exp\{-\sqrt{t}\}$	$(1 + t^2)^{-(n+1)/2}$
Laplace	$(1 + \frac{1}{2}t)^{-1}$	$\exp\{-\sqrt{2} t \}$
logistic	$\frac{2\pi\sqrt{t}}{e^{\pi\sqrt{t}} - e^{-\pi\sqrt{t}}}$	$\frac{e^{-t^2}}{(1+e^{-t^2})^2}$

Table 2.1: Table of selected elliptical distributions.

Remark 2.1. *Table 2.1 provides a short selection of prominent multivariate elliptical distributions, together with their characteristic generators and radial densities. Note that Cauchy distribution is a special case of the t distribution with $\nu = 1$.*

Consider the following linear chance constraint sets

$$M_p^\alpha := \{x \in \mathbb{R}^n \mid \mathbb{P}(\langle q(x), \xi \rangle \leq p) \geq \epsilon\}$$

Here, ξ is an s -dimensional random vector defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $q : \mathbb{R}^n \rightarrow \mathbb{R}^s$ is a mapping from the space of decision vectors to the space of realizations of the random vector. $p \in \mathbb{R}$, $\epsilon \in (0, 1)$.

In the following, the norm induced by a positive definite matrix C is $\|x\|_C = \sqrt{\langle x, Cx \rangle}$. Moreover, for a one-dimensional distribution function F , we define its ϵ -quantile as $F^{-1}(\epsilon) = \inf\{t \mid F(t) \geq \epsilon\}$. Among many properties of elliptical distributions, we notice that the class of elliptical distributions is closed under affine transformation. Then, we have the following lemma (Proposition 2.1 in [49]).

Lemma 2.1 ([49]). *Let q be an arbitrary mapping and let ξ have an elliptically symmetric distribution with parameters Σ , μ , where Σ is positive definite. Denote by ψ its characteristic generator. Then,*

$$M_\epsilon^p := \{x \in \mathbb{R}^n \mid \langle \mu, q(x) \rangle + \Psi^{(-1)}(\epsilon) \|q(x)\|_\Sigma \leq p\}$$

where Ψ is one-dimensional distribution function induced by the characteristic function $\phi(t) = \psi(t^2)$.

Theorem 2.10 ([49]). *In addition to the setting of Lemma 2.1, let one of the following assumptions hold true.*

- q is affine linear
- q has nonnegative, convex components, μ_i for $i = 1, \dots, s$ and all elements of Σ are nonnegative.

Then, M_p^α is convex for all $p \in \mathbb{R}$ and all $\epsilon > 0.5$. If moreover, the random vector ξ in Lemma 2.1 has a strictly positive density, then M_ϵ^p is convex for all $p \in \mathbb{R}$ and all $\epsilon \geq 0.5$.

As generalization, Henrion and Strugarek [50] further proved the convexity of joint chance constraints, where random vectors appear separated from decision vectors. More precisely,

$$M(\epsilon) = \{x \in \mathbb{R}^n | \mathbb{P}(\xi \leq q(x)) \geq \epsilon\}, \quad (2.5)$$

where $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is some vector-valued mapping. With $F : \mathbb{R}^m \rightarrow \mathbb{R}$ denoting the distribution function of ξ , the same set can be rewritten as

$$M(\epsilon) = \{x \in \mathbb{R}^n | F(q(x)) \geq \epsilon\}.$$

We focus on conditions on F and g such that $M(\epsilon)$ becomes a convex set for all $\epsilon \geq \epsilon^*$, where $\epsilon^* < 1$.

Definition 2.5. *We call a function $f : \mathbb{R} \rightarrow \mathbb{R}$ r -decreasing for some $r \in \mathbb{R}$, if it is continuous on $(0, \infty)$ and if there exists some $t^* > 0$ such that the function $t^r f(t)$ is strictly decreasing for all $t > t^*$.*

Lemma 2.2. *Let $F : \mathbb{R} \rightarrow [0, 1]$ be a distribution function with $(r+1)$ -decreasing density f for some $r > 0$. Then, the function $z \mapsto F(z^{-1/r})$ is concave on $(0, (t^*)^{-r})$, where t^* refers to Definition 2.5. Moreover, $F(t) < 1$ for all $t \in \mathbb{R}$.*

Theorem 2.11 ([50]). *For (2.5), we make the following assumptions for $i = 1, \dots, m$:*

1. *There exists $r_i > 0$ such that the components g_i are $(-r_i)$ -concave.*
2. *The components ξ_i of ξ are independently distributed with (r_i+1) -decreasing densities f_i .*

Then, $M(\epsilon)$ is convex for all $\epsilon > \epsilon^* := \max\{F_i(t_i^*) | 1 \leq i \leq m\}$, where F_i denotes the distribution function of ξ_i and the t_i^* refers to Definition 2.5 in the context of f_i being $(r_i + 1)$ -decreasing.

2.2 Reformulations and approximations for chance constraints

2.2.1 Individual chance constraints

As shown in [49], for elliptically distributed random vector $\xi \sim \text{Ellip}(\mu, \Sigma; \phi)$, consider the following chance constraint:

$$S(\alpha) := \{x \in \mathbb{R}^n \mid \mathbb{P}(\xi^T x \leq h) \geq \alpha\}. \quad (2.6)$$

From Lemma 2.1, we have

$$S(\alpha) := \left\{ x \in \mathbb{R}^n \mid \mu^T x + \Psi^{(-1)}(\alpha) \sqrt{x^T \Sigma x} \leq h \right\},$$

where Ψ is the one-dimensional distribution function induced by the characteristic function $\phi(t) = \psi(t^2)$.

For some elliptical distributions, the inverse function $\Psi^{(-1)}(\alpha)$ can be written in a closed form. Thus, the set $S(\alpha)$ can be reformulated as the following examples.

Example 2.1. (Cauchy distribution) Suppose that ξ follows a Cauchy distribution with parameters μ, Σ . The inverse function of cumulative distribution function of Cauchy distribution is as follows:

$$\Psi^{-1}(\alpha) = \tan \left[\pi \left(\alpha - \frac{1}{2} \right) \right], \quad \alpha \in (0, 1).$$

The set $S(\alpha)$, defined in (2.6), can be rewritten as

$$S(\alpha) = \left\{ x \in X \mid \mu^T x + \tan \left[\pi \left(\alpha - \frac{1}{2} \right) \right] \sqrt{x^T \Sigma x} \leq h \right\}.$$

Since $\tan \left[\pi \left(\alpha - \frac{1}{2} \right) \right] \geq 0$ when $\alpha \geq \frac{1}{2}$, the set $S(\alpha)$ is convex if $\alpha \geq \frac{1}{2}$. While the set $S(\alpha)$ is concave if $\alpha < \frac{1}{2}$.

Example 2.2. (Laplace distribution) Suppose that ξ follows a Laplace distribution with parameters μ, Σ . The inverse function of cumulative distribution function of Laplace distribution is as follows:

$$\Psi^{-1}(\alpha) = \begin{cases} \ln(2\alpha), & \alpha \in [0, \frac{1}{2}], \\ -\ln[2(1-\alpha)], & \alpha \in [\frac{1}{2}, 1]. \end{cases}$$

The set $S(\alpha)$, defined in (2.6), can be rewritten as

$$S(\alpha) = \begin{cases} \left\{ x \in X \mid \mu^T x + \ln(2\alpha) \sqrt{x^T \Sigma x} \leq h \right\}, & \alpha \in [0, \frac{1}{2}], \\ \left\{ x \in X \mid \mu^T x - \ln[2(1-\alpha)] \sqrt{x^T \Sigma x} \leq h \right\}, & \alpha \in [\frac{1}{2}, 1]. \end{cases}$$

It is easy to see that when $\alpha \in [\frac{1}{2}, 1]$, the set $S(\alpha)$ is convex since $-\ln[2(1-\alpha)] \geq$

0. Otherwise, the set $S(\alpha)$ is concave.

Example 2.3. (Logistic distribution) Suppose that ξ follows a Logistic distribution with parameters μ, Σ . The inverse function of cumulative distribution function of Logistic distribution is as follows:

$$\Psi^{-1}(\alpha) = \ln \left(\frac{\alpha}{1 - \alpha} \right), \quad \alpha \in (0, 1).$$

The set $S(\alpha)$, defined in (2.6), can be rewritten as

$$S(\alpha) = \left\{ x \in X \mid \mu^T x + \ln \left(\frac{\alpha}{1 - \alpha} \right) \sqrt{x^T \Sigma x} \leq h \right\}.$$

Since $\ln \left(\frac{\alpha}{1 - \alpha} \right) \geq 0$ when $\alpha \geq \frac{1}{2}$, the set $S(\alpha)$ is convex if $\alpha \geq \frac{1}{2}$. While the set $S(\alpha)$ is concave if $\alpha < \frac{1}{2}$.

Besides elliptical distributions, Calafiore and El Ghaoui [15] reformulated individual chance constraints under Q -radial distribution.

Definition 2.6. A random vector $d \in \mathbb{R}^n$ has a Q -radial distribution with defining function $g(\cdot)$ if $d - E\{d\} = Q\omega$ where $Q \in \mathbb{R}^{n \times \nu}$, $\nu \leq n$, is a fixed, full-rank matrix and $\omega \in \mathbb{R}^\nu$ is a random vector with probability density f_ω that depends only on the norm of ω ; i.e., $f_\omega(\omega) = g(\|\omega\|)$. The function $g(\cdot)$ that defines the radial shape of the distribution is named the defining function of d .

The family of Q -radial distributions includes all probability densities whose level sets are ellipsoids with shape matrix $Q \succ 0$ and may have any radial behavior. Another notable example is the uniform density over an ellipsoidal set.

Theorem 2.12 ([15]). For any $\alpha \in [0.5, 1)$, the chance constraint

$$\mathbb{P}\{\xi^T x \leq 0\} \geq \epsilon,$$

where ξ has a Q -radial distribution with defining function $g(\cdot)$ and covariance Γ , is equivalent to the convex second-order cone constraint

$$\kappa_{\alpha,r} \sqrt{x^T \Gamma x} + \xi^T x < 0, \quad (2.7)$$

where $\kappa_{\alpha,r} = v\Psi^{-1}(\alpha)$, with $v := (V_\nu \int_0^\infty r^{\nu+1} g(r) dr)^{-1/2}$, V_ν denotes the volume of the Euclidean ball of unit radius in \mathbb{R}^ν , Ψ the cumulative probability function of density $f(\xi) = S_{\nu-1} \int_0^\infty g(\sqrt{\rho^2 + \xi^2}) \rho^{\nu-2} d\rho$, S_n denotes the surface measure of the Euclidean ball of unit radius in \mathbb{R}^n .

In the case where ξ follows a normal distribution with mean $\hat{\mu}$ and covariance Γ , which is Q -radial with $v = 1$, $Q = \Gamma_f$, $\Gamma = \Gamma_f \Gamma_f^T$, the function $g(r)$ is

$$g(r) = \frac{1}{(2\pi)^{1/2}} e^{-\frac{r^2}{2}}.$$

Therefore, for $\epsilon \in (0, 0.5]$, the parameter $\kappa_{\epsilon, r}$ is given by $\kappa_{\epsilon, r} = \Psi_G^{-1}(1 - \epsilon)$, where Ψ_G is the standard normal cumulative distribution function. Hence, (2.7) can be rewritten as

$$\Psi_G^{-1}(\alpha) \sqrt{x^T \Gamma x} + \xi^T x < 0,$$

which is convex.

2.2.2 Normally distributed joint linear chance constraints

For joint chance constrained problems, Cheng and Lisser [24] considered the following normally distributed linear program with joint chance constraint:

$$\begin{aligned} \min_{\mathbf{x}} \quad & c^T \mathbf{x} \\ \text{s.t.} \quad & \mathbb{P}\{T\mathbf{x} \leq D\} \geq \epsilon, \\ & \mathbf{x} \in \mathbf{X}, \end{aligned} \quad (2.8)$$

where $\mathbf{X} \subset \mathbb{R}_+^n$ is polyhedron, $c \in \mathbb{R}^n$, $D = (D_1, \dots, D_K) \in \mathbb{R}^K$, ϵ is a prespecified confidence parameter, $T = [T_1, \dots, T_K]^T$ is a $K \times n$ random matrix, where T_k , $k = 1, \dots, K$, is a normally distributed random vector in \mathbb{R}^n with mean vector $\mu_k = (\mu_{k1}, \dots, \mu_{kn})^T$ and covariance matrix Σ_k . Moreover, the multivariate normally distributed vectors T_k , $k = 1, \dots, K$, are independent.

Then, a deterministic reformulation of (2.8) can be derived as follows:

$$\begin{aligned} \min_{\mathbf{x}} \quad & c^T \mathbf{x} \\ \text{s.t.} \quad & \mu_k^T \mathbf{x} + F^{-1}(\epsilon^{y_k}) \|\Sigma_k^{1/2} \mathbf{x}\| \leq D_k, k = 1, \dots, K \\ & \sum_{k=1}^K y_k = 1, y_k \geq 0 \\ & \mathbf{x} \in \mathbf{X} \end{aligned} \quad (2.9)$$

Because of the nonelementary function $F^{-1}(\epsilon^{y_k})$, two piecewise linear approximations are applied to $F^{-1}(\epsilon^{y_k})$.

Piecewise tangent approximation of $F^{-1}(\epsilon^z)$

The authors of [24] use the linear tangents $a_s z + b_s$, $s = 1, \dots, S$, between $z_1, z_2, \dots, z_S \in (0, 1]$ to form a piecewise linear function

$$l(z) = \max_{s=1, \dots, S} \{a_s z + b_s\},$$

where

$$a_s = (F^{-1})'(\epsilon^{z_s}) \epsilon^{z_s} \ln \epsilon, \quad b_s = F^{-1}(\epsilon^{z_s}) - a_s z_s.$$

Then, (2.9) can be approximated by the following second order cone programming:

$$\begin{aligned} \min_{\mathbf{x}} \quad & c^T \mathbf{x} \\ \text{s.t.} \quad & \mu_k^T \mathbf{x} + \|\Sigma_k^{1/2} \tilde{z}_k\| \leq D_k, k = 1, \dots, K \\ & \tilde{z}_k \geq a_s x_i + b_s y_{ki}, s = 0, 1, \dots, S, i = 1, \dots, n \\ & \sum_{k=1}^K y_{ki} = x_i, y_{ki} \geq 0, i = 1, \dots, n \\ & \mathbf{x} \in \mathbf{X}, \end{aligned} \quad (2.10)$$

where $a_0 = 0, b_0 = 0$. Additionally, the optimal value of problem (2.10) is a lower

bound of problem (2.9).

Piecewise segment approximation of $F^{-1}(\epsilon^z)$

To get the piecewise segment approximation of $F^{-1}(\epsilon^z)$, the linear segments $\bar{a}_s z + \bar{b}_s$, $s = 1, \dots, S$, between $z_1, z_2, \dots, z_S \in (0, 1]$ are used to form a piecewise linear function

$$\bar{l}(z) = \max_{s=1, \dots, S} \{ \bar{a}_s z + \bar{b}_s \},$$

where

$$\bar{a}_s = \frac{F^{-1}(\epsilon^{z_{s+1}}) - F^{-1}(\epsilon^{z_s})}{z_{s+1} - z_s}, \quad \bar{b}_s = F^{-1}(\epsilon^z) - \bar{a}_s z_s.$$

Then, (2.9) can be approximated by the following second order cone programming:

$$\begin{aligned} \min_{\mathbf{x}} \quad & c^T \mathbf{x} \\ \text{s.t.} \quad & \mu_k^T \mathbf{x} + \|\Sigma_k^{1/2} \tilde{z}_k\| \leq D_k, k = 1, \dots, K \\ & \tilde{z}_k \geq \bar{a}_s x_i + \bar{b}_s y_{ki}, s = 0, 1, \dots, S, i = 1, \dots, n \\ & \sum_{k=1}^K y_{ki} = x_i, y_{ki} \geq 0, i = 1, \dots, n \\ & \mathbf{x} \in \mathbf{X}, \end{aligned} \tag{2.11}$$

where $\bar{a}_0 = 0, \bar{b}_0 = 0$. In addition, the optimal value of problem (2.11) is an upper bound of problem (2.9).

2.2.3 Integer programming approaches for joint linear chance constraints

When a random vector in chance constraints has a finite distribution, Luedtke et. al. [68] considered a joint chance constrained linear programming problem with random right-hand side given by

$$\begin{aligned} \min_{\mathbf{x}} \quad & c^T \mathbf{x} \\ \text{s.t.} \quad & \mathbb{P}\{T\mathbf{x} \geq \xi\} \geq 1 - \epsilon, \\ & \mathbf{x} \in \mathbf{X}, \end{aligned} \tag{2.12}$$

where $\mathbf{X} \subset \mathbb{R}_+^d$ is polyhedron, T is an $m \times d$ matrix, ξ is a random vector in \mathbb{R}^m , $\epsilon \in (0, 1)$ and $c \in \mathbb{R}^d$. Assume that ξ has finite support, that is, there exists vectors, $\xi_i \in \mathbb{R}^m, i = 1, \dots, n$ such that $\mathbb{P}\{\xi = \xi_i\} = \pi_i$ for each i where $\pi_i > 0$ and $\sum_{i=1}^n \pi_i = 1$. Without loss of generality, they assume that $\xi_i \geq 0$. In the following, we assume $\pi_i \leq \epsilon$ for each i .

To formulate (2.12) as a mixed-integer program, we introduce for each i a binary variable z_i , where $z_i = 0$ guarantees that $Tx \geq \xi_i$. Observe that because $\epsilon < 1$ we must have $Tx \geq \xi_i$ for at least one i . Since $\xi_i \geq 0$ for all i , this implies $Tx \geq 0$ in every feasible solution of (2.12). Then, let $v = Tx$, we obtain the

MIP formulation of (2.12):

$$\begin{aligned}
& \min_{\mathbf{x}} && c^T \mathbf{x} \\
& \text{s.t.} && \mathbf{x} \in \mathbf{X}, Tx - v = 0 \\
& && v + \xi_i z_i \geq \xi_i, i = 1, \dots, n \\
& && \sum_{i=1}^n \pi_i z_i \leq \epsilon \\
& && z \in \{0, 1\}^n
\end{aligned} \tag{2.13}$$

The assumption of finite distribution may seem restrictive. However, if the possible values for ξ are generated through Monte Carlo sampling from a general distribution, the resulting problem can be viewed as an approximation of a problem with general distribution. There is theoretical and empirical evidence which demonstrates that such a sample approximation can indeed be used to approximately solve problems with continuous distribution. A standard result will be introduced in subsection 2.2.5.

2.2.4 Convex approximations for chance constraints

An alternative to the above approaches consists in providing tractable convex approximations for chance constraints. Nemirovski and Shapiro [76] studied the following optimization problem:

$$\min_{x \in X} f(x) \quad \text{subject to} \quad \mathbb{P}\{F(x, \xi) \leq 0\} \geq 1 - \epsilon. \tag{2.14}$$

Here ξ is a random vector with probability distribution P supported on a set $\Xi \subset \mathbb{R}^d$, $X \subset \mathbb{R}^n$ is a nonempty convex set, $\epsilon \in (0, 1)$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a real valued convex function, $F = (f_1, \dots, f_m) : \mathbb{R}^n \times \Xi \rightarrow \mathbb{R}^m$, and $\mathbb{P}(A)$ denotes probability of an event A .

The chance constraint of problem (2.14) is equivalent to the constraint

$$p(x) := \mathbb{P}\{F(x, \xi) > 0\} \leq \epsilon.$$

Let $\mathbb{1}_A$ denotes the indicator function of a set A , i.e., $\mathbb{1}_A(z) = 1$ if $z \in A$ and $\mathbb{1}_A(z) = 0$ if $z \notin A$.

Let $\psi : \mathbb{R} \rightarrow \mathbb{R}$ be a nonnegative valued, nondecreasing, convex function satisfying the following property:

$$\psi(z) > \psi(0) = 1 \text{ for any } z > 0.$$

We refer to function $f(z)$ which satisfies the above properties as a (one-dimensional) generating function. It follows from the property that for $t > 0$ and random variable Z ,

$$\mathbb{E}[\psi(tZ)] \geq \mathbb{E}[\mathbb{1}_{[0, +\infty)}(tZ)] = \mathbb{P}\{tZ \geq 0\} = \mathbb{P}\{Z \geq 0\}.$$

By taking $Z = F(x, \xi)$ and changing t to t^{-1} , we have

$$p(x) \leq \mathbb{E}[\psi(t^{-1}F(x, \xi))] \tag{2.15}$$

for all x and $t > 0$. Let

$$\Psi(x, t) = t\mathbb{E}[\psi(t^{-1}F(x, \xi))].$$

Then, we have

$$\inf_{t>0} [\Psi(x, t) - t\epsilon] \leq 0 \quad \text{implies} \quad p(x) \leq \epsilon$$

Therefore, we obtain, under the assumption that X , $f(\cdot)$ and $F(\cdot, \xi)$ are convex, that

$$\min_{x \in X} f(x) \quad \text{subject to} \quad \inf_{t>0} [\Psi(x, t) - t\epsilon] \leq 0 \quad (2.16)$$

gives a convex conservative approximation of the chance constrained problem (2.14).

Since $\psi(0) = 1$ and $\psi(\cdot)$ is convex and nonnegative, we conclude that $\psi(z) \geq \max\{1 + az, 0\}$ for all z , so that the upper bounds (2.15) can be only improved when replacing $\psi(z)$ with the function $\hat{\psi}(z) := \max\{1 + az, 0\}$, which is also a generating function. But the bounds produced by the latter function are, up to scaling $z \leftarrow z/a$, the same as those produced by the function

$$\psi^*(z) := [1 + z]_+, \quad (2.17)$$

where $[a]_+ := \max\{a, 0\}$. That is, from the point of view of the most accurate approximation, the best choice of the generating function ψ is the piecewise linear function ψ^* defined in (2.17). For the generating function ψ^* defined in (2.17), the approximate constraint in problem (2.16) takes the form

$$\inf_{t>0} [\mathbb{E} [[F(x, \xi) + t]_+] - t\epsilon] \leq 0. \quad (2.18)$$

Replacing in the left-hand side $\inf_{t>0}$ with \inf_t , we clearly do not affect the validity of the relation; thus, we can equivalently rewrite (2.18) as

$$\inf_{t \in \mathbb{R}} [\mathbb{E} [-t\epsilon + [F(x, \xi) + t]_+]] \leq 0.$$

In that form, the constraint is related to the definition of conditional value at risk (CVaR). Then, the constraint

$$\text{CVaR}_{1-\epsilon}[F(x, \xi)] \leq 0.$$

defines a convex conservative approximation of the chance constraint in problem (2.14).

2.2.5 Sample average approximations for chance constraints

Besides the above approaches for chance constraints, the sample average approximation (SAA) method is also a computationally tractable approximation for chance constrained problems.

In [96], the authors considered a chance constrained problem of the form

$$\min_{x \in X} f(x) \quad \text{s.t.} \quad p(x) \leq \epsilon, \quad (2.19)$$

where $X \subset \mathbb{R}^n$ is a closed set, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuous function, $\epsilon \in (0, 1)$ is a given tolerance level, and $p(x) = \mathbb{P}\{C(x, \xi) > 0\}$. We assume that ξ is a random vector, whose probability distribution \mathbb{P} is supported on set $\Xi \in \mathbb{R}^d$, and the function $C : \mathbb{R}^n \rightarrow \mathbb{R}$ is a Carathéodory function.

For the sake of simplicity, we assume that the objective function $f(x)$ is given explicitly and only the chance constraints should be approximated. We can write the probability $p(x)$ with the expectation form,

$$p(x) = \mathbb{E}[\mathbb{1}_{(0, \infty)}(C(x, \xi))],$$

and estimate this probability by the corresponding SAA function

$$\hat{p}(x) = \frac{1}{N} \sum_{j=1}^N \mathbb{1}_{(0, \infty)}(C(x, \xi^j)).$$

Consequently we can write the corresponding SAA problem as

$$\min_{x \in X} f(x) \quad \text{s.t.} \quad \hat{p}_N(x) \leq \epsilon. \quad (2.20)$$

Denote ϑ^* and S the optimal value and the optimal solution set of problem (2.19), respectively, and ϑ^N and S^N the optimal value and the optimal solution set of problem (2.20), respectively. Then, we have the following consistency properties of the optimal value $\hat{\vartheta}_N$ and the set \hat{S}_N of optimal solutions of the SAA problem (2.20).

Theorem 2.13 ([96]). *Suppose that X is a compact set, the function $f(x)$ is continuous, $C(x, \xi)$ is a Carathéodory function, the samples are iid, and the following condition holds: (a) there exists an optimal solution \bar{x} to the true problem such that for any $\delta > 0$, there exists an $x \in X$ with $\|x - \bar{x}\| \leq \delta$ and $p(x) < \epsilon$. Then $\hat{\vartheta}_N \rightarrow \vartheta^*$ and $\mathbb{D}(\hat{S}_N, S) \rightarrow 0$ w.p.1 as $N \rightarrow \infty$.*

2.3 Distributionally robust chance constraints

Since in many practical situations, one can only obtain the partial information about probability distributions. Replacing the real distribution by an estimated one, we may obtain an infeasible solution in practice with high probability. Therefore, the distributionally robust chance constrained approaches are proposed.

2.3.1 Distributionally robust individual chance constraints with known mean and covariance

With given mean and covariance, Calafiore and El Ghaoui [15] studied a distributionally robust individual chance constrained problem where the family \mathcal{P} of

probability distributions on the random vector is composed of all distributions with given mean $\boldsymbol{\mu}$ and covariance Σ . The distribution family is denoted by $\mathcal{P} = (\boldsymbol{\mu}, \Sigma)$. Then, the authors of [15] proved the following theorem.

Theorem 2.14 ([15]). *For any $\epsilon \in (0, 0.5]$, the distributionally robust chance constraint*

$$\inf_{\xi \sim (\boldsymbol{\mu}, \Sigma)} \mathbb{P}\{\xi^T x \leq 0\} \geq 1 - \epsilon \quad (2.21)$$

is equivalent to the convex second-order cone constraint

$$\kappa_\epsilon \sqrt{x^T \Sigma x} + \boldsymbol{\mu}^T x \leq 0, \quad \kappa_\epsilon = \sqrt{(1 - \epsilon)/\epsilon}. \quad (2.22)$$

2.3.2 Distributionally robust joint chance constraints with known mean and covariance

As a generalization of distributionally robust individual chance constrained problem, Zymler et al. [108] studied a distributionally robust problem with joint chance constraint.

Let $\boldsymbol{\mu} \in \mathbb{R}^k$ be the mean vector and $\Sigma \in \mathbb{S}^k$ be the covariance matrix of a random vector ξ . Here, \mathbb{S}^k denotes the space of symmetric matrices of dimension k . Furthermore, let \mathcal{P} denote the set of all probability distributions on \mathbb{R}^k that have the same first- and second-order moments $\boldsymbol{\mu}$ and Σ . For notational simplicity, we let

$$\Lambda = \begin{bmatrix} \Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T & \boldsymbol{\mu} \\ \boldsymbol{\mu}^T & 1 \end{bmatrix}.$$

be the second-order moment matrix of ξ .

Then, the following theorem is obtained.

Theorem 2.15 ([108]). *Let $L : \mathbb{R}^k \rightarrow \mathbb{R}$ be a continuous loss function that is either*

- (i) *concave in ξ , or*
- (ii) *(possibly nonconcave) quadratic in ξ .*

Then, the following equivalence holds:

$$\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P} - \text{CVaR}_\epsilon(L(\xi)) \leq 0 \Leftrightarrow \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}(L(\xi) \leq 0) \geq 1 - \epsilon. \quad (2.23)$$

Define the feasible set \mathcal{X}^{JCC} of the distributionally robust joint chance constraint as

$$\mathcal{X}^{JCC} = \left\{ \mathbf{x} \in \mathbb{R}^n : \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P} \left(y_i^0(\mathbf{x}) + \mathbf{y}_i^T \xi \leq 0, \forall i = 1, \dots, m \right) \geq 1 - \epsilon \right\}.$$

The joint chance constraint in \mathcal{X}^{JCC} can be reformulated as

$$\inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P} \left(\max_{i=1, \dots, m} \{ a_i (y_i^0(\mathbf{x}) + \mathbf{y}_i^T \xi) \} \leq 0 \right) \geq 1 - \epsilon$$

for any vector of strictly positive scaling parameters $\mathbf{a} \in \mathcal{A} = \{\mathbf{a} \in \mathbb{R}^m : \mathbf{a} > 0\}$. Thus, for any $\mathbf{a} \in \mathcal{A}$, the requirement

$$\mathbf{x} \in \mathcal{Z}^{JCC}(\mathbf{a}) = \left\{ \mathbf{x} \in \mathbb{R}^n : \sup_{\mathbb{P} \in \mathcal{P}} \text{CVaR}_\alpha \left(\max_{i=1, \dots, m} \{a_i (y_i^0(\mathbf{x}) + \mathbf{y}_i^T \xi)\} \right) \leq 0 \right\}$$

implies $\mathbf{x} \in \mathcal{X}^{JCC}$.

Then, it can be shown that $\mathcal{Z}^{JCC}(\mathbf{a})$ provide a convex approximation for \mathcal{X}^{JCC} with an exact tractable representation in terms of linear matrix inequalities.

Theorem 2.16 ([108]). *For any fixed $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{a} \in \mathcal{A}$, we have*

$$\mathcal{Z}^{JCC}(\mathbf{a}) = \left\{ \mathbf{x} \in \mathbb{R}^n : \begin{array}{l} \exists(\beta, M) \in \mathbb{R} \times \mathbb{S}^{k+1}, \\ \beta + \frac{1}{\epsilon} \langle \Lambda, M \rangle \leq 0, M \succeq 0, \\ M - \begin{bmatrix} \mathbf{0} & \frac{1}{2} a_i \mathbf{y}_i(\mathbf{x}) \\ \frac{1}{2} a_i \mathbf{y}_i^T(\mathbf{x}) & a_i y_i^0(\mathbf{x}) - \beta \end{bmatrix} \succeq 0, \forall i = 1, \dots, m \end{array} \right\}.$$

2.3.3 Distributionally robust joint chance constraints with uncertain mean and covariance

In practice, the models' parameters, such as mean and covariance, themselves are unknown and can only be estimated from data. For this case, distributionally robust optimization problem with uncertain mean and covariance was first proposed and studied by Delage and Ye [27], and was also investigated for distributionally robust chance constrained problem by Cheng et al. [21].

In [21], the authors' interest lies in solving a distributionally robust knapsack problem with chance constraint:

$$(DRSKP) \quad \max_{\mathbf{x}} \quad \inf_{F \in \mathcal{D}} \mathbb{E}_F[u(\mathbf{x}^T \tilde{\mathbf{R}} \mathbf{x})]$$

$$\text{s.t.} \quad \inf_{F \in \mathcal{D}} \mathbb{P}_F(\tilde{\mathbf{w}}_j^T \mathbf{x} \leq d_j \quad \forall j \in \{1, 2, \dots, M\}) \geq 1 - \epsilon, \quad (2.24)$$

$$x_i \in \{0, 1\} \quad \forall i \in \{1, 2, \dots, n\}, \quad (2.25)$$

where $u(\cdot)$ is some concave utility function that captures risk aversion with respect to the total achieved reward. \mathbf{x} is a vector of binary values indicating whether each item is included in the knapsack. $\tilde{\mathbf{R}} \in \mathbb{R}^{n \times n}$ is a random matrix whose (i, j) th term describes the linear contribution to reward of holding both items i and j . $\tilde{\mathbf{w}}_j \in \mathbb{R}^n$ is a random vector of attributes whose total amount must satisfy some capacity constraint d_j .

Distributionally robust one-dimensional knapsack problem

In this case, we have $M = 1$. To ensure that the approximation model obtained can be solved efficiently, a few assumptions are made.

Definition 2.7. *Let ξ be a random vector in \mathbb{R}^m on which $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{w}}$ depend*

linearly:

$$\tilde{\mathbf{R}} = \sum_{i=1}^n \mathbf{A}^{\tilde{\mathbf{R}}\xi_i}, \quad \tilde{\mathbf{w}} = \mathbf{A}^{\tilde{\mathbf{w}}}\xi.$$

Assumption 2.1. The utility function $u(\cdot)$ is piecewise linear, increasing and concave. In other words, it can be represented in the form

$$u(y) = \min_{k \in \{1, 2, \dots, K\}} a_k y + b_k,$$

where $\mathbf{a} \in \mathbb{R}^K$ and $\mathbf{a} \geq 0$.

Assumption 2.2. The distributional uncertainty set accounts for information about the convex support \mathfrak{S} , mean μ in the strict interior of \mathfrak{S} , and an upper bound $\Sigma \succ 0$ on the covariance matrix of the random vector ξ :

$$\mathcal{D}(\mathfrak{S}, \mu, \Sigma) = \left\{ F \left| \begin{array}{l} \mathbb{P}(\xi \in \mathfrak{S}) = 1 \\ \mathbb{E}_F[\xi] = \mu \\ \mathbb{E}_F[(\xi - \mu)(\xi - \mu)^T] \preceq \Sigma \end{array} \right. \right\}.$$

Theorem 2.17 ([21]). Under Assumptions 2.1 and 2.2, $M = 1$ and given that the support of F is ellipsoidal, $\mathfrak{S} = \{\xi | (\xi - \xi_0)^T \Theta (\xi - \xi_0) \leq 1\}$, problem (DRSKP) reduces to the following problem

$$\begin{aligned} & \max_{\mathbf{x}, t, \mathbf{q}, \mathbf{Q}, \mathbf{v}, \mathbf{s}, \bar{t}, \bar{\mathbf{q}}, \bar{\mathbf{Q}}, \bar{\mathbf{s}}} && t - \mu^T \mathbf{q} - (\Sigma + \mu \mu^T) \bullet \mathbf{Q} \\ & \text{s.t.} && \begin{bmatrix} \mathbf{Q} & \frac{\mathbf{q} + a_k \mathbf{v}}{2} \\ \frac{\mathbf{q}^T + a_k \mathbf{v}^T}{2} & b_k - t \end{bmatrix} \succeq -s_k \begin{bmatrix} \Theta & -\Theta \xi_0 \\ -\xi_0^T \Theta & \xi_0^T \Theta \xi_0 - 1 \end{bmatrix} \quad \forall k, \\ & && v_j = A_j^{\tilde{\mathbf{R}}} \bullet (\mathbf{x} \mathbf{x}^T) \quad \forall j \in \{1, 2, \dots, m\}, \\ & && \bar{t} + 2\mu^T \bar{\mathbf{q}} + (\Sigma + \mu \mu^T) \bullet \bar{\mathbf{Q}} \leq \epsilon \bar{s}_1, \\ & && \begin{bmatrix} \bar{\mathbf{Q}} & \bar{\mathbf{q}} \\ \bar{\mathbf{q}}^T & \bar{t} \end{bmatrix} \succeq -\bar{s}_2 \begin{bmatrix} \Theta & -\Theta \xi_0 \\ -\xi_0^T \Theta & \xi_0^T \Theta \xi_0 - 1 \end{bmatrix}, \\ & && \begin{bmatrix} \bar{\mathbf{Q}} & \bar{\mathbf{q}} - \mathbf{A}^{\tilde{\mathbf{w}}T} \mathbf{x} \\ (\bar{\mathbf{q}} - \mathbf{A}^{\tilde{\mathbf{w}}T} \mathbf{x})^T & \bar{t} + 2d - \bar{s}_1 \end{bmatrix} \succeq -\bar{s}_3 \begin{bmatrix} \Theta & -\Theta \xi_0 \\ -\xi_0^T \Theta & \xi_0^T \Theta \xi_0 - 1 \end{bmatrix}, \\ & && \mathbf{Q} \succeq 0, \quad \bar{\mathbf{Q}} \succeq 0, \quad \mathbf{s} \geq 0, \quad \bar{\mathbf{s}} \geq 0, \\ & && x_i \in \{0, 1\} \quad \forall i \in \{1, 2, \dots, n\}. \end{aligned}$$

Alternatively, if the support of F is polyhedral, i.e., $\mathfrak{S} = \{\xi | \mathbf{C}\xi \leq \mathbf{c}\}$ with $\mathbf{C} \in \mathbb{R}^{p \times m}$ and $\mathbf{c} \in \mathbb{R}^p$, then problem (DRSKP) reduces to

$$\begin{aligned} & \max_{\substack{\mathbf{x}, t, \mathbf{q}, \mathbf{Q}, \mathbf{v}, \mathbf{s}, \bar{t}, \bar{\mathbf{q}}, \bar{\mathbf{Q}} \\ \lambda_1, \lambda_2, \dots, \lambda_{K+2}}} && t - \mu^T \mathbf{q} - (\Sigma + \mu \mu^T) \bullet \mathbf{Q} \\ & \text{s.t.} && \begin{bmatrix} \mathbf{Q} & \frac{\mathbf{q} + a_k \mathbf{v} + \mathbf{C}^T \lambda_k}{2} \\ \frac{\mathbf{q}^T + a_k \mathbf{v}^T + \lambda_k^T \mathbf{C}}{2} & b_k - t - \mathbf{c}^T \lambda_k \end{bmatrix} \succeq 0 \quad \forall k \in \{1, 2, \dots, K\}, \end{aligned}$$

$$\begin{aligned}
v_j &= A_j^{\tilde{\mathbf{R}}} \bullet (\mathbf{x}\mathbf{x}^T) \quad \forall j \in \{1, 2, \dots, m\}, \\
\bar{t} + 2\mu^T \bar{\mathbf{q}} + (\Sigma + \mu\mu^T) \bullet \bar{\mathbf{Q}} &\leq s, \\
\begin{bmatrix} \bar{\mathbf{Q}} & \bar{\mathbf{q}} + \frac{1}{2}\mathbf{C}^T \lambda_{K+1} \\ \bar{\mathbf{q}}^T + \frac{1}{2}\lambda_{K+1}^T \mathbf{C} & \bar{t} - \mathbf{c}^T \lambda_{K+1} \end{bmatrix} &\succeq 0, \\
\begin{bmatrix} \bar{\mathbf{Q}} & \bar{\mathbf{q}} - \mathbf{A}^{\tilde{\mathbf{w}}^T} \mathbf{x} + \frac{1}{2}\mathbf{C}^T \lambda_{K+2} \\ (\bar{\mathbf{q}} - \mathbf{A}^{\tilde{\mathbf{w}}^T} \mathbf{x})^T + \frac{1}{2}\lambda_{K+2}^T \mathbf{C} & \bar{t} + 2d - s - \mathbf{c}^T \lambda_{K+2} \end{bmatrix} &\succeq 0, \\
\bar{\mathbf{Q}} \succeq 0, \bar{\mathbf{Q}} \succeq 0, s \geq 0, \lambda_k \geq 0, \forall k \in \{1, 2, \dots, K+2\}, \\
x_i \in \{0, 1\} \quad \forall i \in \{1, 2, \dots, n\},
\end{aligned}$$

where $\lambda_k \in \mathbb{R}^p$ for all k are the dual variables associated with the linear inequalities $\mathbf{C}\xi \leq \mathbf{c}$ for each infinite set of constraints. Finally, if the support of F is unbounded (i.e., $\xi = \mathbb{R}^m$), then problem (DRSKP) reduces to

$$\begin{aligned}
&\max_{\mathbf{x}, t, \mathbf{q}, \mathbf{Q}, \mathbf{v}, \mathbf{s}, \mathbf{z}, \tau} \quad t - \mu^T \mathbf{q} - (\Sigma + \mu\mu^T) \bullet \mathbf{Q} \\
&\text{s.t.} \quad \begin{bmatrix} \mathbf{Q} & \frac{\mathbf{q} + a_k \mathbf{v}}{2} \\ \frac{\mathbf{q}^T + a_k \mathbf{v}^T}{2} & b_k - t \end{bmatrix} \succeq 0 \quad \forall k = \{1, 2, \dots, K\}, \\
&\quad v_j = A_j^{\tilde{\mathbf{R}}} \bullet (\mathbf{x}\mathbf{x}^T) \quad \forall j \in \{1, 2, \dots, m\}, \\
&\quad \mathbf{Q} \succeq 0, \tau_k \geq 0 \quad \forall k = \{1, 2, \dots, K\}, \\
&\quad \begin{bmatrix} \mathbf{0} & \Sigma^{1/2} \mathbf{z} \\ \mathbf{z}^T \Sigma^{1/2} & 0 \end{bmatrix} \succeq \sqrt{\frac{\epsilon}{1-\epsilon}} (\mu^T \mathbf{z} - d) \mathbf{I}, \\
&\quad \mathbf{z} = \mathbf{A}^{\tilde{\mathbf{w}}^T} \mathbf{x}, \\
&\quad x_i \in \{0, 1\} \quad \forall i \in \{1, 2, \dots, n\}.
\end{aligned}$$

Distributionally robust multidimensional knapsack problem

In this case, we consider problem (DRSKP) with $M > 1$, where F now describes the joint distribution of all types of “weights” of all items, $\{\tilde{\mathbf{w}}_j\}_{j=1}^M \sim F$, and \mathcal{D} now describes a set of such joint distributions.

Definition 2.8. Without loss of generality, for all $j = 1, \dots, M$, let ξ_j be a random vector in \mathbb{R}^m on which the $\tilde{\mathbf{w}}_j$ depend linearly and let $\tilde{\mathbf{R}}$ depend linearly on $\{\xi_j\}_{j=1}^M$:

$$\tilde{\mathbf{R}} = \sum_{j=1}^M \sum_{i=1}^m \mathbf{A}_{ji}^{\tilde{\mathbf{R}}} (\xi_j)_i, \quad \tilde{\mathbf{w}}_j = \mathbf{A}_j^{\tilde{\mathbf{w}}_j} \xi_j, \quad j = 1, \dots, M.$$

Assumption 2.3. The distributional uncertainty set accounts for information about the mean μ_j , and an upper bound Σ_j on the covariance matrix of the random vector ξ_j , for each $j = 1, \dots, M$:

$$\mathcal{D}(\mu_j, \Sigma_j) = \left\{ F_j \left| \begin{array}{l} \mathbb{E}_{F_j} [\xi_j] = \mu_j \\ \mathbb{E}_{F_j} [(\xi_j - \mu_j)(\xi_j - \mu_j)^T] \preceq \Sigma_j \end{array} \right. \right\}.$$

Furthermore, the random vectors ξ_i and ξ_j are independent when $i \neq j$. Note that the support of F_j is unbounded, i.e., $\xi = \mathbb{R}^m$.

Theorem 2.18 ([21]). *Under Assumptions 2.2 and 2.3, the following problem is a conservative approximation of problem (DRSKP):*

$$\begin{aligned}
& \max_{\mathbf{x}, t, \mathbf{q}, \mathbf{Q}, \mathbf{v}, \mathbf{y}} && t - \mu^T \mathbf{q} - (\Sigma + \mu \mu^T) \bullet \mathbf{Q} \\
& \text{s.t.} && \begin{bmatrix} \mathbf{Q} & \frac{\mathbf{q} + a_k \mathbf{v}}{2} \\ \frac{\mathbf{q}^T + a_k \mathbf{v}^T}{2} & b_k - t \end{bmatrix} \succeq 0 \quad \forall k = \{1, 2, \dots, K\}, \\
& && v_{(j-1)M+i} = A_{ji}^{\tilde{\mathbf{R}}} \bullet (\mathbf{x} \mathbf{x}^T) \quad \forall j \in \{1, 2, \dots, M\} \forall i \in \{1, 2, \dots, m\}, \\
& && \mathbf{Q} \succeq 0, \\
& && \mu_j^T \mathbf{A}^{\tilde{\mathbf{w}}_j} \mathbf{x} + \sqrt{\frac{p^{y_j}}{1-p^{y_j}}} \|\Sigma_j^{1/2} \mathbf{A}^{\tilde{\mathbf{w}}_j T} \mathbf{x}\|_2 \leq d_2, \\
& && \sum_{j=1}^M y_j = 1, \quad y_j \geq 0, \\
& && x_i \in \{0, 1\} \quad \forall i \in \{1, 2, \dots, n\},
\end{aligned}$$

where $p = 1 - \epsilon$, $\mathbf{q}, \mathbf{v} \in \mathbb{R}^{mM}$, and $\mathbf{Q} \in \mathbb{R}^{mM \times mM}$, and with

$$\mu := \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_M \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} \Sigma_1 & \mathbf{0}_{m,m} & \mathbf{0}_{m,m} & \cdots & \mathbf{0}_{m,m} \\ \mathbf{0}_{m,m} & \Sigma_2 & \mathbf{0}_{m,m} & \cdots & \mathbf{0}_{m,m} \\ \mathbf{0}_{m,m} & \mathbf{0}_{m,m} & \Sigma_3 & \cdots & \mathbf{0}_{m,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{m,m} & \mathbf{0}_{m,m} & \mathbf{0}_{m,m} & \cdots & \Sigma_M \end{bmatrix}.$$

Furthermore, this approximation is exact when $u(\cdot)$ is linear, i.e. the attitude is risk neutral.

2.3.4 Distributionally robust chance constrained problem based on ϕ -divergence

An alternative to moments based uncertainty set of probability distributions is the distance based uncertainty set. Jiang and Guan [57] applied ϕ -divergence to construct the uncertainty set of probability distributions. Then, they considered the following distributionally robust chance constrained problem:

$$\begin{aligned}
& \min_x && \psi(x) \\
& \text{s.t.} && \inf_{\mathbb{P} \in \mathcal{D}_\phi} \mathbb{P}\{C(x, \xi)\} \geq 1 - \epsilon, \\
& && x \in \mathbf{X},
\end{aligned} \tag{2.26}$$

where $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ represents a convex function, $\mathbf{X} \subset \mathbb{R}^n$ represents a tractable bounded convex set, $\xi \in \mathbb{R}^K$ is a random vector defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $\epsilon \in (0, 1)$. \mathcal{D}_ϕ is the uncertainty set based on ϕ -divergence, which is

defined as follows:

$$\mathcal{D}_\phi = \{\mathbb{P} \in \mathcal{M}_+ : D_\phi(f||f_0) \leq d, f = d\mathbb{P}/d\xi\}.$$

Here, \mathcal{M}_+ represents the set of all probability distributions and the divergence tolerance d can be chosen by the decision makers to represent their risk-aversion level or can be obtained from statistical inference. $D_\phi(f||f_0)$ is the ϕ -divergence defined as

$$D_\phi(f||f_0) = \int_{\Omega} \phi\left(\frac{f(\xi)}{f_0(\xi)}\right) f_0(\xi) d\xi,$$

where f and f_0 denote the true density function and its estimate respectively, and $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is a convex function on \mathbb{R}^+ such that

1. $\phi(1) = 0$,
2. $0\phi(x/0) := \begin{cases} x \lim_{p \rightarrow +\infty} \phi(p)/p & \text{if } x > 0, \\ 0 & \text{if } x = 0, \end{cases}$
3. $\phi(x) = +\infty$ for $x < 0$.

