OPTIMAL SCENARIO GENERATION FOR HEAVY-TAILED CHANCE CONSTRAINED OPTIMIZATION

JOSE BLANCHET, FAN ZHANG, AND BERT ZWART

ABSTRACT. We consider a generic class of chance-constrained optimization problems with heavy-tailed (i.e., power-law type) risk factors. In this setting, we use the scenario approach to obtain a constant approximation to the optimal solution with a computational complexity that is uniform in the risk tolerance parameter. We additionally illustrate the efficiency of our algorithm in the context of solvency in insurance networks.

1. Introduction

Chance constrained optimization problems have a rich history in Operations Research. Introduced by Charnes et al. (1958), chance constrained optimization formulations have proved to be versatile in modeling and decision making in a wide range of settings. For example, Prekopa (1970) used these types of formulations in the context of production planning. The work of Bonami and Lejeune (2009) illustrates how to take advantage of chance constrained optimization formulations in the context of portfolio selection. In the context of power and energy control the use of chance constrained optimization is illustrated in Andrieu et al. (2010). These are just examples of the wide range of applications that have benefited (and continue to benefit) from chance constrained optimization formulations and tools.

Consequently, there has been a significant amount of research effort devoted to the solution of chance constrained optimization problems. Unfortunately, however, these types of problems are provably NP-hard in the worst case, see Luedtke et al. (2010). As a consequence, much of the methodological effort has been placed into developing a) solutions in the case of specific models; b) convex and, more generally, tractable relaxations; c) combinatorial optimization tools; d) Monte-Carlo sampling schemes and sample-average approximations. Of course, hybrid approaches are also developed, for example, combining relaxations of type b) with sample-average approximation associated with type d) methods.

Examples of type a) approaches include the study of Gaussian or elliptical distributions when ϕ is affine both in L and x. In this case, the problem admits a conic programming formulation, which can be efficiently solved, see Lagoa et al. (2005). An example of type b) approach is provided in Nemirovski and Shapiro (2006a). This approach introduces a tractable constraint based on probabilistic inequalities (e.g., Chebyshev bounds). Type c) methods are based on branch and bounding algorithms, which connect squarely with the class of tools studied in areas such as integer programming, see Ahmed and Shapiro (2008); Luedtke et al. (2010).

The method we consider in this paper is the scenario approach, which can be placed in the setting of type d) in the above taxonomy. The scenario approach is introduced and studied in Calafiore and Campi (2005) and is further developed in a series of papers, including Calafiore and Campi (2006); Nemirovski and Shapiro (2006b). Following these papers, we consider the following standard family of chance constrained optimization problems:

$$(CCP_\delta) \qquad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathrm{P}(\phi(x,L) > 0) \leq \delta, \\ & x \in \mathbb{R}^{d_x}. \end{array}$$

where $x \in \mathbb{R}^{d_x}$ is a d_x-dimensional decision vector. L is a d_l-dimensional random vector in $\mathbb{R}^{d_l}$. The elements of L are often referred to as risk factors; the function $\phi : \mathbb{R}^{d_x} \times \mathbb{R}^{d_l} \to \mathbb{R}$ is often assumed to be convex in x and often models a cost constraint; the parameter $\delta > 0$ is the risk level of the tolerance.

The scenario approach is the most popular generic method for (approximately) solving chance constrained optimization. The idea is to sample a number N of scenarios and enforce the constraint in all of these scenarios. The intuition is that if for any scenario, say L_i, the constraint $\phi(L_i, x) < 0$ is convex in x, and $\delta > 0$ is small, we expect that by suitably choosing N the constrained regions can be relaxed by enforcing $\phi(L_i, x) < 0$ for all $1 \leq i \leq N$, leading to a good and, in some sense, tractable (if N is of moderate size) approximation of the chance constrained region. Of course, this intuition is correct only when $\delta > 0$ is small and we expect the choice of N to be largely influenced by this asymptotic regime.

By choosing N sufficiently large, the scenario approach allows obtaining both upper and lower bounds which become asymptotically tighter as $\delta \to 0$. In a celebrated paper, Calafiore and Campi (2006) provide rigorous support for this claim. In particular, given a confidence level $\beta \in (0,1)$, if $N \geq (2/\delta) \times \log(1/\beta) + 2d + (2d/\delta) \times \log(2/\delta)$, with probability at least $1 - \beta$, the optimal solution of the scenario approach relaxation is feasible for the original chance constrained problem and, therefore, an upper bound to the problem is obtained.

Unfortunately, the required sample size of N grows with $(1/\delta) \times \log(1/\delta)$ as δ becomes small, limiting the scope of the scenario methods in applications. Motivated by this, Nemirovski and Shapiro (2006b) developed a method that lowers the required sample size to the order of $\log(1/\delta)$, making additional assumptions on the function ϕ (which is taken to be bi-affine), and the risk factors L, which are to be assumed light-tailed. Specifically, the moment generating function $E[\exp(sL)]$ is assumed to be finite in a neighborhood of the origin. No guarantee is given in terms of how far the upper bound is from the optimal value function of the problem as $\delta \to 0$.

In the present paper, we focus on improving the scalability of N in terms of $1/\delta$ for the practically important case of heavy-tailed risk factors. Heavy-tailed distributions appear in a wide range of applications in science, engineering and business, see e.g., Embrechts et al. (2013), Wierman and Zwart (2012), but, in some aspects, are not as well understood as light-tails. One reason is that techniques from convex duality cannot be applied as the moment generating function of L does not exist in a neighborhood of 0. In addition, probabilistic inequalities, exploited in Nemirovski and Shapiro (2006b), do not hold in this setting. Only very recently, a versatile algorithm for heavy-tailed rare event simulation has been developed in Chen et al. (2019).

The main contribution of our paper is an algorithm that provides a sample complexity for N which is bounded in $1/\delta$, assuming a versatile class of heavy-tailed distributions for L. Specifically, we shall assume that L follows a semi-parametric class of models known as multivariate regular variation, which is quite standard in multivariate heavy-tail modeling, cf. Embrechts et al. (2013); Resnick (2013). A precise definition is given in Section 4. Moreover, our estimator is shown to be within a constant factor to the solution to (CCP_δ) with high probability, uniformly as $\delta \to 0$. We are not aware of other approaches that provide a uniform performance guarantee of this type (i.e. a constant approximation) as $\delta \to 0$.

We illustrate our assumptions and our framework with a risk problem of independent interest. This problem consists in computing a collective salvage fund in a network of financial entities whose liabilities and payments are settled in an optimal way using the Eisenberg-Noe model, see Eisenberg and Noe (2001). The salvage fund is computed to minimize its size in order to guarantee a probability of collective default after settlements of less than a small prescribed margin.

The rest of the paper is organized as follows. In Section 2, we introduce the minimal salvage fund problem as a particular application of chance constraint optimization. We employ the minimal salvage fund problem as a running example to provide a concrete and intuitive explanation for the concepts we introduce throughout the paper. Our main algorithmic contribution is given in Section 3. In that section, we introduce the main idea of our method and the basis for its intuition. This intuition is rooted in ideas originating from rare event simulation. Our algorithm requires the construction of several auxiliary functions and sets. How to do this is detailed in Section 4, in which we also present several additional technical assumptions required by our constructions. In Section 4, we also explain that our procedure results in an estimate which is within a constant factor of the optimal solution of the underlying chance constrained problem with high probability as $\delta \to 0$. In Section 5 we show that the assumptions imposed are valid in our motivating example (as well as a second example with quadratic cost structure inside the probabilistic constraint). Numerical results for the salvage fund example are provided in Section 6. Throughout our discussion in each section we present a series of results which summarize the main ideas of our constructions. To keep the discussion fluid, we present the corresponding proofs at the end of the sections, unless otherwise indicated.

Notations: in the sequel, $\mathbb{R}_+ = [0, +\infty)$ is the set of non-negative real numbers, $\mathbb{R}_{++} = (0, +\infty)$ is the set of positive real numbers, and $\overline{\mathbb{R}} = [-\infty, +\infty]$ is the extended real line. A column vector with zeros is denoted by $\mathbf{0}$, and a column vector with ones is denoted by $\mathbf{1}$. For any matrix Q, the transpose of Q is denoted by Q^T; the Frobenius norm of Q is denoted by $\|Q\|_F$. The identity matrix is denoted by I. For two column vectors $x, y \in \mathbb{R}^d$, we say $x \preceq y$ if and only if $y - x \in \mathbb{R}^d_+$. For $\alpha \in \mathbb{R}$ and $x \in \mathbb{R}^d$, we use $\alpha \cdot x$ to denote the scalar multiplication of x with α. For $\alpha \in \mathbb{R}$ and $E \subseteq \mathbb{R}^d$, we define $\alpha \cdot E = \{\alpha \cdot x \mid x \in E\}$. The optimal value of an optimization problem (P) is denoted by $\mathrm{Val}(P)$. We also use Landau's notation. In particular, if $f(\cdot)$ and $g(\cdot)$ are non-negative real valued functions, we write $f(t) = O(g(t))$ if $f(t) \leq c_0 \times g(t))$ for some $c_0 \in (0, \infty)$ and $f(t) = \Omega(g(t))$ if $f(t) \geq g(t))/c_0$ for some $c_0 \in (0, \infty)$.