Definition 2.9. Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be a convex function such that $\phi(1) = 0$ and $\phi(x) = +\infty$ for $x < 0$. Define $\underline{m}(\phi^*) := \sup\{m \in \mathbb{R} : \phi^* \text{ is a finite constant on } (-\infty, m]\}$ and $\overline{m}(\phi^*) := \inf\{m \in \mathbb{R} : \phi^*(m) = +\infty\}$.

Theorem 2.19 ([57]). Let \mathbb{P}_0 represent the probability distribution defined by f_0 . Then the

$$\inf_{\mathbb{P} \in \mathcal{D}_\phi} \mathbb{P}\{C(x, \xi)\} \geq 1 - \epsilon \quad (2.27)$$

is equivalent to

$$\mathbb{P}_0\{C(x, \xi)\} \geq 1 - \epsilon'_+, \quad (2.28)$$

where

$$\epsilon'_+ = 1 - \inf_{\substack{z > 0, z_0 + \pi_z < l_\phi, \\ \underline{m}(\phi^*) \leq z_0 + z \leq \overline{m}(\phi^*)}} \left\{ \frac{\phi^*(z_0 + z) - z_0 - \epsilon z + d}{\phi^*(z_0 + z) - \phi^*(z_0)} \right\},$$

$\epsilon_+ = \max\{\epsilon, 0\}$ for $\epsilon \in \mathbb{R}$, ϕ^* is the conjugate function of ϕ , $l_\phi = \lim_{x \rightarrow +\infty} \phi(x)/x$, and

$$\pi = \begin{cases} -\infty & \text{if } \text{Leb}\{[f_0 = 0]\} = 0, \\ 0 & \text{if } \text{Leb}\{[f_0 = 0]\} > 0 \text{ and } \text{Leb}\{[f_0 = 0] \setminus C(x, \xi)\} = 0, \\ 1 & \text{otherwise,} \end{cases}$$

$\text{Leb}\{\cdot\}$ represents the Lebesgue measure and $[f_0 = 0] := \{\xi \in \Omega : f_0(\xi) = 0\}$.

By applying different ϕ -divergences in Theorem 2.19, the following proposition can be obtained.

Proposition 2.1 ([57]). 1. Suppose that \mathcal{D}_ϕ is constructed by using the χ divergence of order 2 with $\phi(x) = (x - 1)^2$ and $\epsilon < 1/2$. Then

$$\epsilon'_+ = \epsilon - \frac{\sqrt{d^2 + 4d\epsilon(1 - \epsilon)} - (1 - 2\epsilon)d}{2d + 2},$$

2. Suppose that \mathcal{D}_ϕ is constructed by using the variation distance with $\phi(x) = |x - 1|$. Then

$$\epsilon' = \epsilon - \frac{d}{2},$$

3. Suppose that \mathcal{D}_ϕ is constructed by using the KL divergence with $\phi(x) = x \log x - x + 1$. Then

$$\epsilon' = 1 - \inf_{x \in (0,1)} \left\{ \frac{e^{-d} x^{1-\epsilon} - 1}{x - 1} \right\}.$$

2.3.5 Distributionally robust chance constrained problem based on Wasserstein distance

Besides ϕ -divergence based uncertainty set, Wasserstein distance based uncertainty set is another popular distance based uncertainty set. Wasserstein distance is a natural way of comparing two distributions when one is obtained from the other by perturbations.

The type-1 Wasserstein distance $d_W(\mathbb{P}_1, \mathbb{P}_2)$ between two distributions \mathbb{P}_1 and \mathbb{P}_2 on \mathbb{R}^K , equipped with a general norm $\|\cdot\|$, is defined as the minimal transportation cost of redistributing mass from \mathbb{P}_1 to \mathbb{P}_2 in terms that the cost of moving a Dirac point mass from $\boldsymbol{\xi}_1$ to $\boldsymbol{\xi}_2$ is $\|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\|$. Mathematically,

$$d_W(\mathbb{P}_1, \mathbb{P}_2) = \inf_{\mathbb{P} \in (\mathbb{P}_1, \mathbb{P}_2)} \mathbb{E}_{\mathbb{P}}[\|\tilde{\boldsymbol{\xi}}_1 - \tilde{\boldsymbol{\xi}}_2\|],$$

where $\tilde{\boldsymbol{\xi}}_1 \sim \mathbb{P}_1$, $\tilde{\boldsymbol{\xi}}_2 \sim \mathbb{P}_2$, and $(\mathbb{P}_1, \mathbb{P}_2)$ represents the set of all distributions on $\mathbb{R}^K \times \mathbb{R}^K$ with marginals \mathbb{P}_1 and \mathbb{P}_2 . The Wasserstein ambiguity set $\mathcal{F}(\theta)$ is then defined as a ball of radius $\theta \geq 0$ with respect to the Wasserstein distance, centered at a prescribed reference distribution $\hat{\mathbb{P}}$:

$$\mathcal{F}(\theta) = \{\mathbb{P} \in \mathcal{P}(\mathbb{R}^K) | d_W(\mathbb{P}, \hat{\mathbb{P}}) \leq \theta\}.$$

If only a finite training dataset $\{\hat{\boldsymbol{\xi}}_i\}_{i \in [N]}$ is available, a natural choice for $\hat{\mathbb{P}}$ is the empirical distribution $\hat{\mathbb{P}} = \frac{1}{N} \sum_{i=1}^N \delta_{\hat{\boldsymbol{\xi}}_i}$, which represents the uniform distribution on the training samples.

Assume that $\hat{\mathbb{P}}$ is the empirical distribution as above, Chen et al. [19] study distributionally robust chance constrained programs of the form

$$\begin{aligned} (DRCCW) \quad & \min_{\boldsymbol{x} \in \mathcal{X}} \quad \boldsymbol{c}^T \boldsymbol{x} \\ & \text{s.t.} \quad \mathbb{P}[\tilde{\boldsymbol{\xi}} \in \mathcal{S}(\boldsymbol{x})] \geq 1 - \epsilon \quad \forall \mathbb{P} \in \mathcal{F}(\theta), \end{aligned}$$

where the goal is to find a decision \boldsymbol{x} from within a compact polyhedron $\mathcal{X} \in \mathbb{R}^L$ that minimizes a linear cost function $\boldsymbol{c}^T \boldsymbol{x}$ and ensures that the exogenous random vector $\tilde{\boldsymbol{\xi}}$ falls within a decision-dependent safety set $\mathcal{S}(\boldsymbol{x}) \subset \mathbb{R}^K$ with high probability $1 - \epsilon$ under every distribution $\mathbb{P} \in \mathcal{F}(\theta)$.

Theorem 2.20 ([19]). *The chance constrained program (DRCCW) is equivalent to*

$$\begin{aligned} \min_{\mathbf{s}, t, \mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \epsilon N t - \mathbf{e}^T \mathbf{s} \geq \theta N \\ & \mathbf{dist}(\tilde{\boldsymbol{\xi}}_i, \bar{\mathcal{S}}(\mathbf{x})) \geq t - s_i \quad \forall i \in [N] \\ & \mathbf{s} \geq \mathbf{0}, \mathbf{x} \in \mathcal{X}, \end{aligned}$$

where $\mathbf{0}$ and \mathbf{e} respectively correspond to the zero vector and the vector of all ones, $[N] = \{1, 2, \dots, N\}$ represents the set of all integers up to N , $\mathbf{dist}(\tilde{\boldsymbol{\xi}}_i, \bar{\mathcal{S}}(\mathbf{x}))$ denotes the distance of $\tilde{\boldsymbol{\xi}}_i$ to the set $\bar{\mathcal{S}}(\mathbf{x})$, and $\bar{\mathcal{S}}(\mathbf{x})$ denotes the closed complement of $\mathcal{S}(\mathbf{x})$.

With Theorem 2.20, the following corollaries can be obtained.

Corollary 2.1 ([19]). *Given $\mathbf{A} \in \mathbb{R}^{L \times K}$, $\mathbf{a} \in \mathbb{R}^L$, $\mathbf{b} \in \mathbb{R}^K$, $b \in \mathbb{R}$. Assume that $\mathbf{A}^T \mathbf{x} \neq \mathbf{b}$ for all $\mathbf{x} \in \mathcal{X}$. For the safety set $\mathcal{S}(\mathbf{x}) = \{\boldsymbol{\xi} \in \mathbb{R}^K | (\mathbf{A}\boldsymbol{\xi} + \mathbf{a})^T \mathbf{x} \leq \mathbf{b}^T \boldsymbol{\xi} + b\}$, problem (DRCCW) is equivalent to the mixed integer conic program*

$$\begin{aligned} \min_{\mathbf{q}, \mathbf{s}, t, \mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \epsilon N t - \mathbf{e}^T \mathbf{s} \geq \theta N \|\mathbf{b} - \mathbf{A}^T \mathbf{x}\|_* \\ & (\mathbf{b} - \mathbf{A}^T \mathbf{x})^T \tilde{\boldsymbol{\xi}}_i + b - \mathbf{a}^T \mathbf{x} + M q_i \geq t - s_i \quad \forall i \in [N] \\ & M(1 - q_i) \geq t - s_i \quad \forall i \in [N] \\ & \mathbf{q} \in \{0, 1\}^N, \mathbf{s} \geq \mathbf{0}, \mathbf{x} \in \mathcal{X}, \end{aligned}$$

where $\|\cdot\|_*$ denotes the dual norm of a general norm $\|\cdot\|$ and M is a suitably large (but finite) positive constant.

Corollary 2.2 ([19]). *Given $\mathbf{a}_m \in \mathbb{R}^L$, $\mathbf{b}_m \in \mathbb{R}^K$, $b_m \in \mathbb{R}$, $m \in [M]$. For the safety set $\mathcal{S}(\mathbf{x}) = \{\boldsymbol{\xi} \in \mathbb{R}^K | \mathbf{a}_m^T \mathbf{x} < \mathbf{b}_m^T \boldsymbol{\xi} + b_m \quad \forall m \in [M]\}$, where $\mathbf{b} \neq \mathbf{0}$ for all $m \in [M]$, the chance constrained problem (DRCCW) is equivalent to the mixed integer conic program*

$$\begin{aligned} \min_{\mathbf{p}, \mathbf{q}, \mathbf{s}, t, \mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \epsilon N t - \mathbf{e}^T \mathbf{s} \geq \theta N \\ & p_i + M q_i \geq t - s_i \quad \forall i \in [N] \\ & M(1 - q_i) \geq t - s_i \quad \forall i \in [N] \\ & \frac{\mathbf{b}_m^T \tilde{\boldsymbol{\xi}}_i + b_m - \mathbf{a}_m^T \mathbf{x}}{\|\mathbf{b}_m\|_*} \geq p_i \quad \forall m \in [M], \forall i \in [N] \\ & \mathbf{q} \in \{0, 1\}^N, \mathbf{s} \geq \mathbf{0}, \mathbf{x} \in \mathcal{X}, \end{aligned}$$

where M is a suitably large (but finite) positive constant.

2.4 Conclusion

In this chapter, we introduce some standard results of chance constrained problems covering convexity, reformulation/approximation approaches for chance constraints and distributionally robust chance constraints. These results have important influences on development of chance constrained problems. As this kind of problems are widely applied in many practical areas, such as finance, energy problems, water resources, telecommunication, production and inventory, the theory and application of chance constrained problems are still important issues which attract a lot of attentions from different researchers.

Chapter 3

Geometric Chance Constrained Programs

In this chapter, we first review a work of Liu et al. [65] about geometric programs with joint chance constraints, where the stochastic parameters are normally distributed and independent of each other. Then, we extend the work in [65] and discuss joint rectangular geometric chance constrained programs with elliptical distribution.

3.1 Stochastic geometric optimization with joint chance constraints

In this section, we review the work of Liu et al. [65] about geometric programs with joint chance constraints. When the stochastic parameters are normally distributed and independent of each other, the problem is approximated by using piecewise linear functions, and the approximation problem is transformed into a convex geometric program. The authors proved that this approximation method provides a lower bound and designed an algorithm under the sequential convex optimization scheme to find an upper bound. Finally, numerical tests are carried out with a stochastic shape optimization problem.

3.1.1 Introduction

Geometric programming is an important topic in operations research where the objective function and the constraints of the corresponding optimization problems have a special form. Geometric optimization has been studied for several decades, it was introduced by Duffin et al. in the late 1960s [30]. Applications of geometric programming can be found in several surveys papers, namely Boyds et al. [10], Ecker [33] and Peterson [83]. A numerous practical problems can be formulated as geometric programs, e.g., electrical circuit design problems [10], information theory [26], queue proportional scheduling in fading broadcast channels [95], mechanical engineering problems [101], economic and managerial problems [69], nonlinear network problems [60]. A geometric program can be formulated

as

$$(GP) \quad \min_t g_0(t) \text{ s.t. } g_k(t) \leq 1, \quad k = 1, \dots, K, \quad t \in \mathbb{R}_{++}^M \quad (3.1)$$

with

$$g_k(t) = \sum_{i \in I_k} c_i \prod_{j=1}^M t_j^{a_{ij}}, \quad k = 0, \dots, K. \quad (3.2)$$

Usually, $c_i \prod_{j=1}^M t_j^{a_{ij}}$ is called a monomial, c_i needs to be non-negative and $g_k(t)$ is called a posynomial. We denote by Q the number of monomials in (3.1), and $\{I_k, k = 0, \dots, K\}$ is the disjoint index sets of $\{1, \dots, Q\}$.

Geometric programs are not convex with respect to t whilst they are convex with respect to $\{z : z_j = \log t_j, j = 1, \dots, M\}$. Hence, interior point method can be efficiently used to solve geometric programs.

In real world applications, some of the coefficients in (3.1) may not be known precisely. Hence, the stochastic geometric programming is proposed to model geometric problem with random parameters. For instance, in [31, 92], individual probabilistic constraints have been used to control the uncertainty level of the constraints in (3.1):

$$\mathbb{P} \left(\sum_{i \in I_k} c_i \prod_{j=1}^M t_j^{a_{ij}} \leq 1 \right) \geq 1 - \epsilon_k, \quad k = 1, \dots, K, \quad (3.3)$$

where ϵ_k is the tolerance probability for the k -th constraint in (3.2).

In this section, we furthermore consider the following joint probabilistic constrained stochastic geometric programs

$$(SGP) \quad \min_{t \in \mathbb{R}_{++}^M} \mathbb{E} \left[\sum_{i \in I_0} c_i \prod_{j=1}^M t_j^{a_{ij}} \right] \quad (3.4)$$

$$\text{s.t.} \quad \mathbb{P} \left(\sum_{i \in I_k} c_i \prod_{j=1}^M t_j^{a_{ij}} \leq 1, \quad k = 1, \dots, K \right) \geq 1 - \epsilon. \quad (3.5)$$

Unlike [31, 92], we require that the overall probability of meeting the K geometric constraints is above a certain probability level $1 - \epsilon$, where $\epsilon \in (0, 0.5]$.

3.1.2 Stochastic geometric optimization under Gaussian distribution

We suppose that the coefficients $a_{ij}, i \in I_k, \forall k, j = 1, \dots, M$, are deterministic and the parameters $c_i, i \in I_k, \forall k$ are normally distributed and independent of each other, i.e., $c_i \sim N(E_{c_i}, \sigma_i^2)$ [31]. Moreover, we assume that $E_{c_i} \geq 0$. As c_i

are independent of each other, constraint (3.5) is equivalent to

$$\prod_{k=1}^K \mathbb{P} \left(\sum_{i \in I_k} c_i \prod_{j=1}^M t_j^{a_{ij}} \leq 1 \right) \geq 1 - \epsilon. \quad (3.6)$$

By introducing auxiliary variables $y_k \in \mathbb{R}_+$, $k = 1, \dots, K$, (3.6) can be equivalently transformed into

$$\mathbb{P}(\sum_{i \in I_k} c_i \prod_{j=1}^M t_j^{a_{ij}} \leq 1) \geq y_k, \quad k = 1, \dots, K, \quad (3.7)$$

and

$$\prod_{k=1}^K y_k \geq 1 - \epsilon, \quad 1 \geq y_k \geq 0, \quad k = 1, \dots, K. \quad (3.8)$$

It is easy to see that for independent normally distributed $c_i \sim N(E_{c_i}, \sigma_i^2)$ [31], constraint (3.7) is equivalent to

$$\sum_{i \in I_k} \mathbb{E}_{c_i} \prod_{j=1}^M t_j^{a_{ij}} + \Phi^{-1}(y_k) \sqrt{\sum_{i \in I_k} \sigma_i^2 \prod_{j=1}^M t_j^{2a_{ij}}} \leq 1, \quad k = 1, \dots, K. \quad (3.9)$$

Here, $\Phi^{-1}(y_k)$ is the quantile of the standard normal distribution $N(0, 1)$. However, biconvex inequalities (3.9) are still very hard to solve within an optimization problem [41].

Standard variable transformation

The standard variable transformation $r_j = \log(t_j)$, $j = 1, \dots, M$ and $x_k = \log(y_k)$, $k = 1, \dots, K$ applied to (3.8) and (3.9) leads to the following constraints:

$$\sum_{i \in I_k} \mathbb{E}_{c_i} \exp \left\{ \sum_{j=1}^M a_{ij} r_j \right\} + \sqrt{\sum_{i \in I_k} \sigma_i^2 \exp \left\{ \sum_{j=1}^M (2a_{ij} r_j + \log(\Phi^{-1}(e^{x_k})^2)) \right\}} \leq 1, \quad k = 1, \dots, K, \quad (3.10)$$

$$\sum_{k=1}^K x_k \geq \log(1 - \epsilon), \quad x_k \leq 0, \quad k = 1, \dots, K. \quad (3.11)$$

$\Phi^{-1}(\cdot)$ is also called the probit function and can be expressed in terms of the inverse error function:

$$\Phi^{-1}(y_k) = \sqrt{2} \operatorname{erf}^{-1}(2y_k - 1), \quad y_k \in (0, 1).$$

The inverse error function is a nonelementary function which can be represented by the Maclaurin series:

$$\operatorname{erf}^{-1}(z) = \sum_{p=0}^{\infty} \frac{\lambda_p}{2p+1} \left(\frac{\sqrt{\pi}}{2} z \right)^{2p+1},$$

where $\lambda_0 = 1$ and $\lambda_p = \sum_{i=0}^{p-1} \frac{\lambda_i \lambda_{p-1-i}}{(i+1)(2i+1)} > 0$, $p = 1, 2, \dots$. Thus, we know that $\Phi^{-1}(y_k)$ is convex for $1 > y_k \geq 0.5$, and concave for $0 > y_k \leq 0.5$. Moreover, $\Phi^{-1}(y_k)$ is always monotonic increasing.

Under constraint (3.11), we have $0.5 \leq 1 - \epsilon \leq y_k = e^{x_k} < 1$. Hence, we can only focus on the right tail part of $\Phi^{-1}(e^{x_k})$. This means the feasible set constrained by both (3.10) and (3.11) is convex. However, as $\Phi^{-1}(\cdot)$ is nonelementary, we still need to approximate it for practical use. Unlike the approximation method in [21], we approximate $\log(\Phi^{-1}(e^{x_k})^2)$ rather than $\Phi^{-1}(y_k)$ by a piecewise linear function.

Approximation of $\log(\Phi^{-1}(e^{x_k})^2)$

We choose S different linear functions:

$$F_s(x_k) = d_s x_k + b_s, \quad s = 1, \dots, S,$$

such that

$$F_s(x_k) \leq \log(\Phi^{-1}(e^{x_k})^2), \quad \forall x_k \in [\log(1 - \epsilon), 0], \quad s = 1, \dots, S. \quad (3.12)$$

$\log(\Phi^{-1}(e^{x_k})^2)$ is then approximated by a piecewise linear function

$$F(x_k) = \max_{s=1, \dots, S} F_s(x_k). \quad (3.13)$$

Constraints (3.12) and (3.13) guarantee that $F(x_k)$ provides a lower bound of $\log(\Phi^{-1}(e^{x_k})^2)$.

For a practical use, we can choose the tangent lines of $\log(\Phi^{-1}(e^{x_k})^2)$ at different points in $[\log(1 - \epsilon), 0]$, say $\xi_1, \xi_2, \dots, \xi_S$. Then, we have

$$d_s = \frac{2e^{\xi_s} (\Phi^{-1})^{(1)}(e^{\xi_s})}{\Phi^{-1}(e^{\xi_s})} \quad (3.14)$$

and

$$b_s = -d_s \xi_s + \log(\Phi^{-1}(e^{\xi_s})^2), \quad s = 1, \dots, S. \quad (3.15)$$

Remark 3.1. *By changing the variables into y_k , we can see that the piecewise linear approximation of $\log(\Phi^{-1}(e^{x_k})^2)$ with respect to x_k is equivalent to the piecewise power function approximation $\max_{s=1, \dots, S} e^{\frac{b_s}{2}} y_k^{\frac{d_s}{2}}$ for $\Phi^{-1}(y_k)$.*

Convex geometric approximation

After obtaining the piecewise linear approximation of $\log(\Phi^{-1}(e^{x_k})^2)$, we can replace $\log(\Phi^{-1}(e^{x_k})^2)$ by $F(x_k)$ in (3.13).

Theorem 3.1. *Using the piecewise linear function $F(x_k)$, we have the following convex approximation of the stochastic geometric program with a joint probabilistic*

constraint (SGP):

$$(SGP_A) \quad \min_{r \in \mathbb{R}^M, x \in \mathbb{R}^K} \sum_{i \in I_0} \mathbb{E}_{c_i} \exp \left\{ \sum_{j=1}^M a_{ij} r_j \right\} \quad (3.16)$$

$$\sum_{i \in I_k} \mathbb{E}_{c_i} \exp \left\{ \sum_{j=1}^M a_{ij} r_j \right\} + \sqrt{\sum_{i \in I_k} \sigma_i^2 \exp \left\{ \sum_{j=1}^M (2a_{ij} r_j + d_s x_k + b_s) \right\}} \leq 1, \quad s = 1, \dots, S, \quad k = 1, \dots, K, \quad (3.17)$$

$$\sum_{k=1}^K x_k \geq \log(1 - \epsilon), \quad x_k \leq 0, \quad k = 1, \dots, K. \quad (3.18)$$

The optimal value of the approximation problem (3.16)-(3.18) is a lower bound of problem (GP). Moreover, when S goes to infinity, (SGP_A) is a reformulation of problem (SGP).

We call this approximation a piecewise linear approximation. As problem (SGP_A) is convex, it can be solved efficiently by interior point methods.

Proof. From (3.7) and (3.8), we know that $0 \leq y_k \leq 1$. Moreover, we have from (3.8) that for any k , $y_k \geq \prod_{k=1}^K y_k \geq 1 - \epsilon \geq 0.5$ and hence $x_k \geq \log(1 - \epsilon) \geq \log(0.5)$. Then, from the convexity of $\log(\Phi^{-1}(e^{x_k})^2)$ with respect to x_k in $[\log(1 - \epsilon), 0)$, we have

$$\log(\Phi^{-1}(e^{x_k})^2) \geq d_s x_k + b_s, \quad \forall x_k \in [\log(1 - \epsilon), 0), \quad s = 1, \dots, S. \quad (3.19)$$

where d_s and b_s are defined in (3.14) and (3.15). As $y_k = e^{x_k}$, (3.19) is equivalent to

$$\Phi^{-1}(y_k) \geq e^{\frac{b_s}{2}} y_k^{\frac{d_s}{2}}, \quad \forall y_k \in [1 - \epsilon, 1), \quad s = 1, \dots, S. \quad (3.20)$$

Furthermore, constraint (3.17) is equivalent to

$$\sum_{i \in I_k} \mathbb{E}_{c_i} \prod_{j=1}^M t_j^{a_{ij}} + \left(\max_{s=1, \dots, S} e^{\frac{b_s}{2}} y_k^{\frac{d_s}{2}} \right) \sqrt{\sum_{i \in I_k} \sigma_i^2 \prod_{j=1}^M t_j^{2a_{ij}}} \leq 1, \quad k = 1, \dots, K. \quad (3.21)$$

As $e^{\frac{b_s}{2}} y_k^{\frac{d_s}{2}} \leq \Phi^{-1}(y_k)$, $\forall s$, any feasible solution for (3.9) is feasible for (3.17) or (3.21). From the equivalence between (3.5) and (3.9) under the Gaussian distribution assumption, the optimal solution of (SGP) is feasible for (SGP_A). This means that the approximation problem (SGP_A) provides a lower bound for the original problem (SGP).

Moreover, the S tangent functions in (3.19) are chosen differently and for $x_k \in \{\xi_1, \dots, \xi_S\}$, $\log(\Phi^{-1}(e^{x_k})^2) = d_s x_k + b_s$. Hence, when S goes to infinity, inequality (3.19) becomes tight. Furthermore, from the biconvexity of (3.9) and (3.21), we know the distance between the sets constrained by (3.9) and (3.21) is small enough when S goes to infinity. As the objective functions of (SGP) and

(SGP_A) are the same, i.e., (SGP_A) is a reformulation of (SGP) when S goes to infinity. \square

Sequential convex approximation

In order to come up with an upper bound of the joint probabilistic constrained problem (SGP) , which is equivalent to minimize (3.4) subject to the constraints (3.8) and (3.9), we use the popular sequential convex approximation. The basic idea of the sequential convex approximation consists in decomposing the original problem into subproblems where a subset of variables is fixed alternatively. For our problem, we first fix $y = y^n$ and update t by solving

$$(SQ_1) \quad \min_{t \in \mathbb{R}_{++}^M} \sum_{i \in I_0} \mathbb{E}_{c_i} \prod_{j=1}^M t_j^{a_{ij}} \quad (3.22)$$

$$\text{s.t.} \quad \sum_{i \in I_k} \mathbb{E}_{c_i} \prod_{j=1}^M t_j^{a_{ij}} + \Phi^{-1}(y_k^n) \sqrt{\sum_{i \in I_k} \sigma_i^2 \prod_{j=1}^M t_j^{2a_{ij}}} \leq 1, k = 1, \dots, K \quad (3.23)$$

and then fix $t = t^n$ and update y by solving

$$(SQ_2) \quad \min_{y \in \mathbb{R}_+^K} \sum_{k=1}^K \phi_k y_k \quad (3.24)$$

$$\text{s.t.} \quad y_k \leq \Phi \left(\frac{1 - \sum_{i \in I_k} \mathbb{E}_{c_i} \prod_{j=1}^M (t_j^n)^{a_{ij}}}{\sqrt{\sum_{i \in I_k} \sigma_i^2 \prod_{j=1}^M (t_j^n)^{2a_{ij}}}} \right), k = 1, \dots, K. \quad (3.25)$$

$$\prod_{k=1}^K y_k \geq 1 - \epsilon, y_k \geq 0, k = 1, \dots, K. \quad (3.26)$$

Here, ϕ_k is a chosen searching direction. The sequential approximation is given by the following algorithm:

Algorithm 1 Sequential convex approximation

Initialization:

Choose an initial point y^0 of y feasible for (3.8). Set $n = 0$.

Iteration:

while $n \geq 1$ and $\|y^{n-1} - y^n\|$ is small enough **do**

- Solve problem (SQ_1) ; let t^n , θ^n and v^n denote an optimal solution of t , an optimal solution of the Lagrangian dual variable θ and the optimal value, respectively.

- Solve problem (SQ_2) with $\phi_k = \theta_k^n \cdot (\Phi^{-1})'(y_k^n) \sqrt{\sum_{i \in I_k} \sigma_i^2 \prod_{j=1}^M (t_j^n)^{2a_{ij}}}$; let \tilde{y} denote an optimal solution.

- $y^{n+1} \leftarrow y^n + \tau(\tilde{y} - y^n)$, $n \leftarrow n + 1$. Here, $\tau \in (0, 1)$ is the step length.

end while

Output: t^n, v^n

Theorem 3.2. *Algorithm 1 converges in a finite number of iterations and the returned value v^n is an upper bound for the optimal value of problem (SGP).*

Proof. For given feasible y^n , problem (SQ₁) is bounded from below because of the positivity of t and E_c . Moreover, it has at least a feasible solution $t_j = \min \left\{ \left(\frac{1}{2Q \max_i \{E_{c_i}\}} \right)^{\frac{1}{M \max_i \{a_{i,j}\}}}, \left(\frac{1}{2\Phi^{-1}(y^n) \sqrt{Q \max_i \{\sigma_i^2\}}} \right)^{\frac{1}{M \max_i \{a_{i,j}\}}} \right\}$. Hence, the optimal value sequence $\{v^n\}$ is bounded.

Both t^{n-1} and t^n are feasible for problem (SQ₁) with $y = y^n$. Hence, we know $\{v^n\}$ is non-increasing. Together with the boundness, this means the sequential convex approximation algorithm converges.

Moreover, at each iteration, y^n and t^n give a feasible solution for problem (3.4) subject to constraints (3.8) and (3.9). As (3.5) is equivalent to (3.8) and (3.9), the optimal values of the approximation problems with fixed y^n is an upper bound for the original problem (SGP). \square

Problems (SQ₁) and (SQ₂) are both geometric programs, hence they can be transformed into a convex programming problem, and solved by interior point methods.

3.1.3 Numerical experiments

We test the performance of our approximations by considering a joint probabilistic constrained stochastic shape optimization problem.

The shape optimization problem is widely applied in the design and construction of structural mechanics and in the optimal control of distributed parameter systems [97]. The shape optimization problem consists in finding a geometry of the structure which minimizes a cost functional w.r.t. given constraints [10]. In this subsection, we consider a shape optimization problem in which we maximize the volume of a box-shaped structure with height h , width w and depth ζ . We have a joint probabilistic constraint on the total wall area $2(hw + h\zeta)$, and the floor area $w\zeta$. Here the limits on the total wall area and the floor area are considered as random variables. Moreover, there are some lower and upper bounds on the aspect ratios h/w and w/ζ . This example is a generalization of the shape optimization problem with random parameters [10]. It can be formulated in the standard form of a geometric program as follows:

$$(SCP) \quad \min_{h,w,\zeta} \quad h^{-1}w^{-1}\zeta^{-1} \quad (3.27)$$

$$\text{s.t.} \quad \mathbb{P}((2/A_{wall})hw + (2/A_{wall})h\zeta \leq 1, (1/A_{flr})w\zeta \leq 1) \geq 1 - \epsilon, \quad (3.28)$$

$$\alpha h^{-1}w \leq 1, (1/\beta)hw^{-1} \leq 1, \quad (3.29)$$

$$\gamma w\zeta^{-1} \leq 1, (1/\delta)w^{-1}\zeta \leq 1. \quad (3.30)$$

We set $\alpha = \gamma = 0.5$, $\beta = \delta = 2$, $\epsilon = 5\%$, and assume $1/A_{wall} \sim N(0.005, 0.01)$ and $1/A_{flr} \sim N(0.01, 0.01)$.

We use Algorithm 1 to compute an upper bound for problem (SCP), the numerical results are given in Table 1. We solve five piecewise linear approximation problems to compute five lower bounds for problem (SCP).

The first column in Table 1 gives the number of segments S used in the piecewise linear approximation for solving (SGP_A). The second and third columns give the number of variables and the number of constraints of the five problems, respectively. Notice that the number of variables is composed by the number of decision variables and the slack variables. Similarly, the number of constraints concerns both the original constraints and the added constraints. The fourth and the fifth columns give the lower bounds and the CPU time for the five approximation problems, respectively. Algorithm 1 converges within 7 outer iterations, the corresponding upper bound and the CPU time are given by columns 6 and 7 respectively. We use Sedumi solver from CVX package [42] to solve the approximation problems with Matlab R2012b, on a PC with a 2.6 Ghz Intel Core i7-5600U CPU and 12.0 GB RAM. For better illustration, we compute gaps of the piecewise linear approximation bounds, which are the percentage differences between these lower bounds and the upper bound, and show them in the last column.

Table 3.1: Computational results of piecewise linear approximations and Algorithm 1.

S	Var. Num.	Con. Num.	Lower bound	CPU(s)	Upper bound	CPU(s)	Gap(%)
1	133	60	0.232	0.5955	0.256	5.5274	9.655
2	184	91	0.234	0.6272	0.256	5.5274	8.789
5	283	153	0.241	0.9480	0.256	5.5274	6.044
10	513	273	0.252	1.3554	0.256	5.5274	1.713
20	973	513	0.256	1.9986	0.256	5.5274	0

From Table 3.1, we can see that as the number of segments S increases, the gap of the corresponding piecewise linear approximation becomes smaller. When the number of segments is equal to 20, the gap is tight.

Although the problem size is increasing with the number of segments, the CPU time does not increase proportionally. This means our approximation approaches together with an increasing number of segments keeps a good structure of the problem to be solved.

3.2 Joint rectangular geometric chance constrained programs

In this section, we extend the results in section 3.1 to joint rectangular geometric chance constrained programs. When the stochastic parameters are elliptically distributed and pairwise independent, we present a reformulation of the joint rectangular geometric chance constrained programs. As the reformulation is not convex, we propose new convex approximations based on the variable transformation together with piecewise linear approximation method. Numerical results show that our approximations are tight.

3.2.1 Introduction

A rectangular geometric program can be formulated as

$$(GP) \quad \min_{t \in \mathbb{R}_{++}^M} g_0(t) \\ \text{s.t.} \quad \alpha_k \leq g_k(t) \leq \beta_k, \quad k = 1, \dots, K,$$

where $\alpha_k, \beta_k \in \mathbb{R}$, $k = 1, \dots, K$ and

$$g_k(t) = \sum_{i=1}^{I_k} c_i^k \prod_{j=1}^M t_j^{a_{ij}^k}, \quad k = 0, \dots, K. \quad (3.31)$$

Usually, $c_i^k \prod_{j=1}^M t_j^{a_{ij}^k}$ is called a monomial where c_i^k , $i = 1, \dots, I_k$, $k = 0, \dots, K$, are nonnegative and $g_k(t)$, $k = 0, \dots, K$, are called posynomials.

We require that $0 < \alpha_k < \beta_k$, $k = 1, \dots, K$. When $\alpha_k \leq 0$, $k = 1, \dots, K$, the rectangular geometric program is equivalent to the geometric program discussed in [30, 83, 10].

Both geometric programs and rectangular geometric programs are not convex with respect to $[t_1, t_2, \dots, t_M]^\top$. However, geometric programs are convex with respect to $[\log(t_1), \log(t_2), \dots, \log(t_M)]^\top$. Hence, interior point methods can be used to efficiently solve geometric programs. To the best of our knowledge, there is no variable transformation method to derive a convex equivalent reformulation for a rectangular geometric program.

Stochastic geometric programming is used to model geometric problems when some of the parameters are not known precisely. Stochastic geometric programs with individual chance constraints are discussed in [31] and [92] where the authors showed that an individual chance constraint is equivalent to several deterministic constraints involving posynomials and common additional slack variables. In their work, the parameters $a_{ij}^k, \forall k, i, j$, are deterministic and $c_i^k, \forall k, i$, are uncorrelated normally distributed random variables. When $a_{ij}^k \in \{0, 1\}, \forall k, i, j$ and $\sum_j a_{ij}^k = 1, \forall k, i$, stochastic geometric programs are equivalent to stochastic linear programs.

In the rest of this chapter, we consider the following joint rectangular geometric chance constrained programming problem

$$(SGP) \quad \min_{t \in \mathbb{R}_{++}^M} \mathbb{E} \left[\sum_{i=1}^{I_0} c_i^0 \prod_{j=1}^M t_j^{a_{ij}^0} \right] \quad (3.32) \\ \text{s.t.} \quad \mathbb{P} \left(\alpha_k \leq \sum_{i=1}^{I_k} c_i^k \prod_{j=1}^M t_j^{a_{ij}^k} \leq \beta_k, \quad k = 1, \dots, K \right) \geq 1 - \epsilon. \quad (3.33)$$

Here $1 - \epsilon$ is a prespecified probability with $\epsilon < 0.5$, $a_{ij}^k, k = 0, 1, \dots, K$, $i = 1, \dots, I_k$, $j = 1, \dots, M$, are given parameters and $c_i^k, k = 0, 1, \dots, K$, $i = 1, \dots, I_k$, are random parameters with non-negative mean values.

3.2.2 Elliptically distributed stochastic geometric problems

In this section, we consider the joint rectangular geometric chance constrained programs under the elliptical distribution assumption.

Assumption 3.1. $c^k = [c_1^k, c_2^k, \dots, c_{I_k}^k]$ follows a multivariate elliptical distribution $Ellip_{I_k}(\mu^k, \Gamma^k, \varphi_k)$ with a mean vector (location parameter) $\mu^k = [\mu_1^k, \mu_2^k, \dots, \mu_{I_k}^k]^\top \geq 0$, and a scale matrix Γ^k . The element in the i th row and the p th column of Γ^k is $\sigma_{i,p}^k$, $i, p = 1, \dots, I_k$, $k = 0, 1, \dots, K$. We require that $\mu_i^k > 0$, $i = 1, \dots, I_k$, and $\sigma_{i,p}^k \geq 0$, $i, p = 1, \dots, I_k$, $k = 0, 1, \dots, K$. Moreover, we assume that c^k , $k = 1, \dots, K$ are pairwise independent.

Definition 3.1. A L -dimensional random vector ξ follows an elliptical distribution $Ellip_L(\mu, \Gamma, \varphi)$ if its characteristic function is given by $\mathbb{E}e^{iz^\top \xi} = e^{iz^\top \mu} \varphi(z^\top \Gamma z)$ where φ is the characteristic generator function, μ is the location parameter, and Γ is the scale matrix.

Elliptical distributions include normal distribution with $\varphi(t) = \exp\{-\frac{1}{2}t\}$, student's t distribution with $\varphi(t)$ varying with its degree of freedom [61], Cauchy distribution with $\varphi(t) = \exp\{-\sqrt{|t|}\}$, Laplace distribution with $\varphi(t) = (1 + \frac{1}{2}t)^{-1}$, and logistic distribution with $\varphi(t) = \frac{2\pi\sqrt{t}}{e^{\pi\sqrt{t}} - e^{-\pi\sqrt{t}}}$. The mean value of an elliptical distribution $Ellip_L(\mu, \Gamma, \varphi)$ is μ , and its covariance matrix is $\frac{E(r^2)}{\text{rank}(\Gamma)}\Gamma$, where r is the random radius [39].

Proposition 3.1 (Embrechts et al., 2005). *If a L -dimensional random vector ξ follows an elliptical distribution $Ellip_L(\mu, \Gamma, \varphi)$, then for any $(N \times L)$ -matrix A and any N -vector b , $A\xi + b$ follows an N -dimensional elliptical distribution $Ellip_N(A\mu + b, A\Gamma A^\top, \varphi)$.*

Moreover, we have some restrictions on ϵ as follows.

Assumption 3.2. *We assume that*

- * $\frac{\phi'_{\varphi_k}(\Phi_{\varphi_k}^{-1}(1-\epsilon))}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(1-\epsilon))} \Phi_{\varphi_k}^{-1}(1-\epsilon) < -1$, $k = 1, \dots, K$,
- * $(\Phi_{\varphi_k}^{-1}(1-\epsilon))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k > 0$, $i, p = 1, \dots, I_k$, $k = 1, \dots, K$,
- * $2\sigma_{i,p}^k \left(1 - \frac{\phi'_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z))}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z))} \Phi_{\varphi_k}^{-1}(z)\right) \left((\Phi_{\varphi_k}^{-1}(z))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k\right) - (2\sigma_{i,p}^k \Phi_{\varphi_k}^{-1}(z))^2 \geq 0$,
for $z \in [1-\epsilon, 1]$, $i, p = 1, \dots, I_k$, $k = 1, \dots, K$.

Here, $\Phi_\varphi(\cdot)$ is the distribution function of an univariate standard elliptical distribution $Ellip_1(0, 1, \varphi)$. We assume that the elliptical distribution $Ellip_1(0, 1, \varphi)$ associated with φ has a density function $\phi_\varphi(\cdot)$, and $\Phi_\varphi(\cdot)$ is continuous and increasing. We denote that, $\Phi_\varphi^{-1}(\cdot)$ is the inverse function of $\Phi_\varphi(\cdot)$, i.e., the quantile of the standard elliptical distribution. We assume $\Phi_\varphi^{-1}(\cdot)$ to be continuous and $\phi_\varphi(\cdot)$ is positive a.s., and denote the first order derivative of $\phi_\varphi(\cdot)$ by $\phi'_\varphi(\cdot)$.

For some distributions in elliptical distribution group, such as a normal distribution, Assumption 3.2 has some simpler equivalent formulation.

Proposition 3.2. For an univariate standard normal distribution $N(0, 1)$, $\epsilon \leq 1 - \Phi_{\varphi_k}(1)$ implies $\frac{\phi'_{\varphi_k}(\Phi_{\varphi_k}^{-1}(1-\epsilon))}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(1-\epsilon))} \Phi_{\varphi_k}^{-1}(1-\epsilon) < -1$.

Proof. We denote $\Phi_{\varphi_k}^{-1}(1-\epsilon)$ by ς for short. The left hand side of the conclusion can be rewritten as $\varsigma \frac{\phi'_{\varphi_k}(\varsigma)}{\phi_{\varphi_k}(\varsigma)}$. We have

$$\varsigma \frac{\phi'_{\varphi_k}(\varsigma)}{\phi_{\varphi_k}(\varsigma)} = \varsigma \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{\varsigma^2}{2}}(-\varsigma)}{\frac{1}{\sqrt{2\pi}} e^{-\frac{\varsigma^2}{2}}} = -\varsigma^2.$$

Moreover, $\epsilon \leq 1 - \Phi_{\varphi_k}(1)$ implies $\varsigma \geq 1$, which furthermore implies $\varsigma \frac{\phi'_{\varphi_k}(\varsigma)}{\phi_{\varphi_k}(\varsigma)} < -1$. \square

Theorem 3.3. Given Assumption 3.1, the joint rectangular geometric chance constrained programs (SGP) can be equivalently reformulated as

(SGP_r)

$$\min_{t \in \mathbb{R}_{++}^M} \sum_{i=1}^{I_0} \mu_i^0 \prod_{j=1}^M t_j^{a_{ij}^0} \quad (3.34)$$

$$\text{s.t.} \quad \Phi_{\varphi_k}^{-1}(z_k^+) \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma_{i,p}^k \prod_{j=1}^M t_j^{a_{ij}^k + a_{pj}^k}} - \sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^M t_j^{a_{ij}^k} \leq -\alpha_k, \quad k = 1, \dots, K, \quad (3.35)$$

$$\Phi_{\varphi_k}^{-1}(z_k^-) \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma_{i,p}^k \prod_{j=1}^M t_j^{a_{ij}^k + a_{pj}^k}} + \sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^M t_j^{a_{ij}^k} \leq \beta_k, \quad k = 1, \dots, K, \quad (3.36)$$

$$z_k^+ + z_k^- - 1 \geq y_k, \quad 0 \leq z_k^+, z_k^- \leq 1, \quad k = 1, \dots, K, \quad (3.37)$$

$$\prod_{k=1}^K y_k \geq 1 - \epsilon, \quad 0 \leq y_k \leq 1, \quad k = 1, \dots, K. \quad (3.38)$$

Proof. As c_i^k are pairwise independent, constraint (3.33) is equivalent to

$$\prod_{k=1}^K \mathbb{P} \left(\alpha_k \leq \sum_{i=1}^{I_k} c_i^k \prod_{j=1}^M t_j^{a_{ij}^k} \leq \beta_k \right) \geq 1 - \epsilon. \quad (3.39)$$

By introducing auxiliary variables $y_k \in \mathbb{R}_+$, $k = 1, \dots, K$, (3.39) can be equivalently written as

$$\mathbb{P} \left(\alpha_k \leq \sum_{i=1}^{I_k} c_i^k \prod_{j=1}^M t_j^{a_{ij}^k} \leq \beta_k \right) \geq y_k, \quad k = 1, \dots, K, \quad (3.40)$$

and

$$\prod_{k=1}^K y_k \geq 1 - \epsilon, \quad 0 \leq y_k \leq 1, \quad k = 1, \dots, K. \quad (3.41)$$

(3.40) is also equivalent to

$$\mathbb{P} \left(\sum_{i=1}^{I_k} c_i^k \prod_{j=1}^M t_j^{a_{ij}^k} \geq \alpha_k \right) + \mathbb{P} \left(\sum_{i=1}^{I_k} c_i^k \prod_{j=1}^M t_j^{a_{ij}^k} \leq \beta_k \right) - 1 \geq y_k, \quad k = 1, \dots, K. \quad (3.42)$$

Let $z_k^+, z_k^- \in \mathbb{R}_+$, $k = 1, \dots, K$, be two additional auxiliary variables. Constraint (3.42) can be equivalently expressed by

$$\mathbb{P} \left(\sum_{i=1}^{I_k} c_i^k \prod_{j=1}^M t_j^{a_{ij}^k} \geq \alpha_k \right) \geq z_k^+, \quad k = 1, \dots, K, \quad (3.43)$$

$$\mathbb{P} \left(\sum_{i=1}^{I_k} c_i^k \prod_{j=1}^M t_j^{a_{ij}^k} \leq \beta_k \right) \geq z_k^-, \quad k = 1, \dots, K, \quad (3.44)$$

$$z_k^+ + z_k^- - 1 \geq y_k, \quad 0 \leq z_k^+, z_k^- \leq 1, \quad k = 1, \dots, K. \quad (3.45)$$

From Proposition 3.1, we know that $\sum_{i=1}^{I_k} c_i^k \prod_{j=1}^M t_j^{a_{ij}^k}$ follows an elliptical distribution $Ellip_1(\sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^M t_j^{a_{ij}^k}, \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma_{i,p}^k \prod_{j=1}^M t_j^{a_{ij}^k + a_{pj}^k}, \varphi_k)$. By using the quantile function $\Phi_{\varphi_k}^{-1}(z_k^+)$, we can equivalently rewrite constraint (3.43) as

$$-\sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^M t_j^{a_{ij}^k} + \Phi_{\varphi_k}^{-1}(z_k^+) \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma_{i,p}^k \prod_{j=1}^M t_j^{a_{ij}^k + a_{pj}^k}} \leq -\alpha_k, \quad k = 1, \dots, K,$$

and rewrite constraint (3.44) as

$$\sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^M t_j^{a_{ij}^k} + \Phi_{\varphi_k}^{-1}(z_k^-) \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma_{i,p}^k \prod_{j=1}^M t_j^{a_{ij}^k + a_{pj}^k}} \leq \beta_k, \quad k = 1, \dots, K.$$

This gives the equivalent reformulation of the joint chance constrained problem. As $c^k \sim Ellip_{I_k}(\mu^k, \Gamma^k, \varphi_k)$, its expected value is μ^k . Hence, from the additivity property of the expectation operator, we can get the equivalent reformulation of the objective function. \square

In (SGP_r) , both constraints (3.35) and (3.36) are nonconvex constraints. In the next section, we propose inner and outer convex approximations of the feasible solution set.