2. Minimal Salvage Fund

Suppose that there are d entities or firms, which we can interpret as (re)insurance firms. Let $L = (L_1, \ldots, L_d)$ denotes the vector of incurred losses by each firm, where L_i denotes the total incurred loss that entity i is responsible to pay. We assume that L follows a multivariate heavy-tailed distribution, in a way made precise later on. Let $Q = (Q_{i,j} : i, j \in \{1, \ldots, d\})$ be a deterministic matrix where $Q_{i,j}$ denotes the amount of money received by entity j when entity i pays one dollar. We assume that $Q_{i,j} \geq 0$ and $\sum_{j=1}^d Q_{i,j} < 1$. Let $x = (x_1, \ldots, x_d)$ denotes the total amount that the salvage fund allocated to each entity, and $y^* = (y_1^*, \ldots, y_d^*)$ denotes the total amount paid by each entity in equilibrium. The equilibrium payment is determined by the following optimization problem:

$$y^* = y^*(x, L) = \arg\max\{\mathbf{1}^T y \mid 0 \preceq y \preceq L, \quad \left(I - Q^T\right) y \preceq x\}.$$

In words, the system maximizes the payments subject to the constraint that nobody pays more than what they have (in equilibrium), and nobody pays more than what they owe. Notice that $y^* = y^*(x, L)$ is also a random variable (the randomness comes from L) satisfying $\mathbf{0} \preceq y^* \preceq L$. Suppose that entity i bankrupts if the deficit $L_i - y_i^* \geq m$, where $m \geq 0$ is a given constant. We are interested in finding the minimal amount of salvage fund that ensures no bankruptcy happens with probability at least $1 - \delta$. The problem can be formulated as a chance constraint programming problem as follows

$$\begin{array}{lll} & \text{minimize} & \mathbf{1}^T x \\ (1) & \text{subject to} & \mathrm{P}(\|L - y^*(x, L)\|_\infty > m) \leq \delta, \\ & & x \in \mathbb{R}^d_{++}. \end{array}$$

Now we write the problem (1) into standard form. Note that $\|L - y^*(x, L)\|_\infty > m$ if and only if $\phi(x, L) > 0$, where $\phi(x, L)$ is defined as follows

$$\phi(x, L) := \min\{b - m \mid (L - y) \preceq b \cdot \mathbf{1}, \ \left(I - Q^T\right) y \preceq x, \ y \succeq \mathbf{0}\}.$$

Therefore, problem (1) is equivalent to

$$\begin{array}{lll} & \text{minimize} & \mathbf{1}^T x \\ (2) & \text{subject to} & \mathrm{P}(\phi(x, L) > 0) \leq \delta, \\ & & x \in \mathbb{R}^d_{++}. \end{array}$$

As mentioned in the introduction, a popular approach to solve the chance constraint problem proceeds by using the scenario approach developed by Calafiore and Campi (2006). They suggest to approximate the probabilistic constraint $\mathrm{P}(\phi(x, L) > 0) \leq \delta$ by N sampled constraints $\phi(x, L^{(i)}) \leq 0$ for $i = 1, \ldots, N$, where $\{L^{(1)}, \ldots, L^{(N)}\}$ are independent samples. Instead of solving the original chance constraint problem (CCP_δ), which is usually intractable, we turn to solve the following optimization problem

$$(SP_N) \qquad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \phi(x, L^{(i)}) \leq 0, \quad i = 1, \dots, N, \\ & x \in \mathbb{R}^{d_x}. \end{array}$$

The total sample size N should be large enough to ensure the feasible solution to the sampled problem (SP_N) is also a feasible solution to the original problem (CCP_δ) with a high confidence level. According to Calafiore and Campi (2006), for any given confidence level parameter $\beta \in (0, 1)$, if

$$N \geq \frac{2}{\delta} \log \frac{1}{\beta} + 2d + \frac{2d}{\delta} \log \frac{2}{\delta},$$

then any feasible solution to the sampled optimization problem (SP_N) is also a feasible solution to (CCP_δ) with probability at least $1 - \beta$. However, when δ is small, the total number of sampled constraints is of order $\Omega((1/\delta) \log(1/\delta))$, which could be a problem for implementation. For example, as we shall see in Section 6, when $\beta = 1 - 10^{-6}$, $d = 15$ and $\delta = 10^{-3}$, the number of sampled constraints N is required to be larger than 2×10^5. In contrast, our method only requires to sample 10^3 constraints.

3. General Algorithmic Idea

To facilitate the development of our algorithm, we introduce some additional notation and a desired technical property. As we shall see, if the technical property is satisfied, then there is a natural way to construct a scenario approach based algorithm that only requires $O(1)$ of total sampled constraints.

Let $F_\delta \subseteq \mathbb{R}^{d_x}$ denote the feasible region of the chance constraint optimization problem (CCP_δ), i.e.,

$$F_\delta := \{x \in \mathbb{R}^{d_x} \mid \mathrm{P}(\phi(x, L) > 0) \leq \delta\}. \tag{3}$$

Here, the subscript δ is involved to emphasize that the feasible region F_δ is parametrized by the risk level δ. For any fixed $x \in \mathbb{R}^{d_x}$, let $V_x := \{L \in \mathbb{R}^{d_l} \mid \phi(x, L) > 0\}$ denote the *violation event at* x.

Property 1. For any $\delta > 0$, there exist a set $O_\delta \subseteq \mathbb{R}^{d_x}$, and an event $C_\delta \subseteq \mathbb{R}^{d_l}$ that satisfy the following statements.

a) The feasible set F_δ is a subset of O_δ.
b) The event C_δ contains the violation event V_x for any $x \in O_\delta$.
c) There exist a constant $M > 0$ independent of δ such that $\mathrm{P}(L \in C_\delta) \leq M \cdot \delta$.

In the rest of this paper, we will refer to O_δ as *the outer approximation set*, and C_δ as *the uniform conditional event*. A graphical illustration of O_δ and C_δ is shown in Figure 1.

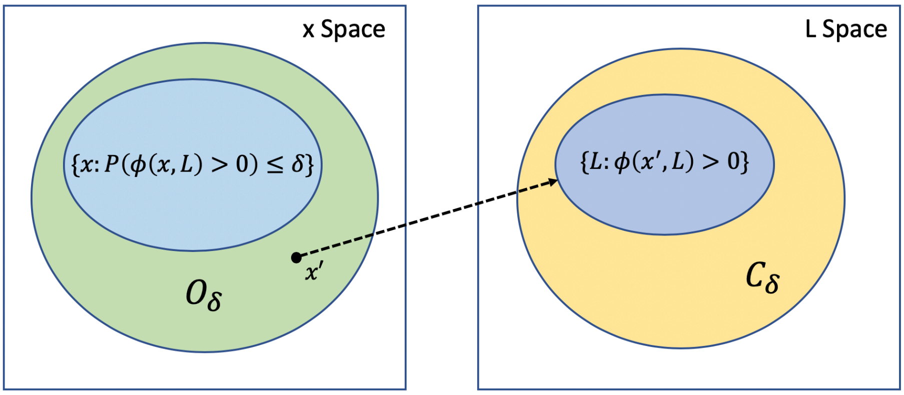

FIGURE 1. Pictorial illustration of O_δ and C_δ.

Now, given O_δ and C_δ that satisfies Property 1, we define the conditional sampled problem $(CSP_{\delta,N'})$:

$$(CSP_{\delta,N'}) \qquad \begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & \phi(x, L_\delta^{(i)}) \leq 0, & i = 1, \dots, N'. \\ & x \in O_\delta. & \end{array}$$

where $L_\delta^{(i)}$ are i.i.d. samples generated from the conditional distribution $(L|L \in C_\delta)$.

We now present our main result. The proof of Theorem 1 will be presented in Section 3.1.

Theorem 1. Suppose that Property 1 is imposed, and let $\beta > 0$ be a given confidence level.

(1) Let N' be any integer that satisfies

$$N' \geq (2M)\log(1/\beta) + 2d + (2dM)\log(2M). \tag{4}$$

With probability at least $1 - \beta$, if the conditional sampled problem $(CSP_{\delta,N'})$ is feasible, then its optimal solution $x_N^* \in F_\delta$ and $\mathrm{Val}\,(CSP_{\delta,N'}) \geq \mathrm{Val}\,(CCP_\delta)$.

(2) Let N' be any integer such that $N' \leq \beta\delta^{-1}\mathrm{P}(L \in C_\delta)$. Assume that the chance constraint problem (CCP_δ) is feasible. Then, with probability at least $1 - \beta$, $\mathrm{Val}\,(CCP_\delta) \geq \mathrm{Val}\,(CSP_{\delta,N'})$.

Remark 1. Note that the lower bound given in (4) is independent of δ. Therefore, Theorem 1 shows that the chance constraint problem (CCP_δ) can be approximated by $(CSP_{\delta,N'})$ with sample complexity bounded uniformly as $\delta \to 0$, as long as Property 1 is satisfied.

Remark 2. Efficiently generating samples of $(L|L \in C_\delta)$ when $\delta \to 0$ requires rare event simulation techniques. For example, when L is light-tailed, exponential tilting can be applied to achieve $O(1)$ sample complexity uniformly in δ; when L is heavy-tailed, with the help of specific problem structure, one can apply importance sampling, see Blanchet and Liu (2010), or Markov Chain Monte Carlo, see Gudmundsson and Hult (2014), to design an efficient sampling scheme. The specific structure of our salvage fund example results in C_δ being the complement of a box, which makes the sampling very tractable if the element of L are independent.

Even if the aforementioned rare event simulation techniques are hard to apply in practice, we can still apply a simple acceptance-rejection procedure to sample the conditional distribution $(L|L \in C_\delta)$. It cost $O(1/\delta)$ samples of L on average to get one sample of $(L|L \in C_\delta)$, since $\mathrm{P}(L \in C_\delta) = O(\delta)$. Consequently, the total complexity for generating $L_\delta^{(i)}, i = 1, \ldots, N'$ and solving $(CSP_{\delta,N'})$ is $O(1/\delta)$, which is still much more efficient than the scenario approach in Calafiore and Campi (2006), which requires computational complexity $O(((1/\delta)\log(1/\delta))^3)$ for solving a linear programming problem with $O((1/\delta)\log(1/\delta))$ sampled constraints by the interior point method.

Although Property 1 seems to be restrictive at first glance, we are still able to construct the sets O_δ and C_δ for a rich class of functions $\phi(L, x)$, including the constraint function for the minimal salvage fund problem. As we shall see in the proof of Theorem 1, once O_δ and C_δ are constructed the sampled problem $(CSP_{\delta,N'})$ is a tractable approximation to the problem (CCP_δ). We explain how to construct the sets O_δ and C_δ in the next section under some additional assumptions. These assumptions relate in particular to the distribution of L. It turns out that, if L is heavy-tailed, the construction of O_δ and C_δ becomes tractable.