3.2.3 Convex approximations of constraints (3.35) and (3.36)

Convex approximations of constraint (3.35)

We first denote

$$w^k = \left[\prod_{j=1}^M t_j^{a_{1j}^k}, \dots, \prod_{j=1}^M t_j^{a_{I_k j}^k} \right] \in \mathbb{R}^{I_k}, \quad k = 1, \dots, K.$$

Constraint (3.35) can be reformulated as

$$-(\mu^k)^\top w^k + \Phi_{\varphi_k}^{-1}(z_k^+) \sqrt{(w^k)^\top \Gamma_k w^k} \leq -\alpha_k, \quad k = 1, \dots, K. \quad (3.46)$$

As we assume that $\epsilon \leq 0.5$ and c^k follows a symmetric distribution, it is easy to see that $(\mu^k)^\top w^k - \alpha_k \geq 0$. Hence, (3.46) is equivalent to

$$(\Phi_{\varphi_k}^{-1}(z_k^+))^2 ((w^k)^\top \Gamma_k w^k) \leq ((\mu^k)^\top w^k - \alpha_k)^2, \quad k = 1, \dots, K,$$

which can be reformulated as

$$(w^k)^\top ((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \Gamma_k - \mu^k (\mu^k)^\top) w^k + 2\alpha_k (\mu^k)^\top w^k \leq \alpha_k^2, \quad k = 1, \dots, K. \quad (3.47)$$

As $w^k = \left[\prod_{j=1}^M t_j^{a_{1j}^k}, \dots, \prod_{j=1}^M t_j^{a_{I_k j}^k} \right]$, $k = 1, \dots, K$, constraint (3.47) is equivalent to

$$2\alpha_k \sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^M t_j^{a_{ij}^k} + \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} ((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k) \prod_{j=1}^M t_j^{a_{ij}^k + a_{pj}^k} \leq \alpha_k^2, \quad k = 1, \dots, K. \quad (3.48)$$

From (3.37) and (3.38), we know that $z_k^+ \geq 1 - \epsilon \geq 0.5$. Moreover, we know from Assumption 3.2 that

$$(\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k \geq (\Phi_{\varphi_k}^{-1}(1 - \epsilon))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k > 0,$$

for all $z_k^+ \in [1 - \epsilon, 1)$, $i, p = 1, \dots, I_k$, $k = 1, \dots, K$. Hence, given Assumption 3.2, we can apply the standard variable transformation $r_j = \log(t_j)$, $j = 1, \dots, M$, to (3.48). Therefore, we have an equivalent formulation of (3.48)

$$2\alpha_k \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} + \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) \right. \\ \left. + \log((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k) \right\} \leq \alpha_k^2, \quad k = 1, \dots, K. \quad (3.49)$$

Proposition 3.3. *Given Assumption 3.2, $f_{i,p,k}(z_k^+) = \log((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)$ is monotone increasing and convex for $z_k^+ \in [1 - \epsilon, 1)$, $i, p = 1, \dots, I_k$, $k = 1, \dots, K$.*

Proof. From the continuity and monotonicity of $\Phi_{\varphi_k}(\cdot)$ and the positiveness of $\phi_{\varphi_k}(\cdot)$, we know $\Phi_{\varphi_k}(\cdot)$ is strictly monotone a.s.. Hence, from the formula for the derivative of the inverse function, we have that $\Phi_{\varphi_k}^{-1}(z)$ is differential and its derivative is $\frac{1}{\frac{d}{dx}\Phi_{\varphi_k}(x)} = \frac{1}{\phi_{\varphi_k}(x)}$, where $x = \Phi_{\varphi_k}^{-1}(z)$. Furthermore, we have that $f_{i,p,k}(z_k^+)$ is differential and its first and second order derivatives are

$$\begin{aligned} f'_{i,p,k}(z_k^+) &= \frac{df_{i,p,k}(z_k^+)}{dz_k^+} = \frac{\sigma_{i,p}^k ((\Phi_{\varphi_k}^{-1}(z_k^+))^2)' }{(\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k} = \frac{2\sigma_{i,p}^k \Phi_{\varphi_k}^{-1}(z_k^+) (\Phi_{\varphi_k}^{-1}(z_k^+))'}{(\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k} \\ &= \frac{2\sigma_{i,p}^k \Phi_{\varphi_k}^{-1}(z_k^+)}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^+)) ((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)}, \\ f''_{i,p,k}(z_k^+) &= \frac{d^2 f_{i,p,k}(z_k^+)}{d(z_k^+)^2} = \frac{2\sigma_{i,p}^k (\Phi_{\varphi_k}^{-1}(z_k^+))'}{(\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^+))) ((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)} \\ &\quad - \frac{2\sigma_{i,p}^k \Phi_{\varphi_k}^{-1}(z_k^+) (\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^+)))' ((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)}{(\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^+)))^2 ((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)^2} \\ &= \frac{2\sigma_{i,p}^k \left(1 - \frac{\phi'_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^+))}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^+))} \Phi_{\varphi_k}^{-1}(z_k^+) \right) ((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k) - (2\sigma_{i,p}^k \Phi_{\varphi_k}^{-1}(z_k^+))^2}{(\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^+)))^2 ((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)^2}. \end{aligned}$$

From Assumption 3.2, we know that $\frac{df_{i,p,k}(z_k^+)}{dz_k^+} \geq 0$ and $f_{i,p,k}(z_k^+)$ is monotone increasing for $z_k^+ \in [1 - \epsilon, 1)$, $i, p = 1, \dots, I_k$, $k = 1, \dots, K$. From Assumption 3.2, we know that $\frac{d^2 f_{i,p,k}(z_k^+)}{d(z_k^+)^2} \geq 0$ and $f_{i,p,k}(z_k^+)$ is convex for $z_k^+ \in [1 - \epsilon, 1)$, $i, p = 1, \dots, I_k$, $k = 1, \dots, K$. \square

Thanks to the convexity and the monotonicity of $f_{i,p,k}(z_k^+)$, we use the piecewise linear approximation method to find a lower approximation of $f_{i,p,k}(z_k^+)$ [65]. Then, we propose a piecewise linear approximation method to find an lower approximation of $f_{i,p,k}(z_k^+)$.

We choose S different linear functions:

$$F_{s,i,p,k}^L(z_k^+) = d_{s,i,p,k} z_k^+ + b_{s,i,p,k}, \quad s = 1, \dots, S,$$

which are the tangent segments of $f_{i,p,k}(z_k^+)$ at given points in $[1 - \epsilon, 1)$, e.g., $\xi_1, \xi_2, \dots, \xi_S$. Here, we choose $\xi_S = 1 - \delta$, where δ is a very small positive real number. We have

$$d_{s,i,p,k} = f'_{i,p,k}(\xi_s)$$

and

$$b_{s,i,p,k} = f_{i,p,k}(\xi_s) - f'_{i,p,k}(\xi_s) \xi_s, \quad s = 1, \dots, S.$$

Then, we use the piecewise linear function

$$F_{i,p,k}^L(z_k^+) = \max_{s=1, \dots, S} F_{s,i,p,k}^L(z_k^+)$$

to approximate $f_{i,p,k}(z_k^+)$.

Proposition 3.4. Given Assumption 3.2, $F_{i,p,k}^L(z_k^+) \leq f_{i,p,k}(z_k^+)$, $\forall z_k^+ \in [1-\epsilon, 1)$.

Proof. The proof can be drawn from the convexity of $f_{i,p,k}(z_k^+)$ shown in Proposition 3.3. \square

We use the piecewise linear function $F_{i,p,k}^L(z_k^+)$ to replace $\log((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)$ in (3.49). Hence, we have the following convex approximation of constraint (3.35):

$$\begin{cases} 2\alpha_k \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} + \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) \right. \\ \left. + \omega_{i,p,k}^L \right\} \leq \alpha_k^2, \quad k = 1, \dots, K, \\ d_{s,i,p,k} z_k^+ + b_{s,i,p,k} \leq \omega_{i,p,k}^L, \quad s = 1, \dots, S, \quad i, p = 1, \dots, I_k, \quad k = 1, \dots, K. \end{cases} \quad (3.50)$$

As $\log((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)$ is convex, (3.50), together with (3.36)-(3.38), provides an outer approximation of the feasible set of (SGP).

To get an upper approximation of the function $f_{i,p,k}(z_k^+)$, we sort $1-\epsilon, \xi_1, \xi_2, \dots, \xi_S$ in the increasing order and denote the sorted array by $\tilde{\xi}_1, \tilde{\xi}_2, \dots, \tilde{\xi}_{S+1}$. The segments

$$F_{s,i,p,k}^U(z_k^+) = \tilde{d}_{s,i,p,k} z_k^+ + \tilde{b}_{s,i,p,k}, \quad s = 1, \dots, S,$$

form a piecewise linear function

$$F_{i,p,k}^U(z_k^+) = \max_{s=1, \dots, S} F_{s,i,p,k}^U(z_k^+).$$

Here,

$$\tilde{d}_{s,i,p,k} = \frac{f_{i,p,k}(\tilde{\xi}_{s+1}) - f_{i,p,k}(\tilde{\xi}_s)}{\tilde{\xi}_{s+1} - \tilde{\xi}_s}$$

and

$$\tilde{b}_{s,i,p,k} = -\tilde{d}_{s,i,p,k} \tilde{\xi}_s + f_{i,p,k}(\tilde{\xi}_s), \quad s = 1, \dots, S.$$

Using the piecewise linear function $F_{i,p,k}^U(z_k^+)$ leads to the following convex approximation of constraint (3.35):

$$\begin{cases} 2\alpha_k \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} + \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) \right. \\ \left. + \omega_{i,p,k}^U \right\} \leq \alpha_k^2, \quad k = 1, \dots, K, \\ \tilde{d}_{s,i,p,k} z_k^+ + \tilde{b}_{s,i,p,k} \leq \omega_{i,p,k}^U, \quad s = 1, \dots, S, \quad i, p = 1, \dots, I_k, \quad k = 1, \dots, K. \end{cases} \quad (3.51)$$

As $\log((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)$ is convex, (3.51), together with (3.36)-(3.38), pro-

vides an inner approximation of the feasible set of (SGP).

Convex approximation of constraint (3.36)

In constraint (3.36), the terms $\sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma_{i,p}^k \prod_{j=1}^M t_j^{a_{ij}^k + a_{pj}^k}}$ and $\sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^M t_j^{a_{ij}^k}$ are convex with respect to $r_j = \log(t_j)$, $j = 1, \dots, M$.

From [50], we know that $\Phi_{\varphi_k}^{-1}(z_k^-)$ is convex with respect to z_k^- on $[1 - \epsilon, 1)$, if ϕ_{φ_k} is 0-decreasing with some threshold $t^*(0) > 0$, and $\epsilon < 1 - \Phi_{\varphi_k}(t^*(0))$. The definition of r -decreasing and $t^*(0)$ can be found in [50]. Hence, given some conditions, constraint (3.36) is a biconvex constraint on $[1 - \epsilon, 1)$. One can use the sequential convex approach to solve this problem. However, $\Phi_{\varphi_k}^{-1}(z_k^-)$ cannot be expressed explicitly, it is still not easy to compute the optimal z_k^- with fixed t_j . In this section, we use the piecewise linear approximation method proposed in [65], and its modified approximation method to find tight lower and upper bounds of (3.36).

We first make the standard variable transformation $r_j = \log(t_j)$, $j = 1, \dots, M$ to (3.36) in order to get an equivalent formulation of (3.36)

$$\begin{aligned} & \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) + 2 \log(\Phi_{\varphi_k}^{-1}(z_k^-)) + \log(\sigma_{i,p}^k) \right\}} \\ & + \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} \leq \beta_k, \quad k = 1, \dots, K. \end{aligned} \quad (3.52)$$

Lemma 3.1. *Given Assumption 3.2, $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ is monotone increasing and convex on $[1 - \epsilon, 1)$.*

Proof. The first order and second order derivatives of $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ are

$$\frac{d(\log(\Phi_{\varphi_k}^{-1}(z_k^-)))}{dz_k^-} = \frac{1}{\Phi_{\varphi_k}^{-1}(z_k^-) \phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^-))},$$

and

$$\frac{d^2(\log(\Phi_{\varphi_k}^{-1}(z_k^-)))}{d(z_k^-)^2} = - \frac{1 + \Phi_{\varphi_k}^{-1}(z_k^-) \frac{\phi'_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^-))}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^-))}}{(\Phi_{\varphi_k}^{-1}(z_k^-) \phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^-)))^2}.$$

From Assumption 3.2, we know that

$$1 + \Phi_{\varphi_k}^{-1}(1 - \epsilon) \frac{\phi'_{\varphi_k}(\Phi_{\varphi_k}^{-1}(1 - \epsilon))}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(1 - \epsilon))} < 0.$$

As $1 + \Phi_{\varphi_k}^{-1}(z_k^-) \frac{\phi'_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^-))}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z_k^-))}$ is monotone decreasing with respect to z_k^- in $[1 - \epsilon, 1)$, we can find that the second order derivative of $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ is larger than or equal to zero on $z_k^- \in [1 - \epsilon, 1)$, which implies the convexity of $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ on $[1 - \epsilon, 1)$, $k = 1, \dots, K$. \square

From the convexity and monotonicity, we can use the piecewise linear approximation methods introduced in the last section to find tight piecewise linear approximations for $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$. We choose S different linear functions:

$$G_{s,k}^L(z_k^-) = l_{s,k}z_k^- + q_{s,k}, \quad s = 1, \dots, S,$$

which are the tangent segments of $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ at $\xi_1, \xi_2, \dots, \xi_S$, respectively. We have

$$l_{s,k} = \frac{1}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(\xi_s))\Phi_{\varphi_k}^{-1}(z_k^-)}$$

and

$$q_{s,k} = \log(\Phi_{\varphi_k}^{-1}(\xi_s)) - \frac{\xi_s}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(\xi_s))\Phi_{\varphi_k}^{-1}(z_k^-)}, \quad s = 1, \dots, S.$$

Then, we use the piecewise linear function

$$G_k^L(z_k^-) = \max_{s=1, \dots, S} G_{s,k}^L(z_k^-)$$

to approximate $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$, and derive the following convex approximation of (3.52):

$$\left\{ \begin{array}{l} \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} + \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) + 2\tilde{\omega}_k^L + \log(\sigma_{i,p}^k) \right\}} \\ \leq \beta_k, \quad k = 1, \dots, K, \\ l_{s,k}z_k^- + q_{s,k} \leq \tilde{\omega}_k^L, \quad s = 1, \dots, S, \quad k = 1, \dots, K. \end{array} \right. \quad (3.53)$$

As $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ is convex, (3.53) together with (3.35), (3.37)-(3.38) provides an outer approximation of the feasible set of (SGP).

Moreover, we use the segments

$$G_{s,k}^U(z_k^-) = \tilde{l}_{s,k}z_k^- + \tilde{q}_{s,k}, \quad s = 1, \dots, S,$$

between $\tilde{\xi}_1, \tilde{\xi}_2, \dots, \tilde{\xi}_{S+1}$ to form a piecewise linear function

$$G_k^U(z_k^-) = \max_{s=1, \dots, S} G_{s,k}^U(z_k^-).$$

Here,

$$\tilde{l}_{s,k} = \frac{\log(\Phi_{\varphi_k}^{-1}(\tilde{\xi}_{s+1})) - \log(\Phi_{\varphi_k}^{-1}(\tilde{\xi}_s))}{\tilde{\xi}_{s+1} - \tilde{\xi}_s},$$

and

$$\tilde{q}_{s,k} = -\tilde{l}_{s,k}\tilde{\xi}_s + \log(\Phi_{\varphi_k}^{-1}(\tilde{\xi}_s)), \quad s = 1, \dots, S.$$

Using the piecewise linear function $G_k^U(z_k^-)$ to replace $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ in (3.52) gives the following convex approximation of the constraint (3.52):

$$\left\{ \begin{array}{l} \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} + \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) + 2\tilde{\omega}_k^U + \log(\sigma_{i,p}^k) \right\}} \\ \leq \beta_k, \quad k = 1, \dots, K, \\ \tilde{l}_{s,k} z_k^- + \tilde{q}_{s,k} \leq \tilde{\omega}_k^U, \quad s = 1, \dots, S, \quad k = 1, \dots, K. \end{array} \right. \quad (3.54)$$

As $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ is convex, (3.54) provides an inner approximation.

Main result

Theorem 3.4. *Given Assumptions 3.1 and 3.2, we have the following convex approximations for the joint rectangular geometric chance constrained programs (SGP):*

$$\begin{aligned} (SGP_L) \quad & \min_{\substack{r, z^+, z^-, \\ x, \omega^L, \tilde{\omega}^L}} \sum_{i=1}^{I_0} \mu_i^0 \exp \left\{ \sum_{j=1}^M a_{ij}^0 r_j \right\} \\ & \text{s.t.} \quad d_{s,i,p,k} z_k^+ + b_{s,i,p,k} \leq \omega_{i,p,k}, \quad s = 1, \dots, S, \\ & \quad \quad i, p = 1, \dots, I_k, \quad k = 1, \dots, K, \\ & \quad \quad 2\alpha_k \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} + \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) \right. \\ & \quad \quad \left. + \omega_{i,p,k}^L \right\} \leq \alpha_k^2, \quad k = 1, \dots, K, \\ & \quad \quad \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) + 2\tilde{\omega}_k^L + \log(\sigma_{i,p}^k) \right\}} \\ & \quad \quad + \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} \leq \beta_k, \quad k = 1, \dots, K, \\ & \quad \quad l_{s,k} z_k^- + q_{s,k} \leq \tilde{\omega}_k^L, \quad s = 1, \dots, S, \quad k = 1, \dots, K, \\ & \quad \quad z_k^+ + z_k^- - 1 \geq e^{x_k}, \quad 0 \leq z_k^+, z_k^- \leq 1, \quad k = 1, \dots, K, \\ & \quad \quad \sum_{k=1}^K x_k \geq \log(1 - \epsilon), \quad x_k \leq 0, \quad k = 1, \dots, K. \end{aligned}$$

$$\begin{aligned} (SGP_U) \quad & \min_{\substack{r, z^+, z^-, \\ x, \omega^U, \tilde{\omega}^U}} \sum_{i=1}^{I_0} \mu_i^0 \exp \left\{ \sum_{j=1}^M a_{ij}^0 r_j \right\} \\ & \text{s.t.} \quad \tilde{d}_{s,i,p,k} z_k^+ + \tilde{b}_{s,i,p,k} \leq \omega_{i,p,k}^U, \quad s = 1, \dots, S, \\ & \quad \quad i, p = 1, \dots, I_k, \quad k = 1, \dots, K, \end{aligned}$$

$$2\alpha_k \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} + \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) \right. \\ \left. + \omega_{i,p,k}^U \right\} \leq \alpha_k^2, \quad k = 1, \dots, K,$$

$$\sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \exp \left\{ \sum_{j=1}^M (a_{ij}^k r_j + a_{pj}^k r_j) + 2\tilde{\omega}_k^U + \log(\sigma_{i,p}^k) \right\}} \\ + \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^M a_{ij}^k r_j \right\} \leq \beta_k, \quad k = 1, \dots, K,$$

$$\tilde{l}_{s,k} z_k^- + \tilde{q}_{s,k} \leq \tilde{\omega}_k^U, \quad s = 1, \dots, S, \quad k = 1, \dots, K,$$

$$z_k^+ + z_k^- - 1 \geq e^{x_k}, \quad 0 \leq z_k^+, z_k^- \leq 1, \quad k = 1, \dots, K,$$

$$\sum_{k=1}^K x_k \geq \log(1 - \epsilon), \quad x_k \leq 0, \quad k = 1, \dots, K.$$

The optimal value of the approximation problem (SGP_L) is a lower bound of problem (SGP). The optimal value of the approximation problem (SGP_U) is an upper bound of problem (SGP). Moreover, if the feasible set of (SGP) is nonempty, both (SGP_L) and (SGP_U) are reformulations of (SGP) when S goes to infinity.

Proof. (SGP_L) is obtained from the reformulation (SGP_r) of (SGP) and the two outer approximations (3.50) and (3.53). Besides, we transform the variable y into $x_k = \log(y_k)$, $k = 1, \dots, K$. The outer approximations guarantee that the feasible region of (SGP_L) contains the feasible region of (SGP). Hence the optimal value of (SGP_L) is a lower bound of problem (SGP).

Meanwhile, (SGP_U) is obtained from the reformulation (SGP_r) of (SGP), the two inner approximations (3.51) and (3.54) and the variable transformation $x_k = \log(y_k)$, $k = 1, \dots, K$. The inner approximations guarantee that the feasible region of (SGP_U) is contained in the feasible region of (SGP). Hence the optimal value of (SGP_U) is an upper bound of problem (SGP).

Finally, when S goes to infinity, $G^L(z_k^-)$ uniformly converges to $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ for $z_k^- \in [1 - \epsilon, 1)$ due to the convexity of the two terms given in Assumption 3.2 and the property of the piece-wise linear approximation. Hence, the left hand side term of (3.53) converges to the left hand side term of (3.52) uniformly. Given Assumption 3.2, left hand side terms of both (3.52) and (3.53) are convex. Moreover, both feasible sets are closed and the feasible set constrained by (3.52) are nonempty. Hence, the convex feasible set constrained by (3.53) converges to the convex feasible set constrained by (3.52) when S is large enough. Similarly, when S goes to infinity, $G^U(z_k^-)$ uniformly converges to $\log(\Phi_{\varphi_k}^{-1}(z_k^-))$ for $z_k^- \in [1 - \epsilon, 1)$, and the feasible set constrained by (3.54) converges to the feasible set constrained by (3.52). Moreover, both $F^L(z_k^+)$ and $F^U(z_k^+)$ uniformly converge to $\log((\Phi_{\varphi_k}^{-1}(z_k^+))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k)$. From the convexity of the feasible sets constrained by (3.49), (3.50) and (3.51), we know the feasible sets constrained by (3.50) and by (3.51) both converge to the feasible set constrained by (3.49). By taking the

intersection of the feasible sets constrained by (3.50) and by (3.53), we can find that the distance between the feasible sets of (SGP_L) and (SGP) converges to zero when S goes to infinity. Similarly, by taking the intersection of the feasible sets constrained by (3.51) and by (3.54), we can find that the distance between the feasible sets of (SGP_U) and (SGP) converges to zero when S goes to infinity. This means both (SGP_L) and (SGP_U) are reformulations of (SGP) when S goes to infinity. \square

Both (SGP_L) and (SGP_U) are convex programming problems. Interior point methods can be used to solve them efficiently.

3.2.4 Numerical experiments

We test the performances of our approximation methods by considering the following stochastic rectangular shape optimization problem with a joint chance constraints:

$$\begin{aligned} \min_{x_1, \dots, x_m} \quad & \prod_{i=1}^m x_i^{-1} \\ \text{s.t.} \quad & \mathbb{P} \left[\alpha_{wall} \leq \sum_{j=1}^{m-1} \left(\frac{m-1}{A_j} x_1 \prod_{i=2, i \neq j}^m x_i \right) \leq \beta_{wall}, \alpha_{flr} \leq \frac{1}{A_{flr}} \prod_{j=2}^m x_j \leq \beta_{flr} \right] \\ & \geq 1 - \epsilon, \\ & x_i x_j^{-1} \leq \gamma_{i,j}, \forall i \neq j. \end{aligned}$$

Here, $1/A_{flr}$ and $1/A_j$, $j = 1, \dots, m-1$, are considered as random variables. We assume $1/A_{flr}$ to be independent of $1/A_j$, $j = 1, \dots, m-1$. This example is a generalization of the shape optimization problem with random parameters in [65]. In all following experiments, we set $\gamma_{i,j} = 2, \forall i \neq j$, $\alpha_{wall} = 1$, $\beta_{wall} = 2$, $\alpha_{flr} = 1$, $\beta_{flr} = 2$ and $\epsilon = 5\%$.

In order to test the tightness of the piecewise linear approximation, we first consider three different examples in the elliptical distribution group.

Example 3.1. We set $m = 10$. We let $1/A_{flr}$ follow a normal distribution $N(0.02, 0.02)$, and the random vector $[1/A_1, 1/A_2, \dots, 1/A_{m-1}]$ follow a $(m-1)$ -dimensional normal distribution, such that $\mathbb{E}[1/A_j] = 0.01$, $j = 1, \dots, m-1$, and $\text{cov}[1/A_i, 1/A_j] = 0.01$, $i, j = 1, \dots, m-1$. Moreover, $1/A_j$ and $1/A_{flr}$ are independent, $j = 1, \dots, m-1$.

Example 3.2. We set $m = 15$. We let $1/A_{flr}$ follow a logistic distribution with location parameter 0.02 and scale parameter 0.078, i.e., $\mathbb{E}[1/A_{flr}] = 0.02$ and $\text{Var}[1/A_{flr}] = 0.02$. We let $[1/A_1, 1/A_2, \dots, 1/A_{m-1}]$ follow a $(m-1)$ -dimensional logistic distribution, such that $\mathbb{E}[1/A_j] = 0.01$, $j = 1, \dots, m-1$, and $\text{cov}[1/A_i, 1/A_j] = 0.01$, $i, j = 1, \dots, m-1$. Moreover, $1/A_j$ and $1/A_{flr}$ are independent, $j = 1, \dots, m-1$.

Example 3.3. We set $m = 20$. We let $1/A_{flr}$ follow a Student's t distribution, with location parameter 0.02, scale parameter 0.02 and degree of freedom 4. We

let $[1/A_1, 1/A_2, \dots, 1/A_{m-1}]$ follow a $(m-1)$ -dimensional Student's t distribution with location vector μ , scale matrix Γ and degree of freedom 4, where $\mu_j = 0.01$, $j = 1, \dots, m-1$, and $\Gamma_{i,j} = 0.01$, $i, j = 1, \dots, m-1$. Moreover, $1/A_j$ and $1/A_{flr}$ are independent, $j = 1, \dots, m-1$.

We should verify the satisfaction of the three conditions in Assumption 3.2 before we make the upper and lower approximations. We set three indexes:

$$\text{Index 1} = \max_k \frac{\phi'_{\varphi_k}(\Phi_{\varphi_k}^{-1}(1-\epsilon))}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(1-\epsilon))} \Phi_{\varphi_k}^{-1}(1-\epsilon),$$

$$\text{Index 2} = \min_{i,p,k} (\Phi_{\varphi_k}^{-1}(1-\epsilon))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k,$$

$$\text{Index 3} = \min_{i,p,k,z \in [1-\epsilon, 1]} 2\sigma_{i,p}^k \left(1 - \frac{\phi'_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z))}{\phi_{\varphi_k}(\Phi_{\varphi_k}^{-1}(z))} \Phi_{\varphi_k}^{-1}(z) \right) \left((\Phi_{\varphi_k}^{-1}(z))^2 \sigma_{i,p}^k - \mu_i^k \mu_p^k \right) - (2\sigma_{i,p}^k \Phi_{\varphi_k}^{-1}(z))^2.$$

We compute the three indexes for the three examples, shown in Table 3.2.

Table 3.2: Verification on the satisfaction of Assumption 3.2.

	Index 1	Index 2	Index 3
Example 3.1	-2.7055	0.027	0.0009
Example 3.2	-2.6500	0.0263	0.0009
Example 3.3	-2.6594	0.0453	0.0015

From the negativeness of Index 1 and the positiveness of Index 2 and Index 3, we can verify that Assumption 3.2 holds for all the three examples.

Then we use the proposed convex approximations, (SGP_U) and (SGP_L) , to obtain lower and upper bounds for the stochastic rectangular shape optimization problem. For each example, we let the number of segments S go from 1 to 2000 and solve 10 groups of (SGP_U) and (SGP_L) . Computation results of Example 3.1 are shown in Table 3.3, computation results of Example 3.2 are shown in Table 3.4, and computation results of Example 3.3 are shown in Table 3.5,

The first column in the three Tables gives the number of segments S used in (SGP_U) and (SGP_L) . The second and third columns give the numbers of variables and the numbers of constraints of (SGP_U) , respectively. The sixth and seventh columns give the numbers of variables and the numbers of constraints of (SGP_L) , respectively. The fourth and the fifth columns give the upper bounds and the CPU times of (SGP_U) , respectively. The eighth and the ninth columns give the lower bounds and the CPU times of (SGP_L) , respectively. We use SDPT3 solver from CVX package to solve the approximation problems with Matlab R2012b, on a PC with a 2.6 Ghz Intel Core i7-5600U CPU and 12.0 GB RAM. For better illustration, we compute the gaps of the two piecewise linear approximation bounds, which are the percentage differences between the lower bound and the corresponding upper bound, and show them in the last column.

Table 3.3: Computational results of approximations for normal distribution in Example 3.1.

S	Var.	Con.	UB	CPU(s)	Var.	Con.	LB	CPU(s)	Gap(%)
1	681	216	8.1507	13.3390	681	216	6.6180	10.6090	23.1583
2	685	216	8.1368	13.8255	685	216	7.4902	10.7009	8.6325
5	697	216	8.1311	14.7254	697	216	8.0253	13.4361	1.3186
10	717	216	8.1311	15.5340	717	216	8.1116	14.6222	0.2403
20	757	216	8.1303	15.7734	757	216	8.1279	14.9508	0.0300
100	1077	216	8.1303	17.9095	1077	216	8.1303	17.8283	0.0000
200	1477	216	8.1303	18.0155	1477	216	8.1303	18.3668	0.0000
500	2677	216	8.1303	20.4507	2677	216	8.1303	20.2324	0.0000
1000	4677	216	8.1303	24.5032	4677	216	8.1303	24.3255	0.0000
2000	8677	216	8.1303	33.3699	8677	216	8.1303	33.8009	0.0000

Table 3.4: Computational results of approximations for Logistic distribution in Example 3.2.

S	Var.	Con.	UB	CPU(s)	Var.	Con.	LB	CPU(s)	Gap(%)
1	1526	466	20.3036	27.0223	1526	466	17.0253	25.1127	19.2557
2	1530	466	19.8479	32.8251	1530	466	18.7557	25.9661	5.8232
5	1542	466	19.6918	35.1827	1542	466	19.5681	32.7463	0.6320
10	1562	466	19.6740	36.5728	1562	466	19.6583	35.3287	0.0800
20	1602	466	19.6701	39.4038	1602	466	19.6681	42.9635	0.0100
100	1922	466	19.6701	39.2123	1922	466	19.6701	41.7810	0.0000
200	2322	466	19.6701	39.2405	2322	466	19.6701	40.4707	0.0000
500	3522	466	19.6701	41.2078	3522	466	19.6701	40.4667	0.0000
1000	5522	466	19.6701	46.4123	5522	466	19.6701	46.5576	0.0000
2000	9522	466	19.6701	54.0321	9522	466	19.6701	53.9358	0.0000

Table 3.5: Computational results of approximations for Student's t distribution in Example 3.3.

S	Var.	Con.	UB	CPU(s)	Var.	Con.	LB	CPU(s)	Gap(%)
1	2721	816	57.2656	73.8665	2721	816	43.3974	51.5729	31.9562
2	2725	816	49.5113	74.7218	2725	816	46.5860	57.3962	6.2793
5	2737	816	47.4843	91.6425	2737	816	47.2853	79.1193	0.4209
10	2757	816	47.3326	80.6797	2757	816	47.2853	82.1539	0.1001
20	2797	816	47.3042	94.3056	2797	816	47.2853	85.8753	0.0400
100	3117	816	47.2901	87.8563	3117	816	47.2901	91.1235	0.0000
200	3517	816	47.2901	86.2648	3517	816	47.2901	84.5257	0.0000
500	4717	816	47.2901	89.6698	4717	816	47.2901	87.5018	0.0000
1000	6717	816	47.2901	95.3347	6717	816	47.2901	92.2044	0.0000
2000	10717	816	47.2901	106.3244	10717	816	47.2901	104.3638	0.0000

From the three tables, we can see that as the number of segments S increases, the gap of the corresponding piecewise linear approximation bounds for the problem becomes smaller. For all the three examples, the gap is tight when the number of segments is equal to or larger than 100. Notice that the CPU time does not increase proportionally with the increase of S . When $S = 2000$, the CPU time is less than twice of the CPU time when $S = 2$.

We further study the effect on the approximation precision and CPU time of mean value, increasing variance and m .

We use Example 3.1 as the basic setting. We set $\mathbb{E}[1/A_{flr}] = \mu_{flr}$, $\mathbb{E}[1/A_j] = \mu_A$, $j = 1, \dots, m - 1$, and let them vary from 0.02 to 0.2 and from 0.01 to 0.1, respectively. All other parameters are fixed at the same values as Example 3.1. We should verify the satisfaction of the conditions in Assumption 3.2. Like what we do in Table 3.2, we compute three indexes for each group of the distributions with different mean values, and show them in Table 3.6. From the table, we can easily see that Assumption 3.2 holds for all the groups of distributions with different mean values.

Table 3.6: Verification on the satisfaction of Assumption 3.2 with increasing mean value.

μ_A	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
μ_{flr}	0.02	0.04	0.06	0.08	0.1	0.12	0.14	0.16	0.18	0.2
Index 1	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055
Index 2	0.0270	0.0267	0.0262	0.0255	0.0246	0.0235	0.0222	0.0207	0.0190	0.0171
Index 3	0.0009	0.0009	0.0009	0.0008	0.0007	0.0007	0.0006	0.0004	0.0003	0.0002

Then we use the proposed convex approximations, (SGP_U) and (SGP_L) , to obtain lower and upper bounds for the stochastic rectangular shape optimization problem. We set the number of segments S to be 200. Computation results, including sizes of programming problems, lower/upper bounds and CPU times, and bound gaps, are shown in Table 3.7. From Table 3.7, we can observe that upper bound and lower bound are tight under all the settings with different mean values. The CPU times are almost the same when we let the mean values vary.

Table 3.7: Computational results with increasing mean value.

μ_A	μ_{flr}	Var.	Con.	UB	CPU(s)	Var.	Con.	LB	CPU(s)	Gap(%)
0.01	0.02	1477	216	8.1303	18.3817	1477	216	8.1303	19.5148	0
0.02	0.04	1477	216	8.6504	18.7815	1477	216	8.6504	19.2838	0
0.03	0.06	1477	216	9.1724	21.7914	1477	216	9.1724	22.3960	0
0.04	0.08	1477	216	9.6978	22.3719	1477	216	9.6978	22.5001	0
0.05	0.1	1477	216	10.3026	18.5951	1477	216	10.3026	21.5380	0
0.06	0.12	1477	216	11.9317	18.6117	1477	216	11.9317	22.3053	0
0.07	0.14	1477	216	13.9698	18.1727	1477	216	13.9698	21.3361	0
0.08	0.16	1477	216	16.1110	18.1408	1477	216	16.1110	21.4588	0
0.09	0.18	1477	216	18.3055	18.3133	1477	216	18.3055	21.5131	0
0.1	0.2	1477	216	20.5385	18.2731	1477	216	20.5385	18.7971	0

Nextly, we let the variance and covariance vary. In detail, we use Example 3.1 as the basic setting, set $\text{Var}[1/A_{flr}] = \sigma_{flr}^2$, $\text{Cov}[1/A_i, 1/A_j] = \sigma_A^2$, $i, j = 1, \dots, m - 1$, and let σ_{flr}^2 and σ_A^2 vary from 0.02 to 0.2 and from 0.01 to 0.1,

respectively. We verify the satisfaction of the conditions in Assumption 3.2 by computing the three indexes shown in Table 3.8. Then we set $S = 200$ and use (SGP_U) and (SGP_L) , to obtain lower and upper bounds for the stochastic rectangular shape optimization problem. Computation results, including sizes of programming problems, lower/upper bounds and CPU times, and bound gaps, are shown in Table 3.9. From Table 3.9, we can observe that upper bound and lower bound are tight under all the settings with different variances and covariances. The CPU times are almost the same when we let variances and covariances vary.

Table 3.8: Verification on the satisfaction of Assumption 3.2 with increasing variance.

σ_A^2	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
σ_{flr}^2	0.02	0.04	0.06	0.08	0.1	0.12	0.14	0.16	0.18	0.2
Index 1	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055	-2.7055
Index 2	0.0270	0.054	0.0811	0.1081	0.1352	0.1622	0.1893	0.2163	0.2434	0.2705
Index 3	0.0009	0.0037	0.0083	0.0147	0.023	0.0332	0.0452	0.0590	0.0747	0.0922

Table 3.9: Computational results with increasing variance.

σ_A^2	σ_{flr}^2	Var.	Con.	UB	CPU(s)	Var.	Con.	LB	CPU(s)	Gap(%)
0.01	0.02	1477	216	8.1303	18.4132	1477	216	8.1303	19.2847	0
0.02	0.04	1477	216	11.7271	18.8420	1477	216	11.7271	19.8981	0
0.03	0.06	1477	216	14.5661	18.7148	1477	216	14.5661	19.2016	0
0.04	0.08	1477	216	17.0083	18.6797	1477	216	17.0083	19.3238	0
0.05	0.1	1477	216	19.2036	18.5938	1477	216	19.2036	22.4753	0
0.06	0.12	1477	216	21.2424	18.7255	1477	216	21.2424	22.0337	0
0.07	0.14	1477	216	23.1872	18.1202	1477	216	23.1872	21.7534	0
0.08	0.16	1477	216	25.0757	18.5851	1477	216	25.0757	21.698	0
0.09	0.18	1477	216	26.9343	18.5935	1477	216	26.9343	18.8600	0
0.1	0.2	1477	216	28.7719	18.3161	1477	216	28.7719	19.0957	0

Table 3.10: Computational results for increasing m .

m	Var.	Con.	UB	CPU(s)	Var.	Con.	LB	CPU(s)	Gap(%)
3	880	32	1.4430	4.5124	880	32	1.4430	5.9128	0
5	982	66	1.6864	7.4550	982	66	1.6864	7.9741	0
7	1138	114	3.5648	10.7008	1138	114	3.5648	11.1916	0
10	1477	216	8.1303	19.1872	1477	216	8.1303	20.4025	0
20	3517	816	36.4121	86.0015	3517	816	36.4121	95.0357	0

Finally, we test the effect of m . In detail, we use Example 3.1 with different m . For each group, we let $1/A_{flr}$ follow a normal distribution $N(0.02, 0.02)$, and the random vector $[1/A_1, 1/A_2, \dots, 1/A_{m-1}]$ follow a $(m-1)$ -dimensional normal distribution, with $\mathbb{E}[1/A_j] = 0.01$, $j = 1, \dots, m-1$, and $\text{Cov}[1/A_i, 1/A_j] = 0.01$, $i, j = 1, \dots, m-1$. As the parameters of each component are set the same, the satisfaction of the conditions in Assumption 3.2 can be observed from Example 3.1. Then we set $S = 200$ and use (SGP_U) and (SGP_L) , to obtain lower and upper bounds for the stochastic rectangular shape optimization problem. Computation results, including sizes of programming problems, lower/upper bounds and CPU

times, and bound gaps, are shown in Table 3.10. From Table 3.10, we can again observe that the upper bound and the lower bound are both tight for all the examples with different sizes. However, the CPU time increases faster than a linear way with respect to m . It increases even faster when m is larger.

3.3 Conclusion

In this chapter, we first review a work of Liu et al. [65] about geometric programs with joint chance constraints. As an extension, we then discuss joint rectangular geometric chance constrained programs. We propose tight convex approximations for them under elliptical distributions and some conditions on the parameters.

This chapter corresponds to a paper submitted to Engineering and Optimization.

Chapter 4

Bounds for Chance Constrained Problems

In this chapter, we develop four upper bounds for individual and joint chance constraints with independent matrix vector rows. The deterministic approximations of the probability constraints are based on the one-side Chebyshev inequality, Chernoff inequality, Bernstein inequality and Hoeffding inequality, respectively. Various sufficient conditions under which the aforementioned approximations are convex and tractable are derived. Therefore, we approximate the chance constrained problems as tractable convex optimization problems based on piecewise linear and tangent approximations allowing to reduce further the computational complexity. Finally, numerical results on randomly generated data are discussed allowing to identify the tight deterministic approximations.

4.1 Introduction

In this chapter, we consider the following chance constrained linear program:

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \mathbb{P} \{ \Xi x \leq H \} \geq \alpha, \\ & x \in X, \end{aligned} \tag{4.1}$$

where $H = (h_1, \dots, h_K) \in \mathbb{R}^K$, $\Xi = [\xi_1, \dots, \xi_K]^T$ is a $K \times n$ random matrix, where ξ_k , $k = 1, \dots, K$ is a random vector in \mathbb{R}^n . \mathbb{P} is a probability measure, x is a decision vector, set $X \subseteq \mathbb{R}_+^n$, $c \in \mathbb{R}^n$ and $0 < \alpha < 1$ a prespecified confidence parameter.

An individual chance constrained problem can be written as

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \mathbb{P} \{ \xi^T x \leq h \} \geq \alpha, \\ & x \in X. \end{aligned} \tag{4.2}$$

One of the main goals in problems with an individual chance constraint is to come up with a deterministic equivalent problem such that the feasible set $S(\alpha) = \{x \in X : \mathbb{P}\{\xi^T x \leq h\} \geq \alpha\}$ of problem (4.2) is convex.

For instance, if we consider the case of a multivariate normally distributed vectors ξ with mean $\bar{\xi} = \mathbb{E}(\xi)$ and positive definite variance-covariance matrix Σ , the following relations hold true:

$$\mathbb{P}(\xi^T x \leq h) \geq \alpha, \quad (4.3)$$

$$\begin{aligned} & \Updownarrow \\ \bar{\xi}^T x + F^{-1}(\alpha) \|\Sigma^{1/2} x\| & \leq h, \end{aligned} \quad (4.4)$$

where $F^{-1}(\cdot)$ is the inverse of F , the standard normal cumulative distribution function. The same scheme can be applied to elliptical distributions, e.g., Laplace distribution, t-Student distribution, Cauchy distribution, Logistic distribution. [24, 65].

When the probability distributions are not elliptical or not known in advance, lower and upper bounds on $\mathbb{P}\{\xi^T x \leq h\} \geq \alpha$, can be very useful.

4.2 Chebyshev and Chernoff Bounds

In the following, we provide bounds based on deterministic approximations of probabilistic inequalities such as the one side Chebyshev and Chernoff inequalities.

4.2.1 Chebychev bounds

We consider the *one-side Chebyshev* inequality [64, 85]. We assume that ξ has finite second moments and denote by $\sigma_\xi^2 = \text{Var}(\xi)$ the variance of ξ , and $\bar{\xi} = E(\xi)$ the mean of ξ . The one-side Chebyshev inequality is given by

$$\mathbb{P}(\xi - \bar{\xi} \geq h) \leq \frac{\sigma_\xi^2}{\sigma_\xi^2 + h^2}. \quad (4.5)$$

For the individual chance constrained problem, we have the following results:

Theorem 4.1. *Assume that ξ has finite first and second moments. Under one-sided Chebyshev inequality (4.5), we have an inner approximation of problem (4.2) as follows*

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \bar{\xi}^T x + \sqrt{\frac{\alpha}{1-\alpha}} \|\Sigma^{1/2} x\| \leq h, \\ & x \in X, \end{aligned} \quad (4.6)$$

Moreover, problem (4.6) is a convex problem.

Proof. First, we note that

$$\mathbb{P}(\xi^T x \leq h) \geq \alpha, \quad (4.7)$$

$$\begin{aligned} & \Downarrow \\ \mathbb{P}(\xi^T x > h) & \leq 1 - \alpha, \quad (4.8) \end{aligned}$$

$$\begin{aligned} & \Downarrow \\ \mathbb{P}(\xi^T x - \bar{\xi}^T x > h - \bar{\xi}^T x) & \leq 1 - \alpha. \quad (4.9) \end{aligned}$$

Then, we apply (4.5) to (4.9):

$$\mathbb{P}(\xi^T x - \bar{\xi}^T x > h - \bar{\xi}^T x) \leq P(\xi^T x - \bar{\xi}^T x \geq h - \bar{\xi}^T x) \leq \frac{\sigma_\xi^2}{\sigma_\xi^2 + (h - \bar{\xi}^T x)^2}, \quad (4.10)$$

where $\sigma_\xi^2 = x^T \Sigma x$ with variance-covariance matrix Σ . If $\frac{\sigma_\xi^2}{\sigma_\xi^2 + (h - \bar{\xi}^T x)^2} \leq 1 - \alpha$, then (4.7) will be satisfied. Therefore,

$$\frac{x^T \Sigma x}{x^T \Sigma x + (h - \bar{\xi}^T x)^2} \leq 1 - \alpha, \iff \frac{\alpha}{1 - \alpha} x^T \Sigma x \leq (h - \bar{\xi}^T x)^2,$$

which is equivalent to

$$\sqrt{\frac{\alpha}{1 - \alpha}} \|\Sigma^{1/2} x\| \leq h - \bar{\xi}^T x. \quad (4.11)$$

□

In the following, we extend our results to the case of independent joint chance constraints.

Remark 4.1. *For the sake of clarity, we replace the constraints $x \in X$ by $x \in \mathbb{R}_+^n$ for the joint chance constraints case in the remainder of this chapter. Notice that additional constraints could be considered if their logarithm transformation preserves the convexity. See [82] for examples preserving the convexity.*

If we assume that ξ_k , $k = 1, \dots, K$ are multivariate normally distributed independent row vectors with mean vector $\mu_k = (\mu_{k1}, \dots, \mu_{kn})$, and covariance matrix Σ_k , we can derive a deterministic reformulation of problem (4.1). $\mathbb{P}\{\Xi x \leq H\} \geq \alpha$ is equivalent to

$$\prod_{k=1}^K \mathbb{P}\{\xi_k^T x \leq h_k\} \geq \alpha = \prod_{k=1}^K \alpha^{y_k}, \quad (4.12)$$

with $\sum_{k=1}^K y_k = 1$, $y_k \geq 0$, $k = 1, \dots, K$ and $y = (y_1, \dots, y_K)^T$.

We provide now an upper bound to problem (4.1) based on the *one-side Chebyshev* inequality. We assume that ξ_k , $k = 1, \dots, K$, has finite second moments. Let Σ_k denote the covariance matrix of ξ_k and $\bar{\xi}_k = \mathbb{E}(\xi_k)$ denote its mean, we have

Theorem 4.2. *Based on one-side Chebyshev inequality, an upper bound for problem (4.1) can be obtained by solving the following deterministic equivalent problem*

$$\begin{aligned}
& \min_{x,y} c^T x \\
& \text{s.t.} \quad \bar{\xi}_k^T x + \sqrt{\frac{\alpha^{y_k}}{1-\alpha^{y_k}}} \left\| \Sigma_k^{1/2} x \right\| \leq h_k, \quad k = 1, \dots, K, \\
& \quad \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad x \in X. \quad (4.13)
\end{aligned}$$

Proof. From the reformulation (4.12) and the proof of Theorem 4.1, we can immediately obtain the conclusion. \square

Assumption 4.1. $X = \mathbb{R}_+^n \cap \mathbf{L}$, \mathbf{L} is selected such that $Z = \{z \in \mathbb{R}^n : z_j = \ln(x_j), j = 1 \dots, n, x \in \mathbf{L}\}$ is convex.

Problem (4.13) is not convex but biconvex ([41]) because of the first group of constraints. To come-up with a tractable convex reformulation, with Assumption 4.1, we use the following logarithmic transformation $z = \ln x$.

In this case, problem (4.13) can be reformulated as follows

$$\begin{aligned}
& \min_{z,y} c^T e^z \\
& \text{s.t.} \quad \bar{\xi}_k^T e^z + \left\| \Sigma_k^{1/2} e^{\ln\left(\sqrt{\frac{\alpha^{y_k}}{1-\alpha^{y_k}}}\right) \cdot e_n + z} \right\| \leq h_k \quad k = 1, \dots, K, \\
& \quad \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad z \in Z, \quad (4.14)
\end{aligned}$$

where e_n is an $n \times 1$ vector of ones

We now prove that problem (4.14) is convex for all $\alpha \in [0, 1]$.

Lemma 4.1. *Given sets X, Y, Z where X, Y are convex. Let $f : X \rightarrow Y$ be a convex function, $g : Y \rightarrow Z$ be a nonincreasing concave function. Then, we have $g \circ f : X \rightarrow Z$ is a concave function.*

Proof. Since f is convex, we have that for $\lambda \in [0, 1]$ and $x_1, x_2 \in X$, $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$. Therefore, as $g : Y \rightarrow Z$ be a nonincreasing concave function, we have

$$g \circ f(\lambda x_1 + (1-\lambda)x_2) \geq g(\lambda f(x_1) + (1-\lambda)f(x_2)) \geq \lambda g \circ f(x_1) + (1-\lambda)g \circ f(x_2).$$

This conclude this lemma. \square

Assumption 4.2. *For each $k = 1, \dots, K$, all the components of $\bar{\xi}_k$ and Σ_k are non-negative.*

Theorem 4.3. *If Assumption 4.2 holds, then problem (4.14) is convex for all $\alpha \in [0, 1]$.*

Proof. To show the convexity of problem (4.14), we firstly need to show the convexity of $\ln\left(\sqrt{\frac{\alpha^{y_k}}{1-\alpha^{y_k}}}\right)$. As $\ln\left(\sqrt{\frac{\alpha^{y_k}}{1-\alpha^{y_k}}}\right) = \frac{1}{2}(y_k \ln \alpha - \ln(1 - \alpha^{y_k}))$, we can deduce the convexity of function $\ln\left(\sqrt{\frac{\alpha^{y_k}}{1-\alpha^{y_k}}}\right)$ if $\ln(1 - \alpha^{y_k})$ is concave.

Since $\log(1 - p)$ is decreasing and concave with respect to p and α^{y_k} is convex with respect to y_k . we have that $\ln(1 - \alpha^{y_k})$ is concave with respect to y_k as shown by Lemma 4.1.

Since the norm is a convex function and it is also a nondecreasing function on nonnegative space, the composition function $\left\|\Sigma_k^{1/2} e^{\ln\left(\sqrt{\frac{\alpha^{y_k}}{1-\alpha^{y_k}}}\right)+z}\right\|$ is a convex function. The term $\bar{\xi}_k^T e^z$ is a convex function because $\bar{\xi}_k \geq 0$. Hence, the problem (4.14) is convex for all $\alpha \in [0, 1]$. \square

4.2.2 Chernoff bounds

We consider now the *Chernoff bound*:

$$\mathbb{P}(\xi \geq h) \leq \frac{\mathbb{E}(e^{t\xi})}{e^{th}}, \quad (4.15)$$

where $\mathbb{E}(e^{t\xi})$ is the moment generating function of the random variable ξ and $t > 0$. $\bar{\xi}$ is the mean of ξ and $\sigma^2 = \text{Var}(\xi)$ is the variance.

First, we proof the convexity of $\mathbb{E}(e^{t\xi^T x})$.

Lemma 4.2. *For any $t > 0$, $\mathbb{E}(e^{t\xi^T x})$ is a convex function with respect to x .*

Proof. Since $e^{t\xi^T x}$ is convex with respect to $x \in X$, we have that for $\lambda \in [0, 1]$ and $x_1, x_2 \in X$, $e^{t\xi^T(\lambda x_1 + (1-\lambda)x_2)} \leq \lambda e^{t\xi^T x_1} + (1-\lambda)e^{t\xi^T x_2}$. Therefore,

$$\mathbb{E}(e^{t\xi^T(\lambda x_1 + (1-\lambda)x_2)}) \leq \mathbb{E}(\lambda e^{t\xi^T x_1} + (1-\lambda)e^{t\xi^T x_2}) = \lambda \mathbb{E}(e^{t\xi^T x_1}) + (1-\lambda)\mathbb{E}(e^{t\xi^T x_2}).$$

\square

Theorem 4.4. *If ξ follows a normal distribution with mean vector $\bar{\xi}$ and variance-covariance matrix Σ , under Chernoff bound, problem (4.2) can be formulated as follows*

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \bar{\xi}^T x + \sqrt{2 \ln \frac{1}{(1-\alpha)}} \|\Sigma^{1/2} x\| \leq h, \\ & x \in X \end{aligned} \quad (4.16)$$

Moreover, problem (4.16) is a convex problem.

Proof. First, we have from (4.8)

$$\mathbb{P}(\xi^T x \leq h) \geq \alpha \iff \mathbb{P}(\xi^T x \geq h) \leq 1 - \alpha.$$

This implies

$$\mathbb{P}(\xi^T x \geq h) \leq \frac{\mathbb{E}(e^{t\xi^T x})}{e^{th}}. \quad (4.17)$$

Given $t > 0$, if we choose $\frac{\mathbb{E}(e^{t\xi^T x})}{e^{th}} \leq 1 - \alpha$, then we get an upper bound to problem (4.2) with feasible region

$$\bar{S}(\alpha) = \left\{ x \in X \subseteq \mathbb{R}_+^n \mid \mathbb{E}(e^{t\xi^T x}) \leq (1 - \alpha)e^{th} \right\}, \quad (4.18)$$

which is convex as $\mathbb{E}(e^{t\xi^T x})$ is convex with respect to x , as shown by Lemma 4.2.

If ξ is a normal distribution with mean $\bar{\xi}$ and variance-covariance Σ , i.e. $\xi \sim N(\bar{\xi}, \Sigma)$ then in (4.18) we have $\mathbb{E}(e^{t\xi^T x}) = e^{t\bar{\xi}^T x} \cdot e^{\frac{1}{2}x^T \Sigma x t^2}$. The feasible region $\bar{S}(\alpha)$ can be written as:

$$\bar{S}(\alpha) = \left\{ x \in X \subseteq \mathbb{R}_+^n \mid \exists t > 0 : \frac{1}{2}x^T \Sigma x t^2 + t\bar{\xi}^T x - th \leq \ln(1 - \alpha) \right\}. \quad (4.19)$$

The set (4.19) is equivalent to:

$$\inf_{t > 0} \left\{ \frac{1}{2}x^T \Sigma x t^2 + t\bar{\xi}^T x - th \right\} \leq \ln(1 - \alpha). \quad (4.20)$$

The first derivative of $\frac{1}{2}x^T \Sigma x t^2 + t\bar{\xi}^T x - th$ equal to zero with respect to t at $\frac{h - \bar{\xi}^T x}{x^T \Sigma x}$. If $h - \bar{\xi}^T x < 0$, the constraint (4.20) is equivalent to $0 < \ln(1 - \alpha)$, which is impossible. Hence, let $h - \bar{\xi}^T x \geq 0$. Therefore (4.19) is equivalent to

$$\bar{S}(\alpha) = \left\{ x \in X \subseteq \mathbb{R}_+^n \mid -(h - \bar{\xi}^T x)^2 \leq 2 \ln(1 - \alpha) x^T \Sigma x \right\}, \quad (4.21)$$

which is equivalent to the following convex set:

$$\bar{S}(\alpha) = \left\{ x \in X \subseteq \mathbb{R}_+^n \mid h - \bar{\xi}^T x \geq \sqrt{2 \ln \frac{1}{(1 - \alpha)}} \|\Sigma^{1/2} x\| \right\}. \quad (4.22)$$

□

We extend our results to the case of independent joint chance constraints.

If we assume that ξ_k , $k = 1, \dots, K$ are multivariate normally distributed independent row vectors with mean vector $\bar{\xi}_k = (\bar{\xi}_{k1}, \dots, \bar{\xi}_{kn})^T$ and covariance matrix Σ_k , we can derive a deterministic reformulation of problem (4.1) based on (4.12).

We consider now an upper bound to problem (4.1) based on *Chernoff bound*.

Theorem 4.5. *If ξ_k , $k = 1, \dots, K$ are independent with each other and normally distributed with mean vector $\bar{\xi}_k$ and covariance matrix Σ_k , based on Chernoff bound, an upper bound for problem (4.1) can be obtained by solving the following*

problem

$$\begin{aligned}
& \min_{z,y} c^T x \\
& \text{s.t.} \quad \bar{\xi}_k^T x + \sqrt{2 \ln \left(\frac{1}{1 - \alpha^{y_k}} \right)} \|\Sigma_k^{1/2} x\| \leq h_k, \quad k = 1, \dots, K, \\
& \quad \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad x \in X.
\end{aligned} \tag{4.23}$$

Proof. First, we note that

$$\mathbb{P}(\xi_k^T x \leq h_k) \geq \alpha^{y_k} \iff \mathbb{P}(\xi_k^T x \geq h_k) \leq 1 - \alpha^{y_k}.$$

Chernoff bound leads to

$$\mathbb{P}(\xi_k^T x \geq h_k) \leq \frac{\mathbb{E}(e^{t\xi_k^T x})}{e^{th_k}}, \quad k = 1, \dots, K, \tag{4.24}$$

with $t > 0$. An upper bound to problem (4.1) is then obtained by solving the following problem with $t > 0$:

$$\begin{aligned}
& \min_{z,y} c^T x \\
& \text{s.t.} \quad \mathbb{E}(e^{t\xi_k^T x}) \leq (1 - \alpha^{y_k}) e^{th_k}, \quad k = 1, \dots, K, \\
& \quad \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad x \in X.
\end{aligned} \tag{4.25}$$

However, if the probability distributions of ξ_k , $k = 1, \dots, K$ are not known, the main difficulty of the model (4.25) is given by the computation of $\mathbb{E}(e^{t\xi_k^T x})$. On the other hand, if we assume ξ_k , $k = 1, \dots, K$ are normal distributions with mean $\bar{\xi}_k$ and variance-covariance Σ_k , i.e. $\xi_k \sim N(\bar{\xi}_k, \Sigma_k)$, then we have that $\mathbb{E}(e^{t\xi_k^T x}) = e^{t\bar{\xi}_k^T x} \cdot e^{\frac{1}{2}x^T \Sigma_k x t^2}$, $k = 1, \dots, K$. Consequently problem (4.25) can be written as

$$\begin{aligned}
& \min_{z,y} c^T x \\
& \text{s.t.} \quad \frac{1}{2}x^T \Sigma_k x t^2 + t\bar{\xi}_k^T x - th_k \leq \ln(1 - \alpha^{y_k}), \quad k = 1, \dots, K, \\
& \quad \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad x \in X.
\end{aligned} \tag{4.26}$$

Similarly to the individual chance constraint case, we have:

$$\begin{aligned}
& \min_{z,y} c^T x \\
& \text{s.t.} \quad h_k - \bar{\xi}_k^T x \geq \sqrt{2 \ln \left(\frac{1}{1 - \alpha^{y_k}} \right)} \|\Sigma_k^{1/2} x\|, \quad k = 1, \dots, K, \\
& \quad \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad x \in X.
\end{aligned}$$

□

Similarly, problem (4.23) is not a convex optimization problem. Therefore, with Assumption 4.1, we apply the transformation $z = \ln x$ and get:

$$\begin{aligned}
& \min_{z,y} c^T e^z \\
& \text{s.t. } \bar{\xi}_k^T e^z + \left\| \sum_k^{1/2} e^{\ln\left(\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}\right)+z} \right\| \leq h_k \quad k = 1, \dots, K, \\
& \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad z \in Z,
\end{aligned} \tag{4.27}$$

Moreover, if $\bar{\xi}_k \geq 0$, $k = 1, 2, \dots, K$, and the function $\ln\left(\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}\right)$ is convex, then problem (4.27) is convex. The following lemma shows the convexity of $\ln\left(\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}\right)$.

Lemma 4.3. *If $\alpha \geq 1 - e^{-1} \approx 0.6321$, then $\ln\left(\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}\right)$ is convex with respect to y_k .*

Proof. By the convexity theorem of composite function, we only need to prove the convexity of $\ln\left(\sqrt{2\ln\left(\frac{1}{1-p}\right)}\right)$ with respect to p , since the convexity of composite function $\ln\left(\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}\right)$ is implied when $\ln\left(\sqrt{2\ln\left(\frac{1}{1-p}\right)}\right)$ is nondecreasing and convex. As $\ln\left(\sqrt{2\ln\left(\frac{1}{1-p}\right)}\right)$ is monotone, we need to show the convexity of $\ln\left(\sqrt{2\ln\left(\frac{1}{1-p}\right)}\right)$. We can notice that $\ln\left(\sqrt{2\ln\left(\frac{1}{1-p}\right)}\right) = \frac{1}{2} \ln\left(2\ln\left(\frac{1}{1-p}\right)\right)$. Therefore, we only need to focus on the convexity of $\ln\left(2\ln\left(\frac{1}{1-p}\right)\right)$.

The second order derivative of $\ln\left(2\ln\left(\frac{1}{1-p}\right)\right)$ can be written as

$$\begin{aligned}
& - \left[(1-p)^{-2} (\ln(1-p))^{-1} + (1-p)^{-2} (\ln(1-p))^{-2} \right], \\
& = - (1-p)^{-2} (\ln(1-p))^{-2} (\ln(1-p) + 1).
\end{aligned}$$

Then, $\ln\left(2\ln\left(\frac{1}{1-p}\right)\right)$ is convex if and only if

$$\ln(1-p) + 1 \leq 0.$$

Therefore, we have $p \geq 1 - e^{-1}$.

As α^{y_k} is convex with respect to y_k and $\alpha^{y_k} \geq \alpha$ for any $0 \leq y_k \leq 1$, if $\alpha \geq 1 - e^{-1}$, then the function $\ln\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}$ is convex. \square

Therefore, when $c \geq 0$, $\alpha \geq 1 - e^{-1}$, $\bar{\xi}_k \geq 0$, $k = 1, 2, \dots, K$, problem (4.27) is convex.

4.3 Bernstein and Hoeffding Bounds

Bernstein and Hoeffding bounds are considered as exponential type estimates of probabilities. These inequalities are frequently used for investigating the law of large numbers for instance. They are also often used in statistics and probability theory. In this section, we investigate these bounds for the case of individual and joint chance constraints.

4.3.1 Bernstein bounds

In this section, we consider *Berstein bound* [85]. We assume that the mean and the range parameters for all independent components ξ_i of the random vector ξ are known, i.e. $l_i \leq \xi_i \leq u_i$, and $\mathbb{E}(\xi_i) = \bar{\xi}_i$, for $i = 1, \dots, n$. Then, the respective values for the random variable $\xi_i x_i$ are $l'_i = l_i x_i$, $u'_i = u_i x_i$, and $\bar{\xi}'_i = \bar{\xi}_i x_i$. From Bernstein-type exponential estimate, we have that

$$e^{-g^* a} \prod_{i=1}^n \left\{ \frac{u_i - \bar{\xi}_i}{u_i - l_i} e^{g^* l_i x_i} + \frac{\bar{\xi}_i - l_i}{u_i - l_i} e^{g^* u_i x_i} \right\} \leq \alpha \quad (4.28)$$

with arbitrary constant $g^* > 0$ implies $\mathbb{P}(\sum_{i=1}^n \xi_i x_i \geq a) \leq \alpha$.

Theorem 4.6. *An upper bound for problem (4.2) can be obtained by solving the following problem*

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \sum_{i=1}^n \ln \left\{ \frac{u_i - \bar{\xi}_i}{u_i - l_i} e^{g^* l_i x_i} + \frac{\bar{\xi}_i - l_i}{u_i - l_i} e^{g^* u_i x_i} \right\} \leq \ln(1 - \alpha) + g^* h, \\ & x \in X \end{aligned} \quad (4.29)$$

with arbitrary $g^* > 0$.

Proof. Applying Bernstein inequality to the chance constraint in (4.2), and passing to the logarithm both sides, we obtain the conclusion. \square

We provide now an upper bound to problem (4.1) based on the *Berstein bound*. We assume that the mean and the range parameters for all independent components $(\xi_k)_i$ of the random vectors ξ_k are known, i.e. $(l_k)_i \leq (\xi_k)_i \leq (u_k)_i$, and $\mathbb{E}[(\xi_k)_i] = (\bar{\xi}_k)_i$, for $k = 1, \dots, K$ and $i = 1, \dots, n$. Then, the respective values for the random variable $(\xi_k)_i x_i$ are $(l'_k)_i = (l_k)_i x_i$, $(u'_k)_i = (u_k)_i x_i$, and $(\bar{\xi}'_k)_i = (\bar{\xi}_k)_i x_i$.

Theorem 4.7. *With the assumption of ξ mentioned above, an upper bound to*

problem (4.2) can be obtained by solving the following problem

$$\begin{aligned}
& \min_{x,y} c^T x \\
& \text{s.t.} \quad \sum_{i=1}^n \ln \left\{ \frac{(u_k)_i - (\bar{\xi}_k)_i}{(u_k)_i - (l_k)_i} e^{g_k^*(l_k)_i x_i} + \frac{(\bar{\xi}_k)_i - (l_k)_i}{(u_k)_i - (l_k)_i} e^{g_k^*(u_k)_i x_i} \right\} \leq g_k^* h_k + \\
& \quad \ln(1 - \alpha^{y_k}), \quad k = 1, \dots, K, \\
& \quad \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad x \in X.
\end{aligned} \tag{4.30}$$

with arbitrary $g^* > 0$.

Proof. According to Bernstein-type exponential estimate, we have

$$e^{-g_k^* a_k} \prod_{i=1}^n \left\{ \frac{(u_k)_i - (\bar{\xi}_k)_i}{(u_k)_i - (l_k)_i} e^{g_k^*(l_k)_i x_i} + \frac{(\bar{\xi}_k)_i - (l_k)_i}{(u_k)_i - (l_k)_i} e^{g_k^*(u_k)_i x_i} \right\} \leq \alpha^{y_k}, \quad k = 1, \dots, K, \tag{4.31}$$

with arbitrary $g_k^* > 0$ which implies $\mathbb{P}(\sum_{i=1}^n (\xi_k)_i x_i \geq a_k) \leq \alpha^{y_k}$, $k = 1, \dots, K$.

We note that

$$\mathbb{P}(\xi_k^T x \leq h_k) \geq \alpha^{y_k}, \quad k = 1, \dots, K, \tag{4.32}$$

\Downarrow

$$\mathbb{P}\left(\sum_{i=1}^n (\xi_k)_i x_i \geq h_k\right) \leq 1 - \alpha^{y_k}, \quad k = 1, \dots, K. \tag{4.33}$$

Constraints (4.33) can be approximated by

$$\sum_{i=1}^n \ln \left\{ \frac{(u_k)_i - (\bar{\xi}_k)_i}{(u_k)_i - (l_k)_i} e^{g_k^*(l_k)_i x_i} + \frac{(\bar{\xi}_k)_i - (l_k)_i}{(u_k)_i - (l_k)_i} e^{g_k^*(u_k)_i x_i} \right\} \leq \ln(1 - \alpha^{y_k}) + g_k^* h_k \tag{4.34}$$

for any $g_k^* > 0$, $k = 1, \dots, K$. □

From Proposition 4.1 in [85] and the concavity of function $\ln(1 - \alpha^{y_k})$, problem (4.30) is convex.

4.3.2 Hoeffding bounds

We consider now an approximation based on *Hoeffding inequality* ([85]):

$$\mathbb{P}\left(\frac{\xi^T e_n}{n} - \frac{\bar{\xi}^T e_n}{n} \geq h\right) \leq e^{\frac{-2n^2 h^2}{\sum_{i=1}^n (u_i - l_i)^2}}, \tag{4.35}$$

with the range parameters for all independent components ξ_i of the random vector ξ , i.e. $l_i \leq \xi_i \leq u_i$, for $i = 1, \dots, n$, $\bar{\xi} = \mathbb{E}(\xi)$ and $e_n \in \mathbb{R}^n$ is a vector with all elements equal to 1.

Theorem 4.8. *With the assumption of ξ mentioned above, an upper bound for problem (4.2) can be obtained by solving the following convex programming problem*

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \bar{\xi}^T x + \frac{\sqrt{2}}{2} \sqrt{-\ln(1-\alpha)} \|Mx\| \leq h, \\ & x \in X \end{aligned} \quad (4.36)$$

where $M = \text{diag}(u - l)$, $u = (u_1, \dots, u_n)^T$, $l = (l_1, \dots, l_n)^n$.

Proof. We note that

$$\mathbb{P}(\xi^T x \leq h) \geq \alpha, \quad (4.37)$$

$$\begin{aligned} & \Downarrow \\ \mathbb{P}(\xi^T x - \bar{\xi}^T x \geq h - \bar{\xi}^T x) & \leq 1 - \alpha. \end{aligned} \quad (4.38)$$

Then, we apply (4.35) to (4.38) and get:

$$\mathbb{P}(\xi^T x - \bar{\xi}^T x \geq h - \bar{\xi}^T x) \leq e^{\frac{-2(h - \bar{\xi}^T x)^2}{\sum_{i=1}^n (u_i - l_i)^2 x_i^2}}. \quad (4.39)$$

If

$$e^{\frac{-2(h - \bar{\xi}^T x)^2}{\sum_{i=1}^n (u_i - l_i)^2 x_i^2}} \leq 1 - \alpha, \quad (4.40)$$

then (4.37) will be satisfied. Logarithmic transformation of (4.40) leads to

$$\frac{-2(h - \bar{\xi}^T x)^2}{\sum_{i=1}^n (u_i - l_i)^2 x_i^2} \leq \ln(1 - \alpha), \quad (4.41)$$

which can be written as

$$h - \bar{\xi}^T x \geq \frac{\sqrt{2}}{2} \sqrt{-\ln(1-\alpha)} \|Mx\|, \quad (4.42)$$

where $M = \text{diag}(u - l)$ and then (4.42) is a convex inequality. \square

Theorem 4.9. *Assume that the mean and the range parameters for all independent components $(\xi_k)_i$ of the random vectors ξ_k are known, i.e. $(l_k)_i \leq (\xi_k)_i \leq (u_k)_i$, for $k = 1, \dots, K$ and $i = 1, \dots, n$. An approximation of problem (4.1) based on Hoeffding's inequality can be given by*

$$\begin{aligned} \min_{x,y} \quad & c^T x \\ \text{s.t.} \quad & \bar{\xi}_k^T x + \frac{\sqrt{2}}{2} \sqrt{\ln\left(\frac{1}{1-\alpha y_k}\right)} \|M_k x\| \leq h_k, \quad k = 1, \dots, K, \\ & \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad x \in X, \end{aligned} \quad (4.43)$$

where $M_k = \text{diag}(u_k - l_k)$, $u_k = ((u_k)_1, \dots, (u_k)_n)^T$, $l_k = ((l_k)_1, \dots, (l_k)_n)^T$, $k = 1, \dots, K$.

Proof. With almost the same proof method of Theorem 4.8, the conclusion can be obtained. \square

Additionally, with Assumption 4.1, an equivalent *upper bound* for problem (4.1) based on Hoeffding inequality can be obtained by applying the following transformation $z = \ln x$:

$$\begin{aligned} \min_{z, y} \quad & c^T e^z \\ \text{s.t.} \quad & \xi_k^T e^z + \frac{1}{2} \|M_k e^{\ln\left(\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}\right)+z}\| \leq h_k, \quad k = 1, \dots, K, \\ & \sum_{k=1}^K y_k = 1, \quad y_k \geq 0, \quad k = 1, \dots, K, \quad z \in Z. \end{aligned} \quad (4.44)$$

From Lemma 4.3, function $\ln\left(\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}\right)$ is convex, when $\alpha \geq 1 - e^{-1}$. Hence, if $c \geq 0$, $\alpha \geq 1 - e^{-1}$, problem (4.44) is convex.

4.4 Computational Results

Although the problems which give upper bounds obtained by Chebyshev inequality, Chernoff inequality, Bernstein inequality and Hoeffding inequality, respectively, for problem (4.1) are convex under some conditions, they are still hard to solve directly by current tools because of the following terms: $\ln\left(\sqrt{\frac{\alpha^{y_k}}{1-\alpha^{y_k}}}\right)$ and $\ln\left(\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}\right)$.

In the following, we denote these functions by $\Upsilon(y_k)$. We propose piecewise linear approximations for $\Upsilon(y_k)$ based on tangent and segment approximations.

Tangent approximation

We choose S different linear functions:

$$l_s(y_k) = a_s y_k + b_s, \quad s = 1, \dots, S,$$

such that

$$l_s(y_k) \leq \Upsilon(y_k), \quad \forall y_k \in [\rho, 1), \quad k = 1, \dots, K.$$

Here $\rho \geq 0$ is a constant such that $\Upsilon(y_k)$ is convex on $[\rho, 1)$. Then, $\Upsilon(y_k)$ can be approximated by the following piecewise linear function

$$l(y_k) = \max_{s=1, \dots, S} l_s(y_k),$$

which provides a lower approximation for $\Upsilon(y_k)$.

In order to achieve the expected precision, we set $l_s(y_k)$ as the tangent line of $\Upsilon(y_k)$ at S points τ_1, \dots, τ_S with $\tau_s \in [\rho, 1)$, $s = 1, \dots, S$. Then, we have

$$a_s = \left. \frac{d\Upsilon(y_k)}{dy_k} \right|_{y_k=\tau_s}, \quad b_s = \Upsilon(\tau_s) - a_s \tau_s.$$

Thanks to these piecewise linear approximations for $\Upsilon(y_k)$, we have the following results:

Theorem 4.10. *Under the aforementioned convexity conditions, if we replace in problems (4.14), (4.27), (4.44) $\Upsilon(y_k)$ by $l(y_k)$, we obtain their convex approximations. The optimum values of the approximation problems are lower bounds for problems (4.14), (4.27), (4.44), respectively. Moreover, the approximation problems become an equivalent reformulation of problems (4.14), (4.27), (4.44) when S goes to infinity.*

Proof. As the approximation problems are obtained by relaxing some constraints in problems (4.14), (4.27), (4.44), it is easy to see that the optimal values of the approximation problems are lower bounds for problems (4.14), (4.27), (4.44), respectively.

We know under convexity conditions for problems (4.14), (4.27), (4.44), $\Upsilon(y_k)$ is convex for each problem. As the S tangent functions are selected differently, when S goes to infinity, the constraints in the approximation problems are equivalent to the constraints in problems (4.14), (4.27), (4.44), respectively. As the original problems and the corresponding approximation problems are all convex programs, the approximation problems become an equivalent reformulation of problems (4.14), (4.27), (4.44), respectively, when S goes to infinity. \square

Segment approximation

In order to come up with conservative bounds for the optimum values of problems (4.14), (4.27), (4.44), we use the linear segments $\bar{a}_s y_k + \bar{b}_s$, $s = 1, \dots, S$, between $\tau_1, \tau_2, \dots, \tau_{S+1} \in [\rho, 1)$ to construct a piecewise linear function

$$\bar{l}(y_k) = \max_{s=1, \dots, S} \{ \bar{a}_s y_k + \bar{b}_s \}, \quad (4.45)$$

where

$$\bar{a}_s = \frac{\Upsilon(\tau_{s+1}) - \Upsilon(\tau_s)}{\tau_{s+1} - \tau_s}, \quad \bar{b}_s = \Upsilon(\tau_s) - \bar{a}_s \tau_s, \quad s = 1, \dots, S.$$

Using the piecewise linear function $\bar{l}(y_k)$ to replace $\Upsilon(y_k)$ in problems (4.14), (4.27), (4.44), gives the corresponding approximation problems.

Similar to Theorem 4.10, we can derive the following result for the linear approximation:

Theorem 4.11. *Under the aforementioned convexity conditions, if we replace in problems (4.14), (4.27), (4.44) $\Upsilon(y_k)$ by $\bar{l}(y_k)$, we obtain the convex approximations of these problems. The optimum values of the approximation problems are an upper bound for problems (4.14), (4.27), (4.44), respectively. Moreover, the approximation problems become an equivalent reformulation of problems (4.14), (4.27), (4.44), respectively, when S goes to infinity.*

The proof of this theorem follows the same pattern as the proof of Theorem 4.10.

Numerical experiments

The four bounds have been implemented and compared under Matlab environment using CVX software, a modeling system for constructing and solving convex programs. We run the bounds for 100 instances randomly generated with the following characteristics: in the individual chance constraint (4.2) we set $n = 10$, $h = 0.5$, $\alpha = 0.95$, the constraint $x \in X \in \mathbb{R}_+^n$ is given by $\sum_{i=1}^n x_i = 1$, c is a random vector from a uniform distribution in the interval $[0, 1]^n$, ξ^T is uniformly generated in the interval $[0, 10]^n$ and σ_ξ^2 is uniformly generated in the interval $[0, 1]^{n \times n}$. In the following, we assume that the random variable ξ is distributed according to a normal distribution with mean $\bar{\xi}$ and variance σ_ξ^2 generated as described above. This will allow us to make a fair comparison of the bounds with the exact SOCP reformulation. For joint chance constrained problem(4.1), we set $n = 10$, $K = 5$ and $h_i = 0.5$, $i = 1, \dots, K$. The other parameters are the same as those in the individual case.

As upper and lower bounds for random vector ξ are needed in Bernstein and Hoeffding bounds, 5000 samples following a normal distribution with mean $\bar{\xi}$ and variance σ_ξ^2 described above are generated. And the maximal value of these 5000 samples is selected as upper bound while the minimal value is selected as lower bound

Numerical results for the individual chance constraint case are reported in Figure 4.1: Figure 4.1(a) shows a comparison of the objective function values of the four bounds (Chebyshev, Chernoff, Bernstein and Hoeffding) with the exact SOCP reformulation for 100 different randomly generated instances, while Figure 4.1(b) shows the corresponding box-and-whisker plots where the extrema of the box represents the 1/4 and 3/4 quartiles, the band inside the box is the median and the whiskers the minimum and maximum obtained from 100 randomly generated instances.

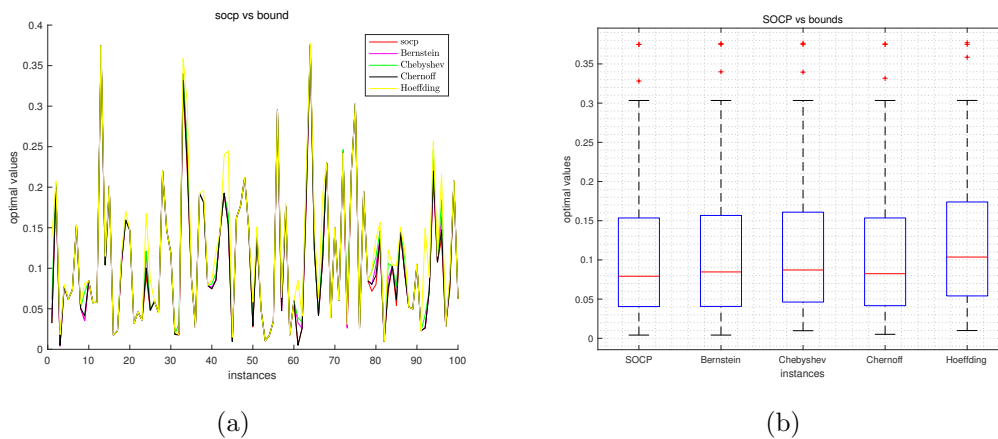


Figure 4.1: (a) Comparison of SOCP versus bounds for individual chance constraint. (b) box-and-whisker plots from the same 100 randomly generated instances of (a).

The average CPU times and gaps for different bounds and SOCP are shown in Table 4.1 in second and third rows, respectively. Results from Figure 4.1 and

Table 4.1 show that the best upper bound is obtained by Chernoff inequality, followed by Bernstein which in all the instances outperforms one-side Chebyshev and Hoeffding ones. The worst bound is in all the cases obtained by Hoeffding approximation. Notice that the CPU times for Chebyshev, Chernoff, Bernstein, Hoeffding and SOCP are relatively comparable. The reason for the good performance of Chernoff bound compared to the others, could be due to the explicit assumption that the moment generating function of the random variable ξ follows a normal distribution, information not taken into account in all the other approaches. Very good is the performance of the Bernstein bound, considering that only the mean and the range of the random variable are known.

Table 4.1: Average results over 100 instances of different bounds and SOCP for individual chance constrained problem.

Bound	Chebyshev	Chernoff	Bernstein	Hoeffding	SOCP
CPU time(s)	0.9815	0.9727	1.2761	1.0850	0.9932
Gap (%)	7.55	1.58	2.8	13.64	-

Numerical results for the joint chance constraints case are reported in Figure 4.2: Figure 4.2(a) shows a comparison of the objective function values of the four bounds (Chebyshev, Chernoff, Bernstein and Hoeffding) with the SOCP reformulation for 100 different randomly generated instances, where SOCP, Hoeffding and Chernoff are obtained both with tangent and segment approximations with $S = 50$. Figure 4.2(b) shows the corresponding box-and-whisker plots where the extrema of the box represents the 1/4 and 3/4 quartiles, the band inside the box is the median and the whiskers the minimum and maximum obtained from 100 randomly generated instances.

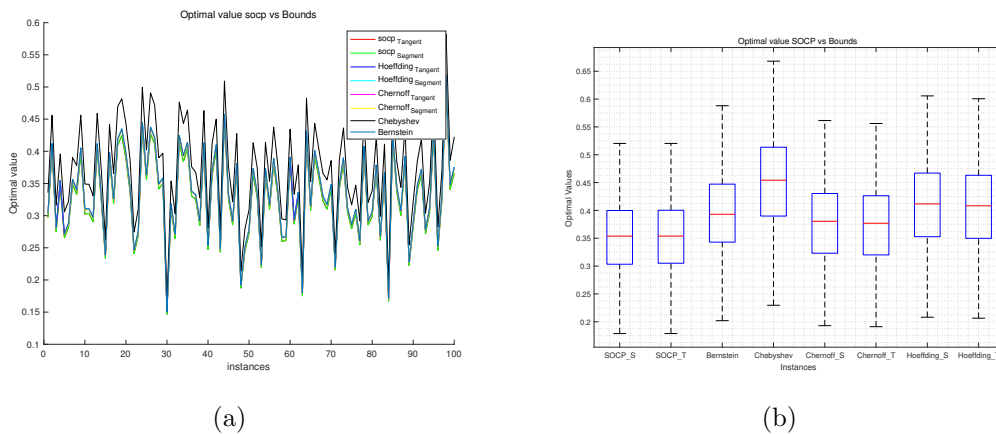


Figure 4.2: (a) Comparison of SOCP versus bounds for the joint chance constrained problem. (b) Box-and-whisker plots from the same 100 randomly generated instances of (a).

Table 4.2 shows the average CPU time (second row) for different bounds and SOCP and the average gap (third row) between different bounds and SOCP for

the joint chance constraints case. Results from Figure 4.2 and Table 4.2 show that Chebyshev bound always provide the worst bound, while there are not much differences among other bounds. As in the individual chance constraint case, the best bound is the Chernoff one followed by the Bernstein's bound. Notice that the CPU times for Chebyshev, Chernoff, Bernstein, Hoeffding and SOCP are relatively comparable.

Table 4.2: Average results of different bounds and SOCP for the joint chance constrained problem.

Bound	Chebyshev	Chernoff	Bernstein	Hoeffding	SOCP
CPU time(s)	2.9932	2.8255	3.2030	2.2845	2.2551
Gap (%)	12.30	3.69	3.91	4.01	-

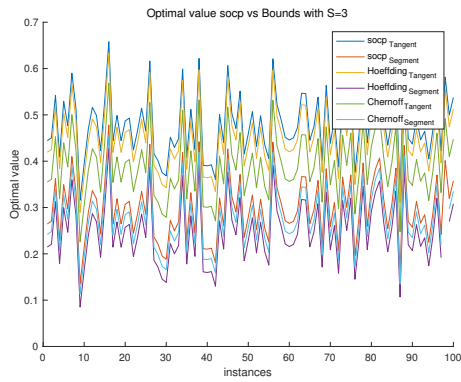
Notice that in the upper bound problems obtained by Chernoff and Hoeffding inequalities for problem (1), we use tangent and segment approximations, as described above, to approximate the complex terms: $\ln\left(\sqrt{2\ln\left(\frac{1}{1-\alpha^{y_k}}\right)}\right)$ and $\ln\left(\sqrt{2\ln\left(\frac{1}{1-y_k}\right)}\right)$, respectively.

A sensitivity analysis of the segment and tangent approximations on the number of segments $S = 3, 10, 20, 50$ is reported in Figure 4.3 with corresponding box-and-whisker plots in Figure 4, for 100 instances randomly generated. Results show that when $S = 50$, there's almost no gap between the results obtained by tangent and segment approximations, respectively. And for each S , Chernoff inequality and Hoeffding inequality always provide an upper bound for the result obtained by SOCP.

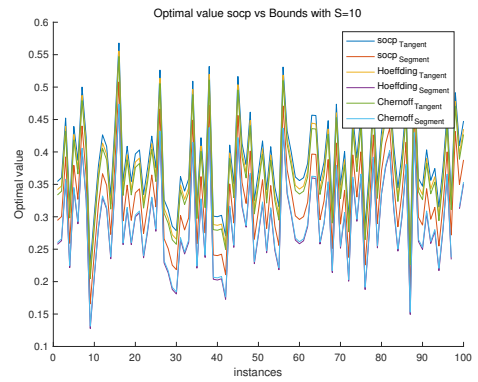
4.5 Conclusion

In this chapter, we propose deterministic approximations for individual and joint chance constraints with independent matrix vector rows. The bounds are based on classical inequalities from probability theory such as the one-side Chebyshev inequality, Bernstein inequality, Chernoff inequality and Hoeffding inequality and allow to reformulate the problem in in a tractable convex way. Approximations based on piecewise linear and tangent are also provided in case of Chernoff and Hoeffding inequalities allowing to reduce the computational complexity of the problem. Finally numerical results on randomly generated data are provided allowing to identify that the Chernoff bound provides the tighter deterministic approximation while the Chebyshev bound, requiring the knowledge of the first and second moments, is very loose both for individual and joint chance constrained problems. Remarkable is also the performance of Bernstein's bound, considering that only the mean and the range of the random variables are assumed to be known. In terms of CPU times all the considered bounds are relatively comparable.

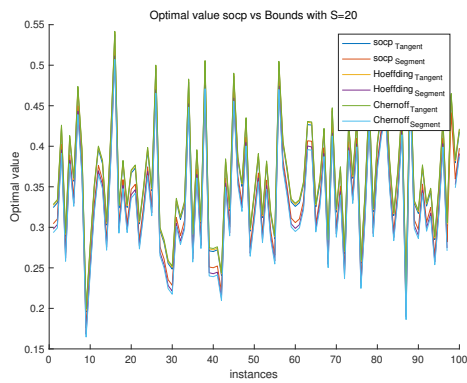
This chapter correpnds to the reference [81].



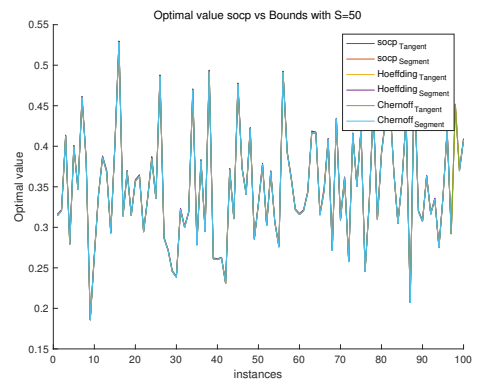
(a)



(b)

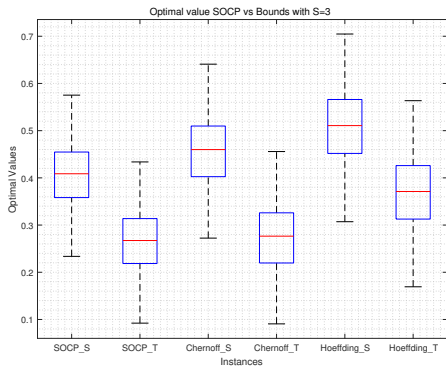


(c)

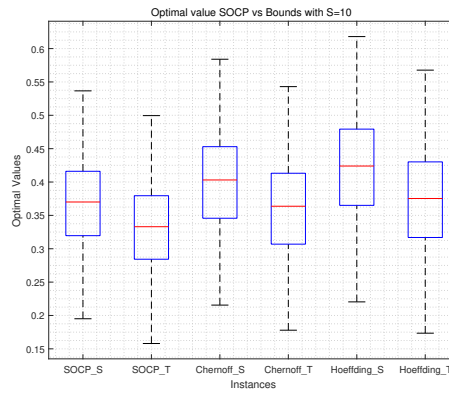


(d)

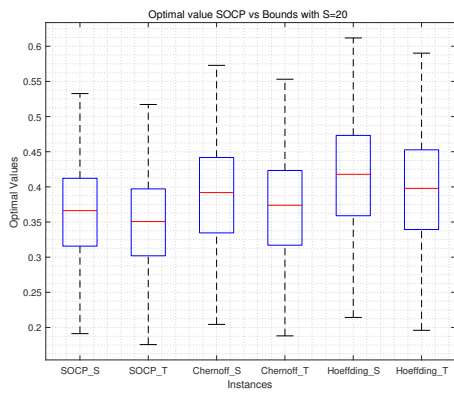
Figure 4.3: Comparison of SOCP versus bounds for the joint chance constrained problem with different values of S for segment and tangent approximations.



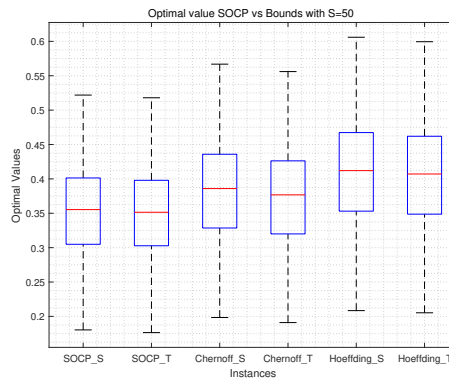
(a)



(b)



(c)



(d)

Figure 4.4: Box-and-whisker plots corresponding to Figure 4.3 for different values of $S = 3, 10, 20, 50$ with segment and tangent approximations.

Chapter 5

Data-Driven Robust Chance Constrained Problems: A Mixture Model Approach

This chapter discusses the mixture distribution based data-driven robust chance constrained problem. We construct a data-driven mixture distribution based uncertainty set from the perspective of simultaneously estimating higher order moments. Then, we derive a reformulation of the data-driven robust chance constrained problem. As the reformulation is not a convex programming, we propose new and tight convex approximations based on the piecewise linear approximation method under certain conditions. For the general case, we propose a DC approximation to derive an upper bound and a relaxed convex approximation to derive a lower bound for the optimal value of the original problem, respectively. We establish the theoretical foundation for these approximations. Finally, the results of simulation experiments show that the proposed approximations are practical and efficient.

5.1 Mixture Distribution Based Uncertainty Set

As a preparation for our later construction of the data-driven robust chance constrained problem, we first define the uncertainty set under the mixture distribution framework, and then consider the moments estimation from sample data to determine the parameters of the data-driven uncertainty set.

We assume that the n -dimensional random vector $\boldsymbol{\xi}$ has a continuous probability density function. For notational clarity, we use f to denote the distribution of $\boldsymbol{\xi}$ and $f(\cdot)$ to denote the probability density function of $\boldsymbol{\xi}$.

To ensure a reliable estimation of the real distribution, we assume that $\boldsymbol{\xi}$ follows a finite normal mixture distribution. This means that $f(\cdot)$ can be explicitly defined by

$$f(\cdot) = \sum_{m=1}^M w_m f_m(\cdot), \quad (5.1)$$

here $f_m(\cdot)$ is the probability density function of the m -th component. In this

chapter, we assume that the m -th component follows a n -dimensional normal distribution with the mean vector being $\boldsymbol{\mu}_m = (\mu_{m1}, \dots, \mu_{mn})^\top$ and the covariance matrix being $\Sigma_m = \{\sigma_{ij}^m\}$, which is positive semidefinite, $m = 1, \dots, M$. $w_m \in \mathbb{R}_+$ is the weight assigned to the m -th component, $m = 1, \dots, M$, with $\sum_{m=1}^M w_m = 1$.

As the exact distribution of the random vector is never completely known in practice, it is reasonable to assume certain kind of uncertainties when modeling the distribution. One advantage of the normal mixture distribution is that it can fairly represent different types of distributions with respective means and variances. Many popular distributions, such as elliptical distribution, skew-t distribution and generalized hyperbolic distribution, can be viewed as (infinite) normal mixture distributions (see McNeil et al.[74]). Considering both the description accuracy and the tractability, we assume in the rest of this chapter that all the component distributions of the assumed mixture distribution are completely specified. Therefore, establishing a confidence region of the original distribution is equivalent to constructing a confidence region of the mixture weights. For this reason, we introduce the following data-driven uncertainty set:

Definition 5.1. *A mixture distribution based uncertainty set is expressed as*

$$D = \left\{ \sum_{m=1}^M w_m f_m(\cdot) : \mathbf{w} \in W \right\},$$

where $W = \left\{ \mathbf{w} : \|U_k \mathbf{w}\| \leq \gamma_k, k = 1, \dots, I, \sum_{m=1}^M w_m = 1, \mathbf{w} \geq 0 \right\}$. I is an positive integer. For $k = 1, \dots, I$, the constraint $\|U_k \mathbf{w}\| \leq \gamma_k$ corresponds to the k -th order moment constraint, $\gamma_k \in \mathbb{R}_{++}$ is a given reference scalar parameter, $U_k \in \mathbb{R}^{n \times M}$ is the reference matrix given as

$$U_k = \left(\bar{\boldsymbol{\mu}}^{(k)} - \boldsymbol{\mu}_1^{(k)}, \dots, \bar{\boldsymbol{\mu}}^{(k)} - \boldsymbol{\mu}_M^{(k)} \right) \in \mathbb{R}^{n \times M},$$

where $\boldsymbol{\mu}_m^{(k)} = \left(\mu_{m,1}^{(k)}, \dots, \mu_{m,n}^{(k)} \right)^\top$, $\mu_{m,j}^{(k)} = \int \xi_j^k f_m(\boldsymbol{\xi}) d\boldsymbol{\xi}$, $j = 1, \dots, n$, $m = 1, \dots, M$, and $\bar{\boldsymbol{\mu}}^{(k)}$ denotes the empirical estimation of $\sum_{m=1}^M w_m \boldsymbol{\mu}_m^{(k)}$.