3.1. **Proof of Theorem 1.** If Property 1 is satisfied, (CCP_δ) is equivalent to

$$\begin{array}{lll} & \text{minimize} & c^T x \\ (5) & \text{subject to} & \mathrm{P}(\phi(x, L) > 0 \mid L \in C_\delta) \leq \delta / \mathrm{P}(L \in C_\delta), \\ & & x \in O_\delta \subseteq \mathbb{R}^{d_x}. \end{array}$$

Let $\delta' := \delta/\mathrm{P}(L \in C_\delta) \geq 1/M$ denote the risk level in the equivalent problem (5). The sampled optimization problem related to problem (5) is given by

$$\begin{array}{lll} & \text{minimize} & c^T x \\ (CSP_{\delta,N'}) & \text{subject to} & \phi(x, L_\delta^{(i)}) \leq 0, \quad i = 1, \ldots, N', \\ & & x \in O_\delta, \end{array}$$

where the $L_\delta^{(i)}$ are independently sampled from $\mathrm{P}(\cdot \mid L \in C_\delta)$. Notice that

$$N' \geq 2M \log \frac{1}{\beta} + 2d + 2dM \log 2M \geq \frac{2}{\delta'} \log \frac{1}{\beta} + 2d + \frac{2d}{\delta'} \log \frac{2}{\delta'}.$$

According to (Calafiore and Campi, 2006, Corollary 1 and Theorem 2), with probability at least $1 - \beta$, if the sampled problem $(CSP_{\delta,N'})$ is feasible, then the optimal solution to problem $(CSP_{\delta,N'})$ is feasible to the chance constraint problem (5), thus it is also feasible for (CCP_δ). The proof of the first statement is complete.

Now we turn to prove the second statement. Note that the equivalence between (CCP_δ) and (5) is still valid, so it is sufficient to compare the optimal values of (5) and $(CSP_{\delta,N'})$. By applying (Calafiore and Campi, 2006, Theorem 2) again, we have with probability at least $1-\beta$ the value of $(CSP_{\delta,N'})$ is smaller or equal than the optimal value of

$$\text{(6)} \qquad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathrm{P}(\phi(x,L)>0 \mid L\in C_\delta) \le 1-(1-\beta)^{1/N'}, \\ & x\in O_\delta \subseteq \mathbb{R}^{d_x}. \end{array}$$

The proof is complete by using $1-(1-\beta)^{1/N'} \ge \beta/N' \ge \frac{\delta}{\mathrm{P}(L\in C_\delta)}$. So, using Val for "value of", Val (6) $\le$ Val (5) $=$ Val (CCP_δ).

4. Constructing outer approximations and summary of the algorithm

In this section, we present two methods for the construction of O_δ and C_δ satisfying Property 1. We mostly focus on our "scaling method" which is presented in Section 4.1. This scaling procedure is facilitated by the fact that L is assumed to heavy-tailed distribution with scale-free properties. After showing the construction of the outer sets under the scaling method, we summarize the algorithm at the end of Section 4.1. We supply a lower bound guaranteeing a constant approximation for the output of the algorithm in Section 4.2. Our second method for outer approximation constructions is summarized in Section 4.3. This method is simpler to apply because is based on linear approximations, however, it is somewhat less powerful because it assume that $\phi(x,L)$ is jointly convex.

4.1. Scaling Method.

We are now ready to state our assumption on the distribution of L. We assume that the distribution of L is of multivariate regular variation, a definition that we review first. For background, we refer to Resnick (2013). Let $\mathcal{M}_+(\overline{\mathbb{R}}^{d_l}/\{\mathbf{0}\})$ denote all Radon measures on the space $\overline{\mathbb{R}}^{d_l}/\{\mathbf{0}\}$ (recall that a measure is Radon if it assigns finite mass to all compact sets). If $\mu_n(\cdot),\mu(\cdot)\in\mathcal{M}_+(\overline{\mathbb{R}}^{d_l}/\{\mathbf{0}\})$, then μ_n converges to μ vaguely, denoted by $\mu_n \xrightarrow{v} \mu$, if for all compactly supported continuous functions $f:\overline{\mathbb{R}}^{d_l}/\{\mathbf{0}\}\to\mathbb{R}_+$,

$$\lim_{n\to\infty}\int_{\overline{\mathbb{R}}^{d_l}/\{\mathbf{0}\}} f(x)\mu_n(dx) = \int_{\overline{\mathbb{R}}^{d_l}/\{\mathbf{0}\}} f(x)\mu(dx).$$

L is *multivariate regularly varying* with *limit measure* $\mu(\cdot)\in\mathcal{M}_+(\overline{\mathbb{R}}^{d_l}/\{\mathbf{0}\})$ if

$$\frac{\mathrm{P}(x^{-1}L\in\cdot)}{\mathrm{P}(\|L\|_2>x)} \xrightarrow{v} \mu(\cdot), \qquad \text{as } x\to\infty.$$

Assumption 1. L is multivariate regularly varying with limit measure $\mu(\cdot)\in\mathcal{M}_+(\overline{\mathbb{R}}^{d_l}/\{\mathbf{0}\})$.

We give some intuition behind this definition. Write L in terms of polar coordinates, with R the radius and Θ a random variable taking values on the unit sphere. The radius $R=\|L\|_2$ has a one-dimensional regularly varying tail (i.e. we can write $\mathrm{P}(R>x)=L(x)x^{-\alpha}$ for a slowly varying function L and $\alpha>0$). The angle Θ, conditioned on R being large, converges weakly (as $R\to\infty$) to a limiting random variable. The distribution of this limit can be expressed in terms of the measure μ. For another recent application of multivariate regular variation in operations research, see Kley et al. (2016).

We proceed to analyze the feasible region F_δ when $\delta \to 0$. Intuitively, if the violation probability $\mathrm{P}(\phi(x,L) > 0)$ has a strictly positive lower bound in any compact set, then F_δ will ultimately be disjoint with the compact set when $\delta \to 0$. Thus, the set F_δ is expelled to infinity when $\delta \to 0$ in this case. F_δ is moving towards the direction that $\phi(x,L)$ becomes small such that the violation probability becomes smaller. For instance, if x is one dimensional and $\phi(x,L)$ is increasing in x, then F_δ is moving towards the negative direction. Consider the minimal salvage fund problem as another example, in which the amount of minimal salvage fund $\min_{i=1}^d x_i \to +\infty$ as $\delta \to 0$.

Now we begin to construct the outer approximation set O_δ. To this end, we need to introduce an auxiliary function which we shall call a *level function*.

Definition 1. We say that $\pi : \mathbb{R}^{d_x} \to [0, +\infty]$ is a level function if

(1) for any $\alpha \geq 0$ and $x \in \mathbb{R}^{d_x}$, we have $\pi(\alpha \cdot x) = \alpha \cdot \pi(x)$,
(2) $\lim_{\delta \to 0} \inf_{x \in F_\delta} \pi(x) = +\infty$.

We also define the *level set* $\Pi = \{x \in \mathbb{R}^{d_x} \mid \pi(x) = 1\}$.

As F_δ is moving to infinity the level function is helpful to characterize the 'moving direction' of F_δ as well as the correct rate of scaling as δ becomes small. As we shall see in the proof of Lemma 2, for any δ small enough we can choose some α_δ and define

$$O_\delta := \bigcup_{\alpha \geq \alpha_\delta} (\alpha \cdot \Pi) \supseteq F_\delta.$$

To construct O_δ, we first select the level set Π, and then derive the scaling rate of α_δ.

The level function π and the shape of Π should be chosen in accordance with the moving direction of F_δ to reduce the size of O_δ, in order to achieve better sample complexity. For example, when $\phi(x,L) = -\|x\|^2 - L$, the level function π can be chosen as the Euclidean norm and Π can be chosen as the unit sphere in $\mathbb{R}^{d_x}$. For the minimal salvage fund problem, the level function can be chosen as $\pi(x) = \min_{i=1}^d x_i + \infty \cdot I(x \notin \mathbb{R}_{++}^{d_x})$ in accordance with our intuition that $\min_{i=1}^d x_i \to \infty$, and the level set can be chosen as $\Pi = \{x \in \mathbb{R}^{d_x} \mid \min_{i=1}^d x_i = 1\}$. Therefore, it is natural to impose the following assumption about the existence of the level function.

Assumption 2. There exist a level function π and a level set Π.

To analyze the asymptotic shape of the uniform conditional event C_δ, we connect the asymptotic distribution of L to the asymptotic distribution of $\phi(x,L)$. We pick a continuous non-decreasing function $h : \mathbb{R}_{++} \to \mathbb{R}_{++}$ such that $\lim_{\alpha \to +\infty} h(\alpha) = +\infty$ to characterize the scaling rate of L. In addition, we pick another positive function $r : \mathbb{R}_{++} \to \mathbb{R}_{++}$ to characterize the scaling rate of $\phi(\alpha \cdot x, h(\alpha) \cdot L)$. Intuitively, the scaling function $r(\cdot)$ and $h(\cdot)$ should ensure the condition that $\{\frac{1}{r(\alpha)}\phi(\alpha \cdot x, h(\alpha) \cdot L)\}_{\alpha \geq 1}$ is tight. For the minimal salvage fund problem with fixed δ, as the deficit $\phi(x,L)$ is asymptotically linear with respect to the salvage fund x and the loss L, we can simply pick $r(\alpha) = h(\alpha) = \alpha$ in this problem. We next introduce two auxiliary functions Ψ_+ and Ψ_-.

Definition 2. Let $\Psi_+ : \mathbb{R}^{d_l} \to \mathbb{R}$, $\Psi_- : \mathbb{R}^{d_l} \to \mathbb{R}$ be two Borel measurable functions. We say Ψ_+ (resp. Ψ_-) is the asymptotic uniform upper (resp. lower) bound of $\frac{1}{r(\alpha)}\phi(\alpha \cdot x, h(\alpha) \cdot l)$ over the level set $x \in \Pi$ if for any compact set $K \subseteq \mathbb{R}^{d_l}$,

$$\liminf_{\alpha\to\infty} \inf_{l\in K} \left(\Psi_+(l) - \sup_{x\in\Pi} \left[\frac{1}{r(\alpha)} \phi(\alpha \cdot x, h(\alpha) \cdot l) \right] \right) \geq 0, \tag{7a}$$

$$\limsup_{\alpha\to\infty} \sup_{l\in K} \left(\Psi_-(l) - \inf_{x\in\Pi} \left[\frac{1}{r(\alpha)} \phi(\alpha \cdot x, h(\alpha) \cdot l) \right] \right) \leq 0. \tag{7b}$$

In Section 5, we show for the salvage fund example how Ψ_+ and Ψ_- can be written as maxima or minima of affine functions. Here, we employ the functions Ψ_+ and Ψ_- to define the event $C_{\varepsilon,-}$ and $C_{\varepsilon,+}$, which serve as the inner and outer approximation of the event $\cup_{x\in\Pi} V_x$, where $V_x = \{l \in \mathbb{R}^{d_l} \mid \phi(x, l) > 0\}$ is the violation event at x.

Definition 3. For $\varepsilon > 0$, let $C_{\varepsilon,+}$ (resp. $C_{\varepsilon,-}$) be the ε-outer (resp. inner) approximation event

$$C_{\varepsilon,+} := \{l \in \mathbb{R}^{d_l} \mid \Psi_+(l) \geq -\varepsilon\}, \tag{8a}$$

$$C_{\varepsilon,-} := \{l \in \mathbb{R}^{d_l} \mid \Psi_-(l) \geq +\varepsilon\}. \tag{8b}$$

We now define $O_\delta := \bigcup_{\alpha\geq\alpha_\delta} \alpha \cdot \Pi$. The following property ensures that the shape of Π is appropriate and α_δ is large enough, hence O_δ is an outer approximation of F_δ.

Property 2. There exist δ_0 such that for any $\delta < \delta_0$, we have an explicitly computable constant α_δ that satisfies

$$\mathrm{P}(\|L\|_2 > h(\alpha_\delta)) = O(\delta) \qquad \text{and} \qquad F_\delta \subseteq \bigcup_{\alpha\geq\alpha_\delta} \alpha \cdot \Pi = O_\delta.$$

If the violation probability is easy to analyze, we will directly derive the expression of α_δ and verify Property 2. Otherwise, we resort to Lemma 2, which provides a sufficient condition of Property 2 by analyzing the asymptotic distribution. The proof of Lemma 2 is deferred to Section 4.1.1.

Lemma 2. Suppose that Assumptions 1 and 2 hold. If there exists an asymptotic uniform lower bound function $\Psi_-(\cdot)$ as given in (7b) and $\varepsilon > 0$ such that $\mu(C_{\varepsilon,-}) > 0$, then Property 2 is satisfied.

We impose the following Assumption 3 on the asymptotic uniform upper bound $\Psi_+(\cdot)$ so that we can employ the multivariate regular variation of L to estimate $\mathrm{P}(L \in \alpha \cdot C_{\varepsilon,+})$ for large scaling factor α.

Assumption 3. There exist an event $S \subseteq \mathbb{R}^{d_l}$ with $\mu(S^c) < \infty$ such that

$$S \subseteq \alpha \cdot S, \qquad \Psi_+(l) \leq \Psi_+(\alpha \cdot l), \qquad \forall l \in S,\ \alpha \geq 1.$$

In addition, there exist some $\varepsilon > 0$ such that $C_{\varepsilon,+}$ is bounded away from the origin, i.e., $\inf_{l\in C_{\varepsilon,+}} \|l\|_2 > 0$.

For the minimal salvage fund problem, since the deficit function $\phi(x, L)$ is coordinatewise nondecreasing with respect to the loss vector L, it is reasonable to assume that its asymptotic bound $\Psi_+(\cdot)$ is also coordinatewise nondecreasing. For this example, the closed form expression of $\Psi_+(\cdot)$ and the detailed verification of all the assumptions are deferred to Proposition 7. Our next result summarizes the construction of the outer approximation sets. The proof of Theorem 3 is deferred to Section 4.1.2.

Theorem 3. Suppose that Property 2 and Assumption 3 are imposed. Then there exist $\delta_0 > 0$ such that the following sets

$$O_\delta = \bigcup_{\alpha \geq \alpha_\delta} \alpha \cdot \Pi, \qquad C_\delta = h(\alpha_\delta) \cdot \big(C_{\varepsilon,+} \cup K^c \cup S^c\big) \tag{9}$$

satisfy Property 1 for all $\delta < \delta_0$. Here, S is given in Assumption 3 and K is a ball in $\mathbb{R}^{d_l}$ with $\mu(K^c) < \infty$.

With the aid of Theorem 1 and 3, we provide Algorithm 1 for approximating (CCP_δ) in which the sampled optimization problem is bounded in $1/\delta$.

Algorithm 1: Scenario Approach with Optimal Scenario Generation

input : Constraint function ϕ, risk tolerance parameter δ, confidence level β, all the elements and constants appearing in Property 2 and Assumption 3.

1 Compute the expression of sets O_δ and C_δ by (9);
2 Compute required number of samples N' by (4);
3 **for** $i = 1, \ldots, N'$ **do**
4   Sample $L_\delta^{(i)}$ using acceptance-rejection or importance sampling.
5 **end**
6 Solve the conditional sampled problem $(CSP_{\delta,N'})$.

In Section 4.2, our objective is to show that the output of the previous algorithm is guaranteed to be within a constant factor of the optimal solution to (CCP_δ) with high probability, uniformly in δ.

4.1.1. *Proof of Lemma 2.* We will derive an expression of α_δ to ensure that $F_\delta \subseteq \bigcup_{\alpha \geq \alpha_\delta} \alpha \cdot \Pi$ for δ small enough. Because of Assumption 2, for any $\alpha_0 > 0$ there exist some δ small enough such that $F_\delta \subseteq \bigcup_{\alpha \geq \alpha_0} \alpha \cdot \Pi$. Therefore, it suffices to prove that F_δ and $\bigcup_{\alpha < \alpha_\delta} \alpha \cdot \Pi$ are disjoint. In other words,

$$\mathrm{P}\left(\phi(\alpha \cdot x, L) > 0\right) > \delta, \quad \forall \alpha < \alpha_\delta, x \in \Pi, \delta < \delta_0. \tag{10}$$

Let ε be a positive number such that $\mu(C_{\varepsilon,-}) > 0$. Pick the set K in (7b) as a compact set such that $0 < \mu(K \cap C_{\varepsilon,-}) < \infty$. It follows from the inequality (7b) that there exist a constant α_1 such that

$$\Psi_-(l) - \varepsilon \leq \inf_{x \in \Pi} \left[\frac{1}{r(\alpha)} \phi(\alpha \cdot x, h(\alpha) \cdot l)\right] \qquad \forall l \in K,\ \alpha > \alpha_1 \tag{11}$$

Therefore, for any $\alpha \geq \alpha_1$ we have,

$$\begin{aligned} \mathrm{P}\left(\min_{x\in\Pi} \phi(\alpha\cdot x, L) > 0\right) &= \mathrm{P}\left(\min_{x\in\Pi} \frac{1}{r(\alpha)}\phi(\alpha\cdot x, L) > 0\right) \\ \text{(Due to (11))} &\geq \mathrm{P}\left(G(L/h(\alpha)) \geq \varepsilon; L/h(\alpha) \in K\right) \\ &= \mathrm{P}\left(L \in h(\alpha)\cdot(K\cap C_{\varepsilon,-})\right). \end{aligned} \tag{12}$$

Recall that L is regularly varying from Assumption 1,

$$\lim_{\alpha\to\infty} \frac{\mathrm{P}\left(L \in h(\alpha)\cdot(K\cap C_{\varepsilon,-})\right)}{\mathrm{P}(\|L\|_2 > h(\alpha))} = \mu(K\cap C_{\varepsilon,-}).$$

Therefore, there exist a number α_2 such that

$$\mathrm{P}\left(L \in h(\alpha)\cdot(K\cap C_{\varepsilon,-})\right) \geq \frac{1}{2}\mathrm{P}(\|L\|_2 > h(\alpha))\mu(K\cap C_{\varepsilon,-}), \qquad \forall \alpha \geq \alpha_2. \tag{13}$$

Note that the right hand side of (13) is nondecreasing in α. Thus, if

$$\delta_1 := \frac{1}{2}\mathrm{P}\big(\|L\|_2 > h(\alpha_2)\big)\mu(K\cap C_{\varepsilon,-}),$$

for any $\delta \leq \delta_1$ there exist α_δ satisfying

$$\frac{1}{2}\mathrm{P}(\|L\|_2 > h(\alpha_\delta))\mu(K\cap C_{\varepsilon,-}) = \delta. \qquad \forall \alpha, \delta \quad s.t. \quad \alpha_2 \leq \alpha < \alpha_\delta, 0 < \delta \leq \delta_1. \tag{14}$$

Substituting (14) into (12), we have

$$\begin{aligned} &\mathrm{P}\left(\phi(x, L) > 0\right) \geq \mathrm{P}\left(\min_{x\in\Pi} \phi(\alpha\cdot x, L) > 0\right) > \delta. \\ &\forall \alpha, x, \delta \quad s.t. \quad \max(\alpha_1,\alpha_2) \leq \alpha < \alpha_\delta, x\in\Pi, 0<\delta\leq\delta_1. \end{aligned}$$

Moreover, Assumption 2 guarantees the existence of δ_2 such that

$$\mathrm{P}\left(\phi(\alpha\cdot x, L) > 0\right) > \delta, \quad \forall \alpha < \max(\alpha_1,\alpha_2), x\in\Pi, \delta<\delta_2.$$

Consequently (10) is proved with $\delta_0 = \min(\delta_1,\delta_2)$.