Mixture distribution has already been studied in robust statistics and used in modeling the distribution of financial data, see, for example, Hall et al. [45], Peel and McLachlan [80] and Zhu and Fukushima [107].

The above uncertainty set is based on the moment estimation. It can incorporate the information of the arbitrary order moments of $\boldsymbol{\xi}$, which is different from the uncertainty set in Zhu et al.[106] and Gupta [43], where the uncertainty sets were defined with only first two order moments constraints, or specified from the viewpoint of Bayesian learning. And our uncertainty set focuses on the moments and includes higher order moments, which are often concerned in practical problems, like financial management.

Mixture distribution can efficiently describe the shape of a distribution. Eisenberger [34] showed that the density of a mixture of two univariate normal components with means and variances being μ_1, σ_1^2 and μ_2, σ_2^2 , respectively, is unimodal

if $(\mu_1 - \mu_2)^2 < \frac{27\sigma_1^2\sigma_2^2}{4(\sigma_1^2 + \sigma_2^2)}$. For the case that $(\mu_1 - \mu_2)^2 > \frac{27\sigma_1^2\sigma_2^2}{4(\sigma_1^2 + \sigma_2^2)}$, there exist values of the weights for which the density function is bimodal. For the unimodal case, first two order moments can describe the location and dispersion, while the density shape can be described through concepts like skewness and kurtosis [93]. For the multimodal case, a probability distribution can be characterized by its enough moments [73]. Then, the multimodality can be described efficiently with enough moments. For the above reasons, we use the mixture distribution with uncertain weights to construct the uncertainty set and use moments to estimate the size of uncertainty set. Our approach is a complementary to the existing uncertainty sets which focus on the unimodal case [63].

We will show that the parameters $U_k, \gamma_k, k = 1, \dots, I$, of the uncertainty set D can be determined by a set of independently observed samples from the underlying distribution of $\boldsymbol{\xi}$, and the confidence region defined by $U_k, \gamma_k, k = 1, \dots, I$, is assured with high probability to contain the first four componentwise moments of $\boldsymbol{\xi}$.

5.1.1 Data-Driven Confidence Region for Moments

In order to get the confidence region for moments of $\boldsymbol{\xi}$, we use a measurable function $\pi(\boldsymbol{\xi})$ to denote a vector valued or matrix valued mapping of $\boldsymbol{\xi}$, while the norm $\|\cdot\|$ with respect to $\pi(\boldsymbol{\xi})$ is the Euclidean norm when $\pi(\boldsymbol{\xi})$ is vector valued, and is the Frobenius norm when $\pi(\boldsymbol{\xi})$ is matrix valued. Suppose that we have a sample set S of $\boldsymbol{\xi}$ with N observations $\boldsymbol{\xi}^{(l)}, l = 1, \dots, N$. Then we denote $\bar{\pi}_S = \frac{1}{N} \sum_{l=1}^N \pi(\boldsymbol{\xi}^{(l)})$.

To derive the basic conclusion of this subsection, we utilize an inequality known as the ‘‘independent bounded differences inequality’’, which is established by McDiarmid in [44].

Lemma 5.1. *Let $\mathbf{X} = (X_1, \dots, X_n)$ be a family of independent random variables with X_k taking values in a set A_k for each $k, 1 \leq k \leq n$. Suppose that the real-valued function Γ defined on $\prod_{k=1}^n A_k$ satisfies*

$$|\Gamma(\mathbf{x}) - \Gamma(\mathbf{x}')| \leq c_k, k = 1, \dots, n,$$

whenever the vectors \mathbf{x} and \mathbf{x}' differ only in the k -th coordinate. Then for any $\tau \geq 0$,

$$\mathbb{P}\{\Gamma(\mathbf{X}) - \mathbb{E}[\Gamma(\mathbf{X})] \geq \tau\} \leq \exp\left(\frac{-2\tau^2}{\sum_{k=1}^n c_k^2}\right).$$

With this lemma, we have the following theorem for chosen $\epsilon \in]0, 1[$.

Theorem 5.1. *Let S be a sample set with N observations which are generated independently at random according to the distribution of $\boldsymbol{\xi}$ whose density function is defined by (5.1), and suppose that there exists an $R < \infty$ such that $\mathbb{P}\{\|\pi(\boldsymbol{\xi})\| \leq R\} \geq 1 - \epsilon$. Then with probability at least $1 - \Delta$, we have*

$$\|\bar{\pi}_S - \mathbb{E}_{\boldsymbol{\xi}}[\pi(\boldsymbol{\xi})]\| \leq \frac{1}{\sqrt{N}} \left(2\rho_M + R\sqrt{2 \ln \frac{1-\epsilon}{\Delta-\epsilon}} \right), \quad (5.2)$$

where $\epsilon < \Delta < 1$, $\rho_M^2 = \max_{m=1, \dots, M} \left\{ \int \|\pi(\boldsymbol{\xi})\|^2 f_m(\boldsymbol{\xi}) d\boldsymbol{\xi} \right\}$ and $\rho_M \geq 0$.

Proof. Let $g(S) = \|\bar{\pi}_S - \mathbb{E}_{\boldsymbol{\xi}}[\pi(\boldsymbol{\xi})]\|$ denote the measure of the estimation accuracy based on the sample S .

By Lemma 2.3.1 in [100], we have

$$\mathbb{E}_S[g(S)] = \mathbb{E}_S[\|\bar{\pi}_S - \mathbb{E}_{\boldsymbol{\xi}}[\pi(\boldsymbol{\xi})]\|] \leq 2\mathbb{E}_{\mathbf{r}, S}[\|\frac{1}{N} \sum_{l=1}^N r_l \pi(\boldsymbol{\xi}^{(l)})\|],$$

where $\mathbf{r} = (r_1, \dots, r_N)^\top$ and $r_l, l = 1, \dots, N$, are Rademacher variables which have 50% chance of being +1 or -1 and are independent of S . It is then known from the Jensen's inequality that

$$\begin{aligned} \mathbb{E}_{\mathbf{r}, S}[\|\frac{1}{N} \sum_{l=1}^N r_l \pi(\boldsymbol{\xi}^{(l)})\|] &= \frac{1}{N} \mathbb{E}_{\mathbf{r}, S}[\|\sum_{l=1}^N r_l \pi(\boldsymbol{\xi}^{(l)})\|] \\ &= \frac{1}{N} \mathbb{E}_{\mathbf{r}, S} \left[\left(\left\langle \sum_{l=1}^N r_l \pi(\boldsymbol{\xi}^{(l)}), \sum_{\hat{l}=1}^N r_{\hat{l}} \pi(\boldsymbol{\xi}^{(\hat{l})}) \right\rangle \right)^{1/2} \right] \\ &\leq \frac{1}{N} \left(\mathbb{E}_{\mathbf{r}, S} \left[\sum_{l, \hat{l}=1}^N r_l r_{\hat{l}} \langle \pi(\boldsymbol{\xi}^{(l)}), \pi(\boldsymbol{\xi}^{(\hat{l})}) \rangle \right] \right)^{1/2}, \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product with respect to $\pi(\boldsymbol{\xi})$. When $\pi(\boldsymbol{\xi})$ is a vector, $\langle \pi(\boldsymbol{\xi}^{(l)}), \pi(\boldsymbol{\xi}^{(\hat{l})}) \rangle = (\pi(\boldsymbol{\xi}^{(l)}))^\top \pi(\boldsymbol{\xi}^{(\hat{l})})$. Otherwise, $\langle \cdot, \cdot \rangle$ denotes the Frobenius inner product. Thus, by utilizing the fact that $r_l r_{\hat{l}}, l \neq \hat{l}$, takes the two possible values $\{+1, -1\}$ with equal probabilities, we have

$$\left(\mathbb{E}_{\mathbf{r}, S} \left[\sum_{l, \hat{l}=1}^N r_l r_{\hat{l}} \langle \pi(\boldsymbol{\xi}^{(l)}), \pi(\boldsymbol{\xi}^{(\hat{l})}) \rangle \right] \right)^{1/2} = \left(\mathbb{E}_S \left[\sum_{l=1}^N \|\pi(\boldsymbol{\xi}^{(l)})\|^2 \right] \right)^{1/2}.$$

Therefore,

$$\begin{aligned} \mathbb{E}_S[g(S)] &\leq \frac{2}{N} \left(\mathbb{E}_S \left[\sum_{l=1}^N \|\pi(\boldsymbol{\xi}^{(l)})\|^2 \right] \right)^{1/2} \\ &= \frac{2}{N} (N \mathbb{E}_{\boldsymbol{\xi}} [\|\pi(\boldsymbol{\xi})\|^2])^{1/2} = \frac{2}{\sqrt{N}} (\mathbb{E}_{\boldsymbol{\xi}} [\|\pi(\boldsymbol{\xi})\|^2])^{1/2}, \end{aligned}$$

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\xi}} [\|\pi(\boldsymbol{\xi})\|^2] &= \int \sum_{m=1}^M \|\pi(\boldsymbol{\xi})\|^2 w_m f_m(\boldsymbol{\xi}) d\boldsymbol{\xi} \\ &= \sum_{m=1}^M w_m \int \|\pi(\boldsymbol{\xi})\|^2 f_m(\boldsymbol{\xi}) d\boldsymbol{\xi} = \sum_{m=1}^M w_m \rho_m^2. \end{aligned}$$

Here, $\rho_m^2 = \int \|\pi(\boldsymbol{\xi})\|^2 f_m(\boldsymbol{\xi}) d\boldsymbol{\xi}$ and $\rho_m \geq 0$. Let $\rho_M^2 = \max_{m=1, \dots, M} \{\rho_m^2\}$. Then, we have $\mathbb{E}_{\boldsymbol{\xi}} [\|\pi(\boldsymbol{\xi})\|^2] \leq \rho_M^2$ and $\mathbb{E}_S[g(S)] \leq \frac{2\rho_M}{\sqrt{N}}$.

On the other hand, we can obtain a sample \hat{S} by replacing the $\boldsymbol{\xi}^{(l)}$ of the sample S with $\hat{\boldsymbol{\xi}}^{(l)}$ for some $l \in \{1, 2, \dots, N\}$. Then we have

$$\begin{aligned} |g(S) - g(\hat{S})| &= \left| \|\bar{\pi}_S - \mathbb{E}_{\boldsymbol{\xi}}[\pi(\boldsymbol{\xi})]\| - \|\bar{\pi}_{\hat{S}} - \mathbb{E}_{\boldsymbol{\xi}}[\pi(\boldsymbol{\xi})]\| \right| \\ &\leq \|\bar{\pi}_S - \bar{\pi}_{\hat{S}}\| = \frac{1}{N} \|\pi(\boldsymbol{\xi}^{(l)}) - \pi(\hat{\boldsymbol{\xi}}^{(l)})\|. \end{aligned} \quad (5.3)$$

We get from (5.3) the following inequality with probability at least $1 - \epsilon$,

$$|g(S) - g(\hat{S})| \leq \frac{2R}{N},$$

since $\mathbb{P}\{\|\pi(\boldsymbol{\xi})\| \leq R\} \geq 1 - \epsilon$.

Consequently, we have for any $\tau > 0$ that

$$\begin{aligned} &\mathbb{P}\left\{g(S) \leq \tau + \frac{2\rho_M}{\sqrt{N}}\right\} \\ &\geq \mathbb{P}\left\{g(S) \leq \tau + \frac{2\rho_M}{\sqrt{N}} \mid |g(S) - g(\hat{S})| \leq \frac{2R}{N}\right\} \cdot \mathbb{P}\left\{|g(S) - g(\hat{S})| \leq \frac{2R}{N}\right\} \\ &\geq \mathbb{P}\left\{g(S) - \mathbb{E}_S[g(S)] \leq \tau \mid |g(S) - g(\hat{S})| \leq \frac{2R}{N}\right\} (1 - \epsilon) \\ &\geq (1 - \exp\left(-\frac{2N\tau^2}{4R^2}\right))(1 - \epsilon). \end{aligned}$$

The last inequality is deduced from Lemma 5.1 with $c_k = \frac{2R}{N}$, $k = 1, \dots, N$. Let $\Delta = 1 - (1 - \exp\left(-\frac{2N\tau^2}{4R^2}\right))(1 - \epsilon)$, from which we have $\tau = \frac{R}{\sqrt{N}} \sqrt{2 \ln \frac{1 - \epsilon}{\Delta - \epsilon}}$, and thus the inequality (5.2) holds with probability at least $1 - \Delta$. \square

With respect to the sample S , we denote, for $k = 1, \dots, I$,

$$\begin{aligned} \pi^{(k)}(\boldsymbol{\xi}) &= (\xi_1^k, \dots, \xi_n^k)^\top, \quad \boldsymbol{\mu}^{(k)} = \mathbb{E}_{\boldsymbol{\xi}}[\pi^{(k)}(\boldsymbol{\xi})], \quad \bar{\boldsymbol{\mu}}_S^{(k)} = \frac{1}{N} \sum_{l=1}^N \pi^{(k)}(\boldsymbol{\xi}^{(l)}), \\ \rho_M^{(k)} &= \max_{m=1, \dots, M} \left\{ \int \left(\sum_{j=1}^n \xi_j^{2k} \right) f_m(\boldsymbol{\xi}) d\boldsymbol{\xi} \right\} \\ &= \max_{m=1, \dots, M} \left\{ \sum_{j=1}^n \int \xi_j^{2k} f_m(\boldsymbol{\xi}) d\boldsymbol{\xi} \right\}. \end{aligned} \quad (5.4)$$

Then, we can derive the following corollary from Theorem 5.1.

Corollary 5.1. *Under the same assumptions as those in Theorem 5.1, for $k = 1, \dots, I$, the following inequality holds with probability at least $1 - \Delta$:*

$$\|\bar{\boldsymbol{\mu}}_S^{(k)} - \boldsymbol{\mu}^{(k)}\| \leq \frac{1}{\sqrt{N}} \left(2\rho_M^{(k)} + R^k \sqrt{2 \ln \frac{1 - \epsilon}{\Delta - \epsilon}} \right), \quad (5.5)$$

where $\rho_M^{(k)}$, $1 \leq k \leq I$, are defined in (5.4) and Δ is the same as that in Theorem 5.1.

Proof. For $k = 1, \dots, I$, consider the mapping $\tilde{\pi} : \boldsymbol{\xi} \mapsto \pi^{(k)}(\boldsymbol{\xi})$. Then we have

$$\|\tilde{\pi}(\boldsymbol{\xi})\| = \sqrt{\sum_{m=1}^n \xi_m^{2k}} \leq \sqrt{\left(\sum_{m=1}^n \xi_m^2\right)^k} = (\|\boldsymbol{\xi}\|)^k.$$

Therefore, by applying Theorem 5.1 to the mapping $\tilde{\pi}$, the inequality (5.5) follows since $\mathbb{P}\{\|\tilde{\pi}(\boldsymbol{\xi})\| \leq R^k\} \geq \mathbb{P}\{\|\boldsymbol{\xi}\| \leq R\} \geq 1 - \epsilon$. \square

The determination of parameters $\rho_M^{(k)}$, $1 \leq k \leq I$, in Theorem 5.1 amounts to calculating $\int \xi_j^{2k} f_m(\boldsymbol{\xi}) d\boldsymbol{\xi}$, $j = 1, \dots, n$, $m = 1, \dots, M$, which are actually the $2k$ -th order origin moments of the j -th random variable whose density function is $f_m(\cdot)$. For $m = 1, \dots, M$, as $f_m(\cdot)$ is the density function of a given n -dimensional normal distribution with the mean vector being $\boldsymbol{\mu}_m$ and the covariance matrix being Σ_m , $\int \xi_j^{2k} f_m(\boldsymbol{\xi}) d\boldsymbol{\xi}$, $j = 1, \dots, n$, can be easily computed with the help of the moment generation function of normal distribution.

All the above derivations assume that one can pre-specify a ball which contains the support with probability at least $1 - \epsilon$. To make the results more tractable, we need to estimate the radius R from the samples. Concretely, we have the following conclusion.

Corollary 5.2. *Let S be an N sample set, with observations $\boldsymbol{\xi}^{(l)}$, $l = 1, \dots, N$, generated independently at random according to the distribution of $\boldsymbol{\xi}$ whose density function is defined by (5.1), and the support of $\boldsymbol{\xi}$ is contained in a ball with probability at least $1 - \epsilon$. Let $\bar{\boldsymbol{\mu}}_S^{(k)} = \frac{1}{N} \sum_{l=1}^N \pi^{(k)}(\boldsymbol{\xi}^{(l)})$ and*

$$\bar{R} = \max \{ \|\boldsymbol{\xi}^{(l)}\|, l = 1, \dots, N \}$$

be a sample-based estimation for the radius R of the ball. If

$$N \geq \frac{2}{\epsilon} \ln \frac{1}{\Delta} + 2n + \frac{2n}{\epsilon} \ln \frac{2}{\epsilon},$$

then with probability at least $(1 - \Delta)^2$, the inequality (5.5) holds with R replaced with \bar{R} .

Proof. Consider the following optimization problem:

$$\min_R R \quad \text{s.t.} \quad \mathbb{P}\{\|\boldsymbol{\xi}\| \leq R\} \geq 1 - \epsilon.$$

We use the scenario approach to bound the above chance constraint such that it holds with probability $1 - \Delta$. Specifically, the scenario approach replaces the chance constraint with the following deterministic constraints:

$$\|\boldsymbol{\xi}^{(l)}\| \leq R, \quad l = 1, \dots, N. \quad (5.6)$$

This means that we have $\bar{R} = \max \{ \|\boldsymbol{\xi}^{(l)}\|, l = 1, \dots, N \}$ to be the optimum value under constraints (5.6). We know from Corollary 1 in [14] that, if $N \geq \frac{2}{\epsilon} \ln \frac{1}{\Delta} + 2n + \frac{2n}{\epsilon} \ln \frac{2}{\epsilon}$, the original chance constraint will hold with probability no smaller than $1 - \Delta$, or more formally,

$$\mathbb{P} \{ \mathbb{P} \{ \|\boldsymbol{\xi}\| \leq \bar{R} \} \geq 1 - \epsilon \} \geq 1 - \Delta.$$

Here, the inner probability is with respect to $\boldsymbol{\xi}$, while the outer probability is with respect to \bar{R} , since \bar{R} is also random due to the randomness of the sample S from a statistical point of view.

Given that the event $A_0: \mathbb{P} \{ \|\boldsymbol{\xi}\| \leq \bar{R} \} \geq 1 - \epsilon$ occurs, we conclude from Corollary 5.1 that with probability at least $1 - \Delta$, the inequality (5.5) holds. The probability of the event \mathcal{E} that constraint (5.5) holds is necessarily at least $(1 - \Delta)^2$:

$$\mathbb{P} \{ \mathcal{E} \} \geq \mathbb{P} \{ \mathcal{E} | A_0 \} \mathbb{P} \{ A_0 \} \geq (1 - \Delta)(1 - \Delta) = (1 - \Delta)^2,$$

which completes the proof. \square

5.1.2 Data-Driven Confidence Region for Mixture Weights

With the above results, we can concretely specify the uncertainty set D in Definition 5.1 by using historical samples. To this end, we denote $\boldsymbol{\mu}_m^{(k)} = \left(\mu_{m,1}^{(k)}, \dots, \mu_{m,n}^{(k)} \right)^\top$, here $\mu_{m,j}^{(k)} = \int \xi_j^k f_m(\boldsymbol{\xi}) d\boldsymbol{\xi}, j = 1, \dots, n, m = 1, \dots, M$, and denote the empirical estimation of $\boldsymbol{\mu}^{(k)}$ by $\bar{\boldsymbol{\mu}}^{(k)}, k = 1, \dots, I$.

As $\boldsymbol{\xi}$ follows the mixture distribution whose density function is defined by (5.1), we have $\boldsymbol{\mu}^{(k)} = \sum_{m=1}^M w_m \boldsymbol{\mu}_m^{(k)}$, and the support of $\boldsymbol{\xi}$ can be contained in a ball with any specified probability $1 - \epsilon$. Therefore, by applying Corollary 5.2, we have

$$\|\bar{\boldsymbol{\mu}}^{(k)} - \sum_{m=1}^M w_m \boldsymbol{\mu}_m^{(k)}\| \leq \gamma_k, k = 1, \dots, I, \quad (5.7)$$

hold with probability at least $(1 - \Delta)^2$. Here

$$\gamma_k = \frac{1}{\sqrt{N}} \left(2\rho_M^{(k)} + \bar{R}^k \sqrt{2 \ln \frac{1 - \epsilon}{\Delta - \epsilon}} \right), k = 1, \dots, I. \quad (5.8)$$

To derive a proper presentation for the uncertainty set, we now show that the confidence region for mixture weights can be described as the intersection of some ellipsoidal sets. Concretely, for $k = 1, \dots, I$, we have

$$\|\bar{\boldsymbol{\mu}}^{(k)} - \sum_{m=1}^M w_m \boldsymbol{\mu}_m^{(k)}\| \leq \gamma_k \Leftrightarrow \left\| \sum_{m=1}^M w_m (\bar{\boldsymbol{\mu}}^{(k)} - \boldsymbol{\mu}_m^{(k)}) \right\| \leq \gamma_k \Leftrightarrow \|U_k \mathbf{w}\| \leq \gamma_k, \quad (5.9)$$

where

$$U_k = \left(\bar{\boldsymbol{\mu}}^{(k)} - \boldsymbol{\mu}_1^{(k)}, \dots, \bar{\boldsymbol{\mu}}^{(k)} - \boldsymbol{\mu}_M^{(k)} \right) \in \mathbb{R}^{n \times M}. \quad (5.10)$$

By now, we have obtained an easily implemented representation for the mixture distribution based uncertainty set D given in Definition 5.1, with γ_k and $U_k, k = 1, \dots, I$, being defined in (5.8) and (5.10), respectively.

From (5.8) and the first inequality in (5.9), we see that when N increases, the distance between the estimated moments and true moments becomes smaller. This means that the estimated moments become more accurate and the uncertainty set D becomes smaller when N increases, which is consistent with intuition. Moreover, as N increases, the estimated values of moments tend to their true values. All these observations tell us that when N becomes large, the true values of weights \mathbf{w} are contained in the uncertainty set, which implies the feasibility of our obtained uncertainty set.

5.2 Data-Driven Robust Chance Constrained Problems

With the data-driven uncertainty set specified in Definition 5.1, we now consider the correspondingly data-driven robust chance constrained problem:

$$(\mathbf{P}) \quad \min_{\mathbf{z}} g(\mathbf{z}), \text{ s.t. } \inf_{f \in D} \mathbb{P}_f \{ \boldsymbol{\xi}^\top \mathbf{z} \leq d \} \geq 1 - \alpha, \mathbf{z} \in \mathbf{Z},$$

where $\boldsymbol{\xi} \in \mathbb{R}^n$ is a random vector defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbf{Z} \subseteq \mathbb{R}^n$ is a deterministic closed and convex set, $g(\mathbf{z}) : \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuous function, $d \in \mathbb{R}$ is a constant. f is the joint probability distribution function of $\boldsymbol{\xi}$. D , defined in Definition 5.1, is an uncertainty set of the joint probability distribution f and $\alpha \in]0, 1[$ is the tolerance probability, say, 0.01 or 0.05.

5.2.1 Reformulation of Problem (P)

As $\boldsymbol{\xi}$ follows a mixture distribution f in the uncertainty set D , we have that

$$\mathbb{P}_f \{ \boldsymbol{\xi}^\top \mathbf{z} \leq d \} = \int_{\boldsymbol{\xi}^\top \mathbf{z} \leq d} \sum_{m=1}^M w_m f_m(\boldsymbol{\xi}) d\boldsymbol{\xi} = \sum_{m=1}^M w_m \int_{\boldsymbol{\xi}^\top \mathbf{z} \leq d} f_m(\boldsymbol{\xi}) d\boldsymbol{\xi} = \mathbf{p}(\mathbf{z})^\top \mathbf{w},$$

where $\mathbf{p}(\mathbf{z}) = (p_1(\mathbf{z}), \dots, p_M(\mathbf{z}))^\top$, $p_m(\mathbf{z}) = \int_{\boldsymbol{\xi}^\top \mathbf{z} \leq d} f_m(\boldsymbol{\xi}) d\boldsymbol{\xi}$. Therefore,

$$\inf_{f \in D} \mathbb{P}_f \{ \boldsymbol{\xi}^\top \mathbf{z} \leq d \} = \inf_{\mathbf{w} \in W} \mathbf{p}(\mathbf{z})^\top \mathbf{w}. \quad (5.11)$$

To derive the reformulation of problem (P), we need the following assumption.

Assumption 5.1. *Assume that for any $\mathbf{z} \in \mathbf{Z}$, the Slater's conditions hold for problem (5.11).*

According to the definition of W , we have from the duality theory in [5] that

the dual problem of problem (5.11) is:

$$\begin{aligned}
(\mathbf{D}) \quad & \sup_{\beta, \boldsymbol{\eta}, \boldsymbol{\varphi}_k} \beta - \boldsymbol{\gamma}^\top \boldsymbol{\eta} \\
\text{s.t.} \quad & \beta \mathbf{e} + \sum_{k=1}^I U_k^\top \boldsymbol{\varphi}_k \leq \mathbf{p}(\mathbf{z}), \tag{5.12}
\end{aligned}$$

$$\|\boldsymbol{\varphi}_k\| \leq \eta_k, \quad k = 1, \dots, I, \tag{5.13}$$

where $\mathbf{e} = (1, 1, \dots, 1)^\top \in \mathbb{R}^M$, $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_I)^\top$, $\beta \in \mathbb{R}$, $\boldsymbol{\eta} \in \mathbb{R}^I$, $\boldsymbol{\varphi}_k \in \mathbb{R}^n$, $k = 1, \dots, I$.

For notational brevity, we denote the m -th row of U_k by $\mathbf{u}_{k,m}$, $m = 1, \dots, M$. Then, constraint (5.12) can be rewritten as

$$\beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq p_m(\mathbf{z}), \quad m = 1, \dots, M. \tag{5.14}$$

In addition,

$$p_m(\mathbf{z}) = \int_{\boldsymbol{\xi}^\top \mathbf{z} \leq d} f_m(\boldsymbol{\xi}) d\boldsymbol{\xi} = \Phi \left(\frac{d - \boldsymbol{\mu}_m^\top \mathbf{z}}{\sqrt{\mathbf{z}^\top \Sigma_m \mathbf{z}}} \right),$$

here $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution $N(0, 1)$. Consequently, (5.14) is equivalent to

$$d - \boldsymbol{\mu}_m^\top \mathbf{z} - \Phi^{-1} \left(\beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \right) \sqrt{\mathbf{z}^\top \Sigma_m \mathbf{z}} \geq 0, \quad m = 1, \dots, M,$$

where $\Phi^{-1}(\cdot)$ is the inverse function of $\Phi(\cdot)$. As the strong duality holds for the dual representation of $\min_{f \in D} \mathbb{P}_f \{ \boldsymbol{\xi}^\top \mathbf{z} \leq d \}$ under Assumption 5.1, problem (\mathbf{P}) can be equivalently reformulated as the following optimization problem by introducing auxiliary variables $\boldsymbol{\theta} = (\theta_1, \dots, \theta_M)^\top \in \mathbb{R}^M$:

$$\begin{aligned}
(\tilde{\mathbf{P}}) \quad & \min_{\mathbf{z}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}_k, \boldsymbol{\theta}} g(\mathbf{z}) \\
\text{s.t.} \quad & \beta - \boldsymbol{\gamma}^\top \boldsymbol{\eta} \geq 1 - \alpha, \tag{5.15}
\end{aligned}$$

$$\beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq \theta_m, \quad m = 1, \dots, M, \tag{5.16}$$

$$d - \boldsymbol{\mu}_m^\top \mathbf{z} - \Phi^{-1}(\theta_m) \sqrt{\mathbf{z}^\top \Sigma_m \mathbf{z}} \geq 0, \quad m = 1, \dots, M, \tag{5.17}$$

$$\|\boldsymbol{\varphi}_k\| \leq \eta_k, \quad k = 1, \dots, I, \tag{5.18}$$

$$\mathbf{z} \in \mathbf{Z}.$$

To be more precise, we have the following theorem:

Theorem 5.2. *Suppose that Assumption 5.1 holds. Then if $(\mathbf{z}^*, \beta^*, \boldsymbol{\eta}^*, \boldsymbol{\varphi}_k^*, \boldsymbol{\theta}^*)$ is an optimal solution to problem $(\tilde{\mathbf{P}})$, \mathbf{z}^* is an optimal solution to problem (\mathbf{P}) ; conversely, if \mathbf{z}^* is an optimal solution to problem (\mathbf{P}) , $(\mathbf{z}^*, \beta^*, \boldsymbol{\eta}^*, \boldsymbol{\varphi}_k^*, \boldsymbol{\theta}^*)$ is*

an optimal solution to problem $(\tilde{\mathbf{P}})$, here $(\beta^*, \boldsymbol{\eta}^*, \boldsymbol{\varphi}_k^*)$ is an optimal solution to problem (\mathbf{D}) and $\boldsymbol{\theta}^* = \Phi\left(\frac{d - \boldsymbol{\mu}_m^\top \mathbf{z}}{\sqrt{\mathbf{z}^\top \Sigma_m \mathbf{z}}}\right)$.

Proof. Let $(\mathbf{z}^*, \beta^*, \boldsymbol{\eta}^*, \boldsymbol{\varphi}_k^*, \boldsymbol{\theta}^*)$ be an optimal solution to problem $(\tilde{\mathbf{P}})$. By the weak duality theorem, we have

$$\inf_{f \in D} \mathbb{P}_f \{ \boldsymbol{\xi}^\top \mathbf{z} \leq d \} = \inf_{\mathbf{w} \in W} \mathbf{p}(\mathbf{z})^\top \mathbf{w} \geq \beta^* - \boldsymbol{\gamma}^\top \boldsymbol{\eta}^* \geq 1 - \alpha,$$

where the last inequality follows from (5.15). This implies that \mathbf{z}^* is feasible for problem (\mathbf{P}) . In the following, we prove that it is also optimal to problem (\mathbf{P}) .

Suppose that, on the contrary, \mathbf{z}^* is not optimal to problem (\mathbf{P}) , i.e., there exists an optimal solution $\bar{\mathbf{z}}^*$ such that $g(\bar{\mathbf{z}}^*) < g(\mathbf{z}^*)$. Let $(\bar{\beta}^*, \bar{\boldsymbol{\eta}}^*, \bar{\boldsymbol{\varphi}}_k^*)$ be the corresponding optimal solution to problem (\mathbf{D}) and $\bar{\boldsymbol{\theta}}^* = \mathbf{p}(\bar{\mathbf{z}}^*) = \Phi\left(\frac{d - \boldsymbol{\mu}_m^\top \bar{\mathbf{z}}^*}{\sqrt{(\bar{\mathbf{z}}^*)^\top \Sigma_m \bar{\mathbf{z}}^*}}\right)$.

From Assumption 5.1 and the strong conic duality theory in [5], we have

$$\bar{\beta}^* - \boldsymbol{\gamma}^\top \bar{\boldsymbol{\eta}}^* = \inf_{\mathbf{w} \in W} \mathbf{p}(\bar{\mathbf{z}}^*)^\top \mathbf{w} \geq 1 - \alpha.$$

This and other constraints in problems (\mathbf{P}) and (\mathbf{D}) mean that $(\bar{\mathbf{z}}^*, \bar{\beta}^*, \bar{\boldsymbol{\eta}}^*, \bar{\boldsymbol{\varphi}}_k^*, \bar{\boldsymbol{\theta}}^*)$ is feasible for problem $(\tilde{\mathbf{P}})$. This contradicts the fact that $(\mathbf{z}^*, \beta^*, \boldsymbol{\eta}^*, \boldsymbol{\varphi}_k^*, \boldsymbol{\theta}^*)$ is an optimal solution to problem $(\tilde{\mathbf{P}})$ since $g(\bar{\mathbf{z}}^*) < g(\mathbf{z}^*)$. Therefore, \mathbf{z}^* is an optimal solution to problem (\mathbf{P}) .

On the other hand, let \mathbf{z}^* be an optimal solution to problem (\mathbf{P}) , $(\beta^*, \boldsymbol{\eta}^*, \boldsymbol{\varphi}_k^*)$ be an optimal solution to problem (\mathbf{D}) and $\boldsymbol{\theta}^* = \Phi\left(\frac{d - \boldsymbol{\mu}_m^\top \mathbf{z}}{\sqrt{\mathbf{z}^\top \Sigma_m \mathbf{z}}}\right)$. Then, $(\mathbf{z}^*, \beta^*, \boldsymbol{\eta}^*, \boldsymbol{\varphi}_k^*, \boldsymbol{\theta}^*)$ must constitute an optimal solution to problem $(\tilde{\mathbf{P}})$. Otherwise, there would exist an optimal solution $(\bar{\mathbf{z}}^*, \bar{\beta}^*, \bar{\boldsymbol{\eta}}^*, \bar{\boldsymbol{\varphi}}_k^*, \bar{\boldsymbol{\theta}}^*)$ to problem $(\tilde{\mathbf{P}})$ such that $g(\bar{\mathbf{z}}^*) < g(\mathbf{z}^*)$. And we know from the first part of the proof that $\bar{\mathbf{z}}^*$ must be an optimal solution to problem (\mathbf{P}) , which contradicts the fact that \mathbf{z}^* is an optimal solution to problem (\mathbf{P}) since $g(\bar{\mathbf{z}}^*) < g(\mathbf{z}^*)$. \square

Problem $(\tilde{\mathbf{P}})$ is a non-convex programming problem due to the third group of constraints (5.17). We will derive its convex approximations in subsection 5.2.2, and its DC programming approximations in subsection 5.2.3.

5.2.2 Convex Approximation

In this part, we will show that, under certain conditions, problem $(\tilde{\mathbf{P}})$ can be reformulated as a convex programming problem. Liu et al [65] proposed tight convex approximations to a kind of joint geometric chance constrained problems. We extend their method and propose tight convex approximations to problem $(\tilde{\mathbf{P}})$ under some conditions. To this end, we need the following assumption:

Assumption 5.2. Denote $\mathbf{e}^t = (e^{t_1}, \dots, e^{t_n})^\top$ for $\mathbf{t} = (t_1, \dots, t_n)^\top$. Assume that

- 1) $\alpha \leq 1 - \Phi(1)$,

- 2) the function $h(\mathbf{t}) = g(e^{\mathbf{t}})$ is convex with respect to \mathbf{t} ,
- 3) $\mu_{mj} \geq 0$, $\sigma_{jl}^m \geq 0$, $j = 1, \dots, n$, $l = 1, \dots, n$, $m = 1, \dots, M$,
- 4) $\mathbf{Z} = \mathbb{R}_+^n \cap \mathbf{L}$, \mathbf{L} is selected such that $\tilde{\mathbf{L}} = \{\tilde{\mathbf{z}} \in \mathbb{R}^n : \tilde{z}_j = \log(z_j), j = 1, \dots, n, \mathbf{z} \in \mathbf{L}\}$ is closed and convex.

To illustrate the possible selections of \mathbf{L} , we consider the following two examples.

Example 5.1. Let $\mathbf{L} = \{\mathbf{z} \in \mathbb{R}_{++}^n : \mathbf{c}^\top \mathbf{z} \leq h\}$, here $\mathbf{c} = (c_1, \dots, c_n)^\top \in \mathbb{R}_+^n$ and $h \in \mathbb{R}_+$ are constants. Then, $\tilde{\mathbf{L}} = \{\tilde{\mathbf{z}} \in \mathbb{R}^n : \sum_{j=1}^n c_j e^{\tilde{z}_j} \leq h\}$ is a closed and convex set.

Example 5.2. Define $\mathbf{L} = \{\mathbf{z} \in \mathbb{R}_{++}^n : f(\mathbf{z}) \leq h\}$, here $f : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$ is a non-decreasing convex function and $h \in \mathbb{R}$ is a constant. Then, $\tilde{\mathbf{L}} = \{\tilde{\mathbf{z}} \in \mathbb{R}^n : f(e^{\tilde{\mathbf{z}}}) \leq h\}$ is a closed and convex set. Here $e^{\tilde{\mathbf{z}}} = (e^{\tilde{z}_1}, \dots, e^{\tilde{z}_n})$.

Then, we have the following lemma.

Lemma 5.2. Constraints (5.15), (5.16) and (5.18) imply that

- (a) if $\alpha \leq \frac{1}{2}$, $\Phi^{-1}(\theta_m) \geq 0$, $m = 1, \dots, M$,
- (b) if $\alpha \leq 1 - \Phi(1)$, the function $\log(\Phi^{-1}(\theta_m))$ is convex with respect to θ_m , $m = 1, \dots, M$.

Proof. (a) Firstly, we determine the range of θ_m , $m = 1, \dots, M$. We have from (5.15) and (5.16) that

$$1 - \alpha + \boldsymbol{\gamma}^\top \boldsymbol{\eta} + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq \theta_m, \quad m = 1, \dots, M.$$

Hence,

$$L_m := \min_{\boldsymbol{\eta}, \boldsymbol{\varphi}_k} 1 - \alpha + \boldsymbol{\gamma}^\top \boldsymbol{\eta} + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \quad (5.19)$$

s.t. $\|\boldsymbol{\varphi}_k\| \leq \eta_k, \quad k = 1, \dots, I,$

is smaller than or equal to θ_m , $m = 1, \dots, M$. Problem (5.19) can be decomposed into I subproblems, and correspondingly $L_m = \sum_{k=1}^I L_{m,k}$, where

$$L_{m,k} = \min_{\eta_k, \boldsymbol{\varphi}_k} \gamma_k \eta_k + \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \quad (5.20)$$

s.t. $\|\boldsymbol{\varphi}_k\| \leq \eta_k.$

Since $\gamma_k > 0$, $k = 1, \dots, I$, it is easy to deduce that $L_{m,k} = 0$, and $L_m = 1 - \alpha$. Hence, $\theta_m \geq 1 - \alpha$, $m = 1, \dots, M$. Therefore, if $\alpha \leq \frac{1}{2}$, we have $\Phi^{-1}(\theta_m) \geq \Phi^{-1}(1 - \alpha) \geq 0$, $m = 1, \dots, M$.

(b) The second order derivative function of $\log(\Phi^{-1}(\theta_m))$ is $-\frac{\phi(y)+y\phi'(y)}{(y\phi(y))^2\phi(y)}$, where $y = \Phi^{-1}(\theta_m)$, $\phi(\cdot)$ is the probability density function of the standard normal distribution $N(0, 1)$ and $\phi'(\cdot)$ is its first order derivate.

$-\frac{\phi(y)+y\phi'(y)}{(y\phi(y))^2\phi(y)} \geq 0$ if and only if $\log(\Phi^{-1}(\theta_m))$ is convex. Moreover, we have that

$$\begin{aligned} -\frac{\phi(y) + y\phi'(y)}{(y\phi(y))^2\phi(y)} &\geq 0 \Leftrightarrow \phi(y) + y\phi'(y) \leq 0 \\ \Leftrightarrow \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} + y\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}(-y) &\leq 0 \\ \Leftrightarrow 1 - y^2 \leq 0 &\Leftrightarrow \theta_m \geq \Phi(1) \text{ or } \theta_m \leq \Phi(-1). \end{aligned}$$

As $\theta_m \geq 1 - \alpha$, $m = 1, \dots, M$, we can conclude that if $\alpha \leq 1 - \Phi(1)$, the function $\log(\Phi^{-1}(\theta_m))$ is convex with respect to θ_m , $m = 1, \dots, M$. \square

From Lemma 5.2, we have that $\Phi^{-1}(\theta_m) \geq 0$ holds for $\alpha \leq \frac{1}{2}$. Hence, the inequality constraint (5.17) can be rewritten as

$$d - \boldsymbol{\mu}_m^\top \mathbf{z} - \sqrt{\sum_{j=1}^n \sum_{l=1}^n \sigma_{jl}^m z_j z_l (\Phi^{-1}(\theta_m))^2} \geq 0, \quad m = 1, \dots, M. \quad (5.21)$$

By introducing $t_m = \log(z_m)$, $m = 1, \dots, M$, we have that (5.21) is equivalent to

$$\sqrt{\sum_{j=1}^n \sum_{l=1}^n \sigma_{jl}^m e^{t_j+t_l+2\log(\Phi^{-1}(\theta_m))}} + \sum_{j=1}^n \mu_{mj} e^{t_j} - d \leq 0, \quad m = 1, \dots, M.$$

By now, problem $(\tilde{\mathbf{P}})$ can be reformulated as

$$\begin{aligned} (\mathbf{P}^C) \quad &\min_{\mathbf{t}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}_k, \boldsymbol{\theta}} g(\mathbf{e}^{\mathbf{t}}) \\ \text{s.t.} \quad &\sqrt{\sum_{j=1}^n \sum_{l=1}^n \sigma_{jl}^m e^{t_j+t_l+2\log(\Phi^{-1}(\theta_m))}} \\ &\quad + \sum_{j=1}^n \mu_{mj} e^{t_j} - d \leq 0, \quad m = 1, \dots, M, \\ &(5.15), (5.16), (5.18), \mathbf{t} \in \tilde{\mathbf{L}}. \end{aligned}$$

From Lemma 5.2, we have that $\log(\Phi^{-1}(\theta_m))$ is convex with respect to θ_m , $m = 1, \dots, M$, when $\alpha \leq 1 - \Phi(1)$. From Assumption 5.2, we have that $\alpha \leq 1 - \Phi(1)$ and $\mu_{mj} \geq 0$, $\sigma_{jl}^m \geq 0$, $j = 1, \dots, n$, $l = 1, \dots, n$, $m = 1, \dots, M$. Hence, $\sqrt{\sum_{j=1}^n \sum_{l=1}^n \sigma_{jl}^m e^{t_j+t_l+2\log(\Phi^{-1}(\theta_m))}}$ is convex with respect to $(\mathbf{t}, \boldsymbol{\theta}_m)$ under Assumption 5.2. This means that, under Assumption 5.2, problem (\mathbf{P}^C) is a convex programming problem.

However, it is still difficult to directly solve problem (\mathbf{P}^C) because of the nonelementary function $\log(\Phi^{-1}(\theta_m))$. We thus propose two piecewise linear

approximations to $\log(\Phi^{-1}(\theta_m))$, and then derive two approximations to problem $(\mathbf{P}^{\mathbf{C}})$, which provide lower and upper bounds to the optimum value of problem $(\mathbf{P}^{\mathbf{C}})$, respectively.

Lower Approximation

We choose K different linear functions:

$$l_i(\theta_m) = a_i\theta_m + b_i, \quad i = 1, \dots, K,$$

such that

$$l_i(\theta_m) \leq \log(\Phi^{-1}(\theta_m)), \quad \forall \theta_m \in [1 - \alpha, 1[, \quad i = 1, \dots, K.$$

Then $\log(\Phi^{-1}(\theta_m))$ can be approximated by the following piecewise linear function

$$l(\theta_m) = \max_{i=1, \dots, K} l_i(\theta_m),$$

which provides a lower approximation to $\log(\Phi^{-1}(\theta_m))$.

In order to achieve a satisfactory accuracy, we set $l_i(\theta_m)$ as the tangent line of $\log(\Phi^{-1}(\theta_m))$ at K points τ_1, \dots, τ_K with $\tau_i \in [1 - \alpha, 1[, \quad i = 1, \dots, K$. Then, we have

$$a_i = (\Phi^{-1}(\tau_i)\phi(\Phi^{-1}(\tau_i)))^{-1}, \quad b_i = \log(\Phi^{-1}(\tau_i)) - a_i\tau_i.$$

Remark 5.1. *In general, we can choose the K tangent points uniformly distributed in the interval $[1 - \alpha, 1[$.*

To demonstrate the reasonability of the above piecewise linear approximation to $\log(\Phi^{-1}(\theta_m))$, we first introduce some notations. For sets $A, B \subset \mathbb{R}^n$, let $\text{dist}(x, A) = \inf_{x' \in A} \|x - x'\|$, and $\mathbb{D}(A, B) = \sup_{x \in A} \text{dist}(x, B)$ denote the deviation of the set A from the set B (see Shapiro et al. [96]). We denote the feasible solution set, the optimal solution set and the optimal value to problem $(\mathbf{P}^{\mathbf{C}})$ by $\tilde{\mathbf{Z}}$, $\tilde{\mathbf{O}}$ and \tilde{v} , respectively. Then for the above piecewise linear approximation to $\log(\Phi^{-1}(\theta_m))$, we have the following theorem.

Theorem 5.3. *Under Assumption 5.2, replacing $\log(\Phi^{-1}(\theta_m))$ in problem $(\mathbf{P}^{\mathbf{C}})$ by $l(\theta_m)$ generates the following convex approximation to problem $(\mathbf{P}^{\mathbf{C}})$:*

$$\begin{aligned}
(\mathbf{P}_L^{\mathbf{C}}) \quad & \min_{\mathbf{t}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}_k, \boldsymbol{\theta}} g(e^{\mathbf{t}}) \\
\text{s.t.} \quad & \sqrt{\sum_{j=1}^n \sum_{l=1}^n \sigma_{jl}^m e^{t_j + t_l + 2(a_i\theta_m + b_i)}} \\
& + \sum_{j=1}^n \mu_{mj} e^{t_j} - d \leq 0, \quad m = 1, \dots, M, i = 1, \dots, K, \\
& (5.15), (5.16), (5.18), \mathbf{t} \in \tilde{\mathbf{L}}.
\end{aligned}$$

The optimum value of problem $(\mathbf{P}_L^{\mathbf{C}})$ is a lower bound to that of problem (\mathbf{P}) . What's more, we have $\tilde{\mathbf{Z}} = \lim_{K \rightarrow \infty} \tilde{\mathbf{Z}}_K^L$, $\lim_{K \rightarrow \infty} \mathbb{D}(\tilde{\mathbf{O}}_K^L, \tilde{\mathbf{O}}) = 0$ and $\tilde{v} = \lim_{K \rightarrow \infty} \tilde{v}_K^L$, here

$\bar{\mathbf{Z}}_K^L$, $\bar{\mathbf{O}}_K^L$ and \bar{v}_K^L are the feasible solution set, the optimal solution set and the optimal value to problem (\mathbf{P}_L^C) , respectively.

Proof. As problem (\mathbf{P}_L^C) is deduced by relaxing constraints (5.17) in problem (\mathbf{P}^C) , it is obvious that the optimal value of problem (\mathbf{P}_L^C) is a lower bound to that of problem (\mathbf{P}^C) , which is equivalent to problem (\mathbf{P}) .