4.1.2. *Proof of Theorem 3.* We construct the uniform conditional event C_δ that contains all the V_x for $x \in O_\delta$. Due to the definition (7) and $\lim_{\delta\to 0}\alpha_\delta = \infty$, there exist δ_0 such that for all $\delta < \delta_0$,

$$\Psi_+(l) + \varepsilon \geq \sup_{x\in\Pi}\left[\frac{1}{r(\alpha)}\phi(\alpha\cdot x, h(\alpha)\cdot l)\right] \qquad \forall l \in K,\ \alpha > \alpha_\delta. \tag{15}$$

Notice that for any $x \in O_\delta$, there exist an $\alpha_x \geq \alpha_\delta$ such that $x \in \alpha_x \cdot \Pi$. Consequently, it follows from (15) that

$$\phi(x,l) > 0 \Longrightarrow \Psi_+\Big(\frac{l}{h(\alpha_x)}\Big) \geq -\varepsilon, \qquad \forall x \in O_\delta, l \in h(\alpha_x)\cdot K.$$

Applying Assumption 3 yields that

$$\Psi_+\Big(\frac{l}{h(\alpha_\delta)}\Big) \geq \Psi_+\Big(\frac{l}{h(\alpha_x)}\Big) \geq -\varepsilon, \qquad \forall x \in O_\delta, l \in h(\alpha_x)\cdot(K\cap S).$$

Recall that K is a ball in $\mathbb{R}^{d_l}$ (thus $K \subseteq (h(\alpha_x)/h(\alpha_\alpha))\cdot K$) and that $S \subseteq (h(\alpha_x)/h(\alpha_\alpha))\cdot S$ from Assumption 3, it turns out that $h(\alpha_\delta)\cdot(K\cap S) \subseteq h(\alpha_x)\cdot(K\cap S)$. Consequently, whenever $l \in V_x$ for some $x \in O_\delta$, we either have $l \in h(\alpha_x)\cdot(K\cap S)$ implying $\Psi_+\Big(\frac{l}{h(\alpha_\delta)}\Big) \geq -\varepsilon$, or we have $l \in \big(h(\alpha_x)\cdot(K\cap S)\big)^c \subseteq \big(h(\alpha_\delta)\cdot(K\cap S)\big)^c$. Summarizing these two scenarios,

$$\begin{aligned}\bigcup_{x\in O_\delta} V_x &\subseteq \{l \in \mathbb{R}^{d_l} \mid \Psi_+\Big(\frac{l}{h(\alpha_\delta)}\Big) \geq -\varepsilon\} \bigcup \big(h(\alpha_\delta)\cdot(K\cap S)\big)^c \\ &= h(\alpha_\delta)\cdot\big(C_{\varepsilon,+}\cup K^c \cup S^c\big).\end{aligned}$$

Thus, we define the conditional set C_δ as

$$C_\delta := h(\alpha_\delta)\cdot\big(C_{\varepsilon,+}\cup K^c \cup S^c\big).$$

It remains to analyze the probability of the uniform conditional event C_δ. As L is multivariate regularly varying,

$$\lim_{\delta\to 0}\frac{\mathrm{P}\left(L \in C_\delta\right)}{\mathrm{P}(\|L\|_2 > h(\alpha_\delta))} = \mu(C_{\varepsilon,+}\cup K^c\cup S^c).$$

Recalling, $\mathrm{P}(\|L\|_2 > h(\alpha_\delta)) = O(\delta)$ and invoking Property (2), we get

$$\limsup_{\delta\to 0}\delta^{-1}\mathrm{P}(L\in C_\delta) < \infty.$$

Hence, the proof is complete.

4.2. **Constant Approximation Guarantee.** We shall work under the setting of Theorem 3, so we enforce Property 2 and Assumptions 3. We want to show that there exist some constant $\Lambda > 1$ independent of δ, such that $\mathrm{Val}\left(CCP_\delta\right) \leq \mathrm{Val}\left(CSP_{\delta,N'}\right) \leq \Lambda\times\mathrm{Val}\left(CCP_\delta\right)$ with high probability. This indicates that our result guarantees a constant approximation to (CCP_δ) for regularly varying distributions (under our assumptions) in $O(1)$ sample complexity when $\delta\to 0$ with high probability.

Note that $(CSP_{\delta,N'}) \leq \Lambda\times\mathrm{Val}\left(CCP_\delta\right)$ is meaningful only if $\mathrm{Val}\left(CCP_\delta\right) > 0$. We assume that the outer approximation set is good enough such that the following natural assumption is valid.

Assumption 4. There exist $\delta > 0$ such that $\min_{x\in O_\delta} c^T x > 0$.

The previous assumption will typically hold if c has strictly positive entries. Theorem 3 and the form of O_δ guarantee that the norm of the optimal solution of $(CSP_{\delta,N'})$ grows in proportion to α_δ, so we also assume the following scaling property for $\phi(x,l)$.

Assumption 5. There exist a function $\phi_{\lim} : (\mathbb{R}^{d_x}/\{\mathbf{0}\})\times(\mathbb{R}^{d_l}/\{\mathbf{0}\}) \to \mathbb{R}$ such that for every compact set $E \subseteq \mathbb{R}^{d_l}/\{\mathbf{0}\}$, we have

$$\lim_{\alpha\to\infty}\sup_{l\in E}\left|\frac{1}{r(\alpha)}\phi(\alpha\cdot x, h(\alpha)\cdot l) - \phi_{\lim}(x,l)\right| = 0.$$

In addition, $\phi_{\lim}(x,l)$ is continuous in l.

Assumption 5 is satisfied by the salvage fund problem (1), since in this case we have $\phi_{\text{lim}}(x,l) = \phi(x,l) - m$ such that $|\alpha^{-1}\phi(\alpha \cdot x, \alpha \cdot l) - \phi_{\text{lim}}(x,l)| \le \alpha^{-1}m$ and $|\phi_{\text{lim}}(x,l) - \phi_{\text{lim}}(x,l')| \le \|l - l'\|_1$.

We define the following optimization problem, which will serve as an asymptotic upper bound of $(CSP_{\delta,N'})$ in stochastic order when $\delta \to 0$:

$$(CSP_{\text{lim},N'}) \qquad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \phi_{\text{lim}}(x, L_{\text{lim}}^{(i)}) \le 0, \quad i = 1,\dots,N', \\ & x \in \bigcup_{\alpha \ge 1} \alpha \cdot \Pi, \end{array}$$

where $L_{\text{lim}}^{(i)}$ are i.i.d. samples from a random variable L_{lim}, whose distribution is characterized by $\mathrm{P}(L_{\text{lim}} \in (C_{\varepsilon,+} \cup K^c \cup S^c)) = 1$ and $\mathrm{P}(L_{\text{lim}} \in E) = \mu(E)/\mu(C_{\varepsilon,+} \cup K^c \cup S^c)$ for all measurable set $E \subseteq C_{\varepsilon,+} \cup K^c \cup S^c$.

Theorem 4. Let $\beta > 0$ be a given confidence level and N' be a fixed integer that satisfies (4). If Assumptions 4 and 5 are enforced, and $(CSP_{\text{lim},N'})$ satisfies Slater's condition with probability one, then there exist $\delta_0 > 0$ and $\Lambda > 0$ such that

$$\mathrm{P}\Big(\mathrm{Val}\,(CCP_\delta) \le \mathrm{Val}\,(CSP_{\delta,N'}) \le \Lambda \times \mathrm{Val}\,(CCP_\delta) \Big) \ge 1 - 2\beta, \qquad \forall \delta < \delta_0.$$

The Slater's condition can be verified directly on the problem $(CSP_{\text{lim},N'})$. This condition is satisfied in the salvage fund problem by standard linear programming duality. We now proceed with the proof of Theorem 4.

4.2.1. *Proof of Theorem 4.* Using Theorem 1, we immediately have

$$\mathrm{P}(\mathrm{Val}\,(CCP_\delta) \le \mathrm{Val}\,(CSP_{\delta,N'})) \ge 1 - \beta,$$

it remains to show that there exist $\Lambda > 0$ such that

$$\mathrm{P}(\mathrm{Val}\,(CSP_{\delta,N'}) \le \Lambda \times \mathrm{Val}\,(CCP_\delta)) \ge 1 - \beta.$$

For simplicity, in the proof we will use L_δ as a shorthand for $(L | L \in C_\delta)$, the random variable with conditional distribution of L given $L \in C_\delta$. By a scaling of x by a factor α_δ in $(CSP_{\delta,N'})$, we have an equivalent optimization problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \frac{1}{r(\alpha_\delta)}\phi(\alpha_\delta \cdot x, L_\delta^{(i)}) \le 0, \quad i = 1,\dots,N', \\ & x \in \bigcup_{\alpha \ge 1} \alpha \cdot \Pi. \end{array} \tag{16}$$

where $L_\delta^{(i)}$ are i.i.d. samples from L_δ. Notice that $\mathrm{Val}\,(CSP_{\delta,N'}) = \alpha_\delta \times \mathrm{Val}\,(16)$.

For any compact set $E \subseteq C_\delta$, since L is multivariate regularly varying,

$$\lim_{\delta \to 0} \mathrm{P}((h(\alpha_\delta))^{-1} L_\delta \in E) = \lim_{\delta \to 0} \frac{\mathrm{P}(L \in (h(\alpha_\delta) \cdot E))}{\mathrm{P}(L \in C_\delta)} = \frac{\lim_{\delta \to 0} \frac{\mathrm{P}(L \in (h(\alpha_\delta) \cdot E))}{\mathrm{P}(\|L\|_2 > h(\alpha_\delta))}}{\lim_{\delta \to 0} \frac{\mathrm{P}(L \in C_\delta)}{\mathrm{P}(\|L\|_2 > h(\alpha_\delta))}} = \frac{\mu(E)}{\mu(C_{\varepsilon,+} \cup K^c \cup S^c)}.$$

Thus $(h(\alpha_\delta))^{-1} L_\delta \xrightarrow{v} L_{\text{lim}}$. As the limiting measure is a probability measure, the family $\{h(\alpha_\delta))^{-1} L_\delta \mid \delta > 0\}$ is tight and consequently $(h(\alpha_\delta))^{-1} L_\delta \xrightarrow{d} L_{\text{lim}}$ follows directly from the vague convergence, see Resnick (2013). Consequently, since all the samples are i.i.d, we also have

$$(h(\alpha_\delta))^{-1} \cdot (L_\delta^{(1)}, \ldots, L_\delta^{(N')}) \xrightarrow{d} (L_{\lim}^{(1)}, \ldots, L_{\lim}^{(N')}).$$