Let

$$T^* = \left\{ (\mathbf{t}, \boldsymbol{\theta}) : \sqrt{\sum_{j=1}^n \sum_{l=1}^n \sigma_{jl}^m e^{t_j+t_l+2\log(\Phi^{-1}(\theta_m))}} + \sum_{j=1}^n \mu_{mj} e^{t_j} - d \leq 0, m = 1, \dots, M \right\},$$

$$T_K^L = \left\{ (\mathbf{t}, \boldsymbol{\theta}) : \sqrt{\sum_{j=1}^n \sum_{l=1}^n \sigma_{jl}^m e^{t_j+t_l+2(a_i\theta_m+b_i)}} + \sum_{j=1}^n \mu_{mj} e^{t_j} - d \leq 0, m = 1, \dots, M, i = 1, \dots, K \right\},$$

and $\mathbf{Z}_0 = \tilde{\mathbf{Z}} \setminus T^*$. Because $\tilde{\mathbf{Z}} = \mathbf{Z}_0 \cup T^*$ and $\bar{\mathbf{Z}}_K^L = \mathbf{Z}_0 \cup T_K^L$, we just need to verify $T^* = \bigcup_{K=1}^{\infty} \text{cl}T_K^L$ to establish the rest conclusions.

Since for any K , $T_K^L \supset T_{K+1}^L$, by Exercise 4.3 in [94], we have $\lim_{K \rightarrow \infty} T_K^L = \bigcap_{K=1}^{\infty} \text{cl}T_K^L$. It is obvious that $T^* \subset \bigcap_{K=1}^{\infty} \text{cl}T_K^L$. Then we will show that $T^* \supset \bigcap_{K=1}^{\infty} \text{cl}T_K^L$.

For any $\boldsymbol{\tau} \in \bigcap_{K=1}^{\infty} \text{cl}T_K^L$, we have $\boldsymbol{\tau} \in \text{cl}T_K^L, \forall K$. If $\boldsymbol{\tau} \notin T^*$, since T^* is closed and convex, $\boldsymbol{\tau}$ can be separated strictly from the set T^* by a hyperplane. This implies that there must be a \hat{K} such that $t \notin \text{cl}T_{\hat{K}}^L$, which is contrary with $t \in \text{cl}T_K^L, \forall K$.

Thus we get $T^* = \bigcap_{K=1}^{\infty} \text{cl}T_K^L$. Therefore, we have $\tilde{\mathbf{Z}} = \lim_{K \rightarrow \infty} \bar{\mathbf{Z}}_K^L$.

Denote $\mathbf{x} = (\mathbf{t}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}_k, \boldsymbol{\theta})$. Let $\bar{g}(\mathbf{x}) = g(e^{\mathbf{t}}) + I_{\tilde{\mathbf{Z}}}(\mathbf{x})$ and $\bar{g}_K^L(\mathbf{x}) = g(e^{\mathbf{t}}) + I_{\bar{\mathbf{Z}}_K^L}(\mathbf{x})$, where $I_A(\mathbf{x}) = 0$ if $\mathbf{x} \in A$, otherwise $I_A(\mathbf{x}) = +\infty$. As $\tilde{\mathbf{Z}} = \lim_{K \rightarrow \infty} \bar{\mathbf{Z}}_K^L$, by Proposition 7.4(f) in [94], we have that $I_{\bar{\mathbf{Z}}_K^L}(\cdot)$ epi-converges to $I_{\tilde{\mathbf{Z}}}(\cdot)$ as $K \rightarrow +\infty$. Since $g(e^{\mathbf{t}})$ is continuous, we then have that $\bar{g}_K^L(\cdot)$ epi-converges to $\bar{g}(\cdot)$. As $\tilde{\mathbf{Z}}$ and $\bar{\mathbf{Z}}_K^L$ are closed and convex sets, we have that $\bar{g}_K^L(\cdot)$ and $\bar{g}(\cdot)$ are lower semi-continuous. Then, by Theorem 7.33 in [94], we have $\tilde{v} = \lim_{K \rightarrow \infty} \bar{v}_K^L$ and $\limsup_{K \rightarrow \infty} \bar{\mathbf{O}}_K^L \subset \tilde{\mathbf{O}}$. It can thus be deduced from the discussions in Example 4.13 in [94] that $\lim_{K \rightarrow \infty} \mathbb{D}(\bar{\mathbf{O}}_K^L, \tilde{\mathbf{O}}) = 0$. \square

Upper Approximation

In order to come up with an upper bound to the optimum value of problem (\mathbf{P}^C) , we use the linear segments $\bar{a}_i\theta_m + \bar{b}_i$, $i = 1, \dots, K$, between $\tau_1, \tau_2, \dots, \tau_{K+1} \in [1 - \alpha, 1[$ to form a piecewise linear function

$$\bar{l}(\theta_m) = \max_{i=1, \dots, K} \{ \bar{a}_i\theta_m + \bar{b}_i \},$$

where

$$\bar{a}_i = \frac{\log(\Phi^{-1}(\tau_{i+1})) - \log(\Phi^{-1}(\tau_i))}{\tau_{i+1} - \tau_i}, \quad \bar{b}_i = \log(\Phi^{-1}(\tau_i)) - \bar{a}_i \tau_i, \quad i = 1, \dots, K.$$

Using the piecewise linear function $\bar{l}(\theta_m)$ to replace $\log(\Phi^{-1}(\theta_m))$ in problem $(\mathbf{P}^{\mathbf{C}})$ gives the following approximation problem:

$$(\mathbf{P}_{\mathbf{U}}^{\mathbf{C}}) \quad \min_{\mathbf{t}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}_k, \boldsymbol{\theta}} g(e^{\mathbf{t}})$$

$$\text{s.t.} \quad \sqrt{\sum_{j=1}^n \sum_{l=1}^n \sigma_{jl}^m e^{t_j + t_l + 2(\bar{a}_i \theta_m + \bar{b}_i)}} + \sum_{j=1}^n \mu_{mj} e^{t_j} - d \leq 0, \quad m = 1, \dots, M, i = 1, \dots, K,$$

$$(5.15), (5.16), (5.18), \mathbf{t} \in \tilde{\mathbf{L}}.$$

Similar to the proof of Theorem 5.3, we can establish the following conclusion:

Theorem 5.4. *Under Assumption 5.2, problem $(\mathbf{P}_{\mathbf{U}}^{\mathbf{C}})$ is a conic programming and the optimum value of problem $(\mathbf{P}_{\mathbf{U}}^{\mathbf{C}})$ is an upper bound to that of problem (\mathbf{P}) . What's more, we have $\tilde{\mathbf{Z}} = \lim_{K \rightarrow \infty} \bar{\mathbf{Z}}_K^{\mathbf{U}}, \lim_{K \rightarrow \infty} \mathbb{D}(\bar{\mathbf{O}}_K^{\mathbf{U}}, \tilde{\mathbf{O}}) = 0$ and $\tilde{v} = \lim_{K \rightarrow \infty} \bar{v}_K^{\mathbf{U}}$, here $\bar{\mathbf{Z}}_K^{\mathbf{U}}, \bar{\mathbf{O}}_K^{\mathbf{U}}$ and $\bar{v}_K^{\mathbf{U}}$ are the feasible solution set, the optimal solution set and the optimal value to problem $(\mathbf{P}_{\mathbf{U}}^{\mathbf{C}})$, respectively.*

5.2.3 DC reformulation

Assumption 5.2 might not always hold in practice. Therefore, in this subsection, we assume $g(\mathbf{z})$ is convex and bounded from below on the feasible solution set of problem $(\tilde{\mathbf{P}})$ instead of Assumption 5.2, and then demonstrate that problem (\mathbf{P}) is still solvable, since it can be equivalently represented as a DC programming.

Let $p_m = \Phi^{-1}(\theta_m)$ and introduce auxiliary variables $\sigma_m, m = 1, \dots, M$, problem $(\tilde{\mathbf{P}})$ can be rewritten as

$$(\mathbf{P}^{\mathbf{DC}}) \quad \min_{\mathbf{z}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}, \mathbf{p}, \boldsymbol{\sigma}} g(\mathbf{z})$$

$$\text{s.t.} \quad \beta - \boldsymbol{\gamma}^{\top} \boldsymbol{\eta} \geq 1 - \alpha, \quad (5.22)$$

$$\beta + \sum_{k=1}^I \mathbf{u}_{k,m}^{\top} \boldsymbol{\varphi}_k \leq \Phi(p_m), \quad m = 1, \dots, M, \quad (5.23)$$

$$\|\boldsymbol{\varphi}_k\| \leq \eta_k, \quad k = 1, \dots, I, \quad (5.24)$$

$$p_m \sigma_m \leq d - \boldsymbol{\mu}_m^{\top} \mathbf{z}, \quad m = 1, \dots, M, \quad (5.25)$$

$$\left\| \Sigma_m^{\frac{1}{2}} \mathbf{z} \right\| \leq \sigma_m, \quad m = 1, \dots, M, \quad (5.26)$$

$$\mathbf{z} \in \mathbf{Z},$$

here $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_M)^{\top}$. In problem $(\mathbf{P}^{\mathbf{DC}})$, the constraints (5.25) can be rewrit-

ten as

$$\frac{1}{2}(p_m + \sigma_m)^2 - \frac{1}{2}(p_m^2 + \sigma_m^2) \leq d - \boldsymbol{\mu}_m^\top \mathbf{z}, \quad m = 1, \dots, M,$$

which are DC constraints. Meanwhile, the function $\Phi(p_m)$ in (5.23) can also be written as a DC function, which means the inequality constraints (5.23) are also DC constraints. Therefore, problem (\mathbf{P}^{DC}) is essentially a DC programming.

We know from Lemma 5.2 that $p_m = \Phi^{-1}(\theta_m) \geq 0$ when $\alpha < \frac{1}{2}$. This implies that the second group of constraints (5.23) become convex constraints, since the function $\Phi(\cdot)$ is concave on positive axis. Thus the number of the DC constraints in problem (\mathbf{P}^{DC}) is reduced. On the other hand, we always have $\alpha < \frac{1}{2}$ in practice. Hence, we only consider problem (\mathbf{P}^{DC}) with $\alpha < 0.5$ in the rest of this chapter. Similar to that in subsection 5.2.2, we approximate $\Phi(p_m)$ by a lower piecewise linear function.

Upper bound

From Lemma 5.2, we have that $\Phi(p_m)$ is concave when $\alpha < 0.5$. Therefore, we can obtain an upper bound approximation of problem $(\tilde{\mathbf{P}})$ by replacing $\Phi(p_m)$ with its lower piecewise linear function.

Concretely, we construct a piecewise linear function through $K \geq 2$ different points $\{\tau_1 = \Phi^{-1}(1 - \alpha), \tau_2, \dots, \tau_K\}$ which are arranged in an ascending order, as follows:

$$\tilde{l}(p_m) = \min_{i=1, \dots, K} \left\{ \tilde{a}_i p_m + \tilde{b}_i \right\}, \quad (5.27)$$

where

$$\tilde{a}_i = \begin{cases} \frac{\Phi(\tau_{i+1}) - \Phi(\tau_i)}{\tau_{i+1} - \tau_i}, & 1 \leq i < K \\ 0, & i = K \end{cases}; \quad \tilde{b}_i = \Phi(\tau_i) - \tilde{a}_i \tau_i.$$

By replacing $\Phi(p_m)$ with $\tilde{l}(p_m)$ in problem (\mathbf{P}^{DC}) , we get the following DC upper bound approximation:

$$\begin{aligned} (\mathbf{P}_U^{\text{DC}}) \quad & \min_{\mathbf{z}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}, \mathbf{p}, \boldsymbol{\sigma}} \quad g(\mathbf{z}) \\ \text{s.t.} \quad & \beta - \boldsymbol{\gamma}^\top \boldsymbol{\eta} \geq 1 - \alpha, \\ & \beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq \tilde{a}_i p_m + \tilde{b}_i, \quad m = 1, \dots, M, i = 1, \dots, K, \\ & \|\boldsymbol{\varphi}_k\| \leq \eta_k, \quad k = 1, \dots, I, \\ & \frac{1}{2}(p_m + \sigma_m)^2 - \frac{1}{2}(p_m^2 + \sigma_m^2) \leq d - \boldsymbol{\mu}_m^\top \mathbf{z}, \quad m = 1, \dots, M, \\ & \left\| \sum_m \frac{1}{2} \mathbf{z} \right\| \leq \sigma_m, \quad m = 1, \dots, M, \\ & \mathbf{z} \in \mathbf{Z}. \end{aligned}$$

We denote the feasible solution sets of problem (\mathbf{P}^{DC}) and problem $(\mathbf{P}_U^{\text{DC}})$ by $\tilde{\mathbf{Z}}_{DC}$ and $\tilde{\mathbf{Z}}_K$, respectively. Let \tilde{O}_{DC} and \tilde{v}_{DC} be the optimal solution set and the optimal value of problem (\mathbf{P}^{DC}) , \bar{O}_K and \bar{v}_K be the optimal solution set and optimal value of problem $(\mathbf{P}_U^{\text{DC}})$, respectively. Then, similar with Theorem 5.4, we have the following theorem.

Theorem 5.5. *The optimal value of problem $(\mathbf{P}_U^{\text{DC}})$ is an upper bound for the optimal value of problem $(\tilde{\mathbf{P}})$. What's more, we have $\tilde{\mathbf{Z}}_{DC} = \lim_{K \rightarrow \infty} \tilde{\mathbf{Z}}_K$, $\lim_{K \rightarrow \infty} \mathbb{D}(\bar{O}_K, \tilde{O}_{DC}) =$*

0 and $\tilde{v}_{DC} = \lim_{K \rightarrow \infty} \bar{v}_K$.

Proof. Let

$$T_{DC}^* = \left\{ \mathbf{p} : \beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq \Phi(p_m), m = 1, \dots, M \right\},$$

$$\bar{T}_K = \left\{ \mathbf{p} : \beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq a_i p_m + b_i, m = 1, \dots, M, i = 1, \dots, K \right\},$$

and $\mathbf{Z}_0 = \tilde{\mathbf{Z}}_{DC} \setminus T_{DC}^*$. It is obvious that $\bar{T}_K \subset T_{DC}^*$. This means that the optimal value of problem $(\mathbf{P}_U^{\text{DC}})$ is an upper bound for the optimal value of problem (\mathbf{P}^{DC}) , which is equivalent to problem $(\tilde{\mathbf{P}})$.

As $\tilde{\mathbf{Z}}_{DC} = \mathbf{Z}_0 \cup T_{DC}^*$ and $\tilde{\mathbf{Z}}_K = \mathbf{Z}_0 \cup \bar{T}_K$, to establish the rest conclusions, we just need to verify $T_{DC}^* = \bigcup_{K=1}^{\infty} \text{cl} \bar{T}_K$.

Since for any K , $\bar{T}_K \subset \bar{T}_{K+1}$, by Exercise 4.3 in [94], we have $\lim_{K \rightarrow \infty} \bar{T}_K = \text{cl} \bigcup_{K=1}^{\infty} \bar{T}_K$. It is obvious that $T_{DC}^* \supset \text{cl} \bigcup_{K=1}^{\infty} \bar{T}_K$. So we only need to show that $T_{DC}^* \subset \text{cl} \bigcup_{K=1}^{\infty} \bar{T}_K$.

As T_{DC}^* is closed and convex, we have $\text{cl}(\text{int} T_{DC}^*) = T_{DC}^*$. Therefore, for any $t \in \text{int} T_{DC}^*$, there exists a \tilde{K} such that $t \in \bar{T}_{\tilde{K}}$. This implies $\text{int} T_{DC}^* \subset \text{cl} \bigcup_{K=1}^{\infty} \bar{T}_K$,

which means $T_{DC}^* \subset \text{cl} \bigcup_{K=1}^{\infty} \bar{T}_K$. Then we conclude $T_{DC}^* = \text{cl} \bigcup_{K=1}^{\infty} \bar{T}_K$. Therefore,

we have $\tilde{\mathbf{Z}}_{DC} = \lim_{K \rightarrow \infty} \tilde{\mathbf{Z}}_K$.

As $\lim_{K \rightarrow \infty} \tilde{\mathbf{Z}}_K = \tilde{\mathbf{Z}}_{DC}$, the rest of the conclusions can be established by using the same argument as that in the proof of Theorem 2 in [53]. \square

SCA Algorithm for DC programming

In this part, we discuss how to solve the DC programming problem $(\mathbf{P}_U^{\text{DC}})$. DC programming problems have been studied extensively in recent years. Hong et al.[53] proposed an efficient method to solve the optimization problem with one DC constraint. We extend this method so that it can be used to solve problem $(\mathbf{P}_U^{\text{DC}})$ which has K DC constraints.

The basic idea of the SCA algorithm in [53] is to convexify the DC constraint via a first-order Taylor approximation. Specially, for each DC constraint in problem $(\mathbf{P}_U^{\text{DC}})$, we can use the first-order Taylor expansion

$$\frac{1}{2} (\hat{p}_m^2 + \hat{\sigma}_m^2) + \hat{p}_m (p_m - \hat{p}_m) + \hat{\sigma}_m (\sigma_m - \hat{\sigma}_m)$$

at any feasible point with $(\hat{p}_m, \hat{\sigma}_m)$ to approximate $\frac{1}{2} (p_m^2 + \sigma_m^2)$. With the con-

vexity of the function $\frac{1}{2}(p_m^2 + \sigma_m^2)$, we have

$$\frac{1}{2}(p_m^2 + \sigma_m^2) \geq \frac{1}{2}(\hat{p}_m^2 + \hat{\sigma}_m^2) + \hat{p}_m(p_m - \hat{p}_m) + \hat{\sigma}_m(\sigma_m - \hat{\sigma}_m),$$

and the original constraint is replaced by

$$\frac{1}{2}(p_m + \sigma_m)^2 - \frac{1}{2}(\hat{p}_m^2 + \hat{\sigma}_m^2) - \hat{p}_m(p_m - \hat{p}_m) - \hat{\sigma}_m(\sigma_m - \hat{\sigma}_m) \leq d - \boldsymbol{\mu}_m^\top \mathbf{z}.$$

Let $\mathbf{CP}((\hat{\mathbf{p}}, \hat{\boldsymbol{\sigma}}))$ denote the following optimization problem:

$$\begin{aligned} \min_{\mathbf{z}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}_k, \mathbf{p}, \boldsymbol{\sigma}} \quad & g(\mathbf{z}) \\ \text{s.t.} \quad & \beta - \boldsymbol{\gamma}^\top \boldsymbol{\eta} \geq 1 - \alpha, \\ & \beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq a_i p_m + b_i, \quad m = 1, \dots, M, i = 1, \dots, K, \\ & \|\boldsymbol{\varphi}_k\| \leq \eta_k, \quad k = 1, \dots, I, \\ & \frac{1}{2}(p_m + \sigma_m)^2 - \hat{p}_m(p_m - \hat{p}_m) - \hat{\sigma}_m(\sigma_m - \hat{\sigma}_m) \\ & \quad \leq d - \boldsymbol{\mu}_m^\top \mathbf{z} + \frac{1}{2}(\hat{p}_m^2 + \hat{\sigma}_m^2), \quad m = 1, \dots, M, \\ & \left\| \sum_m \frac{1}{2} \mathbf{z} \right\| \leq \sigma_m, \quad m = 1, \dots, M, \\ & \mathbf{z} \in \mathbf{Z}. \end{aligned}$$

Here, $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_M)^\top$ and $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_1, \dots, \hat{\sigma}_M)^\top$. And the feasible set to problem $(\mathbf{P}_U^{\text{DC}})$ is denoted by $\bar{\mathbf{Z}}_{(\hat{\mathbf{p}}, \hat{\boldsymbol{\sigma}})}$. Then for any feasible point with $(\hat{\mathbf{p}}, \hat{\boldsymbol{\sigma}})$ of problem $(\mathbf{P}_U^{\text{DC}})$, problem $\mathbf{CP}((\hat{\mathbf{p}}, \hat{\boldsymbol{\sigma}}))$ is a convex conservative approximation to problem $(\mathbf{P}_U^{\text{DC}})$. Therefore, we can repeat solving problem $\mathbf{CP}((\hat{\mathbf{p}}, \hat{\boldsymbol{\sigma}}))$ at the newly obtained solution, which leads to the following algorithm.

Algorithm 2 SCA method for solving problem $(\mathbf{P}_U^{\text{DC}})$

Step 1. Choose an initial feasible point with $(\mathbf{p}^0, \boldsymbol{\sigma}^0)$, set $j = 0$.

Step 2. Solve problem $\mathbf{CP}((\mathbf{p}^j, \boldsymbol{\sigma}^j))$ and obtain $(\mathbf{p}^{j+1}, \boldsymbol{\sigma}^{j+1})$.

Step 3. Stop if the optimal solution of problem $\mathbf{CP}((\mathbf{p}^j, \boldsymbol{\sigma}^j))$ is a KKT point of problem $(\mathbf{P}_U^{\text{DC}})$. Otherwise, let $j = j + 1$ and go to **Step 2**.

Algorithm 2 is easy to implement since we only need to solve at each iteration the convex optimization problem $\mathbf{CP}((\mathbf{p}^j, \boldsymbol{\sigma}^j))$, which is actually a SOCP and can thus be solved in polynomial time.

We say the Slater's condition holds at the feasible point with $(\hat{\mathbf{p}}, \hat{\boldsymbol{\sigma}})$ if the interior of the feasible solution set of problem $\mathbf{CP}((\hat{\mathbf{p}}, \hat{\boldsymbol{\sigma}}))$ is nonempty. The Slater's condition is one of the most commonly used constraint qualifications for convex optimization. Then we have the following theorem.

Theorem 5.6. *Suppose that $\{\mathbf{z}^j, \beta^j, \boldsymbol{\eta}^j, \boldsymbol{\varphi}_k^j, \mathbf{p}^j, \boldsymbol{\sigma}^j\}$ is a sequence of solutions generated by 2. Then*

- (a) $\{\mathbf{z}^j, \beta^j, \boldsymbol{\eta}^j, \boldsymbol{\varphi}_k^j, \mathbf{p}^j, \boldsymbol{\sigma}^j\}$ is contained in the feasible solution set of problem $(\mathbf{P}_U^{\text{DC}})$ and $\{g(\mathbf{z}^j)\}$ is a convergent non-increasing sequence.

(b) If a cluster point of $\{\mathbf{z}^j, \beta^j, \boldsymbol{\eta}^j, \boldsymbol{\varphi}_k^j, \mathbf{p}^j, \boldsymbol{\sigma}^j\}$ satisfies the Slater's condition, the cluster point is a KKT point of problem $(\mathbf{P}_U^{\text{DC}})$, whose corresponding objective function value also provides an upper bound for the optimal value of problem (\mathbf{P}) .

Proof. (a) For any $j \geq 0$, $\{\mathbf{z}^{j+1}, \beta^{j+1}, \boldsymbol{\eta}^{j+1}, \boldsymbol{\varphi}_k^{j+1}, \mathbf{p}^{j+1}, \boldsymbol{\sigma}^{j+1}\}$ is an optimal solution to problem $\mathbf{CP}((\mathbf{p}^j, \boldsymbol{\sigma}^j))$. Then, it is also a feasible solution to problem $(\mathbf{P}_U^{\text{DC}})$, since the feasible solution set of problem $\mathbf{CP}((\mathbf{p}^j, \boldsymbol{\sigma}^j))$ is a subset of the feasible solution set of problem $(\mathbf{P}_U^{\text{DC}})$.

As $\{\mathbf{z}^{j+1}, \beta^{j+1}, \boldsymbol{\eta}^{j+1}, \boldsymbol{\varphi}_k^{j+1}, \mathbf{p}^{j+1}, \boldsymbol{\sigma}^{j+1}\}$ is an optimal solution to problem $\mathbf{CP}((\mathbf{p}^j, \boldsymbol{\sigma}^j))$ and $\{\mathbf{z}^j, \beta^j, \boldsymbol{\eta}^j, \boldsymbol{\varphi}_k^j, \mathbf{p}^j, \boldsymbol{\sigma}^j\}$ is a feasible solution to problem $\mathbf{CP}((\mathbf{p}^j, \boldsymbol{\sigma}^j))$, we have $g(\mathbf{z}^{j+1}) \leq g(\mathbf{z}^j)$. This shows that $\{g(\mathbf{z}^j)\}$ is non-increasing. Because the continuous function $g(\mathbf{z})$ is bounded from below on the feasible solution set, $\{g(\mathbf{z}^j)\}$ is thus bounded from below, which implies that $\{g(\mathbf{z}^j)\}$ is a convergent non-increasing sequence.

(b) Let

$$g_1^m := \frac{1}{2} (p_m + \sigma_m)^2 + \boldsymbol{\mu}_m^\top \mathbf{z} - d$$

and

$$g_2^m := \frac{1}{2} (p_m^2 + \sigma_m^2),$$

for $m = 1, \dots, M$, and we denote the mappings $\mathbf{g}_1 = (g_1^1, \dots, g_1^M)^\top$ and $\mathbf{g}_2 = (g_2^1, \dots, g_2^M)^\top$. Then, the DC constraints in problem $(\mathbf{P}_U^{\text{DC}})$ can be rewritten as

$$\mathbf{g}_1 - \mathbf{g}_2 \leq 0.$$

With this representation, we can then use the same argument as that in the proof of Property 3 in [53] to prove the conclusion (b), which is thus omitted. □

The property (a) in 5.6 shows that we always search for better solutions in the feasible region, which is similar to the framework of classical interior-point methods. It also shows that we make improvement at each iteration and the sequence of objective function values converges to a certain value. The second property (b) ensures that all accumulation points of the sequence of solutions generated are KKT points of problem $(\mathbf{P}_U^{\text{DC}})$. If problem $(\mathbf{P}_U^{\text{DC}})$ has a unique KKT point or has only one KKT point that is better than the initial solution, 2 guarantees to converge to a global optimal solution of problem $(\mathbf{P}_U^{\text{DC}})$.

Lower bound

In order to evaluate how conservative the solution of the above upper bound approximation is when it is compared to the true solution of problem $(\tilde{\mathbf{P}})$, we need a procedure that could provide a lower bound for the optimal value of problem $(\tilde{\mathbf{P}})$. For this purpose, we apply the tangent approximation technique

where the function $\Phi(p_m)$ is approximated by its upper bound tangent piecewise linear function.

From Lemma 5.2, we have that $\Phi(p_m)$ is concave when $\alpha < \frac{1}{2}$. We select K different tangent functions:

$$l_i(p_m) = \bar{a}_i p_m + \bar{b}_i, \quad i = 1, \dots, K,$$

with

$$\bar{a}_i = \phi(\tau_i), \quad \bar{b}_i = \Phi(\tau_i) - \bar{a}_i \tau_i,$$

where τ_1, \dots, τ_K are K points with $\tau_i \in [\Phi^{-1}(1 - \alpha), \infty)$, $1 \leq i \leq K$. Therefore, we have

$$l(p_m) = \min_{i=1, \dots, K} l_i(p_m) \geq \Phi(p_m), \quad \forall p_m \in [\Phi^{-1}(1 - \alpha), \infty).$$

With the above tangent piecewise linear approximation $l(p_m)$ of $\Phi(p_m)$, we obtain the following lower bound approximation of problem (\mathbf{P}^C) .

$$\begin{aligned} (\mathbf{P}_{L_0}^{\text{DC}}) \quad & \min_{\mathbf{z}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}, \mathbf{p}, \boldsymbol{\sigma}} \quad g(\mathbf{z}) \\ \text{s.t.} \quad & \beta - \boldsymbol{\gamma}^\top \boldsymbol{\eta} \geq 1 - \alpha, \\ & \beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq \bar{a}_i p_m + \bar{b}_i, \quad m = 1, \dots, M, i = 1, \dots, K, \\ & \|\boldsymbol{\varphi}_k\| \leq \eta_k, \quad k = 1, \dots, I, \\ & p_m \sigma_m \leq d - \boldsymbol{\mu}_m^\top \mathbf{z}, \quad m = 1, \dots, M, \\ & \left\| \sum_m^{\frac{1}{2}} \mathbf{z} \right\| \leq \sigma_m, \quad m = 1, \dots, M, \\ & \mathbf{z} \in \mathbf{Z}. \end{aligned}$$

By approximating the biconvex term $p_m \sigma_m$ with Cheng and Lisser's method in [24], we have the following conclusion.

Theorem 5.7. *Replacing $\Phi(p_m)$ in problem (\mathbf{P}^C) by $l(p_m)$, we get the following convex approximation of problem (\mathbf{P}^{DC}) :*

$$\begin{aligned} (\mathbf{P}_L^{\text{DC}}) \quad & \min_{\mathbf{z}, \beta, \boldsymbol{\eta}, \boldsymbol{\varphi}, \mathbf{p}, \boldsymbol{\sigma}} \quad g(\mathbf{z}) \\ \text{s.t.} \quad & \beta - \boldsymbol{\gamma}^\top \boldsymbol{\eta} \geq 1 - \alpha, \\ & \beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq \bar{a}_i p_m + \bar{b}_i, \quad m = 1, \dots, M, i = 1, \dots, K, \\ & \|\boldsymbol{\varphi}_k\| \leq \eta_k, \quad k = 1, \dots, I, \\ & y_m \leq d - \boldsymbol{\mu}_m^\top \mathbf{z}, \quad m = 1, \dots, M, \\ & \left\| \sum_m^{\frac{1}{2}} \mathbf{z} \right\| \leq \sigma_m, \quad m = 1, \dots, M, \\ & \Phi^{-1}(1 - \alpha) \sigma_m \leq y_m, \quad m = 1, \dots, M, \\ & \kappa_m p_m \leq y_m, \quad m = 1, \dots, M, \\ & \mathbf{z} \in \mathbf{Z}, \end{aligned}$$

where $\kappa_m = \min_{\mathbf{z} \in \mathbf{Z}} \left\| \sum_m^{\frac{1}{2}} \mathbf{z} \right\|$. Then, the optimal value of problem $(\mathbf{P}_L^{\text{DC}})$ is a lower bound to that of problem (\mathbf{P}) .

Proof. Let

$$T^* = \left\{ \mathbf{p} : \beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq \Phi(p_m), m = 1, \dots, M \right\},$$

and

$$\bar{T}_S = \left\{ \mathbf{p} : \beta + \sum_{k=1}^I \mathbf{u}_{k,m}^\top \boldsymbol{\varphi}_k \leq \bar{a}_i p_m + \bar{b}_i, m = 1, \dots, M, i = 1, \dots, K \right\}.$$

It is obvious that $T^* \subset \bar{T}_S$.

In addition, for $m = 1, \dots, M$, we replace $p_m \sigma_m$ by y_m . As $p_m \geq \Phi^{-1}(1 - \alpha)$, the constraint $y_m = p_m \sigma_m$ can be relaxed by $\Phi^{-1}(1 - \alpha) \sigma_m \leq y_m$ and $\kappa_i p_m \leq y_m$, where $\kappa_m = \min_{\mathbf{z} \in \mathcal{Z}} \left\| \Sigma_m^{\frac{1}{2}} \mathbf{z} \right\|$. Therefore, the optimal value of problem $(\mathbf{P}_L^{\text{DC}})$ is a lower bound to that of problem (\mathbf{P}) . \square

5.3 Numerical Results

In this section, we apply the proposed data-driven robust chance constrained optimization problem (\mathbf{P}) to two practical problems. The data is generated randomly. We use the CVX package to solve all the approximation problems in this section with Matlab R2015a, on a Laptop with a 2.60 GHz Intel Core m5-4300M CPU and 4.0 GB RAM.

5.3.1 Convex approximation

In this subsection, we test the performance of our convex approximations by considering a data-driven robust production problem with the Cobb-Douglas utility.

Cobb-Douglas utility function delicately captures the practical characteristics of different multi-factor input-output contexts besides production. Here, we generalize the model in [32] by maximizing the Cobb-Douglas utility with chance constraints. Suppose that there are n goods, and the random price of good j is ξ_j , $j = 1, \dots, n$. Suppose the distribution of $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$ is contained in the uncertainty set D where $M = 3$ and $I = 4$. Let $\mathbf{z} = (z_1, \dots, z_n) \geq 0$ denote the quantities of n goods purchased in the market. Then, the agent solves the following optimization problem

$$\begin{aligned} (\mathbf{P}_P) \quad & \min_{\mathbf{z}} \quad \prod_{j=1}^n z_j^{-r_j} \\ & \text{s.t.} \quad \min_{f \in D} \mathbb{P}_f \{ \boldsymbol{\xi}^\top \mathbf{z} \leq W \} \geq 1 - \alpha, \\ & \quad \mathbf{z} \geq 0, \end{aligned}$$

where $n = 10$, $\alpha = 0.1$, $r_j = 0.1, j = 1, \dots, 10$, $W = 10000$. The problem (\mathbf{P}_P) can be approximated by piecewise linear approximation problems (\mathbf{P}_L^C) and (\mathbf{P}_U^C) , respectively, to find the lower bound and upper bound to its optimal value, respectively.

To generate the necessary data, we adopt the following framework in [106] to determine the parameters of the component distributions $f_1(\cdot)$, $f_2(\cdot)$ and $f_3(\cdot)$ in the uncertainty set D :

$$y_k^m = a_k^m + b_k^i y_{\text{base}}^m + e_k^m, \quad m = 1, \dots, 3, \quad k = 1, \dots, 10,$$

where y_k^m represents the random realization of the k th component of the random vector, corresponding to the m th component distribution $f_m(\cdot)$; a_k^m and b_k^m are constants which are randomly selected in $[0, 1]$ and $[1, 1.5]$, respectively; $\mathbf{y}_{\text{base}} = (y_{\text{base}}^1, y_{\text{base}}^2, y_{\text{base}}^3)^\top$ denotes the base random vector, which is normally distributed with the mean vector being $E(\mathbf{y}_{\text{base}}) = (5, 10, 15)$ and the covariance matrix being $\text{Cov}(\mathbf{y}_{\text{base}}) = \text{diag}(0.2, 0.4, 0.6)$; and e_k^m is a normally distributed residual term with zero mean and the variance randomly selected in $[0, 2 \times 10^{-2}]$, which is independent of \mathbf{y}_{base} .

In this simulation, 5000 samples are randomly generated according to the mixture distribution with the mixture weight vector being $\mathbf{w} = (0.3, 0.4, 0.3)$. With the generated data and the parameters of the component distributions $f_1(\cdot)$, $f_2(\cdot)$ and $f_3(\cdot)$, and setting $\epsilon = 0.02$ and $\Delta = 0.03$, we can determine the values of parameters $U_k, \gamma_k, k = 1, 2, 3, 4$, in the uncertainty set D given in 5.1. Under 5.2, the problem $(\mathbf{P}_{\mathbf{P}})$ is actually a convex programming problem, which can be approximated by problem $(\mathbf{P}_{\mathbf{L}}^{\mathbf{C}})$ to find a lower bound to its optimal value and by problem $(\mathbf{P}_{\mathbf{U}}^{\mathbf{C}})$ to find an upper bound to its optimal value, respectively.

By setting K to 3, 5, 10, and 20, respectively, we solve problem $(\mathbf{P}_{\mathbf{L}}^{\mathbf{C}})$ and problem $(\mathbf{P}_{\mathbf{U}}^{\mathbf{C}})$ four times and find four groups of respective lower bounds and upper bounds. The results are presented in Table 1. The first column in 5.1 gives the value of K . The second and third columns give the lower bound and the corresponding CPU time in seconds to solve problem $(\mathbf{P}_{\mathbf{L}}^{\mathbf{C}})$. The fourth and fifth columns give the upper bound and the CPU time in seconds to solve problem $(\mathbf{P}_{\mathbf{U}}^{\mathbf{C}})$. For better illustration, we compute the relative difference between the lower bound and the respective upper bound, which is shown in the last column of 5.1.

Table 5.1: Computational results of convex reformulation problems.

K	lower bound	time(s)	upper bound	time(s)	Gap(%)
3	0.01262726	4.1763	0.01388198	3.6563	9.04
5	0.01291510	6.4696	0.01359047	6.5156	4.97
10	0.01310726	8.0135	0.01326751	8.7344	1.21
20	0.01314893	11.7500	0.01314893	12.5469	0 ($< 10^{-8}$)

From Table 5.1, we can see that as K increases, the difference between the lower bound and corresponding upper bound monotonically decreases. When the number of tangent points is 20, these bounds become tight. Meanwhile, the solution times in the third and fifth columns tell us that our convex approximation problems can be solved within a few seconds. These results demonstrate that, with a suitable size of linear approximations, our convex approximations can provide a very good approximate solution quickly.

5.3.2 DC approximation

In this subsection, we examine the performance of our DC approximation by considering a data-driven robust portfolio selection problem under the safety-first principle.

Suppose there are n risky securities and one riskless security in the stock market. The random loss of the j th stock is $\xi_j, 1 \leq j \leq n$, and the proportion of the wealth invested in the j th asset is $z_j, 1 \leq j \leq n$. Let $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^\top$, $\mathbf{z} = (z_1, \dots, z_n)^\top$. Therefore, the loss of the portfolio \mathbf{z} can be determined as $\boldsymbol{\xi}^\top \mathbf{z}$. In addition, we assume that the distribution of $\boldsymbol{\xi}$ is contained in the uncertainty set D with $M = 3$. Then the data-driven robust portfolio selection model under the safety-first principle can be established as follows:

$$\begin{aligned}
 (\mathbf{P}_S) \quad & \min_{\mathbf{z}, R_L} R_L \\
 \text{s.t.} \quad & \min_{f \in D} \mathbb{P}_f \{ \boldsymbol{\xi}^\top \mathbf{z} \leq R_L \} \geq 1 - \alpha, \\
 & \mathbf{e}_n^\top \mathbf{z} \leq 1, \mathbf{z} \geq 0.
 \end{aligned}$$

where $n = 10$, $\alpha = 0.1$, \mathbf{e}_n denotes the 10-dimensional vector with all ones.

The data generation procedure is almost the same as that in 5.3.1. The only difference is that the mean vector and covariance matrix of \mathbf{y}_{base} are set as $E(\mathbf{y}_{\text{base}}) = (-1, -1.5, -2)$ and $\text{Cov}(\mathbf{y}_{\text{base}}) = \text{diag}(0.02, 0.04, 0.06)$, respectively. Similarly, 1000 samples are randomly generated, and we can then obtain the values of parameters $U_k, \gamma_k, k = 1, 2, 3, 4$. It is obvious that Assumption 5.2 does not hold for problem (\mathbf{P}_S) . Therefore, we solve problems $(\mathbf{P}_U^{\text{DC}})$ and $(\mathbf{P}_L^{\text{DC}})$ to find an upper bound and a lower bound for the optimal value of problem (\mathbf{P}_S) .

For the piecewise linear approximations, we set the number of points K to 3, 5, 10, 15, 20, 30, 40, 60, 80 respectively, and denote the corresponding point sets as $\mathbf{K}_3, \mathbf{K}_5, \mathbf{K}_{10}, \mathbf{K}_{15}, \mathbf{K}_{20}, \mathbf{K}_{30}, \mathbf{K}_{40}, \mathbf{K}_{60}, \mathbf{K}_{80}$ which satisfy $\mathbf{K}_3 \subset \mathbf{K}_5 \subset \mathbf{K}_{10} \subset \mathbf{K}_{15} \subset \mathbf{K}_{20} \subset \mathbf{K}_{30} \subset \mathbf{K}_{40} \subset \mathbf{K}_{60} \subset \mathbf{K}_{80}$.

The first column in Table 5.2 gives the value of K . The second and third columns give the lower bound and the CPU time in seconds. The fourth and fifth columns provide the upper bound and the CPU time in seconds. For better illustration, we compute the percentage difference between the lower bound and the corresponding upper bound, which are shown in the last column of Table 5.2.

We can see from Table 5.2 that our DC algorithms can find lower and upper bounds quickly. What's more important, as K increases, the gap between the lower bound and corresponding upper bound becomes smaller and smaller. When the number of tangent points reaches 80, the gap is very small and doesn't decrease significantly. Hence, the respective bounds provide a very good approximation to the optimal return value of (\mathbf{P}_S) .

All the above results demonstrate the practicality and efficiency of the algorithms proposed in this chapter for the solution of data-driven robust chance constrained problems.

Table 5.2: Computational results of DC reformulation problems.

K	lower bound	time(s)	upper bound	time(s)	Gap(%)
3	-0.822313	0.4212	-0.608109	3.2136	35.22
5	-0.822312	0.5460	-0.613951	3.3400	33.94
10	-0.822311	0.5928	-0.631720	3.6052	30.17
15	-0.822311	0.7020	-0.655087	3.7768	25.53
20	-0.822310	0.7176	-0.708550	3.9484	16.06
30	-0.822310	0.9204	-0.780821	5.9256	5.31
40	-0.822309	1.2012	-0.822271	7.8222	0.0046
60	-0.822309	1.6068	-0.822276	11.6689	0.0040
80	-0.822309	1.7004	-0.822276	13.1197	0.0040

5.4 Conclusion

In this chapter, we discuss a data-driven robust chance constrained problem. Starting from the available data, we construct a mixture distribution based uncertainty set of the ambiguous underlying distribution. And the uncertainty set takes the first four moments into consideration simultaneously. In addition to the robustness, the mixture distribution employed in our model enable us to reformulate the chance constraint. Moreover, we propose tight convex approximations for the problem under some conditions on the parameters for the uncertainty set. And we propose a tight DC approximation for the upper approximation and a relaxed convex approximation for the lower approximation in the general case. Finally, simulation experiments are carried out to illustrate the practicality and efficiency of our approach.

This chapter correrponds to the reference [20].

Chapter 6

Chance Constrained Stochastic Game Theory

In this chapter, we consider an n -player non-cooperative game with continuous strategy sets. The strategy set of each player contains a set of stochastic linear constraints. We model the stochastic linear constraints of each player as a joint chance constraint. We assume that the row vectors of a matrix defining the stochastic constraints of each player are pairwise independent. We model the stochastic constraints with normal distribution, elliptical distribution and distributionally robustness, respectively. Under certain conditions we show the existence of a Nash equilibrium for this game.

6.1 The model

We consider an n -player non-cooperative game. Let $I = \{1, 2, \dots, n\}$ be the set of players. A generic element of the set I is denoted by i . The payoffs of player i is defined by a function $u_i : \mathbb{R}_{++}^{m_1} \times \mathbb{R}_{++}^{m_2} \times \dots \times \mathbb{R}_{++}^{m_n} \rightarrow \mathbb{R}$, where $\mathbb{R}_{++}^{m_i}$ ($\mathbb{R}_+^{m_i}$) denotes the positive (non-negative) orthant of \mathbb{R}^{m_i} . The set $X^i \subset \mathbb{R}_{++}^{m_i}$ denotes the set of all strategies of player i . We assume X^i to be a convex and compact set. The product set $X = \prod_{i \in I} X^i$ denotes the set of all strategy profiles. We denote the set of all vectors of strategies of all the players except player i by $X^{-i} = \prod_{j=1, j \neq i}^n X^j$. The generic elements of X^i , X^{-i} , and X are denoted by x^i , x^{-i} , and x respectively. We define (y^i, x^{-i}) to be a strategy profile where player i chooses a strategy y^i and each player $j \in I$, $j \neq i$, chooses a strategy x^j . We consider the case where the strategies of player i are further constrained by the following stochastic linear constraints

$$A^i x^i \leq b^i, \tag{6.1}$$

where $A^i = [A_1^i, A_2^i, \dots, A_{K_i}^i]^T$ is a $K_i \times m_i$ random matrix, and $b^i \in \mathbb{R}^{K_i}$; T denotes the transposition. For each $k = 1, 2, \dots, K_i$, A_k^i is the k^{th} row of A^i . We consider the case where the constraints of player i given by (6.1) are jointly satisfied with at least a given probability level. Let α_i be a given probability level of player i . We formulate the stochastic linear constraints (6.1) as a joint chance

constraint given by

$$\mathbb{P}_{F_i}\{A^i x^i \leq b^i\} \geq \alpha_i, \quad (6.2)$$

where \mathbb{P} is a probability measure, F_i is the distribution of A^i . Therefore, for an $\alpha_i \in [0, 1]$, the feasible strategy set of player i is defined by

$$S_{\alpha_i}^i = \{x^i \in X^i \mid \mathbb{P}_{F_i}\{A^i x^i \leq b^i\} \geq \alpha_i\}, \quad i \in I.$$

We assume that the set $S_{\alpha_i}^i$ is non-empty, and the probability distribution of the random matrix A^i and the probability level vector $(\alpha_i)_{i \in I}$ are known to all the players. Then, the above chance-constrained game is a non-cooperative game with complete information. A strategy profile x^* is said to be a Nash equilibrium of a chance-constrained game if and only if for each $i \in I$

$$u_i(x^{i*}, x^{-i*}) \geq u_i(x^i, x^{-i*}), \quad \forall x^i \in S_{\alpha_i}^i.$$

Assumption 6.1. *For each player i , $i \in I$, the following conditions hold.*

1. *The payoff function $u_i(\cdot, x^{-i})$ is a concave function of x^i for every $x^{-i} \in X^{-i}$.*
2. *The payoff function $u_i(\cdot)$ is a continuous function of x .*

Such game theoretic situations arise in a renewable energy markets based on wind turbine and solar panels with n -players. Each player aims at maximizing his payoff subject to technical, operational and budget constraints. Both energy technologies are highly concerned by uncertainties. For the wind turbine case, the constraints are related to tensile strength, tip deflection rate, blade natural frequency, turbulence, turbine size. For the solar energy, the constraints are related to generation sites, storage and inter-regional power transmission, the size of the panels. Such constraints could be modeled as chance constraints and considered either as individual chance constraints or joint chance constraints. And it is well known that joint chance constraints are highly reliable compared to individual ones

6.2 Existence of Nash equilibrium with normal distribution

We consider the case where for each $i \in I$, the row vector A_k^i , $k = 1, 2, \dots, K_i$, follows a multivariate normal distribution with mean $\mu_k^i = (\mu_{k1}^i, \mu_{k2}^i, \dots, \mu_{km_i}^i)$ and a covariance matrix Σ_k^i , i.e., $A_k^i \sim \mathcal{N}(\mu_k^i, \Sigma_k^i)$. We assume Σ_k^i to be a positive definite matrix. Moreover, the row vectors are also independent. In this case we have the following results.

Lemma 6.1. *For each $i \in I$, let the row vector $A_k^i \sim \mathcal{N}(\mu_k^i, \Sigma_k^i)$ with positive definite covariance matrix Σ_k^i , $k = 1, 2, \dots, K_i$. Moreover, the row vectors of A^i , $i \in I$, are independent. Then, $S_{\alpha_i}^i$, $i \in I$, is a convex set when*

$\alpha_i > F(\max\{\sqrt{3}, \bar{v}_i\})$, where $F(\cdot)$ is the one-dimensional standard normal distribution function,

$$\bar{v}_i = \max_{k=1,2,\dots,K_i} 4\lambda_{\max}^{i,k} \left(\lambda_{\min}^{i,k} \right)^{-\frac{3}{2}} \|\mu_k^i\|,$$

and $\lambda_{\max}^{i,k}, \lambda_{\min}^{i,k}$ refer to the largest and smallest eigenvalues of Σ_k^i ; $\|\cdot\|$ denotes the Euclidean norm.

Proof. The proof follows from Theorem 5.1 of [50]. \square

Remark 6.1. If the value of \bar{v}_i is smaller than $\sqrt{3}$, $S_{\alpha_i}^i$ is a convex set when $\alpha_i > F(\sqrt{3}) \approx 0.958$.