Now we define a family of deterministic optimization problem, denoted by $(DP(l_1, \cdots, l_{N'}))$, which is parameterized by $(l_1, \cdots, l_{N'})$ as follows,

$$(DP(l_1, \cdots, l_{N'})) \qquad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \phi_{\lim}(x, l_i) \leq 0, \quad i = 1, \ldots, N', \\ & x \in \bigcup_{\alpha \geq 1} \alpha \cdot \Pi. \end{array}$$

Then, there exist a compact set $E_1 \subseteq \mathbb{R}^{d_l \times N'}$ such that:

(1) Problem $(DP(l_1, \cdots, l_{N'}))$ satisfies Slater's condition if $(l_1, \cdots, l_{N'}) \in E_1$;
(2) $\mathrm{P}((h(\alpha_\delta))^{-1} \cdot (L_\delta^{(1)}, \ldots, L_\delta^{(N')}) \in E_1) \geq 1 - \beta$ for all $\delta > 0$;

For every $(l_1, \cdots, l_{N'}) \in E_1$ and $\epsilon > 0$, due to the Slater's condition, there exist a feasible solution $x \in \bigcup_{\alpha \geq 1} \alpha$ such that $\sup_{j=1,\ldots,N'} \phi_{\lim}(x, l_j) < -\epsilon$. Since $\phi_{\lim}(x, l)$ is continuous in l, there exist an open neighborhood U around $(l_1, \cdots, l_{N'})$ such that $\sup_{(l_1,\ldots,l_{N'}) \in U} \sup_{j=1,\ldots,N'} \phi_{\lim}(x, l_j) < -\epsilon/2$. Notice that such feasible solution x and neighborhood U exist for every $(l_1, \cdots, l_{N'}) \in E_1$. Since E_1 is compact, and all the open neighborhoods U form an open cover for the compact set E_1. Now let $\{U_i\}_{i=1}^m$ be a finite open cover of E_1 and let $\{x_i\}_{i=1}^m$ be the corresponding feasible solutions. Due to Assumption 5, there exist $\delta_1 > 0$ such that for all $\delta < \delta_1$ we have

$$\sup_{(l_1,\ldots,l_{N'}) \in E_1} \sup_{i=1,\ldots,m} \sup_{j=1,\ldots,N'} \left| \frac{1}{r(\alpha_\delta)} \phi(\alpha_\delta \cdot x_i, h(\alpha_\delta) \cdot l_j) - \phi_{\lim}(x_i, l_j) \right| < \epsilon/2. \tag{17}$$

Therefore by the triangle inequality, it follows that if $\delta < \delta_1$,

$$\sup_{(l_1,\ldots,l_{N'}) \in U_i} \sup_{j=1,\ldots,N'} \frac{1}{r(\alpha_\delta)} \phi(\alpha_\delta \cdot x_i, h(\alpha_\delta) \cdot l_j) < 0.$$

Consequently, x_i is a feasible solution for optimization problem (16) if

$$(h(\alpha_\delta))^{-1} \cdot (L_\delta^{(1)}, \ldots, L_\delta^{(N')}) \in U_i,$$

which further implies that $\alpha_\delta^{-1} \times \mathrm{Val}\,(CSP_{\delta,N'}) \leq c^T x_i$. As a result, we have

$$\mathrm{Val}\,(CSP_{\delta,N'}) \leq \alpha_\delta \times \max_{i=1,\ldots,m} c^T x_i, \quad \text{if} \quad (h(\alpha_\delta))^{-1} \cdot (L_\delta^{(i)}, \ldots, L_\delta^{(N')}) \in E_1.$$

Note that $\mathrm{Val}\,(CCP_\delta) \geq \inf_{x \in O_\delta} c^T x = \alpha_\delta \times \inf\{c^T x \mid x \in \bigcup_{\alpha \geq 1} \alpha \cdot \Pi\}$. Therefore, let

$$\Lambda = \big(\inf\{c^T x \mid x \in \bigcup_{\alpha \geq 1} \alpha \cdot \Pi\}\big)^{-1} \times \big(\max_{i=1,\ldots,m} c^T x_i\big) > 0$$

It follows that

$$\mathrm{P}\Big(\mathrm{Val}\,(CSP_{\delta,N'}) \leq \Lambda \times \mathrm{Val}\,(CCP_\delta)\Big) \geq \mathrm{P}\Big((h(\alpha_\delta))^{-1} \cdot (L_\delta^{(1)}, \ldots, L_\delta^{(N')}) \in E_1\Big) \geq 1 - \beta.$$

The statement is concluded by using the union bound, combining the lower bound together with the upper bound implied by Theorems 1 and 3, hence obtaining factor 2β.

4.3. **Linear Approximation Method.** Suppose that the constraint function $\phi(x, L)$ is jointly convex in L and x, and L is multivariate regularly varying. We will develop a simpler method in this section to construct the outer approximation set O_δ and the uniform conditional event C_δ.

Assumption 6. The constraint function $\phi(x, L) : \mathbb{R}^{d_x} \times \mathbb{R}^{d_l} \to \mathbb{R}$ has a compact zero sublevel set $Z_\phi := \{(x, L) \in \mathbb{R}^{d_x} \times \mathbb{R}^{d_l} \mid \phi(x, L) \leq 0\}$. In addition, $\phi(x, L)$ is convex and twice continuously differentiable.

The following lemma shows that $\phi(x, L)$ can be uniformly approximated by finitely many linear functions over the zero sublevel set. This is crucial in the construction of O_δ and C_δ.

Lemma 5. If Assumption 6 holds, there exist a function $\phi_-(x, L) : \mathbb{R}^{d_x} \times \mathbb{R}^{d_l} \to \mathbb{R}$ such that

(1) $\phi_-(x, L) \leq \phi(x, L), \forall (x, L) \in \mathbb{R}^{d_x} \times \mathbb{R}^{d_l}$;
(2) there exist some constant C such that $\phi_-(x, L) + C \geq \phi(x, L), \forall (x, L) \in Z_\phi$;
(3) there exist $a_i \in \mathbb{R}^{d_l}$, $b_i \in \mathbb{R}^{d_x}$ and $c_i \in \mathbb{R}$ for $i = 1, \ldots, N$ such that the function $\phi_-(x, L)$ can be written as
$$\phi_-(x, L) = \max_{i=1,\ldots,N} \quad \left(a_i^T L + b_i^T x + c_i\right).$$

With the aid of Lemma 5, we are now ready to provide our main result in this section to fully summarize the construction of O_δ and C_δ.

Theorem 6. If Assumptions 1 and 6 hold, we can construct O_δ and C_δ that satisfy Property 1 as

(18)
$$O_\delta := \bigcap_{i=1}^{N} \{x \in \mathbb{R}^{d_x} \mid b_i^T x + c_i + \bar{F}^{-1}_{a_i^T L}(\delta) \leq 0\}, \qquad C_\delta := \bigcup_{i=1}^{N} \{L \in \mathbb{R}^{d_l} \mid a_i^T L + C > \bar{F}^{-1}_{a_i^T L}(\delta)\},$$

where $F^{-1}_{a_i^T L}(\delta) = \inf\{x \in \mathbb{R} \mid \mathrm{P}(x > a_i^T L) \leq \delta\}$.

4.3.1. *Proof of Lemma 5.* Without loss of generality, assume that R is an integer such that
$$Z_\phi = \{(x, L) \in \mathbb{R}^{d_x} \times \mathbb{R}^{d_l} \mid \phi(x, L) \leq 0\} \subseteq [-R, R]^{(d_x + d_l)}.$$

Let $N = (2R + 1)^{(d_x + d_l)}$, and let $(x_i, L_i), i = 1, \ldots, N$ be the integer lattice points in $[-R, R]^{(d_x + d_l)}$. In addition, let $a_i = \frac{\partial \phi}{\partial L}(x_i, L_i)$, $b_i = \frac{\partial \phi}{\partial x}(x_i, L_i)$ and $c_i = \phi(x_i, L_i) - \frac{\partial \phi}{\partial L}(x_i, L_i)^T L_i - \frac{\partial \phi}{\partial x}(x_i, L_i)^T x_i$ for $i = 1, \ldots, N$. Since the function $\phi(x, L)$ is convex, we can invoke the supporting hyperplane theorem to deduce that $a_i^T L + b_i^T x + c_i \leq \phi(x, L)$, and consequently $\phi_-(x, L) \leq \phi(x, L)$. Thus, for an arbitrary point $(x, L) \in Z_\phi$, there exist a lattice point (x_i, L_i) such that $\|(x, L) - (x_i, L_i)\|_2 \leq \sqrt{d_x + d_l}/2$. Next, since $\phi(x, L)$ is twice continuously differentiable, the gradient $\nabla \phi(x, L)$ is Lipschitz over Z_ϕ

with Lipschitz constant denoted by M_ϕ. Therefore, for any $(x, L) \in Z_\phi$,

$$\phi(x, L) - \phi_-(x, L) \leq \min_{i=1,\dots,N} \big(\phi(x, L) - (a_i^T L + b_i^T x + c_i)\big) \leq \frac{1}{4} M_\phi^2 \sqrt{d_x + d_l}.$$

The proof is now complete.

4.3.2. *Proof of Theorem 6.* Since $\phi_-(x, L) \leq \phi(x, L)$, the probability constraint $\mathrm{P}(\phi(x, L) > 0) \leq \delta$ implies that $\mathrm{P}(\phi_-(x, L) > 0) \leq \delta$, which further implies $\mathrm{P}(a_i^T L + b_i^T x + c_i > 0) \leq \delta$ for $i = 1, \dots, N$. Therefore, we have $-b_i^T x - c_i \geq \bar{F}^{-1}_{a_i^T L}(\delta)$ for $i = 1, \dots, N$, so $F_\delta \subseteq O_\delta$.

Then, consider $x \in O_\delta$ and $L \in V_x = \{L \in \mathbb{R}^{d_l} \mid \phi(x, L) > 0\}$. It follows from Lemma 5 that $\phi(x, L) > 0$ implies that $\phi_-(x, L) + C > 0$. Thus, there exist an index i such that $a_i^T L + b_i^T x + c_i + C > 0$. As $x \in O_\delta$ implies that $b_i^T x + c_i + \bar{F}^{-1}_{a_i^T L}(\delta) \leq 0$, so

$$a_i^T L - \bar{F}^{-1}_{a_i^T L}(\delta) + C \geq a_i^T L + b_i^T x + c_i + C > 0.$$

Therefore, the condition set C_δ can be constructed as

$$C_\delta := \bigcup_{i=1}^N \{L \in \mathbb{R}^{d_l} \mid a_i^T L + C > \bar{F}^{-1}_{a_i^T L}(\delta)\}.$$

Thus, as the distribution $a_i^T L$ is regularly varying in dimension one for each i, we have $\limsup_{\delta \to 0} \delta^{-1} \mathrm{P}(L \in C_\delta) \leq N$, completing the proof.

5. Verifying the assumptions in examples

In this section, we illustrate the verification of the elements required to apply our algorithm. The verification process, the main point of the section, is presented as the proof of propositions which state validity of the approach, so for this section these proofs are presented in the main body.

5.1. **The minimal salvage fund.**

Proposition 7. The minimal salvage fund problem (2) satisfies all assumptions required by Theorem 3. So it can be solved using the scenario approach within constant sample complexity.

Proof of Proposition 7. Assumption 1 follows directly from the assumptions of the example. Now we turn to verify Assumption 2 with the level set $\Pi = \{x \in \mathbb{R}^d \mid \min_{i=1}^d x_i = 1\}$. We start by deriving a lower bound of $\phi(x, L)$. We first write the dual problem of $\phi(x, L)$, since any feasible solution to the dual problem provides a lower bound for $\phi(x, L)$. Notice that the dual problem of $\phi(x, L)$ is

$$\begin{array}{ll} \text{maximize} & \kappa^T L - \beta^T x - m \\ \text{subject to} & (I - Q)\beta - \kappa \succeq \mathbf{0}, \kappa^T \mathbf{1} = 1, \kappa, \beta \succeq \mathbf{0}. \end{array} \tag{19}$$

Let $\mathbf{e}_i$ be the unit vector on the ith coordinate and η be a small real number such that $\beta = \eta \cdot \mathbf{e}_i$ and $\kappa = c_i \mathbf{1} + \eta Q \mathbf{e}_i - \eta \mathbf{e}_i$ is dual feasible, where $c_i = (1 + \eta - \eta \mathbf{1}^T Q \mathbf{e}_i)/d$. Consequently, $c_i \mathbf{1}^T L + \eta \mathbf{e}_i^T Q^T L - \eta L_i - \eta x_i \leq \phi(x, L) - m$ for any indices $i \in \{1, \dots, d\}$.

As a result, $x \in F_\delta$ implies that $\mathrm{P}(c_i \mathbf{1}^T L + \eta \mathbf{e}_i^T Q^T L - \eta L_i > \eta x_i + m) \leq \delta$, which further implies

$$\lim_{\delta \to 0} \inf_{x \in F_\delta} x_i = \infty.$$

We can conclude that Assumption 2 is also verified.

Now we construct the uniform asymptotic bounds Ψ_+ and Ψ_-. The scaling rate functions are chosen as $r(\alpha) = \alpha$ and $h(\alpha) = \alpha$ so that $\frac{1}{r(\alpha)}\phi(\alpha \cdot x, h(\alpha) \cdot L) = \psi(x, L)$. Using the lower bound of $\phi(x, L)$ given in the above paragraph,

$$\begin{aligned} \inf_{x \in \Pi} \phi(x, L) &\geq \inf_{x \in \Pi} \max_{i=1,\ldots,d} c_i \mathbf{1}^T L + \eta \mathbf{e}_i^T Q^T L - \eta L_i - \eta x_i - m \\ &\geq \min_{i=1,\ldots,d} c_i \mathbf{1}^T L + \eta \mathbf{e}_i^T Q^T L - \eta L_i - \eta - m. \end{aligned}$$

Let $\varepsilon < 1/2$ be an arbitrary small number, and we pick the scaling function as $h(\alpha) = \alpha$ and $r(\alpha) = \alpha$. Using the above lower bounded of $\inf_{x \in \Pi} \phi(x, L)$ and further noting that $\phi(x, L) \leq \max_{i=1,\ldots,d}\{L_i - x_i\} - m$ because $y = x$ and $b = \max_{i=1,\ldots,d}\{L_i - x_i\}$ is primal feasible, the asymptotic uniform bound $\Psi_+(L)$ and $\Psi_-(L)$ can be chosen as

$$\Psi_+(L) = \max\{\max_{i=1}^{d}(L_i - 1), -2\varepsilon\} - m, \qquad \Psi_-(L) = \min_{i=1,\ldots,d} c_i \mathbf{1}^T L + \eta \mathbf{e}_i^T Q^T L - \eta L_i - \eta - m.$$

When η is small we have $\Psi_-(L) \approx \mathbf{1}^T L - m$, hence the set $C_{\varepsilon,-}$ satisfies $\mu(C_{\varepsilon,-}) > 0$. Consequently Property 2 is verified due to Lemma 2. In addition, $C_{\varepsilon,+} = \{l \in \mathbb{R}^{d_l} \mid \max_{i=1}^d l_i \geq m + 1 - \varepsilon\}$, which is bounded away from the origin. Thus Assumption 3 is verified with $S = \mathbb{R}^d$. □

5.2. **Quadratic Model.** In this section, we consider a model with a quadratic control term in x. Suppose that the constraint function $\phi(x, L) : \mathbb{R}^{d_x} \times \mathbb{R}^{d_l} \to \mathbb{R}$ is defined as

$$\phi(x, L) = x^T Q x + x^T A L, \tag{20}$$

where $Q \in \mathbb{R}^{d_x \times d_x}$ is a symmetric matrix and $A \in \mathbb{R}^{d_x \times d_l}$ is a matrix with $\operatorname{rank}(A) = d_x$, i.e. there exist $\sigma > 0$ such that $\|A^T x\|_2 \geq \sigma \|x\|_2$.

Proposition 8. Consider the chance constraint optimization model with constraint function defined as (20).

(1) If Q is a positive semi-definite matrix and L has a positive density, there exist some δ such that the problem is infeasible.
(2) If Q has a negative eigenvalue and L is multivariate regularly varying, the model satisfies all the assumptions required by Theorem 3.

Proof. Proof of Proposition 8. For the first statement, since $x^T Q x \geq 0$ and $A^T x \in \mathbb{R}^{d_l}$, and invoking the assumption that L has a positive density,

$$\min_{y \in \mathbb{R}^{d_l}/\{\mathbf{0}\}} P(y^T L > 0) \geq \min_{y : \|y\|_2 = 1} P(y^T L > 0) > 0.$$

For the second statement, Assumption 1 is easy to verify. Notice that $\alpha^{-2}\phi(\alpha \cdot x, \alpha \cdot L) = \phi(x, L)$ for all $\alpha > 0$, so we pick the scaling rate function as $h(\alpha) = \alpha$ and $r(\alpha) = \alpha^2$.

Let $\lambda_{\max}$ denote the maximal eigenvalue of Q, and $\lambda_{\min}$ denote the minimal eigenvalue of Q. The rest of the proof will be divided into two cases.

Case 1 ($\lambda_{\max} < 0$): We pick the level set as $\Pi = \{x \in \mathbb{R}^{d_x} \mid \|x\|_2 = 1\}$. Since $\lim_{\delta\to 0}\inf_{x\in F_\delta}\|x\|_2 = \infty$, Assumption 2 is verified. Next, we directly show Property 2 instead of using Lemma 2. For any $x \in \alpha\cdot\Pi$ we have

$$\begin{aligned}
\min_{x\in\alpha\cdot\Pi} \mathrm{P}(x^TQx + x^TAL > 0) &\geq \min_{x\in\Pi} \mathrm{P}\big(\alpha\lambda_{\min} + x^TAL > 0\big) \\
&= \min_{x\in\Pi} \mathrm{P}\left(\frac{x^TAL}{\|A^Tx\|_2} > \frac{-\alpha\lambda_{\min}}{\|A^Tx\|_2}\right) \\
&\geq \min_{z\in\Pi} \mathrm{P}\left(z^TL > -\alpha\sigma^{-1}\lambda_{\min}\right) \\
\text{(Apply Lemma 9)} \quad &\geq \min_{i=1,\ldots,2d_l} \mathrm{P}(L \in -\alpha\sigma^{-1}\lambda_{\min}S_i).
\end{aligned}$$

Thus, α_δ can be chosen such that $\alpha_\delta = O(\delta)$, and $\min_{i=1,\ldots,2d_l}\mathrm{P}(L \in -\alpha\sigma^{-1}\lambda_{\min}S_i) > \delta$. As a result, Property 2 is verified. We next turn to derive the asymptotic uniform bound Ψ_+. Observing that

$$\sup_{x\in\Pi}\phi(x,L) \leq \lambda_{\max} + \|A\|_F\|L\|_2,$$

we define $\Psi_+(L) := \lambda_{\max} + \|A\|_F\|L\|_2$. Assumption 3 now follows from the definition of Ψ_+.

Case 2 ($\lambda_{\max} \geq 0$): The level set Π is chosen as an unbounded set $\Pi = \{x \in \mathbb{R}^{d_x} \mid x^TQx = -\|x\|_2\}$ and we have $\min_{x\in\Pi}\|x\|_2 = 1/|\lambda_{\min}|$. For any $x \in \alpha\cdot\Pi$ we have

$$\begin{aligned}
\min_{x\in\alpha\cdot\Pi} \mathrm{P}(x^TQx + x^TAL > 0) &\geq \min_{x\in\Pi} \mathrm{P}\big(x^TAL > \alpha\big), \\
&= \min_{x\in\Pi} \mathrm{P}\left(\frac{x^TAL}{\|A^Tx\|_2} > \frac{\alpha}{\|A^Tx\|_2}\right) \\
&\geq \min_{z:\|z\|=1} \mathrm{P}\left(z^TL > -\alpha\sigma^{-1}\lambda_{\min}\right) \\
\text{(Apply Lemma 9)} \quad &\geq \min_{i=1,\ldots,2d_l} \mathrm{P}(L \in -\alpha\sigma^{-1}\lambda_{\min}S_i).
\end{aligned}$$

Thus we can pick an α_δ that satisfies Property 2. Now note that $\sup_{x\in\Pi}\phi(x,L)$ is bounded by

$$\sup_{x\in\Pi}\phi(x,L) \leq \sup_{x\in\Pi}\|x\|_2(\|AL\|_2 - 1) \leq -\frac{1}{2}|\lambda_{\min}|^{-1}\cdot I(\|AL\|_2 \leq 1/2) + \infty\cdot I(\|AL\|_2 > 1)$$

so we can pick $\Psi_+(L) := -\frac{1}{2}|\lambda_{\min}|^{-1}\cdot I(\|AL\|_2 \leq 1/2) + \infty\cdot I(\|AL\|_2 > 1)$. Consequently Assumption 3 follows immediately.