Lemma 6.2. For each $i \in I$, let the row vector $A_k^i \sim \mathcal{N}(\mu_k^i, \Sigma_k^i)$ with positive definite covariance matrix Σ_k^i , $k = 1, 2, \dots, K_i$. Moreover, the row vectors of A^i , $i \in I$, are independent. Then, $S_{\alpha_i}^i$, $i \in I$, is a compact set provided

$$\alpha_i > \min_{k=1,2,\dots,K_i} F(\|(\Sigma_k^i)^{-1/2} \mu_k^i\|).$$

Proof. The proof directly follows from Theorem 2.3 of [49]. \square

For each $i \in I$, the set of best response strategies of player i for a fixed strategy profile x^{-i} of other players is given by

$$\mathcal{B}_i^{\alpha_i}(x^{-i}) = \{\bar{x}^i \in S_{\alpha_i}^i \mid u_i(\bar{x}^i, x^{-i}) \geq u_i(x^i, x^{-i}), \forall x^i \in S_{\alpha_i}^i\}.$$

Denote, $S_\alpha = \prod_{i \in I} S_{\alpha_i}^i$, where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$. Let $\mathcal{P}(S_\alpha)$ be the power set of S_α . Then, for an $\alpha \in [0, 1]^n$, define a set-valued map

$$G^\alpha : S_\alpha \rightarrow \mathcal{P}(S_\alpha)$$

such that

$$G^\alpha(x) = \prod_{i \in I} \mathcal{B}_i^{\alpha_i}(x^{-i}).$$

A point x is said to be a fixed point of $G^\alpha(\cdot)$ if $x \in G^\alpha(x)$. It is clear that a fixed point of $G^\alpha(\cdot)$ is a Nash equilibrium of a chance-constrained game.

Theorem 6.1. Consider an n -player non-cooperative game where the payoff function of player i , $i \in I$, satisfies the Assumption 6.1. The stochastic linear constraints of player i are jointly satisfied with at least a given probability $\alpha_i \in [0, 1]$. For each $i \in I$, let the row vector $A_k^i \sim \mathcal{N}(\mu_k^i, \Sigma_k^i)$ with positive definite covariance matrix Σ_k^i , $k = 1, 2, \dots, K_i$. Moreover, the row vectors of A^i , $i \in I$, are independent. Then, there exists a Nash equilibrium for a chance-constrained game for all $\alpha \in (\hat{\alpha}_1, 1] \times (\hat{\alpha}_2, 1] \times \dots \times (\hat{\alpha}_n, 1]$, where for each $i \in I$

$$\hat{\alpha}_i = \max\{\bar{\alpha}_i, \tilde{\alpha}_i\},$$

and

$$\bar{\alpha}_i = F\left(\max\{\sqrt{3}, \bar{v}_i\}\right), \text{ where } \bar{v}_i = \max_{k=1,2,\dots,K_i} 4\lambda_{\max}^k \left(\lambda_{\min}^k \right)^{-\frac{3}{2}} \|\mu_k^i\|,$$

$$\tilde{\alpha}_i = \min_{k=1,2,\dots,K_i} F \left(\left\| (\Sigma_k^i)^{-1/2} \mu_k^i \right\| \right).$$

Proof. Fix $\alpha \in (\hat{\alpha}_1, 1] \times (\hat{\alpha}_2, 1] \times \dots \times (\hat{\alpha}_n, 1]$. To show the existence of a Nash equilibrium for a chance-constrained game, it is enough to show that $G^\alpha(\cdot)$ has a fixed point. We show that $G^\alpha(\cdot)$ satisfies all the conditions of Kakutani fixed point theorem [58] as given below:

- (i) S_α is a non-empty, convex, and compact subset of a finite dimensional Euclidean space.
- (ii) $G^\alpha(x)$ is a non-empty and convex set for all $x \in S_\alpha$.
- (iii) The graph of set-valued map $G^\alpha(\cdot)$, defined by $\{(x, y) \mid y \in G^\alpha(x)\}$, is a closed subset of $S_\alpha \times S_\alpha$: if $(x_n, \bar{x}_n) \rightarrow (x, \bar{x})$ with $\bar{x}_n \in G^\alpha(x_n)$ for all n , then $\bar{x} \in G^\alpha(x)$.

S_α is a non-empty set because for each $i \in I$, $S_{\alpha_i}^i$ is a non-empty set. The convexity of S_α follows from Lemma 6.1 and the compactness of S_α follows from Lemma 6.2. To show the condition (ii), it is enough to show that $\mathcal{B}_i^{\alpha_i}(x^{-i})$ is a non-empty and convex set for all $i \in I$. The set $\mathcal{B}_i^{\alpha_i}(x^{-i})$ is non-empty because $u_i(\cdot, x^{-i})$ is a continuous function of x^i and $S_{\alpha_i}^i$ is a compact set. The set $\mathcal{B}_i^{\alpha_i}(x^{-i})$ is convex because $u_i(\cdot, x^{-i})$ is a concave function of x^i . Since, the payoff function $u_i(\cdot)$, $i \in I$, is a continuous function of x , then the closed graph condition can be proved using the similar arguments given in the proof of Theorem 3.2. [98] (see also Theorem 4.4 of [3]). \square

From Theorem 6.3, a Nash equilibrium for a chance-constrained game exists for sufficiently large values for α_i , $i \in I$. Therefore, in most of the cases we do not have an answer for the existence of a Nash equilibrium for a chance-constrained game defined in Section 6.1. In order to answer this question we first propose a new reformulation for (6.2).

Under independent and normally distributed assumption on matrix A^i , we have the following equivalent deterministic reformulation for the joint chance-constraint (6.2)

$$Q_{\alpha_i}^i = \begin{cases} (\mu_k^i)^T x^i + F^{-1} \left(\alpha_i^{z_k^i} \right) \left\| (\Sigma_k^i)^{1/2} x^i \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i, & (i) \\ \sum_{k=1}^{K_i} z_k^i = 1, & (ii) \\ z_k^i \geq 0, \quad \forall k = 1, 2, \dots, K_i, & (iii) \end{cases}$$

where $F^{-1}(\cdot)$ is a quantile function for a standard normal distribution [24]. For an $\alpha_i \in [0.5, 1]$, the set $Q_{\alpha_i}^i$ is a bi-convex set as it is a convex set in x^i (resp. $(z_k^i)_{k=1}^{K_i}$) for a fixed $(z_k^i)_{k=1}^{K_i}$ (resp. x^i).

We propose a new convex reformulation of the set $Q_{\alpha_i}^i$. Let $\alpha_i \in [0.5, 1]$, $i \in I$. For $0 \leq z_k^i \leq 1$, $F^{-1} \left(\alpha_i^{z_k^i} \right) \geq 0$ for all $\alpha_i \geq 0.5$. Therefore, the constraint (i) of

the set $Q_{\alpha_i}^i$ can be written as

$$(\mu_k^i)^T x^i + \left\| (\Sigma_k^i)^{1/2} \left(F^{-1} \left(\alpha_i^{z_k^i} \right) x^i \right) \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i. \quad (6.3)$$

We use a change of variables technique under logarithmic transformation [65]. The logarithmic transformation is well defined because $X^i \subset \mathbb{R}_{++}^{m_i}$. We transform the vector $x^i \in X^i$ into a vector $y^i \in \mathbb{R}^{m_i}$, where $y_j^i = \log x_j^i$, $j = 1, 2, \dots, m_i$. Then, constraint (6.16) can be written as

$$(\mu_k^i)^T e^{y^i} + \left\| (\Sigma_k^i)^{1/2} e^{\log F^{-1}(\alpha_i^{z_k^i}) \mathbb{1}_{m_i+y^i}} \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i,$$

where $\mathbb{1}_{m_i}$ is an $m_i \times 1$ vector of ones, and $e^{y^i} = (e^{y_1^i}, \dots, e^{y_{m_i}^i})^T$ and $e^{\log F^{-1}(\alpha_i^{z_k^i}) \mathbb{1}_{m_i+y^i}} = (e^{\log F^{-1}(\alpha_i^{z_k^i})+y_1^i}, \dots, e^{\log F^{-1}(\alpha_i^{z_k^i})+y_{m_i}^i})^T$. Therefore, we have the following deterministic reformulation for (6.2)

$$\tilde{Q}_{\alpha_i}^i = \begin{cases} (\mu_k^i)^T e^{y^i} + \left\| (\Sigma_k^i)^{1/2} e^{\log F^{-1}(\alpha_i^{z_k^i}) \mathbb{1}_{m_i+y^i}} \right\| \leq b_k^i, & \forall k = 1, 2, \dots, K_i, & (i) \\ \sum_{k=1}^{K_i} z_k^i = 1 & & (ii) \\ z_k^i \geq 0, \quad \forall k = 1, 2, \dots, K_i. & & (iii) \end{cases}$$

Let Y^i be an image of X^i under logarithmic function. Since, the logarithmic function is continuous and X^i is a compact set, Y^i is also a compact set. Broadly speaking, the convexity may not be preserved under logarithmic transformation. In this section, we consider the sets X^i for which the sets Y^i remain convex. We give hereafter few examples of convex sets X^i which are invariant under logarithmic transformation.

Example 6.1. Consider a set

$$X^i = \left\{ x^i \in \mathbb{R}_{++}^{m_i} \mid c^T x^i \leq h \right\},$$

where $c = (c_1, \dots, c_{m_i})^T \in \mathbb{R}_+^{m_i}$ and $h \in \mathbb{R}_+$ are all constant. Then,

$$Y^i = \left\{ y^i \in \mathbb{R}^{m_i} \mid \sum_{j=1}^{m_i} c_j e^{y_j^i} \leq h \right\},$$

is a convex set.

Example 6.2. Consider a set

$$X^i = \left\{ x^i \in \mathbb{R}_{++}^{m_i} \mid \sum_{j=1}^{m_i-1} x_j^i \leq x_{m_i}^i \right\}.$$

Then, Y^i can be reformulated as

$$Y^i = \left\{ y^i \in \mathbb{R}^{m_i} \mid \sum_{j=1}^{m_i-1} e^{y_j^i - y_{m_i}^i} \leq 1 \right\},$$

which is also a convex set.

Example 6.3. Consider a set

$$X^i = \left\{ x^i \in \mathbb{R}_{++}^{m_i} \mid \sum_{j=1}^{m_i} (x_j^i)^2 \leq h \right\},$$

where $h \in \mathbb{R}_{++}$ is a constant. Then,

$$Y^i = \left\{ y^i \in \mathbb{R}^{m_i} \mid \sum_{j=1}^{m_i} e^{2y_j^i} \leq h \right\},$$

is a convex set.

Example 6.4. Consider a set

$$X^i = \left\{ x^i \in \mathbb{R}_{++}^{m_i} \mid f(x^i) \leq h \right\},$$

where $f : \mathbb{R}_{++}^{m_i} \rightarrow \mathbb{R}_{++}$ is a log-convex and non-decreasing function of x^i , and $h \in \mathbb{R}_{++}$ is a constant. Then,

$$Y^i = \left\{ y^i \in \mathbb{R}^{m_i} \mid f(e^{y^i}) \leq h \right\},$$

is a convex set.

The reformulation of feasible strategy set $S_{\alpha_i}^i$ of player i , $i \in I$, is given by

$$\tilde{S}_{\alpha_i}^i = \left\{ (y^i, z^i) \in Y^i \times \mathbb{R}^{K_i} \mid (y^i, z^i) \in \tilde{Q}_{\alpha_i}^i \right\}.$$

Assumption 6.2. For each player i , $i \in I$, $u_i(\cdot, x^{-i})$ is a non-increasing function for every $x^{-i} \in X^{-i}$.

Assumption 6.3. For each $i \in I$ and $k = 1, 2, \dots, K_i$, all the components of Σ_k^i and μ_k^i are non-negative.

Under Assumption 6.3, we show that the set $\tilde{S}_{\alpha_i}^i$ is convex. It is enough to show that constraint (i) of $\tilde{Q}_{\alpha_i}^i$ is convex. We present Lemma 6.3 on the composition of convex functions. It is used to prove Lemma 6.4 which is the key to prove the convexity of the sets $\tilde{S}_{\alpha_i}^i$, $i \in I$.

Lemma 6.3. *Let $\Gamma_1, \Gamma_2 \subseteq \mathbb{R}$ and Γ_1, Γ_2 are convex sets. Suppose $f_1 : \Gamma_1 \rightarrow \Gamma_2$ is a convex function on Γ_1 , and $f_2 : \Gamma_2 \rightarrow \mathbb{R}$ is a non-decreasing and convex function on Γ_2 . Then, the composition function $f_2 \circ f_1$ is a convex function on Γ_1 .*

Proof. Consider any two points $x_1, x_2 \in \Gamma_1$ and $\lambda \in [0, 1]$. Since, f_1 is a convex function on Γ_1 , we have

$$f_1(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f_1(x_1) + (1 - \lambda)f_1(x_2).$$

Using the non-decreasing and convexity properties of f_2 on Γ_2 , we have

$$(f_2 \circ f_1)(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda(f_2 \circ f_1)(x_1) + (1 - \lambda)(f_2 \circ f_1)(x_2).$$

Hence, $f_2 \circ f_1$ is a convex function on Γ_1 . □

Lemma 6.4. *For each $i \in I$ and $k = 1, 2, \dots, K_i$, $\log F^{-1}(\alpha_i^{z_k^i})$ is a convex function of z_k^i on $[0, 1]$ for all $\alpha_i \in [F(1), 1]$ where $F(1) \approx 0.84$.*

Proof. Let $g_1 : [0, 1] \rightarrow [F(1), 1]$ such that $g_1(z_k^i) = \alpha_i^{z_k^i}$, and $g_2 : [F(1), 1] \rightarrow \mathbb{R}$ such that $g_2(p) = \log F^{-1}(p)$ be two functions. Then, the function composition $(g_2 \circ g_1)(z_k^i) = \log F^{-1}(\alpha_i^{z_k^i})$. From Lemma 6.3, $\log F^{-1}(\alpha_i^{z_k^i})$ is a convex function of z_k^i if $g_1(\cdot)$ is a convex function and $g_2(\cdot)$ is convex and non-decreasing function on their respective domains.

Since, $0 \leq z_k^i \leq 1$ and $F(1) \leq \alpha_i \leq 1$, then $\alpha_i^{z_k^i} \geq \alpha_i$. Therefore, the function $g_1(\cdot)$ is well defined and it is also a convex function of z_k^i . The function $g_2(\cdot)$ is a non-decreasing function because the quantile function $F^{-1}(\cdot)$ as well as $\log(\cdot)$ are non-decreasing functions. The only thing remains to show is that $\log F^{-1}(p)$ is a convex function. It is enough to show that the second order derivative of $\log F^{-1}(p)$ is non-negative. The second order derivative of $\log F^{-1}(p)$ can be written as

$$-\frac{\psi(y) + y\psi'(y)}{(y\psi(y))^2\psi(y)}.$$

where $y = F^{-1}(p)$, and $\psi(\cdot)$ is the probability density function of a standard normal distribution, and $\psi'(\cdot)$ is the first order derivate of $\psi(\cdot)$. We have

$$\psi(y) + y\psi'(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}(1 - y)(1 + y) \leq 0,$$

because $1 + y = 1 + F^{-1}(p) \geq 0$ and $1 - y = 1 - F^{-1}(p) \leq 0$ for all $p \in [F(1), 1]$. Hence, the second order derivative of $\log F^{-1}(p)$ is non-negative for all $p \in [F(1), 1]$. Therefore, $\log F^{-1}(\alpha_i^{z_k^i})$ is a convex function of z_k^i on $[0, 1]$ for all $\alpha_i \in [F(1), 1]$. □

Lemma 6.5. For each $i \in I$, let the convex set X^i be such that Y^i is a convex set. Let Assumption 6.3 hold. Then, the set $\tilde{S}_{\alpha_i}^i$, $i \in I$, is a convex set for all $\alpha_i \in [F(1), 1]$.

Proof. Fix $i \in I$ and $\alpha_i \in [F(1), 1]$. Then, $\left((\Sigma_k^i)^{1/2} e^{\log F^{-1}(\alpha_i^{z_k^i})} \mathbb{1}_{m_i+y^i} \right)$ is an $m_i \times 1$ vector. Each component of the vector is a non-negative linear combination of the convex functions. Hence, it is a vector of non-negative convex functions. The Euclidean norm is a convex function and it is also a non-decreasing function in each argument when the arguments are non-negative. Therefore, the composition function $\left\| (\Sigma_k^i)^{1/2} e^{\log F^{-1}(\alpha_i^{z_k^i})} \mathbb{1}_{m_i+y^i} \right\|$ is a convex function. The term $(\mu_k^i)^T e^{y^i}$ is a convex function because $\mu_k^i \geq 0$. Hence, the constraints

$$(\mu_k^i)^T e^{y^i} + \left\| (\Sigma_k^i)^{1/2} e^{\log F^{-1}(\alpha_i^{z_k^i})} \mathbb{1}_{m_i+y^i} \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i,$$

are convex. It is easy to see that the other constraints of $\tilde{S}_{\alpha_i}^i$ are convex. Hence, $\tilde{S}_{\alpha_i}^i$ is a convex set. \square

Theorem 6.2. Consider an n -player non-cooperative game where the payoff function of player i , $i \in I$, satisfies the Assumptions 6.1 and 6.2. The stochastic linear constraints of each player are jointly satisfied with at least a given probability $\alpha_i \in [0, 1]$. For each $i \in I$, let the row vector $A_k^i \sim \mathcal{N}(\mu_k^i, \Sigma_k^i)$ where mean vector μ_k^i and positive definite covariance matrix Σ_k^i , $k = 1, 2, \dots, K_i$, satisfies Assumption 6.3. Moreover, the row vectors of A^i , $i \in I$, are independent. Then, there exists a Nash equilibrium of a chance-constrained game for all $\alpha \in [F(1), 1]^n$.

Proof. Let $\alpha \in [F(1), 1]^n$. For each $i \in I$, define a composition function $C_i = -u_i \circ d_i$, where $d_i : \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \dots \times \mathbb{R}^{m_n} \rightarrow \mathbb{R}_{++}^{m_1} \times \mathbb{R}_{++}^{m_2} \times \dots \times \mathbb{R}_{++}^{m_n}$, such that

$$d_i(y^1, y^2, \dots, y^n) = (e^{y^1}, e^{y^2}, \dots, e^{y^n}).$$

Then, define a best response set for player i for a fixed $y^{-i} \in Y^{-i}$

$$\tilde{\mathcal{B}}_i^{\alpha_i}(y^{-i}) = \left\{ (\bar{y}^i, \bar{z}^i) \in \tilde{S}_{\alpha_i}^i \mid C_i(\bar{y}^i, y^{-i}) \leq C_i(y^i, y^{-i}), \forall (y^i, z^i) \in \tilde{S}_{\alpha_i}^i \right\}.$$

Denote $\tilde{S}_\alpha = \prod_{i \in I} \tilde{S}_{\alpha_i}^i$. Then, define a set-valued map

$$\tilde{G}^\alpha : \tilde{S}_\alpha \rightarrow \mathcal{P}(\tilde{S}_\alpha)$$

such that

$$\tilde{G}^\alpha(y, z) = \prod_{i \in I} \tilde{\mathcal{B}}_i^{\alpha_i}(y^{-i}).$$

It follows from Lemma 6.5 that \tilde{S}_α is a convex set. It is also a closed and bounded set. Both the functions $u_i(\cdot)$ and $d_i(\cdot)$ are continuous functions. Therefore, the

composition function $C_i(\cdot)$ is also a continuous function. Since, $\tilde{S}_{\alpha_i}^i$ is a compact set, then the best response set $\tilde{\mathcal{B}}_i^{\alpha_i}(y^{-i})$ is non-empty. The function $d_i(\cdot)$ is a convex function and $-u_i(\cdot, x^{-i})$ is a convex and non-decreasing function of x^i . Then, the composition function $C_i(\cdot, y^{-i})$ is a convex function of y^i which, in turn, implies that the best response set $\tilde{\mathcal{B}}_i^{\alpha_i}(y^{-i})$ is a convex set. Hence, $\tilde{G}^\alpha(y, z)$ is a non-empty and convex set for each (y, z) . The closed graph condition for $\tilde{G}^\alpha(\cdot)$ follows from the continuity of the functions $C_i(\cdot)$, $i \in I$. Therefore, from Kakutani fixed point theorem there exists a fixed point (y^*, z^*) for the set-valued map $\tilde{G}^\alpha(\cdot)$. Then, for each $i \in I$

$$C_i(y^{i*}, y^{-i*}) \leq C_i(y^i, y^{-i*}), \forall (y^i, z^i) \in \tilde{S}_{\alpha_i}^i.$$

Under the hypothesis of theorem, $\tilde{S}_{\alpha_i}^i$ is a reformulation of $S_{\alpha_i}^i$, where $x^i = e^{y^i}$. This implies

$$u_i(x^{i*}, x^{-i*}) \geq u_i(x^i, x^{-i*}), \forall x^i \in S_{\alpha_i}^i.$$

Hence, x^* is a Nash equilibrium of a chance-constrained game for all $\alpha \in [F(1), 1]^n$. \square

6.3 Existence of Nash equilibrium with elliptical distribution

In this section, we will prove the existence of Nash equilibrium with elliptical distribution. The definition of elliptical distribution is shown in subsection 2.1 and density function is given by (2.4).

Then, we assume that the row vector A_k^i , $k = 1, 2, \dots, K_i$ all follow some elliptical distributions, that is, $A_k^i \sim \text{Ellip}(\mu_k^i, \Sigma_k^i; \phi_k^i)$, $k = 1, \dots, K$, where ϕ_k^i is the characteristic function as defined in subsection 2.1. Moreover, $A_{k_r}^i$ and $A_{k_s}^i$ are independent of each other when $k_r \neq k_s$. In this case, with the help of Theorem 3 in [22], we can reformulated the joint chance-constraint (6.2) as

$$Q_{\alpha_i}^i = \begin{cases} (i) (\mu_k^i)^T x^i + (F_k^i)^{-1} \left(\alpha_{z_k^i} \right) \| (\Sigma_k^i)^{1/2} x^i \| \leq b_k^i, \forall k = 1, 2, \dots, K_i \\ (ii) \sum_{k=1}^{K_i} z_k^i = 1 \\ (iii) z_k^i \geq 0, \forall k = 1, 2, \dots, K_i, \end{cases} \quad (6.4)$$

where F_k^i is one-dimensional distribution function induced by the characteristic function $\phi_k^i(t) = \varphi_k^i(t^2)$, φ_k^i is the characteristic generator function.

The reformulation of feasible strategy set $S_{\alpha_i}^i$ of player i , $i \in I$, is given by

$$\hat{S}_{\alpha_i}^i = \left\{ (x^i, z^i) \in X^i \times \mathbb{R}_+^{K_i} \mid (x^i, z^i) \in Q_{\alpha_i}^i \right\}.$$

Denote $\hat{S}_\alpha = \prod_{i \in I} \hat{S}_{\alpha_i}^i$.

To prove the convexity of set $\hat{S}_{\alpha_i}^i$, we introduce the following definition.

Definition 6.1. ([50]) A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is called r -decreasing for some $r \in \mathbb{R}$ with 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 following proposition shows the thresholds of some elliptical distributions.

Proposition 6.1. ([22]) 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 ν degrees of freedom, for $0 < r < \nu + 1$ with the threshold $t^*(r) = \sqrt{\frac{r\nu}{\nu+1-r}}$;
3. Laplace distribution, for all $r > 0$ with the threshold $t^*(r) = \frac{r\sigma}{\sqrt{2}}$.

With the definition of threshold, we have the following theorem about the convexity of $\widehat{S}_{\alpha_i}^i$:

Theorem 6.3. For each $i \in I$, $\widehat{S}_{\alpha_i}^i$ is a convex set when

1. the densities associated with F_k^i are $(r_k^i + 1)$ -decreasing with thresholds $t_k^{i*}(r_k^i + 1) > 0$ for some $r_k^i > 1$, or at least 2-decreasing if $\mu_k^i = 0$;
2. the probability level α_i satisfies

$$\alpha_i > \max_k \begin{cases} F_k^i \left(\max \left\{ t_k^{i*}(r_k^i + 1), \frac{r_k^i + 1}{r_k^i - 1} (\lambda_{min}^k)^{-\frac{1}{2}} \|\mu_k^i\| \right\} \right), & \mu_k^i \neq 0 \\ F_k^i(t_k^{i*}(2)), & \mu_k^i = 0 \end{cases}$$

Proof. The proof follows directly from [22]. □

Similar with the case of normal distribution, we have the following theorem about the existence of Nash equilibrium with elliptical distribution. And the proof is exactly the same with the proof of Theorem 6.3.

Theorem 6.4. Consider an n -player non-cooperative game where the payoff function of player i , $i \in I$, satisfies the Assumption 6.1. The stochastic linear constraints of each player are jointly satisfied with at least a given probability $\alpha_i \in [0, 1]$. Let the row vector $A_{k_r}^i \sim \text{Ellip}(\mu_k^i, \Sigma_k^i; \phi_k^i)$, $k = 1, \dots, K_i$, of random matrix A^i . Moreover, $A_{k_r}^i$ and $A_{k_s}^i$ are independent of each other when $k_r \neq k_s$. Then, there always exists a Nash equilibrium for a chance-constrained game for all $\alpha \in (\bar{\alpha}_1, 1] \times (\bar{\alpha}_2, 1] \times \dots \times (\bar{\alpha}_n, 1]$, where for each $i \in I$

$$\bar{\alpha}_i = \max_k \begin{cases} F_k^i \left(\max \left\{ t_k^{i*}(r_k^i + 1), \frac{r_k^i + 1}{r_k^i - 1} (\lambda_{min}^k)^{-\frac{1}{2}} \|\mu_k^i\| \right\} \right), & \mu_k^i \neq 0, \\ F_k^i(t_k^{i*}(2)), & \mu_k^i = 0. \end{cases} \quad (6.5)$$

From Theorem 6.4, a Nash equilibrium for a chance-constrained game exists for very high values for α_i , $i \in I$. This is because the reformulation $\widehat{S}^i(\alpha_i)$ of the joint chance constraint (6.2) is not a convex set. Therefore, for most of the cases we do not have an answer for the existence of a Nash equilibrium for a

chance-constrained game. In order to answer this question we propose a new reformulation for (6.2).

From the constraints (i) in (6.4), we can find that the feasible strategy set $S^i(\alpha_i)$ is nonconvex in general. In the following, we will consider a convex reformulation of set $S^i(\alpha_i)$.

When $\alpha \geq 0.5$, we have $(F_k^i)^{-1}(\alpha_i^{z_k^i}) \geq 0$ where $0 \leq z_k \leq 1$. Therefore, the first group constraints in (6.4) can be reformulated as

$$(\mu_k^i)^T x + \left\| (\Sigma_k^i)^{\frac{1}{2}} \left((F_k^i)^{-1}(\alpha_i^{z_k^i}) x^i \right) \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K \quad (6.6)$$

We transform the vector $x^i \in X^i$ into a vector $y^i \in \mathbb{R}^{m_i}$, where $y_j^i = \ln x_j^i$, $j = 1, 2, \dots, m_i$. Then, constraint (6.6) can be written as

$$(\mu_k^i)^T e^{y^i} + \left\| (\Sigma_k^i)^{1/2} e^{\log(F_k^i)^{-1}(\alpha_i^{z_k^i}) \cdot \mathbb{1}_{m_i} + y^i} \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i,$$

where $\mathbb{1}_{m_i}$ is an $m_i \times 1$ vector of ones, and $e^{y^i} = (e^{y_1^i}, \dots, e^{y_{m_i}^i})^T$ and

$$e^{\log(F_k^i)^{-1}(\alpha_i^{z_k^i}) \cdot \mathbb{1}_{m_i} + y^i} = \left(e^{\log(F_k^i)^{-1}(\alpha_i^{z_k^i}) + y_1^i}, \dots, e^{\log(F_k^i)^{-1}(\alpha_i^{z_k^i}) + y_{m_i}^i} \right)^T.$$

Therefore, we have the following deterministic reformulation for (6.4)

$$\tilde{Q}_{\alpha_i}^i = \begin{cases} (i) \quad (\mu_k^i)^T e^{y^i} + \left\| (\Sigma_k^i)^{1/2} e^{\log(F_k^i)^{-1}(\alpha_i^{z_k^i}) \cdot \mathbb{1}_{m_i} + y^i} \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i, \\ (ii) \quad \sum_{k=1}^{K_i} z_k^i = 1 \\ (iii) \quad z_k^i \geq 0, \quad \forall k = 1, 2, \dots, K_i. \end{cases} \quad (6.7)$$

Let $Y^i = \{y^i \in \mathbb{R}^{m_i} \mid y_j^i = \ln x_j^i, x^i = (x_1^i, \dots, x_{m_i}^i)^T, x^i \in X^i\}$. The set Y^i is an image of X^i under logarithmic function. Since, the logarithmic function is continuous and X^i is a compact, Y^i is also compact set. The reformulation of feasible strategy set $S_{\alpha_i}^i$ of player i , $i \in I$, is given by

$$\tilde{S}_{\alpha_i}^i = \{(y^i, z^i) \in Y^i \times \mathbb{R}^{K_i} \mid (y^i, z^i) \in \tilde{Q}_{\alpha_i}^i\}.$$

Assumption 6.4. For each $i \in I$ and $k = 1, 2, \dots, K_i$, all the components of Σ_k^i and μ_k^i are non-negative.

Under Assumption 6.4, we show that the set $\tilde{S}_{\alpha_i}^i$ is convex. It is enough to show that the constraint (i) of $\tilde{Q}_{\alpha_i}^i$ is convex.

Lemma 6.6. $\log (F_k^i)^{-1}(\alpha_i^{z_k})$ is convex, if α_i satisfies

$$r_k^i \left((F_k^i)^{-1}(\alpha_i^{z_k}) \right) + (F_k^i)^{-1}(\alpha_i^{z_k}) \cdot (r_k^i)' \left((F_k^i)^{-1}(\alpha_i^{z_k}) \right) \leq 0, \quad \forall 0 \leq z_k \leq 1 \quad (6.8)$$

where r_k^i is the radial density corresponding to F_k^i , $(r_k^i)'$ is the first order derivative function of r_k^i , $k = 1, \dots, K$.

Proof. As $\alpha_i^{z_k}$ is a convex function, with Lemma 6.3 and the monotonicity of function $\log (F_k^i)^{-1}$, we only need to show the convexity of $\log (F_k^i)^{-1}$, which is equivalent with

$$-\frac{\psi_k^i(y) + y(\psi_k^i)'(y)}{(y\psi_k^i(y))^2\psi_k^i(y)} \geq 0.$$

where $y = (F_k^i)^{-1}(p)$, $\psi_k^i(\cdot)$ is the probability density function corresponding to F_k^i , $(\psi_k^i)'(\cdot)$ is the derivate of $\psi_k^i(\cdot)$. Then, $\log (F_k^i)^{-1}(p)$ is convex if and only if $\psi_k^i(y) + y(\psi_k^i)'(y) \leq 0$. Therefore, we have

$$r_k^i \left((F_k^i)^{-1}(\alpha_i^{z_k}) \right) + (F_k^i)^{-1}(\alpha_i^{z_k}) \cdot (r_k^i)' \left((F_k^i)^{-1}(\alpha_i^{z_k}) \right) \leq 0.$$

Since $\alpha_i^{z_k}$ is convex, if (6.8) is satisfied, we have $\log (F_k^i)^{-1}(\alpha_i^{z_k})$ is convex. \square

Example 6.5. (Normal distribution) For normal distribution, the radial density function can be expressed as $r_k^i(z) = \exp \left\{ -\frac{1}{2}z^2 \right\}$. Then inequality (6.8) can be written as $\alpha_i^{z_k} \geq F_k^i(1)$, $\forall 0 \leq z_k \leq 1$. Since $\alpha_i^{z_k} \geq \alpha_i$, $\forall 0 \leq z_k \leq 1$, then from Lemma 6.6, we have when $\alpha_i \geq F_k^i(1)$, $\log (F_k^i)^{-1}(\alpha_i^{z_k})$ is convex.

Example 6.6. (t distribution) For t distribution with degree ν , the radial density function can be expressed as $g(z) = \left(1 + \frac{1}{\nu}z^2\right)^{-(1+\nu)/2}$. Then inequality (6.8) can be written as $\alpha_i^{z_k} \geq F_k^i(1)$, $\forall 0 \leq z_k \leq 1$. Then from Lemma 6.6 we have when $\alpha_i \geq F_k^i(1)$, $\log (F_k^i)^{-1}(\alpha_i^{z_k})$ is convex.

Example 6.7. (Cauchy distribution) For Cauchy distribution, the radial density function can be expressed as $g(z) = (1 + z^2)^{-1}$. Then inequality (6.8) can be written as $\alpha_i^{z_k} \geq F_k^i(1)$, $\forall 0 \leq z_k \leq 1$. Then from Lemma 6.6 we have when $\alpha_i \geq F_k^i(1)$, $\log (F_k^i)^{-1}(\alpha_i^{z_k})$ is convex.

Example 6.8. (Laplace distribution) For Laplace distribution, the radial density function can be expressed as $g(z) = e^{-\sqrt{2}|z|}$. Then inequality (6.8) can be written as $\alpha_i^{z_k} \geq F_k^i\left(\frac{\sqrt{2}}{2}\right)$, $\forall 0 \leq z_k \leq 1$. Then from Lemma 6.6 we have when $\alpha_i \geq F_k^i\left(\frac{\sqrt{2}}{2}\right)$, $\log (F_k^i)^{-1}(\alpha_i^{z_k})$ is convex.

With Lemma 6.6, we can deduce the following lemma:

Lemma 6.7. For each $i \in I$, let the convex set $X^i \subset \mathbb{R}_+^{m_i}$ be such that Y^i is a convex set. Let Assumption 6.4 holds. Then, the set $\tilde{S}_{\alpha_i}^i$, $i \in I$, is a convex set for all α_i satisfying the conditions in Lemma 6.6.

Proof. Fix $i \in I$ and α_i satisfies the conditions in Lemma 6.6. Then,

$$\left((\Sigma_k^i)^{1/2} e^{\log(F_k^i)^{-1}(\alpha_i^{z_k^i})} \cdot \mathbb{1}_{m_i+y^i} \right)$$

is an $m_i \times 1$ vector. Each component of the vector is a linear combination of the convex functions where the coefficients are nonnegative. The norm is a convex function and it is also a nondecreasing function when the arguments are nonnegative. Therefore, The composition function $\left\| (\Sigma_k^i)^{1/2} e^{\log(F_k^i)^{-1}(\alpha_i^{z_k^i})} \cdot \mathbb{1}_{m_i+y^i} \right\|$ is a convex function. The term $(\mu_k^i)^T e^{y^i}$ is a convex function because $\mu_k^i \geq 0$. Hence, the constraints

$$(\mu_k^i)^T e^{y^i} + \left\| (\Sigma_k^i)^{1/2} e^{\log(F_k^i)^{-1}(\alpha_i^{z_k^i})} \cdot \mathbb{1}_{m_i+y^i} \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i,$$

are convex. It is easy to see that the other constraints of $\tilde{S}_{\alpha_i}^i$ are convex. Therefore, $\tilde{S}_{\alpha_i}^i$ is a convex set. \square

Hence, similar with Theorem 6.2, we have the following theorem about the existence of Nash equilibrium for elliptically distributed case.

Theorem 6.5. *Consider an n -player non-cooperative game where the payoff function of player i , $i \in I$, satisfies the Assumption 6.1. The stochastic linear constraints of each player are jointly satisfied with at least a given probability $\alpha_i \in [0, 1]$. Let the row vector $A_k^i \sim \text{Ellip}(\mu_k^i, \Sigma_k^i; \phi_k^i)$, $k = 1, \dots, K_i$, of random matrix A^i with the mean vector μ_k^i and the covariance matrix Σ_k^i which satisfy Assumption 6.4. Moreover, $A_{k_r}^i$ and $A_{k_s}^i$ are independent of each other when $k_r \neq k_s$. Then, there always exists a Nash equilibrium for a chance-constrained game for all α_i satisfying the conditions in Lemma 6.6.*

6.4 Existence of Nash equilibrium for distributionally robust model

As mentioned in section 6.1, we formulate the stochastic linear constraints (6.1) as a joint chance constraint given by (6.2):

$$\mathbb{P}_{F_i} \{A^i x^i \leq b^i\} \geq \alpha_i.$$

However, in practical, the assumption of full knowledge of the distribution F_i fails. In this case, we take the uncertainty of F_i into consideration. The only knowledge we have is that F_i is in some uncertainty set D_i . Therefore, the stochastic linear constraints (6.1) can be formulated as

$$\inf_{F_i \in D_i} \mathbb{P}_{F_i} \{A^i x^i \leq b^i\} \geq \alpha_i, \quad (6.9)$$

Therefore, for an $\alpha_i \in [0, 1]$, the feasible set of player i , $i \in I$, is defined by

$$S_{\alpha_i}^i = \left\{ x^i \in X^i \mid \inf_{F_i \in \mathcal{D}_i} \mathbb{P}_{F_i} \{ A^i x^i \leq b^i \} \geq \alpha_i \right\}. \quad (6.10)$$

For each $i \in I$, we assume $S_{\alpha_i}^i$ to be a nonempty set, and the uncertainty set D_i and the probability level vector $(\alpha_i)_{i \in I}$ are known to all the players. Then, above chance-constrained game is a non-cooperative game.

In the following, we consider four kinds of popular uncertainty sets: ϕ -divergence uncertainty set and three moment based uncertainty sets. With these four uncertainty sets of distribution of A^i , we will show the existence of Nash equilibrium of the non-cooperative game.

6.4.1 ϕ -divergence uncertainty set

In this subsection, we consider the case that the uncertainty sets $\mathcal{D}_i, i \in I$ are all defined as

$$\mathcal{D}_i = \{ f_i \mid D_\varphi(f_i \| f_i^0) \leq \epsilon_i \}, \quad (6.11)$$

where

$$D_\varphi(f_i \| f_i^0) = \int_{\Omega} \varphi \left(\frac{f_i(\xi)}{f_i^0(\xi)} \right) f_i^0(\xi) d\xi,$$

f_i^0 denotes the estimated density function of A^i . Furthermore, the corresponding estimated distribution of row vector A_k^i , $k = 1, 2, \dots, K_i$, is a multivariate normal distribution with mean $\mu_k^i = (\mu_{k1}^i, \mu_{k2}^i, \dots, \mu_{km_i}^i)$ and a covariance matrix Σ_k^i . $A_{k_r}^i$ and $A_{k_s}^i$ are independent of each other when $k_r \neq k_s$. Moreover, $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is a convex function on \mathbb{R}_+ such that

1. $\varphi(1) = 0$,
2. $0\varphi(x/0) := \begin{cases} p \lim_{p \rightarrow +\infty} \varphi(p)/p & \text{if } x > 0, \\ 0 & \text{if } x = 0, \end{cases}$
3. $\varphi(x) = +\infty$ for $x < 0$.

By using Theorem 1 in Jiang and Guan [57], we have

Theorem 6.6. *For player i , with the uncertainty set (6.11), the constraint (6.9) is equivalent to*

$$\mathbb{P}_{f_i^0} \{ A^i x^i \leq b^i \} \geq \alpha'_i, \quad (6.12)$$

where

$$\alpha'_{i0} = \inf_{\substack{z > 0, z_0 + \pi_{\alpha_i} z < l_\varphi, \\ \underline{m}(\varphi^*) \leq z_0 + z \leq \bar{m}(\varphi^*)}} \left\{ \frac{\varphi^*(z_0 + z) - z_0 - (1 - \alpha_i)z + \epsilon_i}{\varphi^*(z_0 + z) - \varphi^*(z_0)} \right\},$$

$\alpha'_i = \min\{\alpha'_{i0}, 1\}$ for $\alpha'_{i0} \in \mathbb{R}$, φ^* is the conjugate function of φ , $l_\varphi = \lim_{x \rightarrow +\infty} \varphi(x)/x$, $\underline{m}(\varphi^*) = \sup\{m \in \mathbb{R} : \varphi^* \text{ is a finite constant on } (-\infty, m]\}$ and $\bar{m}(\varphi^*) = \inf\{m \in$

$R: \varphi^* = +\infty\}$, and

$$\pi_{\alpha_i} = \begin{cases} -\infty & \text{if } \text{Leb}\{[f_0 = 0]\} = 0, \\ 0 & \text{if } \text{Leb}\{[f_0 = 0]\} > 0 \text{ and } \text{Leb}\{[f_0 = 0][A^i x^i \leq b^i]\} = 0, \\ 1 & \text{otherwise,} \end{cases}$$

$\text{Leb}\{\cdot\}$ represents the Lebesgue measure and $[f_0 = 0] := \{\xi \in \Omega : f_0(\xi) = 0\}$.

From the Proposition 2-4 in Jiang and Guan [57], we have the following proposition.

Proposition 6.2. 1. Suppose that D_φ is constructed by using the χ divergence of order 2 with $\varphi(x) = (x - 1)^2$ and $\alpha_i \geq 1/2$. Then

$$\alpha'_i = \alpha_i + \frac{\sqrt{\epsilon_i^2 + 4\epsilon_i\alpha_i(1 - \alpha_i)} - (2\alpha_i - 1)\epsilon_i}{2\epsilon_i + 2},$$

2. Suppose that D_φ is constructed by using the variation distance with $\varphi(x) = |x - 1|$. Then

$$\alpha'_i = \alpha_i + \frac{\epsilon_i}{2},$$

3. Suppose that D_φ is constructed by using the KL divergence with $\varphi(x) = x \log x - x + 1$. Then

$$\alpha'_i = \inf_{x \in (0,1)} \left\{ \frac{e^{-\epsilon_i x^{\alpha_i}} - 1}{x - 1} \right\}.$$

With Theorem 6.6, we can rewrite the set $S_{\alpha_i}^i$ as

$$S_{\alpha_i}^i = \left\{ x^i \in X^i \mid \mathbb{P}_{f_i^0} \{A^i x^i \leq b^i\} \geq \alpha'_i \right\}, \quad (6.13)$$

where α'_i is defined in Theorem 6.6. As the estimated distribution of row vector A_k^i is a multivariate normal distribution, the existence of Nash equilibrium can be derived directly from the case of normal distribution in section 6.2.

Theorem 6.7. Consider an n -player non-cooperative game where the payoff function of player i , $i \in I$, satisfies the Assumptions 6.1 and 6.2. The stochastic linear constraints of each player are jointly satisfied with at least a given probability $\alpha_i \in [0, 1]$ and uncertainty set D_i defined as (6.11). For each $i \in I$, mean vector μ_k^i and positive definite covariance matrix Σ_k^i , $k = 1, 2, \dots, K_i$, satisfies Assumption 6.3. Moreover, the row vectors of A^i , $i \in I$, are independent. Then, there exists a Nash equilibrium of a chance-constrained game for all $\alpha' \in [F(1), 1]^n$, where $\alpha' = (\alpha'_1, \dots, \alpha'_n)$ and $\alpha'_i, i = 1, \dots, n$, are defined in Theorem 6.6.

6.4.2 Moments based uncertainty set I

In this section, we consider the case that the uncertainty sets D_i , $i \in I$ are all defined as

$$D_i = \left\{ F_i \left| \begin{array}{l} \mathbb{E}[A_k^i] = \mu_k^i, \quad k = 1, \dots, K_i \\ \mathbb{E}[(A_k^i - \mu_k^i)(A_k^i - \mu_k^i)^T] = \Sigma_k^i, \quad k = 1, \dots, K_i \\ A_k^i \text{ and } A_j^i \text{ are independent when } k \neq j \end{array} \right. \right\} \quad (6.14)$$

In this case, from [15], we have the following deterministic reformulation for distributionally robust joint chance-constraint (6.9)

$$Q_{D_i}^i(\alpha_i) = \left\{ \begin{array}{l} (i) \quad (\mu_k^i)^T x^i + \sqrt{\frac{\alpha_i^{z_k^i}}{1 - \alpha_i^{z_k^i}}} \left\| (\Sigma_k^i)^{1/2} x^i \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i \\ (ii) \quad \sum_{k=1}^{K_i} z_k^i = 1 \\ (iii) \quad z_k^i \geq 0, \quad \forall k = 1, 2, \dots, K_i. \end{array} \right. \quad (6.15)$$

The reformulation of feasible strategy set $S^i(\alpha_i)$ of player i , $i \in I$, is given by

$$\bar{S}_{\alpha_i}^i = \left\{ (x^i, z^i) \in X^i \times \mathbb{R}_+^{K_i} \mid (x^i, z^i) \in Q_{D_i}^i(\alpha_i) \right\}.$$

Denote $\bar{S}_\alpha = \prod_{i \in I} \bar{S}_{\alpha_i}^i$.

To propose a convex reformulation of joint chance-constraint (6.9), we start with the deterministic reformulation (6.15). Therefore, the constraint (i) of (6.15) can be written as

$$(\mu_k^i)^T x^i + \left\| (\Sigma_k^i)^{1/2} \left(\sqrt{\frac{\alpha_i^{z_k^i}}{1 - \alpha_i^{z_k^i}}} \cdot x^i \right) \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i. \quad (6.16)$$

We use a change of variables technique under logarithmic transformation. The logarithmic transformation is well defined because $X^i \subset \mathbb{R}_+^{m_i}$. We transform the vector $x^i \in X^i$ into a vector $y^i \in \mathbb{R}^{m_i}$, where $y_j^i = \ln x_j^i$, $j = 1, 2, \dots, m_i$. Then, constraint (6.16) can be written as

$$(\mu_k^i)^T e^{y^i} + \left\| (\Sigma_k^i)^{1/2} e^{\frac{1}{2} \left(z_k^i \log \alpha_i - \log(1 - \alpha_i^{z_k^i}) \right) \cdot \mathbb{1}_{m_i} + y^i} \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i,$$

where $\mathbb{1}_{m_i}$ is an $m_i \times 1$ vector of ones, $e^{y^i} = \left(e^{y_1^i}, \dots, e^{y_{m_i}^i} \right)^T$ and

$$\begin{aligned} & e^{\frac{1}{2} \left(z_k^i \log \alpha_i - \log(1 - \alpha_i^{z_k^i}) \right) \cdot \mathbb{1}_{m_i} + y^i} \\ &= \left(e^{\frac{1}{2} \left(z_k^i \log \alpha_i - \log(1 - \alpha_i^{z_k^i}) \right) + y_1^i}, \dots, e^{\frac{1}{2} \left(z_k^i \log \alpha_i - \log(1 - \alpha_i^{z_k^i}) \right) + y_{m_i}^i} \right)^T. \end{aligned}$$

Therefore, we have the following deterministic reformulation for (6.9)

$$\tilde{Q}_{D_i}^i(\alpha_i) = \begin{cases} (i) \left\| (\mu_k^i)^T e^{y^i} + \left\| (\Sigma_k^i)^{1/2} e^{\frac{1}{2} \left(z_k^i \log \alpha_i - \log \left(1 - \alpha_i^{z_k^i} \right) \right)} \cdot \mathbf{1}_{m_i + y^i} \right\| \right\| \leq b_k^i, \\ (ii) \sum_{k=1}^{K_i} z_k^i = 1 \\ (iii) z_k^i \geq 0, \forall k = 1, 2, \dots, K_i. \end{cases} \quad (6.17)$$

Let $Y^i = \left\{ y^i \in \mathbb{R}^{m_i} \mid y_j^i = \ln x_j^i, x^i = (x_1^i, \dots, x_{m_i}^i)^T, x^i \in X^i \right\}$. The set Y^i is an image of X^i under logarithmic function. Since, the logarithmic function is continuous and X^i is a compact, Y^i is also compact set. The convexity need not to be preserved under logarithmic transformation. We consider the set X^i for which the set Y^i remains convex. Such sets indeed exists.