The following lemma is used in the proof of Proposition 8.

Lemma 9. There exist sets $S_1,\ldots,S_{2d_l} \subseteq \mathbb{R}^{d_l}$ with positive Lebesgue measure such that for any $z \in \mathbb{R}^{d_l}$ with $\|z\|_2 = 1$, there exist some $S_i \subseteq \{l \in \mathbb{R}^{d_l} | z^Tl > 1\}$.

Proof of Lemma 9. Let $\mathbf{e}_i$ denote the unit vector on the ith coordinate in $\mathbb{R}^{d_l}$ for $i = 1, \ldots, d_l$. Fix $z = (z_1, \ldots, z_{d_l}) \in \mathbb{R}^{d_l}$ with $\|z\|_2 = 1$, define θ_i be the angle between z and $\mathbf{e}_i$, which satisfies $\cos(\theta_i) = z^T e_i$. Since we have $\sum_{i=1}^n \cos(\theta_i)^2 = 1$, so there exist some i such that $\cos(\theta_i)^2 \geq 1/n$, thus $z_i \in [-1, -1/\sqrt{n}] \cup [1/\sqrt{n}, 1]$. Then, define

$$S_{2i-1} = \{l = (l_1, \ldots, l_{d_l}) \in \mathbb{R}^{d_l} \mid l_i > 0,\ l_i^2 \geq (n-1)\sum_{j \neq i} l_j^2\},$$

$$S_{2i} = \{l = (l_1, \ldots, l_{d_l}) \in \mathbb{R}^{d_l} \mid l_i < 0,\ l_i^2 \geq (n-1)\sum_{j \neq i} l_j^2\}.$$

we have either $S_{2i-1} \subset \{l \in \mathbb{R}^{d_l} | z^T l > 1\}$ or $S_{2i} \subset \{l \in \mathbb{R}^{d_l} | z^T l > 1\}$. Thus the proof is complete. □

6. Numerical Experiment

In this section we conduct a numerical experiment for the minimal salvage fund problem. In the experiment we pick $d = 15$ and $Q = (Q_{i,j} : i, j \in \{1, \cdots, d\})$ where $Q_{i,j} = 1/d$ if $i \neq j$ and otherwise $Q_{i,j} = 0$. In addition, L_i are i.i.d. Pareto random variables with cumulative distribution function $P(L_i > l) = (1/l)$. The limit measure $\mu(\cdot)$ for $L = (L_1, \ldots, L_d)$ is supported on the positive axes: we have $\mu(A) = 0$ for every measurable set A satisfying $A \subseteq (0, \infty]^d$; and for any index i, $\mu(\{L_i > l, L_j = 0 \text{ for } j \neq i\}) = (dl)^{-1}$. The confidence level is chosen as $\beta = 1 - 10^{-6}$.

The numerical experiment is conducted using a Laptop with 2.2 GHz Intel Core i7 CPU and the sampled linear programming problem is solved using CVXPY (Diamond and Boyd (2016)) with Gurobi, cf. Gurobi Optimization (2019). Our experimental results are given in Table 1 and 2. When $\delta \leq 10^{-3}$, solving the optimization problem $(CSP_{\delta,N'})$ costs much more time than simulating $L_\delta^{(i)}$, despite that a simple acceptance rejection scheme is applied to sample $L_\delta^{(i)}$ in our experiments.

	$\delta = 0.1$	$\delta = 0.01$	$\delta = 0.001$
Scenario Approach [Calafiore and Campi (2006)]	1206	18689	255689
This Paper	926	1893	2047

Table 1. Required number of samples for the chance constraint problem.

	$\delta = 0.1$	$\delta = 0.01$	$\delta = 0.001$
Scenario Approach [Calafiore and Campi (2006)]	3.768	107.1	3837
This Paper	3.351	6.299	7.050

Table 2. CPU time(s) for solving the chance constraint problem

Remark 3 (Explicit Computation of O_δ and C_δ)**.** In order to apply the algorithm in practice, one needs to compute the explicit expression of the set O_δ and C_δ. In the numerical experiment, we pick $\eta = 1/d$, and $c_i = d^{-1} + d^{-3}$. Suppose that $\{L_i, i =$

$1, \ldots, d\}$ are i.i.d. random variables with Pareto distribution that satisfies $P(L_i > l) = (1/l)^\alpha, \forall l \geq 1$. Then, we have

$$\begin{aligned}
&\mathrm{P}(c_i \mathbf{1}^T L + \eta \mathbf{e}_i^T Q^T L - \eta L_i > \eta x_i + m) \leq \delta,\ \forall x \in F_\delta \\
\Longrightarrow &\mathrm{P}((1 + d^{-1} + d^{-2})\mathbf{1}^T L - (1 + d^{-1})L_i > x_i + d \cdot m) \leq \delta,\ \forall x \in F_\delta \\
\Longrightarrow &\mathrm{P}((1 + d^{-1} + d^{-2})L_i > x_i + d \cdot m) \leq \delta,\ \forall x \in F_\delta \\
\Longrightarrow &x_i > (1 + d^{-1} + d^{-2})\delta^{-1/\alpha} - d \cdot m \\
\Longrightarrow &\alpha_\delta = (1 + d^{-1} + d^{-2})\delta^{-1/\alpha} - d \cdot m.
\end{aligned}$$

Thus, we have $O_\delta = \{x \in \mathbb{R}^d \mid \min_{i=1}^d x_i \geq \alpha_\delta\}$. In addition, we pick $\varepsilon = 0.1$, $S = \mathbb{R}^d$, K large enough such that $K^c \subset C_{\varepsilon,+}$, so it follows that

$$C_\delta = h(\alpha_\delta) \cdot \big(C_{\varepsilon,+} \cup K^c \cup S^c\big) = \{l \in \mathbb{R}^d \mid \max_{i=1}^d l_i \geq \alpha_\delta \cdot (m + 0.9)\}.$$

Acknowledgement

The authors are grateful to Alexander Shapiro for valuable feedback. The research of BZ is supported by NWO grant 639.033.413. The research of JB is supported by NSF grants 1915967, 1820942, 1838576 as well as DARPA award N660011824028.

References

Ahmed, S. and Shapiro, A. (2008). Solving chance-constrained stochastic programs via sampling and integer programming. In *State-of-the-Art Decision-Making Tools in the Information-Intensive Age*, pages 261–269. INFORMS.

Andrieu, L., Henrion, R., and Römisch, W. (2010). A model for dynamic chance constraints in hydro power reservoir management. *European Journal of Operational Research*, 207(2):579–589.

Blanchet, J. and Liu, J. (2010). Efficient importance sampling in ruin problems for multidimensional regularly varying random walks. *Journal of Applied Probability*, 47(2):301–322.

Bonami, P. and Lejeune, M. A. (2009). An exact solution approach for portfolio optimization problems under stochastic and integer constraints. *Operations Research*, 57(3):650–670.

Calafiore, G. and Campi, M. C. (2005). Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25–46.

Calafiore, G. and Campi, M. C. (2006). The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753.

Charnes, A., Cooper, W. W., and Symonds, G. H. (1958). Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3):235–263.

Chen, B., Blanchet, J., Rhee, C.-H., and Zwart, B. (2019). Efficient rare-event simulation for multiple jump events in regularly varying random walks and compound poisson processes. *Mathematics of Operations Research*, 44(3):919–942.

Diamond, S. and Boyd, S. (2016). CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5.

Eisenberg, L. and Noe, T. H. (2001). Systemic risk in financial systems. *Management Science*, 47(2):236–249.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media.

Gudmundsson, T. and Hult, H. (2014). Markov chain monte carlo for computing rare-event probabilities for a heavy-tailed random walk. *Journal of Applied Probability*, 51(2):359–376.

Gurobi Optimization, L. (2019). Gurobi optimizer reference manual.

Kley, O., Klüppelberg, C., and Reinert, G. (2016). Risk in a large claims insurance market with bipartite graph structure. *Operations Research*, 64(5):1159–1176.

Lagoa, C. M., Li, X., and Sznaier, M. (2005). Probabilistically constrained linear programs and risk-adjusted controller design. *SIAM Journal on Optimization*, 15(3):938–951.

Luedtke, J., Ahmed, S., and Nemhauser, G. L. (2010). An integer programming approach for linear programs with probabilistic constraints. *Mathematical Programming*, 122(2):247–272.

Nemirovski, A. and Shapiro, A. (2006a). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996.

Nemirovski, A. and Shapiro, A. (2006b). Scenario approximations of chance constraints. In *Probabilistic and Randomized Methods for Design under Uncertainty*, pages 3–47. Springer.

Prekopa, A. (1970). On probabilistic constrained programming. In *Proceedings of the Princeton Symposium on Mathematical Programming*, volume 113, page 138. Princeton, NJ.

Resnick, S. I. (2013). *Extreme values, regular variation and point processes*. Springer.

Wierman, A. and Zwart, B. (2012). Is tail-optimal scheduling possible? *Operations Research*, 60(5):1249–1257.

Department of Management Science and Engineering, Stanford University, Stanford, CA 94305

E-mail address: jose.blanchet@stanford.edu

Department of Management Science and Engineering, Stanford University, Stanford, CA 94305

E-mail address: fzh@stanford.edu

Centrum Wiskunde & Informatica, Amsterdam 1098 XG, Netherlands, and Eindhoven University of Technology.

E-mail address: bert.zwart@cwi.nl