The reformulation of feasible strategy set $S_{\alpha_i}^i$ of player i , $i \in I$, is given by

$$\tilde{S}_{\alpha_i}^i = \left\{ (y^i, z^i) \in Y^i \times \mathbb{R}^{K_i} \mid (y^i, z^i) \in \tilde{Q}_{D_i}^i(\alpha_i) \right\}.$$

Assumption 6.5. For each $i \in I$ and $k = 1, 2, \dots, K_i$, all the components of Σ_k^i and μ_k^i are non-negative.

Under Assumption 6.5, we show that the set $\tilde{S}_{\alpha_i}^i$ is convex. It is enough to show that the constraint (i) of $\tilde{Q}_{D_i}^i(\alpha_i)$ is convex.

Lemma 6.8. Let $f : X \rightarrow Y$ be a nonincreasing concave function, $g : Z \rightarrow X$ be a convex function. Then we have that $f(g) : Z \rightarrow Y$ is a concave function.

Proof. To show that $f(g)$ is a concave function, it is enough to show that the second order derivative of $f(g)$ is nonpositive. The second order derivative of $f(g)$ can be written as

$$f''(g)(g')^2 + f'(g)g''.$$

Since f is nonincreasing and concave, g is convex, we have $f''(g) \leq 0$, $f'(g) \leq 0$ and $g'' \geq 0$. Therefore, $f''(g)(g')^2 + f'(g)g'' \leq 0$, which means $f(g)$ is concave. \square

With Lemma 6.8, we can deduce the following lemma:

Lemma 6.9. For each $i \in I$, let the convex set $X^i \subset \mathbb{R}_+^{m_i}$ be such that Y^i is a convex set. Let Assumption 6.5 holds. Then, the set $\tilde{S}_{\alpha_i}^i$, $i \in I$, is a convex set for all $\alpha_i \in [0, 1]$.

Proof. Fix $i \in I$. Then, $\left((\Sigma_k^i)^{1/2} e^{\frac{1}{2} \left(z_k^i \log \alpha_i - \log \left(1 - \alpha_i^{z_k^i} \right) \right)} \cdot \mathbf{1}_{m_i + y^i} \right)$ is an $m_i \times 1$ vector. Since $\log(1-p)$ is decreasing and concave with respect to p and $\alpha_i^{z_k^i}$ is convex with respect to z_k^i , from Lemma 6.8, we have that $\log \left(1 - \alpha_i^{z_k^i} \right)$ is concave with respect to z_k^i . Each component of the vector is a linear combination of the convex functions where the coefficients are nonnegative. The norm is a convex

function and it is also a nondecreasing function when the arguments are nonnegative. Therefore, The composition function $\left\| \left(\sum_k^i \right)^{1/2} e^{\frac{1}{2} \left(z_k^i \log \alpha_i - \log \left(1 - \alpha_i^{z_k^i} \right) \right) \cdot \mathbb{1}_{m_i + y^i}} \right\|$ is a convex function. The term $(\mu_k^i)^T e^{y^i}$ is a convex function because $\mu_k^i \geq 0$. Hence, the constraints

$$(\mu_k^i)^T e^{y^i} + \left\| \left(\sum_k^i \right)^{1/2} e^{\frac{1}{2} \left(z_k^i \log \alpha_i - \log \left(1 - \alpha_i^{z_k^i} \right) \right) \cdot \mathbb{1}_{m_i + y^i}} \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i,$$

are convex. It is easy to see that the other constraints of $\tilde{S}^i(\alpha_i)$ are convex. Therefore, $\tilde{S}_{\alpha_i}^i$ is a convex set. \square

Then, with the same proof statement of Theorem 6.2, we get the following theorem about existence of Nash equilibrium:

Theorem 6.8. *Consider an n -player non-cooperative game where the payoff function of player i , $i \in I$, satisfies the Assumptions 6.1 and 6.2. The stochastic linear constraints of each player are jointly satisfied with at least a given probability $\alpha_i \in [0, 1]$ and uncertainty set D_i defined as (6.14). The mean vectors μ_k^i and the covariance matrixes Σ_k^i in (6.14) satisfy Assumption 6.5. Then, there always exists a Nash equilibrium for a chance-constrained game for all $\alpha \in [0, 1]^n$.*

6.4.3 Moments based uncertainty set II

In this section, we consider the case that the uncertainty sets D_i , $i \in I$ are all defined as

$$D_i = \left\{ F_i \left| \begin{array}{l} \mathbb{E}[A_k^i] = \mu_k^i, \quad k = 1, \dots, K_i \\ \mathbb{E}[(A_k^i - \mu_k^i)(A_k^i - \mu_k^i)^T] \preceq \Sigma_k^i, \quad k = 1, \dots, K_i \\ A_k^i \text{ and } A_j^i \text{ are independent when } k \neq j \end{array} \right. \right\} \quad (6.18)$$

In this case, from [21], we have the following deterministic reformulation for distributionally robust joint chance-constraint (6.9)

$$Q_{D_i}^i(\alpha_i) = \begin{cases} (i) \quad (\mu_k^i)^T x^i + \sqrt{\frac{\alpha_i^{z_k^i}}{1 - \alpha_i^{z_k^i}}} \left\| \left(\sum_k^i \right)^{1/2} x^i \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i \\ (ii) \quad \sum_{k=1}^{K_i} z_k^i = 1 \\ (iii) \quad z_k^i \geq 0, \quad \forall k = 1, 2, \dots, K_i. \end{cases} \quad (6.19)$$

The reformulation of feasible strategy set $S_{\alpha_i}^i$ of player i , $i \in I$, is given by

$$\bar{S}_{\alpha_i}^i = \left\{ (x^i, z^i) \in X^i \times \mathbb{R}_+^{K_i} \mid (x^i, z^i) \in Q_{D_i}^i(\alpha_i) \right\}.$$

Denote $\bar{S}_\alpha = \prod_{i \in I} \bar{S}_{\alpha_i}^i$, which is same with the case in subsection 6.4.2. Then, the existence of Nash equilibrium is exactly same with the case in subsection 6.4.2.

Theorem 6.9. *Consider an n -player non-cooperative game where the payoff function of player i , $i \in I$, satisfies the Assumptions 6.1 and 6.2. The stochastic linear constraints of each player are jointly satisfied with at least a given probability $\alpha_i \in [0, 1]$ and uncertainty set D_i defined as (6.20). The mean vectors μ_k^i and the covariance matrixes Σ_k^i in (6.14) satisfy Assumption 6.5. Then, there always exists a Nash equilibrium for a chance-constrained game for all $\alpha \in [0, 1]^n$.*

6.4.4 Moments based uncertainty set III

In this section, we consider the case that the uncertainty sets D_i , $i \in I$ are all defined as

$$D_i = \left\{ F_i \left| \begin{array}{l} (\mathbb{E}[A_k^i] - \mu_k^i)^\top (\Sigma_k^i)^{-1} (\mathbb{E}[A_k^i] - \mu_k^i) \leq \gamma_{k1}^i, \quad k = 1, \dots, K_i \\ \mathbb{E}[(A_k^i - \mu_k^i)(A_k^i - \mu_k^i)^\top] \preceq \gamma_{k2}^i \Sigma_k^i, \quad k = 1, \dots, K_i \\ A_k^i \text{ and } A_j^i \text{ are independent when } k \neq j \end{array} \right. \right\} \quad (6.20)$$

By applying the Corollary 4 in Yang and Xu [103], we have the following deterministic reformulation for distributionally robust joint chance-constraint (6.9)

$$Q_{D_i}^i(\alpha_i) = \left\{ \begin{array}{l} (i) \quad k_2 \cdot \langle Q_k^i, \Sigma_k^i \rangle + \frac{\alpha_i^{z_k^i}}{1 - \alpha_i^{z_k^i}} \cdot r_k^i + (\mu_k^i)^\top x^i \\ \quad \quad \quad + \sqrt{k_1} \left\| (\Sigma_k^i)^{1/2} x^i \right\| \leq b_k^i, \quad \forall k = 1, 2, \dots, K_i \\ (ii) \quad \begin{pmatrix} Q_k^i & -\frac{x^i}{2} \\ -\frac{x^i}{2}^\top & r_k^i \end{pmatrix} \succeq 0, \\ (iii) \quad \sum_{k=1}^{K_i} z_k^i = 1 \\ (iv) \quad z_k^i \geq 0, \quad \forall k = 1, 2, \dots, K_i. \end{array} \right. \quad (6.21)$$

The reformulation of feasible strategy set $S_{\alpha_i}^i$ of player i , $i \in I$, is given by

$$\bar{S}_{\alpha_i}^i = \left\{ (x^i, z^i, r^i, Q^i) \in X^i \times \mathbb{R}_+^{K_i} \times \mathbb{R}_+^{K_i} \times \mathbb{S}_+^{K_i} \mid (x^i, z^i, r^i, Q^i) \in Q_{D_i}^i(\alpha_i) \right\}.$$

Here, \mathbb{S}_+ denote the set of semidefinite matrixes. Denote $\bar{S}_\alpha = \prod_{i \in I} \bar{S}_{\alpha_i}^i$.

To propose a convex reformulation of joint chance-constraint (6.9), we start with the deterministic reformulation (6.21). Therefore, because of the nonnega-

tivity of r_k^i , the constraints (i) and (ii) of (6.21) can be written as

$$\begin{aligned}
(i) \quad & k_2 \cdot \langle Q_k^i, \Sigma_k^i \rangle + \frac{\alpha_i^{z_k^i}}{1 - \alpha_i^{z_k^i}} + (\mu_k^i)^T \chi_k^i \\
& + \sqrt{k_1} \left\| (\Sigma_k^i)^{1/2} \chi_k^i \right\| \leq b_k^i (r_k^i)^{-1}, \quad \forall k = 1, 2, \dots, K_i \\
(ii) \quad & \begin{pmatrix} Q_k^i & -\frac{\chi^i}{2} \\ -\frac{\chi^i}{2}^\top & 1 \end{pmatrix} \succeq 0, \\
(iii) \quad & r_k^i \chi_k^i = x^i.
\end{aligned} \tag{6.22}$$

We use a change of variables technique under logarithmic transformation. The logarithmic transformation is well defined because $X^i \subset \mathbb{R}_+^{m_i}$. We transform the vector $x^i \in X^i$ into a vector $y^i \in \mathbb{R}^{m_i}$, where $y_j^i = \ln x_j^i$, $j = 1, 2, \dots, m_i$. And $r_k^i = e^{\gamma_k^i}$, $\chi_k^i = e^{\beta_k^i}$, where $k = 1, 2, \dots, K_i$. Therefore, we have the following deterministic reformulation for (6.9)

$$\tilde{Q}_{D_i}^i(\alpha_i) = \begin{cases} (i) \quad k_2 \cdot \langle Q_k^i, \Sigma_k^i \rangle + \frac{\alpha_i^{z_k^i}}{1 - \alpha_i^{z_k^i}} + (\mu_k^i)^T e^{\beta_k^i} \\ \quad + \sqrt{k_1} \left\| (\Sigma_k^i)^{1/2} e^{\beta_k^i} \right\| \leq b_k^i e^{-\gamma_k^i}, \quad \forall k = 1, 2, \dots, K_i \\ (ii) \quad \begin{pmatrix} Q_k^i & -\frac{e^{\beta_k^i}}{2} \\ -\frac{e^{\beta_k^i}}{2}^\top & 1 \end{pmatrix} \succeq 0, \\ (iii) \quad \gamma_k^i + \beta_k^i = y^i \\ (iv) \quad \sum_{k=1}^{K_i} z_k^i = 1 \\ (v) \quad z_k^i \geq 0, \quad \forall k = 1, 2, \dots, K_i. \end{cases} \tag{6.23}$$

Let $Y^i = \{y^i \in \mathbb{R}^{m_i} \mid y_j^i = \ln x_j^i, x^i = (x_1^i, \dots, x_{m_i}^i)^T, x^i \in X^i\}$, The reformulation of feasible strategy set $S_{\alpha_i}^i$ of player i , $i \in I$, is given by

$$\tilde{S}_{\alpha_i}^i = \left\{ (y^i, z^i, \gamma^i, \beta^i, Q^i) \in Y^i \times \mathbb{R}_+^{K_i} \times \mathbb{R}^{K_i} \times \mathbb{R}^{K_i} \times \mathbb{S}_+^{K_i} \mid (y^i, z^i, \gamma^i, \beta^i, Q^i) \in \tilde{Q}_{D_i}^i(\alpha_i) \right\}.$$

Assumption 6.6. For each $i \in I$ and $k = 1, 2, \dots, K_i$, b_k^i are negative.

Since $\frac{p}{1-p}$ is a convex and increasing function on $[0, 1)$ and $\alpha_i^{z_k^i}$ is convex for $\forall \alpha \in [0, 1]$, from Lemma 6.3, we have that function $\frac{\alpha_i^{z_k^i}}{1 - \alpha_i^{z_k^i}}$ is convex with respect to z_k^i for $\forall \alpha \in [0, 1]$. Hence, with Assumptions 6.2, 6.5 and 6.6, we have every constraints in $\tilde{Q}_{D_i}^i(\alpha_i)$ are convex. Hence, the set $\tilde{S}_{\alpha_i}^i$, $i \in I$, is a convex set for all $\alpha_i \in [0, 1]$.

Then, with the exactly the same proof of Theorem 6.2, we get the following theorem about existence of Nash equilibrium:

Theorem 6.10. *Consider an n -player non-cooperative game where the payoff function of player i , $i \in I$, satisfies the Assumptions 6.1 and 6.2. The stochastic linear constraints of each player are jointly satisfied with at least a given probability $\alpha_i \in [0, 1]$ and uncertainty set D_i defined as (6.11). The mean vectors μ_k^i and the covariance matrixes Σ_k^i in (6.11) satisfy Assumption 6.5. And b_k^i satisfy Assumption 6.6. Then, there always exists a Nash equilibrium for a chance-constrained game for all $\alpha \in [0, 1]^n$.*

6.5 Conclusion

In this chapter, we modeled an n -player non-cooperative game with continuous strategy sets containing stochastic linear constraints by chance constrained problem. And the row vectors in chance constraints of each player are pairwise independent. We considered the existence of Nash equilibrium for the chance constrained stochastic game in three cases: normal distribution, elliptical distribution and distributionally robustness, respectively. Under certain conditions we showed the existence of Nash equilibrium for this stochastic game in all above three cases, respectively.

The work about stochastic games with normal distribution corresponds to the reference [82].

Chapter 7

Conclusions and Prospects

7.1 Conclusions

Chance constrained problem was first introduced in 1959 [17]. Since then the theory and applications of chance constrained problem have been developed. Up to now, there are still some important issues about chance constrained problems unsolved. Therefore, the topics around the theory and applications of chance constrained problem are still attractive. Based on the above statements, this dissertation has been dedicated to addressing the chance constrained problems and their applications. And the following research results were obtained:

(1) In Chapter 3, we reviewed the work about geometric programs with joint chance constraints, where the stochastic parameters are normally distributed and independent of each other. As an extension, we considered a joint rectangular geometric chance constrained programs under elliptical distribution with independent components. By using the standard variable transformation, we derived a convex reformulation of rectangular geometric chance constrained programs. As there was a quantile function of elliptical distribution in the reformulation which is nonelementary, we proposed new tight convex approximations based on the variable transformation together with piecewise linear approximation method.

(2) In Chapter 4, we developed upper bounds for linear individual and joint chance constrained problems with independent matrix vector rows. The uncertainty was considered in the coefficient matrix. The deterministic approximations of probability inequalities were based on the one-side Chebyshev inequality, Chernoff inequality, Bernstein and Hoeffding inequalities. We derived various sufficient conditions related to the confidence parameter value under which the aforementioned approximations are convex and tractable. The approximations could be computed under different assumptions: more specifically Chebyshev inequality requires the knowledge of the first and second moments of the random variables while Bernstein and Hoeffding ones, their mean and support. On the contrary, Chernoff inequality required only the moment generating function of the random variables. Approximations based on piecewise linear and tangent were also provided in order to reduce further the computational complexity of the problem. Finally, numerical results on randomly generated data were discussed.

(3) In Chapter 5, we considered a mixture distribution based data-driven un-

certainty set which can characterize higher order moments information. When we can only estimate the proposed uncertainty set from data, we proposed a data-driven approach, which is based on confidence region for higher order moments. Such a data-driven uncertainty set can more efficiently match non-Gaussian characters of real random variables. Then, we used the proposed data-driven mixture distribution based uncertainty set in a distributionally robust individual linear chance constrained problem. We showed that, under certain conditions of parameters, the distributionally robust chance constrained problem can be reformulated as a convex programming problem, with the data-driven mixture distribution based uncertainty set. As the convex equivalent reformulation contains a quantile function, we further proposed two approximations leading to tight upper and lower bounds. Moreover, under much weaker conditions, we showed that, the distributionally robust chance constrained problem can be reformulated as a DC programming problem. We proposed a sequence convex approximation method to find a tight upper bound, and used relax convex approximation method to find a lower bound.

(4) In Chapter 6, we considered an n-player non-cooperative game with continuous strategy sets. We investigated the case where the strategy sets are stochastic in nature. We assumed that the strategy set of each player contains a set of stochastic linear constraints. We formulated the stochastic linear constraints of each player as a joint chance constraint. We assumed that the row vectors of the matrix defining stochastic linear constraints are independent and each row vector follows a multivariate normal distribution. Under certain conditions, we proposed a new convex reformulation for the joint chance constraints in this case. We showed that there always exists a Nash equilibrium of such a chance constrained game if the payoff function of each player satisfies certain assumptions.

7.2 Prospects

In this dissertation, we discuss some research fields about chance constrained problems, including chance constrained stochastic games, bounds for chance constraints, joint rectangular geometric chance constrained programs and data-driven chance constrained problems. However, there are still some open issues and future works about chance constrained problems worth studying:

(1) In Chapter 3, we reformulated a joint rectangular geometric chance constrained programs under elliptical distribution with independent components and certain conditions about parameters. Therefore, finding tight convex approximations with weaker conditions or under other distributions, such as the generalized hyperbolic distributions, are still open questions for joint rectangular geometric chance constrained programs.

(2) In Chapter 4, we developed upper bounds for linear individual and joint chance constrained problems with independent matrix vector rows. Then, proposing a bound for joint chance constrained problem with dependent matrix vector rows is an interesting issue. What's more, another future work will be devoted on the application of the bounds addressed in Chapter 4 to more general stochastic optimization problems with chance constraints.

(3) In Chapter 5, we focused on a data-driven robust chance constrained problem. The proposed approaches in Chapter 5 just focus on the individual linear chance constraints. Therefore, application for practical problems of our approaches is limited. Different approaches for joint nonlinear chance constrained problems still have not been proposed. A further research topic is how to apply our approaches to data-driven robust joint nonlinear chance constrained problems.

(4) In Chapter 6, we discuss the existence of Nash equilibrium with certain conditions under elliptical distribution and distributionally robust framework. How to characterize the Nash equilibria under elliptical distribution and distributionally robust framework still needs to be solved. In addition, proving Nash equilibrium with weaker conditions or under different distributions, such as generalized mixture distributions, is also an open question.

Bibliography

- [1] van Ackooij, W.: Eventual convexity of chance constrained feasible sets. *Optimization* **64**(5), 1263–1284 (2015). DOI 10.1080/02331934.2013.855211
- [2] Bakir, I., Boland, N., Dandurand, B., Erera, A.: Scenario set partition dual bounds for multistage stochastic programming: A hierarchy of bounds and a partition sampling approach. http://www.optimization-online.org/DB_FILE/2016/01/5311.pdf (2016)
- [3] Basar, T., Olsder, G.J.: *Dynamic non-cooperative game theory*, 2nd edn. SIAM, Philadelphia, PA (1999)
- [4] Ben-Tal, A., den Hertog, D., De Waegenaere, A., Melenberg, B., Rennen, G.: Robust solutions of optimization problems affected by uncertain probabilities. *Management Science* **59**, 341–357 (2013). DOI 10.1287/mnsc.1120.1641
- [5] Ben-Tal, A., El Ghaoui, L., Nemirovski, A.: *Robust Optimization*. Princeton University Press (2009)
- [6] Ben-Tal, A., Nemirovski, A.: Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming* **88**, 411–424 (2000). DOI 10.1007/PL00011380
- [7] Bertsimas, D., Gupta, V., Kallus, N.: Data-driven robust optimization. *Mathematical Programming* pp. 1–58 (2017). DOI 10.1007/s10107-017-1125-8
- [8] Birge, J.R., Louveaux, F.: *Introduction to Stochastic Programming*. Springer, New York, NY (2011)
- [9] Borell, C.: Convex set functions ind-space. *Periodica Mathematica Hungarica* **6**(2), 111–136 (1975). DOI 10.1007/BF02018814
- [10] Boyd, S., Kim, S.J., Vandenberghe, L., Hassibi, A.: A tutorial on geometric programming. *Optimization and Engineering* **8**(1), 67–127 (2007)
- [11] Brascamp, H.J., Lieb, E.H.: On extensions of the brunn-minkowski and prékopa-leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of Functional Analysis* **22**(4), 366–389 (1976). DOI 10.1016/0022-1236(76)90004-5

- [12] Calafiore, G., Campi, M.: Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming* **102**(1), 25–46 (2005). DOI 10.1007/s10107-003-0499-y
- [13] Calafiore, G., Dabbene, F.: *Probabilistic and Randomized Methods for Design under Uncertainty*. Springer London (2006)
- [14] Calafiore, G.C., Campi, M.C.: The scenario approach to robust control design. *IEEE Transactions on Automatic Control* **51**, 742–753 (2006). DOI 10.1109/TAC.2006.875041
- [15] Calafiore, G.C., El Ghaoui, L.: On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications* **130**, 1–22 (2006). DOI 10.1007/s10957-006-9084-x
- [16] Campi, M., Garatti, S.: The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization* **19**(3), 1211–1230 (2008). DOI 10.1137/07069821X
- [17] Charnes, A., Cooper, W.W.: Chance-constrained programming. *Management Science* **6**, 73–79 (1959). DOI 10.1287/mnsc.6.1.73
- [18] Chen, L., He, S., Zhang, S.: Tight bounds for some risk measures, with applications to robust portfolio selection. *Operations Research* **59**, 847–865 (2011). DOI 10.1287/opre.1110.0950
- [19] Chen, Z., Kuhn, D., Wiesemann, W.: Data-driven chance constrained programs over Wasserstein balls. http://www.optimization-online.org/DB_HTML/2018/06/6671.html (2018)
- [20] Chen, Z., Peng, S., Liu, J.: Data-driven robust chance constrained problems: A mixture model approach. *Journal of Optimization Theory and Applications* **179**, 1065–1085 (2018). DOI 10.1007/s10957-018-1376-4
- [21] Cheng, J., Delage, E., Lisser, A.: Distributionally robust stochastic knapsack problem. *SIAM Journal on Optimization* **24**, 1485–1506 (2014). DOI 10.1137/130915315
- [22] Cheng, J., Houda, M., Lisser, A.: Second-order cone programming approach for elliptically distributed joint probabilistic constraints with dependent rows. http://www.optimization-online.org/DB_FILE/2014/05/4363.pdf (2014)
- [23] Cheng, J., Houda, M., Lisser, A.: Chance constrained 0–1 quadratic programs using copulas. *Optimization Letters* **9**(7), 1283–1295 (2015). DOI 10.1007/s11590-015-0854-y
- [24] Cheng, J., Lisser, A.: A second-order cone programming approach for linear programs with joint probabilistic constraints. *Operations Research Letters* **40**(5), 325–328 (2012). DOI 10.1016/j.orl.2012.06.008

- [25] Cheng, J., Lisser, A.: A completely positive representation of 0–1 linear programs with joint probabilistic constraints. *Operations Research Letters* **41**(6), 597–601 (2013). DOI 10.1016/j.orl.2013.08.008
- [26] Chiang, M., Boyd, S.: Geometric programming duals of channel capacity and rate distortion. *IEEE Transactions on Information Theory* **50**(2), 245–258 (2004).
- [27] Delage, E., Ye, Y.: Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research* **58**, 595–612 (2010). DOI 10.1287/opre.1090.0741
- [28] Dentcheva, D., Prekopa, A., Ruszczyński, A.: Concavity and efficient points of discrete distributions in probabilistic programming. *Mathematical Programming* **89**, 55–77 (2000). DOI 10.1007/PL00011393
- [29] Duffie, D., Pan, J.: An overview of value at risk. *The Journal of Derivatives* **4**, 7–49 (1997). DOI 10.3905/jod.1997.407971
- [30] Duffin, R.J., Peterson, E.L., Zener, C.: *Geometric programming: theory and application*. Wiley, New York, United States (1967)
- [31] Dupačová, J.: Stochastic geometric programming: Approaches and applications. In: *Conference on Mathematical Methods in Economics*, pp. 163–169. Kostelec nad Cernými Lesy (2009)
- [32] Eaves, B.C.: Finite solution of pure trade markets with cobb-douglas utilities. *Economic Equilibrium: Model Formulation and Solution* **23**, 226–239 (1985). DOI 10.1007/BFb0121035
- [33] Ecker, J.: *Geometric programming: methods, computations and applications*. *SIAM Review* **22**, 338–362 (1980).
- [34] Eisenberger, I.: Genesis of bimodal distributions. *Technometrics* **6**(4), 357–363 (1964)
- [35] El Ghaoui, L., Oks, M., Oustry, F.: Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research* **51**, 543–556 (2003). DOI 10.1287/opre.51.4.543.16101
- [36] Erdogan, E., Iyengar, G.: Ambiguous chance constrained problems and robust optimization. *Mathematical Programming* **107**, 37–61 (2006). DOI 10.1007/s10107-005-0678-0
- [37] Escudero, L.F., Garín, A., Merino, M., Pérez, G.: The value of the stochastic solution in multistage problems. *TOP* **15**(1), 48–64 (2007). DOI 10.1007/s11750-007-0005-4
- [38] Esfahani, P.M., Kuhn, D.: Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations (2016)

- [39] Fang, K.T., Kotz, S., Ng, K.W.: Symmetric Multivariate and Related Distributions. Chapman and Hall, London (1990)
- [40] de Farias, D.P., Van Roy, B.: On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research* **29**(3), 462–478 (2004). DOI 10.1287/moor.1040.0094
- [41] Gorski, J., Pfeuffer, F., Klamroth, K.: Biconvex sets and optimization with biconvex functions: a survey and extensions. *Mathematical Methods of Operations Research* **66**(3), 373–407 (2007)
- [42] Grant, M., Boyd, S., Ye, Y.: CVX: Matlab software for disciplined convex programming (2008).
- [43] Gupta, V.: Near-optimal ambiguity sets for distributionally robust optimization. http://www.optimization-online.org/DB_FILE/2015/07/4983.pdf (2015)
- [44] Habib, M., McDiarmid, C., Ramirez-Alfonsin, J., Reed, B.: Probabilistic Methods for Algorithmic Discrete Mathematics. Springer, Berlin (1998)
- [45] Hall, J.A., Brorsen, W., Irwin, S.H.: The distribution of futures prices: A test of the stable paretian and mixture of normals hypotheses. *The Journal of Financial and Quantitative Analysis* **24**, 105–116 (1989). DOI 10.2307/2330751
- [46] Hanasusanto, G.A., Kuhn, D., Wallace, S.W., Zymler, S.: Distributionally robust multi-item newsvendor problems with multimodal demand distributions. *Mathematical Programming* **152**, 1–32 (2015). DOI 10.1007/s10107-014-0776-y
- [47] Hanasusanto, G.A., Roitch, V., Kuhn, D., Wiesemann, W.: A distributionally robust perspective on uncertainty quantification and chance constrained programming. *Mathematical Programming* **151**, 35–62 (2015). DOI 10.1007/s10107-015-0896-z
- [48] Hanasusanto, G.A., Roitch, V., Kuhn, D., Wiesemann, W.: Ambiguous joint chance constraints under mean and dispersion information. http://www.optimization-online.org/DB_FILE/2015/11/5199.pdf (2016)
- [49] Henrion, R.: Structural properties of linear probabilistic constraints. *Optimization* **56**, 425–440 (2007). DOI 10.1080/02331930701421046
- [50] Henrion, R., Strugarek, C.: Convexity of chance constraints with independent random variables. *Computational Optimization and Applications* **41**, 263–276 (2008). DOI 10.1007/s10589-007-9105-1
- [51] Henrion, R., Strugarek, C.: Convexity of chance constraints with dependent random variables: The use of copulae. In: M. Bertocchi, G. Consigli, M.A.H. Dempster (eds.) *Stochastic Optimization Methods in Finance and Energy: New Financial Products and Energy Market Strate-*

- gies, pp. 427–439. Springer New York, New York, NY (2011). DOI 10.1007/978-1-4419-9586-5_17
- [52] Hoeffding, W.: Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* **58**(301), 13–30 (1963). DOI 10.1080/01621459.1963.10500830
- [53] Hong, L.J., Yang, Y., Zhang, L.: Sequential convex approximations to joint chance constrained programs: A monte carlo approach. *Operations Research* **59**, 617–630 (2011). DOI 10.1287/opre.1100.0910
- [54] Houda, M.: Weak structural dependence in chance-constrained programming. In: *Proceedings of the 26th International Conference on Mathematical Methods in Economics*, pp. 198–203. Technical University of Liberec, Faculty of Economics, Liberec (2008)
- [55] Hu, Z., Hong, L.J.: Kullback-Leibler divergence constrained distributionally robust optimization. http://www.optimization-online.org/DB_FILE/2012/11/3677.pdf (2012)
- [56] Jagannathan, R.: Chance-constrained programming with joint constraints. *Operations Research* **22**, 358–372 (1974). DOI 10.1287/opre.22.2.358
- [57] Jiang, R., Guan, Y.: Data-driven chance constrained stochastic program. *Mathematical Programming* **158**, 291–327 (2016). DOI 10.1007/s10107-015-0929-7
- [58] Kakutani, S.: A generalization of Brouwer’s fixed point theorem. *Duke Mathematical Journal* **8**(3), 457–459 (1941)
- [59] Kataoka, S.: A stochastic programming model. *Econometrica* **31**, 181–196 (1963). DOI 10.2307/1910956
- [60] Kim, S.J., Boyd, S.P., Yun, S., Patil, D.D., Horowitz, M.A.: A heuristic for optimizing stochastic activity networks with applications to statistical digital circuit sizing. *Optimization and Engineering* **8**(4), 397–430 (2007)
- [61] Kotz, S., Nadarajah, S.: *Multivariate T-Distributions and Their Applications*. Cambridge University Press, Cambridge (2004)
- [62] Lagoa, C.M., Li, X., Sznaier, M.: Probabilistically constrained linear programs and risk-adjusted controller design. *SIAM Journal on Optimization* **15**, 938–951 (2005). DOI 10.1137/S1052623403430099
- [63] Li, B., Jiang, R., Mathieu, J.L.: Ambiguous risk constraints with moment and unimodality information. http://www.optimization-online.org/DB_FILE/2016/09/5635.pdf (2016)
- [64] Lin, Z., Bai, Z.: *Probability Inequalities*. Science Press Beijing and Springer-Verlag, Berlin Heidelberg (2010)

- [65] Liu, J., Lisser, A., Chen, Z.: Stochastic geometric optimization with joint probabilistic constraints. *Operations Research Letters* **44**, 687–691 (2016). DOI 10.1016/j.orl.2016.08.002
- [66] Lubin, M., Bienstock, D., Vielma, J.P.: Two-sided linear chance constraints and extensions (2015)
- [67] Luedtke, J., Ahmed, S.: A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization* **19**, 674–699 (2008). DOI 10.1137/070702928
- [68] Luedtke, J., Ahmed, S., Nemhauser, G.L.: An integer programming approach for linear programs with probabilistic constraints. *Mathematical Programming* **122**, 247–272 (2010). DOI 10.1007/s10107-008-0247-4
- [69] Luptáčik, M.: Geometric programming: Method and applications. *OR Spektrum* **2**(3), 129–143 (1981)
- [70] Maggioni, F., Allevi, E., Bertocchi, M.: Bounds in multistage linear stochastic programming. *Journal of Optimization Theory and Applications* **163**(1), 200–229 (2014). DOI 10.1007/s10957-013-0450-1
- [71] Maggioni, F., Allevi, E., Bertocchi, M.: Monotonic bounds in multistage mixed-integer stochastic programming. *Computational Management Science* **13**(3), 423–457 (2016). DOI 10.1007/s10287-016-0254-5
- [72] Maggioni, F., Pflug, G.: Bounds and approximations for multistage stochastic programs. *SIAM Journal on Optimization* **26**(1), 831–855 (2016). DOI 10.1137/140971889
- [73] Mccaughey, J.L.: *Dynamics of Markets: The New Financial Economics*. Cambridge University Press, New York (2009)
- [74] McNeil, A.J., Frey, R., Embrechts, P.: *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, New Jersey (2005)
- [75] Miller, B.L., Wagner, H.M.: Chance constrained programming with joint constraints. *Operations Research* **13**, 930–945 (1965). DOI 10.1287/opre.13.6.930
- [76] Nemirovski, A., Shapiro, A.: Convex approximations of chance constrained programs. *SIAM Journal on Optimization* **17**, 969–996 (2006). DOI 10.1137/050622328
- [77] Nemirovski, A., Shapiro, A.: Scenario approximations of chance constraints. *Probabilistic and Randomized Methods for Design under Uncertainty* pp. 3–47 (2006). DOI 10.1007/1-84628-095-8_1

- [78] Pagnoncelli, B.K., Ahmed, S., Shapiro, A.: Sample average approximation method for chance constrained programming: Theory and applications. *Journal of Optimization Theory and Applications* **142**(2), 399–416 (2009). DOI 10.1007/s10957-009-9523-6
- [79] van de Panne, C., Popp, W.: Minimum-cost cattle feed under probabilistic protein constraints. *Management Science* **9**(3), 405–430 (1963). DOI 10.1287/mnsc.9.3.405
- [80] Peel, D., McLachlan, G.J.: Robust mixture modeling using the t distribution. *Statistics and Computing* **10**, 339–348 (2000). DOI 10.1023/A:1008981510081
- [81] Peng, S., Lisser, A., Maggioni, F.: Bounds for probabilistic constrained problems. http://www.optimization-online.org/DB_FILE/2018/10/6884.pdf(2018)
- [82] Peng, S., Singh, V.V., Lisser, A.: General sum games with joint chance constraints. *Operations Research Letters* **46**(5), 482 – 486 (2018)
- [83] Peterson, E.L.: Geometric programming. *SIAM Review* **18**(1), 1–51 (1976)
- [84] Pflug, G.C., Pichler, A.: *Multistage Stochastic Optimization*. Springer International Publishing (2014)
- [85] Pintér, J.: Deterministic approximations of probability inequalities. *Mathematical Methods of Operations-Research* **33**(4), 219–239 (1989). DOI 10.1007/BF01423332
- [86] Popescu, I.: Robust mean-covariance solutions for stochastic optimization. *Operations Research* **55**, 98–112 (2007). DOI 10.1287/opre.1060.0353
- [87] Prékopa, A.: On probabilistic constrained programming. In: H.W. Kuhn (ed.) *Proceedings of the Princeton Symposium on Mathematical Programming*, pp. 113–138. Princeton University Press, Princeton, NJ (1970)
- [88] Prékopa, A.: Logarithmic concave measures with application to stochastic programming. *Acta Scientiarum Mathematicarum* **32**, 301–316 (1971)
- [89] Prékopa, A.: *Stochastic Programming*. Springer Netherlands (1995)
- [90] Prékopa, A., Yoda, K., Subasi, M.M.: Uniform quasi-concavity in probabilistic constrained stochastic programming. *Operations Research Letters* **39**(3), 188–192 (2011). DOI 10.1016/j.orl.2011.03.007
- [91] Prékopa, A., Vizvári, B., Badics, T.: Programming under probabilistic constraint with discrete random variable. In: F. Giannessi, S. Komlósi, T. Rapszák (eds.) *New Trends in Mathematical Programming: Homage to Steven Vajda*, pp. 235–255. Springer US, Boston, MA (1998). DOI 10.1007/978-1-4757-2878-1_18

- [92] Rao, S.S.: Engineering Optimization: Theory and Practice, 3 edn. Wiley (2009)
- [93] Ray, S., Lindsay, B.G.: The topography of multivariate normal mixtures. *The Annals of Statistics* **33**(5), 2042–2065 (2005)
- [94] Rockafellar, R.T., Wets, R.J.B.: Variational Analysis. Springer-Verlag, Berlin Heidelberg (1998)
- [95] Seong, K., Narasimhan, R., Cioffi, J.M.: Queue proportional scheduling via geometric programming in fading broadcast channels. *IEEE Journal on Selected Areas in Communications* **24**(8), 1593–1602 (2006).
- [96] Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on Stochastic Programming: Modeling and Theory. SIAM, Philadelphia (2009)
- [97] Sokolowski, J., Zolesio, J.P.: Introduction to Shape Optimization. Springer, Berlin Heidelberg (1992).
- [98] Singh, V.V., Jouini, O., Lissner, A.: Existence of Nash equilibrium for chance-constrained games. *Operations Research Letters* **44**, 640–644 (2016)
- [99] Szántai, T.: Evaluation of a special multivariate gamma distribution function. In: A. Prékopa, R.J.B. Wets (eds.) *Stochastic Programming 84 Part I*, pp. 1–16. Springer, Berlin, Heidelberg (1986). DOI 10.1007/BFb0121111
- [100] van der Vaart, A.W., Wellner, J.A.: Weak Convergence and Empirical Processes: With Applications to Statistics, 2th edn. Springer, New York (1996)
- [101] Wiebking, R.D.: Optimal engineering design under uncertainty by geometric programming. *Management Science* **23**(6), 644–651 (1977)
- [102] Xie, W., Ahmed, S.: On deterministic reformulations of distributionally robust joint chance constrained optimization problems. http://www.optimization-online.org/DB_FILE/2016/09/5639.pdf (2017)
- [103] Yang, W., Xu, H.: Distributionally robust chance constraints for non-linear uncertainties. *Mathematical Programming* **155**, 231–265 (2016). DOI 10.1007/s10107-014-0842-5
- [104] Zhang, Y., Jiang, R., Shen, S.: Distributionally robust chance-constrained bin packing (2016)
- [105] Zhao, C., Guan, Y.: Data-driven risk-averse stochastic optimization with Wasserstein metric. http://www.optimization-online.org/DB_FILE/2015/05/4902.pdf (2015)
- [106] Zhu, S., Fan, M., Li, D.: Portfolio management with robustness in both prediction and decision: A mixture model based learning approach. *Journal of Economic Dynamics and Control* **48**, 1–25 (2014). DOI 10.1016/j.jedc.2014.08.015

- [107] Zhu, S., Fukushima, M.: Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research* **57**, 1155–1168 (2009). DOI 10.1287/opre.1080.0684
- [108] Zymler, S., Kuhn, D., Rustem, B.: Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming* **137**, 167–198 (2013). DOI 10.1007/s10107-011-0494-7

Titre : Optimisation stochastique avec contraintes en probabilités et applications

Mots clés : Contraintes en probabilités, Programmation convexe, Distributionnellement robuste, Distributions mixtes, Théorie des jeux stochastiques

Résumé : L'incertitude est une propriété naturelle des systèmes complexes. Les paramètres de certains modèles peuvent être imprécis ; la présence de perturbations aléatoires est une source majeure d'incertitude pouvant avoir un impact important sur les performances du système. L'optimisation sous contraintes en probabilités est une approche naturelle et largement utilisée pour fournir des décisions robustes dans des conditions d'incertitude. Dans cette thèse, nous étudierons systématiquement les problèmes d'optimisation avec contraintes en probabilités dans les cas suivants :

En tant que base des problèmes stochastiques, nous passons d'abord en revue les principaux résultats de recherche relatifs aux contraintes en probabilités selon trois perspectives: les problèmes liés à la convexité en présence de contraintes probabilistes, les reformulations et les approximations de ces contraintes, et les contraintes en probabilités dans le cadre de l'optimisation distributionnellement robuste. Pour les problèmes d'optimisation géométriques stochastiques, nous étudions les programmes avec contraintes en probabilités géométriques rectangulaires jointes. A l'aide d'hypothèses d'indépendance des variables aléatoires elliptiquement distribuées, nous déduisons une reformulation des programmes à contraintes géométriques rectangulaires jointes. Comme la reformulation n'est pas convexe, nous proposons de nouvelles approximations convexes basées sur la transformation des variables ainsi que des méthodes d'approximation linéaire par morceaux. Nos résultats numériques montrent que nos approximations sont asymptotiquement serrées.

Lorsque les distributions de probabilité ne sont pas connues à l'avance ou que la reformulation des contraintes probabilistes est difficile à obtenir, des bornes obtenues à partir des contraintes en probabilités peuvent être très utiles. Par conséquent, nous développons quatre bornes supérieures pour les contraintes probabilistes individuelles, et jointes dont les vecteur-lignes de la matrice des contraintes sont indépendantes. Sur la base de l'inégalité unilatérale de Chebyshev, de l'inégalité de Chernoff, de l'inégalité de Bernstein et de l'inégalité de Hoeffding, nous proposons des approximations déterministes des contraintes probabilistes. En outre, quelques conditions suffisantes dans lesquelles les approximations susmentionnées sont convexes et solvables de manière efficace sont déduites. Pour réduire davan-

tage la complexité des calculs, nous reformulons les approximations sous forme de problèmes d'optimisation convexes solvables basés sur des approximations linéaires et tangentielles par morceaux. Enfin, des expériences numériques sont menées afin de montrer la qualité des approximations déterministes étudiées sur des données générées aléatoirement.

Dans certains systèmes complexes, la distribution des paramètres aléatoires n'est que partiellement connue. Pour traiter les incertitudes complexes en termes de distribution et de données d'échantillonnage, nous proposons un ensemble d'incertitude basé sur des données obtenues à partir de distributions mixtes. L'ensemble d'incertitude basé sur les distributions mixtes est construit dans la perspective d'estimer simultanément des moments d'ordre supérieur. Ensuite, à partir de cet ensemble d'incertitude, nous proposons une reformulation du problème robuste avec contraintes en probabilités en utilisant des données issues d'échantillonnage. Comme la reformulation n'est pas un programme convexe, nous proposons des approximations nouvelles et convexes serrées basées sur la méthode d'approximation linéaire par morceaux sous certaines conditions. Pour le cas général, nous proposons une approximation DC pour dériver une borne supérieure et une approximation convexe relaxée pour dériver une borne inférieure pour la valeur de la solution optimale du problème initial. Nous établissons également le fondement théorique de ces approximations. Enfin, des expériences numériques sont effectuées pour montrer que les approximations proposées sont pratiques et efficaces.

Nous considérons enfin un jeu stochastique à n joueurs non-coopératif. Lorsque l'ensemble de stratégies de chaque joueur contient un ensemble de contraintes linéaires stochastiques, nous modélisons les contraintes linéaires stochastiques de chaque joueur sous la forme de contraintes en probabilité jointes. Pour chaque joueur, nous supposons que les vecteurs lignes de la matrice définissant les contraintes stochastiques sont indépendants les uns des autres. Ensuite, nous formulons les contraintes en probabilité dont les variables aléatoires sont soit normalement distribuées, soit elliptiquement distribuées, soit encore définies dans le cadre de l'optimisation distributionnellement robuste. Sous certaines conditions, nous montrons l'existence d'un équilibre de Nash pour ces jeux stochastiques.



Title : Chance constrained problem and its applications

Keywords : Chance constraints, Convex programming, Distributionally robust, Mixture distribution, Stochastic game theory

Abstract : Chance constrained optimization is a natural and widely used approaches to provide profitable and reliable decisions under uncertainty. And the topics around the theory and applications of chance constrained problems are interesting and attractive. However, there are still some important issues requiring non-trivial efforts to solve. In view of this, we will systematically investigate chance constrained problems from the following perspectives.

As the basis for chance constrained problems, we first review some main research results about chance constraints in three perspectives: convexity of chance constraints, reformulations and approximations for chance constraints and distributionally robust chance constraints.

For stochastic geometric programs, we formulate consider a joint rectangular geometric chance constrained program. With elliptically distributed and pairwise independent assumptions for stochastic parameters, we derive a reformulation of the joint rectangular geometric chance constrained programs. As the reformulation is not convex, we propose new convex approximations based on the variable transformation together with piecewise linear approximation methods. Our numerical results show that our approximations are asymptotically tight.

When the probability distributions are not known in advance or the reformulation for chance constraints is hard to obtain, bounds on chance constraints can be very useful. Therefore, we develop four upper bounds for individual and joint chance constraints with independent matrix vector rows. Based on the one-side Chebyshev inequality, Chernoff inequality, Bernstein inequality and Hoeffding inequality, we propose deterministic approximations for chance constraints. In addition, various sufficient conditions under which the aforementioned approximations are convex and tractable are derived. To reduce further computational complexity, we reformulate the approximations as

tractable convex optimization problems based on piecewise linear and tangent approximations. Finally, based on randomly generated data, numerical experiments are discussed in order to identify the tight deterministic approximations.

In some complex systems, the distribution of the random parameters is only known partially. To deal with the complex uncertainties in terms of the distribution and sample data, we propose a data-driven mixture distribution based uncertainty set. The data-driven mixture distribution based uncertainty set is constructed from the perspective of simultaneously estimating higher order moments. Then, with the mixture distribution based uncertainty set, we derive a reformulation of the data-driven robust chance constrained problem.

As the reformulation is not a convex program, we propose new and tight convex approximations based on the piecewise linear approximation method under certain conditions. For the general case, we propose a DC approximation to derive an upper bound and a relaxed convex approximation to derive a lower bound for the optimal value of the original problem, respectively. We also establish the theoretical foundation for these approximations. Finally, simulation experiments are carried out to show that the proposed approximations are practical and efficient.

We consider a stochastic n-player non-cooperative game. When the strategy set of each player contains a set of stochastic linear constraints, we model the stochastic linear constraints of each player as a joint chance constraint. For each player, we assume that the row vectors of the matrix defining the stochastic constraints are pairwise independent. Then, we formulate the chance constraints with the viewpoints of normal distribution, elliptical distribution and distributionally robustness, respectively. Under certain conditions, we show the existence of a Nash equilibrium for the stochastic game.

